

Școala doctorală în domeniul Științe medicale

Cu titlu de manuscris
C.Z.U: 616.94-037:004.8(043.2)

IAPĂSCURTĂ Victor

**PREZICEREA TIMPURIE A SEPSISULUI CU AJUTORUL UNEI
APLICAȚII PROPRII ELABORATE ÎN BAZA ÎNVĂȚĂRII
AUTOMATE (INTELIGENȚĂ ARTIFICIALĂ)**

321.19 – ANESTEZOLOGIE ȘI TERAPIE INTENSIVĂ

Teză de doctor în științe medicale

Chișinău, 2023

Teza a fost elaborată la Catedra de anesteziologie și reanimatologie nr.1 “Valeriu Ghereg”, IP Universitatea de Stat de Medicină și Farmacie “Nicolae Testemițanu” din Republica Moldova

Conducător

Belîi Adrian,

dr. hab. șt. med, profesor universitar

Membrii comisiei de îndrumare:

Cojocaru Svetlana,

dr. hab. informatică, prof. cercetător,
membru corespondent al AȘ a Moldovei

Ciobanu Nelly,

dr. șt. fizico-matematice, conf. univ.

Savan Veaceslav,

dr. șt. med.

Susținerea va avea loc la 27 septembrie, 2023, 14:00 în incinta USMF ”Nicolae Testemițanu”, bd. Ștefan cel Mare și Sfânt, 165, biroul 205 în ședința Comisiei de susținere publică a tezei de doctorat, aprobată prin decizia Consiliului Științific al Consorțiului din 25 mai, 2023 (*proces verbal nr 5*).

Componența Comisiei de susținere publică a tezei de doctorat:

Președinte:

Cojocaru Victor,

dr. hab. șt. med, prof. univ.

Membri:

Arnaut Oleg,

dr. hab. șt. med, prof. univ.

Baltaga Ruslan,

dr. șt. med, conf. univ.

Belîi Adrian,

dr. hab. șt. med, profesor universitar

Cobeț Valeriu,

dr. hab. șt. med., prof. univ.

Cojocaru Svetlana,

dr. hab. informatică, prof. cercetător, m.c. AȘ

Cornogolub Alexandru,

dr. șt. med., conf. univ.

Autor

Iapăscurtă Victor

© Iapăscurtă Victor, 2023

© Iapăscurtă Victor, 2023

CUPRINS

LISTA ABREVIERILOR	6
LISTA TABELELOR	8
LISTA FIGURILOR	10
INTRODUCERE	13
1. APLICAREA TEHNOLOGIILOR DE ÎNVĂȚARE AUTOMATĂ / INTELIGENȚĂ ARTIFICIALĂ ÎN ANESTEZIOLOGIE-TERAPIE INTENSIVĂ ȘI ÎN MANAGEMENTUL SEPSISULUI	19
1.1. Tehnologii de învățare automată / inteligență artificială utilizate în anesteziologie - terapie intensivă	19
1.2. Conceptul de sepsis. Evoluția conceptului și conținutul lui modern	25
1.3. Situația actuală în abordarea problemei sepsisului cu utilizarea tehnologiilor intelectuale în managementul clinic al stării de sepsis	27
1.4. Sinteza Capitolului 1.	47
2. METODOLOGIA CERCETĂRII ÎN STUDIILE BAZATE PE ÎNVĂȚAREA AUTOMATĂ / INTELIGENȚA ARTIFICIALĂ. MATERIALE ȘI METODE DE CERCETARE ÎN STUDIUL CURENT	49
2.1. Designul general al cercetării	49
2.2. Descrierea materialului clinic	53
2.3. Aspecte metodologice importante pentru investigație, procesarea datelor și analiza rezultatelor.	56
2.4. Sinteza Capitolului 2	74
3. ETAPELE DE PROCESARE A DATELOR	75
3.1. Reconstruirea datelor prin rezolvarea problemei valorilor lipsă	75
3.2. Complexitatea algoritmică Kolmogorov-Chaitin ca metrică și metodă de procesare a datelor	82
3.3. Procesarea datelor cu crearea caracteristicilor pentru a fi utilizate în sistemul de învățare automată	84
3.4. Sinteza Capitolului 3	88
4. ÎNVĂȚAREA AUTOMATĂ. REZULTATE INTERMEDIARE ȘI DISCUȚIA LOR	89
4.1. Etapa de învățare automată	89
4.2. Elaborarea, validarea și testarea sistemului de învățare automată creat	92
4.3. Incorporarea sistemului creat într-o aplicație pentru utilizare clinică	96
4.4. Utilizarea aplicației pentru precizie continuă a riscului dezvoltării sepsisului.	98
4.5. Sinteza Capitolului 4	100

5. EXPLICABILITATEA MODELELOR DE INTELIGENȚĂ ARTIFICIALĂ. DISCUȚII ȘI DIRECȚII DE CERCETARE DE VIITOR	102
5.1. Sumar al lucrului de cercetare efectuat. Discuția rezultatelor. Limitări	102
5.2. Aspecte de explicabilitate a modelelor de inteligență artificială	110
5.3. Transferabilitatea sistemelor de inteligență artificială. Direcții de cercetare de viitor	121
5.4. Sinteza Capitolului 5	123
CONCLUZII GENERALE	125
RECOMANDĂRI	126
BIBLIOGRAFIE	127
ANEXE	143
DECLARAȚIA PRIVIND ASUMAREA RĂSPUNDERII	168
CV-ul AUTORULUI	169

LISTA ABREVIERILOR

ATI	- Anestezie-terapie intensivă
ACC	- Analiza completă a cazului (abordare statistică)
AUC ROC	- Aria de sub curba ROC
AutoML	- Învățare automată automatizată (<i>eng.</i> Automated Machine Learning)
BDM	- <i>eng.</i> Block decomposition method – metodă matematică
BN (NB)	- <i>eng.</i> Bayesian Network/ Naive Bayes - algoritmi
CA	- Complexitate algoritmică
Curba-PR	- Curba Precision-Recall
DRF	- <i>eng.</i> Distributed Random Forest (pădure aleatorie) - algoritm
EtCO₂	- Presiunea CO ₂ la sfârșitul expirului
FC	- Frecvența cardiacă (<i>eng.</i> HR)
FDA	<i>eng.</i> Food and Drug Administration, S.U.A.
FR	- Frecvența respirației (<i>eng.</i> RR, Resp)
FiO₂	- Frația oxigenului inspirat
GBM	- <i>eng.</i> Gradient Boosting Machine (stimularea coborârii gradientului) - algoritm
GLM	- <i>eng.</i> Generalized Linear Model – model în baza regresiei liniare
GPowerWin	- Aplicație software statistică
H2O.ai	- Platformă pentru învățarea automată
IA	- Inteligență artificială (<i>eng.</i> AI, Artificial Intelligence)
ICE	- <i>eng.</i> Individual Conditional Expectations (plot) – tip de diagramă
Î	- Interval de încredere
ÎA	- Învățare automată (<i>eng.</i> ML, Machine Learning)
ÎC	- Învățare prin consolidare (<i>eng.</i> Reinforcement ML)
ÎNS	- Învățare (automată) nesupravegheată, (<i>eng.</i> Non-supervised ML)
ÎP	- Învățare profundă (<i>eng.</i> Deep Learning, DL) – varietate ÎA
ÎS	- Învățare (automată) supravegheată, (<i>eng.</i> Supervised ML)
Kolmogorov-Chaitin	- Metrică a complexității algoritmice/metodă matematică
LIME	<i>eng.</i> Local Interpretable Model-agnostic Explanations
LOCF	- <i>eng.</i> Last Observation Carried Forward (metodă statistică)
MAD	- Metode cu arbori de decizie (<i>eng.</i> Decision Tree) - algoritm

MVS	- Metode cu vectori de suport (<i>eng.</i> Support Vector Machine) - algoritm
PAD	- Presiunea arterială diastolică (<i>eng.</i> DBP)
PAS	- Presiunea arterială sistolică (<i>eng.</i> SBP)
PD	- <i>eng.</i> Partial Dependence (plot) – tip de diagramă
Python	- Limbaj de programare
PROBAST	- <i>eng.</i> Prediction model Risk Of Bias ASsessment Tool - ghid
R	- Limbaj de programare
RLin	- Regresie liniară
RLog	- Regresie logistică
RNA	- Rețea neuronală artificială (<i>eng.</i> Artificial Neural Network) - algoritm
RNR	- Rețea neuronală recurentă (<i>eng.</i> Recurrent Neural Network) - algoritm
RNC	- Rețea neuronală convoluțională (<i>eng.</i> Convolutional Neural Network) - algoritm
Se	- Sensibilitate
SHAP	- <i>eng.</i> SHapley Additive exPlanations – tip de diagramă/metodă
SIRS	- <i>eng.</i> Systemic Inflammatory Response Syndrome
SÎA / SIA	- Sistem de învățare automată / sistem de inteligență artificială
SOFA	- <i>eng.</i> Sequential Organ Failure Assessment (scor)
Sp	- Specificitate
STARD	- <i>eng.</i> Standards for Reporting of Diagnostic Accuracy Study - ghid
TRIPOD	- <i>eng.</i> Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis - ghid
Temp	- Temperatura corpului (uman)
UTI	- Unitate (secție) de terapie intensivă
Wolfram	- Limbaj de programare
XAI	- <i>eng.</i> E(X)plainable Artificial Intelligence

LISTA TABELELOR

Tabelul 1.1	Direcții și aplicații clinice ale învățării automate / inteligenței artificiale în ATI	23
Tabelul 1.2	Date sumare despre performanța predicției sepsisului cu modelele ÎA (pentru ultimii 5 ani)	42
Tabelul 2.1	Analiza comparativă a terminologiei ÎA versus modelarea statistică tradițională	50
Tabelul 2.2	Caracterizarea generală a seturilor de date utilizate	54
Tabelul 3.1	Aspectul inițial al datelor unui pacient (p000009)	75
Tabelul 3.2	Caracteristicile distribuției și normalitatea datelor (pentru fiecare variabilă)	80
Tabelul 3.3	Aspectul datelor aceluiași pacient (p000009) după reconstrucție	82
Tabelul 3.4	Matricea de corelare a parametrilor fiziologici în sepsis	86
Tabelul 3.5	Matricea de corelare a parametrilor fiziologici la non-septici	87
Tabelul 4.1	Metrica comparativă a două sisteme de prezicere a sepsisului (InSight versus sistemul creat în cadrul cercetării curente)	90
Tabelul 4.2	Formatul final al datelor transmise algoritmului de învățare automată	91
Tabelul 4.3	Performanța sistemului în baza algoritmului GBM	92
Tabelul 4.4	Robustețea modelului în funcție de volumul datelor utilizate pentru training și calitatea lor	93
Tabelul 4.5	Matricea de confuzie la validarea multiplă (x 10) a modelului	94
Tabelul 4.6	Matricea de confuzie la testarea modelului pe setul 1 (A)	94
Tabelul 4.7	Matricea de confuzie la testarea modelului pe setul 2 (B)	95
Tabelul 5.1	Performanța modelelor finale după aria de sub curba ROC (AUC)	103
Tabelul 5.2	Caracteristica comparativă a 4 sisteme create în baza acelorași date	105
Tabelul 5.3	Informație sumară despre sisteme similare cu o performanță mai înaltă decât a celui creat în studiul curent (ultimii trei ani)	107
Tabelul A.3.1	Matricea de corelare a parametrilor fiziologici în sepsis (set B)	146
Tabelul A.3.2	Matricea de corelare a parametrilor fiziologici la non-septici	147
Tabelul A.4.1	Matricea de corelare a parametrilor fiziologici în pre-sepsis (set A)	148
Tabelul A.4.2	Matricea de corelare a parametrilor în sepsis confirmat clinic	149
Tabelul A 5.1	Caracteristicile setului “Sepsis”	150
Tabelul A 5.2	Caracteristicile setului “Non-sepsis”	150
Tabelul A 5.3	Testul Kolmogorov-Smirnov cu două eșantioane: sepsis vs non-sepsis	150
Tabelul A 6.1	Caracteristicile setului “Pre-sepsis”	151
Tabelul A 6.2	Caracteristicile setului “Non-sepsis”	151

Tabelul A 6.3	Testul Kolmogorov-Smirnov cu două eșantioane: pre-sepsis vs non-sepsis	151
Tabelul A.8.1	Rezultatele orare ale prezicerii riscului de dezvoltare a sepsisului la 10 pacienți septici timp de 9 ore consecutive	154
Tabelul A.8.2	Rezultatele orare ale prezicerii riscului de dezvoltare a sepsisului la 10 pacienți fără sepsis timp de 9 ore consecutive	154
Tabelul A 10	Performanța a șapte modele antrenate pe date din setul A în cadrul învățării automate automatizate (AutoML) în baza platformei H2O (etapa de cross-validare)	159

LISTA FIGURILOR

Figura 1.1	Frecvența publicațiilor (PubMed) în domeniul inteligenței artificiale/ învățării automate în terapie intensivă / anesteziologie pe perioada 1980 – 2021	20
Figura 1.2	Ponderea principalelor domenii clinice vizate de aplicațiile de inteligența artificială (I) și a tipurilor de învățare automată (II) în ATI	22
Figura 1.3	Frecvența publicațiilor (PubMed) în domeniul inteligenței artificiale/ învățării automate și sepsis pe perioadă 1990 – 2021	27
Figura 1.4	(a) Partiția spațiului de caracteristici cu ajutorul liniilor paralele axelor de coordonate; (b) Un arbore decizional care determină dacă un pacient este septic urmând definiția Sepsis-3	33
Figura 1.5	Mașina cu vectori de suport cu utilizarea a două caracteristici: temperatura corpului și frecvența contracțiilor cardiace	34
Figura 1.6	Regresia logistică	36
Figura 1.7	O rețea neuronală artificială cu șase predictorii și un strat ascuns	38
Figura 1.8	Schema unei rețele recurente	39
Figura 1.9	Un model naïve Bayes	41
Figura 2.1	Designul general al studiului	51
Figura 2.2	Elementele cheie pentru raportarea designului și a rezultatelor studiului curent.	52
Figura 2.3	Densitatea (prezența) datelor: Primii 20 parametri fiziologici și de laborator cu cea mai înaltă densitate a datelor (în procente)	55
Figura 2.4	Sumarul caracteristicilor datelor utilizate în studiu	56
Figura 2.5	Diagrama de tip CONSORT pentru studiul curent	58
Figura 2.6	Elementele cheie pentru raportarea trainingului și validării modelului	61
Figura 2.7	Partea incipientă a unui arbore decizional, construit în baza datelor din studiu	62
Figura 2.8	Principalele etape ale fluxului de date în cercetarea curentă	66
Figura 3.1	Reprezentarea grafică a distribuției valorilor parametrilor de interes	78
Figura 3.2	Algoritmul pentru reconstrucția datelor în cazul lipsei valorilor	81
Figura 3.3	Formatul datelor la 3 etape succesive de prelucrare a lor	84
Figura 3.4	Corelarea parametrilor fiziologici de interes în sepsis	85
Figura 3.5	Corelarea parametrilor fiziologici de interes în cazurile non-septice	86
Figura 4.1	Etapele și elementele principale în crearea unui sistem de învățare automată	89
Figura 4.2	Etapele fluxului de procesare a datelor la crearea unui SÎA	89
Figura 4.3	Relația dintre setul de testare și celelalte date (p-u training și validare)	90

Figura 4.4	Diminuarea erorii și performanța în procesul trainingului (în dependență de numărul de arbori)	94
Figura 4.5	Curba ROC la testarea modelului pe setul-test 1 (A)	95
Figura 4.6	Curba ROC la testarea pe setul-test 2 (B)	95
Figura 4.7	Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: introducerea datelor și vizualizarea lor în formă de tabel și grafică	96
Figura 4.8	Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: obținerea rezultatului predicției și vizualizarea dinamicii riscului	97
Figura 4.9	Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: introducerea datelor unui pacient nou	97
Figura 4.10	Secvența temporală a evenimentelor relevante în cazul unui pacient cu sepsis la utilizarea sistemului de predicție timpurie a sepsisului	98
Figura 4.11	Dinamica riscului sepsisului estimată de sistemul creat la 10 pacienți cu sepsis timp de 11 ore consecutive	99
Figura 4.12	Dinamica riscului sepsisului estimată de sistemul creat la 10 pacienți neseptici timp de 11 ore consecutive	99
Figura 5.1	Importanța pentru predicție a primelor 10 variabile în cazul unui model GBM	116
Figura 5.2	Importanța pentru predicție a variabilelor pentru cinci modele	117
Figura 5.3	Diagrama PD pentru diferența de temperatură	118
Figura 5.4	Diagrama PD pentru diferența de temperatură în cazul a trei modele	118
Figura 5.5	Diagrama ICE pentru diferența de temperatură în cazul unui model GBM cu performanță înaltă	119
Figura 5.6	Diagrama SHAP-sumar pentru unul din modelele GBM cu performanță înaltă	120
Figura A.3.1	Corelarea parametrilor fiziologici în sepsis (setul B)	146
Figura A.3.2	Corelarea parametrilor fiziologici în cazurile non-septice (set B)	147
Figura A.4.1	Corelarea parametrilor fiziologici în pre-sepsis (setul A)	148
Figura A.4.2	Corelarea parametrilor în cazurile septice, confirmate clinic	149
Figura A.7.1	Curbele de învățare și cea obținută la cross-validare	152
Figura A.7.2	Curba ROC și aria de sub ea în rezultatul trainingului modelului	152
Figura A.7.3	Curba AUC-PR	153
Figura A.7.4	Corelarea performanței modelelor construite în baza diferitor algoritmi	153
Figura A.11.1	Diagrama PD (dependența parțială) pentru diferența de temperatură dintre ora a doua și ora precedentă în cazul unui model GBM	160
Figura A.11.2	Diagrama PD (dependența parțială) pentru diferența de temperatură dintre ora a treia de observație și ora precedentă în cazul unui model GBM	160
Figura A.11.3	Diagrama PD pentru diferența SpO ₂ în cazul unui model GBM	161

Figura A.11.4	Diagrama ICE pentru diferența de temperatură între ora a treia de observație și ora a doua în cazul unui model GBM cu performanță înaltă	161
Figura A.11.5	Diagrama ICE pentru diferența de temperatură între ora a doua de observație și ora precedentă în cazul unui model GBM cu performanță înaltă	162
Figura A.11.6	Diagrama de corelare (după performanță) între modele din clase diferite	162
Figura A.11.7	Diagrama de corelare (după performanță) între modele din aceeași clasă	163

INTRODUCERE

Actualitatea și importanța problemei abordate. Pot fi evidențiate două aspecte importante care conturează direcția de cercetare reflectată în lucrare:

(a) Problema sepsisului ca varietate de stare critică, deseori dificil de diagnosticat la timp, rezultatele la distanță fiind influențate de timpul când se începe tratamentul, iar aceasta, la rândul său, determină mortalitatea (Strandberg G. et al., 2020; Evans L. et al. 2021), care în timp s-a redus neînsemnat (Stevenson E. et al., 2014) și

(b) Apariția unui nou actor – a tehnologiilor de inteligență artificială (IA), care deseori se declară drept panacee pentru mai multe probleme, inclusiv medicale.

Cercetarea actuală încearcă să evalueze posibilitatea utilizării acestor tehnologii pentru rezolvarea unor aspecte importante în managementul acestui grup de pacienți prin prezicerea timpurie (câteva ore până la debut) a stării de sepsis, ce ar reduce timpul până la inițierea tratamentului.

Descrierea situației în domeniu și identificarea problemei de cercetare. În ciuda unei mari mortalități asociate (Strandberg G. et al., 2020) și a costurilor ridicate ale tratamentului, sepsisul rămâne dificil de diagnosticat și de tratat (Stevenson E. et al., 2014; Moor M. et al., 2022). Cercetările anterioare au subliniat importanța recunoașterii în timp util a sepsisului pentru îmbunătățirea rezultatelor și reducerea costurilor asociate tratamentului (Burdick H. et al., 2020).

Republica Moldova urmează tendințele mondiale, problema sepsisului fiind una importantă. Spre exemplu, conform OMS și studiilor autohtone (Friptu V., Hodoroșea S. 2014; Donduci Iu., 2018) sepsisul este a doua cea mai importantă cauză a mortalității materne.

Recent au fost propuse noi definiții destinate să îmbunătățească recunoașterea clinică a sepsisului (criteriile “Sepsis-3”, 2016), deoarece utilizarea anterioară a screening-ului pe baza sindromului de răspuns inflamator sistemic (SIRS) a fost găsită ca fiind nespecifică (Shankar-Hari M., 2016). Datele din literatura medicală au arătat că diagnosticarea și inițierea tratamentului precoce poate reduce riscul apariției unor rezultate adverse ca consecințe ale sepsisului (Damiani E. et al., 2015; Shimabukuro D. et al, 2017). Prin urmare, recunoașterea precoce a sepsisului și identificarea pacienților cu risc sporit de dezvoltare a sepsisului este esențială pentru tratamentul eficient. Instrumentele de screening cel mai des utilizate în condiții clinice pentru identificarea pacienților septici includ SOFA (Sequential Organ Failure Assessment), criteriile SIRS (Systemic Inflammatory Response Syndrome), MEWS (Modified Early Warning Score).

În pofida unor limitări, aceste sisteme de scor au stabilit valori de performanță și servesc ca comparatori importanți pentru sistemele de detectare și predicție a sepsisului recent dezvoltate și pentru evaluarea eficienței lor clinice (Finkelsztein E. et al., 2017; Shimabukuro D. et al, 2017).

Un supliment recent la aceasta colecție de instrumente este reprezentat de sistemele bazate pe învățarea automată (SÎA) (Burdick H. et al., 2020; Evans L et al. 2021). Deși SÎA reprezintă suplimente destul de noi în domeniul managementului clinic al sepsisului, ele au potențialul de a îmbunătăți rezultatele pacienților prin avertizarea în avans a debutului iminent al sepsisului, astfel asistând medicul în procesul de luare a deciziilor. În prezent în domeniul cercetării sepsisului, sistemele de asistență decizională bazate pe învățare automată reprezintă o direcție cu dezvoltare rapidă (Delahanty, R. et al., 2019; Burdick, H. et al., 2020; Evans L et al. 2021; Wu M et al. 2021; Moor M. et al. 2021).

Scopul studiului. Evaluarea fezabilității tehnologiilor IA în managementul pacientului critic din unitatea de terapie intensivă cu risc de a dezvolta sepsis, cu elaborarea unui sistem cu abilități discriminative (sepsis versus non-sepsis), care ar permite prezicerea timpurie a dezvoltării sepsisului.

Obiectivele studiului

1. Evaluarea utilizării tehnologiilor de inteligență artificială și, în special, a învățării automate, ca una din tehnologiile de bază ale IA la etapa actuală, utilizate în anestezie-terapie intensivă.
2. Evaluarea utilizării SIA în managementul pacienților cu sepsis, în stare critică.
3. Identificarea unui set de date în volum suficient pentru crearea unui eventual sistem de prezicere timpurie a sepsisului.
4. Analiza exploratorie a datelor clinice și de laborator și procesarea lor în modul necesar pentru crearea unui sistem de discriminare/prezicere.
5. Crearea unui sistem de tipul unei aplicații practice, care ar permite prezicerea dezvoltării sepsisului la pacientul din secțiile de terapie intensivă.

Ipoteze de cercetare:

- Detectarea precoce a sepsisului și recunoașterea mai exactă a pacienților cu risc ridicat de dezvoltare a sepsisului este esențială pentru tratamentul eficient. Instrumentele de screening utilizate uzual în condiții clinice pentru identificarea pacienților septici (SOFA, criteriile SIRS, MEWS) nu posedă proprietăți de predicție timpurie. În acest context pentru prezicerea timpurie a sepsisului ar fi binevenite alte unelte, iar instrumentele descrise ar

putea servi ca comparatori importanți pentru sistemele de detectare și predicție a sepsisului eventual create și pentru evaluarea efectului lor asupra rezultatelor clinice (Finkelsztejn E. et al., 2017)

- Analiza și prelucrarea unor volume suficiente de date ar putea contribui la crearea unor sisteme de precizie timpurie a sepsisului (cel puțin 3 ore până la debut). Un așa sistem, eventual creat și livrat sub formă de aplicație software pentru utilizare clinică în Republica Moldova, ar servi pentru suportul decizional în ATI.

Metodologia cercetării științifice. Cercetarea este un studiu retrospectiv al unui set de date ce conține 40336 pacienți/cazuri, dintre care 7,26% sunt pacienți cu sepsis. Restul – 92.74% reprezintă pacienți internați în secțiile de terapie intensivă cu alte diagnostice (non-sepsis). Analiza exploratorie și procesarea datelor sunt în mare măsură determinate de pașii ulteriori ai cercetării, ce presupun crearea în final a unui sistem de inteligență artificială cu abilități de precizie a unui eveniment (sepsis). Pentru analiza și prelucrarea datelor, inclusiv și cea statistică, și procesarea lor ulterioară s-a utilizat mediul integrat de dezvoltare RStudio în baza limbajului de programare R (<http://www.R-project.org/>), care a fost și limbajul de interacțiune cu platforma H2O (<https://www.h2o.ai/>) pentru învățarea automată propriu-zisă. Calculele de înaltă intensitate au fost realizate utilizând Amazon Web Services (AWS EC2 <https://aws.amazon.com/ec2/>). Pentru evaluarea performanței modelului creat și raportarea rezultatelor s-a utilizat ghidul TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis), (Collins, G. et al., 2015, 2021) și principiile PROBAST (Prediction model Risk Of Bias ASsessment Tool), (Wolff, R. et al., 2019, 2021). Aplicația software pentru utilizare practică este creată cu ajutorul pachetului “Wave” (0.0.14) în Python (3.8.5).

Noutatea și originalitatea științifică. În baza analizei unui set larg de date (40366 cazuri, dintre care – 2932 diagnosticate cu sepsis) și prelucrare a lor cu utilizarea unui nou algoritm de restabilire a datelor-lipsă și utilizarea metricii complexității algoritmice s-a creat un sistem de suport decizional pentru precizarea timpurie a sepsisului.

Problema științifică importantă soluționată în teză. Sepsisul reprezintă o problemă actuală în serviciul medical/serviciul terapiei intensive, iar diagnosticarea lui precoce poate duce la sporirea ratei de succes a tratamentului, diminuarea mortalității și reducerea costului asistenței livrate acestui grup de pacienți, în special în cazurile complexe. Rezultatul obținut care contribuie la soluționarea unei probleme științifice importante constă în elaborarea unui sistem în baza învățării automate, care sporește eficiența managementului clinic al pacienților cu sepsis.

Semnificația teoretică a cercetării. A fost explorată și confirmată posibilitatea utilizării conceptelor și metricii dinamicii algoritmice în reprezentarea datelor medicale, inclusiv și a celor în formă de serii temporale, care descriu starea clinică a pacientului. Această reprezentare s-a dovedit a fi reușită și la etapele ulterioare de procesare a datelor – în construirea sistemului de prezicere, iar performanța înaltă a sistemului este un argument suplimentar. Aspectele ce se referă la valoarea predictivă a unor parametri clinici, care au fost elucidate în studiu, ar putea contribui la o înțelegere mai bună a problemei sepsisului ca fenomen medical.

Valoarea aplicativă a lucrării. Aplicația software creată, în care este integrat sistemul de prezicere timpurie a sepsisului elaborat, poate asista medicul din terapie intensivă în procesul de luare a deciziilor, în special în cazurile de sepsis mai complexe și în deosebi în situațiile ambigue. Metodele propuse pentru reconstrucția și reprezentarea datelor pot facilita, diversifica și înviora activitatea cercetătorilor în domeniu.

Rezultatele științifice principale înaintate spre susținere:

1. Algoritmul de reconstrucție a datelor reprezentate prin serii temporale ale parametrilor fiziologici, care conțin valori lipsă;
2. Metoda de selectare a datelor pentru crearea unui sistem de prezicere a stării pacientului (pe exemplul pacienților cu risc de a dezvolta sepsis);
3. Metoda de prelucrare a datelor prin calcularea complexității algoritmice cu ajutorul metodei de decompoziție în blocuri (BDM);
4. Sistemul automat pentru prezicerea sepsisului încorporat într-o aplicație software pentru utilizare clinică;

Implementarea rezultatelor științifice.

- La etapa finală a fost creată o aplicație software pentru a fi transmisă în Clinica ATI a Institutului de medicină urgentă, Chișinău. Aplicația are o interfață grafică care permite introducerea datelor (frecvența cardiacă, SpO₂, temperatura, PA sistolică, PA diastolică, frecvența respirației) unui pacient concret pe o perioadă de trei ore cu posibilitatea vizualizării lor grafice și obținerea rezultatului prezicerii, adică a riscului de dezvoltare a sepsisului cu un orizont de 4 ore.
- O parte din rezultatele cercetării sunt încorporate în cursul “Advanced perioperative monitoring. Elements of computational medicine” (pentru studenții a. VI la USMF “Nicolae Testemițanu”)
- Rezultatele de etapă ale cercetării prezentate la:

- ✓ Cursurile Comitetului European pentru Educare în Anestezie, ediția a XIII-a, Modulul 1, 7-9 februarie, 2018,
 - ✓ Cursurile Comitetului European pentru Educare în Anestezie, ediția a XV-a, Modulul 3, 10-11 decembrie, 2020,
 - ✓ Cursurile Comitetului European pentru Educare în Anestezie, ediția a XVI-a, Modulul 5, 9-11 decembrie, 2021,
 - ✓ 5th International Conference on Nanotechnologies and Biomedical Engineering, November 3-5, 2021,
 - ✓ The 12th International Conference on Electronics, Communications and Computing, 20-21 october, 2022,
 - ✓ Conferința științifică anuală a USMF “N. Tetseanu”. Cercetarea în biomedicină și sănătate: calitate, excelență și performanță, Chișinău, 19-21 octombrie, 2022
- Utilizarea BDM pentru analiza și reprezentarea datelor medicale în formă de serii temporale a fost prezentată la conferința AUTOMATA 2020, Stockholm, Sweden, 10-12 august, 2020
 - Activitate de brevetare: rezultatele cercetării oficializate prin două certificate de invenție (nr.18 din 12.12.2022 și nr.19 din 13.12.2022) și acte de implementare.

Publicații la tema tezei. La tema tezei sunt publicate 16 lucrări științifice care sunt parafrazate în conținutul tezei, observându-se astfel importanța teoretică și valoarea aplicativă a lucrării. Trei publicații – în reviste/culegeri cu factor de impact SCOPUS, publicații în calitate de singur autor – 6, prim-autor – 7.

Sumarul compartimentelor tezei. Teza este expusă pe 107 pagini de text de baza, structurat în 5 capitole, concluzii și recomandări practice, urmate de lista a 274 surse bibliografice și 12 anexe. Materialul iconografic include 52 figuri și 32 tabele.

Capitolul 1 include o trecere în revistă a datelor din literatura de specialitate pe trei direcții, care stau la baza compartimentelor acestui capitol și care conturează câmpul cercetării: (a) Tehnologiile de învățare automată / inteligență artificială utilizate în anesteziologie - terapie intensivă la etapa actuală; (b) Conceptul de sepsis, evoluția conceptului și conținutul lui modern; (c) Situația actuală în abordarea problemei sepsisului cu utilizarea tehnologiilor intelectuale în managementul clinic al stării de sepsis. Sunt descrise cele mai frecvent utilizate modele/algoritmi de învățare automată, în special pentru diagnosticarea/prezicerea sepsisului. Este prezentată performanța sistemelor create în baza modelelor menționate în ultimii cinci ani. La final este prezentată o analiză a principiilor și cerințelor de raportare a rezultatelor în studiile bazate pe învățarea automată/inteligența artificială (TRIPOD-AI, PROBAST-AI, STARD-AI).

În *Capitolul 2* sunt prezentate designul studiului; o analiză comparativă a terminologiei ÎA versus modelarea statistică tradițională; descrierea datelor clinice, utilizate în studiu; descrierea aplicării principiilor TRIPOD și PROBAST în studiile în baza IA, inclusiv criteriile de selectare/excludere a cazurilor; extrapolarea la studiul curent a principiilor de raportare TRIPOD-AI, PROBAST-AI, CONSORT-AI, SPIRIT-AI, STARD-AI; descrierea detaliată, inclusiv cu utilizarea formalismului matematic al algoritmului GBM (Gradient Boosting Machine), care este metoda de bază pentru ÎA în studiul curent. Sunt descrise particularitățile fluxului de date la diferite etape ale cercetării și metricele de performanță utilizate pentru evaluarea rezultatelor și raportare.

În *Capitolul 3* sunt descrise etapele de procesare a datelor, începând cu reconstrucția datelor prin restabilirea valorilor lipsă în datele inițiale. Este descrisă analiza datelor prin prisma diferitor statistici (ex. distribuția datelor pentru fiecare parametru fiziologic, ulterior utilizat la crearea modelului de predicție). Este descris algoritmul de restabilire a datelor elaborat în studiu și rezultatele aplicării lui pe datele cercetate. Este prezentată complexitatea algoritmică Kolmogorov – Chaitin ca metodă de procesare a datelor și rezultatele aplicării ei. Sunt descrise unele rezultate obținute în procesul prelucrării și analizei, care ar avea importanță pentru o înțelegere mai bună a sepsisului ca fenomen medical.

Capitolul 4 descrie învățarea automata propriu-zisă, rezultatele acestei etape și discuția lor. Este descris procesul de creare a unui sistem de învățare automata prin prisma componentelor și etapelor principale și particularitățile creării sistemului în cercetarea curentă. Este descris sistemul de referință (InSight), raportat în literatură ca unul din cele mai performante în prezicerea sepsisului, cu care se compara sistemul creat. Este prezentat și discutat formatul final al datelor înainte de a fi utilizate pentru ÎA. Sunt descrise și comentate etapele de creare, validare și testare a sistemului de predicție a sepsisului, elaborat în studiul curent, inclusiv testarea lui pe un set de date extern (validare externă). Este descrisă și ilustrată incorporarea sistemului creat într-o aplicație software pentru utilizare clinică. La final este descrisă utilizarea aplicației pentru prezicerea continue a riscului de sepsis și analizată performanța sistemului.

Capitolul 5 este axat pe problema explicabilității modelelor IA și direcțiile de cercetare de viitor. Începe cu un sumar al cercetării curente. Sunt descrise și discutate limitările sistemului elaborat. Este prezentată analiza comparativă a sistemului creat versus alte sisteme similare de înaltă performanță, descrise în literatură. Sunt descrise aspectele de explicabilitate a modelelor IA, ca factor-cheie în adoptarea clinică, acestea fiind extrapolate la sistemul elaborat în studiu. Sunt descrise potențialele direcții de cercetare de viitor. Fiecare capitol este finalizat cu o sinteză a celor descrise în capitolul respectiv.

Cuvinte cheie: sepsis, model, inteligență artificială, învățare automată, complexitate algoritmică, metoda de decompoziție în blocuri, sisteme de suport decizional, sisteme de prezicere.

1. APLICAREA TEHNOLOGIILOR DE ÎNVĂȚARE AUTOMATĂ / INTELIGENȚĂ ARTIFICIALĂ ÎN ANESTEZIOLOGIE-TERAPIE INTENSIVĂ ȘI ÎN MANAGEMENTUL SEPSISULUI

1.1. Tehnologii de învățare automată / inteligență artificială utilizate în anesteziologie - terapie intensivă

Etapă actuală de dezvoltare a științei și practicii medicale se caracterizează printr-o pătrundere “explozivă” a tehnologiilor de învățare automată / inteligență artificială, așteptându-se un beneficiu incontestabil și imediat atât pentru pacient, cât și pentru clinician. O teză doctorală din 2022 [1], ce are tangențe cu cercetarea curentă, conturează starea lucrurilor din acest domeniu în felul următor : *“Științele vieții din era digitală sunt conduse de moneda sa cea mai fundamentală și de neînlocuit: datele. Apariția algoritmilor de “Big data” și de învățare automată (ÎA) a promis că va revoluționa științele biomedicale și practica medicală prin intermediul diagnosticului automat, al subtipării bolilor, bazate pe date, și al tratamentelor personalizate. Cu toate acestea, în timp ce ÎA în sănătate a devenit un domeniu vibrant, în multe cazuri traducerea în practică s-a dovedit a fi mai provocatoare decât se aștepta, sau pentru a spune mai direct: revoluția este încă în așteptare”*.

Conform Food and Drug Administration, S.U.A., *“Pe măsură ce tehnologia continuă să avanseze fiecare aspect al îngrijirii sănătății, software-ul care încorporează inteligența artificială (IA), și în special subsetul de IA cunoscut sub numele de învățare automată (ÎA), a devenit o parte importantă a unui număr tot mai mare de dispozitive medicale. Unul dintre cele mai mari beneficii potențiale ale ML constă în capacitatea sa de a crea perspective noi și importante din cantitatea mare de date generate în timpul furnizării de îngrijiri medicale în fiecare zi”* [2].

O prognoză din 2018 a plasat valoarea pieței IA pentru asistența medicală în creștere de la 2 miliarde de dolari în 2018 la 36 de miliarde de dolari până în 2025, cu o rată de creștere anuală de 50% [3]. Conform unei prognoze mai recente (2021) în baza aceleiași surse, se estimează că inteligența artificială pe piața sănătății va crește de la 6,9 miliarde USD în 2021, la 67,4 miliarde USD până în 2027, cu o creștere anuală de 46,2% din 2021 până în 2027 [4].

Conform unui articol recent [5], aplicațiile comerciale ale inteligenței artificiale (AI) și învățării automate au înregistrat recent progrese remarcabile, în special în domenii precum recunoașterea imaginilor, procesarea limbajului natural, traducerea limbajului, analiza textelor și autoînvățarea. Cu toate acestea, tehnologiile descrise au atins cea mai mare performanță în aplicațiile cu o singură sarcină, în care pot fi tolerate rezultatele imperfecte și erorile ocazionale. Practica anesteziei-terapiei intensive (ATI) este diferită. Ea încorporează cerințe pentru fiabilitate

ridicată, restricții temporale de interpretare, acțiune/răspuns fizic, mai degrabă decât orice act cognitiv. Tranziția la medicina bazată pe dovezi și “Big Data” este un eveniment relativ recent. În schimb, medicii din serviciul anesteziologie-terapie intensivă (ATI) s-au bazat mult timp pe fluxuri personalizate de date cuantificate pentru a-și îngriji pacienții, iar progresele în monitorizare și volumul acestor date au susținut îmbunătățirile dramatice ale siguranței pacienților în domeniul ATI [6]. Specialiștii respectivi deseori acționează în circumstanțe “de muchie” a cuplului “cauză și efect”: de obicei, deciziile nu pot fi amânate, iar erorile de judecată pot fi adesea de neevitat. Un articol de sinteză din 2020 [7] indică următoarele domenii din specialitatea ATI cu cea mai înaltă rată de utilizare a tehnologiilor IA: (1) monitorizarea profunzimii anesteziei, (2) controlul anesteziei, (3) prezicerea evenimentelor și a riscurilor, (4) ghidarea cu ultrasunet, (5) managementul durerii și (6) logistica sălii de operație, la care mai pot fi suplimentate prezicerea mortalității, suportul decizional în ventilația mecanică, prezicerea sepsisului [8].

Analiza dinamicii în timp a frecvenței publicațiilor (PubMed) la intersecția domeniilor învățare automată / inteligență artificială cu domeniile anesteziologie / terapie intensivă pentru perioada 1980 – 2021 (Fig. 1.1) denotă o creștere exponențială a numărului publicațiilor începând cu anul 2018.

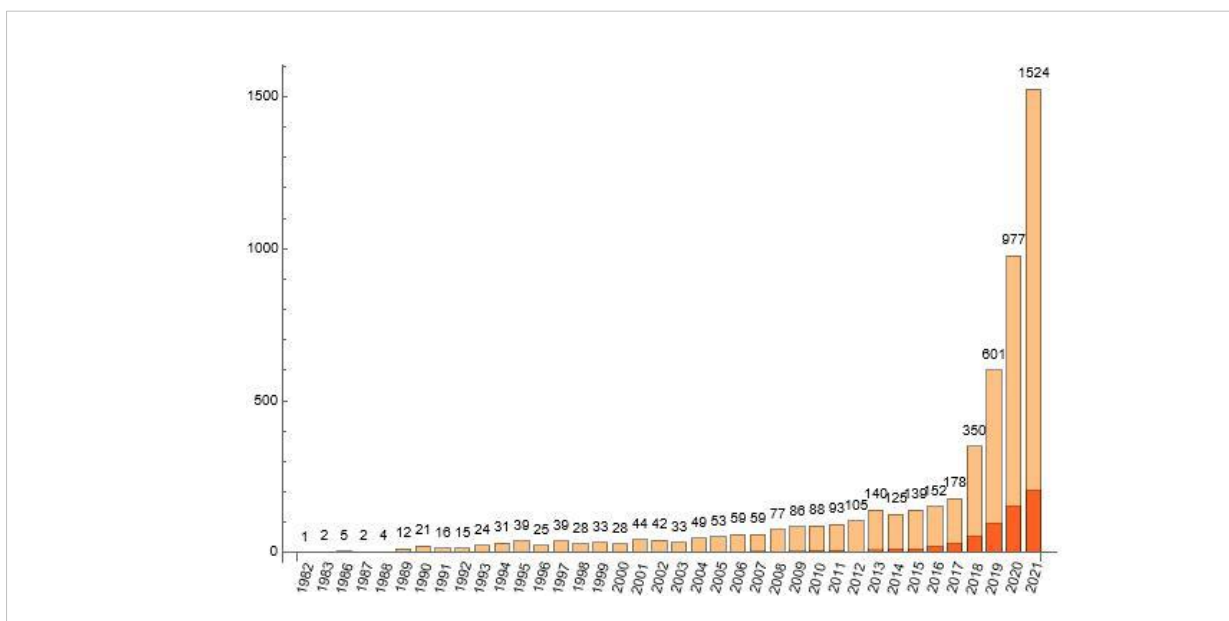


Figura 1.1. Frecvența publicațiilor (PubMed) în domeniul inteligenței artificiale/ învățării automate în terapie intensivă / anesteziologie pe perioadă 1980 – 2021

Notă: Cu culoare mai întunecată este reprezentată frecvența surselor ce vizează sepsisul.

O sinteză sistematică din 2022 [9], care include o evaluare a unui număr larg de surse ce descriu domeniul în perioada 1980 – 2020, ilustrează importanța dezvoltării sistemelor inteligente

de sprijinire a deciziilor clinice bazate pe învățarea automată și menționează că aceste sisteme pot ajuta în mod direct clinicienii la diagnosticarea cu acuratețe, prognozarea rezultatelor, identificarea evenimentelor cu risc sau la luarea unei decizii terapeutice.

Astfel, această lucrare descrie următoarele aspecte/direcții de aplicare a IA în ATI (în paranteze este indicată ponderea lucrărilor publicate, caracteristice pentru un anumit subdomeniu): 4 direcții clinice: (a) monitorizare, detecție și diagnostic (13,2%); (b) identificarea precoce a evenimentelor clinice (33,0%); (c) predicția rezultatului și evaluarea prognosticului (47,6%), și (d) decizii de tratament (6,2%). Din cele 97 de lucrări selectate în final, 82 (84,5%) studii au folosit datele pacienților adulți, 9 (9,3%) studii au folosit datele de la copii și adolescenți și 6 (6,2%) studii au folosit datele de la nou-născuți. O mare parte de studii - 65 (67,0%) au folosit date dintr-un singur centru, iar 32 (33,0%) studii au folosit seturi de date multicentrice; 88 (90,7%) studii au folosit învățarea supravegheată, 3 (3,1%) studii au folosit învățarea nesupravegheată și 6 (6,2%) studii au folosit învățarea prin întărire. Categoriile de variabile clinice, începând cu cele mai frecvent utilizate, au fost datele demografice (n=74), valorile de laborator (n=59), semnele vitale (n=55), scorurile (n=48), parametrii ventilatori (n=43), comorbiditățile (n=27), medicamentele (n=18), rezultatele (n=14), balanțele hidrice (n=13), intervențiile non medicale (n=10), simptomele (n=7) și istoricul medical (n= 4). Măsurile de evaluare adoptate cel mai frecvent pentru studiile de modelare a datelor clinice au inclus aria sub curba ROC (n=61), sensibilitatea (n=51), specificitatea (n=41), acuratețea (n=29) și valoarea predictivă pozitivă (n=23).

O analiză mai detaliată a celor 4 direcții enumerate mai sus relevă următoarele aplicații clinice: (I) Monitorizarea, detectarea și diagnosticarea (13 studii) include următoarele aspecte: (a) Monitorizarea indicatorilor fiziologici [10-13]; (b) Anomaliile ventilației mecanice (în special, asincronia pacientului cu ventilatorul) [14-16]; (c) Monitorizarea electroencefalografică (la pacienți cu epilepsie, encefalopatie metabolică, hipoxie cerebrală, ischemie și hematom subdural în leziunile cerebrale traumatice) [17-20]; (d) Detectarea infecției [21-23]; (II) Identificarea timpurie sau predicția evenimentelor clinice (32 studii), care includ: (a) Predicția leziunilor renale acute [24-26], (b) Predicția sepsisului și a infecției [27-37], (c) Predicția bolilor respiratorii și a ventilației mecanice [38-45], (d) Predicții și evaluări ale altor condiții clinice (10 studii, vizând prognoza unor parametri fiziologici, ca debitul urinar, glucoza serică, acidul lactic, nivelul anemiei, etc.) [46-54]; (III) Evaluarea rezultatelor și evaluarea prognostică (46 lucrări), ce include: (a) Rezultatele pacienților cu boli specifice (mortalitatea în sepsis [55-59], leziunea renală acută [60-62], leziunea hepatică [63-65], cancerul [66, 67], bolile cardiace [68, 69], patologia cerebrală [70, 71], pulmonară și altele [72-77], (b) Prognosticul tratamentului și evaluarea timpului de ședere în unitatea ATI [78-89]; (IV) Deciziile de tratament (6 studii) exprimate prin: (a) Abordarea

dozei optime de medicament (vasopresină, heparină, morfină) [90-93] și (b) Definirea momentului optim pentru detubare [94, 95].

Din cele trei tipuri de învățare automată de bază întâietatea o deține învățarea supravegheată (91%), urmată de cea prin consolidare (6%) și învățarea nesupravegheată (3%). Cele expuse sunt rezumate în diagramele de mai jos:

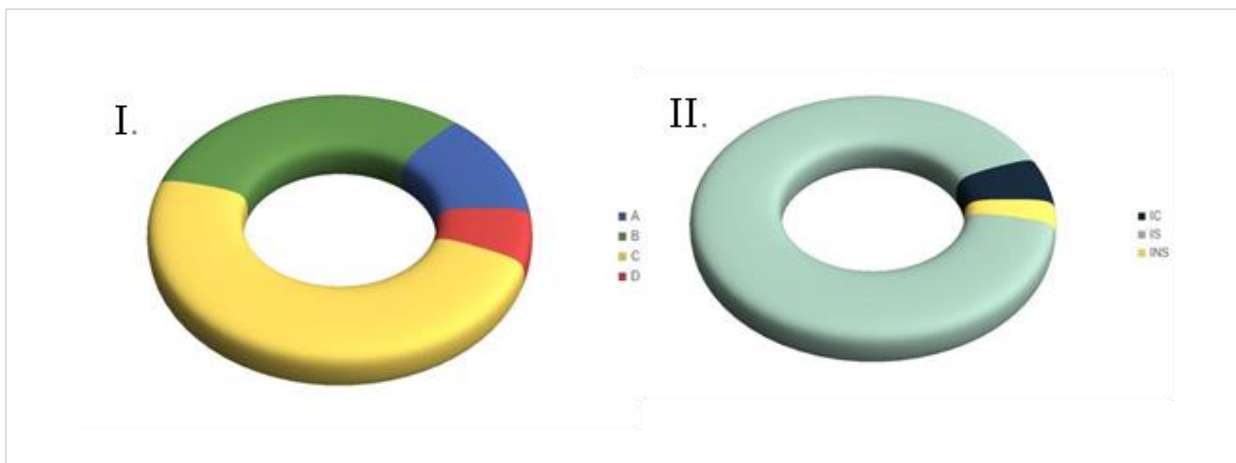


Figura 1.2. Ponderele principalelor domenii clinice vizate de aplicațiile de inteligența artificială (I) și a tipurilor de învățare automată (II) în ATI

Notă: A - monitorizare, detectare și diagnostic; B – identificare precoce a evenimentelor clinice; C – predicția rezultatului și evaluarea pronosticului; D – decizii de tratament. IS – învățare supravegheată, IC – învățare prin consolidare, INS – învățare nesupravegheată

Un sumar al aplicațiilor incluse în acest studiu este accesibil în Tabelul 1.1, iar o descriere mai detaliată a tehnicilor folosite este prezentă în următorul compartiment al tezei, unde se descriu metodele utilizate în studiul curent în comparație cu cele utilizate în alte studii.

Din cele 643 de lucrări extrase inițial din baza de date PubMed conform criteriilor de includere în acest studiu au fost selectate 97. Aceasta, dintr-o parte, pare să îngusteze spectrul de tehnici/direcții descrise în studiu, din altă parte, caracterizează nivelul de credibilitate al studiilor neincluse ca fiind unul sub anumite limite. Totuși, lucrarea include principalele aplicații AI utilizate în ATI până la momentul scrierii articolului (2022), aceasta fiind în concordantă și cu rezultatele altor sinteze sistematice ce vizează domeniul utilizării IA în ATI [3, 96, 97].

Tabelul 1.1. Direcții și aplicații clinice ale învățării automate / inteligenței artificiale în ATI

Direcția/Domeniul clinic de aplicare	Aplicații	Surse bibliografice
Monitorizare, detectare și diagnostic	<ul style="list-style-type: none"> - Monitorizarea indicatorilor fiziologici - Anomaliile ventilației mecanice (în special, asincronia pacient-ventilator) - Monitorizare electroencefalografică (la pacienți cu epilepsie, encefalopatie metabolică, hipoxie cerebrală, și ischemie și hematom subdural în leziunile cerebrale traumatice) - Detectarea infecției 	<ul style="list-style-type: none"> [10-13] [14-16] [17-20] [21-23]
Identificarea timpurie sau predicția evenimentelor clinice	<ul style="list-style-type: none"> - Predicția leziunilor renale acute - Predicția sepsisului și a infecției - Predicția bolilor respiratorii și a ventilației mecanice - Predicția și evaluarea altor condiții clinice (proгноza unor parametri fiziologici: debitul urinar, glucoza serică, acidul lactic, nivelul anemiei, etc.) 	<ul style="list-style-type: none"> [24-26] [1, 21, 22, 27-37, 97, 115, 120-124, 126, 129, 130] [38-45] [46-54]
Predicția rezultatului și evaluarea pronosticului	<ul style="list-style-type: none"> - Rezultatele pacienților cu boli specifice: <ul style="list-style-type: none"> ~ mortalitatea în sepsis ~ leziunea renală acută ~ leziunea hepatică ~ cancer ~ boli cardiace ~ patologii cerebrale ~ patologii pulmonară și altele - Prognosticul tratamentului și evaluarea timpului de ședere în unitatea ATI 	<ul style="list-style-type: none"> [55-59] [60-62] [63-65] [66, 67] [68, 69] [70, 71] [72-77] [78-89]
Decizii de tratament	<ul style="list-style-type: none"> - Definirea dozei optime de medicament (vasopresină, heparină, morfină) - Identificarea momentului optim pentru detubare - Alegerea tacticii optime de tratament al sepsisului 	<ul style="list-style-type: none"> [90-93] [94, 95] [92, 131]
Altele	<ul style="list-style-type: none"> ~ Monitorizarea profunzimii anesteziei ~ Ghidarea cu ultrasunet, ~ Managementul durerii, etc. 	<ul style="list-style-type: none"> [7]

În concluzie, vizând sinteza sistematică menționată și alte lucrări similare, se poate de menționat că identificarea timpurie, predicția rezultatelor clinice și evaluarea prognosticului a contribuit la aproximativ 80% din studii, iar aplicațiile de sprijinire a deciziilor clinice bazate pe învățarea automată în unitatea de terapie intensivă ar putea sprijini luarea în timp util a deciziilor [13], iar transformarea datelor în perspective mai acționabile sau reguli clinice bazate pe dovezi [88], ajută la diagnosticarea bolii [29], prezic rezultatele adverse înainte ca acestea să apară [63], permit evaluarea continuă a răspunsului pacientului la intervențiile de îngrijire critică [78], sporesc calitatea managementului în situații complexe și asigură mai bune decizii de tratament [90], reduc în cele din urmă povara clinicienilor [51] și oferă medicilor libertatea timpului, care este redirecționat spre asigurarea asistenței emoționale și îngrijirilor umane [98].

Merită să fie menționate două aspecte importante suplimentare la cele descrise. Aceste sunt explicabilitatea și transferabilitatea sistemelor IA. Prima este crucială pentru adoptarea de către comunitatea medicală a acestor sisteme, iar a doua poate influența rezultatele finale la etapa de utilizare practică în unitatea medicală țintă.

În acest context, medicina este foarte interesată de inteligența artificială explicabilă (XAI). Explicabilitatea modelelor de inteligență artificială este, astfel, un subiect important care a primit recent o atenție considerabilă în literatura de specialitate (Holzinger, 2019; Phillips et al., 2021). Din punct de vedere tehnic, problema explicabilității presupune gestionarea incertitudinii în lumea reală și explicarea cum un model IA a ajuns la anumite concluzii (rezultate). XAI lucrează cu implementarea transparenței și trasabilității metodelor de învățare automată statistică, în special a învățării profunde. Se remarcă faptul că inteligența artificială explicabilă, așa cum există în prezent, trebuie îmbunătățită prin suplimentarea ei cu aspecte de cauzalitate pentru a ajunge la medicina explicabilă.

Prin transferabilitatea sistemelor IA se subînțelege utilizarea unui sistem bazat pe date de la un spital la altul. În această situație, sistemul funcționează de obicei mai prost în circumstanțe noi. Astfel, este necesară o validare externă independentă pentru a stabili generalizabilitatea modelului la diverse cohorte înainte ca aceste modele de predicție a riscului să poată fi aplicate în mediul clinic (Cho et al. 2020). Cu toate acestea, atunci când sunt utilizate pe eșantioane externe, majoritatea modelelor de predicție a riscului propuse recent au performanțe slabe (Siontis, Tzoulaki, Castaldi, 2015; Chen et al., 2022)

Aceste aspecte sunt discutate în detaliu în ultimul capitol al tezei.

1.2. Conceptul de sepsis. Evoluția conceptului și conținutul lui modern

Bolile infecțioase au fost o cauză principală a mortalității de-a lungul istoriei civilizației umane. Sepsisul se referă la o complicație potențial fatală a unei infecții severe, de obicei bacteriene, și a fost definit recent ca fiind un răspuns dereglat al gazdei la infecție [99]. În recurs istoric, de exemplu, Moartea Neagră (ciuma bubonică) a distrus aproximativ o treime din populația Europei medievale și Asiei, cercetătorii antici fiind deja conștienți de magnitudinea problemei și având adesea temeri temeinice [100].

Cuvântul “sepsis” provine din greacă (σηψις) însemnând „descompunerea materiei organice animale sau vegetale în prezența bacteriilor” [101]. În context medical cuvântul “sepsis” a fost întâlnit pentru prima dată în poeziile lui Homer cu aproximativ 2700 de ani în urmă, unde apare ca derivat din verbul “sepo” (σηπω) care se traduce prin „putrezesc” [101]. Chiar dacă sepsisul a păstrat o prezență de-a lungul secolelor, de exemplu prin apariția în scrierile lui Hipocrate și Galen, numai odată cu apariția “teoriei germenilor” din anii 1800, originile sale microbiene au fost clar conturate [100].

În timp ce reprezentarea antică a sepsisului era caracterizată în mare parte prin “putrezirea cărnii”, răni și febră, apariția microbiologiei și imunologiei moderne a permis cercetătorilor să ajungă la o imagine mult mai rafinată a sepsisului. În primul rând, aceasta a implicat clarificarea că infecțiile bacteriene contagioase erau fundamentul unei complicații septice. Apoi, pe parcursul secolului al XX-lea, au fost obținute perspective și mai profunde asupra proceselor moleculare ale sepsisului, dezvăluind rolurile citokinelor și ale sistemului de coagulare [100].

La începutul anilor 1990, prima definiție internațională de consens a sepsisului, “Sepsis-1”, a descris sepsisul ca o inflamație sistemică indusă de infecție [102]. Ulterior, datorită faptului că inflamația sistemică nu este suficient de specifică pentru sepsis, grație prezenței ei și la pacienții fără infecție care au un prognostic mai bun, a fost propus conceptul “Sepsis-2” pentru a extinde prima definiție, luând în considerare în continuare parametrii inflamatori, hemodinamici și de disfuncție de organ [103]. Conceptul recent propus a urmărit creșterea specificității prin includerea indicatorilor leziunii de organ. Cu toate acestea, criteriile anterioare pentru definirea sepsisului au fost păstrate în uz, ceea ce a condus la confuzii, de exemplu, dacă să se diagnosticheze „sepsis” după noile criterii de diagnostic sau „sepsis sever” după cele vechi [104].

O reconsiderare a sepsisului a avut loc odată cu apariția conceptului “Sepsis-3”, propus în 2016 [99], unde sepsisul a fost definit ca un răspuns dereglat al gazdei la infecție, care pune viața în pericol. Acesta concept este operaționalizat prin evaluarea insuficienței secvențiale de organ (SOFA) [105], unde criteriile clinice de diagnostic ale sepsisului cuprind o creștere acută cu cel

puțin două puncte a scorului SOFA combinat cu o infecție confirmată sau suspectată [99]. Cercetarea din cadrul tezei utilizează cea mai recentă definiție, "Sepsis-3".

Ultima definiție are și limitările sale. În comparație cu definițiile anterioare, "Sepsis-3" prin designul său implică o disfuncție a organelor, făcând-o ca definiție mai restrânsă, care descrie o cohortă severă de pacienți. În continuare, ca o limitare relevantă pentru Sepsis-2 și Sepsis-3, este cerința pentru infecția suspectată (sau dovedită) ce presupune o dependență puternică a acestor definiții de intervențiile de diagnostic (de exemplu, prelevarea hemoculturii).

În sfârșit, complexitatea conceptului "Sepsis-3", care utilizează scorul SOFA, face ca această definiție să fie cea mai aplicabilă în unitatea de terapie intensivă. Cu toate acestea, definirea sepsisului în situații mai puțin exprimate clinic rămâne provocatoare.

Chiar dacă sepsisul a îmbrăcat multe fețe pe parcursul secolelor, o proprietate esențială a fost păstrată. De fapt, unul care era deja cunoscut cu aceasta este Niccolo Machiavelli, un scriitor politic din secolul al XVI-lea. Unul din citatele lui celebre afirmă: „După cum spun medicii despre febra agitată, că la începutul bolii este greu de detectat, dar ușor de tratat, dar cu timpul, având în vedere că nu a fost nici detectată, nici tratată la început, devine ușor de detectat, dar dificil de tratat" [106].

În ciuda progreselor inovatoare ale medicinei moderne, cum ar fi terapiile antimicrobiene și imunomodulatoare sau monitorizarea avansată, clinicienii de astăzi încă se confruntă cu aceeași situație. În stadiile incipiente ale sepsisului el este greu de identificat, leziunile organelor fiind reversibile, astfel încât o terapie antimicrobiană eficientă să rezulte în vindecare completă. Cu toate acestea, fiecare oră de intervenție întârziată duce la o creștere măsurabilă a mortalității [107].

În pofida deceniilor de cercetare, sepsisul rămâne o problema de sănătate publică importantă cu o mortalitate semnificativă, morbiditate și costuri înalte asociate tratamentului [108-110]. Există tot mai multe dovezi că managementul sepsisului necesită un diagnostic precoce urmat de o inițiere rapidă a unui tratament antimicrobian eficient [111].

Cu toate acestea, recunoașterea pacienților cu sepsis în stadiile incipiente, unde afectarea organelor este deocamdată reversibilă, este o sarcină notoriu de dificilă pentru clinicieni. În timp ce semnele timpurii pot fi deocamdată vagi și nespecifice, provocarea clinicianului fiind determinată de varietatea mare a diagnosticelor diferențiale și ca rezultat de prea multe ori există o întârziere în inițierea unei terapii de resuscitare și antimicrobiene potențial salvatoare [112].

1.3. Situația actuală în abordarea problemei sepsisului cu utilizarea tehnologiilor intelectuale în managementul clinic al stării de sepsis

În ultimele decenii, clinicienii și cercetătorii au căutat markeri biologici care ar permite o recunoaștere precoce a sepsisului, deși succesele fiind modeste [113, 114]. Acum, în epoca revoluției digitale în curs de desfășurare în asistența medicală, condusă de colectarea de rutină a datelor în sistemele de gestionare a datelor pacienților și bazele de date electronice ale dosarelor de sănătate, există o nouă speranță în abordarea problemei și recunoașterea timpurie prin descoperirea markerilor digitali de sepsis. În acest scop s-a propus extragerea multitudinii datelor pacientului folosind învățarea automată pentru a valorifica semnalele timpurii, care sunt predictive pentru o iminentă stare de sepsis.

Studiile timpurii care au folosit ÎA la pacienții cu sepsis s-au concentrat pe prezicerea rezultatelor clinice legate de sepsis, cum ar fi mortalitatea [115, 116]. Chiar dacă un punct final ca mortalitatea este clar definit și pare a fi o alegere naturală pentru o țintă de predicție clinică, ne putem întreba în ce măsură ar fi de fapt predicția precisă a mortalității la pacienții cu sepsis acționabilă clinic prin faptul că sprijină clinicienii în luarea deciziilor. În special, din moment ce managementul sepsisului este o situație extrem de sensibilă de timp, în care fiecare oră de întârziere a antibioticoterapiei și resuscitării cu fluide poate duce la leziuni ale organelor potențial ireversibile, există un argument practic puternic în favoarea prezicerii dacă și când va apărea exact sepsisul, decât să se prevadă rezultatul final al acestuia. Interesul din ultimii ani fata de subiect este ilustrat și de dinamica numărului de publicații în domeniu (Fig.1.3.)

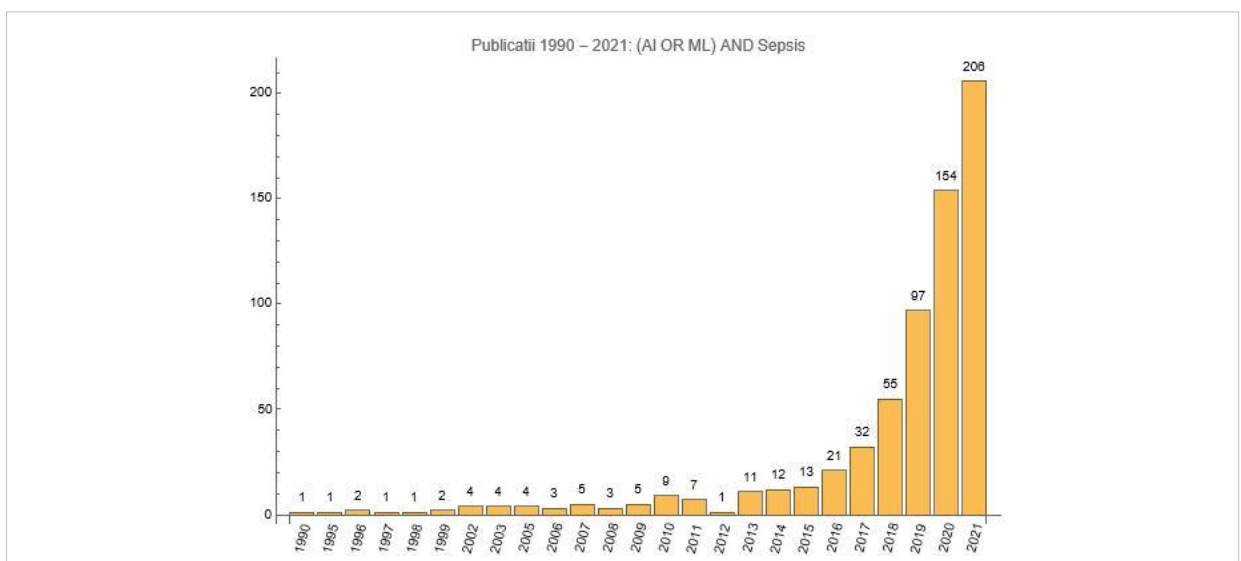


Figura 1.3. Numărul publicațiilor (PubMed) în domeniul inteligenței artificiale/învățării automate dedicate sepsisului în perioadă 1990 – 2021

Propulsat de o disponibilitate nevăzută anterior a bazelor de date mari, cum ar fi MIMIC [117, 118], au fost efectuate studii inițiale care au considerat predicția sepsisului ca o problemă de

avertizare timpurie. Henry și colab. [119] a dezvoltat un scor de avertizare în timp real (TREWScore) pentru predicția șocului septic cu un model de hazard proporțional Cox. Apoi, a fost dezvoltat un model de risc numit „In Sight” pentru a prezice apariția viitoare a sepsisului [34]. Majoritatea acestor studii se bazează pe tehnicile convenționale de modelare statistică. De exemplu, Calvert și colab. [120] utilizează un set de nouă parametri vitali.

În anii următori atenție considerabilă capătă învățarea automată profundă [121, 122]. Astfel, având în vedere dispozițiile de mai sus și nevoia clinică urgentă de avertismente timpurii și precise, predicția sepsisului a atras atenția comunității de învățare automată.

În Ghidul internațional pentru managementul sepsisului și al șocului septic din 2021 [123] în premieră este menționat potențialul rol al învățării automate, care *“poate îmbunătăți performanța instrumentelor de screening [al sepsisului], iar într-o meta-analiză pe un lot de 42623 pacienți din șapte studii dedicate prezicerii sepsisului spitalicesc, aria de sub curba ROC este de 0,89 (95% ÎÎ, 0,86-0,92); sensibilitatea - 81% (95% ÎÎ, 80-81) și specificitatea - 72% (95% ÎÎ, 72-72), fiind mai mari pentru învățarea automată decât aria de sub curba ROC pentru instrumentele tradiționale de screening, cum ar fi SIRS (0,70), MEWS (0,50), și SOFA (0,78)”* [97].

În plan de implementare, învățarea automată este o disciplină a domeniului IA axată pe a face mașinile capabile să îndeplinească sarcini fără a fi programate în mod explicit pentru ele. Pentru a face acest lucru, trebuie antrenați algoritmi, ceea ce, în funcție de algoritm, poate fi realizat prin analiza eșantionului, sau antrenarea bazată pe date, sau prin dezvoltarea iterativă a unei strategii de rezolvare a problemelor, bazată pe recompense sau pedepse. Dintre diferitele tipuri de strategii de învățare, una dintre cele mai utilizate este învățarea supravegheată. Aici, obiectivul este de a crea un model capabil să prezică o valoare de ieșire având în vedere un set de variabile de intrare.

De exemplu, pentru a prezice starea malignă (vs. benignă) a unui țesut dat având în vedere unele variabile morfologice, datele pacientului sau chiar imaginea specimenului. Dacă variabila de ieșire este limitată la un set cunoscut de valori, sarcina se numește clasificare, în timp ce dacă predicția ar trebui să fie orice valoare numerică dintr-un interval, obiectivul este de a face o regresie [124]. În învățarea supervizată, algoritmi au nevoie de un set de eșantioane rezolvate, adică date care includ valorile variabilelor de intrare împreună cu rezultatul corect. Odată antrenați, acești algoritmi pot prezice rezultatul (eticheta) unor seturi noi de variabile de intrare.

În continuare, sunt explicați pe scurt principalii algoritmi de ÎA supravegheată utilizați în studiile descrise în literatură [7, 15, 22, 28, 31-36, 55, 97, 115, 116, 120, 121, 124, 129-132, 134,

135, 137, 139]. Elementele de formalism matematic utilizate în continuare sunt bazate pe [125] și [126], dacă nu este specificat altfel.

Urmând afirmația că învățarea supravegheată este cel mai bine înțeleasă [127] și un element crucial în acest caz este învățarea dintr-un set de date de antrenament care include un set limitat de vectori (fiecare descriind o pereche de intrare-ieșire: ex. starea pacientului – diagnostic/pronostic, unde fiecare intrare corespunde unei etichete). Sarcina de învățare constă în identificarea funcției care descrie cu un grad suficient de acuratețe relația dintre intrare și ieșire, apoi această funcție învățată este utilizată pentru a prezice rezultatul în cazul unor date noi care nu au participat la învățare.

Dacă notăm cu X spațiul vectorial care include toate intrările posibile și cu Y toate ieșirile/etichetele posibile putem presupune că există o distribuție de probabilitate necunoscută pe spațiul produs $Z = X \times Y$, astfel încât $p(z) = p(\vec{x}, y)$. În acest caz, setul de date de antrenament, care constă din n eșantioane din această distribuție, poate fi scris după cum urmează:

$$S = \{(\vec{x}_1, y_1), \dots, (\vec{x}_n, y_n)\} = \{\vec{z}_1, \dots, \vec{z}_n\}, \quad (1.1)$$

unde \vec{x}_i este un vector din setul de training și y_i este eticheta lui.

În acest caz, problema constă în identificarea unei funcții $f : X \rightarrow Y$, astfel încât $f(\vec{x}) \sim y$. Fie \mathcal{H} spațiul funcțiilor $f : X \rightarrow Y$, numit spațiu de ipoteze, pe care algoritmul va efectua căutarea, folosind funcția de pierdere $L(f(\vec{x}), y)$, care reprezintă diferența dintre valoarea prezisă $f(\vec{x})$ și valoarea adevărată y . Pentru a aborda acest lucru mai formal, implicând noțiunea de probabilitate, presupunem că există o distribuție comună de probabilitate $P(x, y)$ peste X și Y , iar setul de date de antrenament este o colecție obținută prin eșantionare aleatorie din setul părinte. Ipoteza unei distribuții comune de probabilitate ne permite să modelăm incertitudinea în predicții (de exemplu, din zgomotul din date) deoarece y nu este o funcție deterministă a lui x , ci mai degrabă o variabilă aleatoare cu distribuție condiționată $P(y|x)$ pentru un x fix.

Atunci riscul (R) asociat cu funcția de pierdere poate fi definit ca așteptarea matematică (E) a unei ipoteze $h(x)$, din spațiul ipotezelor \mathcal{H} [128]:

$$R(h) = E[L(h(x), y)] = \int L(h(x), y) dP(x, y). \quad (1.2)$$

În acest context, scopul final al unui algoritm de învățare este găsirea unei ipoteze h^* în clasa fixă de funcții \mathcal{H} , pentru care riscul $R(h)$ este minim:

$$h^* = \underset{h \in \mathcal{H}}{\operatorname{argmin}} R(h). \quad (1.3)$$

Tipurile de funcții de pierdere utilizate pot fi diverse, în funcție de mai mulți factori (de exemplu, tipul datelor (numerice versus categorice), tipul problemei (predicție vs clasificare), etc.). Multe probleme de diagnostic/prognostic medical (acesta este și cazul cercetării curente) reprezintă probleme de clasificare dihotomică (prezența vs absența bolii). În acest caz, în termeni generali, funcția de pierdere este de tipul:

$$L(\hat{y}, y) = \begin{cases} 0 & \text{dacă } \hat{y} = y \\ 1 & \text{dacă } \hat{y} \neq y \end{cases} \quad (1.4)$$

În descrierea ce urmează, sunt utilizate următoarele notații: o variabilă de intrare este de obicei desemnată prin simbolul X . Dacă X este un vector, componentele sale pot fi accesate de indicele X_j . Ieșirile cantitative sunt notate cu Y și ieșirile calitative cu G (pentru grup). Literele mari sunt folosite atunci când se referă la aspectele generice ale unei variabile. Valorile observate sunt scrise cu litere mici (de exemplu, a i -a valoare observată a lui X este scrisă ca x_i). Matricele sunt reprezentate cu majuscule aldine; de exemplu, un set de N p -vectori de intrare x_i , $i = 1, \dots, N$ este reprezentat de matricea X de dimensiuni $N \times p$. În general, vectorii nu vor fi aldine, decât atunci când au N componente; un vector p de intrări x_i pentru a i -a observație diferă de un N -vector x_i constând din toate observațiile asupra variabilei X_i . Deoarece toți vectorii sunt presupuși a fi vectori coloană, al i -lea rând al lui X este x_i^T , adică transpunerea vectorială a lui x_i . Pentru un G cu două clase, ținta este codificată binar ca Y și apoi tratată ca o ieșire cantitativă. Predicțiile \hat{Y} vor fi de obicei în intervalul $[0, 1]$ și putem atribui eticheta de clasă \hat{G} în funcție de pragul $\hat{y} > 0.5$. Această abordare se generalizează și la rezultatele calitative de nivel K (de exemplu, în cazul variabilelor categoricale).

Regresia liniară

Modelele liniare au fost un pilon al statisticilor timp de câteva decenii și rămân unul dintre cele mai importante instrumente care servesc drept bază sau tangentă la alte metode [125, 126]. În general, un model liniar modelează dependența dintre un răspuns (ieșire) y și un vector de intrare x ca o funcție liniară:

$$y = x^T \beta + \beta_0 + \epsilon, \quad (1.5)$$

unde, β este vectorul parametrului, β_0 reprezintă termenul de interceptare (ordonata la bază) și $\epsilon \sim \mathcal{N}(0, \sigma^2)$ este o variabilă aleatorie gaussiană care este zgomotul din model.

Având un vector de intrări $X^T = (X_1, X_1, \dots, X_p)$ putem prezice ieșirea Y folosind modelul:

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^p X_j \hat{\beta}_j. \quad (1.6)$$

Potrivirea modelului la un set de date de antrenament se poate face prin mai multe metode, însă cea mai populară este metoda celor mai mici pătrate. În această abordare, alegem coeficienții β pentru a minimiza suma reziduală a pătratelor (RSS):

$$RSS(\beta) = \sum_{i=1}^N (y_i - x_i^T \beta)^2, \quad (1.7)$$

unde $RSS(\beta)$ este o funcție pătratică a parametrilor și, prin urmare, minimumul său există întotdeauna, dar poate să nu fie unic. Soluția este cel mai ușor caracterizată în notație matriceală:

$$RSS(\beta) = (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta), \quad (1.8)$$

unde \mathbf{X} este o matrice $N \times P$ cu fiecare rând un vector de intrare și \mathbf{y} este un N -vector de ieșiri din setul de antrenament.

Dacă extrapolăm la un context de diagnostic dihotomic, în care cazurile care dau un răspuns pozitiv Y (prezența bolii) sunt codificate ca 1, iar cele cu răspuns negativ (absența bolii) – ca 0, atunci valorile ajustate \hat{Y} sunt convertite într-o variabilă de clasă \hat{G} ajustată conform regulii:

$$\hat{G} = \begin{cases} \text{maladia PREZENTĂ} & \text{dacă } \hat{Y} > 0.5, \\ \text{maladia ABSENTĂ} & \text{dacă } \hat{Y} \leq 0.5 \end{cases} \quad (1.9)$$

Metode bazate pe arbori decizionali

Metodele bazate pe arbori decizionali (MAD) împart spațiul caracteristic într-un set de dreptunghiuri și apoi potrivesc un model simplu (cum ar fi o constantă) în fiecare dintre ele. Să considerăm o problemă de regresie cu răspuns continuu Y și intrări X_1 și X_2 , fiecare luând valori în intervalul unitar. Panoul (a) din Fig. 1.4 prezintă o partiție a spațiului caracteristic prin linii care sunt paralele cu axele de coordonate. În fiecare element de partiție putem modela Y cu o constantă diferită. Mai întâi împărțim spațiul în două regiuni și modelăm răspunsul prin media lui Y în fiecare regiune. Apoi, una sau ambele dintre aceste regiuni sunt împărțite în alte două regiuni și acest proces este continuat, până când se aplică o regulă de oprire. Rezultatul acestui proces este o partiție în M regiuni. Pentru fiecare regiune se aplică procedura de regresie. Modelul de regresie corespunzător prezice Y cu o constantă c_m în regiunea R_m :

$$\hat{f}(X) = \sum_{m=1}^M c_m I\{(X_1, X_2) \in R_m\}, \quad (1.10)$$

unde X_1, X_2 sunt variabile din setul de caracteristici, iar I este riscul (de predicție greșită sau de clasificare greșită). Dacă folosim ca criteriu minimizarea sumei pătratelor $\sum (y_i - f(x_i))^2$ putem observa că cel mai reușit c_m este media lui y_i în regiunea R_m :

$$\hat{c}_m = \text{ave}(y_i | x_i \in R_m). \quad (1.11)$$

Un aspect important este dimensiunea arborelui, care este un parametru de reglare care guvernează complexitatea modelului, iar dimensiunea optimă a arborelui este aleasă adaptiv din date. O abordare ar fi împărțirea nodurilor arborelui numai dacă scăderea sumei pătratelor din cauza împărțirii depășește un anumit prag. Strategia preferată este să crească un arbore mare T_0 , oprind procesul de împărțire numai atunci când este atinsă o dimensiune minimă a nodului. Apoi, acest arbore mare este „tăiat” folosind abordarea cost-complexitate, care constă în definirea unui subarbore $T \subset T_0$ ca orice arbore care poate fi obținut prin tăierea T_0 , adică prin eliminarea oricărui număr de noduri interne (non-terminale) ale acestuia. Criteriul pentru „tăiere” este compromisul dintre dimensiunea arborelui și performanța modelului obținut din „tăiere”.

Cele descrise se referă în primul rând la o problemă de regresie. Pentru sarcinile de clasificare acest lucru este de asemenea valabil, cu unele diferențe în împărțirea nodurilor și strategia de „tăiere”. Să considerăm ca țintă un rezultat de clasificare care ia valorile $1, 2, \dots, K$. Într-un nod m , reprezentând o regiune R_m cu N_m observații, fie

$$\hat{p}_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k) \quad (1.12)$$

proporția de observații de clasa k în nodul m .

Atribuim observațiile de la nodul m clasei $k(m) = \text{argmax}_k \hat{p}_{mk}$, clasa majoritară din nodul m . Drept criterii de optimizare pot servi eroarea de clasificare greșită, indicele Gini, cross-entropia sau devianța, etc.

Astfel, un algoritm cu arbore de decizie (AD) are structura internă a unui arbore grafic, în care fiecare nod de ramură evaluează o variabilă a eșantionului, fiecare muchie reprezintă un rezultat posibil al evaluării iar fiecare frunză reprezintă un rezultat final posibil. Fig. 1.4 (b) reprezintă un AD simplu care determină dacă un pacient este septic urmând definiția Sepsis-3. Structura arborelui, inclusiv numărul de noduri, starea evaluată pe fiecare nod și valoarea fiecărei muchii, este determinată în timpul procesului de antrenament. Un tip special de AD este Random

Forest (RF), care nu este strict un arbore, ci un set de AD care utilizează diferite variabile ale eșantioanelor. Rezultatul final al unui RF este o medie a rezultatelor arborilor interni, de obicei modul, pentru clasificare și media, pentru regresie. Clasificatoarele RF au în general performanțe ridicate, sunt ușor de configurat și este posibil de obținut informații utile pentru interpretarea clasificatorilor antrenați, cum ar fi importanța variabilelor.

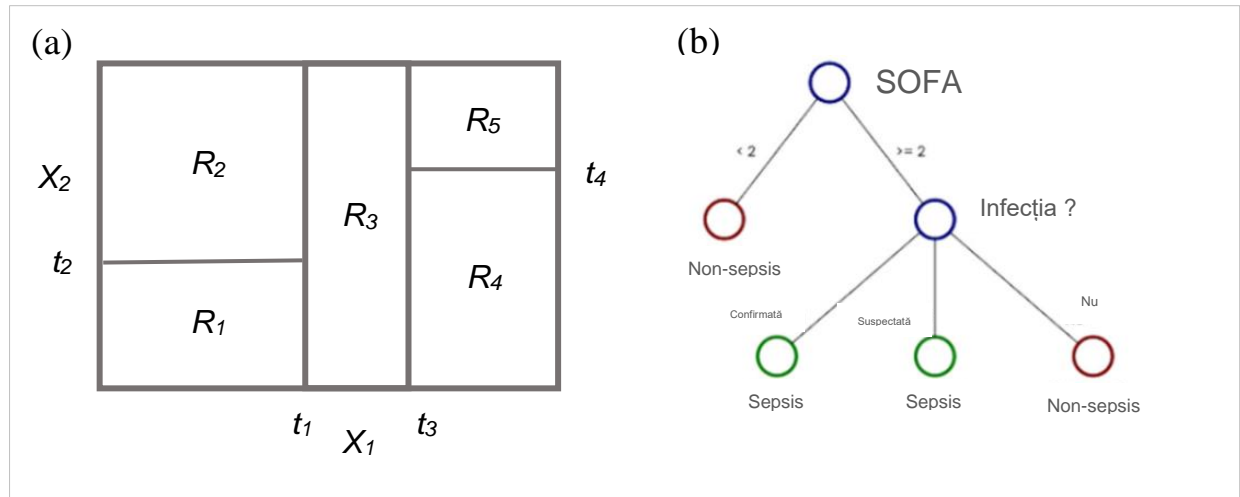


Figura 1.4. (a) Partiția spațiului de caracteristici cu ajutorul liniilor paralele axelor de coordonate. În fiecare partiție putem modela Y cu o constantă diferită (preluat din [125]); (b) Un arbore decizional care determină dacă un pacient este septic urmând definiția Sepsis-3 (preluat din [124] cu modificări)

Metode cu vectori de suport

Mașina cu vectori de suport (MVS) este o abordare a clasificării care a fost dezvoltată în comunitatea informatică în anii 1990 și de atunci a crescut în popularitate. Această abordare include o colecție de metode (clasificator de marjă maximă, clasificator cu vectori de suport și însăși mașina cu vectori de suport). Conform [125] acești clasificatori, cu toate că reprezintă entități distincte, de obicei extensii ale versiunilor mai timpurii, sunt adesea denumite „mașini cu vectori de suport”.

Un concept central în acest domeniu este noțiunea de hiperplan. Într-un spațiu p –dimensional, un hiperplan este un subspațiu aflat în plan de dimensiunea planului $p - 1$. De exemplu, în două dimensiuni, un hiperplan este un subspațiu plat unidimensional - cu alte cuvinte, o linie. În trei dimensiuni, un hiperplan este un subspațiu plat bidimensional, adică un plan. În $p > 3$ dimensiuni, poate fi greu de vizualizat un hiperplan, dar noțiunea de subspațiu plat ($p - 1$) – dimensional încă se aplică. Definiția matematică a unui hiperplan este destul de simplă. În p dimensiuni, un hiperplan este definit prin ecuația

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = 0 \quad (1.13)$$

în sensul că dacă un punct $X = (X_1, X_2, \dots, X_p)^T$ în spațiul p – dimensional (adică un vector de lungime p) satisface (1.13), atunci X se află pe hiperplan. În caz contrar, se află de o parte sau cealaltă a hiperplanului (Fig. 1.5).

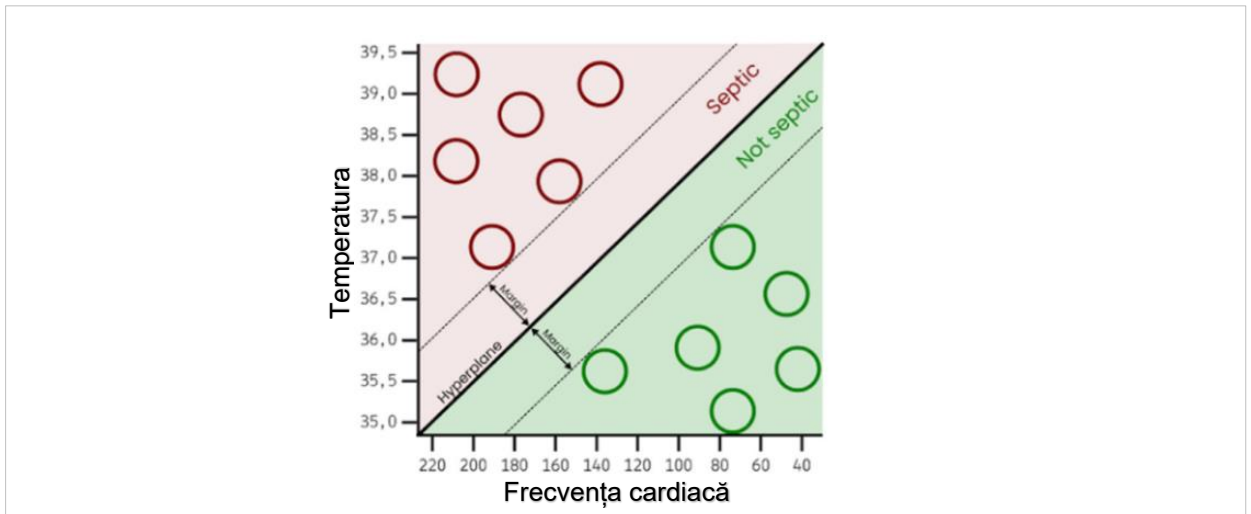


Figura 1.5. Mașina cu vectori de suport cu utilizarea a două caracteristici: temperatura corpului și frecvența contracțiilor cardiace (preluat din [124] cu modificări)

Să presupunem că avem o matrice X de date $n \times p$ constând din n observații în spațiul p – dimensional,

$$x = \begin{pmatrix} X_{11} \\ \vdots \\ X_{1p} \end{pmatrix}, \dots, x_n = \begin{pmatrix} X_{n1} \\ \vdots \\ X_{np} \end{pmatrix}, \quad (1.14)$$

și că aceste observații aparțin a două clase, adică $y_1, \dots, y_n \in \{-1, 1\}$, unde -1 reprezintă o clasă și 1 cealaltă clasă. Având în vedere o observație-test, un vector p de caracteristici observate $x^* = (x_1^* \dots x_p^*)^T$, scopul nostru este să dezvoltăm un clasificator bazat pe datele de antrenament care va clasifica corect observația-test folosind valorile caracteristicilor sale. Putem eticheta o observație dintr-o clasă ca $y_i = 1$ și pe cea din clasa opusă ca $y_i = -1$. Atunci un hiperplan de separare are proprietatea că

$$\begin{aligned} \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} &> 0 \text{ dacă } y_i = 1, \text{ și} \\ \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} &< 0 \text{ dacă } y_i = -1. \end{aligned} \quad (1.15)$$

Adică clasificăm observația-test x^* pe baza semnului lui $f(x^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_p x_p^*$. Dacă $f(x^*)$ este pozitiv, atunci atribuim observația-test la clasa 1 , iar dacă $f(x^*)$ este negativ, atunci o atribuim clasei -1 . Putem folosi și mărimea lui $f(x^*)$. Dacă $f(x^*)$ este departe de zero, atunci aceasta înseamnă că x^* se află departe de hiperplan și astfel putem fi încrezători în

alocarea clasei pentru x^* . Pe de altă parte, dacă x^* este aproape de zero, atunci x^* este situat lângă hiperplan și, prin urmare, suntem mai puțin siguri de alocarea clasei pentru x^* .

Pe baza celor descrise, folosind produsul interior a doi vectori (ca funcție scalară a acestor vectori, egal cu produsul mărimii lor și cosinusul unghiului dintre ei) putem defini un clasificator cu vectori de suport:

$$f(x) = \beta_0 + \sum_{i=1}^n \alpha_i \langle x, x_i \rangle, \quad (1.16)$$

unde există n parametri $\alpha_i, i = 1, \dots, n$, câte unul pentru fiecare observație de antrenament.

Pentru a estima parametrii $\alpha_1, \dots, \alpha_n$ și β_0 , avem nevoie doar de $\binom{n}{2}$ produse interne $\langle x_i, x_{i'} \rangle$ între toate perechile de observații de antrenament. Notăția $\binom{n}{2}$ înseamnă $n(n-1)/2$ și dă numărul de perechi dintr-o mulțime de n elemente.

Acum să presupunem că produsul interior este înlocuit cu o generalizare a produsului interior de forma $K(x_i, x_{i'})$, unde K este o funcție la care ne vom referi ca un nucleu. Un nucleu este o funcție de nucleu care cuantifică asemănarea a două observații. De exemplu, am putea pur și simplu să luăm

$$K(x_i, x_{i'}) = \sum_{j=1}^p x_{ij} x_{i'j}, \quad (1.17)$$

care este un clasificator liniar al vectorului de suport al nucleului. Cu toate acestea, nucleele pot fi diferite (polinomial, radial, etc.).

Metodele cu vectori de suport (MVS) tratează fiecare probă ca pe un punct într-un spațiu n -dimensional și încearcă să găsească un hiperplan care separă eșantioanele din două condiții diferite. Din acest motiv, MVS sunt considerați clasificatori binari, deoarece pot clasifica datele doar cu două ieșiri posibile. Cu toate acestea, există mai multe strategii pentru a utiliza MVS în probleme cu mai mult de două ieșiri posibile prin combinarea mai multor clasificatoare MVS. Fig. 1.5 reprezintă o MVS simplă cu două variabile de intrare (adică temperatura și frecvența contracțiilor cardiace). În timpul antrenamentului, parametrii hiperplanului sunt ajustați pentru a maximiza marja dintre hiperplan și eșantioanele cu ieșiri diferite. Clasificatoarele MVS au în general performanțe bune, cu toate acestea, sunt complexe de configurat și clasificatorii instruiți sunt dificil de interpretat.

Regresia logistică

Regresia logistică (RLog) își ia numele de la funcția logistică pe care se bazează:

$$g(z) = \frac{1}{1 + e^{-z}}. \quad (1.18)$$

La fel ca MVS, algoritmi RLog sunt considerați clasificatori binari, dar, în acest caz, rezultatul va fi apropiat sau egal cu 0 pentru o condiție și aproape sau egal cu 1 pentru cealaltă condiție. Fig. 1.6 reprezintă un algoritm RLog simplu cu o singură variabilă de intrare (adică frecvența cardiacă).

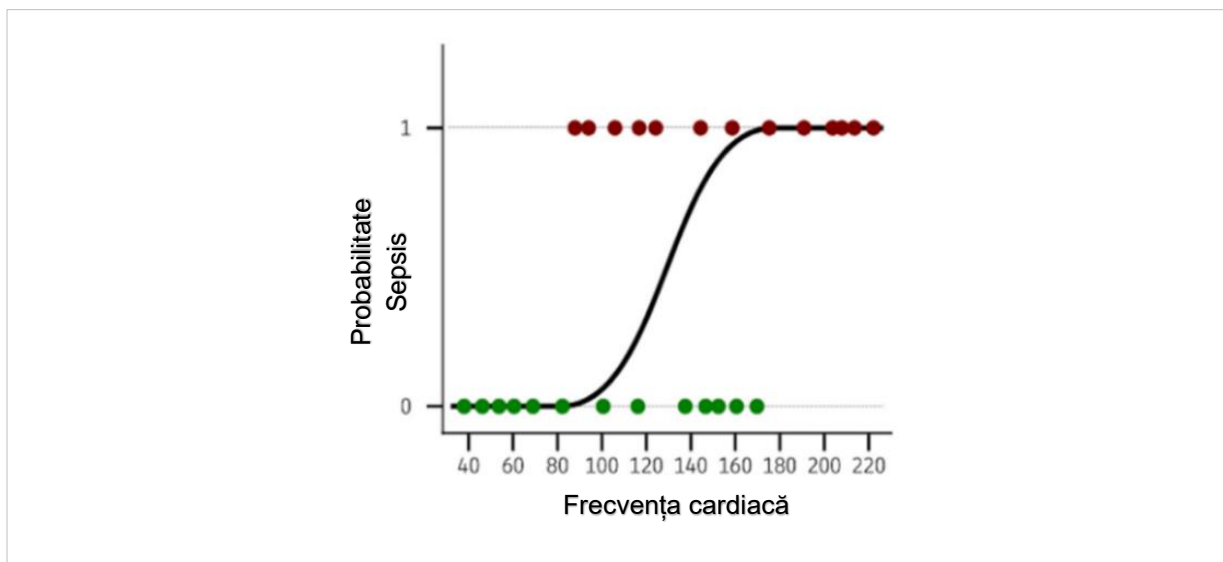


Figura 1.6. **Regresia logistică. Cu roșu sunt prezentate cazurile septice, iar cu verde – cele non-septice** (preluat din [124] cu modificări)

În acest exemplu, probabilitatea de sepsis crește odată cu frecvența cardiacă, totuși această variabilă nu este suficientă pentru a prezice sepsisul și unele mostre sunt clasificate greșit. Există o RLog generalizată pentru probleme cu condiții multiple, numită Regresie logistică multinomială. În timpul procesului de antrenament, coeficienții funcției utilizate pentru a genera ieșirea sunt ajustați, astfel încât valoarea de ieșire să fie cât mai aproape posibil de 0 sau 1 pentru fiecare condiție. În pofida numelui său, algoritmul RLog este folosit în probleme de clasificare. Clasificatoarele în baza RLog funcționează bine în problemele liniare separabile și sunt ușor de configurat și antrenat. Cu toate acestea, clasificatorii antrenați sunt greu de interpretat.

Rețele neuronale artificiale (RNA)

O rețea neuronală folosește un vector de intrare p cu variabile $X = (X_1, X_2, \dots, X_p)$ și construiește o funcție neliniară $f(X)$ pentru a prezice răspunsul Y , similar altor modele descrise, diferența fiind în structura modelului [125]. Figura 1.7 prezintă o rețea neuronală cu feed-forward simplă pentru modelarea unei rețele neuronale feed-forward cu răspuns cantitativ folosind $p = 6$ predictorii. În terminologia rețelelor neuronale, cele șase caracteristici X_1, \dots, X_6 alcătuiesc

unitățile stratului de intrare. Conexiunile indică faptul că fiecare dintre intrările din stratul de intrare este conectată la fiecare dintre cele K unități ascunse. Modelul rețelei neuronale are forma:

$$f(X) = \beta_0 + \sum_{k=1}^K \beta_k h_k(X) = \beta_0 + \sum_{k=1}^K \beta_k g \left(w_{k0} + \sum_{j=1}^p w_{kj} X_j \right). \quad (1.19)$$

Ea este construit în două etape. Primele K activări $A_k, k = 1, \dots, K$, în stratul ascuns sunt calculate ca funcții ale caracteristicilor de intrare X_1, \dots, X_p :

$$A_k = h_k(X) = g \left(w_{k0} + \sum_{j=1}^p w_{kj} X_j \right), \quad (1.20)$$

unde $g(z)$ este o funcție de activare neliniară care este specificată în prealabil. Putem considera fiecare funcție A_k o transformare diferită $h_k(X)$ a caracteristicilor originale. Aceste K activări din stratul ascuns sunt apoi trecute la stratul de ieșire, rezultând

$$f(X) = \beta_0 + \sum_{r=1}^K \beta_r A_r, \quad (1.21)$$

care este practic un model de regresie cu $K = 3$. Parametrii β_0, \dots, β_K și w_{10}, \dots, w_{Kp} sunt estimați din date. Funcția de activare poate fi diferită (sigmoidă, tangentă, etc.) dar cea mai frecvent utilizată în stadiul modern este *ReLU* (unitate liniară rectificată), care are forma

$$g(z) = (z)_+ = \begin{cases} 0 & \text{dacă } z < 0 \\ z & \text{altfel} \end{cases}. \quad (1.22)$$

Montarea unei rețele neuronale necesită estimarea parametrilor necunoscuți din (1.19). Pentru un răspuns cantitativ, pierderea de eroare pătrată este de obicei utilizată, astfel încât parametrii să fie aleși pentru a minimiza

$$\sum_{i=1}^n (y_i - f(x_i))^2. \quad (1.23)$$

Actualmente sunt răspândite rețelele neuronale cu mai multe niveluri. În acest caz, se generează un număr de modele liniare egal cu numărul de neuroni ascunși.

Stratul de ieșire poate folosi diferite funcții pentru a furniza predicția finală, cea mai frecvent utilizată fiind funcția *softmax*. În acest caz, în loc să selectăm o clasă de bază, tratăm toate K clasele simetric și presupunem că pentru $k = 1, \dots, K$,

$$Pr(Y = k | X = x) = \frac{e^{\beta_{k0} + \beta_{k1}x_1 + \dots + \beta_{kp}x_p}}{\sum_{l=1}^K e^{\beta_{l0} + \beta_{l1}x_1 + \dots + \beta_{lp}x_p}}. \quad (1.24)$$

Astfel, coeficienții pentru toate clasele K sunt estimați și putem rescrie sub forma de raport log-odds:

$$\log\left(\frac{\Pr(Y = k|X = x)}{\Pr(Y = k'|X = x)}\right) = (\beta_{k0} - \beta_{k'0}) + (\beta_{k1} - \beta_{k'1})x_1 + \dots + (\beta_{kp} - \beta_{k'p})x_p \quad (1.25)$$

Rețelele neuronale artificiale (RNA) sunt inspirate de modul în care funcționează neuronii dintr-un creier uman sau biologic. Un RNA este un grup interconectat de noduri (neuroni artificiali) separate în cel puțin trei straturi: (a) intrare, care primește datele eșantionului, (b) ascuns, care transformă valorile de intrare și (c) ieșire, care oferă predicția finală pentru fiecare probă. Cu toate acestea, este foarte obișnuit să existe mai mult de un singur strat ascuns, deoarece mai multe straturi cresc capacitatea RNA de a învăța probleme mai complexe. Fig. 1.7 reprezintă o RNA simplă cu șase neuroni de intrare, fiecare corespunzând cu o variabilă, un strat ascuns cu trei neuroni și un strat de ieșire cu un singur neuron care emite probabilitatea de sepsis.

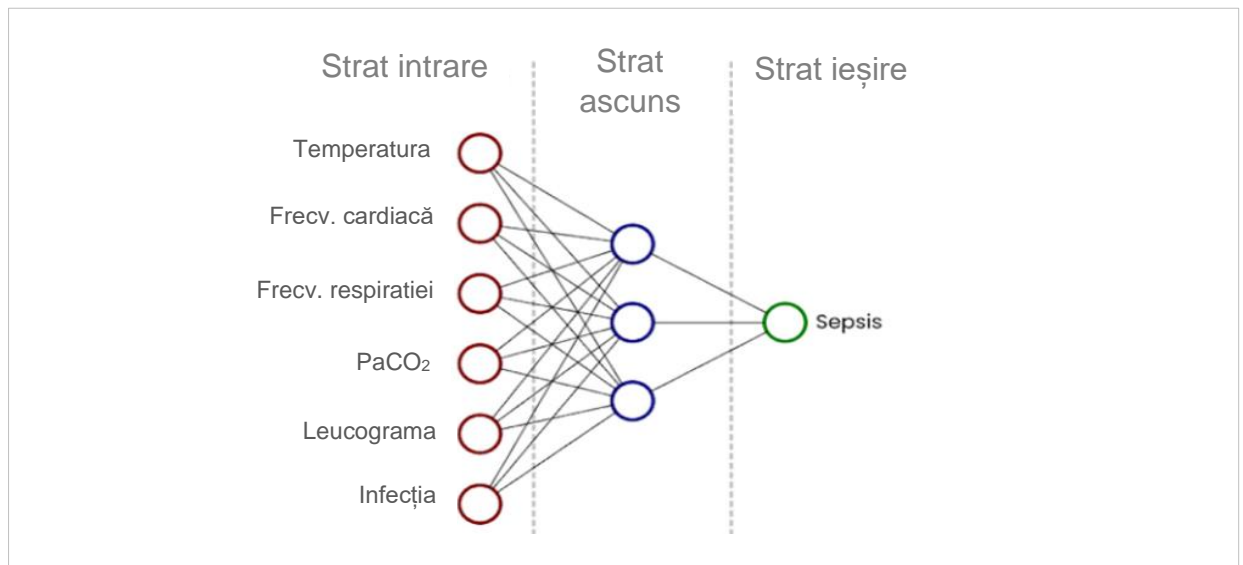


Figura 1.7. **O rețea neuronală artificială cu șase predictor și un strat ascuns.**
(preluat din [124] cu modificări)

După cum se poate observa, nodurile de pe fiecare strat sunt conectate cu toate nodurile de pe stratul următor. Aceste conexiuni sunt de obicei inițializate cu o pondere aleatorie. Stratul de ieșire este de obicei configurat să aibă un nod pentru fiecare condiție posibilă de ieșire, iar nivelul de activare al fiecăruia dintre aceste noduri pentru o probă corespunde cu probabilitatea ca eșantionul să aparțină condiției asociate cu nodul. În timpul antrenamentului, ponderile conexiunii sunt ajustate, astfel încât valorile de intrare să poată fi transformate în rezultatul așteptat prin aplicarea succesivă a transformărilor matematice bazate pe ponderea conexiunii. Clasificatoarele

RNA au în general performanțe bune, cu toate acestea, sunt complexe de configurat, costisitoare de antrenat și greu de interpretat.

Rețele neuronale recurente (RNR)

Acest tip de rețele este potrivit pentru date secvențiale, iar pentru construirea de modele predictive necesită un tratament special [125]:

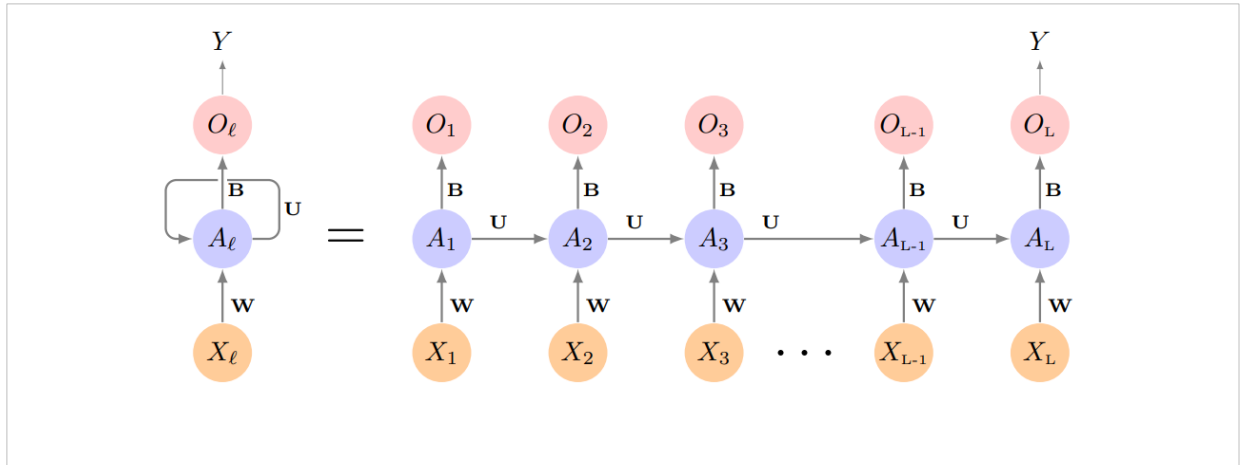


Figura 1.8. Schema unei rețele recurente (reprodus din [211])

Figura 1.8 ilustrează structura unei RNR de bază cu o secvență $X = \{X_1, X_2, \dots, X_L\}$ ca intrare, o ieșire simplă Y și o secvență de straturi ascunse $\{A_\ell\}_1^L = \{A_1, A_2, \dots, A_L\}$. Fiecare X_ℓ este un vector; de exemplu, într-un document X_ℓ ar putea reprezenta o codificare one-hot pentru al ℓ -lea cuvânt bazat pe dicționarul de limbă pentru corpus. Pe măsură ce secvența este procesată câte un vector X_ℓ la un moment dat, rețeaua actualizează activările X_ℓ din stratul ascuns, luând ca intrare vectorul X_ℓ și vectorul de activare $X_{\ell-1}$ din pasul anterior din secvență. Fiecare X_ℓ se alimentează în stratul de ieșire și produce o predicție O_ℓ pentru Y . O_ℓ , ultimul dintre acestea, este cel mai relevant.

Într-o formă mai detaliată, să presupunem că fiecare vector X_ℓ al secvenței de intrare are p componente $X_\ell^T = \{X_{\ell 1}, X_{\ell 2}, \dots, X_{\ell p}\}$, iar stratul ascuns este format din K unități $X_\ell^T = \{X_{\ell 1}, X_{\ell 2}, \dots, X_{\ell K}\}$. Dacă reprezentăm colecția de $K \times (p + 1)$ de ponderi comune w_{kj} pentru stratul de intrare printr-o matrice \mathbf{W} și, în mod similar, \mathbf{U} este o matrice $K \times K$ a ponderilor u_{ks} pentru straturile ascunse, iar \mathbf{B} este un vector $K + 1$ al ponderilor β_k pentru stratul de ieșire. Atunci

$$A_{\ell k} = g \left(w_{k0} + \sum_{j=1}^p w_{kj} X_{\ell j} + \sum_{j=1}^p u_{ks} A_{\ell-1, s} \right) \quad (1.26)$$

iar ieșirea O_{ℓ} se calculează ca

$$O_{\ell} = \beta_0 + \sum_{k=1}^K \beta_k A_{\ell k} \quad (1.27)$$

pentru un răspuns cantitativ, sau cu o funcție suplimentară sigmoidală de activare pentru un răspuns binar. Sunt utilizate aceleași ponderi \mathbf{W} , \mathbf{U} și \mathbf{B} pe măsură ce fiecare element din secvență este procesat, adică nu sunt funcții ale lui ℓ . Pe măsură ce se trece de la început până la sfârșit, activările A_{ℓ} acumulează o istorie a ceea ce a fost obținut anterior, astfel încât contextul învățat să poată fi folosit pentru predicție.

Montarea rețelelor neuronale este considerată o sarcină complexă. În modelul descris de ecuația (1.19) parametrii sunt $\beta = (\beta_0, \beta_1, \dots, \beta_K)$, precum și fiecare dintre $w_k = (w_{k0}, w_{k1}, \dots, w_{kp})$, $K = 1, \dots, K$. Având în vedere observațiile (x_i, y_i) , $i = 1, \dots, n$, am putea potrivi modelul prin rezolvarea unei probleme neliniare cu cele mai mici pătrate:

$$\underset{\{w_r\}_1^K, \beta}{\text{minimize}} \frac{1}{2} \sum_{i=1}^n (y_i - f(x_i))^2, \quad (1.28)$$

unde

$$f(x_i) = \beta_0 + \sum_{k=1}^K \beta_k g \left(w_{k0} + \sum_{j=1}^p w_{kj} x_{ij} \right). \quad (1.29)$$

RNR suferă de problema dispariției (diminuării până la 0) gradientului. Versiunea Long Short-Term Memory (LSTM) depășește această problemă prin introducerea unor elemente suplimentare în arhitectura RNR (poarta de intrare, poarta de ieșire, poarta de uitare și memorycell).

Algoritmii bayesieni

Algoritmii bayesieni, cum ar fi Naïve Bayes (NB) sau Bayesian Network (BN), se bazează pe teorema lui Bayes, care descrie probabilitatea unui eveniment, pe baza cunoștințelor anterioare a condițiilor care ar putea fi legate de eveniment. Acest tip de algoritmi folosesc datele de antrenament pentru a estima probabilitățile fiecărei ieșiri posibile pe baza valorilor variabilelor eșantioanelor de antrenament. NB presupune independența variabilelor, care nu este cerută de BN.

Algoritmii bayesieni sunt utilizați în principal în probleme bazate pe text, deși pot fi utilizați în multe alte domenii. Sunt ușor de configurat, iar clasificatorii antrenați sunt interpretabili.

Prin prisma formalismului matematic, modelele bayesiene folosesc inferența bayesiană pentru ponderea ipotezelor diagnostice. Aceasta oferă o ponderare care încorporează estimări ale incertitudinii, care, la rândul lor, sunt determinanți importanți ai deciziilor de diagnostic. Ipotezele sunt ponderate folosind probabilitatea posterioară a ipotezei în baza dovezilor, calculată prin regula lui Bayes, și probabilitățile condiționate cunoscute. De exemplu, pentru constatarea f și ipoteza H , posteriorul se calculează după cum urmează:

$$P(H|f) = \frac{P(f|H)P(H)}{P(f)}, \quad (1.30)$$

unde probabilitatea condiționată $P(H|f)$ este probabilitatea de a observa o constatare f , având în vedere că ipoteza H este cunoscută a fi adevărată, iar $P(H)$ este probabilitatea anterioară ca ipoteza să fie adevărată (de exemplu, prevalența unei anumite boli din populație) și același lucru pentru $P(f)$. Aceste probabilități condiționate și probabilități anterioare sunt codificate direct în formula cauzală a modelului fiziopatologic și pot fi învățate din mai multe surse de date, inclusiv date epidemiologice și cunoștințe de specialitate. Cel mai simplu exemplu de rețea bayesiană, încă utilizată pe scară largă pentru diagnosticare, sunt rețelele Bayes naive. Acestea folosesc ipoteza că constatările sunt independente condiționat, având în vedere ipoteza, oferind o formă simplificată a modelului de probabilitate condiționată care poate fi calculată cu ușurință pentru constatări multiple,

$$P(f_1, \dots, f_N|H) = \prod_{i=1}^N P(f_i|H). \quad (1.31)$$

Această ipoteză a independenței poate fi rezumată ca un model grafic (Fig. 1.9) [129].

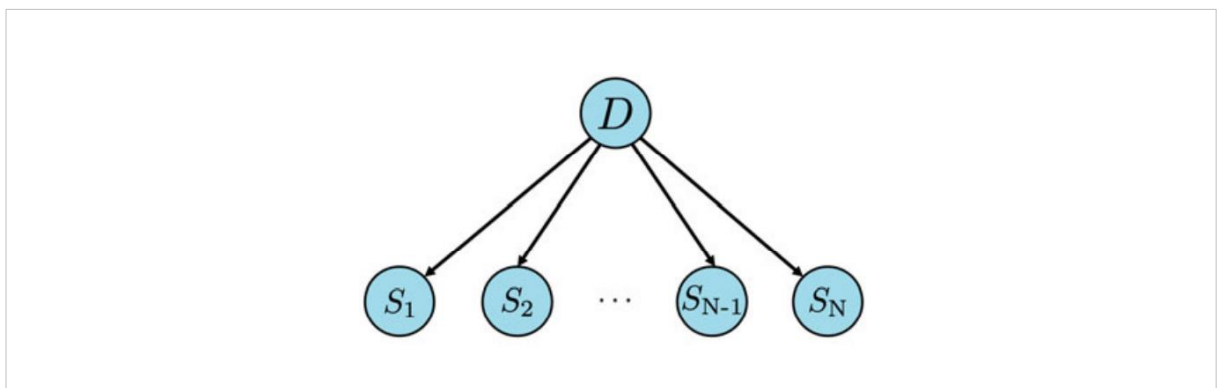


Figura 1.9. Un model naiv Bayes. Constatările sunt simptome S_1, S_2, \dots, S_N și modelul reprezintă o singură ipoteză (maladie - D). Preluat din [129]

Algoritmii bayesieni sunt ușor de configurat, iar clasificatorii antrenați sunt interpretabili.

Performanța prognostică variază în dependență de mai mulți factori, inclusiv algoritmul/modelul utilizat. Tabelul 1.2. alcătuit în baza [130-139] prezintă un sumar al performanței modelelor raportate în literatura de specialitate în pentru perioada 2018-2022, fiind incluse studiile care raportează performanța modelelor în baza ariei de sub curba ROC (AUROC).

Tabelul 1.2. Date sumare despre performanța predicției sepsisului cu modele IA (pentru ultimii 5 ani)

Studiul	Anul publicării	Algoritm(i) utilizat(i)	Performanța predicției (AUROC)	Orizontul predicției (ore)
Nemati et al.	2018	Weibull-Cox proportional hazards	0.79	12
			0.82	4
Mao et al.	2018	InSight Algorithm	0.92	-
Kaji et al.	2018	LSTM	0.88	“Ziua următoare”
Khojandi et al.	2018	RF	0.90	-
Faisal et al.	2018	Logistic Regression	0.78	4
Bloch et al.	2019	SVM-RBF	0.81	5
			0.87	1
Li et al.	2019	CNN+RNN	0.75	12
Scherpf et al.	2019	RNN	0.76	12
			0.81	3
Moor et al.	2019	MGP-TSN	0.91	-
Schamoni et al.	2019	Non-linear ordinal regression	0.84	4
Van Wyk F. et al.	2019	RF	0.79	-
Liu et al.	2019	XGBoost	0.92	-
Lauritsen et al.	2020	CNN-LSTM	0.75	24
			0.88	0.25
Li et al.	2020	LightGBM	0.85	6
Yang et al.	2020	XGBoost	0.85	1
Helguera-Repetto et al.	2020	Artificial Neural Network	0.94	-
Cooper et al.	2020	Logistic regression	0.86	-
Mohammed et al.	2020	Support Vector Machine	0.78	-
Bedoya et al.	2020	Multi-output Gaussian Process and RNN	0.88	-
Kok et al.	2020	Temporal Convolution Network	0.98	-
Lee et al.	2020	Graph Convolutional Network	0.78	-

Wickramaratne et al.	2020	Bi-Directional Gated Recurrent Units	0.97	-
Wardi et al.	2021	Modified Weibull-Cox proportional hazards mode	0.83	-
Misra et al.	2021	RF	0.95	-
Rafiei et al.	2021	ANN	0.92	4
Shashikumar et al.	2021	DeepANN	0.87	4
Pei-Chen Lin et al.	2021	XGBoost	0.86	-
Wang Z, Yao B.	2021	Multi-Branching Temporal Convolutional Network	0.89	-
Persson, I. et al.	2021	CNN	0.90	-
Kim et al.	2022	Genetic Algorithms	0.94	3
Camacho-Cogollo et al.	2022	SVM, ADA-Boost, ANN, RF, DT, Ensemble	0.91 (Ensemble) 0,92 (XGBoost)	-
Honoré et al.	2022	Logistic Regression (LR), Naive Bayes (NB), Multi-layer perceptron (MLP), Gaussian mixture models (GMM),	0.55(GMM) 0.87 (NB)	-

Notă: AUROC – aria de sub curba ROC; LSTM - long short-term memory; CNN – convolutional neural network; SVM-RBF - support vector machines with radial basis function; CNN-LSTM - convolutional neural network + long short-term memory; RNN - recurrent neural network; GB - Gradient boosting; GBT - Gradient boosting tree; RF – random forest; ANN - artificial neural network; GBM – Gradient boosting machine; DT – decision tree; Ensemble – ansamblu de modele.

În contextul descris, o analiză sistematică [130] arată că: (a) definiția sepsisului nu este consecventă între studii; (b) sursele de date și metodele de preprocesare a datelor, modelele de învățare automată, ingineria caracteristicilor și tipurile de includere variază foarte mult între studii; (c) cu cât este mai aproape de debutul sepsisului, cu atât valoarea AUROC este mai mare; (d) îmbunătățirea AUROC se datorează în primul rând utilizării învățării automate ca instrument de inginerie a caracteristicilor; (e) rețelele neuronale profunde cuplate cu criteriile de diagnostic Sepsis-3 tind să dea rezultate mai bune pe datele în formă de serii temporale colectate de la pacienții cu sepsis. Stabilirea unor criterii de evaluare și standarde de raportare facilitează dezvoltarea de modele de învățare automată îmbunătățite pentru aplicații clinice.

Conform [140] modelele care prezic rezultatele clinice sunt abundente în literatura medicală și sunt clasificate în linii mari ca acelea care estimează probabilitatea prezenței unui anumit rezultat (diagnostic) sau dacă un anumit rezultat (de exemplu, eveniment) va apărea în viitor (prognostic) [141]. În mod tradițional, aceste modele (denumite în continuare modele de predicție) au fost dezvoltate folosind metode bazate pe regresie, de obicei regresia logistică pentru rezultatele pe termen scurt și regresia Cox pentru rezultatele pe termen lung [142]. Numeroase recenzii au

observat că studiile, care descriu dezvoltarea și validarea (inclusiv actualizarea) unui model de predicție, nu reușesc adesea să raporteze informații cheie pentru a ajuta cititorii să judece metodele și să aibă o imagine completă, transparentă și clară a acurateții predictive a modelului și a altor detalii relevante [143, 144].

Pentru a diminua efectul absenței unei raportări complete și cuprinzătoare, care limitează capacitatea de utilizare a constatărilor studiilor, de exemplu, în studiile de validare, studiile de sinteză a dovezilor sau în practica zilnică, în 2015 a fost publicat Ghidul pentru raportarea transparentă a unui model de predicție multivariabilă de prognostic sau diagnostic individual (TRIPOD) [141].

De la publicarea TRIPOD, inteligența artificială (IA) și în special învățarea automată, abordările de predicție clinică au evoluat și au căpătat popularitate tot mai mare, odată cu creșterea rapidă a numărului de publicații în domeniile respective. Acest lucru este evident într-un recent studiu, care revizuieste modelele de predicție legate de COVID-19, în care 57 (din 107 studii incluse) au folosit metode de învățare automată [145]. Învățarea automată, o ramură a IA, reprezintă în linii mari metode de analiză a datelor care învață din date fără a fi programată în mod explicit, cu modele identificate pe baza datelor în sine. Ele sunt adesea descrise ca având flexibilitate pentru a capta asocieri complexe, în special în date mari și nestructurate. În timp ce marea majoritate a elementelor din Ghidul TRIPOD sunt relevante pentru studiile modelelor de predicție bazate pe învățarea automată, există și unele aspecte care nu sunt acoperite.

Dintr-o parte, datorită complexității lor, aceste modele de predicție sunt de obicei considerate a fi „cutie neagră”, spre deosebire de modelele mai simple (de exemplu, modele în baza regresiei liniare), în care modelul complet poate fi prezentat în mod transparent (de exemplu, ca o ecuație care conține toți coeficienții de regresie). Din altă parte, în timp ce multe metode de învățare automată au origini în literatura statistică, au apărut două culturi de modele de predicție: cele din mediul statistic/epidemiologic și cele din informatică/știința datelor [145]. Deși există o suprapunere clară între aceste două abordări ele au generat terminologii diferite și uneori contradictorii.

Datorită noutății relative a aplicării metodelor de învățare automată la modelarea predicției clinice, există puține informații despre calitatea raportării acestor studii. Cu toate acestea, puținele recenzii care au examinat caracterul raportării, au cuantificat-o ca una slabă [146, 147]. Ca răspuns la aceste preocupări, este necesară îndrumare pentru a ajuta autorii să-și descrie pe deplin studiul modelului de predicție atunci când au fost utilizate metode de învățare automată [140]. Astfel, grupul TRIPOD a inițiat un proiect internațional pentru a dezvolta o extindere a TRIPOD bazată pe consens, cu un accent special pe raportarea studiilor care întreprind dezvoltarea, validarea sau

actualizarea unui model de predicție de diagnostic sau prognostic, folosind tehnici de învățare automată - denumit în continuare TRIPOD – AI [140].

Extensia TRIPOD-AI, care cuprinde o listă de verificare și un document însoțitor de elaborare și explicație, se așteaptă să ofere cercetătorilor, autorilor, editorilor, utilizatorilor și altor părți interesate studii de modele de predicție bazate pe învățarea automată, cu îndrumări privind setul minim de elemente de raportat, cu exemple detaliate de raportare bună pentru fiecare articol. De asemenea se așteaptă că raportarea completă va permite ca studiile să fie înțelese, replicate și utilizate.

Evaluarea critică a calității metodei de studiu este, de asemenea, o componentă crucială a medicinei bazate pe dovezi. În acest context, în 2019, a fost publicat Instrumentul de evaluare a riscului de părtinire a modelului de predicție (PROBAST) [140, 148] pentru a ajuta o varietate de părți interesate, inclusiv, de exemplu, recenzori sistematici, cercetători, editori de reviste, recenzori de manuscrise și factori de decizie implicați în elaborarea ghidurilor clinice, să evalueze critic proiectarea studiului, efectuarea și analiza studiilor modelului de predicție. PROBAST cuprinde patru domenii (participanți, predictorii, rezultat și analiză) și conține 20 de întrebări de semnalizare pentru a facilita evaluarea riscului de părtinire. În mod clar, evaluarea și raportarea riscului de părtinire sunt intrinsec legate, prin aceea că evaluarea riscului de părtinire se bazează pe ceea ce a fost raportat în studiul primar. În timp ce PROBAST este relevant pentru studiile modelelor de predicție care utilizează învățarea automată, au apărut diferite abordări ale dezvoltării și validării modelelor și terminologiei, iar capacitatea de a evalua critic aceste studii este crucială înainte de a fi implementate [149]. Astfel, în paralel cu dezvoltarea TRIPOD-AI, se propune de a fi dezvoltat, PROBAST-AI, un instrument de evaluare a riscului de părtinire în studiile modelelor de predicție multivariabile bazate pe învățarea automată [140].

Atât TRIPOD-AI, cât și PROBAST-AI se concentrează pe rapoartele de cercetare sau eforturile în care un model de predicție multivariabilă este dezvoltat (sau actualizat) sau validat (testat) folosind orice tehnica de învățare automată. Conform publicațiilor originale TRIPOD și PROBAST, un model de predicție multivariabil este definit ca orice combinație sau ecuație a doi sau mai mulți predictorii care urmează să fie utilizați pentru predicții individualizate pentru a estima probabilitatea unui individ de a avea (diagnostic) sau de a dezvolta (prognoză) un anumit rezultat sau stare a sănătății. Predictorii pot avea orice formă și pot rezulta din istoricul pacientului, examenul fizic, testele de diagnostic, prognostic sau de monitorizare și din tratamentele primite. Rezultatele pot avea, de asemenea, orice formă (dihotomică, categorială, continuă) și de orice fel, cum ar fi o anumită afecțiune sau tulburare prezentă sau absentă (rezultat diagnostic sau clasificare), rezultate de prognostic pe termen scurt (de exemplu, mortalitate în spital sau

complicații postoperatorii), și rezultate prognostice pe termen lung, cum ar fi apariția complicațiilor tratamentului la un an, apariția metastazelor la 5 ani sau supraviețuirea pe tot parcursul vieții.

Conform publicațiilor originale, TRIPOD-AI și PROBAST-AI urmează să abordeze, de asemenea, studii de model de predicție din toate mediile de îngrijire medicală (sănătate publică, îngrijire primară, secundară, terțiară și azil) și toate populațiile țintă corespunzătoare.

În sumar, același studiu [140] concludă că numărul de studii de modele de predicție care utilizează metode de învățare automată crește rapid, inclusiv modele de predicție dezvoltate, validate sau actualizate. Este important să se asigure că detaliile cheie sunt raportate, astfel încât cititorii să poată evalua calitatea studiului și să interpreteze concluziile acestuia, inclusiv modelul de predicție dezvoltat, validat sau actualizat, pentru a îmbunătăți implementarea acestora în cercetările ulterioare (de exemplu, studii de validare), proiecte de sinteză a dovezilor (de exemplu, revizuirii sistematice ale modelelor de predicție) și în practica zilnică de către profesioniștii din domeniul sănătății.

Cele descrise până aici se referă în primul rând la studiile diagnostice. Un aspect important al acestor studii este raportarea rezultatelor lor. Pentru a îmbunătăți calitatea raportării studiilor de acuratețe a diagnosticului, a fost elaborată declarația “Standards for Reporting Diagnostic Accuracy” (STARD). În 2015 a fost publicată o listă actualizată de 30 de elemente esențiale care ar trebui incluse în fiecare raport al unui studiu de acuratețe a diagnosticului. Această actualizare încorporează dovezi recente despre sursele de părtinire și variabilitatea acurateței diagnosticului și are scopul de a facilita utilizarea STARD și poate ajuta la îmbunătățirea completitudinii și transparenței în raportarea studiilor de acuratețe a diagnosticului [150].

STARD 2015 nu abordează problemele și provocările ridicate de intervențiile centrate pe inteligența artificială (AI). Astfel, în 2021 a fost inițiată o versiune specifică pentru IA a listei de verificare STARD (STARD-AI), care se concentrează pe raportarea studiilor de acuratețe a testelor de diagnosticare care implică tehnologiile de ÎA/IA., care este în curs de dezvoltare [151] (mai multe detalii referitor la aceste aspecte vor fi prezentate în capitolul următor).

În contextul celor expuse anterior se poate de sumarizat că datele din literatură arată că depistarea precoce sau predicția sepsisului pot reduce timpul până la administrarea antibioticelor [152], iar intervenția timpurie la rândul sau s-a dovedit a reduce rata mortalității în acest grup de pacienți [153]. Spre exemplu, utilizarea unui sistem de învățare automată (SÎA) din studiul [154] a fost asociat cu o reducere cu 39,5% a mortalității spitalicești ($p < 0,001$), o reducere cu 32,3% a sejurului ($p < 0,001$) și o reducere cu 22,7% ratei reinternărilor timp de 30 de zile ($p < 0,001$).

Deși sistemele de învățare automată (SÎA) reprezintă suplimentele destul de noi la domeniul managementului clinic al sepsisului, algoritmiile de învățare automată au potențialul de a îmbunătăți în mod semnificativ rezultatele pacienților prin avertizarea în avans a debutului iminent al sepsisului. SÎA -urile cu abilități de predicție a sepsisului pot servi, de asemenea, pentru a permite medicilor clinici să aibă încredere în diagnosticul de sepsis într-o varietate de cazuri ambigui, inclusiv cazuri în care rezultatele culturii pozitive nu sunt disponibile [155] și în cazurile de prezentare clinică atipică, în special în rândul pacienților mai în vârstă, care cuprind o majoritate de cazuri de sepsis [156]. Prin urmare, sistemele de asistență decizională bazate pe învățare automată reprezintă un domeniu important de investigare pentru cercetarea sepsisului [130-139, 152, 157, 158].

1.4. Sinteza Capitolului 1.

Acest capitol este axat pe trecerea în revistă a datelor din literatura de specialitate pe trei direcții, care stau la baza compartimentelor, care conturează sferile cercetării:

- Tehnologiile de învățare automată/inteligență artificială utilizate în anesteziologie - terapie intensivă la etapa actuală;
- Conceptul de sepsis, evoluția conceptului și conținutul lui modern;
- Situația actuală în abordarea problemei sepsisului cu utilizarea tehnologiilor intelectuale în managementul clinic al stării de sepsis.

Dinamica publicațiilor în literatura de specialitate a lucrărilor ce țin de direcția de cercetare curentă confirmă dezvoltarea expansivă a uneltelor pentru managementul diferitor aspecte de diagnostic și tratament în ATI bazate pe IA (Fig. 1.1) și, în particular, ce țin de managementul sepsisului (Fig. 1.3).

Structura ponderii utilizării tehnologiilor de învățare automată în acest domeniu este:

- Învățarea supravegheată – 91%
- Învățarea prin consolidare – 6 %
- Învățarea nesupravegheată – 3%

În pofida progresului în medicină, diagnosticarea sepsisului rămâne în continuare o provocare. Cu toate că în unele clinici mai continue să fie folosită abordarea Sepsis-2 (în baza criteriilor SIRS), abordarea recomandată este Sepsis-3 (în baza SOFA) – din 2016. Majoritatea lucrărilor din ultimii cinci ani care folosesc ÎA au la bază definiția Sepsis-3.

În Ghidul internațional pentru managementul sepsisului și al socului septic din 2021 în premieră este menționat potențialul rol al învățării automate, care poate îmbunătăți performanța instrumentelor de screening. Conform acestui ghid performanța sistemelor de predicție a sepsisului este în jur de 89% după aria de sub curba ROC.

Conform publicațiilor, algoritmi de ÎA cei mai des utilizați în ATI în general și în managementul sepsisului, în particular, sunt:

- a. Regresia liniară
- b. Regresia logistică
- c. Metodele bazate pe arbori decizionali
- d. Metodele cu vectori de suport
- e. Rețelele neuronale artificiale (multistrat)
- f. Rețelele neuronale recurente
- g. Rețelele neuronale convoluționale
- h. Algoritmii bayesieni

Orizontul de predicție cu utilizarea IA variază de la mai mult de 24 ore - la 0,25 ore până la detectarea sepsisului prin metode tradiționale (ex. SOFA). Un orizont de 3-4 ore poate fi considerat optim pentru a iniția în timp util tratamentul (antibacterian).

Performanța modelelor de predicție a sepsisului raportate în literatura de specialitate în ultimii cinci ani variază între 55% și 98%, majoritatea lucrărilor raportând o performanță mai mare de 80%.

Este recomandată planificarea și realizarea proiectelor de cercetare, precum și raportarea rezultatelor în baza ghidurilor TRIPOD, PROBAST, STARD, etc. În ultimii ani aceste ghiduri au fost extinse astfel ca să acopere și studiile în baza IA.

2. METODOLOGIA CERCETĂRII ÎN STUDIILE BAZATE PE ÎNVĂȚAREA AUTOMATĂ / INTELIGENȚA ARTIFICIALĂ. MATERIALE ȘI METODE DE CERCETARE ÎN STUDIUL CURENT

2.1. Designul general al cercetării

Conform abordărilor metodologice și statistice tradiționale studiul curent, după produsul finit – un sistem de predicție timpurie a sepsisului – este o cercetare ce vizează un sistem de predicție multivariabil pentru prezicerea unui eveniment (sepsis). Metodologia și regulile de raportare a rezultatelor în acest caz sunt descrise de mai multe documente cu caracter de recomandare/reglementare [141, 142, 148, 151, 160, 161], în special, de ghidul TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis, 2015) [141] și principiile PROCAST (Prediction model Risk Of Bias ASsessment Tool, 2019) [160] și altele [151, 161].

De la publicarea TRIPOD, inteligența artificială (IA) și, în special, învățarea automată (ÎA), utilizate cu scop de predicție clinică au evoluat și au câpătat tot mai mare popularitate, odată cu creșterea rapidă a numărului de publicații privind inteligența artificială și învățarea automată [162-167]. Învățarea automată poate fi descrisă în linii mari ca metode de analiză a datelor care învață din date fără a fi programată în mod explicit, cu modele identificate pe baza datelor în sine. Ele sunt adesea descrise ca având flexibilitate în captarea asocierilor complexe, în special în date cu volum mare și nestructurate, ce condiționează complexitate înaltă în modelare. În timp ce marea majoritate a elementelor din ghidul TRIPOD sunt relevante pentru studiile modelelor de predicție bazate pe învățarea automată, există unele provocări unice cu învățarea automată, care nu sunt surprinse de acest ghid. Datorită complexității lor, aceste modele de predicție sunt de obicei considerate a fi „cutie neagră”, spre deosebire de modelele bazate pe regresie, în care modelul complet poate fi prezentat în mod transparent (de exemplu, ca o ecuație care conține toți coeficienții de regresie).

De asemenea, în timp ce multe metode de învățare automată au origini în literatura statistică, au apărut două culturi de modele de predicție: (a) cele din mediul statistic/epidemiologic și (b) cele din informatică/știința datelor [168]. Deși există o suprapunere clară, abordările diferite în dezvoltarea modelului, validarea și actualizarea lui au generat terminologii diferite și uneori contradictorii [169]. O analiză comparativă a celor mai des utilizați termeni este prezentată în Tabelul 2.1 (după [169] cu suplimentări).

Tabelul 2.1. Analiza comparativă a terminologiei ÎA vs modelarea statistică tradițională

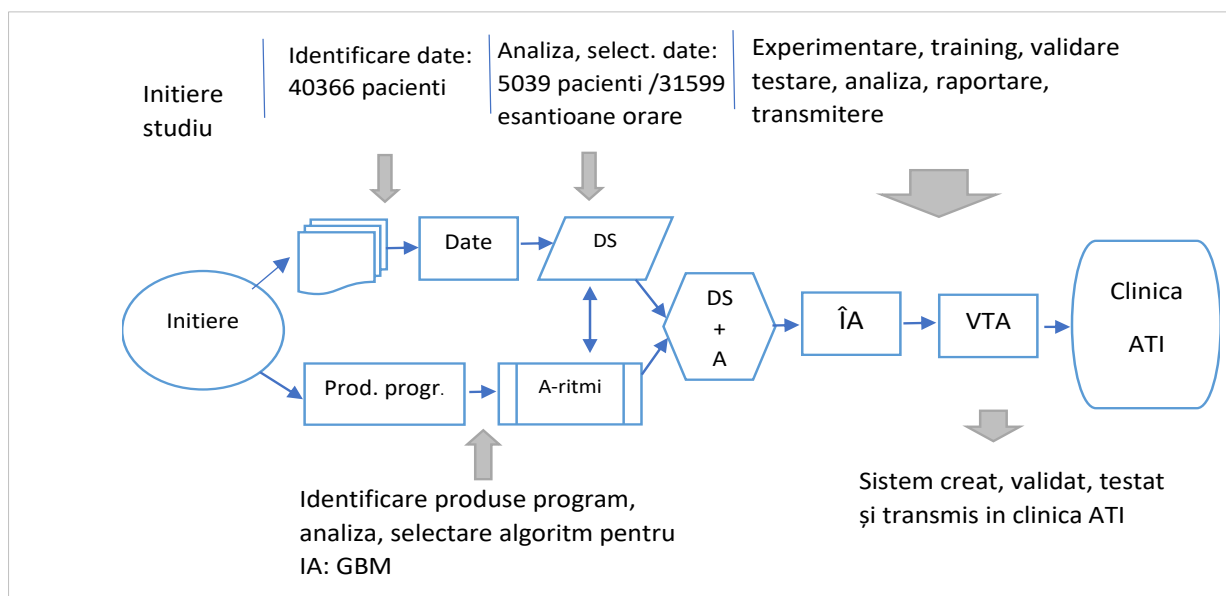
Termen ÎA	Concept statistic comparabil
Caracteristică	Variabilă independentă (explicativă)/predictor
Etichetă/răspuns/rezultat	Variabilă dependentă
Selecție a caracteristicii(-lor)	Selecție variabilă(-le)
Caracteristici extrase	Variabile latente sau transformate
Optimizare model	Potrivire/ajustare model
Training sau învățare	Estimarea modelului
Ponderi (weights)	Parametri
Tensor	Matrice sau vector de ordin superior
Măricea de confuzie/tabel de clasificare greșită	Tabel de contingență (ex. 2x2)
Recall (sensibilitate)	Sensibilitate
Precizie	Valoare predictivă pozitivă

Pentru a face față acestei provocări, grupul TRIPOD a inițiat un proiect internațional (în curs de realizare) pentru a dezvolta o extindere a TRIPOD bazată pe consens, cu accent special pe raportarea studiilor care întreprind dezvoltarea, validarea sau actualizarea unui model de predicție de diagnostic sau prognostic, folosind tehnici de învățare automată - denumită în continuare TRIPOD – AI [140]. Cercetarea curentă tinde să îmbine abordarea tradițională cu cea caracteristică pentru ÎA, utilizând recomandările din [169].

În contextul abordării tradiționale, cercetarea curentă este un studiu observațional - analitic retrospectiv cu câteva etape în baza unui set inițial de date ce descriu 40366 pacienți, inclusiv 2932 de pacienți cu sepsis dintr-o bază de date publică [170]. Datele inițiale conțin multe valori lipsă (până la 80.9% date). După analiza exploratorie pentru crearea modelului final au fost selectați 5039 pacienți, dintre care 1703 – cu sepsis. Criteriile de selectare au inclus: (a) prezența a cel puțin 7 ferestre de observare orare până la apariția/diagnosticarea sepsisului (4 ore – orizontul de prognozare, 3 ore – pentru evaluarea dinamicii parametrilor care descriu pacientul) și (b) densitatea maxim posibilă a datelor (pentru pacienții non-septici – prezența a 7 valori consecutive pentru fiecare din parametrii de interes). În final au fost generate 31599 de eșantioane orare, care sunt utilizate pentru ÎA. Datele pacienților septici cu valori lipsă au fost reconstruite conform unui algoritm elaborat în cadrul studiului curent. Aceste date au fost utilizate pentru crearea unui sistem de prognozare timpurie a sepsisului.

Elaborarea sistemului a urmat etapele tradiționale în crearea unui sistem de învățare automată: training, cross-validare, testare pe date noi, care nu au fost utilizate la etapele precedente.

Sunt analizate rezultatele cercetării bazate pe performanța sistemului, descrise concluziile finale și recomandări practice. Sistemul creat, sub forma de aplicație va fi transmis în Clinica ATI a Institutului de medicină urgentă, Chișinău pentru calibrare și dezvoltare de viitor.



Notă: DS - date selectate; produsele-program reprezintă algoritmi (A-ritmi) utilizați pentru învățarea automată, VTA – validare, testare, analiză, ÎA – învățare automată

Figura 2.1. Designul general al studiului

Conform [169] la proiectarea studiului inițial se articulează clar întrebarea de cercetare astfel ca să încadreze toate alegerile ulterioare cu privire la pregătirea datelor, selecția metodei, rezultatele și evaluarea performanței, precum și interpretarea (Fig. 2.2).

Când se efectuează analiza în baza ÎA, se fac 2 ipoteze: (1) rezultatele dorite pot fi generate utilizând datele de intrare și (2) datele disponibile conțin informațiile necesare pentru a obține rezultatul dorit. Este important să se țină cont de aceste ipoteze atunci când se iau în considerare datele de intrare, metoda ÎA și arhitectura generală de analiză care vor fi utilizate pentru a aborda întrebarea de cercetare.

Clasificarea întrebării de cercetare și a analizei după tipul de problemă (de exemplu, clasificare binară/multiclasă în baza seriilor temporale sau analiză secvențială versus explorare sau categorizare) este importantă deoarece acești factori determină metodele ÎA adecvate, având în vedere datele disponibile și rezultatul dorit. De exemplu, tehnicile supravegheate pot fi utilizate

numai dacă este disponibil un răspuns etichetat corespunzător. Acesta este și cazul cercetării curente.

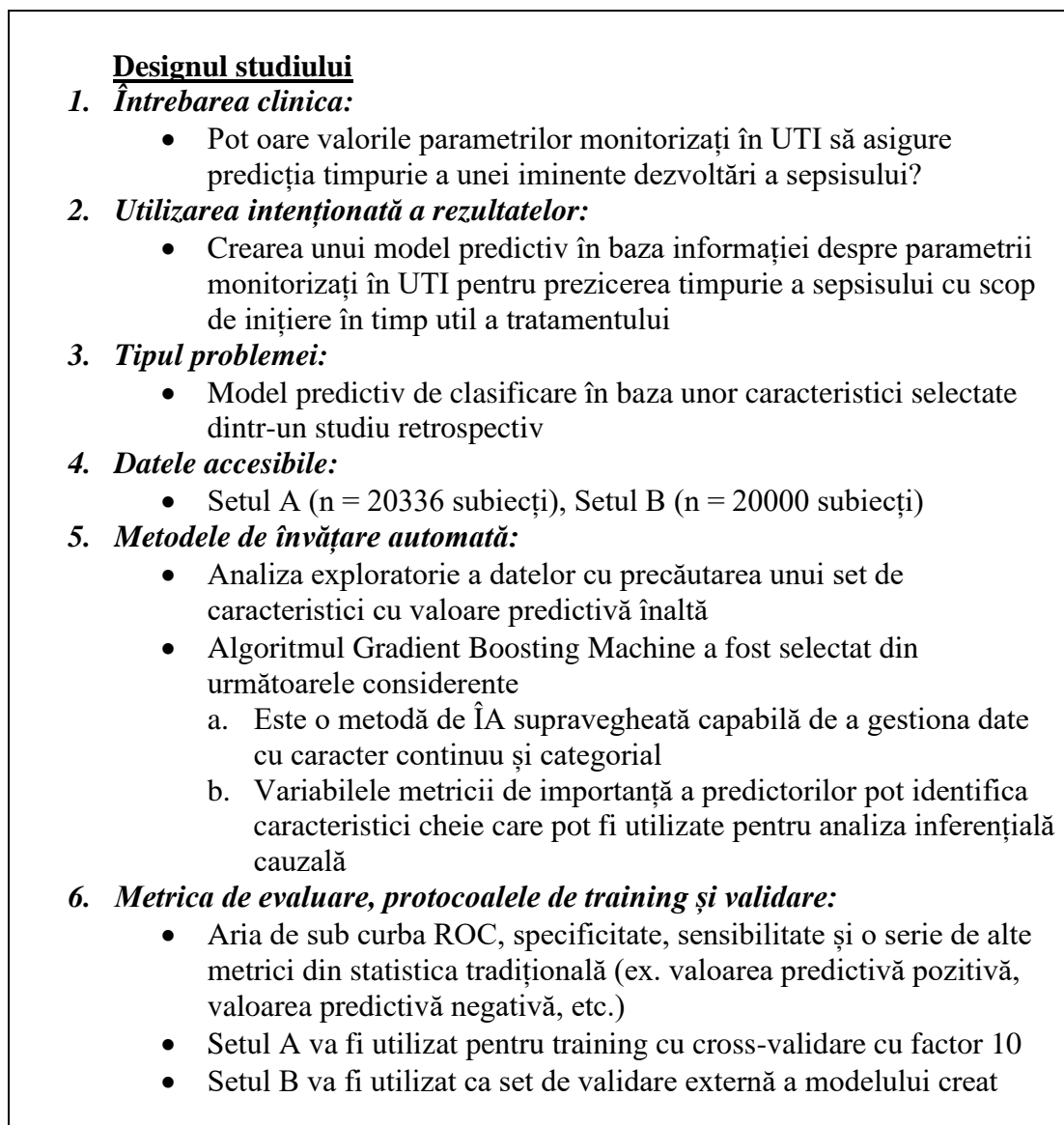


Figura 2.2. **Elementele cheie pentru raportarea designului și a rezultatelor studiului curent.** UTI – unitatea de terapie intensivă

După definirea întrebării de cercetare, tipul problemei trebuie să fie clar precizat: în cazul de față – clasificare binară. Astfel, scopul principal al studiului curent este de a utiliza date în formă de serii temporale multimodale generate în rezultatul monitorizării pacientului în UTI pentru o problemă de clasificare binară (sepsis versus non-sepsis) cu scop de predicție precoce a sepsisului.

Raționamentul utilizării metodelor ÎA în studiul curent este avantajul lor marcat de a face în general posibilă predicția precoce a sepsisului, față de metodele clinice tradiționale (ex. SOFA, SIRS, etc.), care au fost elaborate pentru altă sarcină – confirmarea diagnosticului de sepsis. În cercetarea clinică, raționamentul utilizării ÎA derivă și din avantajele pe care le asigură aceasta în

cazul datelor complexe (ex. serii temporale multimodale). Din cauza barierelor practice în calea implementării clinice a modelelor bazate pe ÎA – inclusiv nevoia frecventă de interpretabilitate a modelului în luarea deciziilor clinice – justificarea utilizării lor este de importanță majoră [171].

În cazul cercetării curente metoda GBM (Gradient Boosting Machine) bazată pe utilizarea arborilor decizionali/de clasificare este o opțiune superioară datorită capacității sale de a identifica relații neliniare într-o varietate de tipuri de date cu dimensiuni mari. Pentru ÎA supravegheată în general, și pentru metoda utilizată în cercetarea curentă în particular, este important să se descrie etichetele pentru predicție – în seturile utilizate în studiu cazurile septice sunt etichetate cu “1”, iar cele non-septice – cu “0”, pentru fiecare observație orară. Descrierea sursei și caracteristicile datelor sunt prezentate în continuare.

2.2. Descrierea materialului clinic

Cercetarea a fost aprobată de Comitetul de Etică a Cercetării USMF „Nicolae Testemițanu”, aviz favorabil nr. 3 din 18.03.2022.

Setul de date utilizat în teză reprezintă date de acces public din cadrul “Early Prediction of Sepsis from Clinical Data: the PhysioNet Computing in cardiology Challenge 2019” [170]. Aceste date sunt accesibile conform licenței ODS (Open Database License) (<https://physionet.org/content/challenge-2019/1.0.0>). Datele provin din două sisteme de spital din S.U.A., distincte din punct de vedere geografic: Centrul medical Beth Israel Deaconess (setul A) și Spitalul universitar Emory (setul B). Aceste date au fost colectate în ultimul deceniu cu aprobarea comitetelor de etică instituționale, au fost identificate și etichetate folosind criteriile clinice Sepsis-3 [172]. Datele constau dintr-o combinație de rezumate ale semnelor vitale la fiecare oră, valori de laborator și descrieri statice ale pacientului. În special, datele conțin 40 de variabile clinice: 9 variabile de semne vitale și suport vital (FiO_2), 26 de variabile de laborator și 6 variabile demografice și logistice. În total, aceste date includ peste 1,5 milioane de ferestre orare și peste 10.4 milioane de puncte de date (variabile fiziologice și de laborator “non-missing”). Datele extrase din sistemele electronice ale spitalelor respective au trecut printr-o serie de etape de preprocesare înainte de analiza formală. Toate caracteristicile pacientului au fost condensate în coșuri orare, simplificând dezvoltarea și testarea modelului; de exemplu, mai multe măsurări ale frecvenței cardiace într-o fereastră orară de timp au fost rezumate ca măsurarea ritmului cardiac mediu. Astfel, datele reprezintă serii temporale multimodale. O serie temporală (ST) este o secvență de puncte de date (măsurări) care are o ordonare temporală naturală. Una din cele mai simple reprezentări matematice ale ST este:

$$Y = (Y_t: t \in T), \quad (2.1)$$

unde T este [y[setul de indexare]].

O definiție mai desfășurată a ST este prezentată în Anexa 1.

Tabelul. 2.2. Caracterizarea generală a seturilor de date utilizate

Parametru	Setul A	Setul B
Numărul total de pacienți	20336	20000
Numărul de pacienți cu sepsis	1790	1142
Vârsta, ani (media, ds)	62.62 (16.24)	60,65 (16,67)
Vârsta %		
18-50 ani	21.44	25,89
51-60 ani	18.54	20.13
61-70 ani	22.41	24.94
71-80 ani	22.67	18.65
>80 ani	14.94	10.40
Repartizare pe sexe, femei/bărbați	8502 (41,80%) 11834 (58,2%)	9268 (46,34%) 10732 (53,66%)
Durata observării (ore în ATI) (media, ds)	38.86 (22.31)	38.10 (23,28)
Parametrii clinici monitorizați	<ul style="list-style-type: none"> • 8 semne vitale și 1 parametru de suport vital • 25 parametri de laborator • 6 parametri demografici și de logistică 	
Prevalența sepsisului %	8,80	5,71
Densitatea datelor % (parametri fiziologici și rezultate laborator)	20,61	19,11
Numărul de observații/ferestre orare	790215	741952
Numărul de intrări de date/celule cu valori 'non-missing'	10296016	9831918

Nota: Conform sursei [170] grupul de **parametri “semne vitale și suport vital”** include: frecvența cardiacă (HR), saturația sângelui periferic cu O_2 (SpO_2), temperatura (Temp), presiunea arterială sistolică (SBP), presiunea arterială medie (MAP), presiunea arterială diastolică (DBP), frecvența respirație (Resp), capnometrie ($EtCO_2$), fracția O_2 inspirat (FiO_2); **parametrii de laborator:** excesul de baze (BE), bicarbonatul (HCO_3), pH, presiunea parțială a CO_2 în sângele arterial ($PaCO_2$), saturația cu O_2 a sângelui arterial (SaO_2), aspartat-transaminaza (AST), urea (BUN), fosfataza alcalină (Alkalinephos), calciul seric (Ca), clorul seric (Chloride), creatinina serică (Creat), bilirubina directă (DB), glucoza (Gluc), lactatul seric (Lact), magneziul seric (Mg), fosfații serici (Phos), potasiul seric (K), bilirubina totală (TB), troponina-I (Trop_I), hematocritul (Hct), hemoglobina (Hgb), timpul tromboplastinei parțial activate (aPTT), numărul de leucocite (WBC), fibrinogenul (Fibr), numărul de trombocite (Plat); **parametrii demografici și de logistica:** vârsta, sexul, timpul de internare (AdmT), durata aflării în terapie intensivă (ICULOS), identificatorul secției (ICU Unit), marcaj sepsis (SepsisLabel); ds – deviația standard

În Anexa 2 este prezentat aspectul unui subset de date inițiale (pacientul 000034) în format “.csv”.

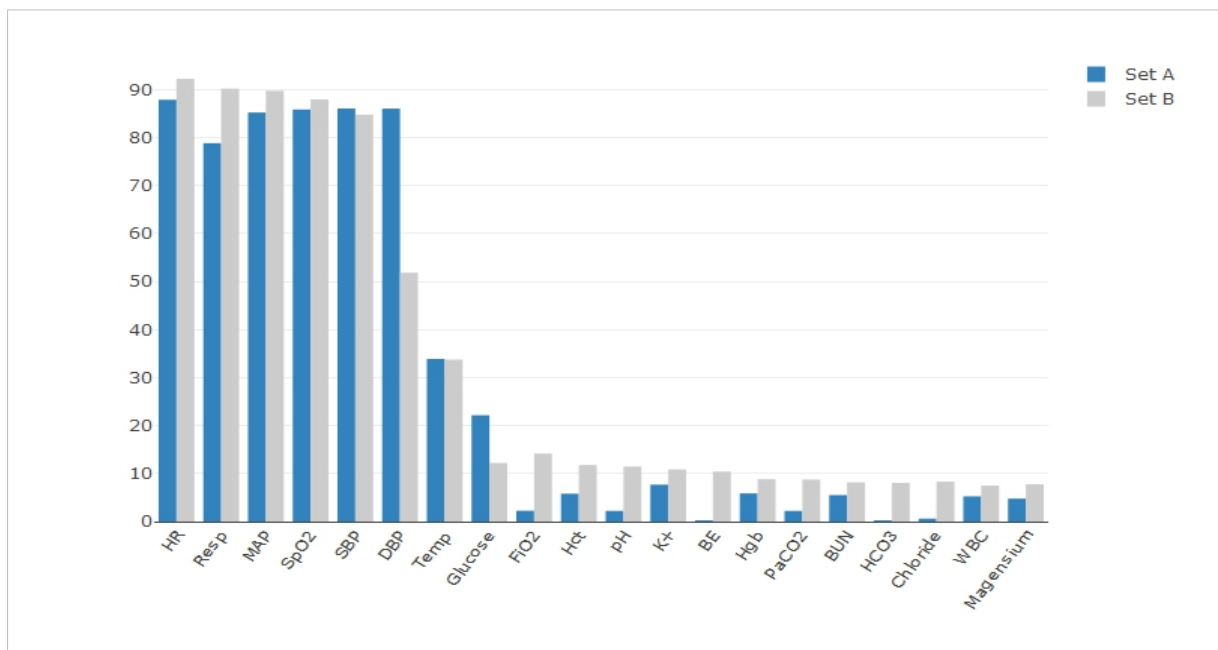


Figura 2.3. Densitatea (prezența) datelor: Primii 20 parametri fiziologici și de laborator cu cea mai înaltă densitate a datelor (în procente).

Notă: Densitatea celorlalți parametri este între 0,55% (Fibrinogen) și 6,83% (Calciu). Marcajul parametrilor (pe axa X) – conform notei la tabelul 2.2.

Sursele de date utilizate pentru analiza ÎA sunt de obicei mari (multe puncte de date) și complexe (multe tipuri diferite de date) și, prin urmare, pot fi dificil sau imposibil de revizuit manual. Descrierea datelor care trebuie utilizate este esențială pentru a evalua calitatea, fiabilitatea acestora pentru a produce rezultatul dorit, potențiala generalizare a oricăror constatări și, în special, reproductibilitatea (Fig. 2.4). Datele trebuie descrise în detaliu în ceea ce privește sursa lor, populația de studiu, instrumentele (colecții standardizate de măsurători), măsurători (definiții specifice variabilelor, numele câmpului și valori, cum ar fi vârsta, sexul masculin, etc.), informații referitor la colectare (detalii instrument/măsurare), etc. și, de asemenea, intervalele de timp de colectare, definiția variabilelor și redundanța, precum și tipul de date (de exemplu, binare, continue, etc.) [173]. Suplimentar la informația prezentată anterior, aceste aspecte ale datelor utilizate în cercetarea curentă sunt sumarizate în Fig. 2.4.

Sursele de date si tehnicile de preprocesare a lor (datele pentru training și validare)

1. Populațiile:

- Setul A – subiecți cu vârsta de 18 – 89 ani din Centrul medical Beth Israel Deaconess, S.U.A.
- Setul B – subiecți cu vârsta de 18 – 100 ani din Spitalul universitar Emory, S.U.A.

2. Sursele și procesul de achiziționare a datelor:

- Date de monitoring fiziologic standard în UTI (FC, FR, SpO₂, temperatura corpului, PA S&D, etc.)
- Date de laborator (hemoleucograma, fermenți hepatici, etc.)
- Înregistrări în fișele de observație (vârsta, sex, diagnoză)

3. Structura și tipul datelor:

- Continui (datele din monitoring și cele de laborator)
- Categoriale (datele demografice, etichetarea cazurilor)

4. Diferența dintre seturile de evaluare si validare:

- Repartizare pe sexe %, (femei/bărbați): Setul A – 41,8/58,2
Setul B – 46,3/53,7
- Prevalența sepsisului: Setul A – 8,8%, setul B – 5,71%

5. Preprocesarea datelor:

- **Agregarea datelor:**
 - Toate caracteristicile pacientului au fost condensate în coșuri orare, de exemplu, mai multe măsurări ale frecvenței cardiace într-o fereastră orară de timp au fost rezumate ca măsurarea ritmului cardiac mediu.
- **Gestionarea datelor-lipsă:**
 - Conform algoritmului de restabilire a datelor elaborat în cadrul studiului curent
- **Transformarea datelor:**
 - Estimarea complexității algoritmice, calculul diferenței valorilor parametrilor de interes între două eșantioane orare consecutive
- **Sursa de etichetare a datelor:**
 - Diagnosticul stabilit conform criteriilor Sepsis-3 (pentru sepsis)

6. Linkul la sursa de date:

<https://physionet.org/content/challenge-2019/1.0.0>

UTI – unitatea de terapie intensivă, FC – frecvența contracțiilor cardiace, FR – frecvența respirației, SpO₂ – saturația cu oxigen a sângelui periferic, PA S&D – presiunea arterială sistolică și diastolică

Figura 2.4. Sumarul caracteristicilor datelor utilizate în studiu

2.3. Aspecte metodologice importante pentru investigație, procesarea datelor și analiza rezultatelor.

Datele inițiale sunt cercetate cu ajutorul metodelor statistice simple/standard: medie, deviație standard, mediană, raport procentual. Analiza datelor selectate pentru crearea nemijlocită a

sistemului mai include estimarea distribuției lor, care servește și la elaborarea algoritmului de reconstrucție a datelor/valorilor lipsă. Logica și metodele de investigare și analiză a datelor sunt influențate de etapele ulterioare și scopul final – elaborarea sistemului de clasificare/prezicere a sepsisului, ceea ce în acest domeniu tradițional poartă denumirea de “Analiza exploratorie a datelor”.

Colectarea și preprocesarea datelor cuprind până la 80% din procesul de analiză și raportarea acestor pași asigură transparența și reproductibilitatea studiului. Toată manipularea datelor înainte de efectuarea analizei principale se recomandă să fie explicată în detaliu, inclusiv gestionarea datelor lipsă, transformările, ingineria caracteristicilor (crearea manuală a caracteristicilor din date brute folosind expertiza în domeniu sau crearea bazată pe date a componentelor principale sau a variabilelor latente), extragerea caracteristicilor (transformarea datelor brute în forma dorită, cum ar fi vectorizarea/crearea de tensori din imagini sau text) și selecția caracteristicilor (selecție manuală folosind expertiza în domeniu sau reducerea dimensionalității bazată pe date prin grupare, învățare semi supravegheată sau alte metode). Pașii de preprocesare pot include metode statistice și ÎA care suplimentează analiza generală.

Multe seturi de date pot avea valori inexacte sau lipsă, iar tehnicile utilizate pentru a gestiona valorile lipsă și valorile aberante, de asemenea, se recomandă să fie evaluate și raportate (Fig. 2.4). Discuția despre abordările lipsei datelor în ÎA este în mare parte aceeași pentru orice metodă de analiză, inclusiv și cea din studiul curent. Datele lipsă pot fi clasificate ca aleatoare versus nealeatoare, lipsă, dar aplicabile sau lipsă și inaplicabile. Multe metode ÎA necesită seturi complete de date pentru fiecare caz, ceea ce poate reduce dimensiunea eșantionului, mai ales atunci când se ia în considerare un număr mare de variabile. Dacă lipsa datelor limitează numărul de caracteristici sau mostre utilizabile, metodele de restabilire a datelor sau ÎA care tolerează datele lipsă pot fi utilizate în cazurile în care datele lipsesc la întâmplare. Expertiza clinică a unui expert în domeniu poate identifica variabile cu utilitate clinică marginală care pot fi excluse. Dacă este cazul, trebuie prezentate criteriile utilizate pentru eliminarea datelor. Ar trebui să fie prezentată justificarea utilizării restabilirii datelor, inclusiv metoda selectată, iar validarea rezultatelor generate cu ajutorul datelor restabilite trebuie raportată pentru a determina eficacitatea și posibilul impact asupra rezultatelor.

În cercetarea biomedicală și clinică este obișnuit să existe mai multe variabile explicative decât eșantioane. Acest lucru scade capacitatea de a detecta predictorii importanți, în timp ce crește paradoxal probabilitatea unor constatări false. Modelele ÎA sunt mai potrivite pentru a analiza seturi de date cu mult mai multe variabile decât subiecți în comparație cu multe abordări statistice tradiționale, dar totuși funcționează cel mai bine atunci când redundanța între variabile este redusă

la minimum și dimensiunile eșantionului sunt suficiente pentru a crea modele eficiente. În cercetarea curentă numărul de eșantioane (31599) este cu mult mai mare decât numărul de caracteristici (14).

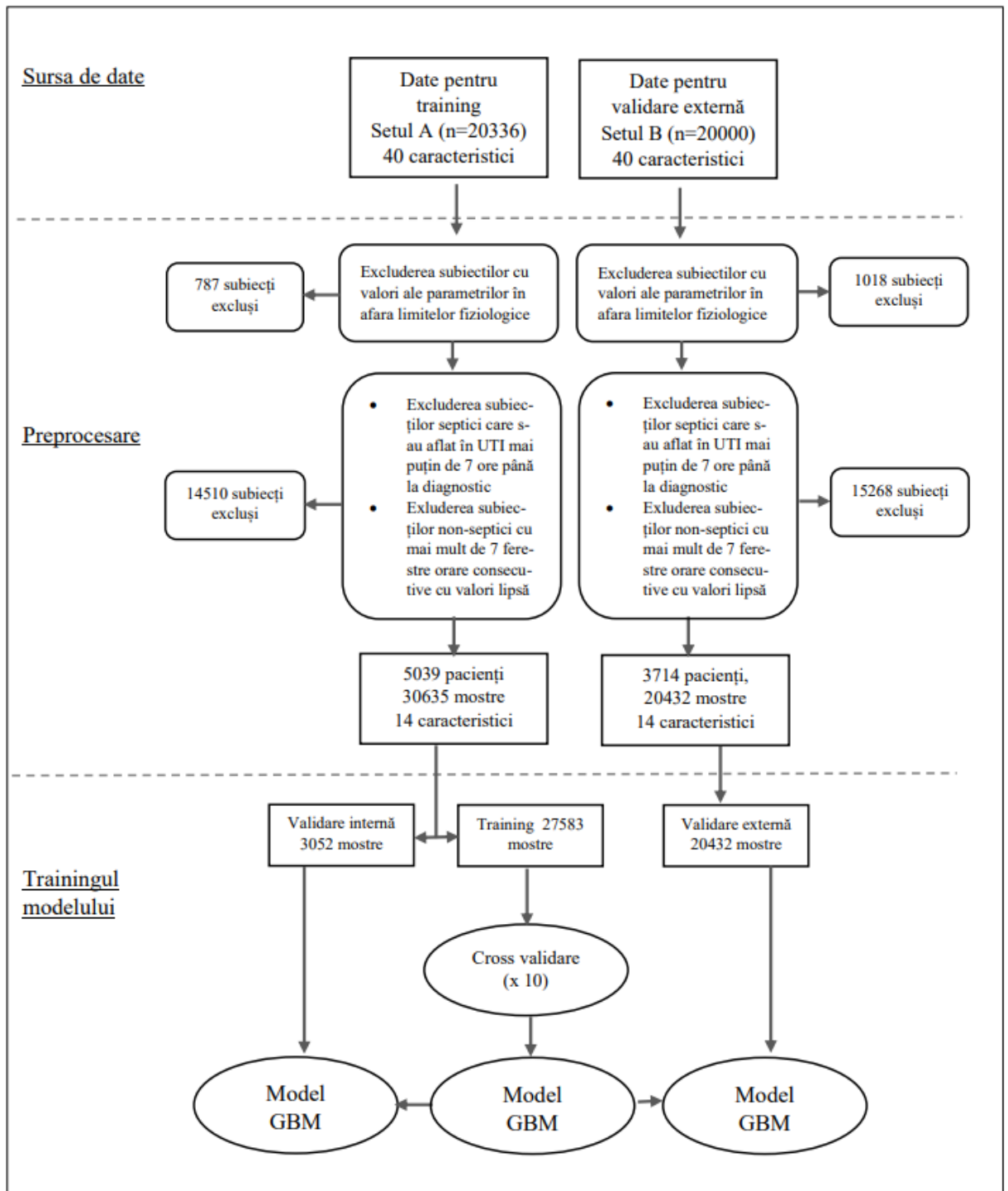


Figura 2.5. Diagrama de tip CONSORT pentru studiul curent

O varietate de metode manuale și tehnici bazate pe date sunt disponibile atunci când se selectează caracteristicile de inclus într-un model. Expertiza în domeniu poate fi folosită pentru a crea un subset de variabile suspectate a avea semnificație clinică. Redundanța variabilelor poate fi, de asemenea, eliminată prin testarea corelației sau prin metode statistice pentru a identifica grupuri de variabile înrudite și a selecta un grup de variabile reprezentative pentru a fi analizate.

Tehnicile de reducere a dimensionalității bazate pe date, cum ar fi analiza componentelor principale, analiza discriminantă liniară, gruparea nesupravegheată și rețelele neuronale artificiale, pot fi, de asemenea, utilizate pentru a extrage un set mai mic de caracteristici proiectate, reprezentative pentru datele de bază pentru analize suplimentare, deși cu riscul unei anumite pierderi de interpretabilitate. Selectarea caracteristicilor bazate pe date este adesea efectuată în contextul unei variabile de răspuns specificate, folosind metode supravegheate, în care mai multe variabile sunt introduse într-un model și evaluate pentru importanța lor în precizarea răspunsului. Acest lucru poate fi eficient pentru identificarea potențialilor factori de risc folosind expertiza de domeniu sau eliminarea redundanței. Deoarece selecția caracteristicilor depinde uneori de valorile variabilei rezultat, aceasta trebuie efectuată fără acces la setul de testare, sau se compromite capacitatea de a obține o estimare imparțială a erorii. Indiferent de metoda utilizată, caracteristicile selectate pentru a fi utilizate în analiza ÎA și baza de selecție se recomandă de a fi raportate pentru toate seturile de date. În plus, numărul de subiecți excluși din analiză trebuie precizat clar. Aceasta poate lua forma unei diagrame CONSORT (Consolidated Standards of Reporting Trials) [174], în care, începând cu întreaga cohortă disponibilă, se indică numărul de pacienți excluși în baza fiecărui criteriu, abordarea utilizată la fiecare etapă și subiecții/eșantioanele rezultante sunt clar delimitate. O asemenea diagramă pentru studiul curent este prezentată în Fig. 2.5.

Se recomandă utilizarea a cel puțin 3 seturi de date într-o analiză ÎA (Fig. 2.5): (a) o parte din setul de date primare care este utilizată pentru a dezvolta modelul(ele) (setul de date pentru training), (b) o parte a datelor primare care este utilizată pentru evaluarea și selectarea dintre modele, cum ar fi diferite valori ale hiperparametrilor (set de date de validare) și (c) unul sau mai multe seturi de date utilizate pentru a evalua modelul după ce acesta a fost complet dezvoltat (set de date de testare/validare externă). Termenii validare și testare, sunt uneori folosiți interschimbabil în literatura clinică, dar se recomandă de făcut distincție dintre validare și seturi de date de testare, deoarece acestea servesc unor scopuri diferite. Datele din seturile de testare nu ar fi trebuit să fi fost folosite în nicio parte a dezvoltării modelului, inclusiv normalizarea datelor, selecția caracteristicilor sau reglarea hiperparametrilor, astfel încât să se evite supra ajustarea și estimările eronat optimiste ale performanței modelului. Seturile de date de testare pot reprezenta o partiție a datelor originale și, astfel, oferă perspective asupra procesului de formare a modelului.

Ele pot fi, de asemenea, derivate dintr-o sursă de date externă complet separată, care să permită evaluarea atât a trainingului modelului, cât și dacă relațiile deduse dintre variabilă și răspuns sunt transferabile în noi setări. Evaluarea performanței modelului pe seturi de testare externe este, de asemenea, puțin probabil să fie influențată de părtinirile apărute în timpul selecției populației de studiu utilizate în dezvoltarea modelului și atenuează riscul ca investigatorul să fi fost influențat de proprietățile setului original de date, înainte de a se împărți în seturi pentru instruire, validare și testare. Cu toate acestea, atunci când un set de date de testare extern nu este disponibil sau când nu este necesară generalizarea la alte seturi de date, evaluarea modelului poate fi limitată la un set de date de training și un set de date pentru testare internă. În studiul curent se utilizează atât testarea internă, cât și validarea/testarea externă.

Indiferent de număr, este esențial să se raporteze rezultatele evaluării pentru toate seturile de date. Evaluarea ar trebui să fie efectuată în contextul măsurilor de performanță definite în timpul proiectării studiului (de exemplu, aria de sub curba ROC), evaluarea ar trebui să fie inclusă în rezultatele analizei și valorile ar trebui raportate pentru toate datele pentru training, validare și seturile de date de testare. În cazurile în care au fost utilizate protocoale de training, cum ar fi bootstrapping sau cross-validarea, raportarea unei valori medii între iterații poate fi suficientă pentru training și raportare de validare internă.

Arhitectura de calcul (de exemplu, viteza procesorului, numărul de nuclee și memorie), tehnicile de înaltă performanță (de exemplu, paralelizarea) pentru implementare, software-ul, pachetele și versiunea lor trebuie raportate pentru toate analizele pentru a facilita reproductibilitatea (Fig. 2.6). Metoda ÎA, hiperparametrii aferenți și justificarea utilizării metodei trebuie raportate pentru orice abordare utilizată în timpul analizei. Se recomandă, de asemenea, de raportat și așa aspecte precum, tehnici de optimizare, cum ar fi hiperparametrii (de exemplu, costul, numărul de arbori și epocile), modificările arhitecturii modelului (de exemplu, adâncimea arborelui sau numărul de straturi) sau implementarea tehnicilor de ansamblu. Implementarea tehnicilor de generalizare (de exemplu, regularizarea L1 sau L2) și hiperparametrii însoțitori (de exemplu, rata abandonului) ar trebui să fie prezentate în mod explicit în anexele suplimentelor de date și în codul de analiză partajat. Parametrii deosebit de importanți pentru analiză sau rezultate ar trebui descriși, inclusiv motivele pentru selectarea acelor parametri. Reprezentarea variabilelor identificate ca importante este crucială.

Dezvoltarea modelului și validarea lui

1. Hardware, software, pachetele utilizate

- AMD Athlon 200 GE, 3.20 GHz, 16 GB RAM
- R versiunea 4.0.5
- h2o 3.32.1.1 (set de algoritmi, inclusiv GBM)
- h2owave 0.0.14
- Wolfram Engine 12.3
- Python 3.8.5

2. Trainingul și evaluarea modelului

- Trainingul realizat în baza setului A
- Aria de sub curba ROC optimizată în timpul trainingului și ajustării
- Protocol de training cu cross-validare (x 10) repetată

3. Parametrii/Hiperparametrii modelului

- Numărul de arbori – 252
- Mărimea modelului – 192608 octeți
- Adâncime medie – 6
- Număr mediu de frunze – 56,14682

4. Caracteristicile selectate

- 6 parametri fiziologici (descriși mai sus) în baza cărora sunt create 2 matrice, pentru care se calculează complexitatea algoritmică și diferența dintre valorile parametrilor în 2 eșantioane orare consecutive = vectori cu lungimea 14 L

5. Metoda de validare și metrica pentru performanță

- Validare externă (setul B)
- Aria de sub curba ROC

6. Reproductibilitate

- Produsul-program pentru etapa de ÎA este posibil de vizualizat și descărcat în formă de Wolfram Notebook
<https://notebookarchive.org/2022-03-b3yfik0>

Figura 2.6. Elementele cheie pentru raportarea trainingului și validării modelului

În baza modelului cel mai performant, dezvoltat în studiul curent, sunt arborii decizionali, care presupun stabilirea celui mai reușit parametru pentru inițierea scindării utilizând un anumit prag (în cazul nostru este diferența de temperatură dintre a treia observație orară și cea de a doua – Diff_Temp_2, și pragul este $\geq 0,095$ grade Celsius). Acesta este “trunchiul” arborelui de la care pornesc “ramuri”, bazate pe scindări a altor parametri după aceleași reguli, cu praguri specifice pentru fiecare ram. În final, după scindări multiple se ajunge la “frunze”, care reprezintă grupurile în baza cărora se obține clasificarea distinctă dorită: „sepsis” sau „non-sepsis”.

Figura 2.7 ilustrează partea incipientă a unui așa arbore în cazul setului de date utilizat în cercetarea curentă. În partea de sus este “trunchiul”, de la care pornesc “ramurile” în baza unor praguri de scindare. În partea de jos a diagramei sunt “frunzele” la care se ajunge în urma scindării. Sunt ilustrați primii pași ai scindării în baza a numai 2 caracteristici (din 14) - diferențele de

temperatură dintre ora a treia și a doua, și dintre ora a doua și cea precedentă: Diff_Temp_2 și Diff_Temp_1. Arborele final are cu mult mai multe „ramuri” și numărul de „frunze” este mai mare de 55, ce face dificilă vizualizarea în formatul obișnuit.

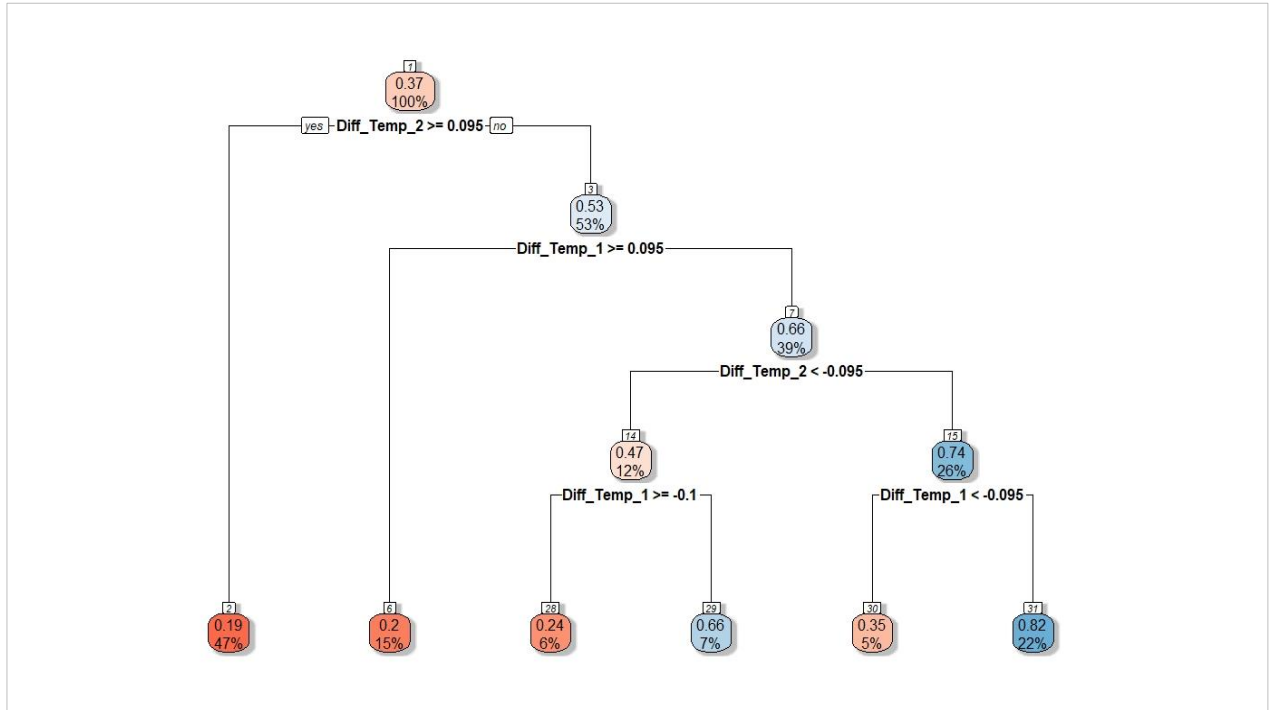


Figura 2.7. Partea incipientă a unui arbore decizional, construit în baza datelor din studiu.

Algoritmul utilizat în acest caz (Gradient Boosting Machine) construiește arbori multipli (până la 252 arbori) și în final îi selectează pe cei mai reușiți/permanenți din punct de vedere a clasificării.

O descriere generală a acestui proces utilizând formalismul matematic (după [175]) este:

Să presupunem că un agent de alocare A are N opțiuni sau strategii din care să aleagă; le numerotăm folosind numerele întregi $1, \dots, N$. La fiecare pas de timp $t = 1, 2, \dots, T$, agentul A decide asupra unei distribuții p^t din spațiul strategiilor; adică $p_i^t \geq 0$ este suma alocată strategiei i , și $\sum_{i=1}^N p_i^t = 1$. Fiecare strategie i suferă apoi o pierdere l_i^t . Pierderea suferită de A este atunci $\sum_{i=1}^N p_i^t l_i^t = p^t \cdot l^t$, adică pierderea medie a strategiilor în raport cu regula de alocare aleasă de A . Numim această funcție de pierdere *pierderea mixtă*. Să presupunem că pierderea suferită de orice strategie este mărginită astfel încât, fără pierderea generalității, $l_i^t x \in [0, 1]$. În afară de această condiție, nu facem nicio presupunere despre forma vectorilor de pierdere l^t , sau despre modul în care sunt ei generați. Scopul algoritmului A este de a minimiza pierderile sale cumulate în raport cu pierderea suferită de cea mai bună strategie. Adică, A încearcă să-și minimizeze *pierderea netă*

$$L_A = \min_i L_i, \quad (2.2)$$

unde

$$L_A = \sum_{t=1}^T p^t \cdot l^t \quad (2.3)$$

este pierderea totală cumulativă suferită de algoritmul A la primele încercări T și

$$L_i = \sum_{t=1}^T l^t \quad (2.4)$$

este pierderea cumulată a strategiei i .

Operația de alocare descrisă aici este echivalentă cu operația de scindare descrisă mai sus în cazul construirii unui arbore, iar scopul algoritmului este de a ajunge la arborele/colecția de arbori cu cea mai mică pierdere, care în principiu sunt cei mai performanți. Descrierea mai detaliată a algoritmului este prezentată în continuare.

Algoritmul de stimulare a (coborârii) gradientului – Gradient Boosting (dupa [176])

Acest algoritm este una din metodele de baza utilizate în studiul curent. Dacă notăm cu X mulțimea vectorilor care descriu cazurile clinice/pacienții și cu Y - etichetele/ diagnosticurile (sepsis sau non-sepsis), atunci sarcina este de a găsi așa posibilități, încât în rezultat să obținem etichetări (preziceri a diagnosticului), realizate de algoritm/sistemul de predicție, care sunt cât mai aproape de realitate.

Algoritmul dat este un algoritm de stimulare care optimizează riscul empiric prin coborârea gradientului – prin selecția celei mai abrupte pante în spațiul funcțional. În general, problema de optimizare pentru estimarea funcției de regresie $f(\cdot)$ a unui model statistic, relaționând variabilele predictoare X cu rezultatul Y , poate fi exprimată ca:

$$\hat{f}(\cdot) = \underset{f(\cdot)}{\operatorname{argmin}} \{E_{X,Y}[p(Y, f(X))]\} \quad (2.5)$$

unde $p(\cdot)$ denotă o funcție de pierdere. Cea mai comună funcție de pierdere este pierderea L_2 , $p(y, f(\cdot)) = (y - f(\cdot))^2$, ceea ce duce la regresia clasică cu cele mai mici pătrate a mediei: $f(x) = E(Y|X = x)$. În practică, cu un eșantion de învățare cu observațiile $(y_1, x_1), \dots, (y_n, x_n)$ minimizăm riscul empiric:

$$\hat{f}(\cdot) = \left\{ \frac{1}{n} \sum_{i=1}^n p(y_i; f(x_i)) \right\} \quad (2.6)$$

Ideea fundamentală de stimulare a gradientul constă în potrivirea modelului la vectorul de gradient negativ $\mathbf{u}^{[m]}$ al funcției de pierdere $\rho(y, \hat{f}(x))$ evaluată la iterația anterioară $m - 1$:

$$\mathbf{u}^{[m]} = \left(u_i^{[m]} \right)_{i=1, \dots, n} = \left(-\frac{\partial}{\partial f} p(y_i, f) \Big|_{f=\hat{f}^{[m-1]}(\cdot)} \right)_{i=1, \dots, n} \quad (2.7)$$

În cazul pierderii L_2 , $\rho(y, f(\cdot)) = \frac{1}{2}(y - f(\cdot))^2$ duce pur și simplu la reajustarea reziduurilor $y - f(\cdot)$. În fiecare iterație de amplificare m , modelul de bază se potrivește direct cu erorile făcute în iterația anterioară $y - f(\cdot)[m - 1]$. Ținând cont de acest principiu, devine evident că algoritmul de stimulare a gradientului sporește performanța unui simplu model de bază, mutând iterativ atenția către observații problematice care sunt „dificil” de prezis. Astfel, amplificarea gradientului identifică observații dificile prin reziduuri mari calculate în iterațiile anterioare. În continuare este descrierea algoritmului pe pași:

Inițializare

- (1) Setează contorul de iterație $m = 0$. Inițializați predictorul aditiv $\hat{f}^{[0]}$ cu o valoare de pornire, de ex. $\hat{f}^{[0]} := (0)_{i=1, \dots, n}$. Specificați un set de modele de bază $h_1(x_1), \dots, h_p(x_p)$.

Potrivirea gradientul negativ

- (2) Setează $m := m + 1$.
- (3) Calculați vectorul de gradient negativ \mathbf{u} al funcției de pierdere evaluată la iterația anterioară:

$$\mathbf{u}^{[m]} = \left(u_i^{[m]} \right)_{i=1, \dots, n} = \left(-\frac{\partial}{\partial f} p(y_i, f) \Big|_{f=\hat{f}^{[m-1]}(\cdot)} \right)_{i=1, \dots, n}.$$

- (4) Ajustați vectorul de gradient negative $\mathbf{u}^{[m]}$ separat la fiecare model de bază:

$$u^{[m]} \xrightarrow{\text{model de baza}} \hat{h}_j^{[m]}(x_i) \quad \text{pentru } j = 1, \dots, p.$$

Actualizarea unei componente

- (5) Selectați componenta j^* care se potrivește cel mai bine vectorului de gradient negativ:

$$j^* = \operatorname{argmin}_{1 \leq j \leq p} \sum_{i=1}^n \left(u_i^{[m]} - \hat{h}_j^{[m]}(x_i) \right)^2.$$

- (6) Actualizați predictorul aditiv \hat{f} cu aceasta componentă

$$\hat{f}^{[m]}(\cdot) = \hat{f}^{[m-1]}(\cdot) + sl \cdot \hat{h}_{j^*}^{[m]}(x_{j^*}),$$

unde sl este o lungime mică a pasului ($0 < sl \ll 1$). O valoare tipică în practică este 0,1.

Repetare

Repetăți pașii (2) până la (6) până când $m = m_{stop}$.

Conform celor descrise mai sus, este important de rezervat un set de testare, care nu a fost utilizat în timpul dezvoltării modelului, pentru a evita supra ajustarea modelului (Fig. 2.6). În cazurile în care se utilizează un set de date extern independent, trebuie descrisă metoda de selectare a setului de testare. Motivul pentru selecția setului de date de testare externă ar trebui să includă o evaluare a similitudinii dintre seturile de training (utilizat pentru a crea modelul), validare (utilizat pentru a evalua performanța modelului și optimizarea hiperparametrilor) și testare (utilizat pentru a testa și valida modelul final). De asemenea, se recomandă de raportat disponibilitatea, manipularea și transformarea caracteristicilor, inclusiv etichetele de răspuns utilizate în trainingul modelului ÎA. Modelele ÎA supravegheate trebuie evaluate din cauza pericolului semnificativ de supraajustare și lipsă de reproductibilitate, care nu poate fi dedusă din evaluările de performanță ale setului de date de training singur. Orice măsurători suplimentare de performanță utilizate pentru evaluarea modelului ar trebui să fie, de asemenea, descrise în mod clar, împreună cu rațiunea selectării acestora. Ar trebui să se acorde atenție performanței probabilistice și valorilor determinate de praguri, precum și modului în care aceste măsuri pot fi afectate de seturi de date dezechilibrate.

Reproductibilitatea rezultatelor în analiza bazată pe ÎA este o problemă importantă și discutată pe larg, în parte, din cauza lipsei de reproductibilitate a multor rezultate a ÎA [177]. În mod ideal, trebuie să fie posibilă inspectarea și executarea tuturor părților analizei (inclusiv preprocesarea datelor) în medii accesibile publicului în resurse bazate pe web și cloud. Astfel de medii ar include seturile complete de date, pachetele software necesare și codul de analiză pentru o transparență deplină și pentru a permite cititorilor să interacționeze cu fluxul de lucru de analiză. Din motive logistice, inclusiv restricții de utilizare a datelor și licențiere pentru software, acest lucru ar putea să nu fie posibil, caz în care autorii ar trebui să furnizeze cât mai multe informații posibil pentru a permite cititorilor să recreeze analiza autorilor.

Diferențele subtile dintre pachetele software în implementarea algoritmilor complecși (de exemplu, R versus SAS) pot produce rezultate diferite, în ciuda utilizării de metode comparabile și seturi de date identice. În consecință, identificarea clară a pachetelor software și a versiunilor utilizate este de asemenea importantă pentru a sprijini reproductibilitatea cercetării. Aceasta poate include link-uri către documentația software-ului părinte, precum și pachete sau proceduri individuale de analiză. Există un sprijin din ce în ce mai mare pentru utilizarea software-ului disponibil gratuit, cum ar fi R, Python, Tensorflow sau XGBoost, pentru analiza ÎA, dar platformele software proprietare sunt încă utilizate pe scară largă și ar trebui descrise într-un mod similar cu software-ul disponibil gratuit.

Ca și în cazul tuturor analizelor, autorii trebuie să păstreze codul de execuție complet pentru a reproduce rezultatele raportate. Codul de analiză arhivat ar trebui să includă parametri de mediu, pachete de analiză, seturi de date, versiuni, apeluri la funcții de analiză, inclusiv toți parametrii și generarea de rezultate, cum ar fi tabele și figuri. Raportarea elementelor cheie și a detaliilor explicate aici ar trebui să fie încorporată în comentarii pentru a conecta o publicație la codul de analiză însoțitor și pentru a îmbunătăți reproductibilitatea. Acolo unde datele pot fi partajate liber, un cititor interesat ar trebui să poată executa întreaga analiză fără o manipulare semnificativă a codului. Acolo unde datele nu pot fi partajate în mod liber, autorii pot lua în considerare afișarea rezultatelor pașilor intermediari. Un cod minim ar trebui să fie furnizat într-o resursă accesibilă pe web, cum ar fi GitHub, Open Science Framework sau Precision Medicine Platform, cu excepția cazului în care sunt prezentate motive valabile, cum ar fi includerea proprietății intelectuale protejate. Deoarece în studiul curent sunt utilizate date de acces public [170] și majoritatea softului utilizat este de asemenea de acces liber, informația despre aspectele menționate este accesibilă (ex. <https://notebookarchive.org/2022-03-b3yfik0>).

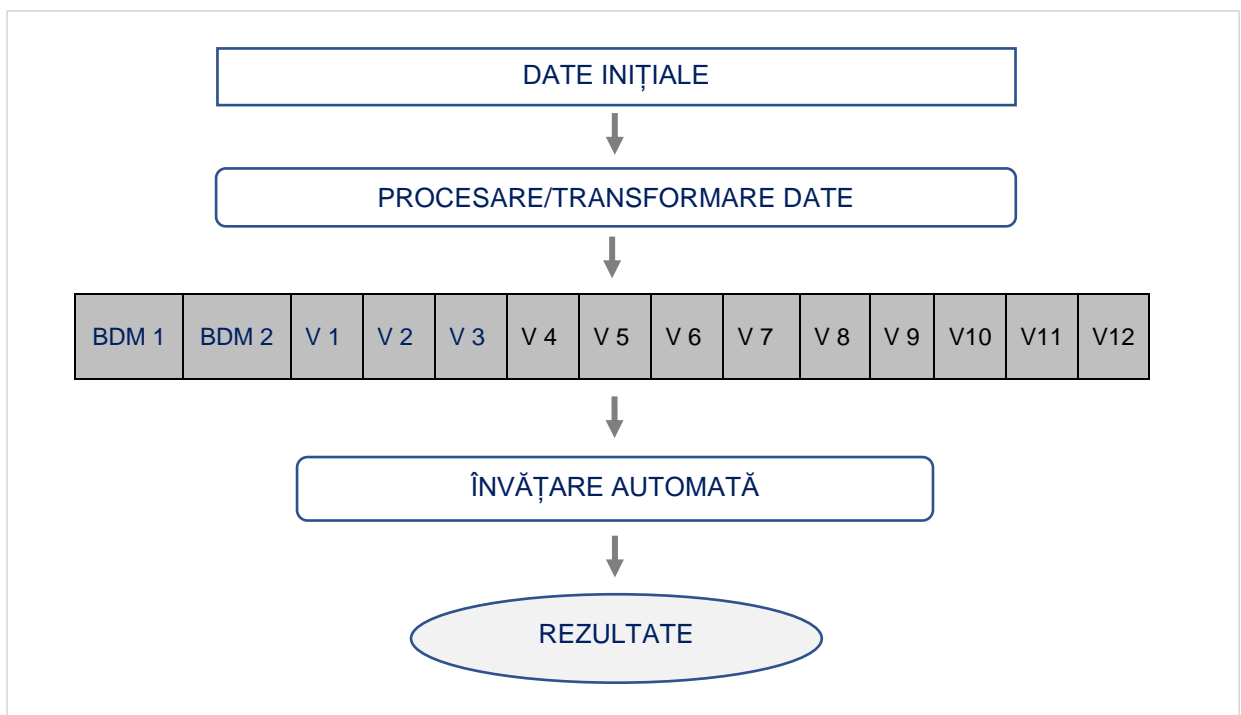


Figura 2.8. **Principalele etape ale fluxului de date în cercetarea curentă.**

Notă: Este hașurat vectorul care constă din 14 caracteristici utilizate pentru învățarea automată propriu-zisă. BDM (1,2) – complexitatea algoritmică estimată în baza celor șase parametri fiziologici selectați, V(1-12) – diferențele dintre valorile parametrilor în trei observații orare consecutive.

Indiferent de metodele ÎA utilizate în timpul analizei, rezultatele ar trebui interpretate clinic și în contextul metricilor de evaluare definite în designul studiului. Interpretarea clinică ar trebui să fie discutată cu privire la rezultatele analizei, inclusiv valorile de evaluare, în plus față de potențialul de implementare și traducere în asistența medicală. Caracteristicile cu impact mare (de exemplu, predictorii importanți sau caracteristici clinice definitorii cheie) ar trebui să fie prezentate într-un format rezumat sau tabel, împreună cu o justificare narativă pentru concentrarea asupra acestor variabile (de exemplu, dinamica temperaturii în cazul unui proces infecțios). Spre deosebire de tabelele descriptive menționate mai devreme, aceste rezultate pot susține inferența directă în ceea ce privește importanța variabilelor selectate în analiza primară.

O privire generală asupra fluxului pe care-l urmează datele în studiul curent este prezentată de Fig. 2.8, iar particularitățile procesării/transformărilor ilustrate sunt descrise în detaliu în capitolele ce urmează, unde se încearcă să se respecte principiile metodologice descrise în acest capitol.

Astfel, dimensiunea și complexitatea datelor clinice se extind rapid, necesitând utilizarea ÎA pentru analiză. Pentru a ajuta la reproductibilitate, precum și la înțelegerea și utilizarea mai largă a ÎA în comunitatea de cercetare clinică, metodele ÎA și rezultatele analizei ar trebui să fie prezentate într-o modalitate succintă, standardizată și relaționabilă pentru a facilita evaluarea de către colegi și publicul general. Multe concepte importante pentru raportarea statisticilor convenționale sunt, de asemenea, importante și în ÎA, cu câteva completări importante. Accesul la detaliile software și codul de analiză, în mod ideal într-un spațiu de lucru funcțional, disponibil publicului, sunt însoțitori esențiali pentru o lucrare cu scop de a asigura transparența și pentru a promova reproductibilitatea.

Un aspect separat în cercetare, inclusiv dezvoltarea modelelor de predicție/diagnostic, este stabilirea mărimii eșantionului utilizat pentru dezvoltarea modelului. În cazul abordării tradiționale sunt accesibile mai multe modalități de determinare a volumului eșantionului, cea mai simplă fiind produsul numărului de variabile independente și 10 [178-182]. În cazul ÎA se iau în considerație volume mai mari (începând de la câteva mii pentru un model cât de cât robust) și la moment nu sunt reguli exacte de calcul a mărimii eșantionului, reprezentând un concept aliniat la “Big data” [178]. În mod brut aceasta depinde de performanța finală a modelului. Pentru a stabili o limită inferioară a eșantionului au fost utilizate trei metode:

Metoda 1 (după [178])

Pentru un rezultat binar, un interval de încredere de 95% pentru proporția generală a rezultatului (φ) este

$$\varphi \pm 1.96 \sqrt{\frac{\hat{\varphi}(1-\hat{\varphi})}{n}}, \quad (2.8)$$

de unde marja absolută a erorii (δ) este $1.96 \sqrt{\frac{\hat{\varphi}(1-\hat{\varphi})}{n}}$. Astfel, pentru a urmări o estimare precisă a probabilității generale a rezultatului în populația țintă, pe baza proporției de rezultat anticipat ($\hat{\varphi}$) și a marjei de eroare dorite, dimensiunea eșantionului necesară este calculată ca

$$n = \left(\frac{1.96}{\delta}\right)^2 \cdot \hat{\varphi} \cdot (1 - \hat{\varphi}). \quad (2.9)$$

Pentru o marjă de eroare de 0,05 și o proporție de rezultat anticipat în populația de studiu de 0,1, este necesar un număr de 138 de participanți (adică aproximativ $138 \times 0,1 \approx 14$ participanți cu diagnosticul „sepsis”)

$$n = \left(\frac{1.96}{0,05}\right)^2 \cdot 0,1 \cdot (1 - 0,1) = 138,3 \quad (2.10)$$

Metoda 2 (după [182])

Volumul eșantionului calculat cu GPowerWin 3.1.9.7 este de 1073 pacienți, utilizând cea mai mică mărime a efectului (0,1 – care generează cel mai mare volum al eșantionului) după cum urmează:

Analysis:	A priori: Compute required sample size
Input:	Tail(s) = One
Effect size p	= 0.1
α err prob	= 0.05
Power (1- β err prob)	= 0.95
Output:	Noncentrality parameter $\delta = 3.2921701$
Critical t	= 1.6462776
Df	= 1071
Total sample size	= 1073
Actual power	= 0.9500394

Metoda 3 (dupa [178])

Calcularea dimensiunii eșantionului trebuie să vizeze o magnitudine mică a micșorării necesare a efectului predictor (pentru a minimiza potențiala supraajustare a modelului)

$$n = \frac{P}{(S-1) \cdot \ln\left(1 - \frac{R^2_{CS}}{S}\right)} = \frac{14}{(0,9-1) \cdot \ln\left(1 - \frac{0,1}{0,9}\right)} = 1189 \quad (2.11)$$

Factorul de contracție pentru S este recomandat [23] să fie $\leq 10\%$, astfel ca S să fie $\geq 0,9$. P este numărul de caracteristici (în cazul nostru = 14). R_{CS}^2 se recomandă să fie cel puțin 0,1 (cu mărirea lui volumul eșantionului se micșorează – ex. la un $R_{CS}^2 = 0,2$, mărirea eșantionului este de 558 participanți). Pentru estimarea mărimii eșantionului prin această metodă a fost utilizat pachetul „pmsampsize” (R) versiunea 1.1.2.

Volumul eșantionului în studiul curent (5039 pacienți/31599 mostre) este superior celor calculate.

Inteligența artificială (IA) este în mod obișnuit citată ca o inovație perturbatoare iminentă [183] în sectorul sănătății. Dacă este utilizată cu succes, IA are potențialul de a aborda (a) rata mare de erori medicale evitabile, (b) ineficiența fluxului de lucru și (c) ineficiența de asistență medicală modernă [184]. La etapa actuală majoritatea intervențiilor AI apropiate de traducere în practică sunt în domeniul diagnosticului medical [96].

În paradigma actuală, investigațiile diagnostice necesită interpretarea în timp util de către un clinician expert pentru a genera un diagnostic și pentru a dirija ulterior tratamentul. Cu toate acestea, problema recurentă cu sistemul actual este că serviciile de diagnosticare sunt inundate cu volume mari de muncă, care depășesc adesea capacitatea forței de muncă [185], pandemia COVID-19 fiind un caz ilustrativ. Pentru a rezolva acest lucru sunt propuși algoritmi de diagnostic în baza IA, care s-au poziționat ca dispozitive medicale care pot obține o acuratețe diagnostică comparabilă cu cea a unui clinician expert, diminuând în același timp utilizarea resurselor sistemului de sănătate. Deși această schimbare de paradigmă poate părea iminentă, este important că multe dintre dovezile care susțin algoritmi de diagnostic au fost elaborate în absența ghidurilor de raportare specifice IA. Fără această îndrumare, și într-un domeniu relativ în curs de dezvoltare, părțile cheie interesate par a fi nereușit plasate pentru a evalua calitatea și a compara acuratețea diagnosticului între studiile științifice. Încă până la emergența IA pentru a reglementa aceste aspecte a fost propus ghidul “Standards for Reporting of Diagnostic Accuracy Study” (STARD 2003), reactualizat în 2015.

Cu toate acestea, în versiunea sa actuală, STARD 2015 nu este conceput pentru a aborda problemele și provocările ridicate de modalitățile bazate pe inteligență artificială [151]. Problemele includ interpretarea metodologică neclară (de exemplu, etapele de preprocesare a datelor, alegerile de dezvoltare a modelului și utilizarea seturilor de date de validare externe), lipsa nomenclaturii standardizate (de exemplu, definiția variată a termenului „validare”), precum și utilizarea de metrice pentru raportarea rezultatelor mai puțin familiare, etc. Până când aceste probleme vor fi abordate, realizarea de evaluări cuprinzătoare ale acestor tehnologii și a potențialelor beneficii translaționale (în practică) rămâne limitată. Pentru a aborda aceste

probleme, a fost propus un ghid STARD specific IA (STARD-AI), care își propune să se concentreze pe raportarea studiilor de acuratețe a testelor de diagnostic IA [151].

Obiectivul STARD-AI este de a ajuta la raportarea cuprinzătoare a cercetărilor care utilizează tehnici de IA pentru a evalua acuratețea și performanța testelor de diagnostic. Dacă se pune accentul pe dezvoltarea, validarea sau actualizarea unui model de predicție multivariabilă care produce o probabilitate individualizată de a dezvolta o afecțiune (de exemplu, predicția timpului până la eveniment), orientările de raportare TRIPOD-AI pot fi mai adecvate [151].

De obicei, studiile de acuratețe a testelor de diagnostic compară rezultatele testelor între participanții care sunt fie cu, fie fără o afecțiune țintă. Datele de la participanții la studiu sunt supuse evaluării printr-un test de referință, care este conceput pentru a identifica o anumită condiție țintă. Estimările de performanță se bazează de obicei pe o comparație între rezultatele testului și rezultatele standard de referință din aceeași cohortă de participanți.

Un număr semnificativ de studii diagnostice în baza IA contemporane includ informații legate atât de dezvoltarea, cât și de testarea (validarea) testelor centrate pe IA. Alte subiecte cheie luate în considerare următoarele: (a) metode de preprocesare a datelor, (b) metode de dezvoltare a testelor de referință în baza IA (de exemplu, partiția setului de date, calibrarea modelului, criteriile de oprire la antrenament, utilizarea de seturi de validare externă), (c) metrici de corectitudine, (d) metrici de performanță non-standard, (e) explicabilitate, etc.

Cu toate ca cercetarea curentă este în esență un studiu preclinic, pentru ea sunt importante și aspectele/recomandările care se referă la studiile clinice, în special la etapele următoare de calibrare și dezvoltare în condiții clinice în ATI în Republica Moldova. În principiu careva dintre aceste aspecte au fost de acum descrise (ex. Fig. 2.4).

Conform [186], protocolul de studiu clinic este un document esențial elaborat de investigatorii studiului detaliind a priori rațiunea, metodele propuse și planurile pentru modul în care va fi desfășurat un studiu clinic. Acest document cheie este utilizat de către recenzori externi (agenții de finanțare, organisme de reglementare, comitete de etică a cercetării, editori de reviste, evaluatori colegi, consilii de evaluare instituționale și, din ce în ce mai mult, publicul larg) pentru a înțelege și interpreta raționamentul, rigoarea metodologică și considerațiile etice ale procesului. În plus, protocoalele de studiu oferă un punct de referință comun pentru a sprijini echipa de cercetare în realizarea unui studiu de înaltă calitate. În ciuda importanței lor, calitatea și completitudinea protocoalelor de studii publicate sunt variabile [187]. Declarația SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) a fost publicată în 2013 pentru a oferi îndrumări pentru conținutul minim de raportare al unui protocol de studiu clinic și a fost aprobată pe scară largă ca standard internațional. Declarația SPIRIT, publicată în 2013, oferă

îndrumări minime aplicabile pentru toate intervențiile din studiile clinice, dar recunoaște că anumite intervenții pot necesita extinderea sau elaborarea acestor elemente [187]. AI este un domeniu de interes enorm, cu motive puternice pentru a accelera noile intervenții până la publicare, implementare și plasarea pe piață. În timp ce sistemele IA au fost cercetate de ceva timp, progresele recente în învățarea profundă și rețelele neuronale au câștigat un interes considerabil pentru potențialul lor în aplicațiile de sănătate. Exemple de astfel de aplicații ale acestora sunt variate și includ sisteme IA pentru screening și triaj, diagnostic, prognostic, suport decizional și recomandare de tratament. Cu toate acestea, în cele mai recente cazuri, majoritatea dovezilor publicate au constat în validare in-silico, în faza timpurie. S-a recunoscut că cele mai recente studii IA sunt raportate în mod inadecvat, iar orientările existente de raportare nu acoperă pe deplin sursele potențiale de părtinire specifice sistemelor IA [188].

SPIRIT-AI (ca parte a inițiativei SPIRIT-AI și CONSORT-AI) este o inițiativă internațională susținută de SPIRIT și de Rețeaua EQUATOR (Îmbunătățirea calității și transparenței cercetării în domeniul sănătății) pentru a extinde declarația existentă SPIRIT 2013 acolo unde este necesar, pentru a dezvolta ghiduri specifice AI bazate pe consens. Ea este complementară declarației CONSORT-AI, care își propune să promoveze raportarea de înaltă calitate a studiilor IA. Această declarație de consens descrie metodele utilizate pentru a identifica și evalua elementele candidate și pentru a obține consens. În plus, oferă și lista completă de verificare SPIRIT-AI, inclusiv elemente noi și explicațiile care le însoțesc.

Extensia SPIRIT-AI [186] recomandă ca, împreună cu articolele SPIRIT 2013 existente, să fie abordate 15 articole (12 extensii și 3 elaborări) pentru protocoalele de testare ale intervențiilor IA. Aceste elemente au fost considerate suficient de importante pentru protocoalele de studii clinice pentru intervențiile IA, încât ar trebui raportate în mod obișnuit în plus față de elementele de bază ale listei de verificare SPIRIT 2013. Astfel, această extensie explică o serie de termeni relevanți (inteligentă artificială, învățare automată, intervenție în baza IA, ajustare fină a modelului, medii de dezvoltare și operațional, eroare de performanță, etc.) și propune o listă de verificare, cu elemente importante pentru aspectul IA.

CONSORT-AI (ca parte a inițiativei SPIRIT-AI & CONSORT AI) [189] este o inițiativă internațională susținută de CONSORT și de Rețeaua EQUATOR pentru a evalua declarația existentă CONSORT 2010 și a extinde sau a elabora acest ghid, acolo unde este necesar, pentru a sprijini raportarea studiilor clinice, pentru intervenții IA [190]. Ea este complementară declarației SPIRIT-AI, care își propune să promoveze raportarea de înaltă calitate a protocolului pentru studiile IA.

Cercetarea curentă tinde să urmeze recomandările, care sunt deja definite. Spre exemplu, descriind detaliat tehnicile de preprocesare a datelor (în capitolul următor), folosirea scorului SOFA în calitate de test de referință (utilizat la etichetarea cazurilor în baza de date inițială), descrierea principiilor de partiție a datelor, folosirea unui set separat de date pentru validare externă, descrierea metricilor pentru evaluarea performanței, etc. Atenție specială este acordată explicabilității modelelor create în cadrul studiului (ex. diagramele de pondere a predictorilor, diagramele de dependență parțială, diagramele SHAP, etc.). Toate acestea sunt considerate în capitolele ce urmează, în special, când sunt prezentate rezultatele.

Datorită importanței cruciale a metricilor care sunt utilizate pentru evaluarea performanței produsului finit – al sistemului elaborat și deoarece cercetarea curentă folosește o îmbinare de asemenea metrici care vin atât din domeniul ÎA, cât și din statistica tradițională, aceste metrici sunt descrise mai detaliat în continuare:

- *Sensibilitate (Recall/TPR)* = $\frac{TP}{(TP+FN)}$,

unde *TPR* – true positive rate (rata adevărat pozitivă), *TP* – true positive (rezultate adevărat pozitive), *FN* – false negative (rezultate fals negative);

- *Specificitate (True Negative Rate)* = $\frac{TN}{(TN+FP)}$,

unde *TNR* – true negative rate (rata adevărat negativă), *TN* – true negative (rezultate adevărat negative), *FP* – false positive (rezultate adevărat pozitive);

- *Acuratete diagnostica* = $\frac{(TP+TN)}{(TP+TN+FP+FN)}$,

unde *TP* – true positive (rezultate adevărat pozitive), *TN* – true negative (rezultate adevărat negative), *FN* – false negative (rezultate fals negative), *FP* – false positive (rezultate adevărat pozitive);

- *Precizia (Rata predictivă pozitivă)* = $\frac{TP}{(TP+FP)}$,

unde *TP* – true positive (rezultate adevărat pozitive), *FP* – false positive (rezultate adevărat pozitive);

- *Rata predictivă negativă* = $\frac{TN}{(TN+FN)}$,

unde *TN* – true negative (rezultate adevărat negative), *FN* – false negative (rezultate fals negative);

- *Rata fals pozitivă (False Positive Rate)* = $\frac{FP}{(FP+TN)}$,

unde FP – false positive (rezultate adevărat pozitive), TN – true negative (rezultate adevărat negative);

- *Rata fals negativă (False Negative Rate)* = $\frac{FN}{(FN+TP)}$,

unde FN – false negative (rezultate fals negative), TP – true positive (rezultate adevărat pozitive);

- *Raport de probabilitate pozitiv (Positive likelihood ratio (LR +))* = $\frac{TPR}{FPR}$,

unde TPR – true positive rate (rata adevărat pozitivă), FPR – false positive rate (rata fals pozitivă);

- *Raport de probabilitate negativ (Negative likelihood ratio (LR -))* = $\frac{FNR}{TNR}$,

unde FNR – false negative rate (rata fals negativă), TNR – true negative rate (rata adevărat negativă);

- *Raportul șanselor (DOR)* = $\frac{LR+}{LR-}$,

unde DOR - Diagnostic odds ratio, $LR+$ - raport de probabilitate pozitiv, $LR-$ - raport de probabilitate negativ;

- *Analiza ROC (ROC - Receiver operating characteristic – reprezentare grafică a curbei ROC, ca raport dintre FPR (rata fals pozitivă) pe axa X și TPR (rata adevărat pozitivă) pe axa Y , unde FPR (rata fals pozitivă) = $FP/(FP + TN)$, și TPR (rata adevărat pozitivă) = $TP/(TP + FN)$, inclusiv aria de sub curba ROC ;*

- *Analiza PR (PR - Precision-recall – reprezentare grafică a curbei PR, ca raport dintre sensibilitate/recall pe axa X și valoarea predictivă pozitivă/precizie pe axa Y , care este o metodă de elecție în cazul seturilor de date dezechilibrate;*

- *Matricea de confuzie (cu indicarea pentru fiecare clasă a numărului de cazuri prezise corect și incorect, eroarea și rata erorii);*

- *Analiza importanței variabilelor (identificarea și prezentarea grafică a celor mai importanți pentru model predictor);*

- *Rata și traiectoria diminuării erorii (reprezentarea grafică a dinamicii erorii în cadrul învățării/training-ului modelului);*

Un rol special în procesarea datelor îl au metodele de estimare a complexității algoritmice, care vin din domeniul dinamicii informației algoritmice. Ele sunt în detaliu descrise în capitolul următor.

2.4. Sinteza Capitolului 2

Capitolul 2 începe cu prezentarea designului studiului (Fig. 2.1), care în cazul cercetării în baza IA poate avea elemente specifice. Datorită faptului că multe metode de învățare automată au origini în literatura statistică, au apărut două culturi de modele de predicție:

- (a) cele din mediul statistic/epidemiologic și
- (b) cele din informatică/știința datelor;

Astfel, au apărut două abordări diferite, inclusiv și la nivel de terminologie. O analiză comparativă a terminologiei IA versus modelarea statistică tradițională este prezentată în Tab. 2.1.

Setul inițial de date pentru cercetare include 40336 cazuri/pacienți, dintre care 2932 sunt cu sepsis. Datele inițiale reprezintă serii temporale cu o frecvență orară de eșantionare pentru 8 parametri vitali (ex. frecvența cardiacă, presiunile arteriale sistolică și diastolică, saturația sângelui periferic cu O₂, temperatura corpului, fracția O₂ inspirat, etc) și 36 parametri de laborator și logistica. Pentru cercetarea ulterioară au fost selectați 6 parametri vitali (frecvența cardiacă, presiunile arteriale sistolică și diastolică, saturația sângelui periferic cu O₂, frecvența respirației și temperatura corpului).

Criteriile de includere sunt descrise printr-o diagramă de tip CONSORT (Fig. 2.5). În final au fost selectate 5039 cazuri (care au livrat 30635 mostre, fiecare incluzând valorile celor șase parametri pentru trei ore consecutive).

Un aspect important este scindarea setului de date în: (a) date pentru trainingul modelului – în cazul studiului curent 27583 mostre și (b) setul test - 3052 mostre. Aceste seturi nu se intersectează pe tot parcursul studiului până la momentul testării finale a modelului elaborat.

Datorită cerințelor specifice pentru așa studii s-a efectuat extrapolarea principiilor de management a studiului, inclusiv raportarea rezultatelor TRIPOD-AI, PROBAST-AI, CONSORT-AI, SPIRIT-AI, STARD-AI.

Algoritmul de bază utilizat la elaborarea modelului final este GBM (Gradient Boosting Machine), care este descris prin prisma formalismului matematic.

Sunt descrise cerințele față de dezvoltarea modelului și validarea lui (Fig. 2.6) și particularitățile fluxului de date la diferite etape ale cercetării (Fig. 2.8).

În final sunt descrise metricile utilizate pentru evaluarea diferitor aspecte ale performanței modelului (15 la număr)

3. ETAPELE DE PROCESARE A DATELOR

Pentru a fi utilizate pentru învățarea automată (ÎA) datele necesită o prelucrare specială, care este descrisă mai jos. Una din etapele incipiente ale acestui proces include scindarea datelor inițiale în setul pentru training și setului de date pentru testare. Prelucrarea datelor se face separat pentru fiecare din aceste seturi. Primul set este ulterior folosit pentru elaborarea modelului de ÎA, iar cel de al doilea - pentru validarea lui. Similar cu cercetarea tradițională, în cea bazată pe ÎA, atât volumul cât și calitatea datelor utilizate are o importanță majoră. Gestionarea aspectelor de calitate a datelor prezente se începe, când este cazul, cu restabilirea valorilor lipsă, etapă descrisă în continuare.

3.1. Reconstruirea datelor prin rezolvarea problemei valorilor lipsă

După cum a fost menționat anterior, în datele inițiale unele valori ale parametrilor lipsesc (Tabelul 3.1.).

Tabelul 3.1. Aspectul inițial al datelor unui pacient (p000009)

FCC	SpO2	Temp	PAS	PAM	PAD	Resp	EtCO ₂	...	Vârsta	Sex	...
NA	NA	NA	NA	NA	NA	NA	NA	...	27.92	1	...
117	99	NA	116	97	81	20	NA	...	27.92	1	...
NA	NA	NA	NA	NA	NA	NA	NA	...	27.92	1	...
NA	NA	NA	NA	NA	NA	NA	NA	...	27.92	1	...
NA	NA	NA	NA	NA	NA	NA	NA	...	27.92	1	...
NA	NA	NA	NA	NA	NA	NA	NA	...	27.92	1	...
120	100	36	118	84	64	30	NA	...	27.92	1	...
...

Notă: FCC – frecvența contracțiilor cardiace, SpO₂ – saturația cu O₂ a sângelui periferic, PAS/PAD/PAM – presiunea arterială sistolică/diastolică/medie, Resp – frecvența respirației, EtCO₂ – indice capnometric, NA (not available) – valori lipsă. Cu “...” sunt marcate coloanele și rândurile omise. Acesta este un fragment al setului de date ale pacientului respectiv ce include primele 7 ferestre orare (rânduri). Setul complet are 258 rânduri (adică ferestre orare/observații) și 40 coloane (parametrii monitorizați).

Datele lipsă sunt adesea inevitabile în cercetare, dar potențialul lor de a influența rezultatele cercetării este rareori discutat. În acest sens, se consideră importantă natura „lipsei”: la întâmplare versus nu la întâmplare.

Din păcate, destul de des nu este posibil să se facă distincția între lipsa la întâmplare și lipsa nu la întâmplare folosind datele observate. Întrucât setul de date utilizat în acest studiu nu este însoțit de o explicație cu privire la natura lipsei, se presupune că acestea lipsesc la întâmplare (de

exemplu, în timpul în care pacientul este supus unei proceduri medicale care necesită senzorii utilizați pentru colectarea datelor să fie îndepărtați, funcționarea defectuoasă a echipamentului, etc.).

Deși recuperarea datelor poate fi posibilă independent de natura lipsei, distorsiunea introdusă de procedura de recuperare se consideră a fi mai mică în cazul lipsei de date la întâmplare [191].

Există multe metode pentru a trata valorile lipsă din date. Cele mai multe dintre aceste metode sunt adecvate pentru date statice, iar setul de instrumente potrivite pentru date continue, sau serii temporale, este de dimensiuni mai mici. Cea mai simplă abordare este de a șterge observațiile cu valori lipsă, cunoscută sub numele de analiză completă a cazului (CCA) [192], dar în multe cazuri acest lucru poate fi greu de aplicat, mai ales când volumul de date disponibile este limitat. Această problemă prezintă un interes deosebit atunci când se utilizează datele în scopuri de învățare automată, mai ales când se folosesc seturi dezechilibrate, unde fiecare observație din clasa minoritară este importantă.

Metodele de recuperare a datelor pot fi împărțite în mod convențional în:

- Metode de restabilire unică - înlocuirea unui punct de date lipsă cu o singură valoare, care provine de obicei din valorile observate de la același subiect (ultima observație efectuată înainte (last observation carried forward, LOCF), observație de referință efectuată înainte (base observation carried forward, BOCF) și următoarea observație efectuată înapoi (next observation carried backward, NOCB)) sau din alte surse (de exemplu, prin valorile medii, restabilirea cu ajutorul regresiei, etc.);
- Metode de restabilire multiplă - prin crearea mai multor seturi diferite de date generate plauzibil și prin combinarea adecvată a rezultatelor obținute din fiecare dintre ele. Există o serie de pachete statistice în acest scop (de exemplu, MICE, Amelia, HMISC în R, etc.)

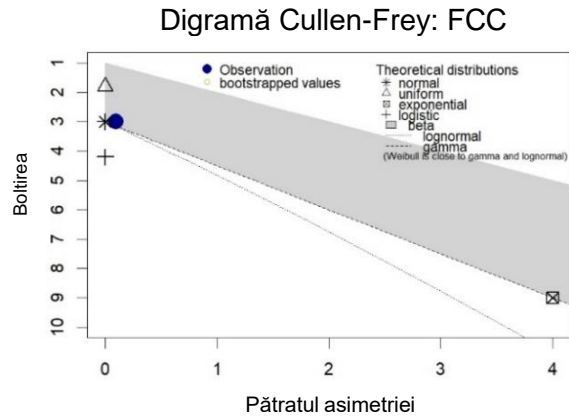
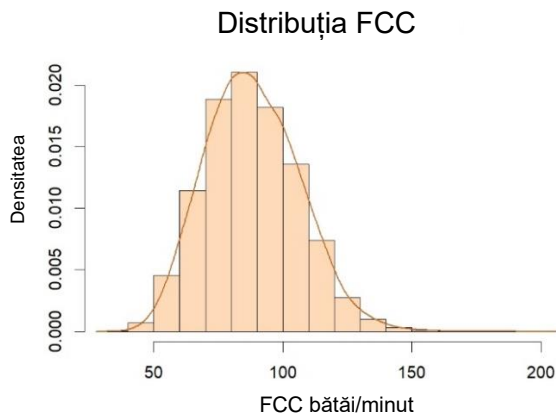
Conform Ghidului privind datele lipsă în studiile clinice de confirmare [193], *„dacă valorile lipsă sunt tratate prin simpla excludere a oricăror pacienți cu valori lipsă din analiză, aceasta va duce la o reducere a numărului de cazuri disponibile pentru analiză și, prin urmare, în mod normal, rezultă o reducere a puterii statistice. În mod clar, cu cât este mai mare numărul de valori lipsă, cu atât este mai mare probabilitatea reducerii puterii. Prin urmare, ar trebui depuse toate eforturile pentru a minimiza cantitatea de date lipsă [și pentru a selecta o metodă de restabilire adecvată]. Din păcate, nu există o abordare metodologică pentru gestionarea valorilor lipsă care să fie universal acceptată în toate situațiile”*.

Ca atare, modul în care se reduce la minimum cantitatea de date lipsă și modul în care datele lipsă vor fi tratate în analiză sunt aspecte critice care trebuie luate în considerare atunci când se planifică un proces.

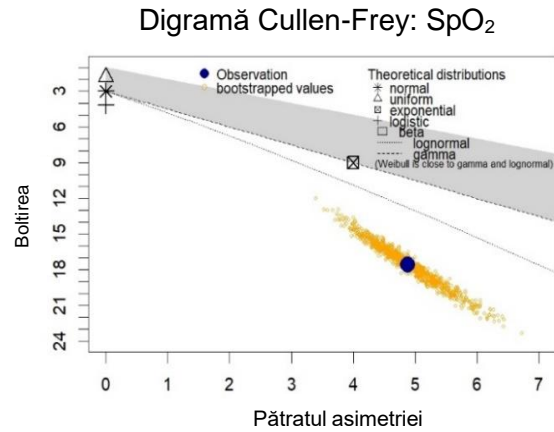
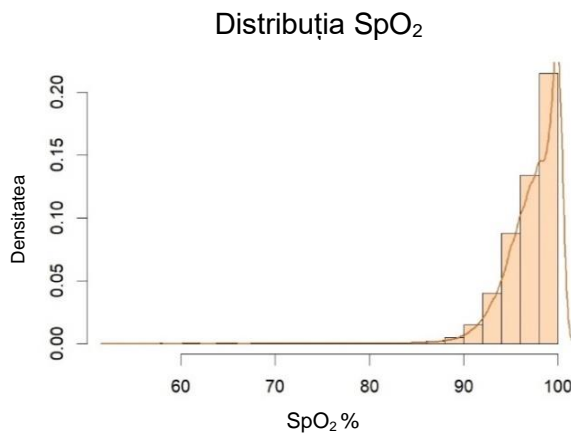
În majoritatea lucrărilor ce au de a face cu date-lipsă în seturi de date similare [34, 154] valoarea lor se restabilește prin metoda “last observation carried forward” (LOCF), care presupune sortarea setului de date (ex. valorile unui parametru/coloană – vezi mai sus) și crearea unui set ordonat. Ulterior se determină prima valoare lipsă și se utilizează valoarea precedentă pentru a o “restabili”. Procesul se repetă pentru următoarea valoare lipsă până când toate valorile lipsă sunt “restabilite”. Considerând procesele fiziologice drept fenomene cu caracter continuu (de obicei, reprezentate prin serii temporale – acesta este și cazul datelor studiate) se poate presupune că și șirurile numerice create din valori ale parametrilor care reprezintă aceste procese moștenesc proprietatea de “continuitate” (contra variantei discrete sau aleatorii). Metoda LOCF cumva ignoră “principiul de continuitate”. De aceea s-a decis de a include principiul de “continuitate”, propunând următoarea procedură: (1) fiecare coloană este evaluată în privința datelor lipsă (NA); (2) prima și ultima valoare (pe rânduri) este “restabilită” conform valorii din cea mai apropiată celulă; (3) în final se efectuează interpolarea, cu calcularea valorilor între două valori prezente cu respectarea tendinței (creștere sau descreștere). Pentru interpolare este utilizată funcția `{na.approx()}` (pachetul “zoo” în R) [194, 195]. În cazul când toate valorile lipsesc s-a procedat în felul următor: (1) din setul integral s-au extras valorile prezente pentru fiecare parametru separat pe clase (ex. pacienți septici din setul A). (2) au fost extrase valorile prezente (“non-missing”) și s-a determinat distribuția datelor pentru fiecare parametru de interes (ex. frecvența cardiacă (HR), presiunea arterială sistolică (SBP), etc). Au fost identificate caracteristicile acestor distribuții și în baza lor au fost generate distribuții similare utilizând pachetul “sn” în R [196]. (3) Valorile lipsă au fost restabilite în acest caz cu valori extrase din distribuțiile generate.

Deoarece unul din pașii descriși mai sus presupune utilizarea datelor generate, s-a considerat oportun de a acorda atenție specială calității acestor date prin investigarea inițială a caracteristicilor datelor originale, ca mai apoi ele să fie translate datelor generate. Pentru determinarea tipului de distribuție a datelor au fost utilizate metode grafice (histograma cu curba densității și diagramele Cullen-Frey – Fig. 3.1) și statistice (în primul rând pentru evaluarea normalității distribuției și estimarea caracteristicilor distribuției: media, deviația standard, mediana, asimetria, boltirea, etc. – Tab. 3.2).

A) Frecvența cardiacă (FCC)



B) Saturația sângelui periferic cu oxigen (SpO₂)



C) Presiunea arterială sistolică (PAS)

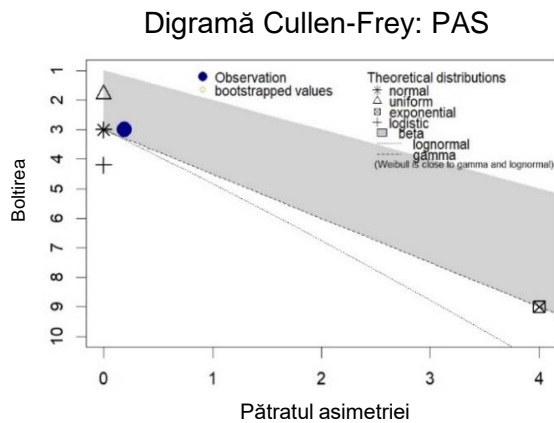
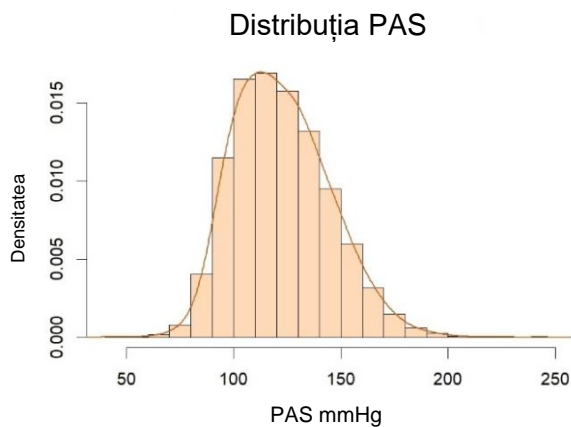
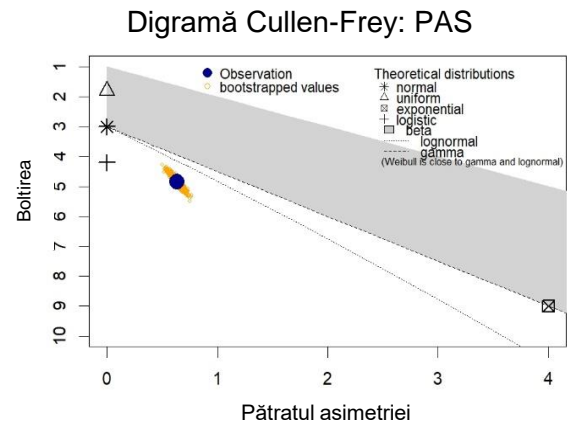
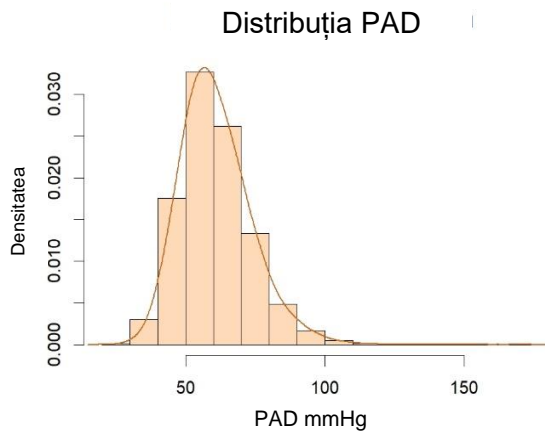


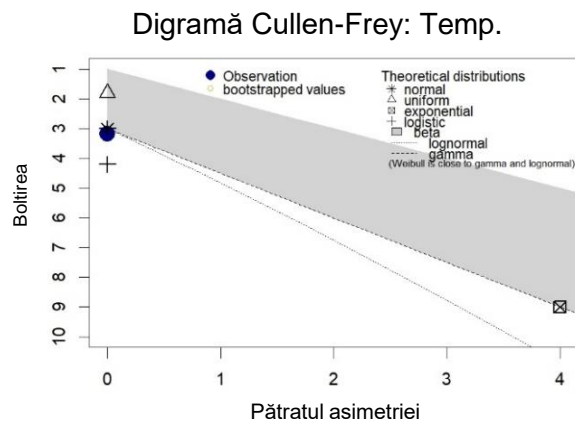
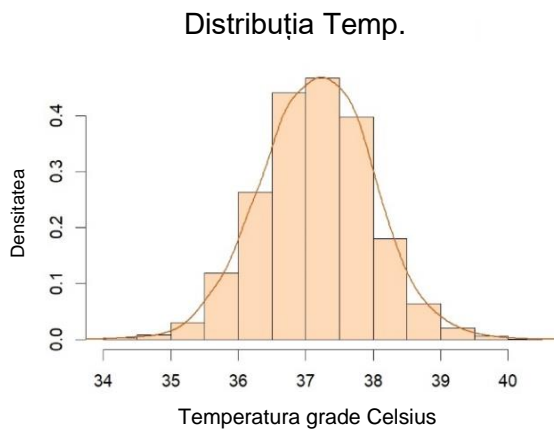
Figura 3.1. Reprezentarea grafică a distribuției valorilor parametrilor de interes.

Notă: Pe stânga – histogramele și curbele densității; pe dreapta – diagramele Cullen-Frey (datele observate sunt reprezentate de cerculețul de culoare albastră, datele obținute prin bootstrapping – cu culoare galbenă)

D) Presiunea arterială diastolică (PAD)



E) Temperatura (Temp.)



F) Frecvența respirației (Resp.)

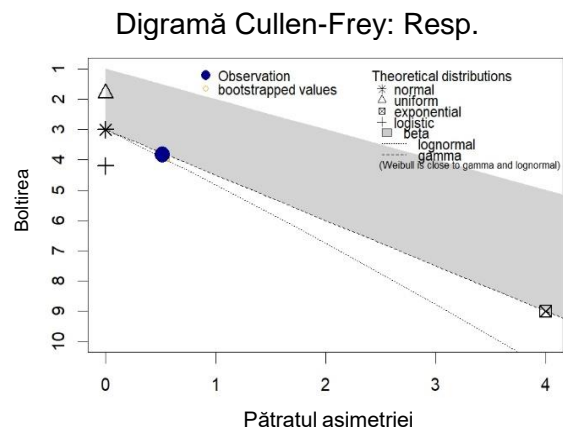
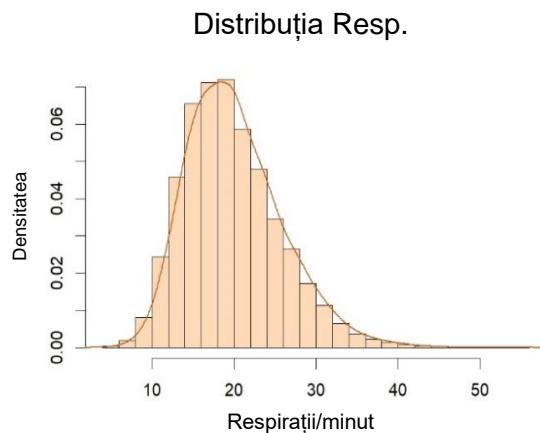


Figura 3.1 (continuare). **Reprezentarea grafică a distribuției valorilor parametrilor de interes**

Notă: Pe stânga – histogramele și curbele densității; pe dreapta – diagramele Cullen-Frey (datele observate sunt reprezentate de cerculețul de culoare albastră, datele obținute prin bootstrapping – cu culoare galbenă)

Tabelul 3.2. Caracteristicile distribuției și normalitatea datelor (pentru fiecare variabilă)

Variabilă	Caracteristicile distribuției	Testul Shapiro - Wilk		Testul Kolmogorov - Smirnov		Comentariu (în baza graficelor Cullen-Frey și testelor)
		W	valoare p	D	valoare p	
Frecvența contracțiilor cardiace (FCC)	min: 37.5; max: 200 mediana: 87 media: 88.34028 dev. standard: 18.074 asimetria: 0.3088 boltirea: 3.001856	0.99309	5.65e-13	1	< 2.2e-16	Distribuție mai apropiată de normală
Saturația sângelui periferic cu O ₂ (SpO ₂)	min: 53; max: 100 mediana: 98 media: 97.40984 dev. standard: 2.712 asimetria: -2.209675 boltirea: 17.56789	0.8391	< 2.2e-16	1	< 2.2e-16	Distribuție de tip lognormală
Presiunea Arterială sistolică (PAS)	min: 41; max: 245 mediana: 120.5 media: 122.4947 dev. standard: 26.067 asimetria: 0.4379288 boltirea: 2.9993063	0.98518	< 2.2e-16	1	< 2.2e-16	Distribuție apropiată de normală
Presiunea Arterială diastolică (PAD)	min: 22; max: 173 mediana: 59.5 media: 60.87633 dev. standard: 12.905 asimetria: 0.7970014 boltirea: 4.83886	0.97034	< 2.2e-16	1	< 2.2e-16	Distribuție apropiată de lognormală
Temperatura (Temp.)	min: 34; max: 40.5 mediana: 37.17 media: 37.1695 dev. standard: 0.819 asimetria: -0.0094542 boltirea: 3.184942	0.9984	5.922e-05	1	< 2.2e-16	Distribuție mai aproape de normală
Frecvența Respirației (Resp.)	min: 5; max: 55 mediana: 19.5 media: 20.13853 dev. standard: 5.821 asimetria: 0.7194058 boltirea: 3.824321	0.96689	< 2.2e-16	1	< 2.2e-16	Distribuție apropiată de lognormală

Unele histograme în figurile 3.1 și 3.2 ar indica aparent la distribuții aproape de normale (coloana de figuri din stânga), însă diagramele Cullen-Frey indică la normalitate mai certă

(aproape fără asimetrie) numai în cazul parametrului “Temperatura” (cerculețul albastru, care reprezintă datele observate, se suprapune pe steluță, care indică la o distribuție normală), alte distribuții fiind aproape de log normale sau normale. Sumarul analizei este ilustrat de Tabelul 3.2, unde sunt prezentate și rezultatele testelor Kolmogorov-Smirnov (cu un eșantion) și Shapiro – Wilk. Conform rezultatelor testelor (W/D aproape de 1 (sau 1) și valorile $p \ll 0.05$) practic nici una dintre distribuții nu este „strict” normală.

Astfel, schema algoritmului utilizat pentru reconstrucția datelor este următoarea [197]:

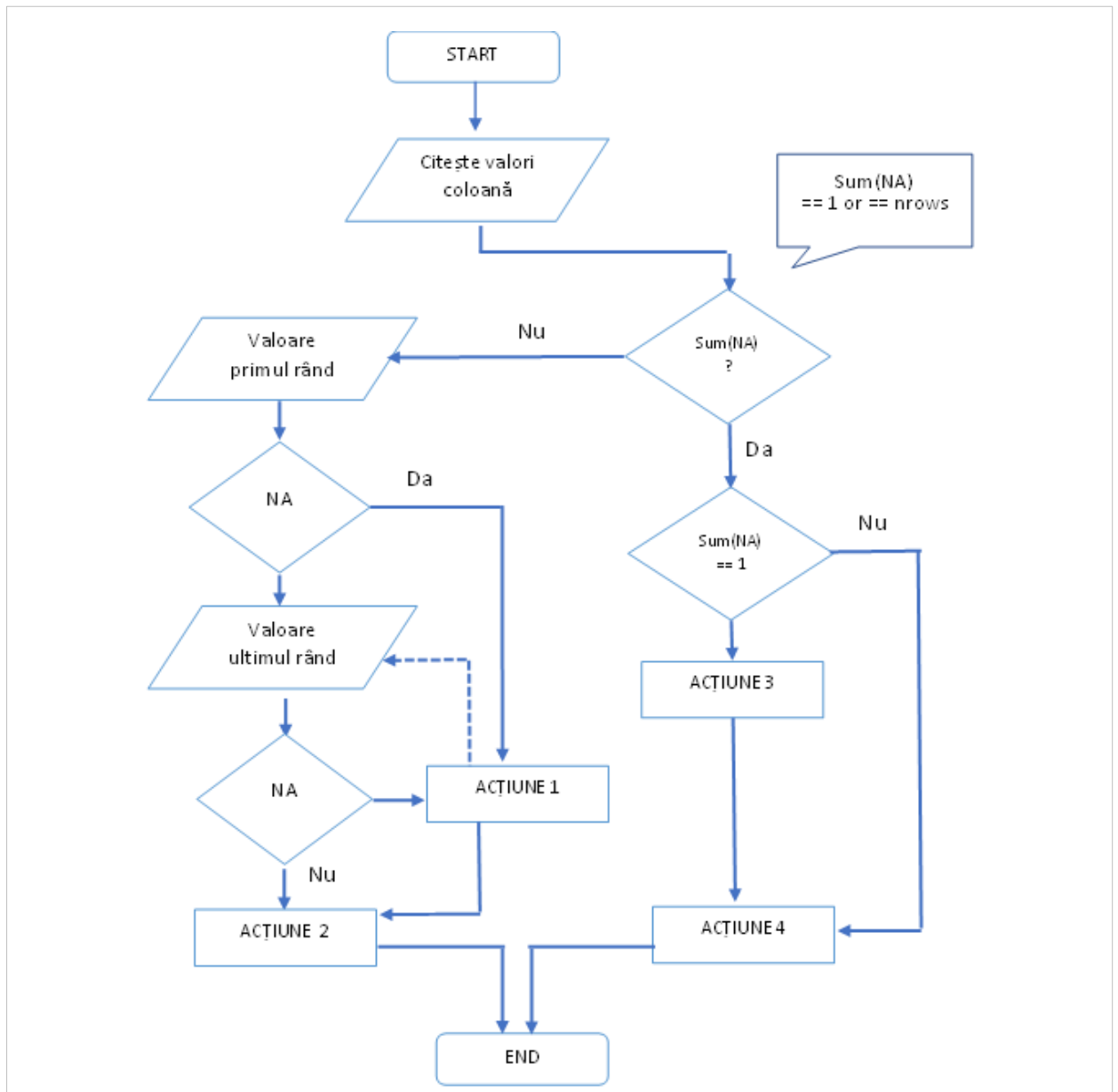


Figura 3.2. Algoritm pentru reconstrucția datelor în cazul lipsei valorilor.

Notă: Acțiunea 1 – înlocuirea valorii-lipsă cu cea mai apropiată valoare (locație); Acțiunea 2 – interpolarea valorilor lipsă prin procedura descrisă în text; Acțiune 3 – ștergerea valorii; Acțiune 4 – umplerea celulelor cu valori din distribuțiile generate.

Algoritmul descris a fost aplicat pe datele din setul A (setul pentru training si cel pentru testare) și setul B. După aplicarea procedurii de reconstrucție datele capătă următorul aspect:

Tabelul 3.3. Aspectul datelor aceluiași pacient (p000009) după reconstrucție

FCC	SpO ₂	Temp	PAS	PAD	Resp
117	99	36	116	81	20
117	99	36	116	81	20
118	99	36	116	78	22
119	99	36	117	74	24
119	100	36	117	71	26
120	100	36	118	67	28
120	100	36	118	64	30

Notă: FCC – frecvența contracțiilor cardiace, SpO₂ – saturația cu O₂ a sângelui periferic, PAS/PAD – presiunea arterială sistolică/diastolică, Resp – frecvența respirației. Sunt prezentate numai valorile parametrilor (în coloane) supuși reconstrucției conform procedurii descrise. Setul de date este cel descris anterior in Tabelul 3.1.

3.2. Complexitatea algoritmică Kolmogorov-Chaitin ca metrică și metodă de procesare a datelor

Complexitatea algoritmică (CA) este abordată in cadrul acestui studiu [198] prin prisma dinamicii informației algoritmice (AID, algorithmic information dynamics). AID [199] este un domeniu emergent al științei complexității bazat pe teoria informației algoritmice (AIT, algorithmic information theory), care cuprinde literatura de specialitate bazata pe conceptul de complexitate Kolmogorov–Chaitin și concepte conexe precum probabilitatea algoritmică, compresia, inferența optimă, distribuția universală și altele. Centrală pentru AIT este definiția complexității algoritmice (complexitatea Kolmogorov–Chaitin sau a mărimii programului - Kolmogorov, 1965; Chaitin, 1969) [200]:

$$K_T(s) = \{ |p|, x \mid T(p) = s \}, \quad (3.1)$$

adică lungimea celui mai scurt program p care generează șirul s și care rulează pe o mașină Turing universală T . AID tinde să identifice soluții la întrebări fundamentale despre cauzalitate: de ce un anumit ansamblu de circumstanțe duce la un anumit rezultat. În acest aspect CA, diferă în esență de statisticile tradiționale.

Ca știință aplicată, AID este un nou tip de calcul discret bazat pe “computer science” și care vizează studierea cauzalității prin generarea de modele mecanice care să ajute la găsirea primelor principii ale fenomenelor fizice, construind următoarea generație de învățare automată.

În setul de instrumente AID, există un instrument special pentru furnizarea de estimări fiabile pentru funcțiile non-computabile, și anume calculatorul on-line al complexității algoritmice (OACC) [201], care oferă posibilitatea estimării CA și a probabilității algoritmice (PA) pentru șiruri numerice scurte și lungi și pentru structuri bidimensionale mai bine decât oricare alt instrument tradițional. Metodele tradiționale în esența lor nu sunt concepute pentru a capta careva conținut algoritmic dincolo de tiparele statistice simple.

OACC utilizează metoda BDM (block decomposition method) [199, 201], care se bazează pe probabilitatea algoritmică definită prin metoda teoremei de codificare (CTM) [199]:

$$BDM = \sum_{i=1}^n CTM(block_i) + \log_2(|block_i|). \quad (3.2)$$

OACC este disponibil ca versiune online [200], precum și pachete autonome în R și o serie de alte limbaje de programare (Matlab, C, Wolfram, etc.) și este utilizat pentru calculele respective pentru scopul lucrării curente. Deoarece intensitatea acestor calcule este destul de înaltă (datorită volumului de date și însăși esența calculelor) utilizarea OACC în forma descrisă este insuficient de productivă. Pentru accelerarea acestor calcule a fost extras nucleul pachetului respectiv [202], care poate fi aplicat pe structuri bidimensionale (matrice) și integrat în programul care efectuează procesarea fluxului de date în cercetarea curentă. Procesul calculării valorii CA/BDM este ilustrat de Figura 3.3.

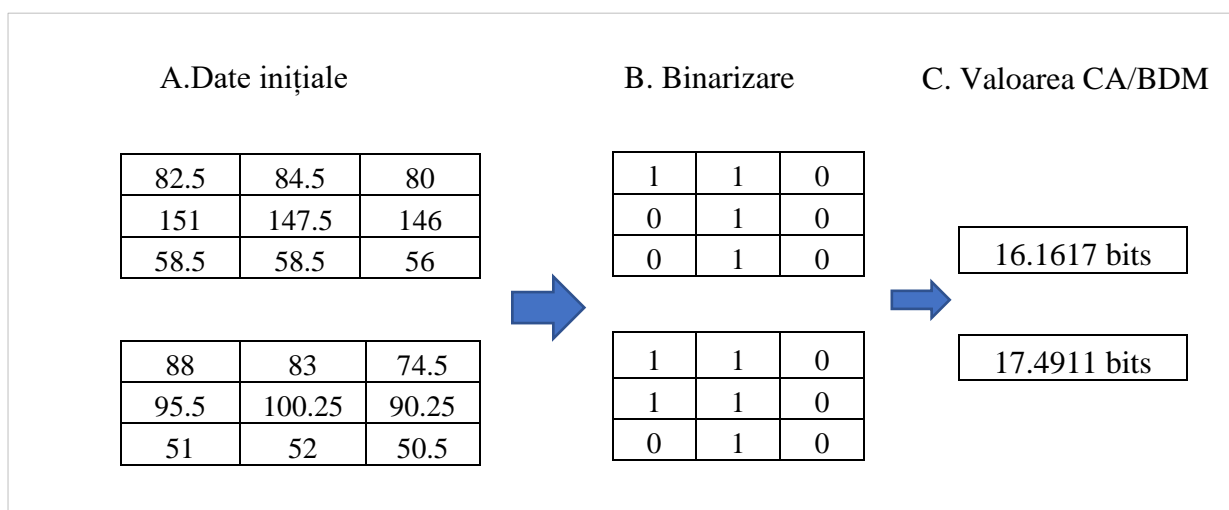


Figura. 3.3. **Formatul datelor la 3 etape succesive de prelucrare a lor.**

Notă: A- formatul inițial, B – binarizare (după valori-prag pe rânduri), C- valoarea CA calculată pentru fiecare matrice. Datele inițiale reprezintă valori ale parametrilor circulatori (FCC, PAS, PAD), structurate în matrice cu dimensiunile 3 x 3 (pe rânduri).

3.3. Procesarea datelor cu crearea caracteristicilor pentru a fi utilizate în sistemul de învățare automată

Pregătirea datelor pentru a fi livrate unui model de învățare automată este o etapă crucială, de care în ultimă instanță poate depinde performanța modelului. În afară de unele principii generale (pentru evitarea scurgerii de date, părtinire, etc.), la moment nu sunt metode universale prestabilite pentru acest proces. De obicei, forma cea mai reușită se stabilește prin încercări. În cazul curent s-ar părea că o oarecare importanță ar avea-o relevanța patofiziologică a unui parametru anumit, însă această abordare s-a dovedit a fi mai puțin productivă. Au fost încercate o serie de variante.

La început au fost utilizați parametri care ar avea importanță în diagnosticarea clinică a sepsisului, și în special cei utilizați pentru SOFA. Această abordare a dat rezultate minore (aria de sub curba ROC, $AUC = 0.61$). Apoi s-a încercat setul integral de parametri clinici (36 parametri), ce a condiționat o performanță a modelului departe de ideală ($AUC = 0.65$). Cei mai reușiți s-au dovedit a fi 6 parametri și dinamica lor sub formă de diferența a valorilor orare timp de 3 ore. La selecția formatului datelor utilizate pentru învățarea automată (trainingul modelului final) a contribuit și principiul “de rând cu valoarea la moment a unui parametru fiziologic, de importanță majoră, deseori, este și dinamica lui în timp”. De asemenea s-a tins spre utilizare în calitate de predictor a unor indici/parametri care se înregistrează cu ușurință/standard, inclusiv și în secțiile ATI din Republica Moldova.

În general, regulile pentru selecția caracteristicilor/variabilelor au fost mai în detaliu elaborate pentru modelele de regresie, care până nu demult erau prevalente în cercetarea biomedicală [204, 205]. Recent, parțial condiționat de dezvoltarea ÎA, au început să fie utilizate o gamă mai largă de modele [205] (ex. bazate pe învățarea profundă, abordarea Bayesiană, arbori decizionali, etc). În cazul modelelor de regresie abordarea cea mai des utilizată este prin analiza corelațională cu estimarea coeficienților de corelare, a puterii corelării, inclusiv corelarea parțială. Astfel, se selectează variabilele cu cea mai evidentă corelare.

Cu toate că în cazul ÎA în ultimă instanță sistemul/algorithmul de învățare automată selectează cele mai importante variabile, de exemplu în cazul arborilor decizionali ea este pusă în vârful arborelui și celelalte variabile – în descendența importanței, contribuind la formarea ramurilor, rămâne totuși important care variabile sunt inițial livrate algoritmului.

În cazul cercetării curente, reieșind din caracterul datelor primare, s-a decis de utilizat variabilele cu cele mai puține valori lipsă. Deoarece una din metodele de transformare a datelor este BDM (pe tensori bidimensionali/matrice – varianta optimă selectată – 2 asemenea matrice (3 x 3) și un interval de timp optim pentru estimarea prezicerii – 3 ore, a fost necesară selectarea a 6 asemenea variabile (cate trei pentru fiecare matrice x 3 ore). Astfel, selecția finală a

variabilelor/parametrilor fiziologici include: frecvența cardiacă (FCC), saturația sângelui periferic cu oxigen (SpO₂), presiunea arterială sistolică (PAS) și diastolică (PAD), temperatura corpului (Temp), frecvența respirației (Resp) – în total 6 la număr.

Cu toate ca cercetarea curentă are ca scop crearea în final a unui sistem cu abilități discriminative (sepsis vs non-sepsis) și algoritmul principal utilizat este GBM, analiza rezultatelor la anumite etape a inclus și proceduri de regresie (în special la compararea performanței diferitor algoritmi de ÎA). Ținând cont ca aceste rezultate ar putea furniza unele perspective pentru o înțelegere mai profundă a stării de sepsis ca fenomen medical, rezultatele analizei corelaționale a celor șase parametri fiziologici de interes și vârsta sunt prezentate mai jos.

Pentru analiza corelațională au fost utilizate datele inițiale separate în 2 clase – “Sepsis” și “Non-sepsis” (până la crearea seturilor pentru training și testare), folosind numai cazurile complete (după excluderea valorilor lipsă). Au fost analizate 4 seturi de date cu cazuri complete: (a) setul integral parvenit de la pacienții septici; (b) setul integral non-septic; (c) setul septic, care include numai observațiile când sepsisul este confirmat în baza criteriilor Sepsis-3 și (d) un set numit “Pre-septic”, care include observațiile timp de până la 10 ore în raport cu prima oră, când sepsisul a fost confirmat prin metode tradiționale/Sepsis-3.

A) Sepsis (setul septic integral, cazuri complete)

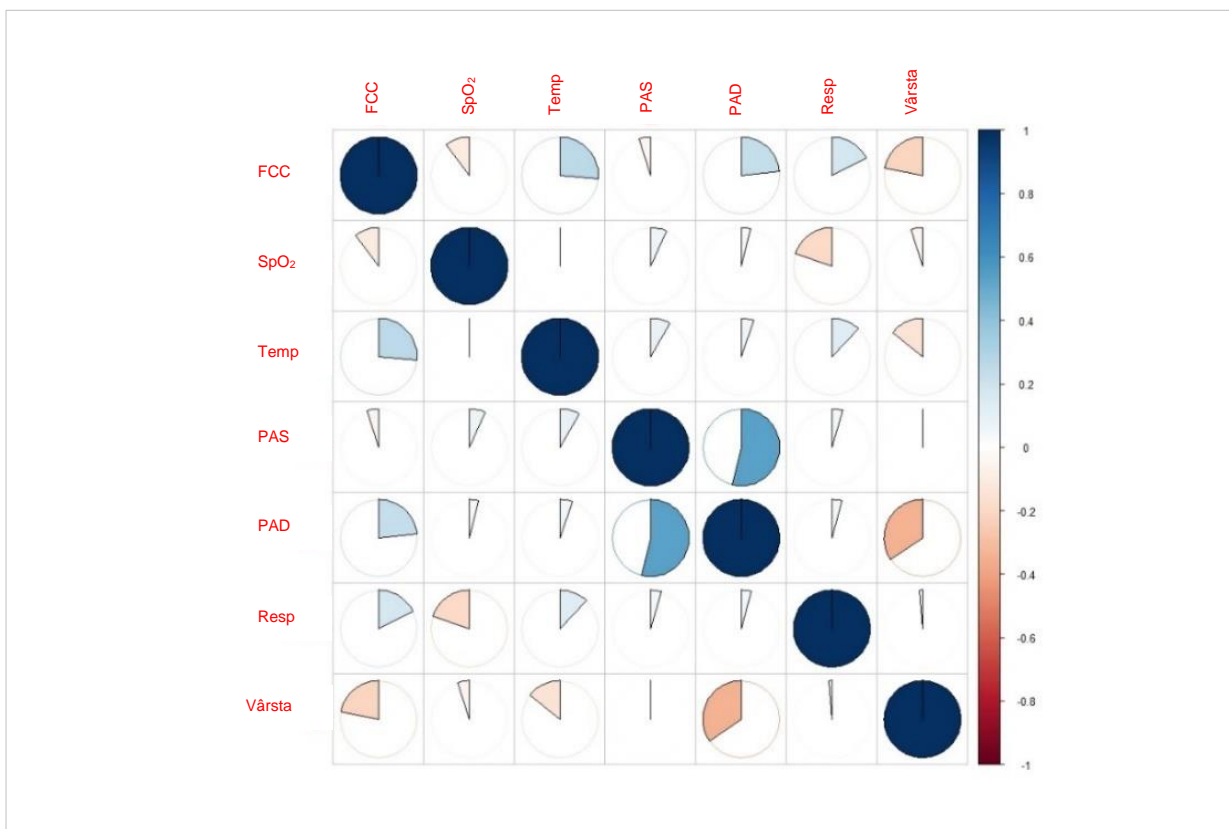


Figura. 3.4. Corelarea parametrilor fiziologici de interes în sepsis.

Tabelul 3.4. Matricea de corelare a parametrilor fiziologici în sepsis

	FCC	SpO ₂	Temp	PAS	Pad	Resp	Vârsta
FCC	1.00	-.10	.26	-.05	.23	.17	-.22
SpO ₂	-.10	1.00	.00	.07	.04	-.20	-.05
Temp	.26	.00	1.00	.08	.05	.12	-.14
PAS	-.05	.07	.08	1.00	.54	.05	.00
PAD	.23	.04	.05	.54	1.00	.04	-.34
Resp	.17	-.20	.12	.05	.04	1.00	-.01
Vârsta	-.22	-.05	-.14	.00	-.34	-.01	1.00

B) Non – sepsis (setul non-septic integral, cazuri complete)

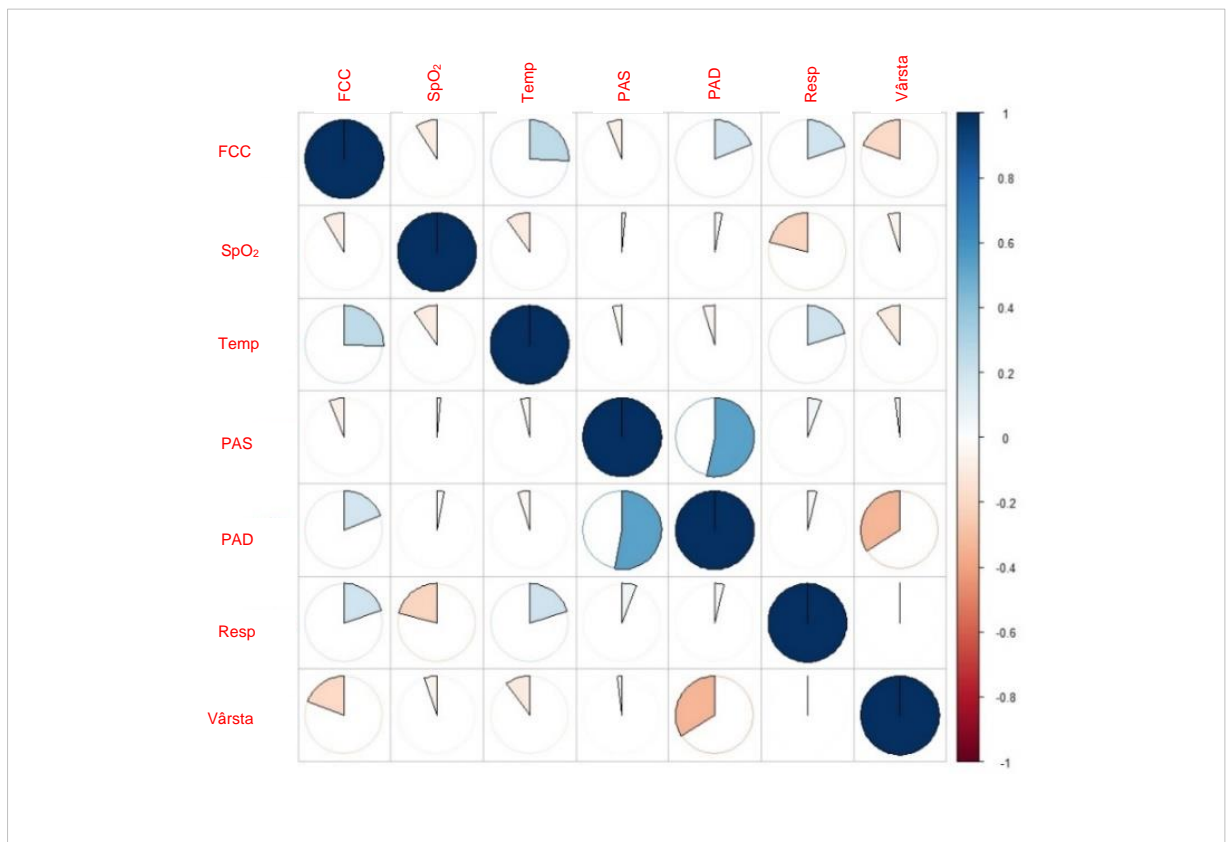


Figura 3.5. Corelarea parametrilor fiziologici de interes in cazurile non-septice.

Tabelul 3.5. **Matricea de corelare a parametrilor fiziologici la non-septici**

	FCC	SpO ₂	Temp	PAS	PAD	Resp	Vârsta
FCC	1.00	-.09	.26	-.06	.19	.20	-.19
SpO ₂	-.09	1.00	-.10	.01	.03	-.21	-.05
Temp	.26	-.10	1.00	-.04	-.05	.20	-.10
PAS	-.06	.01	-.04	1.00	.53	.06	-.02
PAD	.19	.03	-.05	.53	1.00	.04	-.34
Resp	.20	-.21	.20	.06	.04	1.00	-.01
Vârsta	-.19	-.05	-.10	-.02	-.34	-.01	1.00

După cum se poate observa din figurile și tabelele de mai sus, în starea de sepsis vs stările non-septice corelarea dintre parametri este diferită, în unele cazuri – diferă atât magnitudinea, cât și direcția/semnul corelării. Spre exemplu, în stările non-septice corelarea dintre temperatură și presiunea arterială sistolică (SBP) se traduce printr-un coeficient de corelare de -0.04. Însă în sepsis observăm un coeficient de corelare de 0.08 (modificare de 2 ori a mărimii coeficientului și inversia semnului). Corelarea dintre temperatură și presiunea diastolică este însoțită de inversia semnului, iar cea dintre temperatură și frecvența respirației (Resp) – modificarea magnitudinii, etc.

În diferită măsură diferențele sunt valabile pentru majoritatea coeficienților. O situație similară se observă și la pacienții din setul B. Informație mai detaliată pentru acest set este prezentată în Anexa 3.

Au fost, de asemenea, analizate aceleași aspecte referitor la seturile (c) și (d). În Anexa 4 sunt descrise rezultatele analizei corelaționale pentru aceste seturi. Modificările coeficienților sunt prezente, însă modeste.

Deosebirea dintre seturile de date de la pacienți septici vs pacienți non-septici este evaluată și cu ajutorul softului mai specializat – pachetul “semiArtificial” în R [206], care permite de a determina similaritatea dintre seturi. Rezultatele sunt prezentate în Anexele 5 și 6. În acest caz disimilaritate evidentă, statistic semnificativă (prin prisma testului Kolmogorov-Smirnov și distanței Hellinger) între seturi nu a fost identificată, ce ar presupune o potențială dificultate a problemei de discriminare între seturi, cel puțin prin metode liniare, spre exemplu, cu ajutorul metodelor de regresie.

3.4. Sinteza Capitolului 3

În acest capitol sunt descrise etapele de procesare a datelor, începând cu analiza exploratorie a lor. Analiza datelor a inclus diferite aspecte statistice, inclusiv evaluarea distribuției/normalității pentru fiecare parametru din cei șase selectați la etapa precedentă (Fig. 3.1 și Tab. 3.2).

Deoarece datele inițiale conțin valori-lipsă a fost necesară reconstrucția lor cu restabilirea valorilor-lipsă. Pentru aceasta a fost elaborat un algoritm de restabilire (Fig. 3.2.), care a fost aplicat pe datele din studiu. Rezultatul aplicării este exemplificat de Tab. 3.3.

O metodă-cheie de reprezentare și procesare a datelor în cadrul studiului curent este prin prisma complexității algoritmice Kolmogorov – Chaitin, tehnica și rezultatul aplicării acestei abordări fiind ilustrat de Fig. 3.3.

Astfel, formatul final al datelor utilizate pentru învățarea automată reprezintă vectori cu lungimea 14 (șiruri numerice alcătuite din 14 valori), unde primele două valori reprezintă complexitatea algoritmică a 2 matrice (obținute prin gruparea celor 6 parametri de interes, câte 3 observații orare consecutive – convențional matricea respirator-metabolică și matricea circulatorie), iar celelalte 12 valori ale vectorului reprezintă dinamica valorii fiecăruia din cei 6 parametri timp de 3 ore, de asemenea, consecutive [207].

Una din etapele de procesare și analiză a datelor a inclus analiza corelațională a parametrilor de interes în setul de cazuri non-septice versus cazurile septice. Această analiză a relevat diferențe între coeficienții de corelare și magnitudinea lor între seturi, ce s-ar alinia cu conceptul de sepsis ca insuficiență multiorganică – varietate a stării critice, care se caracterizează prin ineficacitatea mecanismelor de autoreglare și control, aspect fiind discutat în următoarele capitole ale tezei.

4. ÎNVĂȚAREA AUTOMATĂ. REZULTATE ȘI DISCUȚIA LOR

4.1. Etapa de învățare automată

Procesul de creare a unui sistem de învățare automată (SÎA) este mai mult sau mai puțin standard. Figurile care urmează descriu componentele și etapele procesului (reprodus cu modificări din [208]).

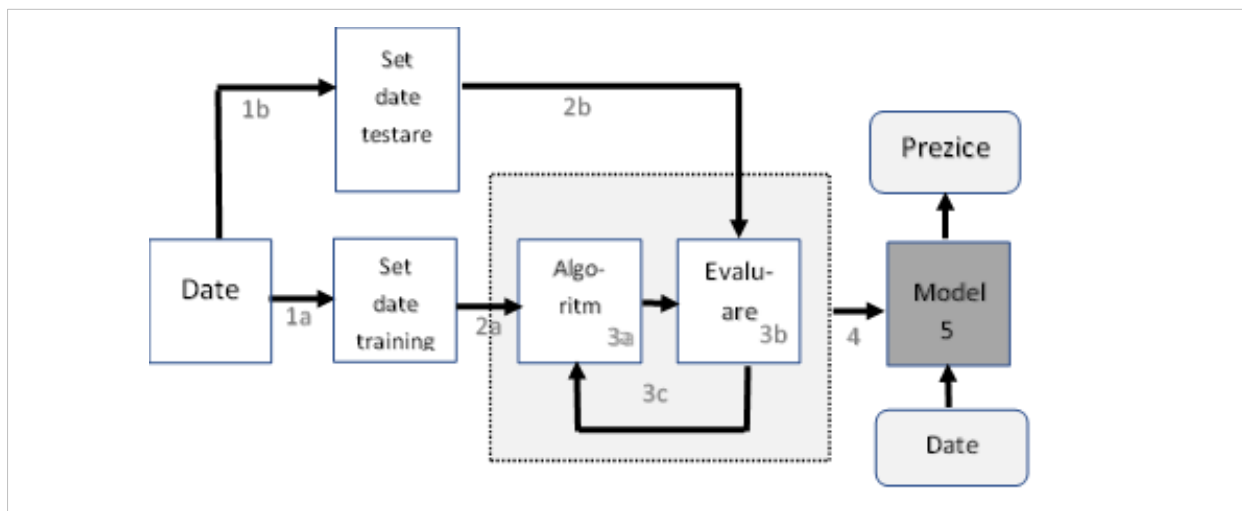


Figura 4.1. Etapele și elementele principale în crearea unui sistem de învățare automată: 1a, 1b, --, 5 – etapele creării sistemului

Etapele procesării datelor sunt aliniate procesului de creare a unui sistem de învățare automată și sunt schematic ilustrate în figura 4.2.

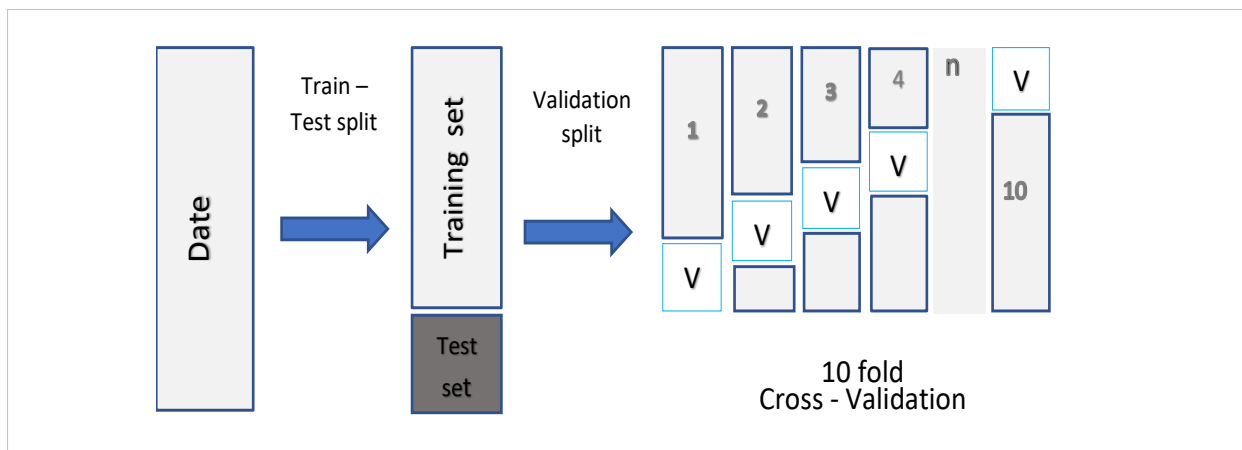


Figura 4.2. Etapele fluxului de procesare a datelor la crearea unui SÎA.

Notă: Inițial datele sunt împărțite în setul pentru training (care ulterior va participa la învățare și cross-validare) și setul – test.

Un moment important este separarea de la început (după procesarea preliminară a datelor) a setului pentru testare, care în cazul modelului de bază în studiul curent este format din 3162 mostre,

numite în continuare specimene (1173 sepsis și 1989 non-sepsis), care nu au participat la training și validare.

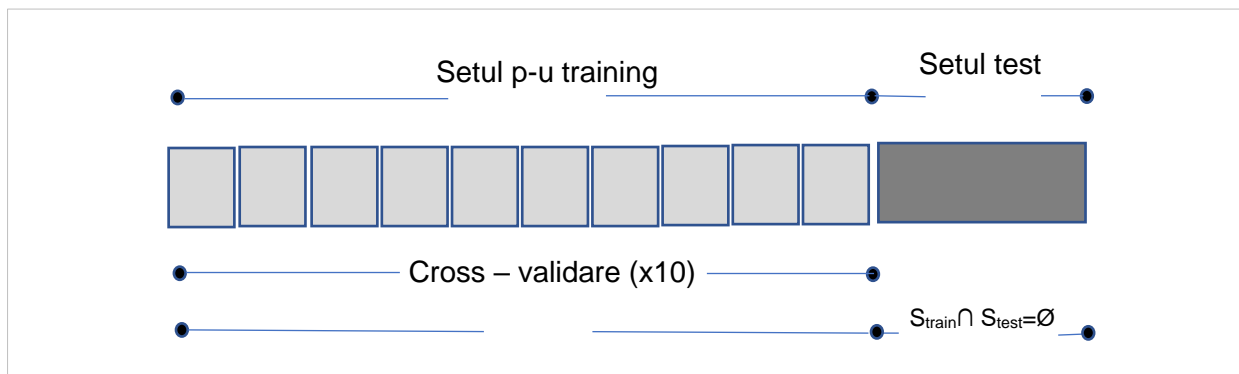


Figura 4.3. Relația dintre setul de testare și celelalte date (p-u training și validare)

Ca punct de orientare pentru crearea sistemului de prezicere a sepsisului în cadrul cercetării curente a fost sistemul InSight, raportat în literatura de specialitate ca fiind unul dintre cele mai performante în domeniu [34, 154]. Folosind informația destul de generală din [34] acest sistem a fost replicat în limbajul de programare R și testată performanța lui pe setul de date din studiul curent. Au fost obținute rezultate ce corespund celor raportate de autori. În continuare, aceste rezultate au servit drept jaloane în crearea și evaluarea sistemului propriu. În tabelul 4.1 este prezentată metrica comparativă a performanței acestor două sisteme la testarea lor pe setul-test.

Tabelul 4.1. **Metrica comparativă a două sisteme de prezicere a sepsisului (InSight vs sistemul creat în cadrul cercetării curente)**

Metrica	Sistemul InSight	Sistemul propriu
AUC	0.914 (95% IÎ: 0.902 – 0.926)	0.926 (95% IÎ: 0.919 – 0.939)
Acuratețe diagnostică	0.925	0.927
Sensibilitate (Recall)	0.888	0.924
Specificitate (Specificity)	0.941	0.929

Notă: Metricile utilizate sunt tradiționale pentru așa sisteme. AUC – aria de sub curba ROC: rata fals pozitivă (pe axa X) vs sensibilitate (pe axa Y). Sunt indicate limitele pentru un interval de încredere de 95%.

Performanța acestor două sisteme este destul de apropiată. O sensibilitate mai înaltă în cazul sistemului creat ar indica la un sistem mai echilibrat, care puțin mai bine detectează cazurile cu sepsis, decât sistemul InSight. Sistemul creat diferă de InSight după datele inițiale folosite pentru

învățarea automată, metodele de procesare a datelor și formatul datelor finale, care în ultimă instanță sunt livrate pentru învățare și validare. Spre exemplu, datele finale în cazul sistemului creat reprezintă vectori cu lungimea 14 (șiruri numerice alcătuite din 14 numere/valori) comparativ cu 30, în cazul InSight, adică o reducere a dimensionalității mai mult de 2 ori. Aspectul datelor finale folosite la crearea sistemului este ilustrat în tabelul 4.2.

Tabelul 4.2. **Formatul final al datelor transmise algoritmului de învățare automată**

Cat	BDM-MR	BDM-CIRC	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
0	16.956	17.491	-4.5	1.5	0	-1.5	-2.5	-3	2	-1.5	0.28	-3.5	0	0.5
0	18.328	14.815	-8.5	0	0.5	-10	-1.5	0	-5	0	0.1	4.75	1	0
1	18.456	15.942	1	0	0.2	28	10	9	4	-1	0.1	-13	-3	0
0	17.855	16.900	6	0	0.2	-15	-6.5	-0.5	-16	0	0.2	21	8.5	0.5
1	16.956	16.956	3	-1	0	-5	-2	0	2	-1	0.1	4	1	0

Notă: Acestea sunt datele finale care descriu 5 specimene: 2 cu sepsis și 3 – cu alte patologii. “CAT”- este marcajul fiecărui specimen (1 – sepsis, 0 – alte patologii); “BDM-MR/BDM-CIRC” – reprezintă valorile BDM pentru grupurile de date respirator-metabolice și cele care se referă la sistemul circulator; “V1 – V12” reprezintă dinamica celor 6 parametri fiziologici (FCC, SpO₂, Temp, PAS, PAD, Resp) pe o perioadă de 3 ore.

Datele finale organizate în formatul descris au fost în ultimă instanță livrate unui model în baza algoritmului GBM (Gradient Boosting Machine - algoritm în baza arborilor decizionali), care a efectuat predicția prin clasificare (sepsis vs non-sepsis). Hiperparametrii optimi (rata învățării, numărul de arbori, profunzimea, nr. de ramificații) au fost identificați inițial prin învățare automată automatizată (AutoML) utilizând platforma H2O [209]. Acești hiperparametri au fost apoi justificați în contextul datelor prezente printr-o căutare grilă grosieră cu validare multiplă (10 folds cross-validation).

Conform terminologiei statistice tradiționale sistemul creat este un model de predicție multivariabil (cu derivatele a 6 predictorii) în baza clasificării binomiale. Conform principiilor PROBAST (2019) [160] – este un model de predicție prognostică (unde în calitate de factori de prognozare servesc 6 parametri fiziologici și dinamica lor), iar conform ghidului TRIPOD (2015) [141] – cercetarea actuală este de “Tip 3”, ce presupune dezvoltarea unui model predictiv cu utilizarea unui set de date și evaluarea performanței lui utilizând date separate, ultima fiind numită și “validare externă”.

4.2. Elaborarea, validarea și testarea sistemului de ÎA creat

Sistemul a fost creat conform principiilor și urmând pașii descriși anterior. S-a efectuat validarea lui (10 folds cross-validation) și testarea pe 2 seturi de date: setul-test din setul A și setul B integral. Ultimul a fost prelucrat folosind aceleași metode de procesare a datelor.

S-a experimentat cu mai mulți algoritmi, inclusiv GLM (Generalized Linear Model), bazat pe regresia liniară, cel mai performant model din această clasă demonstrând o performanță de 70.3% după aria de sub curba ROC (Anexele 7 și 10). Cea mai înaltă performanță a avut-o algoritmul GBM, în baza arborilor decizionali. Descrierea ce urmează se referă la acest algoritm.

Tabelul 4.3. Performanța sistemului în baza algoritmului GBM

Metrică performanță	Cross-validare (x10)	Setul test-1/ validare internă	Setul test-2/ validare externă
<i>Sensibilitate/</i> Recall/True Positive Rate (TPR)	0.885 (0.016)	0.916 (0.898 – 0.932)	0.713 (0.701 – 0.725)
<i>Specificitate/</i> True Negative Rate (TNR)	0.951 (0.009)	0.942 (0.931 – 0.952)	0.913 (0.908 – 0.918)
<i>Acuratețe diagnostică</i>	0.928 (0.006)	0.940 (0.931 – 0.948)	0.902 (0.897 – 0.906)
<i>Valoarea predictivă pozitivă/</i> Positive predictive value (PPV)	0.915 (0.91-0.92)	0.605 (0.561 – 0.647)	0.332 (0.320 – 0.344)
<i>Valoare predictivă negativă/</i> Negative predictive value (NPV)	0.931 (0.93-0.93)	0.992 (0.987 – 0.993)	0.981 (0.981 – 0.982)
<i>Rata fals pozitivă/</i> False Positive Rate (FPR)	0.048	0.058	0.087
<i>Rata fals negativă/</i> False Negative Rate (FNR)	0.121	0.084	0.287
<i>Raport de probabilitate pozitiv/</i> Positive likelihood ratio (LR+)	18.31 (17.1-19.6)	15.85 (13.26 – 18.94)	8.20 (7.76 – 8.66)
<i>Raport de probabilitate negativ/</i> Negative likelihood ratio (LR-)	0.13 (0.12-0.13)	0.09 (0.07 – 0.11)	0.31 (0.30 – 0.33)
<i>Raportul șanselor/</i> Diagnostic odds ratio (DOR)	140.85	176.11	26.45
<i>Analiza ROC (AUC)</i>	0.970 (0.004)	0.929 (95% ÎĪ:0.919 – 0.939)	0.813 (95% ÎĪ: 0.807 - 0,820)
<i>Analiza PR (prAUC)</i>	0.959 (0.004)	0.864	0.647

Notă: Metricile pentru cross-validare reprezintă valorile medii și deviația standard (în paranteze), majoritatea furnizate de platforma H2O (cu deviația standard), unele calculate suplimentar (cu ÎĪ). Pentru validarea internă și externă în paranteze sunt prezentate valorile pentru un interval de încredere de 95% și sunt calculate cu ajutorul MedCalc [206] și pachetului “pROC” în R [207]. Setul B (pentru validare externă) nu conține cazurile cu parametri pentru care toate valorile lipsesc. Prevalența sepsisului pentru setul A este 8.8%, iar pentru setul B – 5.71%.

Setul de date de bază utilizat pentru trainingul modelului a inclus 27583 specimene (sepsis – 9636, non-sepsis – 17947). Pentru setul test (A) au fost rezervate 3052 specimene (sepsis – 1063, non-sepsis – 1989). Setul B a fost utilizat pentru generarea celui de al doilea set-test – pentru validare externă, incluzând 20432 specimene (sepsis – 5383 și non-sepsis – 15049).

Performanța sistemului (la cross-validare, testare pe primul set-test (validare internă) și la validare externă – pe cel de al doilea set-test) este descrisă în tabelul 4.3.

Curba PR are pe axa X sensibilitatea și pe axa Y – precizia/valoarea predictivă pozitivă. Acest tip de analiză este utilă în cazurile în care există un dezechilibru în observațiile între cele două clase. Mai exact, există multe exemple de non sepsis (clasa 0) și mai puține de sepsis (clasa 1). Se consideră că această abordare furnizează un tablou mai realist când este necesar de focusat pe abilitatea sistemului de a distinge anume cazurile din clasa minoritară (sepsis).

Pentru estimarea robusteții sistemului în privința degradării performanței lui pe măsura diminuării dimensiunilor setului pentru training, ultimul a fost gradual diminuat de la 100% (setul pentru training inițial) la 25%. Diminuarea dimensiunilor datelor s-a efectuat prin eșantionare aleatorie din speciemenle inițiale. Setul-test a rămas același. Tabelul ce urmează conferă detalii în această privință.

Tabelul 4.4. Robustețea modelului prin dependența de volumul datelor utilizate pentru training și calitatea lor

Model	Sensibilitate	Specificitate	Acuratețe	AUC (95% ÎÎ)
Model cu 100% date	0.916	0.942	0.940	0.929 (0.919 – 0.939)
Model cu 75% date	0.907	0.939	0.928	0.923 (0.913 – 0.934)
Model cu 50% date	0.908	0.934	0.925	0.921 (0.911 – 0.931)
Model cu 25% date	0.894	0.929	0.917	0.912 (0.901 – 0.923)
Model antrenat pe date unde toate coloanele cu toate valorile lipsă au fost înlăturate	0.921	0.926	0.925	0.924 (0.914 – 0.935)

După cum se vede din tabel diminuarea dimensiunilor setului pentru training duce la o diminuare graduală, însă neînsemnată a performanței (de la 92,6% la 91,2% după AUC). Ultimul rând al tabelului descrie performanța unui model creat separat în baza datelor inițiale din care au fost excluse subseturile/cazurile cu toate datele lipsă în cel puțin o coloană. Aceasta poate furniza informație despre potențiala părtinire, care ar putea fi introdusă în cazul reconstrucției datelor în

coloanele cu toate datele lipsă. Deoarece performanța acestui model este destul de aproape de cea a modelului de bază (AUC de 92,4% vs 92,9%) s-a presupus ca părănirea daca și este prezentă, este minoră.

În continuare sunt prezentate rezultatele de relevanță obținute la diferite etape de training, validare și testare a modelului final.

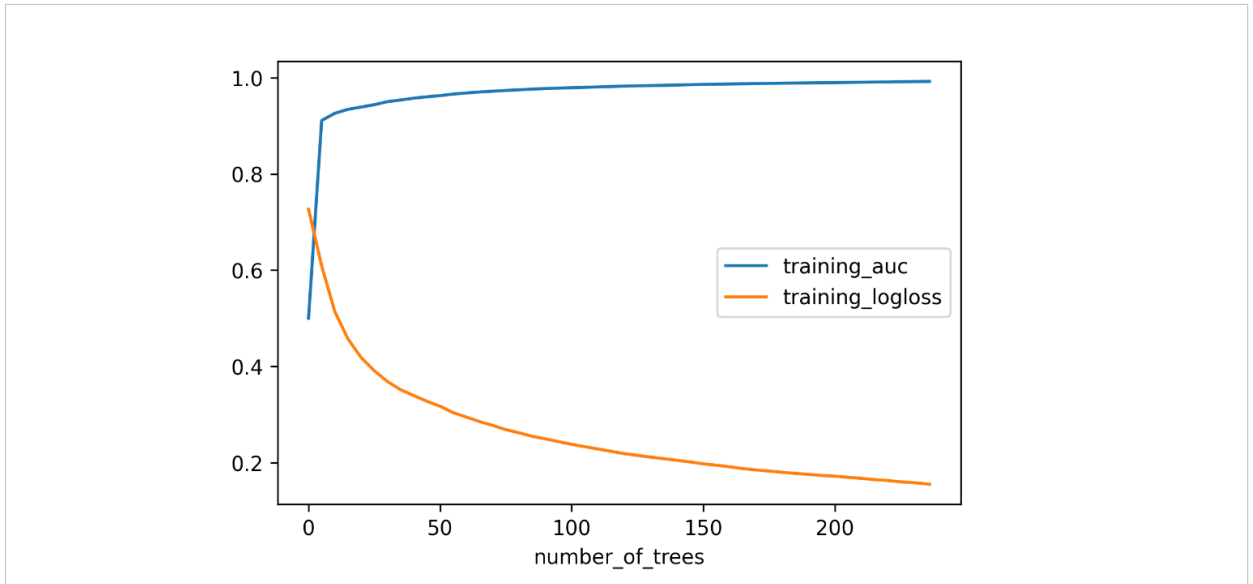


Figura 4.4. Diminuarea erorii (logloss - curba orange) și performanța (AUC – curba albastră) în procesul trainingului (în dependență de numărul de arbori)

Tabelul 4.5. Matricea de confuzie la validarea multiplă (x 10) a modelului

	0	1	Eroare	Rata erorii
0	17085	862	0.048030	=862/17947
1	1164	8472	0.120797	=1164/9636
Total	18249	9334	0.073451	=2026/27583

Nota: Pe verticală sunt prezentate clasele actuale (1-sepsis, 0 – non-sepsis), la intersecție – clasele prezise). Ultimele două coloane prezintă eroarea și rata erorii.

Tabelul 4.6. Matricea de confuzie la testarea modelului pe setul 1 (A)

	0	1	Eroare	Rata erorii
0	1874	89	0.045339	=89/1963
1	115	974	0.105601	=115/1089
Total	1989	1063	0.066841	=204/3052

Tabelul 4.7. Matricea de confuzie la testarea modelului pe setul 2 (B)

	0	1	Eroare	Rata erorii
0	13740	1544	0.101021	=1544/15284
1	1309	3838	0.254323	=1309/5147
Total	15049	5382	0.139682	=2854/20432

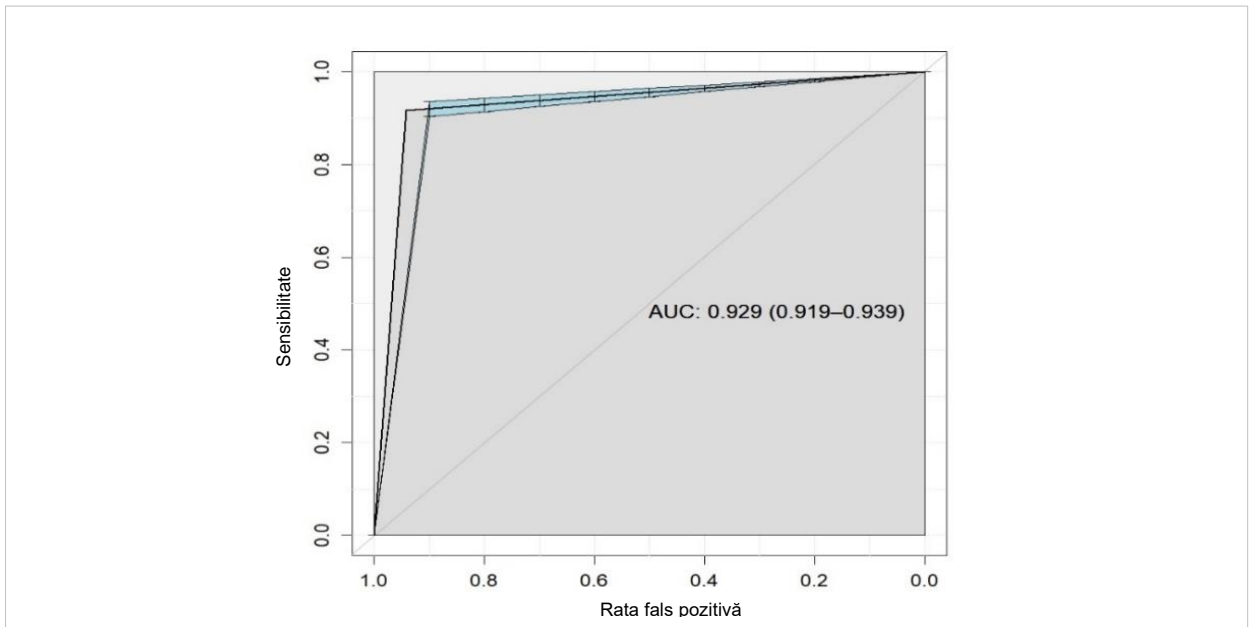


Figura. 4.5. Curba ROC la testarea modelului pe setul-test 1 (A)

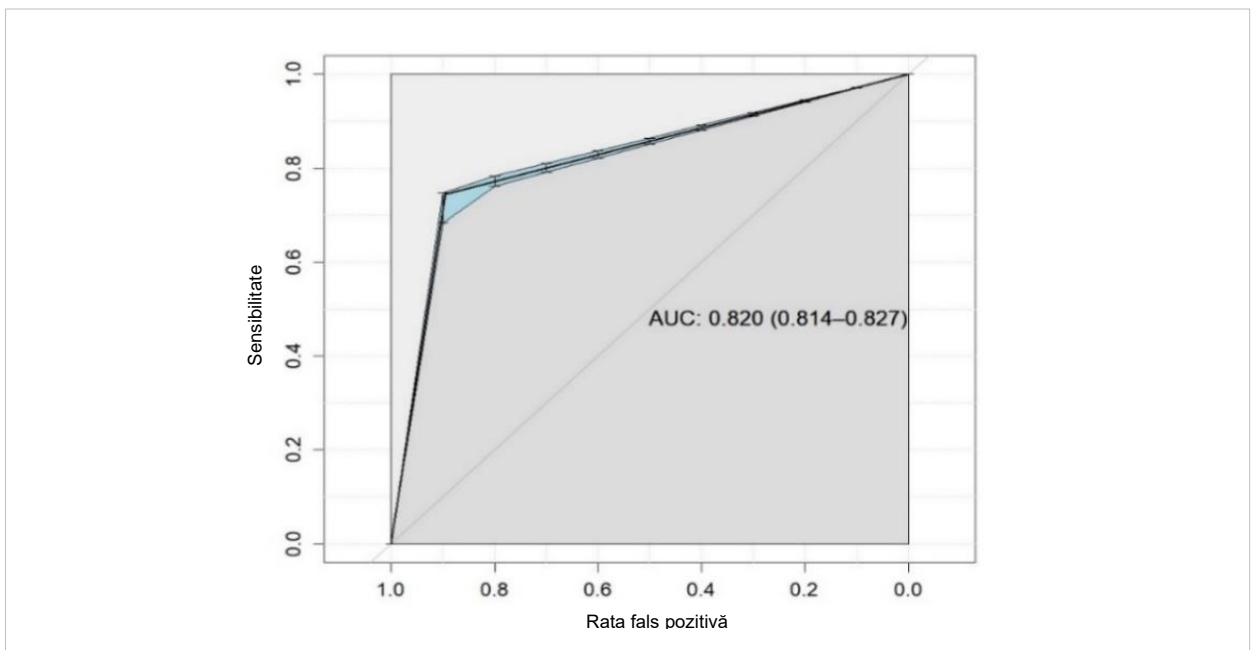


Figura. 4.6. Curba ROC la testarea pe setul-test 2 (B) a modelului antrenat pe date din setul A fără coloanele cu toate valorile lipsă

Parametrii de performanță ilustrați mai sus reprezintă metrici tradiționale pentru așa sisteme. Anexa 7 vine cu informație suplimentară ce descrie această etapă. O importanță aparte o are ponderea parametrilor, determinată în rezultatul învățării și validării. Analiza ei ar putea să contribuie la o înțelegere mai profundă a sepsisului. Aceasta este descris în compartimentul discuții”.

4.3. Incorporarea sistemului creat într-o aplicație pentru utilizare clinică

Pentru utilizarea clinică a sistemului, el a fost “oformat” ca o aplicație software (creată cu ajutorul pachetului “Wave” în Python [212]), care este comodă de utilizat de către medicul ATI. Mai jos este ilustrată interfața grafică a acestei aplicații.

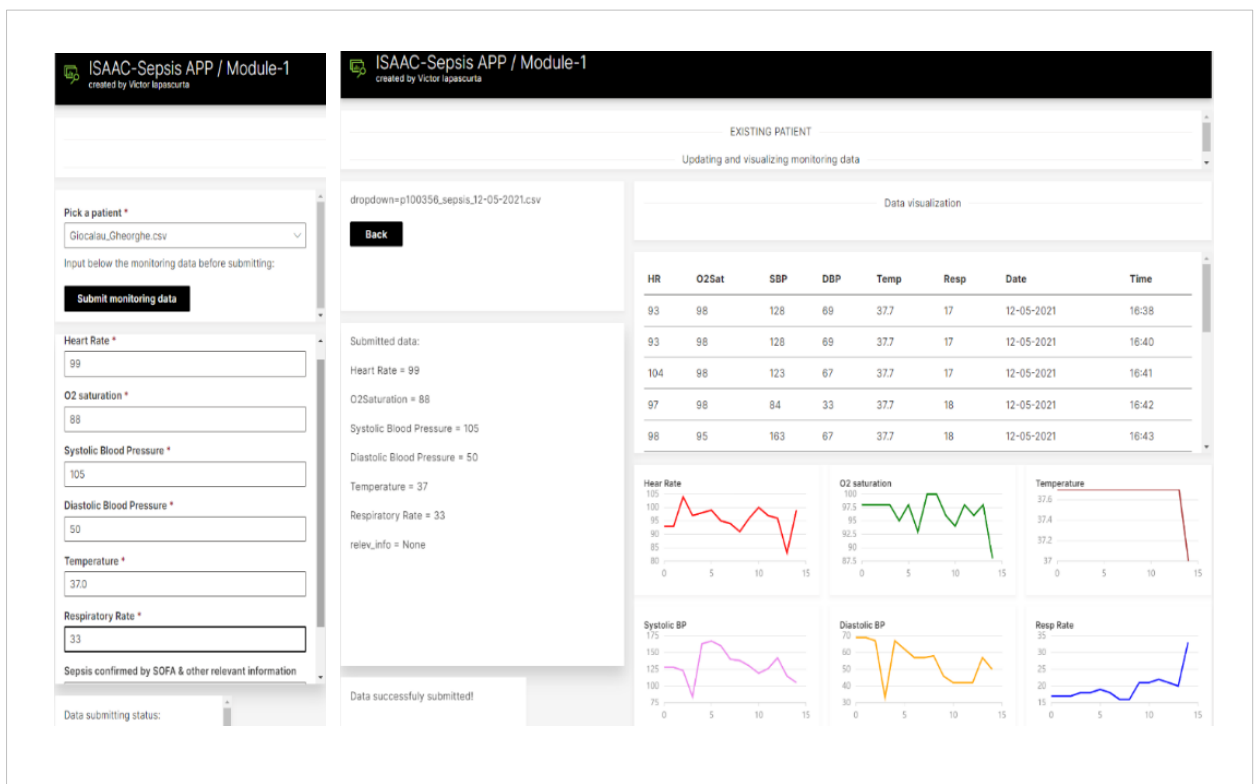


Figura 4.7. Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: introducerea datelor și vizualizarea lor în formă de tabel și grafică.

S-a tins spre comoditatea aplicației pentru utilizator cu posibilitatea de a introduce datele la interval de o oră (care pot fi la necesitate descărcate ca fișier “.csv”) iar în cazul unor date acumulate anterior – este posibilă importarea datelor cu vizualizarea lor.

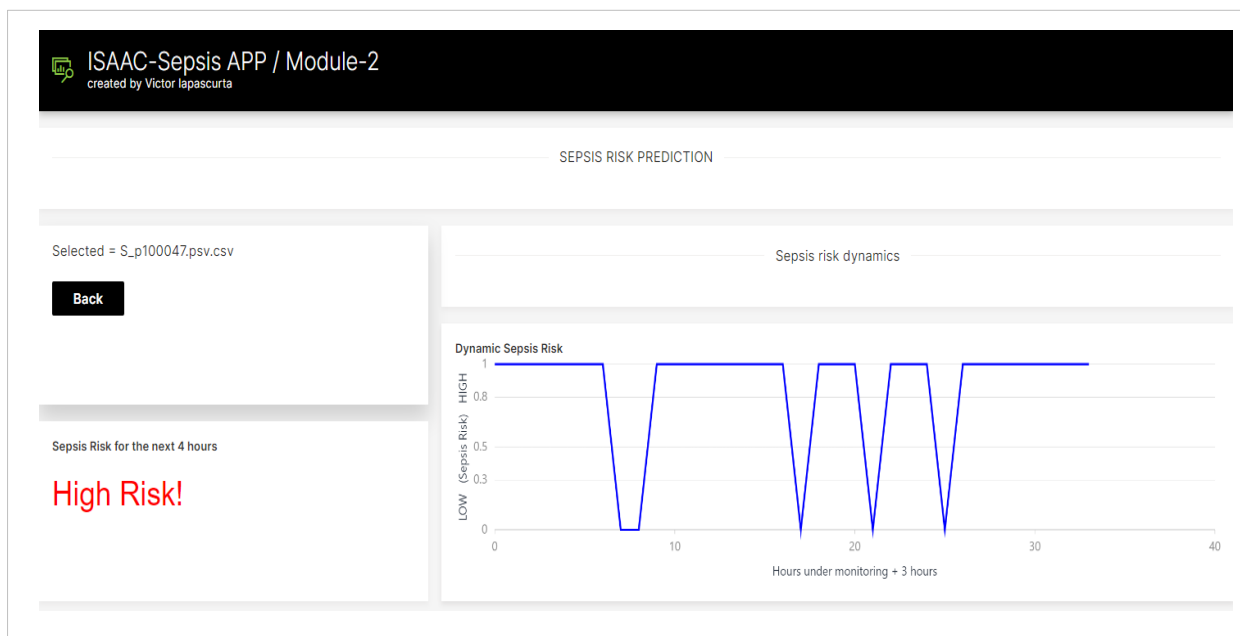


Figura 4.8. Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: obținerea rezultatului predicției și vizualizarea dinamicii riscului.

The screenshot shows the 'NEW PATIENT' form in the ISAAC-Sepsis APP / Module-3. The form is divided into two main sections. The left section, 'Input patient information', contains fields for 'Patient surname *' (Ciocalau), 'Patient first name *' (Gheorghe), 'Date of birth' (Thu Mar 08 1973), and 'Gender' (Male selected). The right section contains fields for 'Heart Rate *' (88), 'O2 saturation *' (93), 'Systolic Blood Pressure *' (110), 'Diastolic Blood Pressure *' (65), 'Temperature *' (37.9), and 'Respiratory Rate *' (24). Below these is a 'Diagnosis & other relevant information' field containing 'Hemicolnectomie la 15.05.2022'. At the bottom, there are 'Verify' and 'Submit data' buttons.

Figura 4.9. Interfața grafică a aplicației pentru utilizare clinică a sistemului de prezicere timpurie a sepsisului: introducerea datelor unui pacient nou

Această aplicație poate fi utilizată în felul următor: (1) se introduc datele ce descriu cazul/pacientul (cei 6 parametri) – pot fi introduse orar sau sub formă de tabel (când sunt date pentru mai multe intervale orare); (2) datele pot fi vizualizate grafic în formă de serii temporale. Odată cu acumularea a 3 ferestre orare automat se determină riscul de sepsis la pacientul vizat (“Risc înalt” vs “Risc neînsemnat”); (3) pe măsura avansării timpului după fiecare oră starea pacientului (în privința riscului sepsisului) este reevaluată cu afișarea gradului de risc pentru fiecare evaluare. Dinamica riscului poate fi vizualizată într-o fereastră separată.

4.4. Utilizarea aplicației pentru prezicere continuă a riscului sepsisului.

Pentru evaluarea performanței sistemului creat în privința urmării dinamicii riscului sepsisului în timp, prin eșantionare aleatorie au fost selectați cate 10 pacienți din setul test cu sepsis (confirmat prin metode tradiționale) și din setul cu alte patologii. Pentru uniformitate pentru fiecare caz selectat s-a extras o perioadă de observație de 11 ore. În cazul pacienților cu sepsis această perioadă include 8 ore până la confirmarea clinica a sepsisului și 3 ore – după. Pentru pacienții non-septici s-au extras 11 ore fără valori lipsă a parametrilor de interes.

Conceptul urmării riscului de sepsis în dinamica este ilustrat de Fig. 4.10

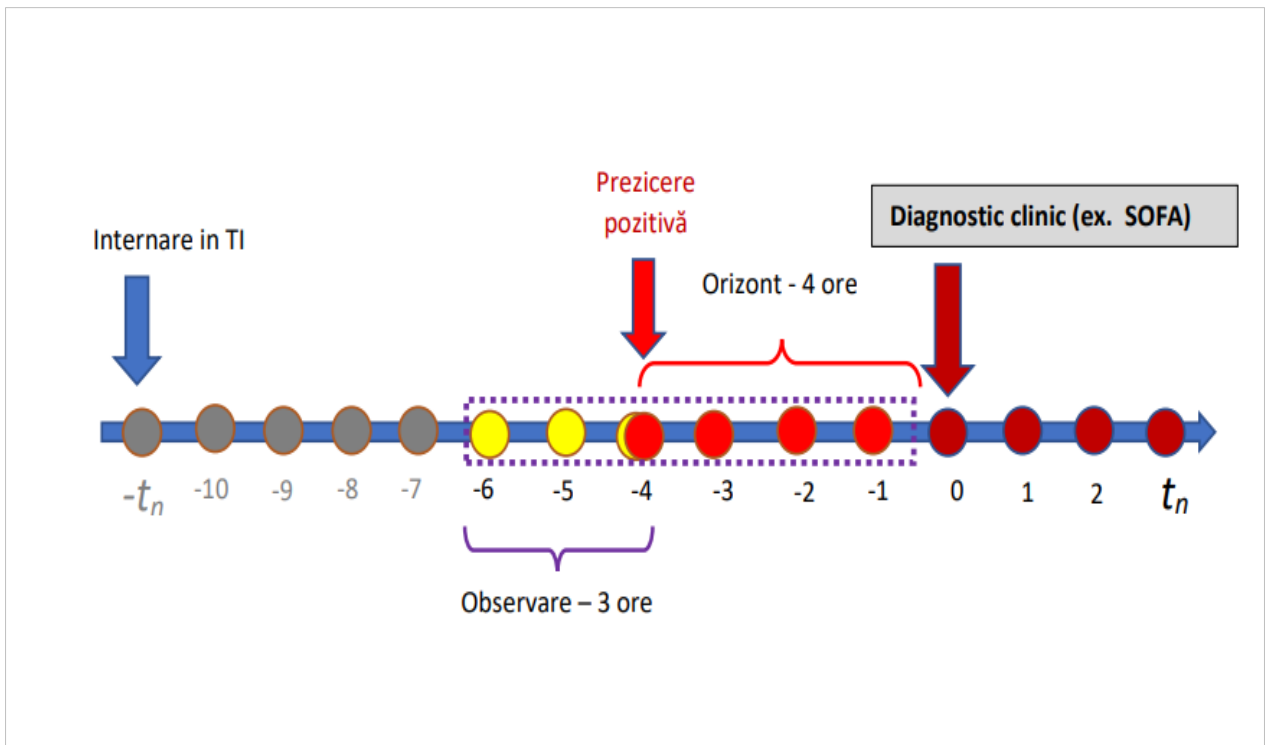


Figura 4.10. Secvența temporală a evenimentelor relevante în cazul unui pacient cu sepsis la utilizarea sistemului de predicție timpurie a sepsisului (intervalul de timp: $-t_n \rightarrow t_n$)

Rezultatul aplicării sistemului este ilustrat de Fig. 4.11 și 4.12. Anexa 8 prezintă această informație în forma tabelar-digitală.

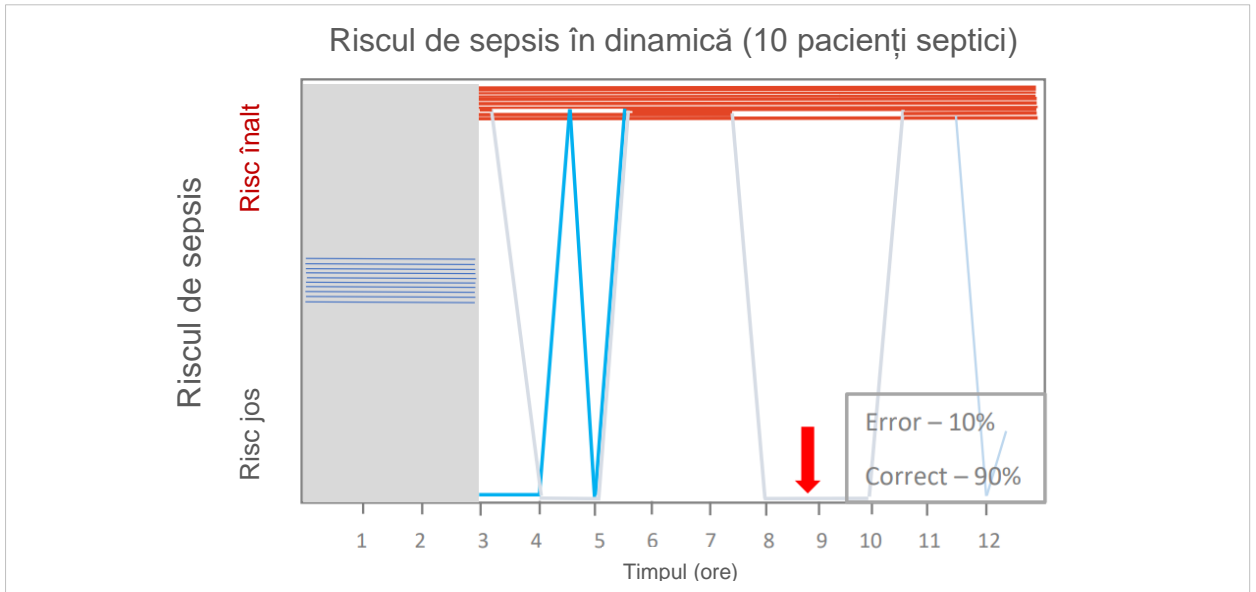


Figura 4.11. **Dinamica riscului sepsisului estimată de sistemul creat la 10 pacienți cu sepsis timp de 11 ore consecutive.**

Liniile roșii din partea de sus a diagramei reprezintă gradul de risc înalt estimat. Liniile gri și cele de culoare albastru deschis – estimările eronate. Săgeata roșie indică momentul confirmării sepsisului prin metode tradiționale. Zona gri din stânga reprezintă primele trei ore de aplicare a sistemului pentru a obține prima predicție, ulterior prezicerea se face orar.

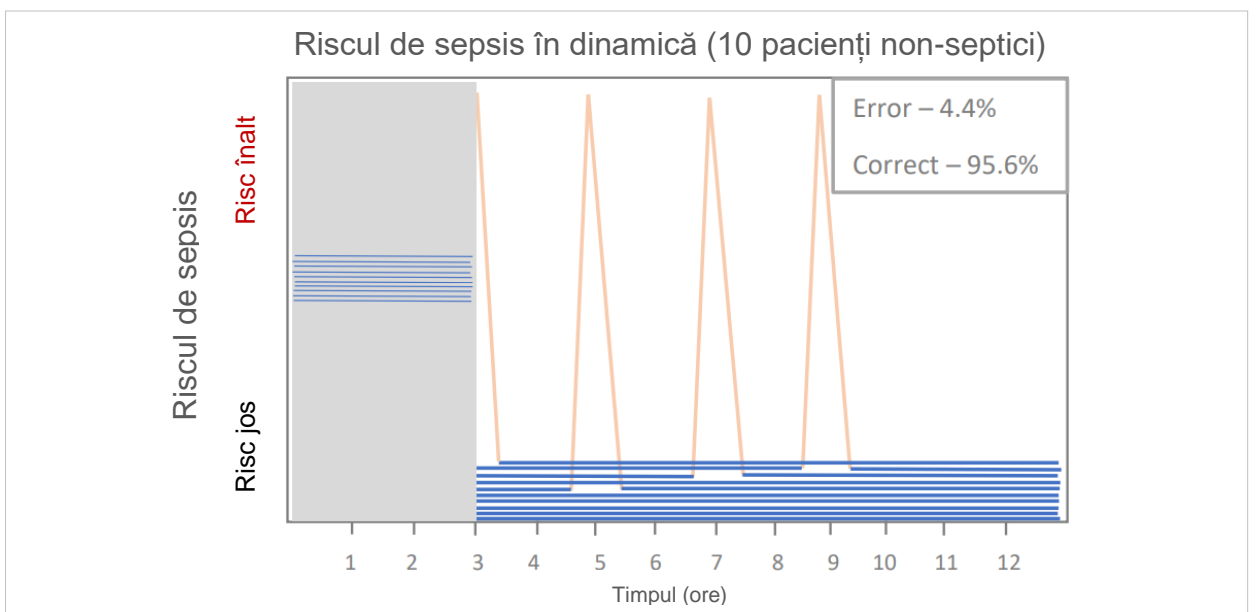


Figura 4.12. **Dinamica riscului sepsisului estimată de sistemul creat la 10 pacienți non-septici timp de 11 ore consecutive.**

Liniile albastre din partea de jos a diagramei reprezintă gradul de risc jos estimat. Cu culoare orange sunt reprezentate estimările eronate. Zona gri din stânga reprezintă primele trei ore de aplicare a sistemului pentru a obține prima predicție, ulterior prezicerea se face orar.

4.5. Sinteza Capitolului 4

Capitolul 4 descrie învățarea automată propriu-zisă, componentele și etapele procesului de învățare.

Urmând recomandările pentru realizarea acestei etape, datele folosite sunt scindate în setul pentru training și setul-test, care se prelucrează separat. Însăși procesul urmează etapele tradiționale în acest caz:

- Training
- Cross-validare
- Testare (pe setul-test, care nu a participat la training).

Drept referință este utilizat sistemul InSight, raportat în literatură ca unul dintre cele mai performante în domeniu. Acest sistem a fost replicat în limbajul de programare R și testat pe datele din studiul curent. Performanța acestor două sisteme (sistemul propriu și InSight) după aria de sub curba ROC este destul de apropiată (Tab. 4.1):

- InSight – 91,4%
- Sistemul propriu – 92,9%

În urma experimentării cu mai mulți algoritmi, cel mai performant s-a dovedit a fi GBM (Gradient Boosting Machine), care a condiționat performanța de predicție descrisă (92,9%), iar cel mai puțin performant - modelul în baza regresiei liniare (GLM – Generalized Linear Model) – 70,3%. Tabelul 4.3. prezintă aspectele de performanță după 12 metrici. Performanța modelului pe setul-test extern (date din alt spital) a fost de 81,3%.

Pentru analiza robusteții sistemului elaborat volumul datelor utilizate la crearea lui a fost gradual micșorat de la 100% la 25%, ce a condiționat o reducere neînsemnata a performanței (după aria de sub curba ROC) de la 92,9% (setul de training complet) la 91,2% (o pătrime din volumul inițial al acestui set).

Varianta finală a sistemului de predicție a fost oformată ca o aplicație software pentru utilizare clinică, cu o interfață grafică ilustrată de Fig. 4.7 – 4.9.

Abilitatea de evaluare a riscului de sepsis a aplicației software obținute a fost testată în dinamică, cu utilizarea parametrilor de intrare a eșantioanelor orare consecutive. Pentru aceasta au fost aleator selectați 10 pacienți septici (inclusiv perioade când sepsisul nu era confirmat prin metodele clinice tradiționale, ex. SOFA) și 10 pacienți non-septici, reprezentând cazuri care nu au participat la trainingul sistemului. Rezultatele au fost următoarele:

- Pacienți septici – din 90 de preziceri orare consecutive, 9 au fost prezise eronat – adică ca non-septice, ce reprezintă o rată a erorii de 10%
- Pacienți non-septici - din 90 de preziceri consecutive, 4 au fost eronat etichetate ca septice, cu o rată a erorii de 4,4%.

Aceste rezultate sunt în concordanță cu metricile de performanță descrise în Tab. 4.3. Informație mai detaliată despre aceste aspecte este accesibilă în Anexele 2-7.

5. EXPLICABILITATEA MODELELOR IA. DISCUȚII ȘI DIRECȚII DE CERCETARE DE VIITOR

5.1. Sumar al lucrului de cercetare efectuat. Discuția rezultatelor. Limitări

Prelucrarea datelor în cazul învățării automate este ghidată de reguli similare cu cele aplicate cercetării tradiționale, însă sunt și o serie de principii/recomandări specifice pentru acest domeniu. Anexa 9 descrie aceasta mai în detaliu. Un sumar al principalilor pași în analiza și prelucrarea datelor în cazul studiului curent este:

1. Excluderea subseturilor/cazurilor care conțin artefacte (adică valori ale parametrilor fiziologici de interes, care iese în afara limitelor (de obicei, fiziologice), spre exemplu, din cauza greșelilor de introducere a datelor). Astfel, din setul A au fost excluse 787 cazuri/pacienți (sepsis – 87, non-sepsis – 700), iar din setul B – 1018 cazuri.
2. Din seturile *non-septice* au fost extrase cazurile cu 7 și mai multe observații consecutive ale parametrilor de interes fără valori lipsă. Aceste observații au constituit baza datelor utilizate în continuare ca cazuri “non-sepsis”: setul A – 3336 cazuri, setul B – 2820 cazuri.
3. Scindare (prin eșantionare aleatorie) a datelor cu crearea setului pentru training (90% date – 1533 cazuri/pacienți cu sepsis și 3003 – non-sepsis) și setului-test (10% date – 170 cazuri/pacienți cu sepsis și 333 – non-sepsis). Setul pentru training a mai fost scindat (de asemenea prin eșantionare aleatorie) în seturi mai mici (75%, 50% și 25% din mărimea inițială), care mai târziu au fost utilizate pentru evaluarea robusteții modelului.
4. Restabilirea datelor lipsă prin interpolare (funcție creată în baza pachetului “zoo”, R) cu excepția coloanelor cu toate valorile lipsă (pentru fiecare set separat)
5. Excluderea subseturilor/cazurilor septice cu valori lipsă în mai mult de o coloană cu următoarele rezultate: set sepsis pentru training – 1431 subseturi, test-set – 159 subseturi). A fost creat și un set cu excluderea tuturor subseturilor/cazurilor cu toate valorile lipsă pentru oarecare din cei 6 parametri (set sepsis pentru training – 1073 subseturi; set-test – 128 subseturi)
6. Restabilirea datelor în coloanele cu toate valorile lipsă prin extragere și înlocuire a valorilor lipsă cu date din distribuții similare celor în datele prezente (create cu pachetul “sn”, R)
7. În baza datelor restabilite (vezi pașii precedenți 4 și 6) s-a efectuat generarea specimenelor finale, care reprezintă vectori cu lungimea 14L (primele două elemente reprezentând complexitatea algoritmică prin metoda BDM (block decomposition method) a 2 matrice, create prin gruparea celor 6 parametri în 2 grupuri și utilizând 3 valori orare pentru fiecare

parametru (se utilizează varianta binarizată a matricelor), iar celelalte 12 elemente ale vectorului fiind reprezentate de diferența dintre valorile celor 6 parametri în trei coșuri orare consecutive). La acest pas prelucrarea datelor se face separat pe seturi și la seturile care reprezintă pacienții septici se mai adaugă un element în vector – eticheta “1”, iar pentru cazurile non-septice – “0”. Astfel, numărul caracteristicilor finale este 14, plus eticheta.

8. Seturile finale au următoarele dimensiuni (număr de specimene): (a) modelul de bază A – setul pentru training - 27583 (sepsis – 9636, non-sepsis – 17947) și setul-test – 3052 (sepsis – 1063, non-sepsis – 1989); (b) setului A, fără cazurile cu parametri pentru care toate valorile lipsesc – setul pentru training – 25457 (sepsis – 7510, non-sepsis – 17947) și setul-test – 2866 (sepsis – 877, non-sepsis – 1989); (c) setul B, în calitate de set-test – 20432 (sepsis – 5383, non-sepsis – 15049).
9. Inițial modelele au fost construite cu ajutorul învățării automate automatizate (AutoML). Ulterior hiperparametrii optimi finali au fost justificați prin căutare grilă grosieră cu cross-validare multiplă (x 10). Selectarea celui mai performant model se face în baza parametrilor de performanță, în primul rând în baza ariei de sub curba ROC (AUC) la testarea pe setul test.
10. Performanța modelelor principale (după curba ROC – AUC) este ilustrată în Tab. 5.1.

Tabelul 5.1. Performanța modelelor finale după aria de sub curba ROC (AUC)

Model	AUC
Modelul-1 (de baza, antrenat și testat pe setul A)	0.929 (0.919 – 0.939)
Modelul-2 (antrenat și testat pe setul A, cu excluderea cazurilor în care cel puțin o coloana de date conține toate valorile lipsă)	0.924 (0.914 – 0.935)
Modelul-1 (testat pe setul B)	0.813 (0.807 - 0,820)
Modelul 2 (testat pe setul B)	0,820 (0.814 – 0.827)

Notă: În paranteze sunt indicate valorile pentru un interval de încredere de 95%.

De regulă, lucrul cu datele se începe cu analiza exploratorie a datelor pentru a investiga particularitățile și calitatea lor și a identifica unele aspecte care ar putea fi de folos la etape mai avansate în prelucrarea/transformarea datelor.

Datele inițiale ce conțin valori lipsă au fost restabilite cu ajutorul metodelor descrise anterior. Deoarece datele au un caracter continuu, reprezentând serii temporale, este puțin probabil, ca

valorile lipsă situate între valori prezente, să difere radical de cele prezente. Astfel, părtinirea introdusă la interpolarea datelor este considerată minimă. Pentru investigarea mărimii unei potențiale părtiniri introduse în cazul restabilirii valorilor în coloane unde toate datele lipsesc a fost creat un model (Modelul-2) pe subsetul de date, care exclude așa cazuri. Performanța acestui model (0.924) este destul de aproape de performanța modelului de baza (Modelul-1, 0.929), ce aparent indică la o părtinire, chiar dacă și prezentă, totuși neînsemnată. Aceeași situație este identificată și la testarea acestor modele pe setul-test extern B.

Analiza corelațională între parametri, efectuată separat pentru seturile ce reprezintă cazurile septice și cele non-septice, a relevat două fenomene, aparent importante: (a) există deosebiri statistic semnificative între mărimea coeficienților de corelare, iar uneori și diferențe în semnul corelării, și (b) mărimea coeficienților de corelare este totuși mică.

Daca privim această situație (adică prezența diferenței dintre coeficienți și a semnelui corelării între cazurile septice și cele non-septice) prin prisma conceptului de “Stare critică” [213], unde unul din componentele centrale este alterarea marcată a mecanismelor de autoreglare și control, atunci diferența coeficienților (între cazurile septice și non-septice) ar putea fi interpretată ca rezultat al modificării funcționării acestor mecanisme. Aceasta se aliniază și cu conceptul curent de sepsis (Sepsis-3), ca răspuns dereglat la infecție, iar cele descrise mai sus ar putea ilustra o verigă patofiziologică importantă în instaurarea insuficienței de organe, care în sepsis are un caracter multiplu.

Mărimea neînsemnată a coeficienților de corelare ar presupune că utilizarea modelelor de regresie liniară cu scopuri discriminative (sepsis vs non-sepsis) în cazul datelor curente ar fi mai puțin efectivă, fapt confirmat la etapa de învățare automată, când modelele de regresie au demonstrat o performanță modestă (70,3%) comparativ cu alte modele (Anexa 10, ultimul model în listă).

Selectarea parametrilor cei mai reușiți pentru o prezicere maximal efectivă a sepsisului și forma de reprezentare a datelor au fost realizate pe cale experimentală. Conform celor descrise anterior, inițial au fost selectați 16 parametri, care cumva sunt legați cu parametrii utilizați pentru definirea sepsisului (Sepsis – 3 cu utilizarea SOFA: frecvența respirației, parametrii ce descriu funcția ficatului, nr. de trombocite, etc.). Apoi a fost utilizat tot spectrul de parametri clinici (36 parametri). În ambele cazuri a fost utilizată procedura de calculare a complexității algoritmice (BDM). O varianta de arhitectură utilizată a fost și ansamblul de modele (stacked ensemble). Rezultatele acestor încercări sunt rezumate în tabelul 5.2.

Tabelul 5.2. Caracteristica comparativă a 4 sisteme create în baza acelorași date

Caracteristici și parametri de performanță	Model BDM -16 parametri	Model BDM - 36 parametri	Model Stacked Ensemble	Modelul final
Algoritmul utilizat	XRT	DNN/MLP	Stacked Ensemble	GBM
Performanța modelului (după ROC AUC)	0.61	0.65	0.85	0.93

Notă: Sunt prezentați algoritmi cu cea mai mare performanță în cazul respectiv: XRT – o varietate a algoritmului “Random Forest”; DNN/MLP – “deep neural network/multilayer perceptron”; Stacked Ensemble – ansamblu de modele; GBM – “gradient boosting machine”.

În ultimă instanță a fost selectată varianta cu cei 6 parametri (FCC, SpO₂, Temp, PAS, PAD, Resp) - semne vitale și dinamica lor ca forma cu cea mai mare valoare predictivă.

Dacă analizăm indicatorii de performanță a sistemului elaborat în studiul curent separat (de exemplu, numai sensibilitatea care, la setul A este de 91,6% - destul de înaltă) am putea ignora tabloul complet. Valoarea predictivă pozitivă (VPP), care denota probabilitatea prezenței de facto a sepsisului la pacienții la care a fost prezis un risc înalt este modestă (60,5% la setul A și 33,2% la setul B). Aceasta în mare măsură este condiționată de prevalența mică a sepsisului, indice care participă la calculul acestor metrici (prevalența în seturile inițiale: set A – 8,8%; set B – 5,71%).

Specificitatea înaltă a sistemului de predicție (94,2% (A) și 90,2% (B)) ar permite cu un grad înalt de certitudine de exclus sepsisul în cazul testului negativ (predicția unui risc de sepsis jos). Aceasta este în concordanță și cu valoarea predictivă negativă (VPN) înaltă (respectiv, 99,2% și 98,1%), care denotă probabilitatea absenței sepsisului la pacienții testați/preziși ca negativi/cu risc de sepsis jos. Fiind pentru ambele seturi înaltă, aceasta în careva măsură echilibrează efectul VPP.

Celelalte metrici se aliniază cu cele descrise. Astfel, raportul de probabilitate pozitiv mai mare decât 10 în cazul setului A (15.85) și în intervalul 5-10 (8.20) la setul B, indică la o creștere mare a probabilității sepsisului la pacienții testați pozitiv în primul caz și una moderată în cel de al doilea. Iar raportul de probabilitate negativ mai mic de 0.1 (set A), și situat în intervalul 0.2 – 0.5 (set B), indică la o diminuare convingătoare a probabilității bolii în cazul setului A și una mică pentru setul B, când predicția este negativă. Raportul șanselor întregeste tabloul.

Dacă calculăm indicii dependenți de prevalență (VPP, VPN, acuratețea diagnostică), utilizând prevalența specimenelor septice (de exemplu, în cazul setului A - 35.14%) în seturile de date care sunt utilizate nemijlocit pentru ÎA (vs prevalența în setul inițial – 8.8%), obținem valori

„mai optimiste”: VPP – 89.43% (Î 95%: 87.62 – 91.12); VPN – 95.41% (Î 95%: 94.45 – 96.2); acuratețe – 93.32% (Î 95%: 92.37 – 94.18).

O metrică importantă pentru practician este rata rezultatelor fals pozitive, RFP (alarme false), care, când este înaltă, poate duce la surmenarea personalului. În cazul setului A ea este de 5,8%. Un aspect similar este relatat de rata fals negativă, RFN (rata de omitere a cazurilor pozitive), care pentru setul A este de 8,4%. Aceste cifre sunt apropiate de cele obținute la evaluarea abilității sistemului de a prezice riscul de sepsis în dinamică (vezi compartimentul 4.4).

Pentru setul B rata alarmelor false este de 8,7%, iar RFN este de 28,7%. Mai jos (când se discută aspectul de transferabilitate) sunt prezentate câteva soluții pentru gestionarea acestei probleme, care ar putea fi unul din obiectivele cercetărilor de viitor.

De rând cu metricile enumerate, ce descriu performanța modelului, se folosește și aria de sub curba ROC (Receiver Operating Characteristic), AUC-ROC, care este o metrică mai complexă, ilustrând relația dintre sensibilitate și rata fals pozitivă, luând în considerație anumite valori-prag. Cu toate că nu există intervale de valori stricte pentru evaluare, valorile cuprinse între 0.9 și 1.0 se consideră excelente și cele în intervalul 0.8 – 0.9 bune/acceptabile. Pentru setul A valoarea acestei metrici aparține primului interval, iar pentru setul B – celui de al doilea.

Compararea performanței diferitor modele/sisteme deseori poate fi dificilă, în deosebi când ele vin din diferite studii. Obstacole pentru comparare pot fi multiple: informație insuficientă referitor la designul și metodologia studiilor, orizont de prognozare diferit, metrici diferite de raportare, absența validării, în special a validării externe, etc. Pe măsura elaborării unor principii/recomandări standard o parte din aceste probleme se rezolvă, însă acest proces este la început de cale.

Anterior performanța modelului creat în acest studiu a fost comparată cu performanța sistemului InSight (vezi Tab. 4.1). O sursă credibilă pentru comparare poate fi “Ghidul internațional pentru managementul sepsisului și al socului septic” din noiembrie, 2021 [123], unde în premieră este menționat potențialul rol al învățării automate, care “poate îmbunătăți performanța instrumentelor de screening”, în baza unei meta-analize a 42623 de pacienți din șapte studii pentru previziunea sepsisului spitalicesc, care raportează o arie de sub curba ROC de 0,89 (95% Î, 0,86-0,92); sensibilitate - 81%; (95% Î, 80-81) și specificitate - 72% (95% Î, 72-72). Performanța sistemului creat în studiul curent este superioară după fiecare din metricile de performanță enumerate.

Încercarea de a compara sistemul creat cu sisteme descrise în literatură în ultimii 3 ani, care raportează o performanță mai superioară relevă următoarele:

Tabelul. 5.3. Informație sumară despre sisteme similare cu o performanță mai înaltă decât a celui creat în studiul curent (ultimii trei ani)

Nr.	Autor (i)	Anul publicării	AUC ROC/ acuratețe diagnostică*	Număr caracteristici/parametri fiziologici	Sursa
1.	Sadasivun S. et al.	2022	0.93*	ECG + caracteristici demografice	[214]
2.	Misra D. et al	2021	0.9483	15	[215]
3.	Wickramaratne D., Mahmud S.	2020	0.97	36	[216]
4.	Kok C. et al.	2020	0.98	40	[217]
5.	Helguera-Repetto, A. et al.	2020	0.944	25	[218]

În cazul sistemelor Nr. 3 (Wickramaratne, 2020) și 4 (Kok, 2020), care raportează cea mai înaltă performanță, din cauza informației accesibile insuficiente (designului studiului și, în particular, referitor la detaliile validării) este dificil de făcut o concluzie definitivă. În ambele cazuri sistemele utilizează mult mai mulți parametri clinici (36 și 40) decât sistemul creat (6 parametri). Dacă comparăm performanța lor raportată (AUC ROC 0.97 și 0.98) cu performanța sistemului creat în cercetarea noastră, în cadrul validării – 0.97 (vezi tabelul 5.1, a doua coloană), atunci în această privință ele sunt destul de aproape.

Sistemul Nr. 2 (Misra, 2021), care raportează destul de detaliat diverse aspecte ale cercetării, este de asemenea greu de al compara din așa motive: (a) ca metodă de referință este utilizată cea bazată pe principiile sepsis-2/SIRS; (b) este de fapt focusat pe predicția socului septic (vs sepsis); (c) este mai greu de închipuit cum o sensibilitate de 0.8392 și o specificitate de 0.8814 pot fi asociate cu AUC ROC de 0.9483 (fără transformări/manipulări suplimentare ale datelor).

Sistemul Nr.5 (Helguera-Repetto, 2020), care este detaliat descris în [218], este elaborat în baza unui grup de cazuri limitat (132 cazuri non-sepsis și 106 cazuri de sepsis), având un grup de studiu diferit (nou-născuți), decât cercetarea curentă. Respectiv, si caracteristicile clinice utilizate sunt diferite, atât după caracterul lor (materne și cele care se referă la nou-născut), cât și după număr (25).

Sistemul Nr. 1 (Sadasivun, 2022) este cel mai recent și este bazat pe o tehnologie inginerescă nouă, la moment în faza experimentală (rețele neuronale artificiale analoge în baza IMC, in-memory computing). Sursa [214] livrează informație cu caracter mai mult tehnic, decât medical. Studiul este realizat pe un lot de 514 pacienți cu sepsis și 451 – non-septici din același spital, fără validare externă. Metrica de bază raportată este acuratețea diagnostică (care este

indicată în tabel). Una din configurațiile tehnice descrise este raportată cu o acuratețe de 93%, sensibilitate – 90%, specificitate – 95% și AUC ROC – 99%. Similar cu sistemul Nr. 2, este dificil de a răspunde la întrebarea cum sensibilitatea și specificitatea raportată ar putea fi asociate cu o așa valoare a AUC ROC.

Una din principalele limitări ale sistemului de predicție creat în cadrul cercetării curente este legată de un subiect care la moment este în dezbatere. O meta-analiză recentă [219] a indicat că majoritatea pacienților cu COVID-19 grav bolnavi (78%) au îndeplinit criteriile Sepsis 3.0 pentru sepsis/șoc septic cu sindrom de detresă respiratorie acută (ARDS), ca fiind cea mai frecventă disfuncție de organ (88%). Astfel, este sugestiv ca COVID-19 la pacienții spitalizați să fie considerat în mod inerent ca sepsis. Această percepție nu este împărtășită pe scară largă, iar opiniile variate despre COVID-19 și sindroamele de sepsis tulbură înțelegerea patofiziologiei lor. Având în vedere că definiția Sepsis 3.0 este relativ incluzivă, este imperativ să înțelegem caracteristicile fenotipice similare și distinctive ale ambelor afecțiuni, pentru a maximiza efectul tratamentului.

Aproape trei ani de cercetare au aruncat careva lumină asupra complexității răspunsului imuno-inflamator la COVID-19 [220, 221]. Este evident că eticheta captivantă „furtună de citokine” ar trebui să fie retrogradată la o „burniță de citokine”, deoarece nivelurile de citokine proinflamatorii circulante (de exemplu IL-6, IL-8, TNF) sunt la o fracțiune din concentrațiile înregistrate într-un sepsis/șoc septic non-SARS CoV-2 [222]. Spre deosebire de răspunsul sistemic, compartimentul pulmonar la pacienții cu COVID-19 grav bolnavi suferă de obicei o inflamație robustă și prelungită. La nivel de management al COVID-19, nu există o strategie dominantă, care să difere dramatic (în afară de antimicrobiene/antivirale) de pachetul de tratament stabilit pentru sepsis de către orientările National Institutes of Health din S.U.A. O excepție importantă este eficacitatea diferită a glucocorticoizilor (GC). În timp ce ghidurile actuale pentru sepsis prezintă o recomandare slabă pentru GC, utilizarea lor pentru pneumonia severă SARS-CoV-2 este fără echivoc benefică. Mecanismele biologice din spatele acestei disparități ar trebui elucidate, deoarece motivele care stau la bază ar putea stimula o renaștere a GC în sepsisul bacterian și terapia intensivă în general [223].

Există o paralelă izbitoare între sepsisul bacterian și fenotipurile COVID-19 - sechelele pe termen lung. În ambele grupuri de pacienți, externarea din spital nu echivalează cu recuperarea completă, dar este frecvent urmată de consecințe prelungite, incapacitante. În timp ce în sepsisul bacterian, complicațiile post-externare sunt denumite sindrom post-sepsis și/sau inflamație persistentă, imunosupresie și sindrom de catabolism (PICS), la pacienții infectați cu SARS-CoV-2, ele sunt cunoscute ca „COVID de lungă durată”. Sindromul “COVID de lungă durată” nu este foarte diferit de sindromul post-sepsis. Cele mai frecvente simptome persistente includ oboseală,

dureri musculare, dereglarea somnului, tulburări cardiace și cognitive (de exemplu, aritmii, pierderea memoriei pe termen scurt) [224]. În mod remarcabil, există o diferență nouă, tulburătoare: spre deosebire de sepsis, COVID-ul de lungă durată este frecvent diagnosticat la pacienți infectați cu SARS-CoV-2 în formă ușoară (adică fără spitalizare) [225]. Prezența „fenotipului lung” în ambele boli indică puternic o dereglare severă și prelungită a sistemului imunoinflamator (cu caracteristici clare de imunosupresie) și a homeostaziei. În contextul manifestărilor severe ale COVID-19, care diminuează lent, ar trebui să ne concentrăm din nou asupra sechelelor pe termen lung pentru a evalua un risc potențial de creștere a consecințelor cronice.

Având în vedere că sepsisul bacterian și paralelele COVID-19 se împletesc puternic cu contrastele, este esențial de a le disecționa cu atenție în piese definite și gestionabile de dovezi fiziopatologice înainte de a se recomanda orice acțiune terapeutică ulterioară. Conform [219], la fel de important este de evitat transferul reflex de preconcepții gata de utilizat (de exemplu, „furtuna de citokine”) de la o boală existent, în timp ce avem de-a face cu orice entitate nouă. Studiile preclinice bine concepute pot ajuta la o verificare validă translațională a practic oricărui dintre conceptele de mai sus la o fracțiune de timp/costurile necesare pentru executarea unui studiu clinic.

Cercetarea curentă folosește date, care au fost colectate până la pandemia COVID-19. Aceste date se referă la sepsisul bacterian. Pentru a ajusta sistemul creat la sepsisul viral în general, și cel cauzat de coronavirusi, în particular, vor fi necesare cercetări suplimentare, care ar trebui să țină cont de eventualele deosebiri esențiale (daca acestea vor fi clar articulate) începând de la nivelul fiziopatologic până la cel clinic, aceasta fiind una din posibilele direcții de cercetare și, la necesitate ajustare/îmbunătățire a modelului.

Alte limitări ale sistemului creat în cadrul cercetării curente sunt:

- O limitare potențial remediabilă ar fi diferența între datele utilizate în cercetare și datele locale (în unitatea de terapie țintă, unde ar putea fi implementat sistemul). În acest caz o soluție ar putea fi învățarea prin transfer, descrisă către sfârșitul lucrării curente.
- Limitări legate de diferența de abordări a pacienților septici, care fiind în esență similare și aliniată la conceptele moderne, pot varia în detalii. Aici sunt mai mulți factori care pot influența această diferență: dispozitive de monitoring diferite, precum și diferența în tehnicile de monitoring și tratament, etc.

- Limitări de ordin “statistic”, generate de ultimele două limitări, care ar putea genera distribuții diferite ale parametrilor de interes, pentru care o soluție ar fi construirea unui sistem “de novo” în baza datelor locale din unitatea țintă.

5.2. Aspecte de „explicabilitate” a modelelor IA

Inteligența artificială explicabilă (XAI) atrage mult interes în medicină. De aici, un aspect important recent discutat frecvent în literatură ține de „explicabilitatea” modelelor ÎA [226, 227].

Din punct de vedere tehnic, problema „explicabilității” este la fel de veche ca ÎA însăși, iar IA clasică a reprezentat abordări comprehensibile și retrasabile. Cu toate acestea, slăbiciunea lor era în a face față incertitudinilor lumii reale. Prin introducerea învățării probabilistice, aplicațiile au devenit din ce în ce mai de succes, dar din ce în ce mai opace. XAI se ocupă de implementarea transparenței și a trasabilității metodelor statistice de învățare automată, în special învățarea profundă (ÎP). Se menționează că este nevoie de a depăși IA explicabilă, existentă în forma actuală. Pentru a ajunge la un nivel de medicină explicabilă avem nevoie de cauzalitate. În același mod în care capacitatea de utilizare include măsurători pentru calitatea utilizării, „cauzabilitatea” cuprinde măsurători pentru calitatea explicațiilor. Lucrarea [228] oferă câteva definiții necesare pentru a discrimina „explicabilitatea” de cauzalitate. Principala contribuție a acestei lucrări este noțiunea de „cauzabilitate”, care se diferențiază de capacitatea de explicație prin aceea că cauzabilitatea este o proprietate a unei persoane, în timp ce „explicabilitatea” este o proprietate a unui sistem.

Medicina ca domeniu de aplicare este printre cele mai mari provocări ale IA/ÎA/ÎP. În suportul decizional în medicină clinicianul se confruntă cu incertitudine, cu seturi de date probabilistice, necunoscute, incomplete, dezechilibrate, eterogene, zgomotoase, eronate, inexacte și lipsă, în spații de dimensiuni arbitrar înalte [229, 230]. Adesea pur și simplu seturile suficient de mari de date lipsesc [231]. Un obiectiv important al medicinei viitoare este modelarea complexității pacienților pentru a adapta deciziile medicale, practicile de îngrijire și tratament la pacientul individual [232]. Acest lucru ridică provocări în special în integrarea, fuziunea și maparea diferitelor date distribuite și eterogene până la analiza vizuală a acestor date [229]. În consecință, IA explicabilă în contextul medicinei trebuie să țină cont de faptul că diverse date pot contribui la un rezultat relevant. Acest lucru necesită ca profesioniștii medicali să aibă posibilitatea de a înțelege cum și de ce a fost luată o anumită decizie de către mașină [234].

„Explicabilitatea” este cel puțin la fel de veche ca IA în sine și mai degrabă o problemă care a fost cauzată de aceasta. În zilele de pionierat ale IA, metodele de raționament erau logice și simbolice. Aceste abordări au avut succes, dar numai într-un spațiu de domenii foarte limitat și cu aplicabilitate practică extrem de limitată. Un exemplu tipic este MYCIN [235], care a fost un

sistem-expert dezvoltat pentru a identifica bacteriile care cauzează infecții severe și pentru a recomanda antibiotice. MYCIN nu a fost niciodată utilizat în rutina clinică, poate din cauza caracterului său de sine stătător și a efortului mare de a menține baza de cunoștințe. Cu toate acestea, aceste sisteme IA timpurii au fost concepute prin efectuarea unei forme de inferență logică asupra simbolurilor, care pot fi citite de om și au fost capabile să ofere posibilitatea de a urmări pașii lor de inferență [236, 237].

Conform [228] există trei tipuri de explicații: (1) o explicație “peer-to-peer”, așa cum este efectuată în rândul medicilor în timpul raportării medicale; (2) o explicație educațională așa cum se realizează între profesori și elevi; (3) o explicație științifică în sensul strict al teoriei științei [238] și lucrarea respectivă face referință la primul tip de explicație.

În medicină, există o cerere din ce în ce mai mare pentru abordări AI, care nu numai că funcționează bine, dar sunt demne de încredere, transparente, interpretabile și explicabile pentru un expert uman [239]. Sunt necesare metode și modele pentru a reactiva procesul de luare a deciziilor de către mașină, pentru a reproduce și pentru a înțelege atât procesul de învățare, cât și cel de extragere a cunoștințelor. Acest lucru este important, deoarece pentru sprijinirea deciziei este necesar să înțelegem cauzalitatea reprezentărilor învățate [240, 241].

În plus, „explicabilitatea” IA ar putea contribui la creșterea încrederii profesioniștilor din domeniul medical în viitoarele sisteme IA. Cercetarea în vederea construirii de sisteme explicabile pentru aplicare în medicină necesită menținerea unui nivel ridicat de performanță de învățare pentru o serie de tehnici de IA și de interacțiune om-calculator. Există o tensiune inerentă între performanța IA (acuratețea predictivă) și „explicabilitate”. Adesea, cele mai performante metode, cum ar fi ÎP, sunt cele mai puțin transparente, iar cele care oferă o explicație clară (de exemplu, arbori de decizie) sunt mai puțin precise [242].

În medicină, IA explicabilă este necesară pentru multe scopuri, inclusiv educația medicală, cercetarea și luarea deciziilor clinice [243]. Dacă profesioniștii medicali sunt completați de sisteme IA sofisticate și, în unele cazuri, sistemele IA chiar joacă un rol important în procesul de luare a deciziilor, experții umani trebuie să aibă în continuare mijloacele pentru a înțelege și a urmări procesul decizional al mașinii.

În același timp, deși se presupune adesea că oamenii sunt întotdeauna capabili să-și explice deciziile, adesea aceasta nu este așa. Uneori, experții nu sunt în măsură să ofere o explicație bazată pe diversele surse eterogene și vaste de informații diferite. În consecință, IA explicabilă solicită încredere, siguranță, securitate, confidențialitate, etică, corectitudine și încredere [244] și aduce interacțiunea om-IA într-un focus nou și important [245]. Toate aceste aspecte împreună sunt

cruciale pentru aplicabilitatea în medicină în general și pentru viitoarea medicină personalizată, în special [246].

În acest context par a fi binevenite câteva definiții pentru a explica esența „explicabilității” prin prisma termenului „cauzabilitate” în contrast cu termenul bine-cunoscut „cauzalitate” și a unei scurte evaluări a unor modele explicabile curente. Un aspect important este și necesitatea urgentă a unei scale de „cauzabilitate” a sistemelor pentru a măsura calitatea unei explicații, care trebuie să includă și aspectele sociale ale comunicării umane [247].

Într-o lume ideală, atât afirmațiile umane, cât și cele ale mașinii ar fi identice și congruente cu adevărul de bază, care este definit pentru mașini și oameni în mod egal. Cu toate acestea, în lumea reală ne confruntăm cu două probleme:

(a) Adevărul de bază nu poate fi întotdeauna bine definit, mai ales atunci când se pune un diagnostic medical.

(b) Modelele umane (științifice) se bazează adesea pe cauzalitate ca scop final pentru înțelegerea mecanismelor subiacente și, deși corelația este acceptată ca bază pentru decizii, este privită ca un pas intermediar. În contrast, algoritmiile IA de succes de astăzi se bazează de obicei pe modele probabilistice și oferă doar o bază brută pentru stabilirea în continuare a modelelor cauzale. Atunci când discutăm despre explicabilitatea unei declarații furnizată de mașină, ne propunem, prin urmare, să distingem între:

- Explicabilitate - în sens tehnic evidențiază părți relevante pentru decizie din reprezentările utilizate ale algoritmilor și părțile active în modelul algoritmic, care fie contribuie la precizia modelului pe setul de antrenament, fie la o predicție specifică pentru o anumită observație și nu se referă la un model uman explicit.
- Cauzabilitate - ca măsura în care o explicație a unei declarații către un expert uman atinge un nivel specificat de înțelegere cauzală cu eficacitate, eficiență și satisfacție într-un context de utilizare specificat.

Putem distinge două tipuri de IA explicabilă, care pot fi denumite cu denumiri latine folosite în drept [248]:

(i) explicabilitate *post-hoc* = „(lat.) după aceasta”, care are loc după evenimentul în cauză; de exemplu, explicarea a ceea ce prezice modelul în ceea ce privește ce este ușor de interpretat;

(ii) explicabilitate *ante-hoc* = „(lat.) înainte de aceasta”, care are loc înaintea evenimentului în cauză; de exemplu, încorporarea explicabilității direct în structura unui model IA, explicabilitatea prin proiectare.

Sistemele „post-hoc” urmăresc să ofere explicații locale pentru o anumită decizie și să o facă reproductibilă la cerere (în loc să explice comportamentul întregului sistem). Un exemplu reprezentativ este explicațiile locale interpretabile model-agnostice (LIME) [249], care este un sistem model-agnostic, unde $x \in \mathbb{R}^d$ este reprezentarea originală a unei instanțe care este explicată și $x' \in \mathbb{R}^{d'}$ este folosit pentru a desemna un vector pentru reprezentarea sa interpretabilă (de exemplu, x poate fi un vector caracteristic care conține înglobări de cuvinte, x' fiind punga de cuvinte). Scopul este de a identifica un model interpretabil peste reprezentarea interpretabilă care este local fidel clasificatorului. Modelul explicativ este g :

$$g : \mathbb{R}^{d'} \rightarrow \mathbb{R}, g \in G, \quad (5.1)$$

unde G este o clasă de modele potențial interpretabile, cum ar fi modele liniare, arbori de decizie sau liste de reguli; dat un model $g \in G$, acesta poate fi vizualizat ca o explicație pentru expertul uman [249].

Sistemele „ante-hoc” sunt interpretabile prin proiectare către abordări tip „cutie de sticlă/transparentă” (versus „cutie neagră/netransparentă”) [250]. Exemplele tipice includ regresia liniară, arbori de decizie și sisteme de inferență neclare. Acestea din urmă au o tradiție lungă și pot fi proiectate din cunoștințe de specialitate sau din date și oferă — din punctul de vedere al interacțiunii om-AI — un cadru bun pentru interacțiunea dintre cunoștințele expertului uman și cunoștințele ascunse în date [251]. Un alt exemplu este [252], unde au fost aplicate modele aditive generalizate de înaltă performanță cu interacțiuni perechi problemelor din domeniul medical, producând modele inteligibile, care pot descoperi modele surprinzătoare în datele care împiedicaseră anterior modele complexe învățate să fie aplicate în acest domeniu; important este că au demonstrat scalabilitatea unor astfel de metode la seturi mari de date care conțin sute de mii de pacienți și mii de atribute, rămânând în același timp inteligibile și oferind o acuratețe comparabilă cu cele mai bune metode ÎA (neinteligibile). Un exemplu suplimentar de metode „ante-hoc” poate fi văzut în [253], unde este descris un cadru pentru explicarea vizuală a deciziilor oricărui clasificator care este formulat ca model aditiv și care a arătat cum se implementează acest cadru în contextul a trei modele: Bayes naiv, mașini cu vectori de suport și regresia logistică într-o aplicație de bioinformatică [254].

O direcție aparent importantă este cercetarea în baza modelelor cauzale structurale [255]. IA actuală funcționează fie într-un mod statistic, fie fără model. Acest lucru implică limite severe atât asupra eficacității, cât și asupra performanței. Astfel de sisteme nu pot raționa cu privire la intervenții și retrospecție și, prin urmare, nu pot servi pentru o IA puternică [255]. Pentru a atinge inteligența la nivel uman, IA are nevoie de îndrumarea unui model de realitate, similar cu cele

utilizate în sarcinile de inferență causală. În consecință, lucrarea [228] își propune să: (1) să dezvolte noi tehnici de vizualizare care pot fi instruite de experți medicali, deoarece aceștia pot explora factorii explicativi de bază ai datelor și (2) să oficializeze un model causal structural de luare a deciziilor umane.

Comunitatea de interacțiune om-calculator a stabilit o serie de metode de utilizare [256]. Similar acestor metodologii, metode și teste de utilizare, este nevoie de dezvoltarea unor metodologii, metode și teste de „cauzabilitate”, care se bazează pe principii și teorii științifice clare ale cauzalității, pentru a stabili „cauzabilitatea” ca domeniu științific care va deveni necesar odată cu utilizarea sporită a IA. La fel cum măsurile de utilizare asigură „calitatea utilizării” [257], măsurile de cauzalitate trebuie să asigure „calitatea explicațiilor”.

O abordare promițătoare este în baza ierarhiei cauzale cu trei straturi de Pearl (2018) [241, 255]:

- Nivelul 1: Asocierea $P(y|x)$ cu activitatea tipică de „văz” și întrebări inclusiv „Cum mi-ar schimba dacă văd X încrederea în Y ?”
- Nivelul 2: Intervenția $P(y|do(x), z)$ cu activitatea tipică de „a face” și întrebări inclusiv „Ce se întâmplă dacă fac X ?”, în cazul nostru de utilizare de mai sus, aceasta a fost întrebarea „ce se întâmplă dacă medicul recomandă tratamentul X – pacientul se va vindeca?”
- Nivelul 3: Contrafacticele $P(y_x|x', y')$ cu activitatea tipică de „retrospecție” și întrebări inclusiv „A fost Y cauza pentru X ?”

În sumar [228] conchide că IA este deja una dintre tehnologiile cheie din economia noastră, care va aduce schimbări similare cu introducerea motorului cu abur sau a electricității. Cu toate acestea, preocupările cu privire la potențiala pierdere a controlului în relația om-AI sunt în creștere. Probleme precum conducerea autonomă și luarea deciziilor neclare a vehiculului, de exemplu, în cazuri extreme, cu puțin timp înainte unei coliziuni accidentale, au făcut de multă vreme subiectul dezbaterii publice. Același lucru este valabil și pentru întrebarea în ce măsură IA poate sau ar trebui să sprijine deciziile medicale sau chiar să le ia ea însăși. În multe cazuri, va fi necesar să se înțeleagă cum a fost luată o decizie de mașină și să se evalueze calitatea explicației.

În timp ce soluțiile bazate pe reguli ale inteligenței artificiale timpurii din anii 1950 reprezentau abordări inteligibile „cutie de sticlă”, slăbiciunea lor constă în abordarea incertitudinilor lumii reale. Multe probleme din viața noastră de zi cu zi nu pot fi reprezentate de reguli formale, matematice ale logicii. Eșecul unor astfel de algoritmi de a rezolva probleme care

sunt relativ simple pentru oameni, cum ar fi limbajul natural, recunoașterea fețelor sau înțelegerea unei glume, a dus în cele din urmă la „iarna AI” în anii 1980. Numai prin triumful metodelor de învățare probabilistică și statistică în legătură cu succesul rețelelor neuronale artificiale („învățare profundă”) aplicațiile IA au devenit din ce în ce mai de succes.

Astăzi, algoritmiile ÎP sunt foarte utili în viața de zi cu zi: conducerea autonomă, recunoașterea feței, înțelegerea vorbirii, sistemele de recomandare, etc. și funcționează deja foarte bine. Cu toate acestea, este foarte dificil pentru oameni să înțeleagă cum acești algoritmi ajung la o decizie. În cele din urmă, acestea sunt așa-numitele modele „cutie neagră”. Problema este că, chiar dacă înțelegem principiile și teoriile matematice care stau la baza, unor astfel de modele le lipsește o reprezentare declarativă explicită a cunoștințelor. Soluțiile IA timpurii (numite la acea vreme sisteme-expert) aveau scopul încă de la început de a face soluțiile inteligibile, înțelese și astfel explicabile, ceea ce era posibil și în problemele foarte restrâns definite. IA explicabilă nu este doar utilă și necesară, ci reprezintă și o oportunitate uriașă pentru soluțiile de IA în general. Opacitatea generalizată a IA poate fi astfel redusă și se poate construi încrederea necesară. Exact acest lucru poate promova și acceptarea de către viitorii utilizatori a IA în mod durabil.

Principala problemă a celor mai de succes sisteme IA actuale, subliniată recent de Pearl (2018), este că acestea funcționează pe un model statistic, sau fără model, ceea ce implică limitări severe ale performanței lor. Astfel de sisteme nu sunt capabile să înțeleagă contextul, prin urmare nu pot raționa intervențiile și retrospectivă. Cu toate acestea, astfel de abordări au nevoie de îndrumarea unui model uman similar cu cele utilizate în cercetarea cauzalității [241, 255] pentru a răspunde la întrebarea „De ce?”. În acest context, stabilirea „cauzabilității” ca domeniu științific solid ar putea ajuta.

În continuarea este o încercare de a extrapola cele descrise mai sus la cercetarea curentă. Conform Fig. 5.1, care ilustrează importanța comparativă a 10 predictorilor se vede că de importanță majoră este anume dinamica acestor parametri, cea mai importantă fiind dinamica temperaturii.

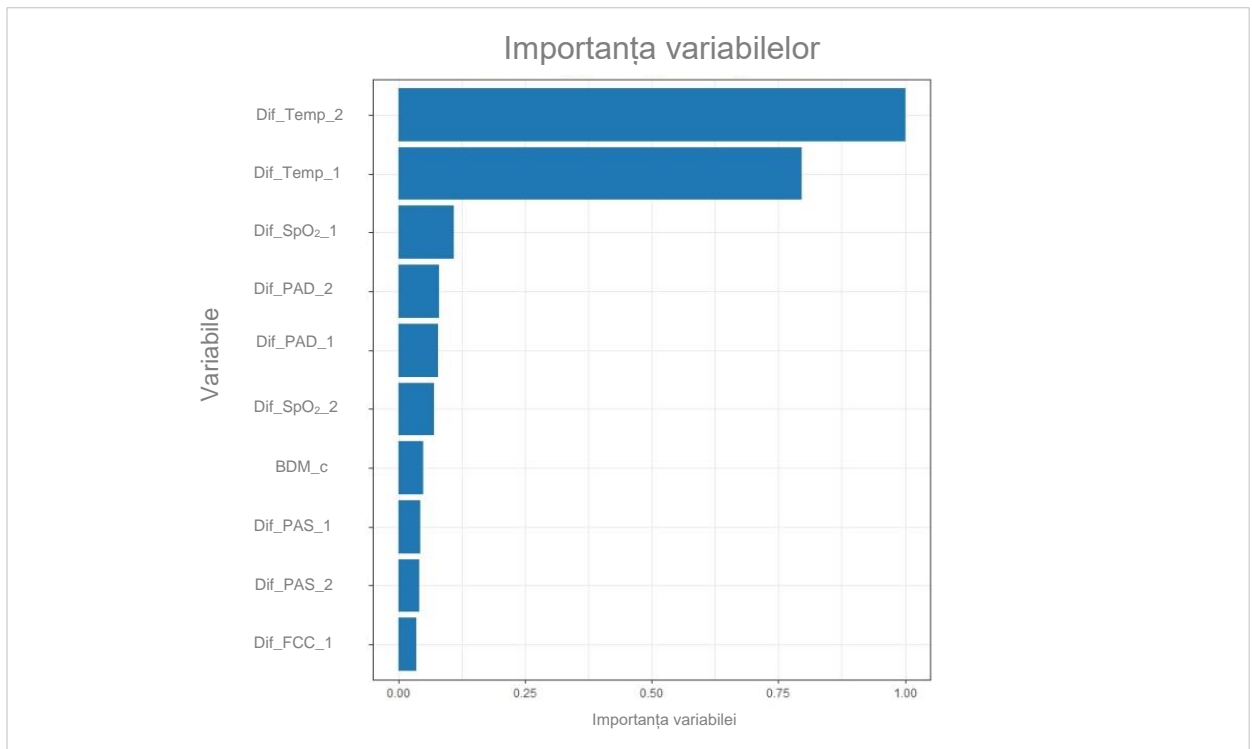


Figura 5.1. Importanța pentru predicție a primelor 10 variabile în cazul unui model GBM.

Notă: Prefixul Dif_ denota diferența valorilor parametrului între trei eșantioane orare consecutive. BDM este complexitatea algoritmică calculată pe matrice 3x3.

Acest fapt este ușor detectabil de către sistem. Orientarea atenției practicianului asupra acestui moment ar putea avea un efect favorabil asupra dezvoltării abilităților predictive ale medicului. Aceasta necesită cercetări suplimentare.

Analiza importanței variabilelor pentru un ansamblu de modele (Fig. 5.2) poate livra informație suplimentară despre aceste modele și a le compara între ele.

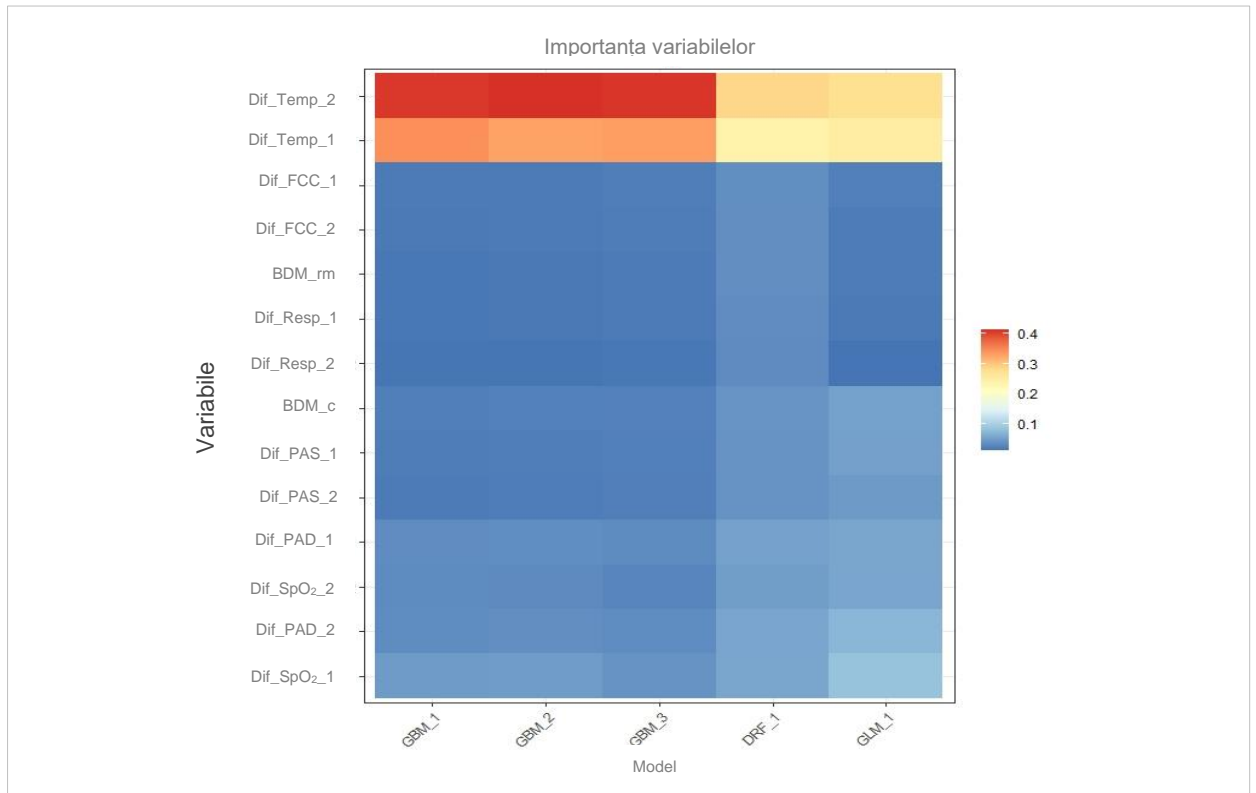


Figura 5.2. **Importanța pentru predicție a variabilelor pentru cinci modele (GBM1/2/3, DRF, GLM)**

Notă: Prefixul *Dif_* denotă diferența valorilor parametrului între trei eșantioane orare consecutive. *BDM* este complexitatea algoritmică calculată pe matrice 3×3 .

Din Fig. 5.2. se poate observa că importanța variabilelor este diferită pentru diferite modele, cu similarități în cazul modelelor ce aparțin aceleși clase (ex. modelele GBM 1, 2 și 3). Această informație poate servi la selectarea unui anumit model în cazul unui predictor de interes special.

O privire mai profundă asupra unui model anumit poate fi obținută cu metodele PD (partial dependence) [258, 259], ICE (individual conditional expectation) [259, 260], LIME (Local Interpretable Model-agnostic Explanations) [259], SHAP (SHapley Additive exPlanations) [135, 259, 261], etc.

Graficul dependenței parțiale (PDP) oferă o reprezentare grafică a efectului marginal al unei variabile asupra răspunsului. Efectul unei variabile este măsurat în dependență de modificarea răspunsului mediu (pe axa Y).

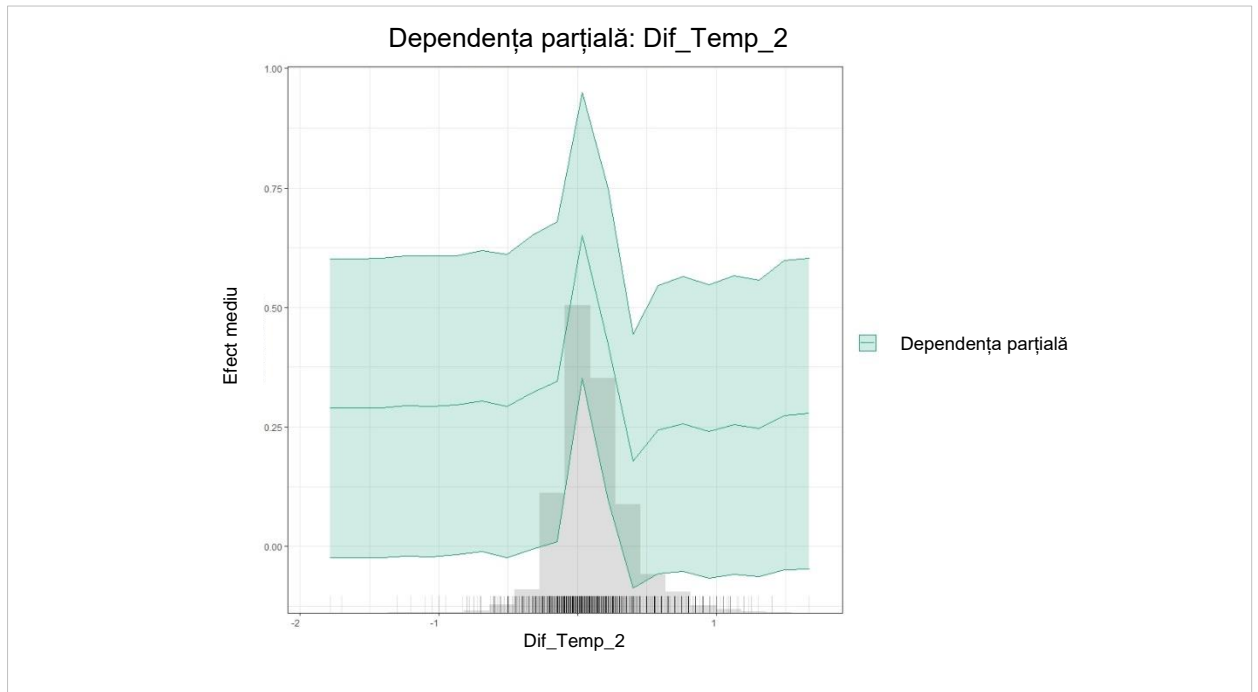


Figura 5.3. **Diagrama PD (dependență parțială) pentru diferența de temperatură**

Diagrama 5.3 prezintă efectul diferenței de temperatură asupra răspunsului. Dacă urmărim de la stânga la dreapta pe axa X (diferența de temperatură de la -2 la 3 grade C) se observă ca valorile mici ale diferenței au cea mai mare valoare predictivă. Aceste diferențe și sunt cele mai des observate în setul de date din cercetarea curentă (codul-bară din partea de jos a diagramei, care arată frecvența observațiilor).

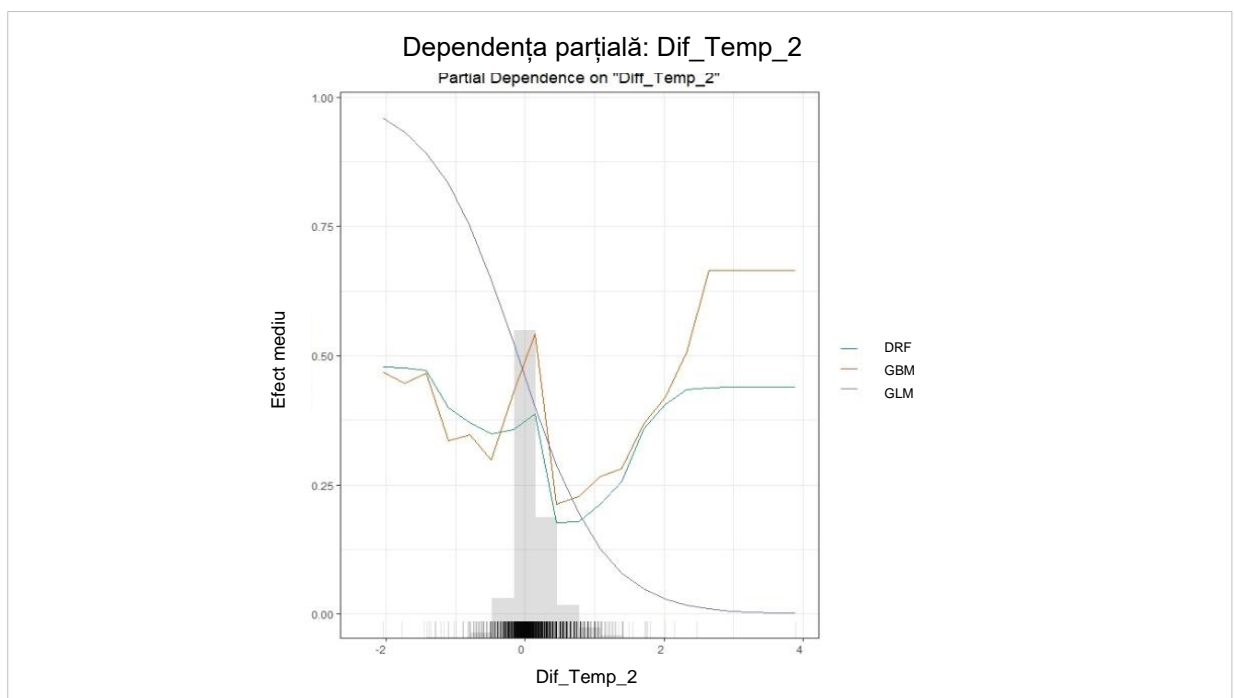


Figura 5.4. **Diagrama PD (dependență parțială) pentru diferența de temperatură în cazul a trei modele (DRF, GBM și GLM) aplicate pe aceleași date.**

Diagrama 5.4. care este asemănătoare cu diagrama precedentă, ilustrează efectul predictor al diferenței de temperatură în cazul a trei modele diferite. După cum reiese din Fig. 5.4 valoarea predictivă a diferenței de temperatură este maximală pentru modelele GBM (Gradient Boosting Machine) și DRF (Distributed Random Forest), când această diferență este în limitele 0.1 – 0.2 grade (această situație fiind observată în majoritatea cazurilor – densitatea observațiilor este ilustrată de barele din partea de jos a diagramei), urmată de o descreștere abruptă a importanței către o diferență de 0.5 grade, ca mai apoi să crească gradual cu atingerea unui nou maximum la o diferență de aproximativ 2.6 grade. Pentru o diferență negativă situația este aproape opusă (în oglindă). Acest tipar este destul de aproape pentru modele GBM și DRF, ambele fiind modele în baza arborilor decizionali. Pentru modelul GLM (Generalized Linear Model), bazat pe regresia liniară, tabloul este esențial diferit – maximumul importanței revine diferenței negative de temperatură (-2 grade), așa cazuri în setul cercetat fiind destul de puține/rare, cu o descreștere lineară spre valori pozitive ale acestei diferențe.

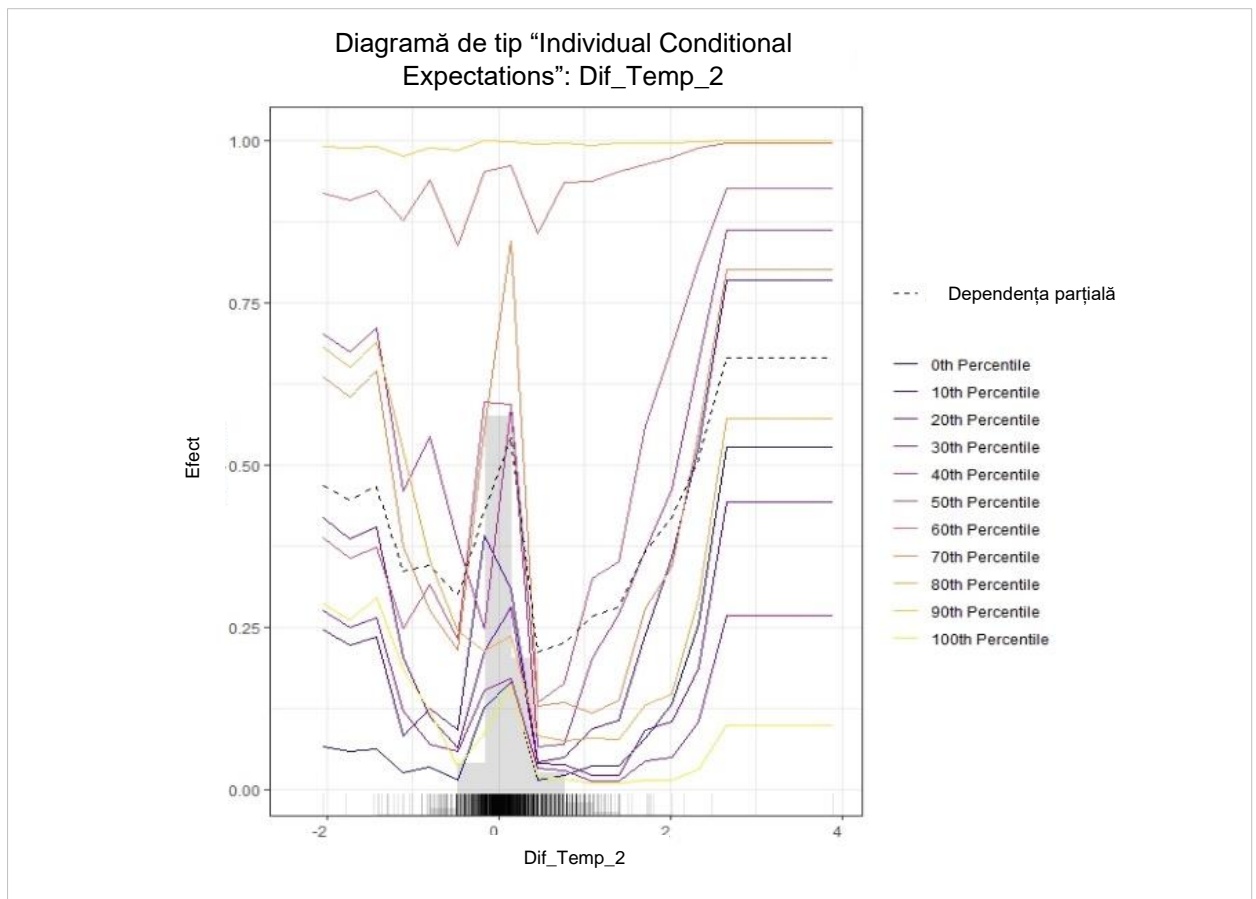


Figura 5.5. Diagrama ICE (individual conditional expectation) pentru diferența de temperatură în cazul unui model GBM cu performanță înaltă.

Diagrama de aşteptări condiționate individuale (ICE), de asemenea, oferă o reprezentare grafică a efectului marginal al unei variabile asupra răspunsului. Diagramele ICE sunt similare diagramelor de dependență parțială (PDP). PDP arată efectul mediu al unei caracteristici, în timp ce diagrama ICE arată efectul pentru o singură instanță. Această funcție reprezintă grafic efectul pentru fiecare decil. Spre deosebire de PDP, diagramele ICE pot oferi mai multe informații, mai ales atunci când există o interacțiune mai puternică a caracteristicilor.

Esența diagramei 5.5 este aproape de esența diagramelor precedente (5.3 și 5.4) în cazul unui model GBM. Aici este ilustrată importanța diferenței de temperatură în dependență de locul pe care îl ocupă ea în cazul ordonării distribuției diferenței sub formă de percentile. În diagramă percentilul 70 este valoarea diferenței cu cel mai detaliat și înalt grad de discriminare.

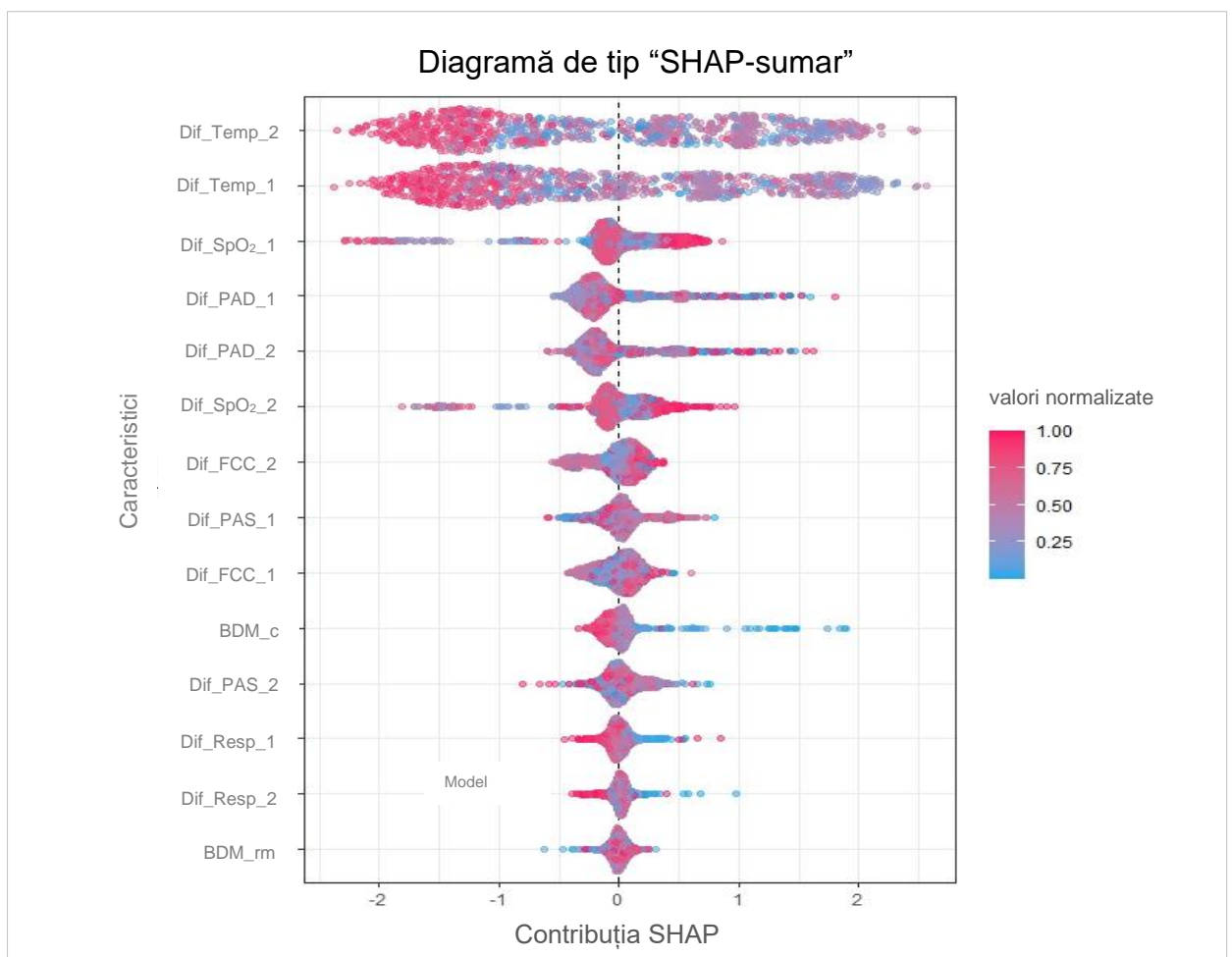


Figura 5.6. Diagrama SHAP-sumar pentru unul din modelele GBM cu performanță înaltă

Diagrama 5.6 (SHAP-sumar) îmbină importanța caracteristicilor (poziția caracteristicii pe axa Y; cu cât mai sus – cu atât mai important) cu efectele caracteristicilor (impactul asupra predicției - pe axa X; spre dreapta de linia verticală “0” – în favoarea sepsisului, iar spre stânga –

în defavoarea sepsisului și cu cât mai departe de “0” – cu atât impactul este mai mare). Culoarea reprezintă valoarea normalizată a caracteristicii de la mică la mare. Astfel, în cazul temperaturii, o diferență de mărime cât mai mare (spre exemplu, când temperatura corpului crește esențial și rapid) se traduce printr-un risc de sepsis jos. Mergând de la contrariu, dinamica modestă a temperaturii ar putea denota situația când forțele de apărare ale organismului pacientului septic sunt epuizate, nefiind în stare să asigure o reacție adecvată (la agentul patogen), exprimată în acest caz prin reacția piretică.

Deosebit de ilustrativă este situația cu complexitatea algoritmică estimată în baza parametrilor circulatori (BDM_c). Valori mari ale complexității algoritmice (punctele de culoare roșie) sunt asociate cu un risc de sepsis scăzut, și, invers, o complexitate algoritmică joasă (punctele de culoare albastră) – cu risc înalt.

În Anexa 11 sunt prezentate mai multe exemple de diagrame PD, ICE pentru caracteristicile de importanță pentru modelele create și, de asemenea, diagrame de corelare dintre predicțiile modelelor. Pentru ultimele se folosește frecvența predicțiilor identice. În acest caz, în mod implicit, modelele sunt ordonate după asemănarea lor, estimată prin clasterizare ierarhică.

5.3. Transferabilitatea sistemelor IA. Direcții de cercetare de viitor

Un aspect separat discutat în literatură îl constituie transferabilitatea SÎA, adică folosirea unui sistem creat în baza datelor dintr-un spital în alt spital. De obicei, în acest caz performanța sistemului în condiții noi este mai mică, potențialele cauze ale acestui fenomen fiind descrise în continuare.

Învățarea automată a apărut ca un instrument promițător pentru detectarea precoce a apariției sepsisului pe baza înregistrărilor medicale electronice, a datelor de laborator și a semnalelor biomedicale [28, 33, 119, 139, 264-267]. Mai multe studii prospective au arătat că implementarea algoritmilor de predicție a sepsisului bazați pe învățarea automată poate reduce mortalitatea spitalicească și durata aflării pacientului în unitatea ATI [152, 268]. Cu excepția performanței excelente de predicție, traducerea acestor modele de predicție a riscului în practica clinică necesită validare externă independentă pentru a determina generalizarea modelului la diferite cohorte [269]. Cu toate acestea, majoritatea modelelor de predicție a riscului propuse recent au performanțe slabe atunci când sunt aplicate la eșantioane externe [270].

Reantrenarea modelului de predicție pe seturi de date locale în centrul medical țintă ar putea îmbunătăți acuratețea predictivă în situații specifice [263, 271]. Pe baza aceluiași obiectiv, s-a raportat că învățarea prin transfer îmbunătățește performanța modelului atunci când setul de date este mic în centrul medical țintă [34, 272]. În plus, interpretabilitatea modelelor de învățare

automată reflectă măsura în care procesul decizional al modelului poate fi înțeles și acceptat în practica clinică. Lipsa de interpretabilitate pentru majoritatea modelelor de predicție disponibile este în prezent o barieră majoră în calea adoptării lor clinice [263, 273, 274].

Lucrarea [263], care descrie mai multe detalii ale învățării prin transfer, folosește pentru predicția sepsisului modele create în baza algoritmilor LightGBM și MLP (multilayer perceptron) pe baza considerațiilor integrate ale performanței, inclusiv acuratețea, aria de sub curba ROC (AUROC), sensibilitatea și viteza de inferență. Conform acestui studiu procesul de învățare prin transfer poate fi împărțit în două etape. În primul rând, se efectuează normalizarea datelor, astfel încât toate caracteristicile să fie scalate între 0 și 1. În al doilea rând, se efectuează transferul parametrilor modelului. Mai exact, pentru LightGBM, care a fost antrenat pe un set de date inițiale, se efectuează reglaj fin în baza altui set de date. De asemenea, a fost folosită o metodă de învățare prin ansamblu pentru a integra modelele LightGBM și MLP folosind media de rezultatelor inferenței. Modelul de ansamblu este folosit pentru predicția practică a sepsisului în condiții noi.

Performanța modelelor de predicție a sepsisului a fost evaluată pe baza acurateței, ariei de sub curba ROC, sensibilității și specificității pe setul de testare. Pragul implicit de clasificare este 0,5. În acest studiu, modelele bazate pe GBM (XGBoost și LightGBM) le-au depășit pe altele.

Pentru validarea externă a modelelor LightGBM și MLP, s-a evaluat performanțele lor de predicție pe setul de date de la centrul medical țintă. Aria de sub curba ROC a fost 0,82-0,86 (pe un set de predicție) și, respectiv, 0,84-0,87 (pe alt set) cu un orizont de predicție de 1-5 ore precedente sepsisului, indicând performanța substanțial înrăutățită a acestor modele atunci când sunt aplicate cohortelor independente externe (vs 0.98 pe setul de antrenament).

Pentru a îmbunătăți performanța de predicție aceste modele au fost reantrenate folosind învățarea prin transfer și tehnica de învățare prin ansamblu. În acest studiu învățarea prin transfer a îmbunătățit performanța de predicție la implementarea modelelor derivate din setul de date publice (inițial pentru training) către spitalul țintă, AUC-urile finale ale modelului de predicție a sepsisului fiind 0,94.

De asemenea s-a efectuat evaluarea performanței de predicție în lumea reală (în spital). Fiecare pacient a fost etichetat de modificarea scorului SOFA și de examinarea medicului pentru infecție. În acest caz performanța (după AUCROC) a fost 0,86-0,90.

Învățarea automată este considerată o metodă promițătoare pentru predicția sepsisului în ATI iar diagnosticul precoce și managementul în timp util al pacienților septici pot îmbunătăți în mod eficient prognosticul [154].

Conform aceluiași studiu [264], generalizarea este obstacolul major în calea implementării învățării automate în practica medicală. Dimensiunea suficient de mare a datelor este crucială

pentru formarea modelului de învățare automată pentru a obține performanțe bune. Mai mult, performanța modelului derivat dintr-o singură cohortă este întotdeauna mai mică atunci când este aplicat la cohorte externe independente din cauza diferențelor de rasă, mediu medical, tip de boală și severitate a bolii în diferite cohorte. În plus, pe măsură ce se adaugă noi teste și tehnici, noi funcții pot ajuta sarcina de predicție, dar includerea directă în modelele de învățare automată existente nu este de obicei fezabilă. În așa situații implementarea tehnicii de învățare prin transfer pentru a îmbunătăți performanța modelelor în centrele medicale țintă pare a fi o alternativă [263].

În studiul citat procesul de învățare prin transfer a îmbunătățit în mod eficient performanța de predicție ale modelelor LightGBM și MLP antrenate pe un set de date și apoi reglate fin cu utilizarea datelor din cohorta țintă. Prin urmare, învățarea prin transfer ar putea fi o strategie promițătoare și fezabilă pentru a menține eficiența implementării trans-centrice a modelelor de învățare automată. Mai mult, învățarea prin transfer a fost aplicată în domenii similare sau sarcini similare în mai multe domenii medicale, reducând cerințele de dimensiune ale setului de date țintă și îmbunătățind viteza de antrenament și performanța de predicție [275-278].

Deci, soluții pentru rezolvarea problemei cu performanța în acest caz sunt propuse cel puțin două: (a) crearea unui model nou în baza datelor noi, folosind cunoștințele acumulate la elaborarea modelelor similare în alte studii (adică replicarea unui studiu de acum realizat pe date noi); (b) învățarea prin transfer, când un model de acum creat este suplimentar antrenat pe date noi.

Eventualele direcții de cercetare pe viitor (ca continuarea a cercetării curente) ar include: (a) calibrarea sistemului de prezicere a sepsisului creat în studiul curent în condiții clinice în Republica Moldova; (b) la necesitate, îmbunătățirea performanței sistemului prin învățarea prin transfer în baza datelor locale; (c) studierea interacțiunii specialistului ATI cu sistemul în condiții clinice.

5.4. Sinteza Capitolului 5

Capitolul 5 este axat pe problema explicabilității modelelor IA și direcțiile de cercetare de viitor și începe cu un sumar al cercetării curente pe etape.

Analiză comparativă a sistemului elaborat în cercetare este realizată prin compararea lui cu cinci sisteme similare cu performanță înaltă, raportate în literatură în ultimii trei ani.

Este descris aspectul de explicabilitate a sistemelor IA, ca factor-cheie pentru adoptarea acestor sisteme de către comunitatea medicală. După o descriere a elementelor conceptuale, explicabilitatea sistemului elaborat este ilustrată și discutată prin prisma următoarelor aspecte:

- Ponderea variabilelor în emiterea unei anumite predicții (Fig. 5.1 și 5.2);
- Dependența parțială (eng. Partial dependence) a predicției de valoarea unui anumit predictor (Fig. 5.3 și 5.4);

- Aspectul de așteptări individuale condiționale (eng, Individual conditional expectations – Fig. 5.5.)
- Diagramele SHAP (eng. SHapley Additive exPlanations – Fig. 5.6), care evaluează contribuția fiecărui predictor la decizia/predicția realizată de model

Printre limitările sistemului una dintre cele mai importante pare a fi condiționată de faptul ca datele, în baza cărora a fost elaborat sistemul de predicție curent, nu includ cazuri de sepsis viral, aspect discutat recent în literatura de specialitate. Odată ce se va ajunge la certitudine referitor la similaritățile și deosebirile dintre sepsisul bacterian și cel viral, cel mai probabil că sistemul va necesita ajustare, care poate fi una din direcțiile de dezvoltare de viitor. Alte limitări se referă la caracteristicile datelor pe care a fost creat sistemul vs datele cazurilor din secțiile țintă, unde el ar putea fi potențial utilizat (ex. pacienți diferiți, abordări/tehnici/echipament de monitoring diferite, etc.).

Astfel, eventualele direcții de cercetare pe viitor ar include: (a) calibrarea sistemului de prezicere a sepsisului creat în studiul curent în condiții clinice în Republica Moldova; (b) îmbunătățirea performanței sistemului prin învățarea prin transfer în baza datelor locale; (c) studierea interacțiunii specialistului ATI cu sistemul în condiții clinice.

Informație mai detaliată despre unele aspecte descrise în acest capitol este prezentată în Anexele 9-11.

CONCLUZII GENERALE

1. Fenomenul pătrunderii tehnologiilor IA în ATI este unul relativ nou, însă de proporții și cu perspective greu de prezis. Aspectele activității medicului anesteziolog-reanimatolog, care poate beneficia de aceste tehnologii sunt numeroase, pornind de la diagnosticul/monitorizarea/predicția diferitor fațete ale stării critice. Conform literaturii de specialitate, cea mai frecvent utilizată la etapa actuală tehnologie de IA este învățarea automată supravegheată, care se regăsește în peste 90% de lucrări.
2. Algoritmul elaborat și utilizat în cercetarea curentă pentru restabilirea datelor biomedicale continui cu valori lipsă, parvenite de la pacienți cu sepsis, furnizează rezultate plauzibile, fapt confirmat de performanța înaltă a sistemului de predicție. În baza rezultatelor obținute (modificarea neînsemnată a performanței sistemului la testarea lui pe date-restabilite cu algoritmul vs cazuri complete) se poate de afirmat că dacă și există părtinire introdusă de algoritm, ea este minoră.
3. Metoda de reprezentare și procesare a datelor în formă de serii temporale a parametrilor fiziologici la pacienții septici, cu ajutorul metricii de complexitate algoritmică Kolmogorov - Chaitin, care a fost în premieră utilizată în asemenea scop în cercetarea curentă, este o metodă reușită de prelucrare a datelor pentru învățarea automată, fapt demonstrat de rezultatele finale obținute.
4. Cei șase parametri fiziologici, reprezentați prin datele - valori în timp ale lor, au fost suficienți pentru crearea unui sistem de predicție performant, utilizând metodele de procesare descrise mai sus. De importanță este și aceia ca parametrii respectivi sunt de rutină monitorizați în ATI, inclusiv și în Republica Moldova.
5. Performanță înaltă (mai mare de 92% vs 89% în ultimul Ghid pentru managementul sepsisului și șocului septic, 2021) a sistemului de predicție creat face rațională cercetarea de mai departe cu calibrarea sistemului în UTI țintă din Moldova și posibilitatea ulterioară de utilizare în practica clinică pentru predicția timpurie a unui risc iminent de dezvoltare a sepsisului cu inițierea în timp util a tratamentului (în primul rând aspectele care se referă la infecție). Aceasta poate avea un impact pozitiv asupra rezultatelor tratamentului, concomitent asigurând suport decizional pentru medicul anesteziolog-reanimatolog.

RECOMANDĂRI PRACTICE

1. Atunci când cercetarea are de a face cu date lipsă și în special în cazul datelor reprezentate prin serii temporale, care în ATI sunt ubicuitare, utilizarea algoritmului de reconstrucție-restabilire a datelor ar putea facilita lucrul cercetătorului și contribui la amplificarea performanței lui.
2. Complexitatea algoritmică ca o posibilă metrică a proceselor ce derulează în subiectul de studiu și reprezentarea datelor obținute cu ajutorul ei (spre ex. cu ajutorul metodei de decompoziție în blocuri – BDM) are un potențial încă puțin explorat. De aceea experimentarea cu aceste instrumente chiar și în studii în afara domeniului IA este binevenită. Ea ar putea reuși să suplimenteze și chiar întrecă metodele statistice tradiționale.
3. În pofida faptului că procesul de raționament al clinicianului probabil diferă esențial de logica computațională, iar performanțele de calcul nu pot fi comparate cu cele ale unui calculator, atragerea atenției asupra dinamicii factorilor cu cea mai mare valoare predictivă determinați de model (ex. dinamica subtilă a temperaturii, a saturației sângelui periferic cu oxigen – SpO₂, a presiunii arteriale diastolice, etc) ar putea fortifica abilitățile de prognozare (subconștiente) ale specialistului uman în cadrul managementului clinic al pacienților cu sepsis. În acest context importante sunt și aspectele legate de explicabilitatea modelelor.
4. La elaborarea unor sisteme de suport decizional similar, pentru utilizare practică în ATI este recomandabil de folosit ghidul TRIPOD și principiile PROBAST, care de fapt sunt aplicabile și pentru studii ce nu țin de IA, de rând cu raționamentul clinic și bunele practici ale medicinei.
5. Numărul de tehnologii și dispozitive în baza învățării automate și IA va crește și în acest context este necesară o atitudine corectă a practicianului față de aceste dispozitive. Una reușită pare a fi acceptarea lor în calitate de asistenți, și nicidecum - rivali, cu atât mai mult ca cel care ia decizia finală și o realizează, cel puțin la moment și în viitorul previzibil, rămâne să fie specialistul uman.

BIBLIOGRAFIE

1. MOOR, M. *Machine Learning on Clinical Time Series: Classification and Representation Learning*: doctoral thesis in med. sciences. ETH Zurich, 2022. 188 p.
2. *Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices*, [online]. FDA, 2022 [citat 10.04.2022]. Disponibil: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
3. NAGENDRAN, M. et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. In: *BMJ*. 2020, 368: m689 ISSN 1756-1833.
4. *ReportLinker. Artificial intelligence in healthcare market by offering, technology, end-use application, end user and geography – global forecast to 2027* [online]. 2022 [citat 10.04.2022]. Disponibil: <https://www.reportlinker.com/p04897122/Artificial-Intelligence-in-Healthcare-Market-by-Offering-Technology-Application-End-User-Industry-and-Geography-Global-Forecast-to.html>
5. CONNOR, W. Artificial Intelligence and Machine Learning in Anesthesiology. In: *Anesthesiology*. 2019, 131:1346–59 ISSN 1365-2044.
6. EICHHORN, J. et al. Standards for patient monitoring during anesthesia at Harvard Medical School. In: *JAMA*. 1986, 256:1017–20 ISSN 0098-7484.
7. HASHIMOTO, D. Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations. In: *Anesthesiology*. 2020,132(2):379-394. ISSN 1365-2044.
8. LOVEJOY, C. et al. Artificial intelligence in the intensive care unit. In: *Critical Care*. 2019, 23:7 ISSN 1364-8535.
9. HONG, N. et al. State of the Art of Machine Learning–Enabled Clinical Decision Support in Intensive Care Units: Literature Review. In: *JMIR Med Inform*. 2022, 10(3): e28781. ISSN 1438-8871.
10. ESHELMAN, L. et al. Development and evaluation of predictive alerts for hemodynamic instability in ICU patients. In: *AMIA Annu Symp Proc*. 2008, 06:379-383 ISSN 1067-5027.
11. QUINN, J., WILLIAMS, C., MCINTOSH, N. Factorial switching linear dynamical systems applied to physiological condition monitoring. In: *IEEE Trans Pattern Anal Mach Intell*. 2009, 31(9):1537-1551. ISSN 1939-3539.
12. CHARBONNIER, S. On line extraction of temporal episodes from ICU high-frequency data: a visual support for signal interpretation. In: *Comput Methods Programs Biomed*. 2005, 78(2):115-132. ISSN 1872-7565.
13. ZHANG, Y., SZOLOVITS, P. Patient-specific learning in real time for adaptive monitoring in critical care. In: *J Biomed Inform*. 2008, 41(3):452-460. ISSN 1532-0464
14. KWOK, H., LINKENS, D., MAHFOUF, M., MILLS G. Adaptive ventilator FiO2 advisor: use of non-invasive estimations of shunt. In: *Artif Intell Med*. 2004, 32(3):157-169. ISSN 1873-2860.
15. REHM, G. et al. Creation of a robust and generalizable machine learning classifier for patient ventilator asynchrony. In: *Methods Inf Med*. 2018, 57(4):208-219. ISSN 0026-1270.
16. GHOLAMI, B. et al. Replicating human expertise of mechanical ventilation waveform analysis in detecting patient-ventilator cycling asynchrony using machine learning. In: *Comput Biol Med*. 2018, 01;97:137-144. ISSN 1879-0534.

17. KOOLEN, N. et al. Automated classification of neonatal sleep states using EEG. In: *Clin Neurophysiol.* 2017, 128(6):1100-1108. ISSN 1388-2457
18. FARZANEH, N. et al. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. In: *Diagnostics.* 2020, 10(10):773. ISSN 2075-4418
19. GOLMOHAMMADI, M. et al. Automatic analysis of EEGs using big data and hybrid deep learning architectures. In: *Front Hum Neurosci.* 2019, 13:76, ISSN 1662-5161
20. SORANI, M. et al. New approaches to physiological informatics in neurocritical care. In: *Neurocrit Care.* 2007, 7(1):45-52, ISSN 1556-0961.
21. RUECKEL, J. et al. Artificial intelligence algorithm detecting lung infection in supine chest radiographs of critically ill patients with a diagnostic accuracy similar to board-certified radiologists. In: *Crit Care Med.* 2020, 48(7):e574-e583, ISSN 0090-3493.
22. CALVERT, J. et al. High-performance detection and early prediction of septic shock for alcohol-use disorder patients. In: *Ann Med Surg (Lond).* 2016, 8:50-55, ISSN 2049-0801
23. TIMSIT, J., PERNER, A. Sepsis: find me, manage me, and stop me!. In: *Intensive Care Med.* 2016, 24;42(12):1851-1853, ISSN 0342-4642
24. SOLIMAN, I. et al. The predictive value of early acute kidney injury for long-term survival and quality of life of critically ill patients. In: *Crit Care.* 2016, 03;20(1):242, ISSN 1364-8535.
25. SUN, M. et al. Early prediction of acute kidney injury in critical care setting using clinical notes and structured multivariate physiological measurements. In: *Stud Health Technol Inform.* 2019, 21;264:368-372, ISSN 0926-9630.
26. SANCHEZ-PINTO, L., KHEMANI, R. Development of a prediction model of early acute kidney injury in critically ill children using electronic health record data. In: *Pediatr Crit Care Med.* 2016, 17(6):508-515, ISSN 1529-7535.
27. FADLALLA, A., GOLOB, J., CLARIDGE, J. Enhancing the fever workup utilizing a multi-technique modeling approach to diagnose infections more accurately. In: *Surg Infect (Larchmt).* 2012, 13(2):93-101, ISSN 1096-2964.
28. NEMATI, S. et al. An interpretable machine learning model for accurate prediction of sepsis in the ICU. In: *Crit Care Med.* 2018, 46(4):547-553, ISSN 0090-3493.
29. KAM, H., KIM, H. Learning representations for the early detection of sepsis with deep neural networks. In: *Comput Biol Med.* 2017, 01;89:248-255, ISSN 1879-0534.
30. KAJI, D. et al. An attention based deep learning model of clinical events in the intensive care unit. In: *PLoS One.* 2019, 14(2):e0211057, ISSN 1932-6203.
31. SCHERPF, M., GRÄBER, F., MALBERG, H., ZAUNSEDER, S. Predicting sepsis with a recurrent neural network using the MIMIC III database. In: *Comput Biol Med.* 2019, 113:103395, ISSN 1879-0534.
32. WANG, S., WU, F., WANG, B. Prediction of severe sepsis using SVM model. In: *Adv Exp Med Biol.* 2010, 680:75-81, ISSN 0065-2598.
33. DESAUTELS, T. et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. In: *JMIR Med Inform.* 2016, 30;4(3):e28. ISSN 1438-8871.
34. MAO, Q. et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. In: *BMJ Open.* 2018, 26;8(1):e017833, ISSN 2044-6055.

35. METSVAHT, T. et al. Clinical parameters predicting failure of empirical antibacterial therapy in early onset neonatal sepsis, identified by classification and regression tree analysis. In: *BMC Pediatr.* 2009, 24;9:72, ISSN 1471-2431.
36. MANI, S. et al. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. In: *J Am Med Inform Assoc.* 2014, 21(2):326-336, ISSN 1067-5027.
37. SHAHIN, J. et al. FIRE Study Investigators. Predicting invasive fungal disease due to *Candida* species in non-neutropenic, critically ill, adult patients in United Kingdom critical care units. In: *BMC Infect Dis.* 2016, 09;16:480, ISSN 1471-2334.
38. MESSINGER, A. et al. Novel pediatric-automated respiratory score using physiologic data and machine learning in asthma. In: *Pediatr Pulmonol.* 2019, 54(8):1149-1155, ISSN 2949-8074.
39. SAUTHIER, M., JOUVET, P., NEWHAMS, M., RANDOLPH, A. Machine learning predicts prolonged acute hypoxemic respiratory failure in pediatric severe influenza. In: *Crit Care Explor.* 2020 2(8):e0175, ISSN 2639-8028.
40. LE, S. et al. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS). In: *J Crit Care.* 2020, 60:96-102, ISSN 1557-8615.
41. HSU, J. et al. Clinical verification of a clinical decision support system for ventilator weaning. In: *Biomed Eng Online.* 2013;12 Suppl 1:S4, ISSN 1475-925X.
42. MIU, T. et al. Predictors of reintubation in critically ill patients. In: *Respir Care.* 2014, 59(2):178-185, ISSN 0020-1324.
43. ISBISTER, G., DUFFULL, S. Quetiapine overdose: predicting intubation, duration of ventilation, cardiac monitoring and the effect of activated charcoal. In: *Int Clin Psychopharmacol.* 2009, 24(4):174-180, ISSN 0268-1315.
44. GHAZAL, S. et al. Using machine learning models to predict oxygen saturation following ventilator support adjustment in critically ill children: A single center pilot study. In: *PLoS One.* 2019, 14(2):e0198921, ISSN 1932-6203.
45. RODRÍGUEZ, A. et al. Grupo Español de Trabajo Gripe A Grave (GETGAG)/Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias (SEMICYUC) Working Group, 2009-2015 H1N1 SEMICYUC Working Group investigators. Risk factors for noninvasive ventilation failure in critically ill subjects with confirmed influenza infection. In: *Respir Care.* 2017, 11;62(10):1307-1315, ISSN 0020-1324.
46. LIN, P. et al. A machine learning approach for predicting urine output after fluid administration. In: *Comput Methods Programs Biomed.* 2019, 177:155-159, ISSN 1872-7565.
47. PAPPADA, S. et al. An artificial neural network-based predictive model to support optimization of inpatient glycemic control. In: *Diabetes Technol Ther.* 2020, 22(5):383-394, ISSN 1520-9156.
48. MAMANDIPOOR, B., MAJD, M., MOZ, M., OSMANI, V. Blood lactate concentration prediction in critical care. In: *Stud Health Technol Inform.* 2020, 16;270:73-77, ISSN 0926-9630
49. SU, L. et al. Toward optimal heparin dosing by comparing multiple machine learning methods: retrospective study. In: *JMIR Med Inform.* 2020, 22;8(6):e17648, ISSN 1438-8871.
50. YU, L., ZHANG, Q., BERNSTAM, E., JIANG, X. Predict or draw blood: an integrated method to reduce lab tests. In: *J Biomed Inform.* 2020, 104:103394, ISSN 1532-0464.

51. CISMONDI, F. et al. Reducing unnecessary lab testing in the ICU with artificial intelligence. In: *Int J Med Inform.* 2013, 82(5):345-358, ISSN 1872-8243.
52. LI, K. et al. A machine learning-based model to predict acute traumatic coagulopathy in trauma patients upon emergency hospitalization. In: *Clin Appl Thromb Hemost.* 2020, 26:1076029619897827, ISSN 1076-0296.
53. OH, S. et al. Automatic delirium prediction system in a Korean surgical intensive care unit. In: *Nurs Crit Care.* 2014, 19(6):281-291, ISSN 1478-5153.
54. MILBRANDT, E. et al. Predicting late anemia in critical illness. In: *Crit Care.* 2006, 10(1):R39, ISSN 1364-8535.
55. PUSKARICH, M. A decision tree incorporating biomarkers and patient characteristics estimates mortality risk for adults with septic shock. In: *Evid Based Nurs.* 2015, 18(2):42, ISSN 1367-6539.
56. WONG, H. et al. A multibiomarker-based outcome risk stratification model for adult septic shock. In: *Crit Care Med.* 2014, 42(4):781-789, ISSN 0090-3493.
57. JAIMES, F., FARBIARZ, J., ALVAREZ, D., MARTÍNEZ, C. Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. In: *Crit Care.* 2005, 9(2):R150-R156, ISSN 1364-8535.
58. RIBAS RIPOLL V., VELLIDO, A., ROMERO, E., RUIZ-RODRÍGUEZ, J. Sepsis mortality prediction with the Quotient Basis Kernel. In: *Artif Intell Med.* 2014, 61(1):45-52, ISSN 1873-2860.
59. SHA, Y., VENUGOPALAN, J., WANG, M. A novel temporal similarity measure for patients based on irregularly measured data in electronic health records. In: *ACM BCB.* 2016, 2016:337-344, ISBN 978-1-4503-4225-4.
60. YANG, T. et al. Biomarkers upon discontinuation of renal replacement therapy predict 60-day survival and renal recovery in critically ill patients with acute kidney injury. In: *Hemodial Int.* 2018, 22(1):56-65, ISSN 1492-7535.
61. XU, Z. et al. Stratified mortality prediction of patients with acute kidney injury in critical care. In: *Stud Health Technol Inform.* 2019, 21;264:462-466, ISSN 0926-9630.
62. TRONGTRAKUL, K. et al. Acute kidney injury risk prediction score for critically-ill surgical patients. In: *BMC Anesthesiol.* 2020, 03;20(1):140, ISSN 1471-2253.
63. BERNAL, W. et al. Development and validation of a dynamic outcome prediction model for paracetamol-induced acute liver failure: a cohort study. In: *Lancet Gastroenterol Hepatol.* 2016, 1(3):217-225, ISSN 2468-1253.
64. LINDENMEYER, C. et al. LIV-4: A novel model for predicting transplant-free survival in critically ill cirrhotics. In: *World J Hepatol.* 2020, 27;12(6):298-311, ISSN 1948-5182.
65. BALEKIAN, A., GOULD, M. Predicting in-hospital mortality among critically ill patients with end-stage liver disease. In: *J Crit Care.* 2012, 27(6):740.e1-740.e7, ISSN 1557-8615.
66. SANTOS, H. et al. Machine learning to predict 30-day quality-adjusted survival in critically ill patients with cancer. In: *J Crit Care.* 2020, 55:73-78, ISSN 1557-8615.
67. VINCENT, F. et al. GrrrOH: Groupe de recherche respiratoire en réanimation en Onco-Hématologie (Group for respiratory research in intensive care in Onco-Hematology, <http://www.grrroh.com/>). In-hospital and day-120 survival of critically ill solid cancer patients after discharge of the intensive care units: results of a retrospective multicenter study-A Groupe de recherche respiratoire en réanimation en Onco-Hématologie (Grrr-OH) study. In: *Ann Intensive Care.* 2018, 27;8(1):40, ISSN 2110-5820.
68. LEE, S. et al. Quantitative EEG predicts outcomes in children after cardiac arrest. In: *Neurology.* 2019,14;92(20):e2329-e2338, ISSN 1432-1459.

69. MURTUZA, B. et al. The importance of blood lactate clearance as a predictor of early mortality following the modified Norwood procedure. In: *Eur J Cardiothorac Surg*. 2011, 40(5):1207-1214, ISSN 1010-7940.
70. GRACIA ARNILLAS, M. et al. Impact of adrenomedullin levels on clinical risk stratification and outcome in subarachnoid haemorrhage. In: *Eur J Clin Invest*. 2020, 50(11):e13318, ISSN 0014-2972.
71. HAVEMAN, M. et al. Predicting outcome in patients with moderate to severe traumatic brain injury using electroencephalography. In: *Crit Care*. 2019, 11;23(1):401, ISSN 1364-8535.
72. WILDMAN, M. et al. Predicting mortality for patients with exacerbations of COPD and asthma in the COPD and asthma outcome study (CAOS). In: *QJM*. 2009, 102(6):389-399, ISSN 1460-2393.
73. DALY, K., BEALE, R., CHANG, R. Reduction in mortality after inappropriate early discharge from intensive care unit: logistic regression triage model. In: *BMJ*. 2001, 26;322(7297):1274-1276, ISSN 1756-1833.
74. HERNÁNDEZ-TEJEDOR, A. et al. EPIPUSE study group. Evolution and prognosis of long intensive care unit stay patients suffering a deterioration: a multicenter study. In: *J Crit Care*. 2015, 30(3):654.e1-654.e7, ISSN 1557-8615.
75. JI, S., SMITH, R., HUYNH, T., NAJARIAN, K. A comparative analysis of multi-level computer-assisted decision making systems for traumatic injuries. In: *BMC Med Inform Decis Mak*. 2009, 14;9:2, ISSN 1472-6947.
76. CHE, Z., PURUSHOTHAM, S., KHEMANI, R., LIU, Y. Interpretable deep models for ICU outcome prediction. In: *AMIA Annu Symp Proc*. 2016, 2016:371-380. ISSN 1559-4076
77. EBADOLLAHI, S. et al. Predicting patient's trajectory of physiological data using temporal trends in similar patients: a system for near-term prognostics. In: *AMIA Annu Symp Proc*. 2010, 13;2010:192-196. ISSN 1559-4076
78. CASTIÑEIRA, D. et al. Adding continuous vital sign information to static clinical data improves the prediction of length of stay after intubation: a data-driven machine learning approach. In: *Respir Care*. 2020, 65(9):1367-1377, ISSN 0020-1324.
79. MUELLER, M. et al. Parameter selection for and implementation of a web-based decision-support tool to predict extubation outcome in premature infants. In: *BMC Med Inform Decis Mak*. 2006, 01;6:11, ISSN 1472-6947.
80. MUELLER, M. et al. Predicting extubation outcome in preterm newborns: a comparison of neural networks with clinical expertise and statistical modeling. In: *Pediatr Res*. 2004, 56(1):11-18, ISSN 1530-0447.
81. DUNNING, J., AU, J., KALKAT, M., LEVINE, A. A validated rule for predicting patients who require prolonged ventilation post cardiac surgery. In: *Eur J Cardiothorac Surg*. 2003, 24(2):270-276, ISSN 1010-7940.
82. MANJI, R. et al. Long-term outcome and predictors of noninstitutionalized survival subsequent to prolonged intensive care unit stay after cardiac surgical procedures. In: *Ann Thorac Surg*. 2016, 101(1):56-63, ISSN 0003-4975.
83. BRANDI, S., TROSTER, E., CUNHA, M. Length of stay in pediatric intensive care unit: prediction model. In: *Einstein (Sao Paulo)*, 2020;18:eAO5476, ISSN 1679-4508.
84. MCWILLIAMS, C. et al. Towards a decision support tool for intensive care discharge: machine learning algorithm development using electronic healthcare data from MIMIC-III and Bristol, UK. In: *BMJ Open*. 2019, 07;9(3):e025925, ISSN 2398-8703.

85. LIN, Y. et al. Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. In: *PLoS One*. 2019, 14(7):e0218942, ISSN 1932-6203.
86. CZEITER, E. et al. CENTER-TBI Participants and Investigators. Blood biomarkers on admission in acute traumatic brain injury: relations to severity, CT findings and care path in the CENTER-TBI study. In: *EBioMedicine*. 2020, 56:102785, ISSN 2352-3964.
87. YIN, W. et al. The utilization of critical care ultrasound to assess hemodynamics and lung pathology on ICU admission and the potential for predicting outcome. In: *PLoS One*. 2017, 12(8):e0182881, ISSN 1932-6203.
88. SHICKEL, B. et al. DeepSOFA: a continuous acuity score for critically ill patients using clinically interpretable deep learning. In: *Sci Rep*. 2019, 12;9(1):1879, ISSN 2045-2322.
89. MCRAE, M. et al. Clinical decision support tool and rapid point-of-care platform for determining disease severity in patients with COVID-19. In: *Lab Chip*. 2020, 21;20(12):2075-2085, ISSN 1473-0189.
90. SRINIVASAN, S., DOSHI-VELEZ, F. Interpretable batch IRL to extract clinician goals in ICU hypotension management. In: *AMIA Jt Summits Transl Sci Proc*. 2020, 2020:636-645. ISSN 2153-4063.
91. NEMATI, S., GHASSEMI, M., CLIFFORD, G. Optimal medication dosing from suboptimal clinical examples: a deep reinforcement learning approach. In: *Annu Int Conf IEEE Eng Med Biol Soc*. 2016, 2016:2978-2981, ISSN 0589-1019.
92. KOMOROWSKI, M. et al. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. In: *Nat Med*. 2018, 24(11):1716-1720, ISSN 1340-3443.
93. LOPEZ-MARTINEZ, D. et al. Deep reinforcement learning for optimal critical care pain management with morphine using dueling double-deep Q networks. In: *Annu Int Conf IEEE Eng Med Biol Soc*. 2019, 2019:3960-3963, ISSN 2154-2287.
94. YU, C., LIU, J., ZHAO, H. Inverse reinforcement learning for intelligent mechanical ventilation and sedative dosing in intensive care units. In: *BMC Med Inform Decis Mak*. 2019, 09;19(Suppl 2):57, ISSN 1472-6947.
95. YU, C., REN, G., DONG, Y. Supervised-actor-critic reinforcement learning for intelligent mechanical ventilation and sedative dosing in intensive care units. In: *BMC Med Inform Decis Mak*. 2020, 09;20(Suppl 3):124, ISSN 1472-6947.
96. BENJAMENS, S., DHUNNOO, P., MESKÓ, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. In: *NPJ Digital Medicine* (2020) 3:118. ISSN 2398-6352.
97. ISLAM, M. et al. Prediction of sepsis patients using machine learning approach: A meta-analysis. In: *Computer Methods Programs Biomed*, 2019, 170; 1-9, ISSN 1872-7565.
98. OEYEN, S. et al. Development of a prediction model for long-term quality of life in critically ill patients. In: *J Crit Care*. 2018, 43:133-138, ISSN 1557-8615.
99. SINGER, M. et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). In: *JAMA*, 2016, 315:8, pp. 801–810. ISSN 0098-7484.
100. FUNK, D., PARRILLO, J., KUMAR, A. Sepsis and septic shock: a history. In: *Critical care clinics*. 2009, 25:1, pp. 83–101. ISSN 0749-0704.
101. GEROULANOS, S., DOUKA, T. Historical perspective of the word “sepsis”. In: *Intensive Care Medicine*. 2006, 32:12, pp. 2077–2087. ISSN 0342-4642.
102. BONE, R. et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. In: *Chest*. 1992, 101:6, pp. 1644–1655. ISSN 1931-3543.

103. LEVY, M. et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. In: *Intensive Care Medicine*. 2003, 29:4, pp. 530–538. ISSN 0342-4642.
104. GÜL, F., ARSLANTAŞ, M., CINEL, I., KUMAR, A. Changing definitions of sepsis. In: *Turkish journal of anaesthesiology and reanimation*. 2017, 45:3, p. 129. ISSN 1304-0871.
105. VINCENT, J. et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. In: *Intensive Care Medicine*. 1996, 22:7, pp. 707–710. ISSN 0342-4642.
106. SKINNER, Q. Machiavelli: the prince. Cambridge University Press, 1988. 152 p. ISBN 9780521349932
107. FERRER, R. et al. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. In: *Critical Care Medicine*. 2014, 42:8, pp. 1749–1755. ISSN 0090-3493.
108. DELLINGER, R. et al. Surviving Sepsis Campaign: International Guidelines for Management of Severe Sepsis and Septic Shock 2012. In: *Critical Care Medicine*. 2013, 41:2, pp. 580–637. ISSN 0090-3493.
109. HOTCHKISS, R. et al. Sepsis and septic shock. In: *Nature Reviews Disease Primers* 2, 2016. ISSN 2056-676X
110. KAUKONEN, K. et al. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000-2012. In: *JAMA*. 2014, 311:13, pp. 1308–1316. ISSN 0098-7484.
111. RHODES, A. et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. In: *Intensive Care Medicine*. 2017, 43:3, pp. 304–377. ISSN 0342-4642.
112. MARIK, P., FARKAS, D. The changing paradigm of sepsis: early diagnosis, early antibiotics, early pressors, and early adjuvant treatment. In: *Critical care medicine*. 2018, 46:10, pp. 1690–1692. ISSN 0090-3493.
113. BOZZA, F. et al. Cytokine profiles as markers of disease severity in sepsis: a multiplex analysis. In: *Critical Care*. 2007, 11:2, R49. ISSN 1364-8535.
114. TSALIK, E. et al. Discriminative value of inflammatory biomarkers for suspected sepsis. In: *The Journal of Emergency Medicine*. 2012, 43:1, pp. 97–106. ISSN 1090-1280.
115. GULTEPE, E. et al. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. In: *Journal of the American Medical Informatics Association*. 2014, 21:2, pp. 315–325. ISSN 1067-5027
116. RIBAS, V. et al. Severe sepsis mortality prediction with relevance vector machines. In: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011. IEEE*, 2011, pp. 100–103. ISSN 1557-170X.
117. GOLDBERGER, A. et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. In: *Circulation*. 2000, 101:23, e215–e220. ISSN 0009-7322.
118. JOHNSON, A. et al. MIMIC-III, a freely accessible critical care database. In: *Scientific Data* 3, 2016. ISSN 2052-4463.
119. HENRY, K., HAGER, D., PRONOVOST, P., SARIA, S. A targeted real-time early warning score (TREW Score) for septic shock. In: *Science Translational Medicine*. 2015, 7:299. ISSN 1946-6234

120. CALVERT, J. et al. A computational approach to early sepsis detection. In: *Computers in Biology and Medicine*. 2016, 74, pp. 69–73. ISSN 1879-0534.
121. FUTOMA, J., HARIHARAN, S., HELLER, K. Learning to detect sepsis with a multitask Gaussian process RNN classifier. In: *International Conference on Machine Learning*. PMLR. 2017, pp. 1174–1182. ISBN 9781510855144
122. SHI, J. et al. DI++: A deep learning system for patient condition identification in clinical notes. In: *Artif Intell Med*, 2022 Jan;123:102224. ISSN 1873-2860.
123. EVANS, L. et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock 2021. In: *Intensive Care Medicine*, ISSN 0342-4642.
124. OCAMPO-QUINTERO, N. Enhancing sepsis management through machine learning techniques: A review. In: *Medicina Intensiva*, nr 46, 2022, pp. 140-156. ISSN 1578-6749.
125. JAMES, G., WITTEN, D., HASTIE, T., TIBSHIRANI, R. *An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics)*, 2nd ed, Springer, 2021, 440 p. ISBN 978-1461471370
126. HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J. *The Elements of statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed, Springer, 2017, 767 p. ISBN 978-0387848570.
127. Course 9.520 on Statistical Learning Theory and Applications, Instr: Poggio, T., Rosasco, L., Frogner, C., Canas, G.D, [online]. Stanford University, 2012 [citat: 30.12.2022]. Disponibil: <https://www.mit.edu/~9.520/spring12/index.html>
128. VAPNIK, V. Principles of Risk Minimization for Learning Theory. In: *Advances in neural information processing*, 1991. ISBN 1-55860-222-4.
129. BUCHARD, A., RICHENS, J.G. Artificial Intelligence for Medical Decisions. In: *Springer Nature Switzerland AG 2021 N. Lidströmer, H. Ashrafian (eds.), Artificial Intelligence in Medicine*. 2021, p. 1914. ISBN 978-3030645724.
130. HONG-FEI, DENG et al. Evaluating machine learning models for sepsis prediction: A systematic review of methodologies. In: *iScience* 25, 103651, January 21, 2022, pp. 1-18. ISSN 0258-9004.
131. MOOR, M. et al. Early Prediction of Sepsis in the ICU Using Machine Learning: A Systematic Review. In: *Review. Front. Med.* 2021, 8:607952, pp.1-18. ISSN 2296-858X.
132. WU, M., DU, X., GU, R., WEI, J. (2021) Artificial Intelligence for Clinical Decision Support in Sepsis. In: *Front. Med.* 8:665464. pp. 1-9. ISSN 2296-858X.
133. KIM, J.K. et al. Early Prediction of Sepsis Onset Using Neural Architecture Search Based on Genetic Algorithms. In: *Int. J. Environ. Res. Public Health* 2022, 19, 2349. ISSN 1661-7827.
134. HASSAN, N. et al. Preventing sepsis; how can artificial intelligence inform the clinical decision-making process? A systematic review. In: *Int J Med Inform.* 2021 Jun; 150:104457. ISSN 1872-8243.
135. CAMACHO-COGOLLO, J.E. et al. Machine Learning Models for Early Prediction of Sepsis on Large Healthcare Datasets. In: *Electronics* 2022, nr. 11, 1507, pp. 2-23. ISSN 2079-9292.
136. HONORÉ, A. et al. Classification and feature extraction for neonatal sepsis detection, TechRxiv. Preprint. 2022
137. PEI-CHEN, LIN et al. Machine Learning Model to Identify Sepsis Patients in the Emergency Department: Algorithm Development and Validation. In: *J. Pers. Med.* 2021, 11, 1055. pp. 1-12. ISSN 2075-4426.

138. ZEKAI, WANG, BING, YAO. Multi-Branching Temporal Convolutional Network for Sepsis Prediction. In: *IEEE Journal of Biomedical and Health Informatics*, 2021, pp. 1-12. ISSN 2168-2194.
139. PERSSON, I. et al. 2021, A Machine Learning Sepsis Prediction Algorithm for Intended Intensive Care Unit Use (NAVOY Sepsis): Proof-of-Concept Study. In: *JMIR Formative Research*, February 2021, nr. 5(9), pp. 1-11. ISSN 2561-326X.
140. COLLINS, G. et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. In: *BMJ Open* 2021; 11:e048008, pp. 1-7. ISSN 2398-8703.
141. COLLINS, G. et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD Statement. In: *BMC Medicine*. 2015, 13:1, ISSN 1741-7015.
142. MOONS, K.G.M. et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. In: *Heart* 2012;98:683–90. ISSN 1355-6037.
143. COLLINS, G.S. et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. In: *BMC Med Res Methodol* 2014;14:40. ISSN 1471-2288.
144. BOUWMEESTER, W. et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012;9:e1001221. ISSN 1549-1277.
145. BREIMAN, L. Statistical modeling: the two cultures. In: *Stat Sci*, 2001;16:199–231. ISSN 0883-4237.
146. SHILLAN, D. et al. Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. In: *Crit Care* 2019;23:284. ISSN 1364-8535.
147. CHRISTODOULOU, E. et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. In: *J Clin Epidemiol* 2019;110:12–22. ISSN 0895-4356.
148. MOONS, K.G.M. et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. In: *Ann Intern Med* 2019;170:W1–33. ISSN 0003-4819.
149. FAES, L. et al. A clinician's guide to artificial intelligence: how to critically appraise machine learning studies. In: *Transl Vis Sci Technol* 2020;9:7. ISSN 2164-2591.
150. BOSSUYT, P.M. et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527, pp. 1-9. ISSN 1756-1833.
151. SOUNDERAJAH, V. et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol On behalf of the STARD-AI Steering Committee. In: *BMJ Open*, (11), 2021, 47709 pp. 1-7. ISSN 2398-8703.
152. SHIMABUKURO, D. et al. Effect of a machine learning- based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. In: *BMJ Open Respir Res*. 2017, 4:e000234. ISSN 2052-4439.
153. YEALY, D. et al. A randomized trial of protocol- based care for early septic shock. In: *N Engl J Med*. 2014, 370:1683–93. ISSN 0028-4793.
154. BURDICK, H. et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-

- world patient data from US hospitals. In: *BMJ Health Care Inform.* 2020, 27:e100109. ISSN 2632-1009
155. PHUA, J. et al. Characteristics and outcomes of culture- negative versus culture-positive severe sepsis. In: *Crit Care.* 2013, 17: R202. ISSN 1364-8535.
 156. LAMANTIA, M. et al. Predictive value of initial triage vital signs for critically ill older adults. In: *West J Emerg Med.* 2013, 14:453–60. ISSN 1936-9018.
 157. HORNG, S. et al. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. In: *PLoS One.* 2017, 12:e0174708. ISSN 1932-6203.
 158. DELAHANTY, R. et al. Development and evaluation of a machine learning model for the early identification of patients at risk for sepsis. In: *Ann Emerg Med.* 2019, 73:334–44. ISSN 0196-0644
 159. RAGHU, A. et al. Deep Reinforcement Learning for Sepsis Treatment. arXiv, 2017, <https://doi.org/10.48550/arXiv.1711.09602>
 160. WOLFF, R. et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model. In: *Ann Intern Med.* 2019, 170:51-58. ISSN 0003-4819.
 161. RIVERA, S. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI Extension. In: *BMJ*, 2020, 370, ISSN 1756-1833.
 162. YU, K., BEAM, A., KOHANE, I. Artificial intelligence in healthcare. In: *Nat Biomed Eng*, 2018;2:719–31. ISSN 2157-846X.
 163. BEAM, A., KOHANE, I. Big data and machine learning in health care. In: *JAMA*, 2018;319:1317. ISSN 0098-7484.
 164. GHASSEMI, M. et al. Practical guidance on artificial intelligence for health-care data. In: *Lancet Digit Health*, 2019;1:e157–9. ISSN 2589-7500.
 165. BEAM, A., MANRAI, A., GHASSEMI, M. Challenges to the reproducibility of machine learning models in health care. In: *JAMA*, 2020;323:305. ISSN 0098-7484.
 166. SENDAK, M., D'ARCY, J., KASHYAP, S. A path for translation of machine learning products into healthcare delivery. In: *EMJ Innov*, 2020. ISSN 2513-8634.
 167. WIENS, J. et al. Do no harm: a roadmap for responsible machine learning for health care. In: *Nat Med*, 2019. ISSN 1546-170X.
 168. BREIMAN, L. Statistical modeling: the two cultures. In: *Stat Sci*, 2001;16:199–231. ISSN 0883-4237
 169. STEVENS, L. et al. Recommendations for Reporting Machine Learning Analyses in Clinical Research. In: *Circ Cardiovasc Qual Outcomes*, 2020;13:e006556. ISSN 1941-7705.
 170. REYNA, M. et al. *Early Prediction of Sepsis from Clinical Data: The PhysioNet Computing in cardiology Challenge 2019* (version 1.0.0), [online]. PhysioNet, 2019 [citat: 24.12. 2022]. Disponibil: <https://doi.org/10.13026/v64v-d857>
 171. HUANG, C. et al. Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: a retrospective cohort study. In: *PLoS Med.* 2018;15:e1002703, ISSN 1549-1277.
 172. SINGER, M. et al. The third International consensus definitions for sepsis and septic shock (sepsis-3). In: *JAMA.* 2016, 315:801–10. ISSN 0098-7484.
 173. MORTAZAVI, B. et al. Analysis of machine learning techniques for heart failure readmissions. In: *Circ Cardiovasc Qual Outcomes.* 2016;9:629–640. ISSN 1941-7705.
 174. CONSORT: Transparent Reporting of Trials, <http://www.consort-statement.org/>

175. FREUND, Y., SCHAPIRE, R. A decision-theoretic generalization of on-line learning and an application to boosting. In: *Lecture Notes in Computer Science, Springer*, 1995, (904), p.23-37. ISSN 0302-9743.
176. MAYR, A. et al. The Evolution of Boosting Algorithms – From Machine Learning to Statistical Modelling. In: *Methods Inf Med* 2014; 53(6): 419–427. ISSN 0026-1270.
177. PINEAU, J. The Machine Learning Reproducibility Checklist (v2.0, Apr.7 2020), 2020. [citat: 04.05.2022], Disponibil: <https://www.cs.mcgill.ca/~jpineau/ReproducibilityChecklist.pdf>
178. RILEY, R. et al. Calculating the sample size required for developing a clinical prediction model. In: *BMJ*. 2020;368:m441, ISSN 1756-1833.
179. VAN SMEDEN, M. et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. In: *Stat Methods Med Res*. 2019;28. ISSN 0962-2802.
180. RILEY, R. et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. In: *Stat Med*. 2019;38:1276. ISSN 0277-6715.
181. DE JONG, V. et al. Sample size considerations and predictive performance of multinomial logistic prediction models. In: *Stat Med*. 2019;38:1601. ISSN 0277-6715.
182. FAUL, F., ERDFELDER, E., LANG, A. G., BUCHNER, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. In: *Behavior Research Methods*. 2007, vol. 39, nr. 2, pp. 175–191. ISSN 1554-3528.
183. SOUNDERAJAH, V. et al. Are disruptive innovations recognised in the healthcare literature? A systematic review. In: *BMJ Innov* 2021;7:208–16. ISSN 2055-642X.
184. TOPOL, E.J. High-Performance medicine: the convergence of human and artificial intelligence. In: *Nat Med* 2019;25:44–56. ISSN 1546-170X.
185. WILLIAMS, B.J., BOTTOMS, D., TREANOR, D. Future-proofing pathology: the case for clinical adoption of digital pathology. In: *J Clin Pathol* 2017;70:1010–8. ISSN 0021-9746.
186. CRUZ RIVERA, S. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension, In: *Nature Medicine* | VOL 26 | September 2020 | pp. 1351–1363. ISSN 1078-8956
187. CHAN, A.-W. et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann. Intern. Med.* 158, 200–207 (2013). ISSN 0003-4819
188. LIU, X. et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit. Health* 1, e271–e297 (2019). ISSN 2589-7500
189. LIU, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension, In: *BMJ* 2020;370:m3164, pp. 1-13. ISSN 1756-1833.
190. LIU, X., FAES, L., CALVERT, M.J., DENNISTON, A.K. CONSORT/SPIRIT-AI Extension Group. Extension of the CONSORT and SPIRIT statements. *Lancet* 2019;394:1225. ISSN 0140-6736.
191. STERNE, J. et al. “Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls,” *BMJ*, 2009, 338:b2393. ISSN 1756-1833.
192. JAIN, R., XU, W. Dynamic model updating (DMU) approach for statistical learning model building with missing data, *BMC Bioinformatics* (2021) 22:221. ISSN 1471-2105.

193. *Guideline on Missing Data in Confirmatory Clinical Trials*, 2010 EMA/CPMP/EWP/1776/99 Rev. 1 Committee for Medicinal Products for Human Use (CHMP), [online]. [citat 30.12.2022]. Disponibil: <https://www.ema.europa.eu/en>
194. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
195. ZEILEIS, A., GROTHENDIECK, G. (2005). “zoo: S3 Infrastructure for Regular and Irregular Time Series.” In: *Journal of Statistical Software*, **14**(6), 1–27. ISSN 1548-7660.
196. AZZALINI, A. (2022). The R package{sn}: The skew-normal and related distributions such as the skew-t and the SUN (version 2.0.2). URL: <http://azzalini.stat.unipd.it/SN/>, <https://cran.r-project.org/package=sn>
197. IAPĂSCURTĂ, V. Gestionarea valorilor lipsă în date biomedicale cu caracter continuu. În: Conferința științifică anuală a USMF „N. Testemițanu”. Cercetarea în biomedicină și sănătate: calitate, excelență și performanță, 19-21 octombrie, 2022, Chișinău. MJHS, 29(3), 2022, pp. 250-250. ISSN 2345-1467.
198. IAPĂSCURTĂ, V. Detection of Movement toward Randomness by Applying the Block Decomposition Method to a Simple Model of the Circulatory System. In: *Complex Systems Journal*. 2019, vol. 28(3), pp. 59-77. ISSN 0891-2513.
199. ZENIL, H. et al. A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity. In: *Entropy*. 2018, 20(8), p. 605. ISSN 1099-4300.
200. SOLER-TOSCANO, F., ZENIL, H., DELAHAYE J.-P., GAUVRIT, N. Calculating Kolmogorov Complexity from the Output Frequency Distributions of Small Turing Machines. In: *PLoS ONE*. 2014, 9(5) e96223. ISSN 1932-6203.
201. ZENIL, H., SOLER-TOSCANO, F., DINGLE K., LOUIS A. Correlation of Automorphism Group Size and Topological Properties with ProgramSize Complexity Evaluations of Graphs and Complex Networks. In: *Physica A: Statistical Mechanics and Its Applications*. 404, 2014 pp. 341–358. ISSN 0378-4371.
202. ZENIL, H. et al. The Online Algorithmic Complexity Calculator (OACC) v3.0, *Algorithmic Dynamics Lab, Science for Life Laboratory (SciLifeLab), Unit of Computational Medicine, Center for Molecular Medicine at the Karolinska Institute in Stockholm, Sweden* [software]. 2022 [citat: 24.12.2022]. Disponibil: www.algorithmicdynamics.net/software.html.
203. IAPĂSCURTĂ, V. A less traditional approach to biomedical signal processing for sepsis prediction, In: 5th International Conference on Nanotechnologies and Biomedical Engineering, November 3-5, 2021, Springer IFMBE Proceedings Series, 2022, pp. 215-222. ISBN 978-3-030-92327-3.
204. GILL, N. Generalizations of linear modelling in the biomedical sciences. University of Groningen, *PhD thesis*, 2016, 112 pages
205. BZDOK, D., IOANNIDIS, J. Exploration, inference and prediction in neuroscience and biomedicine, In: *Trends in Neurosciences*, 42(4), 2019, pp. 251-262, ISSN 0378-5912.
206. ROBNIK-SIKONJA, M. (2021). semiArtificial: Generator of Semi-Artificial Data. R package version 2.4.1, URL: <https://CRAN.R-project.org/package=semiArtificial>.
207. IAPĂSCURTĂ, V., Belii, A. Preclinical stage of building a machine learning system for sepsis prediction: a comparative study of four algorithms, In: 5th International 143 Conference on Nanotechnologies and Biomedical Engineering, November 3-5, 2021, Springer IFMBE Proceedings Series. 2022, pp. 448-455. ISBN 978-3-030-92327-3.

208. PANT, A. Workflow of a Machine Learning project, [online], [citat 10.08.2022], disponibil: <https://towardsdatascience.com/workflow-of-a-machine-learning-project-ec1dba419b94>
209. H2O.ai. (2020) h2o: R Interface for H2O. R package version 3.32.1.1.
210. MedCalc Software Ltd. Diagnostic test evaluation calculator. https://www.medcalc.org/calc/diagnostic_test.php, Versiunea 20.113 [accesat 14.08.22]
211. ROBIN, X. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. In: *BMC Bioinformatics*, 12, p. 77. ISSN 1471-2105
212. H2O.ai. (2020) h2o: H2O Wave. Python package version 0.0.14.
213. SILBER, A. Essays on Critical Care Medicine, vol. 1, Petrozavodsk University Press, Petrozavodsk, 1995, 359 p.
214. SADASIVUN, S. et al., Fusion of fully integrated analog machine learning classifier with electronic medical records for real-time prediction of sepsis onset. In: *Nature Scientific Reports*, 2022, 12:5711, ISSN 2045-2322.
215. MISRA, D. et al., Early Detection of Septic Shock Onset Using Interpretable Machine Learners, In: *J Clin Med*, 2021, 10, 301. ISSN 2077-0383.
216. WICKRAMARATNE, D., MAHMUD, S. Bi-Directional Gated Recurrent Unit Based Ensemble Model for the Early Detection of Sepsis, In: *Annu Int Conf IEEE Eng Med Biol Soc*, 2020, pp.70-73. ISSN 0589-1019.
217. KOK, C. et al, Automated prediction of sepsis using temporal convolutional network, In: *Comput Biol Med*, 2020, 127:103957. ISSN 1879-0534.
218. HELGUERA-REPETTO, A. et al., Neonatal Sepsis Diagnosis Decision-Making Based on Artificial Neural Networks, In: *Frontiers in Pediatrics*, 8:525, 2020, ISSN 2296-2360.
219. KARAKIKE, E. et al. Coronavirus disease 2019 as cause of viral sepsis: a systematic review and meta-analysis. *Crit Care Med*. 2021;49(12):2042–2057. ISSN 0090-3493.
220. BATRA, R. et al. Multiomic comparative analysis of COVID-19 and bacterial sepsis-induced ARDS. *PLoS Pathog*. 2022;18(9):e1010819. ISSN 1553-7366.
221. OSUCHOWSKI, M.F. et al. The COVID-19 puzzle: deciphering pathophysiology and phenotypes of a new disease entity. *Lancet Respir Med*. 2021;9(6):622–642. ISSN 2213-2600.
222. STOLARSKI, A.E., KIM, J., ZHANG, Q., REMICK, D.G. Cytokine drizzle-the rationale for abandoning “Cytokine Storm”. *Shock*. 2021;56(5):667–672. ISSN 1073-2322
223. WINKLER, M.S. et al. Renaissance of glucocorticoids in critical care in the era of COVID-19: ten urging questions. *Crit Care*. 2022;26(1):308. ISSN 1364-8535.
224. HUANG, L. et al. Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study. *Lancet Respir Med*. 2022;10(9):863–876. ISSN 2213-2600.
225. PUNTMANN, V.O. et al. Long-term cardiac pathology in individuals with mild initial COVID-19 illness. *Nat Med*. 2022;28(10):2117–2123. ISSN 1078-8956.
226. LAURITSEN, S. Explainable artificial intelligence model to predict acute critical illness from electronic health records, In: *Nature Communications*, 11:3852, 2020, ISSN 2041-1723.
227. PHILLIPS, P. et al., Four Principles of Explainable Artificial Intelligence, U.S. Department of Commerce, National Institute of Standards and Technology, 2021, 43 p., [online] National Institute of Standards and Technology Interagency or Internal Report

- 8312 Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8312, 43 pages (September 2021) [citat 30.12.2022]. Disponibil: <https://doi.org/10.6028/NIST.IR.8312>
228. HOLZINGER, A. Causability and explainability of artificial intelligence in medicine. In: *WIREs Data Mining Knowl Discov.* 2019;9:e1312. pp. 1-12. ISSN 1942-4795.
 229. HOLZINGER, A., DEHMER, M., JURISICA, I. (2014). Knowledge discovery and interactive data mining in bioinformatics—State-of-the-art, future challenges and research directions. *BMC Bioinformatics*, 15, 11. ISSN 1471-2105.
 230. LEE, S., HOLZINGER, A. (2016). Knowledge discovery from complex high dimensional data. In S. Michaelis, N. Piatkowski, and M. Stolpe (Eds.), *Solving large scale learning tasks. Challenges and algorithms, Lecture notes in artificial intelligence, LNAI 9580* (pp. 148–167). Heidelberg, Germany: Springer. ISSN 0302-9743
 231. HOLZINGER, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3, 119–131. ISSN 2198-4018.
 232. HOLZINGER, A. (2014). Trends in interactive knowledge discovery for personalized medicine: Cognitive science meets machine learning. *IEEE Intelligent Informatics Bulletin*, 15, 6–14. ISSN 2168-8222.
 233. TURKAY, C., JEANQUARTIER, F., HOLZINGER, A., HAUSER, H. (2014). On computationally-enhanced visual analysis of heterogeneous data and its application in biomedical informatics. In A. Holzinger & I. Jurisica (Eds.), *Interactive knowledge discovery and data mining: State-of-the-art and future challenges in biomedical informatics. Lecture notes in computer science LNCS 8401* (pp. 117–140). Heidelberg, Germany: Springer. ISSN 0302-9743.
 234. HOLZINGER, A., BIEMANN, C., PATTICHIS, C.S., KELL, D.B. (2017). What do we need to build explainable AI systems for the medical domain? arXiv:1712.09923.
 235. SHORTLIFFE, E.H., BUCHANAN, B.G. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23, 351–379. ISSN 0025-5564.
 236. JOHNSON, W.L. (1994). Agents that learn to explain themselves. Paper presented at 12th national conference on artificial intelligence (AAAI '94) (pp. 1257–1263). AAAI. ISBN 9780262510783.
 237. LACAVE, C., DIEZ, F.J. (2002). A review of explanation methods for Bayesian networks. *The Knowledge Engineering Review*, 17, 107–127. ISSN 1469-8005.
 238. POPPER, K. (1935). *Die Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft.* Wien, Austria: Springer-Verlag. 248 p. ISBN 978-3-7091-2021-7.
 239. HUDEC, M., BEDNROV, E., HOLZINGER, A. (2018). Augmenting statistical data dissemination by short quantified sentences of natural language. *Journal of Official Statistics (JOS)*, 34, 981. ISSN 0282-423X.
 240. GERSHMAN, S.J., HORVITZ, E.J., TENENBAUM, J.B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349, 273–278. ISSN 0036-8075.
 241. PEARL, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge, England: Cambridge University Press. 2009. 484 p. ISBN 978-0521895606.
 242. BOLOGNA, G., HAYASHI, Y. (2017). Characterization of symbolic rules embedded in deep dimlp networks: A challenge to transparency of deep learning. *Journal of Artificial Intelligence and Soft Computing Research*, 7, 265–286. ISSN 2449-6499.
 243. HOLZINGER, A. (2018). From machine learning to explainable AI. Paper presented at *2018 World symposium on digital intelligence for systems and machines (DISA)*. IEEE, 2018, ISBN 9781538651032.

244. KIESEBERG, P., WEIPPL, E., HOLZINGER, A. (2016). Trust for the doctor-in-the-loop. In *European Research Consortium for Informatics and Mathematics (ERCIM) news: Tackling big data in the life sciences* Vol. 104, pp. 32–33. ISSN 0926-4981.
245. MILLER, T., HOWE, P., SONENBERG, L. (2017) Explainable AI: Beware of inmates running the asylum or: How i learnt to stop worrying and love the social and behavioural sciences. arXiv:1712.00547.
246. HAMBURG, M.A., COLLINS, F.S. (2010). The path to personalized medicine. *New England Journal of Medicine*, 363, 301–304. ISSN 0028-4793.
247. MILLER, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. ISSN: 0004-3702
248. FELLMETH, A.X., HORWITZ, M. (2009). Guide to Latin in international law. Oxford, England: Oxford University Press. 298 p. ISBN 978-0195369380.
249. RIBEIRO, M.T., SINGH, S., GUESTRIN, C. (2016a). Model-agnostic interpretability of machine learning. arXiv:1606.05386.
250. HOLZINGER, A. et al. (2018). Interactive machine learning: Experimental evidence for the human in the algorithmic loop. *Applied Intelligence*, 1–14. ISSN 1573-7497.
251. GUILLAUME, S. (2001). Designing fuzzy inference systems from data: An interpretability-oriented review. *IEEE Transactions on Fuzzy Systems*, 9, 426–443. ISSN 1063-6706.
252. CARUANA, R. et al. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. Paper presented at *21th ACM SIGKDD international conference on knowledge discovery and data mining (KDD '15)* (pp. 1721–1730). ACM. ISBN 978-1450340229.
253. POULIN, B. et al. (2006). Visual explanation of evidence with additive classifiers. In *National conference on artificial intelligence* (pp. 1822–1829). Cambridge, MA: MIT Press. ISBN 978-1-57735-281-5
254. SZAFRON, D. et al. (2004). Proteome analyst: Custom predictions with explanations in a web-based tool for high-throughput proteome annotations. *Nucleic Acids Research*, 32, W365–W371. ISSN 0305-1048.
255. PEARL, J., MACKENZIE, D. (2018). The book of why. New York, NY: Basic Books. 432 p. ISBN 978-0465097609
256. HOLZINGER, A. (2005). Usability engineering methods for software developers. *Communications of the ACM*, 48, 71–74. ISSN 0001-0782
257. BEVAN, N. (1995). Measuring usability as quality of use. *Software Quality Journal*, 4, 115–130. ISSN 1573-1367.
258. FRIEDMAN, J. “Greedy function approximation: A gradient boosting machine.” In: *Annals of statistics*, 29(5), 2001, pp. 1189-1232, ISSN 0090-5364.
259. MOLNAR, C. ”Interpretable machine learning. A Guide for Making Black Box Models Explainable”, 2019. [online] <https://christophm.github.io/interpretable-ml-book/>.
260. GOLDSTEIN, A. et al. “Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation.” In: *Journal of Computational and Graphical Statistics*, 24(1), 2015, pp 44-65. ISSN 1061-8600.
261. LUNDBERG, S., LEE, S. Unified Approach to Interpreting Model Predictions, In: *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017. ISBN 9781510860964

262. ZECH, R. et al. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study, In: *PLoS Med* 15(11):e1002683, 2018. ISSN 1549-1277.
263. CHEN, Q. et al. Transferability and interpretability of the sepsis prediction models in the intensive care unit. In: *BMC Medical Informatics and Decision Making* (2022) 22:343, pp. 1-10. ISSN 1472-6947.
264. Moor, M. et al. Early recognition of sepsis with Gaussian process temporal convolutional networks and dynamic time warping. *Mach Learn Healthc Conf PMLR*. 2019;106:2–26. ISSN 2640-3498.
265. BARTON, C. et al. Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. *Comput Biol Med*. 2019;109:79–84. ISSN 1879-0534.
266. ASUROGLU, T., OGUL, H. A deep learning approach for sepsis monitoring via severity score estimation. *Comput Methods Programs Biomed*. 2021;198:105816. ISSN 1872-7565.
267. ROSNATI, M., FORTUIN, V. MGP-AttTCN: an interpretable machine learning model for the prediction of sepsis. *PLoS ONE*. 2021;16:e0251248. ISSN 1932-6203.
268. MCCOY, A., DAS, R. Reducing patient mortality, length of stay and readmissions through machine learning-based sepsis prediction in the emergency department, intensive care unit and hospital floor units. *BMJ Open Qual*. 2017;6:e000158. ISSN 2399-6641.
269. CHO, K-J. et al. Detecting patient deterioration using artificial intelligence in a rapid response system. *Crit Care Med*. 2020;48(4):e285. ISSN 0090-3493.
270. SIONTIS, G.C., TZOULAKI, I., CASTALDI, P.J., IOANNIDIS, J.P. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol*. 2015;68(1):25–34. ISSN 0895-4356.
271. LEE, J., MASLOVE, D.M. Customization of a severity of illness score using local electronic medical record data. *J Intensive Care Med*. 2017;32(1):38–47. ISSN 1525-1489.
272. WARDI, G. et al. Predicting progression to septic shock in the emergency department using an externally generalizable machine-learning algorithm. *Ann Emerg Med*. 2021;77(4):395–406. ISSN 0196-0644.
273. ZHANG, D. et al. An interpretable deep-learning model for early prediction of sepsis in the emergency department. *Patterns*. 2021;2:100196. ISSN 2666-3899.
274. GANDIN, I., SCAGNETTO, A., ROMANI, S., BARBATI, G. Interpretability of time-series deep learning models: a study in cardiovascular patients admitted to Intensive care unit. *J Biomed Inform*. 2021;121:103876. ISSN 1532-0464.
275. MORID, M.A., BORJALI, A., DEL FIOLE, G. A scoping review of transfer learning research on medical image analysis using ImageNet. *Comput Biol Med*. 2021;128:104115. ISSN 1879-0534.
276. TSENG, K-K. et al. Cross-domain transfer learning for ppg diagnosis algorithm. *Biosensors*. 2021;11:127. ISSN 2079-6374.
277. LIU, K. et al. Development and validation of a personalized model with transfer learning for acute kidney injury risk estimation using electronic health records. *JAMA Netw Open*. 2022;5:e2219776. ISSN 2574-3805.
278. CHEN, H. et al. Forecasting adverse surgical events using self-supervised transfer learning for physiological signals. *NPJ Digit Med*. 2021;4:167. ISSN 2398-6352.

ANEXE

Anexa 1.	Definirea unei serii temporale	144
Anexa 2.	Aspectul inițial al datelor din studiu	145
Anexa 3.	Aspecte de corelare dintre parametrii de interes în setul B	146
Anexa 4.	Aspecte de corelare dintre parametrii de interes în seturile “Pre-sepsis” și “Sepsis confirmat” (set A)	148
Anexa 5.	Similaritatea seturilor de date “Sepsis” și “Non-sepsis”	150
Anexa 6.	Similaritatea seturilor de date “Pre-sepsis” și “Non-sepsis”	151
Anexa 7.	Informație suplimentară în privința trainingului și cross-validării modelelor de bază	152
Anexa 8.	Performanța sistemului la prezicerea continuă a riscului apariției sepsisului	154
Anexa 9.	Raportarea rezultatelor studiilor cu utilizarea învățării automate	155
Anexa 10.	Performanța a șapte modele antrenate pe date din setul A în cadrul învățării automate automatizate (AutoML) în baza platformei H2O (etapa de cross-validare)	159
Anexa 11.	Diagrame PD și ICE pentru unele caracteristici utilizate de modele GBM cu performanță înaltă în cadrul cercetării curente și diagrame de corelare între modele	160
Anexa 12.	CertIFICATE de inovator și acte de implementare	164

Anexa 1. Definierea unei serii temporale

Definiție (Seria temporală). Fie $t \in R$ un parametru de timp, și fie $f(t): R \rightarrow R^d$, notăm un proces de generare a datelor cu $d \in N$ și $d \geq 1$. Pentru o mulțime de puncte temporale discrete $\{t_1, \dots, t_l\}$, considerăm $x_i = f(t_i)$ ca un vector de valori de măsurare observate din f la momentul t_i . Colectăm intrările din x_i pentru $i \in \{1, \dots, l\}$, astfel încât x_{ij} să indice valoarea de măsurare a dimensiunii j la momentul t_i . Colectăm atât valorile observate, cât și momentele de timp în vectorii $x = (x_1, \dots, x_{l1}, \dots, x_{ld}, \dots, x_{ld})$ și $t = (t_1, \dots, t_1, \dots, t_l, \dots, t_l)$, unde fiecare dintre intrările $l \cdot d$ ale celor doi vectori reprezintă intrări și ieșiri corespunzătoare ale f . Pentru $x, t \in R^{l \cdot d}$, definim apoi $T = (x, t) \in T$ ca fiind o serie temporală de lungime $len(T) = l$ și dimensiune $dim(T) = d$.

Notă: Definiția este preluată din [1].

Anexa 2. Aspectul inițial al datelor

pacient 000034													
HR	O2Sat	Temp	SBP	MAP	DBP	Resp	EtCO2	BaseExcess	HCO3	FiO2	pH	PaCO2	
1	88	100	36.11	105	74	55	15.5	NA	3	29	1	7.31	62
2	88	98	36.17	103	75.5	58	20	NA	2	NA	0.6	7.31	62
3	88	95	NA	98	73	57	20	NA	3	NA	0.6	7.37	53
4	83.5	97	NA	92.5	69.5	56.5	19.5	NA	NA	NA	0.6	NA	NA
5	80	92	NA	124	89	69	20	NA	4	NA	NA	7.41	49
6	88	93	36.5	90	67	54	27	NA	3	NA	0.5	7.34	59
7	91	92	NA	110	82	64	24	NA	NA	NA	0.7	NA	NA
8	88	91	NA	99	75	59	23	NA	2	NA	NA	7.31	63
9	83	93	NA	91	68	55	18	NA	NA	NA	NA	NA	NA
10	80	94	NA	110	78	59	19	NA	NA	NA	NA	NA	NA
11	80	91	NA	96	68	52	20	NA	NA	NA	NA	NA	NA
12	82	89	NA	98	71	55	22	NA	NA	NA	NA	NA	NA
13	77	88	NA	115	81	60	24	NA	NA	NA	NA	NA	NA

SaO2	AST	BUN	Alkalinephos	Calcium	Chloride	Creatinine	Bilirubin_direct	Glucose	Lactate	Magnesium	Phosphate	Potassium
NA	NA	15	NA	NA	106	0.9	NA	127	NA	NA	NA	3.7
92	NA	NA	NA	NA	NA	NA	NA	116	NA	NA	NA	NA
94	NA	NA	NA	NA	NA	NA	NA	108	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
95	NA	NA	NA	NA	104	NA	NA	125	1.1	2.5	NA	4.2
93	NA	NA	NA	NA	NA	NA	NA	151	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
92	NA	NA	NA	NA	NA	NA	NA	136	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	16	NA	NA	105	1.3	NA	NA	NA	NA	NA	4.2
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Bilirubin_total	TroponinI	Hct	Hgb	PTT	WBC	Fibrinogen	Platelets	Age	Gender	Unit1	Unit2	HospAdmTime	ICULOS	SepsisLabel
NA	NA	27.5	8.5	40.1	15	NA	472	77.28	0	0	1	-22.64	3	0
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	4	0
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	5	0
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	6	1
NA	NA	28.3	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	7	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	8	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	9	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	10	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	11	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	12	1
NA	NA	26.8	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	13	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	14	1
NA	NA	NA	NA	NA	NA	NA	NA	77.28	0	0	1	-22.64	15	1

Notă: Rândurile reprezintă observații orare, iar coloanele – parametri fiziologici, de laborator, demografici și logistici monitorizați: frecvența cardiacă (HR), saturația sângelui periferic cu O₂ (SpO₂), temperatura (Temp), presiunea arterială sistolică (SBP), presiunea arterială medie (MAP), presiunea arterială diastolică (DBP), frecvența respirației (Resp), capnometrie (EtCO₂), fracția O₂ inspirat (FiO₂), excesul de baze (BaseExcess), bicarbonatul (HCO₃), pH, presiunea parțială a CO₂ în sângele arterial (PaCO₂), saturația cu O₂ sângelui arterial (SaO₂), aspartat-transaminaza (AST), ureea (BUN), fosfataza alcalină (Alkalinephos), calciul seric (Ca), clorul seric (Chloride), creatinina serică (Creat), bilirubina directă (DB), glucoza (Gluc), lactatul seric (Lact), magneziul seric (Mg), fosfații serici (Phos), potasiul seric (K), bilirubina totală (TB), troponina-I (TroponinI), hematocritul (Hct), hemoglobina (Hgb), timpul trombolastinei parțial activate (aPTT), numărul de leucocite (WBC), fibrinogenul (Fibrinogen), numărul de trombocite (Platelet), vârsta (Age), sexul (Gender), timpul de internare (HospAdmTime), durata aflării în terapie intensivă (ICULOS), identificatorul secției (Unit # 1 & 2), marcaj sepsis (SepsisLabel);

Anexa 3. Aspecte de corelare dintre parametrii de interes în setul B

1. Sepsis

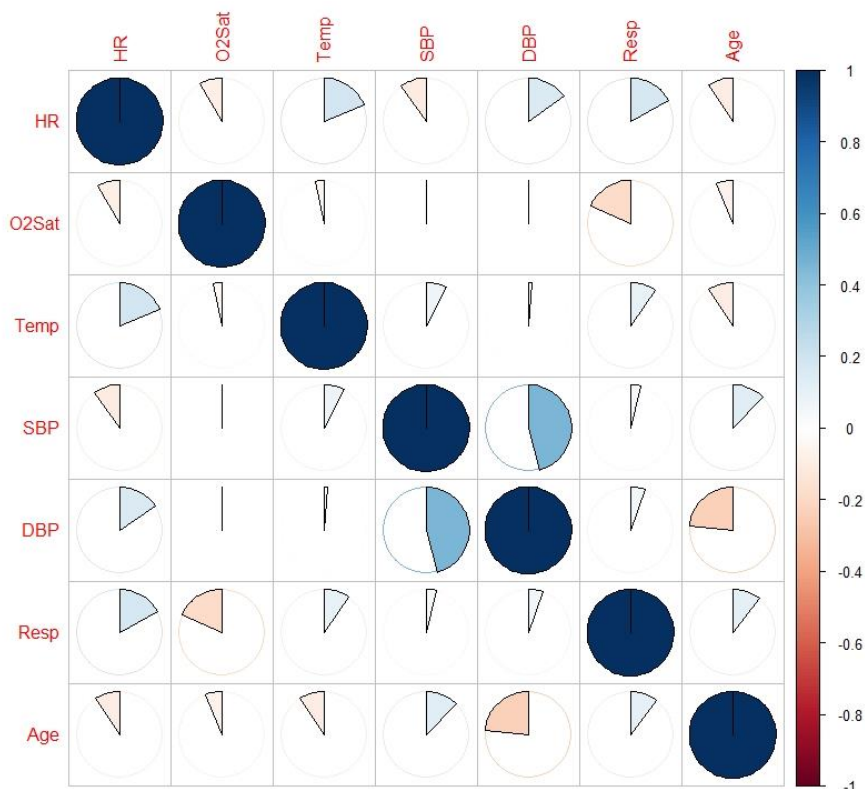


Figura. A.3.1. Corelarea parametrilor fiziologici în sepsis (setul B).

Tabelul A.3.1. Matricea de corelare a parametrilor fiziologici in sepsis (set B)

	FCC	SpO ₂	Temp	PAS	PAD	Resp	Vârsta
FCC	1.00	-.08	.19	-.10	.15	.17	-.09
SpO ₂	-.08	1.00	-.03	.01	.00	-.18	-.06
Temp	.19	-.03	1.00	.08	.01	.10	-.09
PAS	-.10	.01	.08	1.00	.46	.04	.12
PAD	.15	.00	.01	.46	1.00	.06	-.23
Resp	.17	-.18	.10	.04	.06	1.00	.10
Vârsta	-.09	-.06	-.09	.12	-.23	.10	1.00

2. Non-sepsis

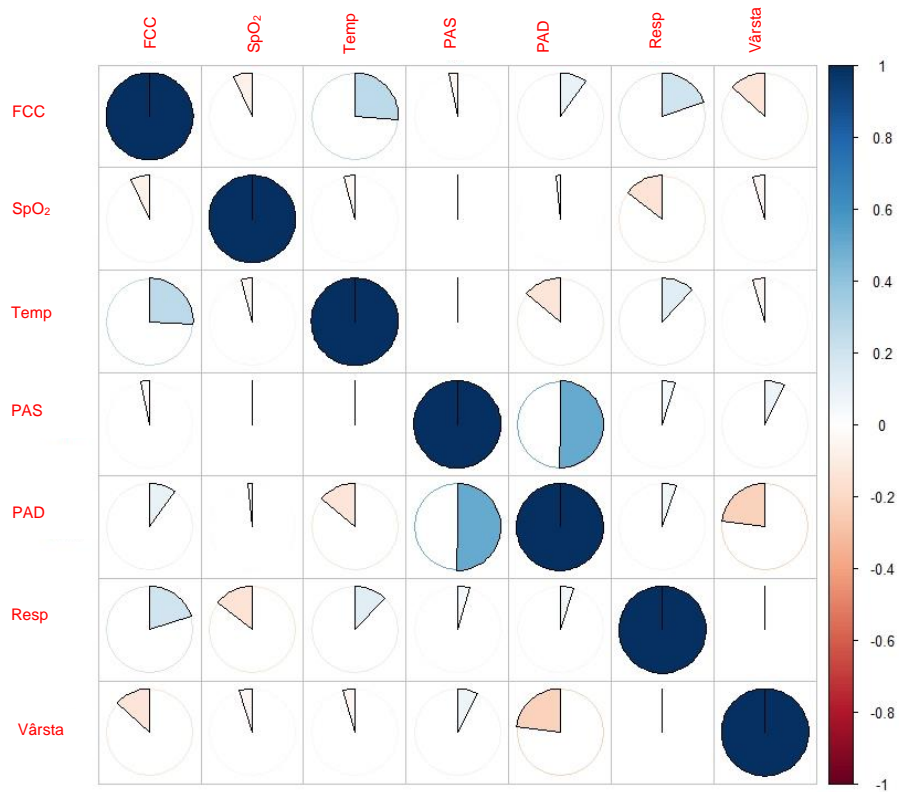


Figura. A.3.2. Corelarea parametrilor fiziologici în cazurile non-septice.

Tabelul A.3.2. Matricea de corelare a parametrilor fiziologici la non-septici

	FCC	SpO ₂	Temp	PAS	DBP	Resp	Vârsta
FCC	1.00	-0.07	.26	-0.03	.10	.20	-0.13
SpO ₂	-0.07	1.00	-0.04	-0.01	.03	-0.15	-0.05
Temp	.26	-0.04	1.00	.00	.50	.05	-0.04
PAS	-0.03	-0.01	.00	1.00	.50	.05	-0.02
DBP	.10	-0.01	-0.14	.50	1.00	.05	-0.23
Resp	.20	-0.15	.12	.05	.05	1.00	.01
Vârsta	-0.13	-0.05	-0.04	.08	-0.23	.01	1.00

Notă:

FCC (HR) – frecvența cardiacă; PAS/PAD (SBP/DBP) – presiunea arterială sistolică/diastolică; SpO₂ (O₂Sat) – saturația sângelui periferic cu O₂; Resp – frecvența respirației; Temp – temperatura corpului

Anexa 4. Aspecte de corelare dintre parametrii de interes în seturile “Pre-sepsis” și “Sepsis confirmat” (set A)

1. Pre-sepsis

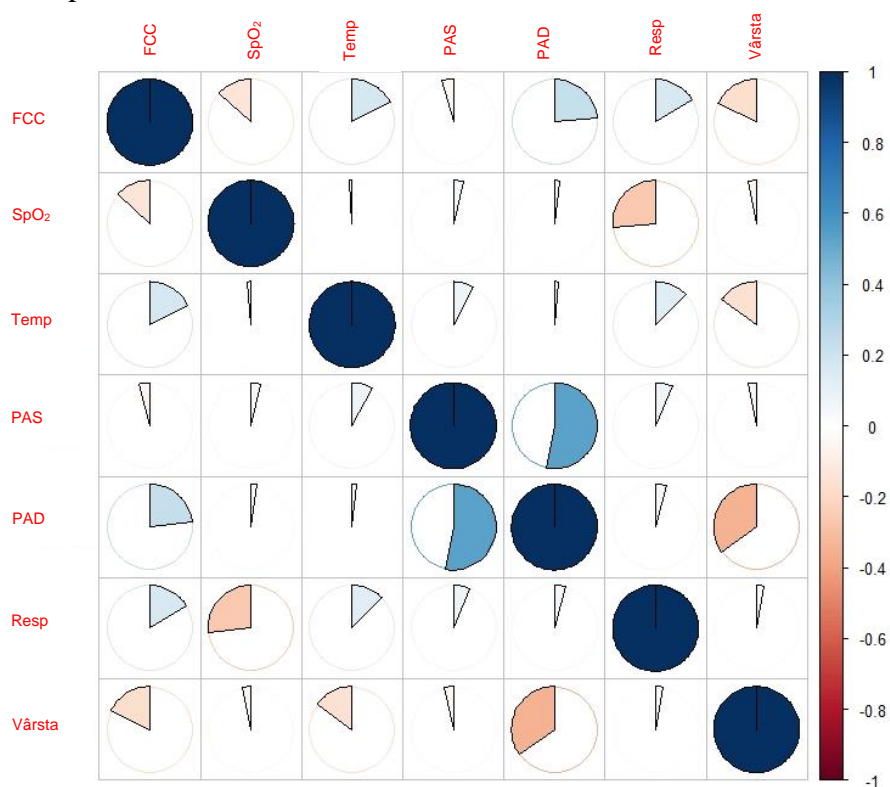


Figura. A.4.1. Corelarea parametrilor fiziologici în pre-sepsis (setul A).

Tabelul A.4.1. Matricea de corelare a parametrilor fiziologici în pre-sepsis (set A)

	FCC	SpO ₂	Temp	PAS	PAD	Resp	Vârsta
FCC	1.00	-.13	.17	-.04	.23	.16	-.18
SpO ₂	-.13	1.00	-.01	.04	.02	-.27	-.03
Temp	.17	-.01	1.00	.08	.01	.12	-.15
PAS	-.04	.04	.08	1.00	.53	.06	-.03
PAD	.23	.02	.01	.53	1.00	.04	-.35
Resp	.16	-.27	.12	.06	.04	1.00	.03
Vârsta	-.18	-.03	-.15	-.03	-.35	.03	1.00

2. Sepsis confirmat clinic

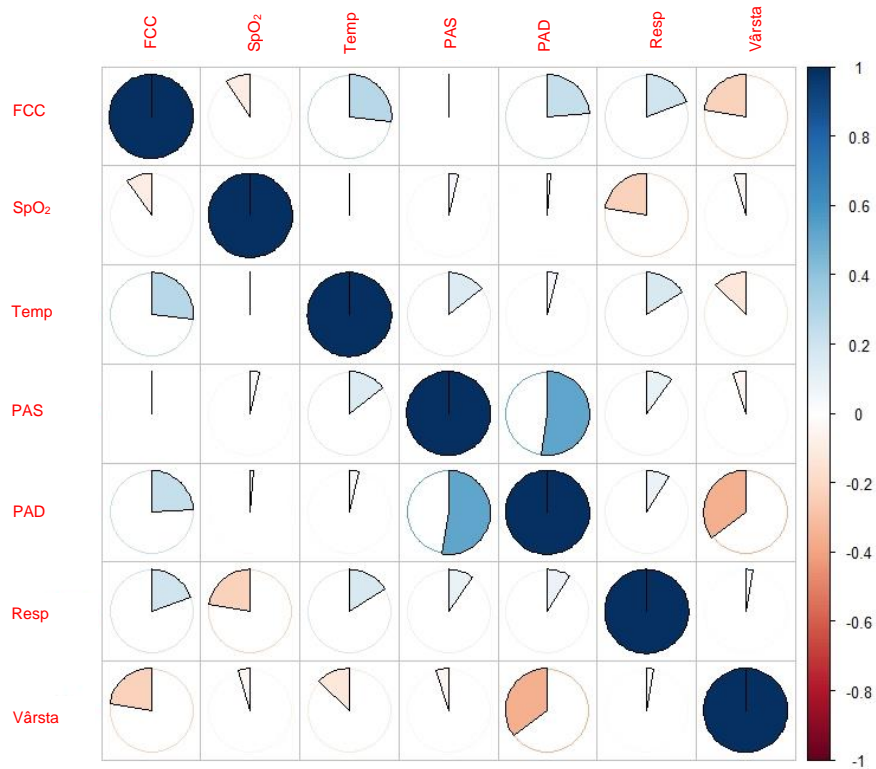


Figura. A.4.2. Corelarea parametrilor în cazurile sepsice, confirmate clinic

Tabelul A.4.2. Matricea de corelare a parametrilor în sepsis confirmat clinic

	FCC	SpO ₂	Temp	PAS	PAD	Resp	Vârsta
FCC	1.00	-0.10	0.27	-0.01	0.24	0.19	-0.22
SpO ₂	-0.10	1.00	0.01	0.04	0.01	-0.22	-0.04
Temp	0.27	0.01	1.00	0.14	0.04	0.16	-0.13
PAS	-0.01	0.04	0.14	1.00	0.53	0.10	-0.05
PAD	0.24	0.01	0.04	0.53	1.00	0.09	-0.35
Resp	0.19	-0.22	0.16	0.10	0.09	1.00	0.03
Vârsta	-0.22	-0.04	-0.13	-0.05	-0.35	0.03	1.00

Notă:

FCC (HR) – frecvența cardiacă; PAS/PAD (SBP/DBP) – presiunea arterială sistolică/diastolică; SpO₂ (O₂Sat) – saturația sângelui periferic cu O₂; Resp – frecvența respirației; Temp – temperatura corpului

Anexa 5. Similaritatea seturilor de date “Sepsis” și “Non-sepsis”

Tabelul A.5.1. Caracteristicile setului “Sepsis”

	FCC	SpO ₂	Temp	PAS	PAD	Resp
media	88.17743475	9.772495e+01	37.24051241	122.70056197	60.4420978	19.65160953
dev. stand.	17.71204178	2.525587e+00	0.79071592	22.14424934	12.5376451	5.66539439
asimetria	0.28586939	-2.615098e+00	-0.12438538	0.48744433	0.8158736	0.69519206
boltirea	-0.14643032	1.944242e+01	0.27785707	-0.03838681	1.9795383	0.77712812
medcouple	0.06666667	-5.144411e-09	-0.07042254	0.12820513	0.11111111	0.09090909

Tabelul A.5.2. Caracteristicile setului “Non-sepsis”

	FCC	SpO ₂	Temp	PAS	PAD	Resp
media	86.10365442	9.769222e+01	37.11116451	118.9268933	58.8555617	17.9609866
dev. stand.	14.99563006	2.508069e+00	0.73900472	20.1061825	11.5852795	5.2064486
asimetria	0.35824979	-2.548373e+00	-0.25115847	0.7005300	0.9397684	0.8138621
boltirea	0.49366161	2.134259e+01	0.35087813	-0.6440265	2.4660501	1.3046556
medcouple	-0.03703704	-1.032181e-09	-0.09876543	0.1478261	0.1851852	0.2000000

Tabelul A.5.3. Testul Kolmogorov-Smirnov cu două eșantioane (valoare p): sepsis vs non-sepsis

FCC	SpO ₂	Temp	PAS	PAD	Resp
0.000000e+00	3.753817e-08	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00

Distanța Hellinger NULL

Notă:

FCC – frecvența cardiacă

SpO₂ – saturația sângelui periferic cu oxigen

Temp – temperatura corpului

PAS/PAD – presiunea arterială sistolică/diastolică

Resp – frecvența respirației

mean – valoare medie

stdev – deviație standard

skewness - asimetrie

kurtosis - boltire

medcouple – mediana diferenței dintre mijlocul părții stângi și drepte a unei distribuții

Anexa 6. Similaritatea seturilor de date “Pre-sepsis” și “Non-sepsis”

Tabelul A.6.1. Caracteristicile setului “Pre-sepsis”

	FCC	SpO ₂	Temp	PAS	PAD	Resp
media	88.72826987	9.765480e+01	37.22507036	122.26075331	60.5815397	19.9015935
dev. stand.	17.76361763	2.477109e+00	0.79504870	22.24266615	12.2644736	5.7383527
asimetria	0.24773320	-1.375806e+00	-0.25953231	0.46022298	0.6250477	0.7108331
boltirea	-0.10349995	2.305686e+00	0.31368631	-0.07141349	0.8245330	0.6812149
medcouple	0.03225806	-5.148756e-09	-0.08256881	0.10000000	0.1428571	0.1666667

Tabelul A.6.2. Caracteristicile setului “Non-sepsis”

	FCC	SpO ₂	Temp	PAS	PAD	Resp
media	86.10365442	9.769222e+01	37.11116451	118.9268933	58.8555617	17.9609866
dev. stand.	14.99563006	2.508069e+00	0.73900472	20.1061825	11.5852795	5.2064486
asimetria	0.35824979	-2.548373e+00	-0.25115847	0.7005300	0.9397684	0.8138621
boltirea	0.49366161	2.134259e+01	0.35087813	-0.6440265	2.4660501	1.3046556
medcouple	-0.03703704	-1.032181e-09	-0.09876543	0.1478261	0.1851852	0.2000000

Tabelul A.6.3. Testul Kolmogorov-Smirnov cu două eșantioane (valoare p): pre-sepsis vs non-sepsis

FCC	SpO ₂	Temp	PAS	PAD	Resp
0.000000e+00	8.069845e-01	0.000000e+00	3.330669e-16	3.624878e-13	0.000000e+00

Distanța Hellinger

NULL

Notă:

FCC – frecvența cardiacă

SpO₂ – saturația sângelui periferic cu oxigen

Temp – temperatura corpului

PAS/PAD – presiunea arterială sistolică/diastolică

Resp – frecvența respirației

mean – valoare medie

stdev – deviație standard

skewness - asimetrie

kurtosis - boltire

medcouple – mediana diferenței dintre mijlocul părții stânga și drepte a unei distribuții

Anexa 7. Informație suplimentară în privința trainingului și cross-validării modelelor de bază

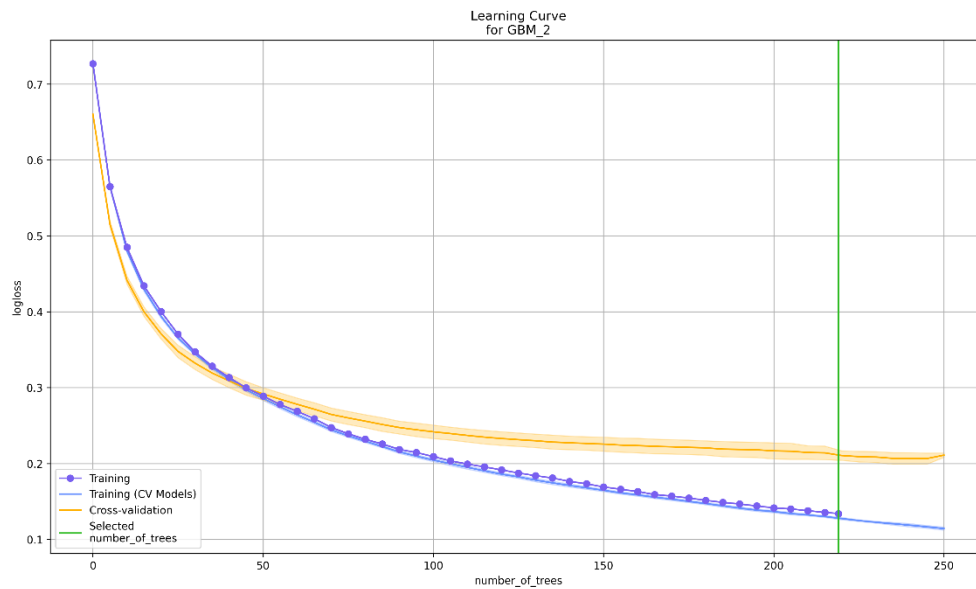


Figura A.7.1. Curbele de învățare (culoare albastră) și cea obținută la cross-validare (culoare orange).

Notă: Ambele curbe ilustrează diminuarea erorii de clasificare (logloss) în dependență de numărul de arbori. Cu culoare verde este indicată selecția finală a acestui număr la trainingul modelului

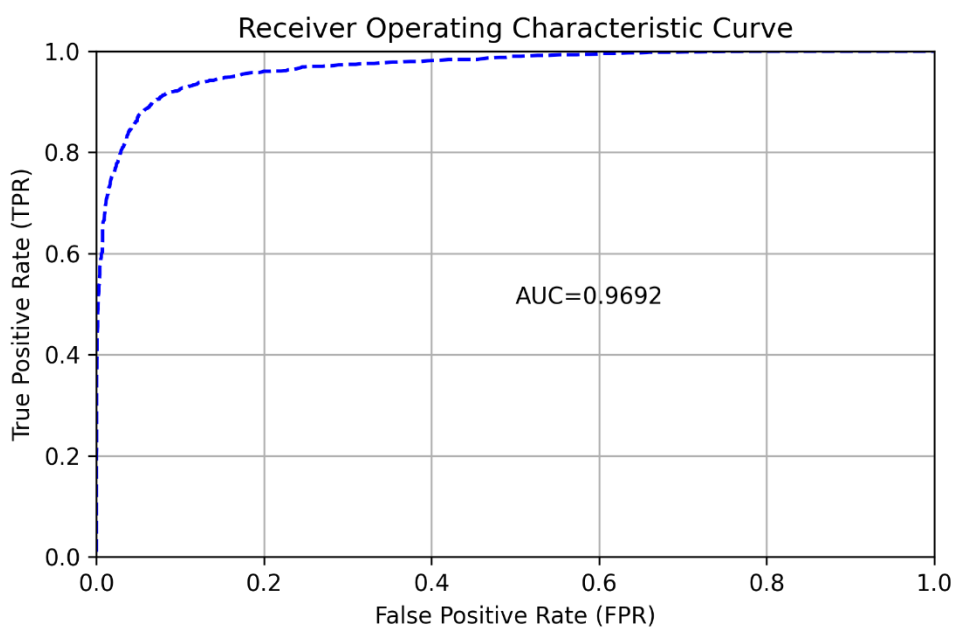


Figura A.7.2. Curba ROC și aria de sub ea în rezultatul trainingului modelului.

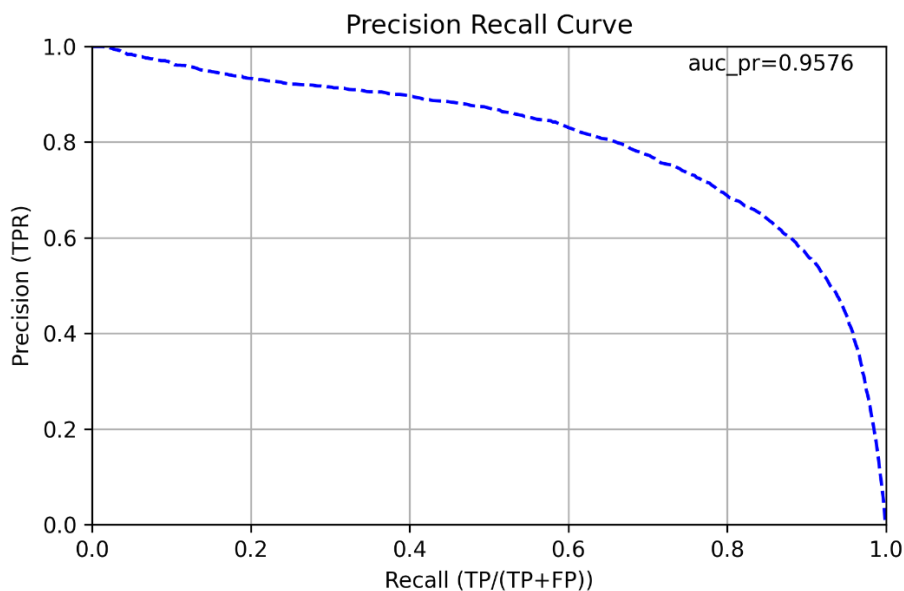


Figura A.7.3. Curba AUC-PR,

Notă: Această curbă este asemănătoare cu curba ROC, însă cu accent pe precizia modelului.

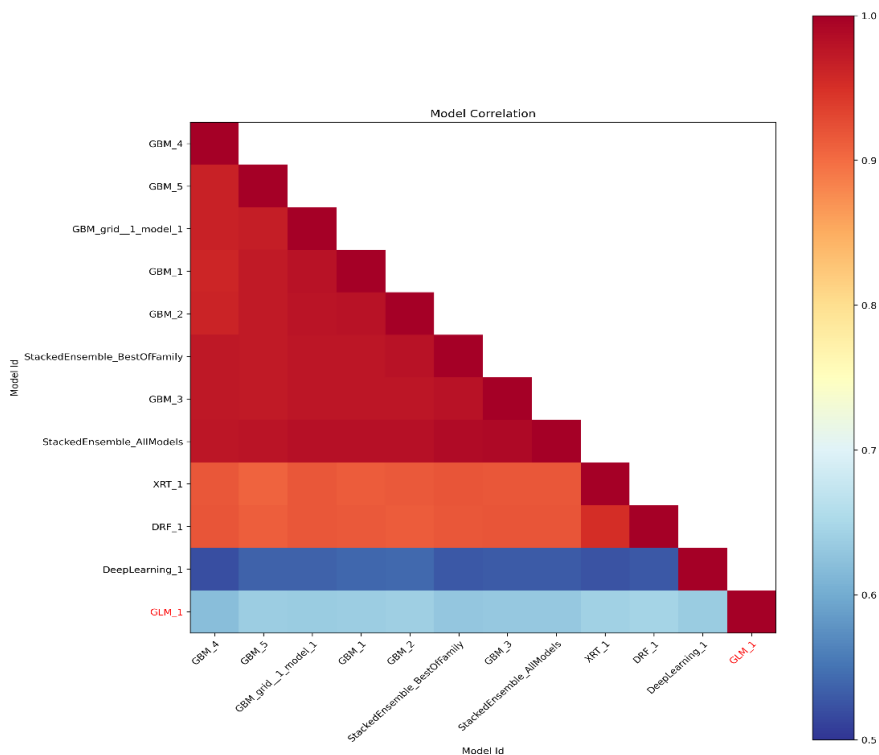


Figura A.7.4. Corelarea performanței modelelor construite în baza diferitor algoritmi.

Notă: Inscripția cu roșu denotă modelul in baza regresiei (GLM), care are una dintre cele mai joase performanțe.

Anexa 8. Performanța sistemului la prezicerea continue a riscului sepsisului

Tabelul A.8.1. Rezultatele orare ale prezicerii riscului de dezvoltare a sepsisului la 10 pacienți septici timp de 9 ore consecutive

Pacient ID	ora 1	ora 2	ora 3	ora 4	ora 5	ora 6	ora 7	ora 8	ora 9
p002706	1	1	1	1	1	1	1	1	1
p002808	1	1	1	1	1	1	1	1	0
p004197	1	1	1	1	1	1	1	1	1
p011056	1	1	1	1	1	1	1	1	1
p011676	1	1	1	1	1	1	1	1	1
p013361	1	1	1	1	1	1	1	1	1
p014518	0	0	1	0	1	1	1	1	1
p015892	1	0	0	1	1	1	0	0	0
p016487	1	1	1	1	1	1	1	1	1
p020538	1	1	1	1	1	1	1	1	1

Notă: 1 – sepsis, 0 – non-sepsis. Sepsisul - diagnosticat prin metode tradiționale la ora a 7-a.

Tabel A.8.2. Rezultatele orare ale prezicerii riscului de dezvoltare a sepsisului la 10 pacienți fără sepsis timp de 9 ore consecutive

Pacient ID	ora 1	ora 2	ora 3	ora 4	ora 5	ora 6	ora 7	ora 8	ora 9
p008368	0	0	0	0	0	0	1	0	0
p012801	0	0	0	0	0	0	0	0	0
p013050	0	0	0	0	0	0	0	0	0
p013661	0	0	0	0	1	0	0	0	0
p014981	1	0	0	0	0	0	0	0	0
p015223	0	0	0	0	0	0	0	0	0
p015886	0	0	0	0	0	0	0	0	0
p017073	0	0	0	0	0	0	0	0	0
p017567	0	0	1	0	0	0	0	0	0
p020600	0	0	0	0	0	0	0	0	0

Notă: 1 – sepsis, 0 – non-sepsis.

Anexa 9. **Raportarea rezultatelor studiilor cu utilizarea învățării automate,**

după Kapoor S., Narayanan A, Leakage and the Reproducibility Crisis in ML-based Science, 2022, arXiv:2207.07048v1

Model Info Sheet

Section 1: Information about paper or report

1) Author(s): Names of the authors of the paper or report

Victor Iapăscurtă

2) Title of the paper or report which introduces the model

Prezicerea timpurie a sepsisului cu ajutorul unei aplicații proprii elaborate în baza învățării automate (inteligență artificială)

3) DOI or permanent link to the paper or report (for example, link to arxiv.org webpage)

<https://notebookarchive.org/2022-03-b3yfik0>

4) License: Under which license(s) are the data and/or model shared?

ODS (Open Database License)

5) Email address of the corresponding author

victor.iapascurta@usmf.md

Section 2: Scientific claim(s) of interest

6) Does your paper make a generalizable claim based on the ML model? If yes, what is the scientific claim? For example, “Our ML model can be used to diagnose Covid-19 using chest radiographs of adult patients”.

Aplicația software creată poate prezice cu un orizont de 4 ore și probabilitate înaltă (mai mare de 90%) starea de sepsis iminentă în baza a 6 parametri clinici monitorizați de rutină la pacienți adulți în secția de terapie intensivă

7) Is the scientific claim made about a distribution or population from which you can sample?

Datele utilizate în cercetare provin de la pacienți adulți (18 – > 100 ani) aflați la tratament în secțiile de terapie intensivă în două spitale din SUA în deceniul precedent (2010 – 2020) cu diferite patologii, inclusiv sepsis (seturile A și B).

Extrapolarea rezultatelor cercetării prin utilizarea aplicației create la pacienții din R. Moldova va necesita calibrarea aplicației la condițiile și specificul secției TI concrete.

8) Does the scientific claim only apply to certain subsets of the distribution mentioned in Q6? For example, “Our model works on chest radiographs of U.S.-based adult patients and might not generalize to radiographs taken in other places or using different machines.”

Aplicația creată poate să nu generalizeze reușit pe pacienți din secțiile din Moldova. Pentru ajustarea aplicației la condițiile locale vor fi necesare cercetări suplimentare și calibrarea aplicației în condițiile concrete ale potențialei utilizări. O soluție pentru ajustarea aplicației ar fi învățarea prin transfer.

Section 3: Train-test split is maintained across all steps in creating the model

9) Train-test split type: How was the dataset split into train and test sets? (For example, cross-validation; separate train and test sets).

Scindarea datelor în seturile pentru training și testare a fost efectuată la începutul cercetării, după excluderea cazurilor care conțin artefacte. Cu toate că s-a experimentat cu diferite proporții de scindare (pentru setul de training – de la 25% până la 90% de date – cu rezultate de performanță a modelelor finale destul de apropiate) aplicația finală a fost creată în baza proporției 90% (training) vs 10% (test). Fiecare din seturi (training și test) a fost procesat separat la toate etapele cercetării.

10) Are there duplicates in the dataset? If yes, explain how duplicates are handled to ensure the train-test split.

Nici la una din etape ale cercetării nu s-a operat cu duplicate ale datelor.

11) In case the dataset has dependencies (e.g., multiple rows of data from the same patient), describe how the dependencies were addressed (for example, using block-cross validation).

Specimenele finale, care sunt utilizate pentru învățarea automată, reprezintă numai pacienți care au fost inițial alocați unui anumit set și pe tot parcursul procesării datelor și creării modelului datele din seturi separate nu se intersectează.

12) List all the pre-processing steps used in creating your model. For example, imputing missing data, normalizing feature values, selecting a subset of rows from the dataset for building the model.

Operațiile de preprocesare au fost efectuate separat pe fiecare set și au inclus:

- a) Excluderea subseturilor/cazurilor care conțin artefacte
- b) Scindare (prin eșantionare aleatorie) a datelor cu crearea setului pentru training (90% date – 1533 cazuri/pacienți cu sepsis și 3003 – non-sepsis) și setului-test (10% date – 173 cazuri/pacienți cu sepsis și 333 – non-sepsis) *NB: setul pentru training a mai fost scindat (de asemenea prin eșantionare aleatorie) în seturi mai mici (75%, 50% și 25% din mărimea inițială) care mai târziu au fost utilizate pentru evaluarea robusteții modelului.*
- c) Restabilirea datelor lipsă prin interpolare (funcție creată în baza pachetului zoo, R) cu excepția coloanelor cu toate valorile lipsă (pentru fiecare set separat)
- d) Excluderea subseturilor/cazurilor septice cu valori lipsă în mai mult de o coloană cu următoarele rezultate: set sepsis pentru training – 1431 subseturi, test-set – 159 subseturi) *NB: a fost creat un set cu excluderea tuturor subseturilor/cazurilor cu toate valorile lipsa*

pentru oarecare din cei 6 parametri(set sepsis pentru training – 1073 subseturi; set-test – 128 subseturi)

- e) Restabilirea datelor în coloanele cu toate valorile lipsă prin extragere și înlocuire a valorilor lipsă cu date din distribuții similare celor în datele prezente (create cu pachetul `sn` în R) *NB: Această etapă nu a fost aplicată seturilor fără toate valorile lipsa într-o coloană/pentru un parametru*

13) How was the train-test split maintained during each pre-processing step? If applicable, use a separate line for each step mentioned in Q14.

Scindarea se păstrează prin prelucrarea separată a datelor din setul pentru training și setul-test

14) List all the modeling steps used in creating your model. For example, feature selection, parameter tuning, model selection.

În baza datelor restabilite (vezi pașii precedenți – 12(c) și 12(e)) s-a efectuat generarea specimenelor finale, care reprezintă vectori cu lungimea 14L (primele două elemente reprezentând complexitatea algoritmică prin metoda BDM (block decomposition method) a 2 matrice, create prin gruparea celor 6 parametri în 2 grupuri și utilizând 3 valori orare pentru fiecare parametru (se utilizează varianta binarizată a matricelor), iar celelalte 12 elemente ale vectorului fiind reprezentate de diferența dintre valorile celor 6 parametri în trei coșuri orare consecutive). La acest pas prelucrarea datelor se face separat pe seturi și la seturile care reprezintă pacienții septici se mai adaugă un element în vector – eticheta 1, iar pentru cazurile non-septice – 0. Astfel, numărul caracteristicilor finale este 14 plus eticheta. Inițial modelele se construiesc cu ajutorul învățării automate automatizate (AutoML). Ulterior hiperparametrii optimi finali sunt justificați prin căutare grilă grosieră cu cross-validare multiplă (x 10). Selectarea celui mai performant model se face în baza AUC la testarea pe setul test.

15) How is the train-test split is observed during each modeling step? If applicable, use a separate line for each step mentioned in Q16.

Prelucrarea separată a seturilor se efectuează pe parcursul tuturor pașilor descriși până acum.

16) List all the evaluation steps used in evaluating model performance. For example, cross-validation, out-of-sample testing.

Performanța modelului a fost evaluată în cadrul cross-validării, pe setul test creat de la început din setul A (validare internă) și pe setul B (validare externă).

17) How is the train-test split observed during each evaluation step? If applicable, use a separate line for each step mentioned in Q18.

Vezi răspunsurile 9, 11, 13 și 14.

Section 4: Test set is drawn from the distribution of scientific interest.

18) Why is your test set representative of the population or distribution about which you are making your scientific claims?

Setul-test A este selectat prin eşantionare aleatorie din setul A integral. Similaritatea seturilor test-A și training-A este investigată prin testul KS și distanța Hellinger, care confirmă această similaritate. Aceiași situație este și cu setul B în calitate de set-test suplimentar.

19) Explain the process for selecting the test set and why this does not introduce selection bias in the learning process.

Vezi răspunsul 18

20) In case your model is used to predict a future outcome of interest using past data, detail how data in the training set is always from a date earlier than the data in the test set.

Cu toate că fiecare specimen reprezintă situația pentru o anumită perioadă de observare (3 ore) el este în ultimă instanță independent de cursul natural al timpului astronomic. Fiecare specimen este precăutat ca având a-priori aceeași probabilitate de a fi clasificat ca septic sau non-septic și în urma aplicării modelului se estimează probabilitatea a-posteriori. Separarea seturilor pentru training și pentru testare și evitarea intersectării acestor seturi la procesarea datelor exclude și intersectarea trecutului și viitorului.

Aspectul de mixare a trecutului și viitorului ar deveni actual la următoarea etapa de dezvoltare a sistemului – la calibrarea/ajustarea lui la condițiile secției de terapie intensivă, unde ar putea fi eventual folosit. Aceasta iese în afara studiului curent.

Section 5: Each feature used in the model is legitimate for the task

21) List the features used in the model, alongside an argument for their legitimacy. A legitimate feature is one that would be available when the model is used in the real world and is not a proxy of the outcome being predicted.

Parametrii fiziologici inițiali (6 parametri: FCC, SpO₂; Temp, PAS, PAD, Resp) în baza cărora se construiește setul de caracteristici (14 caracteristici) utilizat de model, sunt de rutină/standard monitorizați în secțiile de terapie intensivă. La etapa curentă de dezvoltare a medicinei nici unul din acești parametri separat nu este utilizat pentru prezicerea sepsisului cu ajutorul metodelor tradiționale. Scorul qSOFA folosește SBP și Resp, ca 2 dintre cele 3 elemente ale acestui scor. Scorul SOFA folosește SBP și DBP pentru a calcula presiunea medie. Ambele scoruri chiar și dacă folosesc unii din cei 6 parametri de interes în studiul curent, o fac în alt mod decât cercetarea curentă. Ambele scoruri fiind folosite pentru confirmarea diagnozei de sepsis, au o abilitate de prezicere a sepsisului inferioară modelului creat.

Anexa 10. Performanța a șapte modele antrenate pe date din setul A în cadrul învățării automate automatizate (AutoML) în baza platformei H2O (etapa de cross-validare)

	Model ID	AUC	Logloss	AUC PR	Mean per class error	RMSE	MSE
1	GBM_1_AutoML_20220726_140221	0.9702639	0.2027395	0.9588501	0.08441366	0.2404108	0.05779735
2	GBM_2_AutoML_20220726_140221	0.9697746	0.2033073	0.9587876	0.08357230	0.2409972	0.05807963
3	GBM_3_AutoML_20220726_140221	0.9691409	0.2043975	0.9579259	0.08674099	0.2422572	0.05868854
4	GBM_4_AutoML_20220726_140221	0.9676558	0.2089869	0.9565762	0.08573528	0.2446899	0.05987315
5	GBM_5_AutoML_20220726_140221	0.9675202	0.2085878	0.9557443	0.08839656	0.2450129	0.06003132
6	DRF_1_AutoML_20220726_140221	0.9309430	0.3530599	0.9048406	0.14338338	0.3282694	0.10776079
7	GLM_1_AutoML_20220726_140221	0.7025239	0.5925449	0.4993538	0.34323260	0.4518352	0.20415501

Notă:

Model ID – identificatorul modelului

Mean per class error – eroarea medie pe clasă

AUC – aria de sub curba ROC (area under the curve)

RMSE (root mean square error) și MSE (mean square error) - eroare pătratică medie

Logloss – funcția de diminuare a erorii (pierdere)

AUC – PR – curba Precision-Recall

Anexa 11. **Diagrame PD si ICE pentru unele caracteristici utilizate de modele GBM cu performanță înaltă în cadrul cercetării curente si diagrame de corelare între modele.**

Dependența parțială: Dif_Temp_1

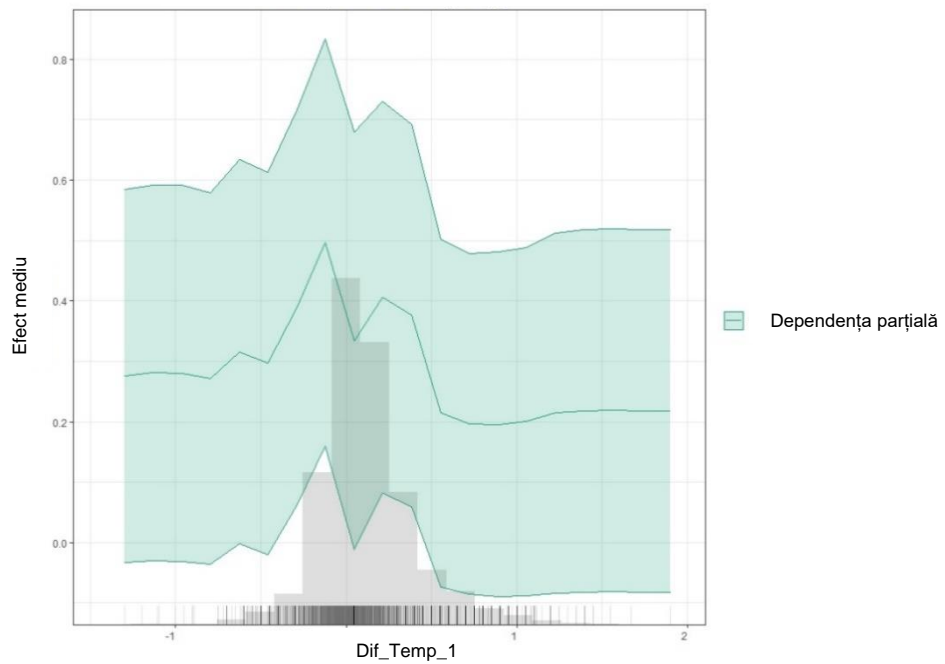


Figura A.11.1. **Diagrama PD (dependență parțială) pentru diferența de temperatură dintre ora a doua de observație si ora precedentă în cazul unui model GBM**

Dependența parțială: Dif_Temp_2

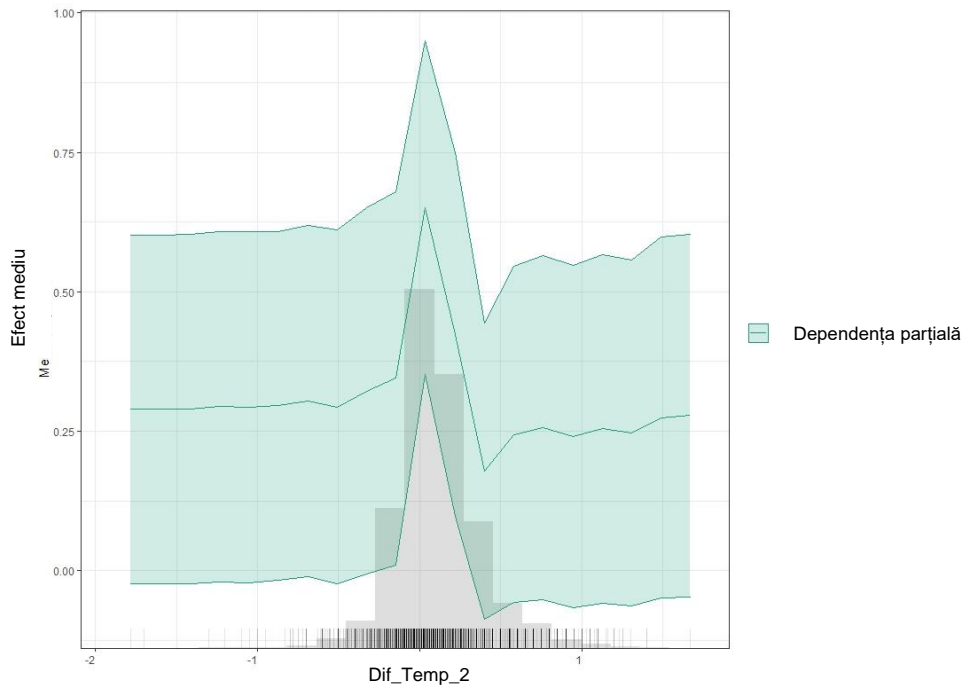


Figura A.11.2. **Diagrama PD (dependență parțială) pentru diferența de temperatură dintre ora a treia de observație si ora precedentă în cazul unui model GBM**

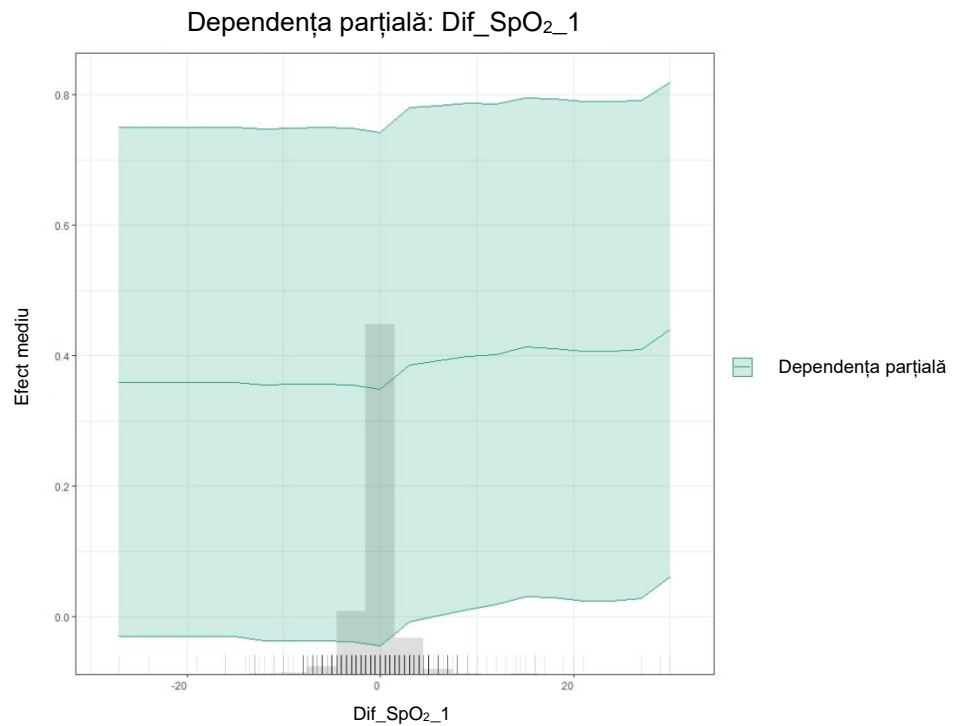


Figura A.11.3. Diagrama PD (dependență parțială) pentru diferența SpO₂ dintre ora a doua de observație și ora precedentă în cazul unui model GBM

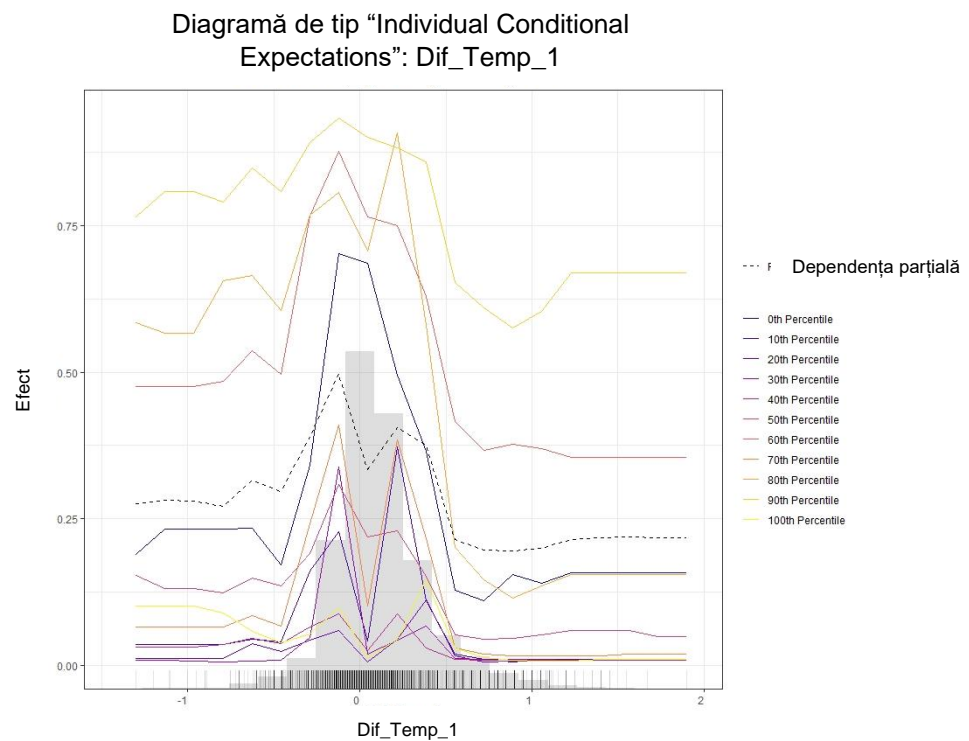


Figura A.11.4. Diagrama ICE pentru diferența de temperatură între ora a treia de observație și ora a doua în cazul unui model GBM cu performanță înaltă.

Diagramă de tip "Individual Conditional Expectations": Dif_Temp_2

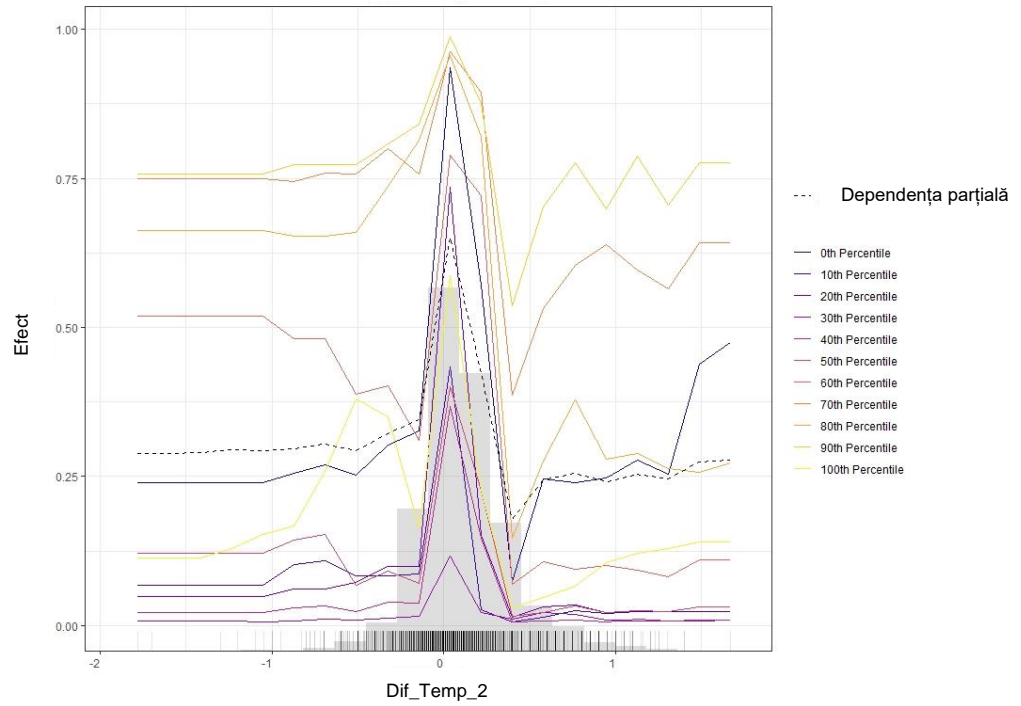


Figura A.11.5. Diagrama ICE pentru diferența de temperatură între ora a doua de observație și ora precedentă în cazul unui model GBM cu performanță înaltă.

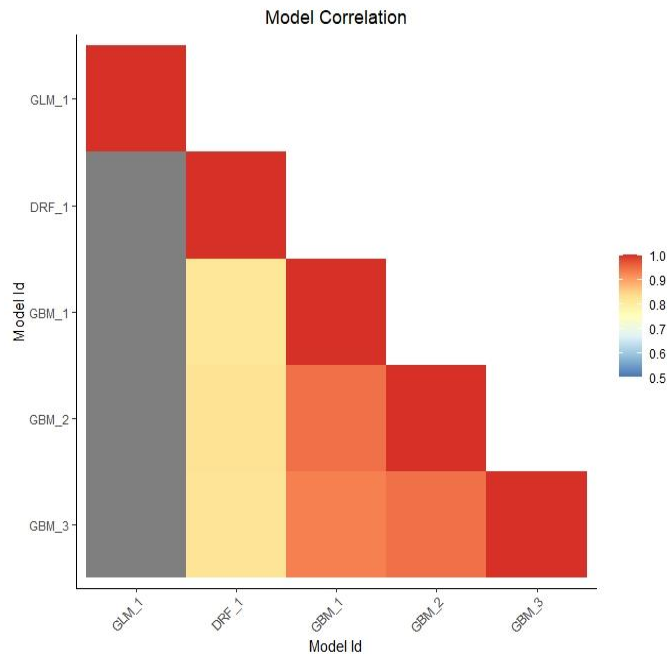


Figura A.11.6. Diagrama de corelare (după performanță) între modele din clase diferite.

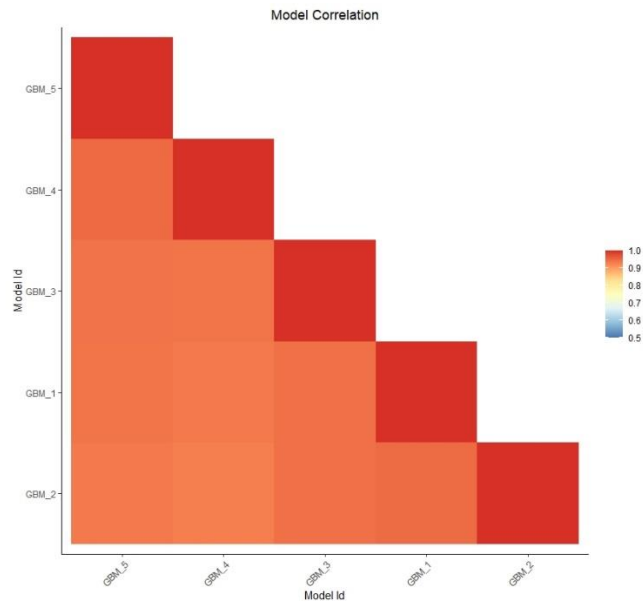


Figura A.11.7. **Diagrama de corelare (după performanță) între modele din aceeași clasă**

Ultimele două diagrame arată corelația dintre predicțiile modelelor. Pentru clasificare se folosește frecvența predicțiilor identice. În mod implicit, modelele sunt ordonate după asemănarea lor (calculată prin clasterizare ierarhică). Ultima diagramă reprezintă corelarea dintre cinci modele din aceeași clasă – GBM (1, 2, 3, 4 și 5)

Anexa 12. **Certificate de inovator. Acte de implementare**



REPUBLICA MOLDOVA
MINISTERUL SĂNĂTĂȚII

INSTITUȚIA MEDICO-SANITARĂ PUBLICĂ
INSTITUTUL DE MEDICINĂ URGENTĂ
(IMSP IMU)
MD-2004, mun. Chișinău, str. T. Ciorba, 1



CERTIFICAT DE INOVATOR

Nr. 18

data 12.12.2022

de înregistrarea inovației în Registrul obiectelor proprietății intelectuale ale IMSP IMU în
conformitate cu art. 16 al Legii nr. 138-XV din 10.05.2001

**IMPLEMENTAREA APLICAȚIEI SOFTWARE ISAAC-SEPSIS ÎN
BAZA ÎNVĂȚĂRII AUTOMATE/INTELIGENȚEI ARTIFICIALE
PENTRU PREZICEREA TIMPURIE A SEPSISULUI**

(denumirea)

IAPĂSCURTĂ Victor, BELÎI Adrian

(coautori)



Director

dr. hab. șt. med., prof. univ. M. CIOCANU

mun. Chișinău



REPUBLICA MOLDOVA
MINISTERUL SĂNĂTĂȚII, MUNCHI ȘI PROTECȚIEI SOCIALE
INSTITUȚIA MEDICO-SANITARĂ PUBLICĂ
INSTITUTUL DE MEDICINĂ URGENTĂ
MD-2004, mun. Chișinău, str. T. Ciorba, 1
tel.: 022-23-78-84, fax: 022-23-53-09,
e-mail: anticamera@urgenta.md
www.urgenta.md



APROB

Director IMSP IMU

Dr. hab șt. med., prof univ.,

M. CIOCANU


ACT DE IMPLEMENTARE

1. **Denumirea ofertei pentru implementare:** "IMPLEMENTAREA APLICAȚIEI SOFTWARE ISAAC-SEPSIS ÎN BAZA ÎNVĂȚĂRII AUTOMATE/INTELIGENȚEI ARTIFICIALE PENTRU PREZICEREA TIMPURIE A SEPSISULUI,,
2. **Autori:** IAPĂSCURTĂ Victor, asitent universitar, doctorand; BELÎI Adrian, dr.hab. șt.med., prof.univ.
3. **Unde și când a fost implementată:** Anul implementării 2019-2022, Catedra de anesteziologie și reanimatologie "V. Ghereg", USMF " Nicolae Testemițanu" în cadrul IMSP IMU; Departamentul Anesteziologie și Trapie Intensivă IMSP IMU.
4. **Rezultatul implimentării:** Ca rezultat al implementării sistemului de predicție timpurie a sepsisului este posibilă prezicerea timpurie a debutului sepsisului, ce permite inițierea în timp util a tratamentului cu antibiotice cu o eventuale rezultate pozitive, precum diminuarea complicațiilor, micșorarea perioadei de internare, etc.
5. **Eficacitatea implementării:** Prezicerea timpurie a sepsisului cu ajutorul aplicației propuse, cu un orizon de predicție de peste 4 ore (pana la posibilitatea identificării cu metodele traditionale) oferă atât beneficii terapeutice pentru pacienti, cât și pentru medici – prin asistarea lor în procesul decizional în cazurile cu grad crescut de incertitudine diagnostică.

Prezenta inovație este implementată conform descrierii în cerere

Persoana responsabilă de implementare,

Șef Catedra de anesteziologie și reanimatologie nr.1 "V. Ghereg",
IP USMF „Nicolae Testemițanu”
dr. hab. șt. med., prof univ.

 Serghei Șandru,



REPUBLICA MOLDOVA
MINISTERUL SĂNĂTĂȚII



INSTITUȚIA MEDICO-SANITARĂ PUBLICĂ
INSTITUTUL DE MEDICINĂ URGENTĂ
(IMSP IMU)
MD-2004, mun. Chișinău, str. T. Ciorba, 1

CERTIFICAT DE INOVATOR

Nr. 19

data 13.12.2022

de înregistrarea inovației în Registrul obiectelor proprietății intelectuale ale IMSP IMU în
conformitate cu art. 16 al Legii nr. 138-XV din 10.05.2001

**IMPLEMENTAREA APLICAȚIEI SOFTWARE SYSTEM -
EXPERT ACIDO-BAZIC ABB 1.1.1 PENTRU
MANAGEMENTUL PACIENȚILOR ÎN STARE CRITICĂ**

(denumirea)

IAPĂSCURTĂ Victor

(coautori)



Director

dr. hab. șt. med., prof. univ. M. CIOCANU

mun. Chișinău



REPUBLICA MOLDOVA
MINISTERUL SĂNĂTĂȚII, MUNCII și PROTECȚIEI SOCIALE
INSTITUȚIA MEDICO-SANITARĂ PUBLICĂ
INSTITUTUL DE MEDICINĂ URGENTĂ
MD-2004, mun. Chișinău, str. T. Ciorba, 1
tel.: 022-23-78-84, fax: 022-23-53-09,
e-mail: anticamera@urgenta.md
www.urgenta.md



APROB

Director IMSP IMU

Dr. hab. șt. med., prof. univ.,

M. CIOCANU

ACT DE IMPLEMENTARE

- 1. Denumirea ofertei pentru implementare:** "IMPLEMENTAREA APLICAȚIEI SOFTWARE SYSTEM - EXPERT ACIDO-BAZIC ABB 1.1.1 PENTRU MANAGEMENTUL PACIENȚILOR ÎN STARE CRITICĂ „
- 2. Autori:** IAPĂSCURTĂ Victor, asistent universitar, doctorand;
- 3. Unde și când a fost implementată:** Anul implementării 2019-2022, Catedra de anesteziologie și reanimatologie "V. Ghereg", USMF "Nicolae Testemițanu" în cadrul IMSP IMU; Departamentul Anesteziologie și Trapie Intensivă IMSP IMU.
- 4. Rezultatul implementării:** Ca rezultat al implementării sistemului-expert de diagnostic al dereglărilor echilibrului acido-bazic se pot îmbunătăți rezultatele activității decizionale a medicului-specialist cu efect potențial pozitiv asupra rezultatelor tratamentului pacientului critic.
- 5. Eficacitatea implementării:** Diagnosticul rapid al dezechilibrilor acido-bazice cu ajutorul sistemului propus, în special la pacientul critic, oferă atât beneficii terapeutice pentru pacienți, cât și pentru medici – prin asistarea lor în procesul decizional în cazurile cu grad crescut de incertitudine diagnostică și restricții de timp.

Prezenta inovație este implementată conform descrierii în cerere

Persoana responsabilă de implementare,

Șef Catedra de anesteziologie și reanimatologie nr. 1 "V. Ghereg",
IP USMF „Nicolae Testemițanu”
dr. hab. șt. med., prof. univ. _____ Serghei Șandru

Declarația privind asumarea răspunderii

Subsemnatul, declar pe răspundere personală că materialele prezentate în teza de doctorat sunt rezultatul propriilor cercetări și realizări științifice. Conștientizez că, în caz contrar, urmează să suport consecințele în conformitate cu legislația în vigoare.

Iapăscurtă Victor

Data:

Informații personale

Nume / Prenume

Adresa

Telefoane

e-mail

Sex

Data nașterii

Naționalitatea

Titlul științifico-didactic

Experiența profesională

Perioada

Iapăscurtă Victor

Str. Tiras 2, or. Chișinău, 2004, Republica Moldova

Mobil: (+373) 69 10 57 13

victor.iapascurta@usmf.md

masculin

27 ianuarie 1966

Republica Moldova

Asistent universitar



septembrie 1994 - curent

Asistent universitar la catedra de Anesteziologie și reanimatologie nr.1 „V. Ghereg” a Universității de medicină și farmacie „N. Testemițanu” din Republica Moldova (bd. Ștefan cel Mare și Sfânt, 165, or. Chișinău, <https://anesteziologie1.usmf.md/ro>)

- Predarea cursului „Anesteziologie-reanimatologie” studentilor anului 5
- Predarea cursului „Medicină perioperatorie” studentilor anului 6 (Imba de studii engleza)

noiembrie 2018 – prezent

Conducătorul grupului de cercetare „AID & (patho)physiological and clinical end models” în cadrul grupului privat de cercetare „Algorithmic Information Dynamics”, Algorithmic Dynamics Lab, Karolinska Institute, Sweden/ Oxford Immune Algorithmics, UK / Algorithmic Nature Group, Paris, France

aprilie 2018 – august 2018

Participant și învingător la Complexity Explorer Spring Challenge 2018, Institutul Santa Fe, New Mexico, S.U.A.

- Elaborarea și implementarea modelelor decizionale în baza programării bazate pe agenți, analiza lor cu scrierea și susținerea lucrării științifice “Decision-making under uncertainty: simple strategy vs bounded rationality and randomness” - locul II
- Participarea la turneul de programare – locul I
<https://www.complexityexplorer.org/challenges/pages/solution-gallery>

decembrie 2006 – noiembrie 2010

Director la Parteneriatul internațional pentru dezvoltare, Managementul operational al programelor de cercetare și asistență în domeniul antreprenoriatului social

septembrie 1994 – decembrie 2009

Medic anesteziolog-reanimatolog la Spitalul clinic municipal de urgență din Chișinău, str. Toma Ciorba 1, or. Chișinău

septembrie 1994 – decembrie 2006

Coordonator de programe, director de programe, președinte la Fundația Medicală din Moldova, str. Toma Ciorba 1, or. Chișinău

Managementul operational al programelor de asistență tehnică în domeniul educației și medicinei în colaborare cu structuri organizaționale guvernamentale și private din S.U.A.

Educație și formare

Perioada

Calificarea/diploma obținută

Numele și tipul instituției de învățământ

septembrie 1987 – iunie 1993

Medic de profil general, diploma E nr. 003648 din 30.06.1993

Universitatea de stat de medicină „N. Testemițanu”, facultatea de medicină generală, or. Chișinău, Republica Moldova

<p>Perioada Calificarea/diploma obținută Numele și tipul instituției de învățământ</p>	<p>august 1993 – iunie 1994 Medic anesteziolog-reanimatolog, certificat nr. 223 din 23.06.1994 Universitatea de stat de medicină și farmacie „N. Testemițanu”, Spitalul clinic municipal de asistență medicală de urgență, or. Chișinău</p>																				
<p>Alte formări/instruiri 01.06-05.10.2018 04.05-05.09.2017 04.05-05.09.2017 01.04-27.08.2015 10.08-12.08.2020</p>	<p>✓ Algorithmic Information Dynamics: A Computational Approach to Causality and Living Systems From Networks to Cells ✓ Curs în domeniul dinamicii algoritmice, Complexity Explorer, Santa Fe Institute, New Mexico, USA</p> <p>✓ Introduction to Agent Based Modeling (2017) ✓ Curs în domeniul modelării și programării bazate pe agenți (NetLogo) Complexity Explorer, Santa Fe Institute, New Mexico, USA</p> <p>✓ Introduction to Dynamical Systems and Chaos (2017) ✓ Curs în domeniul sistemelor dinamice, Complexity Explorer, Santa Fe Institute, New Mexico, USA</p> <p>Sleep Neurobiology, Medicine and Society ✓ Curs în domeniul medicinei somnului de la University of Michigan, USA</p> <p>Participări la conferințe: ✓ AUTOMATA 2020, 10-12 august, Stokholm, Sweden Raport „Block Decomposition Method and Traditional Machine Learning for Epileptic Seizure Prediction”</p>																				
<p>Specialitatea de bază Domenii de cercetare Tema tezei de doctorat</p>	<p>Domeniul de activitate științifică 321.19. Anesteziologie și terapie intensivă; Aplicarea inteligenței artificiale/învățării automate în anesteziologie-terapie intensivă Prezicerea timpurie a sepsisului cu ajutorul unei aplicații proprii în baza învățării automate (inteligență artificială)</p>																				
<p>Participarea la proiecte 2019 2020</p>	<p>Aplicarea dinamicii algoritmice în analiza modelelor fiziologice Aplicarea dinamicii algoritmice în rezolvarea problemelor medicale clinice</p>																				
<p>Autor și coautor</p>	<p>Date statistice privind numărul total de publicații științifice și metodicodidactice ✓ 30 articole și teze</p>																				
<p>Limba maternă</p>	<p>Română</p>																				
<p>Autoevaluare Nivel european (*) Rusa Engleza</p>	<table border="1"> <thead> <tr> <th colspan="2">Înțelegere</th> <th colspan="2">Vorbire</th> <th>Scriere</th> </tr> <tr> <td>Ascultare</td> <td>Citire</td> <td>Participare la conversație</td> <td>Discurs oral</td> <td>Exprimare scrisă</td> </tr> </thead> <tbody> <tr> <td>C2</td> <td>C2</td> <td>C2</td> <td>C2</td> <td>C2</td> </tr> <tr> <td>B2</td> <td>B2</td> <td>B2</td> <td>B2</td> <td>B2</td> </tr> </tbody> </table> <p>(*) Nivelul Cadrului European Comun de Referință Pentru Limbi Străine</p>	Înțelegere		Vorbire		Scriere	Ascultare	Citire	Participare la conversație	Discurs oral	Exprimare scrisă	C2	C2	C2	C2	C2	B2	B2	B2	B2	B2
Înțelegere		Vorbire		Scriere																	
Ascultare	Citire	Participare la conversație	Discurs oral	Exprimare scrisă																	
C2	C2	C2	C2	C2																	
B2	B2	B2	B2	B2																	
<p>Competențe și aptitudini PC</p>	<p>Windows, MS Office applications: Word, Excel, Power Point; Internet, Rstudio, Octave, Wofram, H2O, NetLogo, Python, Sisteme Informaționale Automatizate</p>																				
<p>Competențe organizaționale/manageriale</p>	<p>✓ Capacitate de analiză și sinteză. ✓ Responsabilitate. ✓ Managementul grupurilor și structurilor organizaționale</p>																				
<p>Informații suplimentare</p>	<p>Văduv, 2 copii Permis de conducere - Categoria B</p>																				
<p>Persoane de referință:</p>	<ul style="list-style-type: none"> ▪ Balii Adrian, dr. hab, profesor universitar, adrian.balii@usmf.md, tel. mob. (+373) 79579474. ▪ Sandru Serghei dr. hab., profesor universitar, serghei.sandru@usmf.md, tel. mob. (+373) 79454065. 																				