

# Improved Medical Diagnosis using Wrapper and Filter Techniques of Feature Selection

<sup>1</sup>Sonu Rani, <sup>2</sup>Dharminder kumar, <sup>3</sup>Sunita Beniwal

<sup>1</sup>M. Tech Scholar, Computer Science &Engineering, Guru Jambheshwar University of Science &Technology, Hisar-125001, India

<sup>2</sup>Professor, Computer Science &Engineering, Guru Jambheshwar University of Science & Technology, Hisar-125001, India

<sup>3</sup>Assistant Professor, Computer Science &Engineering, Guru Jambheshwar University of Science & Technology, Hisar-125001, India

**Abstract:-** Data mining field deals with the discovery of knowledge from enormous amount of data. To solve any problem there should be appropriate knowledge about the problem and technique that we are going to use to solve that problem. But there are many areas where problem identification itself takes a lot of time; medical area is one of them in which diagnosis of diseases takes a lot of time. Till then problem (diseases) flourishes to the extent that it cannot be controlled. So there should be some technique that could help in proper and early diagnosis of diseases. Data mining techniques helps here a lot to improve medical diagnosis. The most prevalent technique for this is feature selection. Although there are many feature selection techniques. In our research work we have used feature selection technique on medical data set where each attribute represent a test that is performed for the diagnosis of diseases. For filtering of attributes we have used Relief f attribute Evaluator to check the worthiness of an attribute, to compare the performance we have used multilayer perceptron classifier where comparison has been made on the basis of accuracy and efficacy of classifier.

**Keywords:-** Relief f Attribute evaluator, Multilayer perceptron classifier.

## I. INTRODUCTION

In the era of information industry Data Mining is a novel and promising field. Its superb techniques help in mining of golden nuggets of information from vast amounts of raw data. Thus this field was evolved naturally from the database system technology. And the need arise from the fact that the data is increasing day by day so the situation is "Data Rich Information Poor". This raw data is of no use until and unless it is not converted into information. Researchers are incessantly evaluating tools and technology to mine information from data archives and to turn data into information. To analyse large amount of data in datasets is a big problem. Here the data mining concepts and techniques help to uncover interesting hidden data patterns from huge amount of data that represents some useful information. Data Mining refers to extraction of novel, interesting, useful and valid information from the huge data that makes the task easy and this useful information add to our knowledge base helps in the process of decision making. In this process data mining is most essential step. [1]

There is need to understand very large, complex or information-rich datasets. This is common in all fields like business, science, and bioinformatics, marketing, medical

and engineering. In today's competitive world, the importance is to take out useful knowledge from these data sets which are hidden in these datasets and to act on that knowledge. The process of find out the useful knowledge from the datasets using computer-based methodology and including new techniques is called data mining [2].

Data mining is essential process where intelligent methods are applied extract data patterns. It is the process of discovering interesting pattern and knowledge from large amounts of data. The data source can include database, data warehouses, the web, other repositories, or data that are streamed into the system dynamically. In data mining the data is stored electronically and the search is automated or at least augmented by computer. Data mining is about solving problems by analysing data and applying Data mining Techniques. It is very old discipline but in these days, popular due to the successful applications in telecommunication, marketing and tourism. Apart from these applications, data mining could also be used to detect abnormal behaviour e.g. an intelligence agency could determine or know the abnormal behaviour of its employees by using this technology [3].

## II. FEATURE SELECTION

Feature Selection as the name suggests is the selection of features that explains the characteristics of data sets. Attribute or Feature both carries the same meaning. Each attribute in the data set represents some characteristics of data and each has its own relevance. Relevance of attribute is determined by the task that is performed on the data sets. In feature selection process the features that are relevant to the application domain are retrieved by applying some feature selection techniques. The large dataset contains raw data with many irrelevant attributes. The irrelevant attributes may degrade the performance of data mining tasks and techniques such as classification, clustering etc. So, irrelevant attribute needs to be filtered to increase the efficiency and accuracy of such tasks[4].

Feature selection is an amazing pre-processing technique that can do the task accurately by giving the subset of features that are relevant to specific domain. The aim of attribute selection is to enhance the model performance to provide fast and cost effective models for mining. Broadly feature selection techniques can be categorized into three categories filter approach; wrapper approach; embedded approach[5].Also immense quantities of high dimensional data are accumulated challenging state of art in data mining techniques, here feature selection is an

very essential pre-processing step in successful data mining application which can effectively reduce data dimensionality by removing the irrelevant attributes. The Selection of relevant optimal subset of features may add complexity to the model, so the approach used should be efficient and cost effective also. To improve the performance of the model Pre-processing of raw data is done that improves the efficiency and ease of mining process and also decrease the computational cost dramatically. Selection of relevant Attribute subset reduces the size of data sets by removing redundant and irrelevant attributes from the data set. The best attributes/features are selected by performing some statistical test to determine the significance/relevance of attributes in the specific domain. Many other attribute evaluation measures can be used for this purpose such as information gain, gain ratio, PCA(Principal Component Analysis), Relief Attribute Evaluator etc.[6]

#### A. Filter Approach

In Filter approach of feature selection, Firstly features are selected and then induction step is applied on dataset. This approach does not depend on the mining algorithm that is used to extract information from the data set. The relevance of features in this method is determined by intrinsic properties hidden in raw data. The subset of features selected in this way is given as input to the mining task i.e. classification algorithm. The Pros of this filter approach of feature selection are that they are simple, cost effective and scalable to high dimensional datasets. Cons of this approach are that there is no interaction with mining task that is used for evaluation, thus there is no feature dependencies due to which the performance of mining task get affected.[7]

#### B. Wrapper Approach

As opposite to the filter approach of feature selection this approach makes use of data mining algorithm to check the worthiness for feature subset selection a search method is used. By search method feature subset are generated and then evaluated to check the worthiness of subset selected. When comparing to filter approach the wrapper approach is much slower because the data mining algorithms is applied to each attribute subset generated by the search method. Advantages of wrapper approaches are that there is interaction between feature subset search and the model selected for classifier. Disadvantage of wrapper approach is that it is computationally expensive and also higher risk of over fitting[8]

#### C. Embedded Approach

As the name of the approach the filtering technique is incorporated into the classifier itself. As the approach for selection of optimal subset of features is into the classifier itself, the approach is specific to the data mining learning algorithm. The advantages of both filter and wrapper approach are combined in this interaction with the classification model and also less computationally intensive.

### III. LITERATURE REVIEW

The medical diagnosis needs proficiency as well as experience in dealing with uncertainty. Although in these

days, boundaries of medical science have extremely expanded. To overcome this uncertainty chen et al., [2009] proposed a Semantic Relationship Graph (SRG) to describe the relation between multiple tables and the search for relevant features performed within the relational space. Then they optimize the Semantic Relationship Graph by not doing unnecessary joins and removing irrelevant features and relations [9].

Unleretal. [2010] proposed a hybrid model for feature subset selection. This model integrates the techniques of both filter and wrapper technique of feature subset selection. As filter approach is easy and cost effective method of feature subset selection. In the proposed method they first applied filter approach and then they applied wrapper approach. Thus they presented a new method which reduces the computational cost dramatically. By using this hybrid model they performed feature selection and reduction [10].

Zhang et al., [2010] Rough sets are a powerful mathematical tool for analysing various types of data. Rough set approach to data analysis has many advantages like effective algorithms for finding interesting hidden patterns; to identify relationships that would not be easy by using any statically approach. It allows both qualitative and quantitative data. This approach finds minimal sets of data. They proposed an incremental method for dynamic data mining based on rough set theory. Through rough sets they defined composite information systems that contained attributes of multiple different types, which was liable for feature selection and knowledge discovery [11].

Iet al., [2010] they proposed a distributed and parallel Genetic Algorithm for feature selection. GA is an iterative candidate solution. Each solution is obtained by means of an encoding/decoding mechanism, which enables us to represent the solution as chromosome and vice versa [12].

Hsiao et al., [2010]invented a filter model by integrating three well known methods of feature selection PCA(Principal Component Analysis),Decision trees(CART) and Genetic algorithms(GA).The proposed method filter out irrelevant variables are based on union, intersection, and multi-intersection strategies. For prediction, they proposed the back-propagation neural network for making prediction [13].

Uguzet al., [2011] invented the feature selection approach for text categorization and performed feature selection in to two stages by using only filter method of feature selection. Firstly he apply ranker algorithm to assign rank to each term in the document depending on their importance for classification. For assigning rank he used entropy (IG) Information Gain method and rank in decreasing order of their importance in classification. In the next stage he applied two important well known filter techniques GA (Genetic Algorithm) and (PCA) Principal Component Analysis techniques separately to the terms ranked into the document and after that dimension reduction was carried out for feature subset selection [14].

Qaunzet al., [2012] they proposed a novel feature selection technique for real and synthetic data. In this they use popular sparse coding approach and did not used any classifier. For this they proposed a new feature generation algorithm. Starting with the popular sparse coding approach which learns a set of higher order features for the data and verified the effectiveness of the approach on real and synthetic data [15].

Pachecoet al.,[2013] proposed a novel method NSGAFS(non-dominated sorting genetic algorithm).This method as applied on many different databases and verify the worthiness of the method proposed. This method was proposed to remove the feature selection problem for classification [16].

Sun et al., [2013] they proposed a dynamic weighting-based feature selection algorithm that assign ranks to features based on information metric. In this technique weights are assigned according to their interaction with the selected features. This technique not only selects the most relevant features but also retains the most important intrinsic feature groups. Then the weights of features are updated dynamically when each candidate feature had been selected [17].

Cejudo et al., [2013] They compare several feature selection techniques on Enron dataset and invented ABC-Dynf framework. Using naïve bayes classifier, this classification procedure was conducted. The ABC-DynF framework can work under a dynamic feature set [18].

Bina et al.,[2013]classifier does not perform well if the dataset contains many irrelevant features so they proposed a wrapper classifier for predicting the label of classes. To achieve scalability the relational Naïve Bayes classifier exploits independence assumptions. They introduce a weaker independence assumption to the effect that information from different data tables is independent given the class label [19].

laet al ., [2015] They proposed two feature fusion methods are used in this paper: combination fusion and decision fusion aiming to get comprehensive feature representation and improve prediction performance. Decision fusion of subsets that getting after feature selection obtains excellent prediction performance, which proves feature selection combined with decision fusion is an effective and useful method for the task of HIV-1 protease cleavage site prediction [20].

Peker et al .,[2015]They Proposed Effective feature selection algorithms such as minimum redundancy maximum relevance (mRMR); Relief f; and Sequential Forward Selection (SFS) are preferred at the feature selection stage to select a set of features. These obtained features are used as input parameters of the classification algorithms. At the classification stage, six different classification algorithms such as random forest (RF); feed-forward neural network (FFNN); C4.5 decision tree algorithm (C4.5); support vector machines (SVM); naïve

bayes; and radial basis function neural network (RBF) are preferred to classify the problem [21].

CHD (Coronary heart diseases) is one of the major causes of disability. So karaolis et al [2010] developed a system, targeting in the reduction of CHD events. They investigated the three events for CHD. They used the C4.5 decision tree algorithm for the CHD events using five different splitting criteria. The five different criteria are information gain, gini index, gain ratio, chi-squared statistics and distance measure. Thus any one of the splitting criteria investigated could be used for the datasets. Moreover, the extracted models and rules could help to reduce CHD morbidity and possibly, mortality. For developing future events and selection of therapy decision tree could help in the identification of risk subgroups of subjects [22] [23].

The medical diagnosis process can be interpreted or viewed as a decision making process. So, gudadhe et al., [2010] presented a decision support system for heart diseases classification based on SVM and ANN. A MLPNN (multilayer perceptron neural network) with three layer and trained by back-propagation algorithm is employed to develop a decision support system for diagnosis of heart diseases. It is computationally efficient methods. The experimental result shows that the MLPNN with back propagation (BP) algorithm is better or successfully used for diagnosis of heart diseases than SVM. They used the Cleveland Heart Database. The accuracy of the MLPNN is 97.5% and SVM is 80.41%. So this shows that both the methods show the high accuracy to classify the data. But ANN (Artificial Neural network) classifies the data more accurately as compared to SVM (support Vector Machine) [24].

Campadelli et al., [2005] presented an automatic system detecting lung nodule from Postero Anterior Chest Radiographs. They apply three different and consecutive multi-scale schemes to extract set of candidate regions. They used the SVM classification algorithm to get the best result. The classification was performed by applying NN (Neural Network) with different architecture and SVM with different kernels. But the result obtained with SVM because they are the most robust and promising. This result compare with the results obtained using more complicated sets of feature. So SVM is used to select proper set of feature for better results [25].

For detection and data classification the ANN (Artificial Neural Network) architecture i.e. MLP (Multilayer Perceptron) network is widely used. In MLP network the activation function is most important element. For network performance, the selection of activation function is most important. Therefore Isa et al., [2010] investigate the best activation function in MLP in terms of accuracy performance. Various types of activation function are sigmoid, hyperbolic tangent, neuronal, logarithmic, sinusoidal and exponential. For medical diagnosis in case of breast cancer and thyroid diseases detection, MLP networks are trained using BP (Back Propagation) learning algorithm. They investigate that the hyperbolic tangent function in MLP network had the capability to produce the highest



accuracy for detecting and classifying breast cancer data and for thyroid diseases detection neuronal function is most suitable. The highest accuracy achieved during testing was 94% by neuronal function and the accuracy of hyperbolic tangent is 97.2% [26].

Potdukhe et al., [2009] proposed a system called the Ultrasonic image Analysis is used for classifying liver state. The selected parameter are fed into three different classifier i.e. MLP NN, RBF (Radial Base Function) network, and SVM (Support Vector Machine) for classification of liver diseases. Selection of useful features from this group is important to increase accuracy. This method helps in eliminating the defective influence of inhomogeneous structures in liver classification. MLP NN gives the better result as compared with other classifiers with classification accuracy of 94.44% [27].

Jiang et al., [2010] proposed Liver Cancer identification method based on PSO-SVM. In this method PSO (Particle Swarm Optimization) is used to automatically choose parameters for SVM, and it makes the choice of parameter more objective. In traditional methods parameter are decided on the basis of trial and error. The experimental result shows that the proposed parallel PSO-SVM algorithm improves the prediction accuracy of liver cancer [28].

For improvements in the implementation and performance of classifier for medical diagnosis, there is a need to reduce the data dimensionality which is done by complete feature ranking followed by ranking. So, *Abdel et al., [2005]* described an approach for ranking and features in learning algorithm based on the group method of data handling (GMDH). This feature ranking can be used to determine the optimum feature subset. This approach is used on the two medical diagnosis datasets i.e. breast cancer and heart diseases. They used the ROC (Receiver Operating Characteristics) curve to compare the classifier performance. The result shows that the optimal feature subset giving 56% feature selection. We can also use the other learning algorithms and using this technique with other medical datasets [29].

For automatic diagnosis of hepatoma or liver tumor, *caldeira et al., [2008]* proposed a set of features and computation methods to extract them in order to design a classifier. The *primary liver cancer* or *hepatoma* is one of the most lethal forms of cancer and therefore early detection with non-invasive techniques, such as MRI or ultrasound is desirable. So, they used the Dynamic- Contrast Enhanced MRI as a diagnosis tool to assess the malignancy of the liver cancer or tumor. The classification of the tumor can be based on the mean and variance values of the *Maximum, WashIn* and *Wash Out* rates of the perfusion curves inside the tumor. These rates are adequate discriminative features to automatically classify the tumor with respect to its malignancy. For the experimental purpose they used the two data sets from benign and malign tumors [30].

*Sherbini, et al., [2015]* used the LIBS techniques to diagnosis the liver cancer. LIBS stand for "Laser Induced Breakdown Spectroscopy". It is a useful tool for the analysis

of calcified tissues. The elements which are present in the human liver are i.e. Mg, K, Ca, Na, Fe, Mn and Cu are identified by the LIBS technique. It reduces the standard errors. It is a simple technique of diagnosing malignant cells and tissues. The results obtained from the LIBS-Technique were fed-back to an artificial neural network (ANN) to take a decision about the classification of the cancer [31].

For improving feature selection in medical data classification *ya-ju fan et al., [2010]* proposed a new optimization framework i.e. Support Feature Machine (SFM). SFM is used to find the optimal group feature that shown strong separation between two classes. The proposed framework the proposed SFM framework and its extensions were tested on 5 real medical datasets that are related to the diagnosis of epilepsy, breast cancer, heart disease, diabetes, and liver disorders. The objective of SFM optimization model is to maximize the correctly classified data samples in the training set. The outcome of result is compared with the other optimal feature selection technique i.e. Support Vector Machine (SVM), and Logical Data Analysis (LAD). It gives the better result than the SVM and LAD. The result shows that proposed SFM is fast, scalable and very effective [32].

Cancer is one of the dreadful diseases, which causes considerably death in the humans. There are many techniques are available for cancer detection but none of them give or afford considerably accuracy of detection. So, *rajeswari et al., [2011]* used a new method called Gene expression profiling by microarray. This method is an efficient technique for classification and diagnostic prediction of cancer. For identifying the presence of cancer in human, they used the DNA microarray technique. For experimental purpose liver cancer datasets is used and for implementation MATLAB tool is used [33].

Recent research studies on liver diagnosis indicated K-Nearest Neighbour classifier is to be giving best results with 'India liver patients' data set with all feature set combinations. Performance is better for the India Liver dataset compared to UCLA liver dataset with all the selected algorithms. So, to envisage the reason for this difference *venkataramana et al., [2012]* proposed to analyze the liver patient's populations of both USA and India. For this they used the ANOVA, MANOVA analysis on these data sets in three ways [34].

Knowledge of the understanding of human congenital diseases is complex. Significantly, much of understanding of organ development has arisen from analyses of patients with liver deficiencies. So, *rajeswari et al., [2010]* used the data classification is based on liver disorder the training dataset is developed by collecting data from UCI repository consists of 345 instances with 7 different attributes. The instances in the dataset are pertaining to the two categories of blood tests which are thought to be sensitive to liver disorders that might arise from excessive alcohol consumption mechanisms. Such knowledge also provides a basis, labelled as Low (L), and (H) to represent the profit as 0 and 1 which result in accuracy and time taken to build the algorithm. WEAK tool is used to classify the data and the data is

evaluated using 10-fold cross validation and the results are compared [35].

Data mining techniques on the biological analysis are spreading for most of the areas including the health care and information. So, seker et al., [2013] applied the data mining techniques, such as KNN, SVM, MLP or decision trees over a unique dataset, which is collected from 16,380 analysis results for a year. The results show that there is a correlation among ALT, AST, Billirubin Direct and Billirubin Total down to 15% of error rate. Also the correlation coefficient is up to 93% [36]. In recent years in healthcare sectors, data mining became an ease of use for disease prediction. It is a very challenging task to the researchers to predict the diseases from the voluminous medical databases.

To overcome this issue vijayarani et al., [2015] used the classification algorithms (SVM, Naïve Bayes) to predict liver diseases. These classifier algorithms are compared based on the performance factors i.e. classification accuracy and execution time. From the experimental results it is observed that the SVM is a better classifier for predict the liver diseases [37].

**IV. METHODOLOGY USED**

In Medical diagnosis redundant feature or irrelevant feature degrade the performance of classifier used for classification. To improve diagnosis of diseases in the field of medical era redundant or irrelevant features should be removed. In this work feature selection technique Relief F Attribute Evaluator has been used to evaluate the worthiness of attributes along with Ranker Algorithm that assign rank according to evaluation by the evaluator then subjective measure has been used to select only those attribute having positive rank or by setting a threshold value for the selection of relevant attributes. In this way only relevant attributes has been selected that has relevance to the application and subjective measure ensure that no relevant attribute left which has relevance to the application. After then to evaluate the performance whether increased or decreased classifier has been used. For this purpose Multilayer perceptron Classifier has been used because it has many advantages over other classifier as it used neural network as working principle and neural network is a field of Artificial Intelligence which is currently the main area of research. So for evaluation MLP (Multilayer perceptron Classifier) has been used.

**V. RESULTS AND INTERPRETATIONS**

DATA SET 1: chronic kidney disease No. of attributes 25. For selecting attributes relief attribute evaluator has applied along with ranker algorithm relief f attribute evaluates the worthiness of attribute and ranker algorithm assign rank according to the worthiness of a particular attribute/feature. The attribute ranking is as shown in the Table 1.

Assigned Rank	Attribute Name	Assigned Rank	Attribute Name
0.5056	sg	0.13443	pcv
0.29137	dm	0.12288	pc
0.25912	htn	0.09125	pe
0.17625	rbc	0.087	appet
0.16996	al	0.07	ane
0.13916	hemo	0.04508	rbcc
0.03616	wbcc	0.00756	bu
0.02838	cad	0.00337	bp
0.02825	pcc	0.00151	sod
0.0159	sc	0.00117	su
0.01427	age	-0.0015	ba
0.00924	bgr	-0.01423	pot

Table 1. List of Attributes of Chronic Kidney Data Set along with assigned rank.

Subjective measure is used for selecting the attributes. To measure the performance Multilayer perceptron classifier has been applied before feature selection and after feature selection and then comparison is made between the two as shown in the table Table 2.

Sr. No.	Comparison Parameter	Attribute Before feature Selection (24)	Attribute after feature selection(18)
1	Time taken to build model	14.25sec	7sec
2	Kappa statistics	0.9947	0.9947
3	Mean absolute error	0.0085	0.0078
4	Root Mean Squared error	0.0622	0.0524
5	Relative absolute error	1.81%	1.67%
6	Root relative squared error	12.86%	10.83%

Table 2. Interpreted result on Chronic Kidney Data Set.

DATA SET 2: The name of the dataset is hepatitis that is also taken from UCI machine repository which consist of 155 instances. For experiment purpose the data set is divided into two parts training and test data set.

Assigned Rank	Name of Attribute	Assigned Rank	Name of Attribute
0.06223	PROTIME	-0.00206	MALAISE
0.03731	ALK PHOSPHATE	-0.00369	SPLEEN PALPABLE
0.0359	Class	-0.01067	ASCITES
0.02442	STEROID	-0.01222	VARICES
0.0193	ALBUMIN	-0.01281	BILIRUBIN
0.0126	AGE	-0.01482	FATIGUE
0.01143	LIVER BIG	-0.01792	ANOREXIA
0.00864	SGOT	-0.02777	ANTIVIRALS
0.00453	SEX	-0.02909	SPIDERS
		-0.03689	LIVER FIRM

Table 3. List of Attributes of hepatitis Data Set along with assigned rank.

Subjective measure is used for selecting the attributes. To measure the performance Multilayerperceptron classifier has been applied before feature selection and after feature selection and then comparison is made between the two as shown in the table Table 3.

Sr.No	Comparison Parameter	Attributes before feature selection(19)	Attributes after feature selection(10)
1	Time taken to build model	1.15sec	0.33sec
2	Co-relation coefficient	0.1035	0.2244
3	Mean absolute error	0.6234	0.481
4	Root mean squared error	0.8029	0.6238
5	Relative absolute error	124.32%	95.92%
6	Root relative squared	159.27%	123.75%

Table 4. Interpreted result on Hepatitis Data Set.

From the above table it is clear that time taken to build the model and errors has been reduced to a great extent while co-relation coefficient increased which describes that accuracy of classifier increased by selecting only relevant attributes.

## VI. CONCLUSION

From the above performed experiments it is concluded that performance and accuracy of classifier increased by applying feature selection techniques before applying

classifier which is clear from the results obtained by experiment shown in Table 2 and Table 4.

## REFERENCES

- [1]. J. Han, M. Kamber, and J. Pei, "1 - Introduction," in Data Mining (Third Edition), Boston: Morgan Kaufmann, pp. 1–38,2012.
- [2]. MehmedKantardzic, "Data Mining: Concepts,"inData Mining: Concepts, Models, Methods and Algorithms,2nd Edition, Wiley IEEE Press, pp. 1-25,July 2011.
- [3]. BhavaniThuraisingham, "Introduction," in Data Mining: Technologies, Techniques, Tools and Trends"CRC Press LLC, pp. 1-14, 1999.
- [4]. S. Beniwal and J. Arora, "Classification and feature selection techniques in data mining," In International Journal of Engineering Research & Technology, vol. 1, no. 6, pp. 1-6cc,2012.
- [5]. H. Liu and H. Motoda, Feature Selection for Knowledge Discovery and Data Mining. Springer Science & Business Media, 2012.
- [6]. H. Liu and H. Motoda, Computational Methods of Feature Selection. CRC Press, 2007.
- [7]. W. Awada, T. M. Khoshgoftaar, D. Dittman, R. Wald, and A. Napolitano, "A review of the stability of feature selection techniques for bioinformatics data," in Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on, pp. 356–363,2012.
- [8]. Y. Han, K. Park, and Y.-K. Lee, "Confident wrapper-type semi-supervised feature selection using an ensemble classifier," in Artificial Intelligence, Management Science and Electronic Commerce
- [9]. H. Chen, H. Liu, J. Han, X. Yin, and J. He, "Exploring optimization of semantic relationship graph for multi-relational Bayesian classification," Decis. Support Syst., vol. 48, no. 1, pp. 112–121, 2009.
- A. Unler, A. Murat, and R. B. Chinnam, "mr 2 PSO: a maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification," Inf. Sci., vol. 181, no. 20, pp. 4625–4641, 2011.
- [10]. J. Zhang, T. Li, and H. Chen, "Composite rough sets for dynamic data mining," Inf. Sci., vol. 257, pp. 81–100, 2014.
- [11]. R. Li, J. Lu, Y. Zhang, and T. Zhao, "Dynamic Adaboost learning with feature selection based on parallel genetic algorithm for image annotation," Knowl.-Based Syst., vol. 23, no. 3, pp. 195–201, 2010.
- [12]. C.-F. Tsai and Y.-C. Hsiao, "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches," Decis. Support Syst., vol. 50, no. 1, pp. 258–269, 2010.
- [13]. H. Uğuz, "A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm," Knowl.-Based Syst., vol. 24, no. 7, pp. 1024–1032, 2011.
- [14]. B. Quanz, J. Huan, and M. Mishra, "Knowledge transfer with low-quality data: A feature extraction issue," Knowl. Data Eng. IEEE Trans. On, vol. 24, no. 10, pp. 1789–1802, 2012.

- [15]. S. Casado, F. Angel-Bello, A. Álvarez, and others, "Bi-objective feature selection for discriminant analysis in two-class classification," *Knowl.-Based Syst.*, vol. 44, pp. 57–64, 2013.
- [16]. J. Sun and A. Zhou, "Unsupervised robust Bayesian feature selection," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 558–564, 2014.
- [17]. J. M. Carmona-Cejudo, G. Castillo, M. Baena-García, and R. Morales-Bueno, "A comparative study on feature selection and adaptive strategies for email foldering using the ABC-DynF framework," *Knowl.-Based Syst.*, vol. 46, pp. 81–94, 2013.
- [18]. B. Bina, O. Schulte, B. Crawford, Z. Qian, and Y. Xiong, "Simple decision forests for multi-relational classification," *Decis. Support Syst.*, vol. 54, no. 3, pp. 1269–1279, 2013.
- [19]. H. Liu, X. Shi, D. Guo, Z. Zhao, and others, "Feature Selection Combined with Neural Network Structure Optimization for HIV-1 Protease Cleavage Site Prediction," *BioMed Res. Int.*, 2015.
- [20]. M. Peker, A. Arslan, B. Sen, F. V. Çelebi, and A. But, "A novel hybrid method for determining the depth of anesthesia level: Combining ReliefF feature selection and random forest algorithm (ReliefF+RF)," in *Innovations in Intelligent Systems and Applications (INISTA), 2015 International Symposium on*, pp. 1–8, 2015.
- [21]. Ilias Maglogiannis, Elias Zafiroopoulos, Ioannis Anagnostopoulos, "An Intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers," *Springer Science LLC*, pp. 24–36, 12 July 2007.
- [22]. Minas A. Karaolis, Joseph A. Moutiris, Demetra Hadjipanayi, Constantinos S. Pattichis, "Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining with Decision Tree," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, No. 3, May 2010.
- [23]. MrudulaGudadhe, KapilWankhade, SnehlataDongre, "Decision Support System For Heart Diseases based on Support Vector Machine and Artificial Neural Network," *IEEE International Conference on Computer and Communication Technology (ICCCCT)*, pp. 741-745, 2010.
- [24]. Paola Campadelli, Elena Casiraghi, Giorgio Valentini, "Lung Nodules Detection and Classification," *IEEE*, 2005.
- [25]. I.S. Isa, Z. Saad, S.Omar, M.K. Osman, K.A. Ahmad, H.A. Sakim, "Suitable MLP Network Activation Functions for Breast cancer and Thyroid Diseases Detection," *IEEE Second International Conference on Computational Intelligence, Modeling and Simulation*, pp. 39-44, 2010.
- [26]. M. R. Potdukhe, P.T. Karule, "MLP NN based DSS for analysis of ultrasonic liver image and diagnosis of liver diseases," *IEEE Second International Conference on Emerging Trends in Engineering and Technology (ICETET)*, pp. 67-72, 2009.
- [27]. Huiyan Jiang, Fengzhen Tang, Xiyue Zhang, "Liver Cancer Identification Based on PSO-SVM Model," *IEEE 11<sup>th</sup> International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 2519-2523, 7-10<sup>th</sup> December 2010.
- [28]. R. E. Abdel-Aal, "GMHD-based feature ranking and selection for improved classification of medical data," *Elsevier Science Journals of Biomedical Informatics* 38, pp. 456-468, 2005.
- [29]. Liliana Caldeira, Isabela Silva and Joao Sanches, "Automatic Liver Tumor Diagnosis with Dynamic-Contrast Enhanced MRI," *IEEE*, pp. 2256-2259, 2008.
- [30]. Ashraf M. EL Sherbini, Mohamed M. Hagra, Hania H. Farag, Mohamed R.M. Rizk, "Diagnosis and Classification of Liver Cancer using LIBS Technique and Artificial Neural Network," *International Journal of Science and Research (IJSR)*, Vol. 4, Issue 5, pp. 1153-1158, May 2015.
- [31]. Ya-Ju Fan, Wanpracha Art Chaovalitwongse, "Optimizing Feature Selection to Improve Medical Diagnosis," *Springer Science*, pp. 169–183, 2010.
- [32]. P. Rajeswari, G. Sophia Reena, "Human Liver Cancer Classification using Microarray Gene Expression Data," *International Journal of Computer Applications (IJCA)*, Vol. 34, No.6, pp. 25-37, November 2011.
- [33]. Bendi Venkata Ramana, M. Surendra Prasad Babu, N. B. Venkateswarlu, "A Critical Comparative Study of Liver Patients from USA and INDIA: An Exploratory Analysis," *International Journal of Computer Science Issues (IJCSI)*, Vol. 9, Issue 3, No. 2, pp. 506-516, May 2012.
- [34]. P. Rajeswari, G. Sophia Reena, "Analysis of Liver Disorder Using Data Mining Algorithm," *Global Journal of Computer Science and Technology*, Vol. 10, Issue 14, pp. 48-52, Nov. 2010.
- [35]. Sadi Evren SEKER, Yavuz UNAL, Zeki ERDEM, H. Erdinc KOCER, "Correlation between Liver Analysis Outputs," *Proceedings of the 2013 International Conference on Systems, Control and Informatics*, pp. 217-221, 2013.
- [36]. S. Vijayarani, S. Dhayanand, "Liver Disease Prediction using SVM and Naïve Bayes Algorithms," *International Journal of Science, Engineering and Technology Research (IJSETR)*, Vol. 4, Issue 4, pp. 816-820, April 2015.