Car Evaluation Classification

Zeyad Emad, Mostafa Ahmed, John Hani, Alaa Mohamed

May 2019

Abstract

The main target of the project is to evaluate cars according to their structure information, so we are using a car evaluation dataset from UCI. Using machine learning techniques we were able to train four classifiers with the dataset data, achieving high accurate results.

1 Introduction

The purpose of this project is to evaluate cars according to their acceptability. Four classifiers were used to solve this problem: KNN, Decision Tree, Random Forest and Logistic Regression. Random Forest ended up being the best of all four according to accuracy achieving 0.976.

2 Dataset

The dataset was obtained from UCI which contains 1728 instances and 6 features which are: buying, maint, doors, persons, lug boot and safety. Before sending the data to train the classifier we label encoded all the feature values. Finally, to measure the acceptability of the car, the Class values are: unacc, acc, good and vgood.

3 Experimental Results

We split the dataset into ratios (0.8,0.2) for train and test. In Table 1 we describe the results achieved by the four classifiers: KNN (where k=7), Decision Tree, Random Forest (where the number of trees are 100) and Logistic Regression achieving: 0.930, 0.954, 0.976 and 0.709 in accuracy, respectively, while achieving: 0.929, 0.909, 0.899 and 0.399 in precision, respectively and finally, achieving 0.877, 0.866, 0.944 and 0.351 in recall, respectively. Figure 1 depicts the accuracy results of four classifiers. Figure 2 depicts the F1 score results of four classifiers.

Classifier	Accuracy	Precision	Recall	F1-Score
KNN(k=7)	0.930	0.918	0.825	0.869
Decision Tree	0.954	0.909	0.866	0.886
Random Forest	0.976	0.899	0.944	0.920
Logistic Regression	0.709	0.399	0.351	0.373

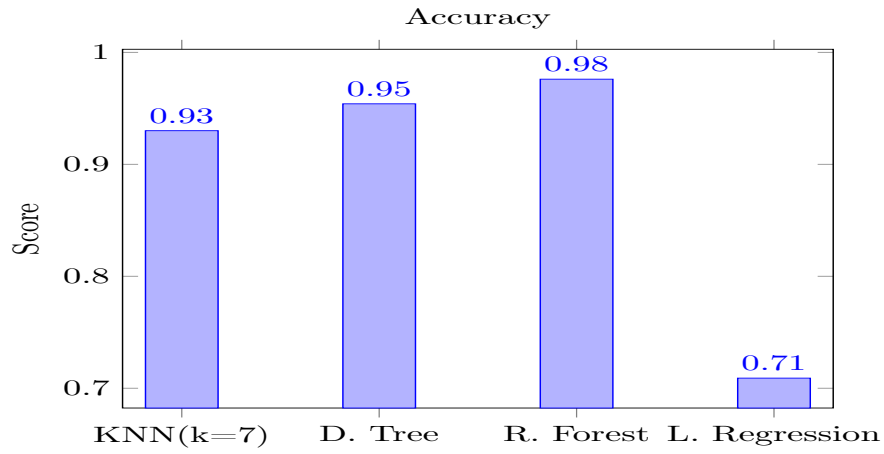Table 1: The Result of all four classifiers

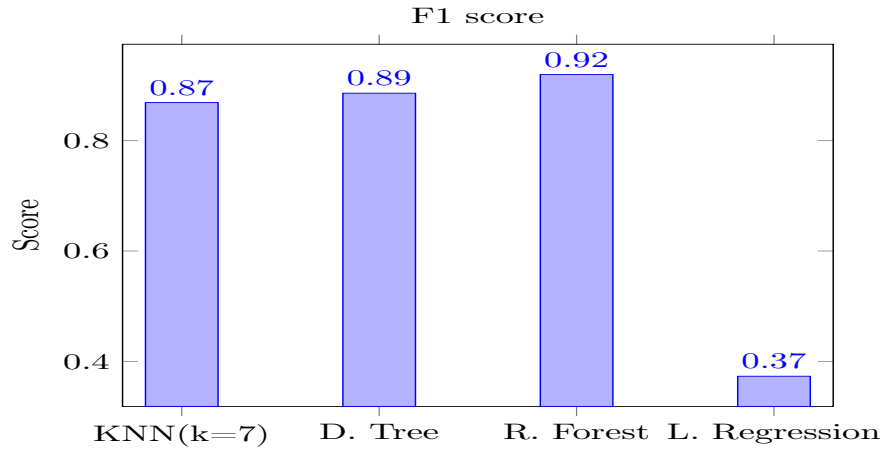Figure 1: COMPARISON BETWEEN ALL FOUR CLASSIFIERS IN TERMS OF ACCURACY

Figure 2: COMPARISON BETWEEN ALL FOUR CLASSIFIERS IN TERMS OF F1 SCORE