Data Structures and Algorithms in Python

Project 4: LLM Applications

Machine learning is a subset of artificial intelligence focused on building systems that learn from data. Unlike traditional software programs that follow explicit instructions to perform tasks, machine learning algorithms use statistical methods to enable computers to improve at tasks with experience. This field encompasses a variety of approaches, including supervised learning, unsupervised learning, and reinforcement learning, each with its unique applications and challenges.

Within the realm of machine learning, Large Language Models (LLMs) represent a significant advancement. These models, which include OpenAI's GPT series and Google's Gemini, are designed to understand, interpret, and generate human language in a way that is both sophisticated and nuanced. By analyzing vast datasets of text, LLMs can compose coherent and contextually relevant responses to prompts, making them powerful tools for a range of applications from automated customer service to content creation and beyond.

This project focuses on exploring the capabilities and implications of LLMs. You will have the opportunity to explore the technological underpinnings of these models, their practical applications, and the ethical considerations they raise. Through research or hands-on programming, you'll gain insights into one of the most exciting frontiers of computer science today.

You will complete this project individually, choosing between two options:

- Option 1: Research Paper
- Option 2: Python Program

Each option is designed to deepen your understanding of LLMs and their place within the broader context of machine learning and artificial intelligence. Detailed requirements and guidance for each option are on the following pages.

Option 1: Research Paper - Exploring Large Language Models

Objective: Research the development, functioning, and societal impact of Large Language Models (LLMs), with a special focus on models such as OpenAI's GPT and Google's Gemini.

This assignment aims to provide you with an understanding of Large Language Models (LLMs), focusing on their development, function, and impact. By researching various aspects of LLMs you will gain insights into the capabilities of these advanced AI systems. You are required to conduct a comprehensive literature review on Large Language Models (LLMs). This review should cover the technological foundations of LLMs, key developments in the field, an overview of notable models such as OpenAI's GPT and Google's Gemini, and the implications of these technologies on society.

Specific Requirements

- **Introduction to LLMs**: Explain what LLMs are, including a brief history of their development. Highlight key milestones in their evolution.
- **How LLMs Work**: Describe the basic principles behind LLMs, including data training processes, neural network architecture, and how these models generate responses.
- **Case Study - Google's Gemini**: Provide a detailed overview of Google's Gemini. Discuss its unique features, capabilities, and how it compares to other LLMs.
- **Impact and Ethics**: Explore the impact of LLMs on various sectors (e.g., education, healthcare, customer service) and discuss ethical considerations, including bias, privacy, and the potential for misuse.
- **Conclusion and Personal Insight**: Conclude with your personal insights on the future of LLMs. Where do you see this technology going, and what potential do you think it holds?

Format

- The review should be between 1500 to 2000 words.
- Use MLA format for citations and references.
- Include at least five reputable sources (e.g., academic journals, official AI research blogs, and tech industry publications).

Rubric Categories

- Content Accuracy and Depth (70%): Demonstrates comprehensive understanding of LLMs, detailed explanation of Google's Gemini, and insightful analysis of impacts and ethical issues.
- Research and Sources (20%): Quality and variety of sources, integration into the text, and proper citation.
- Writing Quality (10%): Logical flow of ideas, clear structure, and seamless transitions between sections. Grammar, punctuation, spelling, and adherence to the specified format.

Submission

- Submit your assignment as a PDF document on CREEKnet.

Option 2: Python Program

Objective: Design and develop an original program leveraging the Google Gemini API to showcase the practical application of LLMs in a creative and impactful manner.

For this option you will create an original computer program that uses the Google Gemini API. You will first propose a project, and once your proposal is approved, you will develop your program.

Proposal

- Submit a proposal outlining your project's goals and steps that need to be completed.
- Proposals will be reviewed for feasibility and relevance.

Ideas for Inspiration

- Text-Based Adventure Game with AI-Powered Characters: Create a game where players interact with characters driven by LLMs. These characters can remember past interactions and respond dynamically, enhancing the storytelling and engagement through personalized narratives.
- Data Processing and Visualization Tool: Develop a tool that uses an LLM to process and classify large datasets (e.g., social media posts, product reviews). Utilize Pandas to display the analysis, offering insights into trends, sentiments, or classifications.
- Interactive Learning Companion: Build an application that serves as a study assistant, using LLMs to generate quizzes, explain complex topics in simpler terms, or provide summaries.
- Story Generation from Images: Design a program that crafts stories or descriptions based on user-uploaded images. By analyzing the images using gemini-pro-vision, the application generates creative narratives or settings.
- Puzzle Game with AI-Generated Challenges: Create a puzzle game where players meet challenges based on criteria generated by an LLM. For example, players might need to submit images that match a certain theme or concept, with the AI giving a score for their submissions.

Rubric Categories

- Project Execution (70%): code quality, functionality, and how well the project meets the proposed objectives.
- Ambition and Creativity (20%): originality and inventive use of LLMs in the project
- Proposal (10%): clarity of objectives, feasibility, and thoroughness. A strong proposal clearly outlines the project goals, methodology, and expected outcomes.

Submission

- Submit your Python program and any necessary data files on CREEKnet.