A COMPREHENSIVE REVIEW ON FEATURE SELECTION AND FEATURE EXTRACTION ON ALZHEIMER'S

A. Mujiba Parvin, S. Manju Priya*

Abstract

In recent years, researchers have proposed and developed many methods and techniques to reduce the high dimensions of data and to attain the required accuracy. To ameliorate the accuracy of learning features as well as to decrease the training time dimensionality reduction is used as a pre-processing step, which can eliminate irrelevant data, noise, and redundant features. Dimensionality reduction (DR) has been performed based on two main methods, which are feature selection (FS) and feature extraction (FE). FS is considered as an important method because data is generated continuously at an ever-increasing rate; some serious dimensionality problems can be reduced with this method, such as decreasing redundancy effectively, eliminating irrelevant `data, and ameliorating result comprehensibility.

Keywords: Dimension Reduction, Dimension Reduction Techniques, Feature Selection, Feature Extraction

I. INTRODUCTION

Now a days, Knowledge is detected by processing and analyzing a large amount of the previously collected data[1].Data generated in a huge volume in different fields, and it is on continuous growth in size, complexity, and dimensionality [2, 3]. A dataset with high dimensionality features its numerous features, but few samples have a direct relation with data mining and machine learning tasks [4, 5]. Therefore, these issues of data become a big challenge for extracting potentially useful and ultimately understandable patterns or information in almost every data mining task. Also, working in high dimensional data increases the difficulty of knowledge discovery and pattern classification because there are a lot of redundant and irrelevant features. Reducing high dimensional datasets to a low dimensional dataset by filter or remove redundant and noise information is a method to solve this problem, and this is known as dimensionality reduction [6].

Dimensionality reduction is a process for decreasing features dimensionality, but the data is still present. In the reduced or low dimension dataset, the crucial features remain even if some particular pattern vanishes [7, 8]. Also, it utilizes to reduce the size of input data and then preserve much variance of essential features compared to the dataset with the larger size. In real-world data, it will become easy to detect and use for data mining applications and gain high accuracy performance [1, 9]. Moreover, the role of dimensionality reduction is to enhance the accuracy and efficiency of the data mining computation, and it is considered as a vital preprocessing step. Furthermore, it provides several advantage such as eliminating irrelevant, redundant patterns in the dataset; as a result, to reduce the time and amount of memory required for processing such data [1, 10]. By reducing the dataset, the quality of data will improve, the algorithm will work efficiently, achieve better accuracy, and pattern design and examination will be clearer for researchers [11]. Additionally, reducing the cost of computing, improving dimensions visualization, and enhancing the results [12, 13].

II. DIMENSIONALITY REDUCTION TECHNIQUES

Dimensionality reduction is the operation of transforming the high dimensional representation of data in low dimensional representations. With the massive growth in

Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
*Corresponding Author

high dimensional data, the use of various dimensionality reduction techniques has become popular in many areas of use. Moreover, several modern approaches are continually emerging. Dimensionality reduction techniques transform the original dataset having high dimensionality and turn it into a new dataset representing low dimensionality while maintaining as much as possible the original meanings of the data. The low dimensional representation of the original data contributes to solving the dimensionality curse problem. The low dimensional data can be easily analyzed, processed, and visualized [14]. Several benefits can be obtained due to applying the dimensionality reduction techniques applied to a dataset. (i) As the number of dimensions comes down, data storage space can be reduced. (ii) It takes less computation time only. (iii) Redundant, irrelevant, and noisy data can be removed. (iv) Data quality can be improved. (v) Some algorithms do not perform well on a greater number of dimensions taken. So, reducing these dimensions helps an algorithm to work efficiently and improves accuracy. (vi) It is challenging to visualize data in higher dimensions. So, reducing the dimension may allow us to design and examine patterns more clearly. (vii) It simplifies the process of classification and also improves efficiency [15, 16]. Generally, the dimensionality reduction techniques can be classified into two main groups, or in other words, the dimensionality reduction is achieved through two different techniques: feature selection and feature extraction. In feature selection, information can be lost since some features should be excluded when the process of feature subset choice by doing this information can be reduced. However, in feature extraction, the dimension can be decreased without losing much initial feature dataset [2, 10, 14, 17]. Table I provides a descriptive summary of the methods of dimension reduction.

A. Feature Selection

Feature selection is utilized to reduce the dimensionality impact on the dataset through finding the subset of feature which efficiently define the data [18, 19]. It selects the important and relevant features to the mining task from the input data and removes redundant and irrelevant features [20, 21]. It is useful for detecting a good subset of features that is appropriate for the given problem [2, 22]. The main purpose of feature selection is to construct a subset of features as small as possible but represents the whole input data vital features [11, 23]. Feature selection provides numerous advantages: reduce the size of data, decrease needed storage, prediction accuracy improvement, over fitting evading, and reduce executing and training time from easily understanding variables. Feature selection algorithm phase is divided into two-phase such as (i) Subset Generation: (ii) Subset Evaluation: In subset Generation, we need to generate subset from the input dataset and to use Subset Evaluations we have to check whether the generated subset is optimal or not [24, 25]. “Fig.1” shows the overall method of the feature selection process.

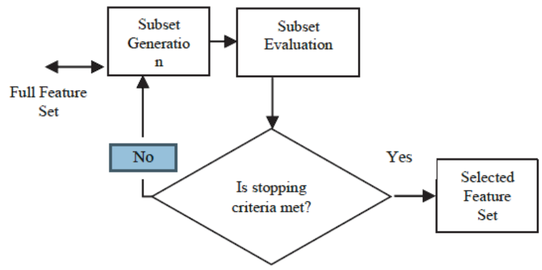

Fig. 1 Process of feature selection

B. Feature Selection Problems

Various issues can benefit from the feature selection techniques application. High dimension, low sample size data are becoming more popular in different fields. Many of the features of these problems do not facilitate an adequate classification. More so, the imbalance problem happens when one of the two classes has more samples than other classes. Many algorithms neglect the minority sample when concentrating on a major sample classification. However, the minority samples are crucial but seldom occurred. Moreover, in machine learning, the shift of the dataset is a popular problem that happens when the joint distribution of inputs and outputs varies between training and test stages. A special case of dataset shift, which happens when only the input distribution changes is called Covariate shift. Furthermore, the reduction of the dimensionality and consequently feature selection is one of the most common techniques of noisy data elimination. Eventually, misclassification costs and test costs are the two most significant kinds of cost in cost-sensitive learning [26,27,28].

C. Feature Selection Methods

Feature selection aims to select a feature subset from the original set of features based on a/the feature's relevance and redundancy. Originally evaluation methods in feature selection are divided into four kinds: filter, wrapper, embedded [10, 14, 18], and hybrid [20, 29]. Recently, another type of evaluation method is developed, i.e., ensemble feature selection [30, 31]. "Fig. 2" depicts the hierarchy of feature selection techniques.

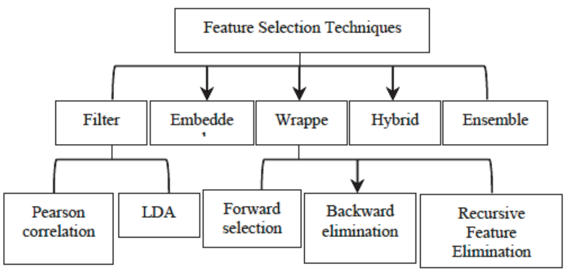

Fig. 2. Hierarchy of Feature Selection Techniques.

Filter is considered the earliest method and also known as an open-loop method. It checks the features relying on the intrinsic characteristics prior to the learning tasks. It mainly measures the feature characteristics depending on four different kinds of measurement criteria, i.e. information, dependency, consistency, and distance [32]. In the filter method, the feature selection process is performed independently of the data mining algorithm. It uses statistical standards for evaluating the ranking of the subset [17, 24]. Moreover, this technique is to perform good performance and high-efficiency computing, easily scalable in high dimensional datasets, and outperformed the wrapper technique. The primary downside of this method is that it neglects the integration between the selected subset and the performance of the induction algorithm [10, 22, 26].

Wrapper or it can be called as close-loop method, wraps the feature selection around the learning algorithm and uses the accuracy of the performance or the error rate of the classification process as a criterion of feature evaluation. By decreasing the estimation error of a specific classifier, it chooses the most discriminative subset of features. The wrapper method performs features election based on the performance of the learning algorithm; it selects the most optimal feature for the prediction algorithm. Hence it achieves better performance and high accuracy compared to the filter algorithm [22,27,33]. The main disadvantage of this approach is computing complexity and more exposure to over fitting in comparison to the filter approach. Most wrapper methods are multivariate; thus, they need extensive computation times to achieve the convergences and can be intractable for large datasets [33, 34]. "Fig. 3" shows the involved steps in the wrapper method.

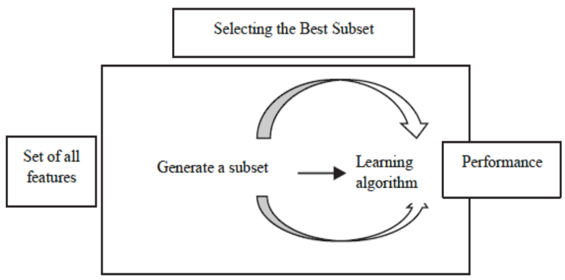

Fig. 3 Wrapper method for feature selection

Hybrid and ensemble methods the recent developments in features election can be represented in the hybrid method[35]. Thus, it can be developed either by integrating two various methods (e.g. wrapper and filter), two methods with the same criteria, or two feature selection approaches. In the hybrid method, the advantages of both methods can be inherited by combining their complementary strengths [36]. The combination of filter and wrapper method is the most common hybrid method [37]. However, ensemble method is a method that aims at building a group of feature subsets and then producing an aggregated result out of the group [38]. This method is depending on various sub sampling techniques where a particular feature selection method is implemented on a variety of subsamples, and the obtained features are merged to create a more stable subset.

D. Feature Extraction Methods

In previous literature, the dim ensionality reduction uses the feature selection methods to select the relevant features have been presented. The remaining aim of this research paper is to review the latest literature related feature extraction and dimensionality reduction techniques.

M. K. Elhadad et al. [39] worked on the emotion recognition task. They used PCA and t-statistical to reduce the dimensionality of extracted features from emotional signals of EEG. The proposed method was applied to the dataset called SJTU emotion EEG. The emotional state with extracted features has been classified by four classifiers: SVM, ANN, LDA, and KNN[40].

Moghaddam et al. [41] proposed a method known as spectral segmentation and integration (SSI) as supervised feature extraction for hyper spectral images. The developed method divided pixels' spectral signature curve to channels. Then a mean weighted operator was used for integration of each channel band in order to extract new features in a very minimal number compared to the original bands. Moreover, the PSO algorithm was used to merge spectral signature curve pixel segments so as to reduce the dimensionality of the image and to increase the class accuracy. In the proposed technique, the SVM was used as a classifier, and two datasets were used. The experimental results confirmed the SSI method outperformed other feature extraction methods such as PCA, SRS, NWFE, DAFE, PCA, SELD, BCC, and CBFE.

A. M. Abdulazeez[46] worked on malignant masses in mammograms based on the feature extraction. The researcher presented Gray Level Co-occurrence Matrix(GLCM)texture feature extraction by three hybrid methods that were used in the proposed method. The three hybrid methods called Wavelet CT1, Wavelet CT2, and ST-GLCM. The interesting point of the image was divided into sub-image then contrast stretching stage was used prior to feature extraction. Then the sub-image has been applied for the methods of feature extraction. Next, the GLCM extracted the seven-feature texture and have been merged with seven statistical features. Moreover, two data sets images were used in this research and SVM classifier. The proposed methods outperformed the multi-resolution feature extractions methods in terms of the number of the extracted feature. Also, in Area under the Curve (AUC) measure, the researcher methods were superior to other feature extraction methods.

Churmonge and Jena [51] proposed a method to address the dimensionality issue based on the clustering combined with correlation filter subset selection. The relevant features were found by the K-means clustering algorithm and redundant features found from clusters and removed by the correlation measure. The presented method used on 8 text and 4 microarray datasets and the Naive Bayes (NB) classifier depended on the classification. Furthermore, the authors compared their method performance with the ReliefF and information gain (IG) feature selection methods relating to the accuracy and computational time. The accuracy of the proposed method outperformed both methods in all datasets except two datasets, and in computational, the proposed was faster than other methods in all datasets.

Tan et al. [52] presented a feature selection method based on the evolutionary algorithm (EA) to reduce the dimensionality of motor imagery brain-computer interface from electroencephalogram (EEG) signals. The subset of important features was generated from each iteration of the EA, while the redundant and insignificant features were eliminated. The experiments were performed in two different datasets: EEG dataset and several machines learning datasets. Also, three classifiers depended: support vector machine (SVM), K-Nearest Neighbor (KNN), and discriminate analysis (DA). Also, the performance of the proposed EA – feature selection method was compared with PCA and independent component analysis (ICA), neighborhood component analysis (NCA), and variable-length particle swarm optimization (VLPSO). The results showed that the introduced methods outperformed all the above methods and could achieve high accuracy even with a small subset of the features.

Hafiz et al. [53] investigated the feature selection issues in the power quality events and proposed a two-dimensional PSO feature selection method. They depended on the two dimensional in order to efficiently guide the search space of the particle swarm. The noise measurement against the reduced subset was studied by the Gaussian. The used induction algorithms in this study were KNN and Naïve Bayes. Moreover, the proposed method performance was compared with the Genetic Algorithm (GA), Ant Colony Optimization (ACO), Binary PSO (BPSO), Catfish BPSO, and Chaotic BPSO (CHBPSO). The results have shown that the presented method could find an important and robust feature subset and achieve better accuracy than the above-mentioned methods.

Han et al. [54] worked on the limitation of the local linear embedding (LLE) method to propose an unsupervised feature selection mechanism. They depended on the low dimensional space learning and graph matrix learning. The experiments performed in 15 datasets. Also, the SVM and decision tree (DT) used as classifiers and. In addition, the presented method compared with eight unsupervised feature selection methods. The results demonstrated that the proposed method accomplished better accuracy regarding the SVM and DT classifier except in two datasets and better stability in both classifiers compared to the 8 feature selection methods.

Niu et al. [55] presented a method to deal with multivariate financial time series nonlinearity inherent to improve the accuracy of forecasting and make the financial decision better. The proposed method involved a feature selection part, deep learning framework, and error correlation part. In the feature selection part, the RReliefF algorithm (which is the enhanced version of the ReliefF) cooperated wrapper-based method to remove the redundant feature. Also, the deep learning part has consisted of long-short term memory (LSTM), gated recurrent unit (GRU), and the optimizer based on adaptive moment estimation (Adam). The deep leaning part was trained based on the subset generated by the first part. Furthermore, the error correlation used to enhance the accuracy of the method. The method performance validated on 16 benchmarks and three datasets, and the results have shown its superiority.

Jain and Singh [56] proposed a hybrid feature selection method that consisted of ReliefF and PCA algorithms. First, the weight for each feature was calculated in the used datasets, and a set of satisfying features was generated by the first algorithm. The second algorithm was applied in the generated set. In the proposed method, two types of datasets were considered (text and microarray), and the experiments performed in ten datasets. The performance of the method evaluated in terms of a number of the selected features and computation time. The results indicated that the presented method could achieve better performance in low and high dimensional datasets and reduced half of the irrelevant and redundant features.

Method	Main concept	Pros	Cons
Feature extraction	Summarize the dataset by creating linear combinations of the features	Preserves the original, relative distance between covers latent structure, objects	Not sufficient enough in the existing of a huge number of irrelevant features
Feature selection	A sub list of relevant features can be selected depending on defined criteria	Strong against irrelevant features	Latent structure does not cover

Table I.The Summary of Dimension Reduction Techniques

III. CONCLUSION

The high dimensionality of data has a direct impact on the learning algorithm, computational time, computer resources (memory), and model accuracy. Therefore, reducing dimensionality and tackling its curse became an exciting topic in search and development areas to provide the most reliable, flexible, and high accurate computerized tools and applications. Hence, several methods and techniques accomplished in the last two decades based on the feature selection and feature extraction.

This paper reviews the most recent studies in several fields such as medical disease analysis, ethnicity identification, emotion recognition, genes classification, text classification, image Steganalysis, data visualization, Hyper spectral images classification, moreover, the details used techniques/algorithms, datasets, classifiers approaches were used by the authors and attained results relating to the accuracy and computational time are summarized for each of the feature selection and feature extraction methods. It is observed that the trend of the researchers for reducing the dimensionality based on the feature selection methods is to use the optimization algorithms, and about half of the reviewed researches were relaying on the different techniques of optimization. Also, the most used classifiers are the SVM and KNN, and the best- achieved accuracy was the SVM algorithm. On the other hand, for feature extraction methods, CNN and DNN techniques take a great role and have been used in 7 methods of the studied research. While the PCA is still a widely used algorithm in the feature extraction works, it has been used in 8 methods. Additionally, the optimized PCA could achieve better performance in terms of accuracy, computational time, and the number of reduced features.

REFERENCES

[1] N. Sharma and K. Saroha, "Study of dimension reduction methodologies in data mining," in International Conference on Computing, Communication & Automation, 2015, pp. 133-137:IEEE.

[2] S. Ayesha, M. K. Hanif, and R. Talib, "Overview and comparative study of dimensionality reduction techniques for high dimensional data," Information Fusion, vol. 59, pp. 44-58,2020.

[3] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. Zeebaree, "CombinationofK-meansclusteringwithGeneticAlgorithm:Areview," International Journal of Applied Engineering Research, vol. 12, no. 24, pp. 14238-14245,2017.

[4] Z. Cheng and Z. Lu, "A novel efficient feature dimensionality reduction method and its applicationinengineering,"Complexity,vo.2018, 2018.

[5]D.A.Zebari,H.Haron,D.Q.Zeebaree,andA.M.Zain,"ASimultaneous Approach for Compression and Encryption Techniques Using Deoxyribonucleic Acid," in 2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), 2019, pp. 1-6:IEEE.

[6] M. Li, H. Wang, L. Yang, Y. Liang, Z. Shang, and H. Wan, "Fast hybrid dimensionality reduction method for classification based on feature selection and grouped feature extraction," Expert Systems with Applications, vol. 150, p. 113277, 2020.

[7] A. P. Pandian, R. Palanisamy, and K. Ntalianis, Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI-2019). Springer Nature,2020.

[8] M.A.Mohammed,B.Al-Khateeb, A.N.Rashid, D.A.Ibrahim, M.K. A. Ghani, and S. A. Mostafa, "Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images,"Computers&ElectricalEngineering,vol.70,pp.871-882,2018.

[9] O. Saini and S. Sharma, "A review on dimension reduction techniques in data mining," Computer engineering and intelligent systems, vol. 9, pp. 7-14,2018.

[10] N. Abd-Alsabour, "On the Role of Dimensionality Reduction," JCP, vol. 13, no. 5, pp. 571-579,2018.

[11]S.Velliangiriand S.Alagumuthukrishnan,"A Review of Dimensionality Reduction Techniques for Efficient Computation," Procedia Computer Science, vol. 165, pp. 104-111, 2019.

[12]W. Wang, W.-g. Shen, Y.-x. Sun, B. Chen, and R. Zhu, "Dimensionality reduction via adjusting data distribution density," in 2018 5th International Conference on Systems and Informatics (ICSAI), 2018, pp. 1052-1055:IEEE.

[13]J.Stuckman,J.Walden,andR.Scandariato,"The effect of dimensionality reduction on software vulnerability prediction models," IEEE Transactions on Reliability, vol. 66, no. 1, pp. 17-37, 2016.

[14]M. Verleysen and D. François, "The curse of dimensionality in data mining and time series prediction," in International Work-Conference on Artificial Neural Networks, 2005, pp. 758-770:Springer.

[15]L. Liu and M. T. Özsu, Encyclopedia of database systems. Springer New York, NY, USA:,2009.

[16]A. Juvonen, T. Sipola, and T. Hämäläinen, "Online anomaly detection using dimensionality reduction techniques for HTTP log analysis," Computer Networks, vol. 91, pp. 46-56,2015.

[17] X. Huang, L. Wu, and Y. Ye, "A Review on Dimensionality Reduction Techniques," International Journal of Pattern Recognition and Artificial Intelligence, vol. 33, no. 10, p. 1950017, 2019.

[18]D. L. Padmaja and B. Vishnuvardhan, "Comparative study of feature subset selection methods for dimensionality reductionon scientific data," in 2016 IEEE 6th International Conference on Advanced Computing (IACC), 2016, pp. 31-34:IEEE.

[19]M. B. Abdulrazzaq and J. N. Saeed, "A Comparison of Three Classification Algorithms for Handwritten Digit Recognition," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 58-63:IEEE.

[20] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," Expert Systems with Applications, vol. 42, no. 5, pp. 2670-2679,2015.

[21] A. S. Eesa, A. M. A. Brifcani, and Z. Orman, "Cuttlefish algorithm-a novel bio-inspired optimization algorithm," International Journal of Scientific & Engineering Research, vol. 4, no. 9, pp. 1978-1986, 2013.

[22] P. Jindal and D. Kumar, "A review on dimensionality reduction techniques," International journal of computer applications, vol. 173, no. 2, pp. 42-46,2017.

[23] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A new feature selection model based on ID3 and bees algorithm for intrusion detection system," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 23, no. 2, pp. 615-622,2015.

[24] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithm: A review," Journal of King Saud University-Computer and Information Sciences, 2019.

[25] S. Visalakshi and V. Radha, "A literature review of feature selection techniques and applications: Review of feature selection in data mining," in 2014 IEEE International Conference on Computational Intelligence and Computing Research, 2014, pp. 1-6: IEEE.

[26] C. M. Teng, "Combining noise correction with feature selection," in International Conference on Data Warehousing and Knowledge Discovery, 2003, pp. 340-349: Springer.

[27] H.Zhao,F.Min,andW.Zhu,"Cost-sensitive features election of numeric data with measurement errors," Journal of Applied Mathematics, vol. 2013,2013.

[28]J. N. Saeed, "A SURVEY OF ULTRASONOGRAPHY BREAST CANCER IMAGE SEGMENTATION TECHNIQUES," Academic Journal of Nawroz University, vol. 9, no. 1, pp. 1-14,2020.

[29] Y. Leung and Y. Hung, "A multiple-filter-multiple-wrapper approach to gene selection and microarray data classification," IEEE/ACM TransactionsonComputationalBiologyandBioinformatics,vol.7,no.1, pp. 108-117,2008.

[30]C.Lazaretal.,"A survey on filter techniques for feature selection in gene expression microarray analysis," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 9, no. 4, pp. 1106-1119, 2012.

[31]M.R.MahmoodandA.M.Abdulazeez,"AComparative Study of a New Hand Recognition Model Based on Line of Features and Other Techniques," in International Conference of Reliable Information and Communication Technology, 2017, pp. 420-432:Springer.

[32]M.DashandH.Liu,"Featureselectionforclassification, "Itelligent data analysis, vol. 1, no. 3, pp. 131-156,1997.

[33]D. Jain and V. Singh, "Feature selection and classification systems for chronic disease prediction: Areview, "EgyptianInformaticsJournal,vol. 19, no. 3, pp. 179-189,2018.

[34]D.Q.Zeebaree, H.Haron, A.M.Abdulazeez, and D.A.Zebri, "Machine learning and Region Growing for Breast Cancer Segmentation," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 88-93:IEEE.

[35]D.Zebari, H.Haron, and S.Zeebaree, "Security Issues in DNA Based on Data Hiding: A Review," International Journal of Applied Engineering Research,vol. 12,no. 24, ISSN, pp. 0973-4562,2017.

[36]M. M. Kabir, M. M. Islam, and K. Murase, "A new wrapper feature selection approach using neural network," Neurocomputing, vol. 73, no. 16-18, pp. 3273-3283,2010.

[37]Y. Peng, Z. Wu, and J. Jiang, "A novel feature selection approach for biomedical data classification," Journal of Biomedical Informatics, vol. 43, no. 1, pp. 15-23,2010.

[38]Q. Shen, R. Diao, and P. Su, "Feature Selection Ensemble," Turing-100, vol. 10, pp. 289-306,2012.

[39]M. K. Elhadad, K. M. Badran, and G. I. Salama, "A novel approach for ontology-based dimensionality reduction for web text document classification," International Journal of Software Innovation (IJSI), vol.5, no. 4, pp. 44-58, 2017.

[40]D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform," in 2019 International Conference on AdvancedScienceandEngineering(ICOASE),2019,pp.100-105:IEEE.

[41]S. H. A. Moghaddam, M. Mokhtarzade, and B. A. Beirami, "A feature extraction method based on spectral segmentation and integration of hyperspectral images," International Journal of Applied Earth Observation and Geoinformation, vol. 89, p. 102097,2020.

[42]D. M. Sulaiman, A. M. Abdulazeez, H. Haron, and S. S. Sadiq, "Unsupervised Learning Approach-Based New Optimization K-Means Clustering for Finger Vein Image Localization," in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 82-87:IEEE.

[43]R.Aziz,C.Verma,andN.Srivastava,"Dimension reduction methods for micro array data: areview, "AIMS. Bioengineering, vol.4, no.1, pp. 179- 197, 2017.

[44] A. S. Eesa, A. M. Abdulazeez, and Z. Orman, "A DIDS Based on The Combination of Cuttle fish Algorithm and Decision Tree, "Science Journal of University of Zakho, vol. 5, no. 4, pp. 313-318,2017.

[45] Z. M. Hira and D. F. Gillies, "A review of feature selection and feature extraction methods applied on microarray data," Advances in bioinformatics, vol.170. 2015,2015.

[46] A. M. Abdulazeez and A. S. Issa, "Intrusion detection system based on neural networks using bipolar input with bipolar sigmoid activation function," AL-Rafidain Journal of Computer Sciences and Mathematics, vol. 8, no. 2, pp. 79-86,2011.

[47]D.A.Zebari, H.Haron, S.R.Zeebaree, and D.Q.Zeebaree, "Multi-Level of DNA Encryption Technique Basedon DNA Arithmetic and Biological Operations," in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 312-317:IEEE.

[48]O. M. S. Hassan, A. M. Abdulazeez, and V. M. TİRYAKİ, "Gait-based human gender classification using lifting 5/3 wavelet and principal component analysis," in 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 173-178:IEEE.

[49] F. P. Shah and V. Patel, "A review on feature selection and feature extraction for text classification," in 2016 International Conference onWireless Communications, Signal Processing and Networking (WiSPNET), 2016, pp. 2264-2268: IEEE.

[50] H. Sadeeq, A. Abdulazeez, N. Kako, and A. Abrahim, "A Novel Hybrid Bird Mating Optimizer with Differential Evolution for Engineering Design

Optimization Problems," in International Conference of Reliable Information and Communication Technology, 2017, pp. 522-534: Springer.

[51] S. Chormunge and S. Jena, "Correlation based feature selection with clustering for high dimensional data," Journal of Electrical Systems and Information Technology, vol. 5, no. 3, pp. 542-549, 2018.

[52] P. Tan, X. Wang, and Y. Wang, "Dimensionality reduction in evolutionary algorithms-based feature selection for motor imagery brain-computer interface," Swarm and Evolutionary Computation, vol. 52, p. 100597, 2020.

[53] F. Hafiz, A. Swain, C. Naik, and N. Patel, "Efficient feature selection of power quality events using two dimensional (2D) particle swarms," Applied Soft Computing, vol. 81, p. 105498, 2019.

[54] X. Han, P. Liu, L. Wang, and D. Li, "Unsupervised feature selection via graph matrix learning and the low-dimensional space learning for classification,"Engineering Applications of Artificial Intelligence, vol. 87, p. 103283, 2020.

[55] T. Niu, J. Wang, H. Lu, W. Yang, and P. Du, "Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting," Expert Systems with Applications, vol. 148, p. 113237, 2020.

[56] D. Jain and V. Singh, "An efficient hybrid feature selection model for dimensionality reduction," Procedia Computer Science, vol. 132, pp. 333- 341, 2018.