Slot-2 report

3. Build the Supervised Learning Models:

→The chosen supervised Learning Models include:

a) Logistic Regression

b) Random forest

c)Gradient Boosting/Ada Boosting

a) Logistic Regression:

Logistic Regression is used for binary classification problem where the goal is to predict a binary outcome that is if the client subscribed a term deposit or not.

b) Random forest:

Random Forest can handle non-linear relationships between features and the target variable. Customer behaviour can be highly non-linear and influenced by various factors which is captured by random forest.

c)Gradient Boosting/Ada Boosting:

As the number of clients who subscribe to term deposits to be lower than those who don't gradient Boosting can mitigate the impact of class imbalance.

4. Choice of Evaluation Metric:

→AUC is chosen to find the probability to discriminate between subscriber and non-subscriber. This metric can be useful when the bank aims to prioritize clients in their marketing campaigns based on their probability scores.

5.Overfitting avoidance mechanism:

→Elastic Net regularization is chosen.

→ The marketing data can have a wide range of features some of them might be highly correlated or less relevant. Elastic Net can help in automatic feature selection by assigning zero coefficients to less important features, reducing model complexity and overfitting.

→ Elastic Net adds penalty terms to the loss function discouraging the model from assigning too much importance to any one feature, making it less prone to overfitting.

6. Model Hyperparameter Tuning:

→Grid Search systematically explores a predefined set of hyperparameter combinations conduct a comprehensive search to find the best set of hyperparameters.

→Since the hyperparameter space is relatively small, Grid Search is a practical and effective approach. You can specify a manageable number of hyperparameter values to search over.

7. Results analysis:

a. Which of the 3 models would you recommend for deployment in the real-world?

Random forest would be recommended for deployment in the real-world as it gives a AUC score with highest accuracy

b. Is any model underfitting?

No model is underfitting.