Brief on Generative Large Language Models

Instructor

Dipanjan Sarkar

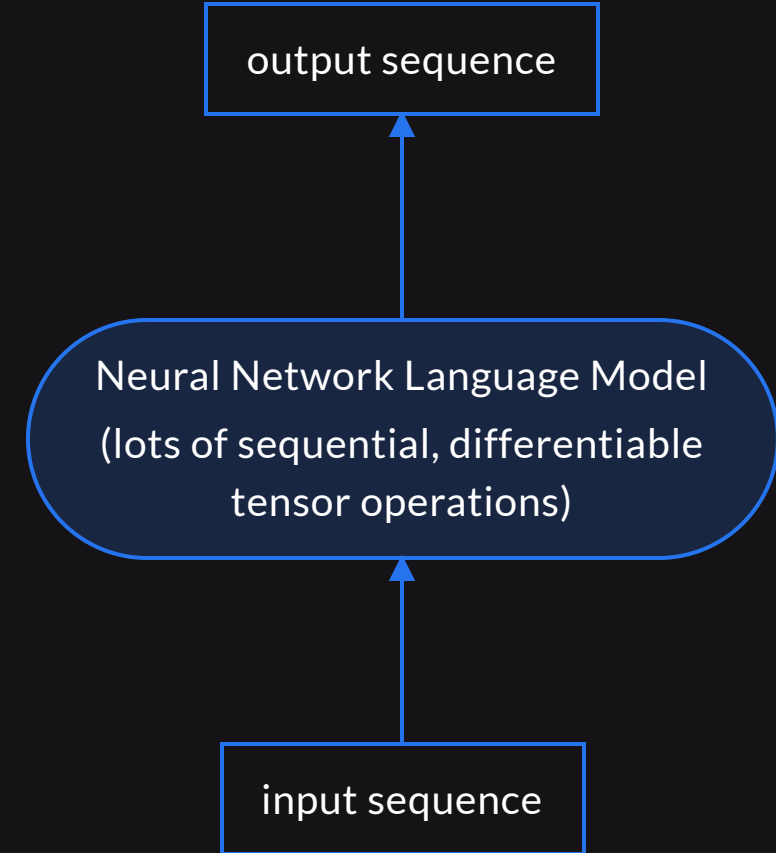Head of Community & Principal AI Scientist at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author

Generative AI for NLP in a Nutshell

A versatile set of models that can be used to process and generate sequential data

output sequence

↑

Neural Network Language Model
(lots of sequential, differentiable tensor operations)

↑

input sequence

Why Generative Language Models?

- Base or foundation Large Language Models learn good representations and patterns on text data e.g. GPT-3 / 3.5 / 4o

- Prompt or instruction fine-tuned models learn how to solve a variety of problems using natural language instructions

- Therefore, you do not need specific models for specific tasks; you can use the same model for different tasks, which is the path toward Artificial General Intelligence.

- The promise: one single model to solve many NLP tasks

Analytics
Vidhya
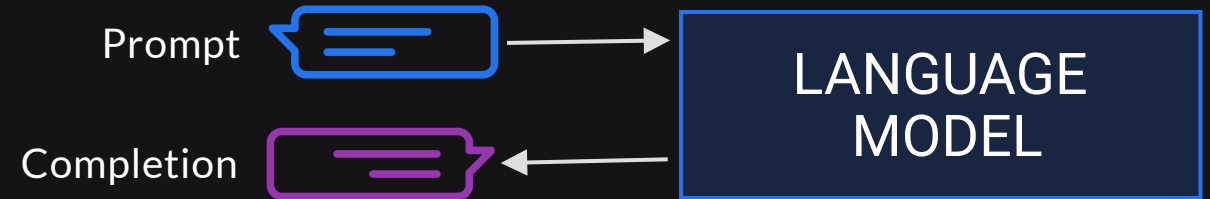
Sequential Transfer Learning for NLP

Sequential transfer learning has led to the biggest improvements so far in the field of NLP
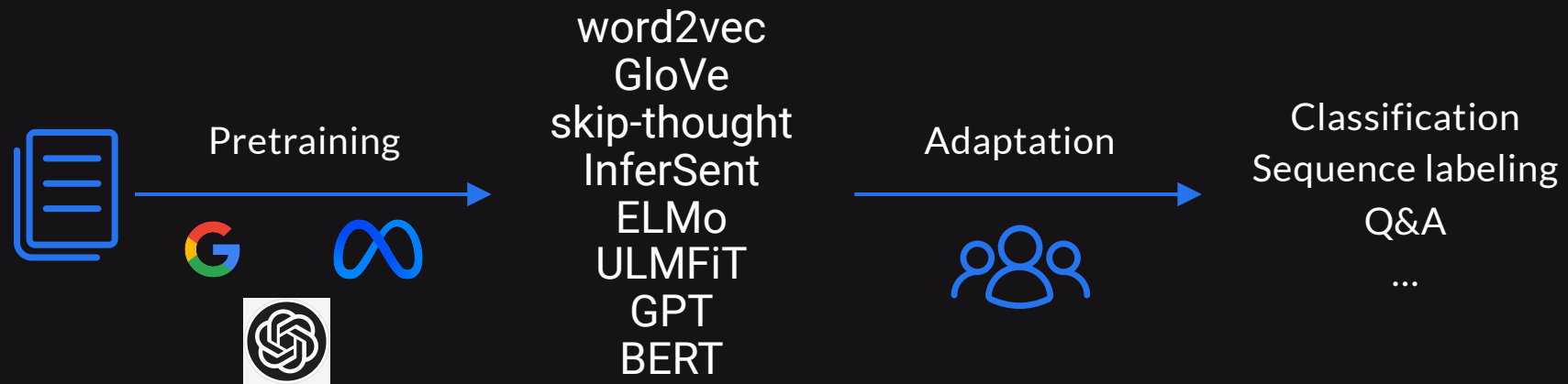
Base vs. Fine-tuned Generative Models

Base Gen AI Model

Predicts next word, based on previous words

Once upon a time, there was a unicorn
that lived in a magical forest with
all her unicorn friends

What is the capital of France?
What is France's largest city?
What is France's population?
What is the currency of France?

Fine-tuned Gen AI Model
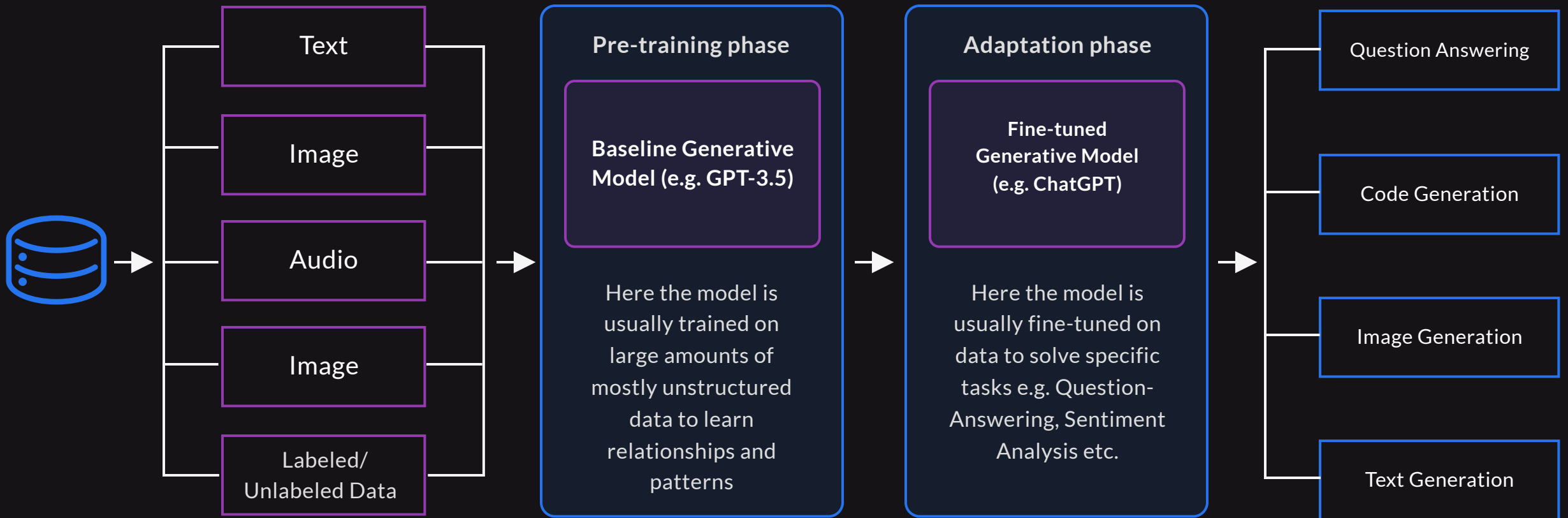
Tries to follow instructions

Fine-tune on prompts (instructions) and
good attempts at following those
instructions (response)
Improve response quality using human
preference alignment methods (RLHF, DPO)

Helpful, Honest, Harmless

What is the capital of France?
The capital of France is Paris.

Analytics
Vidhya

ChatGPT is a fine-tuned LLM

How to use LLMs like ChatGPT?

- The inputs to LLMs are referred to as "prompts"

- Creating a prompt is essentially how you "program" a Large Language Model.

- This is usually done by providing instructions or some examples of how to successfully complete a task

- Systematic ways of creating prompts are known as prompt engineering patterns

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

Andrew: 11 am - 3 pm
Joanne: 12 pm - 2 pm, 3:30 pm - 5 pm
Hannah: 12 pm - 12:30 pm, 4 pm - 6 pm

Common availability for a 30-minute meeting: 12 pm - 12:30 pm

Analytics
Vidhya

Thank You

Analytics Vidhya