Financial: Analysis on Customer data

Rahul Gupta - DSC 680 - Winter 2020

https://github.com/rahulgupta271/DSC680_Project_2_Financial_Analysis_Customer_data

Project 2 Questions

1. **Which Domain is this Project in?**

Financial

2. **Which UI are you using for your Project Development?**

I've been using Python (Jupyter Lab) to explore the data.

3. **What is JupyterLab?**

JupyterLab is a next-generation web-based user interface for Project Jupyter.JupyterLab enables you to work with documents and activities such as Jupyter notebooks, text editors, terminals, and custom components in a flexible, integrated, and extensible manner.

4. **What is CRISP DM?**

The CRoss Industry Standard Process for Data Mining (CRISP-DM) is a process model with six phases that naturally describes the data science life cycle. It's like a set of guardrails to help you plan, organize, and implement your data science project.

5. **Is it a GITHUB link for this project?**

Yes , You can find all the details belongs to this project in GITHUB.

https://github.com/rahulgupta271/DSC680_Project_2_Financial_Analysis_Customer_data

6. **What is module in Python?**

Modules refer to a file containing Python statements and definitions. A file containing Python code, for example: example.py , is called a module.

7. What Python Modules are you using for this project?

- pandas
- numpy
- sklearn
- matplotlib
- seaborn
- joblib
- os

8. What is the business goal for this project?

Goal for this project is to understand the customer behavior analyzing the Data and make some move to help retain the customers.

9. What is model Training?

Training a model simply means learning (determining) good values for all the weights and the bias from labeled examples. In supervised learning, a machine learning algorithm builds a model by examining many examples and attempting to find a model that minimizes loss; this process is called empirical risk minimization.

10. What is the data quality checks?

Data is of high quality when it satisfies the requirements of its intended use for clients, decision-makers, downstream applications and processes. The quality of the data is an important attribute that could drive the value of the data and, hence, impact aspects of the business outcome, such as regulatory compliance, customer satisfaction, or accuracy of decision making.