

Milestone 1 Checklist

This document is supplementary to the project description. Please refer to the project description for the actual deliverables and requirements.

Key dates & logistics:

- **When:** MS1 evaluation will be scheduled **the week after the midterms**.
- **How to reserve a slot:** Evaluation slots are **first-come, first-served**. Details on how and when to reserve a slot will be posted in a separate announcement

What to expect at the evaluation

1. **Every team member must be prepared.**
Any question about any part of the milestone, data cleaning, feature engineering, modeling choices, evaluation metrics, or the inference function, may be directed to any team member. Each member should be able to **explain and justify** every step.
2. **Explain & justify every step in the context of your project.**
For each step you performed (and for the three milestones overall) you must say why that step makes sense *for your dataset and task*, and support your explanation with **evidence** (plots, numbers, statistics, or references to experiments).
3. **Two legitimate ways to motivate a step**
 - **Hypothesis-driven (EDA supported):** e.g., “EDA showed feature X was skewed, e.g., “we applied a log transform because ...” (show the EDA results that led you there.)
 - **Trial-and-error (empirical):** report a clear **baseline** (before the change) and **after** results for the *complete pipeline*; explain why the added step improved or worsened performance.
4. **Compare multiple approaches.**
You must investigate **more than one** approach when trying to improve performance (different features, preprocessing pipelines, models, hyperparameters, etc.), and report comparative results.
5. **Baseline requirement for architecture choices.**
The notebook **must include** a statistical ML model **or** a shallow feed-forward neural network as a baseline. If you choose a different architecture, you **must** implement the baseline models and compare them to justify your chosen architecture.
6. **Order matters — test it.**
Investigate whether the **sequence** of preprocessing/processing steps affects final performance. Report any differences.
7. **Don't overthink!**
No need for an LLM or any generative model for any of the three themes.

Notebook & report requirements

- **Run-all ready:** The notebook must execute from top to bottom with a single **Run all** during evaluation (no hidden cells requiring manual modification).
- **Correct order & documentation:** Steps must follow logical order and include **clear markdown explanations** that justify choices.
- **All plots and Data Engineering answers included:** Every plot, result from data-engineering questions, and the reasoning/justification must appear in the notebook and in the structured report.
- **Inference function:** Provide a function that accepts *raw input*, applies **the exact same preprocessing** used during training, calls the trained model, and returns the output in **natural language** (not just numeric labels). Make sure this function is demonstrated with one or two example inputs in the notebook.

Checklist:

- Data Cleaning
- Data Analysis
 - Plotting the results
- Answering the Data Engineering Questions
 - Plotting the results
 - Answering the questions in a markdown in the notebook, and adding it to the report
- Feature Selection
 - The reason for the chosen features, reporting evidence.
 - Plotting the effect of the features
- Data Pre-processing and Feature Engineering
 - Understanding the step
 - Need for the step
 - Effect on the step on the data distribution, range, values, etc..
- Prediction Model
 - The model choice
 - How it works
 - Its limitations
 - Reporting its performance throughout training and validation
 - Reporting its performance on the unseen test data
 - Plotting the reports
 - Inference Function
- XAI
 - Global Explanation
 - Local Explanation
 - Plotting and explaining the explanations in the report.