



Université d'Aix-Marseille

**UFR des Sciences
Département d'Informatique**

Rapport d'Etude

Spécialité :

Master 1 Informatique

Science et Ingénierie des Données

Thème :

Le modèle « transformer ».

Réalisé par :

M^{lle} ZEMMOURI Yasmine.

M^{lle} BEKHADDA Hadjira.

Responsable :

M. JEGOU Philippe.

Année Universitaire : 2023-2024.

Résumé

Les Transformers jouent un rôle fondamental dans notre quotidien numérique. Ces puissants modèles informatiques, devenus essentiels à notre manière d'interagir avec la technologie moderne, sous-tendent des actions aussi simples que les recherches en ligne et les traductions automatiques, ainsi que des interactions plus avancées avec les assistants vocaux. Ces modèles excellent dans le traitement du langage naturel en capturant des relations complexes entre les mots, et leur capacité à accorder une attention différenciée à divers éléments d'une séquence de mots les rend particulièrement performants dans la résolution de tâches linguistiques. Notre étude est divisée en deux chapitres, explorant l'essence des transformateurs et leurs applications diverses. Alors que nous avons examiné en détail leur fonctionnement, leur histoire, et leur impact sur des domaines spécifiques, la conclusion souligne l'omniprésence croissante de ces modèles, anticipant une intégration encore plus profonde dans notre quotidien à mesure que la technologie continue de progresser. Les transformateurs ne se contentent pas de traiter des données, ils comprennent le langage humain, s'adaptent contextuellement, et promettent de redéfinir les limites de l'intelligence artificielle dans notre vie quotidienne.

Mots Clés : transformers, propagation, mécanisme d'attention, self-attention, NLP, TAL, encodeur, décodeur, apprentissage profond, réseaux de neurones, GPT, BERT.

Abstract

The Transformers play a fundamental role in our digital daily life. These powerful computational models, essential to the way we interact with modern technology, actions as simple as online searches and automatic translations, as well as more advanced interactions with voice assistants. These models excel in natural language processing by capturing complex relationships between words. Their ability to selectively attend to various elements in a sequence of words makes them particularly effective in solving linguistic tasks. Our study is divided into two chapters, exploring the essence of transformers and their diverse applications. While we have examined in detail their architecture, history, and impact on specific domains, the conclusion underscores their growing presence, anticipating an even deeper integration into our daily lives as technology continues to advance. Transformers do not only process data; they understand human language, adapt contextually, and promise to redefine the boundaries of artificial intelligence in our everyday lives.

Keywords : transformers, feedforward network, attention mechanism, self-attention, NLP, TAL, encoder, decoder, deep learning, neural networks, GPT, BERT.

Table des matières

Résumé	1
Introduction Générale	3
1 Qu'est ce que le modèle Transformer ?	4
1.1 Introduction	4
1.2 Historique et contexte du modèle	4
1.3 Architecture et fonctionnement	5
1.3.1 Encodeur-Decodeur	5
1.3.2 Mécanisme d'attention	6
1.3.3 FeedForward Newtork	7
1.4 Entraînement et apprentissage	7
1.5 Conclusion	8
2 Applications du modèle Transformer	9
2.1 Introduction	9
2.2 Génération de texte	9
2.3 Traduction automatique	10
2.4 Résumé automatique	10
2.5 Analyse de sentiments	11
2.6 Conclusion	11
Conclusion Générale	12

Introduction Générale

Dans notre vie quotidienne, il nous est amené à faire une recherche en ligne, que ce soit pour trouver des informations spécifiques, résoudre des problèmes complexes, ou simplement satisfaire notre curiosité.

Ce processus, devenu presque instinctif, repose largement sur les Transformers, ces puissants modèles de langage qui sous-tendent nos interactions avec la technologie moderne. Ils travaillent en coulisses pour analyser le sens de nos requêtes, prenant en compte le contexte et fournissant des résultats pertinents.

Ce sont des modèles d'apprentissage profond qui excellent dans le traitement du langage naturel. Ils fonctionnent en capturant des relations complexes entre les mots dans une phrase, permettant une compréhension plus fine du contexte et de la signification. L'aspect clé des transformateurs est leur capacité à accorder une attention différentielle à différentes parties d'une séquence de mots, ce qui les rend particulièrement performants dans la résolution de tâches linguistiques [5].

Lorsque nous utilisons des services de traduction en ligne par exemple, les Transformers sont en action. Ils dépassent la simple correspondance mot à mot en comprenant le sens global des phrases, prenant en compte les nuances culturelles et contextuelles pour fournir des traductions plus précises et naturelles.

Les assistants vocaux reposent également sur ces modèles pour interpréter et répondre aux requêtes parlées de manière intelligente. Ces derniers traitent la parole comme une séquence de mots et sont capables de déchiffrer les intentions derrière les commandes vocales, offrant ainsi une expérience utilisateur fluide et naturelle.

Dans le cadre d'un sujet d'études, nous nous intéressons principalement sur cet émergent modèle, fournissant une étude complète sur les Transformers. Pour ce faire, le rapport est divisé en deux chapitres détaillés comme suit :

- **Chapitre 1** - Qu'est ce qu'un Transformer ? : introduit l'historique, l'architecture et l'apprentissage de ce modèle.
- **Chapitre 2** - Application du modèle Transformer : présente divers applications de cette technologie, en incluant les modèles les plus utilisés.

Pour finir, une conclusion résume notre étude et discute de quelques perspectives futures.

Chapitre 1

Qu'est ce que le modèle Transformer ?

1.1 Introduction

Les modèles Transformers, une catégorie avancée de réseaux de neurones artificiels, sont devenus particulièrement connus, surtout dans le domaine du traitement du langage naturel, *Natural Language Processing* (NLP), au cours des dernières années. Leurs performances dans une variété de tâches liées au NLP les positionnent comme des composants essentiels de l'intelligence artificielle.

Ce chapitre se consacre à l'exploration des bases de ces modèles, en détaillant leur mécanisme d'attention et leur architecture fondamentale. À travers cette analyse, nous dévoilerons les subtilités de leur fonctionnement, offrant ainsi un aperçu des secrets qui sous-tendent ces avancées marquantes en intelligence artificielle.

1.2 Historique et contexte du modèle

Les Transformers, également connus sous le nom de modèles auto-attentionnels, représentent une innovation majeure dans le domaine de l'apprentissage profond. Le Transformer est un modèle de *Deep Learning* (c'est un réseau de neurones) de type seq2seq qui a la particularité de n'utiliser que le mécanisme d'attention et aucun réseau récurrent (RNN) ou convolutionnel (CNN). Pour rappel, un modèle seq2seq est un modèle qui prend en entrée une séquence (une suite d'éléments du même type) et renvoie une séquence en sortie [5][9].

C'est en 2017, avec la publication de "*Attention is All You Need*" par Vaswani et al. [20], que les Transformers ont été introduits, devenant rapidement l'état de l'art pour de nombreuses applications de NLP, en raison de leurs performances dans l'analyse des dépendances à longue séquence entre les mots d'une phrase.

Cette capacité marque une évolution significative par rapport aux modèles traditionnels tels que les réseaux neuronaux récurrents (RNN) et les mémoires à court terme (LSTM).

L'histoire des Transformers remonte aux années 1990, avec les premières tentatives de Jürgen Schmidhuber avec le *fast weight controller*, qui ont toutefois été inefficaces [20].

Depuis 2017, leur impact s'est étendu, marquant des étapes clés telles que la création de modèles de langage volumineux tels que BERT et GPT-2 en 2019, ainsi que l'émergence de modèles encore plus puissants comme GPT-3 en 2020, consolidant ainsi leur position dominante dans le domaine de l'apprentissage automatique [5].

1.3 Architecture et fonctionnement

L'architecture des Transformers se compose de deux éléments clés : l'encodeur et le décodeur. Le mécanisme d'attention, un des piliers de cette architecture, permet au modèle de pondérer dynamiquement les éléments d'une séquence en fonction du contexte, facilitant ainsi la capture de relations complexes à longue portée. En complément, la couche feedforward apporte non seulement de la non-linéarité au modèle, mais elle joue également un rôle crucial dans la capture de motifs complexes. Ensemble, ces composants confèrent aux Transformers la capacité à comprendre et à générer des séquences de manière contextuelle, établissant ainsi leur position de premier plan dans le traitement du langage naturel et d'autres tâches séquentielles [1][9]. La figure suivante résume l'architecture d'un Transformer [1] :

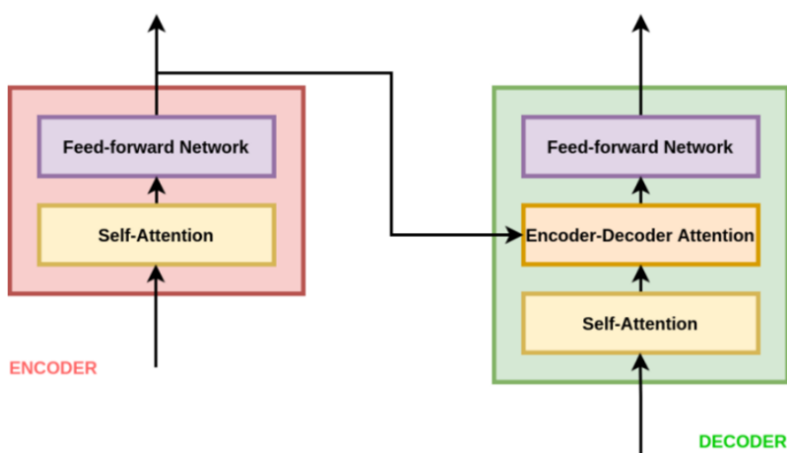


FIG. 1.1 : Architecture d'un Transformer. [1].

1.3.1 Encodeur-Décodeur

Dans le contexte des réseaux neuronaux, un encodeur et un décodeur sont des composants qui interviennent souvent dans les architectures de séquence à séquence.

- **Encodeur dans le Transformer** [9][4][12] : L'encodeur du Transformer est chargé de traiter la séquence d'entrée. Cette séquence passe à travers plusieurs couches d'encodeur, chacune comportant des mécanismes tels que l'auto-attention et des couches feedforward. L'objectif est de capturer les relations complexes et

contextuelles entre les éléments de la séquence d'entrée, produisant ainsi une représentation interne riche.

- **Décodeur dans le Transformer[9][4][12]** : Le décodeur prend la représentation interne générée par l'encodeur et utilise cette information pour générer la séquence de sortie. Comme l'encodeur, le décodeur comporte plusieurs couches, mais avec l'ajout d'un mécanisme d'attention qui prend en compte la sortie générée précédemment. Cela permet au décodeur de se concentrer sur différentes parties de la séquence d'entrée lors de la génération de chaque élément de la séquence de sortie.

Le mécanisme d'attention présent dans les couches d'encodeur et décodeur, a la capacité présente à prendre en compte les relations à long terme et à capturer des dépendances complexes entre les éléments de la séquence. Ce qui a contribué au succès des Transformers. Mais qu'est-ce que le mécanisme d'attention ?

1.3.2 Mécanisme d'attention

Le concept fondamental du modèle Transformer réside dans le fait de conserver de l'interdépendance des mots au sein d'une séquence, mis en œuvre par le mécanisme d'attention au cœur de son architecture [20].

Ce mécanisme mesure la relation entre deux éléments de deux séquences distinctes. Dans le domaine du traitement du langage naturel, l'attention guide le modèle pour concentrer ses ressources sur les mots pertinents de la séquence A lorsqu'il traite un mot de la séquence B [4].

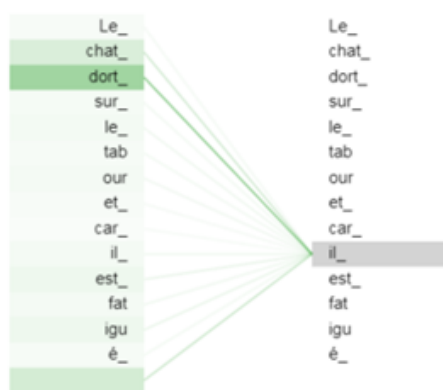


FIG. 1.2 : Exemple du mécanisme d'attention. [4].

Le *self-attention*, une variation du mécanisme d'attention, se concentre sur la détection des liens entre différents mots d'une même séquence pour créer une représentation pertinente [4][12]

Prenons l'exemple de la figure (1.2), « Le chat dort sur le tabouret car il est fatigué ». Pour un être humain, « il » se réfère au chat.

Le mécanisme de self-attention a pour objectif de détecter le lien entre « il » et « chat ».

« Le chat dort sur le tabouret car il est sous la fenêtre », dans ce cas, le mécanisme devra détecter le lien entre « il » et « tabouret ».

Le calcul d'attention spécifique au Transformer est effectué en parallèle par plusieurs blocs d'attention, connus sous le nom de "*Multi-head Attention*". Cette approche vise à éviter un biais total dans la représentation, assurant ainsi que chaque "sous-espace de représentation" contribue de manière équilibrée [12]. Les résultats de ce processus seront l'entrée de couches FeedForward.

1.3.3 FeedForward Newtork

La couche feedforward est responsable de la transformation linéaire des représentations intermédiaires issues du mécanisme d'attention. Cette transformation est ensuite suivie d'une fonction d'activation non linéaire, souvent la fonction ReLU (Rectified Linear Unit) [15].

Cette couche permet au modèle d'introduire de la non-linéarité, ce qui signifie qu'il peut apprendre des relations complexes et non linéaires au sein des données séquentielles. En ajoutant cette capacité d'apprentissage non linéaire, la couche feedforward renforce la capacité du Transformer à saisir des motifs plus complexes et à effectuer des tâches sophistiquées dans le traitement des séquences [20][15].

Prenons un exemple : soit la séquence de mots : "Le chat dort". L'encodeur prend chaque mot de cette séquence et génère des représentations intermédiaires. La couche feedforward intervient ensuite pour appliquer une transformation linéaire à ces représentations.

Par exemple, elle pourrait attribuer des poids différents à chaque mot, accentuant certaines caractéristiques. La fonction d'activation non linéaire (comme ReLU) introduit de la flexibilité, permettant au modèle d'apprendre des relations plus complexes entre les mots [15].

Dans ce cas, la couche feedforward pourrait aider le modèle à comprendre que "dort" est lié à "chat" et à extraire des informations importantes pour la tâche en cours, comme l'objet (chat) et l'action (dormir).

1.4 Entraînement et apprentissage

L'apprentissage dans le domaine de l'intelligence artificielle est un processus itératif et adaptatif. Pendant l'entraînement, le modèle absorbe des connaissances à partir de données d'entraînement, ajustant constamment ses paramètres pour réduire les divergences entre ses prédictions et les étiquettes réelles. Grâce à la rétropropagation de l'erreur et à l'optimisation, le modèle acquiert une compréhension approfondie des tâches spécifiques. L'apprentissage va au-delà de la simple mémorisation des exemples d'entraînement, visant à généraliser ces connaissances pour des prédictions précises sur de nouvelles données [5].

Particulièrement, les Transformers tirent parti de leur capacité à réaliser la majeure partie de leur apprentissage de manière non supervisée. Cette caractéristique est particulièrement marquante en comparaison avec les méthodes antérieures, où les modèles de NLP devaient simultanément apprendre le langage et la tâche spécifique de manière supervisée [21].

Le processus d'entraînement des Transformer se divise généralement en deux étapes distinctes :

- **Pré-entraînement non supervisé [21][6]** : Les Transformers sont initialement pré-entraînés sur de grandes quantités de données non étiquetées, généralement en utilisant des tâches d'apprentissage auto-supervisé. Par exemple, un modèle peut apprendre à prédire le mot suivant dans une phrase après avoir lu les mots précédents, ou à prédire un mot masqué dans une phrase. Ce pré-entraînement vise à permettre au modèle de développer une compréhension statistique de la langue sur laquelle il a été entraîné.
- **Ré-entraînement supervisé [21][6]** : Par la suite, les Transformer subissent une phase de ré-entraînement, souvent appelée fine-tuning, dédiée à une tâche spécifique. Ce processus se base sur l'apprentissage par transfert, où le modèle est ajusté de manière supervisée en utilisant des étiquettes annotées par des humains pour la tâche ciblée. Par exemple, s'il s'agit d'un modèle de traduction, le fine-tuning le préparera à traduire de manière précise d'une langue source à une langue cible.

Ce découpage en deux étapes permet aux Transformer de capitaliser sur les avantages de l'apprentissage non supervisé, en acquérant une connaissance approfondie du langage, suivi d'une spécialisation efficace pour des tâches spécifiques lors du fine-tuning supervisé.

1.5 Conclusion

Dans ce chapitre, nous avons exploré l'histoire, l'architecture et le processus d'entraînement des Transformers. Issus du modèle Transformer en 2017, ces modèles ont redéfini le traitement du langage naturel. Le prochain chapitre se penchera sur leurs applications concrètes, illustrant comment ces modèles ont transformé divers domaines, de la compréhension du langage naturel à la vision par ordinateur, grâce à des exemples concrets.

Chapitre 2

Applications du modèle Transformer

2.1 Introduction

Comme mentionné dans le chapitre précédent, les modèles Transformer sont capables d'apprendre des relations complexes entre les mots d'une séquence. Ce fait les rend très performants pour une variété de tâches NLP, notamment la traduction automatique, la compréhension du langage naturel et la génération de texte. Dans ce chapitre, nous proposons un aperçu complet des applications des modèles transformateurs dans le domaine NLP.

2.2 Génération de texte

La génération de texte a connu un regain d'intérêt ces dernières années, notamment avec la mise à disposition d'API commerciales telles que GPT, Cohere et ChatGPT.

L'objectif consiste à générer de manière automatiquement un nouveau texte, qu'il s'agisse d'un mot, d'une expression, d'une phrase ou d'un texte complet. Cette technique est utilisée dans de nombreuses applications, notamment les chatbots, la création de contenu et la traduction automatique. Le texte généré doit démontrer une compréhension de l'entrée et de la langue générée.

Voici quelques exemples d'applications de modèles Transformers pour la génération de texte :

- Le *Generative Pre-trained Transformer 3* (GPT-3) excelle dans la génération de texte créatif, que ce soit pour la rédaction de contenu, la création de poèmes, ou d'autres formes d'écriture[8]. Le célèbre modèle de langage avancé créé par OpenAI "*ChatGPT*" a été développé en utilisant ce modèle.
- Le transformateur *RoGPT2* (*Romanian GPT2*), une version roumaine du modèle GPT-2, est spécialement conçue pour la génération de texte en langue roumaine.[11]
- *Text-to-Text Transfer Transformer* (T5) est utilisé pour transformer un texte en un résumé, adapté pour la génération automatique de résumés d'articles ou de documents. Par exemple, la génération d'un résumé d'un article de presse.[16]

- Un système a été conçu pour générer de longues phrases et des articles, et pour effectuer une analyse comparative de leurs performances. Ce système utilise les modèles BERT et GPT-2.[14]

2.3 Traduction automatique

L'objectif de la traduction automatique est de générer une traduction précise du sens du texte source dans la langue cible. Les modèles de traduction automatique utilisent une architecture encodeur-décodeur pour capturer le contexte et générer du texte de longueur variable. Ils peuvent être utilisés pour translittérer afin de convertir un texte d'un alphabet à un autre, traduire sans données d'entraînement parallèles et établir des correspondances entre des mots. Parmi les transformers appliqués à la traduction automatique :

- Les modèles à base de transformateurs les plus populaires pour la synthèse et la traduction de texte comprennent GPT-3, BERT, T5 et BART. GPT-3, développé par OpenAI, comme ChatGpt est connu pour ses capacités dans les tâches de traduction. Ces modèles ont démontré leur efficacité. offrant à la fois des approches extractives et abstraites.
- Bard est basé sur LaMDA, un modèle de langage factuel de Google AI, formé sur un ensemble de données massif de texte et de code.

2.4 Résumé automatique

Le but est de générer un court texte à partir d'un fichier ou d'un document à condition que le résumé contienne les idées principales du texte original tout en prenant en considération plusieurs variables telles que la longueur, le style d'écriture et la structure. On peut y trouver différentes variantes :

1. La génération d'un nouveau texte qui capture les informations les plus pertinentes du texte original basées sur les transformateurs tels que T5, BART, Pegasus, à l'aide de mots clés.[17]
2. La génération de résumés significatifs et concis de longs documents, tels que des livres, à l'aide de modèles linguistiques basés sur des transformateurs utilisant le modèle BART.[13]

Le modèle de langage basé sur un transformateur *ChatGPT*, développé par OpenAI, est un outil puissant pour les tâches de traitement du langage naturel, y compris la synthèse de texte.

2.5 Analyse de sentiments

L'analyse de sentiments est une branche NLP. Ce processus consiste à utiliser l'analyse de texte pour identifier, extraire et étudier de manière systématique les états affectifs et les informations subjectives dans les textes [10].

Avec les vastes quantités d'informations disponibles en ligne, l'analyse des sentiments peut s'avérer très efficace pour offrir une meilleure compréhension des attitudes, opinions et émotions du public.

Cette analyse propose différentes approches pour classer les opinions en catégories et détecter des émotions telles que la colère et la joie.[19] Parmi les applications des modèles Transformers dans ce domaine :

- Des chercheurs ont exploité des données textuelles liées aux traumatismes émotionnels, notamment des types de traumatismes simples, et des réactions du corps et du cerveau à des événements stressant utilisant le transformer BERT. [3]
- La détection de sentiment des tweets arabes, utilisant des transformateurs basés sur BERT pour analyser le sentiment dans un contexte spécifique, comme les tweets liés à la COVID-19.[2]
- L'extraction d'informations utiles et subjectives des données textuelles des médias sociaux, permettant une meilleure compréhension des sentiments exprimés dans le contenu des médias sociaux, se basant sur le transformer BERT. [7]
- RoBERTa, une variante BERT fortement optimisée, a été utiliser pour analyser le sentiment dans les discours politiques.[18]

2.6 Conclusion

Nous avons exposé divers domaine d'applications des Transformers, ainsi que différents modèles dont les plus connus, notamment BERT et GPT-3 (utilisé par ChatGpt). Nous avons ainsi montré que les Transformers sont présents dans les taches les plus simples possibles.

Conclusion Générale

Dans ce présent rapport, nous avons introduit le modèle Transformer, son fonctionnement ainsi que ses applications dans différents domaines. Nous avons montré que ce modèle s'appuie principalement sur un mécanisme d'attention, veillant à capturer l'importance de chaque mot.

En conclusion, l'omniprésence des transformateurs dans notre vie quotidienne souligne l'importance croissante de l'intelligence artificielle dans la résolution de tâches complexes. Ces cerveaux informatiques ne se contentent pas de traiter des données, ils comprennent le langage humain et s'adaptent pour fournir des réponses et des solutions contextuellement appropriées.

Alors que cette technologie continue de progresser, il est probable que nous assisterons à une intégration encore plus profonde des transformateurs dans notre routine quotidienne, simplifiant davantage nos interactions avec la technologie et repoussant les limites de ce qui est possible.

Bibliographie

- [1] Jay ALAMMAR. *The Illustrated Transformer. The Illustrated Transformer–Jay Alammār–Visualizing Machine Learning One Concept at a Time.*
- [2] Nora ALTURAYEIF et Hamzah LUQMAN. “Fine-grained sentiment analysis of arabic covid-19 tweets using bert-based transformers and dynamically weighted loss function”. In : *Applied Sciences* 11.22 (2021), p. 10694.
- [3] Gabrielle CHENAIS et al. “Performance en classification de données textuelles des passages aux urgences des modèles BERT pour le français”. In : *Santé & IA, PFIA* (2021).
- [4] K. CORTIAL. *Fonctionnement des transformers et leurs applications.* Accessed : 04 November 2023. OpenStudio. 2022. URL : <https://www.openstudio.fr/2022/02/08/fonctionnement-des-transformers-et-leurs-applications/>.
- [5] Hugging FACE. “How do Transformers work?” In : *Hugging Face NLP Course* (2023). URL : <https://huggingface.co/learn/nlp-course/chapter1/4?fw=pt>.
- [6] Yaser KENESHLOO et al. “Deep reinforcement learning for sequence-to-sequence models”. In : *IEEE transactions on neural networks and learning systems* 31.7 (2019), p. 2469-2489.
- [7] Sayyida Tabinda KOKAB, Sohail ASGHAR et Shehneela NAZ. “Transformer-based deep learning models for the sentiment analysis of social media data”. In : *Array* 14 (2022), p. 100157.
- [8] Junyi LI et al. “Pretrained language models for text generation : A survey”. In : *arXiv preprint arXiv :2201.05273* (2022).
- [9] M. MAXIME. “What is a Transformer?” In : (2020). Accessed : 12 Octobre 2023.
- [10] MONKEYLEARN. *Sentiment Analysis Guide.* Accessed : 16 March 2023, Webography. 2023. URL : <https://monkeylearn.com/sentiment-analysis/>.
- [11] Mihai Alexandru NICULESCU, Stefan RUSSETI et Mihai DASCALU. “RoGPT2: Romanian GPT2 for text generation”. In : *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE. 2021, p. 1154-1161.
- [12] OLIVIER et H. MICHEL. *À la découverte du transformer.* Accessed : 04 November 2023. 2020. URL : <https://ledatascientist.com/a-la-decouverte-du-transformer/>.
- [13] Siddhant PORWAL, Laxmi BEWOOR et Vivek DESHPANDE. “Transformer Based Implementation for Automatic Book Summarization”. In : *arXiv preprint arXiv :2301.07057* (2023).

- [14] Yuanbin QU et al. “A text generation and prediction system : pre-training on new corpora using BERT and GPT-2”. In : *2020 IEEE 10th international conference on electronics information and emergency communication (ICEIEC)*. IEEE. 2020, p. 323-326.
- [15] Kandan RAMAKRISHNAN et al. “Mapping human visual representations by deep neural networks”. In : *Journal of Vision* 16.12 (2016), p. 373-373.
- [16] Quintus ROOS. *Fine-Tuning Pre-Trained Language Models for CEFR-Level and Keyword Conditioned Text Generation : Acomparison between Google’s T5 and OpenAI’s GPT-2*. 2022.
- [17] Mithul Sudharsan RS, PS VIGNESH, SS ASHWIN et al. “Abstractive Text Summarisation Using Keywords with Transformers Model”. In : *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*. IEEE. 2023, p. 1-9.
- [18] Uddagiri SIRISHA et Sai Chandana BOLEM. “Aspect based sentiment & emotion analysis with ROBERTa, LSTM”. In : *International Journal of Advanced Computer Science and Applications* 13.11 (2022).
- [19] Prajval SUDHIR et Varun Deshakulkarni SURESH. “Comparative study of various approaches, applications and classifiers for sentiment analysis”. In : *Global Transitions Proceedings* 2.2 (2021), p. 205-211.
- [20] Ashish VASWANI et al. “Attention is all you need”. In : *Advances in neural information processing systems* 30 (2017).
- [21] Zhilin YANG et al. “Xlnet : Generalized autoregressive pretraining for language understanding”. In : *Advances in neural information processing systems* 32 (2019).