

Objective: Your task is to evaluate your ability to perform exploratory data analysis (EDA) on a mixed-feature dataset and build a simple binary classification neural network using PyTorch. The whole exercise shouldn't take a significant amount of time.

Instructions:

1. Exploratory Data Analysis (EDA):

- **Task:** Conduct EDA on a provided dataset with a mix of numerical and categorical features, missing values, and potential outliers.
- **Dataset:** You will be provided with a dataset with descriptions related to a real-world scenario (loan default prediction) with binary target labels.
- **Steps to Follow:**
 1. Load the training dataset and perform an initial data inspection.
 2. Examine data quality appropriately.
 3. Explore the dataset per your preference
 4. Summarize your findings, visualizations and insights in a Jupyter notebook or a markdown file
- **Git Requirement:** Commit each major step of the EDA process separately, including data cleaning, visualization, and insights, to demonstrate a structured approach.

2. Neural Network Implementation:

- **Task:** Build a neural network of your preference using PyTorch to classify the dataset.
- **Specifications:**
 1. Design a network with at least one hidden layer and a configurable number of neurons.
 2. Optimally choose the activation function and implement a training loop with a loss function like binary cross-entropy.
 3. The neural network model must follow common practices (object-oriented design, forward method, etc.).
 4. You can choose either define classes and functions in the same Jupyter notebook, or in a standalone file or package that can be imported
 5. Follow the best model development practice to finalize the model.
 6. Inference on the test dataset (bad_flag is intentionally left blank)
 7. You may add comments or markdowns to explain why you do certain things (such as why a certain architecture is chosen)
- **Git Requirement:** Commit each major step (e.g., model architecture, training loop, and results) separately to demonstrate incremental progress.

3. Final Deliverables:

- Create a GitHub repository containing:
 1. The EDA script or notebook with visualizations and insights.
 2. The PyTorch model implementation and training script.
 3. A README file explaining the project structure, how to run the code (including environment dependencies), and any assumptions made.
- Share your **GitHub link** and **test result file** with the recruiter within one week of receiving this communication. Please make sure no dataset is in the repo.
- You are encouraged to utilize assisting tools such as IDE, search, LLM, etc., to make your project more refined if you understand and agree with such changes.

Please ensure that each step is clearly documented and that your code is well-organized and easy to follow. Good luck!