CS5011 - Machine Learning
Software Assignment - 2

PROBLEM STATEMENT

The assignment aims at performing classification on the Avila dataset given training and test split, over various classification methods.

TASKS

Two classification models (as specified in Table 1) have to be trained using the training data to predict the class label for the features in the test data. Compare the prediction accuracies between the two prediction models allotted.

INPUT DATA

Avila dataset: Please visit https://bit.ly/2tBceEF for the dataset. The dataset is 10 - dimensional with 10430 and 10437 samples for training and testing respectively. Kindly follow the training and test split already done and saved as separate files on the website itself.

OUTPUT

Compare the classification accuracies for the two methods allotted in Table 1. Justify why one classifier outperforms the other on the given dataset.

HINTS FOR EXCELLENCE

Additional observations and visualizations of the data and the attributes of the trained models will be given extra credit.

Coding language: **Python**

Deadline: 20th April, 2019

Serial No.	Roll No.	Classifier 1	Classifier 2
1	CS18M509	SVM	LDA
2	CS18M510	k-NN	Decision Tree
3	CS18M511	Adaboost	Random Forest
4	CS18M512	LDA	Naive Bayes
5	CS18M513	GMM	k-NN
6	CS18M514	Bayes	Random Forest
7	CS18M515	LDA	Adaboost
8	CS18M516	Decision Tree	SVM
9	CS18M517	Random Forest	Naive Bayes
10	CS18M518	GMM	k-NN
11	CS18M519	LDA	Random Forest
12	CS18M520	k-NN	SVM
13	CS18M522	Decision Tree	LDA
14	CS18M523	Random Forest	k-NN
15	CS18M524	Bayes	GMM
16	CS18M525	SVM	Naive Bayes
17	CS18M526	GMM	Adaboost
18	CS18M527	LDA	Decision Tree
19	CS18M528	GMM	SVM
20	CS18M529	LDA	Adaboost
21	CS18M530	SVM	Naive Bayes
22	CS18M531	Random Forest	Bayes
23	CS18M532	k-NN	GMM
24	CS18M504	GMM	LDA
25	CS18M505	Decision Tree	Random Forest
26	CS18M507	Bayes	SVM
27	CS18M502	Decision Tree	GMM
28	CS18M503	Naive Bayes	LDA