Hackathon Batch 12

Important Notes:

Implement All the algorithms we have learned so far. Don't miss any thing.
This is a precious chance to implement all the things we have learnt so far.
Three Candidate will bag the scholarship

Scholarship is decided upon

1. **The Business insights you derive from EDA(Tableau Chart)**
2. **Preprocessing Steps you have taken**
3. **Implementing all the ML algorithms and ensemble technique we have learnt so far**
4. **The way you tell the story while you present it to me.**

Project – 1
Credit Score Prediction

Data

We provide you with a data set in CSV format.
The data set contains 8,000 train instances and 2000 test instance
There are 304 input features, labeled x001 to x304.

The target variable is labeled y.

Task Create a model to predict the target variable y.

1. A report - A Power point presentation
2. Any custom code you used
3. Instructions for me to run your model on a separate data set

What should be in the report?

1. A consolidated EDA report using Tableau (copy the chart and save it in the PPT which ever is insightful)
2. List of any assumptions that you made
2. Description of your methodology and solution path
3. List of all algorithms and techniques you used
4. List of tools and frameworks you used
5. Results and evaluation of your models
6. Final bar chart with all the model performance with both train and test result RMSE

How to evaluate the model
1. Use the Root Mean Square Error (RMSE).

2. If the absolute error of a prediction is greater than 4.0, I regard the prediction as "wrong". Otherwise, it is "correct".

Project – 2
Credit Card Fraud Detection

Data

We provide you with a data set in CSV format.
The data set contains 2 lakhh+ record train instances and 56 thousand test instance
There are 31 input features, labeled V1 to V28 and Amount .

The target variable is labeled Class.

Task Create a Classification model to predict the target variable Class.

1. A report - A Power point presentation
2. Any custom code you used
3. Instructions for me to run your model on a separate data set

What should be in the report?

1. A consolidated EDA report using Tableau (copy the chart and save it in the PPT which ever is insightful)
2. List of any assumptions that you made
2. Description of your methodology and solution path
3. List of all algorithms and techniques you used
4. List of tools and frameworks you used
5. Results and evaluation of your models
6. Final bar chart with all the model performance with both train and test result F1 Score

How to evaluate the model
1. Use the F1 Score for metrics

Project – 3
Income Prediction

Data
We aim to predict whether an individual's income will be greater than $50,000 per year based on several attributes from the census data. The US Adult Census dataset is a repository of 48,842 entries extracted from the 1994 US Census database.

The target variable is labeled as income.

Convert the column to 1 if income >=50k , else 0

Task Create a Classification model to predict the target variable income.

1. A report - A Power point presentation
2. Any custom code you used
3. Instructions for me to run your model on a separate data set

What should be in the report?

1. A consolidated EDA report using Tableau (copy the chart and save it in the PPT which ever is insightful)
2. List of any assumptions that you made
2. Description of your methodology and solution path
3. List of all algorithms and techniques you used
4. List of tools and frameworks you used
5. Results and evaluation of your models
6. Final bar chart with all the model performance with both train and test result Accuracy

How to evaluate the model

1. Use the Accuracy score