

To, IITD-AIA Foundation of Smart Manufacturing

Subject: Weekly Progress Report for Week 8

Dear sir, Following is the required progress report to the best of my knowledge considering relevant topics to be covered.

What's happened this week:

- SQUAD Data preparation of site for more questions
- Training the data for accuracy
- Deployment of model

My understanding of INTP23-ML-01: CHATBOT for FSM

Scope: In the Chatbot for Site, I will try to scrape out all important info from website and according to the questions of users, it will provide best possible and accurate answers and it should follow the upcoming check points analyze user queries and understand user messages, provide answers to the messages of the user accurately, provide all information about the updated activities and save their time from searching on website and google search the information, provide answers accurately as much as possible, i.e., instead of providing irrelevant information, its' better to admit that bot is not able to understand the question.

Solution: In this, concept of deep learning either Pytorch or TensorFlow will be used along with some other libraries like Beautiful Soup, and one for scrapping the data on the website or using NLP and deep learning technique of transformer for question answering would increase efficiency of Chatbot. Using data of SQUAD, BERT model would be my preference.

Approach: Rule-based approach will be used for most common questions like hi, hello, bye. And web-scraping the data in order to give website related answers. For, questions/queries of users; we will use transformer (BERT model) and SQUAD dataset for pre-training and web-scraped data of IAFSM site for fine-tuning in SQUAD format.

## **WEEKLY PROGRESS:**

### **Day 44 (July 24- Monday)**

I'm working on for increasing the accuracy of my code. For that training the code again and again. Also, in order to deploy the model. I'm learning Gradio. How to provide input, output and how to deploy on hugging spaces etc. Also, I reduced the size of context removed repeated lines and unexplained terms and paragraphs from context in order to increase the accuracy. I tried different-different techniques like fine-tuning, training and reducing context in order to increase the accuracy. I used stride method i.e. stride of 50, max- length of 512 and then I'm asking questions. I got some problem in stride function instead from starting from last 50 words words from end it was taking next 50 words from previous start corrected that.

### **Day 45 (July 25-Tuesday)**

I'm working on for increasing the accuracy of my code. Also, today I started with GUI and applied some Gradio part. We had meeting with mentor and there we got to know about deployment, so now I will start working on deployment part I tried different-different techniques like fine-tuning, training and reducing context in order to increase the accuracy. I used stride method i.e., *stride of 100, max-length of 512 and then I'm asking questions. Now I'm sending inputs to the model using for loop where each chunk of size 512 along with question is sent to tokenizer and model.*

### **Day 46 (July 26- Wednesday)**

I learnt about state in Gradio, it is like chat remain in history and remain visible while we are taking. So now I'm working on that also I will start working on deployment part. I'm done with basic GUI Application, now will work on state maintenance and also I have made final code with proper commenting and making divisions of training and loading chatbot model with different accuracy w.r.t context.

### **Day 47 (July 27- Thursday)**

I applied chat interface using Gradio. Where we can ask a question and asked question and response will remain in history until refreshed. I used local variable so that all the conversation will be visible independently to each human with respect to their asked question i.e., Question ked by person1 will not be visible to person2 but all the question asked by person1 will remain on his/her screen until refreshed. However, if global is used then all question asked at a time will be visible whether it is asked by person1, person2. I'm done with chat interface with the help of Gradio now I'm left with deployment.

### **Day 48 (July 28 -Friday)**

I applied chat interface using Gradio. Where we can ask a question and asked question and response will remain in history until refreshed. Also, done with the deployment of model on hugging -face Spaces. I'm done with chat interface, Deployment.

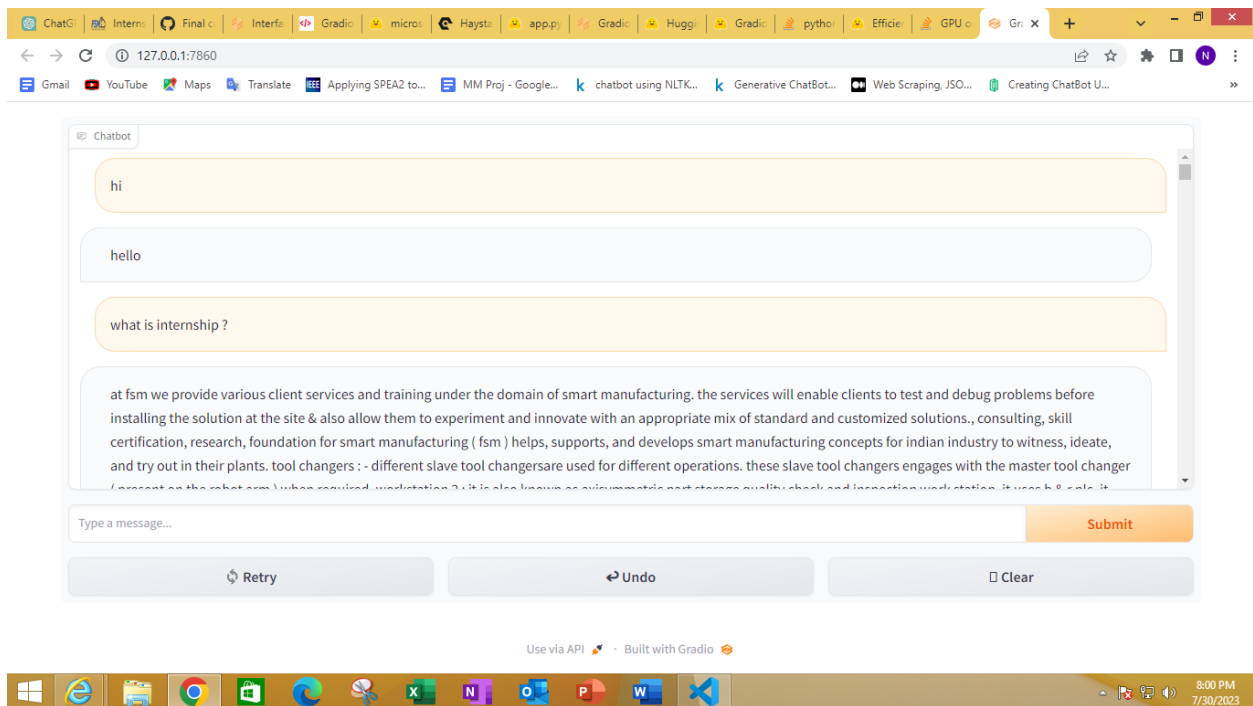
## Day 49 (July 29- Saturday)

I am training the model on more questions by providing questions on IAFSM site data, and preparing questions on the IAFSM using hay stack annotation tool in order to increase the accuracy. I'm training of the model.

## Day 50 (July 30 -Sunday)

Today, I am training the model again and again and now loss function value is 0.048. Also, we had meeting today where I got to know the solution of my answer which was showing some extra data 'output relevance' is some parameter. Now, I will try that parameter in order to fix my error. I'm doing training of the model.

## Screenshot of chat interface:



## Deployment Link:

<https://huggingface.co/spaces/Nickitaa/gradio-chatbot>

(Updated trained model is not uploaded, it is link based on model which was fine-tuned till yesterday.)