

1 Random forests

1.1 Wine-dataset

For the wine dataset, I selected a random subset of size 75% of the original data for training. I did not ensure that the samples in the training set are equally distributed for each label. While this is sub-optimal as mentioned by the instructor on Piazza, I think the split of 33%, 40%, 27% amongst the classes is not substantial to adversely affect the results.