



JOIN LIVE

# NLP & COMPUTER VISION USING TRANSFORMERS

LIVE ONLINE SESSION

ON 30-JUL-2022 @ 7.00 - 08.30 PM IST (SATURDAY)  
@ 9.30 - 11.00 AM EST  
@ 2.30 - 04.00 PM LONDON

JOIN FREE

[HTTPS://FORMS.GLE/CLXRQUIWUDRWAZGJ6](https://forms.gle/CLXRQUIWUDRWAZGJ6)

BY PRASANNA VENKATESH J

AI & CLOUD EVANGELIST

# AGENDA

---

- Introduction to NLP
- Journey towards Transformers
- Live Demo of Hugging faces Transformers
- Introduction to Computer Vision using Transformers
- Live Demo of Computer Vision based classification using Transformers

# INTRODUCTION TO NLP

---

- Natural Language Processing
  - Translation Services
    - Google Translate, etc
  - Personal Assistants
    - Alexa, Google Assistant etc.
  - Chat Bots
    - Siri Chatbots
  - Many More.....

# UNDERSTANDING TECHNOLOGY BEGIND NLP

---

- Computers cannot understand Text?
- Convert Text to Numerical form to enable computers to understand.
  - Embedding.
- Word Embedding
  - Word2Vec
  - Convert list of words to embedding vector.

# WORD2VEC

Token index in  
the vocabulary

39 1592 10 2548 5  
↑ ↑ ↑ ↑ ↑  
**I** saw a cat .

Embedding  
dimension

39



Vocabulary  
size

## PROBLEM - I - WITH WORD EMBEDDING - MISSING CONTEXT

---

- He didn't receive **fair** treatment



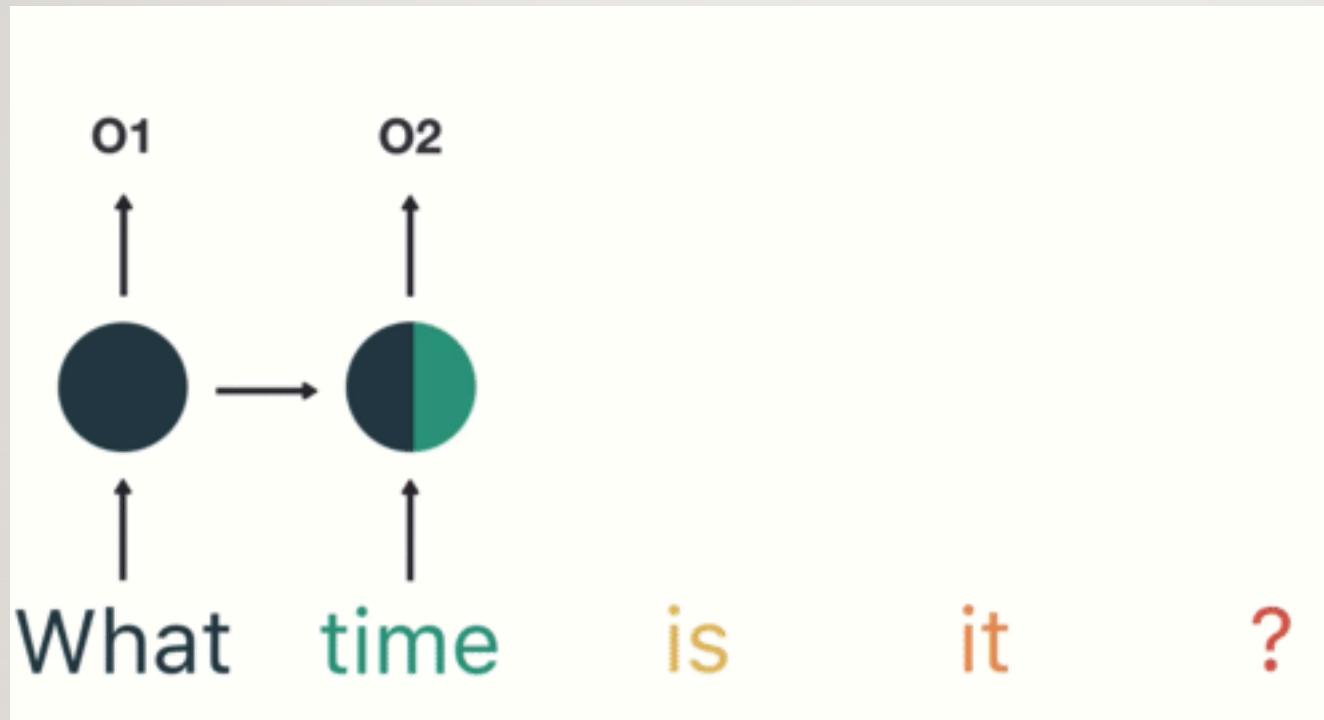
Word2Vec would refer the same Vec for both statement.

Context Information is missing.

- Fun **fair** in chennai city this season

# EXISTING NEURAL NETWORK ARCHITECTURE FOR NLP

---

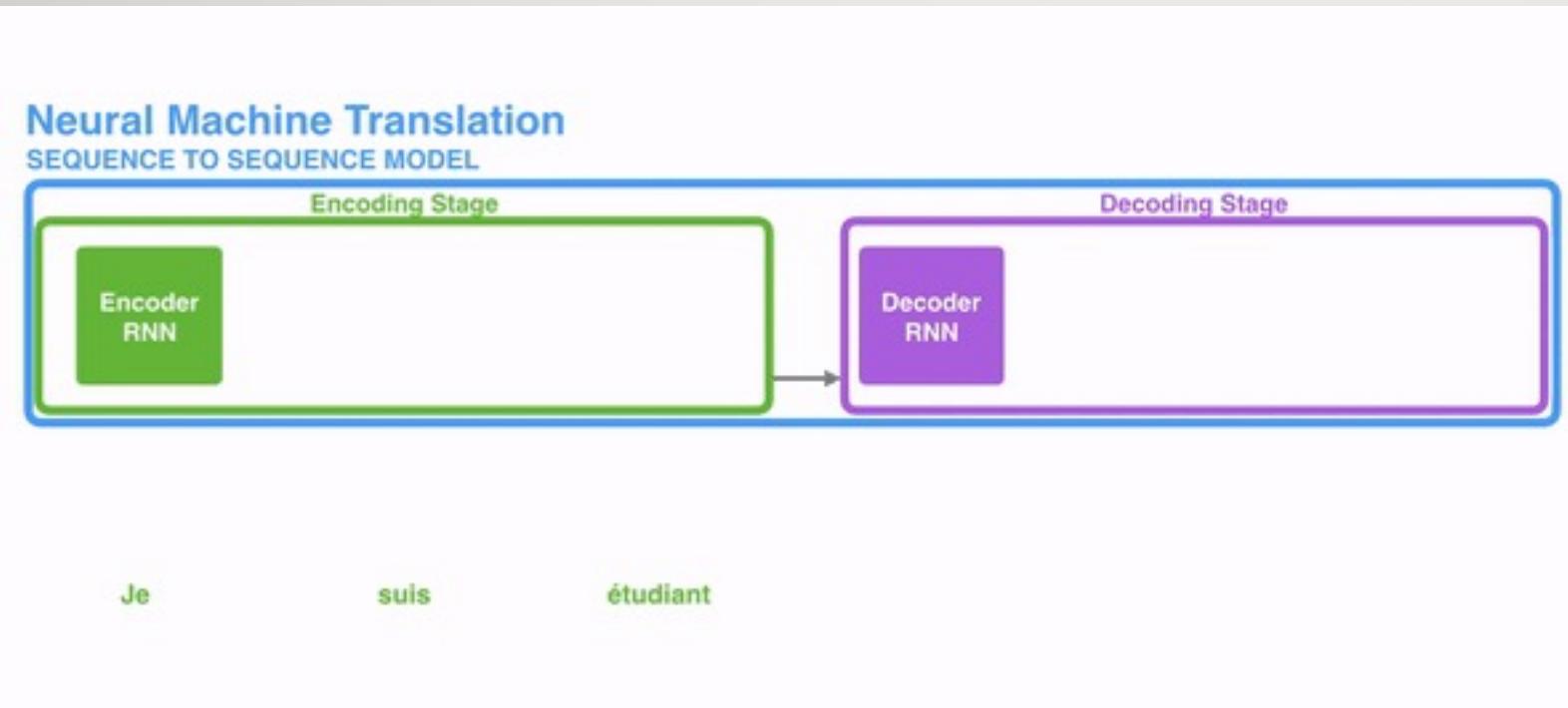


Understanding Recurrent Neural Network

Each word is processing in sequence timestep

# UNDERSTANDING RECURRENT NEURAL NETWORK

---



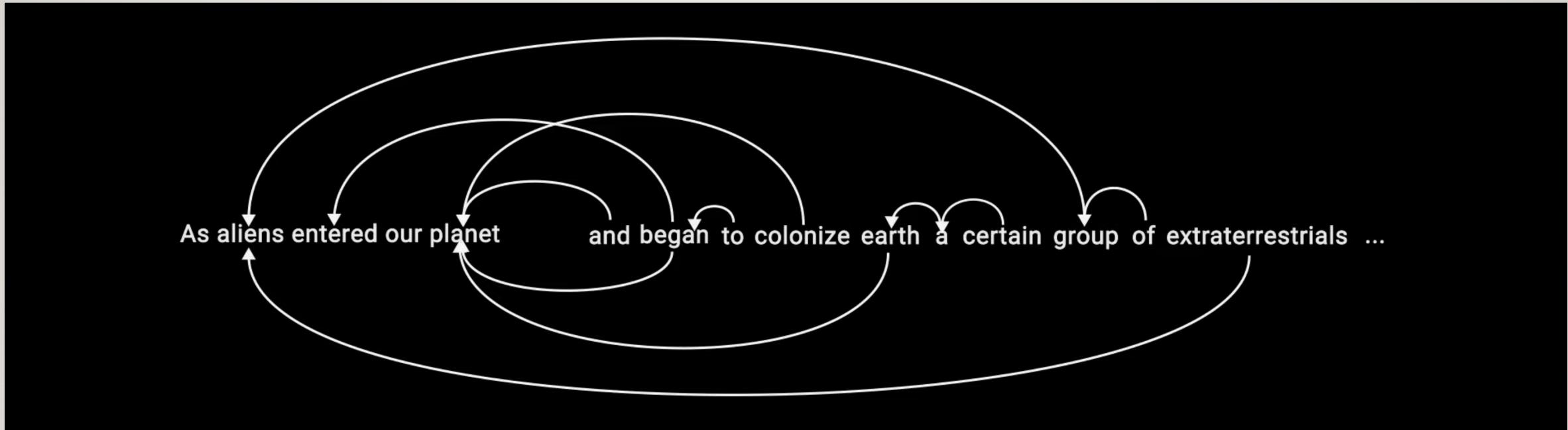
Each Time step learning is given to next time step to enable remembering what is the previous word.

To Improve Efficiency

- LSTM
- GRU etc where developed.

# HOW MUCH CAN BE LEARNED??

---



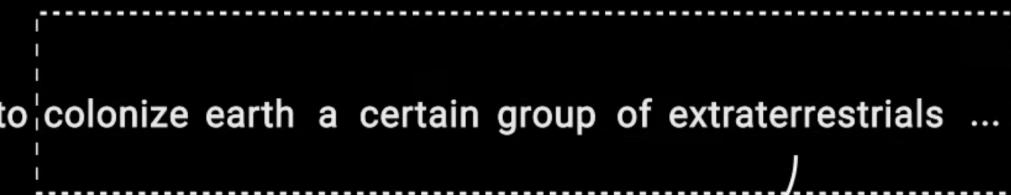
## PROBLEM-2 - RNN AND OTHER VARIENCE LIKE LSTM GRU HAS LIMITATION

---

Recurrent Neural Networks has a short reference window

As aliens entered our planet

and began to colonize earth a certain group of extraterrestrials ...



# GOOGLE RESEARCH PAPER IN 2017 – INTROCUED THE CONCEPT OF TRANSFORMERS

- \*Attention is all you need\*
- Research paper on google experiment on
- Transformers

## Attention Is All You Need

Ashish Vaswani\*  
Google Brain  
avaswani@google.com

Noam Shazeer\*  
Google Brain  
noam@google.com

Niki Parmar\*  
Google Research  
nikip@google.com

Jakob Uszkoreit\*  
Google Research  
usz@google.com

Llion Jones\*  
Google Research  
llion@google.com

Aidan N. Gomez\* †  
University of Toronto  
aidan@cs.toronto.edu

Lukasz Kaiser\*  
Google Brain  
lukaszkaiser@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

### Abstract

The dominant sequence transduction models are based on complex recurrent or

# SOLVING PROBLEM-I - CONTEXT INFORMATION

---

- Encoding with Context and Positional Information

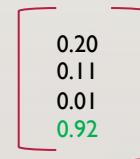
He didn't receive **fair** treatment

**fair**



Fun **fair** in chennai city this season

**fair**



# SOLVING PROBLEM-2- INFINITE REFERENCING USING ATTENTION MECHANISM

---

Attention Mechanism has an infinite reference window

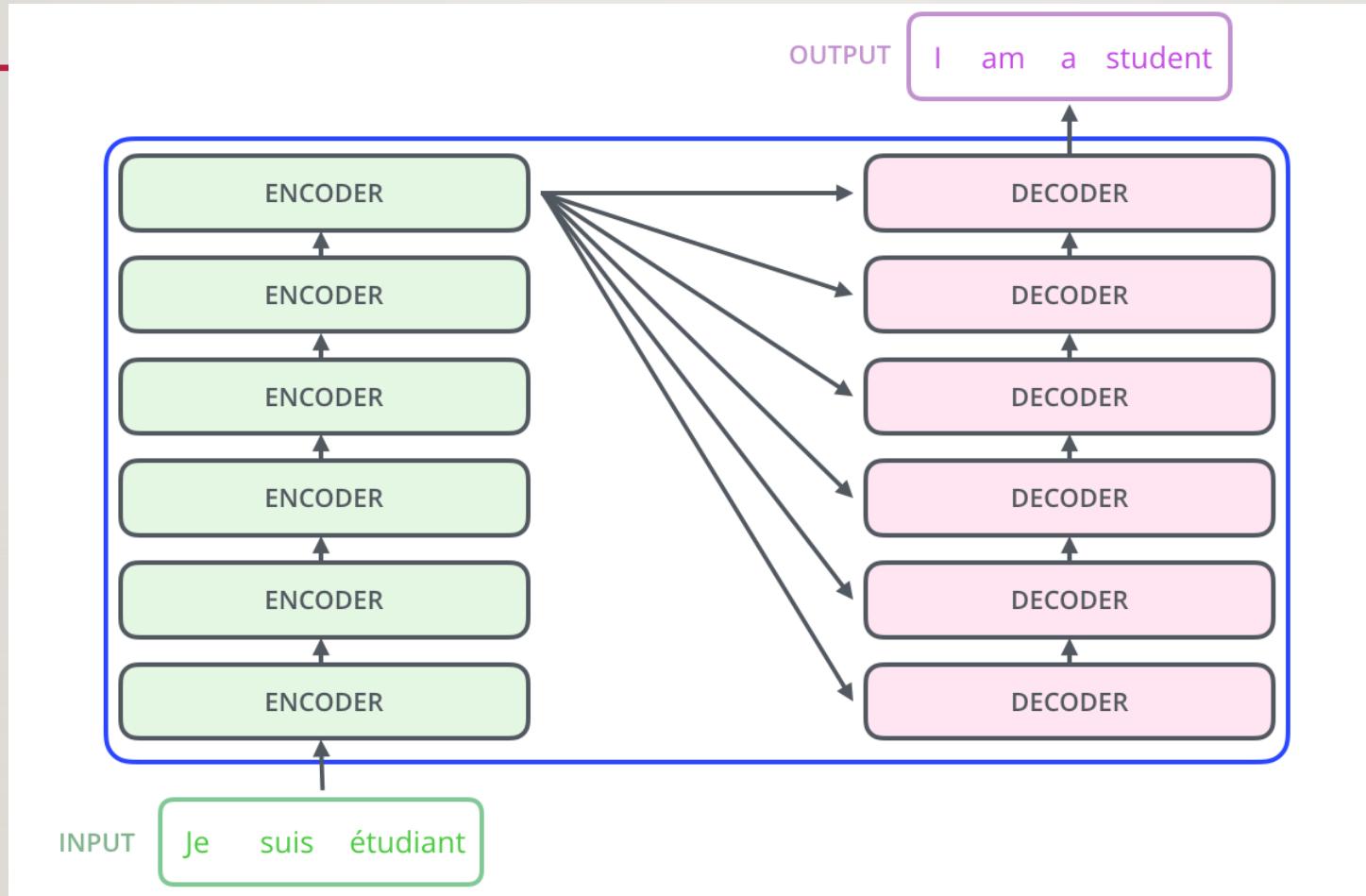
As aliens entered our planet and began to colonize earth a certain group of extraterrestrials ...

# TRANSFORMERS – HOW IT WORKS

---

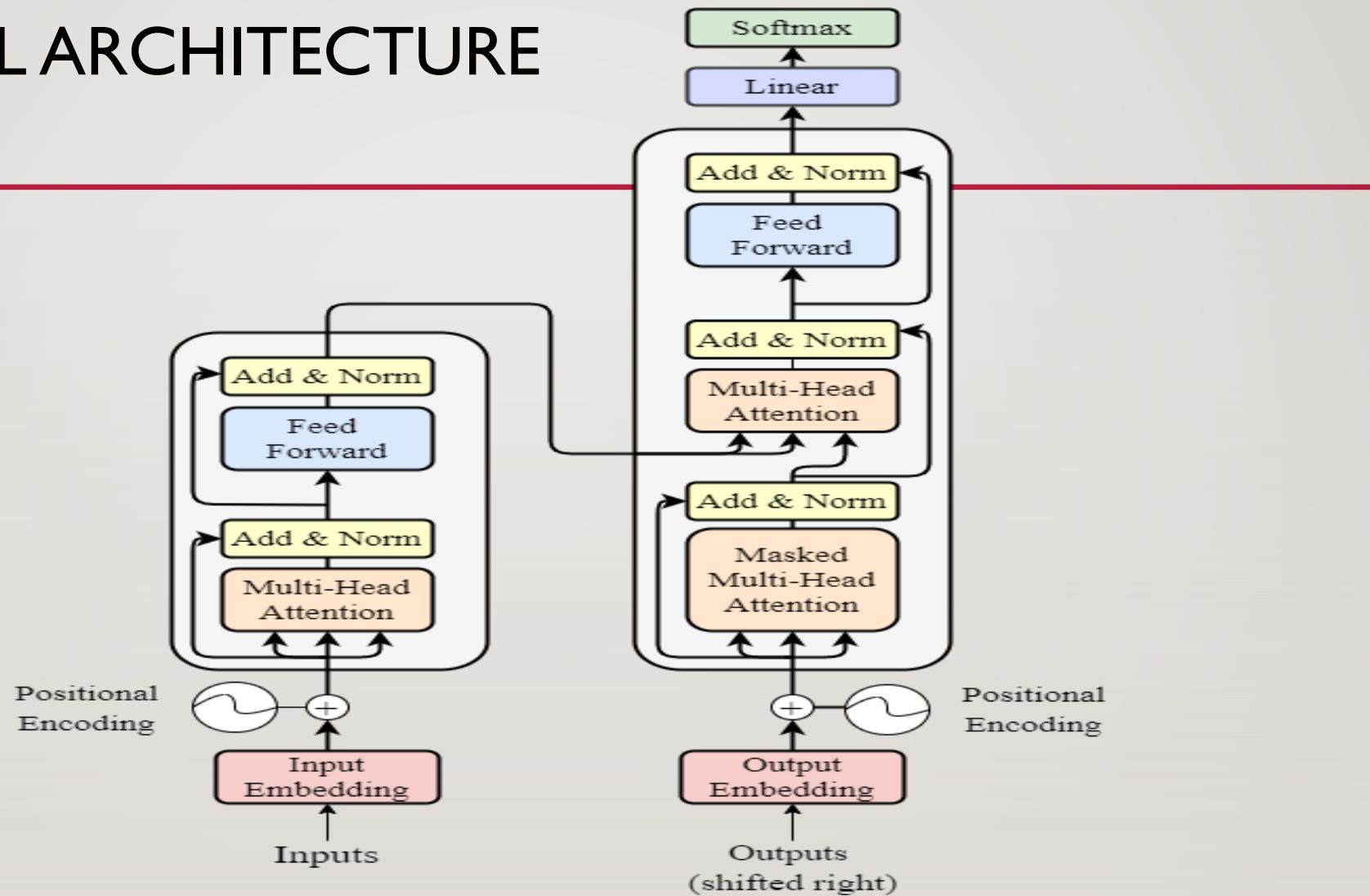


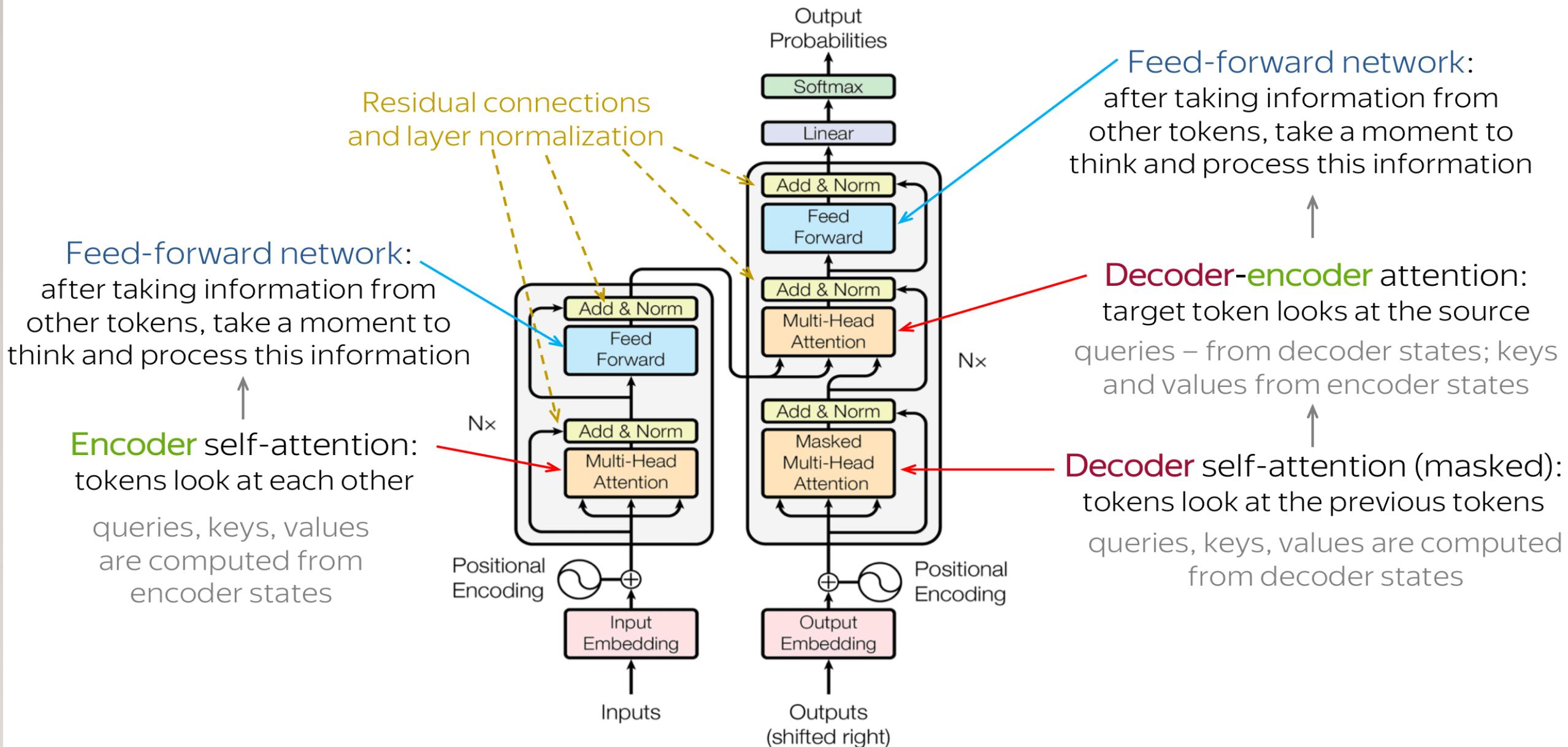
# TRANSFORMERS – HOW IT WORKS



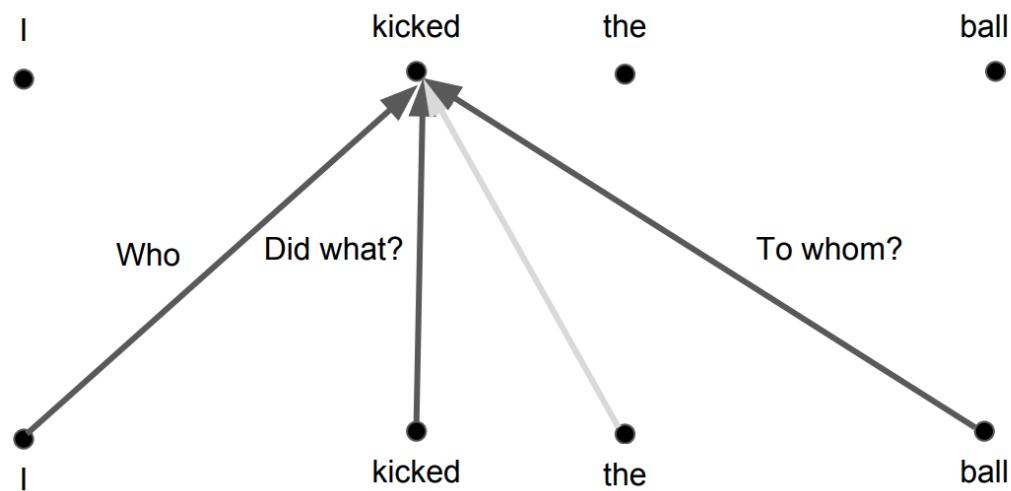
# OVERALL ARCHITECTURE

---





# Self-Attention



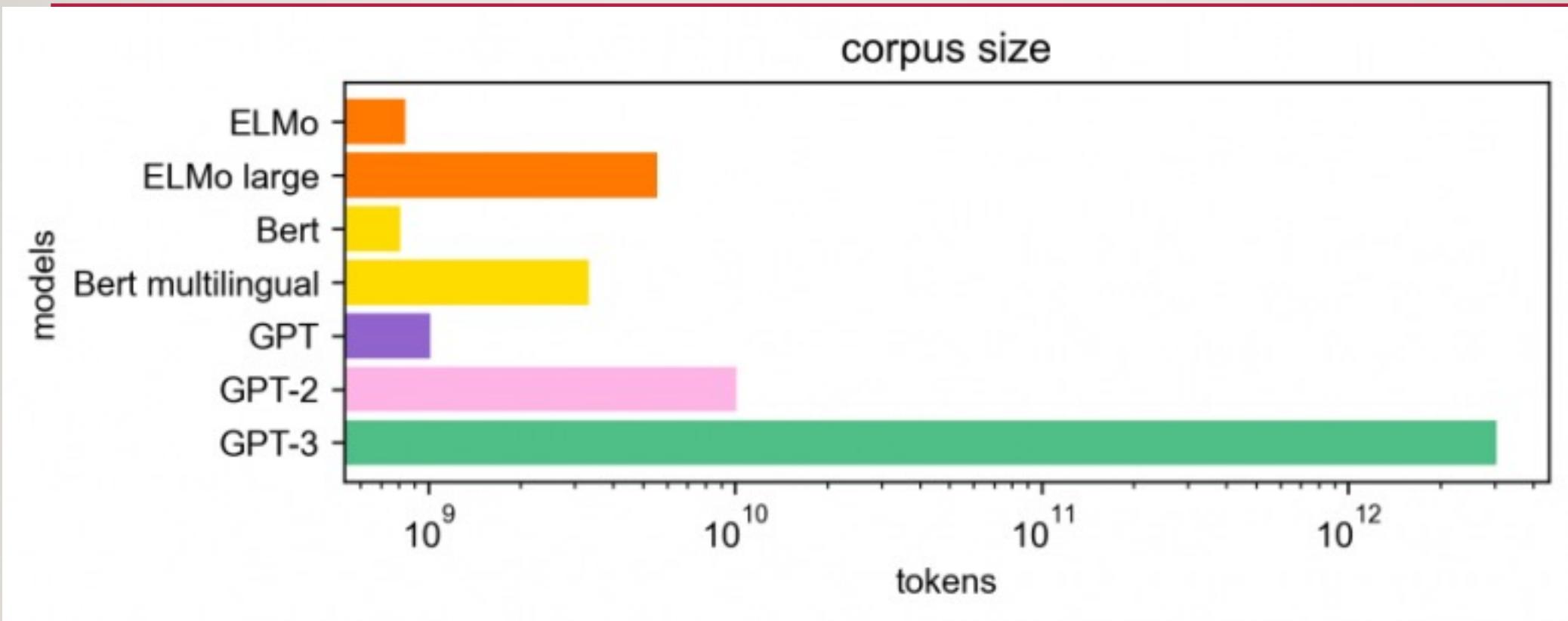
# EVOLUTION OF NLP MODELS



# ELMo – Embedding from Language Models

# BERT - Bidirectional Encoder Representations from Transformers

# TRANSFORMERS - VARIOUS FRAMEWORK



# HUGGING FACES - INTRODUCTION

---

- Company build and provide the Transformers as a simple API to access and operate.
- Also has GPT-2 (Older version of the Transformer)
- Follow the process of:
  - Pipeline - All preprocessing is done at this stage
  - Model – Specific usecases will be used for solving the problem,
  - Marketplace for all companies and individual to create model and share.

---

## TRANSFORMERS LIVE DEMO

GITHUB LINK – [HTTPS://GITHUB.COM/PRASANNAVJ/TRANSFORMERS](https://github.com/prasannavj/transformers)

- TRANSLATION
- TEXT GENERATION
- SENTIMENT ANALYSIS
- SUMMARIZATION
- TEXT CLASSIFICATION
- QUESTION ANSWERING
-

# COMPUTER VISION

---

- Neural network architecture was predominantly used for Image specific activity.
  - CNN – Convolutional Neural Network was prominent for Image based activity.
- 
- Transformers are built for NLP usecases but in 2020 – google researcher used the transformers for Image classification and was successful.
  - There born the Vision Transformers(ViT)

# HUGGING FACES VISION TRANSFORMER MODULE

---

- Similar to NLP we can use the Pipeline and Model for the Vision Transformer for the Image classification activity.
- Due to parallel processing in nature the Vision transformers are much faster.

# VISION TRANSFORMER DEMO

---

- Hugging faces vision transformer demo.
- GitHub Link: <https://github.com/prasannavj/TRANSFORMERS>

# OPENAI GPT3 ACCESS

---

- Open AI provides GPT3 pretrained model on cost basis.
- Free tier available for access and utilizing the same.