Questions and Answers - Chapters 11 to 13

Chapter 11: Multilingual and Multimodal LLMs

True/False:

Q1. Multilingual LLMs use same scripts.

Answer: False

Q2. Multimodal LLMs cannot process video.

Answer: False

Q3. Cross-lingual transfer enables zero-shot tasks.

Answer: True

Q4. Contrastive learning only used in multimodal models.

Answer: False

Q5. Curse of multilinguality reduces model performance.

Answer: True

Multiple Choice:

Q6. Not a pretraining task for multilingual LLMs:

(a) Translation (b) MLM (c) Image captioning (d) Cross-lingual retrieval

Answer: (c) Image captioning

Q7. Instruction tuning improves:

(a) Parameter freezing (b) Following task instructions (c) Output compression (d) Static indexing

Answer: (b) Following task instructions

Q8. Key challenge for multimodal LLMs:

(a) Aligning modalities (b) Reducing parameters (c) Increasing labels (d) Shorter outputs

Answer: (a) Aligning modalities

Q9. Parallel data means:

(a) Augmented prompts (b) Same text in different languages (c) Labeled dataset (d) JSON format

Answer: (b) Same text in different languages

Q10. Oversampling helps:

(a) High-resource language boost (b) Low-resource language balance (c) Increase hallucinations (d) Longer training

Answer: (b) Low-resource language balance

Chapter 12: Responsible AI and Bias Mitigation

True/False:

Q11. Inherent bias reflects social perspectives.

Answer: True

Q12. Responsible AI aims balanced demographic representation.

Answer: True

Q13. Annotator quality check is wasteful.

Answer: False

Q14. Fine-tuning can amplify bias.

Answer: True

Q15. Debiased models never show bias.

Answer: False

Multiple Choice:

Q16. Token blocking applies to:

(a) Prompt-based (b) Training-data based (c) Decoding strategies (d) Retrieval-only

Answer: (a), (b), (c)

Q17. Culture-based info causes:

(a) Overgeneralization HRL (b) No impact (c) Under-generalization LRL (d) Improved embedding

Answer: (a) and (c)

Q18. Counterfactual data augmentation:

(a) Prompt method (b) Training-data method (c) Retrieval method (d) Evaluation-only

Answer: (b) Training-data method

Q19. Wikipedia limits include:

(a) English-dominance (b) Encyclopedic tone (c) Unverified data (d) All of the above

Answer: (d) All of the above

Q20. African-American English (AAE) issues:

(a) Always high-resource (b) Underrepresented (c) Overgeneralized by standard English (d) Never misclassified

Answer: (b) and (c)

Chapter 13: Latest Architectures (LLaMA2, PaLM, Mistral, Mixtral)

True/False:

Q21. Decoder-only models are good encoders.

Answer: False

Q22. LLaMA2 supports longer context.

Answer: True

Q23. DeBERTa uses disentangled attention.

Answer: True

Q24. LLaMA2 applies grouped-query attention.

Answer: True

Q25. ZERO optimizer dynamically manages communication.

Answer: True

Multiple Choice:

Q26. Sliding window attention benefits:

(a) Faster inference (b) Better memory (c) Parallelism (d) More heads

Answer: (a) Faster inference

Q27. Mistral model uses:

(a) Multi-query (b) Random attention (c) Grouped-query attention (d) No attention

Answer: (c) Grouped-query attention

Q28. Models using ZERO:

(a) LLaMA2 (b) GPT-2 (c) Turing-NLG (d) OPT

Answer: (a) and (c)

Q29. Features of Mixtral:

(a) Sparse experts (b) Weight access (c) Private license (d) Dialogue from social media

Answer: (a), (b), and (d)

Q30. Characteristics of PaLM:

(a) Successor to LaMDA (b) Sparse expert model (c) Implements masked modeling (d) Fixed-length prompts

Answer: (a), (b), and (c)

Prepared for the User

Compiled by ChatGPT