Project Name – Santander Customer Transaction Prediction

Deadline - 15 Days

Background -

At Santander, mission is to help people and businesses prosper. We are always looking for ways to help our customers understand their financial health and identify which products and services might help them achieve their monetary goals.

Our data science team is continually challenging our machine learning algorithms, working with the global data science community to make sure we can more accurately identify new ways to solve our most common challenge, binary classification problems such as: is a customer satisfied? Will a customer buy this product? Can a customer pay this loan?

Problem Statement -

In this challenge, we need to identify which customers will make a specific transaction in the future, irrespective of the amount of money transacted.

Data Set :

1) test.csv
2) train.csv

Number of attributes:

You are provided with an anonymized dataset containing numeric feature variables, the binary target column, and a string ID_code column. The task is to predict the value of target column in the test set.

Missing Values: Yes

Evaluation Basis

The project will be evaluated on the following basis:

1. The process of building the model should start from basic (i.e. Use of basic algorithms Like Logistic Regression in case of classification and Linear Regression in case of Regression Problems).

2. The use of basic Algorithms aforementioned in Point 1 are mandatory and play's key role in evaluation of the Projects. After the use of basic algorithm, students are free to use any technique/ algorithms as required.

3. Minimum 3 techniques/ algorithms should be tried out and reasoning behind the use of algorithms and accuracy from this algorithms should be the part of your reports.

4. Every model should be supported by reason of acceptance or rejection. Special emphasis on the reasons why student has picked/dropped an algorithm.

5. For Classification the Accuracy metrics need to be considered are AUC, Precision & Recall. Please ensure to report them in the report. You are free to use other metrics in addition to the aforementioned metrics.

6. The student should revise the concepts before starting the project work.

7. The student should be confident enough to explain every concept that is written in the project report.

8. The questions during the mock sessions would not limit to project. Any topic that is covered in the syllabus can be asked during the mock session.

9. If student unable to explain the project report or during the session, it appears that student has copied the solution then zero marks will be given to the student.

10. The code should be written keeping in mind the code file can be run from DOS prompt.

11. The instructions to run the code file should be submitted with the project report.

Deliverables from Candidate

1. Code written in both R and Python.
2. Comprehensive Project Report with Data Visualization for explaining Features and its importance.
3. Instruction to deploy and run code.
4. Summarize the Understanding of How this project can help the business in achieving the strategic goals.

Always remember these evaluation basis, deliverables, and your deadline. And your aim is to meet the deadline.

Warning - Do not submit incomplete projects or projects that are not running. They will result in a negative skill score. Also, you are not allowed to seek help from a discussion board or any individual at all. Taking such help will be considered plagiarism and will violate the terms and conditions associated with project stage on edwisor.com .