Assignment 3

Transformer-Based Sentiment Analysis (Total: 50 Points)

Objectives

Implement a transformer-based sentiment analysis system using the Hugging Face ecosystem. This task will help you understand fine-tuning, evaluation, and interpretability of transformer models for text classification.

1. Data Collection and Preprocessing (10 Points)

- Choose a well-known sentiment analysis dataset from HuggingFace and perform text cleaning, tokenization, and data splitting into training, validation, and test sets.

2. Model Selection and Fine-Tuning

- Select a pre-trained transformer model (e.g., BERT, RoBERTa) from Hugging Face's model hub and fine-tune it on the chosen dataset using the Trainer API or a custom training loop.
- Experiment with hyperparameters like learning rate, batch size, and number of epochs.

3. Evaluation and Benchmarking (10 Points)

- Evaluate the model using metrics such as accuracy, precision, recall, and F1 score.
- Compare the performance with relevant baselines or published results.

4. Visualization and Interpretability (10 Points)

- Generate visualizations (e.g., attention heatmaps, confusion matrices) to illustrate which tokens contribute most to the model's decisions.
- Analyze the visualizations to explain model behavior and potential limitations.

5. Analysis and Reporting (5 Points)

- Provide concise explanations detailing your methodology, experimental observations, challenges encountered, and insights gained from the visualizations and evaluation.

Named Entity Recognition (NER) with Transformers (Total: 50 Points)

Objectives

Implement a transformer-based NER system using the Hugging Face ecosystem. This task will help you understand fine-tuning transformer models for sequence labeling, manage token-label alignment, and analyze model behavior through visualization and error analysis.

1. Data Preparation (10 Points)

- Select a NER dataset and preprocess it by aligning tokens with entity labels, carefully handling cases of subword tokenization.

2. Model Setup and Fine-Tuning (15 Points)

- Choose an appropriate pre-trained transformer model (e.g., BERT, RoBERTa) from Hugging Face and fine-tune it on the NER dataset.
- Implement strategies to address token-label alignment challenges during training.

3. Evaluation and Metrics (10 Points)

- Evaluate the model using entity-level precision, recall, and F1 scores.
- Perform error analysis to identify common misclassifications and areas for improvement.

4. Visualization and Interpretation (10 Points)

- Create visualizations such as attention maps or confusion matrices to demonstrate how the model distinguishes between entities.
- Use these visualizations to provide insights into model strengths and weaknesses.

5. Analysis and Reporting (5 Points)

- Provide detailed explanations outlining your approach, experimental results, visualizations, and critical reflections on the model's performance and interpretability.

Submission Guidelines

- Submit all code as a Google Colab notebook, ensuring the link is shared in the submission description with **Anyone with the link** access. Additionally, upload the notebook file and include detailed explanations and visualizations, either within the notebook as markdown cells or in a separate PDF report. Avoid sharing .py or ZIP files for this assignment.
- Clearly label each part and question in your submissions.
- Do not attach any datasets; instead, provide the link to them on Hugging Face in your notebook.
- To familiarize yourself with the Hugging Face ecosystem, you may refer to their NLP course.
- **Deadline:** March 19th, 2025

Rubric and Expectations

- **Code Quality and Functionality:** Code should be well-organized, commented, and fully functional. Demonstrate effective use of the Hugging Face ecosystem and best practices in Python.
- **Performance Metrics and Benchmarking:** Accurately report evaluation metrics (accuracy, precision, recall, F1, etc.) and compare results with relevant baselines or published work.
- **Visualization and Interpretability :** Provide clear and insightful visualizations (e.g., attention heatmaps, confusion matrices) that explain the model's inner workings. Interpret these visualizations to demonstrate an in-depth understanding of model behavior.
- **Analysis and Reflection:** Offer a comprehensive analysis discussing methodology, challenges, and insights gained from the experiments. Reflect on the interpretability of transformer models and propose potential improvements.
- **Adherence to Guidelines:** Submissions must follow the provided guidelines, including formatting, labeling, and meeting the deadline.