A SURVEY ON IMPLEMENTATION OF MACHINE LEARNING TECHNIQUES FOR DERMATOLOGY DISEASES CLASSIFICATION

[1]MadhuraRambhajani, [2]Wyomesh Deepanker, [2]Neelam Pathak
[1]M-Tech Scholar, [2]Assistant Professor,
Technocrats Institute of Technology (Excellence), Bhopal, India

ABSTRACT
The availability of various computational and predictions tools helps doctors and practitioners to cope up with their problems in their daily clinical tasks. These computational methods are known as machine learning techniques which is a novel approach to improve its performance through the use of software to mimic the ways by which human learn such as repetition, experience etc. These methods have been successfully implemented in various sectors like finance, health care etc .This paper presents a review on various machine learning techniques such as Data mining, Soft Computing, Hybrid method etc and the survey of application of machine learning techniques for classification of various dermatology diseases in past two decades.

KEYWORDS: *Machine Learning, Health care, statistical, data mining, soft computing, hybrid Classification*

I. INTRODUCTION

Today diseases diagnosis is a very crucial task in medical science. It is necessary to interpret the correct diagnosis of patients with help of clinical investigations and examination. Computer based decision support system can play an important role in accurate diagnosis and cost effective treatment. Now a day's health care domain gathers a bulk amount of information or data regarding clinical examination, patient report, treatment, follow ups, medicine etc is difficult to organize in proper manner. Due to improper organization of the data the quality of decision making is getting affected. This increase in volume of bulk of data requires some way in whichdata can be extracted and processed efficiently. Health care industry today generate a large amount of complex data about patients record, diseases diagnosis, hospital resource , electronic patient records etc There are wide applications in health care sector which are as follows :
(I)Treatment effectiveness
(ii)Health-care management
(ii)Customer relationship management
(iv)Fraud and abuse
(v)Pharmaceutical management
This can be possible with the help of information technology, the use of information technology is being increasingly implemented in health care organization in order to help doctors in their day to day decision making activities. It helps doctors and physicians in diseases management, tests, medications and discovery of patterns and relationships among clinical and diagnosis data and as well as employ machine learning techniques.

Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that one can probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human level. Machine-learning techniques[19] are data-driven approaches that are designed to discover statistical patterns in high-dimensional, multivariate data sets, such as those that are frequently found in electronic health record (EHR) systems. The theme of Machine Learning Techniques is in identification pattern that provides support for predictions and decision making process for diagnosis and treatment planning. The Machine Learning Techniques [19] has been applied with success to different fields other than health care like marketing, banking, Customer relationship management, engineering, crime analysis, mobile computing and various fields of science. Machine learning algorithms have been used in a variety of applications. They have been shown to be of special use in data mining scenarios involving large databases and where the domain is poorly understood and therefore difficult to model by humans. These techniques are able to handle large amounts of data, to integrate data from different sources, and to incorporate background knowledge in the analysis.

In this paper rest of the section is organized is as follow: It first gives detail explanation of machine learning techniques and its different categories. In section 3 we have discussed about contribution done by different researchers in context to differential diagnosis of erythemato squmous diseases and its classification. The article ends by concluding with a summary of investigated methods with their results.

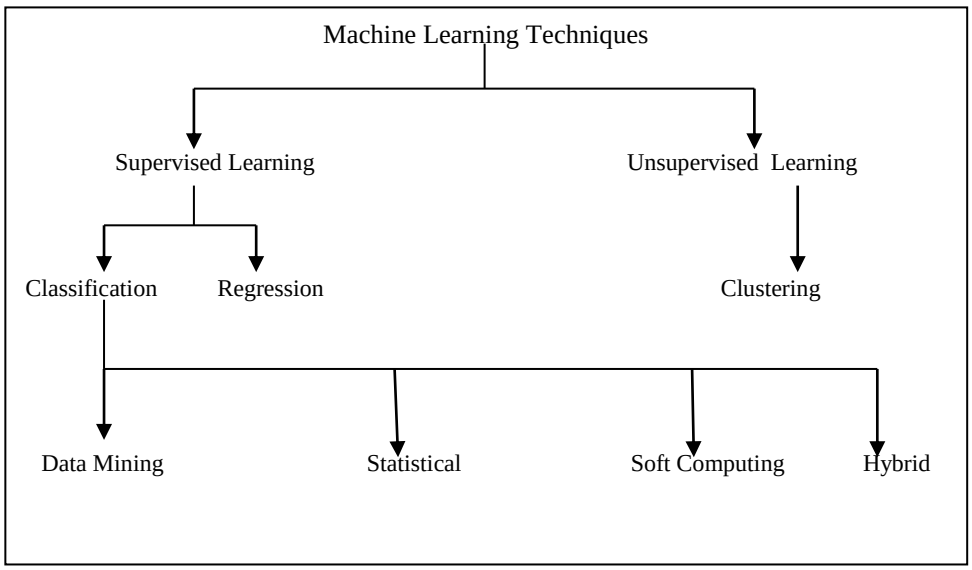

Fig 1: Classification of Machine Learning techniques

II. REVIEW OF MACHINE LEARNING TECHNIQUES

Machine learning [19] is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can teach themselves to grow and change when exposed to new data. Maching learning, concerns the construction and study of systems that can learn from data. For example, a machine learning system could be trained on email messages to learn to distinguish between spam and non-spam messages. After learning, it can then be used to classify new email messages into spam and non-spam folders. Machine learning [6] in terms of ability of a computer program to improve its own performance, in some domain, based on the past experience. As per fig 1 Machine learning algorithms are classified in two types supervised learning and unsupervised learning. Supervised learning based on input-output pairs patterns. In these algorithms aims to predict output values based on input values. Supervised learning mainly focuses on classification and

regression. In unsupervised learning algorithms training examples contain input values in this target values are not associated with input values. Unsupervised learning focuses on clustering approaches. There are various machine learning techniques which have been used in this health care for classification of various diseases which are as follows.

2.1) Data mining: Data mining [3] is the non trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. With the widespread use of databases and the explosive growth in their sizes, organizations are faced with the problem of information overload. The problem of effectively utilizing these massive volumes of data is becoming a major problem or all enterprises.

There are wide ranges of data mining tools which help to analyze patterns, relations from which data can be refined and can be transformed into a new data.

Decision tree [3] is one of the most popular predictive data mining techniques. Decision tree builds classification or regression model in form of tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. Numerous algorithms have been developed, each using one or more different techniques to identify patterns that are useful in generating a concept description.CART (Classification and Regression Technique) [3] is one of the popular methods of building decision tree in the machine learning community. CHAID (Chi-Squared Automation Interaction Detection**)** is a derivative of AID (Automatic Interaction Detection); it attempts to stop growing the tree before over fitting occurs. CHAID avoids the pruning phase. In the standard manner, the decision tree is constructed by partition the data set into two or more data subsets, based on the values of one of the non-class attributes. After the data set is partitioned according to the chosen attributes, each subset is considered for further partitioning using the same algorithm. Each subset is partitioned without regard to any other subset. The process is repeated for each subset until some stopping criteria is met. In CHAID, the number of subsets in a partition can range from two up to the number of distinct values of the splitting attribute.Quinlan introduced the ID3, Iterative Dichotomizer 3, [3] for constructing the decision tree from data. In ID3, each node corresponds to a splitting attribute and each arc is a possible value of that attribute. At each node the splitting attribute is selected to the most informative among the attributes not it considered in the path from the root. Entropy is used to measure how informative is a node. This algorithm uses the criterion of information gain to determine the goodness of a split. The attribute with the greatest information gain is taken as the splitting attribute, and the data set is split for all distinct values of the attribute.C4.5 [3] is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier. C4.5 builds decision trees from a set of training data in the same way as ID3, using the concept of information entropy. C4.5 has additional features such as handling missing values, categorization of continuous attributes, and pruning of decision trees, rule derivation and others. C4.5 constructs a very big tree by considering all attribute values and finalizes the decision rule by pruning. It uses a heuristic approach for pruning based on the statistical significance of splits.

b) Statistical Techniques: Statistical models [19] involving a latent structure often support clustering, classification, and other data mining tasks. Because of their ability to deal with minimal information and noisy labels in a systematic fashion, statistical models of this sort have recently gained popularity, and success stories can be found in a variety of applications; for example, population genetics, scientific publications, words and images, disability analysis, fraud detection, biological sequences & networks. It is concerned with the development of algorithms and techniques that learn from observed data by constructing stochastic models that can be used for making predictions and decisions. Statistical techniques have been widely used in health care for the classification of various diseases Statistical techniques have been widely used in health care for the classification of various diseases like Bayesian net, Support vector machine. The Bayesian approach assumes that pattern possesses random characteristics and they are generated in a random way by some natural phenomena and process. The natural choice for dealing with random and uncertain pattern is to use statistical technique based on probabilistic characteristics of data. The Bayesian method is based on the assumption that the classification of patterns is expressed in probabilistic terms. Support Vector Machine (SVM) [3] is a robust classification and regression technique that perform binary

classification. SVM are also recognized as efficient tools for classification and learning because of two important reasons, first, unlike the other classification methods, SVM minimizes the expected rate rather than minimizing the classification error. Secondly SVM employs the duality theory of mathematical programming to get a dual problem that admits efficient computational methods. SVM is particularly suited to analyzing data with extremely large numbers (for example, thousands) of predictor fields

c) Soft computing techniques: Soft Computing [24] is the fusion of methodologies that were designed to model and enable solutions to real world problems, which are not modeled or too difficult to Model, mathematically. Soft computing is a consortium of methodologies that works synergistically and provides, in one form or another, flexible information processing capability for handling real-life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning and partial truth in order to achieve tractability, robustness and low-cost solutions. The guiding principle is to devise methods of computation that lead to an acceptable solution at low cost, by seeking for an approximate solution to an imprecisely or precisely formulated problem. Artificial Neural Network (ANN) is known as best classifier and is able to mine huge amount of data for classification. They were originally developed in the field of machine learning and the concept of neural network is similar to that human brain. A neural network is composed of a set of elementary computational units, called neurons, connected together through weighted connections. These units are organized in layers so that every neuron in a layer is exclusively connected to the neurons of the preceding layer and the subsequent layer. Every neuron, also called a node, represents an autonomous computational unit and receives inputs as a series of signals that dictate its activation. Following activation, every neuron produces an output signal. Genetic algorithm basic concepts were developed by Holland, while the practicality of using the GA [24] to solve complex problems was demonstrated in. GA starts off with population of randomly generated chromosomes, each representing a candidate solution to the concrete problem being solved and advances towards better chromosomes by applying genetic operators based on the genetic processes occurring in nature. Associated with each chromosome at every generation is a fitness value, which indicates the quality of the solution. Like in nature, the new chromosomes are created using genetic operators such as crossover and mutation.Feature selection [16] is a process of finding the best feature subset from original set offeatures, according to some defined feature selection criterion, withoutfeature constructionor transformation. The main benefits of using feature selection techniques are improved model interpretability, Shorter training time, Enhanced generalization by reduced overfitting.There are various methods of feature selection which have been used in this study like ranker method, forwards sequential search etc.

d) Hybrid Technique: The term hybrid [19] means to combine two or more techniques or approaches to build a new approach. The term hybrid can also be used to refer as Ensemble. The term hybrid is introduced here because there are some criteria in which some properties are required which is not present in which is not present in single method , so in order to obtain some desired output two or more techniques are combined to form an ensemble model or hybrid approach which yields best results in performance. Hybrid techniques have been used in many practical biomedical management, data mining and text mining approaches. For example fuzzy logic has been applied to rule induction and genetic algorithm, genetic algorithm, has been combined with neural network

III. LITERATURE SURVEY

Bruno Fernandez et.al [5] has compared various data mining algorithm such as Bayesian techniques and LMT for classification and association on UCI repository breast cancer dataset, dermatology dataset and vertebral column dataset. After analysis Bayes Net was found to be the best classification algorithm for Dermatology and breast cancer dataset with 97% of classification accuracy. For vertebral column the best algorithm was logistic model tree with 85% of accuracy. H.AltayGuvenir et.al [11] [4] has proposed a new classification algorithm VFI5 and has applied to problem of differential diagnosis of erythemato squamous. There are many authors who have used dermatology dataset from UCI (University of California at Irvine) starting from his work where he applied his newly developed algorithm VFI5.This represents a concept description by a set of feature intervals. The classification of a new instance is based on a voting among the classification made by the values

of each feature separately. All training examples are processed at once. The VFI5 algorithm constructs intervals for each feature from the training examples. For each interval, a single value and the votes of each class in that interval are maintained. Thus, an interval may represent several classes by sorting the vote for each class. This algorithm has obtained 96.25% of classification accuracy. After that [12] he extended his work and presented an expert system incorporating by three classification algorithms: nearest neighbor, naïve Bayesian classifier and VF15.In this they have obtained 99.25% of accuracy.

Dinesh k Sharma et.al [7] have proposed a hybrid ensemble model by combining Support Vector Machine and Artificial Neural network. This ensemble model has been tested on UCI repository Dermatology dataset and has achieved a classification accuracy of 99.25% and 98.99% at training and testing stages separately.Alaa M. Elsayad[2] has developed an ensemble model by combining three techniques c5.0, Multilayer perceptron, Linear Discriminant Analysis and has achieved 98.23% of accuracy at testing stage.

Ubeyli et.al[9] have used a hybrid technique in which he has combined multiclass SVM with error correcting output code (ECOC) for the diagnosis of erythemato squamous diseases. For this erythemato squamous disease dataset has been downloaded from UCI repository consisting of 34 features and six different classes. This model has achieved classification accuracy of 98.32%. Ubeyli [10] have also used combined neural network to guide model selection for the diagnosis of erythemato squamous and has achieved 97.77% of classification accuracy. Ubeyli and Guler [8] have used ANFIS I.e. adaptive neuro fuzzy inference system for diagnosis of erythemato squamous. In this in order to classify six dermatology diseases six ANFIS classifiers are used when the inputs were 34 features defining six disease indications. This methodology has obtained 95.50% of accuracy.

Polat and gunes [14] have used fuzzy weighted pre-processing-Nearest Neighbor based on weighted preprocessing and a decision tree classifier for the diagnosis of erythemato squamous diseases. In this study authors have achieved classification accuracy of 88%, 97.57% and 99% respectively. Polat and Gunes [15] have also used data mining method C4.5 cross validation method i.e. one against all. In this study the features of a patient were represented as a vector of features. C4.5 Decision tree learning is a method in which the learned function is represented by a decision tree. Learned trees can be represented as sets of if-then rules to improve human readability. The aim of C4.5 Decision tree learning is recursively partition data into sub-groups [14]. This method has achieved classification accuracy of 96.71%.

JuanyingXie et.al [26][4]have proposed a hybrid method of Support Vector Machine and f score and sequential forward search for UCI machine learning repository Dermatology dataset. In the process of filter, they calculated the improved *F*-score for each feature, and then sort them in descending order. In this the original *F*-score is a simple filter technique which measures the discrimination of two sets of real numbers In the process of wrapper, a subset of the original training set was generated by including the features with top *N F*-scores, where $N = 1, 2, \ldots,$ m and m was the total number of features. Then a grid search was carried out to find the optimized value. This procedure, using SFS, was carried out until all features appeared in the subset. Finally they obtained the SVM diagnosis model which had 98.61% classification accuracy. With this model they have achieved 98.61% of accuracy with 21 features.

C.C Lopes et.al [6] [4] have used Genetic algorithm and C4.5 decision tree algorithm on UCI machine. In order to discover the rules, a constrained-syntax GP algorithm was developed which was based on some concepts of data mining, particularly with emphasis on discovering the comprehensible knowledge .In this 5 fold cross validation is performed. GP was run once for each class. Once all runs of GP for a given dataset was completed, all the rules found by GP in that experiment were grouped into a rule set. The proposed Gp method obtained 96.64% of accuracy which is considerably better than c4.5.C4.5 have obtained 89.12% of classification accuracy. Akin Qzcift et.al [1] have proposed a hybrid method of genetic algorithm wrapped Bayesian network Feature selection for Erythemato squamous dataset. In GA–BN algorithm, GA makes a heuristic search to find most relevant feature model that increase accuracy of BN algorithm with the use of a10-fold cross-validation strategy. This algorithm produces 99.20% of accuracy.

Mohammad javed Abdi [17] et.al have also proposed a hybrid method based on particle swarm optimization. Support vector machine, and association rules for the diagnosis of erythemato squmous diseases. This model produces 98.91% of classification accuracy with 24 features from the dataset.

Pasi Luukka [23] have used principal component Analysis for the diagnosis of Liver Disorder, Dermatology, Hepatitis. For this he has downloaded all the dataset from UCI machine learning repository. After analysis he found that classification accuracy for dermatology dataset was highest with 97.09%.While classification accuracies for Liver Disorder and Hepatitis is72.27% and 88.94% respectively.

Lukka et.al [21] have used fuzzy similarity based classification method for the diagnosis of erythemato squamous and have obtained 97.02% of accuracy. Lukka [22] have also applied similarity measure for dermatology classification and have obtained 97.80% of accuracy. Nanni[20][4] have also contributed for the squamous diagnosis and used techniques such as LSVM,RS and have used cross validation methods and have produced classification accuracies of 97.22%,97.22%,97.50%, 98.20%, 97.22%, 97.50%, 97.80%, 98.30%. Karbatak and Inee[18] have proposed a hybrid method of association rules and Neural Network for the classification of dermatology diseases and have obtained 98.61% of accuracy.Kalik.b et.al [13] have proposed a system for the diagnosis of erythemato squamous. In this author have used supervised back propagation algorithm which is used to train the network .The average value of classification accuracy was 95%.

Lekkas et.al [5] [25] have used evolving fuzzy classification methods for the diagnosis of erythemato squamous and diabetes. This evolving fuzzy classification methods allows data to be processed in online mode by recursively modifying a fuzzy rule base on a per-sample basis from data streams. They presented a study of semi-supervised evolving fuzzy classification on the diagnostics for two well known medical problems i.e pima Indian diabetes and dermatology dataset. In regards to Indian Pima diabetes dataset an accuracy of 79.37% of accuracy is achieved, while for the erythemato squamous dataset classification accuracy of 97.55% of accuracy is obtained. In both these studies erythemato squamous and pima Indian diabetes dataset have been downloaded from UCI machine learning repository.

IV. DISCUSSION

This section presents a comparative study of data mining applications in classification of dermatology diseases by different researchers. Mainly data mining tools are used to predict the successful results from the data recorded on healthcare problems. Different data mining tools are used to predict the accuracy level for the diagnosis of different dermatology diseases. Here Table 1 shows the comparison of various methods for diagnosis dermatology diseases in terms of accuracy. Bruno Fernandez et.al[5] have compared bayes net and LMT ,in which bayes net approach had better accuracies in comparison to LMT.Ubeyli and Guler[8] have proposed Adaptive neuro fuzzy inference system and have obtained 95.5 accuracy and Ubeyli et.al [10] have used combined neural network and has obtained 97.77 accuracy and they have also proposed another method[9] Support vector machine with Error correcting codes and have obtained 98.32 of accuracy which was considerably better than other two approaches. Lukka et al[21][22] have used fuzzy similarity and similarity measures and has obtained 97.02 and 97.8 of accuracy. Lekkas et.al[25] have also proposed fuzzy methods and has achieved classification accuracy of 97.55%.Pasi Lukka[23] using principle component analysis has obtained 97.09% of accuracy.Karbatak and Inee[18] have used association rules in combination with neural network and has obtained 95% of accuracy and in comparison Dinesh k.Sharma[7] have used artificial neural network with support vector machine has obtained 99.25 % of accuracy.Alaa M. Elsayad[2]have proposed a hybrid method using C5.0, Multilayer perceptron and Linear Discriminanat analysis and has obtained 98.23 % of accuracy and M. javed Abdi et.al[17] have proposed a hybrid method based on particle swarm optimization , support vector machine and association rules and has obtained 98.91% of accuracy, Akin Qzcift et.al[1] have also proposed a hybrid method based on Genetic algorithm and Bayesian network and has obtained 99.2% of accuracy This model is as much as accurate as above two hybrid model. Polat and Gunes[15] have used C4.5 method one against all and has obtained 96.71% of accuracy. H.AltayGuvenir et.al[11] have proposed VFI5 for the differential diagnosis of erythemato squamous and has obtained 96.25% of classification accuracy. Nanni and Xie and Wang [20][26] both used SVM and feature selection methods. The method proposed by achieved 98.61% classification accuracy and its accuracy rates and the model proposed by Nanni are very close. Among all these Guvenir and Emeksiz [12] had the

highest classification accuracy, 99.25%, on the differential diagnosis of erythemato-squamous diseases using voting feature intervals-5 ,Nearest neighbor and naïve bayes.

Table I Results of different methodologies of dermatology diseases

S.no	Author Name	Technique	Outcome	S.no	Author Name	Technique	Outcome
1	Alaa M.Elsayad[2]	c5.0,MLP,LDA	98.23	10	Lukka et.al[21]	Fuzzy similarity	97.02
2	Akin Qzcift et.al [1]	GA-BN	99.2	11	Lukka et.al [22]	Similarity measures	97.8
3	Bruno Fernandez et.al [5]	Bayes net,LMT	97 85	12	Pasi Lukka[23	PCA	97.09
				13	Ubeyli et.al [9]	SVM-ECOC	98.32
4	Dinesh K Sharma et.al[7]	ANN-SVM	99.2	14	Ubeyli et.al [10]	CNN	97.77
5	H.AltayGuvenir et.al[11]	VF[15]	96.25	15	Ubeyli and Guler[8]	ANFIS	95.5
6	Guvenir and Emeksiz[12]	Nearest Neighbour, Naïve Bayes,VFI5	99.25	16	Polat and Gunes[15]	C4.5	96.71
7	M. javed Abdi et.al [17]	PSO-SVM-AR	98.91	17	Lekkas et.al	Fuzzy methods	97.55
8	Nanni[20]	LSVM,RS	97.22-98.30	18	JuanyingXie et.al[26]	SVM-f search SFS	98.61

V. CONCLUSION

In this paper various machine learning techniques used in various health care sectors have been discussed. Due to the large volume and complexity of data there is a need to process theses data. For this medical data mining can help to prepare some methods for diagnosis and decision making activites.In this we have focused on the use of machine learning techniques for dermatology classification in past years. Classification of dermatology diseases is a difficult task because all of them share the same clinical features. There are several researchers who have worked on it using dermatology dataset. In this we have presented some of them with their results.

VI. FUTURE SCOPE

In this paper survey related to the classification of dermatology diseases has been presened.In future survey of other diseases like thyroid, heart diabetes etc can be presented. In future comparison can be made of application of any machine learning technique for various diseases.

REFRENCES

[1]Akin Ozcift, ArifGulten."Genetic algorithm wrapped Bayesian network feature selection applied to differential diagnosis of erythemato squamous diseases", Digital Signal Processing, 23(1), 230-237, jun2013

[2] Alaa .M Elsayad,." Diagnosis of erythemato-squamous diseases using ensemble of data mining methods".ICGST-BIME Journal, 10(1), 13-23,2010.

[3]Arun K. Pujari, Data mining techniques, 4th edition, Universities Press (India) Private Limited,2001.

[4] Barati, E., Saraee, M., Mohammadi, A., Adibi, N, &Ahamadzadeh, M.R." A survey on utilization of datamining approaches for dermatology (Skin) disease prediction," Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Health Informatics (JSHI), March Edition, 1-11,2011.

[5] Bruno Fernandes Chimieski1, Rubem Dutra RibeiroFagundes," Association and Classification Data Mining Algorithms Comparison over Medical Datasets", J. Health Inform. Abril-Junho; 5(2): 44-5,2013.

[6]C.CLopes,H..S.Freitas,A.ABojarczuk."Data mining with Constrained –Syntax Genetic Programming: Applications in Medical Dataset, "in Data Analysis in Medicine and Pharmacology(IDAMAP-2001),a Workshop at Medinfor-2001,London,UK

[7]Dinesh K.Sharma ,Hota H.S. ,"Data Mining techniques for prediction of different categories of dermatology diseases", Academy of information and management science journal , 16 (2), 103-115, 2013.

[8]E.D Ubeyli,I.Guler,"Automatic detection of erythemato squamous diseases using adaptive neuro-fuzzy inference system",Computer in Biology and medicine,35, 421-433,2005
[9]E. D. Ubeyli. Multiclass support vector machines for diagnosis of erythemato-squamous
Diseases. Expert Systems with Applications, 35,1733-1740, 2008.
[10]E.D Ubeyli."Combined neural network for diagnosis of erythemato squamous diseases", Expert system with Applications,36(3),5107-5112,2009
[11]Güvenir, H., Demiröz, G., &Ilter, N. "Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals". Artificial Intelligence in Medicine, 13(3) 147-165,1998.
[12]Guvenir, H., &Emeksiz, N. "An expert system for the differential diagnosis of erythemato-squamous diseases.",Expert Systems with Applications, 18(1) 43–49,2000.
[13]Kalik.B,Harman.G."Computer-aided software for early diagnosis of erythemato-squamous "International Scientific Conference Electronics &Nanotechnology(ELNANO),2013 Conference Proceeding
[14]Kemal Polat and SalihGunes. The e_ect to diagnostic accuracy of decision tree classifier of fuzzy and k-nn based weighted pre-processing methods to diagnosis of erythemato-squamousdiseases. Digital Signal Processing, 16,922-930, 2006
[15]Kemal Polat ,SalihGunes. "A novel hybrid intelligent method based on C4.5 decision tree classifier and one against-all approach for multi-class classification problems". Expert Systems with Applications, 36(2) 1587-1592,2009
[16]Krzysztof Cios,Witoldpedryeyz,RomanSwiniarski."Data Mining Methods for Knowledge Discovery",Kluwar Academic Publishers,1998
[17]Mohammad javedAbdi,DaverGiveki,"automatic detection of erythemato squamous diseases using PSO-SVM based on association rule", Engineering Applications of Artificial Intelligence, 26(1),pp603-608,January 2013.
[18]Murat karabatak,M.CevdetInee."An expert system for detection of breast cancer on association rules and neural network "Expert system with Application, volume 36(2),part2,pp3465-3469,March2009
[19]N.Satyanandam, Dr. Ch. Satyanarayana, Md.Riyazuddin, Amjan.Shaik."Data mining machine learning approaches and medical diagnosis system: A survey", International journal of computer and organization trends, (3),53-60,2012
[20]Nanni, L. "An ensemble of classifiers for the diagnosis of erythemato-squamous diseases.",Neurocomputing,69, 842–845,2006
[21]P. Luukka and T. Lepp • alampi. Similarity classifier with generalized mean applied to medical
data. Computers in Biology and Medicine, 36,1026-1040, 2006
[22]P. Luukka. Similarity classifier using similarity measure derived from yu's norms in classifcation of medical data sets. Computers in Biology and Medicine, 37,1133-1140, 2007.
[23]P.Luukka."A New Nonlinear Fuzzy Robust PCA Algorithm and Similarity Classifier in Classification of Medical Data Sets", International journal of Fuzzy system,13(3),September 2011
[24]S.Rajeasekaran,G.A.VijayalakshmiPai."Neural Network, Fuzzy logic and Genetic algorithm synthesis and applications", PHI learning private limited,2003
[25]Stavros Lekkas,Ludmil Mikhailov,"Evolving fuzzy medical diagnosis of Pima Indian Diabetes and of dermatological diseases."Artificial Intelligence in Medicine, 50, pp117-126, 2010
[26]J.Xie,Ch.Wang ."Using support vector machine with a novel hybrid feature selection method for diagnosis of erythemato squamous diseases." Expert System with Applications, 38(5), pp5809-5815,2011

ABOUT AUTHORS

Madhura Rambhajani is a M-tech scholar in Department of Information technology, Technocrats Institute of Technology, Bhopal, India. She received B.E in 2012 from Guru Ghasidas Vishwavidialaya; Bilaspur (C.G).She has published one paper in international conference paper and one in international conference.

Wyomesh Deepanker currently working as Assistant Professor in Technocrats Institute of Technology, Bhopal, India. His area of specializations soft computing, Data Mining, Bio inspired Computing, Image Processing.

Neelam Pathak currently working as Assistant Professor in Technocrats Institute of Technology, Bhopal, India.Her area of Specialization is Computer network and soft computing.