



Exercícios

☰ Ciclo	Ciclo 02: Aprendizado supervisionado - Classificação
# Aula	16
🕒 Created	@January 5, 2023 4:41 PM
☑ Done	<input type="checkbox"/>
☑ Ready	<input checked="" type="checkbox"/>

Objetivo da Aula:

☐ Exercícios

Conteúdo

▼ Exercícios

- ▼ 1. Rafaça o código de treinamento da aula 12: “K-Nearest Neighbors - Prática” no seu computador usando o Jupyter Notebook ou o Google Colabs.
- ▼ 2. Retreino o algoritmo com os seguintes valores para K: [3, 5, 7, 9, 11, 13, 15, 17, 19 e 21] e anote a acurácia.
- ▼ 3. Qual o problema principal de usar a métrica acurácia? Escreve um exemplo hipotético, no qual o problema acontece.
- ▼ 4. Explique com um pequeno texto ilustrando a diferença entre a métrica de Precision e Recall e mostrando quando usa deve ser escolhida em relação a outra.
- ▼ 5. Escreve um trecho de código que automatize o treinamento do algoritmo K-NN, a fim de encontrar o melhor valor para K, do exercício 2.
- ▼ 6. Escreva um pequeno texto, explicando as 6 denominações da matriz de confusão: P, N, TP, FN, FP e TN

- ▼ 7. No conjunto de dados usado na aula 12: “K-Nearest Neighbors - Prática”, existe alguma variável que fere as premissas do K-NN? Se sim ou não, explique.
- ▼ 8. Faça a seguinte bateria de testes
- ▼ 8.1 Class balanceada originalmente:
 1. Faça a matriz de confusão, calcule a acurácia, recall e precision do conjunto de dados original
 2. Anote os resultados.
 - ▼ 8.2 Classe balanceada:
 1. Mantenha a proporção de 50% das linhas da planilha de dados com exemplos da classe “Conceder” e 50% com a classe “Negar”.
 2. Faça a matriz de confusão, calcule a acurácia, recall e precision.
 3. Anote os resultados.
 - ▼ 8.3 Classe desbalanceada 90/10
 1. Mantenha a proporção de 90% das linhas da planilha de dados com exemplos da classe “Conceder” e 10% com a classe “Negar”.
 2. Faça a matriz de confusão, calcule a acurácia, recall e precision.
 3. Anote os resultados.
 - ▼ 8.4 Refaça o exercícios 8.3 com as seguintes variações: 80/20, 70/30 e 60/40
 - ▼ 8.5 Crie um tabela de comparação entre os resultados das 3 métricas (acurácia, recall e precision) para cada uma das proporções (50/50, 90/10, 80/20, 70/30, 60/40) e responda as seguintes perguntas:
 1. Como a métrica da acurácia se comporta com a variação do desbalanceamento do conjunto de dados?
 2. O que acontece com a métrica “Precision” e “Recall” a medida que os conjunto de dados tendem ao balanceamento de 50/50?
 3. Observando as respostas anteriores, quais são as ações que aumentam ou diminuem a métrica de “Precision” ou “Recall” de um problema de negócio?

▼ 8.6 Escreva um artigo para blog da Comunidade DS, descrevendo o seu aprendizado em relação ao comportamento das métricas de Precision, Recall e Acurácia a partir dos seus experimentos com os dados desbalanceados.