

1. Ce se poate spune despre fazele algoritmului de retro-propagare (backpropagation)?

În prima fază, rețeaua primește vectorul de intrare și propaga semnalul înainte, strat cu strat, până se generează ieșirea. În a doua fază, semnalul de eroare este propagat înapoi, de la stratul de ieșire către stratul de intrare, iar gradientii calculați sunt folosiți pentru ajustarea ponderilor rețelei.

2. Care a fost scopul inițial al sistemelor Lindenmeyer?

Modele matematice pentru creșterea plantelor

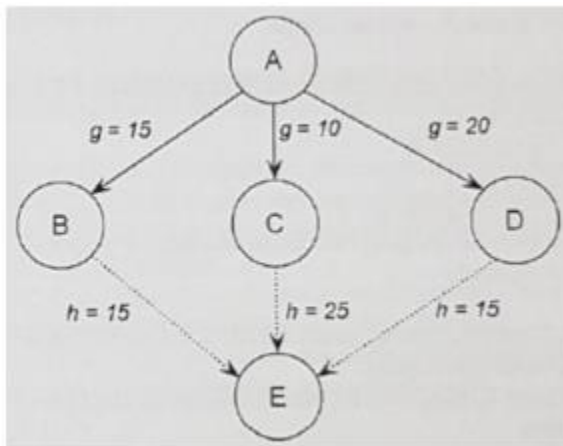
3. Care sunt avantajele planificării cu ordine parțială față de planificarea cu ordine totală?

În planificarea cu ordine parțială, se poate profita de descompunerea problemei: se rezolvă independent subscopurile și apoi se combină subplanurile.

4. Care este ideea de bază a clasificării folosind mașini cu vectori suport?

Maximizarea marginii de separare între cele două clase.

5. După expandarea nodului A, care nod va fi prelucrat primul de către algoritmul de căutare A\*?



Nodul B

6. Care din următoarele metode permite inferențe bazate pe probabilități?

Rețele bayesiene

7. Testul Turing este un exemplu de abordare a inteligenței artificiale din perspectiva:

Acțiunii umane

8. Care din metodele de mai jos se folosește în mod normal pentru rezolvarea unor probleme de optimizare?

Algoritmii evolutivi

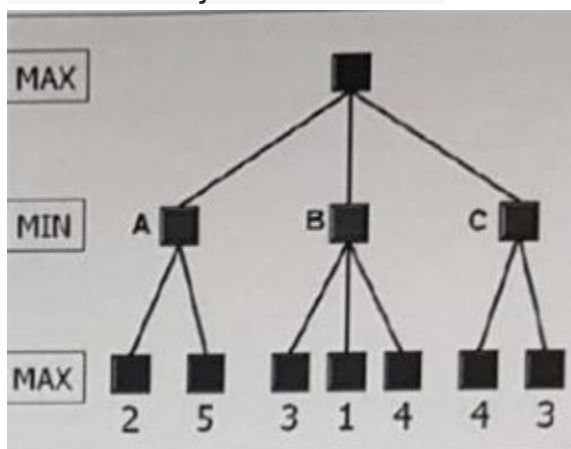
9. Care este scopul unui perceptron cu un singur strat?

Clasificator pentru 2 clase prin gasirea unui hiperplan care le separa

10. Fie un algoritm evolutiv cu codare reală, pentru care se dau parintii  $p1 = (1,2,3,4)$  si  $p2 = (5,2,1,6)$ . Care din cromozomii următori reprezinta un copil valid rezultat prin incrucisare aritmetica cu factorul  $\alpha=0.5$ ?

(3, 2, 2, 5)

11. Pentru arborele din figură, aplicând algoritmul Minimax, care mutare va fi aleasă de jucătorul MAX?



mutarea C

12. Cine a inventat limbajul Lisp în anul 1958:

John Mccarthy

13. Care este caracteristica unei căutări "best-first"?

✓Ordonează nodurile frontierei în ordinea crescătoare a costului estimat până la un nod scop

14. Fie clauzele (a sau b sau c) și (not a sau not b sau d). Care sunt rezolvenții rezultați prin algoritmul de rezoluție propozițională?

✓(b sau c sau notb sau d) și (a sau c sau nota sau d)

15. Care este unificarea expresiilor următoare:  $P(X, g(Y))$  și  $Q(f(Z), g(y))$  ?

Expresiile nu pot fi unificate deoarece P și Q sunt predicate distincte

16. Cum pot fi rezolvate problemele nelineare liniar cu ajutorul mașinilor cu vectori suport?

✓ Se aplică un nucleu liniar care transformă datele într-un spațiu multidimensional, unde acestea devin liniar separabile

17. Cum se poate defini inteligența?

✓ O măsură a capacității de a atinge scopuri într-un mediu complex și dinamic

18. Fie următoarele două propoziții: (A sau B), respectiv (non A sau C). Ce concluzie se poate trage aplicând rezoluția propozițională?

B sau C

19. În cazul unui joc cu factor de ramificare variabil și necesitate de căutare pe multe niveluri, în care dorim totuși să avem la un moment de timp dat o mutare disponibilă, chiar dacă nu mutarea optimă, ce strategie alegem?

Căutarea iterativă în adâncime

20. Care este presupunerea STRIPS care evită problema cadrului (frame problem)?

✓ Orice termen care nu este menționat în efect rămâne neschimbat

21. Care este structura unui perceptron multistrat?

✓ Mai multe straturi de neuroni interconectați, fiecare neuron calculează suma ponderată a intrărilor la care se adăugă un termen constant și apoi rezultatului i se aplică o funcție de activare neliniară

22. Ce se poate spune despre performanțele algoritmului de căutare iterativă în adâncime?

✓ Algoritmul este complet optim dacă un pas are costul 1

23. Ce tip de inferență fuzzy are efectul celei din figură?



Larsen

24. Fie următoarele metode de învățare cu întărire pasivă: Estimarea directă a utilității (EDU), Programarea dinamică adaptivă (PDA) și Învățarea diferențelor temporale (IDT). Care din următoarele afirmații este adevărată?

✓ IDT combină EDU cu corecțiile date de ecuația Bellman

25. Care este expresia utilizată pentru inferența prin enumerare într-o rețea bayesiană?

Se consideră:

X - variabila interogată

y - variabilele neobservate

e - variabilele observate

alfa - coeficientul de normalizare.

$$\begin{aligned} A - P(X, y) &= \alpha \cdot P(X|y) \cdot P(y) \\ B - P(X|y) &= \alpha \cdot \sum_e P(X, e, y) \\ C - P(X, e) &= \alpha \cdot P(X|e) \cdot P(e) \\ D - P(X|e) &= \alpha \cdot \sum_y P(X, e, y) \end{aligned}$$

Răspuns: D

26. Unde a fost realizat Deep Blue, care l-a învins la șah pe Gari Kasparov în 1997?

IBM

27. Ce semnificație are ecuația Bellman?

✓ Utilitatea unei stări este recompensa imediată pentru acea stare plus utilitatea așteptată maximă a stării următoare

28. Fie numărul fuzzy triunghiular A cu centrul  $c = 50$ , lățimea la stânga  $a = 30$  și lățimea la dreapta  $b = 20$ , astfel încât suportul numărului este intervalul  $(20, 70)$ . Care este gradul de apartenență al elementului  $x = 30$  în mulțimea A?

1/3

29. Care este ordinea operatorilor genetici pentru algoritmul de evoluție diferențială?

Mutație, încrucișare, selecție

30. Pentru un algoritm evolutiv, cum funcționează selecția de tip competiție (engl. "tournament")?

✓ Un părinte se alege ca fiind cel mai adaptat individ dintr-un număr de indivizi selectați aleatoriu din populație

31. Care este ideea de bază a algoritmului "particle swarm optimization"?

✓ Viteza fiecărei particule este actualizată pe baza unei componente cognitive, calculată cu ajutorul optimului personal al particulei, și a unei componente sociale, calculată cu ajutorul optimului social, al vecinătății particulei

32. Care dintre următoarele afirmații reprezintă o diferență corectă între sistemele de inferență vagă (fuzzy) Mamdani și Sugeno?

✓ Metoda Mamdani permite descrierea mai intuitivă a cunoștințelor; metoda Sugeno este mai eficientă computațional

33. În algoritmul de căutare pe arbori Monte Carlo (Monte Carlo Tree Search, MCTS), metoda de selecție UCB1 presupune selectarea acțiunii care maximizează următoarea expresie:  $w_i / n_i + \sqrt{2 \ln n / n_i}$ . Ce reprezintă variabilele din această formulă?

✓  $w_i$  este numărul de victorii în fiul  $i$ ,  $n_i$  este numărul de simulări în fiul  $i$ , iar  $n$  este numărul de simulări în nodul curent

34. Dacă:

- $w_i$  este ponderea conexiunii  $i$  la momentul  $k$
- $\alpha$  este rata de învățare
- $y_d$  este ieșirea dorită
- $y$  este ieșirea reală
- $x_i$  este intrarea  $i$

Care este regula de actualizare a ponderilor unui perceptron cu un singur strat?

✓  $w_i(k+1) = w_i(k) + \alpha * (y_d(k) - y(k)) * x_i(k)$

35. Ce proprietăți are algoritmul Kohonen?

✓ Reducerea dimensionalității și păstrarea topologie

36. Care este scopul modelului “boids”?

✓ Simulează comportamentul unui stol de păsări

37. Algoritmul Q-Learning este:

✓ Un algoritm de învățare cu întărire

38. Care sunt nivelurile ierarhiei/piramidei cunoașterii?

✓ Date - Informații - Cunoștințe - Înțelepciune

39. Călirea simulată (simulated annealing) este:

✓ Un algoritm de optimizare

1. Cine a propus modelul perceptronului la sfarsitul anilor 50?

- a) Marvin Minsky
- b) Teuvo Kohonen
- c) **Frank Rosenblatt**

### 3. Functia XOR:

- a) poate fi invatata de un perceptron multistrat
- b) nu poate fi invatata de o retea neuronală
- c) poate fi invatata de un perceptron cu un singur strat

*Poate învăța tot ce poate reprezenta, dar nu poate reprezenta multe funcții (neselectabile liniar, de ex. XOR)*

4. Care este scopul unui perceptron cu un singur strat?

- a) clasificator pentru n clase prin invatarea clasei de apartenenta a fiecărei instante
- b) **clasificator pentru 2 clase prin gasirea unui hiperplan care le separa**
- c) clasificator pentru 2 clase prin gasirea unui hipervolum care le separa

### 5. Care este iesirea unui perceptron cu un singur strat?

- a) functia sigmoida aplicata intrarii cu ponderea cea mai mare
- b) **functia Heaviside aplicata sumei intrarilor ponderate**
- c) functia sigmoida aplicata sumei intrarilor ponderate

6. Dacă  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de învățare,  $t$  este ieșirea dorită,  $a$  este ieșirea reală și  $x_i$  este intrarea  $i$ , care este regula de actualizare a ponderilor unui perceptron cu un singur strat?

- a)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$
- b)  **$w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$**
- c)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$

25. Care este structura unui perceptron multistrat?

- a) mai multe straturi de neuroni interconectați, fiecare din ei calculează suma ponderată a intrărilor, rezultatului  $i$  se aplică o funcție de activare neliniară și apoi se adaugă un termen constant
- b) mai multe straturi de neuroni interconectați, asupra fiecărei intrări a neuronilor se aplică o funcție de activare neliniară, se face suma ponderată a rezultatelor la care se adaugă un termen constant
- c) **mai multe straturi de neuroni interconectați, fiecare din ei calculează suma ponderată a intrărilor la care se adaugă un termen constant și apoi rezultatului  $i$  se aplică o funcție de activare neliniară**

### 26. Ce înseamnă conectarea feedforward a neuronilor într-un perceptron multistrat?

- a) interconectarile formează bucle
- b) neuronii sunt total conectați (fiecare cu fiecare)
- c) interconectarile nu formează bucle

*Perceptronul multistrat este o rețea neuronală cu propagare înainte (.feed-forward.) cu unul sau mai multe straturi ascunse*

27. Cum se numesc rețelele perceptron multistrat în care interconectarile formează una sau mai multe bucle?

- a) **recurente**

- b) feedforward
- c) stratificate

28. Care este expresia functiei sigmoide unipolare (functia logistica)?

- a)  $S(s) = 1 / (1 + \exp(-s))$
- b)  $S(s) = 1 / (1 + \exp(s))$
- c)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$

29. Care este expresia functiei sigmoide bipolare (tangenta hiperbolica)?

- a)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$
- b)  $S(s) = 1 / (1 + \exp(-s))$
- c)  $S(s) = 1 / (\exp(s) + \exp(-s))$

30. Cum se realizeaza antrenarea unui perceptron multistrat?

- a) prin modificarea modului de interconectare a neuronilor
- b) prin modificarea ponderilor conexiunilor si a valorilor prag ale neuronilor**
- c) prin modificarea valorilor neuronilor care retin informatiile

31. Ce sunt straturile ascunse ale unui perceptron multistrat?

- a) straturi interne retelei care nu participa direct la antrenare
- b) straturi interne retelei, nici de intrare si nici de iesire
- c) straturi interne retelei care sunt parcurse uneori de semnale dinspre intrare spre iesire

*Un strat ascuns își .ascunde. iesirea dorită; cunoscând corespondența intrare-iesire a rețelei (.cutie neagră.), nu se poate deduce care trebuie să fie ieșirea dorită a unui neuron dintr-un strat ascuns*

32. De ce perceptronul multistrat nu are functii de activare liniare?

- a) functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat
- b) datorita progresului tehnologic, functiile exponentiale care pot fi utilizate acum pentru perceptronul multistrat erau imposibil de utilizat practic in anii '50 pentru perceptronul cu un singur strat
- c) functiile liniare nu asigura o crestere a puterii de calcul in raport cu rețelele cu un singur strat, deoarece o functie liniara de functii liniare este tot o functie liniara**

33. Cine a introdus in 1986 algoritmul backpropagation ca metoda generala de antrenare a perceptronilor multistrat?

- a) Werbos
- b) Rumelhart, Hinton si Williams**
- c) Bryson si Ho

34. Care este ideea care sta la baza algoritmului backpropagation?

- a) minimizarea erorii (diferentei dintre iesirea dorita si iesirea reala) in raport cu ponderile prin metoda gradientului descendent**
- b) minimizarea ponderilor conexiunilor prin metoda gradientului descendent
- c) minimizarea ponderilor conexiunilor in doua etape: una de propagare inainte si una de propagare inapoi

35. Pentru aplicarea algoritmului backpropagation, care este derivata functiei sigmoide unipolare?

- a)  $S'(s) = (1 - S(s)) * (1 + S(s))$



**b)  $S'(s) = S(s) * (1 - S(s))$**

c)  $S'(s) = 1 - S(s)$

36. Pentru aplicarea algoritmului backpropagation, care este derivata functiei tangenta hiperbolica?

**a)  $S'(s) = (1 - S(s)) * (1 + S(s))$**

b)  $S'(s) = S(s) * (1 - S(s))$

c)  $S'(s) = 1 + S(s)$

37. In algoritmul backpropagation, cum se calculeaza gradientii erorilor pentru neuronii din stratul de iesire (delta\_k), daca S' este derivata functiei de activare, ek(t) este eroarea la momentul t iar wjk(p) sunt ponderile conexiunilor dintre stratul ascuns si cel de iesire?

a)  $\text{delta\_k}(t) = S' / \text{ek}(t)$

**b)  $\text{delta\_k}(t) = S' * \text{ek}(t)$**

c)  $\text{delta\_k}(t) = \text{ek}(t) * wjk(t)$

38. In algoritmul backpropagation, cum se calculeaza gradientii erorilor pentru neuronii din stratul ascuns (delta\_j), daca S' este derivata functiei de activare, delta\_k sunt gradientii erorilor pentru neuronii din stratul de iesire, wjk(t) sunt ponderile conexiunilor dintre stratul ascuns si cel de iesire la momentul t iar l este numarul de iesiri ale rețelei (notam suma cu i de la 1 la n din ai cu  $\sum_{i=1}^n (a_i)$ )?

**a)  $\text{delta\_j}(t) = S' * \sum_{k=1}^l (\text{delta\_k}(t) * wjk(t))$**

b)  $\text{delta\_j}(t) = S' * wjk(t)$

c)  $\text{delta\_j}(t) = S' * \text{delta\_k}(t)$

40. Care este impedimentul cel mai important al algoritmului backpropagation?

a) nu converge pentru probleme complexe de dimensiuni mari

**b) deoarece se bazeaza pe metoda gradientului descendent, poate converge intr-un minim local al functiei de eroare pe care incearca sa o minimizeze**

c) restrictiile impuse pentru valorile ratei de invatare

41. Cum se justifica proprietatea perceptronului multistrat de "aproximator universal"?

a) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie polinomiala

**b) s-a demonstrat ca o retea (posibil infinita) cu un singur strat ascuns este capabila sa aproximeze orice functie continua**

c) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie derivabila datorita algoritmului backpropagation

43. Prin ce se caracterizeaza metoda ratei de invatare adaptive pentru accelerarea invatarii cu algoritmul backpropagation?

**a) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor**

b) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective

c) consta in micșorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

57. Cum pot fi clasificate memoriile asociative din punct de vedere al asociatiilor stocate:

**a) autoasociative si heteroasociative**

b) asociative si neasociative

c) statice si recurente

62. Ce este o memorie autoasociativa?

**a) o memorie asociativa care reda mapari intrare-iesire de forma  $(x_k, x_k)$**

b) o memorie asociativa care reda mapari intrare-iesire de forma  $(y_k, x_k)$

c) o memorie asociativa care reda mapari intrare-iesire de forma  $(x_k, y_k)$

71. Cand este stabila reseaua Hopfield?

a) cand matricea de interconectare este simetrica si functia de activare este continua si monoton descrescatoare

**b) cand matricea de interconectare este simetrica si functia de activare este continua si monoton crescatoare**

c) cand matricea de interconectare este antisimetrica si functia de activare este continua si monoton crescatoare

74. Care sunt pasii principali ai metodologiei de calire simulata?

a) generarea determinista a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru asigurarea convergentei catre solutie

b) generarea aleatorie a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru cresterea distantei intre doua stari succesive acoperind zone din ce in ce mai extinse in spatiul solutiilor

**c) generarea aleatorie a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru stabilizarea intr-o anumita regiune**

83. In contextul clasificarii, care este diferenta dintre un neuron de tip perceptron si un neuron RBF?

a) perceptronul si neuronul RBF se comporta la fel in contextul clasificarii

**b) perceptronul imparte spatiul trasaturilor in hiperplane, neuronul RBF defineste un hipervolum sau o hipersfera**

c) neuronul RBF imparte spatiul trasaturilor in hiperplane, perceptronul defineste un hipervolum sau o hipersfera

85. Care este topologia unei retele RBF?

a) o retea RBF este formata din doua straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni iar al doilea contine neuroni RBF

**b) o retea RBF este formata din trei straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni, al doilea contine neuroni RBF iar al treilea contine neuroni cu functii de activare liniare**

c) o retea RBF este formata din doua straturi: primul contine neuroni RBF care prelucreaza intrarile iar al doilea contine neuroni RBF cu functii de activare liniare care transmit iesirile

91. Ce sunt estimatorii densitatii nucleului (kernel density estimators)?

**a) retele neuronale RBF pentru estimarea functiilor de densitate de probabilitate**

b) perceptroni multistrat pentru estimarea functiilor de densitate de probabilitate

c) retele neuronale stohastice pentru estimarea functiilor de densitate de probabilitate

92. Care sunt avantajele metodelor de clasificare bazate pe nucleu (kernel classifiers)?

a) nu necesita reducerea numarului de nuclee sau ajustarea latimii nucleelor

**b) invatare simpla, posibilitatea convergentei catre clasificatorul bayesian optim, care poate ramane optim fara reantrenare**

c) estimarea exacta a functiilor de densitate de probabilitate ale claselor

1. In 1943, contributia cercetatorilor McCulloch si Pitts in domeniul inteligentei artificiale a fost:

- **un model neuronal**
- un program de sah
- logica vaga (fuzzy)
- logica predicativa

2. Termenul de înțelegenta artificială a fost propus in anul:

- 1946
- **1956 - McCarthy**
- 1966
- 1976

3. Cine a inventat limbajul Lisp in anul 1958:

- Allen Newell
- **John McCarthy**
- Marvin Minsky
- Alain Colmerauer

4. Ce tip de programe de inteligenta artificiala au fost primele utilizate cu succes in **industrie**?

- sistemele de traducere automata
- retelele neuronale
- ?- **sistemele expert**
- algoritmii genetici

5. Ce sistem expert din anii 70 a fost utilizat in diagnosticarea medicala?

- Dendral
- **Mycin**
- Prospector
- XCON

6. Care a fost primul program ce incerca sa treaca testul Turing?

- ?- **Eliza**
- Alice
- Cyc
- NetTalk

7. Cine a redescoperit in anul 1986 algoritmul retro-propagarii (back-propagation)?

- Hopfield
- Hebb
- **Rumelhart, Hinton si Williams**
- Bryson si Ho

8. Unde a fost realizat Deep Blue, care l-a invins pe Gari Kasparov in 1997?

- Stanford
- Princeton
- MIT
- **IBM**

9. In ce tara a fost inventat limbajul Prolog?

- SUA
- Marea Britanie
- Germania
- **Franta**

10. Care este avantajul principal al limbajului Clips (NASA, 1985)?

- **motorul de inferenta**
- functiile de calcul numeric
- facilitatile de programare orientata obiect
- strictetea definirii tipurilor de date

1. Cine a propus modelul perceptronului la sfarsitul anilor 50?

- a) Marvin Minsky
- b) Teuvo Kohonen
- c) **Frank Rosenblatt**

2. Cine sunt autorii cartii "Perceptronii" (1969) in care se evidentia imposibilitatea perceptronului de a invata functii neseperabile liniar?

- a) **Minsky si Papert**
- b) Rosenblatt si Hopfield
- c) Rumelhart si Hinton

3. Functia XOR:

- a) **poate fi invatata de un perceptron multistrat**
- b) nu poate fi invatata de o retea neuronală
- c) poate fi invatata de un perceptron cu un singur strat

4. Care este scopul unui perceptron cu un singur strat?

- a) clasificator pentru n clase prin invatarea clasei de apartenenta a fiecărei instante
- b) **clasificator pentru 2 clase prin gasirea unui hiperplan care le separa**
- c) clasificator pentru 2 clase prin gasirea unui hipervolum care le separa

5. Care este iesirea unui perceptron cu un singur strat?

- a) functia sigmoida aplicata intrarii cu ponderea cea mai mare
- b) **functia Heaviside aplicata sumei intrarilor ponderate**
- c) functia sigmoida aplicata sumei intrarilor ponderate

6. Dacă  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de învățare,  $t$  este ieșirea dorită,  $a$  este ieșirea reală și  $x_i$  este intrarea  $i$ , care este **regula de actualizare** a ponderilor unui perceptron cu un singur strat?

a)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$

**b)  $w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$  // păstrează pasul anterior**

c)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$

7. Care este avantajul perceptronului în comparație cu Adaline?

a) algoritmul perceptronului converge și atunci când clasele nu sunt liniar separabile

**b) perceptronul garantează clasificarea corectă a 2 clase liniar separabile**

c) perceptronul funcționează numai pentru ieșiri binare

8. Care este metoda de antrenare a unui neuron Adaline?

a) backpropagation

b) trial and error

**c) regula delta (least mean square)**

9. Pe ce principiu se bazează algoritmul least mean square?

a) minimizarea pătratelor ponderilor

**b) minimizarea erorii pătratice medii dintre ieșirea dorită și cea reală**

c) minimizarea ieșirii neuronului

10. Dacă  $w_i(k)$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de învățare,  $t$  este ieșirea dorită,  $a$  este ieșirea reală și  $x_i$  este intrarea  $i$ , care este regula de actualizare a ponderilor unui neuron Adaline?

**a)  $w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$**

b)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$

c)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$

11. Care din următoarele caracteristici nu sunt avantaje ale neuronului Adaline în comparație cu perceptronul?

a) Adaline suportă ieșiri reale iar perceptronul numai ieșiri binare

b) Adaline converge către soluție destul de repede, chiar și atunci când clasele nu sunt liniar separabile

**c) Adaline garantează separarea a 2 clase liniar separabile**

12. Care este topologia Madaline?

**a) două straturi de neuroni, primul format din neuroni Adaline cu funcție Heaviside aplicată ieșirilor, iar al doilea cu porți logice**

b) mai multe straturi de neuroni Adaline

c) mai multe straturi de neuroni Adaline cu funcție Heaviside aplicată ieșirilor, împreună cu un strat cu porți logice

13. În topologia Madaline, care este diferența dintre ponderile porții logice și ponderile Adalinelor?

a) nici o diferență, toate ponderile sunt ajustabile

- b) ponderile Adalinelor sunt fixe, iar ponderile portii logice sunt ajustabile
- c) ponderile portii logice sunt fixe, iar ponderile Adalinelor sunt ajustabile**

14. La antrenarea unei retele Madaline, care este ordinea recomandata de prezentare a vectorilor de intrare?

- a) ordine aleatorie, deoarece ordinea ciclica poate impiedica convergenta**
- b) ordine ciclica, deoarece ordinea aleatorie poate impiedica convergenta
- c) nu are nici o importanta ordinea de prezentare

15. Ce inseamna principiul celei mai mici perturbatii in procesul de invatare a unei retele neuronale?

- a) ponderile trebuie ajustate astfel incat sa se reduca eroarea la iesire pentru vectorul de intrare curent cu o perturbatie minima asupra raspunsurilor deja invatate**
- b) ponderile trebuie ajustate astfel incat eroarea la iesire sa fie redusa cu o cantitate cat mai mica astfel incat sa nu fie perturbata retea
- c) ponderile trebuie ajustate cu o cantitate cat mai mica astfel incat sa nu fie perturbata retea

16. Care sunt modalitatile in care algoritmul de invatare Madaline respecta principiul celei mai mici perturbatii?

- a) se alege numarul minim de Adaline care ar putea influenta schimbarea iesirii, se aleg Adalinele cu iesirea negativa iar ponderile sunt modificate in directia vectorului de intrare
- b) se alege numarul maxim de Adaline care ar putea influenta schimbarea iesirii, se aleg Adalinele cu iesirea cat mai mica iar ponderile sunt modificate in directia vectorului de intrare
- c) se alege numarul minim de Adaline care ar putea influenta schimbarea iesirii, se aleg Adalinele cu iesirea cat mai apropiata de zero iar ponderile sunt modificate in directia vectorului de intrare**

17. Care este diferenta dintre retelele Learning Vector Quantization si hartile cu auto-organizare (SOM)? Self Organizing Map

- a) LVQ si SOM reprezinta acelasi lucru
- b) LVQ este o metoda de invatare supervizata**
- c) LVQ este o metoda de invatare nesupervizata

18. Cum se utilizeaza retelele Learning Vector Quantization pentru clasificare?

- a) spatiul vectorilor de intrare se divide in regiuni disjuncte de arii egale, delimitate de hiperplane; un vector de intrare este etichetat cu clasa regiunii in care este inclus
- b) spatiul vectorilor de intrare se divide in regiuni disjuncte, fiecare cu un vector prototip; un vector de intrare este etichetat cu clasa celui mai apropiat vector prototip**
- c) LVQ nu se pot folosi pentru clasificare deoarece sunt metode nesupervizate

19. Cum se utilizeaza retelele Learning Vector Quantization pentru compresie de date?

- a) spatiul vectorilor de intrare se divide in regiuni disjuncte, fiecare cu un vector prototip; un vector de intrare este inlocuit cu indexul vectorului prototip cel mai apropiat**
- b) LVQ nu se pot folosi pentru compresie de date, intrucat nu este o metoda de compresie exacta; ar trebui sa existe la fel de multe prototipuri cat vectori de intrare
- c) spatiul vectorilor de intrare se divide in regiuni disjuncte de arii egale, delimitate de hiperplane; un vector de intrare este inlocuit cu indexul vectorului prototip cel mai apropiat

20. Dacă  $W_i(t)$  este vectorul de ponderi al neuronului câștigător la momentul  $t$ ,  $X_j$  este vectorul de intrare,  $\alpha$  este rata de învățare, cum se realizează actualizarea ponderilor în procesul de învățare al unei rețele Learning Vector Quantization?

a)  $W_i(t+1) = W_i(t) + \alpha * (X_j - W_i(t))$

b)  $W_i(t+1) = W_i(t) - \alpha * (X_j - W_i(t))$  dacă  $X_j$  și  $W_i$  aparțin aceleiași clase și  $W_i(t+1) = W_i(t) + \alpha * (X_j - W_i(t))$  dacă  $X_j$  și  $W_i$  aparțin la clase diferite

**c)  $W_i(t+1) = W_i(t) + \alpha * (X_j - W_i(t))$  dacă  $X_j$  și  $W_i$  aparțin aceleiași clase și  $W_i(t+1) = W_i(t) - \alpha * (X_j - W_i(t))$  dacă  $X_j$  și  $W_i$  aparțin la clase diferite**

21. Ce aduce în plus modelul LQV2 față de LVQ?

a) se consideră și următorul cel mai apropiat vecin, în condițiile în care cel mai apropiat vecin  $W_i$  aparține aceleiași clase cu vectorul de intrare  $X$ , următorul cel mai apropiat vecin  $W_j$  aparține altei clase în comparație cu  $X$  iar  $X$  aparține unei ferestre definite de planul bisector al segmentului

care unește  $W_i$  și  $W_j$

b) se consideră și următorul cel mai apropiat vecin, în condițiile în care cel mai apropiat vecin  $W_i$  și următorul cel mai apropiat vecin  $W_j$  aparțin aceleiași clase cu vectorul de intrare  $X$

**c) se consideră și următorul cel mai apropiat vecin, în condițiile în care cel mai apropiat vecin  $W_i$  aparține altei clase în comparație cu vectorul de intrare  $X$ , următorul cel mai apropiat vecin  $W_j$  aparține aceleiași clase cu  $X$  iar  $X$  aparține unei ferestre definite de planul**

**bisector al segmentului care unește  $W_i$  și  $W_j$**

22. Ce fenomene cognitive consideră Grossberg ca explică modelele Instar și Outstar?

**a) învățare pavloviană, învățare hebbiană, repetarea îmbunătățește rezultatele, repetarea excesivă nu este necesară**

b) la intrări apropiate corespund ieșiri apropiate iar la ieșiri mult diferite corespund ieșiri mult diferite

c) învățare pavloviană, învățare hebbiană, cu cât se repetă mai mult rezultatele sunt mai bune

23. Care sunt avantajele modelului CMAC (cerebellar model articulation controller)?

a) procedura de hashing care mapează memoria virtuală în memorie reală este întotdeauna exactă, fapt ce accelerează învățarea

b) proprietatea de generalizare locală facilitează descoperirea de relații globale în spațiul intrărilor

**c) este un model viabil pentru sarcini de control complexe, algoritmul de învățare este mult mai rapid decât backpropagation**

24. Care este cel mai utilizat tip de rețea neuronală?

a) hartile cu auto-organizare

**b) perceptronul multistrat**

c) memoria asociativă

25. Care este structura unui perceptron multistrat?

a) mai multe straturi de neuroni interconectați, fiecare din ei calculează suma ponderată a

intrarilor, rezultatului i se aplica o functie de activare neliniara si apoi se adauga un termen constant

b) mai multe straturi de neuroni interconectati, asupra fiecarei intrari a neuronilor se aplica o functie de activare neliniara, se face suma ponderata a rezultatelor la care se adauga un termen constant

**c) mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor la care se adauga un termen constant si apoi rezultatului i se aplica o functie de activare neliniara**

26. Ce inseamna conectarea feedforward a neuronilor intr-un perceptron multistrat?

a) interconectarile formeaza bucle

b) neuronii sunt total conectati (fiecare cu fiecare)

**c) interconectarile nu formeaza bucle**

27. Cum se numesc retelele perceptron multistrat in care interconectarile formeaza una sau mai multe bucle?

**a) recurente**

b) feedforward

c) stratificate

28. Care este expresia functiei sigmoide unipolare (functia logistica)?

**a)  $S(s) = 1 / (1 + \exp(-s))$**

b)  $S(s) = 1 / (1 + \exp(s))$

c)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$

29. Care este expresia functiei sigmoide bipolare (tangenta hiperbolica)?

**a)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$**

b)  $S(s) = 1 / (1 + \exp(-s))$

c)  $S(s) = 1 / (\exp(s) + \exp(-s))$

30. Cum se realizeaza antrenarea unui perceptron multistrat?

a) prin modificarea modului de interconectare a neuronilor

**b) prin modificarea ponderilor conexiunilor si a valorilor prag ale neuronilor**

c) prin modificarea valorilor neuronilor care retin informatiile

31. Ce sunt straturile ascunse ale unui perceptron multistrat?

a) straturi interne retelei care nu participa direct la antrenare

**b) straturi interne retelei, nici de intrare si nici de iesire**

c) straturi interne retelei care sunt parcurse uneori de semnale dinspre intrare spre iesire

32. De ce perceptronul multistrat nu are functii de activare liniare?

a) functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat

b) datorita progresului tehnologic, functiile exponentiale care pot fi utilizate acum pentru perceptronul multistrat erau imposibil de utilizat practic in anii '50 pentru perceptronul cu un singur strat



**c) funcțiile liniare nu asigură o creștere a puterii de calcul în raport cu rețelele cu un singur strat, deoarece o funcție liniară de funcții liniare este tot o funcție liniară**

33. Cine a introdus în 1986 algoritmul backpropagation ca metoda generală de antrenare a perceptronilor multistrat?

a) Werbos

**b) Rumelhart, Hinton și Williams**

c) Bryson și Ho

34. Care este ideea care stă la baza algoritmului backpropagation?

**a) minimizarea erorii (diferenței dintre ieșirea dorită și ieșirea reală) în raport cu ponderile prin metoda gradientului descendent**

b) minimizarea ponderilor conexiunilor prin metoda gradientului descendent

c) minimizarea ponderilor conexiunilor în două etape: una de propagare înainte și una de propagare înapoi

35. Pentru aplicarea algoritmului backpropagation, care este derivata funcției sigmoide unipolare?

a)  $S'(s) = (1 - S(s)) * (1 + S(s))$

**b)  $S'(s) = S(s) * (1 - S(s))$**

c)  $S'(s) = 1 - S(s)$

36. Pentru aplicarea algoritmului backpropagation, care este derivata funcției tangente hiperbolice?

**a)  $S'(s) = (1 - S(s)) * (1 + S(s))$**

b)  $S'(s) = S(s) * (1 - S(s))$

c)  $S'(s) = 1 + S(s)$

37. În algoritmul backpropagation, cum se calculează gradientii erorilor pentru neuronii din stratul de ieșire ( $\delta_k$ ), dacă  $S'$  este derivata funcției de activare,  $e_k(t)$  este eroarea la momentul  $t$

iar  $w_{jk}(p)$  sunt ponderile conexiunilor dintre stratul ascuns și cel de ieșire?

a)  $\delta_k(t) = S' / e_k(t)$

**b)  $\delta_k(t) = S' * e_k(t)$**

c)  $\delta_k(t) = e_k(t) * w_{jk}(t)$

38. În algoritmul backpropagation, cum se calculează gradientii erorilor pentru neuronii din stratul ascuns ( $\delta_j$ ), dacă  $S'$  este derivata funcției de activare,  $\delta_k$  sunt gradientii erorilor pentru neuronii din stratul de ieșire,  $w_{jk}(t)$  sunt ponderile conexiunilor dintre stratul ascuns și cel de

ieșire la momentul  $t$  iar  $l$  este numărul de ieșiri ale rețelei (notăm suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum\{i, 1, n\} (a_i)$ )?

**a)  $\delta_j(t) = S' * \sum\{k, 1, l\} (\delta_k(t) * w_{jk}(t))$**

b)  $\delta_j(t) = S' * w_{jk}(t)$

c)  $\delta_j(t) = S' * \delta_k(t)$

39. Ce reprezintă algoritmul backpropagation stohastic?

- a) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face numai pentru anumiti vectori de intrare, dupa o distributie de probabilitate
- b) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor nu se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, ci in timp real, dupa prezentarea fiecarui vector de intrare**
- c) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, si nu in timp real, dupa prezentarea fiecarui vector de intrare

40. Care este impedimentul cel mai important al algoritmului backpropagation?

- a) nu converge pentru probleme complexe de dimensiuni mari
- b) deoarece se bazeaza pe metoda gradientului descendent, poate converge intr-un minim local al functiei de eroare pe care incearca sa o minimizeze**
- c) restrictiile impuse pentru valorile ratei de invatare

41. Cum se justifica proprietatea perceptronului multistrat de "aproximator universal"?

- a) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie polinomiala
- b) s-a demonstrat ca o retea (posibil infinita) cu un singur strat ascuns este capabila sa aproximeze orice functie continua**
- c) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie derivabila datorita algoritmului backpropagation

42. Prin ce se caracterizeaza metoda momentului de accelerare a invatarii cu algoritmul backpropagation?

- a) consta in micsorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie
- b) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor
- c) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective**

43. Prin ce se caracterizeaza metoda ratei de invatare adaptive pentru accelerarea invatarii cu algoritmul backpropagation?

- a) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor**
- b) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective
- c) consta in micsorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

44. Ce se intelege prin capacitatea de a generaliza a unei retele neuronale?

- a) reseaua trebuie sa dea rezultate cat mai bune pentru instantele folosite la antrenare
- b) reseaua trebuie sa dea rezultate bune indiferent de numarul de instante folosite la antrenare
- c) reseaua trebuie sa dea rezultate bune nu numai pentru instantele folosite la antrenare, ci si pentru instante noi**

45. Care ar fi condițiile pentru ca o rețea să generalizeze bine?

**a) să aibă o topologie adecvată cu cât mai puține ponderi și totuși să dea rezultate bune la antrenare, exploatând astfel regularitățile problemei**

b) să aibă un număr suficient de ponderi astfel încât să dea rezultate cât mai bune la antrenare

c) să aibă un număr de ponderi egal cu o zecime din produsul dintre numărul de vectori de antrenare și numărul de ieșiri

46. Cum se realizează validarea încrucișată ("cross-validation")?

a) se validează rezultatele antrenării prin adăugarea unui pas suplimentar în care se evaluează performanțele rețelei pe o mulțime independentă de vectori după încheierea antrenării

b) se efectuează mai multe antrenări succesive pe mulțimi de antrenare diferite și se aleg ponderile care dau cele mai bune rezultate pentru vectorii de validare aparținând intersecției mulțimilor considerate

**c) în timpul antrenării, se analizează performanțele rețelei pe o mulțime de vectori de validare diferită de mulțimea vectorilor de antrenare; antrenarea se oprește în momentul în**

**care eroarea pe setul de validare nu mai scade, chiar dacă eroarea pe setul de antrenare ar putea scădea în continuare**

47. Cum se îmbunătățește capacitatea de generalizare a unui perceptron multistrat prin simplificarea rețelei ("network pruning") și tehnicile constructive?

**a) metoda simplificării rețelei pornește de la o rețea mare și apoi elimină succesiv cele mai puțin importante conexiuni; tehnicile constructive pornesc de la o rețea de dimensiuni mici și**

**ii adaugă unități până când performanțele devin acceptabile**

b) aceste tehnici nu au nici o influență asupra capacității de generalizare a unui perceptron multistrat

c) metoda simplificării rețelei pornește de la o rețea mare și apoi elimină succesiv cele mai puțin importante unități; tehnicile constructive pornesc de la o rețea de dimensiuni mici și ii adaugă conexiuni până când performanțele devin acceptabile

48. În contextul regularizării unui perceptron multistrat, care este rolul termenului de regularizare  $E_{reg}$  din expresia funcției de cost  $E_{total} = E + \lambda * E_{reg}$ ?

a) determină rețeaua să producă funcții de aproximare cât mai instabile (cu variabilitate mare), deoarece acestea au șanse mai mari să găsească soluția

**b) determină rețeaua să producă funcții de aproximare cât mai netede, deoarece acestea sunt mai regulate (au variabilitate mai mică) și deci pot produce generalizări mai bune**

c) scade erorile de antrenare prin adăugarea unei rate de învățare suplimentare  $\lambda$

49. În contextul regularizării unui perceptron multistrat, în expresia funcției de cost  $E_{total} = E + \lambda * E_{reg}$ , să presupunem că termenul de regularizare  $E_{reg}$  este suma pătratelor ponderilor rețelei. Cum se justifică o asemenea alegere?

**a) ponderile mai mici tind să producă funcții care se schimbă mai greu și deci mai netede**

b) această expresie pentru  $E_{reg}$  nu are nici o influență asupra funcției approximate

c) ponderile mai mari tind să producă funcții care se schimbă mai rapid și deci cu șanse mai mari

sa produca functii netede

50. In contextul regularizarii unui perceptron multistrat, in expresia functiei de cost  $E_{total} = E + \lambda * E_{reg}$ , sa presupunem ca termenul de regularizare  $E_{reg}$  este suma valorilor absolute ale ponderilor retelei. Cum se justifica o asemenea alegere?

**a) in componenta gradientilor functiei de cost apare un termen proportional cu semnul ponderilor, ceea ce asigura atat regularizarea, cat si punerea pe 0 a ponderilor mai putin importante**

b) in componenta gradientilor functiei de cost apare un termen proportional cu semnul ponderilor, ceea ce asigura atat regularizarea, cat si punerea in valoare a ponderilor mai mici, care

in mod normal nu si-ar fi putut aduce contributia la solutie

c) in componenta gradientilor functiei de cost apare un termen invers proportional cu semnul ponderilor, ceea ce asigura atat regularizarea, cat si punerea pe 0 a ponderilor mai putin importante

51. Prin ce se caracterizeaza o retea neuronală recurentă secvențială?

**a) rețeaua are "memorie" iar ieșirea dorită pentru fiecare vector de intrare depinde nu numai de acel vector ci și de cei anteriori**

b) rețeaua nu are "memorie" iar ieșirea dorită pentru fiecare vector de intrare depinde numai de acel vector și nu și de cei anteriori

c) rețeaua nu are bucle iar propagarea semnalului se face secvențial, din strat în strat, de la intrare către ieșire

52. Care este diferența dintre variantele algoritmului backpropagation pentru rețele feed-forward și recurente?

a) nu există nici o diferență

**b) în cazul rețelelor recurente, după aplicarea unui vector de intrare și după retropropagarea unui set de erori, rețeaua trebuie lăsată să se stabilizeze**

c) în cazul rețelelor feed-forward, după aplicarea unui vector de intrare și după retropropagarea unui set de erori, rețeaua trebuie lăsată să se stabilizeze

53. Care sunt condițiile care trebuie respectate pentru a garanta stabilitatea la antrenare a unei rețele recurente de tip perceptron multistrat?

a) folosirea unor funcții de activare cât mai simple (de exemplu prag), evitarea simetriei ponderilor și implementarea dinamicii rețelei

**b) folosirea unor neliniarități crescătoare și marginite (de exemplu sigmoide), păstrarea simetriei ponderilor și implementarea dinamicii rețelei**

c) folosirea unor funcții liniare, evitarea simetriei ponderilor și implementarea dinamicii rețelei

54. Cum funcționează rețelele recurente secvențiale de timp discret?

a) intrările rețelei se modifică numai la momente discrete de timp, ieșirile unităților de timp discret și ale unităților instantanee urmăresc imediat variațiile intrărilor lor

b) intrările rețelei se modifică la orice moment de timp, ieșirile unităților de timp discret se modifică numai la momente discrete de timp iar ieșirile unităților instantanee urmăresc imediat variațiile intrărilor lor

**c) intrările rețelei se modifică numai la momente discrete de timp, ieșirile unităților de**

**timp discret se modifica sincron cu intrarile iar iesirile unitatilor instantanee urmaresc imediat variatiile intrarilor lor**

55. De ce trebuie sa existe in fiecare bucla a unei retele recurente de tip perceptron multistrat cel putin o unitate de timp discret?

a) daca bucla ar fi formata numai din unitati instantanee, reseaua ar deveni un perceptron feedforward

b) aceasta conditie nu este obligatorie

**c) daca bucla ar fi formata numai din unitati instantanee, ar aparea o secventa infinita de actualizari**

56. Care este avantajul retelelor recurente secventiale de timp discret?

a) pot fi expandate in retele de tip perceptron cu un singur strat

**b) pot fi reduse la retele feed-forward si antrenate cu algoritmul backpropagation standard**

c) necesita mai putina memorie pentru antrenare

57. Cum pot fi clasificate memoriile asociative din punct de vedere al asociatiilor stocate:

**a) autoasociative si heteroasociative**

b) asociative si neasociative

c) statice si recurente

58. Care este trasatura cea mai importanta a unei memorii asociative cu feedback?

a) sistemul dinamic rezultat trebuie sa aiba un comportament oscilant la inceputul invatarii si stabil spre final

**b) sistemul dinamic rezultat trebuie sa aiba un comportament asimptotic stabil**

c) sistemul dinamic rezultat trebuie sa aiba un comportament asimptotic oscilant

59. Intrarea intr-o memorie asociativa se numeste:

**a) cheie de memorie**

b) model regasit

c) asociatie de memorie

60. Ce trebuie sa realizeze o memorie asociativa?

**a) o mapare  $y_k = G(x_k)$  pentru toate perechile de intrare-iesire  $(x_k, y_k)$ , maparea sa fie toleranta la zgomote si sa poata corecta eventualele erori**

b) o mapare  $y_k = G(x_k)$  pentru toate perechile de intrare-iesire  $(x_k, y_k)$ , maparea sa fie exacta numai pentru perechile specificate la antrenare

c) o mapare  $y_k = G(x_k)$  pentru unele perechi de intrare-iesire  $(x_k, y_k)$ , maparea sa fie toleranta la zgomote si sa poata corecta eventualele erori

61. Ce este o "amintire fundamentala" a unei memorii asociative?

a) o valoare a iesirii de tip atractor pentru o gama larga de intrari

**b) o pereche intrare-iesire  $(x_k, y_k)$  cu care se face antrenarea**

c) cea mai probabila intrare a memoriei asociative

62. Ce este o memorie autoasociativa?

**a) o memorie asociativa care reda mapari intrare-iesire de forma  $(x_k, x_k)$**

- b) o memorie asociativa care reda mapari intrare-iesire de forma  $(y_k, x_k)$
- c) o memorie asociativa care reda mapari intrare-iesire de forma  $(x_k, y_k)$

63. Ce este o "amintire falsa" (spurious memory)?

- a) o asociatie gresita stocata in memorie in mod intentionat pentru a asigura stabilitatea sistemului dinamic
- b) o asociatie gresita care trebuie identificata si eliminata
- c) o asociatie stocata in memorie in mod neintentionat, care nu face parte din multimea de amintiri fundamentale**

64. De ce sunt mai puternice arhitecturile de memorie autoasociativa recurente decat cele feedforward?

- a) arhitecturile recurente au o topologie mult mai complexa decat cele feedforward
- b) o arhitectura feedforward este incapabila de a elimina zgomotul din vectorii de intrare
- c) intr-o arhitectura feedforward, la prezentarea unui vector de intrare numai o parte din eventualul zgomot va fi eliminat**

65. Daca  $X$  sunt vectorii de intrare,  $W$  este matricea de interconectare,  $F$  este vectorul functiilor de activare iar  $t$  sunt momentele de timp discrete, care este ecuatia unei memorii autoasociative dinamice (DAM)?

- a)  $X(t+1) = F(W * X(t))$**
- b)  $X(t+1) = W * F(X(t))$
- c)  $X(t+1) = W * X(t)$

66. Daca  $X$  sunt vectorii de intrare,  $Y$  sunt vectorii de iesire,  $W1$  si  $W2$  sunt matricele de interconectare,  $F$  este vectorul functiilor de activare iar  $t$  sunt momentele de timp discrete, care sunt ecuatiile unei memorii heteroasociative?

- a)  $Y(t+1) = F(W1 * X(t))$  si  $X(t+1) = F(W2 * Y(t))$**
- b)  $X(t+1) = F(W1 * X(t))$  si  $Y(t+1) = F(W2 * Y(t))$
- c)  $Y(t+1) = W1 * F(X(t))$  si  $X(t+1) = W2 * F(Y(t))$

67. Ce caracteristici are o memorie asociativa dinamica de inalta performanta?

- a) bazine mari de atractie in jurul amintirilor fundamentale, un numar relativ mic de amintiri false cu bazine mici de atractie, o stare stabila implicita de indecizie si fara regiuni cu oscilatii**
- b) bazine mici de atractie in jurul amintirilor fundamentale, un numar relativ mic de amintiri false cu bazine mari de atractie, o stare stabila implicita de indecizie si putine regiuni cu oscilatii
- c) capacitate mare de stocare indiferent de restul parametrilor

68. Cum se determina matricea de interconectare  $W$  intr-o memorie asociativa liniara (notam  $A$  transpus cu  $A^T$ )?

- a)  $W = (X * Y)^T$
- b)  $W = Y * X^T$**
- c)  $W = X * Y^T$

69. Care sunt parametrii unei memorii asociative neliniare (notam  $A$  transpus cu  $A^T$  si suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum\{i,1,n\} (a_i)$  )?

a)  $Y = F(W * X^T)$  si  $W = (1/n) * \sum\{k,1,m\} (y_k * x_k)$

**b)  $Y = F(W * X)$  si  $W = (1/n) * \sum\{k,1,m\} (y_k * x_k^T)$**

c)  $Y = F(W * X^T)$  si  $W = (1/n) * \sum\{k,1,m\} (y_k * x_k^T)$

70. Care este reteta de inregistrare a proiectiei (projection recording recipe) pentru o memorie asociativa liniara optima OLAM (notam  $A$  transpus cu  $A^T$  si inversa lui  $A$  cu  $A^{(-1)}$  )?

a)  $W^* = Y * (X^T * X)^{(-1)} * Y^T$

**b)  $W^* = Y * (X^T * X)^{(-1)} * X^T$**

c)  $W^* = X * (Y^T * Y)^{(-1)} * Y^T$

71. Cand este stabila retea Hopfield?

a) cand matricea de interconectare este simetrica si functia de activare este continua si monoton descrescatoare

**b) cand matricea de interconectare este simetrica si functia de activare este continua si monoton crescatoare**

c) cand matricea de interconectare este antisimetrica si functia de activare este continua si monoton crescatoare

72. Care este ecuatia caracteristica modelului Hopfield discret (notam suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum\{i,1,n\} (a_i)$  )?

a)  $x_i(k+1) = \text{sgn}(w_{ij} * x_j(k)) + I_i$

**b)  $x_i(k+1) = \text{sgn}(\sum\{j,1,n\} (w_{ij} * x_j(k))) + I_i$**

c)  $x_i(k+1) = \sum\{j,1,n\} (w_{ij} * x_j(k)) + I_i$

73. Memoria asociativa bidirectionala (BAM) propusa de Kosko, cu constrangerea  $W_2^T = W_1 = W$  ( $W_2^T$  noteaza  $W_2$  transpus):

a) este stabila pentru orice matrice reala de interconectare  $W$  si converge intotdeauna catre minimul functiei sale de energie

b) este stabila pentru orice matrice simetrica de interconectare  $W$  si converge intotdeauna catre minimul functiei sale de energie

**c) este stabila pentru orice matrice reala de interconectare  $W$  si converge intotdeauna catre un minim local al functiei sale de energie**

74. Care sunt pasii principali ai metodologiei de calire simulata?

a) generarea determinista a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru asigurarea convergentei catre solutie

b) generarea aleatorie a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru cresterea distantei intre doua stari succesive acoperind zone din ce in ce mai extinse in spatiul solutiilor

**c) generarea aleatorie a unei noi stari, testarea validitatii acesteia si un program de "racire" pentru stabilizarea intr-o anumita regiune**

75. Care este densitatea de probabilitate folosita de masina Boltzman pentru a genera deplasarea dintre starea veche si cea noua intr-o retea neuronală stohastica?

- a)  $GT(x' | x' = x + X) = \exp(-X^2 / T)$
- b)  $GT(x' | x' = x + X) = (1/2) * \pi * T^{(1/2)} * \exp(-X^2 / T)$**
- c)  $GT(x' | x' = x + X) = 1/2 * \pi * X^2 / T$

76. Care este densitatea de probabilitate Cauchy pentru a genera deplasarea dintre starea veche si cea noua intr-o retea neuronală stohastica?

- a)  $GT(x' | x' = x + X) = T * \pi * T^2 + |X|^2$
- b)  $GT(x' | x' = x + X) = T^2 / \pi * (T + |X|^2)$
- c)  $GT(x' | x' = x + X) = T / \pi * (T^2 + |X|^2)$**

77. Care fenomen natural este analog cu ideea de calire simulata?

- a) evaporarea rapida a apei
- b) inghetarea treptata a apei pentru a se forma o structura cristalina**
- c) inghetarea rapida a apei

78. In calirea simulata, cunoscand diferenta de energie dintre doua stari  $dE = E2 - E1$ , cand este posibila o tranzitie din starea  $E1$  in starea  $E2$ ?

- a) cand  $dE > 0$  sau cu probabilitatea  $PT = 1 / (1 + \exp(-dE / T))$  daca  $dE < 0$
- b) cand  $dE > 0$  sau cu probabilitatea  $PT = 1 + \exp(-dE / T)$  daca  $dE < 0$
- c) cand  $dE < 0$  sau cu probabilitatea  $PT = 1 / (1 + \exp(-dE / T))$  daca  $dE > 0$**

79. In calirea simulata, cand se face o tranzitie intre doua stari  $E1$  si  $E2$ ?

- a) daca  $E1$  si  $E2$  sunt la fel de bune, se face tranzitia in  $E2$  cu o anumita probabilitate, altfel nu se efectueaza tranzitia
- b) daca  $E2$  este o stare mai buna, se trece in ea sigur, daca  $E2$  este mai proasta, se trece in ea cu o anumita probabilitate descrescatoare in timp**
- c) daca  $E2$  este o stare mai buna, se trece in ea cu o anumita probabilitate descrescatoare in timp, daca  $E2$  este mai proasta, nu se efectueaza tranzitia

80. Care este programul de racire pentru calirea simulata folosind un proces aleatoriu gaussian?

- a)  $T = T0 / \exp(1+t)$
- b)  $T = T0 / \log(1+t)$**
- c)  $T = T0 / (1+t)$

81. Cum poate fi accelerat programul de racire in calirea simulata fata de procesul gaussian aleatoriu folosind zgomot colorat Cauchy?

- a)  $T = T0 / (1+t)$**
- b)  $T = T0 / \log(1+t)$
- c)  $T = T0 / \exp(1+t)$

82. Care este relatia dintre zgomot si racire intr-un proces de calire simulata?

- a) zgomotul asigura omogenitatea starilor vizitate iar temperatura determina cat de apropiate pot fi doua stari succesive; pe masura ce temperatura scade, cautarea se va face intr-o zona din ce in ce mai extinsa asigurand gasirea solutiei
- b) zgomotul asigura diversitatea starilor vizitate iar temperatura determina cat de**



**departate pot fi doua stari succesive; pe masura ce temperatura scade, cautarea se va localiza**

**intr-o anumita regiune si se spera ca va converge catre solutie**

c) zgomotul asigura diversitatea starilor vizitate iar temperatura determina cat de departate pot fi doua stari succesive; pe masura ce temperatura creste, cautarea se va localiza intr-o anumita regiune

si se spera ca va converge catre solutie

83. In contextul clasificarii, care este diferenta dintre un neuron de tip perceptron si un neuron RBF?

a) perceptronul si neuronul RBF se comporta la fel in contextul clasificarii

**b) perceptronul imparte spatiul trasaturilor in hiperplane, neuronul RBF defineste un hipervolum sau o hipersfera**

c) neuronul RBF imparte spatiul trasaturilor in hiperplane, perceptronul defineste un hipervolum sau o hipersfera

84. Retelele RBF:

a) pot realiza clasificari dar nu si aproximari functionale

**b) pot realiza atat clasificari cat si aproximari functionale**

c) pot realiza aproximari functionale dar nu si clasificari

85. Care este topologia unei retele RBF?

a) o retea RBF este formata din doua straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni iar al doilea contine neuroni RBF

**b) o retea RBF este formata din trei straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni, al doilea contine neuroni RBF iar al treilea contine neuroni cu functii de activare liniare**

c) o retea RBF este formata din doua straturi: primul contine neuroni RBF care prelucreaza intrarile iar al doilea contine neuroni RBF cu functii de activare liniare care transmit iesirile

86. Care este iesirea unei retele RBF gaussiene, daca notam suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum_{i=1}^n a_i$ ,  $w_j$  sunt ponderile conexiunilor, centrele sunt exemplarele  $x_j$ , dispersiile sunt  $\sigma_j$  iar  $N$  este numarul de vectori de antrenare?

a)  $f(x) = \sum_{j=1}^N (w_j * (-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$

**b)  $f(x) = \sum_{j=1}^N (w_j * \exp(-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$**

c)  $f(x) = \sum_{j=1}^N (w_j * \exp(-1 / \sigma_j * \|x - x_j\|^2))$

87. Ce parametri ai unei retele RBF gaussiene sunt ajustati prin invatare?

a) numai centrele gaussianelor, prin metode supervizate sau nesupervizate

b) numai ponderile  $w_i$  sunt ajustate, ceilalti parametri sunt fiksi

**c) ponderile  $w_i$ , dispersiile  $\sigma_i$  iar numarul de neuroni RBF poate fi redus prin selectia unor exemplare reprezentative din vectorii de antrenare**

88. Ce este o retea RBF generalizata?

**a) o retea in care numarul de unitati RBF este mult redus fata de numarul de vectori de antrenare**

b) o retea RBF care poate realiza si aproximari functionale, nu numai clasificari

c) o retea RBF in care iesirile sunt reale, nu binare

89. Cum se calculeaza ponderile unei retele RBF cu selectie a centrelor?

**a) prin metoda celor mai mici patrate**

b) cu algoritmul backpropagation

c) prin metoda divide et impera

90. Care din urmatoarele metode nu poate fi folosita pentru adaptarea latimii functiei de baza radiala intr-o retea RBF?

a) metoda gradientului descendent al functiei de cost

b) metoda analizei componentelor principale (determinarea valorilor proprii prin procedura Gram-Schmidt)

**c) metoda celor mai mici patrate ortogonale nesupervizate**

91. Ce sunt estimatorii densitatii nucleului (kernel density estimators)?

**a) retele neuronale RBF pentru estimarea functiilor de densitate de probabilitate**

b) perceptroni multistrat pentru estimarea functiilor de densitate de probabilitate

c) retele neuronale stohastice pentru estimarea functiilor de densitate de probabilitate

92. Care sunt avantajele metodelor de clasificare bazate pe nucleu (kernel classifiers)?

a) nu necesita reducerea numarului de nuclee sau ajustarea latimii nucleelelor

**b) invatare simpla, posibilitatea convergentei catre clasificatorul bayesian optim, care poate ramane optim fara reantrenare**

c) estimarea exacta a functiilor de densitate de probabilitate ale claselor

93. Care este topologia unei retele neuronale de energie Coulomb restransa?

**a) trei straturi: un strat de intrare care retransmite intrarile, un strat de prototipuri si un strat de decizie**

b) trei straturi: un strat de intrare cu functii prag aplicate intrarilor, un strat de prototipuri si un strat de decizie

c) trei straturi: un strat de intrare cu functii prag aplicate intrarilor, un strat de prototipuri cu functii de activare sigmoide si un strat de decizie

94. Intr-o retea neuronală de energie Coulomb restransa, care este rolul stratului de prototipuri?

a) prototipizarea vectorilor de intrare prin eliminarea diferentelor semnificative fata de vectorii de antrenare cei mai frecventi

b) clasificarea vectorilor de antrenare printr-o multime de hipervolume cu raza fixe pentru vectorii de intrare cei mai frecventi

**c) aproximarea claselor printr-o superpozitie de hipervolume cu anumite raze in spatiul trasaturilor**

95. Ce se intampla daca intr-o retea neuronală de energie Coulomb restransa raza initiala  $\lambda_0$  este prea mica?

**a) numarul de neuroni prototip devine egal cu numarul de vectori de antrenare**

b) invatarea tinde spre supra-potrivire (over-fitting)

c) invatarea tinde spre sub-potrivire (under-fitting)

96. Care este topologia unui arbore neuronal?

**a) un strat de intrare, un strat de neuroni binari (cu un numar de neuroni egal cu numarul de noduri ale arborelui), un strat de neuroni logici SI (cu un numar de neuroni egal cu numarul de frunze ale arborelui) si un un strat de neuroni logici SAU (cu un numar de neuroni egal cu numarul de clase)**

b) un strat de intrare, un strat de neuroni binari (cu un numar de neuroni egal cu numarul de frunze ale arborelui), un strat de neuroni logici SI (cu un numar de neuroni egal cu numarul de clase) si un un strat de neuroni logici SAU (cu un numar de neuroni egal cu numarul de noduri ale arborelui)

c) un strat de intrare, un strat de neuroni binari (cu un numar de neuroni egal cu numarul de frunze ale arborelui), un strat de neuroni logici SI (cu un numar de neuroni egal cu numarul de noduri ale arborelui) si un un strat de neuroni logici SAU (cu un numar de neuroni egal cu numarul de clase)

1. Cine a propus modelul perceptronului la sfarsitul anilor 50?

a) Marvin Minsky

b) Teuvo Kohonen

**c) Frank Rosenblatt**

2. Cine sunt autorii cartii "Perceptronii" (1969) in care se evidentia imposibilitatea perceptronului de a invata functii neseperabile liniar?

**a) Minsky si Papert**

b) Rosenblatt si Hopfield

c) Rumelhart si Hinton

3. Functia XOR:

**a) poate fi invatata de un perceptron multistrat**

b) nu poate fi invatata de o retea neuronală

c) poate fi invatata de un perceptron cu un singur strat

4. Care este scopul unui perceptron cu un singur strat?

a) clasificator pentru n clase prin invatarea clasei de apartenenta a fiecărei instante

**b) clasificator pentru 2 clase prin gasirea unui hiperplan care le separa**

c) clasificator pentru 2 clase prin gasirea unui hipervolum care le separa

? 5. Care este iesirea unui perceptron cu un singur strat?

a) functia sigmoida aplicata intrarii cu ponderea cea mai mare

**b) functia Heaviside aplicata sumei intrarilor ponderate**

c) functia sigmoida aplicata sumei intrarilor ponderate

6. Daca  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de invatare,  $t$  este iesirea dorita,  $a$  este iesirea reala si  $x_i$  este intrarea  $i$ , care este **regula de actualizare** a ponderilor unui perceptron cu un singur strat?

a)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$

**b)  $w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$  // pastrez pasul anterior**

c)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$

15. Ce inseamna principiul celei mai mici perturbatii in procesul de invatare a unei retele neuronale?

**a) ponderile trebuie ajustate astfel incat sa se reduca eroarea la iesire pentru vectorul de intrare curent cu o perturbatie minima asupra raspunsurilor deja invatate**

b) ponderile trebuie ajustate astfel incat eroarea la iesire sa fie redusa cu o cantitate cat mai mica astfel incat sa nu fie perturbata reseaua

c) ponderile trebuie ajustate cu o cantitate cat mai mica astfel incat sa nu fie perturbata reseaua

24. Care este cel mai utilizat tip de retea neuronală?

a) hartile cu auto-organizare

**b) perceptronul multistrat**

c) memoria asociativa

25. Care este structura unui perceptron multistrat?

a) mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor, rezultatului i se aplica o functie de activare neliniara si apoi se adauga un termen constant

b) mai multe straturi de neuroni interconectati, asupra fiecarei intrari a neuronilor se aplica o functie de activare neliniara, se face suma ponderata a rezultatelor la care se adauga un termen constant

**c) mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor la care se adauga un termen constant si apoi rezultatului i se aplica o functie de activare neliniara**

26. Ce inseamna conectarea feedforward a neuronilor intr-un perceptron multistrat?

a) interconectarile formeaza bucle

b) neuronii sunt total conectati (fiecare cu fiecare)

**c) interconectarile nu formeaza bucle**

27. Cum se numesc retelele perceptron multistrat in care interconectarile formeaza una sau mai multe bucle?

**a) recurente**

b) feedforward

c) stratificate

28. Care este expresia functiei sigmoide unipolare (functia logistica)?

**a)  $S(s) = 1 / (1 + \exp(-s))$**

b)  $S(s) = 1 / (1 + \exp(s))$

c)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$

29. Care este expresia functiei sigmoide bipolare (tangenta hiperbolica)?

**a)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$**

b)  $S(s) = 1 / (1 + \exp(-s))$

c)  $S(s) = 1 / (\exp(s) + \exp(-s))$

30. Cum se realizeaza antrenarea unui perceptron multistrat?

- a) prin modificarea modului de interconectare a neuronilor
- b) prin modificarea ponderilor conexiunilor si a valorilor prag ale neuronilor**
- c) prin modificarea valorilor neuronilor care retin informatiile

31. Ce sunt straturile ascunse ale unui perceptron multistrat?

- a) straturi interne retelei care nu participa direct la antrenare
- b) straturi interne retelei, nici de intrare si nici de iesire**
- c) straturi interne retelei care sunt parcurse uneori de semnale dinspre intrare spre iesire

32. De ce perceptronul multistrat nu are functii de activare liniare?

- a) functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat
- b) datorita progresului tehnologic, functiile exponentiale care pot fi utilizate acum pentru perceptronul multistrat erau imposibil de utilizat practic in anii '50 pentru perceptronul cu un singur strat
- c) functiile liniare nu asigura o crestere a puterii de calcul in raport cu retelele cu un singur strat, deoarece o functie liniara de functii liniare este tot o functie liniara**

33. Cine a introdus in 1986 algoritmul backpropagation ca metoda generala de antrenare a perceptronilor multistrat?

- a) Werbos
- b) Rumelhart, Hinton si Williams**
- c) Bryson si Ho

34. Care este ideea care sta la baza algoritmului backpropagation?

- a) minimizarea erorii (diferentei dintre iesirea dorita si iesirea reala) in raport cu ponderile prin metoda gradientului descendent**
- b) minimizarea ponderilor conexiunilor prin metoda gradientului descendent
- c) minimizarea ponderilor conexiunilor in doua etape: una de propagare inainte si una de propagare inapoi

35. Pentru aplicarea algoritmului backpropagation, care este derivata functiei sigmoida unipolara?

- a)  $S'(s) = (1 - S(s)) * (1 + S(s))$
- b)  $S'(s) = S(s) * (1 - S(s))$**
- c)  $S'(s) = 1 - S(s)$

36. Pentru aplicarea algoritmului backpropagation, care este derivata functiei tangenta hiperbolica?

- a)  $S'(s) = (1 - S(s)) * (1 + S(s))$**
- b)  $S'(s) = S(s) * (1 - S(s))$
- c)  $S'(s) = 1 + S(s)$

37. In algoritmul backpropagation, cum se calculeaza gradientii erorilor pentru neuronii din stratul de iesire (delta\_k), daca S' este derivata functiei de activare,  $e_k(t)$  este eroarea la momentul t

iar  $w_{jk}(p)$  sunt ponderile conexiunilor dintre stratul ascuns si cel de iesire?

a)  $\delta_k(t) = S' / e_k(t)$

**b)  $\delta_k(t) = S' * e_k(t)$**

c)  $\delta_k(t) = e_k(t) * w_{jk}(t)$

39. Ce reprezinta algoritmul backpropagation stochastic?

a) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face numai pentru anumiți vectori de intrare, după o distribuție de probabilitate

**b) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor nu se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, ci in timp real, după prezentarea fiecarui vector de intrare**

c) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, si nu in timp real, după prezentarea fiecarui vector de intrare

40. Care este impedimentul cel mai important al algoritmului backpropagation?

a) nu converge pentru probleme complexe de dimensiuni mari

**b) deoarece se bazeaza pe metoda gradientului descendent, poate converge intr-un minim local al functiei de eroare pe care incearca sa o minimizeze**

c) restrictiile impuse pentru valorile ratei de invatare

41. Cum se justifica proprietatea perceptronului multistrat de "aproximator universal"?

a) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie polinomiala

**b) s-a demonstrat ca o retea (posibil infinita) cu un singur strat ascuns este capabila sa aproximeze orice functie continua**

c) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie derivabila datorita algoritmului backpropagation

42. Prin ce se caracterizeaza metoda momentului de accelerare a invatarii cu algoritmul backpropagation?

a) consta in micșorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

b) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor

**c) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective**

43. Prin ce se caracterizeaza metoda ratei de invatare adaptive pentru accelerarea invatarii cu algoritmul backpropagation?

**a) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor**

- b) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective
- c) consta in micsorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

44. Ce se intelege prin capacitatea de a generaliza a unei retele neuronale?

- a) reseaua trebuie sa dea rezultate cat mai bune pentru instantele folosite la antrenare
- b) reseaua trebuie sa dea rezultate bune indiferent de numarul de instante folosite la antrenare
- c) reseaua trebuie sa dea rezultate bune nu numai pentru instantele folosite la antrenare, ci si pentru instante noi**

45. Care ar fi conditiile pentru ca o retea sa generalizeze bine?

- a) sa aiba o topologie adecvata cu cat mai putine ponderi si totusi sa dea rezultate bune la antrenare, exploatand astfel regularitatile problemei**
- b) sa aiba un numar suficient de ponderi astfel incat sa dea rezultate cat mai bune la antrenare
- c) sa aiba un numar de ponderi egal cu o zecime din produsul dintre numarul de vectori de antrenare si numarul de iesiri

51. Prin ce se caracterizeaza o retea neuronală recurentă secvențială?

- a) reseaua are "memorie" iar iesirea dorita pentru fiecare vector de intrare depinde nu numai de acel vector ci si de cei anteriori**
- b) reseaua nu are "memorie" iar iesirea dorita pentru fiecare vector de intrare depinde numai de acel vector si nu si de cei anteriori
- c) reseaua nu are bucle iar propagarea semnalului se face secvențial, din strat in strat, de la intrare catre iesire

53. Care sunt conditiile care trebuie respectate pentru a garanta stabilitatea la antrenare a unei retele recurente de tip perceptron multistrat?

- a) folosirea unor functii de activare cat mai simple (de exemplu prag), evitarea simetriei ponderilor si implementarea dinamicii retelei
- b) folosirea unor neliniaritati crescatoare si marginite (de exemplu sigmoide), pastrarea simetriei ponderilor si implementarea dinamicii retelei**
- c) folosirea unor functii liniare, evitarea simetriei ponderilor si implementarea dinamicii retelei

57. Cum pot fi clasificate memoriile asociative din punct de vedere al asociatiilor stocate:

- a) autoasociative si heteroasociative**
- b) asociative si neasociative
- c) statice si recurente

85. Care este topologia unei retele RBF?

- a) o retea RBF este formata din doua straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni iar al doilea contine neuroni RBF
- b) o retea RBF este formata din trei straturi: primul contine neuroni simpli care transmit intrarile fara distorsiuni, al doilea contine neuroni RBF iar al treilea contine neuroni cu functii de activare liniare**

c) o retea RBF este formata din doua straturi: primul contine neuroni RBF care prelucreaza intrarile iar al doilea contine neuroni RBF cu functii de activare liniare care transmit iesirile

86. Care este iesirea unei retele RBF gaussiene, daca notam suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum_{i=1}^n (a_i)$ ,  $w_j$  sunt ponderile conexiunilor, centrele sunt exemplarele  $x_j$ , dispersiile sunt  $\sigma_j$  iar  $N$  este numarul de vectori de antrenare?

a)  $f(x) = \sum_{j=1}^N (w_j * (-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$

**b)  $f(x) = \sum_{j=1}^N (w_j * \exp(-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$**

c)  $f(x) = \sum_{j=1}^N (w_j * \exp(-1 / \sigma_j * \|x - x_j\|^2))$

87. Ce parametri ai unei retele RBF gaussiene sunt ajustati prin invatare?

a) numai centrele gaussianelor, prin metode supervizate sau nesupervizate

b) numai ponderile  $w_i$  sunt ajustate, ceilalti parametri sunt fiksi

**c) ponderile  $w_i$ , dispersiile  $\sigma_i$  iar numarul de neuroni RBF poate fi redus prin selectia unor exemplare reprezentative din vectorii de antrenare**

89. Cum se calculeaza ponderile unei retele RBF cu selectie a centrelor?

**a) prin metoda celor mai mici patrate**

b) cu algoritmul backpropagation

c) prin metoda divide et impera

1. Pentru partitionarea instantelor unui nod dintr-un arbore de decizie, care este entropia corespunzatoare situatiei din figura:

Node N2	Count
Class = 0	1
Class = 1	5

**a. 0.65**

b. 1

c. 0

d. 0.35

2. Care este diferenta dintre un proces de decizie Markov si o problema de invatare cu intarire (reinforcement learning)

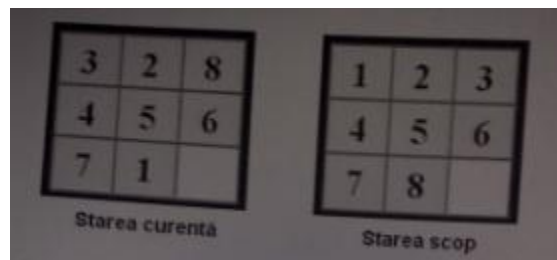
a. Intr-un proces de decizie Markov tactica (policy) optima se invata, iar intr-o problema de invatare cu intarire se calculeaza

b. Intr-un proces de decizie Markov functia recomensa  $R(s)$  este necunoscuta, intr-o problema de invatare cu intarire fiind cunoscuta

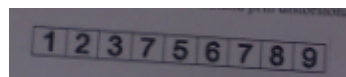
**c. Intr-un proces de decizie Markov modelul de tranzitii  $T(s, a, s')$  este cunoscut, intr-o problema de invatare cu intarire nu**



- d. Nu exista nicio diferenta
3. Metoda epsilon-greedy pentru invatarea cu intarire presupune:
- Alegerea ca actiune urmatoare a actiunii optime cunoscute cu probabilitatea  $\epsilon$  si a unei actiuni aleatoare cu prob  $1-\epsilon$
  - Alegerea ca actiune urmatoare a actiunii optime cunoscute cu probabilitatea  $1-\epsilon$  si a unei actiuni aleatoare cu prob  $\epsilon$**
  - Alegerea ca actiune urmatoare a actiunii cu rangul  $k$  cu probabilitatea  $\epsilon \cdot k$
  - Scaderea progresiva a ratei de exploatare data de probabilitatea  $\epsilon$
4. Folosind euristica distantei Manhattan , fara a include spatiul , care este estimarea in starea curenta a numarului de mutari necesare pentru rezolvare?



- 2
  - 12
  - 8**
  - 9
5. Intr-un algoritm evolutiv, fie individul din figura la care se aplica mutatia bazata pe permutari. Care din urmatoarele variante poate reprezenta doar mutatia prin distorsionare?



- 135426789**
  - 153426789
  - 154326789
  - 123456489
6. Ce tip de restrictor fuzzy este redat in figura?



- a. Puterea
  - b. Intensificarea**
  - c. Dilatarea
  - d. Concentrarea
7. Care din urmatoarele variante NU este o metoda de evitare a repetarii starilor intr-o problema de cautare neinformata ?
- a. Evitarea starilor generate anterior, care necesita memorarea tuturor starilor generate
  - b. Evitarea cailor cu bucle, cand starea unui nod este identica cu starea unui nod de pe calea din starea initiala
  - c. Evitarea starilor corespunzatoare nodurilor aflate pe niveluri inferioare nivelului solutiei**
  - d. Evitarea intoarcerii in starea din care tocmai s-a plecat, cand starea fiului este identica cu starea parintelui
8. Cine a redescoperit in 1986 algoritmul retro-propagarii (back-propagation)?
- a. Hopfield
  - b. Bryson si Ho
  - c. Rumelhart, Hinton si Williams**
  - d. Hebb
9. Care este unificarea expresiilor din figura?
- $$P(f(X,Z),Y)$$

$$P(f(g(Y),Z),Y)$$
- a.  $P / P, (X,Z) / (g(Y),Z)$
  - b. Expresiile nu pot fi unificate
  - c.  $X / g(Y), Y/Y, Z/Z$**
  - d.  $X/Y$  si  $g(W) = W$  oricare ar fi  $W$
10. Care este structura unui perceptron multistrat?

- a. **Mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor la care se adauga un termen constant si apoi rezultatului i se aplica o functie de activare neliniara**
- b. Mai multe straturi de neuroni interconectati, asupra fiecaruia din ei se aplica o functie de activare neliniara, se face suma ponderata a rezultatelor la care se adauga un termen constant
- c. Un strat de functii de baza radiala urmat de un strat cu functii liniare
- d. Mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor, rezultatului i se aplica o functie de activare neliniara si apoi se adauga un termen constant

11. De ce perceptronul multistrat nu are functii de activare liniare?

- a. Functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat
- b. **Functiile liniare nu asigura o crestere a puterii de calcul in raport cu retelele cu un singur strat deoarece o functie liniara de functii liniare este tot o functie liniara**
- c. Datorita progresului tehnologic, functiile exponentiale ce pot fi utilizate acum pentru perceptronul multistrat erau imposibil de utilizat in anii 50 pentru perceptronul cu un singur strat
- d. Perceptronul multistrat are functii de activare liniare

12. Considerand un arbore al unui joc cu un factor de ramificare constant  $b = 2$  si cu adancimea  $d = 4$ , in cazul cel mai favorabil pentru retezarea alfa-beta, care va fi nr de evaluari statice?

- a. 6
- b. 5
- c. **7**
- d. 8

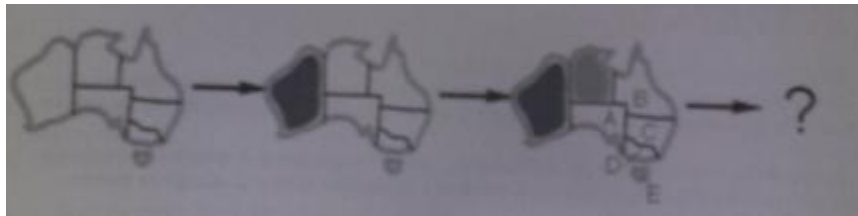
13. Daca  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de invatare,  $t$  este iesirea dorita,  $a$  este iesirea reala si  $x_i$  este intrarea  $i$ , care este regula de actualizare a ponderilor unui perceptron cu un singur strat?

- a.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$
- b.  **$w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$**
- c.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [a(k) - t(k)] * x_i(k)$
- d.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$

14. Pentru algoritmul Graphplan, ce relatie mutex, din cele de mai jos, este incorecta?

- a. Negarea termenilor
- b. Interferenta: un efect al unei actiuni neaga o preconditionie a alteia
- c. **Efecte de inconsecventa: un efect al unei actiuni neaga o preconditionie a alteia**
- d. Necesitati concurente: o preconditionie a unei actiuni neaga o preconditionie a alteia

15. Fie situatia din figura in care se coloreaza harta Australiei cu 3 culori, in ordinea negru, gri inchis, gri deschis, astfel incat doua regiuni vecine sa fie colorate diferit. S-au folosit pana acum doua culori: negru si gri inchis. Folosind euristica celei mai putin constrangatoare valori, care va fi urmatoarea regiune colorata?

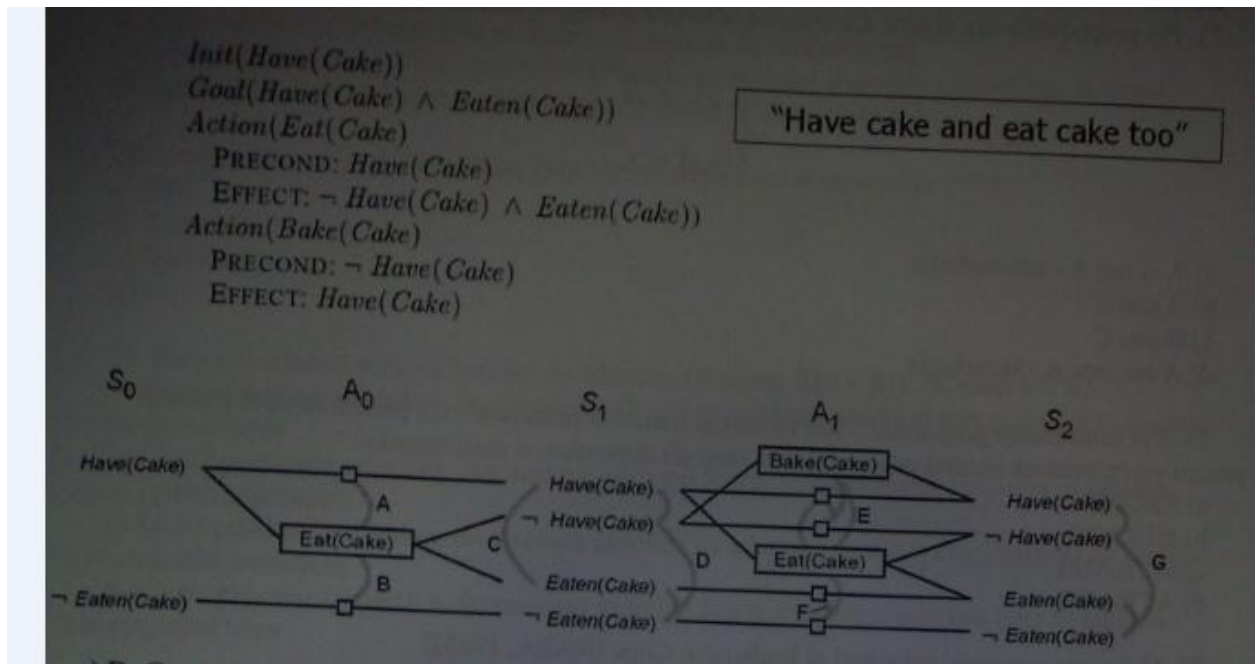


- a. Regiunea E cu negru
- b. Regiunea B cu gri deschis
- c. Regiunea A cu gri deschis
- d. **Regiunea B cu negru**

16. Care din urmatoarele afirmatii reprezinta o diferenta valida intre reprezentarile STRIPS si ADL?

- a. STRIPS permite formule predicative, ADL permite numai formule propozitionale
- b. **STRIPS respecta presupunerea lumii inchise (closed world assumption) – termenii neprecizati sunt falsi. ADL respecta presupunerea lumii deschise (open world assumption) – termenii neprecizati sunt necunoscuti**
- c. STRIPS permite operatorul egalitate, ADL nu
- d. STRIPS permite termeni pozitivi si negativi, ADL permite doar termeni pozitivi

17. Fie problema din figura: “A avea o prajitura si a manca o prajitura”. Care excluziuni mutuale (mutex) sunt incorecte?



- a. D,G
- b. C, F, G
- c. A, D,F
- d. **Toate sunt mutexuri corecte**

18. Care este utilitatea "trucului nucleului" (kernel trick) pentru o masina cu vectori suport (Support Vector Machine) ?

- a. Determina multiplicatorii langrangieni care sunt toti 0 cu exceptia celor corespunzatori vectorilor suport, iar numarul vectorilor suport este mult mai mic decat numarul de dimensiuni al instantelor
- b. **Nu este nevoie sa calculam explicit trasaturile instantelor, iar calcularea nucleului unei perechi de instante este mult mai simpla**
- c. Asigura calcularea marginii optime de clasificare fara a mai utiliza problema duala
- d. Afirma urmatoarele: conditia necesara si suficienta ca un nucleu real sa fie valid este ca matricea nucleului sa fie simetrica si pozitiv definita

19. Care dintre urmatoarele afirmatii reprezinta o diferenta corecta intre sistemele de inferenta Mamdani si Sugeno?

- a. Metoda Mamdani este mai potrivita pentru probleme de control, in special pentru sisteme neliniare dinamice
- b. Metoda Sugeno permite descrierea mai intuitiva a cunostiintelor, metoda Mamdani este mai eficienta computational

- c. Metoda Sugeno este mai potrivita pentru probleme cu numar mic de premise
- d. **Metoda Mamdani permite descrierea mai intuitiva a cunostiintelor, metoda Sugeno este mai eficienta computational**

20. In inferenta fuzzy de tip Mamdani, cum se determina matricea R?

- a.  $R_{ij} = a_j * b_i$
- b.  $R_{ij} = \min(a_j, b_i)$
- c.  $R_{ij} = a_i * b_j$
- d.  **$R_{ij} = \min(a_i, b_j)$**

23. Care este avantajul principal al limbajului CLIPS (NASA 1985)?

- a. **Motorul de inferenta**
- b. Functiile de calcul numeric
- c. Facilitatile OOP
- d. Strictetea definirii tipurilor de date

27. Care este utilizarea corectiei Laplace in metoda de clasificare bayesiana naiva?

- a. Asigura minimizarea arborelui de decizie rezultat
- b. **Evita anularea produsului de probabilitati cauzata de un factor nul**
- c. Evita problemele asociate termenului  $0 * \log(0)$  in calculul entropiei multimii de instante
- d. Accelereaza gasirea instantei celei mai apropiate de instanta de test

28. Unde a fost realizat Deep Blue care l-a invins pe Kasparov in 1997?

- a. **IBM**
- b. Princeton
- c. Stanford
- d. MIT

29. Care este efectul scalarii functiei de adaptare (fitness):  $f'(i) = f(i) - \beta$ , cand  $\beta < 0$ ?

- a. Scade presiunea selectiva, favorizeaza indivizii cu adaptare medie si este potrivita in situatia convergentei lente
- b. Creste presiunea selectiva, favorizeaza indivizii cu foarte adaptati si este potrivita in situatia convergentei premature
- c. **Scade presiunea selectiva, favorizeaza indivizii cu adaptare medie si este potrivita in situatia convergentei premature**
- d. Creste presiunea selectiva, favorizeaza indivizii cu foarte adaptati si este potrivita in situatia convergentei lente

30. Care sunt probabilitatile de selectie ale indivizilor din figura cand se foloseste metoda rangurilor, cu  $\alpha = 0$ , si  $\beta = 2$ ?

Individual	Fitness
A	71
B	7
C	2
D	20

- a. A – 40%, B – 25%, C – 15%, D – 20%
  - b. A – 2%, B – 20%, C – 71%, D – 7%
  - c. A – 71%, B – 7%, C – 2%, D – 20%
  - d. **A – 50%, B – 17%, C – 0%, D – 33%**
31. In procesul de transformare a unei formule predicative in Forma Normal Conjunctiva, care este pasul urmator dupa mutarea tuturor cuantificatorilor la stanga formulei, fara a le schimba ordinea relativa?
- a. **Eliminarea cuantificatorilor existentiali prin Skolemizare**
  - b. Standardizarea variabilelor, astfel incat fiecare cuantificator sa dispuna de propria variabila
  - c. Transformarea formulei intr-o conjunctie de disjunctii
  - d. Eliminarea cuantificatorilor universali
32. Care este starea de echilibru Nash pentru jocul ajutorului social din figura utilizand strategii pure?

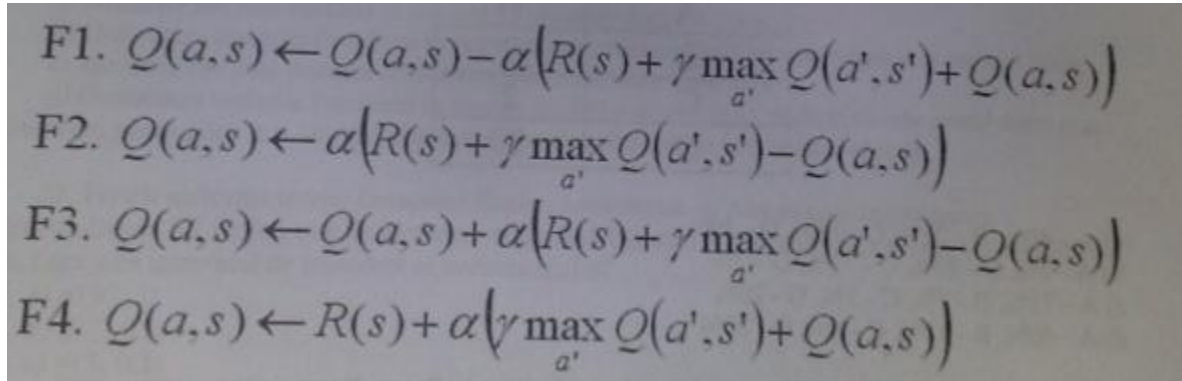
		Pauper	
		Try to work	Be idle
Government	Aid	3, 2	-1, 3
	No aid	-1, 1	0, 0

- a. Starile (3,2) si (0,0)
- b. Starea (3,2)
- c. Un joc cu strategii pure are intotdeauna echilibru Nash, deci oricare stare poate fi de echilibru Nash in functie de probabilitatile selectate pentru actiuni
- d. **Jocul nu are echilibru Nash**

33. Care a fost scopul initial al sistemelor Lindenmeyer?

- a. Modele de auto-organizare pt probleme de gasirea si optimizarea rutei
- b. Modele matematice ale evolutiei celulelor automate
- c. **Modele matematice pentru cresterea plantelor**
- d. Modele matematice pentru simularea unui stol de pasari

34. Care este formula de actualizare in algoritmul Q-Learning?



The image shows four handwritten formulas for Q-learning updates, labeled F1 through F4. Each formula uses  $Q(a, s)$  for the current state-action value and  $Q(a', s')$  for the next state-action value. The learning rate is  $\alpha$ , the discount factor is  $\gamma$ , and the reward is  $R(s)$ .

$$\begin{aligned} \text{F1. } Q(a, s) &\leftarrow Q(a, s) - \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') + Q(a, s) \right) \\ \text{F2. } Q(a, s) &\leftarrow \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') - Q(a, s) \right) \\ \text{F3. } Q(a, s) &\leftarrow Q(a, s) + \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') - Q(a, s) \right) \\ \text{F4. } Q(a, s) &\leftarrow R(s) + \alpha \left( \gamma \max_{a'} Q(a', s') + Q(a, s) \right) \end{aligned}$$

- a. F1
- b. F4
- c. F2
- d. **F3**



1. Pentru partitionarea instantelor unui nod dintr-un arbore de decizie, care este entropia corespunzatoare situatiei din figura:

Node N2	Count
Class = 0	1
Class = 1	5

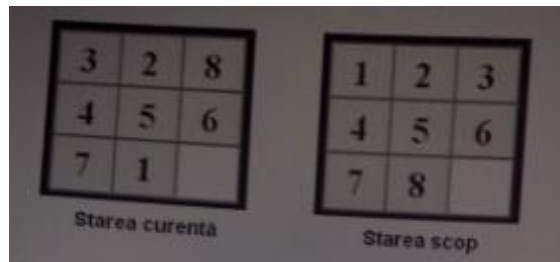
$$\text{Entropy}(t) = - \sum_{i=0}^{c-1} p(i|t) \log_2 p(i|t)$$

$$\text{Gini}(t) = 1 - \sum_{i=0}^{c-1} [p(i|t)]^2$$

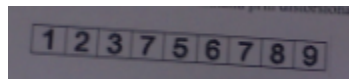
t=6

- a. **0.65**
  - b. 1
  - c. 0
  - d. 0.35
2. Care este diferenta dintre un proces de decizie Markov si o problema de invatare cu intarire (reinforcement learning)
- a. Intr-un proces de decizie Markov tactica (policy) optima se invata, iar intr-o problema de invatare cu intarire se calculeaza
  - b. Intr-un proces de decizie Markov functia recompensa R(s) este necunoscuta, intr-o problema de invatare cu intarire fiind cunoscuta
  - c. **Intr-un proces de decizie Markov modelul de tranzitii T(s, a, s') este cunoscut, intr-o problema de invatare cu intarire nu**
  - d. Nu exista nicio diferenta
3. Metoda epsilon-greedy pentru invatarea cu intarire presupune:
- a. Alegerea ca actiune urmatoare a actiunii optime cunoscute cu probabilitatea eps si a unei actiuni aleatoare cu prob 1-eps
  - b. **Alegerea ca actiune urmatoare a actiunii optime cunoscute cu probabilitatea 1-eps si a unei actiuni aleatoare cu prob eps**
  - c. Alegerea ca actiune urmatoare a actiunii cu rangul k cu probabilitatea eps\*k
  - d. Scaderea progresiva a ratei de exploatare data de probabilitatea eps

4. Folosind euristica distantei Manhattan , fara a include spatiul , care este estimarea in starea curenta a numarului de mutari necesare pentru rezolvare?



- a. 2
  - b. 12
  - c. **8**
  - d. 9
5. Intr-un algoritm evolutiv, fie individul din figura la care se aplica mutatia bazata pe permutari. Care din urmatoarele variante poate reprezenta doar mutatia prin distorsionare?



- a. **135426789**
  - b. 153426789
  - c. 154326789
  - d. 123456489
6. Ce tip de restrictor fuzzy este redat in figura?
- 

- a. Puterea
- b. **Intensificarea**
- c. Dilatarea
- d. Concentrarea

7. Care din urmatoarele variante NU este o metoda de evitare a repetarii starilor intr-o problema de cautare neinformata ?

- a. Evitarea starilor generate anterior, care necesita memorarea tuturor starilor generate
- b. Evitarea cailor cu bucle, cand starea unui nod este identica cu starea unui nod de pe calea din starea initiala
- c. Evitarea starilor corespunzatoare nodurilor aflate pe niveluri inferioare nivelului solutiei**
- d. Evitarea intoarcerii in starea din care tocmai s-a plecat, cand starea fiului este identica cu starea parintelui

8. Cine a redescoperit in 1986 algoritmul retro-propagarii (back-propagation)?

- a. Hopfield
- b. Bryson si Ho
- c. Rumelhart, Hinton si Williams**
- d. Hebb

9. Care este unificarea expresiilor din figura?

$$\begin{aligned} &P(f(X,Z),Y) \\ &P(f(g(Y),Z),Y) \end{aligned}$$

- a.  $P / P, (X,Z) / (g(Y),Z)$
- b. Expresiile nu pot fi unificate
- c.  $X / g(Y), Y/Y, Z/Z$**
- d.  $X/Y$  si  $g(W) = W$  oricare ar fi  $W$

10. Care este structura unui perceptron multistrat?

- a. Mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor la care se adauga un termen constant si apoi rezultatului i se aplica o functie de activare neliniara**
- b. Mai multe straturi de neuroni interconectati, asupra fiecaruia din ei se aplica o functie de activare neliniara, se face suma ponderata a rezultatelor la care se adauga un termen constant
- c. Un strat de functii de baza radiala urmat de un strat cu functii liniare
- d. Mai multe straturi de neuroni interconectati, fiecare din ei calculeaza suma ponderata a intrarilor, rezultatului i se aplica o functie de activare neliniara si apoi se adauga un termen constant

11. De ce perceptronul multistrat nu are functii de activare liniare?

- a. Functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat

- b. **Funcțiile liniare nu asigură o creștere a puterii de calcul în raport cu rețelele cu un singur strat deoarece o funcție liniară de funcții liniare este tot o funcție liniară**
- c. Datorită progresului tehnologic, funcțiile exponentiale ce pot fi utilizate acum pentru perceptronul multistrat erau imposibile de utilizat în anii 50 pentru perceptronul cu un singur strat
- d. Perceptronul multistrat are funcții de activare liniare
12. Considerând un arbore al unui joc cu un factor de ramificare constant  $b = 2$  și cu adâncimea  $d = 4$ , în cazul cel mai favorabil pentru rețezarea alfa-beta, care va fi nr de evaluări statice?
- a. 6
- b. 5
- c. **7**
- d. 8
13. Dacă  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de învățare,  $t$  este ieșirea dorită,  $a$  este ieșirea reală și  $x_i$  este intrarea  $i$ , care este regula de actualizare a ponderilor unui perceptron cu un singur strat?
- a.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$
- b.  **$w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$**
- c.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [a(k) - t(k)] * x_i(k)$
- d.  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$
14. Pentru algoritmul Graphplan, ce relație mutex, din cele de mai jos, este incorectă?
- a. Negarea termenilor
- b. Interferență: un efect al unei acțiuni neagă o condiție a alteia
- c. **Efecte de inconsecvență: un efect al unei acțiuni neagă o condiție a alteia**
- d. Necesități concurente: o condiție a unei acțiuni neagă o condiție a alteia
15. Fie situația din figura în care se colorează harta Australiei cu 3 culori, în ordinea negru, gri închis, gri deschis, astfel încât două regiuni vecine să fie colorate diferit. S-au folosit până acum două culori: negru și gri închis. Folosind euristica celei mai puțin constrângătoare valori, care va fi următoarea regiune colorată?

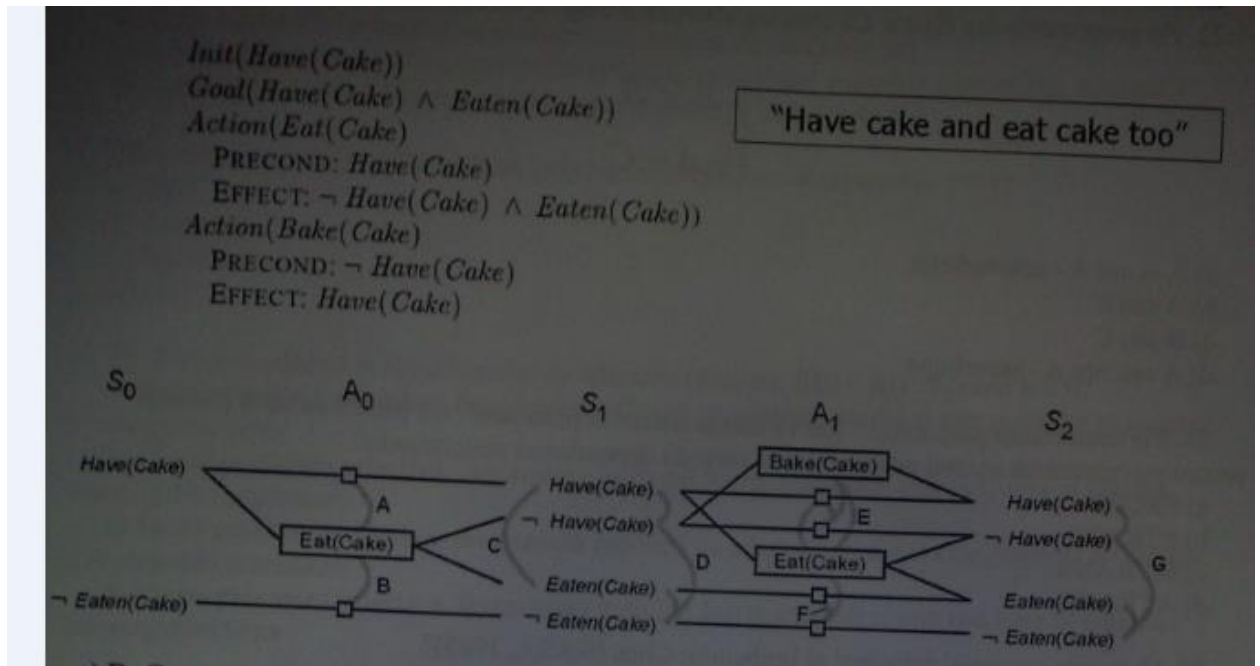


- a. Regiunea E cu negru
- b. Regiunea B cu gri deschis
- c. Regiunea A cu gri deschis
- d. **Regiunea B cu negru**

16. Care din urmatoarele afirmatii reprezinta o diferenta valida intre reprezentarile STRIPS si ADL?

- a. STRIPS permite formule predicative, ADL permite numai formule propozitionale
- b. **STRIPS respecta presupunerea lumii inchise (closed world assumption) – termenii neprecizati sunt falsi. ADL respecta presupunerea lumii deschise (open world assumption) – termenii neprecizati sunt necunoscuti**
- c. STRIPS permite operatorul egalitate, ADL nu
- d. STRIPS permite termeni pozitivi si negativi, ADL permite doar termeni pozitivi

17. Fie problema din figura: “ A avea o prajitura si a manca o prajitura”. Care excluziuni mutuale (mutex) sunt incorecte?



- a. D,G
- b. C, F, G
- c. A, D,F
- d. Toate sunt mutexuri corecte

18. Care este utilitatea "trucului nucleului" (kernel trick) pentru o masina cu vectori suport (Support Vector Machine) ?

- a. Determina multiplicatorii langrangieni care sunt toti 0 cu exceptia celor corespunzatori vectorilor suport, iar numarul vectorilor suport este mult mai mic decat numarul de dimensiuni al instantelor
- b. Nu este nevoie sa calculam explicit trasaturile instantelor, iar calcularea nucleului unei perechi de instante este mult mai simpla
- c. Asigura calcularea marginii optime de clasificare fara a mai utiliza problema duala
- d. Afirma urmatoarele: conditia necesara si suficienta ca un nucleu real sa fie valid este ca matricea nucleului sa fie simetrica si pozitiv definita

19. Care dintre urmatoarele afirmatii reprezinta o diferenta corecta intre sistemele de inferenta Mamdani si Sugeno?

- a. Metoda Mamdani este mai potrivita pentru probleme de control, in special pentru sisteme neliniare dinamice
- b. Metoda Sugeno permite descrierea mai intuitiva a cunostiintelor, metoda Mamdani este mai eficienta computational
- c. Metoda Sugeno este mai potrivita pentru probleme cu numar mic de premise
- d. **Metoda Mamdani permite descrierea mai intuitiva a cunostiintelor, metoda Sugeno este mai eficienta computational**

20. In inferenta fuzzy de tip Mamdani, cum se determina matricea R?

- a.  $R_{ij} = a_j * b_i$
- b.  $R_{ij} = \min(a_j, b_i)$
- c.  $R_{ij} = a_i * b_j$
- d.  **$R_{ij} = \min(a_i, b_j)$**

21. Fie propozitiile din figura. (exemplul e la eliminare in cursul 6 pag 18)

$$\begin{cases} A \vee B \\ \neg A \vee C \end{cases}$$

**c. B sau C**

23. Care este avantajul principal al limbajului CLIPS (NASA 1985)?

- a. **Motorul de inferenta**
- b. Functiile de calcul numeric
- c. Facilitatile OOP
- d. Strictetea definirii tipurilor de date

27. Care este utilizarea corectiei Laplace in metoda de clasificare bayesiana naiva?

- a. Asigura minimizarea arborelui de decizie rezultat
- b. **Evita anularea produsului de probabilitati cauzata de un factor nul**
- c. Evita problemele asociate termenului  $0 * \log(0)$  in calculul entropiei multimii de instante
- d. Accelereaza gasirea instantei celei mai apropiate de instanta de test

28. Unde a fost realizat Deep Blue care l-a invins pe Kasparov in 1997?

- a. **IBM**
- b. Princeton
- c. Stanford

d. MIT

29. Care este efectul scalarii functiei de adaptare (fitness):  $f'(i) = f(i) - \beta$ , cand  $\beta < 0$ ?

- a. Scade presiunea selectiva, favorizeaza indivizii cu adaptare medie si este potrivita in situatia convergentei lente
- b. Creste presiunea selectiva, favorizeaza indivizii cu foarte adaptati si este potrivita in situatia convergentei premature
- c. **Scade presiunea selectiva, favorizeaza indivizii cu adaptare medie si este potrivita in situatia convergentei premature**
- d. Creste presiunea selectiva, favorizeaza indivizii cu foarte adaptati si este potrivita in situatia convergentei lente

30. Care sunt probabilitatile de selectie ale indivizilor din figura cand se foloseste metoda rangurilor, cu  $\alpha = 0$ , si  $\beta = 2$ ?

Individual	Fitness
A	71
B	7
C	2
D	20

- a. A – 40%, B – 25%, C – 15%, D – 20%
- b. A – 2%, B – 20%, C – 71%, D – 7%
- c. A – 71%, B – 7%, C – 2%, D – 20%
- d. **A – 50%, B – 17%, C – 0%, D – 33%**

31. In procesul de transformare a unei formule predicative in Forma Normal Conjunctiva, care este pasul urmator dupa mutarea tuturor cuantificatorilor la stanga formulei, fara a le schimba ordinea relativa?

- a. **Eliminarea cuantificatorilor existentiali prin Skolemizare**
- b. Standardizarea variabilelor, astfel incat fiecare cuantificator sa dispuna de propria variabila
- c. Transformarea formulei intr-o conjunctie de disjunctii
- d. Eliminarea cuantificatorilor universali

32. Care este starea de echilibru Nash pentru jocul ajutorului social din figura utilizand strategii pure?



		Pauper	
		Try to work	Be idle
Government	Aid	3, 2	-1, 3
	No aid	-1, 1	0, 0

- a. Stările (3,2) și (0,0)
- b. Starea (3,2)
- c. Un joc cu strategii pure are întotdeauna echilibru Nash, deci oricare stare poate fi de echilibru Nash în funcție de probabilitățile selectate pentru acțiuni
- d. **Jocul nu are echilibru Nash**

33. Care a fost scopul inițial al sistemelor Lindenmeyer?

- a. Modele de auto-organizare pt probleme de găsirea și optimizarea rutei
- b. Modele matematice ale evoluției celulelor automate
- c. **Modele matematice pentru creșterea plantelor**
- d. Modele matematice pentru simularea unui stol de pasări

34. Care este formula de actualizare în algoritmul Q-Learning?

$$\begin{aligned}
 F1. \quad Q(a, s) &\leftarrow Q(a, s) - \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') + Q(a, s) \right) \\
 F2. \quad Q(a, s) &\leftarrow \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') - Q(a, s) \right) \\
 F3. \quad Q(a, s) &\leftarrow Q(a, s) + \alpha \left( R(s) + \gamma \max_{a'} Q(a', s') - Q(a, s) \right) \\
 F4. \quad Q(a, s) &\leftarrow R(s) + \alpha \left( \gamma \max_{a'} Q(a', s') + Q(a, s) \right)
 \end{aligned}$$

- a. F1
- b. F4
- c. F2
- d. **d.F3**

1. In 1943, contributia cercetatorilor McCulloch si Pitts in domeniul inteligentei artificiale a fost:

- **un model neuronal**
- un program de sah
- logica vaga (fuzzy)
- logica predicativa

2. Termenul de înțelegenta artificială a fost propus in anul:

- 1946
- **1956 - McCarthy**
- 1966
- 1976

3. Cine a inventat limbajul Lisp in anul 1958:

- Allen Newell
- **John McCarthy**
- Marvin Minsky
- Alain Colmerauer

4. Ce tip de programe de inteligenta artificiala au fost primele utilizate cu succes in **industrie**?

- sistemele de traducere automata
- retelele neuronale
- ?- **sistemele expert**
- algoritmi genetici

5. Ce sistem expert din anii 70 a fost utilizat in diagnosticarea medicala?

- Dendral
- **Mycin**
- Prospector
- XCON

6. Care a fost primul program ce incerca sa treaca testul Turing?

- ?- **Eliza**
- Alice
- Cyc
- NetTalk

7. Cine a redescoperit in anul 1986 algoritmul retro-propagarii (back-propagation)?

- Hopfield
- Hebb
- **Rumelhart, Hinton si Williams**
- Bryson si Ho

8. Unde a fost realizat Deep Blue, care l-a invins pe Gari Kasparov in 1997?

- Stanford
- Princeton
- MIT
- **IBM**

9. In ce tara a fost inventat limbajul Prolog?

- SUA
- Marea Britanie
- Germania
- ?- **Franta**

10. Care este avantajul principal al limbajului Clips (NASA, 1985)?

- **motorul de inferenta**
- functiile de calcul numeric
- facilitatile de programare orientata obiect
- strictetea definirii tipurilor de date

1. Cine a propus modelul perceptronului la sfarsitul anilor 50?

- a) Marvin Minsky
- b) Teuvo Kohonen
- c) **Frank Rosenblatt**

2. Cine sunt autorii cartii "Perceptronii" (1969) in care se evidentia imposibilitatea perceptronului de a invata functii neseparabile liniar?

- a) **Minsky si Papert**
- b) Rosenblatt si Hopfield
- c) Rumelhart si Hinton

3. Functia XOR:

- a) **poate fi invatata de un perceptron multistrat**
- b) nu poate fi invatata de o retea neuronală
- c) poate fi invatata de un perceptron cu un singur strat

4. Care este scopul unui perceptron cu un singur strat?

- a) clasificator pentru n clase prin invatarea clasei de apartenenta a fiecărei instante
- b) **clasificator pentru 2 clase prin gasirea unui hiperplan care le separa**
- c) clasificator pentru 2 clase prin gasirea unui hipervolum care le separa

? 5. Care este iesirea unui perceptron cu un singur strat?

- a) functia sigmoida aplicata intrarii cu ponderea cea mai mare
- b) **functia Heaviside aplicata sumei intrarilor ponderate**
- c) functia sigmoida aplicata sumei intrarilor ponderate

6. Dacă  $w_i$  este ponderea conexiunii  $i$  la momentul  $k$ ,  $\alpha$  este rata de învățare,  $t$  este ieșirea dorită,  $a$  este ieșirea reală și  $x_i$  este intrarea  $i$ , care este **regula de actualizare** a ponderilor unui perceptron cu un singur strat?

a)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$

**b)  $w_i(k+1) = w_i(k) + \alpha * [t(k) - a(k)] * x_i(k)$  // păstrez pasul anterior**

c)  $w_i(k+1) = w_i(k) * x_i(k) + \alpha(k) * [t(k) - a(k)] * x_i(k)$

15. Ce înseamnă principiul celei mai mici perturbări în procesul de învățare a unei rețele neuronale?

**a) ponderile trebuie ajustate astfel încât să se reducă eroarea la ieșire pentru vectorul de intrare curent cu o perturbare minimă asupra răspunsurilor deja învățate**

b) ponderile trebuie ajustate astfel încât eroarea la ieșire să fie redusă cu o cantitate cât mai mică astfel încât să nu fie perturbată rețeaua

c) ponderile trebuie ajustate cu o cantitate cât mai mică astfel încât să nu fie perturbată rețeaua

24. Care este cel mai utilizat tip de rețea neuronală?

a) hartile cu auto-organizare

**b) perceptronul multistrat**

c) memoria asociativă

25. Care este structura unui perceptron multistrat?

a) mai multe straturi de neuroni interconectați, fiecare din ei calculează suma ponderată a intrărilor, rezultatului  $i$  se aplică o funcție de activare neliniară și apoi se adaugă un termen constant

b) mai multe straturi de neuroni interconectați, asupra fiecărei intrări a neuronilor se aplică o funcție de activare neliniară, se face suma ponderată a rezultatelor la care se adaugă un termen constant

**c) mai multe straturi de neuroni interconectați, fiecare din ei calculează suma ponderată a intrărilor la care se adaugă un termen constant și apoi rezultatului  $i$  se aplică o funcție de activare neliniară**

26. Ce înseamnă conectarea feedforward a neuronilor într-un perceptron multistrat?

a) interconectarile formează bucle

b) neuronii sunt total conectați (fiecare cu fiecare)

**c) interconectarile nu formează bucle**

27. Cum se numesc rețelele perceptron multistrat în care interconectarile formează una sau mai multe bucle?

**a) recurente**

b) feedforward

c) stratificate

28. Care este expresia funcției sigmoide unipolare (funcția logistică)?

**a)  $S(s) = 1 / (1 + \exp(-s))$**

- b)  $S(s) = 1 / (1 + \exp(s))$
- c)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$

29. Care este expresia functiei sigmoide bipolare (tangenta hiperbolica)?

- a)  $S(s) = (1 - \exp(-2s)) / (1 + \exp(-2s))$
- b)  $S(s) = 1 / (1 + \exp(-s))$
- c)  $S(s) = 1 / (\exp(s) + \exp(-s))$

30. Cum se realizeaza antrenarea unui perceptron multistrat?

- a) prin modificarea modului de interconectare a neuronilor
- b) prin modificarea ponderilor conexiunilor si a valorilor prag ale neuronilor**
- c) prin modificarea valorilor neuronilor care retin informatiile

31. Ce sunt straturile ascunse ale unui perceptron multistrat?

- a) straturi interne retelei care nu participa direct la antrenare
- b) straturi interne retelei, nici de intrare si nici de iesire**
- c) straturi interne retelei care sunt parcurse uneori de semnale dinspre intrare spre iesire

32. De ce perceptronul multistrat nu are functii de activare liniare?

- a) functiile de activare liniare si neliniare determina rezultate echivalente pentru perceptronul multistrat
- b) datorita progresului tehnologic, functiile exponentiale care pot fi utilizate acum pentru perceptronul multistrat erau imposibil de utilizat practic in anii '50 pentru perceptronul cu un singur strat
- c) functiile liniare nu asigura o crestere a puterii de calcul in raport cu retelele cu un singur strat, deoarece o functie liniara de functii liniare este tot o functie liniara**

33. Cine a introdus in 1986 algoritmul backpropagation ca metoda generala de antrenare a perceptronilor multistrat?

- a) Werbos
- b) Rumelhart, Hinton si Williams**
- c) Bryson si Ho

34. Care este ideea care sta la baza algoritmului backpropagation?

- a) minimizarea erorii (diferentei dintre iesirea dorita si iesirea reala) in raport cu ponderile prin metoda gradientului descendent**
- b) minimizarea ponderilor conexiunilor prin metoda gradientului descendent
- c) minimizarea ponderilor conexiunilor in doua etape: una de propagare inainte si una de propagare inapoi

35. Pentru aplicarea algoritmului backpropagation, care este derivata functiei sigmoide unipolara?

- a)  $S'(s) = (1 - S(s)) * (1 + S(s))$
- b)  $S'(s) = S(s) * (1 - S(s))$**
- c)  $S'(s) = 1 - S(s)$

36. Pentru aplicarea algoritmului backpropagation, care este derivata functiei tangenta hiperbolica?

a)  $S'(s) = (1 - S(s)) * (1 + S(s))$

b)  $S'(s) = S(s) * (1 - S(s))$

c)  $S'(s) = 1 + S(s)$

37. In algoritmul backpropagation, cum se calculeaza gradientii erorilor pentru neuronii din stratul de iesire (delta\_k), daca S' este derivata functiei de activare, ek(t) este eroarea la momentul t

iar wjk(p) sunt ponderile conexiunilor dintre stratul ascuns si cel de iesire?

a)  $\text{delta\_k}(t) = S' / \text{ek}(t)$

b)  $\text{delta\_k}(t) = S' * \text{ek}(t)$

c)  $\text{delta\_k}(t) = \text{ek}(t) * \text{wjk}(t)$

39. Ce reprezinta algoritmul backpropagation stochastic?

a) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face numai pentru anumiți vectori de intrare, după o distributie de probabilitate

**b) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor nu se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, ci in timp real, după prezentarea fiecarui vector de intrare**

c) o varianta a algoritmului backpropagation standard in care actualizarea ponderilor se face o singura data, la sfarsitul unei epoci de antrenare, cu ajutorul gradientilor corespunzatori fiecarei conexiune, si nu in timp real, după prezentarea fiecarui vector de intrare

40. Care este impedimentul cel mai important al algoritmului backpropagation?

a) nu converge pentru probleme complexe de dimensiuni mari

**b) deoarece se bazeaza pe metoda gradientului descendent, poate converge intr-un minim local al functiei de eroare pe care incearca sa o minimizeze**

c) restrictiile impuse pentru valorile ratei de invatare

41. Cum se justifica proprietatea perceptronului multistrat de "aproximator universal"?

a) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie polinomiala

**b) s-a demonstrat ca o retea (posibil infinita) cu un singur strat ascuns este capabila sa aproximeze orice functie continua**

c) s-a demonstrat ca un perceptron multistrat poate aproxima orice functie derivabila datorita algoritmului backpropagation

42. Prin ce se caracterizeaza metoda momentului de accelerare a invatarii cu algoritmul backpropagation?

a) consta in micșorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

b) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor

**c) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective**

43. Prin ce se caracterizeaza metoda ratei de invatare adaptive pentru accelerarea invatarii cu algoritmul backpropagation?

**a) consta in utilizarea unei rate de invatare individuale pentru fiecare pondere si adaptarea acestor parametri in fiecare iteratie, in functie de semnele succesive ale gradientilor ponderilor**

b) propune, la ajustarea unei ponderi, adaugarea unui termen proportional cu ultima modificare a ponderii respective

c) consta in micșorarea progresiva a ratei de invatare pe masura ce algoritmul converge catre solutie

44. Ce se intelege prin capacitatea de a generaliza a unei retele neuronale?

a) rețeaua trebuie sa dea rezultate cat mai bune pentru instantele folosite la antrenare

b) rețeaua trebuie sa dea rezultate bune indiferent de numarul de instante folosite la antrenare

**c) rețeaua trebuie sa dea rezultate bune nu numai pentru instantele folosite la antrenare, ci si pentru instante noi**

45. Care ar fi conditiile pentru ca o retea sa generalizeze bine?

**a) sa aiba o topologie adecvata cu cat mai putine ponderi si totusi sa dea rezultate bune la antrenare, exploatand astfel regularitatile problemei**

b) sa aiba un numar suficient de ponderi astfel incat sa dea rezultate cat mai bune la antrenare

c) sa aiba un numar de ponderi egal cu o zecime din produsul dintre numarul de vectori de antrenare si numarul de iesiri

51. Prin ce se caracterizeaza o retea neuronală recurentă secvențială?

**a) rețeaua are "memorie" iar iesirea dorita pentru fiecare vector de intrare depinde nu numai de acel vector ci si de cei anteriori**

b) rețeaua nu are "memorie" iar iesirea dorita pentru fiecare vector de intrare depinde numai de acel vector si nu si de cei anteriori

c) rețeaua nu are bucle iar propagarea semnalului se face secvențial, din strat in strat, de la intrare catre iesire

53. Care sunt conditiile care trebuie respectate pentru a garanta stabilitatea la antrenare a unei retele recurente de tip perceptron multistrat?

a) folosirea unor functii de activare cat mai simple (de exemplu prag), evitarea simetriei ponderilor si implementarea dinamicii retelei

**b) folosirea unor neliniaritati crescatoare si marginite (de exemplu sigmoide), pastrarea simetriei ponderilor si implementarea dinamicii retelei**

c) folosirea unor functii liniare, evitarea simetriei ponderilor si implementarea dinamicii retelei

57. Cum pot fi clasificate memoriile asociative din punct de vedere al asociatiilor stocate:

**a) autoasociative si heteroasociative**

b) asociative si neasociative

c) statice si recurente

85. Care este topologia unei rețele RBF?

a) o rețea RBF este formată din două straturi: primul conține neuroni simpli care transmit intrările fără distorsiuni iar al doilea conține neuroni RBF

**b) o rețea RBF este formată din trei straturi: primul conține neuroni simpli care transmit intrările fără distorsiuni, al doilea conține neuroni RBF iar al treilea conține neuroni cu funcții de activare liniare**

c) o rețea RBF este formată din două straturi: primul conține neuroni RBF care prelucrează intrările iar al doilea conține neuroni RBF cu funcții de activare liniare care transmit ieșirile

86. Care este ieșirea unei rețele RBF gaussiene, dacă notăm suma cu  $i$  de la 1 la  $n$  din  $a_i$  cu  $\sum_{i=1,n} (a_i)$ ,  $w_j$  sunt ponderile conexiunilor, centrele sunt exemplarele  $x_j$ , dispersiile sunt  $\sigma_j$  iar  $N$  este numărul de vectori de antrenare?

a)  $f(x) = \sum_{j=1,N} (w_j * (-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$

**b)  $f(x) = \sum_{j=1,N} (w_j * \exp(-1 / (2 * \sigma_j^2) * \|x - x_j\|^2))$**

c)  $f(x) = \sum_{j=1,N} (w_j * \exp(-1 / \sigma_j * \|x - x_j\|^2))$

87. Ce parametri ai unei rețele RBF gaussiene sunt ajustați prin învățare?

a) numai centrele gaussianelor, prin metode supervizate sau nesupervizate

b) numai ponderile  $w_i$  sunt ajustate, ceilalți parametri sunt fixi

**c) ponderile  $w_i$ , dispersiile  $\sigma_i$  iar numărul de neuroni RBF poate fi redus prin selecția unor exemplare reprezentative din vectorii de antrenare**

89. Cum se calculează ponderile unei rețele RBF cu selecție a centrelor?

**a) prin metoda celor mai mici pătrate**

b) cu algoritmul backpropagation

c) prin metoda divide et impera