

INTELIGENȚĂ ARTIFICIALĂ



Sisteme inteligente

Sisteme care învață singure

– kNN si programare genetică –

Laura Dioșan

Sumar

A. Scurtă introducere în Inteligența Artificială (IA)

C. Sisteme inteligente

■ Sisteme care învață singure

- Arbori de decizie
- Rețele neuronale artificiale
- kNN
- Algoritmi evolutivi
- **Mașini cu suport vectorial**

■ Sisteme bazate pe reguli

■ Sisteme hibride

B. Rezolvarea problemelor prin căutare

■ Definirea problemelor de căutare

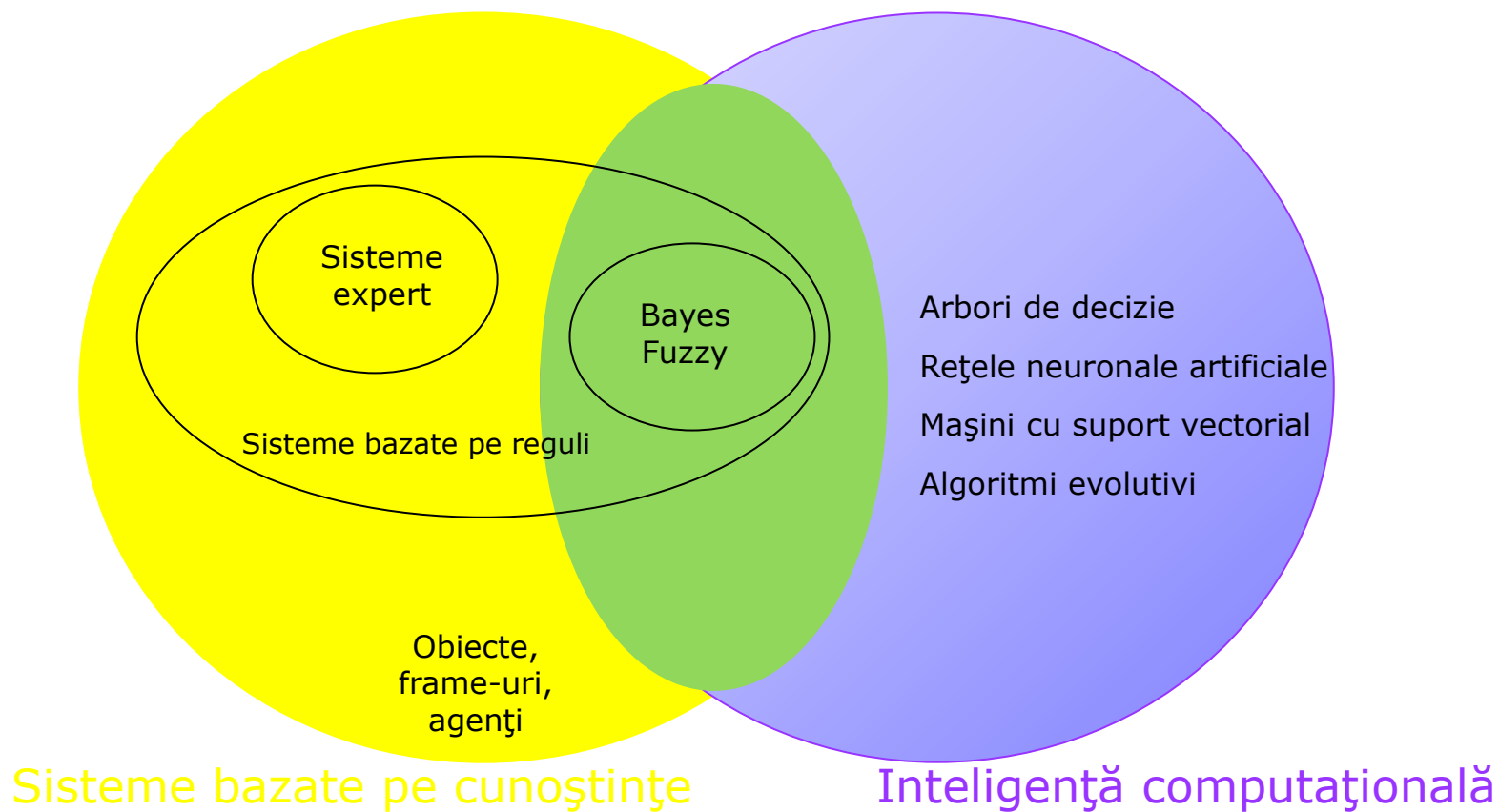
■ Strategii de căutare

- Strategii de căutare neinformate
- Strategii de căutare informate
- Strategii de căutare locale (Hill Climbing, Simulated Annealing, Tabu Search, Algoritmi evolutivi, PSO, ACO)
- Strategii de căutare adversială

Materiale de citit și legături utile

- ❑ capitolul 15 din *C. Groșan, A. Abraham, Intelligent Systems: A Modern Approach, Springer, 2011*
- ❑ Capitolul 9 din *T. M. Mitchell, Machine Learning, McGraw-Hill Science, 1997*
- ❑ capitolul VI (18) din *S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 1995*
- ❑ capitolul 10 și 11 din *C. Groșan, A. Abraham, Intelligent Systems: A Modern Approach, Springer, 2011*
- ❑ capitolul V din *D. J. C. MacKey, Information Theory, Inference and Learning Algorithms, Cambridge University Press, 2003*
- ❑ capitolul 3 din *T. M. Mitchell, Machine Learning, McGraw-Hill Science, 1997*

Sisteme inteligente



Sisteme inteligente – SIS – Învățare automată

□ Tipologie

- În funcție de experiența acumulată în timpul învățării:
 - SI cu învățare supervizată
 - SI cu învățare nesupervizată
 - SI cu învățare activă
 - SI cu învățare cu întărire
- În funcție de modelul învățat (algoritmul de învățare):
 - Rețele neuronale artificiale
 - **Mașini cu suport vectorial (MSV)**
 - Algoritmi evolutivi
 - kNN
 - Arbori de decizie
 - Modele Markov ascunse

Materiale de citit și legături utile

- ❑ capitolul VI (18) din *S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 1995*
- ❑ capitolul 10 și 11 din *C. Groșan, A. Abraham, Intelligent Systems: A Modern Approach, Springer, 2011*
- ❑ capitolul V din *D. J. C. MacKey, Information Theory, Inference and Learning Algorithms, Cambridge University Press, 2003*
- ❑ capitolul 3 din *T. M. Mitchell, Machine Learning, McGraw-Hill Science, 1997*

Sisteme inteligente – SIS – Învățare automată

□ Tipologie

- În funcție de experiența acumulată în timpul învățării:
 - SI cu învățare supervizată
 - SI cu învățare nesupervizată
 - SI cu învățare activă
 - SI cu învățare cu întărire
- În funcție de modelul învățat (algoritmul de învățare):
 - Arbori de decizie
 - Rețele neuronale artificiale
 - **Mașini cu suport vectorial (MSV)**
 - Algoritmi evolutivi
 - Modele Markov ascunse

- Mașini cu suport vectorial (MSV)
 - Definire
 - Tipuri de probleme rezolvabile
 - Avantaje
 - Dificultăți
 - Tool-uri

Sisteme inteligente – SIS – MSV

□ Definire

- Dezvoltate de Vapnik în 1970
- Popularizate după 1992
- Clasificatori liniari care identifică un hiperplan de separare a clasei pozitive de clasa negativă
- Au o fundamentare teoretică foarte riguroasă
- Funcționează foarte bine pentru date de volum mare (analiza textelor, analiza imaginilor)

■ Reamintim

- Problemă de învățare supervizată în care avem un set de date de forma:
 - (x^d, t^d) , cu:
 - $x^d \in \mathbf{R}^m \rightarrow x^d = (x^d_1, x^d_2, \dots, x^d_m)$
 - $t^d \in \mathbf{R} \rightarrow t^d \in \{1, -1\}$, $1 \rightarrow$ clasă pozitivă, $-1 \rightarrow$ clasă negativă
 - cu $d = 1, 2, \dots, n, n+1, n+2, \dots, N$
- Primele n date (se cunosc x^d și t^d) vor fi folosite drept bază de antrenament a MSV
- Ultimele $N-n$ date (se cunosc doar x^d , fără t^d) vor fi folosite drept bază de testare a MSV

Sisteme inteligente – SIS – MSV

□ Definire

- MSV găsește o funcție liniară de forma $f(\mathbf{x}) = \langle \mathbf{w} \cdot \mathbf{x} \rangle + b$, (\mathbf{w} -vector pondere) a.î.

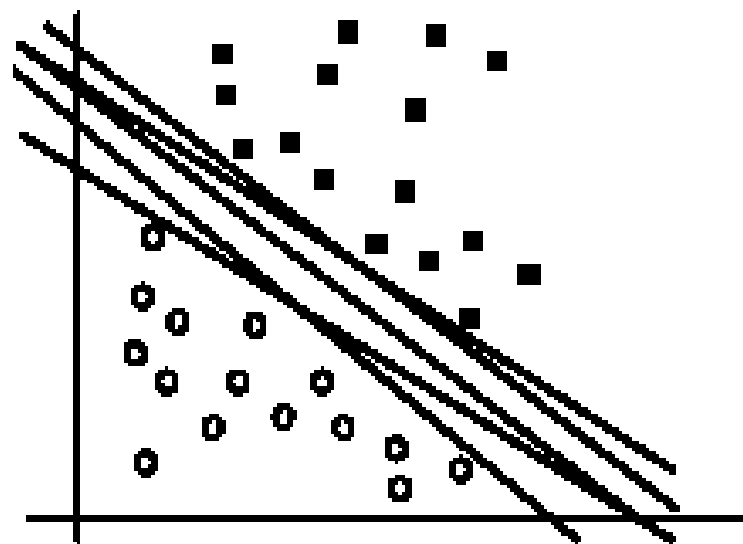
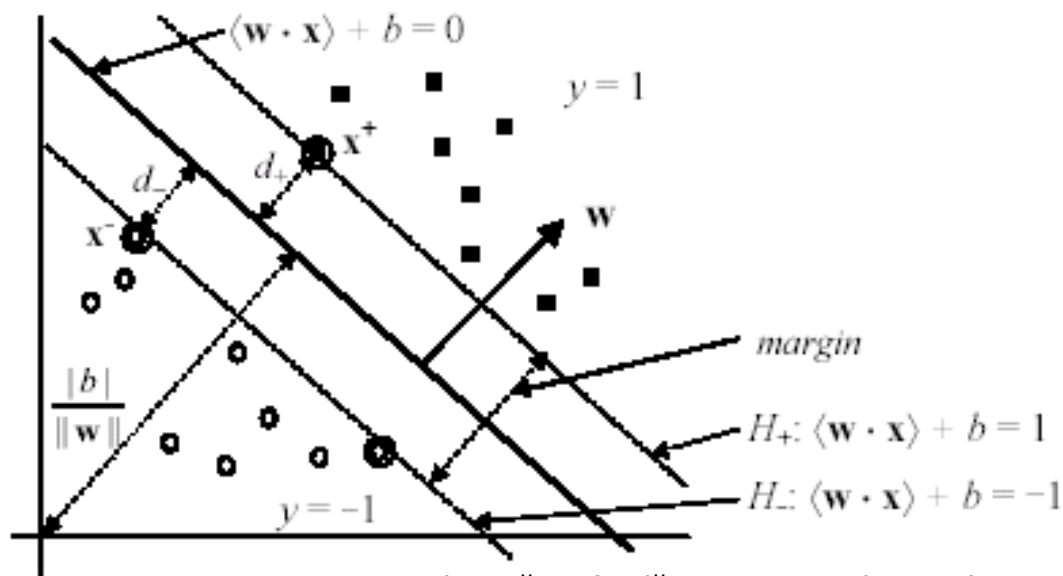
$$y_i = \begin{cases} 1 & \text{if } \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b \geq 0 \\ -1 & \text{if } \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b < 0 \end{cases}$$

- $\langle \mathbf{w} \cdot \mathbf{x} \rangle + b = 0 \rightarrow$ hiperplanul de decizie care separă cele 2 clase

Sisteme inteligente – SIS – MSV

□ Definiere

- Pot exista mai multe hiperplane
 - Care este cel mai bun hiperplan?
- MSV caută hiperplanul cu cea mai largă margine (cel care micșorează eroarea de generalizare)
 - Algoritmul SMO (*Sequential minimal optimization*)



Sisteme inteligente – SIS – MSV

□ Tipuri de probleme rezolvabile

■ Probleme de clasificare → Cazuri de date

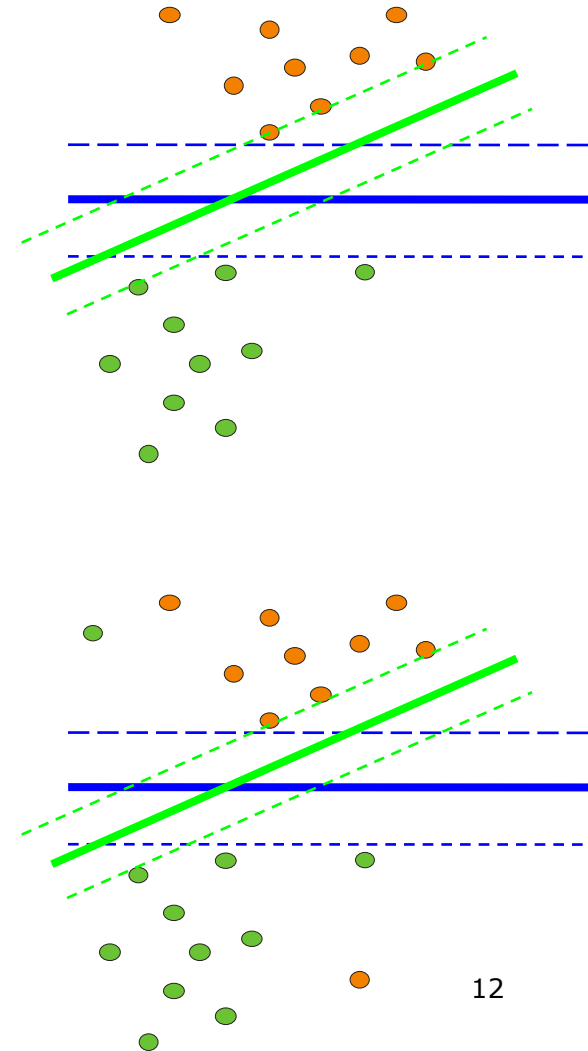
□ Liniar separabile

■ Separabile

- Eroarea = 0

■ Ne-separabile

- Se relaxează constrângerile → se permit unele erori
- C – coeficient de penalizare

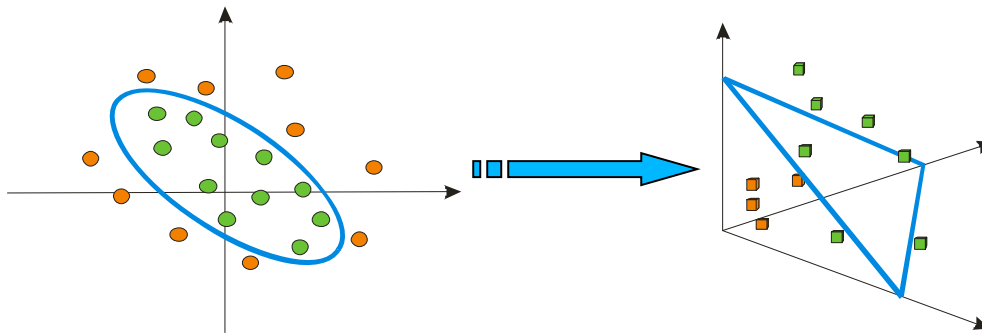


Sisteme inteligente – SIS – MSV

Cazuri de date

□ Non-liniar separabile

- Spațiul de intrare se transformă într-un spațiu cu mai multe dimensiuni (*feature space*), cu ajutorul unei funcții kernel, unde datele devin linear separabile



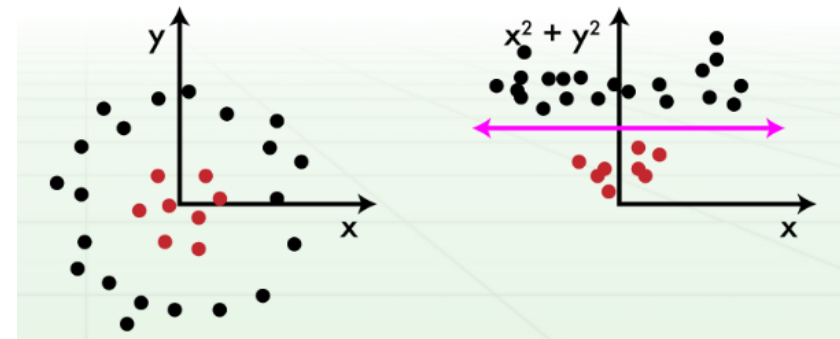
■ Kernele posibile

□ Clasice

- Polynomial kernel: $K(\mathbf{x}_1, \mathbf{x}_2) = (\langle \mathbf{x}_1, \mathbf{x}_2 \rangle + 1)^d$
- RBF kernel: $K(\mathbf{x}_1, \mathbf{x}_2) = \exp(-\sigma \|\mathbf{x}_1 - \mathbf{x}_2\|^2)$

□ Kernele multiple

- Liniare: $K(\mathbf{x}_1, \mathbf{x}_2) = \sum w_i K_i$
- Ne-liniare
 - Fără coeficienți: $K(\mathbf{x}_1, \mathbf{x}_2) = K_1 + K_2 * \exp(K_3)$
 - Cu coeficienți: $K(\mathbf{x}_1, \mathbf{x}_2) = K_1 + c_1 * K_2 * \exp(c_2 + K_3)$



Sisteme inteligente – SIS – MSV

□ Configurarea MSV

■ Parametrii unei MSV

□ Coeficientul de penalizare C

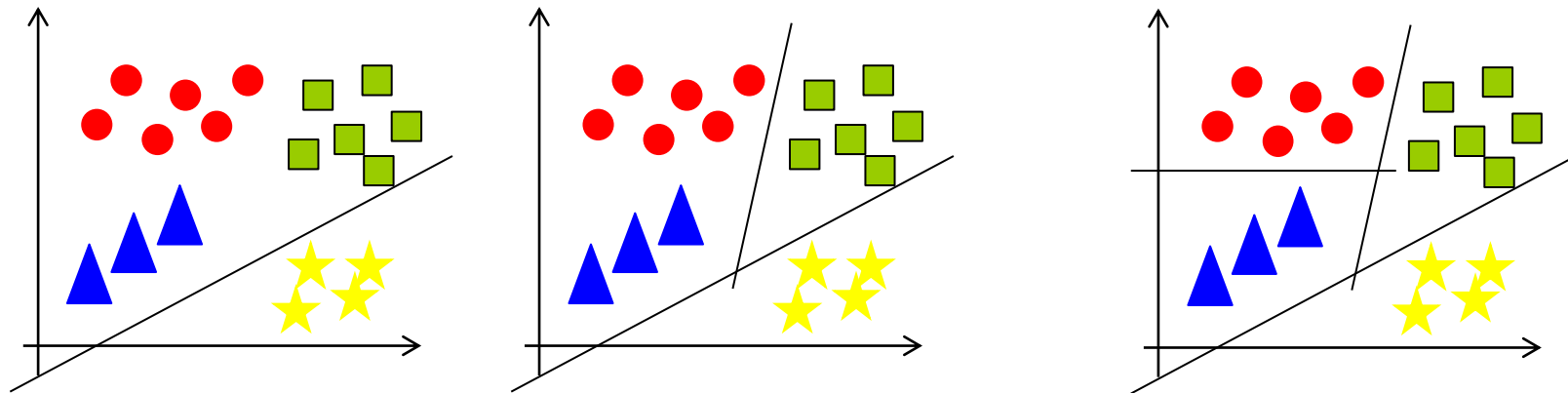
- C – mic → convergență lentă
- C – mare → convergență rapidă

□ Parametrii funcției kernel (care kernel și cu ce parametri)

- Dacă m (nr de attribute) este mult mai mare decât n (nr de instanțe)
 - MSV cu kernel liniar (MSV fără kernel) → $K(\mathbf{x}^{d1}, \mathbf{x}^{d2}) = \mathbf{x}^{d1} \cdot \mathbf{x}^{d2}$
- Dacă m (nr de attribute) este mare, iar n (nr de instanțe) este mediu
 - MSV cu kernel Gaussian $K(\mathbf{x}^{d1}, \mathbf{x}^{d2}) = \exp(-||\mathbf{x}^{d1} - \mathbf{x}^{d2}||^2 / 2\sigma^2)$
 - σ – dispersia datelor de antrenament
 - Attributele instanțelor trebuie normalizate (scalate la (0,1))
- m (nr de attribute) este mic, iar n (nr de instanțe) este mare
 - Se adaugă noi attribute, iar apoi
 - MSV cu kernel liniar

Sisteme inteligente – SIS – MSV

- MSV pentru probleme de clasificare supervizate cu mai mult de 2 clase
 - Una vs. restul (one vs. all)



Sisteme inteligente – SIS – MSV

□ MSV structurate

■ Învăţare automată

□ Normală $f: \mathcal{X} \rightarrow \mathbf{R}$

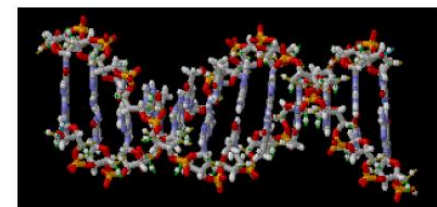
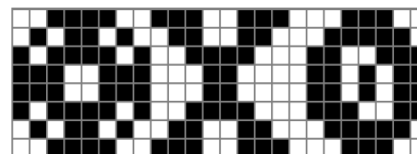
- Intrări de orice fel
- Ieşiri numerice (naturale, întregi, reale)

□ Structurată: $\mathcal{X} \rightarrow \mathcal{Y}$

- Intrări de orice fel
- Ieşiri de orice fel (simple sau structurate)

■ Informaţii structurate

- Texte şi hiper-texte
- Molecule şi structuri moleculare
- Imagini



Sisteme inteligente – SIS – MSV

□ MSV structurate

■ Aplicații

□ Procesarea limbajului natural

- Traduceri automate (ieșiri → propoziții)
- Analiza sintactică și/sau morfologică a propozițiilor (ieșiri → arborele sintactic și/sau morfologic)

□ Bioinformatică

- Predicția unor structuri secundare (ieșirile → grafe bi-partite)
- Predicția funcționării unor enzime (ieșirile → *path*-uri în arbori)

□ Procesarea vorbirii

- Transcrieri automate (ieșiri → propoziții)
- Transformarea textelor în voce (ieșiri → semnale audio)

□ Robotică

- Planificare (ieșirile → secvențe de acțiuni)

Sisteme inteligente – SIS – MSV

□ Avantaje

- Pot lucra cu orice fel de date (liniar separabile sau nu, distribuit uniform sau nu, cu distribuție cunoscută sau nu)
 - Funcțiile kernel care crează noi atribute (features) → straturile ascunse dintr-o RNA
- Dacă problema e convexă oferă o soluție unică → optimul global
 - RNA pot asocia mai multe soluții → optime locale
- Selectează automat mărimea modelului învățat (prin vectorii suport)
 - În RNA straturile ascunse trebuie configurate de către utilizator *a priori*
- Nu învață pe derost datele (overfitting)
 - RNA se confruntă cu problema overfitting-ului chiar și cand modelul se învață prin validare încrucișată

□ Dificultăți

- Doar atribute reale
- Doar clasificare binară
- Background matematic dificil

□ Tool-uri

- LibSVM → <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Weka → SMO
- SVMLight → <http://svmlight.joachims.org/>
- SVM Torch → <http://www.torch.ch/>
- <http://www.support-vector-machines.org/>

Cursul următor

A. Scurtă introducere în Inteligența Artificială (IA)

B. Sisteme inteligente

■ **Sisteme care învață singure**

- Arbori de decizie
- Rețele neuronale artificiale
- Mașini cu suport vectorial
- Algoritmi evolutivi

■ Sisteme bazate pe reguli

■ Sisteme hibride

C. Rezolvarea problemelor prin căutare

■ Definirea problemelor de căutare

■ Strategii de căutare

- Strategii de căutare neinformate
- Strategii de căutare informate
- Strategii de căutare locale (Hill Climbing, Simulated Annealing, Tabu Search, Algoritmi evolutivi, PSO, ACO)
- Strategii de căutare adversarială

Cursul următor – Materiale de citit și legături utile

- ❑ capitolul 15 din *C. Groșan, A. Abraham, Intelligent Systems: A Modern Approach, Springer, 2011*
- ❑ Capitolul 9 din *T. M. Mitchell, Machine Learning, McGraw-Hill Science, 1997*

□ Informațiile prezentate au fost colectate din diferite surse de pe internet, precum și din cursurile de inteligență artificială ținute în anii anteriori de către:

■ Conf. Dr. Mihai Oltean –
www.cs.ubbcluj.ro/~moltean

■ Lect. Dr. Crina Groșan -
www.cs.ubbcluj.ro/~cgrosan

■ Prof. Dr. Horia F. Pop -
www.cs.ubbcluj.ro/~hfpop