**Raisin Classification**

**Objective**

The factors that determine the type of fruit are the external appearance features such a

The external appearance of the fruits is a major determinant of the fruit type. Determining the variety of fruits by looking at their external appearance may necessitate expertise, which is time-consuming and requires great effort. The aim of this study is to classify the types of raisin fruit, that are, Kecimen and Besni by using three different machine learning methods.

**About the dataset**

In accordance with this purpose, the dataset was obtained from the following website - <https://archive.ics.uci.edu/ml/datasets/Raisin+Dataset>.

There are a total of 8 features and 900 samples in this dataset. For this analysis, ‘Class’ was taken as the target feature. This had 2 attributes - the Kecimen and Besni. The other features were Area, Major Axis Length, Minor Axis Length, Eccentricity, Convex Area and Extent Perimeter. Except for the ‘Class’ feature, all the other features were numerical variables.

**Model application**

The processing of the dataset and the application of the machine learning models along with analyses, has been done in the Jupyter notebook, which is also attached with this report.

**Results**

The Raisin dataset did not show a clear-cut pattern in the correlation analysis and also the features generally had a normal distribution.

There were 3 different machine learning models which were fitted and tested on this dataset. Which were Logistic Regression, Random Forest Classifier and XGB Classifier. The accuracy score of the 3 model were generally in the range of 82-84%, with a closer look shows that both Logistic Regression and Random Forest Classifier has the same accuracy score of 83.7%. Which means a simple Logistic Regression model is good enough for this classification.

Hence, it has been concluded that machine learning methods can be applied successfully for the classification of date fruit types.