

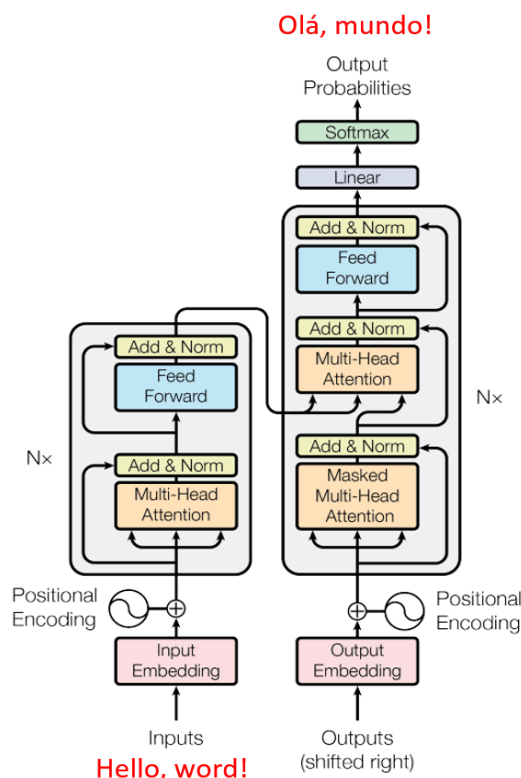
## Relatório 2 - Prática: Introdução a Modelos Transformers (I)

Vitor Eduardo de Lima Kenor

### Descrição da atividade

O primeiro vídeo do card se trata de uma explicação sobre os modelos transformers. No começo do vídeo ela cita alguns modelos transformers conhecidos como BERT, GPT-3 e T5. No vídeo ela explica o que são os modelos transformers, como eles funcionam e por que eles são tão impactantes.

Já sobre a seção oito do curso disponibilizado no card temos 12 vídeos. O primeiro e o segundo vídeo trata sobre uma introdução aos transformers. Nesses vídeos, somos apresentados a topologia de um modelo transformer.



Ele é dividido em duas partes principais, são elas: Encoder e Decoder. Pontos importantes sobre os modelos transformers: Modelo Revolucionário, Utilizado em GPT-3 T5 e BERT, Não usa Recorrência como RNN, Baseado em Mecanismo de Atenção "Attention", Arquitetura ENDEC. O Attention é o fundamento principal de como um transformer funciona, no momento de sua criação quando o modelo está sendo treinado ele aprende a relação entre as palavras, um aprendizado contextual.

Indo para o terceiro vídeo da seção, temos a explicação sobre o transformer BERT. O modelo BERT possui apenas a parte de encode, pois seu objetivo é gerar uma representação da linguagem. É um modelo pré-treinado, o BERT padrão que é o inglês foi treinado através do Toronto BookCorpus e Wikipedia, 4 dias utilizando 16 TPUs. Como treinar um modelo é muito custoso computacionalmente, o que é normal é utilizar esses modelos pré-treinados que são disponíveis e fazer um ajuste fino. O modelo BERT como

padrão é em inglês, mas existe o Multilingual Bert e o BERTimbau como opções para o nosso idioma.

O quarto vídeo apresenta 5 versões diferentes do BERT. A primeira é o ALBERT que é uma versão mais leve, foi criada utilizando uma técnica de redução de parâmetros. RoBERTa implementado com Pytorch, treinado em diferentes tipos de textos. ELECTRA, utiliza uma técnica de substituição de tokens ao invés de máscaras. XLNet, utiliza uma técnica de permutação onde os tokens são previstos de forma aleatória. DistilBert é uma versão menor e mais rápida do BERT.

No quinto vídeo fala sobre duas empresas especializadas em modelos de AI especialmente NLP, são elas a OpenAI e Hugging Face. O Hugging Face fornece APIs e ferramentas para baixar modelos pré-treinados de última geração e ajustá-los ainda mais para maximizar o desempenho. A OpenAI possui a família de modelos GPT, porém esses modelos não são gratuitos.

O sexto vídeo é apenas um tutorial de como usar modelos que estão disponíveis no Hugging Face.

Já no sétimo vídeo através de um Notebook Colab, aprendemos a utilizar um modelo transformer do Hugging Face para responder perguntas dado um contexto.

```
[1] !pip install transformers

[2] import transformers
    from transformers import pipeline

qea = pipeline("question-answering", model="timpal01/mdeberta-v3-base-squad2")

[4] texto = "Capitão América é um super-herói de histórias em quadrinhos americanos publicado pela Marvel Comics. Criado por Joe Simon e Jack Kirby,

[6] pergunta = "Quem criou o personagem chamado de Capitão América"
    resposta = qea(question=pergunta, context=texto)
    resposta

{'score': 0.96441251039505,
 'start': 111,
 'end': 135,
 'answer': ' Joe Simon e Jack Kirby,'}
```

No oitavo vídeo através de um Notebook Colab, aprendemos a utilizar um modelo do Hugging Face para a tarefa de fill-mask.

```
[1] !pip install transformers

[2] import transformers
    from transformers import pipeline

mascarar = pipeline("fill-mask", model="FacebookAI/xlm-roberta-base")

[17] texto = mascarar("Existe uma chance de o copo cair no <mask>.")
    for x in range(len(texto)):
        print(texto[x])

{'score': 0.1463506817817688, 'token': 177122, 'token_str': 'chão', 'sequence': 'Existe uma chance de o copo cair no chão .'}
{'score': 0.07143672555685043, 'token': 84114, 'token_str': 'fogo', 'sequence': 'Existe uma chance de o copo cair no fogo .'}
{'score': 0.04202737659215927, 'token': 175194, 'token_str': 'céu', 'sequence': 'Existe uma chance de o copo cair no céu .'}
{'score': 0.036009132862091064, 'token': 187, 'token_str': 'ar', 'sequence': 'Existe uma chance de o copo cair no ar .'}
{'score': 0.03163788095116615, 'token': 59512, 'token_str': 'sangue', 'sequence': 'Existe uma chance de o copo cair no sangue .'}

```

No nono vídeo através de um Notebook Colab, usamos o pipeline com o modelo padrão para criar um resumidor de textos.

```
[1] !pip install transformers
Mostrar saída oculta

[2] import transformers
    from transformers import pipeline

[4] resumidor = pipeline("summarization")
Mostrar saída oculta

[5] texto = "Tony Stark teve uma relação difícil com seu pai, sendo enviado aos seis anos de idade para um internato, onde ele iria em breve começar a ex

[7] resumo = resumidor(texto, max_length=100)

print(resumo)
[{'summary_text': ' Tony Stark teve uma relação difícil com seu pai, sendo enviado aos seis anos de idade para um internato . Howard Stark, mesmo send
```

O décimo e último vídeo de exemplo, mostra como usarmos o pipeline e algum modelo do Hugging Face para criarmos um gerador de texto.

```
[2] !pip install transformers
Mostrar saída oculta

[3] import transformers
    from transformers import pipeline

[4] gerador = pipeline("text-generation", model="pierreguillou/gpt2-small-portuguese")
Mostrar saída oculta

[6] texto = "Thor nasceu, filho de Odin, rei de Asgard, e da gigante, Jord. Ele tinha um irmão mais novo, Meili."

resultado = gerador(texto)
print(resultado)
, e da gigante, Jord. Ele tinha um irmão mais novo, Meili. Ele foi adotado pela mãe na sua juventude, mas ficou com o pai até a idade de treze anos.']]
```

Nos últimos dois vídeos da seção temos um tutorial de como utilizar um modelo da OpenAI para responder perguntas, se baseando em um contexto que você passa em seu prompt.

## Conclusões

Neste card tivemos desde exemplos teóricos e práticos de como funciona modelos transformers, onde podemos encontrar esses modelos para utilizarmos e como utilizarmos para algumas tarefas. Podemos ver como eles são grandiosos e revolucionários quando se trata de NLP.

## Referências

Vídeo de explicação sobre transformers:

▶ Transformers, explained: Understand the model behind GPT, BERT, and T5

Curso da Udemy:

<https://www.udemy.com/course/formacao-processamento-de-linguagem-natural-nlp/?couponCode=2021PM25>