Task

Input:

You have been given a link to a csv file containing features and labels. The final column represents the label (as a string) and all other columns represent different hand crafted features (floating point values).

Problem Statement:

You have to train different classifiers on this dataset and predict the labels correspoding to the feature vectors.
Not all of the features are contrbuting for an improvement in final classification accuracy. Your task is to try different training algorithms and features selection approaches to show which algorithm suits the task best and which features best define the dataset.
Divide the dataset into train and validation set with 80/20 ratio and observe the validation accuracy for different algorithms.

Output:

Share your code (preferably a jupyter notebook) with some visualization of your resultant accuracies and a brief report over the feature selection process describing which features were most useful and the ones which didn’t contribute much.