AI Engineering for Fine-Tuning and Evaluating the Mistral 7B Language Model - Analytical Data Systems

Over the course of 5 weeks, you'll spend time working on the latest open source generative AI model to fine-tune and evaluate the performance of the Mistral 7B. By leveraging the latest techniques in fine-tuning and developing a structured evaluation methodology, you'll gain hands-on experience in optimizing large language models and assessing their performance using relevant metrics.

Each week, you'll tackle new challenges into various aspects of AI engineering, from researching fine-tuning techniques to implementing evaluation frameworks and visualizing metrics. You'll work with a fixed dataset, split into training, validation, and testing subsets, to ensure a consistent and reliable evaluation process.

By the end of the project, you'll have a comprehensive understanding of the fine-tuning process, evaluation methodologies, and metric reporting. You'll develop skills in collaborating with peers, documenting your work, and presenting your findings in a clear and concise manner.

If you're ready to explore the cutting-edge field of large language models and gain practical experience in AI engineering, then this project is perfect for you. Get ready to challenge yourself, learn from experienced mentors, and showcase your abilities in fine-tuning and evaluating the Mistral 7B model.

Technologies: Python, Mistral 7B Language Model, JSON, Data Visualization Libraries.

Project Overall Goal: To design and build a repeatable and quantitative methodology and python library to compare LLM fine tuning performance based on a fixed data set based on the latest techniques in the generative AI community.

IP Rights: Analytical Data Systems will own all rights to deliverables, including code, data, prompts, and documentation. Feel free to use the know-how to advance your career.

Project Scope:

Week 1: Research and Planning
- Research the latest techniques for fine-tuning LLMs, focusing on Mistral 7B.
- Define a repeatable methodology for fine-tuning, including data preparation, hyperparameter selection, and training process.
- Document the methodology in detail, including any necessary prerequisites and dependencies. This will be updated throughout the project.
- Select a fixed dataset for evaluation and split it into training, validation, and testing subsets.

Week 2: Setup and Baseline Evaluation
● Set up the development environment, including necessary libraries and frameworks in Google Colab.
● Implement a framework to evaluate the performance of the baseline Mistral 7B model on the fixed dataset.
● Run the baseline evaluation and store the results in JSON format in a Google docs folder.
● Analyze the baseline results and identify areas for improvement.

Week 3: Fine-tuning and Evaluation
● Implement the fine-tuning methodology defined in Week 1 using the training subset of the fixed dataset.
● Fine-tune the Mistral 7B model using the selected techniques and hyperparameters.
● Evaluate the fine-tuned model's performance on the validation subset and store the results in JSON format.
● Compare the fine-tuned model's performance with the baseline results and make necessary adjustments.

Week 4: Testing and Metric Reporting
● Evaluate the fine-tuned model's performance on the testing subset and store the results in JSON format.
● Implement a script to read the JSON files containing the evaluation results (baseline and fine-tuned).
● Calculate and report relevant metrics, such as perplexity, accuracy, F1 score, and others depending on the task.
● Visualize the metrics using graphs or charts to facilitate comparison between the baseline and fine-tuned models.

Week 5: Documentation and Presentation
● Prepare a comprehensive report documenting the entire project, including introduction and objectives, methodology for fine-tuning and evaluation, results and analysis, conclusions, and future work.
● Create a presentation summarizing the project, key findings, and lessons learned.
● Present the project to the class or a group of peers and instructors.
● Incorporate feedback and finalize the documentation.

Throughout the project, you will gain experience in researching and implementing fine-tuning techniques, developing evaluation frameworks, and reporting performance metrics. You will also learn to collaborate with your peers, document your work, and present your findings effectively.

By the end of this 5-week project, you will have developed a strong foundation in AI engineering for large language models, equipping you with valuable skills and knowledge that can be applied to future projects and career opportunities in the field of AI and natural language processing.