UE22CS352A – Machine Learning

Hackathon Instructions for Students

Hackathon Timing: The hackathon will take place from 8:30 AM to 4:30 PM. Please ensure you are fully prepared and available for the entire duration.

1. Competition Guidelines:

- Carefully read the instructions provided on the Kaggle competition page. These details are essential for understanding the objectives, rules, and submission requirements.
- All students are required to join the Kaggle competition first. After joining, the team leader must invite the remaining three team members to form a complete team. It is crucial to follow the specified team formation guidelines and naming conventions.

2. Team Collaboration:

- While one team member is actively working on the Kaggle notebook, other members are encouraged to explore alternative models or perform tasks such as data preprocessing, feature engineering, or hyperparameter tuning on their local machines or cloud-based notebooks. This division of tasks can help improve overall performance and efficiency.
- A Google Drive link will be provided, containing the same dataset as the one on Kaggle. Use this resource for data exploration as needed.

3. GPU Configuration:

- Please ensure that the notebook is configured to utilize a **T4 or P100 GPU for optimal performance**. Adjust the runtime settings to enable GPU acceleration by selecting either the **T4 or P100 GPU** under the hardware accelerator options. This setup is crucial as it will significantly enhance the processing speed and efficiency of your computations within the notebook.

3. Starter Notebook:

- A starter notebook would be provided to help you establish a basic outline for your project. This notebook would be available in the Google Drive link that will be shared with the dataset. Use it as a foundation to structure your workflow and guide your approach to data exploration, preprocessing, and model building.
- The required format for final submissions, including specific file structures and naming conventions, is outlined in the **"Submission Format"** section on the Overview tab of the Kaggle competition page. Please refer to this section to ensure that your submission meets all necessary guidelines and avoids penalties.

5. Evaluation Criteria:

- Your evaluation will consider both the overall team performance and individual contributions to the project.
- Each notebook will undergo a plagiarism check. Ensure originality in your work and properly document all code and methodologies used.
- **Pre-trained LLMs (Large Language Models)** are **not permitted** for use in this competition. However, embeddings such as **BERT, GloVe**, and similar pre-trained embeddings are allowed.

- **APIs** are also **not allowed**. All work must be conducted within the competition environment without the use of external API resources.

6. Submissions:

- Teams are allowed a maximum of five submissions throughout the competition. It is recommended to thoroughly review and validate your work before submitting to avoid unnecessary deductions.
- Submission files must be generated directly from a Kaggle notebook. Only submitting CSV files without the supporting notebook will result in no marks being awarded.

7. Google Forms Submission:

- In addition to the Kaggle submission, teams must complete a Google Form submission. The following items are required:
    1. The final notebook file used for the submission.
    2. A PDF document containing deliverables, formatted with the team's naming convention. This document should include:
        - **Data Preprocessing Steps**: Outline the steps taken to clean and prepare the dataset for model training.
        - **Feature Extraction Steps**: Detail the features selected or engineered to enhance model performance.
        - **Comparison Between Early and Late Fusion Results**: Provide a comparative analysis of any fusion techniques used, with insights into their effectiveness.
        - **Model Decisions**: Explain the reasoning behind choosing specific models or methodologies.
        - **Performance Analysis**: Discuss model performance, including key metrics and any insights gained from experimentation.

8. Final Assessment:

- A viva will be conducted, where each team will discuss their approach, choices, and results.

Please ensure that all submissions adhere to these guidelines to avoid penalties. Best of luck, and make the most of this learning experience!