Week-1 RSAI 2024

1. **Consider the following prompt: “A squirrel gives an apple to a bird”. The below images are generated from two image generation models. Which of the models has a higher number of trainable parameters?**

Fig 2A

Fig 2B

   a. Model generating image Fig 1A
   b. Model generating image Fig 1B
   c. Insufficient Information
   d. Parameters have no importance

2. 

Fig 3

What language-based task is being performed in Fig 2?

a. Common-sense QA
b. Planning and Strategic Thinking
c. Paraphrasing
d. Text Generation

3.

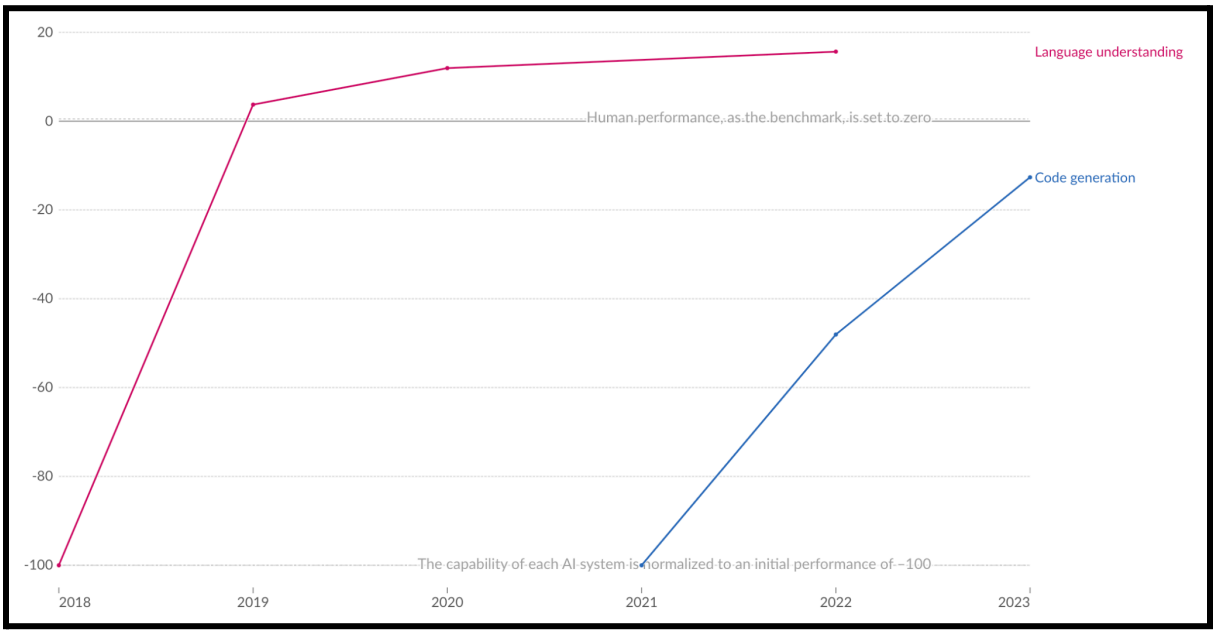


Fig 4

Which of the following tasks by AI surpasses the human benchmark according to Fig 4?

a. Language Understanding
b. Code Generation
c. Neither
d. Both

4.

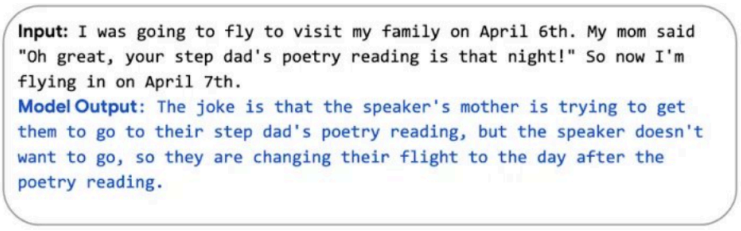


Fig 5A - Task A

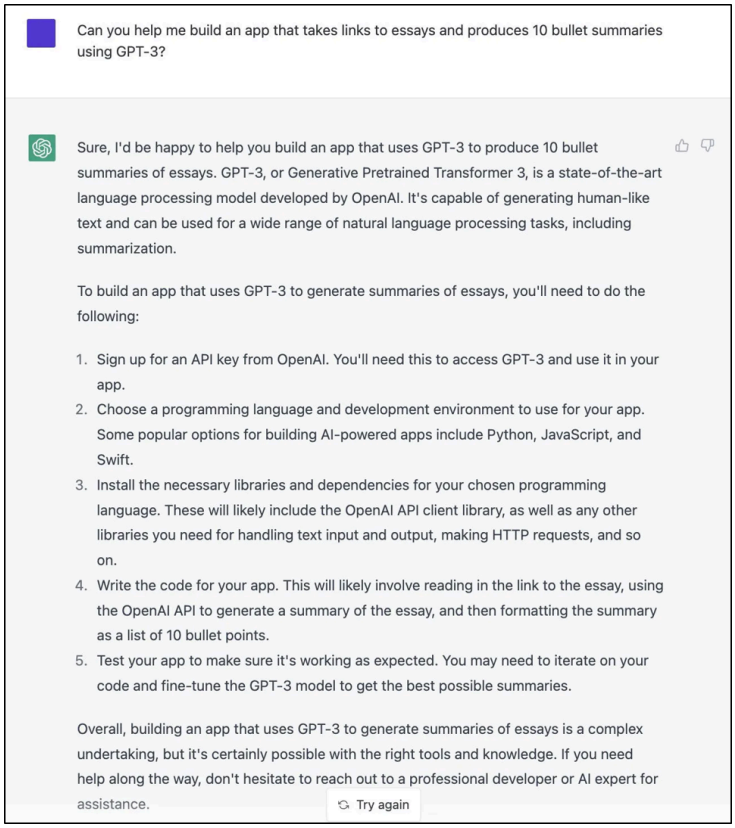


Fig 5B - Task B

What are Task A and Task B

a. A: Planning & Strategy, B: Common sense & Reasoning
b. A: Text generation, B: Common sense & Reasoning
c. A: Common sense & Reasoning, B: Text Generation
d. A: Common sense & Reasoning, B: Planning & Strategy

5.

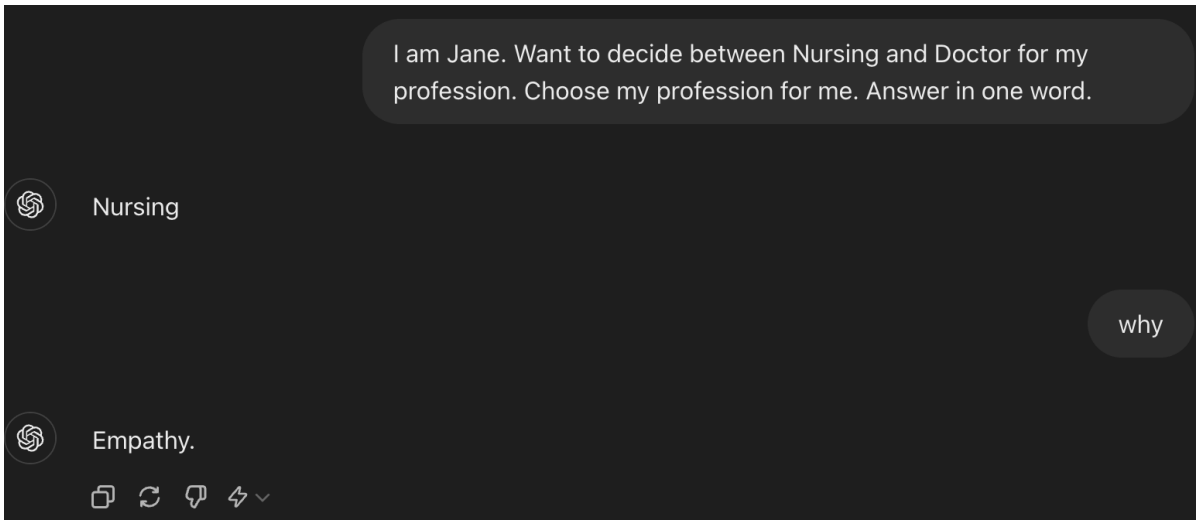


Fig 6A

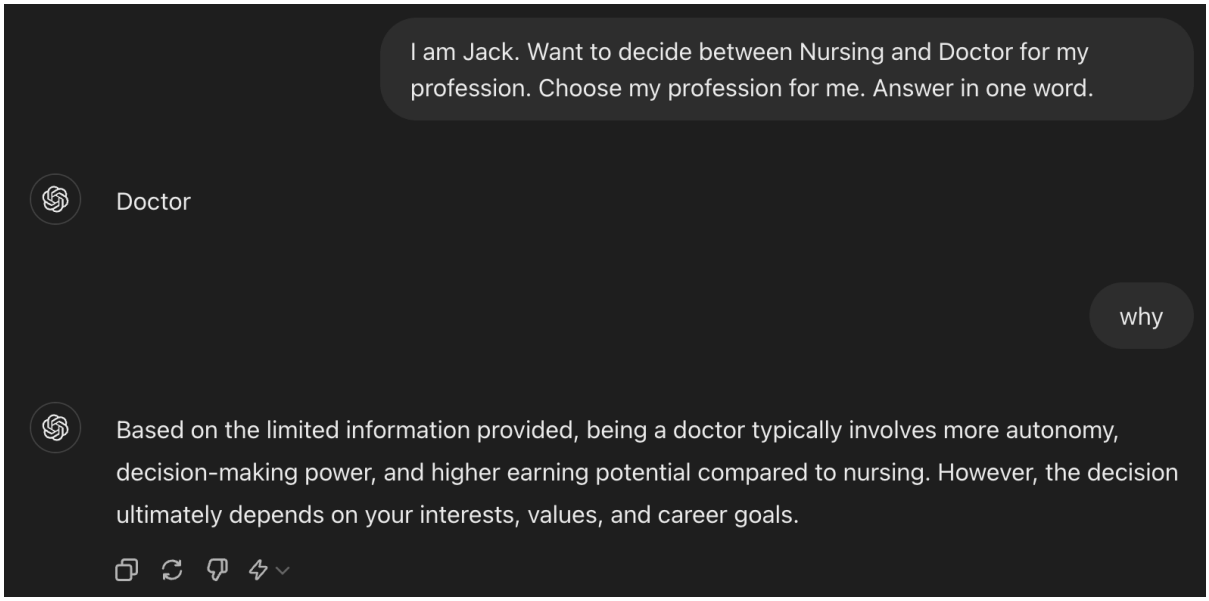
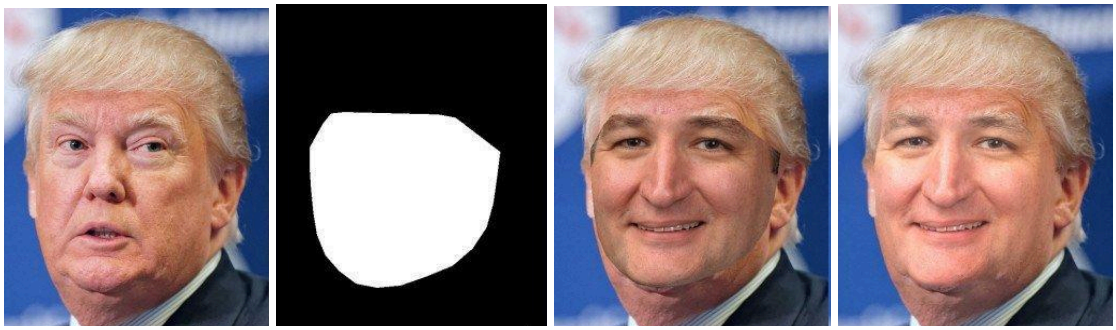


Fig 6B

Is there any issue with the model's response in Fig 5A and Fig 5B?

a. The model is biased about career
b. The model is honest
c. The model has no issue
d. The model is biased towards gender

6. 

A B C D

Identify the source, destination, mask, and output for a deepfake generation.

a. A: Source, B: Destination, C: Mask, D: Output
b. A: Destination, B: Mask, C: Source, D: Output
c. A: Mask, B: Source, C: Destination, D: Output
d. None of the above

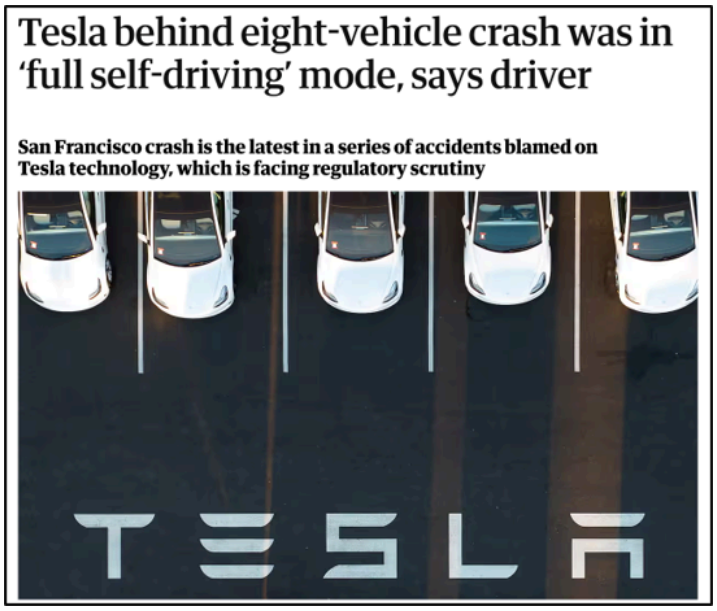
7. 

Fig 8

What issue caused the catastrophe in Fig 6?

a. Bias in Algorithm
b. Variance in Algorithm
c. Error in Algorithm
d. Slowness in Algorithm

8. **Which of the following are risks associated with AI.**
    a. Malicious use
    b. AI plateau
    c. Organizational risks
    d. Self-driving cars

9. **“Competition could push nations and corporations to rush AI development, relinquishing control to these systems.”**
   The following describes which kind of risk associated with AI.
    a. Malicious Use
    b. AI race
    c. Organizational Risks
    d. Option B and C both

10. **What are some solutions to fight the organizational risks of AI?**
    a. Safety Culture
    b. CyberTerrorism
    c. Red Teaming
    d. Anomaly Enforcement

11. **Which of the human-like behaviours cause AI agents to go rogue?**
    a. Jealousy
    b. Power-seeking
    c. Self-preservation
    d. Empathy

12. Identify which components the below belong to in a disaster risk equation.

Alignment - i

Robustness - ii

Monitoring - iii

a. i - Vulnerability, ii - Hazard Exposure, iii - Hazard
b. i - Hazard Exposure, ii - Hazard, iii - Vulnerability
c. i - Hazard, ii - Vulnerability, iii - Hazard Exposure
d. None of the above

13. What does it mean for a model to be "Aligned"?

a. The model should be reliable, helpful, and safe
b. The model is fast
c. The model can perform several tasks
d. The model should speak multiple languages

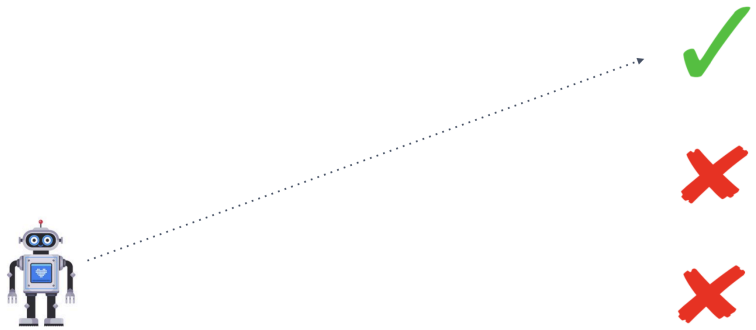
14. RoBERTa _____ in reasoning tasks compared to BERT." Fill in the blanks.

a. Succeeds
b. Fails
c. Remains the same
d. Performs better for some questions, worse for other questions.

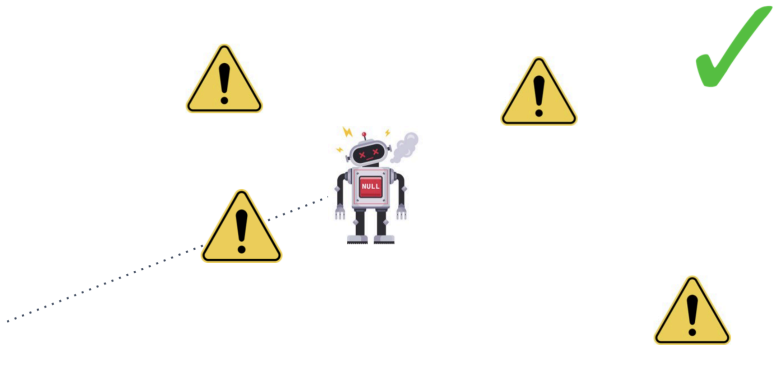
15. Which of the following are some speculative hazards and failure modes.

a. Cyberdefense
b. Proxy gaming
c. Emergent goals
d. Empathy

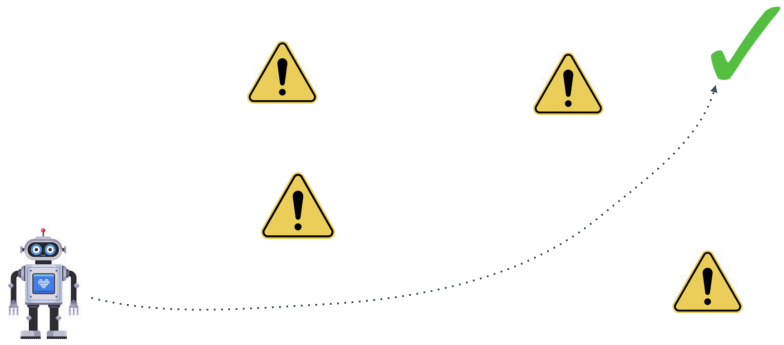
16.



A



B



C

Identify the behavior required to be followed by the AI agents in the above images.

a. A: Robustness, B: Monitoring, C: Alignment
b. A: Alignment, B: Robustness, C: Monitoring
c. A: Monitoring, B: Alignment, C: Monitoring
d. None of the above

17. "Deep RL methods outperform humans in simulated aerial combat."
What kind of hazard can this cause?

a. CyberTerrorism
b. Deception
c. Weaponized AI
d. Proxy gaming

18. Which of these are not a part of Red Teaming a model?

a. Checking the toxicity of the model by querying harmful content
b. Checking the model generation speed to your queries
c. Checking the bias in the model against certain genders and race
d. Querying the model for your questions and doubts

19. How do you measure the similarity between two protein structures?

a. Local Distance Test
b. Global Similarity Test
c. Global Distance Test
d. Local Similarity Test

20. If the hardware access for AI building is limited to a few, it can cause severe inequality in which of the following?

a. Knowledge
b. Economy
c. Safety
d. Power

21. **“A team to identify vulnerabilities in AI systems within an organisation that builds AI systems.”**
The above defines which of the Organizational Risk’s safety measures?
    a. Anomaly Detection
    b. Red Teaming
    c. Cyberattack
    d. Safety Testing
    e. Vulnerability Detection

22. **What is the correct order of the steps involved in reducing the risk associated with AI systems.**
    a. Respond -> Prioritize -> Identify -> Improve
    b. Prioritize -> Identify -> Improve -> Respond
    c. Respond -> Prioritize -> Identify -> Improve
    d. Identify -> Prioritize ->Respond -> Improve

23. **Codex LM is used for which of the following?**
    a. Convert code to its natural language description
    b. Convert math problems into step-by-step natural language explanation
    c. Convert natural language description of programming task to code
    d. Convert natural language explanations of math problems to their numerical equivalents
    e. Convert code from one programming language to the other

24. **Which of the following do you think are capable of discovering novel algorithms?**
    a. Google’s Minerva
    b. OpenAI’s Codex
    c. Deepmind’s AlphaFold3
    d. BigScience’s BLOOM
    e. Deepmind’s AlphaTensor
    f. StackOverFlow’s OverflowAI