

0. Setup

1. Transformer models

Transformers, why are they so damn cool?

Introduction

Natural Language Processing

Transformers, what can they do?

How do Transformers work?

Encoder models

Decoder models

Sequence-to-sequence models

Bias and limitations

Summary

End-of-chapter quiz

2. Using 🤖 Transformers

3. Fine-tuning a pretrained model

4. Sharing models and tokenizers

Agenda

- Roughly 1 Chapter/week
- + Lessons on integrating fastai

Session 1 :

- start looking section 0 and 1
- least FASTAI part how the transformer works ,what kind of NLP tasks it's suited for and how do they work
- burrify level api for doing prediction on huggingface
- Resources and homework (improving how the transformers work)

Resources :

Resources

Study Group registration page: <https://wandb.me/fastai-hf>

Study Group discord: <https://discord.gg/DsnRxSyt>

fastai:

1. The fastai course (<https://course.fast.ai/>)
2. The Walk w/ fastai course (<https://walkwithfastai.com/>)
3. The FastBook (available for purchase or free online via Jupyter notebooks)
4. The FastBook reading/study group form W&B (<http://wandb.me/fastbook>)

fastai + Hugging Face libraries:

1. AdaptNLP (<https://novetta.github.io/adaptnlp/>)
2. FastHugs (<https://github.com/morganmcg1/fasthugs>)
3. Blurr (<https://ohmeow.github.io/blurr/>)

ML/Data Science in general:

1. The Chai Time Data Science podcast (<http://youtube.com/c/chaitimedatascience>)
2. Weights & Biases (<https://wandb.ai/>)

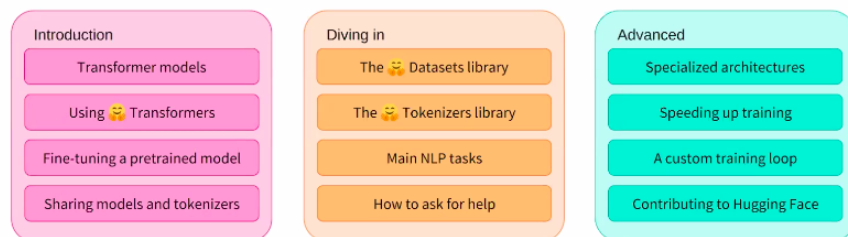
links will be there in resources section

setting up a working environment

for installing transformers

```
pip install transformers[sentencepiece]
```

1. Transformer models: Introduction



What is NLP and general info

What is it?

"NLP is a field of linguistics and machine learning focused on understanding everything related to human language."

What are some common NLP tasks?

1. **Sequence Classification:** Classify whole documents (sentiment, is/not spam, is/not grammatically correct, whether two sentences are logically related)
2. **Token Classification:** Classify individual words (grammatical category like noun, verb, adj.), NER like person, location, organization)
3. **Text Generation:** Complete a prompt with auto-generated text, fill in the blank, given a sentence in one language translate it into another, summarize a larger document
4. **Extractive Q&A:** Given a question and a context, extract the answer based on the information provided in the context

what is transformer and what it can do ?

What is it?

"NLP is a field of linguistics and machine learning focused on understanding everything related to human language."

What are some common NLP tasks?

1. **Sequence Classification:** Classify whole documents (sentiment, is/not spam, is/not grammatically correct, whether two sentences are logically related)
2. **Token Classification:** Classify individual words (grammatical category like noun, verb, adj.), NER like person, location, organization)
3. **Text Generation:** Complete a prompt with auto-generated text, fill in the blank, given a sentence in one language translate it into another, summarize a larger document
4. **Extractive Q&A:** Given a question and a context, extract the answer based on the information provided in the context

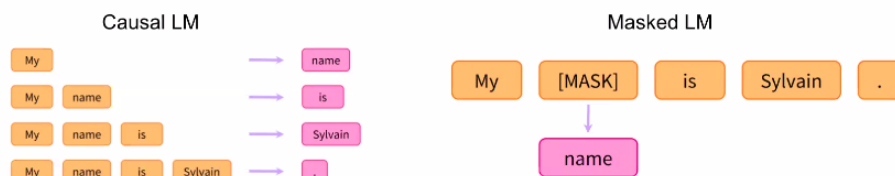
note: while installing with pip you can pass `--qg` so that it won't give more clutter on screen

AutoTokenizer : automatically infer the correct objects based on the model name or paths we are passing

BertTokenizer : use AutoTokenizer it's better cause it automatically detects what kind of problem you are working on

1. Transformer models: How do Transformers work?

"All the Transformer models mentioned above (GPT, BERT, BART, T5, etc.) have been *trained as language models*."



To use these models for specific tasks (like NER, classification, summarization, etc...), we use **transfer learning** where the pre-trained language models are **fine-tuned** "in a supervised way" for the specific task

encoder transformers works through encoding the inputs and builds a representation , convert the text into the self attention as it's main component the output is high level representation of inputs

decoder : generally focused on generating the output (generating higher level representation of what encoder is doing)

What is attention layer :

this will tell model to specific attentionn to certain words in the scentece which we've passed when dealing with the representation of each word

Architecture

1. Transformer models: How do Transformers work?

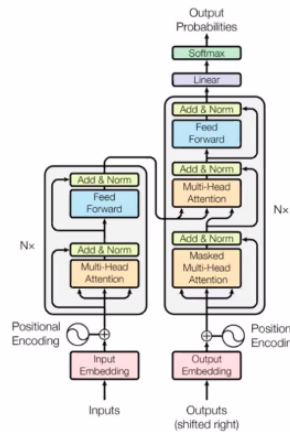
The original architecture

Attention Is All You Need

<https://arxiv.org/abs/1706.03762>

The Illustrated Transformer

<https://jalammar.github.io/illustrated-transformer/>



look at this : <https://theaisummer.com/transformer/>

Jay's blog on transformers : <https://www.youtube.com/watch?v=-QH8fRhqFHM>

Natural language understanding takss : Albert, bert ,destilbert ,etc

Decoder model : CTRL, GPT, Gpt-2 , Transfromer CL

Sequence to sequence model : here the input length is not equal to the output length

eg : BART/mBART,M2M100, MarianMT, Pergasus, PropetNet, T5/mT5

lot fo of the models are having some bias and limitations are mentioned on the website

read more papers slides are shared on email

Homework

Homework

1. Read the **Attention Is All You Need** paper (and Jay Alammar's "*The Illustrated Transformer*")
2. Read the **BERT** paper (and Jay Alammar's "*The Illustrated BERT, ELMo, and co. (How NLP Cracked Transfer Learning)*")
3. Get acquainted with one or more of the fastai Hugging Face libraries; see if you can build a basic classification model of some sort