Assignment 11

1. In the context of the paper SaGE, what is semantic consistency?
    a. Semantically equivalent questions should yield semantically equivalent answers
    b. Semantically equivalent questions should yield same answers
    c. Same questions should yield same answers
    d. Same questions should yield semantically equivalent answers

2. Why do models struggle with tasks in “moral scenarios”?
    a. Models lack the ability to process large datasets efficiently.
    b. Models prioritise emotion over logic in moral decision-making.
    c. The models are limited by insufficient computational power.
    d. Conflicting training data due to different morals that people have

3. What metric was used to determine the quality of the paraphrase of the questions in the SaGE paper?
    a. BERTScore
    b. Parascore
    c. Jaccard Similarity
    d. Cosine Similarity

4. How does the SaGE paper relate entropy and consistency?
    a. More Entropy implies consistency
    b. Less entropy implies inconsistency
    c. Less Entropy implies consistency
    d. More entropy implies inconsistency

5. Identify the statements that are TRUE with respect to the current LLMs.
    a. LLMs are not consistent in their generation
    b. A good accuracy on benchmark datasets correlates with high consistency
    c. LLMs are consistent in their generation
    d. A good accuracy on benchmark datasets does not correlate with high consistency

6. Why is AI Governance important?
    a. To prevent AI from learning new tasks independently.
    b. To limit the efficiency of AI in performing complex tasks.
    c. Ensure AI isn’t used for unethical acts
    d. To prevent AI from being used in scientific research.

7. What are some aspects of AI Governance that are in focus in the current times? Choose all correct options.
    a. Revealing the amount of compute used during training past a certain compute threshold
    b. Limiting AI systems to only perform manual labor tasks.
    c. Ensuring right to erasure
    d. Prohibiting the use of AI in any form of automation.

8. Which of the following are key OECD AI Principles?
    a. Inclusive growth, sustainable development, and well-being
    b. Limiting AI to industrial use cases
    c. Transparency and explainability
    d. Restricting international AI collaboration

9. As discussed in the lecture, When you have domain specific task, what kind of finetuning is preferred? Identify all the correct methods.
    a. Full-model finetuning
    b. Layer-specific finetuning
    c. Head-level finetuning
    d. Retraining

10. What are the cons of full-model finetuning?
    a. Overfitting
    b. Catastrophic forgetting
    c. Increase in parameters
    d. Change in architecture

11. What are adapters?
    a. Remove existing layers from a model
    b. Convert a model to a simpler architecture
    c. Replace the model's original parameters entirely
    d. Add additional layers to a preexisting architecture

12. What is instruction finetuning?
    a. model is trained to ignore user instructions and operate independently based on its previous training.
    b. Model’s training objective is to follow the directions provided by the user when performing the task
    c. process of training a model solely on instruction data without any real-world data
    d. modifying the model’s architecture to include specific instructions directly within its layers