

M1 Informatique – UE Projet

Carnet de bord : les coulisses de la recherche documentaire

Noms, prénoms et spécialité :

OUHENIA Ghiles, DAC
SADI Amayas, DAC
RODRIGUES DA ROCHA Mathilde, DAC

Sujet :

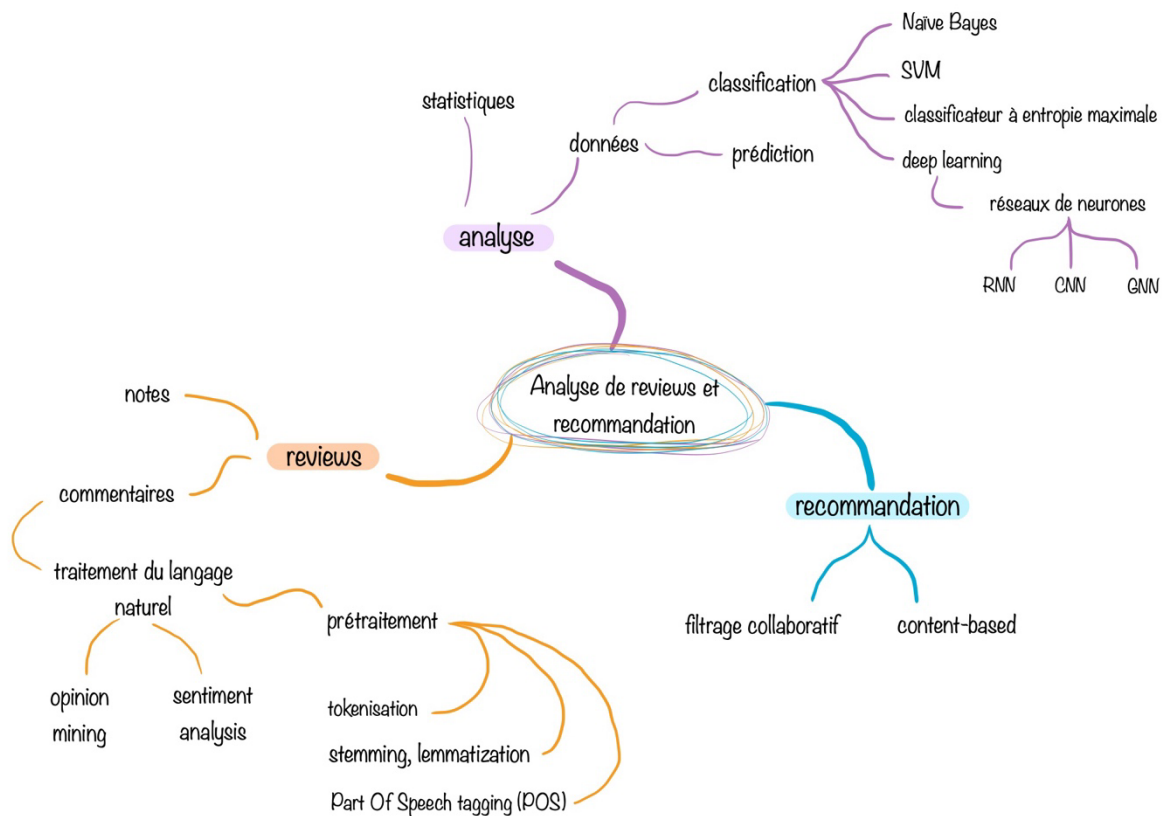
Analyse de Reviews et Recommandation

INTRODUCTION

Notre projet, intitulé « Analyse de reviews et recommandations », vise à exploiter les données provenant d'un site internet recensant un grand nombre de jeux et d'avis associés, rédigés par des utilisateurs. En utilisant des techniques de Traitement du Langage Naturel (NLP), notre objectif est de classer les sentiments des commentaires pour déterminer s'ils sont plutôt positifs ou négatifs. Nous souhaitons également développer un modèle de recommandation de jeux, basé sur les préférences des utilisateurs et des jeux similaires appréciés par d'autres joueurs partageant les mêmes intérêts. Les plateformes d'opinions en ligne sont extrêmement précieuses car elles permettent aux utilisateurs de partager leurs expériences et leurs opinions sur des produits ou des services. L'analyse de sentiments est donc devenue essentielle pour extraire des informations pertinentes et mesurer le ressenti des clients envers une marque ou un produit. Elle permet également de répondre à de nombreux défis, tels que l'amélioration de l'expérience utilisateur, la gestion de la réputation en ligne, la détection des tendances du marché et l'identification des insatisfactions et des problèmes de produits. De même, la recommandation de produits basée sur les préférences des utilisateurs est un enjeu majeur pour de nombreuses entreprises de commerce en ligne. Les modèles de recommandation doivent être capables de prédire les préférences des utilisateurs en fonction de leur historique de navigation et de leurs comportements d'achat, en utilisant des techniques de traitement du langage naturel et de recommandation.

Dans notre projet, nous utiliserons des techniques avancées de NLP telles que l'analyse de sentiments et l'exploration d'opinion pour extraire des informations sémantiques à partir des commentaires des utilisateurs et des techniques de recommandation pour proposer des jeux pertinents aux utilisateurs.

MOTS CLES RETENUS



DESSCRIPTIF DE LA RECHERCHE DOCUMENTAIRE

Pour la recherche documentaire nous avons cherché les mots clés les plus pertinents sur des sites tels que Web of Science ou Science Direct. En effet, c'étaient pour nous les outils de recherche les plus appropriées, nous avons pu trier efficacement les articles notamment en fonction de la date de publication, un critère plutôt important dans notre recherche afin de construire un état de l'art sur les méthodes les plus récentes et de faire ressortir les possibles critiques et améliorations des méthodes un peu plus anciennes. A partir des quelques articles qui ont été sélectionnés dans un premier temps, nous avons pu consulter les articles cités afin d'enrichir notre recherche.

Les articles les plus pertinents à ce stade de notre recherche sont surtout les revues car elles permettent d'avoir une vue d'ensemble sur le sujet en comparant les différents outils de Traitement Automatique du Langage et leur efficacité, en faisant part des améliorations que certains chercheurs ont proposé. Les articles que nous avons sélectionnés ont tous été rédigés par des enseignants chercheurs. Leur niveau de spécialisation est donc très élevé, ce qui fait de ces sources des sources fiables.

Afin de diversifier nos sources d'informations, nous avons consulté des revues scientifiques qui examinent des articles présentant des méthodes utilisées dans le domaine du NLP, des livres qui font un état de l'art complet sur le sujet et détaillent le fonctionnement de toutes les approches, ainsi que des articles scientifiques présentant les résultats d'expérimentations menées avec différents outils.

BIBLIOGRAPHIE PRODUITE DANS LE CADRE DE LA RECHERCHE

- [1] Mohammed Fadhel Aljunid and Manjaiah Dh. 2020. An Efficient Deep Learning Approach for Collaborative Filtering Recommender System. *Procedia Computer Science* 171, (January 2020), 829–836. DOI:<https://doi.org/10.1016/j.procs.2020.04.090>
- [2] Negin Ghasemi and Saeedeh Momtazi. 2021. Neural text similarity of user reviews for improving collaborative filtering recommender systems. *Electronic Commerce Research and Applications* 45, (January 2021), 101019. DOI:<https://doi.org/10.1016/j.elerap.2020.101019>
- [3] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, Association for Computational Linguistics, 79–86. DOI:<https://doi.org/10.3115/1118693.1118704>
- [4] Krupa Patel and Hiren B. Patel. 2020. A state-of-the-art survey on recommendation system and prospective extensions. *Computers and Electronics in Agriculture* 178, (November 2020), 105779. DOI:<https://doi.org/10.1016/j.compag.2020.105779>
- [5] Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor (Eds.). 2011. *Recommender Systems Handbook*. Springer US, Boston, MA. DOI:<https://doi.org/10.1007/978-0-387-85820-3>
- [6] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, ACM, Hong Kong Hong Kong, 285–295. DOI:<https://doi.org/10.1145/371920.372071>
- [7] Sunny Sharma, Vijay Rana, and Vivek Kumar. 2021. Deep learning based semantic personalized recommendation system. *International Journal of Information Management Data Insights* 1, 2 (November 2021), 100028. DOI:<https://doi.org/10.1016/j.jjime.2021.100028>
- [8] Yasir Ali Solangi, Zulfiqar Ali Solangi, Samreen Aarain, Amna Abro, Ghulam Ali Mallah, and Asadullah Shah. 2018. Review on Natural Language Processing (NLP) and Its Toolkits for Opinion Mining and Sentiment Analysis. In *2018 5th IEEE International Conference on Engineering Technologies and Applied Sciences (IEEE Ictetas)*, IEEE, New York. Retrieved February 16, 2023 from <https://www.webofscience.com/wos/woscc/full-record/WOS:000460476900053>
- [9] Shiliang Sun, Chen Luo, and Junyu Chen. 2017. A review of natural language processing techniques for opinion mining systems. *Information Fusion* 36, (July 2017), 10–25. DOI:<https://doi.org/10.1016/j.inffus.2016.10.004>
- [10] Peter D. Turney. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. DOI:<https://doi.org/10.48550/arXiv.cs/0212032>
- [11] Jingjing Wang, Haoran Xie, Fu Lee Wang, Lap-Kei Lee, and Oliver Tat Sheung Au. 2021. Top-N personalized recommendation with graph neural networks in MOOCs. *Computers and Education: Artificial Intelligence* 2, (January 2021), 100010. DOI:<https://doi.org/10.1016/j.caeai.2021.100010>
- [12] Chenyan Zhang, Shan Xue, Jing Li, Jia Wu, Bo Du, Donghua Liu, and Jun Chang. 2023. Multi-Aspect enhanced Graph Neural Networks for recommendation. *Neural Networks* 157, (January 2023), 90–102. DOI:<https://doi.org/10.1016/j.neunet.2022.10.001>

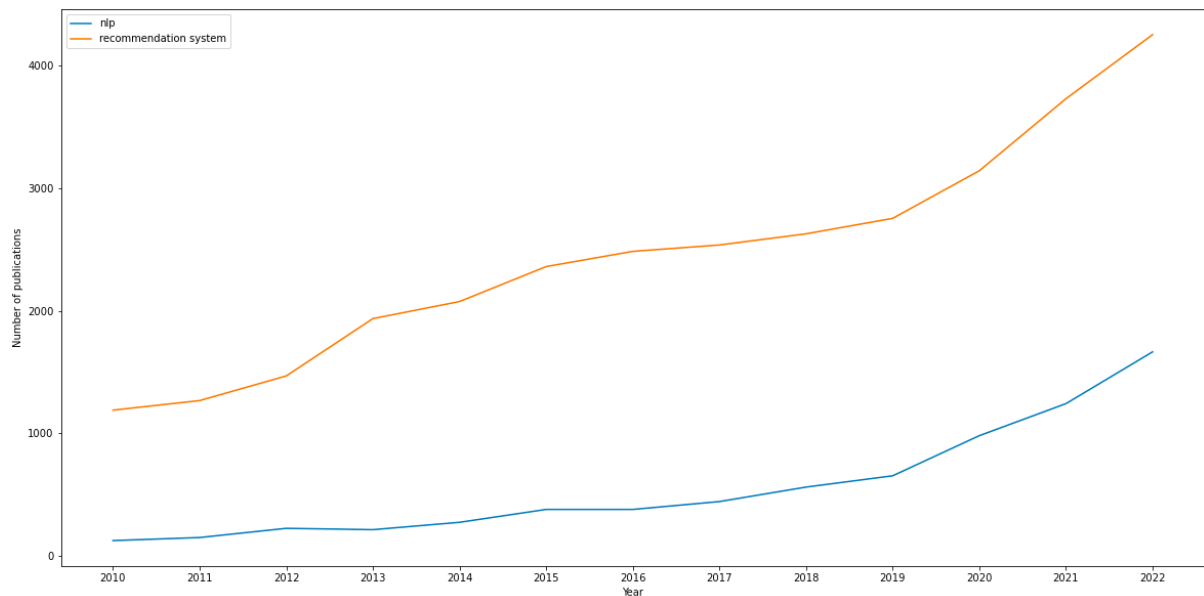
EVALUATION DES SOURCES

L'article de Sun et al. 2017 [9] se trouve être une revue très complète pour notre sujet, du moins concernant l'analyse de sentiment. Nous l'avons trouvée dans les citations d'un article trouvé sur site Web of Science avec les mots clefs suivants : « Review Analysis NLP ». Bien que cet article ait presque 6 ans, il répertorie les modèles de classification les plus couramment utilisés, les différentes approches en expliquant succinctement leur fonctionnement et les améliorations que certains chercheurs ont tenté d'apporter, les techniques de prétraitement ainsi que les outils disponibles pour différents langages de programmation avec une ouverture sur les méthodes de Deep Learning. Cependant, l'article est trop ancien pour un état de l'art sur les méthodes de Deep Learning en NLP, c'est pourquoi nous poursuivons nos recherches sur des articles plus récents.

Le livre de Ricci et al., 2011 [5], « Recommender Systems Handbook » nous a été conseillé par notre enseignant. Il s'agit du premier ouvrage complet entièrement consacré au domaine des systèmes de recommandation et couvrant plusieurs aspects des principales techniques. Le livre décrit en détail les méthodes classiques, ainsi que les extensions et les « nouvelles » approches, du moins qui l'étaient lorsque cet ouvrage fut publié en 2011. Ce que nous apprécions particulièrement dans cet état de l'art, rédigé par des chercheurs d'universités provenant des quatre coins du monde, c'est que toutes les perspectives possibles sont présentées, les différentes méthodes sont détaillées d'un point de vue mathématique, ce qui nous confère de bonnes lignes directrices pour l'implémentation. Cependant, encore une fois, cet article est trop ancien et n'aborde donc pas le sujet du Deep Learning.

La source suivante, trouvée sur science direct en utilisant les mots clefs « recommendation Graph Neural Networks », est un article de Wang et al., 2021 [11] propose de répondre au problème de recommandation d'items en utilisant les Graph Neural Networks, une méthode de Deep Learning émergente. Ils appliquent ce problème à la recommandation de cours en ligne en comparant leurs méthodes avec les méthodes traditionnelles de système de recommandation, les méthodes de Deep Neural Network et les méthodes à base de données structurées en graphe. Leur modèle bat tous les autres, ce qui nous conforte dans l'idée de poursuivre sur cette piste.

BONUS – Évolution du nombre de publications



Nous souhaitons avoir un aperçu de l'évolution du nombre de publications liées à notre sujet de recherche. Pour ce faire, nous avons récupéré ces données sur Science Direct en utilisant les mots clefs « nlp » et « recommendation system », en ajoutant un filtre pour ne cibler que les publications dans le domaine de l'informatique. Les domaines du traitement automatique du langage et des systèmes de recommandation ont connu un essor important à partir de 2019. Cette date marque également le début de la curiosité autour des méthodes de Deep Learning.