Assignment – Machine Learning [Major]

**Answer Sheet**

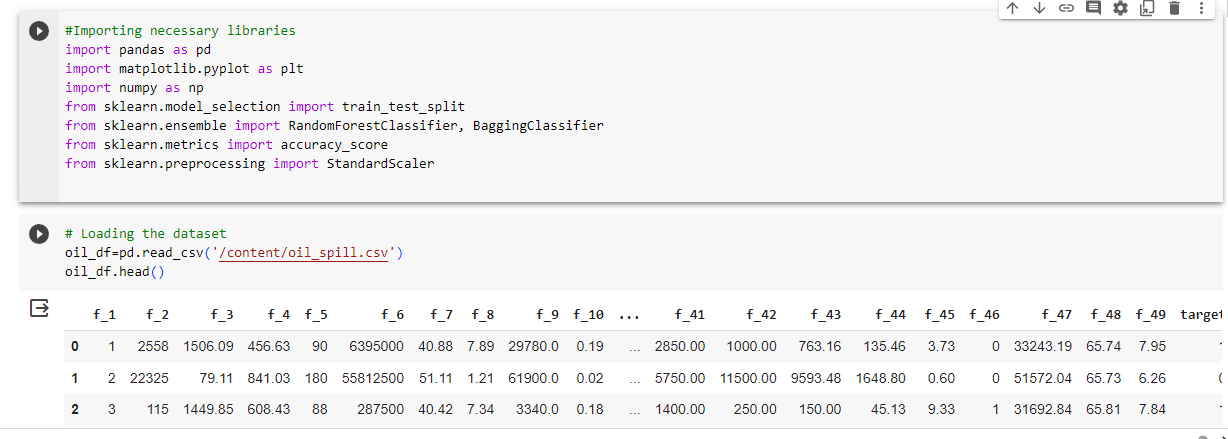
Name: Priyanka Arora

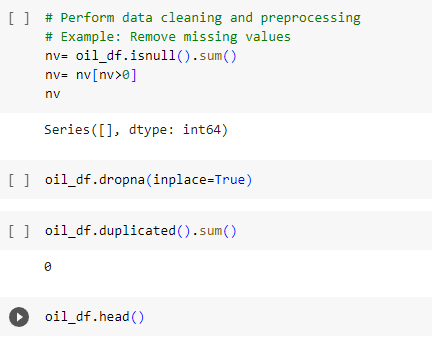
Phone No.: 9149113191

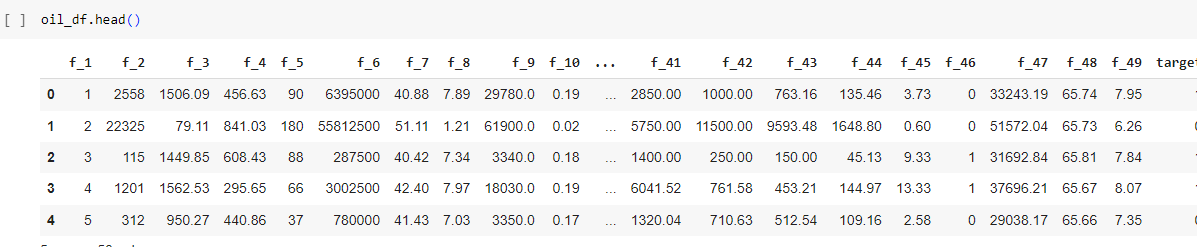
Email: [arora09011987@gmail.com](mailto:arora09011987@gmail.com)

Batch: - 9

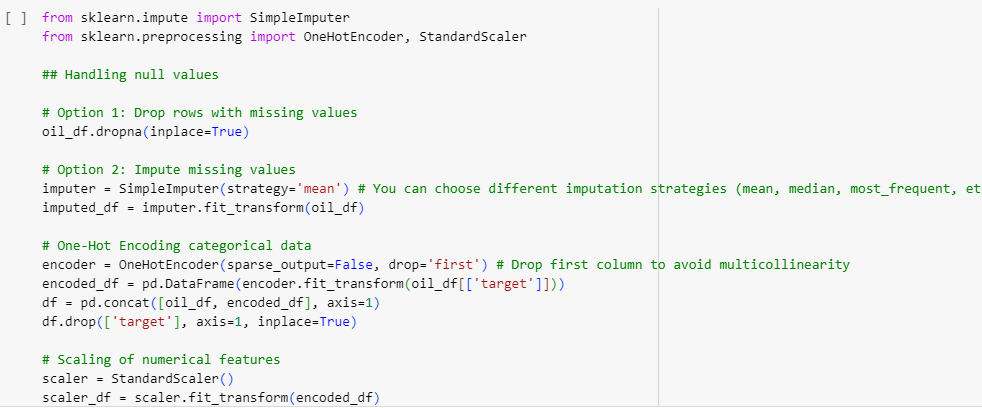
**Question 1. Download the Oil Spill Dataset and perform Data cleaning and Data Pre-Processing if Necessary**

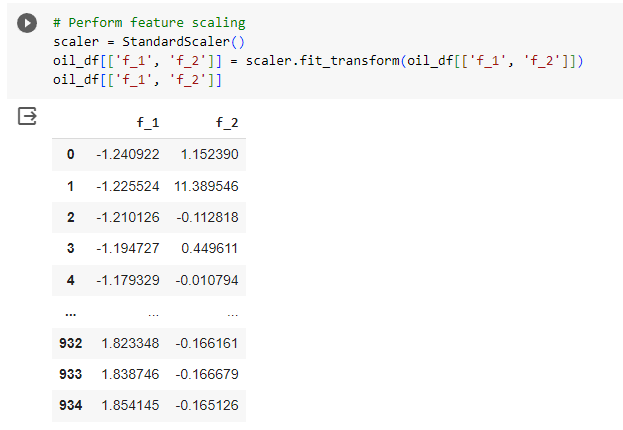


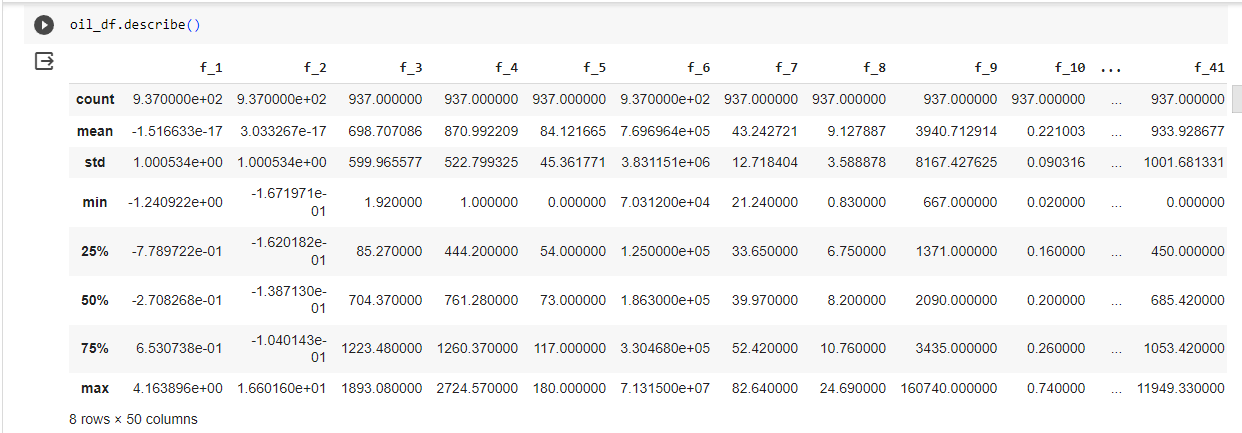




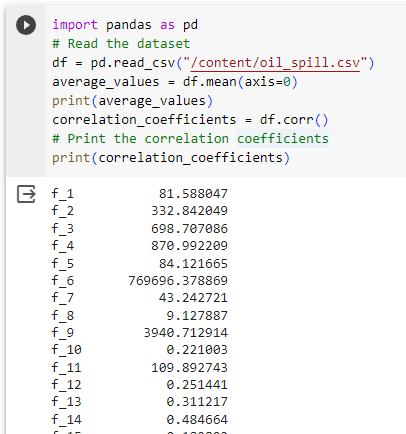
**Question 2) Use various methods such as Handling null values, One-Hot Encoding, Imputation, and Scaling of Data Pre-Processing where necessary.**

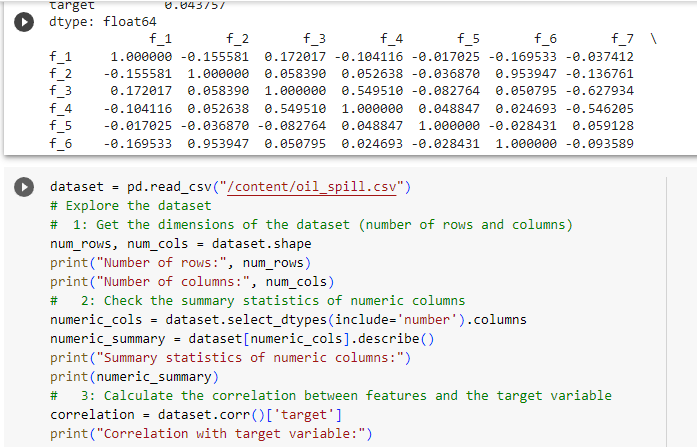
****

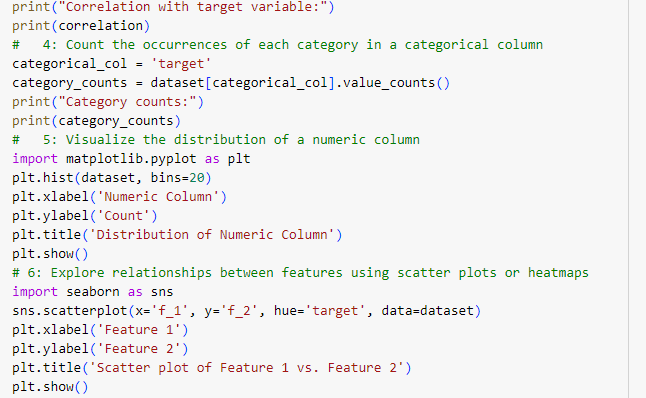
****

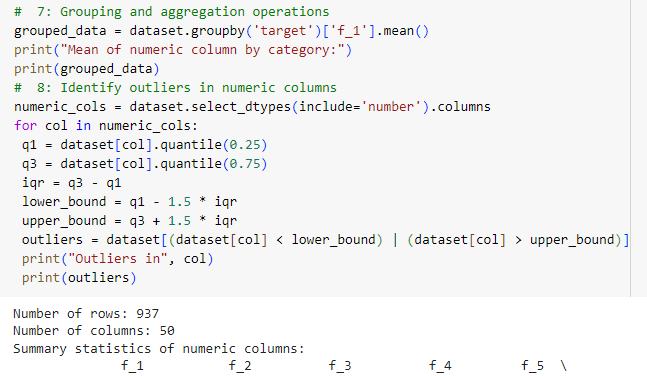
****

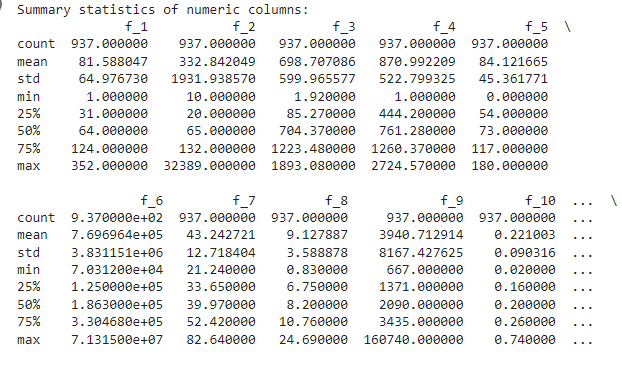
**Q.3. Derive some insights from the dataset.**

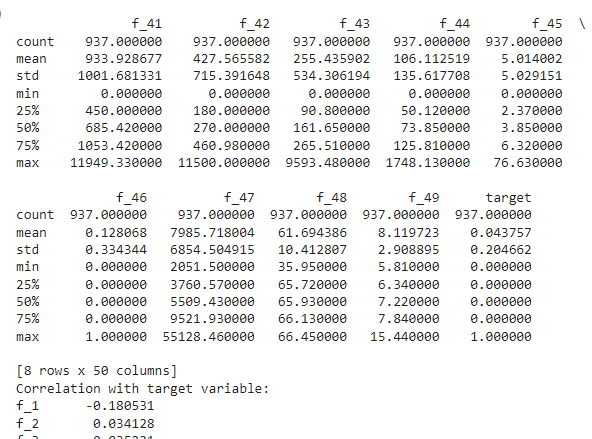
****

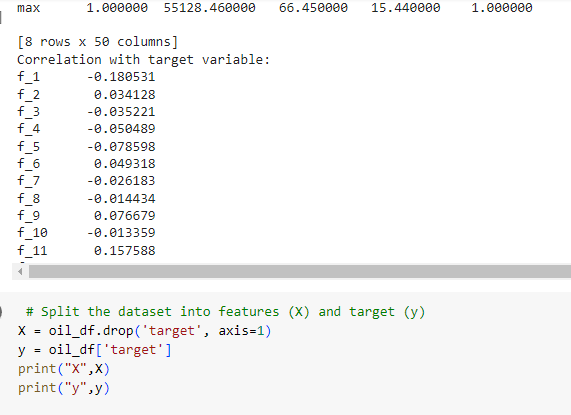
****

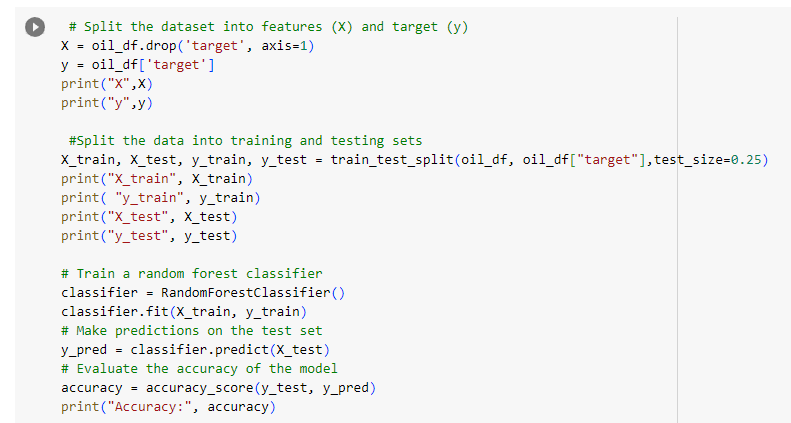
****

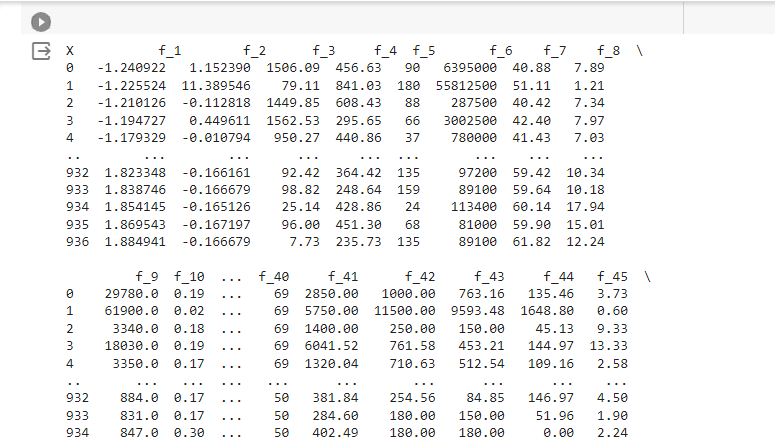
****

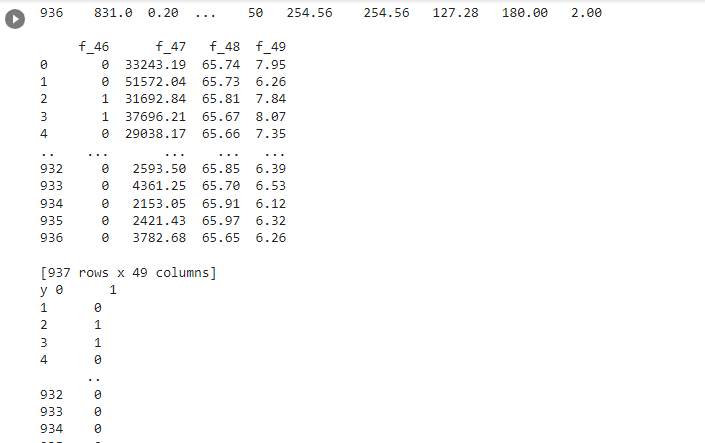
****

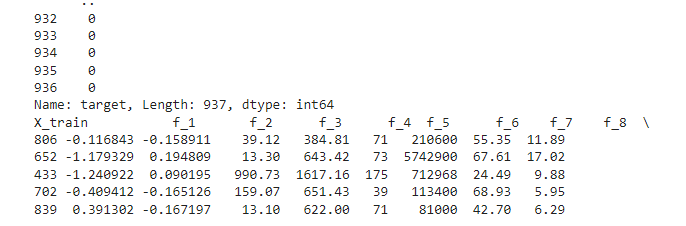
****

****

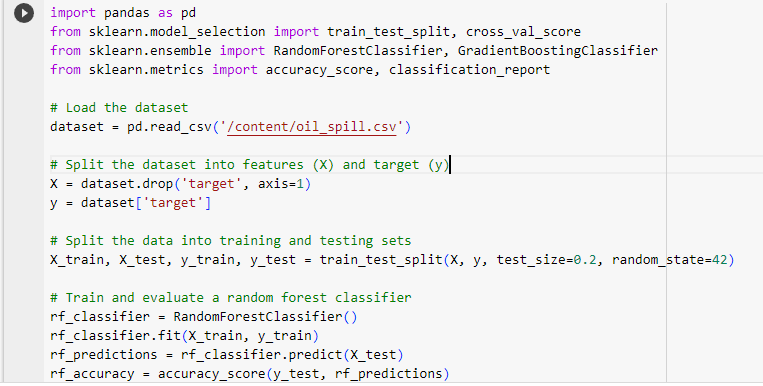
****

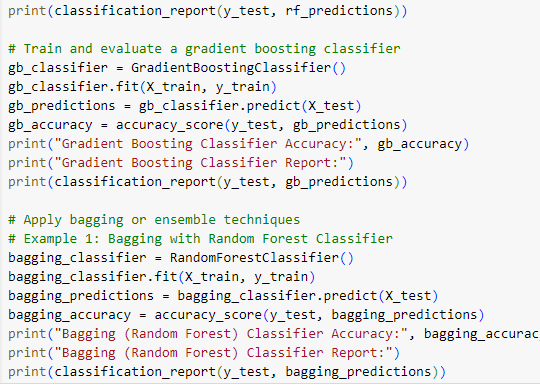
****

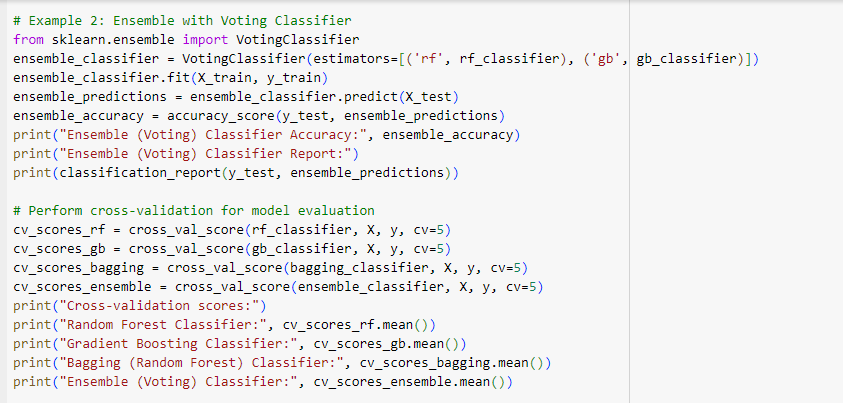
****

****

**Question.4. Apply various Machine Learning techniques to predict the output in target column, make use of Bagging and Ensemble as required and find the best model by evaluating the model using Model evaluation techniques**

****

****

****

Random Forest Classifier Accuracy: 0.973404255319149

Random Forest Classifier Report:

precision recall f1-score support

0 0.98 0.99 0.99 182

1 0.67 0.33 0.44 6

accuracy 0.97 188

macro avg 0.82 0.66 0.72 188

weighted avg 0.97 0.97 0.97 188

Gradient Boosting Classifier Accuracy: 0.9787234042553191

Gradient Boosting Classifier Report:

precision recall f1-score support

0 0.99 0.98 0.99 182

1 0.62 0.83 0.71 6

accuracy 0.98 188

macro avg 0.81 0.91 0.85 188

weighted avg 0.98 0.98 0.98 188

Bagging (Random Forest) Classifier Accuracy: 0.9680851063829787

Bagging (Random Forest) Classifier Report:

precision recall f1-score support

0 0.98 0.99 0.98 182

1 0.50 0.33 0.40 6

accuracy 0.97 188

macro avg 0.74 0.66 0.69 188

weighted avg 0.96 0.97 0.96 188

Ensemble (Voting) Classifier Accuracy: 0.9787234042553191

Ensemble (Voting) Classifier Report:

precision recall f1-score support

0 0.98 1.00 0.99 182

1 1.00 0.33 0.50 6

accuracy 0.98 188

macro avg 0.99 0.67 0.74 188

weighted avg 0.98 0.98 0.97 188

Cross-validation scores:

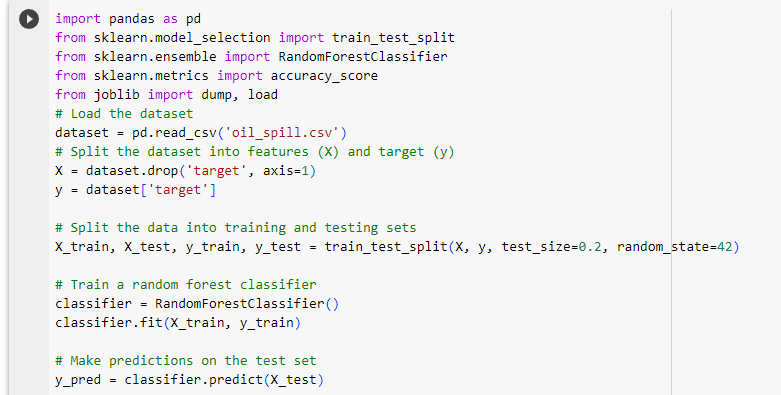
Random Forest Classifier: 0.9391569006712936

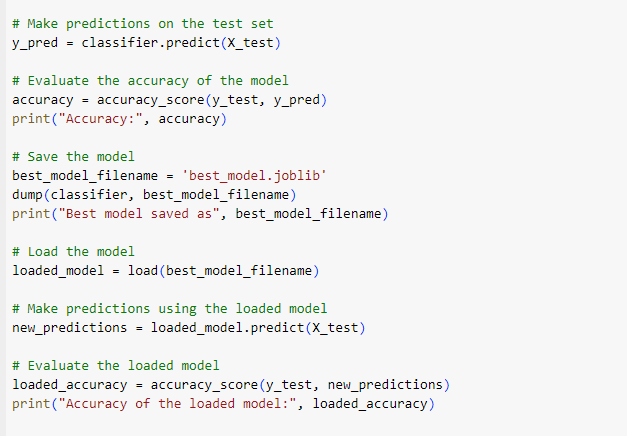
Gradient Boosting Classifier: 0.9198714301968369

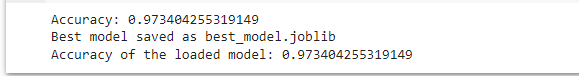
Bagging (Random Forest) Classifier: 0.9412902491751053

Ensemble (Voting) Classifier: 0.9434406644669474

Question 5. Save the best model and Load the model







# ****Q.6. Take the original data set and make another dataset by randomly picking 20 data points from the oil spill dataset and apply the saved model on the same****

