Associate DA Capstone Project

Taste of real-life business problem.

Agenda

- Objective
- Data Information
- General Process
- Presentation Format
- Marking

Objective

- This project will allow you to integrate and apply the knowledge, skills and concepts that you have acquired throughout the academic journey.

- It will allow you to demonstrate your understanding of key concepts, methodologies and best practices.

- It can also help you in professional development by fostering skills such as time management, teamwork and problem solving.

Data Information

- The dataset provided for this project is a subset of data our team worked on for a project of **Bajaj Finserv**. To adhere to privacy polices, certain columns have been deleted from the dataset.

- This dataset contains all the required detail such as Loan type, Loan amount, Interest rate, Loan term, Credit score etc.

- All the information about the features is present in **Data_Dictionary.pdf**

General Process

- Problem Statement
- Research Objective
- Data Understanding and Preparation
- Result Interpretation
- Conclusion

General Process

Problem Statement:

In the financial industry, it is crucial for lenders to assess the creditworthiness of borrowers before granting loans or credit. Identifying potential defaulters, who are at higher risk of failing to repay their debts, can help mitigate financial losses and maintain a healthy lending portfolio. The goal of this project is to develop a predictive model that can accurately classify borrowers as defaulters or non-defaulters based on various financial and demographic factors.

General Process

Research Objective:

- Your task is to create a machine learning model to predict the defaulter and Non-defaulter by analyzing historical data.

- Feel free to explore the dataset and produce the solution as per your understandings.

General Process

Data Understanding and Preparation:

- First step towards the analysis is to understand and prepare the data you have, once you get the idea about the features given in the dataset and the relationship among the variables then you can start preparing the dataset for the data analysis and model building.

- Data preparation is important because you need to transform the data as per the requirements like imputing missing values, changing the data type, deriving a custom column etc.

General Process

Result Interpretation:

- Once the model is ready, you must test it and check its accuracy.
- In the model creation process try different models and techniques to improve the accuracy.
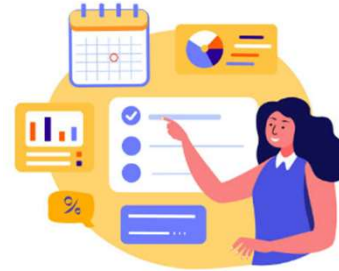- Interpret the result of the models.

General Process

Conclusion:

- Overall accuracy of the model.
- The features in the dataset which are important in predicting the target.
- Any recommendations.

Presentation Format

- Create the model using Jupyter Notebook (IDE).

- It would be better if you include the slides with the full process of your analysis so it looks more professional and would be clear to the audience.

- Along with the PowerPoint Presentation file (.ppt), you must include the Jupyter Notebook file (.ipynb) for the submission.

Marking Criteria

- Problem Statement Understanding   - 10 Marks
- Technical Skills and Proficiency   - 50 Marks
- Result Interpretation and Recommendations   - 20 Marks
- Presentation and Communication   - 20 Marks

Let's do the magic with data?

Thank you !!

I welcome your questions.