

Latihan Logistic Regression

Gunakan dataset 'titanic' dari seaborn.

- Target = survived
- Fitur = pclass, sex, age, fare
 - Isi missing value jika ada
 - Cek multicollinearity
 - Split data menjadi train dan test set dengan proporsi 80:20
 - Buat interpretasi hasil summary
 - Cek dan interpretasi akurasi model terhadap test set

Latihan KNN & Decision Tree

Gunakan dataset 'cancer'

- Untuk simplifikasi, gunakan fitur 'texture_mean' dan 'concave points_mean' saja dengan 20% test set dan random_state = 42.

- Features: 'texture_mean', 'concave points_mean'
- Target: 'diagnosis'

- Temukan nilai K terbaik untuk model KNN, lalu gunakan metrik akurasi untuk mengevaluasi performa model antara KNN tanpa dan dengan scaling (MinMaxScaler).

- Gunakan model decision tree classifier (max_depth = 2, criterion = 'gini'), lalu bandingkan akurasinya terhadap KNN, manakah yang lebih baik?