CSE665: Large Language Models - Midsem

Date of Examination: 28.02.2024 Duration: 60 mins Total Marks: 30 marks

Instructions

- It is a close book examination.
- All questions are mandatory and there is no negative marking.
- MCQ has just one correct answer. Choose the most appropriate response.

Multiple Choice Questions: [10 Marks]

Q1. Large Language Models like GPT-3 have been used to generate text that mimics the writing style of famous authors or personalities. What is this technique called?

a. Language modeling

b. Plagiarism

c. Style transfer

d. Identity theft

Q2. When designing prompts for a language model, which of the following is the most important consideration for addressing bias and fairness issues?

- a. Ensuring that prompts contain explicit constraints on model behavior
- b. Incorporating comprehensive demographic information in prompts
- c. Crafting prompts that are free of cultural and gender-specific references
- d. Implementing post-processing filters to modify model outputs

Q3. Which activation function is commonly used in the recurrent layers of an RNN?

a. ReLU (Rectified Linear Unit)

b. Sigmoid

c. Tanh (Hyperbolic Tangent)

d. Softmax

Q4. How does GPT-3 typically generate text output?

- a. By memorizing and regurgitating predefined responses
- b. By applying deterministic rules and heuristics
- c. By sampling from a probability distribution over words, using a combination of learned patterns and context to generate text
- d. By performing keyword-based searches on the internet

Q5. In transformer-based models, what is the "position encoding" used for?

- a. To encode the geographical positions of data samples
- b. To indicate the order of words in a sequence, as the transformer architecture does not have built-in positional information
- c. To encode the position of the model's parameters
- d. To control the learning rate

Q6. How can domain-specific knowledge contribute to mitigating bias in LLMs?

- a. By introducing domain-specific bias into the training data
- b. By enabling better understanding and identification of biased language
- c. By restricting the model's ability to learn biased language patterns
- d. By reducing the need for bias mitigation techniques in specialized domains
- Q7. What does the Stereotype Score (SS) represent in language models?
 - a. The percentage of instances where the model chooses neutral language over biased language.
 - b. The measure of how often the model selects stereotypical language compared to anti-stereotypical language.
 - c. The degree of accuracy in detecting stereotypes within text.
 - d. The proportion of times the model generates creative language versus conventional language.

Q8. What is the significance of the "pre-training" and "fine-tuning" approach used in Large Language Models?

- a. Pre-training refers to training a model from scratch, while fine-tuning adapts it to specific tasks
- b. Pre-training involves optimizing the model for a specific task, while fine-tuning is a general training phase
- c. Pre-training is the process of compressing the model, while fine-tuning is for model expansion
- d. Pre-training and fine-tuning are synonymous and used interchangeably

Q9. What are the numerical values used to evaluate the outcomes of actions taken by an agent?

a. Observations

b. Rewards

c. Policies

d. Loss functions

Q10. What is "few-shot learning" in the context of LLMs?

- a. Learning with only a small amount of data
- b. Learning with a small number of input features
- c. Learning with a small learning rate
- d. A technique for fine-tuning pre-trained models

Descriptive Questions and Answers: [20 Marks]

Q11. Discuss the concept of bias in Large Language Models (LLMs). Describe its potential impact on

LLM outputs and how bias can be mitigated in LLMs. [5 Marks]

- Q12. Explain Smoothing and Discounting Techniques. [5 Marks]
- Q13. Explain RNN and limitations of RNN overcome by LSTM. [5 Marks]
- Q14. Why is RL using Human Feedback (RLHF) good? [5 Marks]