

Lista de Exercícios de Machine Learning

Matheus Farias

18 de outubro de 2020

Resumo

Esse documento apresenta a lista de exercícios da disciplina de Machine Learning da UFPE 2020.3, feita por Matheus Farias, monitor da disciplina.

Conteúdo

1	Classificação do XOR (2.5 pt)	3
2	Clusterização de Vinhos (2.5 pt)	3
3	Comparação Geométrica (2.5 pt)	4

1 Classificação do XOR (2.5 pt)

Um dos clássicos problemas dado como motivação para o uso de algoritmos de machine learning de predição para funções não lineares é o problema do XOR.

Em lógica digital, a função XOR é descrita de acordo com a tabela verdade da Figura 1

$A \oplus B = C$		
0	0	0
0	1	1
1	0	1
1	1	0

Figura 1: Tabela verdade da função XOR

Neste problema, mostraremos que a classificação da resposta C, dada as entradas A e B não pode ser feita por algoritmos lineares.

a) 0.5 pt

Primeiramente, mostre graficamente os resultados possíveis de C, e observe empiricamente que não é possível, traçando apenas uma reta, clusterizar as respostas.

b) 1.0 pt

Utilizando um algoritmo linear, realize o treinamento de um conjunto de dados de entrada para A e B e mostre que o treinamento através de um algoritmo linear não é efetivo para o problema.

c) 1.0 pt

Utilizando um algoritmo não linear, realize o treinamento e mostre que este foi efetivo para o problema

2 Clusterização de Vinhos (2.5 pt)

Nesse problema, trabalharemos com um dataset de resultados de uma análise química de vinhos cultivados na região da Itália. A tarefa é clusterizar o vinho em três tipos distintos (0, 1 ou 2).

a) 1.5 pt

Realize a clusterização através do algoritmo KNN escolhendo o K através de algum método de otimização. Apresente os resultados da implementação do algoritmo indicando a performance de acerto.

b) 1.0 pt

Realize a clusterização através do algoritmo K-Means. Apresente os centros que convergiram e indique qual centro está mais próximo de qual vinho.

3 Comparação Geométrica (2.5 pt)

Imagine uma situação em que é dado 2 conjuntos de pontos (x, y, z) que definem sólidos geométricos, é pedido que se elabore um algoritmo que determine se as geometrias são semelhantes invariantes por translação, rotação e escala.

Nas Figuras 2 e 3 e são mostradas duas geometrias que visualmente são iguais, invariantes por translação, rotação e escala. A métrica utilizada para comparação é a soma dos erros ao quadrado (SSE), onde o erro é a distância entre os pontos mais próximos do match entre as duas figuras comparadas.

Os pontos são definidos pelos vetores das tabelas abaixo

X1	X2
[0, 0, 0]	[1, 1, 1]
[2, 0, 0]	[3.29, 1, 2.93]
[2, 2, 0]	[4.97, -0.5, 0.94]
[0, 2, 0]	[2.67, -0.5, -0.99]

Tabela 1: Vetores que definem os pontos dos quadrados

X1	X2
[0, 0, 0]	[1, 2, 1.5]
[2, 0, 0]	[2.25, -0.17, 5.83]
[2, 2, 0]	[6.29, -2.16, 3.66]
[0, 2, 0]	[5.04, 0, -0.67]
[1, 1, 1]	[4.98, 1.94, 3.21]

Tabela 2: Vetores que definem os pontos das pirâmides

Faça um algoritmo que, dadas as geometrias X1 e X2 como entrada, tenha como saída o valor SSE, as rotações em cada eixo e o fator de escala para transformar X1 em X2. A avaliação desse problema consistirá em rodar o algoritmo para o quadrado e a pirâmide mostrados, e também outra geometria que será enviada na hora da apresentação.

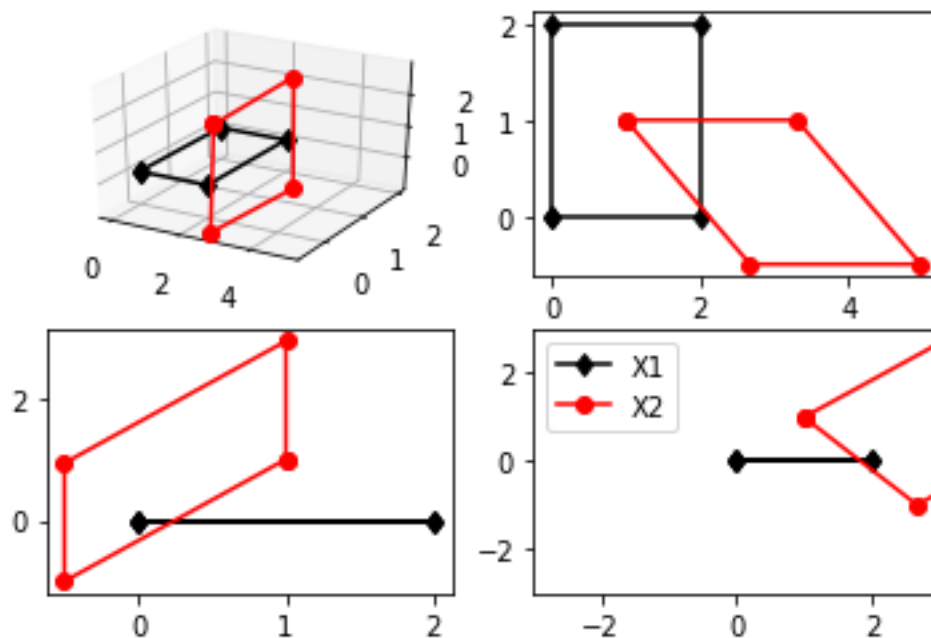


Figura 2: Comparação entre dois quadrados

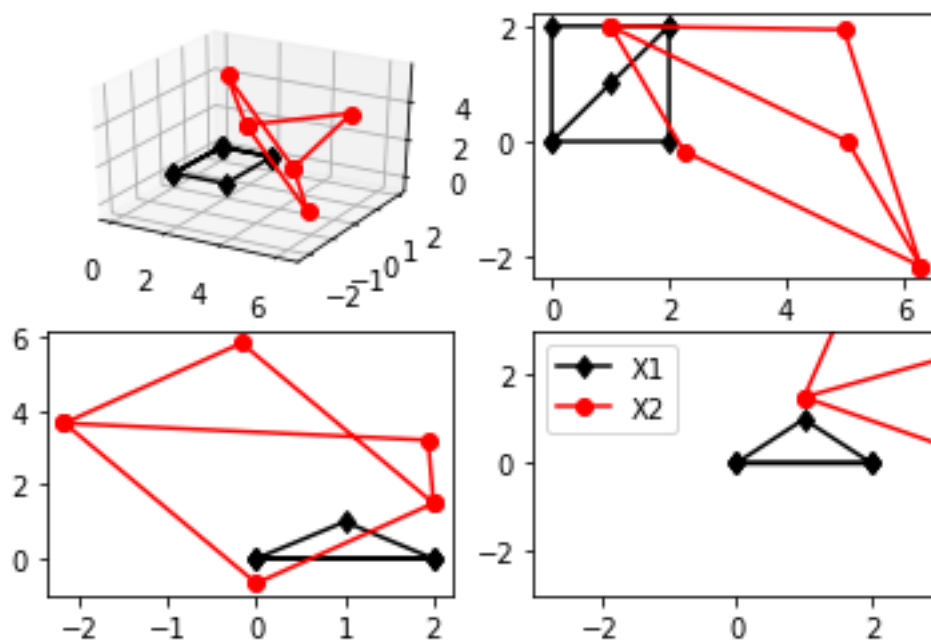


Figura 3: Comparação entre duas pirâmides

Os resultados para o quadrado e a pirâmide são mostrados abaixo ($SSE = 0$ em ambos).

Quadrados

- Rotação em x: 293.86
- Rotação em y: 198.75
- Rotação em z: 216.01
- Escala: 0.66

Pirâmides

- Rotação em x: 72.81
- Rotação em y: 327.76
- Rotação em z: 72.81
- Escala: 0.40