ECE657A Assignment 2
Submitted by :
Amandeep Kaur(21044104)
Bhupesh Dod (21046099)

Question 4 : Random Forest
Solution:

4.1 Abalone Data

4.1a Abalone Raw

The accuracy of Random Forest on training abalone raw with default parameters is 23.54%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 27.18 % with best parameters as 'max_depth': 9, 'n_estimators': 50 and runtime as 0.171 seconds. **Abalone Raw accuracy 27.18 %, 'max_depth': 9, 'n_estimators': 50.**

4.1b Abalone PCA

The accuracy of Random Forest on training abalone PCA with default parameters is 20.00%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 26.79 % with best parameters as 'max_depth': 4, 'n_estimators': 40 and runtime as 0.083 seconds. **Abalone Raw accuracy 26.79 %, 'max_depth': 4, 'n_estimators': 40.**

4.1c Abalone LDA

The accuracy of Random Forest on training abalone LDA with default parameters is 24.21%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 27.94 % with best parameters as 'max_depth': 6, 'n_estimators': 40 and runtime as 0.132 seconds. **Abalone Raw accuracy 27.94 %, 'max_depth': 6, 'n_estimators': 40.**
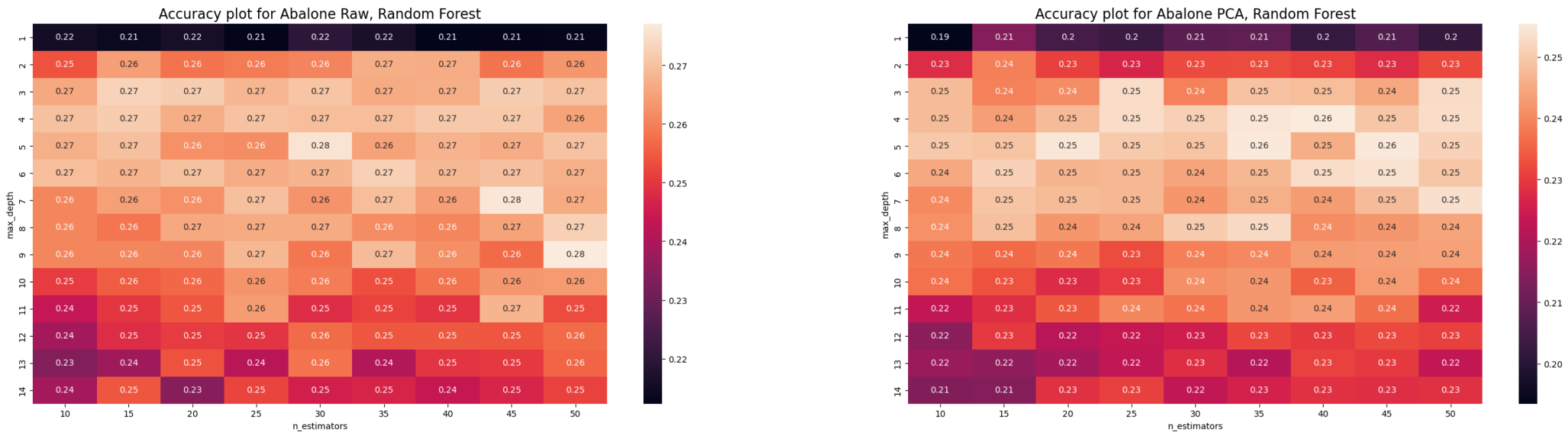
4.1 Wine Data

4.1a Wine Raw

The accuracy of Random Forest on training wine raw with default parameters is 68.49%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 68.92 % with best parameters as 'max_depth': 14, 'n_estimators': 50 and runtime as 0.347 seconds. **Abalone Raw accuracy 68.92%, 'max_depth': 14, 'n_estimators': 50**
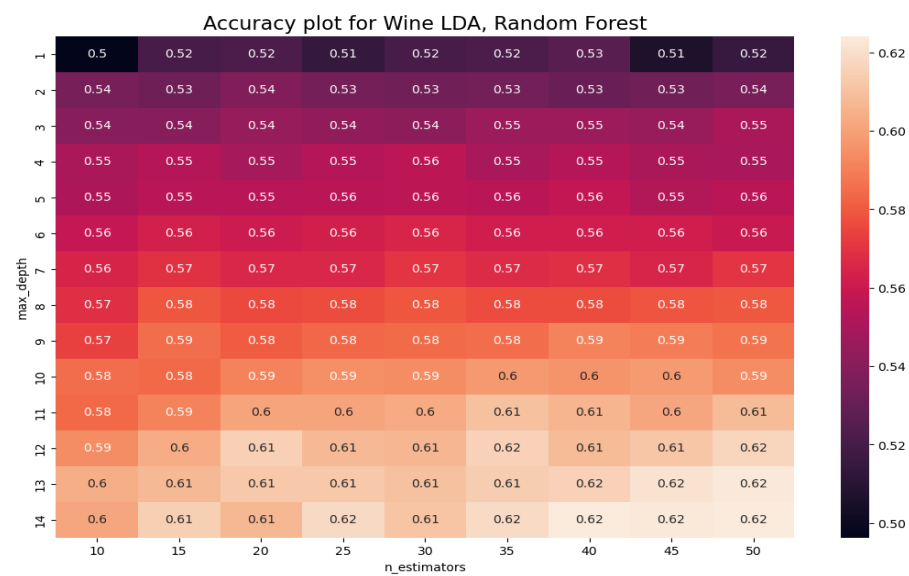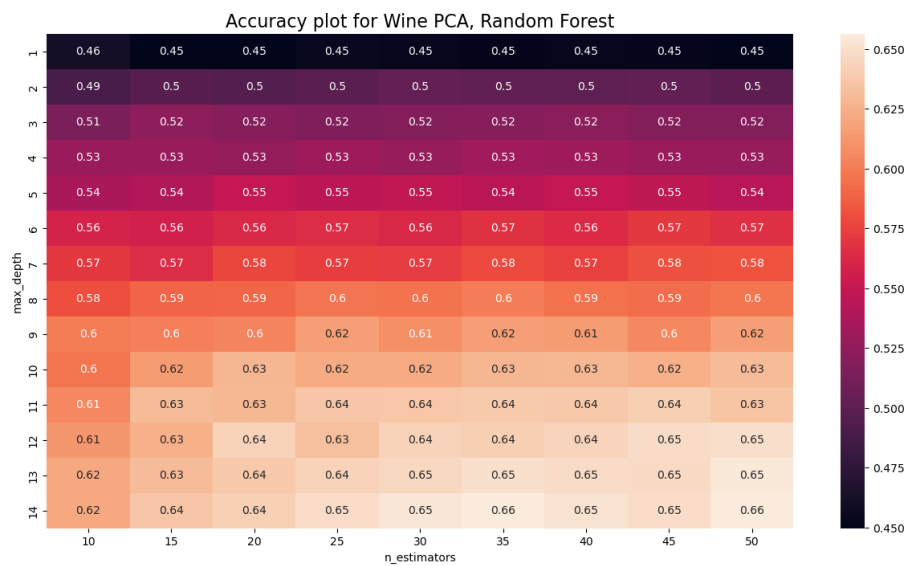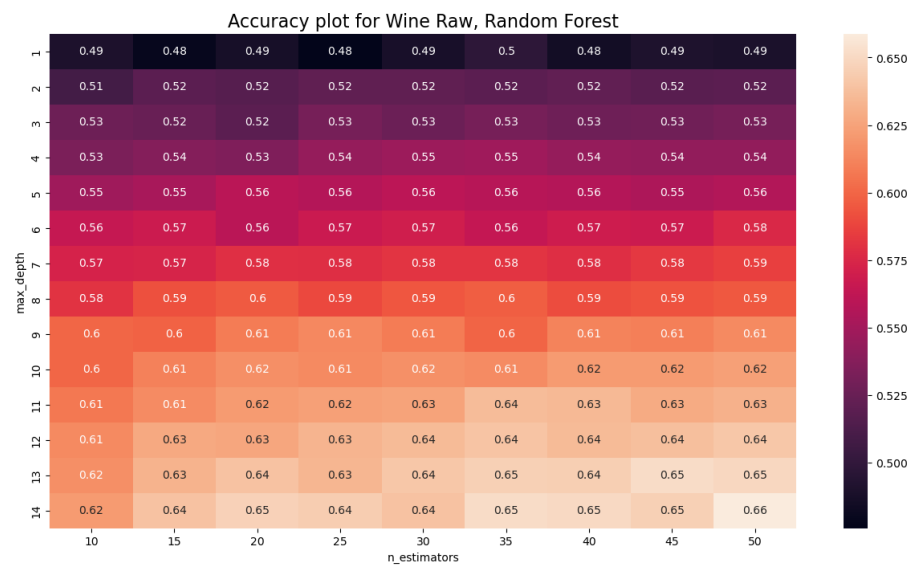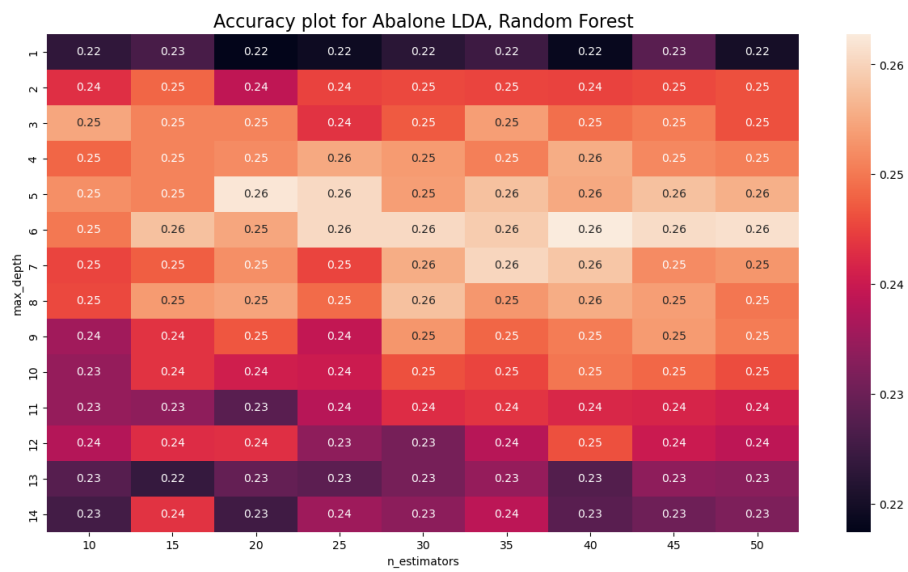
4.1b Wine PCA

The accuracy of Random Forest on wine PCA with default parameters is 67.82%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 67.57 % with best parameters as 'max_depth': 14, 'n_estimators': 50 and runtime as 0.525 seconds. **Abalone Raw accuracy 67.57 %, 'max_depth': 14, 'n_estimators': 50.**

4.1c Wine LDA

The accuracy of Random Forest on wine LDA with default parameters is 66.46%, better than Decision Tree and Naive Bayes. After tuning with GridSearchCV accuracy comes up to 64.49% with best parameters as 'max_depth': 14, 'n_estimators': 40 and runtime as 0.299 seconds. **Abalone Raw accuracy 64.49%, ''max_depth': 14, 'n_estimators': 40.**

Accuracy plot for Abalone Raw, Random Forest

Accuracy plot for Abalone PCA, Random Forest

Accuracy plot for Abalone LDA, Random Forest

Accuracy plot for Wine Raw, Random Forest

Accuracy plot for Wine PCA, Random Forest

Accuracy plot for Wine LDA, Random Forest

These are heatmaps for all the 6 datasets representing the accuracy, max_depth and n_estimators of the random forest. Random forest has worked better on Wine dataset.