Scaling Techniques Documentation

Scaling Techniques Used

To enhance the performance of the machine learning model, StandardScaler from the sklearn.preprocessing module was used in our project to normalise the dataset.

Why we chose StandardScaler

Since we will use the KNN method for clustering, it is important to scale the data because KNN is sensitive to the magnitude of features, and unscaled features could distort the distance calculations.

Code Snippet

```
#Scaling
scaler = StandardScaler()
train_scaled = scaler.fit_transform(train_df)
test_scaled = scaler.transform(test_df)
```