KNN Classification on Titanic Dataset

Objective

This report demonstrates implementation of the K-Nearest Neighbors (KNN) algorithm using the Titanic dataset for classification. The goal is to predict survival of passengers based on features like class, age, fare, etc.

Data Preprocessing

- Removed rows with missing values in essential columns.

- Encoded categorical variables (Sex, Embarked).

- Scaled features using StandardScaler.

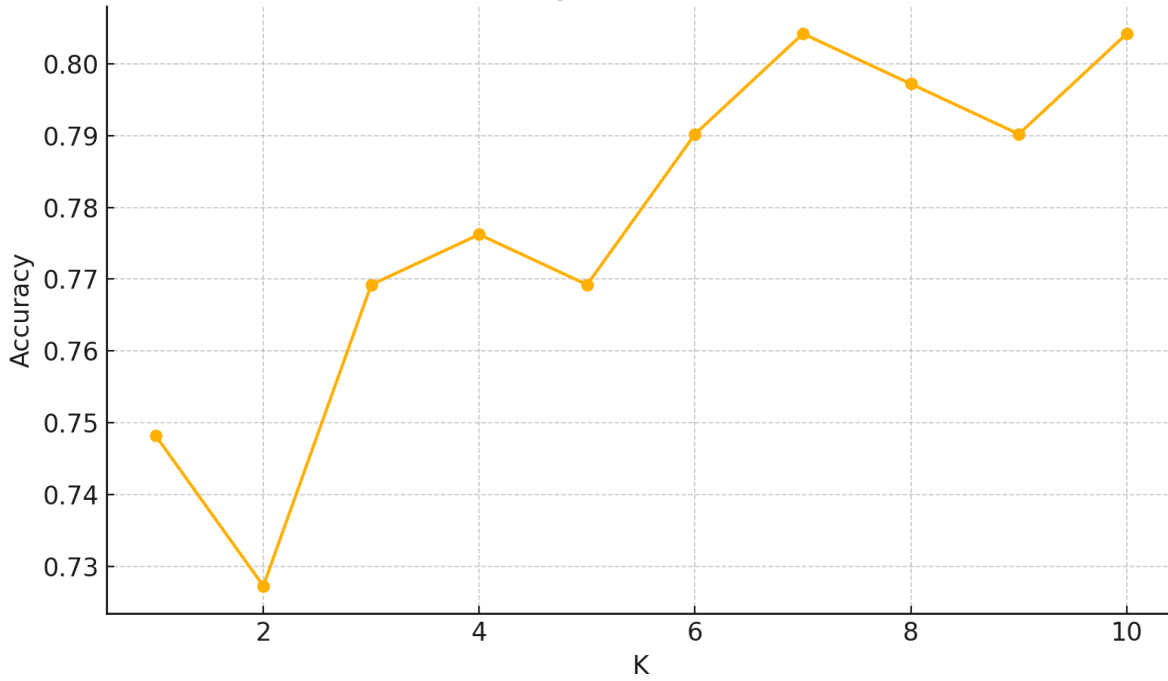- Selected features: Pclass, Sex, Age, SibSp, Parch, Fare, Embarked.

KNN Model Training and Evaluation

- K values from 1 to 10 were tested.

- Best performing K: 7

- Highest Accuracy: 80.42%

Accuracy vs K Plot

KNN Classification on Titanic Dataset

KNN Accuracy for Different K Values

Confusion Matrix (K=7)