Generative AI and its Applications

Hands On -1 Unit -1

Name: Monisha Sharma
SRN: PES2UG23CS906
Class: 6F
Date: 27-01-2026

Comparison of Transformer Models (BERT, RoBERTa, BART)

Task	**Model**	**Success / Failure / Partial**	**Observation (What actually happened?)**	**Why did this happen? (Architectural Reason)**
Text Generation	BERT (bert-base-uncased)	Failure	Could not generate meaningful text; gibberish or empty output.	Encoder-only model trained with Masked Language Modeling, not autoregressive generation.
	RoBERTa (roberta-base)	Failure	Failed to produce coherent sentences, similar to BERT.	Encoder-only architecture optimized for MLM, not designed for generation.
	BART (facebook/bart-base)	Success	Generated coherent continuations of the prompt.	Encoder–decoder (seq2seq) architecture with decoder trained for generation.
Fill-Mask	BERT	Success	Predicted sensible words like create, generate, produce.	Masked Language Modeling is BERT's native pretraining task.
	RoBERTa	Success	Produced strong predictions, often better than BERT.	Improved training strategy over BERT, excels at MLM.
	BART	Partial	Filled the mask but less accurately than BERT or RoBERTa.	Includes MLM but primary strength is seq2seq generation.
Question Answering	BERT	Partial	Sometimes extracted relevant spans but often incomplete.	Base BERT is not fine-tuned for QA datasets like SQuAD.
	RoBERTa	Partial	Occasionally picked relevant terms but inconsistently.	Encoder-only model without QA-specific fine-tuning.
	BART	Partial	Sometimes generated reasonable answers but not always exact.	Seq2seq decoder helps generation, but lack of QA fine-tuning causes inconsistency.