Evaluation Rubric

Criteria	Meets expectations	Does not meet expectations
Data understanding, preparation and EDA (~30%)	All data quality checks are performed, and all data quality issues are addressed in the right way (missing value imputation, removing duplicate data and other kinds of data redundancies, etc.). Explanations for data quality issues are clearly mentioned in comments or in the presentation. Dummy variables are created properly wherever applicable. New metrics are derived if applicable and are used for analysis and modelling.	All quality checks are not done, data quality issues are not addressed correctly to an appropriate level. Dummy variables are not created properly. New metrics are not derived or are not used for analysis. The data is not converted to a clean format

	analysis in Python.	not cleaned using commands in Python.
Model building and evaluation (~40%)	Model parameters are tuned using correct principles and the approach is explained clearly. Both technical and business aspects are considered while building the model. Correct variable selection techniques are used. A reasonable number of different models are attempted and the best one is chosen based on key performance metrics. Model evaluation is done using the correct principles and appropriate evaluation metrics are chosen. The results are at par with the best possible model on the dataset.	Parameters are not tuned enough or tuned incorrectly. Relevant business aspects are not considered while model building. Variable selection techniques are used incorrectly / not conducted. A variety of models are not considered or a sub-optimal one is finalised. The evaluation process deviates from correct model selection principles,

	The model is interpreted and explained correctly. The commented code includes a brief explanation of the important variables and the model in simple terms.	evaluated or are incorrectly evaluated. The results are not at par with the best possible model on the dataset. The model is not interpreted and explained correctly.
Subjective Questions (~10%)	The answer to the subjective questions clear, concise and to the point. No assumptions are made and the reasons behind the answers are explained clearly.	The answers are unnecessarily long and unclear. The assumptions, if any, behind the answers, are not explained and the reasons behind the answers are not given clearly.
Presentation and Recommendations	The presentation has a clear structure, is not too long, and	The presentation lacks structure, is

	results concisely in simple language. The recommendations to solve the problems are realistic, actionable and coherent with the analysis. If any assumptions are made, they are stated clearly.	on the important observations. The language used is complicated for business people to understand. The recommendations to solve the problems are either unrealistic, non-actionable or incoherent with the analysis. Contains unnecessary details or lacks the important ones. Assumptions made, if any, are not stated clearly.
Summary Report (~5%)	The process followed and all the learnings are clearly	The process followed and

	The report is neither too detailed nor too brief. The 500-word word limit is followed.	and the report keeps deviating from it. The report is too brief or too detailed, i.e., it doesn't stick to the 500-word word limit.
Conciseness and readability of the code (~5%)	The code is concise and syntactically correct. Wherever appropriate, built-in functions and standard libraries are used instead of writing long code (if-else statements, for loops, etc.). Custom functions are used to perform repetitive tasks. The code is readable with appropriately named variables and detailed comments are written wherever necessary.	Long and complex code used instead of shorter built-in functions. Custom functions are not used to perform repetitive tasks resulting in the same piece of code being repeated multiple times. Code readability is poor because of vaguely named

		wherever necessary.

Report an error

PREVIOUS
Problem Statement

NEXT
Submission