|  | Meets 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 modeling.    The data is converted to a clean format suitable for analysis 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.    The model is interpreted and explained correctly. The commented code includes a brief explanation of the important variables and the model in simple terms |
| Subjective Questions (10%) | The answer to the subjective questions is clear, concise and to the point.    No assumptions are made and the reasons behind the answers are explained clearly. |
| Presentation and Recommendations (10%) | The presentation has a clear structure, is not too long, and explains the most important 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 |
| Summary Report (5%) | The process followed and all the learnings are clearly mentioned.    The report is neither too detailed nor too brief. The 500-word word limit is followed. |
| 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. |