Power System Technology
ISSN:1000-3673

Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

A New Method for Prediction of Future Links in Social Networks Using Data Preprocessing by Mining Tools

Seyyede Masoume Ahmadi Shakib[1], Davood Karimzadgan Moghadam[2*], Mohammad Reza Sanaei[3]

1- Department of Management and Accounting, Qazvin Branch, Islamic Azad University, Qazvin, Iran.
Email: sm.ahmadishakib@gmail.com
2- Associate Professor and member of the board of directors at Payam Noor University, Tehran, Iran.
Email: d_karimzadgan@pnu.ac.ir (Corresponding author)*
3- Department of Management and Accounting, Qazvin Branch, Islamic Azad University, Qazvin, Iran.
Email: mohamadrezasanaei@gmail.com

Abstract:- Social networks are primarily represented and analyzed in the form of graphs with a large number of vertices and edges in the form of an adjacency matrix. Edges represent relationships between individuals and act as links between vertices. The structural properties of each network are determined by the features of the edges and vertices within it. In this research conducted on various types of social networks data from the Stanford University database, preprocessing method using a competitive colonial algorithm was employed for feature selection operations, selecting features with the highest competence (lowest cost). To evaluate the impact of feature selection on the final output, experiments were conducted with and without feature selection operations using different algorithms commonly used in this field. Valid indices such as accuracy, detection, sensitivity, and major were independently measured on the output results with an average of 10 program executions. Comparing the results between scenarios with and without feature selection showed a significant impact on all final result indicators. Many features in the datasets were either unused or contained minimal information. Not removing these features not only increased computational burden but also affected the accuracy of the output results due to time-consuming executions.

Keywords*: Link prediction, meta-heuristic algorithms, data preprocessing, big data problems, computer social networks.*

1. Introduction

Social networks are considered a form of social media that have enabled a new way of establishing communication and sharing content on the internet [1]. These networks are a popular method for interpreting interactions between people. They can be visualized as graphs; where each node represents a person and the edges indicate the relationship between them.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

These connections are usually formed based on mutual interests. However, social networks are highly dynamic as new edges and nodes are added to the graph over time. Understanding the dynamics that drive the evolution of social networks is a complex issue due to the numerous variable parameters, but a relatively simpler issue is understanding the relationship between two specific nodes [2]. In real life, individuals are not independent of each other. They interact and influence each other mutually. Focusing solely on individual characteristics while disregarding relationships between individuals certainly impacts the accuracy and comprehensiveness of analysis. A social network represents the relationship between social entities (such as each individual in a social group). Analyzing social networks focuses on explaining hidden patterns and the impacts of these relationships, based on the assumption that individuals in social groups are interdependent, not independent units [3].

Predicting relationships in large-scale social networks is challenging due to the dynamic nature of network structures, where links and nodes are added to the network over time. One of the key parameters influencing network dynamics is the type and strength of relationships existing within it [4].

In the following, we will discuss the proposed solution, focusing on utilizing machine learning and metaheuristic optimization algorithms.

2. Challenges and Goals

In this research, a new method for predicting links in large social networks is introduced.

This method solves two basic challenges in this matter.

- Processes very large data and saves useful items (optimizes).

- Provides a powerful predictive model to solve such problems with multiple classes and classifications.

In fact, by using machine methods, it tries to improve the accuracy and accuracy of calculations in a reasonable and optimal time.

3. Comparison of Related Works

In the table below, the researches that have been done so far regarding link prediction are given and their strengths and weaknesses are compared to indicate the necessity of this research and providing a new method.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

Table 1. Summary of various related works with strengths and limitations

Ref./ Year	Proposed method	Strength	Limitations
[5] 2021	pectral clustering	▪ Reduce prediction error	▪ Need initialization
[11] 2021	Algorithm based on tensile contraction	▪ High accuracy	▪ High time complexity
[12] 2021	Neural Networks	▪ Improve prediction accuracy	▪ Need initialization
[13] 2020	Two-phase spatial adaptation mechanism	▪ Suitable for large social networks	▪ Requires prior knowledge
[14] 2020	Iterative degree algorithm	▪ Suitable for multiple networks	▪ Time consuming
[6] 2019	Particle swarm algorithm	▪ Increase the accuracy of the model	▪ Need initialization
[15] 2019	Use the DAG graph	▪ Appropriate time order	▪ Not suitable for large networks
[7] 2018	The criterion of similarity in the distribution of vertices in motifs	▪ Powerful prediction model	▪ The problem of blocking motifs
[8] 2018	Evolutionary algorithm of ants	▪ High algorithm speed ▪ Appropriate Final response	▪ The need for initialization ▪ The possibility of not getting an answer
[9] 2018	Community discovery method	▪ The certainty of the final answer according to the previous information	▪ Requires previous information
[16] 2017	Fast graph clustering	▪ Suitable for large-scale networks	▪ Time consuming model
[17] 2017	Use time the variable	▪ Increasing prediction accuracy using structure and number of common neighbors	▪ Instability of time variable investigation
[10] 2016	k-medoid clustering	▪ High accuracy of the prediction model	▪ Initialization of cluster centers

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

4. Implementation of the proposed method

The proposed method to improve link prediction results is based on machine learning and optimization science, which consists of four different steps as shown in Fig. 1.

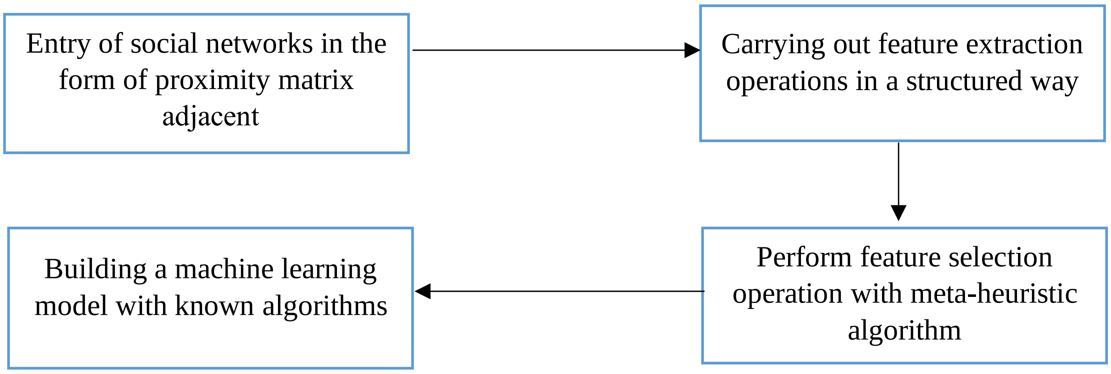

Fig.1. Steps of theproposed method

In summary, the steps of implementing the proposed method, reaching the output, and ultimately evaluating it can be titled as follows:

1. Entering the initial dataset and separating training and testing data:

- The initial dataset is entered into the system and randomly divided into two categories. The training and testing data are separated. The data is split with a 30 to 70 probability. 30% is allocated to testing data and 70% to training data. The training data is used for model building, and the testing data is used for model validation.

2. Extracting features using structural methods:

- Features related to each pair of vertices are extracted using the defined relationships in Table 2.

- The features are normalized to achieve standard values.

3. Applying the colonial competition algorithm to select the best features:

- Form the initial population by randomly selecting features.

- Determine colonies and colonizers based on competence and assign colonies to colonizers.

- If a colony is more competent in selecting features than a colonizer, swap their positions.

- Calculate the total cost of the empire.

- Weak colonies in an empire, in terms of competence in selecting features, should be assigned to stronger colonizers.

- Eliminate weak empires.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

4. Building a classifier model using the training data:

- Data is trained using Support Vector Machine and four other algorithms including Random Forest, Decision Tree, Bayes, and KNN.

5. Using testing data for validation:

- In this section, validation of the results obtained in the fourth step is carried out using the testing data.

Table 2. Structural features related to pairs of social network graph vertices

Feature name	**Formula**
Vertices	$d(v) = \|\rho(v)\|$ (1) $\rho(v) = \{u\|(u.v)\epsilon\ E\ or\ (v.u)\epsilon E\}$ (2)
The following features of the Ross graph	$nh - subgraph(v) = \{(x.y)\epsilon E\|x.y\ \epsilon\ \rho(v)\}$ (3) $subgraph - Edge - Number(v) = \|nh - subgraph(v)\|$ (4) $Densiry - nh - subgraph(v) = {d(v)}/{\|nh - subgraph(v)\|}$ (5)
Features of common friends[1]	$common - friends(u.v) = \|\rho(v) \cap \rho(u)\|$ (6)
Adams/ Adar Index	$adamic - adar - index(u.v) = \sum_{z\epsilon\rho(v)\cap\rho(u)} \frac{1}{\log\ \|\rho(z)\|}$ (7)
The benchmark of all friends	$total - friends(u.v) = \|\rho(v) \cup \rho(u)\|$ (8)
Jacquard coefficient	$jaccard - coefficient(u.v) = \frac{\|\rho(v)\cap\rho(u)\|}{\|\rho(v)\cup\rho(u)\|}$ (9)
Hub promoted index (HPI)	$HPI(u.v) = \frac{\|\rho(v)\cap\rho(u)\|}{\min\{\rho(v).\rho(u)\}}$ (10)
Compressed-up hub index (HDI)[2]	$HDI(u.v) = \frac{\|\rho(v)\cap\rho(u)\|}{\max\{\rho(v).\rho(u)\}}$ (11)
Salton index	$Salton - index(u.v) = \frac{\|\rho(v)\cap\rho(u)\|}{\sqrt{\|\rho(v)\|\times\|\rho(u)\|}}$ (12)

[1] Common Friends
[2] Hub Depressed index

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

Sorensen index	$sorenson - index(u.v) = \frac{\lvert\rho(v)\cap\rho(u)\rvert}{\lvert\rho(v)\rvert+\lvert\rho(u)\rvert}$ (13)
LHN index[3]	$LHN(u.v) = \frac{\lvert\rho(v)\cap\rho(u)\rvert}{\lvert\rho(v)\rvert\times\lvert\rho(u)\rvert}$ (14)
RA index[4]	$RA - index(u.v) = \sum_{z\epsilon\rho(v)\cap\rho(u)} \frac{1}{\rho(z)}$ (15)
Preferential attachment rating[5]	$preferential - attachment(u.v) = \lvert\rho(v)\rvert \times \lvert\rho(u)\rvert$ (16)
The benchmark of friends	$freinds - measure(u.v) = \sum_{x\epsilon\rho(u)} \sum_{y\epsilon\rho(v)} \delta(x.y)$ (17) $\delta(x.y) = \begin{cases} 1 & if\ x = y \\ 0 & otherwise \end{cases}$, $w\ (u) = \frac{1}{\sqrt{1+\lvert\rho(u)\rvert}}$
Edge weight criterion	$w(v) = \frac{1}{\sqrt{1+\lvert\rho(v)\rvert}}$ (18)
Weight gain	$w(u.v) = w(u) + w(v)$ (19)
Weight gain	$w(u.v) = w(u) \times w(v)$ (20)
Clustering coefficient	$clustering - coefficient(u.v) = \frac{\sum_{z\neq u.v} a_{uz} a_{zv}}{\min\{\rho(v).\rho(u)\}}$ (21)
Degree correlation	$degree - correlation(u.v) = \frac{4.\lvert\rho(v)\rvert.\lvert\rho(u)\rvert - \lvert\rho(v)\rvert - \lvert\rho(u)\rvert}{2.\lvert\rho(v)\rvert^2 + 2.\lvert\rho(u)\rvert^2 - \lvert\rho(v)\rvert - \lvert\rho(u)\rvert}$ (22)
FriendLink method	$friend - link(u.v) = \sum_{i=2}^{l} \frac{1}{i-1} . \frac{\lvert paths_{u.v}^{i}\rvert}{\prod_{j=2}^{i} n-j}$ (23)

4.1 First step: Dataset

A dataset is used to refer to data in closely related tables related to a particular experiment or event. Since link prediction methods are dependent on the structure and type of network, it is not possible to provide a dominant link prediction method for all networks. Therefore, the appropriate selection of different data sets can help to solve this defect.

[3] Leichet-Holme-Newman index
[4] Resource Allocation
[5] Preferential attachment Score

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

The datasets used in this research are shown in Table 3.

Table3. Data details used in the research

Network	Header count	Number of edges
City Pension Berlin	484,413	1,751,463
Hep-ph	34,546	421,578
Wiki	167,525	1,164,576
Emails	167	5,784
Movielens	2,625	100,000
Read more	1,574	28,236
Twitter	81,306	1,768,149

4.2 Second step: Extract the feature

Prediction of important links in social networks is crucial because the accuracy of link prediction heavily relies on the features extracted from the raw data. Feature extraction is a process where prominent and distinctive features are identified by performing operations on the data. The aim of feature extraction is to make raw data more usable for subsequent statistical processing. The higher quality of the extracted features, the more powerful the predictive model will be; in other words, this step directly impacts the final outcome. Selecting the appropriate method for feature extraction is also considered a crucial matter [18].

Considering the above, this section presents the features extracted from the social network graph used in this study. If G=(V,E) is a graph representing the topological structure of a social network, each edge of the graph, denoted by e=(u,v)∈E, where u,v∈V, illustrates a connection. Our goal is to create a suitable classification using machine learning techniques so that for every pair of vertices u,v, it can predict whether there is a high probability of a connection between them or not. Therefore, for each candidate edge for classification, a set of features extracted from the topological structure of the network is considered. These features are presented in Table 2.

4.3 Third step: Select the features

Feature selection is a crucial issue in machine learning and statistical pattern recognition. In many applications such as classification, the presence of numerous features, many of which are either redundant or irrelevant, poses a significant challenge. While eliminating these features does not compromise the information, it increases the computational burden for the

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

intended application. Moreover, it results in storing a considerable amount of uninformative data along with useful information [19].

Various solutions and algorithms have been proposed for feature selection, some of which have a history of thirty to forty years. Some algorithms faced high computational costs when initially introduced, which is no longer a concern today due to the emergence of fast computers and large storage resources. However, the rise of very large datasets for new problems highlights the ongoing importance of finding a fast and cost-effective algorithm for this task.

Many classical feature selection methods aim to find the best subset among 2^N candidate subsets. In all these methods, a subset is selected as the answer based on its application and definition type, capable of optimizing an evaluation function. Despite each method striving to select the best features, the vast number of possible solutions, which grow exponentially with N, makes finding the optimal solution challenging and highly costly for average and large N values.

In general, different feature selection methods categorize based on the type of search into various classifications. Some methods explore the entire search space, while others, which can be heuristic or random search, reduce the search space at the cost of losing some performance.

In general, this issue does not have a definitive solution, and so far, no precise method has been proposed to solve it. Various classic approaches have been suggested for these problems, which generally do not provide very suitable and desirable answers; however, intelligent optimization methods can offer significantly better solutions. Therefore, one of the effective and constructive methods in solving problems is selecting features and related issues by using metaheuristic optimization methods and evolutionary algorithms.

The Imperialist Competitive Algorithm (ICA) is a method in the field of evolutionary computations that focuses on finding optimal solutions for various optimization problems. This algorithm presents a mathematical model of social-political evolution processes, offering an algorithm for solving mathematical optimization problems. The main principles of this algorithm are synchronization policy, imperialist competition, and revolution [20]. By mimicking the process of social, economic, and political evolution of countries and mathematically modeling parts of this process, the algorithm provides operators in an organized algorithmic form that can assist in solving complex optimization problems. In essence, this algorithm observes optimization problem solutions in the form of countries and attempts to gradually improve these solutions through a repetitive process, ultimately reaching the optimal solution of the problem. Similar to other evolutionary optimization methods, this algorithm also starts with an initial population. In summary, in this algorithm, each element of the population is referred to as a country. Countries are divided into two categories: colonies and imperialists. Depending on their power, each imperialist brings some of the colonies under

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

their control. The attraction policy and imperialist competition form the core of this algorithm [20].

Contrary to other classic algorithms, which have an exponential state space of 2^N and significantly increase computation time, making computations complex, this method has a polynomial state space of N^2.

To implement the proposed method, the adjacency matrix of the social network graph is first applied as input to the program.

4.3.1 Stages of Colonial Competitive Algorithm

1. Initial parameterization: In this stage, parameters of the algorithm such as initial countries are initialized. Each created node in the network is assigned to one of the countries as the initial location of the countries [21].

In this study, the initial parameters of the colonial competitive algorithm are initialized as follows. It is worth mentioning that the initialization of these parameters can be done in two ways according to the routine of metaheuristic algorithms: following similar articles, which is not highly recommended due to differences in data and their structural effects on results, or trial and error method, which is the most common method considering the uncertain nature of such algorithms. In this study, the trial and error method was employed.

Table4. Determining the initial parameters of the colonial competition algorithm

Parameter	Value
Number of primary population	200
Number of program repetitions	1000
The number of empires	30
The number of colonies	170
Beta parameter (colonies moveto empires)	0.2
Zeta parameter (average colonial power calculation rate)	0.1
Revolution rate	0.1

Then, based on the competence level of each node (in this context, the competence level of each node is the similarity to other nodes), the colonizer and colony are determined.

The competence level of each node is calculated through a competence function. The competence function can be determined in various ways. One of these methods is the random

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

method. In this method, a number of features are randomly selected each time, the algorithm is executed on them, and the output is obtained. Ultimately, the features that achieve the desired output with the least cost are selected. Another method is utilizing relationships. In this method, the intra-class dependencies of the features are calculated, and among them, the features with the least dependency on other features are removed, while the features with the highest interdependency remain. This topic is known as pattern recognition in the field of artificial intelligence.

In this research, for achieving maximum efficiency, the second method is used to determine the competence level of features.

The competence function used in the competitive colonization algorithm in this research determines the value of each feature in the feature vector and ultimately, the most valuable features are used for the final classification operation. Equation 1 represents this function.

$$f(x) = Mu^{+} + Mu^{-}/Var^{+} + Var^{-} \qquad (1)$$

According to Equation 1, Mu^{+} represents the average of positive features, Mu^{-} indicates the average of negative features, Var^{+} denotes the variance of positive features, and Var^{-} represents the variance of negative features.

After substituting the features into the above equation and calculating their competence function, the algorithm execution steps continue.

In Fig. 2., the execution of this algorithm with 50 iterations is shown

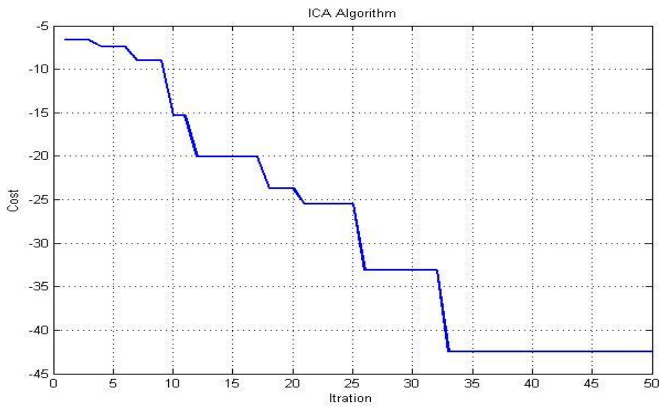

Fig. 2. Movement chart of colonial competition algorithm

In Fig. 2., the X-axis represents the number of program iterations, and the Y-axis represents the calculated cost by the competence function obtained from the selected features in each corresponding iteration. This chart illustrates the best value obtained over all iterations. It is evident that moving the algorithm forward leads to better results.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

2. Colonies are moved towards the colonial power (assimilation policy): Based on the power of each colonial power (competence level), colonies are assigned to the colonizers. Initially, this assignment will be equal among all colonizers. As a result, the nodes in the network are divided into two categories based on competence and form several empires.

3. If a colony within an empire has a lower cost than the imperialist, swap the positions of the colony and the imperialist. After executing the algorithm and desired changes in each country, the competence level of each country will change. If the competence level of a colony exceeds that of the colonizer, the positions of these two will be swapped. Each node, based on competence and similarity level, will change its position within the empire when selected.

4. Calculate the total cost of an empire (considering the cost of the colonizer and its colonies): To understand the power of each empire, the total cost must be calculated. The cost of each empire, consisting of multiple nodes in the network, is calculated.

5. Select a colony from the weakest empire and give it to the empire with the highest probability of acquisition: After calculating the power of each empire, stronger colonizers acquire colonies from weaker colonial empires. These exchanges take place between the nodes in the network.

6. Eliminate weak empires: The weakest empire is removed, and the colonizer of this empire is considered a colony.

7. Repeat these steps until the end of the program iterations: Upon completion of the program iterations, the strongest empire remains, and the node assigned to it is the selected node for clustering.

The flow of this algorithm is illustrated in Fig. 3. [21].

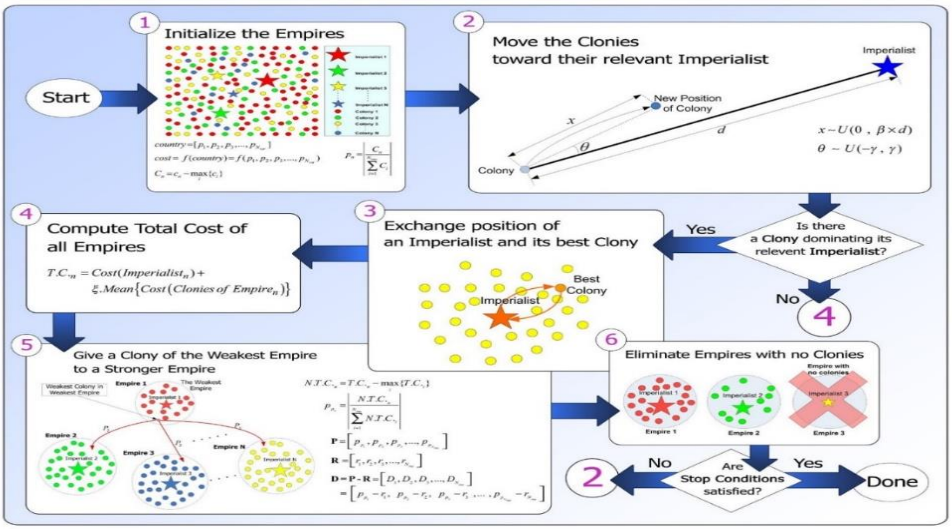

Fig. 3. An example of the order of the colonial competition algorithm

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

4.4 Forth step: Training the system with selected data and features

In this step, the link prediction operation is performed on the datasets based on the obtained features. In this research, this operation was performed once by applying the "feature selection" step and again without applying it. In addition, this method was repeated using several different algorithms that are well-known and widely used in this field, such as: KNN, Random Forest, SVM, Decision Tree, Bayes.

The pseudo-code of the algorithm is as follows:

1. Separate training and testing data
2. Operation ICA Algorithm

- For i=1 to max iteration
- Initialize the parameters used
- Create a Primary Population
- Assimilation: Colony Move toward imperialist
- Revolution: Random Chang occur in the characteristics of some country
- Exchange Between City and Country
- Imperialist Competitions: all imperialist compete to take position of colonies of each others
- End for
- Select best output and use it

3. Train Support Vector Machine and create model
4. Use test data and Validation model
5. Voting Answers for Access Final Solution

5. Result and Analysis

In order to evaluate the effect of feature selection on the final output, the accuracy percentage as the main indicator and the sensitivity, specification, and F-Measure indicators were calculated in each run in order to increase the reliability of the results, in the conditions with feature selection and without it, and the results are shown in Table 5. given.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

Table5. Comparison of results in feature selection mode without selecting a feature

Dataset	Criterion	Accuracy	F-Measure	Spe-cification	Sensitivity
CiteSeer	Feature Selection	87.62	88.59	90	87.22
	No Feature Selection	82.86	85.71	88.53	83.06
Hep-ph	Feature Selection	80.48	85	90	80.48
	No Feature Selection	88.89	84.12	90	88.89
Wiki	Feature Selection	87.06	88.51	85	87.06
	No Feature Selection	84.51	87.74	86	85.59
Emails	Feature Selection	81.48	84.54	88.85	81.11
	No Feature Selection	74.07	78.35	83.08	74.44
Movielens	Feature Selection	91.8	84.95	87.92	82.16
	No Feature Selection	85.25	81.09	82	80.2
US airport	Feature Selection	84.24	82.43	81.11	85.54
	No Feature Selection	79.89	78.03	74.07	83.71

By observing Table 5 and comparing the results between feature selection and non-feature selection scenarios, we find that the use of this operation can significantly impact the final results. In these datasets, there are numerous features, many of which are either irrelevant or carry minimal information. Not removing these features not only increases the computational burden for the problem at hand but also affects the accuracy of the output results due to the time it takes to execute. Furthermore, it leads to storing a large amount of uninformative data along with useful data. Therefore, utilizing this operation will be highly beneficial.

To gain a better understanding of the impact of feature selection operations, two charts have been designed for 5 well-known and widely used classification methods, displaying their impact regardless of the classification algorithm type. Through this experiment, the importance of feature selection as an independent factor in improving prediction processes is well demonstrated. Fig. 4. illustrates the accuracy percentage of each classification method when feature selection is used, while Fig. 5. shows the accuracy percentage of each classification method when feature selection is not used with different datasets.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

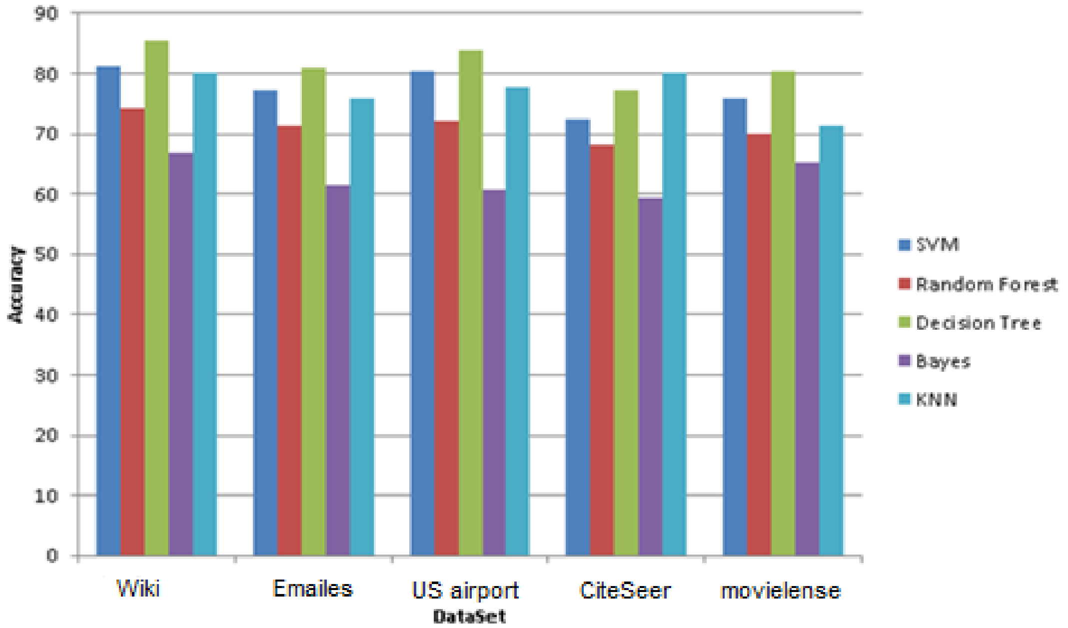

Fig. 4. Percentage of accuracy of different categorization methods if using feature selection operation

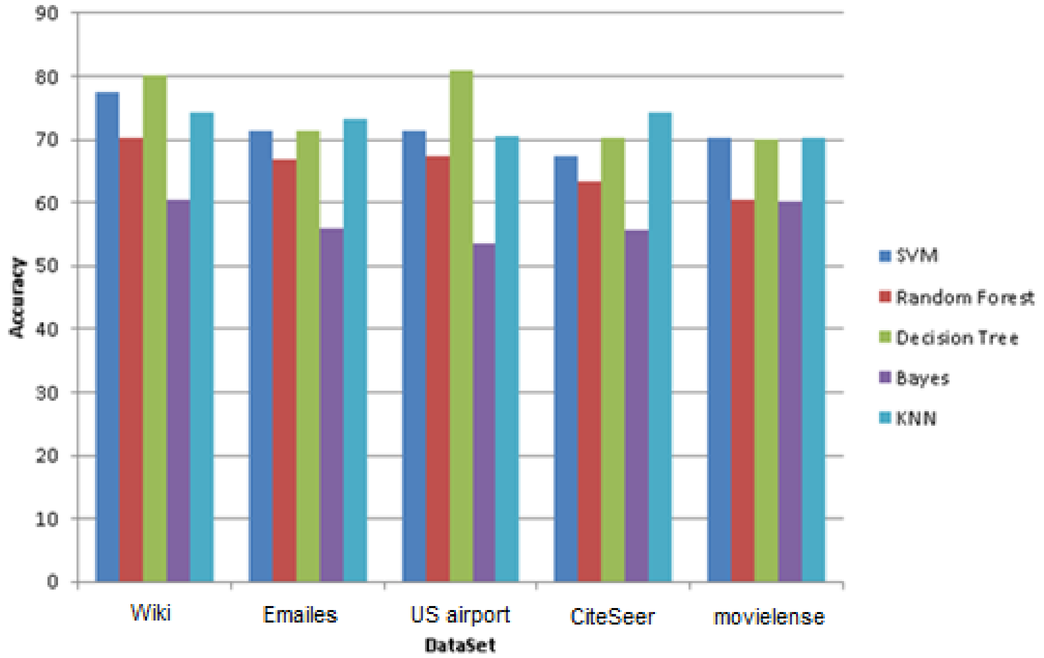

Fig. 5. Percentage of accuracy of different categorization methods if you do not use feature selection operations.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

As shown in Figs. 4. and 5., utilizing feature selection operations is highly effective and beneficial, significantly improving the results. Fig. 4. demonstrates better performance compared to Fig. 5., indicating the superiority of feature selection operations.

6. Conclusion

Due to numerous research studies on link prediction in the field of social networks, each addressing it with different algorithms that have had their own advantages and disadvantages, this research approached this challenge from a different perspective. The vast amount of data processed in this way not only falls within the realm of machine learning but also extends to big data issues. To minimize processing operations, a feature selection method was proposed beyond the algorithms that can be used for link prediction. This method significantly reduces the amount of processed information. Feature selection was carried out among all extracted features using a proposed metaheuristic algorithm, which led to a considerable improvement according to Table 5, as observed in the results. This improvement is noticeable through evaluation metrics such as Sensitivity, Specification, F-Measure, and Accuracy. Furthermore, to examine the impact of this proposal more precisely, a comparison was made with other algorithms including Random Forest, Decision Tree, Bayes, and KNN, the results of which are depicted in Figs. 4. and 5. Surprisingly, it shows that merely implementing this proposal will lead to significant improvements in other existing methods as well.

Refrences

[1] Backstrom, L., and J. Leskovec, Supervised random walks: predicting and recommending links in social networks, WSDM '11: Proceedings of the fourth ACM international conference on Web search and data mining , (2011), 635–644, https://doi.org/10.1145/1935826.1935914.

[2] Kumari, A., et al., Supervised link prediction using structured-based feature extraction in social network. Concurrency and Computation Practice and Experience 34(13) (2020), https://doi.org/10.1002/cpe.5839.

[3] Jagadishwari, V., and I. Accamma, CNJC: A cumulative similarity measure for Link Prediction in social networks, 9th International Conference on Advances in Computing and Communication (ICACC), IEEE. Kochi, India. (2019). https://doi.org/10.1109/ICACC48162.2019.8986211.

[4] Bastami, E., Mahabadi, A., Taghizadeh, E., A gravitation-based link prediction approach in social networks, Swarm and Evolutionary Computation 44 (2019) 176–186 . https://doi.org/10.1016/j.swevo.2018.03.001.

[5] Zare, H., & Shakorzadeh, H., Link prediction in social networks using clustering method with maximum expectation maximization algorithm, Fourth National Conference on

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

Advanced Technologies in Electrical, Computer, and Mechanical Engineering in Iran. (2021). https://civilica.com/doc/1292870.
[6] Piltan, Y., & Mojarrad, M., Introducing a new method for link prediction in social networks based on metaheuristic algorithms. Kahraba Quarterly. 6(25) (2019). https://civilica.com/doc/1442443.
[7] Aghabozorgi, F., and M. R. Khayyambashi, A new similarity measure for link prediction based on local structures in social networks, Physica A: Statistical Mechanics and its Applications. 501(2018), 12-23.
[8] Sherkat, E., Rahgozar, M., Asadpour, M., Ant Colony Approach to Link Prediction in Social Networks, The CSI Journal on Computer Science and Engineering 12 (1) (2014) . https://doi.org/10.1016/j.swevo.2018.03.001.
[9] Moghadam, S. M. J., & Rezaei, S., Link prediction by community detection in social networks, Third National Conference on Computer Engineering, Information Technology, and Data Processing. (2018). https://civilica.com/doc/854073.
[10] Parvazeh, F., A. Harounabadi, and M. A. Naizari, A Recommender System for Making Friendship in Social Networks Using Graph Theory and users profile, Journal of Current Research in Science. 1(2016), 535.
[11] Yan, R., et al., SSDBA: the stretch shrink distance based algorithm for link prediction in social networks, Frontiers of Computer Science, 15(1) (2021), 1-8.
[12] Sharma, A., et al., Link Prediction in Social Network using Artificial Neural Network. Int. J. Comput. Appl. 174 (2021), 26-30.
[13] Chen, H., et al., Multi-level graph convolutional networks for cross-platform anchor link prediction, Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining (2020), 1503–1511. https://doi.org/10.1145/3394486.3403201.
[14] Tang, R., et al., Interlayer link prediction in multiplex social networks: an iterative degree penalty algorithm, Knowledge-Based Systems. 194 (2020). https://doi.org/10.1016/j.knosys.2020.105598.
[15] Ahuja, R., et al., Using hierarchies in online social networks to determine link prediction, Soft Computing and Signal Processing, Springer (2019), 67-76.
[16] Bai, L., X. Cheng, J. Liang, and Y. Guo, Fast graph clustering with a new description model for community detection, Information Sciences. 388- 389(2017), 37-47.
[17] Shang, K., M. Small, X. K. Xu, and W. S. Yan, The role of direct links for link prediction in evolving networks. EPL (Europhysics Letters). 117 (2017), 2002.
[18] Siyari Fard, S., Presenting a new method for link prediction in social networks using deep learning techniques, Third Conference on Electrical, Computer, and Mechanical Engineering, (2020), https://civilica.com/doc/118916.

Power System Technology
ISSN:1000-3673
Received: 06-06-2024 Revised: 15-07-2024 Accepted: 28-08-2024

[19] Shaeghi, H., & A. Ghasemi, Modeling a multi-input multi-output system for simultaneous prediction of price and load in intelligent networks with load management, Computational Intelligence in Electrical Engineering, 6 (4) (2015), 1-87.
[20] Luo, J., J. Zhou, and X. Jiang, A Modification of the Imperialist Competitive Algorithm With Hybrid Methods for Constrained Optimization Problems, Journal of PeerJ Computer Science. 8 (2021), e1075. https://doi.org/10.7717/peerj-cs.1075.
[21] Tao, Xin-rui., J.-Q. Li, T.-H. Huang and P. Duan, Discrete imperialist competitive algorithm for the resource-constrained hybrid flowshop problem with energy consumption, Journal of Complex & Intelligent Systems. 7(2021), 311–326. https://doi.org/10.1007/s40747-020-00193-w.