

Notification No: 550/2023

Date of award: 12-12-2023

Name of the Scholar: Surbhi Kakar

Name of the Supervisor: Prof. Monica Mehrotra

Name of the Department/Centre: Computer Science

Topic of research: Information diffusion in online social networks

Information Diffusion is the passing of information from one node to another in a social network. This research studies retweet diffusion specifically under the realm of information diffusion domain.

Retweet mechanism in the light of information diffusion has several applications in multiple domains ranging from digital marketing, emergency control to fake news identification.

One of the key challenges is the abundance of information. The sheer volume of tweets makes it difficult for users to filter through the noise and identify relevant and reliable information. Additionally, the real-time nature of the platform can result in the rapid spread of misinformation, rumors, and fake news.

Understanding retweet behaviours and the virality of tweets on social media can help us identify influential people who can spread the information at a faster pace. This insight is useful in applications such as viral marketing and emergency response.

Predicting which tweets will be retweeted by a user can also help in providing recommendations to a user to create a personalized experience for them.

This research provides a comprehensive review on the diverse research areas under the realm of information diffusion.

It also identifies value-systems in addition to conventional features used in the literature as a potential feature-set that can improve the performance of retweet prediction models. To accomplish the former, our research also proposes a ValueDiet lexicon in order to label people with their value systems.

Furthermore, this research also proposes a Clustered Bert Model for predicting the rate at which influence diffuses in a social network. The research demonstrates that the Clustered

Bert Model outperforms the traditional models based on numerical, emotions and sentiment features.

Lastly, our research proposes a hybrid recommendation algorithm to solve the problem of information overload in a social network. The research showed that the recommendation algorithm based on Clustered Bert Model outperformed the one based on word2vec model.

The proposed work in this thesis focuses primarily on retweet prediction within the realm of information diffusion. However, it is important to note that other areas of study within information diffusion are equally significant. Exploring these areas would contribute to a more comprehensive understanding of information diffusion dynamics and their applications.

Enhancing the ValueDict lexicon, specifically in the Self Enhancement category, can lead to more accurate and comprehensive value-based analyses. Refining this lexicon would enable researchers to gain a deeper understanding of user behavior and value systems on Twitter, enabling more precise and detailed investigations in future studies.

Investigating how value systems influence retweet popularity and leveraging them in the recommendation process could enhance the accuracy and effectiveness of these models. This represents an important direction for future research in the field of retweet prediction and recommendation algorithms.

One direction for future work could involve using language models such as BERT instead of solely relying on text embeddings generated from the model. Leveraging the power of language models can potentially provide more context-rich and nuanced features, leading to improved prediction performance.

Finally, traditional model improvements such as refining the feature set should also be considered as a direction for future work. Continuously enhancing the feature set used in prediction models can contribute to increased prediction accuracy and reliability.