MACHINE LEARNING

MINI-PROJECT

METRICS USED TO EVALUATE THE PERFORMACE OF THE ALGORITHMS

CLASSIFICATION METRICS:

1) Confusion Matrix : TP,TN,FP,FN

2) Accuracy : Number of correct predictions/Total number of predictions

3) Precision : True Positives/ True Positives+ False Positives

4) Recall : True Positives/ True Positives+False Negatives

5) F1 Score : 2*Precision * Recall / Precision+Recall

6) ROC (Receiver Operating Characteristics Curve)

REGRESSION METRICS:

1) Mean Squared Error(MSE) : $\Sigma(yi - pi)2n$

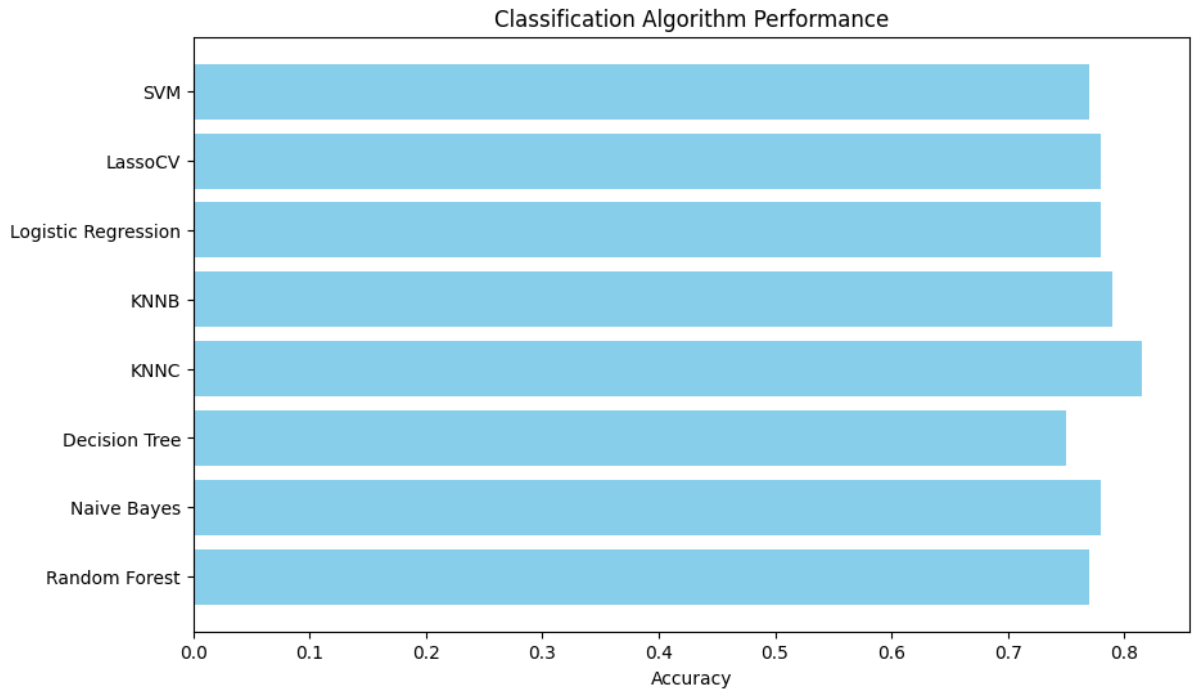
2) Mean Absolute Error(MAE) : $(1/n)\ \Sigma(i=1 \text{ to } n)\ |y\_i - \hat{y}\_i|$

3) R^2 Score : $1 - \sum (y\,i - y\,i\,\hat{}) 2 \sum (y\,i - \bar{y}) 2$

RESULTS:

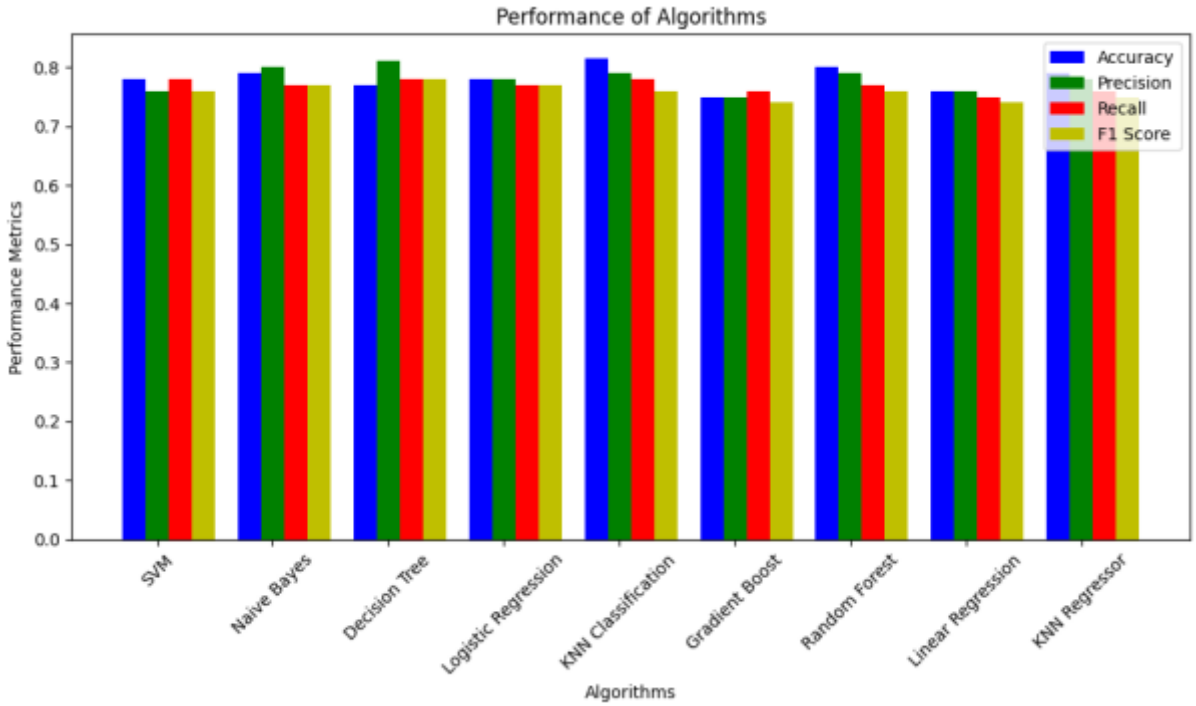
Algorithms / Performance metrics	**Accuracy**	**MSE**
SVM	0.775	-
Naive Bayes	0.782	-
Decision Tree	0.796	-
Logistic Regression	0.787	-
KNN Classifier	0.805	-
Gradient Boost	-	0.12
Random Forest	0.776	-
Linear Regression	-	0.21
KNN Regressor	-	0.21

GRAPH AND INTERPRETAION:



INTERPRETATION

1. **Accuracy:** This metric indicates the overall correctness of the predictions made by each algorithm. Higher accuracy values suggest that the algorithm has made more correct predictions overall. From the chart, we can see that 'Random Forest' has the highest accuracy, followed by 'Naive Bayes' and 'Linear Regression'.
2. **Precision:** Precision measures the proportion of true positive predictions among all positive predictions made by the algorithm. A higher precision value indicates fewer false positives. 'Decision Tree' and 'Random Forest' have the highest precision values among the algorithms.
3. **Recall:** Recall, also known as sensitivity, measures the proportion of true positive predictions among all actual positive instances in the data. Higher recall values indicate fewer false negatives. 'Random Forest' and 'Logistic Regression' have relatively high recall values.
4. **F1 Score:** The F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall. 'Decision Tree' and 'Logistic Regression' have relatively high F1 scores.


Performance of Algorithms
Accuracy
Precision
Recall
F1 Score
Performance Metrics
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
SVM
Naive Bayes
Decision Tree
Logistic Regression
KNN Classification
Gradient Boost
Random Forest
Linear Regression
KNN Regressor
Algorithms