

MES College of Engineering Pune-01
Department of Computer Engineering

| | |
|-----------------------------|----------------------------------|
| Name of Student: | Class: |
| Semester/Year: | Roll No: |
| Date of Performance: | Date of Submission: |
| Examined By: | Subject: LP_VI (EL-V NLP) |
| | |

Assignment No. 4

Aim: Create a transformer from scratch using the Pytorch library.

Theory:

What Is PyTorch, and How Does It Work?

PyTorch is an optimized Deep Learning tensor library based on Python and Torch and is mainly used for applications using GPUs and CPUs. PyTorch is favored over other Deep Learning frameworks like TensorFlow and Keras since it uses dynamic computation graphs and is completely Pythonic. It allows scientists, developers, and neural network debuggers to run and test portions of the code in real-time. Thus, users don't have to wait for the entire code to be implemented to check if a part of the code works or not.

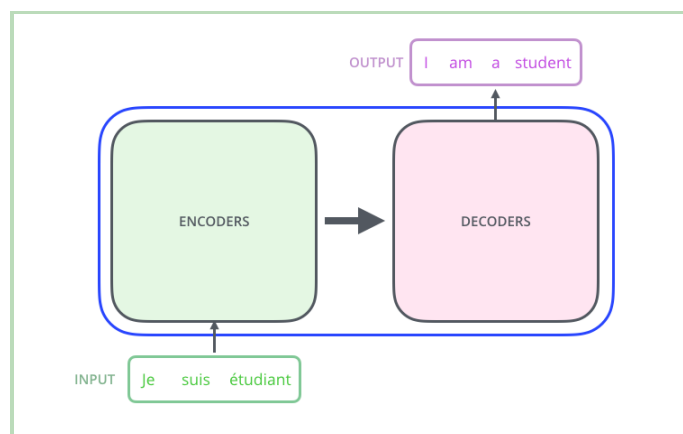
The two main features of PyTorch are:

- Tensor Computation (similar to NumPy) with strong GPU (Graphical Processing Unit) acceleration support
- Automatic Differentiation for creating and training deep neural networks

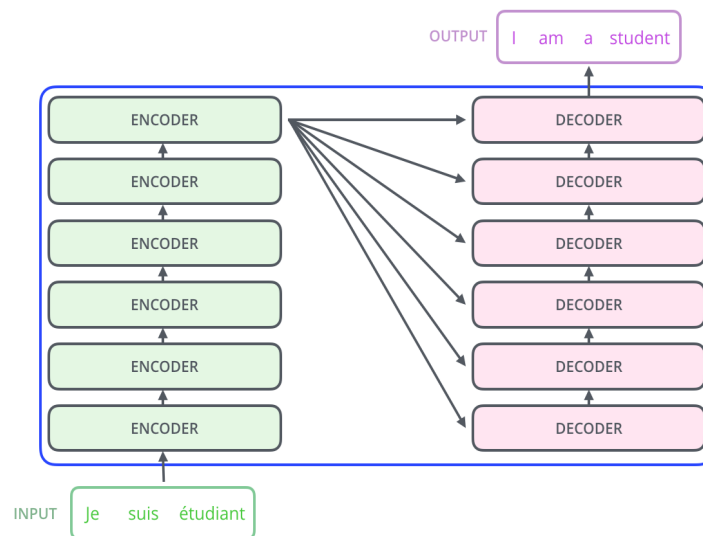
What is a Transformer?

A transformer is a deep learning model that adopts the mechanism of self-attention, differentially weighting the significance of each part of the input (which includes the recursive output) data. It is used primarily in the fields of natural language processing (NLP)^[1] and computer vision (CV).^[2]

Basically transformers have an encoder-decoder architecture.



The above image shows a language translation model from French to English. Actually we can use stack of encoder(one in top of each) and stack of decoders as below:



On a high level, the encoder maps an input sequence into an abstract continuous representation that holds all the learned information of that input. The decoder then takes that continuous representation and step by step generates a single output while also being fed the previous output.

Useful resources:

- <https://www.kaggle.com/code/arunmohan003/transformer-from-scratch-using-pytorch>
- Understanding transformers
 - <https://theaisummer.com/transformer/>
 - <https://jalammar.github.io/illustrated-transformer/>
- Pytorch implementation
 - <https://www.youtube.com/watch?v=U0s0f995w14>

Conclusion:

Questions:

1. What is language modeling? Explain any one language model in detail.
2. What is the transformer model in NLP and how it works?
3. What is topic Modeling?

