



ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO

PQI 3403 Análise de Processos da Indústria Química

Ardson dos Santos Vianna Júnior - ASVJ
e-mail: ardson@usp.br





ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO

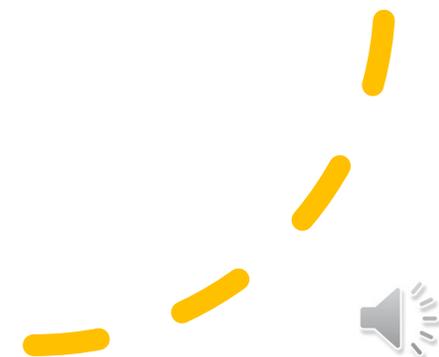
Aula Machine learning

PQI 3403 Análise de Processos da Indústria Química



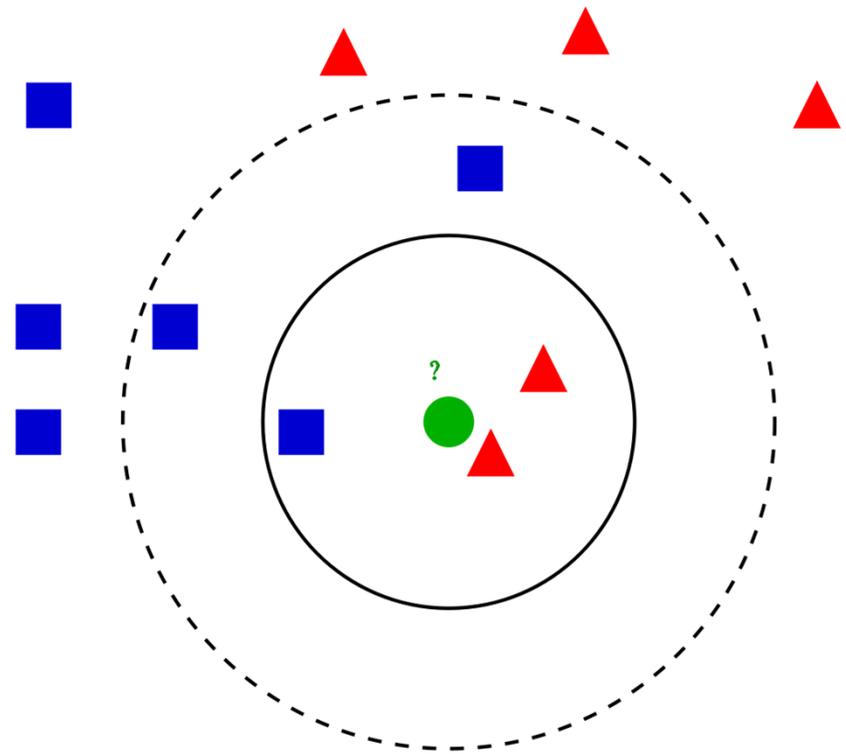
Roteiro

1. Introdução
2. Algoritmo KNN
3. Exemplo
4. Vantagens
5. Conclusão



Algoritmo KNN

- K Nearest neighbors
- K Vizinhos próximos



Algoritmo KNN

- Algoritmo de classificação não paramétrico
- A entrada são os k mais próximos do conjunto de dados
- A saída em um membro que faz parte da classe



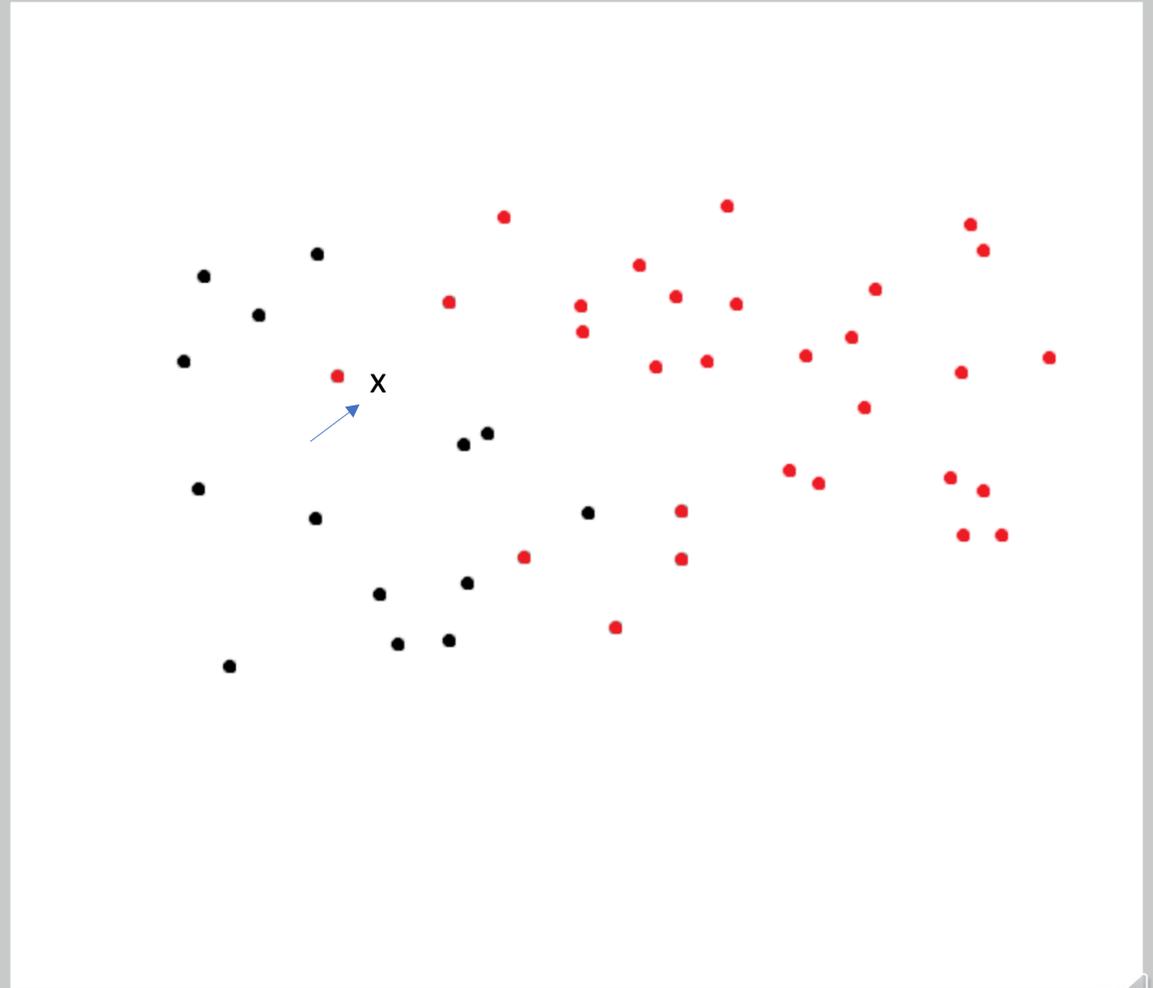
Algoritmo KNN

1. Fase de aprendizagem: identificar a partição que recebeu o conjunto de atributos original
2. Cálculo das distâncias: obter o vetor de posição unidimensional. Métricas de Minkowics
3. Fase de classificação: um ponto, que representa um objeto, é assinalado a uma classe. Esta é escolhida entre as k's vizinhas mais frequentes. A vizinhança é caracterizada a partir das distâncias.



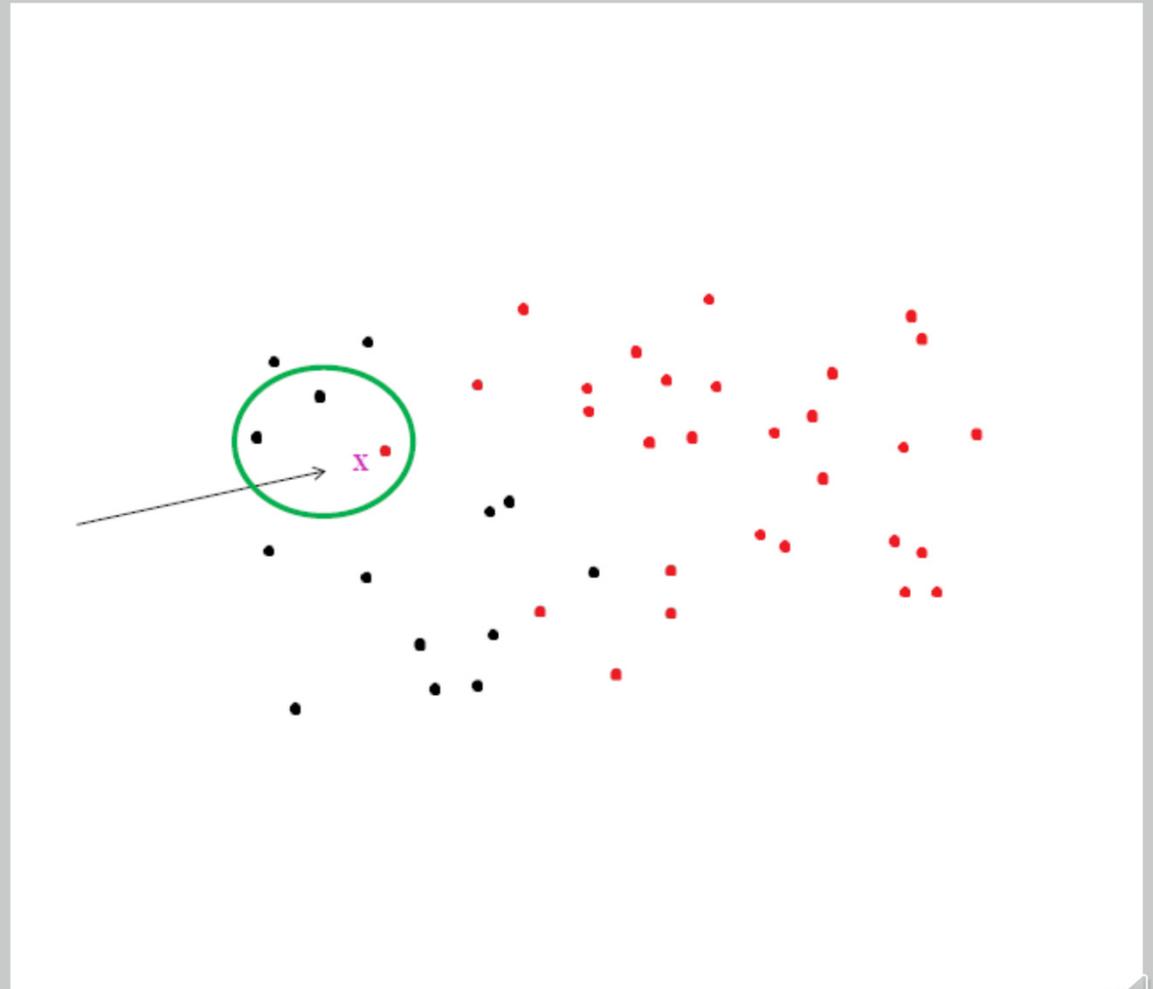
Classificação KNN

- Encontrando o rótulo
 - Encontre o exemplar mais próximo no conjunto de dados (Vermelho)
 - Distância!
 - Predizer este rótulo como o associado



K vizinhos próximos

- Ao considerar mais pontos, a classificação muda
- $K=3$, preto





Ponderação

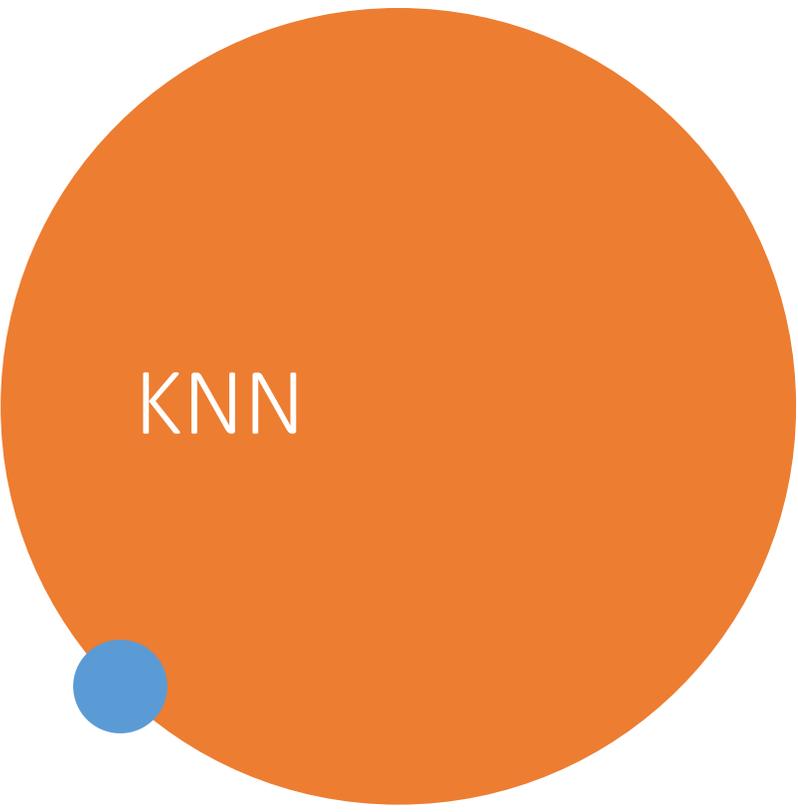
- Usar como peso o inverso da distância
- Normalizar os pesos w_{ni} : $\sum_i w_{ni} = 1$



KNN

- Vantagens
 - Aprendizagem rápida, não precisa de treinamento
 - Não demanda teoria
 - Fácil de explicar o método e os resultados
- Desvantagens
 - Intensivo em memória e as previsões podem demandar tempos longos
 - Não tem um modelo que possa esclarecer no processo de geração de dados





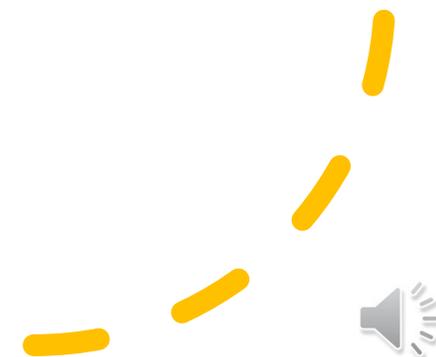
KNN

- 
- Também pode ser usado como regressão
 - O valor de saída é calculado pela média dos valores que o classificam



Conclusão

- Algoritmo KNN
- Algoritmo de classificação não paramétrico
- Tempos computacionais longos



Bibliografia

- <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-867-machine-learning-fall-2006/index.htm>
- Guttag, John. *Introduction to Computation and Programming Using Python: With Application to Understanding Data*. 2nd ed. MIT Press, 2016. ISBN: 9780262529624. Capitulo 24
- <https://www.youtube.com/watch?v=h0e2HAPTGF4>