

1. Defina a tarefa de classificação e dê exemplos de aplicação da mesma.

R.:

A classificação associa um objeto ou evento (entrada) a uma das categorias pertencentes a um conjunto FINITO de possibilidades.

Exemplos:

- E-mails classificados como SPAM ou não;
- Analisar uma imagem e dar como saída se a figura é um carro, tanque de guerra ou moto;
- Diagnóstico médico classificando registros de sintomas em pacientes.

2. Defina “aprendizado indutivo”

R.:

Aprendizado indutivo é baseado em "treinar" o modelo de IA para que ele possa obter conclusões genéricas sobre um conjunto de exemplos.

É caracterizado pelo raciocínio originado em um conceito específico que é generalizado ao longo do processo de aprendizagem.

Com base em exemplos utilizados para treinamento, o modelo é capaz de "gerar conhecimento" e prever eventos futuros.

3. Qual a diferença entre aprendizado supervisionado e não supervisionado?

R.:

Aprendizado supervisionado:

O aprendizado supervisionado consiste em alimentar o algoritmo de aprendizado com um conjunto de de treinamento onde os rótulos das classes associadas são conhecidos; O objetivo é construir um classificador capaz de determinar a classe de exemplos não analisados anteriormente, somente com base no conjunto de treinamento. Para rótulos discretos (classes), o processo é chamado de classificação; para rótulos contínuos, o processo é conhecido como regressão.

Exemplos:

- Classificação e regressão
- Redes Neurais Perceptron Multicamadas
- Somente classificação
- Árvores de decisão
- Naive Bayes
- Redes Neurais Perceptron

Aprendizado não supervisionado:

O indutor (modelo/algoritmo) analisa os exemplos sem informações sobre os rótulos e avalia se (e como) eles podem ser agrupados de alguma maneira (formando clusters, por exemplo). Após a determinação dos agrupamentos, em geral, é necessária uma análise para determinar o que cada agrupamento significa no contexto do problema tratado.

Exemplos:

- K-médias (K-means)
- Técnicas de Agrupamento Hierárquico
- Mapas Auto-organizáveis (Redes Neurais SOM)

4. Considere os seguintes métodos utilizados em aprendizagem de máquina e agrupe nas colunas correspondentes.

- 1. Árvores de decisão**
- 2. Rede Neural Perceptron**
- 3. Rede Neural MLP**

R.:

Regressão:

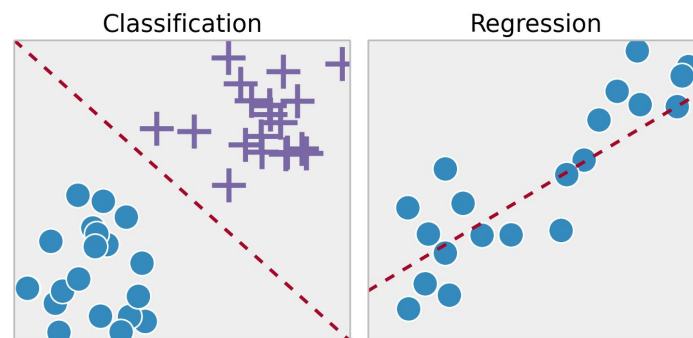
Rede Neural MLP

Classificação:

Rede Neural MLP

Árvores de decisão

Rede Neural Perceptron



5. Considerando que um conjunto de dados rotulados pode ser dividido em conjunto de treinamento e conjunto de teste, comente sobre a utilização dos métodos de amostragem na construção de classificadores.

R.:

Métodos de amostragem são maneiras de testar o classificador, ou seja, “mede” a precisão/acurácia de um modelo. Sua utilidade é ajudar o pesquisador sobre qual modelo melhor se aplica ao seu problema. A diferenciação entre um e outro método de amostragem é o modo que ele divide o conjunto em conjuntos de treinamento e de teste.

Alguns exemplos:

- Holdout
- Amostragem aleatória
 1. r-fold cross-validation
 2. r-fold stratified cross-validation
 3. Leave-one-out
 4. Bootstrap

6. O que são árvores de decisão e a qual tipo de aprendizado essa técnica está associada? Faça o mesmo para as técnicas: Naive Bayes, Rede Neural MLP e Algoritmos Genéticos

R.:

Uma **árvore de decisão** consiste em diversas comparações entre atributos do modelo de modo a chegar a alguma conclusão sobre a classe da entrada que lhe foi apresentado. A árvore lida com valores discretos e é capaz de, no final da classificação, demonstrar o caminho que foi percorrido pela árvore, dando uma “explicação” sobre o porquê do resultado. **Aprendizado simbólico.**

A família de algoritmos **Naive Bayes** tem como base o aprendizado probabilístico, utilizando o teorema de Bayes para a classificação final da entrada. Sua particularidade é que ele “ignora” as possíveis correlações entre atributos (por isso *Naive*, que significa *ingênuo*). Dentre os algoritmos Naive Bayesianos, temos o Naive Bayes Gaussiano, o Naive Bayes Multinomial e o Naive Bayes de Bernoulli. **Aprendizado probabilístico.**

A **Rede Neural Multilayer Perceptron** consiste em uma série de camadas de neurônios cujas ligações entre si possuem um peso (que é ajustado conforme a máquina “aprende”); a primeira camada (onde o dado passa primeiro) é chamada de camada de entrada, após ela, há uma ou mais camadas ocultas/internas e, finalmente, a camada de saída (que possui N neurônios tal que N é o número de classes possíveis para classificação da entrada). **Aprendizado conexonista.**

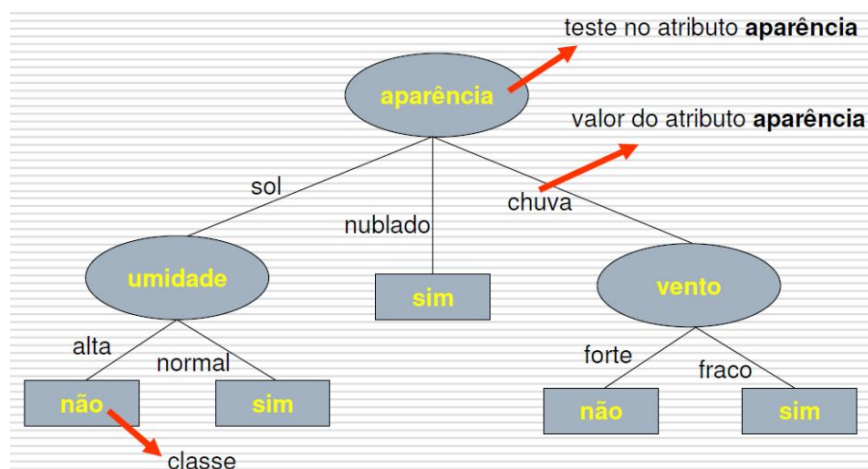
Os **algoritmos genéticos** têm como base a teoria da evolução proposta por Darwin, onde há mutações, avaliação da população/espécie resultando no contexto-problema, outra mutação, outra avaliação e assim por diante. As mutações param somente se algo as indicar para parar, como o índice de acertos esperados ou se o usuário simplesmente mandar o sistema parar. **Aprendizado evolucionista.**

7. Com relação a árvores de decisão, que tipo de informação é encontrada nos “nós internos” e nos “nós folhas”?

R.:

Os nós internos representam testes no valor de um dos atributos do conjunto de exemplos; suas ramificações são rotuladas como os possíveis valores do teste realizado nesse nó (exemplo de um teste no nó: *altura == muito alto*).

Cada nó folha representa a classe a ser retornada/respondida caso aquele nó (folha) seja alcançado.



8. Após o cálculo de entropia de um determinado atributo de um conjunto de dados, obtivemos o valor 0 como resultado. O que podemos afirmar?

R.:

A entropia significa, ligeiramente, “confusão” ou “grau de desconhecimento”. Se a entropia é 0, então os dados estão perfeitamente classificados por este atributo, ou seja, ele é 100% importante para classificar as entradas e suficiente para dar um retorno satisfatório.

9. Considerando que as árvores de decisão trabalham com valores de atributos discretos, o que pode ser feito para aplicar essa abordagem a valores contínuos?

R.:

Uma técnica comum e simples para utilizar árvores de decisão para entradas com valores contínuos é discretizá-los, ou seja, definir intervalos de números que representam valores discretos.

Exemplo:

Temperature: 40 48 60 72 80 90
PlayTennis: No No Yes Yes Yes No

A técnica matematicamente provada sendo a melhor para discretizar valores contínuos é a seguinte:

1. Pega o valor que fica entre os limites da mudança de classes: (48 e 60) (80 e 90) para posteriormente definirmos qual o melhor (de acordo com o seu ganho de informação);
2. Tira a média entre eles para definir os valores limiares: (Temp > 54) (Temp >85)
3. Após o cálculo do ganho de informação, é definido que Temperature > 54 causa o maior ganho e é escolhido como atributo de teste.

10. Descreva, de forma sucinta, como opera o algoritmo de treinamento BackPropagation.

R.:

O algoritmo de BackPropagation faz o caminho inverso de uma rede neural, ou seja, ele se inicia nos neurônios de saída e percorre as camadas ocultas até chegar aos neurônios de entrada; nesse “trajeto”, calcula-se o termo das unidades da camada de saída e utiliza essa informações para calcular os termos de erro dos neurônios da(s) camada(s) oculta(s). Assim, os erros são propagados de volta através da rede.

O critério de parada é “arbitrário”, ou seja, pode ser por um certo número de iterações pré-definido ou pode ser estabelecido um erro (no conjunto de validação ou no de treinamento) limiar (um threshold).