



## Résumé

De nos jours, de plus en plus de personnes se tiennent au courant des nouvelles informations en streaming sur le service de micro-blogging Twitter. Compte tenu du grand nombre de micro-messages partagés via le graphique du réseau de microblogging, les utilisateurs de Twitter se retrouvent face à un nombre immense de tweets, ce qui rend la recherche d'informations appropriées, et la recherche de personnes intéressantes qui sont dans le même domaine d'intérêt une tâche difficile.

Grâce en grande partie au volume et à la modernité sans cesse croissante des données sur Internet, nous sommes en mesure d'accéder à une somme colossale de données du monde entier. Aujourd'hui, le principal défi pour les utilisateurs est de trouver des informations pertinentes à partir d'une source presque interminable et de rechercher des personnes qui se trouvent dans le même domaine d'intérêt. Ce problème a conduit à l'évolution des systèmes de recommandation qui montrent de manière proactive aux utilisateurs des informations liées à leurs intérêts.

Dans cette thèse, trois systèmes de recommandation de followee sont créés. Le premier exploite la méthode d'analyse sémantique des tweets basée sur la solution Babelfy de Babelnet, utilise également le graphe de relation follower/followee pour extraire les candidats. Ensuite, nous avons mis en place un système de recommandation de followee qui améliore le système basé sur la sémantique en extrayant et en analysant les sujets des tweets. Pour les deux approches, le classement des meilleurs utilisateurs suggérés a été effectué en utilisant le TF-IDF.

Le troisième système de recommandation de followee est un système qui utilise des algorithmes de réseau de neuron. Alors qu'il utilise le modèle LDA comme stratégie de modélisation de topiques et le modèle LSTM pour prédire les candidats possibles pour la recommandation, nous avons classé les candidats proposés en fonction de leur popularité.

**Key Words:** Recommendation system, LDA, BERT, ELECTRA, LSTM, Followee, Twitter.

# MICROBLOGGING RECOMMENDER SYSTEMS BASED ON LANGUAGE REPRESENTATION LEARNING, TOPIC MODELS, AND DEEP LEARNING

## Abstract

Nowadays, more people keep up with new streaming information on Twitter micro-blogging service. Considering the huge number of micro-posts shared via the microblogging network graph, Twitter users find themselves in front of many tweets, making seeking appropriate information and searching for interesting people in the same field of interest a challenging task.

Owing to a great extent to the ever-increasing volume and modernity of data on the internet, we can get to a colossal sum of data from around the globe. Nowadays, the critical challenge is for the users to find pertinent information based on a nearly interminable source, and search for people within the same field of interest. This issue has led to the evolution of recommender systems that proactively show users information related to their interests.

In this thesis, three followee recommendation systems are created. The first one leverages the tweets semantics analysis method based on Babelnet solution “Babelfy”, which uses the follower/followee relationship graph to extract candidates.

Then we implemented a followee recommendation system that improves the semantic-based system by extracting and analyzing the tweets topics. For both approaches, ranking the best-suggested users was by utilizing the TF-IDF.

The third followee recommender framework is a system that employs neural network algorithms. Whereas it uses the LDA model as a topic modeling strategy and the LSTM model to predict recommended candidates we ranked the proposed candidates based on their popularity.

**Key Words:** Recommendation system, LDA, BERT, ELECTRA, LSTM, Followee, Twitter.