structured-abstract-ammar-150454388

January 25, 2019

0.1 Context

Pigment skin diagnosis field has been growing due to the importance of early detection of skin conditions like skin cancer, the field has developed various detection methods that include computerised algorithms to help with this.

0.2 Objective

Hence, this project aim is to develop a machine learning model for the diagnosis of pigmented skin lesions. The project objectives include developing and training a models with a diverse and large dataset, and using these models with test data to evaluate their discriminatory and non-discriminatory performance to find a final model.

0.3 Method

To achieve this, the project follows the CRISP DM approach for the data mining process, specifically dividing the process up to stages and reports. The process will be used to construct and evaluate the models.

0.4 Results

The results from the process show that out of the selected three models, the logistic regression seem to perform best in the discriminatory and non-discriminatory categories, getting better accuracies, run times and confusion matricies. However, the results were still poor when it came to the consistency of classification, and the results of the other might have been effected heavily by the difficulty of tuning the SVM and the Random Forest due to poor run times.

0.5 Novelty

Other literature usually covered discriminatory performance of models but glanced over non-discriminatory measures like fitting times and prediction times, which this project considered a priority. Moreover, this makes use of diverse and large dataset with more classes than similar projects which usually cover around 3-2 classes.
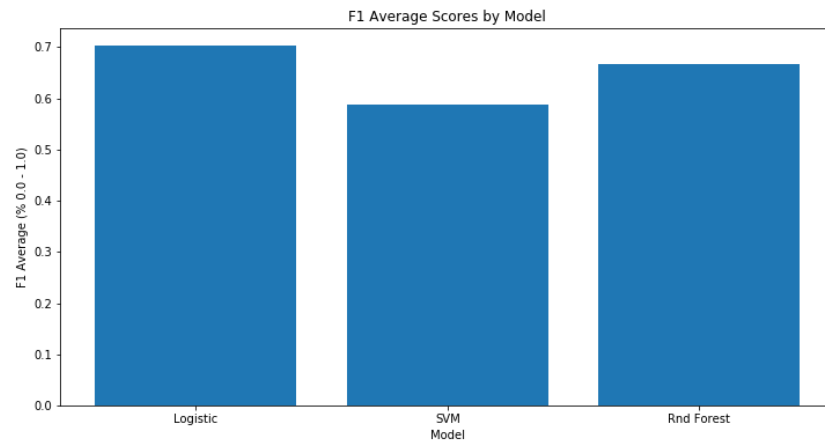
The f1 scores of SVM and Random Forest are not better even after poor run times
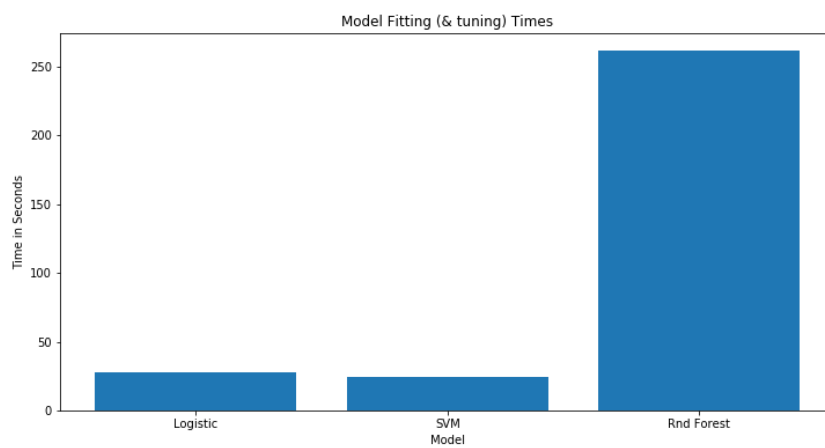
The fit times of Random Forest were slow making tuning difficult