

Social Network Analysis for Precise Friend Suggestion for Twitter by Graph Embedding Using Node2vec

Md.Shahriar Hossain Apu (1901036),Mehedi Hasan Emon (1901031),Mahjabin Ashka(1901048),Arifuzzaman Ripon(1901038)

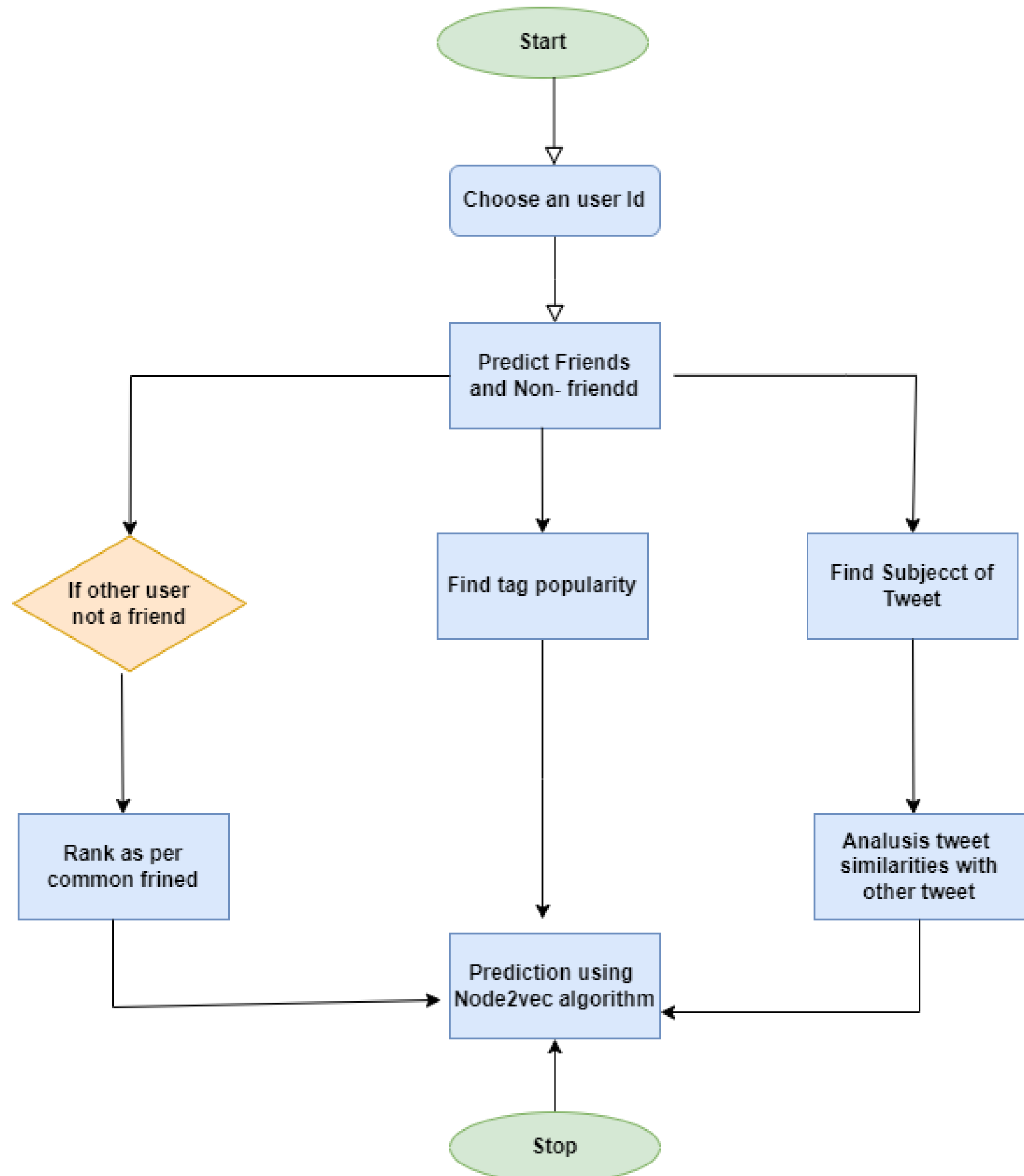
Department of IRE,Bangabandhu Sheikh Mujibur Rahman Digital University



Abstract

The main aim of this paper is to create a friend suggestion algorithm that can recommend new friends to a user on Twitter based on their existing friends and other relevant details. The information used for making these predictions includes the user’s friends, tags, tweets, language spoken, ID, and more. The authors employed supervised learning methods to train their models, specifically using the Node2Vec approach. Node2Vec is a machine learning technique used to learn embeddings of nodes in a network, which can capture the structural characteristics of the social graph. By using Node2Vec, the authors sought to capture the nuanced relationships between users and their friends.The approach is user- centric, focusing on a single user at a time. For each user, their features are compared to those of other users, particularly their non-friends, to suggest potential new friends. This approach takes into account the uniqueness of a user’s social network and community, making personalized friend recommendations based on the learned node embeddings. i

SYSTEM OVERVIEW



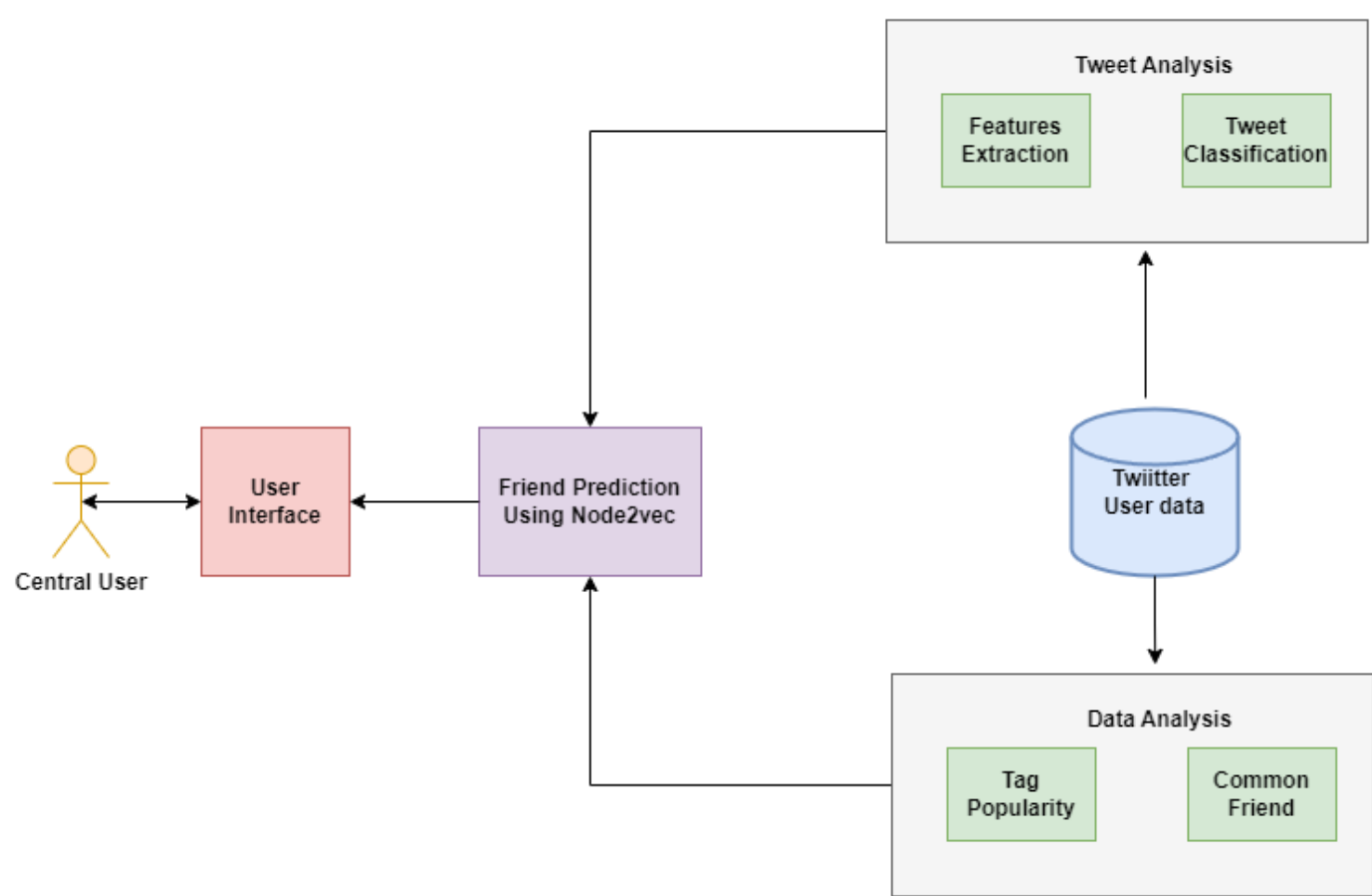
Description of the Dataset

The dataset consists of social media user information and activity attributes, likely from a platform like Twitter. It includes data such as user IDs, screen names, hashtags (tags), profile avatars, follower and friend counts, language preferences, and timestamps indicating when the user was last seen. Additionally, the dataset includes unique tweet IDs and a list of friends (users followed) by each individual. This data is valuable for studying user behavior, network analysis, and content trends on a social media platform.

Table 1. Dataset

Description	Number of users	4000
	Number of distinct tags	483
	Avg. number of followers per user	390 Avg. number of friends per user

Methodology



Node2vec Algorithm

```
node2vecWalk (Graph G' = (V, E, π), Start node u, Length l):  
  
    Initialize walk to [u]  
    for (walk_iter = 1 to l):  
        curr = walk[-1]  
        Vcurr = GetNeighbors(curr, G')  
        s = AliasSample(Vcurr, π)  
        Append s to walk  
    return walk  
  
Learn_Features (Graph G = (V, E, W)):  
  
    Dimensions -> d  
    Walks per node -> r  
    Walk length -> l  
    Context size -> k  
    Return -> p  
    In-out -> q  
  
    π = PreprocessModifiedWeights (G, p, q)  
    G' = (V, E, π)  
  
    Initialize walks = ∅  
  
    for (iter = 1 to r):  
        for (all nodes u ∈ V):  
            walk = node2vecWalk(G', u, l)  
            Append walk to walks  
  
    f = StochasticGradientDescent(k, d, walks)  
    return f
```

Figure 1. Sudo code of how Node2vec Perfrom

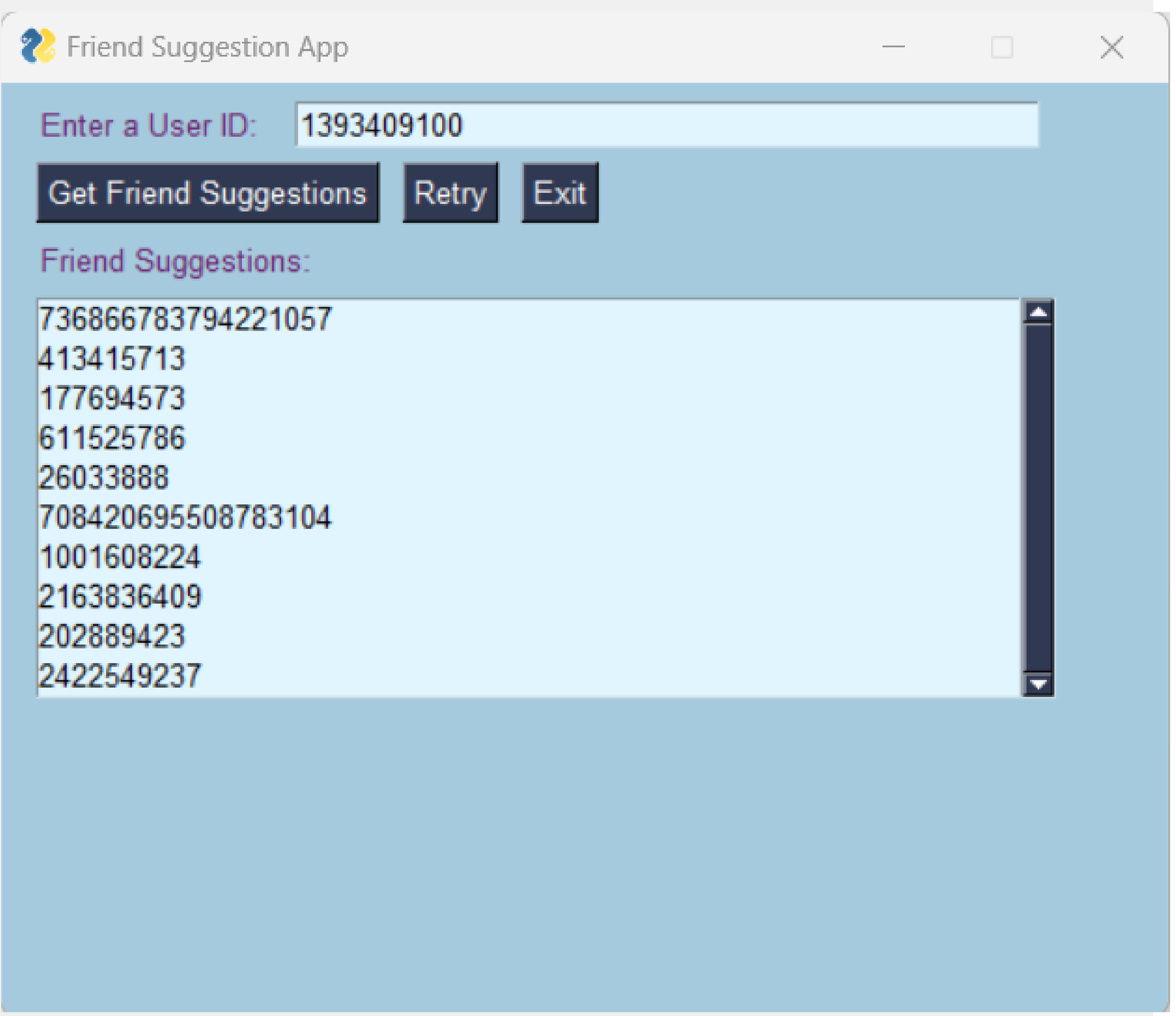
Result and Discussion

Table 2. Model Accuracy

Model Name	Accuracy %
Node2vec	90%
Word2Vec	84%
KNN	80%

System Implementation

User Interface



Conclusion

The research focuses on enhancing the friend suggestion algorithm for Twitter by considering factors beyond mutual friends. It acknowledges that common friends alone don't necessarily imply shared interests and introduces criteria like tweet similarity and tag popularity. By incorporating natural language processing techniques, the algorithm aims to provide more efficient and comprehensive friend recommendations. Additionally, the study attempts to improve the identification of a user's real-life social community based on their Twitter network structure. The proposed algorithm involves three factors and a novel approach to measuring user closeness, ultimately striving to offer a more refined and accurate friend recommendation system on the platform.

References

[1] Tajfar A. H. Nejad A. F. Akbari, F. Graph-based friend recommendation in social networks using artificial bee colony. In Dependable, Autonomic and Secure Computing. DASC., 27(3):379–423, 2013.

[2] Zhang J. Wang L. Hua X. S. Huang, S. Social friend recommendation based on multiple network correlation. IEEE Transactions on Multimedia, 18(2):287–299, 2016.

[3] Song X. Finin T. Tseng B. Java, A. Why we twitter: understanding microblogging usage and communities. Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis.