Introduction of LLM : Assignment

Score: 26 / 38

✓

What is the basic concept behind Large Language Models (LLM)?

The basic concept behind LLM is to train them on a very large dataset using the transformer architecture so that can learn the language. By doing so they become a good reasoning engine. Then the LLM can be fine-tuned for a specific task using a small dataset

Feedback

Great answer! The explanation is clear and demonstrates a good understanding of the basic concept behind Large Language Models (LLM). Well done!

✓

Name one of the widely known LLMs in use today and briefly explain its architecture.

The widely known LLM todays are

1. Open AI ChatGput - > It is based on the decoder architecture of the transformer. It is an autoregressive model fine-tuned for chat conversation using RLHF
2. BERT -> It is based on the encoder architecture of the transformer. It is trained using MLM to predict the next

token. It can be used for sentiment analysis.

Feedback

Great answer! Both LLMs are well explained with their architectures.

How does a Large Language Model like GPT-3 differ from traditional language models?

GPT-3 is an example of a very large language model, containing 175 billion parameters. Traditional language models typically have significantly fewer parameter. LLM require huge compute cost and resources for training. They are based on transformer architecture. They perform considerably well on zero-shot learning.

Feedback

Great Answer! Well done!

Explain the concept of fine-tuning in the context of Large Language Models.

The base model or foundation model is trained on the large dataset to learn the language. This model can be adopted for a specific task using fine-tuning. Fine-tuning involves methods like instruction fine-tuning, peft methods like lora and qlora, soft prompting.

Feedback

The response provides a basic understanding of fine-tuning in the context of Large Language Models. However, it lacks specific examples or details about the methods mentioned. More elaboration and examples would enhance the explanation.

What are some common challenges or limitations associated with Large Language Models?

Too large. Occupies huge memory

They tend to hallucinate if they do not know
the correct answer

Feedback

The answer provided some insight into the challenges of large language models but lacked specific details or examples. Consider including specific examples and elaborating on the challenges for a more comprehensive response.

Describe the concept of zero-shot learning in the context of Large Language Models.

Using a LLM for a task for which it is not
explicitly trained is known as zeor-shot
learning.

Feedback

The concept of zero-shot learning involves using a model for tasks it hasn't been explicitly trained for, but the answer lacks depth and explanation.

What are some of the key applications of Large Language Models in the field of natural language processing?

Text generation, Question and Answering, Sentiment Analysis, Chatbot

Feedback

Great Answer! keep it up

Compare the training process of Large Language Models with traditional language models.

For training LLM we need a huge dataset which needs to be cleaned. Then we need lot of computing power

Feedback

Good response, but could be improved by mentioning the specific steps involved in training large language models and how they differ from traditional language models.

Explain the role of attention mechanisms in Large Language Models like BERT or GPT.

The role of attentino is to learn contextual dependencies among different tokens of the input. How much importance to be given to different tokens

Feedback

The answer explains the role of attention in learning contextual dependencies, but it could be improved by providing specific examples of how attention mechanisms work in large language models like BERT or GPT.

How do Large Language Models handle the problem of context understanding in natural language processing tasks?

LLM can use RAG system for retrieval of relevant information from sources. This relevant information can be passed as contect

Feedback

The response correctly mentions the use of the RAG system for retrieval of relevant information, but it lacks depth in explaining how LLMs handle context understanding in natural language processing tasks.