A Hybrid Contextual Probabilistic Graph Clustering and Link Prediction Model for Complex Social Networking Data

Rajasekhar Nennuri [1], S. Iwin Thanakumar Joseph [2], B. Mohammed Ismail [3], L.V. Narasimha Prasad [4]
[1] KL University, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India.
[2] Assistant Professor, Koneru Lakshmaiah Education Foundation, Guntur District, Andhra Pradesh.
[3] KL University, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India.
[4] Professor, Institute of Aeronautical Engineering, Hyderabad.
[1] rajasekahrrnennuri@gmail.com, [2] iwineee2006@kluniversity.in, [3] aboutismail@gmail.com, [4] lvnprasad@yahoo.com

Article Info
Page Number: 978-999
Publication Issue:
Vol. 71 No. 3s (2022)

Article History
Article Received: 22 April 2022
Revised: 10 May 2022
Accepted: 15 June 2022
Publication: 19 July 2022

Abstract

Probabilistic based graph community detection plays a vital role in the complex social networking datasets. Since, most of the conventional approaches are difficult to predict the new type of link prediction using the standard graph community clustering measures. Also, traditional clustering measures use nearest neighbour measures instead of contextual similarity in order to predict the relationship among the different graph nodes. In order to optimize the contextual node clustering and link prediction, a hybrid dynamic scalable measure is proposed for the community clustering on complex networks. In this work, a hybrid graph clustering and link prediction approaches are proposed on the complex social networking dataset for better decision making patterns. Experimental results prove that the proposed contextual probabilistic graph clustering and link prediction approach has better efficiency than then conventional models on complex social networking datasets.

Keywords: Social network dataset, link prediction, community detection.

1. Introduction

Link prediction anticipates the formation of complex or social links in a network over time. Understanding these connections will aid in understanding the network's evolution over time. It not only aids in the understanding of network evolution, but also in the study of these networks using various parameters. Anthropologists, social scientists, biologists, computer scientists, mathematicians, and researchers from a variety of disciplines have studied complex and social networks to gain a better understanding of them and have proposed new models for link prediction based on topological features of the network graph. In recent years, online social networks have grown in popularity to the point where they have become an inseparable part of our daily lives. The rise in popularity of online social networks was aided by the increased use of mobile devices. Online social networks like Facebook, LinkedIn, Twitter, and Google+ have become increasingly important in our daily lives. The linked formation of social elements, in which social objects such as people, groups, and organizations are represented as nodes of a graph, is known as a social network. The links between the nodes represent the relationship between them [1]. Links are also known as edges, arcs, ties, or bonds, and nodes are also known as actors, points, or vertices. Individuals in a network can have a "non-directed" or "undirected" relationship if the individual I is connected to individual j, implying that j is connected to i. [2]. A directed relationship between individuals in a network is one in which just because individual I is connected to individual j does not mean that j is connected to i. The actor-actor relationship can be directed, as in Facebook or LinkedIn, or undirected, as in Twitter. The majority of research in online social networks is based on network science and graph theory fundamentals. Network science is a broad field that encompasses everything from psychology to biology to social sciences, and it has played a significant role in shaping the research aspects of online social networks. With the rise in popularity of online social networks, new concerns have emerged, such as data privacy and the sensitivity of shared information. Homophily is a feature of online social networks in which like-minded people and actors with similar viewpoints tend to stick together [3][4]. [5] coined the phrase "love of the same" to describe it. Because of their widespread use, online social networks have provided an opportunity to study and research their characteristics on a large scale. The mapping, measuring, and study of relationships flows and social interaction between people, groups, and organizations is known as online social network analysis, and it encompasses a wide range of fields of study, from psychology to biology to information sciences. Human relationships are visualized and mathematically analyzed using social network analysis. It is a methodology developed in the early 1960s by sociologists and social psychology researchers [7]. In the last two decades or so, social network analysis has exploded in popularity, primarily in the fields of sociology and communication science. There are over 50 different social network measures that are commonly used in social network analysis. Finding and evaluating the perfect measure is a problem that has attracted a lot of research attention. Because they connect people all over the world and reach billions of people, real-world social networks are massive. Because the database is so large, new methods for organizing it and managing the relationships between the network's participants are required [8]. The term "social network" refers to the building structure of a group of communities that have joined together and can be developed as groups within which social interactions among actors are both intense and weak [9]. When the relationships between the network's participants become denser, the network becomes a community structure, which is also known as clustering. When the relationships within the network are dense, the entities that are individuals involved in the network automatically fall into the cluster or community. As they describe the interaction and relationship between individuals or entities, social networks can also be considered information networks. In research topics such as sociology, epidemiology, recommendation systems, and criminology, social network analysis aids researchers in discovering methods in social structures and identifying patterns of interaction among entities. Information was collected from social networks using surveys, questionnaires, and personal interviews in the early days of social network analysis. Link prediction predicts links in the network either complex or social that may form over a period of time. Understanding such links will help to understand the evolution of network over a period of time. It not only helps to understand network evolution but also helps to study these networks with different parameters. Anthropologists, social scientists, biologists, computer scientists, mathematicians and researchers from many domains have studied the complex and social networks to understand different perspectives and proposed new models for the link prediction using topological features of the network graph. By predicting links, some of the missing links in the network that might have lost in the due course of network evolution can be identified. Due to war or other natural disasters, human beings (belonging to different races) migrated to new places and started new civilization. Such new civilization might not have any formal links with their ancestors and might have lost links with them and their present generation over a period of time. By link prediction problem such missing links can be identified by proposing suitable complex network models not only in the human evolution network but also in biological networks, disease propagation networks and gene regulatory networks to name a few. Another commercial application of link prediction is in the identification of links among persons who form a link over a period of time. Some of the social media networks like Facebook network, Twitter network and LinkedIn network predict links by involving the additional attributes of the persons in the networks such as the cities they visited, colleges they studied, likes of pages and so on. Friend recommendation in social networks is based on these attributes in addition to the graph topology. Not only unnoticed links in the civilization, in friend network, link prediction is also used in recommending items to users in e-commerce based on the previous purchase and persons profile information such as purchase capability, credit limit, interests and previous search history. Such a method of recommending is well known as recommender systems. Link prediction problem in financial transaction details is used to identify fraudulent transactions and funding for illegal and anti-social activities. If a persons financial transactions are modeled in the form of a network over a due course of time, then the majority of transactions will form a pattern or similar pattern is observed in each interval. Any financial transactions outside these transactions can be identified as an unusual behaviour of that person and countermeasures can be taken to alert the person. Author networks are analyzed to understand connections or interactions between the authors and their co-authors and the journals they are likely to publish in their research findings. By link prediction problem, the interaction among the authors and the authors they collaborate or likely to collaborate and journals or conferences they are likely to communicate can be understood. Over the topological features of the graph, clustering information, edge betweenness centrality and k – path edge connectivity measures were combined to define a new algorithm to predict links on the social network [10]. Weight path based structural link prediction algorithm is proposed by incorporating both cluster information and edge betweenness centrality values. From the experimental values it was concluded that proposed algorithms performs better over the benchmark algorithms in literature. In addition to the existing topological features, user behavior in the network also helps to predict links.

2. Related Works

Online social networks' nodes communicate with one another via some form of online communication. Online social networks are defined as a group of people who communicate via email and another group of people who communicate solely through social networking sites such as Facebook and Myspace. The number of nodes to which a single person is linked varies from person to person. Because the mechanisms of formation of both networks are nearly identical[12], typical properties of underlying offline social networks can be found in online social networks [13]. Individual similarities, just like in an offline social network, make online social interaction more likely. McPherson discovered that there is a direct link between network interaction and network structure [14]. Various researchers have proposed various probabilistic models for determining the likelihood of two people becoming friends [15]. Information was collected from social networks using surveys, questionnaires, and personal interviews in the early days of social network analysis. The structure and evolution of online social networks has simplified the task of gathering data. Web crawlers are now being used to collect information from online social networks [16]. Online social networks' nodes communicate with one another via some form of online communication. Online social networks are defined as a group of people who communicate via email and another group of people who communicate solely through social networking sites such as Facebook and Myspace. The number of nodes to which a single person is linked varies from person to person. Because the mechanisms of formation of both networks are nearly identical, typical properties of underlying offline social networks can be found in online social networks. Individual similarities, just like in an offline social network, make online social interaction more likely. McPherson discovered that there is a direct link between network interaction and network structure [17]. Various researchers have proposed various probabilistic models for determining the likelihood of two people becoming friends [18]. According to the research studies mentioned above, online social networks share many topological properties with complex networks. In online social networks, degree distributions follow a power law distribution, with a heavy tail. At some point, a cut-off may be necessary (usually of few thousands). The majority of online social networks have a small diameter, with a few exceptions. The effective diameter can be used to determine the percentage of nodes that are connected and reachable from one another. Some of the nodes in the social network will be densely populated and dubbed "cores," and these nodes will be in charge of keeping the network together and connecting low-degree nodes to it. The cores are in charge of reducing the network's shortest paths. When compared to the node degree in an offline social network, the node degree in an online social network is very large. The number of friends in an offline social network is usually limited to a few hundreds. Huberman discovered that even though a user in an online social network has a large number of connections, they only interact with a small number of them. The goal of this study is to focus on online social network mining, also known as link prediction, which is applied to undirected online social networks represented by undirected graphs. The main goal of link prediction is to find future missing links between network nodes based on previous relationships, which is what network topology is all about. The limitations of existing state-of-the-art link prediction techniques have also been investigated. Link prediction has been discovered to assist in addressing issues and challenges in a variety of domains, such as future friend recommendation. There are, however, issues with precision and scalability. The diversity of online social networks poses a challenge in terms of precision, because link prediction techniques used to forecast future missing links in one type of network will not accurately predict links between nodes in another type of network. Friendship networks have different properties than co-authorship networks. Furthermore, accurately predicting links between nodes in a sparse network is difficult. Given the widespread use of the internet and social media, the link prediction problem remains a significant issue in the field of social media. The networks that were processed for link prediction were taken one at a time and were also of small size, implying that there were fewer nodes in total. As a result, the techniques developed at the time are now unsuitable for analyzing the current state of network structures. Previous research has not found a significant difference between link detection and link prediction; both appear to be the same[19]. The problems of link detection and prediction are distinct, and our focus is on future link prediction. Along with such flaws, there are a variety of real-time applications that require link predictions to be investigated in order to improve recommendation systems. Link prediction plays an important role in decision-making; depending on the type and behavior of social networks, link prediction analysis can help in a variety of ways, as shown in the use-cases below. As is well known, any network is represented through sociography, where nodes are users of people and interactions between them are represented by links in the popular online social network Facebook. The type of relationship is determined by the interaction context. For example, in a friendship edge, one user initiates a friendship request, which the other node accepts. Predicting friendship links aids in the generation of business revenue through the placement of advertisements on user pages, because knowing who is whose friend will assist the business organization in identifying users who will have a higher probability of connecting new friends across the network. Another scenario for link prediction is in any co-authorship network [20], such as dblp, where authors are nodes and links represent collaboration between them. The interaction between authors, i.e. the link between the nodes, indicates that at least one research document will be published as a result of the collaboration. Although link prediction analysis of such scientific networks has no direct implications, because the network is another example of online social networks, an approximation of link prediction will verify the method in other networks as well. As a result, such networks have piqued the interest of scientists, physicists, and mathematicians for further investigation [21]. The focus of link prediction research has primarily been on improving execution for specific exact frameworks, with little attention paid to link prediction in network models. To begin, the networks' group structure is removed at various resolutions. Then, under various resolutions, a basic recurrence measurable model is connected to determine how frequently a couple of nodes isolated into the same group. The probability of missing connections is calculated in the long run. This calculation is demonstrated by contrasting and other seven link prediction strategies on two different types of networks at different scales. The results show that the methodology performs well in terms of precision and has a lower time complexity than some other calculations that rely on the system's various structures. The results of the tests show that the execution of a few strategies is strongly linked to specific network measurements. The authors discovered how to distinguish between "prediction friendly" networks, for which the vast majority of forecasting strategies produce excellent results, and "prediction unfriendly" networks, for which the vast majority of techniques produce high prediction error. Correlation analysis between network measurements and forecast precision of expectation strategies could form the foundation of a metal earning framework that can recommend the best forecast technique for a given network based on network qualities. The majority of today's strategies rely on the standard neighbour list and its variations. The Pearson correlation coefficient is proposed by Liao et al in their paper to compute the similarity between nodes. When used to compute closeness in higher order paths, this strategy is found to be extremely compelling [22]. The authors combined a relationship-based technique with a resource allocation strategy and discovered that the combined technique outperforms current strategies, especially in sparse networks. According to the findings, the Pearson relationship coefficient is more resistant to noisy data than other methods. An interesting addition would be to investigate the link prediction problem in a noisy environment, such as the watched network with some noisy connections. The relationship-based strategy and alternative strategies can be analyzed, and the robustness of these techniques to noise can be considered methodically. The Pearson relationship introduces a new way of calculating the distance between nodes. In transient dubious networks, [23] proposed a strategy for connection prediction. The anticipating problem is formalized using this technique by outlining an irregular random walk in transient indeterminate networks. Humans (of various races) migrated to new places and established new civilizations as a result of war or other natural disasters. Such a new civilization may have no formal ties to their forefathers and may have lost touch with them and their current generation over time. Such missing links can be identified using the link prediction problem by proposing appropriate complex network models not only in the human evolution network but also in biological networks, disease propagation networks, and gene regulatory networks, to name a few examples. The identification of links among people who form a link over time is another commercial application of link prediction. Some social media networks, such as Facebook, Twitter, and LinkedIn, predict links by incorporating additional attributes of the individuals in the networks, such as cities visited, colleges attended, page likes, and so on. In social networks, friend recommendations are based on these attributes as well as the graph topology. Not only in the civilization, but also in the friend network, link prediction is used in e-commerce to recommend items to users based on previous purchases and personal profile information such as purchase capability, credit limit, interests, and previous search history. Recommender systems are a well-known method of recommending. Financial transaction details are used to identify fraudulent transactions and funding for illegal and anti-social activities using the link prediction problem. When a person's financial transactions are modeled as a network over a period of time, the majority of transactions will form a pattern or a similar pattern in each interval [24]. Any financial transactions that are not part of these transactions can be identified as unusual behavior by that person, and countermeasures can be taken to alert them. Biological networks are studied to better understand the interactions between the various entities involved, such as protein-protein interactions, in order to identify and report existing or new diseases or biological conditions. In addition, understanding the likelihood of developing diseases or abnormal health conditions over time (in future generations) has made link prediction a popular research topic in bioinformatics. Author networks are examined to determine the connections or interactions between authors and their co-authors, as well as the journals in which their research findings are most likely to be published. The interaction between the authors and the authors with whom they collaborate or are likely to collaborate, as well as the journals or conferences with which they are likely to communicate, can be understood using the link prediction problem. Though the link prediction problem appears to be simple, it provides valuable insights into the network, whether social or complex, and the interactions among the entities involved. Large nodes and inherent topology in social networks must be analysed in order to make meaningful inferences about link prediction. It's a difficult task for large graphs because traversing them necessitates a significant amount of computational effort and the use of optimized techniques. Two variants of the firefly link prediction algorithm are proposed, focusing on structural links between network nodes and attribute relationships between attribute nodes and structural nodes that fireflies traverse. Experiments have shown that the proposed algorithms outperform the existing algorithms in the literature. When using fireflies with a normal distribution to traverse a large graph, the search space will be limited by fireflies. To solve this problem, Levy flight is used to guide the behavior of fireflies, ensuring that they search the entire parameter space. In order to improve the accuracy of the link prediction algorithm, node attributes are taken into account in addition to the existing topological features. An augmented graph is one in which existing nodes are connected to additional nodes based on the attributes they possess. To predict links, fireflies will be made to traverse the entire graph, including the augmented nodes. The fireflies will flock to densely connected nodes, which have a higher chance of forming links. An intelligent adaptive local neighbourhood range based link prediction technique is proposed to improve link prediction accuracy and avoid leaving any node that has the potential to form links in the future. The fireflies in the proposed technique will intelligently increase their boundary in the neighbourhood and search for potential links based on a threshold. The exploration boundary will be expanded so that nodes with the potential to form links are not overlooked. The analysis of online social networks aids in the development of valuable inferences on a variety of topics, including people's browsing habits, information diffusion, the discovery of groups/communities in the network, the identification of the most influential people in online social networks, and the degree of influence. Individuals form the vertices, and connections or associations form the edges, hence social networks are typically represented as graphs for ease of analysis. Actors and ties are terms used to describe people. Adjacency list and adjacency matrix are both used to represent social networks [25]. The great majority of parameters/elements in online social networks are ambiguous and not neatly linked to any traits. Identity traits, for example, vary depending on observation, location, and economic well-being, all of which are difficult to define precisely. Traditional logic theory, such as true/false or any parallel rationale, is thus useless in studying online social networks. This chapter attempts to associate a proportion of network measurements in online social network analysis and construct a few inductions between each measure in that regard. The association between two nodes is characterized as a connection in an online social network. There are several OSN measurements that are based on the relationship between nodes in an online social network. In a static network, certain network metrics can't be calculated. Reachability, degree, density, bridge, and connectivity are some of the connection-based measurements available. In online social networks, the degree is the most important connection-based metric. The number of linkages a node has is known as its degree. An in-degree or out-degree can be obtained. The number of links that occur on the node is called in-degree, whereas the number of links that begin at the node is called out-degree. Reachability is the degree to which two nodes in an online social network are connected, regardless of how many hops separate them. It's commonly used to see if two nodes are connected. It is defined as the degree to which any component of a network can communicate with another component of the network. The course taken by the two people can be either well-coordinated or erroneous. In coordinated networks, the reachability will be unique. The fact that node A can be reached from node B does not mean that node B can be reached from node A. The ratio of the real number of nodes in a network to the total number of nodes possible in the network is known as density. The quality of associations between nodes in a network is measured by connectivity. It's calculated by continuously emptying the edges of a graph and checking the time when a node becomes entirely isolated. Different connections to the same node will have better data scattering and more grounded availability. Furthermore, in online social networks, a bridge is a knot that connects two networks or subnetworks. The number of bridges in the system determines how widely data spreads in online informal communities. Radiality is defined as the degree to which a person's affiliation connects with the rest of the network and provides new methods to reach a node. A higher radiality estimate means that it takes less hops for a node to reach any other individual in the online social network. We discovered that, while there are a variety of network topological property indices and link prediction techniques in the literature, no study explores the correlation between topological property indices and link prediction techniques, based on the existing studies on various similarity score heuristics for link prediction. In addition, we discovered that existing similarity score-based link prediction approaches either use node structure or node attribute data. There is no link prediction technique in the literature that uses both node structure and profile information. We also discovered that similarity score-based link prediction systems require a threshold value on which decisions can be made. However, because huge social networks have extremely dynamic structures, determining a threshold value on which link prediction can be conducted is problematic. Each person in an online social network has a unique collection of characteristics, making it difficult for a comprehensive social network to determine a common threshold for determining if two unconnected individuals will be joined (containing thousands of egos). A bipartite graph with two types of nodes, users and products, can be found in an E-Commerce network. If the user has purchased the item, the two are linked. For link prediction, the article discussed the use of a "indirect" feature. For all people that bought a common item, a similar item purchase count is employed as a feature. It was discovered that adding indirect features to the training set enhanced link prediction accuracy. For forecasting the chance of linkages among distinct node pairings, [27] used the logistic regression technique. They used data from CiteSeer articles, AT&T phone calls, and Enron emails to piece together a network. For forecasting the probability of collaborations in DBLP and PubMed research papers, [28] used a local Markov random field model with a logistic regression technique. Using the support vector machine (SVM) approach, [29] predicted links on the Google+ website. By developing a decision tree classifier, [30] tackled the challenge of link prediction in dynamic social networks. For identifying missing links in co-authorship networks, [31] used supervised learning techniques as SVM, Decision Tree, and J48. They discovered that the performance of link prediction is affected by the classification technique used. The authors discovered that using place features boosts the performance of machine learning techniques. [32] investigated the use of supervised machine learning algorithms for yelp.com, a restaurant review website. In a restaurant review network, maximum entropy was utilized to predict links. We discovered that machine learning-based link prediction algorithms are useful for discovering missing linkages in social networks based on previous research. The performance of machine learning-based link prediction algorithms for friendship networks, on the other hand, has not been investigated. As a result, we developed a number of supervised machine learning-based link prediction models for complex networks.

3. Hybrid Filtered based Graph Clustering and Link Prediction Model on Complex OSN Datasets

In this section, a hybrid graph based clustering model is implemented on the online social networking data for link prediction process as shown in figure 1. Various types of social networking complex datasets are used in this framework to find outliers and to transform data. Data classification models are used to find the essential decision-making process patterns after the data filtering operation is completed. Here, statistical metrics are used to compare the proposed model's performance to that of other models. Heterogeneous datasets are used to find the class prediction in the proposed framework, which employs the proposed link prediction model. The input data is initially prepared using the various class attributes. The missing values are filled with mean values after nominal attributes are converted to binary attributes. To improve link prediction rates, these filtered data are fed into classification problems. For the decision making process, various nominal attributes are used as class labels.

Algorithm 1: Complex Social Networking Data Filtering Process

Input: Online social networking datasets OSND={SD-1,SD-2…SD-n}, Attribute: A_T ,Min_c :Minimum frequency count of attribute value that contains class label c,
$Max_{c:}$ Minimum frequency count of attribute value that contains class label c,
M_x: Maximum of attribute values.
M_n: Minimum of attribute values.
1: Read OSN dataset as OSND.
2. To each sample in OSND[]
3. Do
4. To each instance I[i]
5. Do
6. To each attribute in I[i][j]
7. Do
8. If(A_T [i]==Numerical&& A_T [i]==Null)
9. then
10. Replace $A\tau[I]$ using the eq.(1)
11. $A_T[I]=\frac{|A_T[I]-(\mu_x(A_T)|}{2.\sigma_{A\tau}(A_T)}*(Max_n(A_T)/2-Min_n(A_T)/2)$ --(1)
12. End if
13. Done
18 Done

19 done

In the proposed algorithm 1, online social networking data samples are taken as in put for the filtering process. Initially, each data object and the attribute space are used to find the minimum and maximum frequency of the attribute values. If the attribute is numerical then the equation 1 is used to compute the filtered attribute normalized value for the data filtering.

Figure 1; Proposed Framework on Complex Social Networking Data

Algorithm 2: Probabilistic Weighted based Community Clustering
Probabilistic Weighted Measure for Community Detection

In this work, a hybrid probabilistic weighted measure is used to find the key relational graph nodes based on the attributes for community detection process. The centralized mean weighted measure between the attributes is given as,

$$B_f = \text{uniqueCV(D)}; // \text{Unique column values}$$

$$HB_f = \text{Histobins[]} = \text{histogrambin(D)}$$

$$\text{GaussianKernel}: GK(\phi, \theta) = e^{-\theta^2} / (2 * \phi^2)$$

$$\psi = gkv = GK(\sum HB_f, \sum B_f);$$

$$\lambda 1 = \frac{|(\mu_{A1} - \mu_{A2})|}{\psi . \sqrt{\min\{\sigma_{A1}, \sigma_{A2}\}}} \quad \text{-----(1)}$$

$$\lambda 2 = \text{Min}\{|HB_f / (\sum \psi * HB_f)|, \frac{\text{Max(Prob(A1 / Cm))}}{2.|\sum |A1||}, \frac{\text{Max(Prob(A 2 / Cm))}}{2.|\sum |A2||}\}$$

where M_{A1} is the average of the attribute A1 wrt class samples

M_{A2} is the average of the attribute A2 wrt class samples.

Maximized weighted probabilistic measure is given by: MPWM=T=max{ $\lambda 1, \lambda 2$ }

Input: dataset D
Output: Filtered dataset D'

1. Read the input dataset.
2. To each numerical attribute A in the dataset D.
3. Apply algorithm 1 for data filtering.
4. Done
5. If the dataset contains heterogeneous attributes in the list FD[].
6. To each attribute s in FD
7. Do
8. Link prediction using model1.
9. Done
10. To each instance in the local community objects
11. do
12. For each instance O_i in the KNN objects KNN[]
13. Do
14. For each instance O_j in IPG[] // where i!=j
15. Computing the Chebyshev distance N_m^k on the KNN objects.

$$|P(SR_i)| = \sqrt{P(m_1)^2 + P(m_2)^2 \ldots .P(m_n)^2}$$

$$|Q(SR_j)| = \sqrt{Q(n_1)^2 + Q(n_2)^2 \ldots .Q(n_r)^2}$$

$$|P(SR_i).Q(SR_j)| = P(m_1).Q(n_1) + P(m_2).Q(n_2) \ldots + P(m_n).Q(n_r)$$

Proposed Contextual depedency global skyline objects are computed as

$$\text{Contextual dependency ran: CDR} = \frac{\sqrt{P(m_i).Q(n_j) * \cos(|P(m_i)| + |Q(n_j)|)}}{2 * \log(|P(m_i).Q(n_j)|)}; \text{where } i \neq j$$

Contextual similarity ranking is given as

$$CSR = 1 - CDR;$$

Sort all the related neighbor objects with highest similarity as nearest ranked list.

16. Done
17. To each Chebyshev distance objects in the local density modelling, find the k nearest objects in the sorted as
18. N_m^k []=TopKNN(k);

19. Apply local density estimation probability on the filtered local objects.
20. Find the nearest density objects using the proposed probabilistic KNN method.
21. Construct initial graph clustering with k nodes as clusters.
22. Compute local and global density estimation by using the following weighted measures as

$$Dist_c = mean^K + \lambda_1 . \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\phi_i^K / \rho_i^K)}$$

Here, N is total number of filtered knn objects

ρ_i^K is the average of the k-th closest object to instance i.

$\rho_t^K = max_j\{KNN_i(\mathrm{Dist}_{ij})\}$, and

M^K is the average of ρ_t^K ,

calculate average of k-values as $\mathrm{M}^K = \lambda_1 \frac{1}{N}\sum_{i=1}^{N}\rho_t^K$

Prior estimation probability $= \kappa =$ Prob(I,C_k)=Max{Prob(I/C_k); k:k-nearest objects}

Pr*oposed* local density estiamtion is given as

$$\mathrm{P}LDE(v_i) = \frac{\kappa}{\max\{\lambda_1, \lambda_2\}} \mathrm{e}(-\frac{\| \log(v_{ij}) - Dist_c \|^2}{2\sigma^2})$$

23. Repeat this procedure till k graph clusters.
24. Done
25. Apply optimized decision tree classification model for link prediction.

In the proposed approach, in order to increase the classification rate, the boosting mechanism employs a collection of weak classifiers. The method used by Decision Tree is a poor classification for the creation of samples on the Adaboost algorithm in this approach. A novel decision-making hybrid entropy and probabilistic entropy are hybridized using the updated decision tree ranking steps. For example, the low classification error rate classifier is chosen in this algorithm.

26. Done

In the algorithm 2, initially, all the data objects are filtered using the algorihthm1. These filtered data are used for data machine learning and community detection process.

4. Experimental Results

In the proposed system, a graph based community classification framework is designed and implemented for complex social networking data by implementing the community cluster -based machine learning techniques. Each classifier is run independently and the important

ISSN: 2094-0343
2326-9865

features were noted. The common features of all the three classifiers are collected and considered as optimal feature subset. The performance of the selected feature subset was evaluated by Naive Bayes classifier by using the performance metrics like precision, recall, accuracy. The results obtained were compared with different traditional feature selection based classifiers. By comparing the results, it is concluded that the gradient based feature selection model yields better results than the traditional machine learning feature based classification models.

Experimental results are simulated in java environment with third party graph and similarity libraries. To evaluate the performance of proposed model, four datasets such as yelp, football, Zachary and dolphin datasets. In this experimental study, different metrics are used to evaluate the results. Metrics such as density, NMI (normalized mutual information), variation of information(entropy) are used to evaluate the results on the training datasets.

$$Density = \sum \left(e_{ii} - a_i^2 \right)$$

$$NMI(\mathrm{P} |\ \mathrm{Q}) = \frac{\mathrm{e(P)} + e(\mathrm{Q}) - e(\mathrm{P,Q})}{(\mathrm{e(P)} + e(\mathrm{Q}))/2},$$

Where X is the original value and Y is the predicted communities. e(P) and e(Q) are the entropy values of the corresponding communities.

Table 1. Sample Complex Social Networking Dataset

Social networking data are represented by nodes, and mutual links between them are represented by edges. The edges are contained in the edges csv files; nodes are indexed from 0. Experimental data is taken from [https://graphmining.ai/datasets/wikipedia/squirrel.zip] [https://graphmining.ai/datasets/wikipedia/crocodile.zip]. The inclusion of a feature in the feature list indicates the presence of an informative noun in the Wikipedia article's content. We reported the number of nodes and edges for each page-page network, along with some other descriptive statistics.

```
 Proposed Classifier For Link Prediction
==========

Destination = '(-inf-8129.9]'
|   Shortest_Path = '(-inf-1.1]'
|   |   Page_Rank_Dst = '(-inf-0.000196]'
|   |   |   Followees_Dst = '(-inf-39]'
|   |   |   |   Source = '(-inf-8130.5]'
|   |   |   |   |   Page_Rank_Src = '(-inf-0.000196]'
|   |   |   |   |   |   Int_Followers = '(-inf-4.5]'
|   |   |   |   |   |   |   Followers_Src = '(-inf-40.1]'
|   |   |   |   |   |   |   |   Followees_Src = '(-inf-39]'
|   |   |   |   |   |   |   |   |   ID = '(-inf-3532.3]'
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(-inf-40.1]'
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(-inf-7.6]' : 1 (517/31)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(7.6-15.2]' : 1 (1/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(15.2-22.8]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(22.8-30.4]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(30.4-38]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(38-45.6]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(45.6-53.2]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(53.2-60.8]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(60.8-68.4]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Int_Followees = '(68.4-inf)' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(40.1-80.2]' : 1 (5/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(80.2-120.3]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(120.3-160.4]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(160.4-200.5]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(200.5-240.6]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(240.6-280.7]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(280.7-320.8]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(320.8-360.9]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   Followers_Dst = '(360.9-inf)' : 0 (0/0)
|   |   |   |   |   |   |   |   |   ID = '(3532.3-7064.6]'
|   |   |   |   |   |   |   |   |   |   Int_Followees = '(-inf-7.6]'
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(-inf-40.1]' : 1 (535/23)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(40.1-80.2]' : 1 (3/0)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(80.2-120.3]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(120.3-160.4]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(160.4-200.5]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(200.5-240.6]' : 0 (0/0)
|   |   |   |   |   |   |   |   |   |   |   Followers_Dst = '(240.6-280.7]' : 0 (0/0)
```

Figure 2; Loading Facebook OSN Dataset

```
|   |   |   |   |   |   Source = '(-inf-8130.5]' : 1 (205/0)
|   |   |   |   |   |   Source = '(8130.5-16261]' : 0 (1/0)
|   |   |   |   |   |   Source = '(16261-24391.5]' : 0 (0/0)
|   |   |   |   |   |   Source = '(24391.5-32522]' : 0 (0/0)
|   |   |   |   |   |   Source = '(32522-40652.5]' : 0 (0/0)
|   |   |   |   |   |   Source = '(40652.5-48783]' : 0 (0/0)
|   |   |   |   |   |   Source = '(48783-56913.5]' : 0 (0/0)
|   |   |   |   |   |   Source = '(56913.5-65044]' : 0 (0/0)
|   |   |   |   |   |   Source = '(65044-73174.5]' : 0 (0/0)
|   |   |   |   |   |   Source = '(73174.5-inf)' : 0 (0/0)
|   |   |   |   Followers_Dst = '(40.1-80.2]'
|   |   |   |   |   Followees_Dst = '(-inf-39]'
|   |   |   |   |   |   Source = '(-inf-8130.5]' : 1 (108/0)
|   |   |   |   |   |   Source = '(8130.5-16261]' : 0 (1/0)
|   |   |   |   |   |   Source = '(16261-24391.5]' : 0 (0/0)
|   |   |   |   |   |   Source = '(24391.5-32522]' : 0 (1/0)
|   |   |   |   |   |   Source = '(32522-40652.5]' : 0 (1/0)
|   |   |   |   |   |   Source = '(40652.5-48783]' : 0 (0/0)
|   |   |   |   |   |   Source = '(48783-56913.5]' : 0 (1/0)
|   |   |   |   |   |   Source = '(56913.5-65044]' : 0 (3/0)
|   |   |   |   |   |   Source = '(65044-73174.5]' : 0 (1/0)
|   |   |   |   |   |   Source = '(73174.5-inf)' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(39-78]' : 1 (16/0)
|   |   |   |   |   Followees_Dst = '(78-117]' : 1 (3/0)
|   |   |   |   |   Followees_Dst = '(117-156]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(156-195]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(195-234]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(234-273]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(273-312]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(312-351]' : 0 (0/0)
|   |   |   |   |   Followees_Dst = '(351-inf)' : 0 (0/0)
|   |   |   |   Followers_Dst = '(80.2-120.3]' : 1 (24/0)
|   |   |   |   Followers_Dst = '(120.3-160.4]' : 1 (7/0)
|   |   |   |   Followers_Dst = '(160.4-200.5]' : 0 (0/0)
|   |   |   |   Followers_Dst = '(200.5-240.6]' : 0 (0/0)
|   |   |   |   Followers_Dst = '(240.6-280.7]' : 0 (0/0)
|   |   |   |   Followers_Dst = '(280.7-320.8]' : 0 (0/0)
|   |   |   |   Followers_Dst = '(320.8-360.9]' : 0 (0/0)
|   |   |   |   Followers_Dst = '(360.9-inf)' : 0 (0/0)
|   |   |   Page_Rank_Src = '(0.000196-0.000381]' : 1 (20/0)
|   |   |   Page_Rank_Src = '(0.000381-0.000567]' : 1 (7/0)
```

Figure 3; Different Community Clusters in the Visualization Form

```
=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.989    0.000    1.000      0.989    0.994      0.989    0.999     0.999     0
                 1.000    0.011    0.989      1.000    0.994      0.989    0.999     0.999     1
Weighted Avg.    0.994    0.006    0.995      0.994    0.994      0.989    0.999     0.999


=== Confusion Matrix ===

     a     b   <-- classified as
 34932   392 |     a = 0
     0 35324 |     b = 1

Time taken to perform cross-validation: 2.23 seconds


=== Stratified cross-validation ===

Correctly Classified Instances       70258               99.448  %
Incorrectly Classified Instances       390                0.552  %
Kappa statistic                          0.989
Mean absolute error                      0.0103
Root mean squared error                  0.0724
Relative absolute error                  2.0582 %
Root relative squared error             14.4826 %
Total Number of Instances            70648


=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.989    0.000    1.000      0.989    0.994      0.989    0.999     0.999     0
                 1.000    0.011    0.989      1.000    0.995      0.989    0.999     0.998     1
Weighted Avg.    0.994    0.006    0.995      0.994    0.994      0.989    0.999     0.998


=== Confusion Matrix ===

     a     b   <-- classified as
 34935   389 |     a = 0
     1 35323 |     b = 1
```

Figure 4; Dynamic Visualization of Facebook Clusters for OSN

Table 2. Comparative Analysis of Different Graph based Community Clustering Models on Facebook Dataset

Test_Sample	W-cluster	IGC-Cluster	AR-Cluster	Proposed Clusterbased Classifier
NODES-1%	0.942	0.943	0.932	0.981
NODES-2%	0.941	0.944	0.925	0.982
NODES-3%	0.938	0.931	0.922	0.983
NODES-4%	0.937	0.94	0.926	0.984
NODES-5%	0.935	0.944	0.92	0.981
NODES-6%	0.935	0.936	0.942	0.979
NODES-7%	0.937	0.932	0.919	0.98
NODES-8%	0.935	0.937	0.937	0.981
NODES-9%	0.941	0.943	0.923	0.98
NODES-10%	0.94	0.933	0.934	0.983
NODES-11%	0.94	0.94	0.924	0.983
NODES-12%	0.94	0.934	0.944	0.984
NODES-13%	0.933	0.932	0.934	0.981
NODES-14%	0.943	0.942	0.94	0.979
NODES-15%	0.938	0.939	0.942	0.982
NODES-16%	0.944	0.936	0.931	0.983
NODES-17%	0.941	0.936	0.923	0.982
NODES-18%	0.932	0.941	0.925	0.98
NODES-19%	0.94	0.941	0.942	0.984
NODES-20%	0.943	0.935	0.943	0.979

Table 2, represents the comparative analysis of present probabilistic model to the conventional models on facebook dataset. As shown in the table 2, proposed NML has better efficiency than the traditional models on facebook dataset. The NLM value represent the quality of the inter and intra community detection process on the facebook dataset.

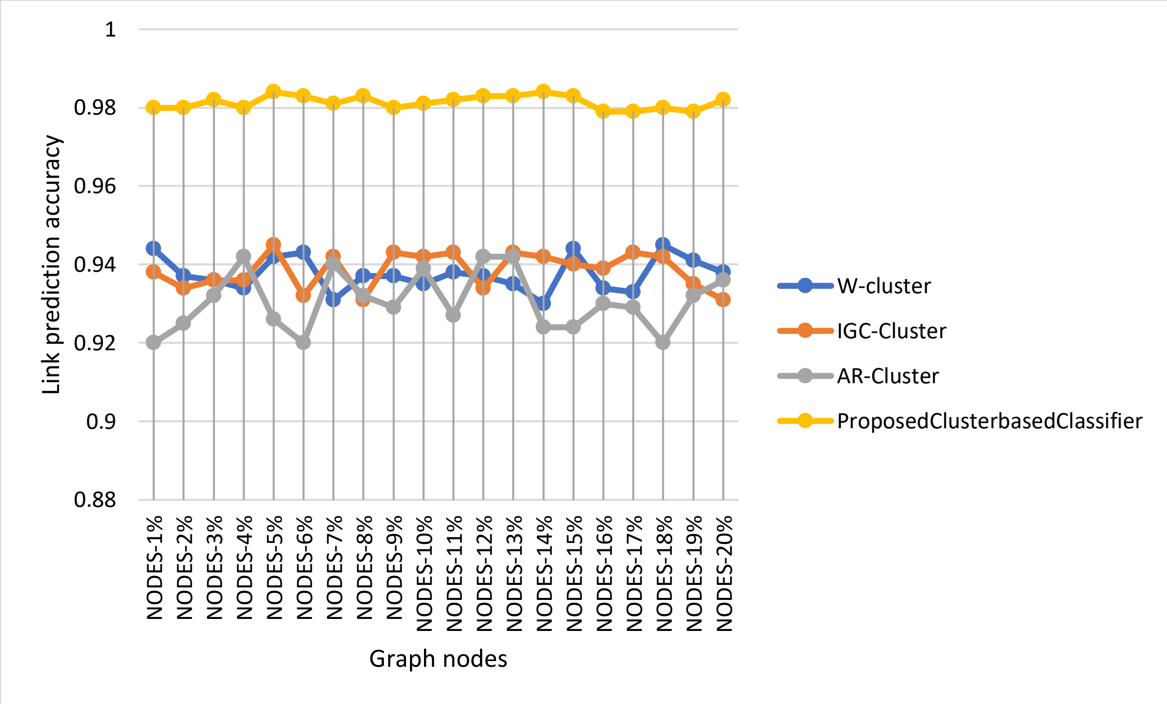

Figure 5; Comparative Analysis of Proposed Graph Cluster based Classification Value to Existing Models on Facebook Data

Figure 5, represents the comparative analysis of present probabilistic graph cluster model to the conventional models on facebook dataset. As shown in the figure 5, proposed entropy value has better efficiency than the traditional models on facebook dataset. The entropy value represents the quality of the inter and intra community detection process on the facebook dataset.

Table 3. Comparative Analysis of Proposed Model Density Value to Existing Models on Yelp Data

Test_Sample	W-cluster	IGC-Cluster	AR-Cluster	Proposed Clusterbased Classifier
NODES-1%	0.932	0.942	0.924	0.981
NODES-2%	0.937	0.944	0.932	0.98
NODES-3%	0.935	0.936	0.94	0.983
NODES-4%	0.942	0.934	0.929	0.981
NODES-5%	0.936	0.942	0.932	0.98
NODES-6%	0.942	0.94	0.919	0.981
NODES-7%	0.942	0.938	0.933	0.984
NODES-8%	0.942	0.937	0.923	0.979
NODES-9%	0.939	0.938	0.926	0.983
NODES-10%	0.936	0.937	0.935	0.983
NODES-11%	0.941	0.935	0.919	0.982
NODES-12%	0.933	0.941	0.926	0.981
NODES-	0.944	0.937	0.923	0.98

13%				
NODES-14%	0.938	0.931	0.941	0.981
NODES-15%	0.942	0.942	0.927	0.984
NODES-16%	0.937	0.933	0.939	0.981
NODES-17%	0.938	0.938	0.928	0.98
NODES-18%	0.936	0.934	0.938	0.984
NODES-19%	0.933	0.932	0.941	0.983
NODES-20%	0.942	0.932	0.928	0.982

Table 3, represents the comparative analysis of present graph cluster based classification model to the conventional models on yelp dataset. As shown in the table 3, proposed density value has better efficiency than the traditional models on yelp dataset. The density value represent the quality of the inter and intra community detection process on the yelp dataset.

Table 4. Performance of Proposed Runtime (MS) to Conventional Models on Different Facebook Network data Samples

Test_Sample	W-cluster	IGC-Cluster	AR-Cluster	Proposed Cluster based Classifier
NODES-1%	4657	4618	4514	3359
NODES-2%	4585	4726	4340	3353
NODES-3%	4735	4675	4315	3405
NODES-4%	4653	4633	4350	3449
NODES-5%	4733	4645	4736	3356
NODES-6%	4630	4624	4442	3407
NODES-7%	4603	4736	4652	3430
NODES-8%	4609	4726	4573	3370
NODES-9%	4622	4535	4717	3363
NODES-10%	4648	4709	4305	3308
NODES-11%	4568	4574	4633	3306
NODES-12%	4653	4748	4708	3342
NODES-13%	4612	4614	4328	3430
NODES-14%	4674	4603	4555	3325
NODES-15%	4566	4682	4766	3383
NODES-16%	4559	4622	4318	3453

NODES-17%	4767	4585	4285	3423
NODES-18%	4763	4559	4480	3361
NODES-19%	4619	4547	4693	3381
NODES-20%	4597	4659	4777	3392

In this table 4, the runtime(ms) comparison of proposed model to the conventional models are presented. From the table, it is noted that the runtime of the probabilistic cluster based machine learning model has better efficiency than the conventional models on the OSN datasets.

5. Conclusion

Dynamic community clustering plays a vital role in the online social networking datasets. Since, most of the traditional community clustering modes are static in nature and only applicable to non-link prediction approach. Also, traditional clustering measures use nearest neighbor measures instead of contextual similarity in order to predict the relationship among the different graph nodes. In order to optimize the contextual node clustering and link prediction, a hybrid dynamic scalable measure is proposed for the community clustering on complex networks. In this work, a hybrid link prediction approach is proposed on the complex social networking data for better decision making patterns. Experimental results prove that the proposed contextual probabilistic graph clustering and link prediction approach has better efficiency than then conventional models on complex social networking datasets.

References

1. Wu, J., Zhang, G., & Ren, Y. (2017). A balanced modularity maximization link prediction model in social networks. Information Processing & Management, 53(1), 295-307.
2. https://doi.org/10.1016/j.ipm.2016.10.001
3. Pandey, B., Bhanodia, P.K., Khamparia, A., & Pandey, D.K. (2019). A comprehensive survey of edge prediction in social networks: Techniques, parameters and challenges. Expert Systems with Applications, 124, 164-181. https://doi.org/10.1016/j.eswa.2019.01.040.
4. Bastami, E., Mahabadi, A., & Taghizadeh, E. (2019). A gravitation-based link prediction approach in social networks. Swarm and evolutionary computation, 44, 176-186.
5. https://doi.org/10.1016/j.swevo.2018.03.001
6. Wang, G., Wang, Y., Li, J., & Liu, K. (2021). A multidimensional network link prediction algorithm and its application for predicting social relationships. Journal of Computational Science, 53.
7. https://doi.org/10.1016/j.jocs.2021.101358
8. Nasiri, E., Berahmand, K., & Li, Y. (2021). A new link prediction in multiplex networks using topologically biased random walks. Chaos, Solitons & Fractals, 151.
9. https://doi.org/10.1016/j.chaos.2021.111230.

10. Nasiri, E., Berahmand, K., Rostami, M., & Dabiri, M. (2021). A novel link prediction algorithm for protein-protein interaction networks by attributed graph embedding. Computers in Biology and Medicine, 137.
11. https://doi.org/10.1016/j.compbiomed.2021.104772
12. Berahmand, K., Nasiri, E., Forouzandeh, S., & Li, Y. (2021). A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. Journal of King Saud University-Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2021.05.006
13. Florentino, É.S., Cavalcante, A.A., & Goldschmidt, R.R. (2020). An edge creation history retrieval based method to predict links in social networks. Knowledge-Based Systems, 205.
14. https://doi.org/10.1016/j.knosys.2020.106268
15. Daud, N.N., Ab Hamid, S.H., Saadoon, M., Sahran, F., & Anuar, N.B. (2020). Applications of link prediction in social networks: A review. Journal of Network and Computer Applications, 166.
16. https://doi.org/10.1016/j.jnca.2020.102716
17. Yang, F., Qiao, Y., Wang, S., Huang, C., & Wang, X. (2021). Blockchain and multi-agent system for meme discovery and prediction in social network. Knowledge-Based Systems, 229.
18. https://doi.org/10.1016/j.knosys.2021.107368
19. Zhang, W., Wu, B., & Liu, Y. (2016). Cluster-level trust prediction based on multi-modal social networks. Neurocomputing, 210, 206-216. https://doi.org/10.1016/j.neucom.2016.01.108
20. Zheng, Y., Hu, R., Fung, S. F., Yu, C., Long, G., Guo, T., & Pan, S. (2020). Clustering social audiences in business information networks. Pattern Recognition, 100.
21. https://doi.org/10.1016/j.patcog.2019.107126
22. Karimi, F., Lotfi, S., & Izadkhah, H. (2021). Community-guided link prediction in multiplex networks. Journal of Informetrics, 15(4). https://doi.org/10.1016/j.joi.2021.101178
23. Gao, H., Li, B., Xie, W., Zhang, Y., Guan, D., Chen, W., & Cai, K. (2021). CSIP: Enhanced Link Prediction with Context of Social Influence Propagation. Big Data Research, 24.
24. https://doi.org/10.1016/j.bdr.2021.100217
25. Chen, L., Gao, M., Li, B., Liu, W., & Chen, B. (2018). Detect potential relations by link prediction in multi-relational social networks. Decision Support Systems, 115, 78-91.
26. https://doi.org/10.1016/j.dss.2018.09.006
27. Verma, A., Sardana, N., & Lal, S. (2020). Developer Recommendation for Stack Exchange Software Engineering Q & A Website based on K-Means clustering and Developer Social Network Metric. Procedia Computer Science, 167, 1665-1674. https://doi.org/10.1016/j.procs.2020.03.377
28. Christoforou, C., Malerou, K., Tsitsas, N. L., & Vakali, A. (2021). DIFCURV: A unified framework for Diffusion Curve Fitting and prediction in Online Social Networks. Array, 12.
29. https://doi.org/10.1016/j.array.2021.100100.
30. Bai, S., Zhang, Y., Li, L., Shan, N., & Chen, X. (2021). Effective link prediction in multiplex networks: A TOPSIS method. Expert Systems with Applications, 177.
31. https://doi.org/10.1016/j.eswa.2021.114973

32. Malla, S., M. J. . Meena, O. . Reddy. R, V. . Mahalakshmi, and A. . Balobaid. “A Study on Fish Classification Techniques Using Convolutional Neural Networks on Highly Challenged Underwater Images”. International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 01-09, doi:10.17762/ijritcc.v10i4.5524.
33. Zhang, Z., Wen, J., Sun, L., Deng, Q., Su, S., & Yao, P. (2017). Efficient incremental dynamic link prediction algorithms in social network. Knowledge-Based Systems, 132, 226-235.
34. https://doi.org/10.1016/j.knosys.2017.06.035
35. Mallek, S., Boukhris, I., Elouedi, Z., & Lefevre, E. (2019). Evidential link prediction in social networks based on structural and social information. Journal of computational science, 30, 98-107.
36. https://doi.org/10.1016/j.jocs.2018.11.009
37. Wang, Z., Liang, J., & Li, R. (2018). Exploiting user-to-user topic inclusion degree for link prediction in social-information networks. Expert Systems with Applications, 108, 143-158.
38. https://doi.org/10.1016/j.eswa.2018.04.034
39. Bütün, E., Kaya, M., & Alhajj, R. (2018). Extension of neighbor-based link prediction methods for directed, weighted and temporal social networks. Information Sciences, 463, 152-165.
40. https://doi.org/10.1016/j.ins.2018.06.051
41. Symeonidis, P., Iakovidou, N., Mantas, N., & Manolopoulos, Y. (2013). From biological to social networks: Link prediction based on multi-way spectral clustering. Data & Knowledge Engineering, 87, 226-242. https://doi.org/10.1016/j.datak.2013.05.008
42. Xu, S., Pi, D., Cao, J., & Fu, X. (2021). Hierarchical temporal–spatial preference modeling for user consumption location prediction in Geo-Social Networks. Information Processing & Management, 58(6). https://doi.org/10.1016/j.ipm.2021.102715
43. Wahid-Ul-Ashraf, A., Budka, M., & Musial, K. (2019). How to predict social relationships—Physics-inspired approach to link prediction. Physica A: Statistical Mechanics and its Applications, 523, 1110-1129.
44. https://doi.org/10.1016/j.physa.2019.04.246
45. Wu, Z., Lin, Y., Zhao, Y., & Yan, H. (2018). Improving local clustering based top-L link prediction methods via asymmetric link clustering information. Physica A: Statistical Mechanics and Its Applications, 492, 1859-1874. https://doi.org/10.1016/j.physa.2017.11.103
46. Pang, S., Wang, J., & Xia, L. (2022). Information matching model and multi-angle tracking algorithm for loan loss-linking customers based on the family mobile social-contact big data network. Information Processing & Management, 59(1). https://doi.org/10.1016/j.ipm.2021.102742
47. Tang, R., Jiang, S., Chen, X., Wang, H., Wang, W., & Wang, W. (2020). Interlayer link prediction in multiplex social networks: an iterative degree penalty algorithm. Knowledge-Based Systems, 194.
48. https://doi.org/10.1016/j.knosys.2020.105598
49. Ma, X., Tan, S., Xie, X., Zhong, X., & Deng, J. (2022). Joint multi-label learning and feature extraction for temporal link prediction. Pattern Recognition, 121.

50. https://doi.org/10.1016/j.patcog.2021.108216
51. Kumar, A., Mishra, S., Singh, S. S., Singh, K., & Biswas, B. (2020). Link prediction in complex networks based on significance of higher-order path index (shopi). Physica A: Statistical Mechanics and its Applications, 545. https://doi.org/10.1016/j.physa.2019.123790
52. Zou, L., Wang, C., Zeng, A., Fan, Y., & Di, Z. (2021). Link prediction in growing networks with aging. Social Networks, 65, 1-7. https://doi.org/10.1016/j.socnet.2020.11.001
53. Wang, X., Chai, Y., Li, H., & Wu, D. (2021). Link prediction in heterogeneous information networks: An improved deep graph convolution approach. Decision Support Systems, 141.
54. https://doi.org/10.1016/j.dss.2020.113448.
55. Alaria, S. K., A. . Raj, V. Sharma, and V. Kumar. "Simulation and Analysis of Hand Gesture Recognition for Indian Sign Language Using CNN". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 10-14, doi:10.17762/ijritcc.v10i4.5556.