

Member names:

1. Pranav Patil (219311292)
2. Siddharth Singhal (229302170)
3. Arnav Kawatra (229301286)
4. Sanidhya Singh (229311232)

Date: 22nd February, 2024

I

Abstract:

Our Large Language Model (LLM) utilizes Recurrent Neural Networks and sequential models like multilayer perceptron's to efficiently cater to students' information needs, consolidating data from multiple sites. Trained via unsupervised learning, it provides coherent responses, showcasing its versatility in various domains. However, ethical considerations and data privacy remain paramount for its responsible deployment and societal impact.

Introduction:

Our Large Language Model (LLM) caters in need of students to find information regarding specific topics without having the need to use a lot of sites. Our model uses Recurrent Neural Network is a type of artificial neural network designed to recognize patterns in sequences of data. Sequential models like multilayer perceptron's which is used in our project are artificial neural networks structured in a sequential layer-by-layer fashion, facilitating the hierarchical learning of complex patterns in data. Layered normalization of input data.

This review aims to explore the architecture, training methodologies, applications, and ethical considerations associated with our LLM. By providing an in-depth analysis of this Project, we seek to elucidate the capabilities, limitations, and societal implications of these powerful AI systems. Additionally, we aim to identify gaps in current research and propose avenues for future investigations.

Body:

Architecture of Large Language Models:

Architecture for this Large language models is based on transformer architectures, which enable them to process and generate text with remarkable fluency and coherence. Large language models also require Attention Mechanism, which are used to detect the relevancy of the position of tokens.

Training Methodology:

Our Large Language Model (LLM) is pre-trained on web links provided for the event employing unsupervised learning. We've standardized and tokenized the data utilizing certain functions to pre-process the input destined for the Large Language Model (LLM). Employing all bestowed data, our Large Language Model (LLM) computes cognitive insights through inquiries.

Applications of our Large Language Model:

Our Large Language Model (LLM) can provide quick, accurate responses to user queries about sites that the LLM Model is trained in. Applications include data analysis and real-time information retrieval for enhanced decision-making and engagement.

Conclusions

In conclusion, large language models (LLMs) trained by scraping data from websites and responding to queries related to that data hold immense potential across diverse applications. They streamline information retrieval processes, enhance user engagement, and facilitate decision-making in various domains. However, careful consideration of ethical implications, data privacy, and the accuracy of responses is crucial to ensure their responsible deployment and beneficial impact on users.

Literature Cited:

1. <https://github.com/karpathy/nanoGPT>