



Ciclul „INTELIGENȚA ARTIFICIALĂ ÎN MEDICINĂ”

Oct. – Dec. 2025

Coordonator ciclu: Prof. dr. Daniel LIGHEZAN

Coordonator științific: Prof. dr. Gheorghe Ioan MIHALAȘ

Tema 2

„Machine Learning”

Prof. Dr. Gheorghe I MIHALAS

Universitatea de Medicină și Farmacie Victor Babeș Timișoara

Centrul de Modelare a Sistemelor Biologice și Analiza Datelor

Laboratorul de Inteligență Artificială

și

Academia de Științe Medicale

Comisia de Informatică Medicală, Inteligență Artificială și Protecția Datelor

„A învăța înseamnă a extrage ordine din haos”

Aforism modern inspirat de gândirea clasică

Planul tematic al ciclului de A.I. - 2025

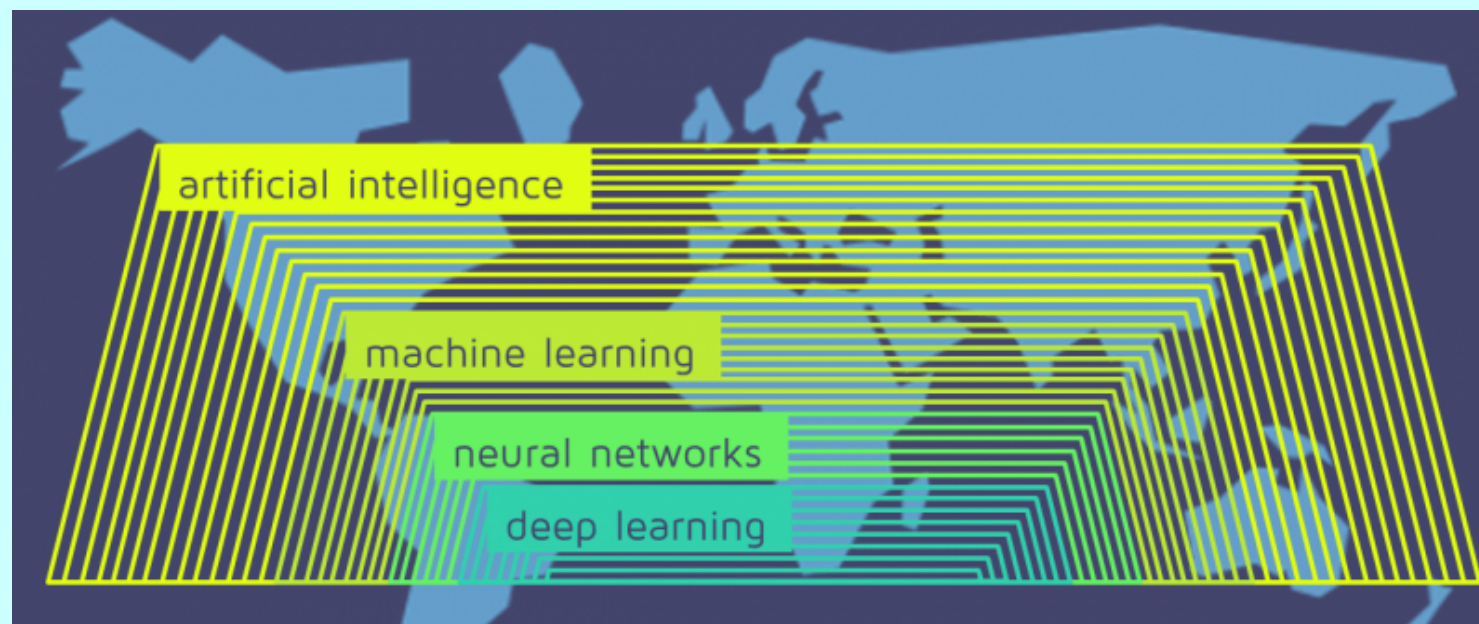


Data	Tema prelegerii
7 oct	Introducere în AI. Definiții, istoric. Abordări cognitive și conexiuniste. Sisteme discriminative și generative. Agenți AI. Aplicații medicale – privire generală, clasificări. Integrarea AI în activitatea medicală.
14 oct	„Machine learning”, învățare supervizată și nesupervizată. Analiza datelor medicale: predicții, clasificări. Date sintetice, simulări. Modelarea proceselor, proiectarea studiilor clinice, evaluare.
28 oct	AI pentru decizia medicală. Reprezentarea cunoștințelor, ontologii medicale, codificări. Sisteme logice, sisteme baysiene. Logica fuzzy. Sisteme expert medicale, decizie asistată. Monitorizarea pacienților.
4 nov	Rețele neuronale: principii, clasificări. Perceptronul, „deep learning”. Aplicații în imagistica medicală, bioinformatică, medicina de precizie. Performanță vs. transparență, criterii de validare.
11 nov	Sisteme generative conversaționale. Prelucrarea limbajului natural, „LLM – large language models”. Construcția și funcționarea unui Chat-bot. Aplicații medicale, chatboți specializați. Prompt engineering. Limite și capcane.
18 nov	AI în învățământul medical și în cercetare. Generare întrebări, cazuri clinice interactive. Utilizarea LLM în redactarea științifică. Construcția și antrenarea de chatboți specializați. Etica utilizării AI în educație și cercetare.
25 nov	Limitele actuale, aspecte etice și juridice. Limite: confidențialitate, bias. Transparență, sisteme explicabile XAI. Tehnologia empatică, riscuri și abuzuri. Competențe AI, responsabilități. Reglementări.

Planul prelegerii

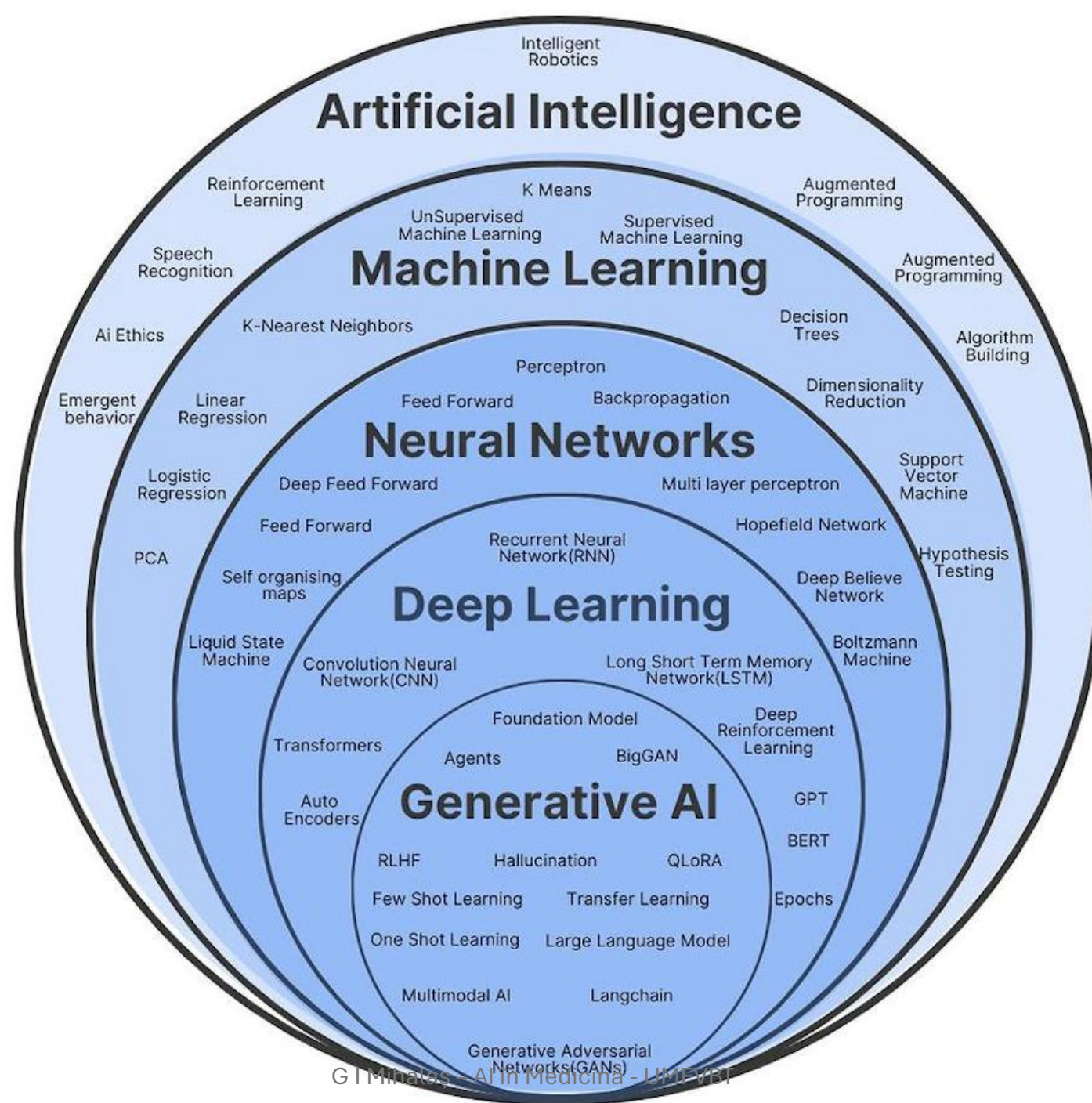
1. **Încadrarea „machine learning” în AI**
 - Structura metodologică a AI
2. **Analiza datelor medicale**
3. **Machine Learning**
 - Tipuri: Învățarea supervizată / Învățarea nesupervizată / „Reinforcement learning”
4. **Aplicații medicale**
5. **Limitări și precauții**
6. **Concluzii și recomandări practice**
 - Alegerea metodei
 - Colaborarea interdisciplinară

Încadrarea „machine learning” în AI

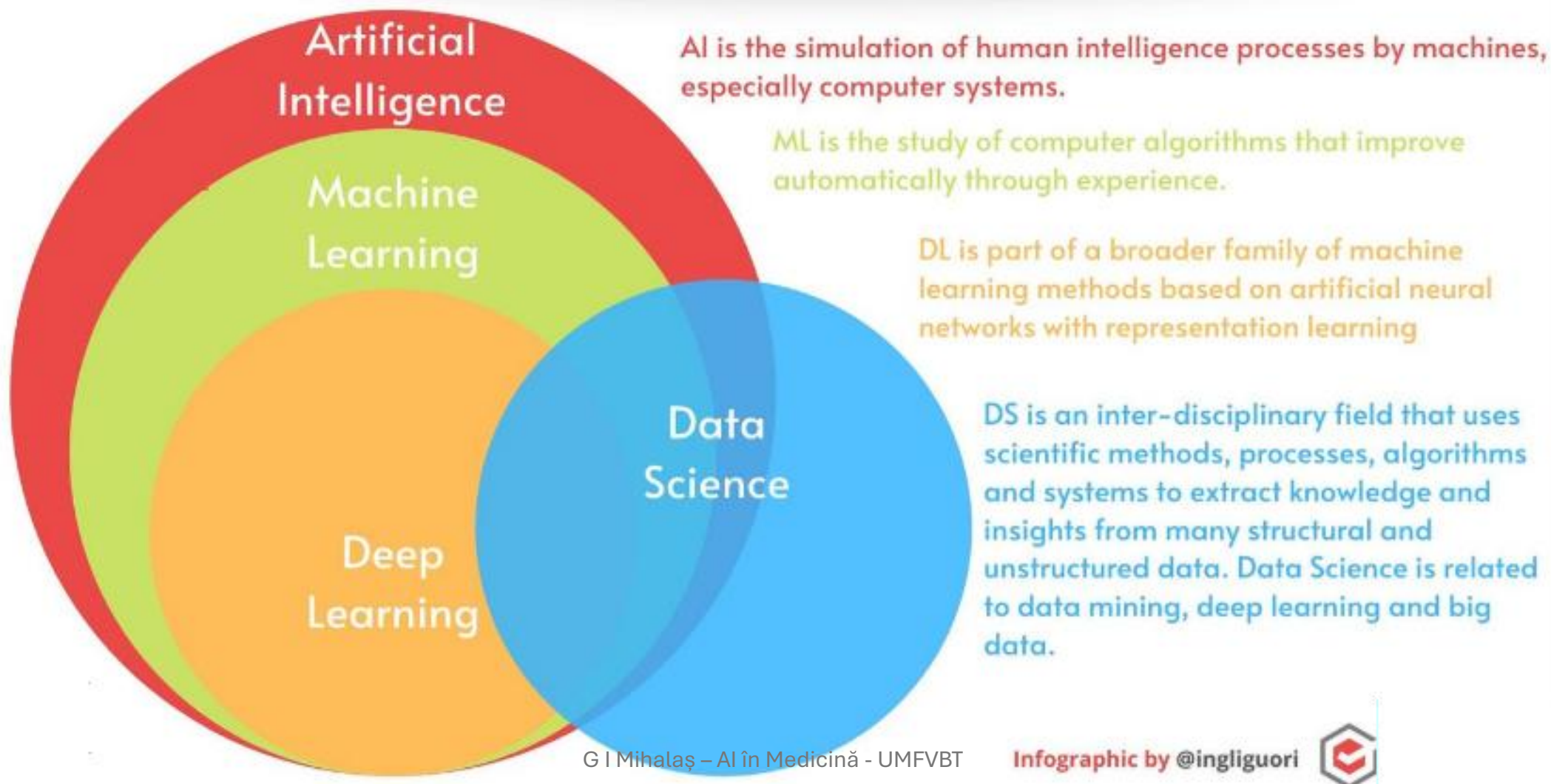


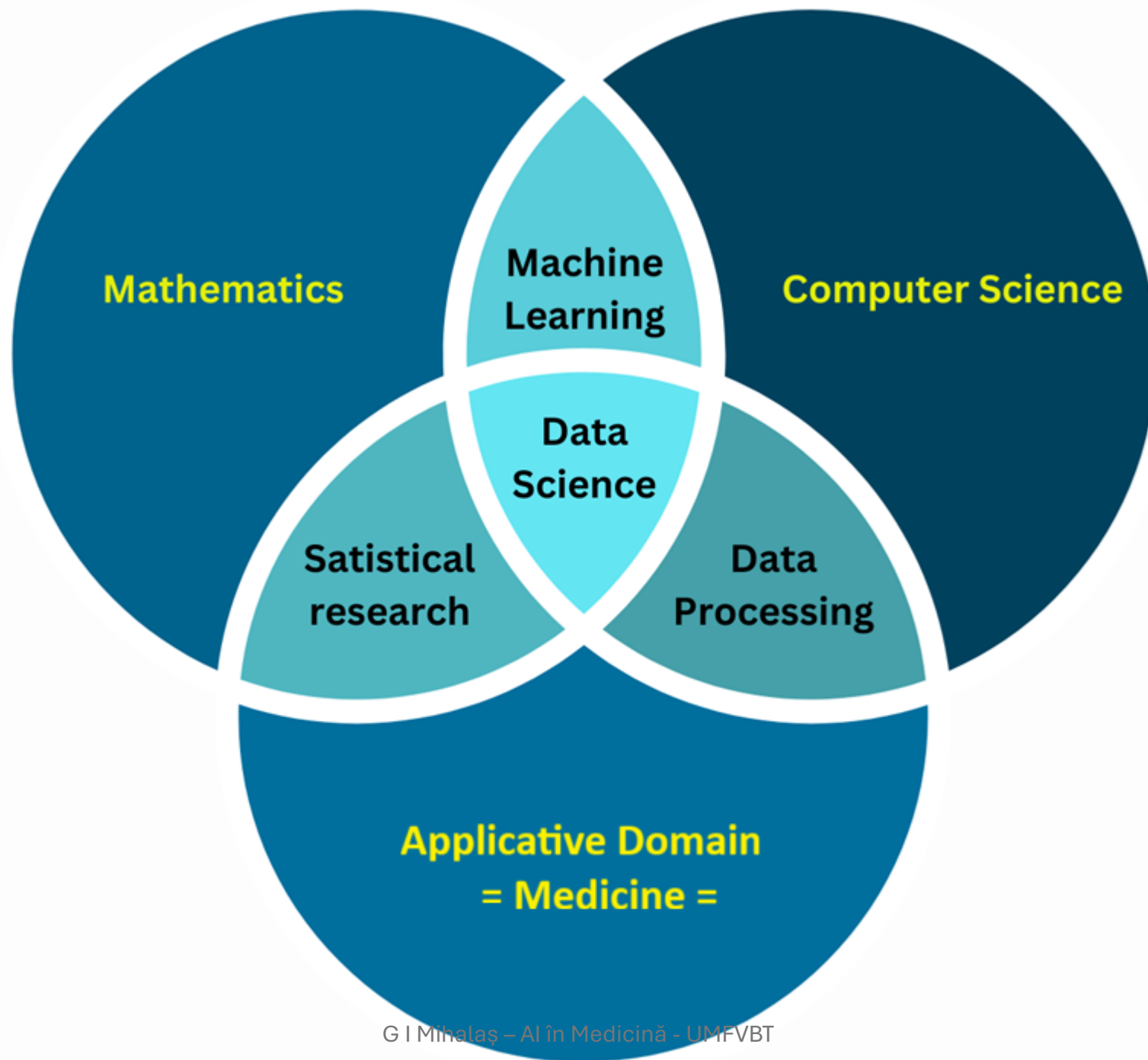
AI – viziune de ansamblu

- AI = totalitatea metodelor care permit rezolvarea inteligentă a problemelor
- Paradigme principale:
 - **Abordarea simbolică** (reguli, formule, logică, ontologii)
 - **Abordarea bazată pe învățare (ML)**
- Legătura dintre AI – ML – NN - DL



RELATIONSHIP BETWEEN ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, DEEP LEARNING AND DATA SCIENCE





Analiza datelor medicale



Analiza datelor în medicină

- ***Tipuri de date medicale:***

- *După sursă:* clinice, biologice, imagistice, administrative etc.
- *După reprezentare:*
 - Structurate: numerice, semnale, genetice (secvențe), calitative (clase)
 - Nestructurate: imagini, text etc.

- ***Etapele analizei:***

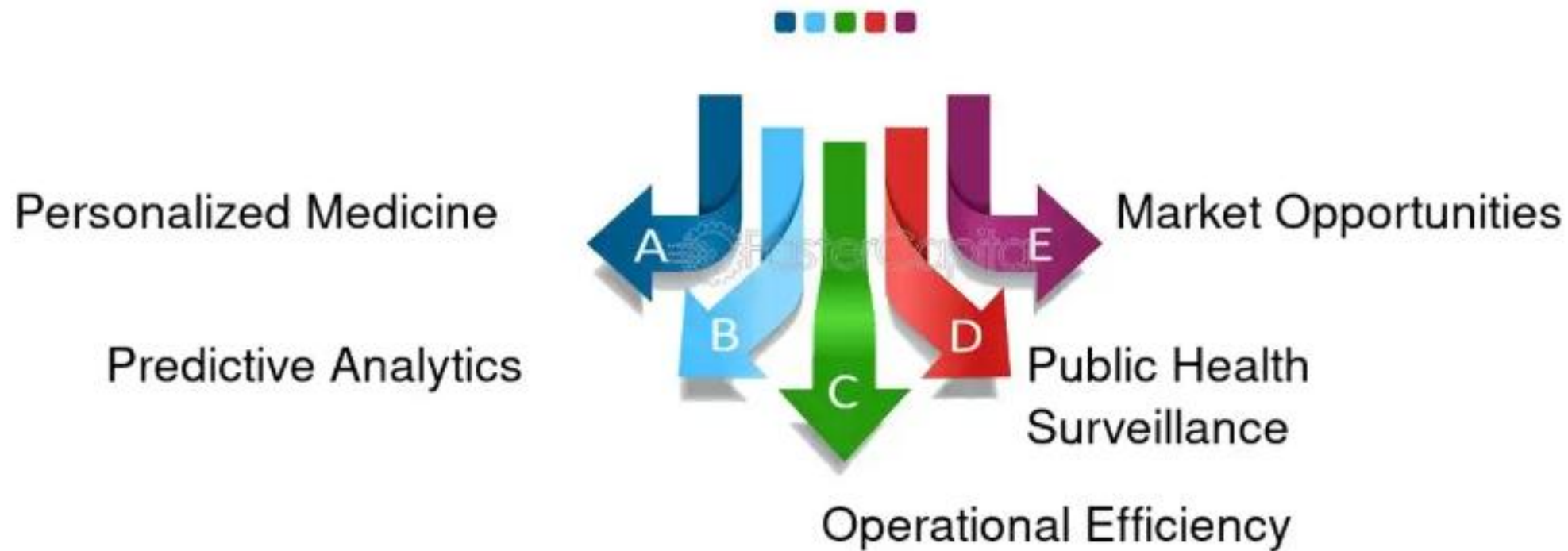
- Colectare, curățare, selecție
- Organizare, stocare
- **Clasificare (Interpretare / decizie – diagnostic, tratament)**
- **Predicție (Analiza riscului)**
- Reprezentare, vizualizare, sonificare, modelare

- ***Exemple clasice de analize statistice***

- Teste statistice, corelații, serii temporale

- ***Limitările statisticii clasice***

The Evolution of Data in Healthcare



Analiza riscului - scoruri



		IMPACT/CONSEQUENCE LEVELS				
		SLIGHT / NEGLIGIBLE	MINOR	MODERATE	MAJOR	CATASTROPHIC
LIKELIHOOD DESCRIPTORS		Injuries requiring no treatment or first aid	Minor injury, first aid only required	Injury requiring medical treatment and some lost time	Serious injury, hospital treatment required	Death or permanent disability
RARE / REMOTE	May happen only in exceptional circumstances	VERY LOW	VERY LOW	LOW	LOW	MODERATE
UNLIKELY	Could happen some time	VERY LOW	LOW	MODERATE	MODERATE	MODERATE
POSSIBLE / OCCASIONALLY	Might occur occasionally	LOW	MODERATE	MODERATE	MODERATE	HIGH
LIKELY	Will probably occur in most circumstances	LOW	MODERATE	MODERATE	HIGH	VERY HIGH
ALMOST CERTAIN	Expected to occur in most circumstances	MODERATE	MODERATE	HIGH	VERY HIGH	VERY HIGH

RISK EXAMPLE	LIKELIHOOD	IMPACT	RISK GRADING
PATIENT INJURY	UNLIKELY	MAJOR	MODERATE

RISK GRADING COLORS				
VERY LOW RISK	LOW RISK	MODERATE RISK	HIGH RISK	VERY HIGH RISK

Analiza riscului

- ***Tranziția medicină curativă – medicină preventivă***
- ***Studii epidemiologice***
 - Metode (*studii transversale, studii pe cohorte, case-control*)
 - Indicatori – *OR, RR, Modelare, Scoruri de risc*
- ***Metode analitice***
 - **Metode statistice**
 - *Regresia liniară – predicții, Regresia logistică – clasificări*
 - *Metode statistice avansate: SVM, Arbori de decizie, Random Forest, K-means*
 - **Metode bazate pe rețele neuronale**
 - (*FFN, CNN, RNN, Transformers*)
- ***Evaluarea statistică***
 - Indicatori (*SN, SP VPP, VPN, AC*)
 - Curbe ROC, AUC

TIPURI DE MACHINE LEARNING

- Învățare supervizată
- Învățare nesupervizată
- Învățarea prin întărire (*Reinforcement learning*)



Învățarea supervizată

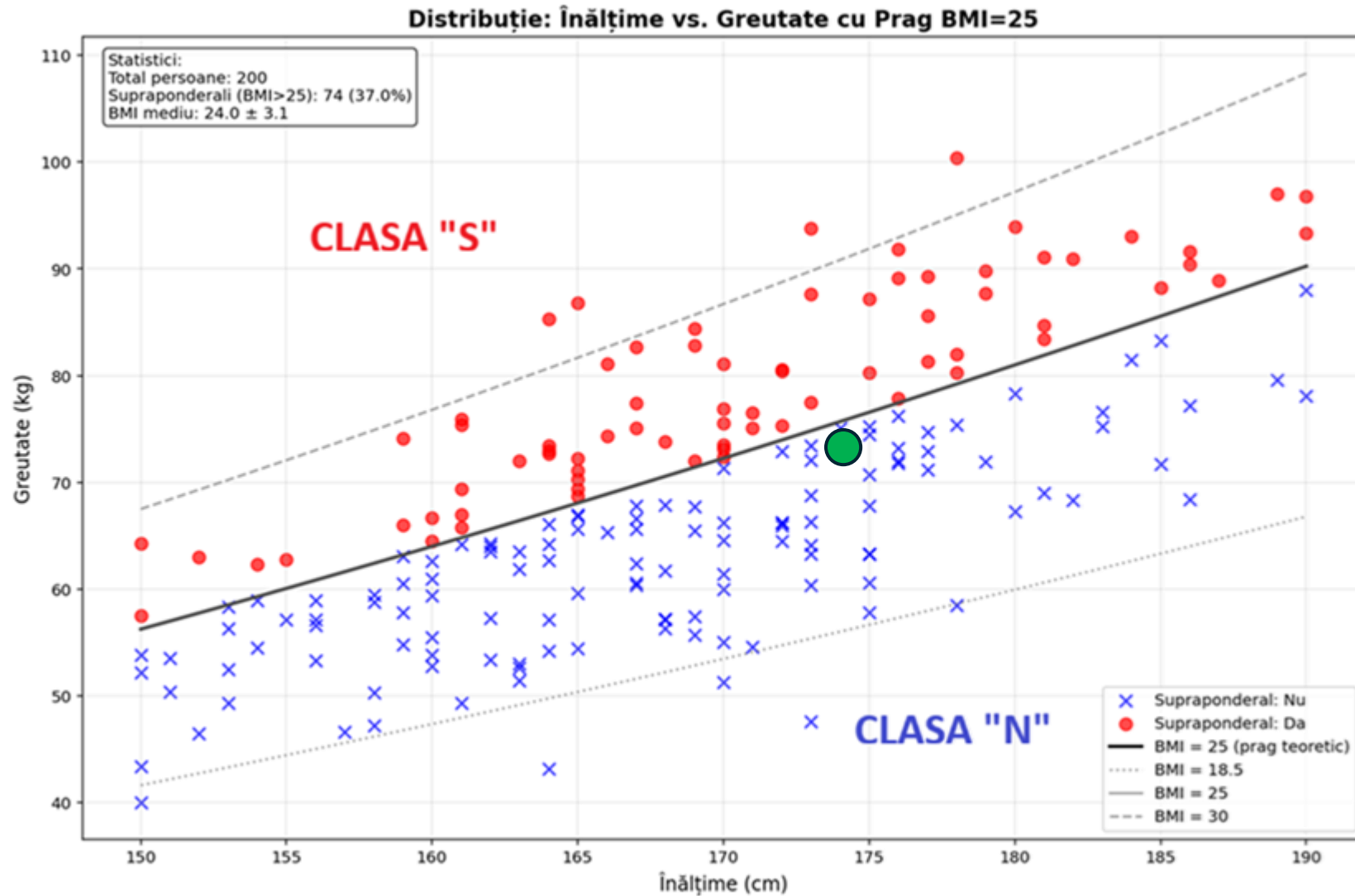
- **Principiu**

- Exemplele aparțin unor clase (disjuncte)
- Fiecare exemplu este definit prin:
 - (a) un set de parametri caracteristici
 - (b) **eticheta** clasei
- Modelul „învață *frontiera de separare*” între clase

- **Algoritmi**

- Regresia logistică
- SVM (Support Vector Machine)
- Arbori de decizie, Random Forest
- Rețele neuronale (FFN – Feed Forward Network)

Exemplu



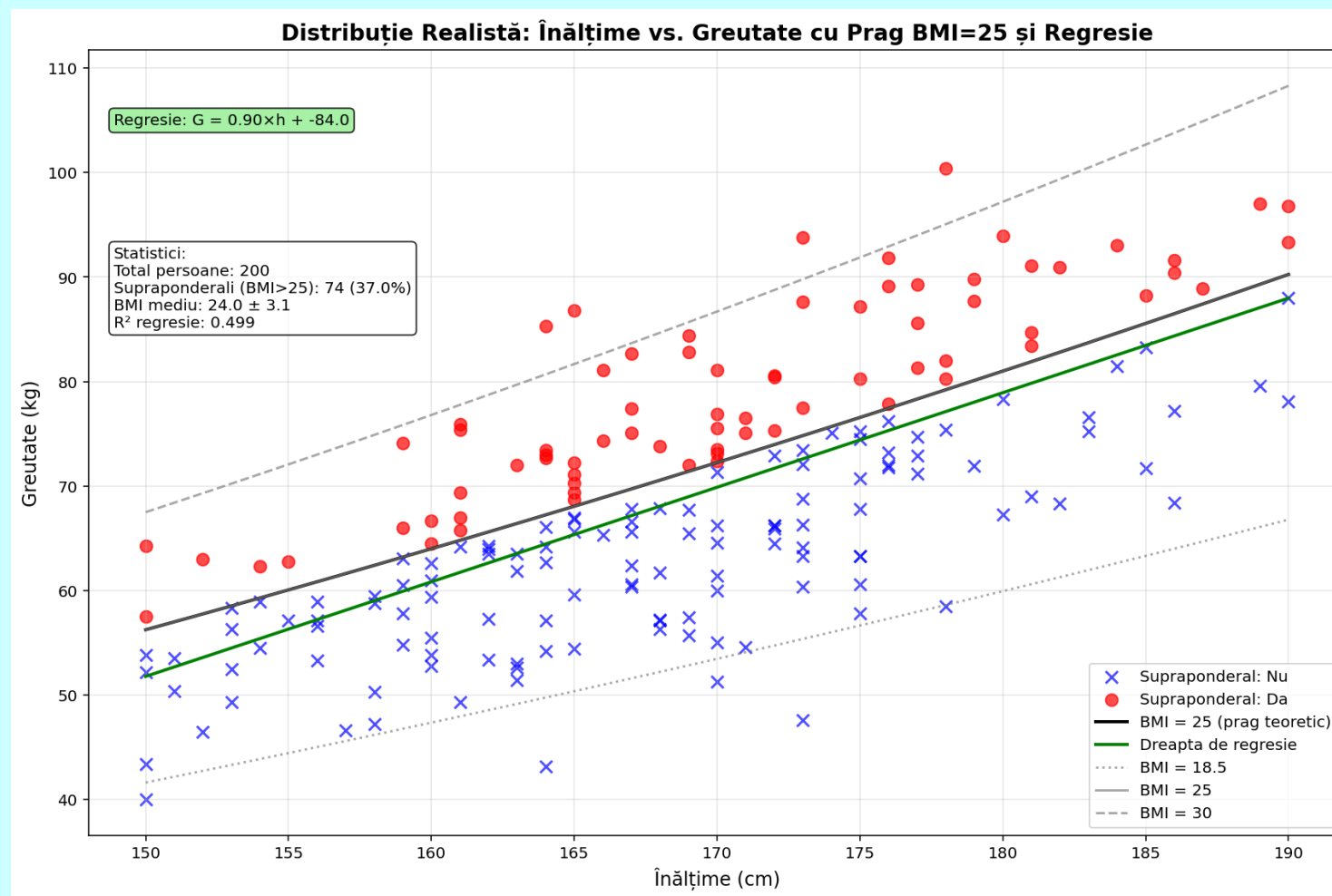
Ce clasă este
un nou caz ?
 $h = 174 \text{ cm}$
 $G = 72 \text{ kg}$

Este în
CLASA N

Verificare:
 $\text{BMI} = 23.79$
 $\text{BMI} < 25$

Exemplu

Linia de separare a claselor $\neq$ Dreapta de regresie



Aplicații medicale pentru învățarea supervizată

- **Clasificări** (*probabilitatea apartenenței la o clasă*)
 - Regresia logistică
 - Metode statistice avansate: SVM, Random Forest
 - Rețele neuronale
 - Ex: clasificare tumori din imagini CT, recunoașterea depresiei din text, etc
- **Predicții** (*estimează o valoare numerică*)
 - Regresia liniară
 - Metode statistice avansate: regresia Ridge, regresia Cox
 - Rețele neuronale
 - Ex: scoruri de risc (de infarct în 5 ani), prognoza creatininei în insuficiența renală

Journal of Thoracic Disease

► J Thorac Dis. 2019 Mar;11(3):950–958. doi: [10.21037/jtd.2019.01.90](https://doi.org/10.21037/jtd.2019.01.90) [↗](#)

Development and validation of a predictive model for the diagnosis of solid solitary pulmonary nodules using data mining methods

[Yangwei Xiang](#)¹, [Yifeng Sun](#)¹, [Yuan Liu](#)², [Baohui Han](#)³, [Qunhui Chen](#)⁴, [Xiaodan Ye](#)⁴, [Li Zhu](#)⁴, [Wen Gao](#)^{1,5},

Methods

A training cohort of 388 consecutive patients with solid SPNs was used to develop a predictive model to evaluate the malignancy of solid SPNs. By using SPSS Modeler, we utilized logistic regression (LR), artificial neural network (ANN), k-nearest neighbor (KNN), random forest (RF), and support vector machines (SVM) classifiers to build predictive models. Another cohort of 200 consecutive patients with solid SPNs was used to verify the accuracy of the predictive model. Predictive performance was evaluated using the area under the receiver operating characteristic curve (AUC).

Prediction of Lung Cancer Using Machine Learning

Avinash S¹, Anu Kiruthiga R², Aditya Saran K P³, Harish A⁴, Dr. D. Rasi⁵

Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India



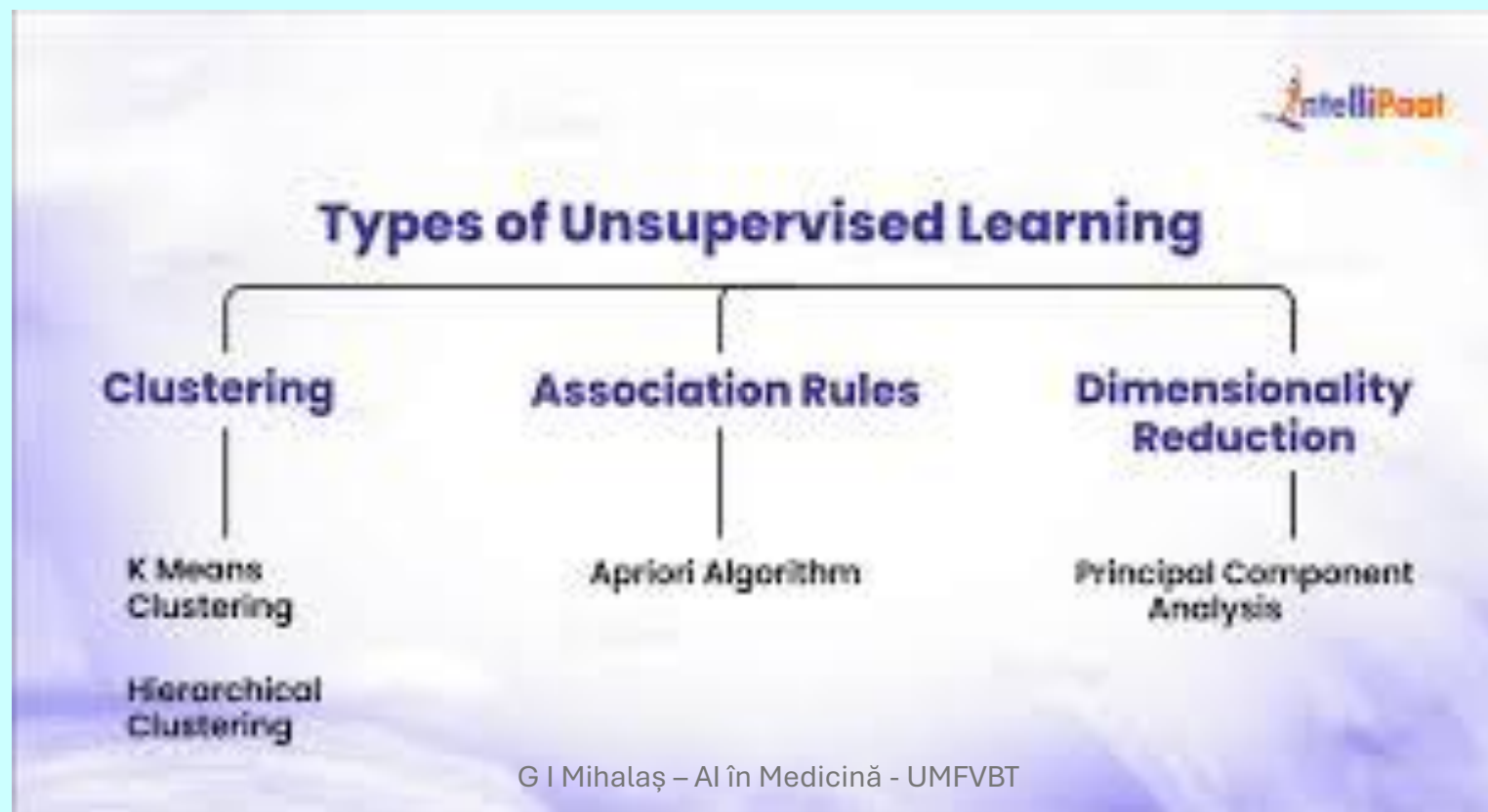
► Front Med (Lausanne). 2025 Mar 4;11:1424750. doi: [10.3389/fmed.2024.1424750](https://doi.org/10.3389/fmed.2024.1424750) [↗](#)

Construction of a risk screening and visualization system for pulmonary nodule in physical examination population based on feature self-recognition machine learning model

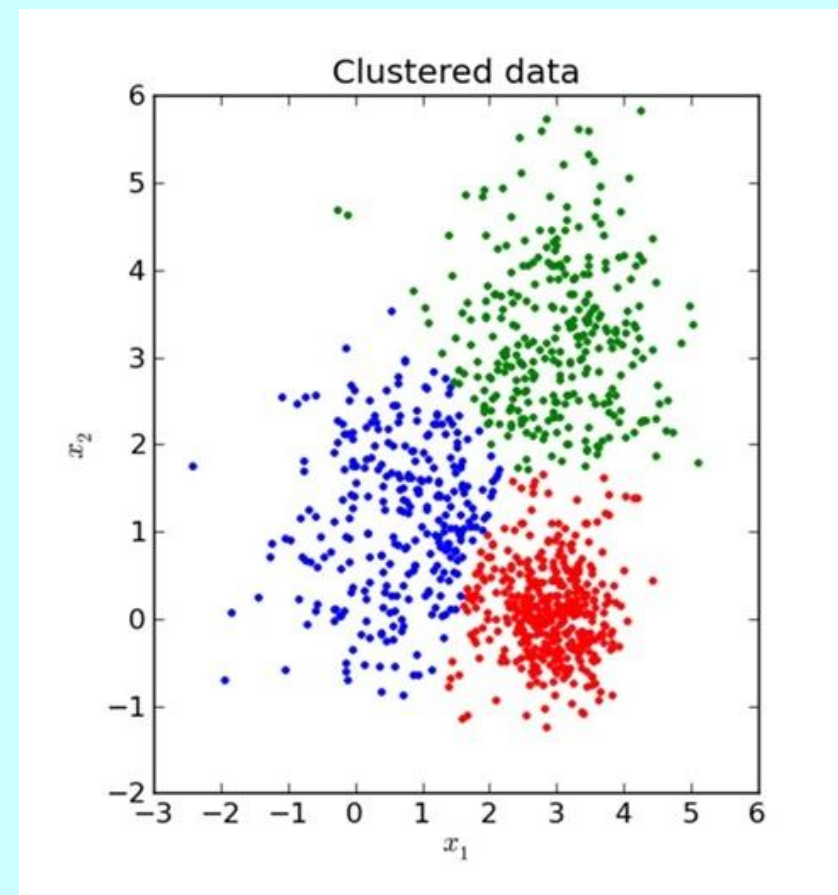
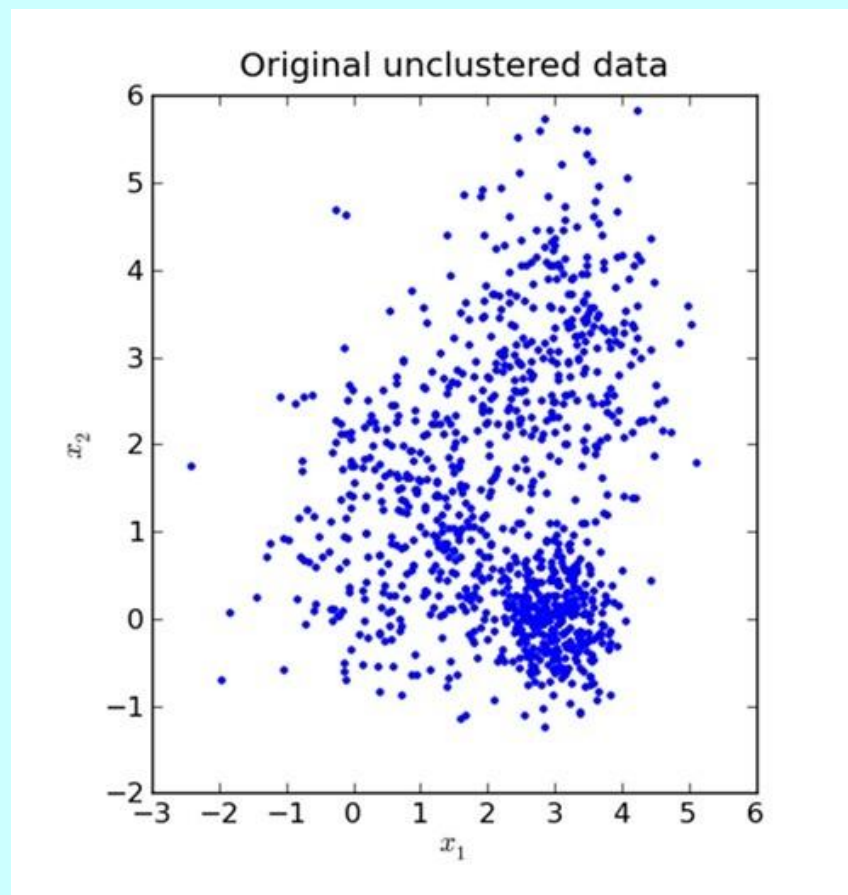
[Fang Tian](#)¹, [Yongchun Lin](#)¹, [Liangjiao Wang](#)¹, [Fei Fang](#)², [Kaiwen Hou](#)^{1,*}

Învățarea nesupervizată

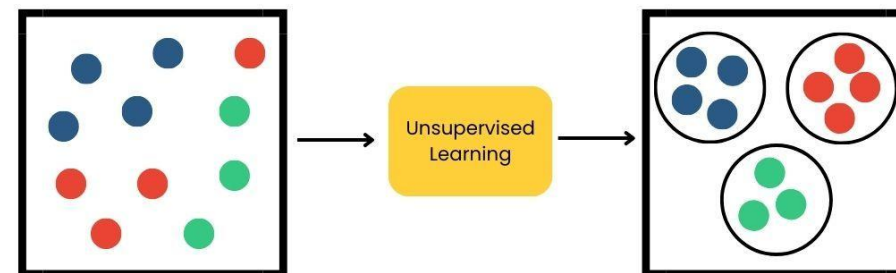
- **Principiu:**
 - Algoritmul primește date fără etichete
 - Scop: Identificarea automată de pattern-uri, grupuri sau structuri în date
- **Metode:** Clustering / Asociere / Reducerea dimensionalității



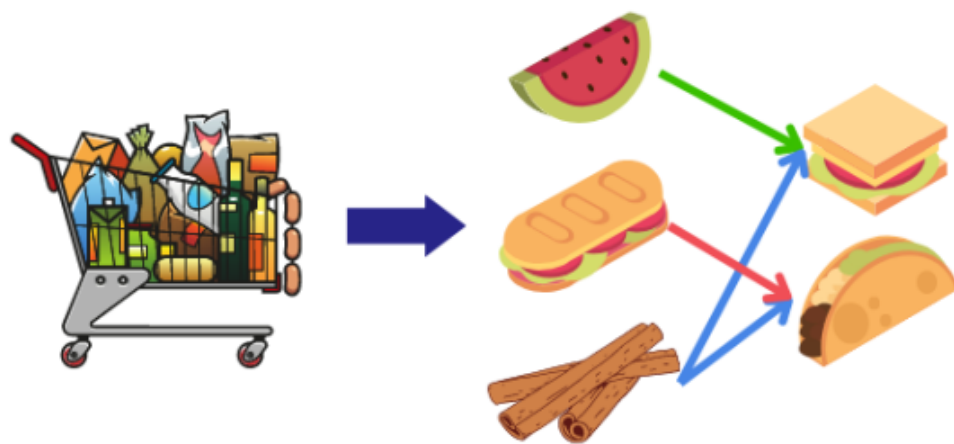
Clustering



Asociere



Association Rule Learning



"93% of people who purchased item A
also purchased item B"

Reducere dimensională

Ex: predicția riscului IMA

- Factori inițiali (8):
 - TAS, LDL-chol, fumat, vârsta, BMI, gly (a-jeun), act.fizică, nivel-stres

<i>Variabile incluse în model</i>	<i>% din variația totală explicată</i>
Toți cei 8 factori (TAS, LDL, fumat, vârstă, BMI, . glicemie, activitate fizică, stres)	100%
TAS, LDL-colesterol, fumat, vârstă	96.6%
TAS, LDL-colesterol, fumat	94.3%
TAS, LDL-colesterol	74.8%

PCA - Analiza Componentelor Principale

- **Principiu:**

- Transformă variabilele inițiale (parțial corelate) într-un set nou de **componente principale** (necorelate) ordonate după importanța lor în explicarea variației din date

- **Avantaje:**

- Reduce numărul de variabile dar păstrează informația esențială

- **Exemple:**

- $BMI = G(kg) / h(m)^2$; (combină h și G)
- $HOMA-IR = (gly \times insulinemia) / 405$; pentru rezistența la insulină
- Scorul APGAR – combină: ritm cardiac, respirație, tonus etc (la nou-născuți)
- eGFR (Estimated Glomerular Filtration Rate): creatinină, vârstă, sex

Aplicații medicale pentru învățarea nesupervizată

- **Identificarea subtipurilor de boli**
 - *(clustering în oncologie, boli rare, COVID-19)*
- **Stratificarea pacienților**
 - *(în funcție de factori de risc, răspuns la tratament)*
- **Detecția valorilor atipice**
 - *(outliers = erori sau cazuri rare)*
- **Reducerea dimensionalității datelor**
 - *(ex: PCA pentru date genetice)*
- **Vizualizarea datelor complexe**
 - *(proiecții 2D pentru explorare clinică)*
- **Explorarea bazelor de date medicale**
 - *(descoperirea de patternuri ascunse)*



ResearchGate

<https://www.researchgate.net> › ... · Traducerea acestei pagini ⋮

(PDF) Novel subgroups of adult-onset diabetes and their ...

A landmark study by **Ahlqvist et al. (2018)** introduced a novel approach to **diabetes** classification using data-driven unsupervised cluster analysis. Conducted ...



The University of North Carolina at Chapel Hill

<https://cdr.lib.unc.edu> › downloads PDF ⋮

Molecular portraits of human breast tumours

de CM Perou · 2000 · Citat de 21376 ori — Perou, C. M. et al. Distinctive gene expression patterns in human mammary ... Nature 403, 503±511 (2000). Supplementary information is available on ...



Cell Press

<https://www.cell.com> › cell-host-microbe › pdfExtended ⋮

Forum - Rethinking "Enterotypes"

de D Knights · 2014 · Citat de 497 ori — It has been suggested that human gut microbiomes fall into three distinct types or "enterotypes" (Arumugam et al., 2011). Although in the ...

„Reinforcement learning”

- **Principiu**

- Agentul învață prin **încercare și eroare**, interacționând cu un mediu, primind **recompense sau penalizări**

- **Metode**

- Q-Learning (*cu sau fără tabel de acțiuni Q [=Quality matrix]*)
- Deep Q-Networks (DQN)

- **Exemple medicale:**

- Ajustarea dozajului insulinei în diabet folosind DQN (ex: Fox et al., 2021)
- Personalizarea chimioterapiei în oncologie



springermedizin.de

<https://www.springermedizin.de> › ... · Traducerea acestei pagini

Role of artificial intelligence in enhancing insulin ...

Liu et al. [33] developed an algorithm based on the **DQN** to recommend the number of oral antidiabetic drugs and insulins for people with T2D. The evaluation of ...

LIMITĂRI ȘI PRECAUȚII



Limitări ale Machine Learning în Medicină

- **Lipsa explicabilității (black-box)**
 - Modelele complexe (ex. rețele neuronale) oferă predicții fără justificare clară → dificultăți în încredere și adoptare clinică.
- **Bias în date și învățare**
 - Date incomplete, dezechilibrate sau eronate duc la decizii părtinitoare (ex: subreprezentarea anumitor grupuri de pacienți).
- **Sensibilitate la zgomot**
 - ML poate „învața greșit” din variații nesemnificative, ducând la suprainvțătare (overfitting) sau concluzii eronate.
- **Validare limitată**
 - Multe modele nu sunt testate pe date independente sau în condiții reale de spital.

Precauții și bune practici

- **Validare riguroasă și transparentă**
 - Cross-validation, testare pe seturi independente, descriere completă a procesului.
- **Colaborare interdisciplinară**
 - Implicarea medicilor în selecția caracteristicilor, interpretarea rezultatelor și evaluarea clinică.
- **Clarificarea scopului**
 - ML este un instrument de suport, nu înlocuiește judecata medicală → trebuie integrat cu discernământ.
- **Respectarea eticii și protecției datelor**
 - Conformitate cu GDPR și bune practici pentru date sensibile → anonimizare, consimțământ, audit.

CONCLUZII

- **Machine Learning** devine o componentă esențială a **analizei datelor medicale** moderne.
- **Alegerea corectă a metodei ML** depinde de tipul de date și scopul analizei (clasificare, predicție, grupare, detecție de anomalii etc.).
- **Datele de calitate** și înțelegerea contextului medical sunt mai importante decât complexitatea modelului.

*„Învățarea este ca o busolă:
te poate ghida, dar nu poate merge în locul tău.”*

Mulumesc !