

1ª LISTA DE EXERCÍCIOS PARA A 2ª PARTE DA DISCIPLINA DE INTELIGÊNCIA ARTIFICIAL

1) Explique a diferença entre incerteza por aleatoriedade e por imprecisão.

Na incerteza por aleatoriedade calculamos a probabilidade, já na incerteza por imprecisão calculamos a possibilidade.

Interteza: "O elemento x é membro do conjunto A "

Aleatoriedade: Probabilidade de ocorrer o conjunto A

Imprecisão: Grau de pertinência ao conjunto fuzzy A

Incerteza por aleatoriedade, não se pode prever com toda certeza o que acontecerá num novo caso.

Incerteza por imprecisão, fuzzy set, trabalho com conjuntos com limites imprecisos.

2) Considere os conjuntos difusos A e B , no intervalo $X=[0; 10]$ de números reais, definidos pelas funções de pertinência

$$A(x) = x/x+2$$

$$B(x) = 2^{(-x)}$$

Resolva:

a) Represente graficamente os conjuntos difusos A e B

b) Obtenha a função de pertinência (graficamente e matematicamente) para os seguintes conjuntos:

b1) A barrado

b2) B barrado

b3) $A \cup B$

b4) $A \cap B$

b5) A barrado $\cap B$

b4) A barrado $\cup B$ barrado

c) Os conjuntos difusos atendem às leis de Morgan?

Sim

3) Qual a diferença entre α -cut e strong α -cut ?

$$\alpha\text{-cut} = \{x \mid A(x) \geq \alpha\}$$

$$\text{strong } \alpha\text{-cut} = \{x \mid A(x) > \alpha\}$$

4) Os machos humanos têm um cromossomo X e um cromossomo Y, enquanto que as fêmeas têm dois cromossomos X, cada cromossomo sendo herdado de um dos pais. A hemofilia é uma doença que apresenta herança recessiva ligada ao cromossomo X, o que significa que um homem que herda o gene que causa a doença no cromossomo X é afetado, enquanto que uma fêmea que transporta o gene em somente um de seus dois cromossomos X não é afetada. A doença é geralmente fatal para mulheres que herdaram dois genes, e isso é muito raro, já que a frequência de ocorrência do gene é reduzida em populações humanas.

Considere uma mulher que tem um irmão afetado, o que implica que a sua mãe deve ser uma portadora do gene da hemofilia com um "bom" e um "mau" gene da hemofilia. Também nos é dito que seu pai não é afetado, assim a própria mulher tem uma chance de 50-50 de ter o gene. A variável de interesse desconhecida, o estado da mulher em relação ao gene da hemofilia, tem apenas dois valores: a mulher é uma portadora do gene ($\theta = 1$) ou não ($\theta = 0$). Com base nas informações fornecidas até agora, a distribuição a priori para a variável desconhecida θ pode ser expressa simplesmente como $\Pr(\theta = 1) = \Pr(\theta = 0) = 1/2$.

Os dados utilizados para atualizar esta informação prévia consistem no estado dos filhos da mulher. Suponhamos que ela tenha tido dois filhos e que nenhum dos dois é afetado pela hemofilia. Seja $y_i = 1$ ou 0 a maneira de indicar se um filho é afetado ou não pela hemofilia. Os resultados dos dois filhos são intercambiáveis e, condicionados à variável desconhecida θ são independentes. Assumimos também que os filhos não são gêmeos idênticos. Diga qual a probabilidade da mulher ser portadora do cromossomo da hemofilia ($\theta = 1$) sabendo que os filhos não apresentam a doença $\Pr(\theta=1 | y_1 = 0 \text{ e } y_2 = 0)$.

5) Em um grupo de 50 turistas temos as seguintes variáveis descritas abaixo:

Nacionalidade	Masculino	Feminino
Brasileira	20	15
Estrangeira	5	10

Ao selecionar aleatoriamente uma pessoa do grupo obtenha as probabilidades de ocorrência dos seguintes eventos:

a) O turista é brasileiro.

$$(20+15) / 50 = 35/50 = 7/10 = 70\%$$

b) O turista é estrangeiro.

$$(5+10) / 50 = 15/50 = 3/10 = 30\%$$

c) O turista é masculino.

$$25/50 = \frac{1}{2} = 50\%$$

d) O turista é feminino.

$$25/50 = \frac{1}{2} = 50\%$$

e) O turista é feminino e brasileiro.

$$15/50 = 3/10 = 30\%$$

f) O turista é feminino e estrangeiro.

$$10/50 = \frac{1}{5} = 20\%$$

g) O turista é masculino e brasileiro.

$$20/50 = \frac{2}{5} = 40\%$$

h) O turista é masculino e estrangeiro

$$5/50 = 1/10 = 10\%$$

i) O turista é feminino ou brasileiro.

Brasileiro: $7/10$

Feminino: $\frac{1}{2}$

$$\text{Feminino ou brasileiro: } 9/10 = 90\%$$

j) O turista é feminino ou estrangeiro.

Feminino: $\frac{1}{2}$

Estrangeiro: $3/10$

$$\text{Feminino ou estrangeiro: } \frac{4}{5} = 80\%$$

k) O turista é masculino ou brasileiro.

Masculino: $\frac{1}{2}$

Brasileiro: $7/10$

$$\text{Masculino ou brasileiro: } 9/10 = 90\%$$

l) O turista é masculino ou estrangeiro.

Masculino: $\frac{1}{2}$

Estrangeiro: $3/10$

$$\text{Masculino ou estrangeiro: } \frac{4}{5} = 80\%$$

m) O turista ser masculino se é brasileiro.

$$20/35 = 4/7 = 57\%$$

n) O turista ser masculino se é estrangeiro.

$$5/15 = \frac{1}{3} = 33,3\%$$

o) O turista ser feminino se é brasileiro.

$$15/35 = 3/7 = 43\%$$

p) O turista ser feminino se é estrangeiro.

$$10/15 = 2/3 = 67\%$$

q) O turista ser brasileiro se é masculino.

$$20/25 = 4/5 = 80\%$$

r) O turista ser estrangeiro se é masculino.

$$5/25 = 1/5 = 20\%$$

s) O turista ser brasileiro se é feminino.

$$15/25 = 3/5 = 60\%$$

t) O turista ser estrangeiro se é feminino.

$$10/15 = 2/3 = 67\%$$

6) Para o diagnóstico de Edemas Subcutâneos Generalizados, tais como Celulite e Alergia, são considerados os seguintes sintomas: Febre, Cor Local, Dor História de Alergia e Prurido.

A seguir são explicitadas as probabilidades a priori $P(H_i)$ e as probabilidades condicionais $P(e/H_i)$.

Tab. 1 - Probabilidade de cada hipótese diagnóstica.

Hipóteses Diagnósticas	$P(H_i)$
Celulite	0,20
Alergia	0,80

Tab. 2 – Probabilidades condicionais de cada sintoma em relação às hipóteses diagnósticas.

Evidências	$P(e_k/Celulite)$	$P(e_k/Alergia)$
Febre		
Sim	0,85	0,5
Não	0,15	0,5
Cor Local		
Roxa	0,98	0,03
Rósea	0,02	0,97
Dor		
Leve	0,03	0,8
Intensa	0,97	0,2
História Alergia Família		

Sim	0,5	0,95
Não	0,5	0,05
Prurido		
Sim	0,5	0,98
Não	0,5	0,02

Com os dados acima:

- Desenvolva um Sistema Especialista Probabilístico (Desenhe a Rede Bayesiana de Causas e Efeitos) para Apoio ao Diagnóstico de Edemas Subcutâneos Localizados.
- Explicita algumas probabilidades da rede bayesiana.

2ª LISTA DE EXERCÍCIOS PARA A 2ª PARTE DA DISCIPLINA DE INTELIGÊNCIA ARTIFICIAL

1. Discorra sobre a capacidade de se extrair explicações para um diagnóstico obtido através de um sistema especialista implementado através de redes neurais.

Redes neurais são ótimos interpoladores, isto é, reconhecem padrões baseando-se em parametros de entrada. Essa mesma rede necessita ser treinada utilizando um conjunto de treinamento, que consiste em exemplos.

Ao entrar com os parametros para um diagnostico, a rede retorna o valor que mais se aproximam dos cenários treinados.

O problema de se usar redes neurais para aplicações como essa é o fato que a rede não retorna um percentual de acerto, se a resposta é precisa ou não. A rede retornará um resultado mesmo se nao tiver sido treinada para os dados de entrada aplicados.

Sendo assim, redes neurais pode não ser o melhor método para diagnóstico médico, principalmente pela dificuldade de treinamento para sistemas reais (com muitas variáveis) e a incerteza se a resposta é confiável ou não (não há retorno de percentual de precisao da resposta).

2. Para que serve o parâmetro taxa de aprendizado no algoritmo de aprendizado backpropagation e o que ocorre se usarmos um valor muito elevado para este parâmetro?

A taxa de aprendizado (η) é um ajuste aplicado ao Delta ou 'passo' ou deslocamento no gradiente-descendente. Se for muito elevado, pode ser difícil se aproximar do erro mínimo, visto que o algoritmo tentará se aproximar por ambos os lados mas não encontrará (ou demorará significativamente) para encontrar o erro que satisfaça a busca.

Já se a taxa de aprendizado for muito pequena, o treinamento pode ser muito lento, de acordo com a configuracao dos erros. (Se for uma superficie mais plana, demora muito caso a taxa de aprendizado seja pequena)

O ideal e aplicar uma taxa de aprendizado adaptativa.

3. Para que serve o parâmetro conhecido como momento no algoritmo de aprendizado backpropagation?

O termo momento (alpha) tenta simular a inércia aplicada ao algoritmo de descida de encosta (gradiente-descendente) com o objetivo de desviar de buracos, isto é, mínimos locais.

4. Analise a capacidade de generalização de uma rede MLP em termos de capacidade de interpolação e extrapolação.

Redes MLP são muito boas para interpolação de dados se os parâmetros forem bem configurados, como número de perceptrons da camada de entrada, camada intermediária e saída, taxa de aprendizagem, função de ativação e método de aprendizado.

Redes MLP são péssimas em extrapolação de dados, sendo difícil prever o comportamento para intervalos não cobertos pelos conjuntos de treinamento.

5. Porque os estudos sobre Redes Neurais Artificiais caíram em descrédito após 1969 e porque eles foram retomados a partir de 1986?

Porque foi concluído que o Perceptron era capaz de distinguir apenas padrões linearmente separáveis, isto causou um “trauma” na comunidade científica e levou ao corte de verbas para as pesquisas em Redes Neurais Artificiais.

Em 1974, Paul Werbos conseguiu o maior progresso em termos de redes neurais desde o perceptron de Rosenblatt: ele lançou as bases do algoritmo de retro-propagação (backpropagation), que permitiu que redes neurais com múltiplas camadas apresentassem capacidade de aprendizado. Em 1982, David Parker desenvolveu um método similar, de forma aparentemente independente. Contudo, a potencialidade deste método tardou a ser reconhecida.

Os primeiros resultados da retomada do desenvolvimento sobre redes neurais foram publicados em 1986 e 1987, através dos trabalhos do grupo PDP (Parallel and Distributed Processing), onde ficou consagrada a técnica de treinamento por backpropagation.

6. Quais os critérios de parada do treinamento backpropagation?

O critério de parada do algoritmo backpropagation não é bem definido, e geralmente é utilizado um número máximo de ciclos. Mas, devem ser considerados a taxa de erro médio por ciclo, e a capacidade de generalização da rede. Pode ocorrer que em um determinado instante do treinamento a generalização comece a degenerar, causando o problema de over-training, ou seja a rede se especializa no conjunto de dados do treinamento e perde a capacidade de generalização.

O algoritmo busca o menor erro possível tendo um erro aceitável como parametro. É também possível limitar a quantidade de épocas em que a rede é treinada. Há também o critério de validação, onde 15-20% dos neuronios sao usados para verificar problemas de overfitting. O treinamento deve ser interrompido quando o erro na validacao começar a subir de forma consistente.

7. Cite algumas aplicações que podem ser feitas com redes neurais e que são bastante complexas de serem implementadas usando outras técnicas de IA.

- aproximação de funções;
- previsão de séries temporais;
- classificações;
- reconhecimento de padrões.

8. Explique no que se constituem os seguintes passos de um Algoritmo Genético Simples.

a) Seleção dos Pais

Os pais são selecionados de acordo com sua adequação. Quanto melhores são os cromossomos, mais chances de serem selecionados. Imagine uma **roleta** onde são colocados todos os cromossomas da população. O lado de cada secção da roleta é proporcional ao valor da adequação de cada cromossoma: quanto maior for esse valor, mais larga a secção. Uma bolinha é lançada na roleta e o cromossomo onde ela para é selecionado. Evidentemente os cromossomas com maiores valores de adequação serão selecionados mais vezes.

b) Recombinação (crossover)

A recombinação, ou *crossing-over* é um processo que imita o processo biológico homônimo na reprodução sexuada: os descendentes recebem em seu código genético parte do código genético do pai e parte do código da mãe. Esta recombinação garante que os melhores indivíduos sejam capazes de trocar entre si as informações que os levam a ser mais aptos a sobreviver, e assim gerar descendentes ainda mais aptos.

c) Mutação

São feitas com probabilidade a mais baixa possível, e tem como objetivo permitir maior variabilidade genética na população, impedindo que a busca fique estagnada em um mínimo local.

d) Cálculo da Função de Aptidão (fitness)

Uma dada população consiste de μ indivíduos. Cada um é caracterizado pelo seu genótipo, consistindo de n genes, que determina de modo não ambíguo a aptidão para a sobrevivência.

9. Explique quais os critérios de parada de um Algoritmo Genético Simples.

O critério de parada do algoritmo genético pode variar de acordo com a opção do usuário. Uma das formas é definir uma quantidade de gerações que devem ser geradas. Outra forma é aplicá-lo até que se encontre uma população onde os indivíduos tenham a função de avaliação que se deseja alcançar.

Normalmente usa-se como critério de parada o número máximo de gerações ou um

tempo limite de processamento. Outro critério usa a idéia de estagnação, ou, seja, para-se o algoritmo quando não se observa melhoria da população depois de várias gerações.

10. Explique as Etapas de um Sistema Especialista Nebuloso desde a etapa de Fuzzificação até a etapa de Defuzzificação.

Passo 1 - Fuzzificar:

Etapa na qual os valores numéricos são transformados em graus de pertinência para um valor lingüístico. Cada valor de entrada terá um grau de pertinência em cada um dos conjuntos difusos.

Passo 2 - Avaliação das regras de inferência:

Cada antecedente (lado if) tem um grau de pertinência. A ação da regra (lado then) representa a saída nebulosa da regra. Durante a avaliação das regras, a intensidade da saída é calculada com base nos valores dos antecedentes e então indicadas pelas saídas nebulosas da regra.

– Alguns métodos de avaliação: MinMax, MaxMin, MaxProduto, MinMin, MaxMedia, MaxMax e Soma dos produtos.

Passo 3 - Defuzzificar:

Processo utilizado para converter o conjunto difuso de saída em um valor crisp correspondente.

– Alguns métodos de defuzzificação:

- Centróide,
- Média dos máximos,
- Distância de Hamming,
- Barras verticais,
- Método da altura, etc.

11. Considerando as variáveis fuzzy de entrada PESO e ALTURA, escreva um conjunto de regras fuzzy para um sistema especialista que selecionasse jogadores de basquete para um time (p.ex.: se o candidato for GIGANTE e PESADO então BOM jogador; se o candidato for BAIXINHO e PESADO então MAU jogador).

12. O que acontece num algoritmo genético simples se a taxa de mutação for colocada igual a zero?

Pode estagnar em um resultado de maximo local, perdendo a capacidade de tirar esta estagnação que no caso do crossover acontece.

13. O que acontece num algoritmo genético simples em que a taxa de recombinação

(crossover) seja 0 e possua uma taxa de mutação elevada?

Resultados inesperados e perda de características importantes das gerações, ou seja, ao achar um caminho que levaria ao bom resultado com a mutação pode-se perder esta característica

14. Conhecendo as técnicas de IA (Simbólica, Neural, Evolucionária e Nebulosa), diga como você poderia analisar um problema a fim de selecionar qual a técnica de IA mais apropriada para tentar achar a solução do problema. Isto é: Que características o problema deve ter que indicam a técnica mais adequada de solução.

3ª LISTA DE EXERCÍCIOS PARA A 2ª PARTE DA DISCIPLINA DE INTELIGÊNCIA ARTIFICIAL

1) Qual a diferença entre a Rede Perceptron e a Adaline?

PERCEPTRON

objetivo:

– Atuar como classificador e como gerador de funções lógicas binárias

características

- aprendizado supervisionado
- representação binária
- apenas uma camada de pesos ajustáveis

Limitações:

não admite mais de uma camada de pesos ajustáveis

- aprendizado nem sempre ocorre
- as duas classes C1 e C2 devem ser linearmente separáveis

ADALINE:

- ADALINE e Perceptron surgiram quase que simultaneamente
- tem seus pesos ajustados em função do erro de sua saída linear, antes da aplicação da função de propagação
- continua com o problema de uma única camada de pesos ajustáveis
- a função de custo a ser minimizada é quadrática nos pesos de entrada, e a minimização é feita pelo método do gradiente

Vantagens:

- Treinamento mais suave que o do Perceptron

Limitações:

- As mesmas do Perceptron
- O treinamento inclusive pode ser muito mais demorado

2) Qual foi a principal dificuldade na migração da Adaline para as MLPs?

3) Calcule os valores de W para uma rede Adaline usando o método analítico de Minimizar o Erro Médio Quadrático para o problema do OU-LÓGICO.

4) Calcule os valores de W para uma rede Adaline usando o método iterativo da Descida da Encosta mais Íngreme (Regra Delta ou Steepest Descent) para o problema do OU-LÓGICO.

5) Porque o modelo de rede neural de McCulloch&Pitts pode ser equiparado a um computador digital? Quais as 5 suposições básicas que suportam o funcionamento deste modelo?

6) Discorra sobre a afirmação: “Uma rede neural direta multicamadas é um dispositivo tolerante a falhas”. Isto é verdadeiro, falso e porque?

Verdadeiro. Redes neurais exibem características similares ao do comportamento dos seres vivos, tais como:

Generalização:

- redes neurais são capazes de generalizar o seu conhecimento a partir de exemplos anteriores
- habilidade de lidar com ruídos e distorções, respondendo corretamente a padrões novos

Robustez e Degradação Gradual:

- a perda de um conjunto de elementos processadores e/ou conexões sinápticas não causa o mal funcionamento da rede neural

7) Dada a equação fundamental do perceptron: 1 se $w_1.x_1 + w_2.x_2 \geq \theta$, 0 caso contrário. Ache os valores de w 's e θ 's para a rede de 2 camadas de perceptrons aprender o ou-exclusivo (XOR).

8) Qual a afirmação comprovada por Minsky & Papert no seu livro Perceptrons de 1969 e que praticamente estagnou as pesquisas em redes neurais artificiais por muitos anos?

Porque foi concluído que o Perceptron era capaz de distinguir apenas padrões linearmente separáveis, isto causou um “trauma” na comunidade científica e levou ao corte de verbas para as pesquisas em Redes Neurais Artificiais.

9) Em que ideia se baseia o algoritmo de aprendizado backpropagation? Para que serve o termo conhecido como momento?

Consiste em analisar os erros dos elementos processadores da camada de saída e retro-propagá-los para as camadas intermediárias (erros são calculados com resultado do treinamento supervisionado).

O método de análise e busca do erro mínimo é a gradiente-descendente.

Momento: (α) simula a inércia na descida de encosta, com o objetivo de escapar de mínimos locais.

10) Comente a seguinte afirmação referente a redes neurais com função de saída do tipo sigmoidal, dizendo se a afirmação é verdadeira ou falsa e porque:

•“ Uma rede neural direta com apenas um neurônio na camada intermediária é equivalente a uma rede neural direta sem camada intermediária e portanto incapaz de resolver o problema do “OU-EXCLUSIVO”.

Verdadeiro

11) Considere uma rede neural direta com neurônios estáticos para reconhecimento de caracteres do alfabeto. Configure esta rede em termos de número de neurônios de entrada, número de neurônios da camada de saída, número de neurônios na camada intermediária, taxa de aprendizado, função de saída dos neurônios, etc. Justifique a sua escolha.

12) Para uma população de 4 indivíduos dada a seguir e uma taxa de cruzamento de 50%, descreva o processo e mostre os elementos da população após fazer a recombinação entre o 3º e 4º bit.

Cromossoma	Valor do Cromossoma	Aptidão (fitness)	% do Total
1	10111	201	51
2	11001	136	35
3	11110	20	5
4	00010	34	9
Total:		391	100

Taxa de cruzamento é 50% então os cromossomos 1 e 3 formam um par

1 0 1 | 1 1 -> 1 0 1 1 0

1 1 1 | 1 0 -> 1 1 1 1 1

População após o cruzamento

Cromossoma	Valor do Cromossoma	Aptidão (fitness)	% do Total
1	10110		
2	11001		
3	11111		
4	00010		

13) Como a taxa de mutação de um AG pode auxiliar a procura por ótimos globais e evitar a estagnação em ótimos locais?

14) Procure na Internet implementações e soluções para o problema da mochila (“knapsack”). Verifique e descreva a forma de codificação do problema em cromossomos.

15) Comente sobre as outras técnicas de computação evolucionária e tente elencar aplicações onde cada uma das técnicas seria mais adequada para solução do problema.

16) O que você entende por algoritmos genéticos?

17) Como funcionam a mutação e o crossover? Quais as vantagens e desvantagens de cada um deles?

Mutação é a alteração de um gen do cromossomo e sua vantagem é que pode achar resultados melhores, pois o processo de recombinação tende a estagnar em um máximo que pode não ser o melhor resultado.

Crossover é a troca de material genético, onde pega-se os pais selecionados e troca-se numa certa altura do cromossomo o material genético deste pais.

18) Faça o cruzamento dos cromossomos abaixo, indique qual o cruzamento que foi feito e faça uma mutação, indicando-a.

Mãe 1 1 1 0 0 0 0 1 1 1 0 1 1 1 1

Pai 1 0 1 0 1 1 0 1 0 1 1 1 0 0 0

Crossover:

1 0 1 0 1 1 0 1 1 1 0 1 1 1 1

1 1 1 0 0 0 0 1 0 1 1 1 0 0 0

7 bit troca do material

mutação

1 1 1 1 0 0 0 1 1 1 0 1 1 1 1

4 bit para 1

19) Encontrar de x para o qual a função $f(x) = x^2 - 3x + 4$ assume o valor mínimo.

a. Assumir que $x \in [10, +10]$

b. Codificar X como vetor binário

c. Criar uma população inicial com 4 indivíduos

d. Aplicar Mutação com taxa de 1%

e. Aplicar Crossover com taxa de 60%

f. Usar seleção pelo método da roleta.

g. Usar 5 gerações.