

IFT 714
Traitement automatique des langues naturelles
Projet : Détection des fake news
Proposition de projet

Anne-Sophia Lim, Yanis Perrin, Caroline Wang

1 Introduction

Internet et les nouveaux médias ont profondément transformé la manière dont nous consommons l'information au XXI^e siècle. Les réseaux sociaux permettent à un grand nombre d'utilisateurs de partager des informations à l'échelle mondiale sans que leur véracité ne soit vérifiée, menant ainsi à de nombreux problèmes. Prenons l'exemple de la COVID-19, qui a fait l'objet de nombreuses discussions et a apporté son lot d'informations erronées, montrant l'importance de vérifier les informations qui circulent (Hamid et al., 2020). C'est pour répondre à cet enjeu que nous menons le projet d'une intelligence artificielle qui permet à partir d'un texte donné en entrée de détecter si celui-ci présente des informations erronées ou non. Cela permettrait à une personne qui doute d'un article, de vérifier sa fiabilité et ainsi d'éviter sa propagation s'il est erroné.

Le problème des fake news réside dans le fait que l'information a tendance à circuler plus vite qu'une information vraie, faisant suite à une vague de réaction plus forte de la population. Prenons pour exemple l'article de California Globe sur le port des masques (cal). Ce dernier affirmait, en se basant sur une étude du CDC, que le port du masque favorisait plutôt que prévenait l'apparition du coronavirus au début de la pandémie de Covid-19. Cette fausse information a engendré la panique et des comportements à risques avant d'avoir pu être corrigée par les responsables de l'étude, soulignant les défis liés à la propagation rapide des fake news.

Ce problème pose un réel défi à toutes personnes et entités, en particulier celles en communication publique. Les plateformes de diffusion d'informations se doivent de vérifier les contenus en développant et utilisant diverses méthodes de détection, comme des analyses textuelles ou des algorithmes utilisant des méthodes de traitement

automatique des langues, qui permet de catégoriser l'importance des différentes informations dans un texte (de Beer and Matthee, 2020) (Aldwairi and Alwahedi, 2018). Une autre méthode consiste à comparer de nouveaux textes avec d'anciennes données vérifiées dans une base de données (Atodiresei et al., 2018). Ces méthodes présentent cependant des limites dans leurs capacités d'analyse dues à la complexité du langage qui ne cesse d'évoluer. De plus, la création d'une base de données de fake news est complexe et peu sont disponibles. Enfin, ces méthodes doivent continuellement être améliorées face à la créativité des individus pour propager des fake news (Preston et al., 2021).

Ce travail a pour objectif d'implémenter puis de tester un classifieur binaire permettant, étant donné un article d'entrée, de déterminer s'il véhicule ou non des informations factices.

Pour ce faire, sera utilisée une base de données composée d'articles représentés a minima par leur titre et leur corps, et étiquetés vrais ou faux. Elle sera divisée en ensembles d'entraînement, de validation et de test, qui devront tous contenir les mêmes proportions d'articles vrais ou faux. Son contenu doit être prétraité et le texte de chaque article doit être nettoyé : les mots communément utilisés et peu représentatifs peuvent être supprimés, et ceux restant sont tokénisés et lemmisés afin de réduire le corps de l'article à un ensemble de mots dans leur forme la plus basique. Certains de ces termes peuvent toutefois demeurer peu pertinents, c'est pourquoi des méthodes de sélection de caractéristiques basées sur leur fréquence pourront être utilisées. Enfin, le modèle sera entraîné à l'aide de l'algorithme passif-agressif, qui a été révélé comme étant l'un des plus performants (T S et al., 2022). Dans la littérature, le corps des articles est utilisé. Le présent objectif est de pouvoir également inclure des informations telles que le titre ou bien l'auteur dans l'entraînement du modèle.

References

- Monther Aldwairi and Ali Alwahedi. 2018. [Detecting fake news in social media networks](#). *Procedia Computer Science*, 141:215–222. The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2018) / The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2018) / Affiliated Workshops.
- Costel-Sergiu Atodiresei, Alexandru Tănăslea, and Adrian Iftene. 2018. [Identifying fake news and fake users on twitter](#). *Procedia Computer Science*, 126:451–461. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 22nd International Conference, KES-2018, Belgrade, Serbia.
- Dylan de Beer and Machdel C. Matthee. 2020. [Approaches to identify fake news: A systematic literature review](#). *Integrated Science in Digital Age 2020*, 136:13 – 22.
- Abdullah Hamid, Nasrullah Shiekh, Naina Said, Kashif Ahmad, Asma Gul, Laiq Hassan, and Ala Al-Fuqaha. 2020. [Fake news detection in social media using graph neural networks and nlp techniques: A covid-19 use-case](#).
- Stephanie J. Preston, Anthony Anderson, David J. Robertson, Mark Shephard, and Narisong Huhe. 2021. [Detecting fake news on facebook: The role of emotional intelligence](#). *PLoS ONE*, 16.
- Steni Mol T S, P. S Sreeja, and R Prabhu Ram. 2022. [Fake news article classification using random forest, passive aggressive, and gradient boosting](#). In *2022 International Conference on Connected Systems Intelligence (CSI)*, pages 1–6.

Contribution des membres

Yanis : Recherche de propositions. Recherche sur les bases de données disponibles. Explication du contexte et de l'importance du projet.

Caroline : Recherche de propositions. Recherche sur les modèles applicables pour le projet. Explication de la méthode adoptée.

Anne-Sophia : Recherche de propositions. Recherche sur les enjeux du sujet. Explication des problématiques posées par le sujet.

Dataset

Le dataset présentée pour notre étude a été trouvé via la task3 : Fake News Detection du CLEF2021, c'est un dataset disponible sur le site internet Kaggle à l'adresse suivante : [https://www.kaggle.com/datasets/](https://www.kaggle.com/datasets/liberoliber/onion-notonion-datasets)

[liberoliber/onion-notonion-datasets](https://www.kaggle.com/datasets/liberoliber/onion-notonion-datasets). Il présente l'ensemble des caractéristiques que nous avons besoin, à savoir : des articles labélisés comme vrai et d'autres comme faux. On dénombre 21417 observations d'articles labélisé comme vrai et 23481 observations d'articles labélisé comme faux pour un nombre total de 44898 observations. Dans les informations que nous disposons, nous avons : le titre de l'article, le texte de l'article, le sujet général de l'article, la date à laquelle a été écrit l'article et enfin le label associé à l'article (vrai ou faux).

Ressources Informatiques

Dans notre projet, python sera le langage de programmation utilisé, celui-ci est le plus utilisé et conseillé dans la manipulation des données. Nous ferons appel à plusieurs librairies python dans ce projet (Numpy, Pandas, PyTorch). Numpy et Pandas nous facilitent la manipulation des données grâce aux diverses méthodes qu'elles présentent. PyTorch nous permettra de créer, d'entraîner et de tester nos modèles. Enfin, nous utilisons Jupyter Notebook qui est une plate-forme informatique interactive basée sur le Web qui nous permet de tester notre code via des cellules de codes. Github sera utilisé pour le partage et la gestion des versions du projet. Le jeu de données étant de taille raisonnable, nous devrions pouvoir faire tourner notre projet sans problème sur nos ordinateurs personnels.