**Applied Machine Learning**

**Lab Report 5**

**Hafiz Ahmad**

**19l-1316**

**Section-8A**

**INTRODUCTION:**

Within machine learning algorithms, supervised learning tasks rely heavily on the use of classifier models. This analysis centers around investigating and looking at three changed classifier models: Guileless Bayes, Choice Trees, and K-Closest Neighbors (KNN) classifiers. Classification, prediction, and pattern recognition are all common uses for these models, which have a wide range of applications.

**OBJECTIVES:**

The goal of this trial is to acquire an exhaustive comprehension of the functioning standards of every classifier, look at their exhibition utilizing exactness measurements, and envision the outcomes.

**Procedure:**

Preprocessing and loading of the dataset: The pandas library is used to load the dataset called "SeoulBikeData.csv." Handling missing values and converting categorical variables into numerical format are two aspects of data preprocessing. The 'Seasons' segment is planned to mathematical qualities (Winter: Spring: 0, 1, Season: 2, Season: 3), whereas the "Functioning Day" and "Holiday" columns are encoded as binary variables (No: 0, Yes: 1).

Dataset Parting: Using the sklearn.model\_selection module's train\_test\_split function, the dataset is divided into training and testing sets. The information is partitioned in a 80:20 proportion, with 80% for preparing and 20% for testing.

Bayes-Naive Classifier: The Naive Bayes classifier makes use of the GaussianNB class from the sklearn.naive\_bayes module. The fit() method is used to train the model on the training data, and the variable y\_pred\_nb stores the predictions for the test data.

Classification by Decision Tree: The Decision Tree classifier makes use of the sklearn.tree module's DecisionTreeRegressor class. The predictions for the test data are stored in the variable y\_pred\_dt following the model's training on the training data.

K-Closest Neighbors (KNN) Classifier: The KNN classifier is implemented using the sklearn.neighbors module's KNeighborsClassifier class. The predictions for the test data are stored in the variable y\_pred\_knn following the model's training on the training data.

Evaluation of Performance: Each classifier's performance is evaluated using the sklearn.metrics module's accuracy\_score function. Naive Bayes (accuracy\_nb), Decision Trees (accuracy\_dt), and KNN (accuracy\_knnn) classifiers all have their respective accuracy values recorded.

Visualization: To show the results, the matplotlib.pyplot library is used to create a scatter plot. The actual values are shown in blue, while the predicted values from each classifier are shown in a different color (red for Naive Bayes, green for Decision Trees, and magenta for KNN) on the scatter plot. The performance and accuracy of each classifier in predicting the target variable are shown in this visualization.

Results: The exactnesses got for every classifier are as per the following:

Accuracy of naive Bayes: 0.03452243958573072

Choice Tree Precision: Accuracy: 0.7905872837463049 KNN 0.018411967779056387 The Decision Tree classifier outperformed the Naive Bayes and KNN classifiers in terms of accuracy in predicting the target variable, as evidenced by these observations.

A screenshot of a computer program

Description automatically generated with medium confidence

A screenshot of a computer code

Description automatically generated with medium confidence

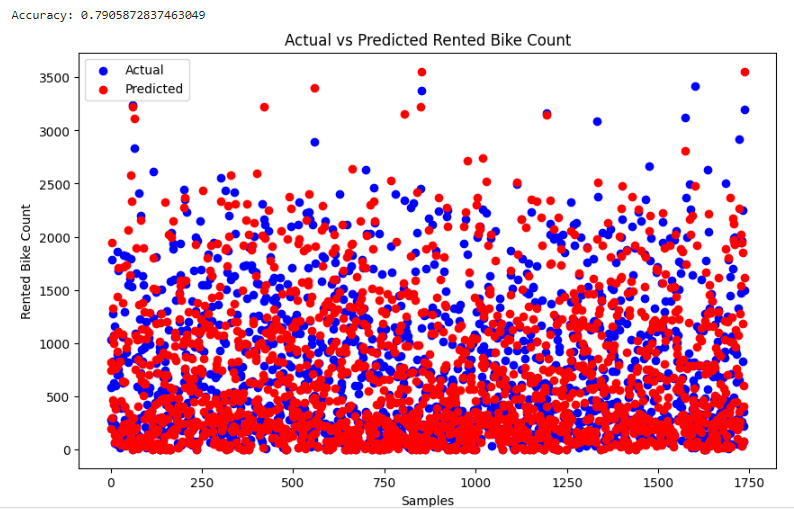
A picture containing text, screenshot, font, number

Description automatically generated

**Naïve Bayes classifier:**

A screen shot of a graph

Description automatically generated with low confidence

**Decision tree classifier:**

**KNN-Classifier:**A screen shot of a graph

Description automatically generated with low confidence

**Application:**

The Naïve Bayes, Decision Trees, and KNN classifiers track down applications in different fields, including:

Email Control: In email filtering systems, Naive Bayes classifiers are used to classify incoming emails based on their content and features as spam or non-spam.

Risk Assessment of Credit: In the banking industry, decision tree classifiers are used to evaluate loan applicants' credit risk by taking into account their income, credit history, and employment status.

Classification of Images: KNN classifiers are utilized in picture acknowledgment undertakings to order pictures into various classifications, like item discovery, facial acknowledgment, and scene acknowledgment.

**Issues:**

No issue was found while performing in the lab.

**Conclusion:**

In conclusion, the dataset used in this experiment was successfully classified using KNN, Decision Trees, and Naive Bayes classifiers. Each classifier's performance in predicting the target variable was examined by evaluating their accuracy and displaying the results. Additionally, these classifiers can be used for a variety of tasks, including image classification, credit risk assessment, and email filtering.