

Exercício 01 - KNN

Avaliação de Desempenho

2 de junho de 2021

Contexto: O KNN é um método de classificação simples baseado em uma métrica de distância e em um hiperparâmetro, o número de vizinhos k , que determina a vizinhança a ser considerada na classificação. Sendo assim, o seu desempenho depende não somente da métrica de distância utilizada, como também do valor de k .

Objetivo: Avaliar o desempenho do KNN quanto à métrica considerada e ao valor de k , utilizando dados sintéticos e conjuntos de dados de problemas reais. O aluno deverá apresentar curvas de desempenho do classificador em função de k para métricas e conjuntos de dados diferentes e avaliar os resultados.

O que deve se entregar:

- Inicialmente, o KNN deve ser avaliado para dados sintéticos de duas dimensões para que a superfície de separação possa ser observada visualmente. Devem ser gerados dados sintéticos para duas classes e duas variáveis e apresentadas as superfícies de separação para diferentes valores de k . Discutir o resultado, observando se houve ou não *over* ou *underfitting* dos dados. Mudar a métrica e observar eventuais alterações na forma da superfície.
- Escolher dois conjuntos de dados de problemas reais e avaliar o desempenho do KNN em função do valor de k . Apresentar curvas de desempenho do KNN em função do k para diferentes métricas em função do valor de k . O desempenho pode ser simplesmente a acurácia do modelo. A curva pode ser obtida variando-se o valor de k dentro de uma faixa de valores. Os conjuntos de dados devem ser de problemas de classificação binária, ou seja, de duas classes.
- Entrega em pdf pelo Moodle.