

# M1 Informatique – UE Projet

## Carnet de bord : les coulisses de la recherche documentaire

*Les éléments que vous indiquez dans ce carnet donneront lieu à une notation*

Noms, prénoms et spécialité :

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

Sujet :

Analyse de Reviews et Recommandation
--------------------------------------

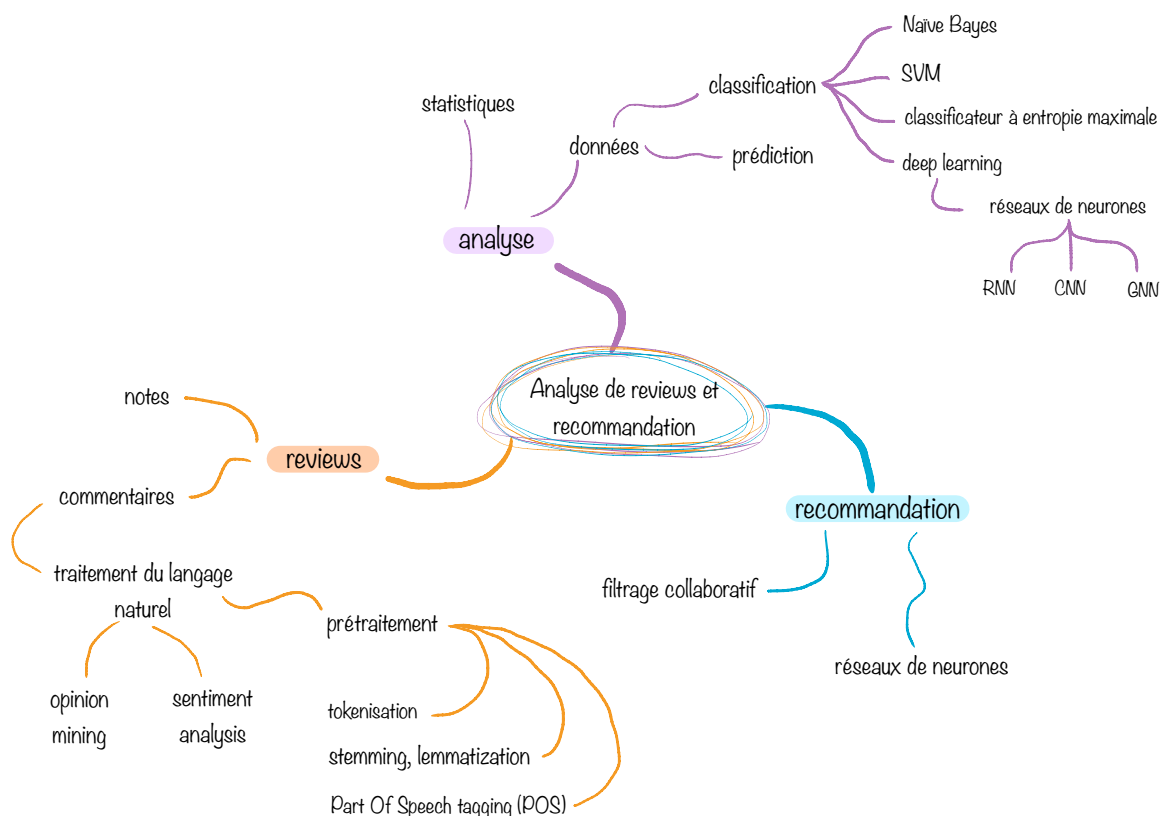
Consigne :

1. **Introduction (5-10 lignes max)** : Décrivez rapidement votre sujet de recherche, ses différents aspects et enjeux, ainsi que l'angle sous lequel vous avez décidé de le traiter.
2. **Les mots clés retenus** : Listez les mots-clés que vous avez utilisés pour votre recherche bibliographique. Organisez-les sous forme de carte heuristique.
3. **Descriptif de la recherche documentaire (10-15 lignes)** : Décrivez votre utilisation des différents outils de recherche (moteurs de recherche, base de donnée, catalogues, recherche par rebond etc.). Comparez ensuite les outils entre eux. A quelles sources vous ont-ils permis d'accéder ? Quelles sont leurs spécificités ? Quel est leur niveau de spécialisation ?
4. **Bibliographie produite dans le cadre du projet** : Utilisez la norme ACM.
5. **Evaluation des sources (5 lignes minimum par source)** : Choisissez 3 sources parmi votre bibliographie, décrivez la manière dont vous les avez trouvées et faites-en une évaluation critique en utilisant les critères vus sur les supports de TDs.

## 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, sur la base de ce qu'il aura appris, de recommander un jeu en fonction de plusieurs critères que nous définirons plus tard. 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.

## MOTS CLES RETENUS



Conditional Random Field (CRF) ?

D'autres mots clés ? CRF ?

## **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 du Langage Naturel 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 provenant d'universités Nord-Américaines ou Chinoises. Le niveau de spécialisation est donc très élevé, ce qui fait de ces sources des sources fiables.

## **BIBLIOGRAPHIE PRODUITE DANS LE CADRE DE LA RECHERCHE**

- [1] 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>

## **EVALUATION DES SOURCES**

L'article de Sun et al. 2017 [1] 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.

TODO : ajouter encore deux articles, de préférence plus récents et pourquoi pas qui parlent de Deep (GNN ?), de recommandation (filtrage collaboratif ?)...