

Frequência 2024

Grupo I

① Erro médio absoluto (MAE): mede a magnitude média dos erros num conjunto de previsões (s/dados). $MAE = \frac{1}{n} \sum_{j=1}^n |g_j - \hat{g}_j|$

Erro médio quadrado (MSE): calcula média das diferenças, ao quadrado, entre os erros num $MSE = \frac{1}{n} \sum_{j=1}^n (g_j - \hat{g}_j)^2$

Raiz quadrada do erro médio quadrado (RMSE): calcula IT da média das diferenças, ao quadrado, entre os erros $RMSE = \frac{1}{n} \sqrt{\sum_{j=1}^n (g_j - \hat{g}_j)^2}$

②

a) Tipo de problema: era de regressão, uma vez que a variável dependente era uma variável contínua "Price", que representa o preço da casa.

b) Redes Neurais: eficaz para obter relações complexas e não lineares entre as variáveis preditoras e o preço. A capacidade de captar padrões foi essencial para melhorar accuracy.

Gradient Boosting Trees: combina várias árvores de decisão fracas que, com o modelo robusto. Elevou performance em problemas regressão, sendo capaz de lidar com tipos de dados e de 1 ano de previsão.

③ assimétrica \rightarrow maioria dos valores à esquerda (valores baixos) e poucos valores elevados (casas com preço à direita).

↳ Igual Legend: Os bins ou casais inteiros iguais nos eixos, como 1 valor baixo a maior parte dos exemplos estão nos bins inferiores e nos bins mais altos que os realmente superiores. Isso é mais representativo real dos dados, com muitos exemplos agrupados em poucos bins e outros praticamente desocupados. Pode comprometer desempenho de algoritmos que dependem de distribuição equilibrada entre classes/algoritmos.

↳ Igual Allure: Cada bin \rightarrow menos nº exemplos. Bins muito pequenos à esquerda e muito grandes à direita. Permite melhorar a distribuição dos dados mas dificulta a interpretação dos resultados.

Grupo II

① unidades computacionais; via de comunicação entre neurónios; o ponto de ligação; dendritos; neurónios

② variáveis contínuas; variáveis binárias

③ nodo; nome; path

④ das variáveis independentes (input); dependentes (target); os dados; incluem

Grupo 4

① $V_i; V_i; F_i; F_i; F_i$

② $F_i; F_i; V_i; F_i; F_i; V_i$

③ $F_i; V_i; F_i; F_i; V_i$ \rightarrow $\begin{matrix} \text{Médias Propriedades} \\ \text{Quanto (200k)} & \text{Taxa (20\%)} \\ \text{C/ 26k} & 14k & 24k & 6k & 0.15k = e1 \\ \text{S/ 60k} & 10k & 20k & 10k & \text{Taxa: } 0.15k = e2 \end{matrix}$

Grupo III

① V, ML ajusta-se pela capacidade dos sistemas aprendem automaticamente a partir de dados, sem serem explicitamente programados para este fim. Sistema ajusta-se com base em dados.

② V, como é o caso de CRISP-DM, inclui etapas de seleção de técnicas de ML mais adequadas, dependendo do problema.

③ F, Ambos desenvolvem algoritmos comensais em regressão, ao agrupar valores contínuos em intervalos.

④ V, podem ser usados tanto para classificação como para regressão.

⑤ V, O fluxo de DSA envolve ajustar os pesos sinápticos entre os neurónios com base nos erros de previsão.

⑥ F, em redes feedforward os neurónios estão organizados em camadas sequenciais e as ligações ocorrem em camadas consecutivas.

Recurso 2024

Grupo 1

① Accuracy: Quantidade de previsões corretas dividido pela quantidade de observações. $A = \frac{TP+TN}{TP+TN+FP+FN}$

Preção: Uma medida de exatidão, determinando a proporção de itens realmente entre todos os itens. $P = \frac{TP}{TP+FP}$

Recall: Uma medida de completude, \hookrightarrow objetos. $A = \frac{TP}{TP+FN}$

②

a) Para tal temos que converter as classes em valores numéricos contínuos. Isto é, convertemos ordinal em escalas numéricas (Alto $\Rightarrow 1$; Baixo $\Rightarrow 2$).

b) Regressão linear (equação linear q melhora ajuste aos dados, valores contínuos) ou Gradient Boosted Trees (GBT) (combina várias árvores de decisão simples para prever um valor contínuo, combinando elas).

Preveremos p.e. ID, podemos interpretar como algo entre baixo e médio ou usar intervalos de decisão para converter de contínuo para classe (ex: $<1.5 \Rightarrow$ baixo; $1.5-2.5 \Rightarrow$ médio; $>2.5 \Rightarrow$ alto)

③ Resampling (Over e Under A): \uparrow artificialmente n° de exemplos de B e C duplicando registros ou gerando ex. sintéticos (SMOTE). Under: exemplos de A (mas pode haver a perda de informações importantes)

\hookrightarrow Criação Class Weights: Atribuir maior peso às classes minoritárias durante o treino, assim elas contribuem a par com B e C. Tem mais penalização (ANA, regressão logística).

\hookrightarrow Augmentação: Se fizer sentido agrupar B e C, pode-se agrupar mesmo só.

Grupo 2

① F_1 : Soft Computing que é tolerante à imprecisão... onde há ambiguidade (ANA). Hard Computing exige dados precisos e bem definidos.

② F_2 : axônio transmite a saída do neurônio, não representa o peso do sinal (via comunicação neurônio).

③ F_3 : difere mais do que apenas no n° de etapas; CRISP-DM (foca-se no usuário) e SEMMA (foca-se na técnica: exploração, modelagem, avaliação).

④ V_1 : ao converter pode-se perder e não duplicar, selecionando assim informações estáveis.

⑤ F_4 : implemente Hold-out Validation.

⑥ F_5 : em caso de um (K-means), existem métricas de distância para atributos binários (Jaccard), logo é possível aplicar.

⑦ V_2 : Problemas de classificação exigem rótulos (class) e apenas supervisão.

⑧ V_3 : o índice foca no número de vezes em que as variáveis

Grupo 3

① V_1 : $V_1; F_1; F_1; F_1; F_1; F_1; F_1; V$
 \rightarrow (base original)
 \rightarrow (base com atributos duplicados)

② V_2 : $F_1; V_1; F_1; F_1; F_1; F_1; V$
 \rightarrow (base com atributos duplicados)

③ (50% dos dados para treino) $V_1; F_1; V_1; V_1; F_1; F_1; F_1; V$
 \rightarrow (base com atributos duplicados)