

Apostila: Aplicação do KNN na Classificação de Risco de Crédito

Prof. Daniel Bergmann

12 de fevereiro de 2025

Sumário

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

1 Introdução

2 Conceitos Fundamentais

3 Exemplo Numérico Passo a Passo

4 Implementação do KNN em Python

5 Conclusão

Introdução

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- A classificação de risco de crédito é essencial na análise financeira.
- Permite avaliar a probabilidade de inadimplência de um solicitante de crédito.
- O modelo **K-Nearest Neighbors (KNN)** é intuitivo e eficaz, especialmente para conjuntos de dados de pequeno a moderado porte.

Objetivos da Apostila

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- Explicar detalhadamente o funcionamento do modelo KNN.
- Apresentar um exemplo numérico passo a passo.
- Demonstrar a implementação prática utilizando Python.

O que é o KNN?

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- Algoritmo de aprendizado supervisionado.
- Baseia-se na similaridade entre as observações.
- Classifica um novo ponto considerando os **K vizinhos mais próximos** no espaço das variáveis.

Funcionamento do KNN na Classificação

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- 1 Escolher um valor para K :** Define o número de vizinhos a serem considerados.
- 2 Medir a distância:** Calcula-se a distância do novo ponto para todos os pontos do conjunto de treinamento (ex.: distância Euclidiana).
- 3 Identificar os K vizinhos:** Seleciona os K pontos mais próximos.
- 4 Determinar a classe:** A classe atribuída é a maioria das classes dos vizinhos selecionados.

Escolha do Número de Vizinhos (K)

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- Valores muito baixos (ex.: $K = 1$ ou $K = 2$) podem tornar o modelo sensível a ruídos.
- Valores muito altos podem suavizar demais a fronteira de decisão, subestimando as diferenças entre classes.

Banco de Dados de Clientes

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

Considere o seguinte banco de dados:

Cliente	Renda (R\$)	Idade	Score de Crédito	Crédito
A	3000	25	600	Bom
B	2500	45	550	Ruim
C	7000	35	750	Bom
D	1500	50	500	Ruim
E	4000	30	650	Bom

Tabela: Dados utilizados para a análise.

Novo Cliente (Cliente X)

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

Características do Cliente X:

- **Renda Mensal:** 3.200 R\$
- **Idade:** 40 anos
- **Score de Crédito:** 630

Cálculo das Distâncias

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

Utiliza-se a **distância Euclidiana**:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

As distâncias entre o Cliente X e os demais clientes são:

- $d(A, X) \approx 202.8$
- $d(B, X) \approx 704.5$
- $d(C, X) \approx 3810.9$
- $d(D, X) \approx 1705.6$
- $d(E, X) \approx 800.3$

Seleção dos K Vizinhos e Classificação

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

Para $K = 3$, os vizinhos mais próximos são:

- Cliente A (202.8) → **Bom**
- Cliente B (704.5) → **Ruim**
- Cliente E (800.3) → **Bom**

Classificação Final: O Cliente X será classificado como **Bom** (maioria: 2 contra 1).

Código Python - Parte 1

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

```
1 import numpy as np
2 import pandas as pd
3 from sklearn.neighbors import KNeighborsClassifier
4
5 # Cria o do banco de dados
6 dados = pd.DataFrame({
7     'Renda': [3000, 2500, 7000, 1500, 4000],
8     'Idade': [25, 45, 35, 50, 30],
9     'Score': [600, 550, 750, 500, 650],
10    'Credito': ['Bom', 'Ruim', 'Bom', 'Ruim', 'Bom']
11 })
12
13 # Dados do novo cliente
14 novo_cliente = np.array([[3200, 40, 630]])
```

Código Python - Parte 2

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

```
1  # Separando variáveis explicativas e alvo
2  X = dados[['Renda', 'Idade', 'Score']]
3  y = dados['Credito']
4
5  # Cria o e treinamento do modelo KNN
6  knn = KNeighborsClassifier(n_neighbors=3)
7  knn.fit(X, y)
8
9  # Previsão para o novo cliente
10 previsao = knn.predict(novo_cliente)
11 print(f'Classificação do Cliente X: {previsao[0]}')
    '
```

Conclusão

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

- O KNN é um modelo intuitivo e eficiente para classificação de risco de crédito.
- Baseia-se na similaridade entre as observações e na escolha adequada de K .
- Ideal para conjuntos de dados de pequeno e moderado porte.

Agradecimentos

Apostila:
Aplicação do
KNN na
Classificação
de Risco de
Crédito

Prof. Daniel
Bergmann

Introdução

Conceitos
Fundamentais

Exemplo
Numérico
Passo a Passo

Implementação
do KNN em
Python

Conclusão

Obrigado!

Dúvidas?

Prof. Daniel Bergmann