**First Slide**

1. Hello everyone and thank you very much for making it to the session!
2. For those of you who do not know me – my name is Ajay Taneja and I work as Senior Data Engineer in the Data Engineering Team
3. And topic of my session today is – I will be speaking with you about the inner working of ChatGPT – that is: I will be talking to you about the model behind ChatGPT, the steps that were involved in the training process of the model, the objective function that’s getting solved in each of the training steps.
4. We will be looking at the mathematics in a highly qualitative sense without going into any rigorous mathematics!

**Second Slide:**

1. With this in mind, I have attempted to organize the talk as follows.
2. Firstly, we will go into the overview of ChatGPT – when I say overview here, I mean we’ll what makes up the ChatGPT, what are the basic ingredients of the model behind ChatGPT and an overview of the 3 step training steps of the model
3. We will then dive into each of the training step a bit more, go into seeing the objective function involved in each of the training step and the problem each of the training step is attempting to solve.

**Third Slide:**

1. I need not tell anyone what ChatGPT is – the whole world knows that. It takes the user prompt and gives you a response. The response generated by that ChatBot are most factual, human like, devoid of any toxic content – it is the best we’ve seen from any ChatBot until now.
2. How is ChatGPT Built? ChatGPT is built on top of GPT and Reinforcement Learning
3. GPT is a Language Model which is built on top of GPT And the entire paradigm of Reinforcement Learning
4. GPT is a language model built on top of the Transformer Neural Network Architecture and
5. We’ve just named some very important concept: Language Model, GPT, Transformer Neural Network Architecture and Reinforcement Learning. Let’s examine these crucial concepts in some details.

**Fourth Slide:**

1. We said GPT is a Language Model
2. What is a Language Model
3. Language Model has an understanding of Language.
4. Has an understanding of context?
5. Understand the probability distribution of a sequence of words
6. Because when humans speak, there is probability distribution of sequence of words
7. And the once we have a language model we can fine tune the language model for a variety of tasks. That could be Question and Answering as ChatGPT does, document classification – engineering simulation report or design report or a data pipeline design report or whatever.

**Fifth slide:**

1. I said GPT is built on top of the Transformer Neural Network Architecture
2. Let us see about the Transformer Neural Network Architecture because it’s the key behind Language Models
3. Transformer Neural Network Architecture comprises of 2 components
4. Encoder and the Decoder