CS631 Cyber Security of Critical Infrastructure
Homework 4 – Report

Sanjeev Lal (19111077)

Avesh Kumar Agrawal (19111020)

Manish Patel (19111051)

Model 1 – Decision Tree

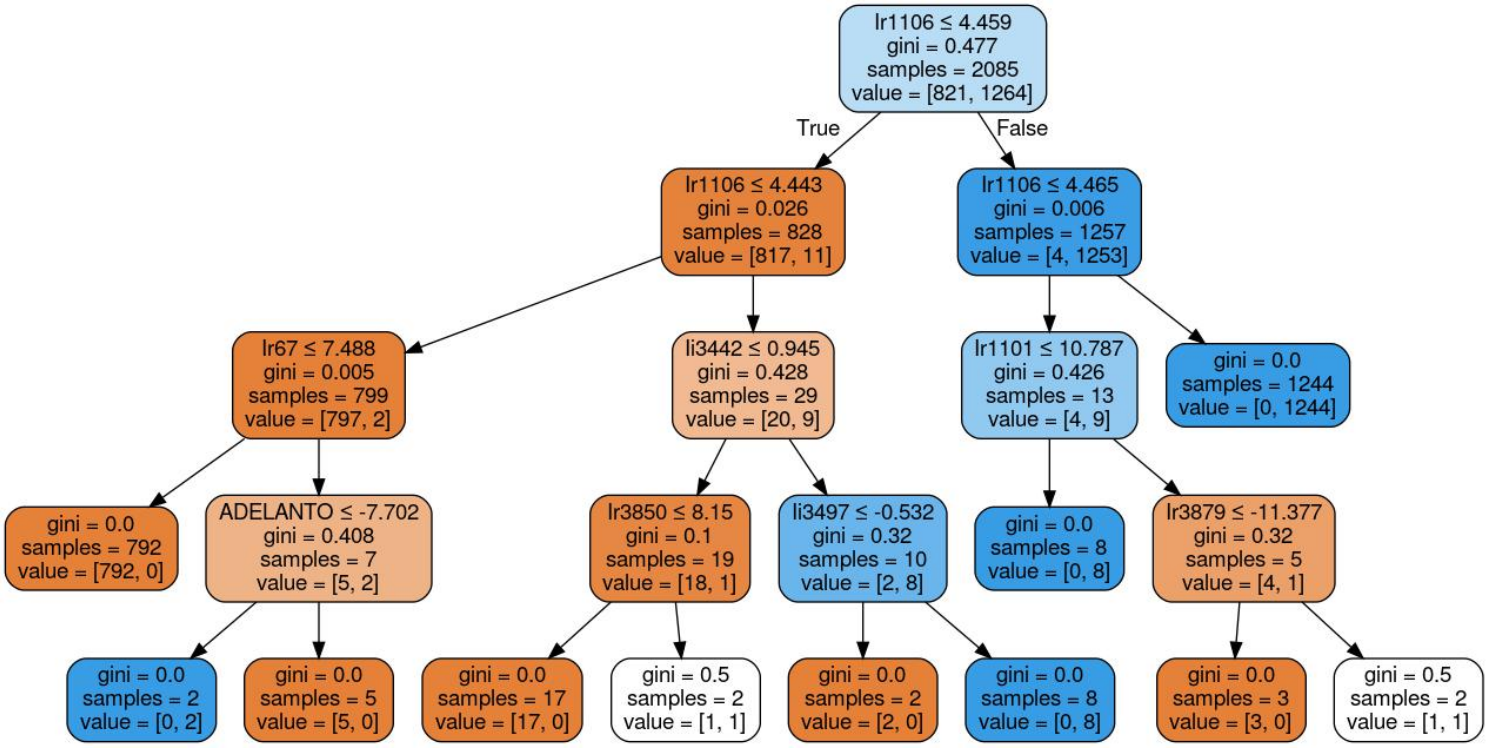
1) We used the built in function of DecisionTreeClasifier() in python.
2) We have trained our model on parameters as criterion='gini', max_depth=6, min_samples_leaf=2 using 50% of the given data.
3) We have validated our model using 25% of the given data.
4) After validation, we have tested our model on the remaining 25% of the given data.

Accuracy	Training time	Testing time
99.61%	0.069 sec	0.008 sec

Confusion matrix

		Actual Label	
		Normal (0)	Stressed (1)
Predicted Label	Normal (0)	410	0
	Stressed (1)	3	625

Decision Tree Diagram



Model 2 – PCA with SVM

1) Principal Component Analysis is used for dimensionality reduction of features. In this model, we used the cumulative variance of 99.90%.
2) We have trained our model on parameters as C=2.5, kernel='linear', gamma='scale' using 50% of the given data.
3) We have validated our model using 25% of the given data.
4) After validation, we have tested our model on the remaining 25% of the given data.

Accuracy	Training time	Testing time	PCA time
98.84%	0.018 sec	0.008 sec	0.102 sec

Confusion matrix

		Actual Label	
		Normal (0)	Stressed (1)
Predicted Label	Normal (0)	399	11
	Stressed (1)	1	627

Conclusion

1) We have found 99.61% accuracy for Decision Tree and 98.84% accuracy for PCA_SVM model. So, for us Decision Tree is performing better than PCA_SVM.
2) But Decision Tree takes more time for training than PCA_SVM.
3) We have also tested PCA with Decision tree. This model was giving bad accuracy.
4) So, **Best** of the above two models is **Decision Tree model**.