AWS Future AI Engineer Intelligent Heavy Machinery Document Querying System Project

Temperature and Top _P Explaination:

- Temperature and `top_p` are crucial parameters that influence the randomness and diversity of AI-generated responses and their values range from 0 to 1. Temperature controls the "creativity" of the model: a higher temperature (e.g., 0.8) makes the output more diverse and unexpected, as the model is more likely to choose less probable words. Conversely, a lower temperature (e.g., 0.2) makes the output more deterministic and focused, sticking to the most probable words, which can be useful for factual or precise responses.

- `Top_p`, also known as nucleus sampling, offers another way to control the diversity by focusing on a cumulative probability. Instead of considering all possible words, `top_p` samples from the smallest set of words whose cumulative probability exceeds the `top_p` value. For example, if `top_p` is set to 0.9, the model will only consider words that account for the top 90% of the probability distribution. This can help prevent the generation of truly nonsense words while still allowing for some variability, making it a more refined control over output diversity compared to temperature in some scenarios.

- One last thing to mention about the project i encountered this issue during my project:

  - Error validating prompt: An error occurred (AccessDeniedException) when calling the InvokeModel operation: User: arn:aws:sts::750960077820:assumed-role/voclabs/user4420338=4c8d97a2-0f8d-11ee-898f-1fe6059501c7 is not authorized to perform: bedrock:InvokeModel on resource: arn:aws:bedrock:us-east-1::foundation-model/anthropic.claude-3-haiku-20240307-v1:0 with an explicit deny in an identity-based policy

- This prevented me from testing and inferring model responses directly throughAWSBedrock's API. The error indicated that the IAM role provided by the Udacity AWS sandbox account had restricted permissions for the `bedrock:InvokeModel` action, explicitly denying access to all Bedrock foundation models, including the Anthropic Claude family.