

Sujet PLDAC : Analyse de reviews et recommandation

Encadrants : Vincent Guigue et Nicolas Baskiotis (`prenom.nom@sorbonne-universite.fr`)

Binôme ou trinôme

A partir d'une base de données de reviews récoltée sur un site spécialisé dans les jeux de société, nous souhaitons analyser les reviews et principalement les sentiments exprimés. La base contient 150 000 avis, 200 000 notes sur plus de 15 000 jeux.

Ce PLDAC vous permettra de travailler sur des données textuelles et de mettre en œuvre d'une part des stratégies de classification de sentiments et d'autre part d'aborder la problématique de la recommandation. Selon l'avancement, il sera également possible d'aborder des approches neuronales modernes pour le traitement de graphe comme les Graph Neural Networks (GNN).

Le plan de travail est le suivant :

- Prise en main des données textuelles
- Construction d'un classifieur de sentiment sur des bases de revues et d'autres sources étiquetées
- Problématique de transfert du classifieur vers les données twitter
- Etude d'un système de recommandation en fonction des notes données par les utilisateurs et les reviews
- Etude de l'apport des GNN dans ce cadre.