NAME – Pranav S

SRN – PES2UG23CS430

SECTION – G

Task	Model	Classification (Success/Failure)	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BERT	Failure	The model failed to extend the prompt meaningfully and did not generate fluent or complete text.	BERT is an encoder-only model and is not trained for autoregressive next-token prediction required for text generation.
	RoBERTa	Failure	Similar to BERT, the model did not generate a coherent continuation of the prompt.	RoBERTa is also encoder-only and lacks a decoder to generate sequential output.
	BART	Success	The model generated a coherent continuation of the input prompt.	BART is an encoder–decoder model designed for sequence-to-sequence generation tasks.
Fill-Mask	BERT	Success	Correctly predicted words such as "Paris", "Lille", and "Lyon" for the masked position.	BERT is trained using Masked Language Modeling (MLM), which directly optimizes it to predict masked tokens.
	RoBERTa	Failure	The task failed with an error stating that no `<mask>` token was found in the input.	RoBERTa requires a specific `<mask>` token format; the failure occurred due to incorrect mask token handling rather than model capability.
	BART	Success	Predicted valid tokens such as "Paris" and "Lyon" for the masked word, though with lower confidence scores.	BART supports masked token prediction but is primarily optimized for sequence-to-sequence generation rather than pure MLM.
QA	BERT	Failure	Returned a partial answer span with low confidence, even though the text appeared related to the context.	The base BERT model is not fine-tuned for Question Answering; its QA head is randomly initialized.
	RoBERTa	Failure	Returned a short answer span ("deepfakes") with low confidence.	RoBERTa-base is not fine-tuned on QA datasets such as SQuAD, leading to weak answer extraction.
	BART	Failure	Returned an incomplete phrase ("such as") instead of a clear answer.	BART-base is not fine-tuned for extractive QA and is primarily designed for generative sequence-to-sequence tasks.