

## Gen-AI Semester-6

**LAB-1**

**Date:** 23/01/2026

**Name:** Ankana Mandal

**SRN:** PES2UG23CS076

**Section:** B

# Documenting the Benchmark Assignment

The objective of the benchmark assignment was to study the behavior of different transformer architectures by forcing them to perform tasks they are not equally suited for. This experiment helps in understanding why model architecture matters in Generative AI.

Three transformer models were used for the benchmark:

- BERT (bert-base-uncased) – an encoder-only model
- RoBERTa (roberta-base) – an optimized encoder-only model
- BART (facebook/bart-base) – an encoder-decoder model

All three models were evaluated on the same set of tasks using Hugging Face pipelines.

## Tasks Performed

### 1. Text Generation

The models were asked to generate text for a given prompt.

It was observed that BERT and RoBERTa failed to generate meaningful text because encoder-only models are not designed for next-token generation. BART, having a decoder component, was able to generate text, though the output quality was limited.

### 2. Fill-Mask (Masked Language Modeling)

In this task, a sentence containing a masked token was provided.

BERT and RoBERTa performed well, correctly predicting suitable words due to their training on masked language modeling. BART showed weaker performance as it is not primarily optimized for this task.

### 3. Question Answering

The models were asked to answer a question based on a given context.

The outputs were inconsistent because the base versions of these models are not fine-tuned specifically for question answering. However, partial or approximate answers were sometimes produced.

## Observations

The benchmark clearly demonstrated that:

- Encoder-only models such as BERT and RoBERTa are best suited for understanding tasks like fill-mask.
- Encoder-decoder models like BART are more flexible for generation-related tasks.
- Using a model for a task it was not designed for leads to poor or failed outputs, which is an expected and valid observation.

## Conclusion

This benchmark assignment highlights the importance of transformer architecture in Generative AI. It shows that model success or failure depends on how well the task aligns with the model's design, reinforcing key concepts covered in Unit 1.