

Group Name: The Awesome Sauce Data Scientist

Name: Aryaman Srivastava

Email: asriva10@stevens.edu

Country: US

College: Stevens Institute of Technology

Specialization: Computer Science

Problem Description:

ABC Bank wants to sell its term deposit product to customers and before launching the product they want to develop a model which helps them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

Business Understanding:

Objective is to create a Machine Learning model to shortlist customers who are more likely to buy their product based on several features, including: age, job, marital status, education, dates of contact, types of loan etc.

Product LifeCycle

7/26/24	Data Intake Report
8/2/24	Data Understanding and Proposal of Approach
8/9/24	Data Cleaning and Transformation
8/16/24	EDA
8/23/24	Business Presentation of EDA
8/30/24	Build and Test Different Models
9/6/24	Final Model and Solution Presentation

Data Intake Report

Name: Final Project

Report date: 7/26/24

Internship Batch: LISUM34

Version: 1.0

Data intake by: Aryaman Srivastava

Data storage location: <https://github.com/AryaSriva/FinalProjectDataGlacier>

Tabular data details:

Total Number of Observations	86401
Total Number of Files	2
Total Number of Features	20
Base format of the file	CSV
Size of the data	10.202 MB

Proposed Approach:

- Remove duplicate rows(rows have all the same values for all the features) and N/A rows(rows with one or more Null values for the features)
- Assumptions
 - All values for the month and day columns are valid months/days
 - The duration feature is expressed in seconds
 - Day of the week values are valid(MTWTFSS)