

CamemBERT: a Tasty French Language Model

Présenté par :

Emir TAS, Shivamshan SIVANESAN, Kishanthan KINGSTON

Encadré par :

Pr. Nicolas OBIN

État de l'art :

CamemBERT: a Tasty French Language Model [1]:

Dans l'article, les auteurs ont utilisé RoBERTa comme architecture et Oscar corpus (partie français seulement) pour l'entraînement du modèle. Ils ont utilisé SentencePiece pour segmenter les mots.

Évaluation: Part-Of-Speech tagging, dependency parsing, Named Entity Recognition, Natural Language Inference.

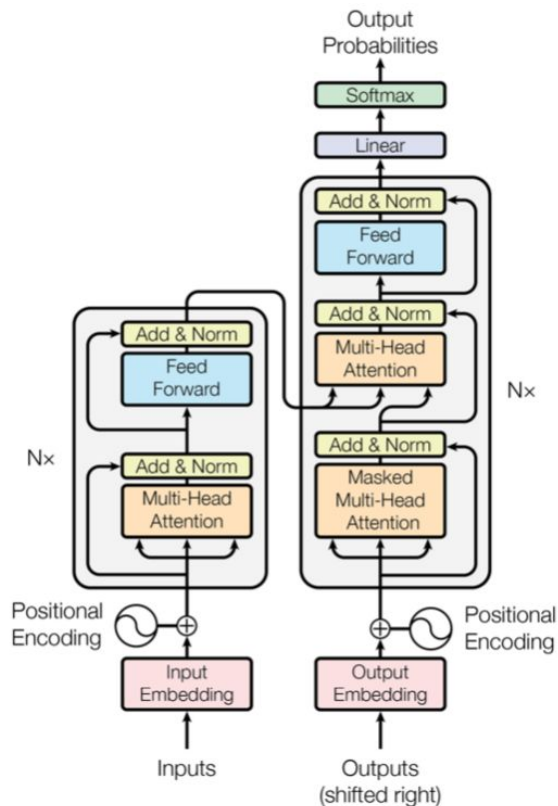
RoBERTa: A Robustly Optimized BERT Pretraining Approach [2]:

RoBERTa est une architecture qui est basée sur l'architecture BERT (Bidirectionnel Encoder Representations from Transformers), qui à son tour est une variante des transformers.

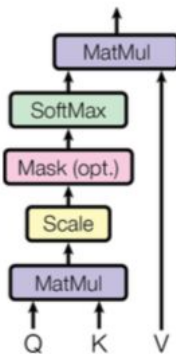
Attention Is All You Need [3]:

Présentation des transformers basé sur le mécanisme de l'attention.

Architecture Transformers [2] :



Scaled Dot-Product Attention



Multi-Head Attention

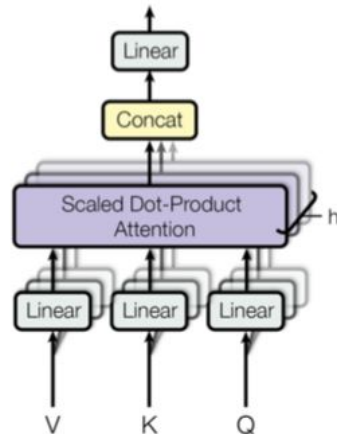


Figure 2: (left) Scaled Dot-Product Attention. (right) Multi-Head Attention consists of several attention layers running in parallel.

Figure 1: The Transformer - model architecture.

Architecture RoBERTa :

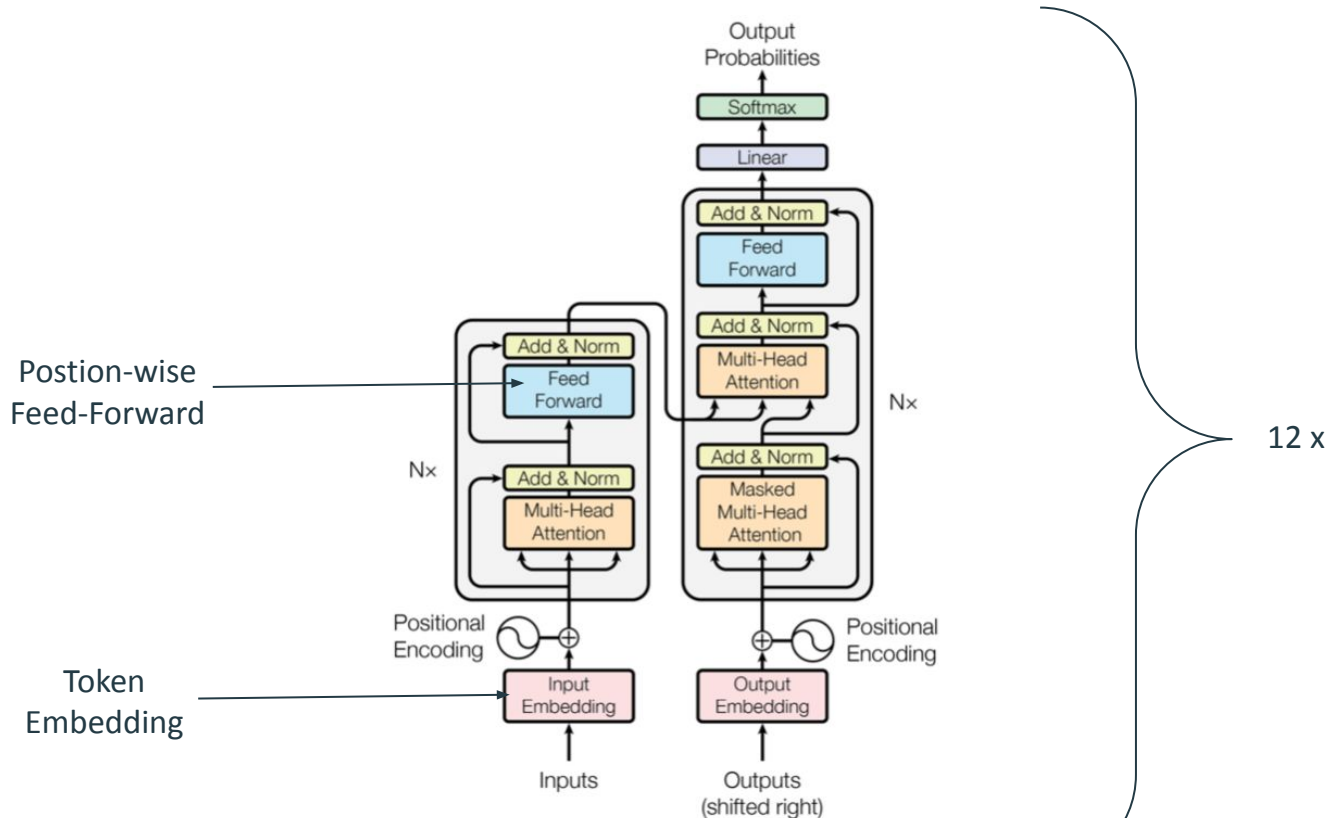


Figure 1: The Transformer - model architecture.

Bibliographie:

[1] Martin, Louis, et al. « CamemBERT: A Tasty French Language Model ». Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2020, p. 7203-19. DOI.org (Crossref), <https://doi.org/10.18653/v1/2020.acl-main.645>.

[2] Y. Liu et al., « RoBERTa: A Robustly Optimized BERT Pretraining Approach ». arXiv, 26 juillet 2019. Consulté le: 22 novembre 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/1907.11692>

[3] A. Vaswani et al., « Attention Is All You Need ». arXiv, 1 août 2023. Consulté le: 22 novembre 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/1706.03762>

