Comparative Analysis of Transformer Models for NLP Tasks

**Business Objective**

Conduct a comparative analysis of different transformer models (such as BERT, RoBERTa, DistilBERT, GPT-2, etc.) on a specific NLP task (e.g., sentiment analysis, text classification, question answering). The goal is to evaluate their performance, efficiency, and applicability to provide insights into the strengths and weaknesses of each model for selecting the most appropriate solution.

**Dataset Details (Choose any 1) or Any of your choice: Download it from web**

Choose a dataset relevant to your selected NLP task. Here are some options:  
1. Sentiment Analysis:  
 - IMDB Movie Reviews Dataset: Large dataset of movie reviews labeled as positive or negative.   
 - Twitter Sentiment Analysis Dataset: Tweets labeled with sentiments.   
2. Text Classification:  
 - AG News Corpus: Over 1 million news articles categorized into classes.  
 - DBpedia Ontology Dataset: Wikipedia articles classified into 14 ontology classes.   
3. Question Answering:  
 - Stanford Question Answering Dataset (SQuAD): Questions posed on Wikipedia articles.

**Acceptance Criterion:**

The final output should be a report that compares the performance, efficiency, and resource usage of selected

transformer models. Key performance metrics will be accuracy, F1-score, training/inference time, and model size.

**Milestones:30 days to complete the Project**

|  |  |  |
| --- | --- | --- |
| **Milestone** | **Duration** | **Task start - End Date** |
| Kick off and Business Objective discussion | 1 day | Discuss project objectives, finalize the NLP task, and select transformer models to be compared |
| EDA & Dataset Preparation | 1 Week | Download and clean the dataset. Preprocess data (tokenization, split into train/validation/test). Perform initial data analysis. |
| Model Implementation & Fine-Tuning | 1 Week | Implement the selected transformer models (BERT, RoBERTa, DistilBERT, GPT-2, etc.) and fine-tune them on the dataset. |
| Model Evaluation & Comparative Analysis | 1 Week | Evaluate models on the test dataset. Compare performance metrics and computational efficiency. |
| Feedback | 1 week | Incorporate feedback, fine-tune models if necessary, and refine documentation. |
| Final Requirements |
| Final presentation | 1 Day | Present the final results, complete documentation, and submit final report and code. |

Protocols:

1. All participants should adhere to agreed timelines and timelines will not be extended.
2. All the documentation – Final presentation and python code to be submitted before the final presentation day.
3. All the participants must attend review meetings.