Dear Participant,

You need to build a classifier model that can determine whether the customer will leave or not based on some given data of a Bank customer.

The case study is from an open-source dataset from Kaggle.

Link to the Kaggle project site:

<https://www.kaggle.com/barelydedicated/bank-customer-churn-modeling>

The points distribution for this case is as follows:

1. Read the dataset. (2 points)

2. Check info and shape of the dataset and understand the dataset. (2 points)

3. Drop the columns which are unique for all users like IDs (3 points)

4. Use describe function to get the summary of the data. (2 points)

5. Check value counts and encode categorical columns (Gender, Geography). (3 points)

6. Check distribution of target column and mention your comments. (3 points)

7. Distinguish the feature and target set (2 points)

8. Divide the data set into training and test sets ( 2 points)

Build models -

9. Standardize the train and test data (3 points)

10. Initialize kNN & build the model using Standardized data. (3 points)

11. Plot elbow plot for odd k values ranging from 1 to 30. (5 points)

12. Mention your insights. (3 points)

13. Build a decision tree model check confusion metrics. (6 points)

14. Build a logistic model and check confusion metrics. (6 points)

15. Document your inferences and findings. (5 points)

Regards,

Program Office