

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

Name: Suvendu Nayak

Email: kanhasuvendu@gmail.com

1. Suvendu Nayak (kanhasuvendu@gmail.com) Contribution:

- Outlining project plan.
- Data wrangling and maintaining data integrity.
- EDA.
- Target Variable.
- Feature Engineering.
- Class Imbalance.
- Training and Testing Model.
- Hyperparameter Tuning.

Please paste the GitHub Repo link and drive link.

Github : <https://github.com/SuvOnGithub/Bank-Marketing-Effectiveness-Prediction>

Drive Link- <https://drive.google.com/drive/folders/1Sbmg6ZChesZkwTW-dwIUdtq6YCOM55hl?usp=sharing>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Data from a marketing campaign run by Portugal is examined. The campaign's aim was to increase customers' subscription rates to fixed-term deposit products. Using knowledge a number of machine learning algorithms are implemented to answer the question: How can banks successfully market these products in the most efficient way possible and with the highest possible rate if success?

We followed step by step process for the project like data collection, data cleaning, EDA, Visualization, Model Training and Testing, Hyperparameter Tuning and Evaluation.

In the first step we collected data and explored data set to get a rough idea about data. In data wrangling process on raw data, formatted data type of columns and added some columns for our analysis. In this data the target variable was not given so we calculated Target variable.

In EDA we divided analysis into several part to get better idea about data like we checked distribution of numerical and categorical data checked relationship between independent and dependent variable.

After that we perform feature engineering, in feature engineering we created some new feature form available features with the goal of simplifying and speeding up data transformation while also improving model performance. Then we used oversampling technique to handle class imbalance to make classes even for training model.

After feature engineering we selected feature to use to train our model and encoded categorical variables as ML model works with numerical data to do computation. We used label encoding and one hot encoding.

As data is ready, we trained different model to check their performances and performed hyperparameter tuning to improve their performances by GridsearchCV technique. Out of all the model XGboost models gives the best performance.