Atividade Avaliativa II

1. Explique a distância de Manhattan.

R: Podemos definir a distância de Manhattan, também conhecida como distância L1, entre dois pontos num espaço eucs é definida como a soma dos comprimentos das projecções do segmento de recta entre os pontos nos eixos de coordenadas. Por exemplo, no plano, a distância de Manhattan entre o ponto P1 com coordenadas (x1, y1) e o ponto P2 em (x2, y2) é, IMG_256

1. Explique a distância euclidiana.

R: A distância euclidiana é uma medida de distância entre dois pontos em um espaço euclidiano. Ela pode ser provada pela aplicação repetida do teorema de Pitágoras. A distância euclidiana entre dois pontos em um espaço euclidiano n-dimensional é definida como a raiz quadrada da soma dos quadrados das diferenças entre as coordenadas dos pontos. Em outras palavras, a distância euclidiana é a distância mais curta entre dois pontos em um espaço euclidiano.

1. Explique a distância de Hamming.

R: A distância de Hamming é uma medida de distância entre duas strings de mesmo comprimento. Ela é definida como o número de posições nas quais as duas strings diferem entre si. Em outras palavras, a distância de Hamming é o número de caracteres que precisam ser substituídos para transformar uma string na outra. [A distância de Hamming é usada em várias áreas, incluindo a teoria da informação, a teoria de códigos e a criptografia](https://pt.wikipedia.org/wiki/Dist%C3%A2ncia_de_Hamming" \t "https://www.bing.com/_blank).

1. Explique o que é aprendizado não-supervisionado.

R: O aprendizado não-supervisionado é uma técnica de aprendizado de máquina em que o algoritmo é alimentado com dados não rotulados e deve encontrar padrões e estruturas por conta própria. Diferentemente do aprendizado supervisionado, em que o algoritmo é alimentado com dados rotulados e deve aprender a mapear entradas para saídas, o aprendizado não-supervisionado é usado para encontrar estruturas e padrões ocultos em dados não rotulados.

1. Explique o que é um cluster.

R: Um cluster é um conjunto de computadores ou dispositivos de armazenamento que trabalham juntos como se fossem um único sistema. [Esses recursos são conectados em rede e trabalham em conjunto para executar tarefas, processar dados e armazenar informações](https://www.zendesk.com.br/blog/cluster-o-que-e/" \t "https://www.bing.com/_blank).Os clusters são usados em várias áreas, incluindo computação de alto desempenho, processamento distribuído, balanceamento de carga, tolerância a falhas e muito mais. [Eles são projetados para melhorar o desempenho, a escalabilidade e a disponibilidade de sistemas de computação, permitindo que os recursos sejam compartilhados e distribuídos de maneira eficiente para atender às demandas da aplicação em tempo real.](https://www.zendesk.com.br/blog/cluster-o-que-e/" \t "https://www.bing.com/_blank)

1. Explique o funcionamento do algoritmo K-Means.

R: O algoritmo K-Means é um algoritmo de agrupamento não supervisionado que divide um conjunto de dados em K clusters. O objetivo do algoritmo é encontrar K centróides que representam os K clusters. O algoritmo começa selecionando aleatoriamente K pontos do conjunto de dados como centróides iniciais. Em seguida, ele atribui cada ponto de dados ao centróide mais próximo. Depois disso, ele recalcula os centróides com base nos pontos de dados atribuídos a eles. Esse processo é repetido até que os centróides não mudem mais ou o número máximo de iterações seja atingido.

1. Explique o que é aprendizado supervisionado.

R: O aprendizado supervisionado é uma técnica de aprendizado de máquina em que o algoritmo é alimentado com dados rotulados e deve aprender a mapear entradas para saídas. À medida que os dados de entrada são inseridos no modelo, ele adapta sua ponderação até que o modelo seja ajustado adequadamente, o que ocorre como parte do processo de validação cruzada.

1. Explique o funcionamento do algoritmo KNN.

R: é um algoritmo de aprendizado de máquina não supervisionado que pode ser usado para classificação e regressão. Ele é usado para prever a classe de um objeto com base nas classes dos objetos vizinhos. O algoritmo funciona encontrando os K objetos mais próximos do objeto de entrada e, em seguida, atribuindo a ele a classe mais comum entre esses objetos. A distância entre os objetos é medida usando uma métrica de distância, como a distância euclidiana ou a distância de Hamming. O valor de K é um parâmetro que pode ser ajustado para melhorar a precisão do modelo. Um valor maior de K pode levar a uma classificação mais precisa, mas também pode levar a uma perda de detalhe.

1. Comente sobre uma área de aplicação da IA na indústria automobilística.

R: A inteligência artificial (IA) tem sido amplamente utilizada na indústria automobilística para melhorar a eficiência da produção, aumentar a segurança e melhorar a experiência do usuário. A IA pode ser usada para melhorar a produção de veículos, acelerar a classificação de dados durante avaliações de risco e avaliações de danos em veículos, e muito mais. A IA também pode ser usada para melhorar a eficiência da cadeia de suprimentos, permitindo que os fabricantes de veículos monitorem cada estágio da jornada de um componente e saibam exatamente quando esperar sua chegada na planta de destino. Além disso, a IA pode ser usada para melhorar a segurança do motorista e dos passageiros, permitindo que os veículos sejam equipados com sistemas de assistência ao motorista, como alertas de colisão, assistência de frenagem e muito mais.

10. Com base no tutorial disponível em: https://ateliware.com/blog/classificacao-knn-k-

means. Implemente uma solução para o dataset Iris.

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

# Carrega o conjunto de dados Iris

iris = pd.read\_csv('https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data', header=None)

# Seleciona as colunas de comprimento e largura da pétala

X = iris.iloc[:, [2, 3]].values

# Executa o algoritmo K-Means para agrupar as flores em clusters

kmeans = KMeans(n\_clusters=3, init='random', n\_init=10, max\_iter=300, tol=1e-04, random\_state=0)

y\_kmeans = kmeans.fit\_predict(X)

# Visualiza os clusters

plt.scatter(X[y\_kmeans == 0, 0], X[y\_kmeans == 0, 1], s=100, c='red', label='Iris-setosa')

plt.scatter(X[y\_kmeans == 1, 0], X[y\_kmeans == 1, 1], s=100, c='blue', label='Iris-versicolour')

plt.scatter(X[y\_kmeans == 2, 0], X[y\_kmeans == 2, 1], s=100, c='green', label='Iris-virginica')

plt.scatter(kmeans.cluster\_centers\_[:, 0], kmeans.cluster\_centers\_[:, 1], s=300, c='yellow', label='Centroids')

plt.title('Clusters de Flores Iris')

plt.xlabel('Comprimento da Pétala')

plt.ylabel('Largura da Pétala')

plt.legend()

plt.show()