RQ2: Evolution of LLMs

- Due Jan 24 at 4:50pm
- Points 4
- Questions 6
- Available Jan 24 at 4:25pm - Jan 24 at 4:50pm 25 minutes
- Time Limit 17 Minutes
- Allowed Attempts Unlimited

Instructions

This quiz has 5 questions and 1 bonus question.

Do not use AI to answer these questions.

Keep the questions concise and less than 5 sentences. Longer answers will be disregarded.

You only have 15 minutes so use your time smartly!

This quiz was locked Jan 24 at 4:50pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	15 minutes	4.4 out of 4

ⓘ Correct answers are hidden.

Score for this attempt: 4.4 out of 4
Submitted Jan 24 at 4:44pm
This attempt took 15 minutes.

PartialQuestion 1
0.8 / 0.8 pts

1. **Evolution of LLMs**
   As discussed in Lecture 2, Pedro Domingos described five tribes of machine learning, each with a unique approach to solving problems. Complete the table below by identifying each tribe and its favored algorithm.

Tribe	Approach	Favored Algorithm
Symbolists	Use symbols, rules, and logic to represent knowledge and draw logical inferences.	Rules and decision trees

Bayesians	Assess the likelihood of occurrence for probabilistic inference.	Naive Bayes or Mar
Connectionists	Recognize and generalize patterns dynamically with weighted neurons.	Neural Networks
Evolutionaries	Generate variations and assess the fitness of each for a purpose.	Genetic Programs
Analogizers	Optimize a function in light of constraints.	Support Vectors

Answer 1:
Naive Bayes or Markov
Answer 2:
Neural Networks
Answer 3:
Genetic Programs
Answer 4:
Support Vectors

Question 2
0.8 / 0.8 pts

LLM Training Challenges
During the lecture, we touched on the challenges of training large language models (LLMs) such as resource demands, massive dataset requirements, and ethical concerns. Pick one, and explain how they influence the scalability and practicality of deploying such models.

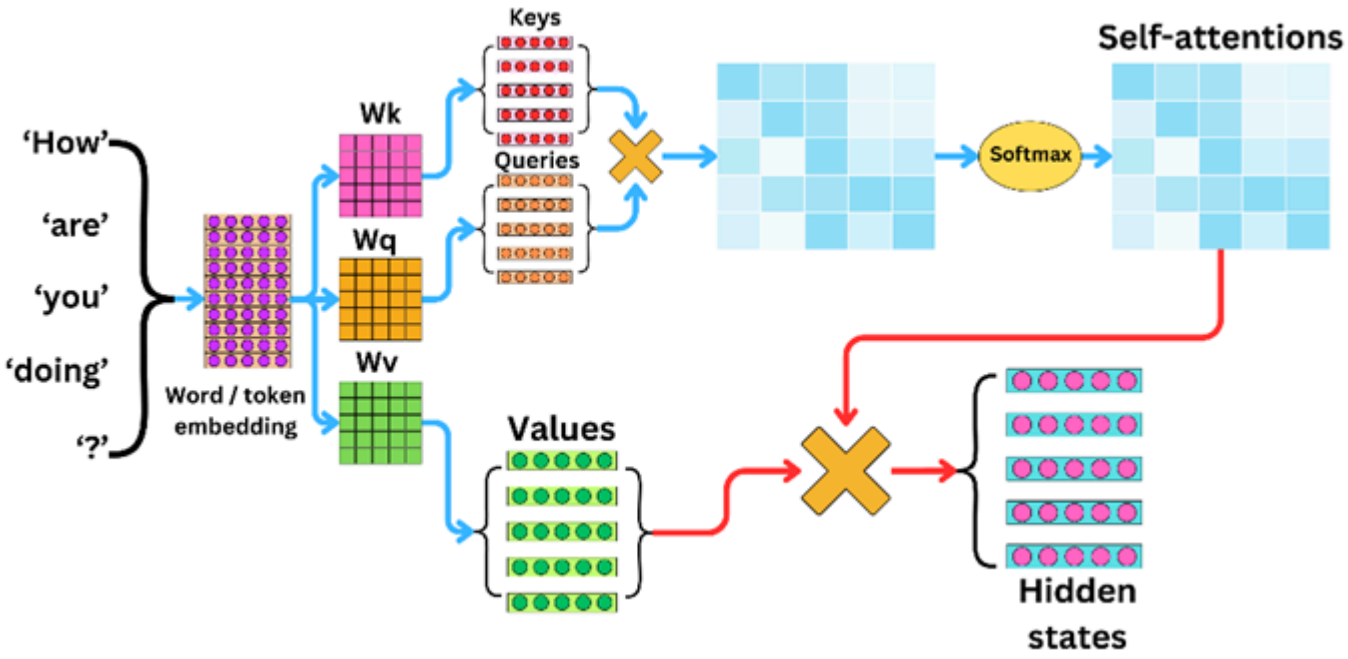
Your Answer:

Training LLMs requires quite a lot of computational power such as clusters of GPUs / TPUs working in parallel, which tends to be quite expensive and consume large amounts of energy. This limits accessibility to smaller organizations or individuals, and thus, only few companies have been able to produce such large language models trained on huge datasets, soemthing which is not practical or scalable for smaller organizations or individuals.

Question 3
0.4 / 0.8 pts

Self-Attention Mechanism
As discussed in class, the self-attention mechanism is a core innovation in transformers. Describe the two main steps involved in self-attention and explain why this mechanism outperforms previous approaches like RNNs.

Hint: First two steps in the following diagram below shows those main steps.



Your Answer:

1. Computing attention scores: for each token in the input, the mechanism computes attention scores with every other token using dot products on query and key vectors.

2. Weighted Sum: scores are transformed into probabilities using the softmax function, and a weighed sum of the value representation is calculated.

You missed a very important point/ step: 2. Generating context vectors: Combines token embeddings weighted by attention scores to capture relationships. Self-attention outperforms RNNs because it evaluates all tokens simultaneously, avoiding sequential bottlenecks and better capturing long-range dependencies.

Question 4
0.8 / 0.8 pts

Ethical Challenges in LLMs
As alluded to in our last lecture, ethical challenges are a critical consideration in deploying LLMs. Discuss a real-world ethical issue related to their use and suggest one practical mitigation strategy.
Hint: The Amazon resume example

Your Answer:

A real world issue is the bias and discrimination issue. An example, the Amazon resume basically shows that the model preferred male candidates over female candidates as the training data reflected societal bias.

A possible strategy could be to use balanced datasets which do not reflect this bias, auditing and intervening in such training data, which could add fairness.

Debiasing was discussed as a mitigation strategy in lecture 2

Question 5
0.8 / 0.8 pts

Transformer Architecture
As we discussed in Lecture 2, positional encodings i.e. encoding which provides information about the order of tokens in a sequence, plays a crucial role in transformer models. Why it is essential and how it differs from standard input embeddings. Provide an example to illustrate your point.

Your Answer:

Transformers treat tokens separately, unlike RNNs which preserve the order, thus require positional information to understand the order of tokens in a sequence. Standard token embeddings represent token meanings, while positional encodings capture each token's position using patterns.

In the sequence, "The mouse ate the cheese", without positional encoding, the transformer can treat it as "cheese ate The cat" which loses context.

Question 6
0.8 / 0 pts

Tokenization (Bonus)
As we explored during Lecture 2, tokenization significantly impacts transformer efficiency. Given the sentence: "Large language models are powerful tools," list the tokens generated by a tokenizer. How does tokenization enhance processing in LLMs?

Your Answer:

["Large", "language", "models", "are", "powerful", "tools"].

Tokenization breaks the text into smaller units, which can help in efficient computation reducing which improves model performance across diverse inputs.

Quiz Score: 4.4 out of 4