Assignment - 1

Feature Names	Decision Tree	Random Forest	XGBoost	Logistic Regression	Linear Regression	Artificial Neural Network	Support Vector Machine	Extra Trees Classifier	Gradient Boosting Classifier
'VWTI', 'SWTI', 'CWTI', 'EI'	0.9818181818181818	0.9954545454545455	0.995454545454555	0.9863636363636363	0.8698345095658888	1.0	0.9954545454545455	0.9954545454545455	0.9909090909090901
'VWTI', 'SWTI', 'CWTI'	0.9772727272727273	0.995454545454555	0.995454545454555	0.9818181818181818	0.8697654431550244	1.0	0.9818181818181818	0.9909090909090901	0.9909090909090901
'VWTI', 'CWTI', 'EI'	0.9590909090909091	0.9681818181818181	0.9772727272727273	0.8681818181818182	0.6017966852050953	0.9772727272727273	0.9636363636363636	0.9681818181818181	0.9681818181818181
'SWTI', 'CWTI'	0.8590909090909091	0.8909090909090909	0.895454545454555	0.7590909090909091	0.30118738376158094	0.8409090909090909	0.8045454545454546	0.9090909090909091	0.8681818181818182
'VWTI', 'EI'	0.8909090909090909	0.9136363636363637	0.9	0.8954545454545455	0.5643929432114736	0.8954545454545455	0.8909090909090909	0.9045454545454545	0.9090909090909091
'VWTI', 'CWTI'	0.8954545454545455	0.9409090909090909	0.9227272727272727	0.9045454545454545	0.5535738816946818	0.9318181818181818	0.9272727272727272	0.9227272727272727	0.9454545454545454
'EI'	0.5636363636363636	0.5681818181818182	0.5636363636363636	0.5227272727272727	0.0009354564051183845	0.5227272727272727	0.5318181818181819	0.5590909090909091	0.5090909090909099
'SWTI'	0.7181818181818181	0.7227272727272728	0.7363636363636363	0.6272727272727273	0.20035249780055941	0.7454545454545455	0.75	0.7227272727272728	0.75
'CWTI'	0.6363636363636364	0.6409090909090909	0.6181818181818182	0.6045454545454545	0.027782882448816666	0.5954545454545455	0.5727272727272728	0.6454545454545455	0.6636363636363637
'VWTI'	0.8636363636363636	0.8681818181818182	0.85	0.8681818181818182	0.5378664272892191	0.8772727272727273	0.8772727272727273	0.8636363636363636	0.8727272727272727
Total Score	8.345454545	8.504545455	8.454545455	8.018181818	4.527488111	8.386363636	8.295454545	8.481818182	8.468181818

From the table given above we can infer that, when all of the features are included most model prediction scores are closer to 1 or 100%. When randomly some of the features are dropped their prediction score also drops. Having said that, the single most feature from the table that seems to be the most reliable predictor of higher score is the ‘VWTI’, because when using only this all models give higher scores compared to any other instances where only a single feature is used.

When looking at the sum of all the predictions made by each model for their given feature set we can definitely say that the highest score is achieved by Random-Forest-Classifier model followed closely by the Extra-Trees-Classifier model and when using all the features we can see that Artificial-Neural-Network gives us the most accurate prediction as the score is 1.0.

In conclusion, if we were to use all the features given in the task we should use the ANN model and if we were to drop some of the features we might get better results from using the Random-Forest model.