

GenAI HandsOn Project

NAME : CHITRA MADARAKHANDI

SRN : PES2UG23CS157

SECTION : C

Poetry & Song Writer

Description :

This project focuses on building a Creative Poetry and Song Generator using Transformer-based language models.

Multiple NLP tasks—Text Generation, Fill-Mask, and Question Answering—were experimentally evaluated using different transformer architectures such as BERT, RoBERTa, BART, and GPT-2.

The main objective was to identify the most suitable model architecture for creative text generation. Based on comparative observations, GPT-2 Medium was selected as the final model for generating theme-based rhyming poems due to its strong generative capabilities.

Table :

Task	Model	Classification (Success / Failure)	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BERT	Failure	Failed to generate continuous or meaningful text.	BERT is an encoder-only model and does not support autoregressive decoding.
Generation	RoBERTa	Failure	Produced incoherent or unusable output for text generation.	RoBERTa is encoder-only and not trained for next-token prediction.

Task	Model	Classification (Success / Failure)	Observation (What actually happened?)	Why did this happen? (Architectural Reason)
Generation	BART	Success	Generated coherent and context-aware text.	BART uses an encoder-decoder architecture designed for generative tasks.
Generation	GPT-2	Success	Generated fluent, creative, and rhyming poems for given themes.	GPT-2 is a decoder-only autoregressive model trained for text generation.
Fill-Mask	BERT	Success	Correctly predicted masked words in sentences.	BERT is trained using Masked Language Modeling (MLM).
Fill-Mask	RoBERTa	Success	Predicted accurate and contextually relevant masked tokens.	RoBERTa improves MLM with dynamic masking and larger datasets.
Fill-Mask	BART	Success	Successfully filled masked tokens with fluent language.	BART learns denoising autoencoding, enabling mask recovery.
QA	BERT	Success	Extracted correct answers from the given context.	BERT is well-suited for extractive question answering tasks.
QA	RoBERTa	Success	Provided more precise answers compared to BERT.	RoBERTa has stronger contextual representations due to improved training.
QA	BART	Success	Generated correct answers but occasionally added extra text.	BART is generative and may produce abstractive-style responses.

Key Observations

- Encoder-only models (BERT, RoBERTa) fail at text generation.
- Decoder-based and encoder-decoder models (GPT-2, BART) succeed in generation tasks.
- BERT and RoBERTa perform best for Fill-Mask and Question Answering.
- GPT-2 produces the most fluent and creative poetic text.
- Model architecture directly impacts task suitability and performance.

Conclusion

The experimental evaluation clearly demonstrates that model architecture plays a critical role in NLP task performance.

While encoder-only models like BERT and RoBERTa are highly effective for understanding-based tasks such as Fill-Mask and Question Answering, they are unsuitable for creative text generation.

Decoder-based models like GPT-2, trained for autoregressive text generation, are ideal for applications such as poetry and song generation. Hence, GPT-2 Medium was selected as the final model for this project.