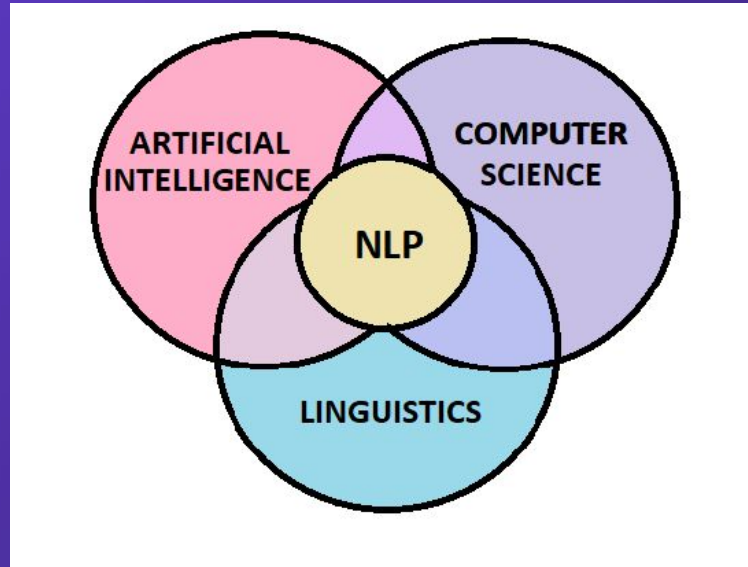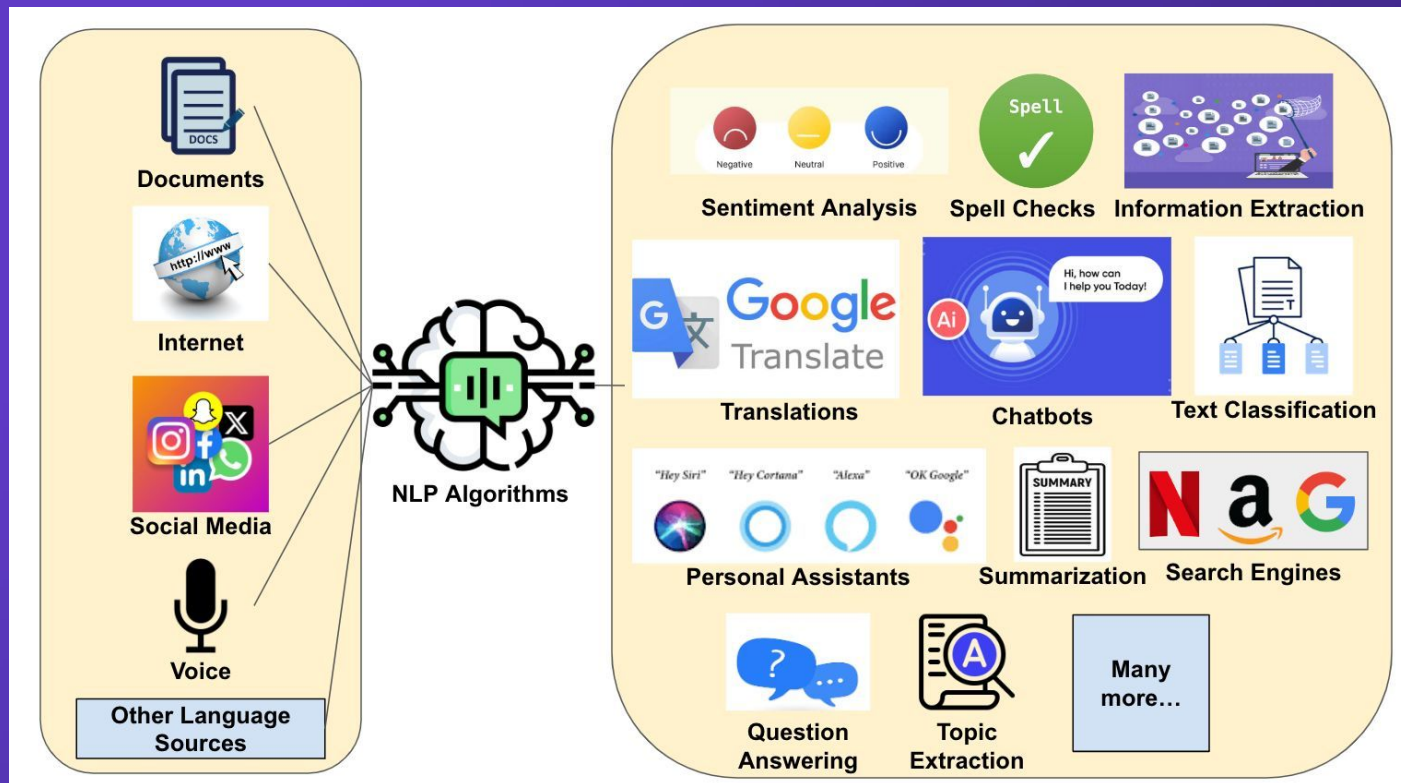# Building your First ASR / NLP Model

Intro to NLP

What is NLP?

Natural language processing (NLP) is a field at the intersection of computer science, artificial intelligence, and linguistics.
It concerns building systems that can process and understand human language.
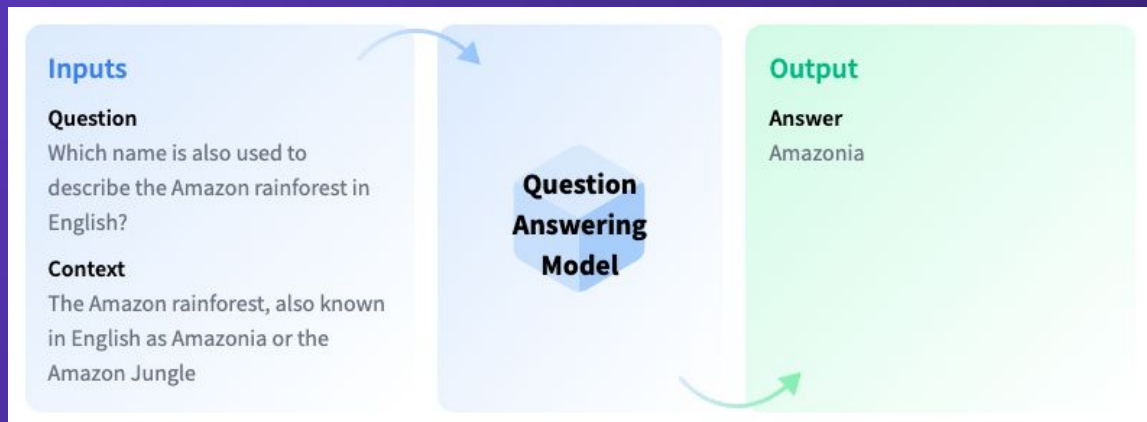
Motivation behind NLP

Introduction to Question Answering

- Question Answering models can retrieve the answer to a question from a given text

- Useful for searching for an answer in a document

Inputs

Question
Which name is also used to describe the Amazon rainforest in English?

Context
The Amazon rainforest, also known in English as Amazonia or the Amazon Jungle

Question Answering Model

Output

Answer
Amazonia

TIL-AI
TODAY I LEARNED AI

What are Pre-trained Models

- Pre-trained models are massive deep learning models trained on enormous text datasets

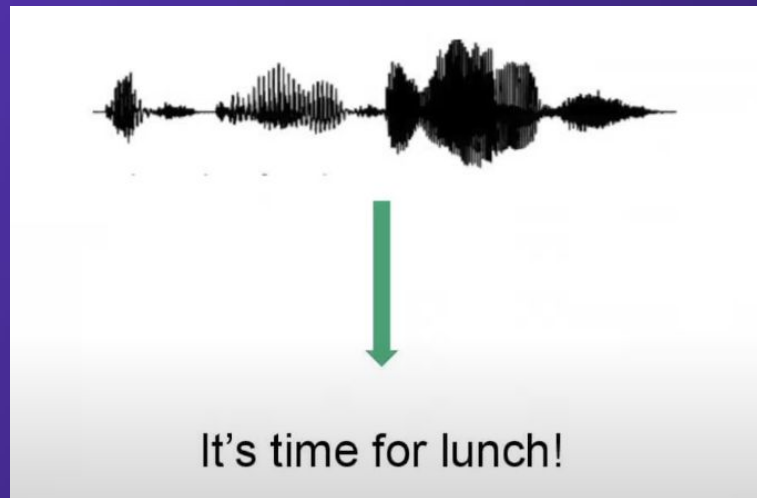- They have learned to represent language effectively
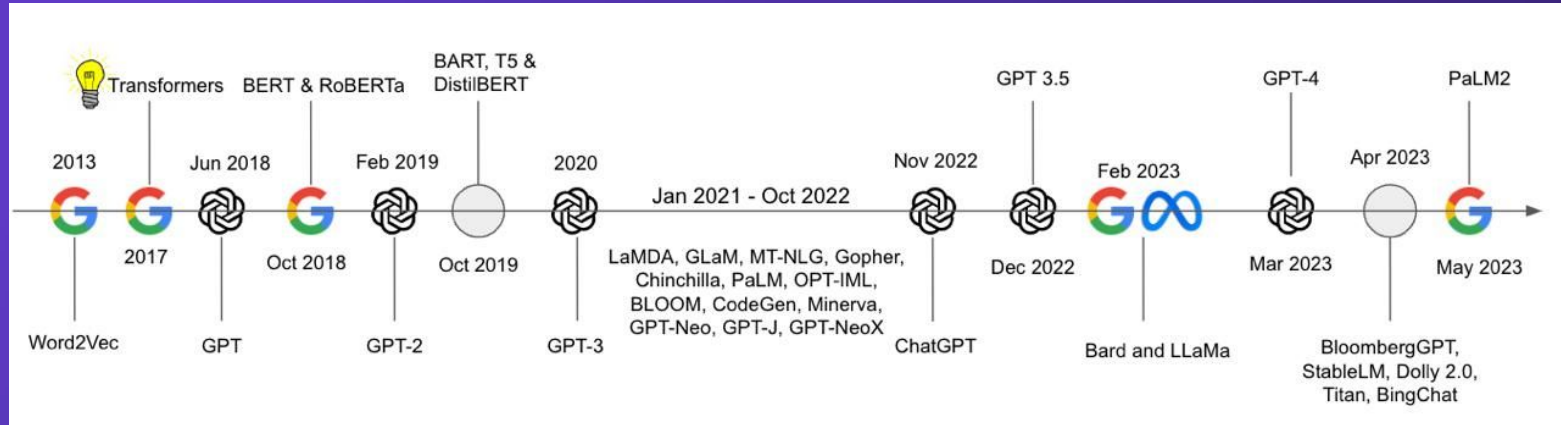
Hands-on activity

Intro to ASR

What is Automatic Speech Recognition (ASR)?

- The process of converting spoken language into written text

- Powered by sophisticated computer algorithms

- Also known as speech-to-text or computer speech recognition

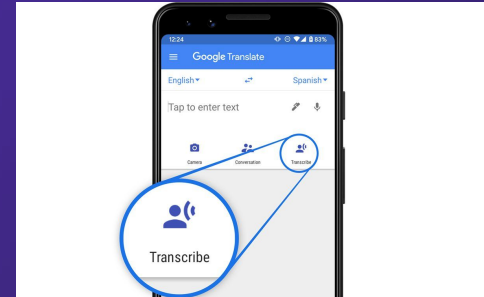It's time for lunch!

Examples of ASR

Virtual Assistants

Vehicle Speech Recognition


Accessibility for Hearing Impaired

Language Translation

How we recognize speech?

Time vs Frequency Domain

- ASR systems rely heavily on transforming speech signals from time domain to frequency domain to reveal the underlying frequencies that make up the speech

- Time domain: Speech is represented as a waveform where the amplitude (intensity) of the signal is plotted over time

- Discrete Fourier Transform (DFT) is used to decompose a time-domain signal into individual frequencies and amplitudes

- The output of the DFT tells us how much energy is present at each frequency in the original speech signal
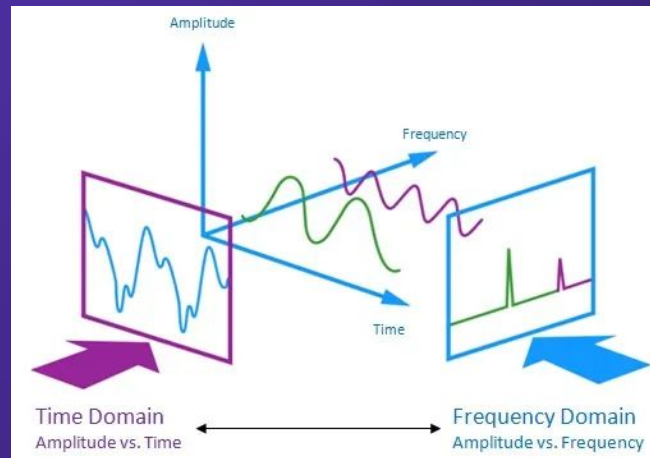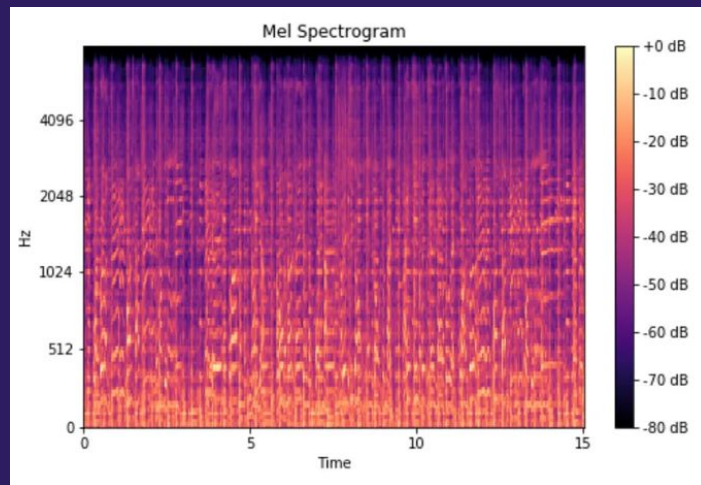
Spectrograms

- Visual representation of the spectrum of frequencies in a sound as they vary with time

- Shows how different frequencies appear, disappear, or change intensity over time in an audio signal
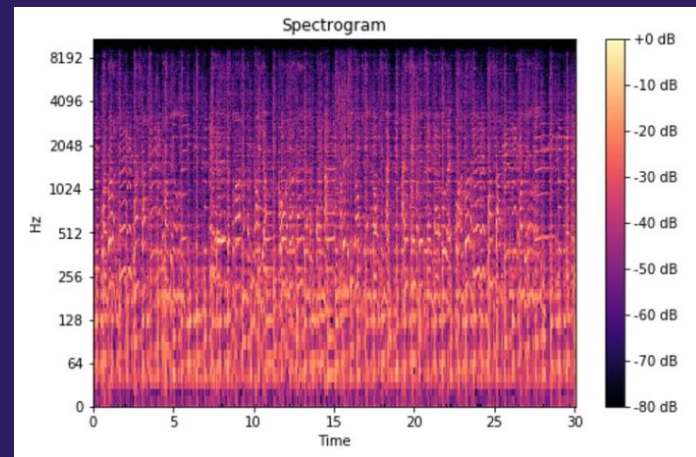
Mel-Spectrograms

- Type of spectrogram where the frequency scale is converted to the Mel scale

- Mel scale more closely approximates human auditory system's response than the linear frequency scale; making it more effective for audio-related tasks in human speech and music

Mel Spectrogram

Hands-on activity