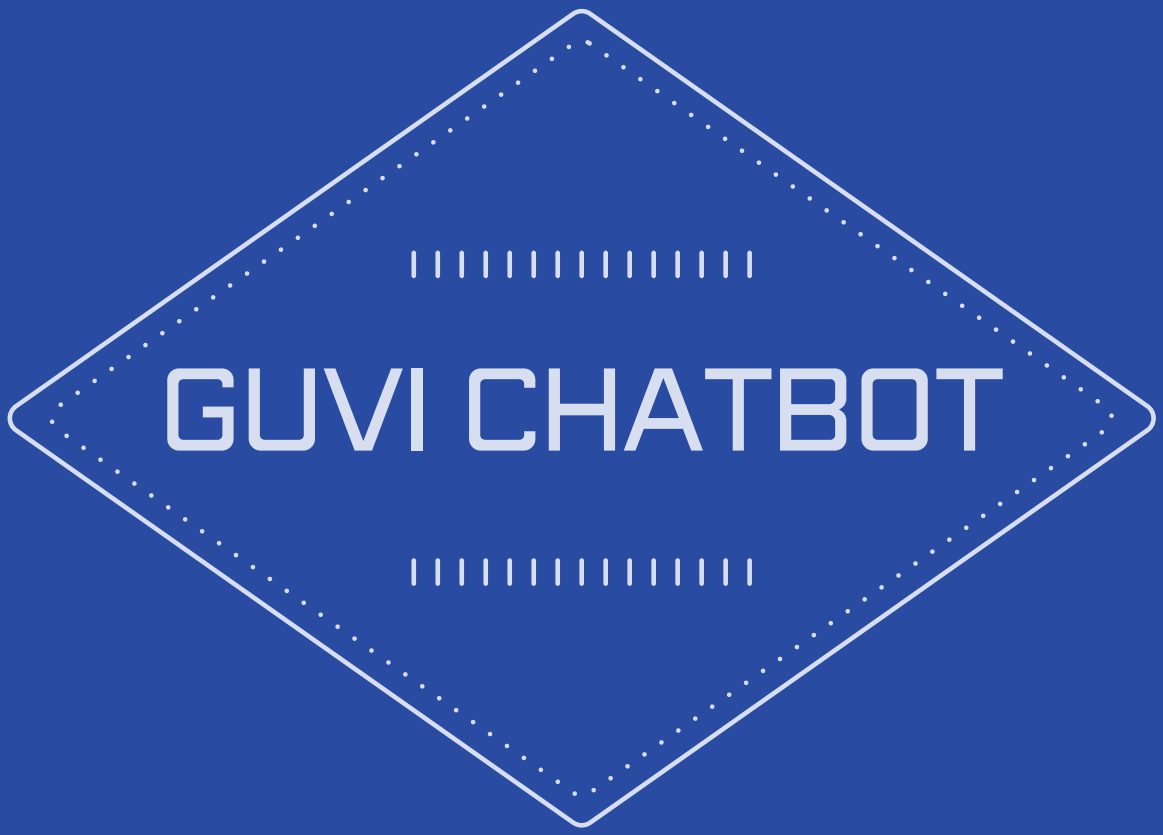


GPT-2 Fine-Tuned Model Deployment on Hugging Face Spaces

GOKUL M | GUVI-GPT | REPORT 2024

INTRODUCTION

PROBLEM STATEMENT :

The task is to deploy a fine-tuned GPT model, trained specifically on GUVI's company data, using Hugging Face services. Students are required to create a scalable and secure web application using Streamlit or Gradio, making the model accessible to users over the internet. The deployment should leverage Hugging Face spaces resources and any database to store the username and login time.

OBJECTIVES :

Deploy a pre-trained or fine-tuned GPT model using Hugging Face Spaces, making it accessible through a web application built with Streamlit or Gradio.

BUSINESS USE CASES :

Customer Support Automation :

Integrate the model with GUVI's customer support system to automate responses to frequently asked questions.

Content Generation for Marketing :

Generate marketing content like blog posts, social media updates, and email newsletters tailored to GUVI's audience.

Educational Assistance for Students :

Implement the model as a virtual teaching assistant within GUVI's educational platform.

Internal Knowledge Base :

Develop a tool for GUVI employees to access company-related information and resources quickly.

Training and Onboarding :

Assist in the training and onboarding process of new employees by providing instant access to training materials and answering common questions.

METHODOLOGY

Data Collection

Data specific to GUVI, including course descriptions, FAQs, and user interactions, was collected for training the model.

Data Preprocessing

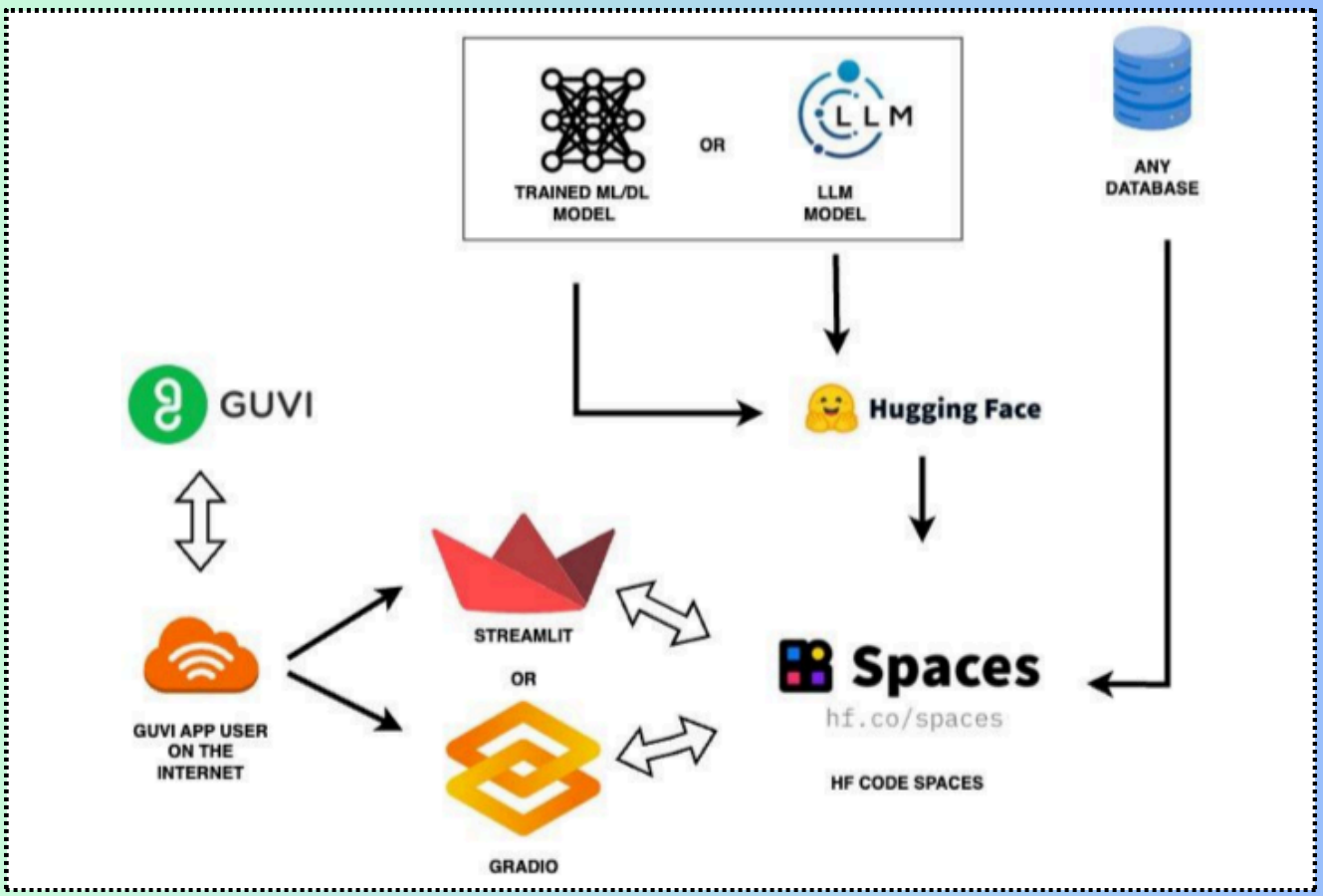
The collected data was cleaned and tokenized to be compatible with the GPT-2 model. Special tokens and padding were added where necessary.

Model Fine-Tuning

The GPT-2 model was fine-tuned using the preprocessed data on a high-performance computing environment. The training parameters were adjusted to optimize model performance.

Deployment

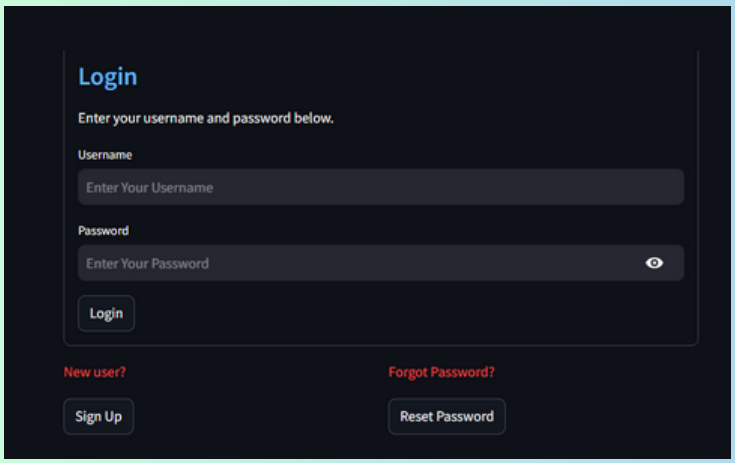
The fine-tuned model was deployed on Hugging Face Spaces, providing an accessible interface for interaction via a web-based application.

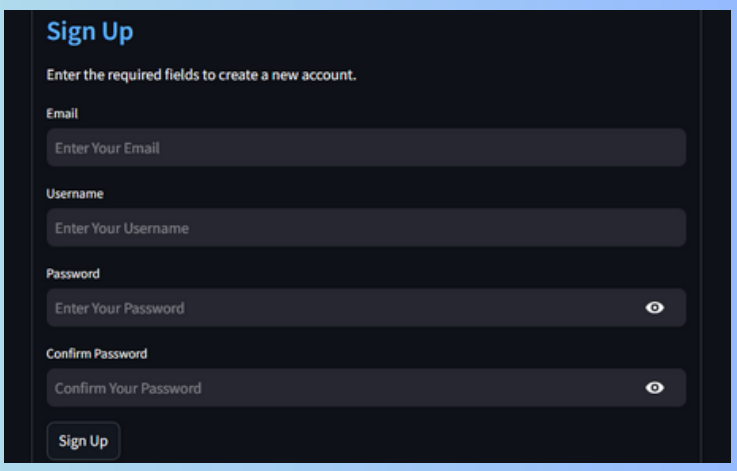


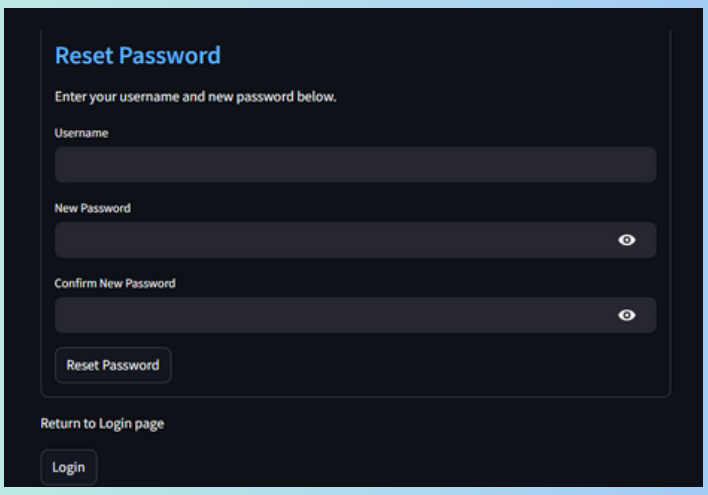
RESULTS

USER AUTHENTICATION :

The authentication system employs **bcrypt** for secure password hashing and integrates functionalities for user signup, login, and password reset. User passwords are securely hashed before storage in **TiDB**, ensuring data privacy and security. The system facilitates seamless user interaction through streamlined login processes and robust password management features, enhancing both security and user experience.







TIDB -CLOUD DATABASE

Result | Chart

id	username	password	email	registered_date	last_login
1	gk	$2b$12$jDSqLjynUoGOdGi9w0BdOOZNzNfdpcp1...	gokul[redacted]@gmail.com	2024-06-20 18:48:23	2024-06-28 12:48:22
30001	qqq	$2b$12$MfgqhdbolcSnp0.qghBSOOQZUCx5SVa...	gk123@gmail.com	2024-06-21 10:48:50	2024-06-21 13:49:38

STREAMLIT INTERFACE

The Streamlit interface features a user-friendly **LLM chatbot UI** and seamlessly guiding users through login, signup, and other functionalities using session state management. The chatbot leverages a fine-tuned language model, ensuring responsive and contextually accurate interactions. This setup prioritizes usability and security, providing a robust platform for personalized user experiences."

DISCLAIMER :

The data used by this model was sourced from various articles and blogs. The generated outputs are subject to the quality and relevance of the input data. Performance may vary and could improve with additional data and fine-tuning.

CONCLUSION :

In conclusion, the deployment of the fine-tuned GPT model using Hugging Face Spaces and Streamlit represents a significant advancement in AI application. This project demonstrates the integration of state-of-the-art natural language processing capabilities into a scalable web application environment.

Key achievements include ensuring robust performance, implementing data security measures such as password encryption for user privacy, and creating a user-friendly interface for seamless interaction.

Future enhancements will focus on optimizing model efficiency and responsiveness based on user feedback, ensuring continued innovation in AI-driven solutions.

REFERENCES :

- Hugging Face Spaces Documentation
- Streamlit documentation
- TiDB Cloud Documentation

CONTACT

gokulgokul6847@gmail.com

Portfolio/Gokul

Github/Gokul

Linkedin/Gokul