AI/ML Researcher Challenge

Challenge Title: Text Classification for Active vs. Passive Voice Detection

Challenge Description:

Objective: Develop a text classification model that can effectively detect whether a given sentence is in the active or passive voice, using a dataset of labelled sentences. This challenge aims to assess your skills in natural language processing, machine learning, and model explainability.

Task:

- Your task is to train a text classification model on this dataset.
- Split the dataset into training (60%), validation (20%), and test (20%) sets.
- Develop a machine learning model capable of accurately classifying sentences as either active or passive.
- Equal importance will be placed on the following aspects:
- Clean Code: Your code should be well-organized, well-documented, and adhere to best practices.
- Explainability: Provide clear explanations of your model's architecture, decisions, and any preprocessing steps. Make it understandable to non-technical stakeholders.
- Metrics: Evaluate your model's performance using appropriate metrics and provide a detailed analysis of its strengths and weaknesses.

- **Completion Criteria**:

1. A well-trained text classification model that can accurately classify sentences as active or passive voice.

2. A clean and well-documented codebase.
3. Comprehensive explanations of your model's architecture and preprocessing steps.
4. Evaluation metrics demonstrating the model's performance on the test set.
5. An analysis of the model's strengths and areas for improvement.
6. A 60:20:20 split of the dataset into training, validation, and test sets.
7. Equal weightage given to clean code, explainability, and metrics in the assessment