



# PROJECT PRESENTATION

iNLP Project Team

Hardik Mittal  
Jainit Bafna  
Suyash Sethia



# PROBLEM STATEMENT

This project involves fine-tuning a generative model to mimic a person's way of speaking by training it on their chat data. The objective is to create a system that can generate text that closely resembles the person's style and patterns of communication.

This approach has potential applications in various fields such as chatbots, virtual assistants, and personalized content creation. The success of the project will depend on the quality and quantity of the training data and the effectiveness of the chosen generative model for the task.





# DATASET USED

<https://github.com/YisiSang/TVSHOWGUESS>

The dataset consists of 5 popular TV series, including *Friends*, *The Big Bang Theory*(TBBT), *The Office*, *Frasier*, and *Gilmore Girls*. For each series, the dataset provides character dialogues and backgrounds descriptions. The characters' dialogues start with the characters' names.

For experiment we have just trained the model on “The Big Bang Theory” specifically for Sheldon, due to his unique personality which helped us in checking if the model was learning his unique character traits





# MODEL USED

<https://huggingface.co/gpt2-medium>

GPT-2 Medium is the 355M parameter version of GPT-2, a transformer-based language model created and released by OpenAI. The model is a pretrained model on English language using a causal language modeling objective.

This model was the perfect choice since we wanted a causal model which had as much parameters as we could possibly fit on Ada.





# FEATURES

- **Transformer:** Fine-tuned GPT2-Medium
- **Froze the first six layers** of the model making only the other six opened to training.
- **Optimizer** :- AdamW
- **Scheduler** for the optimizer
- Integrated **W&B** for better visualization of the model and the hyperparameters.
- Whatever the input is, the model is not breaking



# RESULTS

```
def clean_output(text):  
    """  
    Clean the output text  
    """  
    text = text.replace("<START>", "Question :")  
    text = text.replace("<bot>", "\nSheldon :")  
    text = text.replace("<End>", "")  
    text = text.replace("<pad>", " ")  
    text = text.replace("<npad>", "\n")  
    return text
```

Question : Hello, how are you?  
Sheldon : I am fine.  
Question : What is your name?  
Sheldon : I am Dr. Sheldon Cooper  
Question : Is your name Sheldon? Yes or No?  
Sheldon : Yes, my name is Sheldon.  
Question : Who are your friends?  
Sheldon : My friends are my closest confidantes, my soul mates, my soul mates.  
Question : Where do you work at?  
Sheldon : No, I work at home. I don't have a lot of company.  
Question : What inspired you to pursue a career in physics?  
Sheldon : My innate intelligence and passion for the subject matter led me to pursue a career in physics.  
Question : What is your girlfriend's name?  
Sheldon : My girlfriend's name is Amy Farrah Fowler.

# RESULTS

Sheldon : I'm sorry, I'm just distracted by something that happened over at Penny's.  
You: What happened at Penny's?

Sheldon : I don't know. I just want to be alone.  
You: You sure?

Sheldon : Yes.  
You: Okay I will leave you alone then.

Sheldon : I'm not leaving you alone.  
You: But you wanted to be alone

Sheldon : I did. And I know that moving forward, I want to make sure that everything is on the up and up. I also wanted to apologize for the last few months of your life. I know that you sob like a little girl. And I am one, too.  
You: It is okay, Sheldon

Sheldon : I'm sorry, I'm just excited.  
You: It is fine

Sheldon : I know. I just want you to know that I care.  
You: Thanks! It means a lot.

Sheldon : I know. I'm glad you're happy.  
You: Do you think that MIT is an engineering school?

Sheldon : No.  
You: What do you think about Caltech then?

Sheldon : I think it's a great idea.

# Failed Attempts

- Attempted to use multiple utterances to provide context for the generative model but found that it led to unexpected behavior and the model broke.
- Considered using a separate encoder to handle the context of the chat, but did not have enough time to implement this in the model.
- Tried using GPT-Neo, a variant of the GPT architecture, but found that it performed worse than GPT2 and GPT-Medium models. This may have been due to limitations in the model's size, as the larger models could not be loaded on Ada.
- Encountered challenges integrating "Training Arguments" and "Trainer" into the model due to time constraints.
- The project was likely limited by time constraints, indicating the need for more time and resources to fully explore and implement different approaches to fine-tuning the generative model.





# Literature Survey

- TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents
- Neural Response Generation for Customer Service based on Personality Traits





# THANK YOU

Any questions?