



Universitatea din Timișoara  
Facultatea de matematică și informatică  
Școala Doctorală: Matematică și Informatică  
Domeniul: Informatică

## **TEZĂ DE DOCTORAT**

**Modele de inteligență artificială bazate pe  
arhitecturi multi-agent cu aplicații în  
sănătate**

**Conducător științific:**  
Prof. dr. Viorel Negru

**Doctorand:**  
Kristijan Cincar

**Timișoara 2024**

# Rezumat

Sistemele informatice medicale cuprind platforme complexe de gestionare a datelor de sănătate, aplicații de monitorizare a sănătății și tehnologii de analiză predictivă. Această teză este structurată în trei părți principale, fiecare abordând aspecte importante ale utilizării învățării automate și profunde în imagistica medicală și gestionarea resurselor spitalicești și folosirea arhitecturii multi-agent. În prima parte a tezei, se prezintă o analiză cuprinzătoare a abordărilor actuale în dezvoltarea modelelor de învățare automată și profundă aplicate în imagistica medicală, cu un accent specific pe detectarea polipilor din leziuni colonoscopice. Pentru această sarcină, au fost selectate nouă modele de învățare automată, incluzând Support Vector Machine (SVM), Random Forest (RF), Random Subspace (RS), Decision Tree (DT), Gradient Boosting (GB), AdaBoost (AB) și XGBoost (XGB), precum și Gaussian Naive Bayes (GNB) și Bernoulli Naive Bayes (BNB). Alegerea acestor modele se bazează pe adecvarea lor pentru sarcini de clasificare și pe performanța demonstrată în literatura de specialitate, în diverse aplicații de imagistică medicală. Pentru a evalua performanța acestor modele, s-au utilizat șase seturi de date distincte de imagini medicale. Rezultatele experimentale indică faptul că algoritmul bazat pe Random Subspace (RS) a obținut performanțe mai bune decât ceilalți algoritmi utilizați. În cadrul studiului de învățare profundă, s-au implementat și evaluat trei modele preantrenate de referință: InceptionV3, Convolutional Neural Network (CNN) și Xception. Aceste modele au fost adaptate pentru sarcina specifică de clasificare a imaginilor medicale, iar performanța lor a fost evaluată pe cinci seturi distincte de date medicale. Modelul bazat pe CNN cu Keras a demonstrat performanțe mai bune comparativ cu modelele similare din literatura de specialitate, folosind seturi de date asemănătoare.

În partea a doua a tezei, se propune un sistem inteligent pentru modelarea diverselor secții medicale ale spitalului, care integrează componente inteligente pentru a optimiza rezultatele și nivelul de satisfacție al pacienților. Acest sistem conține o bază de cunoștințe medicale utilizate pentru a descrie în detaliu o gamă largă de boli și afecțiuni, oferind o bază solidă pentru deciziile clinice. De asemenea se extrage regulile de asociere utilizând algoritmul Apriori. Această metodă analizează datele obținute din evaluările de triaj și fișele medicale electronice ale pacienților din secțiile de urgență. Obiectivul principal este identificarea corelațiilor între bolile pacienților și istoricul lor medical, facilitând astfel îmbunătățirea procesului de diagnostic

și tratament.

În cele din urmă, în partea a treia, se prezintă platforma MAS-PatientCare (Sistem Integrat pentru Programarea în spital, pentru colonoscopie și recomandări medicale), o soluție avansată pentru gestionarea procedurilor medicale și îngrijirii pacienților în spitale. Arhitectura multi-agent MAS-PatientCare este proiectată pentru a facilita colaborarea și interacțiunea între diversele componente și agenți implicați în gestionarea procedurilor medicale. Acești agenți sunt responsabili pentru programarea pacienților, gestionarea resurselor, analiza datelor și furnizarea de recomandări medicale. Prin interconectivitatea și comunicarea eficientă între agenți, MAS-PatientCare îmbunătățește calitatea îngrijirii pacienților și eficiența serviciilor prin eficientizarea fluxurilor de lucru și a proceselor spitalicești. Un aspect esențial al sistemului MAS-PatientCare este capacitatea sa de a oferi recomandări personalizate pentru îngrijirea pacienților, bazate pe analiza datelor medicale și preferințele individuale ale acestora. Aceste recomandări sunt generate de agenți specializați în analiza de date care utilizează algoritmi de învățare automată pentru a identifica modele și tendințe relevante în datele pacienților. Prin integrarea acestor tehnologii și metodologii, teza de față oferă o abordare scalabilă și eficientă pentru îmbunătățirea diagnosticării și tratamentului medical, optimizarea resurselor spitalicești și creșterea nivelului de mulțumire al pacienților.

# Abstract

Health information systems comprise complex health data management platforms, health monitoring applications and predictive analytics technologies. This thesis is divided into three main parts, each addressing key aspects of the use of machine and deep learning in medical imaging and hospital resource management, and the use of multi-agent architectures. The first part of the thesis presents a comprehensive review of current approaches to the development of machine and deep learning models applied to medical imaging, with a specific focus on polyp detection in colonoscopic lesions. Eight machine learning models were selected for this task, including Support Vector Machine (SVM), Random Forest (RF), Random Subspace (RS), Decision Tree (DT), Gradient Boosting (GB), AdaBoost (AB) and XGBoost (XGB), as well as Gaussian Naive Bayes (GNB) and Bernoulli Naive Bayes (BNB). The choice of these models is based on their suitability for classification tasks and their performance as reported in the literature on various medical imaging applications. Six different medical image datasets were used to evaluate the performance of these models. Experimental results indicate that the Random Subspace (RS) based algorithm outperformed the other algorithms used. In the deep learning study, we implemented and evaluated three pre-trained benchmark models: InceptionV3, Convolutional Neural Network (CNN) and Xception. These models were adapted to the specific task of medical image classification and their performance was evaluated on five different medical datasets. The CNN-based model with Keras showed superior performance compared to similar models in the literature using similar datasets.

The second part of the thesis proposes an intelligent system for modelling different hospital wards, integrating intelligent components to optimise patient outcomes and satisfaction. This system includes a medical knowledge base, which is used to describe in detail a wide range of diseases and conditions, providing a sound basis for clinical decisions. It also extracts association rules using the Apriori algorithm. This method analyses data from triage assessments and electronic medical records of patients in emergency departments. The main objective is to identify correlations between patients' diseases and their medical history, thus facilitating the improvement of the diagnosis and treatment process.

Finally, the third part presents the MAS-PatientCare (Integrated System for Hospital Scheduling, Colonoscopy and Medical Referrals) platform, an advanced solution for managing me-

dical procedures and patient care in hospitals. The multi-agent architecture MAS-PatientCare is designed to facilitate collaboration and interaction between the various components and agents involved in the management of medical procedures. These agents are responsible for scheduling patients, managing resources, analysing data and making medical recommendations. Through interconnectivity and effective communication between agents, MAS-PatientCare improves the quality of patient care and the efficiency of services by streamlining workflows and hospital processes. A key aspect of the MAS-PatientCare system is its ability to provide personalised recommendations for patient care based on the analysis of medical data and individual patient preferences. These recommendations are generated by specialised data analysis agents that use machine learning algorithms to identify relevant patterns and trends in patient data.

By integrating these technologies and methodologies, this thesis provides a comprehensive and effective approach to improving medical diagnosis and treatment, optimizing hospital resources, and increasing patient satisfaction.

# Cuprins

<b>Listă de figuri</b>	<b>9</b>
<b>Listă de tabele</b>	<b>11</b>
<b>1 Introducere</b>	<b>12</b>
1.1 Context . . . . .	12
1.2 Scopul și obiectivele . . . . .	13
1.2.1 Scopul cercetării . . . . .	13
1.2.2 Obiectivele cercetării . . . . .	14
1.3 Structura lucrării . . . . .	15
1.4 Rezultate . . . . .	18
1.5 Lista lucrărilor științifice . . . . .	21
1.6 Mulțumiri . . . . .	22
<b>2 Sisteme inteligente în sănătate</b>	<b>24</b>
2.1 Context . . . . .	24
2.2 Sistemele cu elemente de inteligență artificială utilizate în domeniul medical . .	25
2.2.1 Spitale Inteligente . . . . .	26
2.2.2 Sisteme bazate pe reguli în domeniul medical . . . . .	27
2.3 Prelucrarea imaginilor medicale . . . . .	29
2.4 Învățarea automată în domeniul medical . . . . .	31
2.5 Învățarea profundă în domeniul medical . . . . .	34
2.6 Modele bazate pe agenți pentru aplicații în domeniul medical . . . . .	38
2.7 Modele de management a spitalelor . . . . .	41
2.8 Seturi de date . . . . .	43
2.9 Sumar . . . . .	45
<b>3 Contribuții proprii în aplicarea modelelor de învățare automată în imagistica medicală</b>	<b>47</b>
3.1 Context . . . . .	47

3.2	Enumerarea etapelor pentru identificarea și clasificarea polipilor . . . . .	47
3.2.1	Preprocesarea imaginilor . . . . .	48
3.2.2	Extragerea caracteristicilor . . . . .	48
3.2.3	Analiza importanței caracteristicilor . . . . .	50
3.2.4	Selectarea metricilor specifice pentru clasificarea multiplă . . . . .	51
3.2.5	Configurarea și antrenarea modelului de clasificare . . . . .	53
3.3	Detectarea polipilor cu modele SVM . . . . .	54
3.3.1	Modele SVM . . . . .	55
3.3.2	Analiza rezultatelor modelelor SVM . . . . .	57
3.4	Detectarea polipilor cu alte modele de învățare automată . . . . .	63
3.4.1	Alte modele de învățare automată . . . . .	63
3.4.2	Analiza rezultatelor modelelor de învățare automată . . . . .	65
3.5	Detectarea polipilor cu modele de învățare profundă . . . . .	72
3.5.1	Modele de învățare profundă . . . . .	72
3.5.2	Analiza rezultatelor modelelor de învățare profundă . . . . .	73
3.6	Rezultate și discuții asupra performanței algoritmilor de învățare automată și profundă . . . . .	82
3.7	Compararea modelelor cu literatura de specialitate . . . . .	87
3.8	Sumar . . . . .	88
<b>4</b>	<b>Sistem inteligent pentru programarea pacienților</b>	<b>90</b>
4.1	Context . . . . .	90
4.2	Sistemul de programare a pacienților . . . . .	91
4.2.1	Descrierea modelului . . . . .	91
4.2.2	Ontologia propusă . . . . .	92
4.2.3	Simularea și experimentele . . . . .	94
4.3	Modele bazate pe reguli de asociere pentru realizarea programărilor în domeniu medical . . . . .	98
4.3.1	Descrierea modelului . . . . .	98
4.3.2	Extragerea regurilor de asociere . . . . .	99
4.3.3	Evaluarea modelului bazat pe reguli extrase . . . . .	104
4.4	Sumar . . . . .	108
<b>5</b>	<b>Sistemul integrat bazat pe arhitectura multi-agent propusă</b>	<b>110</b>
5.1	Context . . . . .	110
5.2	Arhitectura multi-agent . . . . .	111
5.2.1	Prezentare generală . . . . .	111
5.2.2	Tipuri de agenți . . . . .	111

5.3	Interacțiunea agenților și funcționarea sistemului . . . . .	113
5.3.1	Interconectivitatea agenților pentru colonoscopie . . . . .	115
5.3.2	Interconectivitatea agenților pentru programarea pacienților și gestiona- rea resurselor . . . . .	117
5.3.3	Interconectivitatea agenților pentru primire urgență . . . . .	119
5.4	Evaluarea sistemului multi-agent . . . . .	120
5.4.1	Studiu de caz pentru colonoscopie . . . . .	121
5.4.2	Studiu de caz pentru programarea pacienților și definirea numărului op- tim de medici . . . . .	123
5.4.3	Studiu de caz pentru primire urgență . . . . .	125
5.5	Sumar . . . . .	129
<b>6</b>	<b>Concluzii și direcții viitoare</b>	<b>131</b>
	<b>Bibliografie</b>	<b>135</b>



# Listă de figuri

2.1	Exemple de detectare a polipilor . . . . .	30
2.2	Diagrama schematică a modelului Inception-v3, prima parte (a), a două parte (b)	37
2.3	O prezentare de ansamblu asupra sistemelor multi-agent (JADE). . . . .	39
3.1	Importanța caracteristicilor. (a) DecisionTreeClassifier; (b) RandomForestClassifier; (c) XGBClassifier . . . . .	50
3.2	Exemple de polipi gastrointestinali. WL (stânga); NBI (dreapta); (a) adenom (b) hiperplazic și (c) adenom zimțat . . . . .	53
3.3	Modelul SVM - Scorul de acuratețe pentru 50 de rulări . . . . .	56
3.4	Matrice de confuzie pentru toate cele șase seturi de date: (a) performanța SVM-Linear pe [WL-WL]; (b) performanța SVM-Linear pe [NBI-NBI]; (c) performanța SVM-Linear pe [WL-NBI]; (d) performanța SVM-Linear pe [NBI-WL]; (e) performanța SVM-Linear pe [FULL]; (f) performanța SVM-Linear pe [CONCAT] . . . . .	62
3.5	Random Forest - Scorul de acuratețe pentru 50 de rulări: (a) 350 estimatori-entropy (b) 350 estimatori-gini. . . . .	63
3.6	Random Subspace - Scorul de acuratețe pentru 50 de rulări: (a) cu 10 caracteristici (b) cu 5 caracteristici . . . . .	64
3.7	Matrice de confuzie pentru toate cele șase seturi de date: (a) performanța GB100 pe [WL-WL]; (b) performanța XGB200 pe [NBI-NBI]; (c) performanța XGB100 pe [WL-NBI]; (d) performanța RS5 pe [NBI-WL]; (e) performanța RS5 pe [FULL]; (f) performanța XGB100 pe [CONCAT]. . . . .	71
3.8	Matrice de confuzie pentru toate cele cinci seturi de date: (a) performanța InceptionV3 pentru [WL-WL]; (b) performanța InceptionV3 pentru [NBI-NBI]; (c) performanța InceptionV3 pentru [WL-NBI]; (d) performanța InceptionV3 pentru [NBI-WL]; (e) performanța InceptionV3 pentru [FULL]. . . . .	76

3.9	Matrice de confuzie pentru toate cele cinci seturi de date: (a) performanța CNN pentru [WL-WL]; (b) performanța CNN pentru [NBI-NBI]; (c) performanța CNN pentru [WL-NBI]; (d) performanța CNN pentru [NBI-WL]; (e) performanța CNN pentru [FULL]. . . . .	79
3.10	Matrice de confuzie pentru toate cele cinci seturi de date: (a) performanța Xception pentru [WL-WL]; (b) performanța Xception pentru [NBI-NBI]; (c) performanța Xception pentru [WL-NBI]; (d) performanța Xception pentru [NBI-WL]; (e) performanța Xception pentru [FULL]. . . . .	83
4.1	Baza de cunoștințe a sistemului . . . . .	94
4.2	Model de sistem inteligent de programare a spitalului . . . . .	96
4.3	Cele mai frecvente 25 de afecțiuni din setul de date . . . . .	100
4.4	Cele mai frecvente 25 boli/afecțiuni anterioare din setul de date . . . . .	101
4.5	Cele mai frecvente 25 categorii de medicamente din setul de date . . . . .	102
4.6	Niveluri de <i>support</i> și <i>confidence</i> . . . . .	105
4.7	Distribuția regulilor de asociere pentru toate cele trei subseturi: (a) reguli extrase din subsetul CC; (b) reguli extrase din subsetul CC-PMH; (c) extrase reguli din subsetul CC-MEDS. . . . .	106
5.1	Arhitectura multi-agent MAS-PatientCare . . . . .	114
5.2	Componenta MAS-PatientCare de monitorizare colonoscopie . . . . .	116
5.3	Subsistemul de management al spitalului . . . . .	117
5.4	Componenta MAS-PatientCare pentru primire urgență . . . . .	120
5.5	Fluxul pacientului pentru colonoscopie . . . . .	122
5.6	JSON pentru configurația costurilor minime . . . . .	124
5.8	Costului total al sistemului . . . . .	125
5.7	Distribuția secțiunii transversale pentru toate cele trei subseturi: (a) distribuția doctor psihiatru; (b) distribuția doctor endocrinolog; (c) distribuția doctor internist	126
5.9	Planificarea pacienților în unitatea de urgență . . . . .	127
5.10	Fluxul pacienților în cadrul secției de urgență . . . . .	128

# Listă de tabele

2.1	Modelele de învățare profundă Keras . . . . .	35
2.2	Caracteristicile seturilor de date. . . . .	45
3.1	Caracteristicile imagini . . . . .	49
3.2	Modele SVM de învățare automată . . . . .	55
3.3	Analiza rezultatelor modelelor de învățare automată SVM pentru primele 3 seturi de date . . . . .	58
3.4	Analiza rezultatelor modelelor de învățare automată SVM pentru ultimele 3 seturi de date . . . . .	59
3.5	Modele de învățare automată . . . . .	66
3.6	Analiza rezultatelor modelelor de învățare automată pentru primele 3 seturi de date . . . . .	67
3.7	Analiza rezultatelor modelelor de învățare automată pentru ultimele 3 seturi de date . . . . .	68
3.8	InceptionV3 cu cinci seturi de date . . . . .	74
3.9	CNN cu cinci seturi de date . . . . .	78
3.10	Xception cu cinci seturi de date . . . . .	82
3.11	Matrice de confuzie comparativă pentru cele mai bune 10 modelele de învățare automata și învățare profundă: (a) performanța SVM-Linear pe [FULL]; (b) performanța GB100 pe [WL-WL]; (c) performanța XGB200 pe [NBI-NBI]; (d) performanța XGB100 pe [WL-NBI]; (e) performanța RS5 pe [NBI-WL]; (f) performanța RS5 pe [FULL]; (g) performanța XGB100 pe [CONCAT]; (h) InceptionV3 pentru [FULL]; (i) CNN pentru [FULL]; (j)Xception pentru [NBI-WL]. . . . .	85
4.1	Pachetele Python utilizate . . . . .	95
4.2	Setul de afecțiuni . . . . .	99
4.3	Categoriile din setul de date . . . . .	99
4.4	Regulile extrase . . . . .	107