**Report**

Predicting User Types with K-Nearest Neighbors (KNN)

Introduction

The K-Nearest Neighbors (KNN) algorithm is a simple, non-parametric, and versatile supervised machine learning method used for classification and regression tasks. This report presents the application of KNN for predicting user types based on electric vehicle charging patterns.

Dataset Overview

Dataset: Electric Vehicle Charging Patterns

* Features:
  1. Energy Consumed (kWh)
  2. Charging Rate (kW)
  3. Distance Driven (since last charge) (km)
  4. Vehicle Model
  5. Time of Day
* Target: User Type ( Casual Driver, Commuter, Long-Distance Traveler)
* Size: 120 samples

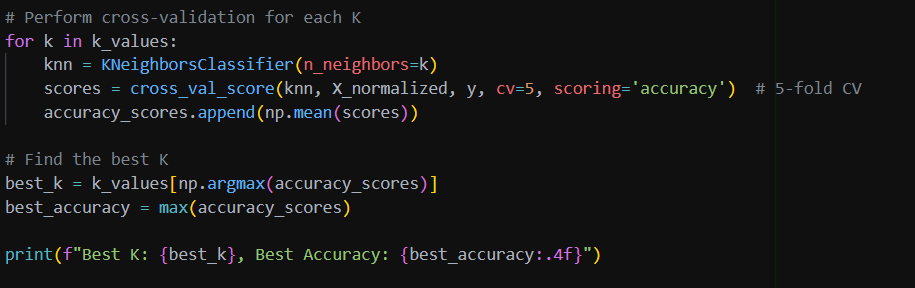
Data Preprocessing

|  |
| --- |
| The following preprocessing steps were applied to prepare the data for modeling: |
| * Handling Missing Data |
| Numeric features (e.g., Energy Consumed, Charging Rate) were imputed using the mean strategy. |
| * Categorical Encoding |
| Features such as Vehicle Model and Time of Day were encoded using Label Encoding and One-Hot Encoding. |
| * Feature Scaling |
| StandardScaler was used to normalize numeric features, ensuring all variables are on the same scale. |
| * Train-Test Split |
| Data was split into 80% training and 20% testing for model evaluation. |

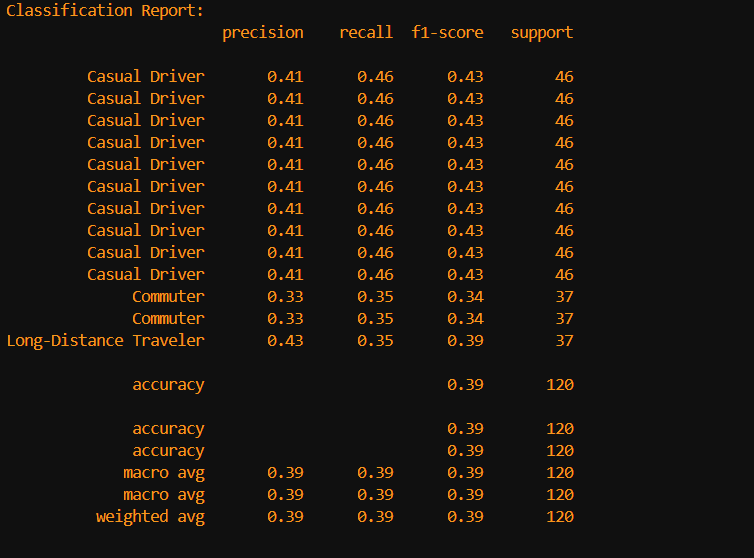
****

Cross-Validation for Optimal

To determine the best 5-fold cross-validation was conducted, evaluating accuracy for values ranging from 1 to 20.

Best: 8 (optimal number of neighbors based on cross-validation results).

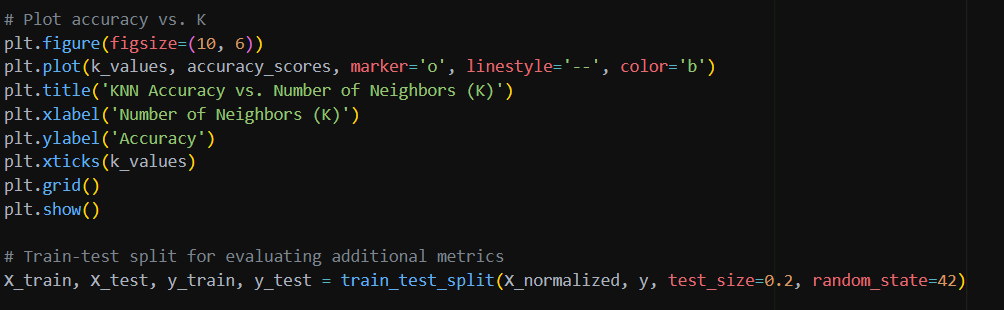
Model Performance Metrics

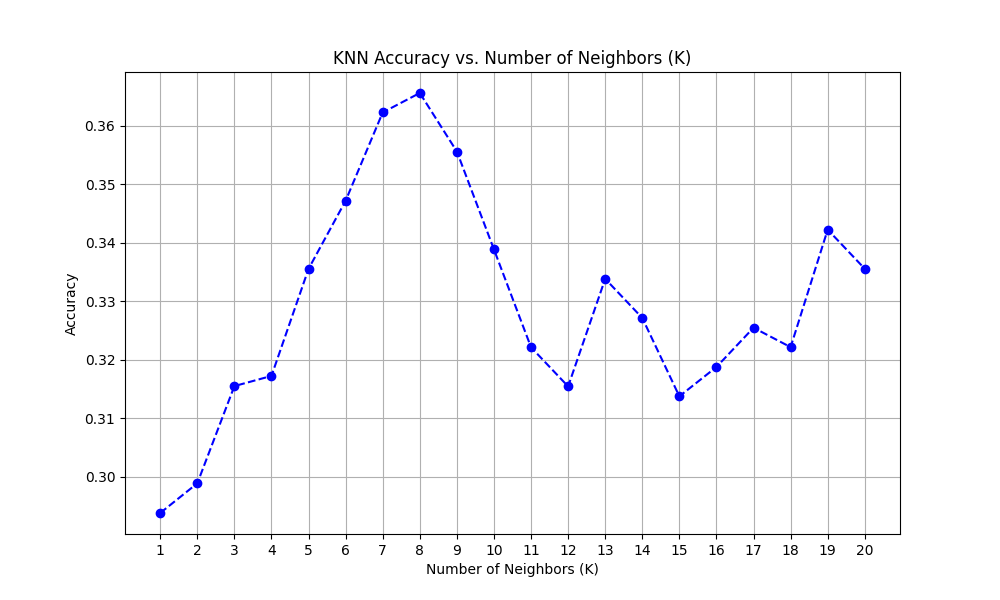


Observations

|  |
| --- |
| Best Performing Class |
| * + The "Casual Driver" class achieved the highest precision (0.41) and recall (0.46), indicating better model performance for this category. |
| Performance Gaps |
| * + The model struggles with the "Commuter" class, which had an F1-score of 0.34, suggesting difficulties in accurately identifying these instances. |
| Overall Accuracy |
| * + At 39%, the model performs better than random guessing but still requires improvement. |

Conclusion

****The K-Nearest Neighbors algorithm was implemented to predict user types for electric vehicle charging patterns. Despite its simplicity and interpretability, the model achieved a modest accuracy of 39%.

****