

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

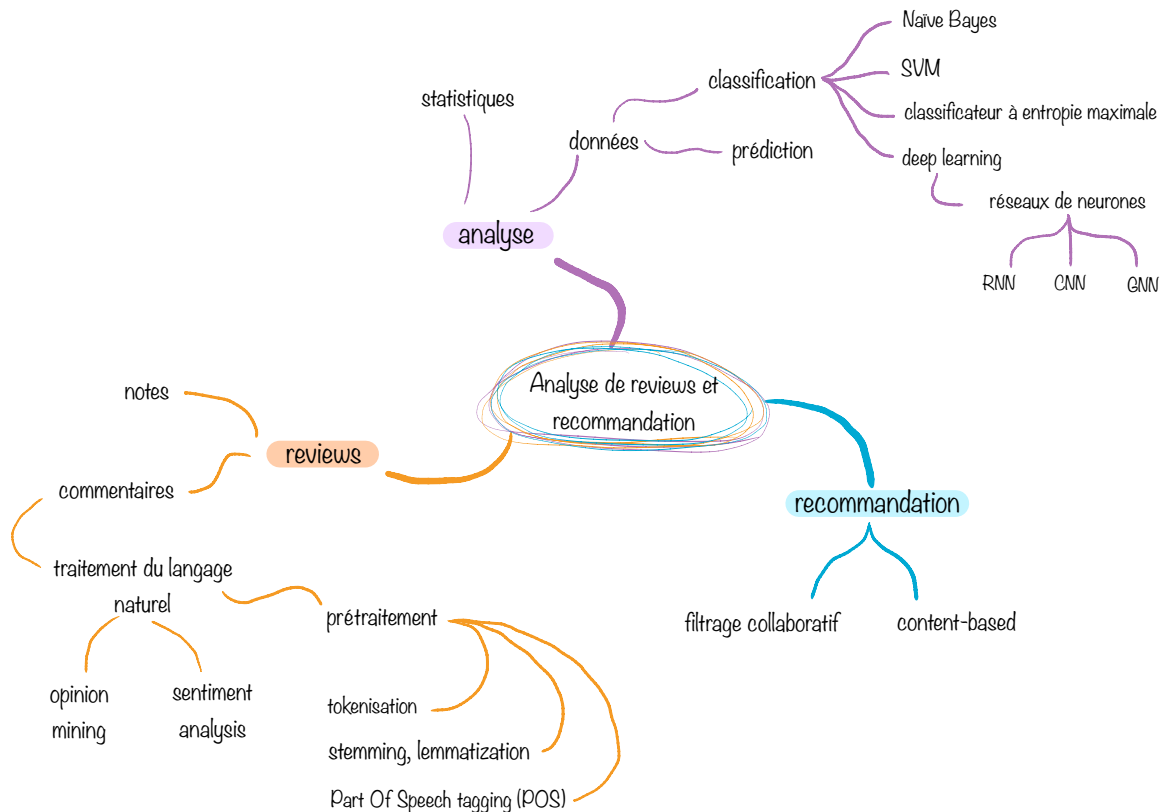
Sujet :

Analyse de Reviews et Recommandation

INTRODUCTION

Notre projet s'intitule « Analyse de reviews et recommandations ». A partir de données provenant d'un site internet qui regroupe grand nombre de jeux et d'avis associés, rédigés par des utilisateurs, ce projet a pour but de faire de la classification de sentiments, c'est-à-dire de déterminer si un commentaire est à valence plutôt positive ou négative et dans un second temps, de faire la recommandation, c'est-à-dire de développer un modèle capable de recommander un jeu sur la base des jeux qu'il a apprécié et que les autres joueurs partageant les mêmes intérêts ont apprécié aussi. Les plateformes permettant aux utilisateurs de partager leurs opinions concernant un produit ou un service sont des denrées très précieuses. Ces opinions véhiculent des émotions et des sentiments, il est donc pertinent aujourd'hui de s'intéresser à l'exploration d'opinion, ou opinion mining et à l'analyse de sentiments afin de répondre à un grand nombre de problèmes, notamment en développant des outils de recommandation. Pour cela nous utiliserons des techniques de NLP, soit de Traitement du Langage Naturel et des techniques de recommandation.

MOTS CLES RETENUS



DESCRIPTIF 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.

Nous avons tenté de faire varier les sources en consultant des articles passant en revue les méthodes utilisées dans le domaine du NLP, des livres faisant un état de l'art complet sur le sujet et détaillant le fonctionnement de tous les approches et des articles scientifiques présentant les résultats d'une expérimentation menée 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] 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>
- [4] 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>
- [5] 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>
- [6] 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>

EVALUATION DES SOURCES

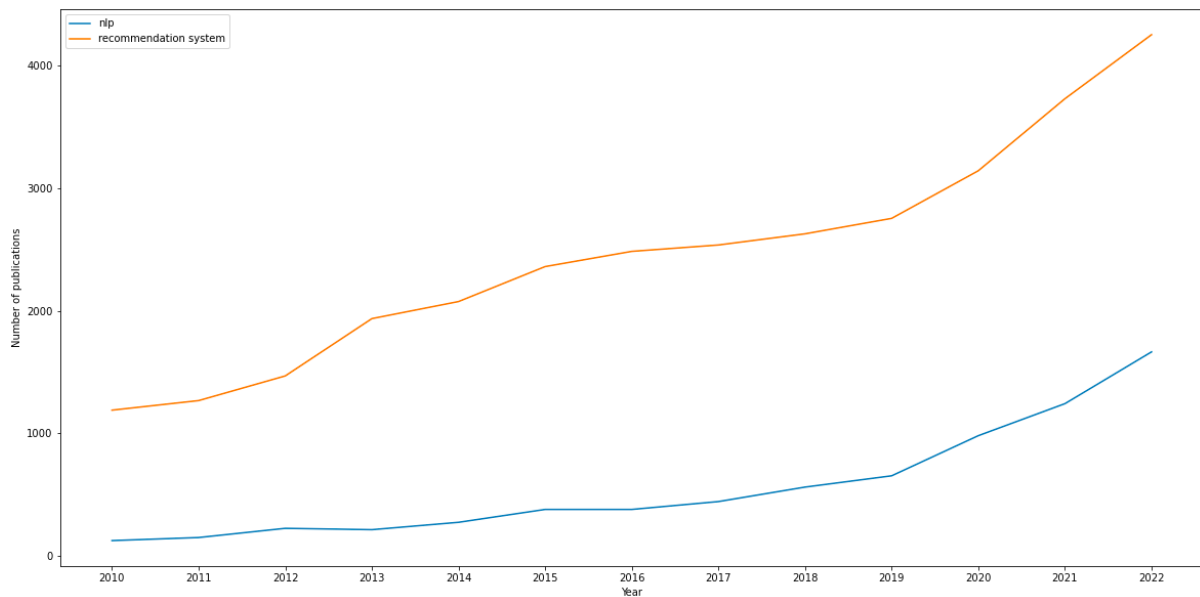
L'article de Sun et al. 2017 [3] se trouve être une revue très complète pour notre sujet, du moins concernant l'analyse de sentiment. Nous l'avons trouvée dans les citations d'un article trouvé sur site Web of Science avec les mots clefs suivants : « Review Analysis NLP ». Bien que cet article ait presque 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 [1], « 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 « recommandation Graph Neural Networks », est un article de Wang et al., 2021 [4] 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 publication



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.