GenAi Hands-on -1

Section:- F

NAME:- Mohit Kumar

SRN:- PES2UG23CS350

6TH SEM

Observation Table

Instructions: After running all experiments above, fill out this table with your observations.

Task	Model	Classification (Success/Failure)	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BERT	Failure	Generated only dots/periods (meaningless)	BERT is Encoder-only; not trained for autoregressive text generation
	RoBERTa	Failure	Generated nothing new, just echoed input	RoBERTa is Encoder-only; lacks decoder for generation
	BART	Success	Generated coherent text continuation	BART is Encoder-Decoder; designed for generation tasks
Fill-Mask	BERT	Success	Predicted "create" (53.97%), "generate" (15.58%)	BERT is trained on Masked Language Modeling (MLM)
	RoBERTa	Success	Predicted "generate" (37.11%), "create" (36.77%)	RoBERTa is optimized for MLM task
	BART	Success	Predicted "create" (7.46%) but lower confidence	BART has masking capability but not optimized for it
QA	BERT	Partial Failure	Wrong answer, very low confidence (0.77%)	Base model not fine-tuned for QA task
	RoBERTa	Partial Success	Correct answer but low confidence (0.90%)	Base model not fine-tuned, but encoder handles context
	BART	Failure	Wrong answer "Generative" (7.13% confidence)	Base model not optimized for extractive QA