Unit 1 Assignment: The Model Benchmark Challenge

Name	Diya D Bhat
SRN	PES2UG23CS183
Section	C

Output Table

Task	Model	Classification	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BERT	Failure	Generated nonsense or repetitive text.	Encoder-only model; not trained for text generation.
Generation	RoBERTa	Failure	Failed to generate meaningful continuation.	Encoder-only architecture; no autoregressive decoding.
Generation	BART	Success	Generated fluent and meaningful text.	Encoder–decoder model designed for generation.
Fill-Mask	BERT	Success	Correctly predicted words like "create".	Trained using Masked Language Modeling (MLM).
Fill-Mask	RoBERTa	Success	Predicted accurate and fluent alternatives.	Optimized MLM training with more data.
Fill-Mask	BART	Partial Success	Predicted reasonable but less precise words.	Not primarily trained for MLM tasks.
QA	BERT	Partial Success	Returned short or incomplete answers.	Not fine-tuned for question answering.
QA	RoBERTa	Partial Success	Slightly better but inconsistent answers.	Base model without QA fine-tuning.
QA	BART	Failure	Output was vague or incorrect.	Not designed for extractive QA tasks.