

Rapport de TIPE

Le traitement du langage naturel par transformers illustré par un exemple pour la classification de texte

Nathan Cerisara, MPI, 2022-2023

16 mai 2023

Résumé

In recent years, there has been a historic advance in artificial intelligence, in many subfields (vision, language, health, ...). And among the major advances in the field of natural language processing, the transformer architecture has played a very important role, as shown by the performance of GPT models that use this architecture.

In this work, I try to understand how this architecture works, while experimenting with an exemple where I train a sentiment classification model based on BERT.

1 L'Architecture Transformer

1.1 Objectifs de cette architecture

1.2 Vision d'ensemble de l'architecture

1.3 La transformation des mots en tokens à l'entrée du modèle

1.4 La partie Encodeur

1.4.1 Sa structure

1.4.2 Les Matrices d'Attention

1.5 La partie Décodeur

1.5.1 Sa structure

1.5.2 Les Matrices de Cross-Attention

2 Preuve

3 Application Personnelle de ce modèle

3.1 Objectif / Rapport à la ville

3.2 Le modèle BERT

3.3 La structure complète du réseau utilisé

3.4 Les données et l'apprentissage

3.5 Les résultats

4 Bilan

5 Annexe