VOLUME XII
ISBN No.: 978-93-94004-01-6
Physical Science

NALLAMUTHU GOUNDER MAHALINGAM COLLEGE

An Autonomous Institution, Affiliated to Bharathiar University, An ISO 9001:2015 Certified Institution,

Pollachi-642001

SUPPORTED BY

ABT Industries Ltd.,
Dairy Division

PROCEEDING

One day International Conference

EMERGING TRENDS IN SCIENCE AND TECHNOLOGY (ETIST-2021)

27th October 2021

Jointly Organized by

Department of Biological Science, Physical Science and Computational Science

NALLAMUTHU GOUNDER MAHALINGAM COLLEGE

An Autonomous Institution, Affiliated to Bharathiar University

An ISO 9001:2015 Certified Institution, Pollachi-642001.

Estd. 1957

Proceeding of the

One day International Conference on

EMERGING TRENDS IN SCIENCE AND TECHNOLOGY (ETIST-2021)

27th October 2021

Jointly Organized by

Department of Biological Science, Physical Science and Computational Science

ISBN No: 978-93-94004-01-6

978- 93- 94004- 01- 6

Nallamuthu Gounder Mahalingam College

An Autonomous Institution, Affiliated to Bharathiar University

An ISO 9001:2015 Certified Institution, 90 Palghat Road, Pollachi-642001.

www.ngmc.org

ABOUT THE INSTITUTION

A nations's growth is in proportion to education and intelligence spread among the masses. Having this idealistic vision, two great philanthropists late. S.P. Nallamuthu Gounder and Late. Arutchelver Padmabhushan Dr.N.Mahalingam formed an organization called Pollachi Kalvi Kazhagam, which started NGM College in 1957, to impart holistic education with an objective to cater to the higher educational needs of those who wish to aspire for excellence in knowledge and values. The College has achieved greater academic distinctions with the introduction of autonomous system from the academic year 1987-88. The college has been Re-Accredited by NAAC and it is ISO 9001 : 2015 Certified Institution. The total student strength is around 6000. Having celebrated its Diamond Jubilee in 2017, the college has blossomed into a premier Post-Graduate and Research Institution, offering 26 UG, 12 PG, 13 M.Phil and 10 Ph.D Programmes, apart from Diploma and Certificate Courses. The college has been ranked within Top 100 (72nd Rank) in India by NIRF 2021.

ABOUT CONFERENCE

The International conference on “Emerging Trends in Science and Technology (ETIST-2021)” is being jointly organized by Departments of Biological Science, Physical Science and Computational Science - Nallamuthu Gounder Mahalingam College, Pollachi along with ISTE, CSI, IETE, IEE & RIYASA LABS on 27th OCT 2021. The Conference will provide common platform for faculties, research scholars, industrialists to exchange and discus the innovative ideas and will promote to work in interdisciplinary mode.

 ISBN No.: 978-93-94004-01-6

Feature Selection and Classification from the Graph using Neural Network based Constructive Learning Approach

Dr. A. SangeethaDevi[1] ,Mrs. A. Kalaivani[2] and Mrs. A. Shanmugapriya[3]

[1] Assistant Professor, Department of Science and Humanities (Mathematics), P. A. College of Engineering and Technology, Pollachi, Tamilnadu, India

[2] Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Tamilnadu, India

[3] Assistant Professor, Department of ECE, Pollachi Institute of Engineering and Technology, Pollachi, Tamilnadu, India

[1]sangeethadevi.a@gmail.com, [2] kalaivanimathsca@gmail.com, [3]priyasamy30@gmail.com

Contact No: 9865265528, 8838918225, 9659835805

Abstract: The real-world knowledge are signified by knowledge graph that gives assistance for diverse applications developed based on artificial intelligence. The knowledge regarding the neighborhood is learned from the entities and relationships of the knowledge graph. Analysis of data on high dimension is a challenging task in many applications and this article addresses the dimensionality by defining a small set of features that signifies the high dimensional data without noticeable or significant loss of data. A learning based unsupervised learning approach that utilizes the neural network concept and it learns the feature from the graph.In this paper, a Constructive Feature Selection Approach using Neural Network (CFSNN) approach identifies the featuresfrom the graph and analyzes the performance of the proposed method. The performance of the CFSNN isevaluated by comparing with the existing classification approaches and uses different datasets. The performance is evaluated using the classification performance metrics and from the observation it is identified that the proposed CFSNN algorithm has best outcome.
Keywords: Graph mining, high dimensional data, feature selection, constructive, and neural network.

1. Introduction

The knowledge graph is a variety of data structure and it utilizes the structural properties of graph [1,2]. The node edge node is used in representing the any semantic network that is concept or entity. The edges in the graph is used in representing the relationship between the entities in the graph. The inadequacy of material in the knowledge graph indicates the incompleteness of information that makes processing of data as imperfect [3]. The sparseness of the data makes the data processing as partial and the process of finishing the missing value is a monotonous process [4]. The accomplishment of generating true relation and identifying the missing values in knowledge graph is a vibrant research area [5].

In today's world decision making and data processing is a complicated process with the constantly growing volume of data [6]. The big data applications are more complicated to process and bigger to manage and hence, traditional approaches are ineffective. The availability of large scale data made numerous challenges in processing and most of the existing algorithms are developed for low-dimensional space of data [7, 8]. As a result of this, application with huge volume of data that is big data based application necessitated effective approaches and the traditional approaches are ineffective in handling. The high dimensionality of the can be handled effectively by the deep learning approaches where the feature selection approaches provide the strategy of preparing the high dimensional data [9].

---

[1] Assistant Professor, Department of Science and Humanities (Mathematics), P. A. College of Engineering and Technology, Pollachi, Tamilnadu, India

[2] Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Tamilnadu, India

[3] Assistant Professor, Department of ECE, Pollachi Institute of Engineering and Technology, Pollachi, Tamilnadu, India

The available feature selection algorithms are developed based on the robust assumption where the features are distributed identically and the features are independent of every other information [10]. The feature selection algorithms disregard the intrinsic dependencies or structure between the features, the particular feature set may not efficiently signify the data [11]. For any occurrence, numerous problematic fields encompasses the feature spaces, which have pairwise dependences of data. In text mining or natural language processing applications, every feature is measured as a term or word, and those that have resemblance with other words are called synonyms [12].

Furthermore, in the field of biology and its applications, certain genes work in groups in which the occurrence of interdependencies amongst genes [13, 14]. The process of analysis high dimensional data is an exciting task in numerous applications and this paper addresses the dimensionality by describing a small set of features that indicates the high dimensional data without significant or noticeable loss of data. The main intent of the proposed approach, Constructive Feature Selection Approach using Neural Network (CFSNN) is categorizing the features. The performance of the proposed work is analyze by the performance metrics namely accuracy, precision, recall, f-measure and error rate.Learning the demonstrations of nodes in a graph or network with conserving definite possessions of the network is constructive for many analysis and it has fascinated substantial consideration in recent years [15].

The rest of the paper is emphasised as follows: existing feature selection and classification approaches are discussed in section 2, detailed the proposed Constructive Feature Selection Approach using Neural Network approach in section 3, investigation of CFSNN is illustrated in section 4 and proposed work is concluded with future idea in section 5.

2. Related Works

The intrinsic features are used in many significant applications and the relationship among the features are retrieved by feature selection algorithms, Classification issue on multivariate variable is rectified by the logistic regression [16] approach. The process is complicated and it doesn’t consider the additional data about the feature. Feature selection plays a significant role in minimizing the high-dimensionality. A graph-based algorithm for feature selection is introduced, which assigns ranks for features and it identifies the most essential ones into indiscriminate set of cues.

Mapping the problematic feature on an empathy graph where significant features are assigned to the nodes and the solution is assumed by assessing the prominence of nodes via some indicators of centrality that is Eigen-vector Centrality (EC). The essence of EC is to evaluate the significance of a feature as a function of the prominence of its neighbors. The process of ranking the central nodes with individual candidate features, which gives prominent classification scheme [17].

In the process of data analysis, unsupervised learning scheme is used on the multi-cluster data that is Multi-Cluster Feature Selection (MCFS). The feature selection is attained via absence of class label and it give relevant information [18]. Long short-term memory (LSTM) is a process of shortening the gradient where this does not do influence the value of LSTM. It can learn to associate the minimal time lags in additional value of 1000 discrete-time steps by implementing persistent error flow via persistent error carousels within distinct units [19].

Convolutional neural networks (CNN) trained on uppermost position of the pre-trained word vectors that is applied for sentence-level classification method. CNN utilises hyperparameter tuning and the static vectors attains exceptional results on multiple benchmarks [20]. A new unsupervised feature selection technique incorporates the k-influence space notion and subspace learning, which map the features onto a weighted graph and assigns the rank to them that is PageRank graph centrality measure. The designing of graph promotes downgrades redundancy, the feature relevance, and it is strong to outliers and cluster imbalances [21]. The shortcomings in the proposed scheme is rectified by the proposed approach.

3. Constructive Feature Selection Approach using Neural Network (CFSNN)

Effective feature selection for graph is accomplished by the wrapper methodin combination with incremental training method to identify a subset of a feature from the available set of features. To attain the best generalization of learning process and the CFSNN automatically decides count of the neurons during the process of feature selection where the hidden neuron determined by the incremental approach. The CFSNN is initialized by the minimum count of neuron and feature. During the incremental process, neurons and features are added whereas simple criteria is used to determine the addition of neuron and features.

Grouping the features

Based on the similarity, features are classified into groups with similar objects and the process is termed as clustering and it huge group of original clusters are needed. For grouping process, threshold value taken from user and handling the overall process is tedious. To overcome the shortcomings, CFSNN aims at identifying the relationship among the features. Through, the identification informative and distinct features for developing the robust feature selection model. The relationship among two variable is determined by the general statistics approach called correlation. In CFSNN approach, Pearson product moment correlation is applied to estimate the correlation measure among diverse features of training set. The coefficient of correlation p_{ij} among two features i and j is equated as,

$$p_{ij} = \frac{\sum_r (y_i - \bar{y}_i)(y_j - \bar{y}_j)}{\sqrt{(y_i - \bar{y}_i)^2}\sqrt{(y_j - \bar{y}_j)^2}}$$

where the values of the features i and j is signified by y_i and y_j respectively. The mean values of y_i and $y_j$is signified by the $\bar{y}_i$ and$\bar{y}_j$ which is averaged over the r value. Existence of exact linear dependency is determined by the complete correlation of values i and j whereby p_{ij}would be 1 or -1. If the values are entirely uncorrelated i and j is assigned with 0. After estimating the all combinations of correlation coefficient, every feature is sorted as descending order. The correlation value of every feature i is estimated by,

$$crln_i = \frac{\sum_{j=1}^{NF} |p_{ij}|}{NF - 1} \quad if\ i\ \neq j$$

where NF is the count of feature and it is incorporated in signifying a given dataset. At the end of the process, CFSNN is categorized into two groups namely NF/2 features of similar (SI) group and NF/2 features of dissimilar (DS) group. The first feature in the group SI is the most correlated and last feature in the group DS is the least correlated in the dataset. The overall process of the proposed CFSNN is described in Figure 1.

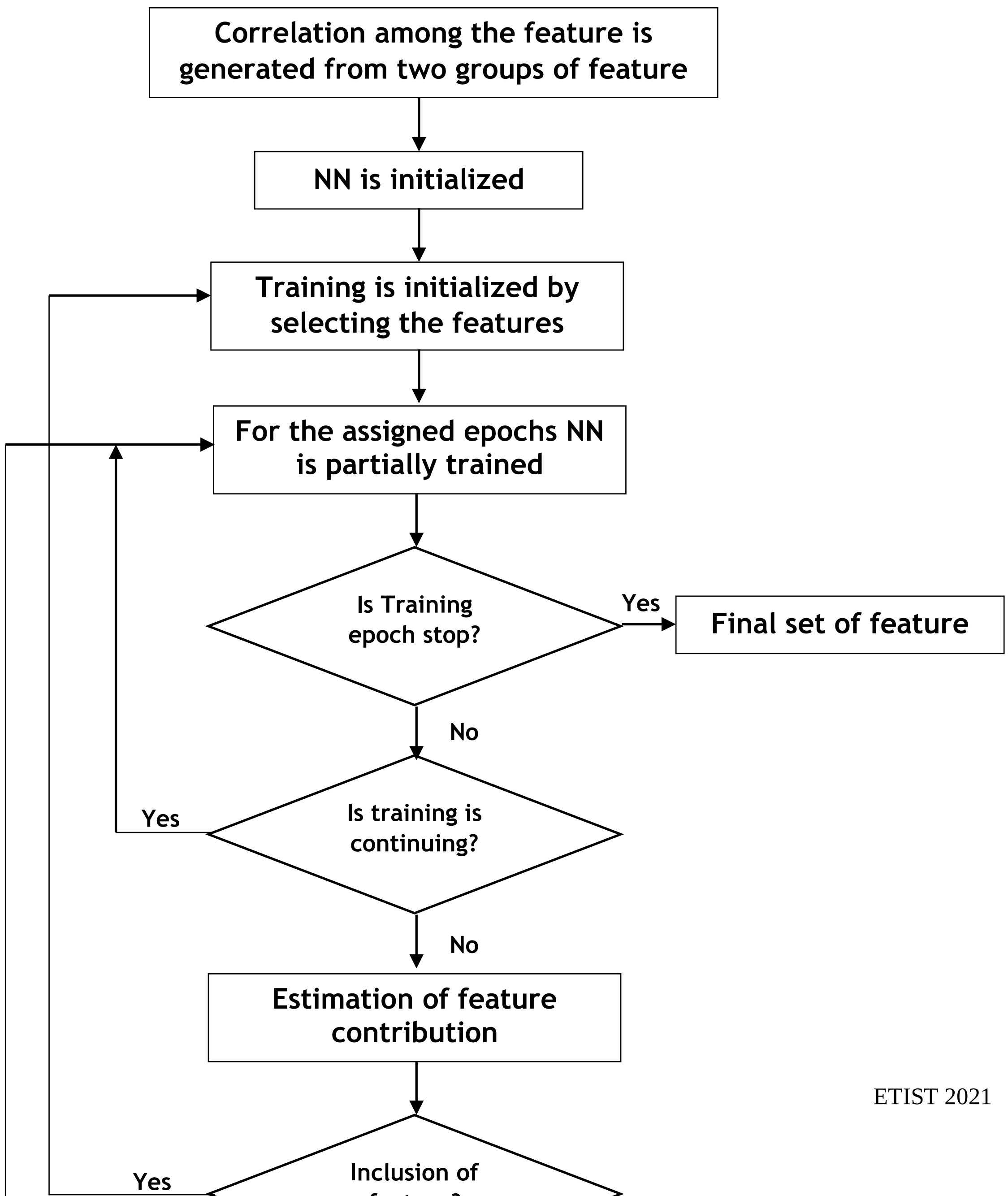

Figure 1. Flowchart for overall training process

NN training process termination

In the process of training hidden neurons and features are added into the proposed model one by one. The occurrence of error in the training process is minimized the training progression. However, the ability of generalization in the NN is enriched by the CFSNN process and the training is not an accurate choice for terminating the NN training process. The general scheme of neural network and training process is illustrated in Figure 2.

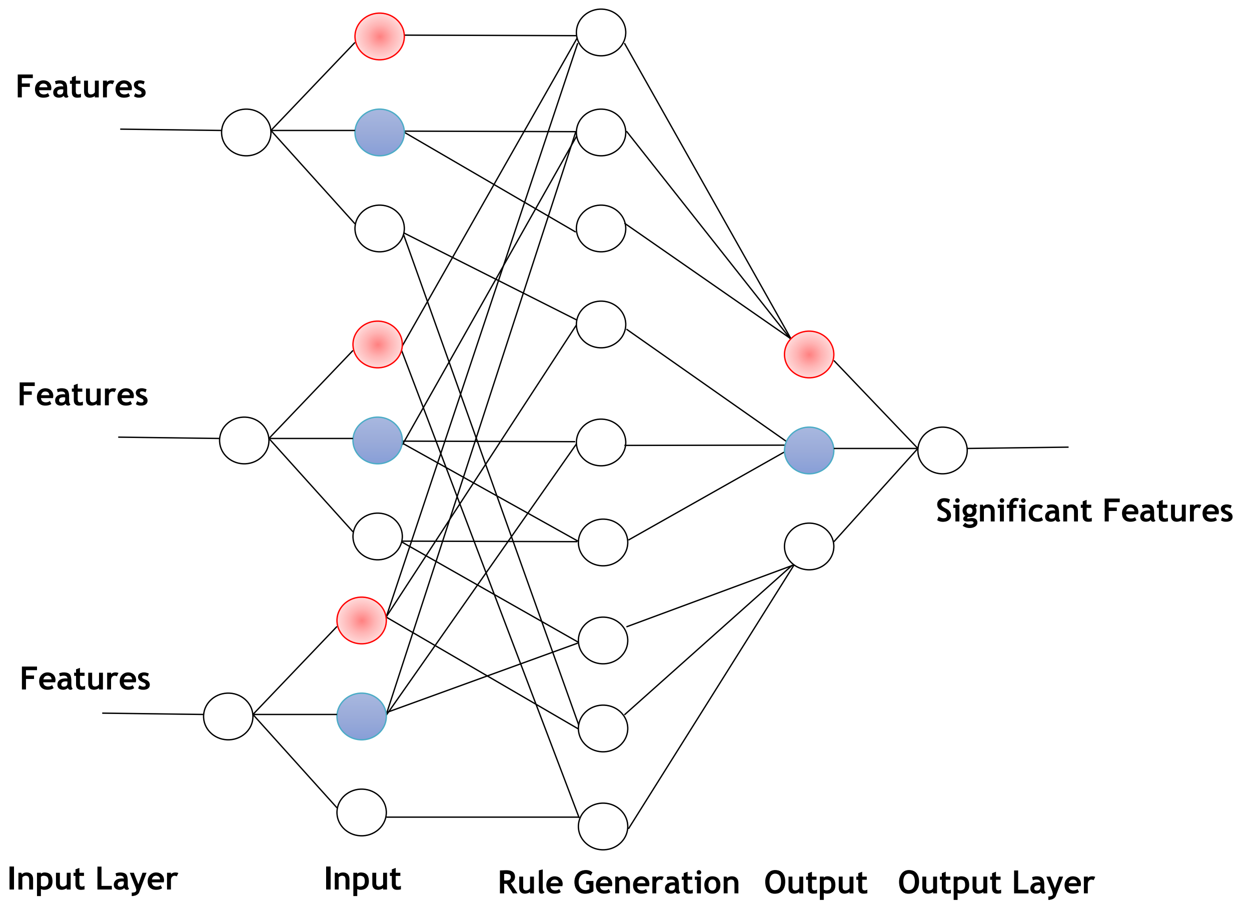

Figure 2. Overall structure of neural network

A separate validation dataset is incorporated for terminate training process. The unbiased estimation is attained by the validation error and it is not used in altering the weights of the NN. The training process is terminated by validation error, which gives best generalization and it is straightforward and simple. After each training epochs termed as strip that measures the error in the validation process. When the validation error is raised above the predefined value then the training is terminated. The validation error is estimated for every successive strips TM for every TMsuccessive time. The criteria of termination is denoted as,

$$T(\tau+i)-T(\tau)> \lambda,\ i=1,2,3\ldots.TM$$

where the values of TM and τ are positive integers and the value is assigned by the user. The training process is terminated once the condition is reached as mentioned in the above equation. If termination condition is not satisfied, though it may holds some hidden neuron and significant features that is added to the NN and then the accuracy is validated for some number of epochs that is equated as,

$$ACY = 100(\frac{R_{ua}}{R_u})$$

where the count of the patterns correctly classified R_{ua} and R_u denotes the whole set of data. If the value is not intensified or raised then the process will terminated automatically.

Selection and Addition of Feature

In the existing neural network, the features are added by the straightforward criteria and it determines the classification accuracy in the validation set. The best generalization is attained by the selection of salient features from the graph. The classification accuracy is described as,

$$ACY(T+\tau) > ACY(T), T = \tau, 2\tau, 3\tau, \ldots\ldots\ldots$$

The proposed approach test criteria for selecting the feature for every τ epoch and add one feature to the subset of feature if the criteria is fulfilled. The classification accuracy is enriched by the feature addition process, which is the significant strategy in the proposed learning approach. The CFSNN improves the network processing power by the feature addition process. The feature selection process is accomplished completely by the selection and addition process until the group reaches the empty set.

4. Result and Discussion

In this section, performance analysis of feature selection algorithm CFSNN and other existing feature selection algorithm. The performance of the algorithm is estimated by the performance metrics namely precision, classification accuracy, normalized mutual information and recall. The experiment is carried over six publicly accessible datasets which are highly used in the analysis of feature selection. The acquired features are passed to the classification phase and the results are compared in this section. The existing classification approaches namely Laplacian Score (LS), ECFS, MCFS, LSTM, CNN and ISGFS, which is compared with the proposed CFSNN.

Precision

The positive analytical value or precision denotes the closeness of the measurement and the relevance among the values identified. The random errors are stated as precision that is determined with the statistical variables. The values of precision and accuracy are similar terms. Typically, binary or decimal digits are used in representing the value of precision. It is measured on the basis of detecting features at True Positive (TP) and False Positive (FP) rates. The value of precision is directly relies on the percent of positive values in the total population. In the process of classification, the precision value is the count of the true positive values (i.e. the count of the item correctly labelled as positive classes). The algorithm with high precision signifies resultant value achieves more needed information than the irrelevant information. The precision value is more in the proposed methodology than the existing approach and achieved the better system performance. It is calculated as

$$Precision = \frac{TP}{TP + FP}$$

Table 1. Precision of feature selection algorithm on different datasets

Dataset	Algorithm						
	LS	ECFS	MCFS	LSTM	CNN	ISGFS	CFSNN
USPS	0.8788	0.8892	0.8952	0.8962	0.9102	0.9512	0.9912
Isolet	0.8879	0.8893	0.8943	0.8971	0.9071	0.9598	0.9978
BaseHock	0.8881	0.8894	0.8948	0.8990	0.9390	0.9678	0.9988
Prostate	0.8882	0.8897	0.8951	0.8991	0.9391	0.9687	0.9989
Yale	0.8884	0.8899	0.8952	0.9012	0.9412	0.9701	1.0991
Relathe	0.8894	0.8991	0.8992	0.9112	0.9512	0.9801	1.1991

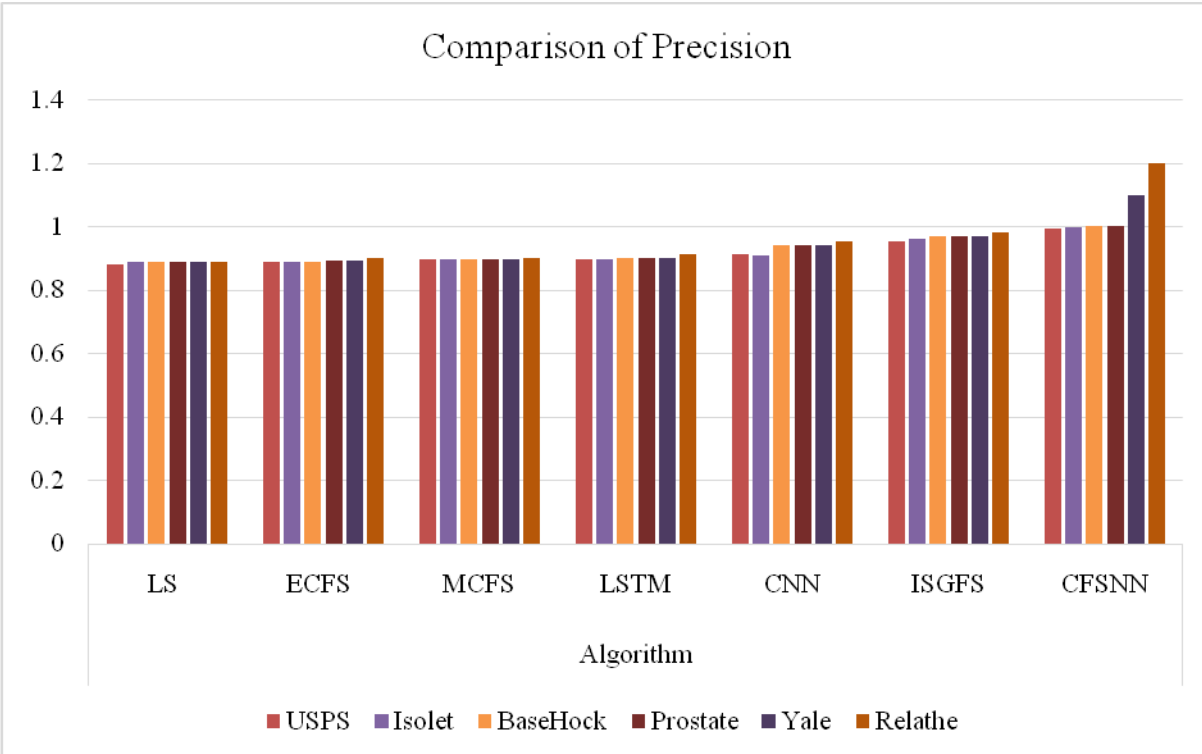

Figure 3. Comparison of Precision

In the Table 1 and Figure 3, the precision values for existing and proposed algorithm for various dataset is compared.From the observation of results it is identified that the proposed algorithm has highest precision.

Classification Accuracy

Accuracy specifies the nearness of the definite value from the classified instances. Accuracy is the depiction of statistical bias and the systematic errors. It is also closeness of an estimation to the true value and also it is the identification (both TP and TN values) amongst the count of the evaluated classes. Occurrence of least accuracy causes variation among the resultant and true resultant value. It is the ratio of accurate detection over the total amount of instances evaluated. It is computed as,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Table 2. Accuracy of feature selection algorithm on different datasets

Dataset	Algorithm						
	LS	ECFS	MCFS	LSTM	CNN	ISGFS	CFSNN
USPS	**0.81**	**0.83**	**0.84**	**0.85**	**0.91**	**0.88**	**1.1**
Isolet	**0.82**	**0.84**	**0.85**	**0.87**	**0.92**	**0.89**	**1.2**
BaseHock	**0.83**	**0.85**	**0.86**	**0.89**	**0.93**	**0.9**	**1.05**
Prostate	**0.84**	**0.86**	**0.87**	**0.91**	**0.94**	**1**	**0.99**

Yale	**0.86**	**0.87**	**0.89**	**0.92**	**0.96**	**0.87**	**1.21**
Relathe	**0.88**	**0.89**	**0.89**	**0.93**	**0.98**	**0.86**	**1.3**

In the Table 2 and Figure 4, the accuracy values for existing and proposed algorithm for various dataset is compared. From the observation of results it is identified that the proposed algorithm has highest accuracy.

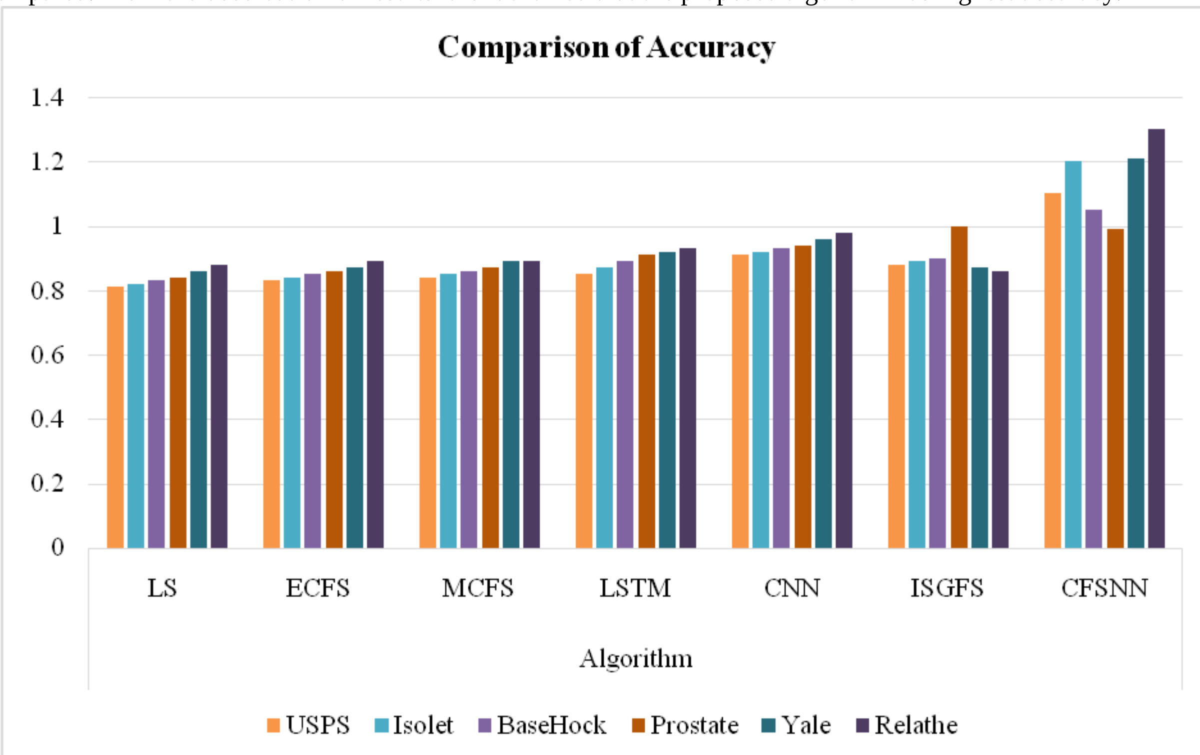

Figure 4. Comparison of Accuracy

Recall

The recall is the fraction of related instances amongst the actually reclaimed instances. The recall is an estimation measure of successful prediction rate and the count of related results are returned as recall.It is measured based on the detection of TP and False Negative (FN) rates. It is calculated as

$$Recall = \frac{TP + TN}{TP + FN}$$

Table 3. Recall of feature selection algorithm on different datasets

Dataset	**Algorithm**						
	LS	**ECFS**	**MCFS**	**LSTM**	**CNN**	**ISGFS**	**CFSNN**
USPS	**0.9799**	**0.9991**	**0.9961**	**0.9961**	**0.9101**	**0.9611**	**0.9911**
Isolet	**0.9979**	**0.9993**	**0.9943**	**0.9971**	**0.9071**	**0.9699**	**0.9979**
BaseHock	**0.9991**	**0.9994**	**0.9949**	**0.9990**	**0.9390**	**0.9679**	**0.9999**

Prostate	**0.9991**	**0.9997**	**0.9981**	**0.9991**	**0.9391**	**0.9697**	**0.9999**
Yale	**0.9994**	**0.9999**	**0.9961**	**0.9011**	**0.9411**	**0.9701**	**1.0991**
Relathe	**0.9994**	**0.9991**	**0.9991**	**0.9111**	**0.9611**	**0.9901**	**1.1991**

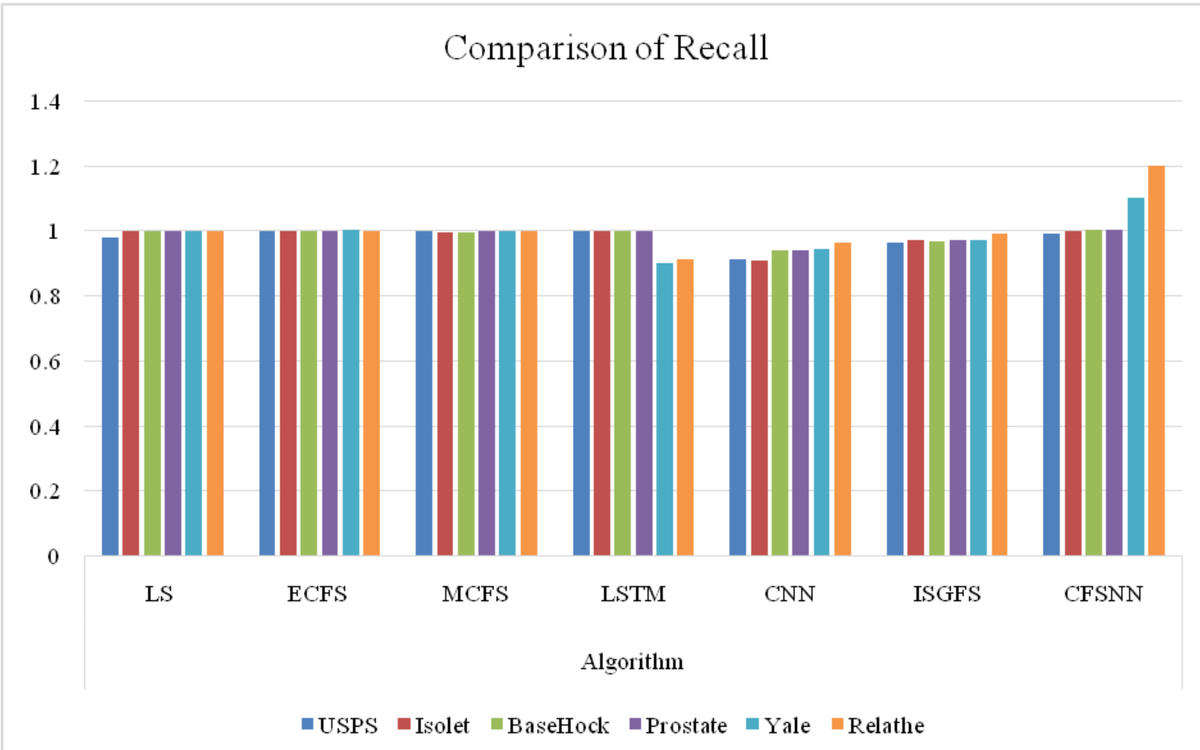

Figure 5. Comparison of Recall

In the Table 3 and Figure 5, the recall values for existing and proposed algorithm for various dataset is compared. From the observation of results it is identified that the proposed algorithm has highest recall.

F-Measure

F-measure or F-score is stated as an accuracy of test in the problem of classification. To compute F-measure, precision and recall value are taken, whereas precision is the count of the true positive values (positive values or correctly classified values) and the recall is the fraction of related instances amongst the actually reclaimed instances (sensitivity or classified instances). Otherwise, it is stated as a harmonic mean of the precision value and recall value. F-measure is chiefly used in the multiclass classification problems and it stabilizes both the precision and recall value.The algorithm with highest precision and recall value results in best f measure. The F-measure value results in a better retrieval of needed information and offers realistic measure of performance of the algorithm. It is computed as,

$$F - Measure = \frac{2. Precision. Recall}{Precision + Recall}$$

Table 4. F-Measure of feature selection algorithm on different datasets

Dataset	**Algorithm**						
	LS	**ECFS**	**MCFS**	**LSTM**	**CNN**	**ISGFS**	**CFSNN**
USPS	**0.7791**	**0.7954**	**0.7971**	**0.914**	**0.7791**	**0.9354**	**0.9563**
Isolet	**0.7793**	**0.7956**	**0.7973**	**0.916**	**0.7793**	**0.9356**	**0.9565**
BaseHock	**0.7795**	**0.7957**	**0.7975**	**0.919**	**0.7795**	**0.9357**	**0.9567**
Prostate	**0.7796**	**0.7957**	**0.7977**	**0.951**	**0.7796**	**0.936**	**0.9567**
Yale	**0.7797**	**0.7959**	**0.7979**	**0.954**	**0.7797**	**0.9363**	**0.9569**
Relathe	**0.7791**	**0.7954**	**0.7971**	**0.914**	**0.7791**	**0.9354**	**0.9563**

In the Table 4 and Figure 6, the F-measure values for existing and proposed algorithm for various dataset is compared. From the observation of results it is identified that the proposed algorithm has highest F-Measure.

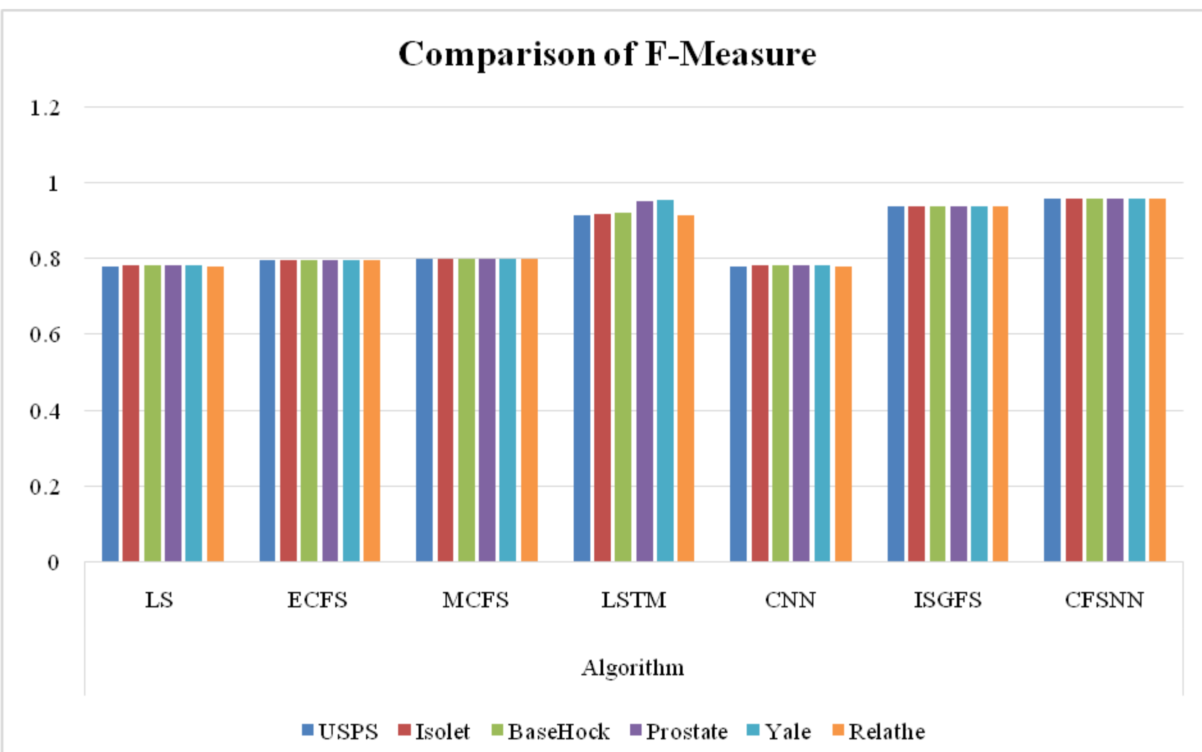

Figure 6. Comparison of F-Measure

Error Rate

In the digital transformation, the occurrence of error is due to the data transformation is due to various factors namely noise, distortion and interference. It is a ratio of performance rate.Error rate is the proportion of sequences that are incorrectly classified by the decision making model. The value of error rate is estimated by summing the FP and FN value that is divided by the sum of TP,TN, FP and FN values. It is measured as:

$$Error\ Rate = \frac{FP + FN}{TP + TN + FP + FN}$$

Table 5. Error Rate of feature selection algorithm on different datasets

Number of Training Epochs	**Dataset**	**Algorithm**						
		LS	**ECFS**	**MCFS**	**LSTM**	**CNN**	**ISGFS**	**CFSNN**
100	**USPS**	**4.3**	**4.1**	**4.0**	**3.9**	**4.3**	**4.1**	**3.5**
200		**3.8**	**3.7**	**3.9**	**3.8**	**3.8**	**3.8**	**3.1**
300		**3.7**	**3.3**	**3.8**	**3.6**	**3.6**	**3.5**	**2.6**
400		**2.8**	**2.7**	**3.5**	**3.3**	**3.5**	**3.1**	**2.5**
100	**Isolet**	**6.7**	**6.1**	**7.0**	**7.9**	**6.7**	**6.1**	**5.5**
200		**7.8**	**7.7**	**7.9**	**7.8**	**7.8**	**7.8**	**5.6**
300		**7.7**	**7.7**	**7.8**	**7.6**	**7.6**	**7.5**	**5.6**
400		**5.8**	**5.7**	**6.5**	**7.7**	**7.5**	**7.1**	**5.1**
100	**BaseHock**	**6.8**	**6.1**	**8.0**	**8.9**	**6.8**	**6.1**	**5.4**
200		**8.8**	**8.4**	**8.9**	**8.8**	**8.8**	**8.8**	**7.6**

300		**8.8**	**8.7**	**8.8**	**8.6**	**8.6**	**8.4**	**7.6**
400		**8.9**	**8.8**	**8.4**	**8.8**	**8.4**	**8.1**	**7.1**
100	**Prostate**	**4.3**	**4.1**	**4.0**	**3.9**	**4.3**	**4.1**	**3.4**
200		**3.8**	**3.8**	**3.9**	**3.8**	**3.8**	**3.8**	**3.1**
300		**3.8**	**3.3**	**3.8**	**3.6**	**3.6**	**3.4**	**3.2**
400		**3.8**	**3.8**	**3.4**	**3.3**	**3.4**	**3.1**	**2.9**
100	**Yale**	**2.3**	**2.1**	**3.0**	**3.9**	**2.3**	**2.1**	**1.8**
200		**3.3**	**3.4**	**3.9**	**3.3**	**3.3**	**3.3**	**1.2**
300		**3.3**	**3.1**	**3.3**	**3.2**	**3.2**	**3.4**	**1.2**
400		**3.9**	**3.3**	**3.4**	**3.3**	**3.4**	**3.1**	**1.1**
100	**Relathe**	**4.2**	**4.1**	**4.0**	**2.9**	**4.2**	**4.1**	**2.4**
200		**2.2**	**2.3**	**2.9**	**2.3**	**2.3**	**2.3**	**2.1**
300		**2.2**	**2.2**	**2.3**	**2.2**	**2.2**	**2.4**	**1.2**
400		**2.1**	**2.1**	**2.4**	**2.2**	**2.4**	**2.1**	**1.9**

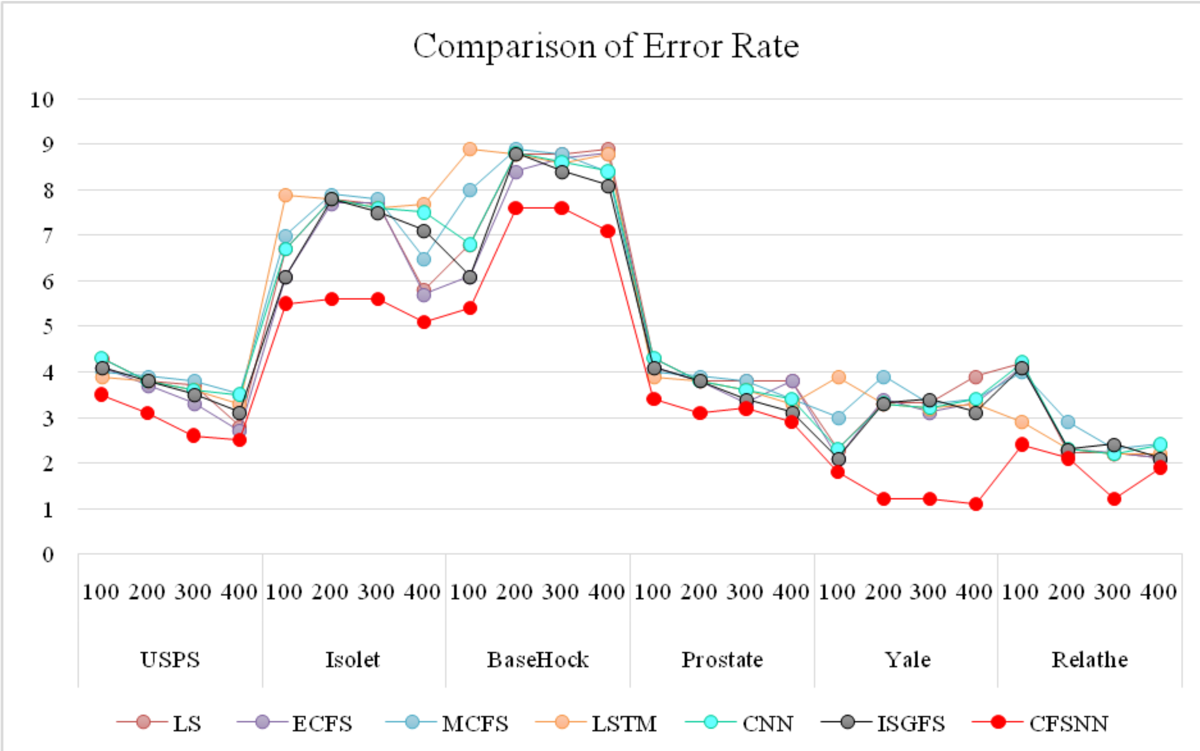

Figure 7. Comparison of Error Rate

In the Table 5 and Figure 7, the error rate for existing and proposed algorithm for various dataset is compared. From the observation of results it is identified that the proposed algorithm has minimum error rate.

5. Conclusion

The main intent of this study is to establish and evaluate a feature selection approach that is based on the learning process. The proposed approach spots the relationship among the feature via neural network and learning process. In CFSNN, a constructive approach is established that incorporated neural network through the process of learning features are elected and classified. To evaluate the performance of the algorithm, the CFSNN is compared with existing classification algorithms and different dataset. The performance is analysed using the performance metrics and from the observation of the investigation it is identified that the proposed CFSNN has better results. The error rate of the proposed approach is highly effective.In future the algorithm can be developed with the threshold assignment process and it can attain better result with user preference.

Reference

1. Kong, X., & Yu, P. S. (2010, July). Semi-supervised feature selection for graph classification. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 793-802).
2. Nie, F., Zhu, W., & Li, X. (2016, February). Unsupervised feature selection with structured graph optimization. In *Proceedings of the Thirtieth AAAI conference on artificial intelligence* (pp. 1302-1308).

3. Zheng, W., Zhu, X., Zhu, Y., Hu, R., & Lei, C. (2018). Dynamic graph learning for spectral feature selection. *Multimedia Tools and Applications*, *77*(22), 29739-29755.

4. Hu, R., Zhu, X., Cheng, D., He, W., Yan, Y., Song, J., & Zhang, S. (2017). Graph self-representation method for unsupervised feature selection. *Neurocomputing*, *220*, 130-137.

5. Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, *50*(6), 1-45.

6. Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. *Data classification: Algorithms and applications*, 37.

7. Kong, X., & Philip, S. Y. (2010, December). Multi-label feature selection for graph classification. In *2010 IEEE International Conference on Data Mining* (pp. 274-283). IEEE.

8. Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70-79.

9. Li, J., & Liu, H. (2017). Challenges of feature selection for big data analytics. *IEEE Intelligent Systems*, *32*(2), 9-15.

10. Doquire, G., &Verleysen, M. (2013). A graph Laplacian based approach to semi-supervised feature selection for regression problems. *Neurocomputing*, *121*, 5-13.

11. Thoma, M., Cheng, H., Gretton, A., Han, J., Kriegel, H. P., Smola, A., ...&Borgwardt, K. (2009, April). Near-optimal supervised feature selection among frequent subgraphs. In *Proceedings of the 2009 SIAM International Conference on Data Mining* (pp. 1076-1087). Society for Industrial and Applied Mathematics.

12. Kong, X., Yu, P. S., Wang, X., & Ragin, A. B. (2013, May). Discriminative feature selection for uncertain graph classification. In *Proceedings of the 2013 SIAM International Conference on Data Mining* (pp. 82-93). Society for Industrial and Applied Mathematics.

13. Dash, M., Liu, H., &Motoda, H. (2000, April). Consistency based feature selection. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 98-109). Springer, Berlin, Heidelberg.

14. Sheikhpour, R., Sarram, M. A., Gharaghani, S., &Chahooki, M. A. Z. (2017). A survey on semi-supervised feature selection methods. *Pattern Recognition*, *64*, 141-158.

15. Zhu, X., Li, X., Zhang, S., Ju, C., & Wu, X. (2016). Robust joint graph sparse coding for unsupervised spectral feature selection. *IEEE transactions on neural networks and learning systems*, *28*(6), 1263-1275.

16. Quanz, B., &Huan, J. (2009, April). Aligned graph classification with regularized logistic regression. In *Proceedings of the 2009 SIAM International Conference on Data Mining* (pp. 353-364). Society for Industrial and Applied Mathematics.

17. Roffo, G., &Melzi, S. (2017). Ranking to learn: Feature ranking and selection via eigenvector centrality. New Frontiers in Mining Complex Patterns. In *Fifth International workshop, nfMCP2016*.

18. Cai, D., Zhang, C., & He, X. (2010, July). Unsupervised feature selection for multi-cluster data. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 333-342).

19. Hochreiter, S., &Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

20. Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.

21. Henni, K., Mezghani, N., &Mitiche, A. (2020). Cluster Density Properties Define a Graph for Effective Pattern Feature Selection. *IEEE Access*, *8*, 62841-62854.