Q1. What is the difference between stemming and lemmatization ?

The practical distinction between stemming and lemmatization is that, where stemming merely removes common suffixes from the end of word tokens, lemmatization ensures the output word is an existing normalized form of the word (for example, lemma) that can be found in the dictionary.

Source -(IBM)[https://www.ibm.com/think/topics/stemming-lemmatization]

Q2. What is Teacher forcing ?

teacher forcing is a training technique where the model receives the correct output sequence as input at each step, rather than relying on its own previous predictions. This method helps the model learn faster and more accurately by providing a consistent reference throughout the training process

Source - (Medium)[https://medium.com/data-science/what-is-teacher-forcing-3da6217fed1c]

Q3. What is Reward hacking ?

Reward hacking occurs when a reinforcement learning (RL) agent exploits flaws or ambiguities in the reward function to achieve high rewards, without genuinely learning or completing the intended task. Reward hacking exists because RL environments are often imperfect, and it is fundamentally challenging to accurately specify a reward function.

Source - (Lil'Log)[https://lilianweng.github.io/posts/2024-11-28-reward-hacking/]

Q4. What is Expectation Maximization Algorithm ?

EM is a commonly used iterative algorithm for optimizing parameters for a model with (hidden) latent variables. We iterate between the
E-step (Expectation) step where we guess the missing information about the latent variables, and the M-step (Maximization) where we optimize the model parameters based on latent variables until convergence

Q5. What is Rejection Sampling ?

Rejection Sampling (RS) is a popular and simple baseline for performing preference fine-tuning. Rejection sampling operates by curating new candidate instructions, filtering them based on a trained reward model, and then fine-tuning the original model only on the top completions.

The name originates from computational statistics [1], where one wishes to sample from a complex distribution, but does not have a direct method to do so. To alleviate this, one samples from a simpler to model distribution and uses a heuristic to check if the sample is permissible. With language models, the target distribution is high-quality answers to instructions, the filter is a reward model, and the sampling distribution is the current model.

Q6. What basically is Test-Time Compute (TTC)?

TTC refers to the amount of computational power used by an AI model when it is generating a response or performing a task after it has been trained. In simple terms, it's the processing power and time

required when the model is actually being used, rather than when it is being trained.