

My Project Review:

Title: A Custom LLM Trained to answer queries related to Modern Web Development

Name : Prithvi Ragavendiran R (CIT Chennai)

Date : 21.02.2024

Abstract:

The Language Learning Model (LLM) developed for the 12-hour Hackathon is the result of diligent efforts aimed at enabling question answering across a diverse range of specialized domains. With a keen eye for detail, a collection of valuable resources spanning JavaScript, React, Node.js, NPM, Mongoose, Express, Next.js, Axios, Tailwind CSS, and Vite was meticulously compiled into a single PDF document. This compilation process involved manual sorting and organizing to ensure that the dataset, consisting of approximately 410 instances, comprehensively covered the key concepts and nuances of each domain.

The LLM's architecture is grounded in a sophisticated LSTM neural network design, carefully tailored to process and comprehend the textual data extracted from the curated PDF corpus. Through the incorporation of bidirectional layers, dropout, and batch normalization techniques, the model undergoes rigorous training over 300 epochs. Additionally, a personalized learning rate scheduler further enhances the training process, fostering efficient convergence towards optimal performance.

Ongoing efforts are dedicated to fine-tuning the LLM, focusing on refining its parameters and optimizing its architecture. These refinements include adjustments to hyperparameters and the integration of transfer learning techniques to enhance the model's capacity to understand intricate concepts and generate contextually relevant responses.

Upon completion, the LLM is poised to serve as a valuable resource for knowledge dissemination and problem-solving within the specified domains. Collaboration with domain experts and rigorous evaluation against real-world scenarios will ensure the model's alignment with the objectives of the Hackathon project, demonstrating its readiness for practical deployment.

Process Flow:

Model Defining:

First up, let's dive into LSTM networks. These are sophisticated neural network architectures designed to tackle the challenge of understanding language patterns over time. We've fine-tuned them, incorporating features like bidirectional layers, dropout regularization, and batch normalization to enhance their ability to capture and retain complex linguistic dependencies.

Dealing with Raw Data:

I Compressed the data related to the domains In a pdf format so it would be much easier for me to handle the data for custom training the model.

Data Extraction through PDF:

Moving on, let's talk about PDFs. These digital documents hold a wealth of information, but extracting relevant text from them can be tricky. We've employed advanced extraction techniques to sift through the content, ensuring we capture key insights across domains like JavaScript, React, and beyond. It's akin to curating a digital library, where each page holds valuable knowledge waiting to be uncovered.

Model Training:

Now, let's blend our extracted text with our LSTM framework. This integration process involves mapping the textual data onto the neural network architecture, setting the stage for comprehensive model training. Through iterative epochs and carefully calibrated learning rate schedules, we've optimized our model's parameters to achieve maximum performance in language comprehension and generation tasks.

Improvements:

Pretrained Models Can be Leveraged and the model needs a lot of training time in my case since it was a custom model and find a better way to curate the datasets.

Conclusion:

In wrapping up, I must say, working with LLMs for the first time has been quite the experience. It's been a journey of patience and perseverance, with lots of trial and error along the way. I must admit, after a few hundred epochs, my model is starting to look a bit... well, let's just say it's not quite there yet. Apparently, it needs to run for thousands of epochs for better performance, but who has the time or resources for that? Certainly not my poor computer, which might just keel over after a few days of non-stop training! Ah, the joys of machine learning. Nonetheless, despite the setbacks, it's been an eye-opening and rewarding experience. I've learned a lot, and I'm excited to see where this journey takes me next