

Project Report

WEB PLATFORM FOR QUOTE GENERATION USING FINE-TUNED DEEPSEEK R1 MODEL

Team Members

Team Member 1

Name : Avhishek Nandi
Enrollment No : 12022002016039
Class Roll No. : 09
Department : CSE(AIML)
Year: 3rd

Team Member 2

Name : Barnali Paul
Enrollment No : 12022002016050
Class Roll No. : 10
Department : CSE(AIML)
Year: 3rd

Web Platform for Quote Generation Using Fine-Tuned DeepSeek R1 Model

Project Overview

This project presents a complete pipeline—from fine-tuning a generative AI model to deploying it on a web-based platform—demonstrating the creative power of artificial intelligence. The application generates original, contextually meaningful quotes based on input text, using a **fine-tuned version of the DeepSeek R1 (7B)** model.

The platform is developed using **Streamlit**, offering an intuitive and responsive interface. The core AI model is fine-tuned on a curated dataset of inspirational quotes, aligning its generative capabilities with human-style writing and emotional tone. This project illustrates the application of **transfer learning** to tailor a powerful language model to a niche task.

Objective

- To fine-tune a large-scale generative model using a dataset of quotes.
- To create a user-facing **Streamlit web application** for real-time quote generation.
- To evaluate and demonstrate the performance of the trained model.
- To illustrate the capabilities of transfer learning in creative language generation.

Tools and Technologies Used

Category	Tools / Libraries
Programming Language	Python

Model & Fine-Tuning	DeepSeek R1 (7B), Hugging Face Transformers, PEFT
Tokenization	DeepSeekTokenizer
Web Development	Streamlit
Training Environment	Google Colab
Evaluation	BLEU, ROUGE, Perplexity

Dataset

The dataset includes thousands of well-structured quotes from various sources. Each entry typically contains:

- A **prompt** (context)
- A **target quote** that aligns thematically with the prompt

This data is used to teach the model how to generate quotes that are stylistically consistent and relevant to the input idea.

```
{'quote': '"I\'m selfish, impatient and a little insecure. I make mistakes, I am out of control and at times hard to handle. But if you can\'t handle me at my worst, then you sure as hell don\'t deserve me at my best."',
  'author': 'Marilyn Monroe',
  'tags': ['best',
    'life',
    'love',
    'mistakes',
    'out-of-control',
    'truth',
    'worst']}
```

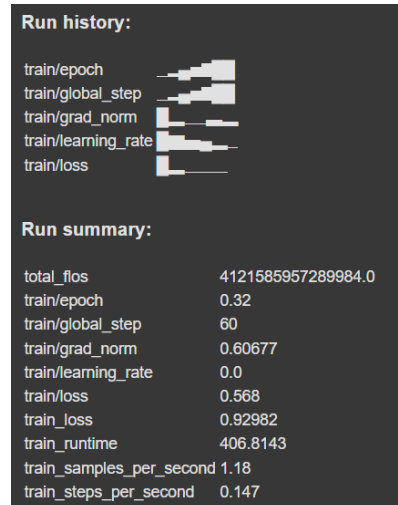
Model and Methodology

1 Base Model: DeepSeek R1 (7B)

DeepSeek R1 is a transformer-based language model with 7 billion parameters. Pre-trained on a large corpus of internet text, it is capable of understanding language structure, semantics, and emotional tone, making it suitable for creative generation tasks.

2 Fine-Tuning Approach: Supervised Fine-Tuning (SFT)

We used **Supervised Fine-Tuning** to adapt the general-purpose model to our specific task of quote generation. The model was trained on structured instruction-response pairs that guide it to produce quotes given a prompt.



3 Framework: Hugging Face + PEFT

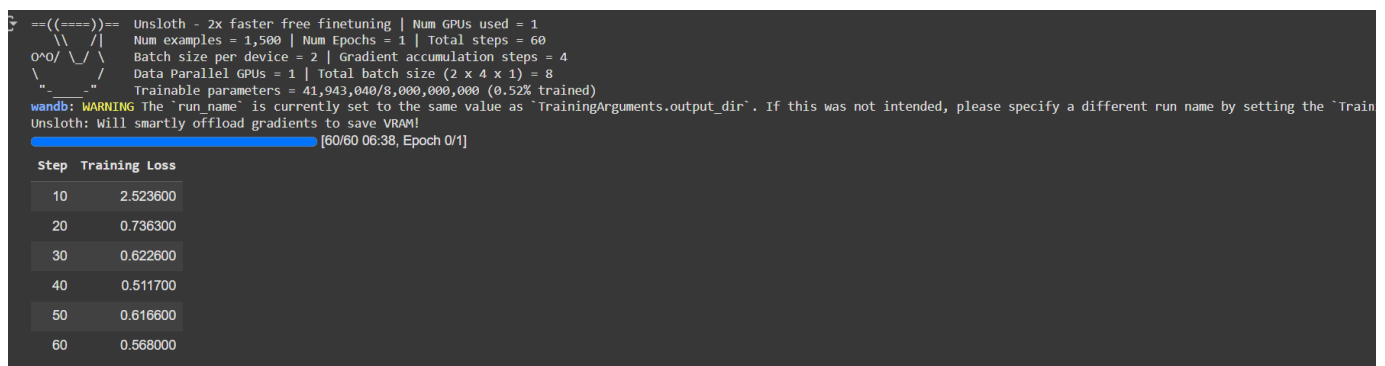
Using the Hugging Face Transformers ecosystem and **Parameter-Efficient Fine-Tuning (PEFT)** techniques allowed us to minimize the computational load while retaining performance. PEFT helps fine-tune only a small portion of the model's parameters, speeding up training and reducing memory usage.

4 Tokenization

The **DeepSeekTokenizer** is used to encode the inputs and decode the outputs, ensuring consistency with the model's pre-training token structure.

5 Loss Function

The **Cross Entropy Loss** function measures the error between the predicted quote and the actual quote, guiding the model to improve output accuracy over time.



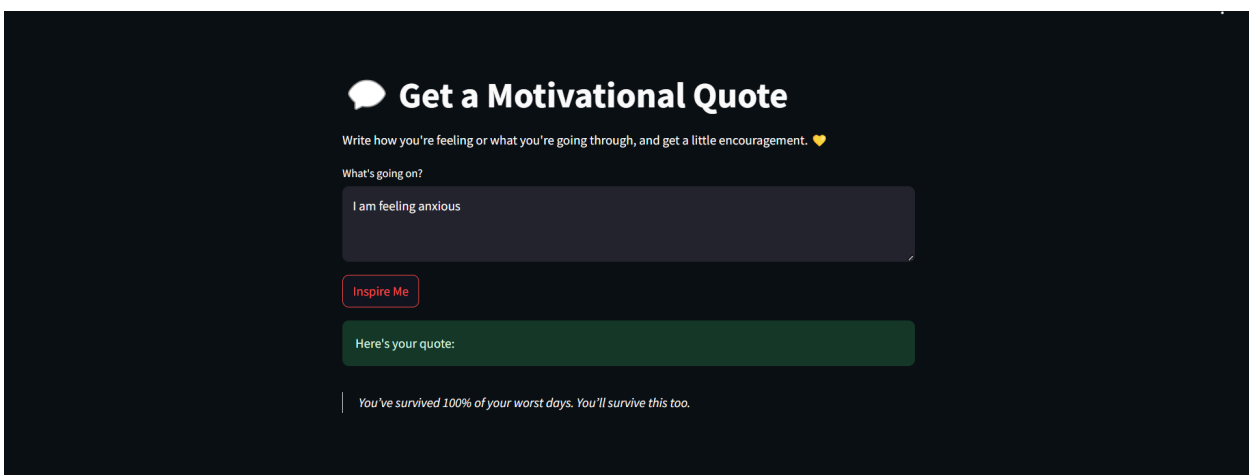
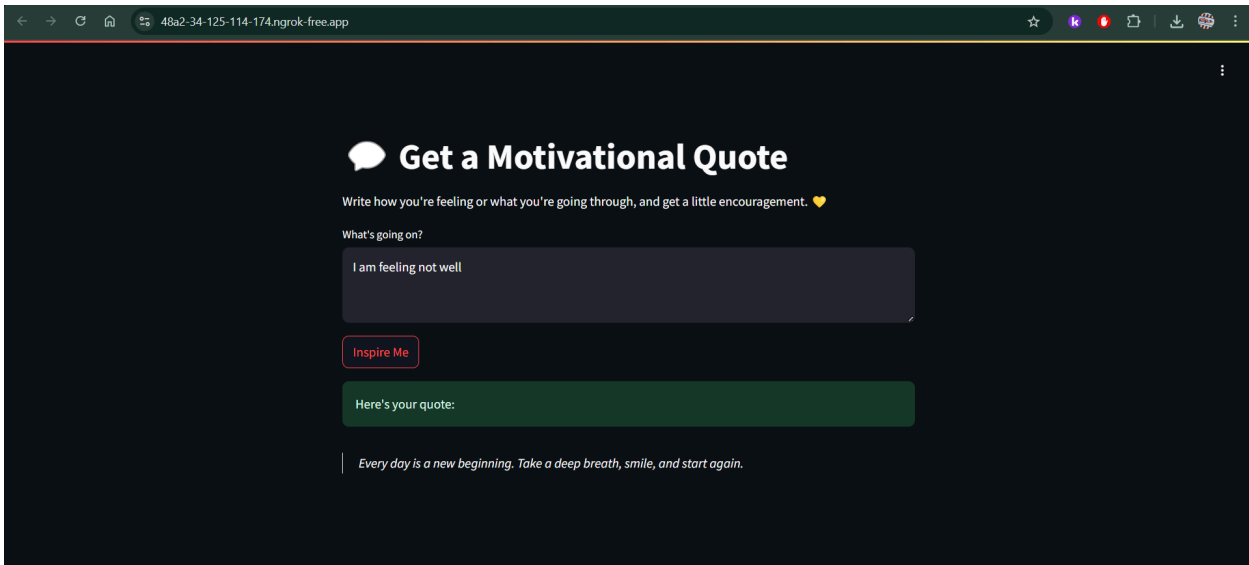
Web Application Interface

The model is integrated into a **Streamlit** application. This web app allows users to:

- Enter a topic or theme
- Generate a relevant quote using the AI model
- View output in real time with a clean and responsive UI

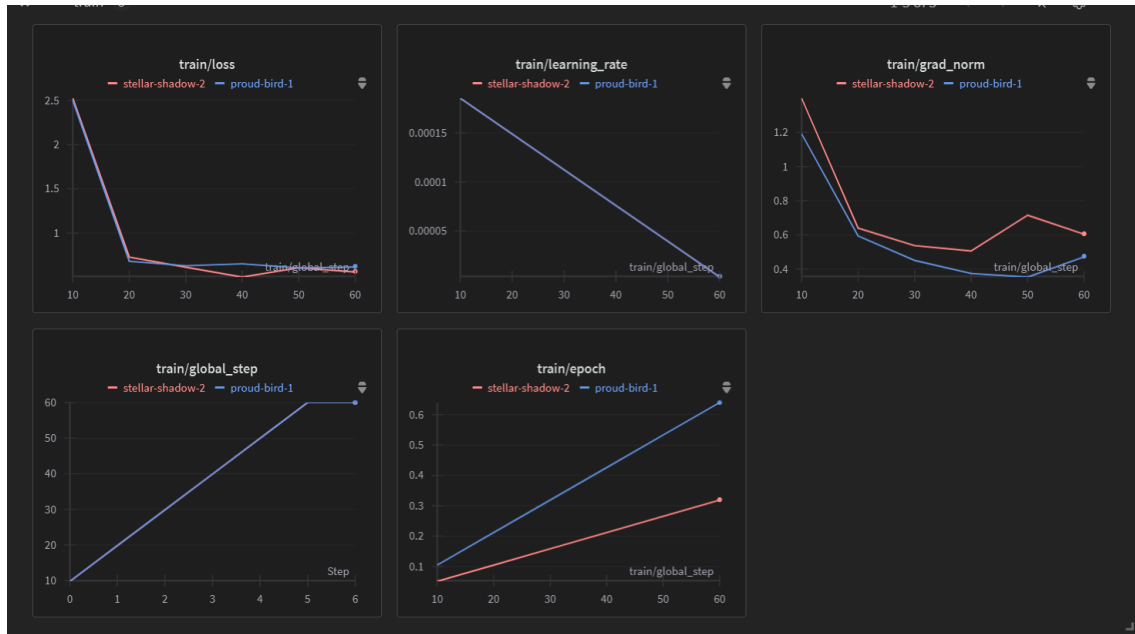
Features:

- Text input box for custom prompts
- Display area for generated quote
- Loading spinner while model processes
- Responsive layout suitable for mobile and desktop



Evaluation Metrics

The model's performance was evaluated both quantitatively and qualitatively.



Automatic Metrics

- **BLEU Score:** Measures n-gram overlap with reference quotes
- **ROUGE Score:** Evaluates recall of key phrases

```
ROUGE: {'rouge1': np.float64(0.1706549713731876), 'rouge2': np.float64(0.03948861680628071),  
BLEU: {'bleu': 0.03818548001849232, 'precisions': [0.14281991138910607, 0.04805914972273567]}
```

Conclusion

This project demonstrates how transfer learning and generative AI can be applied creatively using modern tools. By fine-tuning a pre-trained model and deploying it in a user-friendly interface with Streamlit, we deliver a real-world application that combines machine learning, NLP, and software engineering. It serves as a scalable foundation for more advanced AI-powered creativity tools.