Milestone Submission - Loan Default Prediction

Available From:06 Aug 25 2:25 AM

Due Date:05 Sep 25 2:29 AM

Total Marks:20

Submission Type:File Upload

Need help with your project? Get instant assistance from Glaide!

Description

**Submission Due Date**-

**Submission Guidelines**

1. There are two ways to work on this project:

**i. Full-code way:**The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

**ii. Low-code way:**The low-code way is to use an existing solution notebook template to build the solution and then submit a business report with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their individual learning aspirations and outcomes. The below table elaborates on these two options.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Submission type | Who should choose | What is the same across the two | What is different across the two | Final submission file [IMP] | Submission Format |
| Full-code | Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch | Perform exploratory data analysis to identify insights and recommendations for the problem | Focus on code writing: 10-20% grading on the quality of the final code submitted | Solution notebook from the full-code template submitted in .html format | .html |
| Low-code | Learners who aspire to be in managerial roles in the future-focussed on solution review, interpretation, recommendations, and communicating with business | Focus on business report: 10-20% grading on the quality of the final business report submitted | Business report in .pdf format with problem definition, insights, and recommendations | .pdf |  |

Please follow the below steps to complete the assessment.**Kindly note that if you submit a report along with the notebook, ONLY the report will be evaluated.**Please make sure that all the sections mentioned in the rubric have been covered in your submission.

**i. Full-code version**

* Download the full-code version of the learner notebook.
* Follow the instructions provided in the notebook to complete the project.
* Clearly write down insights and recommendations for the business problems in the comments.
* Submit only thesolution notebook prepared from the learner notebook [format: .html]

**ii.** **Low-code version**

* Download the low-code version of the learner notebook.
* Follow the instructions provided in the notebook to complete the project.
* Prepare a business report with insights and recommendations to the business problem.
* Submit only thepresentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.

3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.

4. Submission will not be evaluated if

* it is submitted post-deadline, or,
* more than 1 file is submitted.

**Best Practices for Full-code submissions**

* The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
* The notebook should be run from start to finish in a sequential manner before submission.
* It is important to remove all warnings and errors before submission.
* The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
* Please refer to the FAQ page for common project-related queries.

**Best Practices for Low-code submissions**

* The report should be made keeping in mind that the audience will be the Data Science lead of a company.
* The key points in the presentation should be the following:
  + Business Overview of the problem and solution approach
  + Key findings and insights which can drive business decisions
  + Business recommendations
  + Focus on explaining the key takeaways in an easy-to-understand manner.
  + The inclusion of the potential benefits of implementing the solution will give you the edge.
* Copying and pasting from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your presentation.
* The report should be submitted as a PDF file (.pdf) and NOT as a .pptx file.
* Please refer to the FAQ page for common project-related queries.

Happy Learning!

Rubric

CriteriaProblem DefinitionThe context: - Why is this problem important to solve? - Provide an additional point of view outside of what has been shared with the problem statement which indicates what is the business problem The objectives: - What is the intended goal? - Provide an additional point of view outside of what has been shared with the problem statement like identifying those customers who will default on loans before approval. The key questions: - What are the key questions that need to be answered? The problem formulation: - What is that we are trying to solve using data science?Points4

CriteriaData ExplorationData Description: - What is the background of this data? What does it contain? - Write the observation from the Data Description such as shape of data, data types of various attributes, missing values,statistical summary. - Write some initial observations on range of attributes, outliers of various attributes. Observations & Insights: - What are some key patterns in the data? - What does it mean for the problem formulation? - Are there any data treatments or pre-processing required? - Does any outlier treatment required? - Any missing value treatment requiredPoints4

CriteriaBuilding ModelsBuilding various models and tuning the model to get better validation - What different techniques should be explored? - How these techniques can be improved? - Why the techniques suggested are the best to explore for the data and problem at hand?Points4

CriteriaComparison of various techniques and their relative performance based on chosen Metric (Measure of success)Measures of success: - What are the key measures of success? - What are the important metrics to consider and why? - How do different techniques perform? - Which one is performing relatively better and why? Is there scope to improve the performance further? - Which variables are significant in predicting the target variable - Are variables still continue to be significant post modelling - What are the most meaningful insights from the data relevant to the problem?Points4

CriteriaProposal for the final solution design- Is the model performance good enough for deployment in production and is it interpretable? - What model do you propose to be adopted? - Why is this the best solution to adopt? - How that solves the problem?Points4