Editor-in-Chief   Dr. Bekir Karlik

INTERNATIONAL JOURNAL OF
ARTIFICIAL INTELLIGENCE AND EXPERT SYSTEMS (IJAE)

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE AND EXPERT SYSTEMS (IJAE)

VOLUME 5, ISSUE 2, 2014

EDITED BY
DR. NABEEL TAHIR

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE AND EXPERT SYSTEMS (IJAE)

CSC Publishers, 2014

EDITORIAL PREFACE

The International Journal of Artificial Intelligence and Expert Systems (IJAE) is an effective medium for interchange of high quality theoretical and applied research in Artificial Intelligence and Expert Systems domain from theoretical research to application development. This is the *Second* Issue of Volume *Five* of IJAE. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJAE emphasizes on efficient and effective Artificial Intelligence, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of Expert Systems. IJAE comprehensively cover the system, processing and application aspects of Artificial Intelligence. Some of the important topics are AI for Service Engineering and Automated Reasoning, Evolutionary and Swarm Algorithms and Expert System Development Stages, Fuzzy Sets and logic and Knowledge-Based Systems, Problem solving Methods Self-Healing and Autonomous Systems etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 5, 2014, IJAE appears with more focused issues related to artificial intelligence and expert system research. Besides normal publications, IJAE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJAE give an opportunity to scientists, researchers, and vendors from different disciplines of Artificial Intelligence to share the ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development. This journal is helpful for the researchers and R&D engineers, scientists all those persons who are involve in Artificial Intelligence and Expert Systems in any shape.

Highly professional scholars give their efforts, valuable time, expertise and motivation to IJAE as Editorial board members. All submissions are evaluated by the International Editorial Board. The International Editorial Board ensures that significant developments in image processing from around the world are reflected in the IJAE publications.

IJAE editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJAE. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJAE provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Artificial Intelligence and Expert Systems (IJAE)

Dr. Alex James
Indian Institute of Information Technology and Management- Kerala
India

Assistant Professor Dr Zubair Baig
King Fahd University
Saudi Arabia

Associate Professor Syed Saad Azhar Ali
Iqra University
Pakistan

Assistant Professor Israel Gonzalez-Carrasco
Universidad Carlos III de Madrid
Spain

Professor Sadiq M. Sait
King Fahd University
Saudi Arabia

Professor Hisham Al-Rawi
University of Bahrain
Bahrain

Dr. Syed Zafar Shazli
Northeastern University
United States of America

Associate Professor Kamran Arshad
University of Greenwich
United Kingdom

Associate Professor, Mashtalir Sergii
Kharkiv National University of Radio Electronics
Ukraine

S.Bhuvaneswari
Pondicherry University
India

Dr Alejandro Rodriguez Gonzalez
Polytechnic University of Madrid
Spain

Assistant Professor, Jose Luis Lopez-Cuadrado
Universidad Carlos III de Madrid
Spain

Assistant Professor, Ilhan AYDIN
Firat University
Turkey

Associate Professor, Afaq Ahmed
Sultan Qaboos University
Oman

Dr. Muhammad Ali Imran
University of Surrey
United Kingdom

TABLE OF CONTENTS

Volume 5, Issue 2, March 2014

Pages

Performance Evaluation of Neural Classifiers Through Confusion Matrices To Diagnose Skin Conditions

İhsan Ömür BUCAK   *iobucak@meliksah.edu.tr*
Faculty of Engineering and Architecture
Department of Electrical-Electronics Engineering
Melikşah University
Talas-Kayseri, 38280, Turkiye

Abstract

In this paper we have aimed to diagnose skin conditions using Artificial Intelligence (AI) based classifier algorithms and do the performance analyses of those presented algorithms through confusion matrices. These algorithms are being used in a large array of different areas including medicine, and display very distinct characteristics in the sense that they are grouped under different categories such as supervised, unsupervised, statistical, or optimization. The objective of this study is to diagnose skin conditions using seven different well-known and popular as well as emerging Artificial Intelligence based algorithms and to help general practitioners and/or dermatologists develop a careful and supportive approach that leads to a probable diagnosis of skin conditions or diseases. These algorithms we chose as neural classifiers include Back-Propagation (BP), Random Forest (RF), Support Vector Machines (SVMs), Linear Vector Quantization (LVQ), Self-Organizing Maps (SOMs), Naïve Bayes, and finally Bayesian Networks. All of these algorithms have been tested and their results of diagnosing skin conditions/diseases by using data set from Dermatology Database have been compared.

Keywords: Artificial Intelligence, Artificial Neural Networks, Medical Diagnosing, Neural Classifiers, Skin Conditions/Diseases, Confusion Matrix, F-score.

1. INTRODUCTION

Diagnosing skin conditions and diseases at an early stage is an extremely important in controlling the increasing global numbers of skin cancer cases each year at an alarming rate. The numbers are devastating worldwide and have doubled in the last 15 years [1]. Globally, every third case diagnosed as a cancer belongs to a skin cancer and, in the USA alone; every fifth American is expected to catch the skin cancer during their life time [1, 2].

If, particularly, the symptoms are not treated at their early stages, they mostly become very dangerous. Furthermore, many symptoms look alike even though they are caused by different diseases. This may even bring about more challenges for dermatologists to make correct diagnosis. Erythema, scaling, definite borders, and size or density of the skin changes are very important clinical features. Modern medicine along with the new technologies comes to dermatologists' rescue in the correct diagnosis of skin conditions in two ways: either from its symptoms or through diagnostic examination. The diagnostic examination will include a range of tests such as skin biopsy, blood test, X-rays, MRI/CT scans, allergy testing, and swab from a sensitive area as well as a physical examination to diagnose the skin condition. Most of these tests are used as standard diagnostic tools to help determine the causes of any skin condition and have been proven to be very effective at diagnosing the skin condition. Skin conditions may also have a genetic connection as the cause of any particular skin condition. Therefore, it is important to share any past history of skin problems, family history with dermatologist at the time of examination as many skin conditions are inherited from family.

A systematic approach to diagnosing skin conditions has also been developed. Five point of descriptive paradigm can be applied to both rashes and lesions. These are distribution, demarcation, color, surface change, and morphological features.

The objective of this paper is to diagnose skin conditions by using Artificial Intelligence based (ANN) seven algorithms and compare their results against training and testing performances. These algorithms, also known as neural classifiers, should be able to learn from the skin data set of very large dermatology database [3], and return the probability of existence of any diagnosed disease by helping dermatologists to improve the diagnosis of dermatological disease.

2. METHODS AND ALGORITHMS

In this paper, neural classifiers are based on the field of machine learning. Machine learning is one of the six disciplines within Artificial Intelligence (AI) and make up most of AI along with natural language processing, knowledge representation, automated reasoning, computer vision, and robotics [4]. Machine learning was first defined by A. Samuel in 1950 as a field that gives computers a specific capability of learning with no explicit programming aid. Russell and Norvig atributes an ability to machine learning for adapting itself to new conditions and for detecting and extrapolating patterns [4]. The field of machine learning usually distinguishes three types of feedback to determine the nature of the learning problem agent faces: supervised, unsupervised, and reinforcement learning [4]. In this paper, most of the algorithms under consideration lie on the first two categories of these three cases. One of the objectives in this study is to compare the supervised and unsupervised learning algorithms.

2.1 Backpropagation Algorithm

Artificial Neural Networks (ANNs) are the most effective approaches for pattern recognition [5-9]. ANNs cannot succeed a medical doctor, but they can help the doctor in identification of problem or diagnosis of the disease [10]. Backpropagation (a.k.a propagation of error) is a familiar method used to teach Artificial Neural Networks upon how to accomplish a given task. Bryson and Ho described it first in 1969 [11], but the work of Rumelhart and his colleagues led it to gain recognition in 1986, and the algorithm started thereafter to evolve in the research area of artificial neural networks [12].

Backpropagation algorithm trains the artificial neural network by presenting a set of exemplary input/output pairs to the network. The algorithm adjusts the internal weights of the network iteratively to minimize the mean squared error between the network's predicted outcome and the true result as the set of different training examples in the form input/output pairs are introduced to the system. The weight adjustment will yield an average response since the error is acquired over the entire training set. For the network to be able to accurately categorize any input set, it is significant that the training set contain examples that represent the full range of possible inputs. The new input pattern is expected to be accurately categorized by the network even if the new input sets and any of the training patterns do not seem alike.

Backpropagation is a supervised learning method whose objective is to find a mapping function that establishes the best mathematical relation between a set of inputs and its correct output. Single layer neural networks also known as perceptrons can easily learn some input and output patterns. However, these single layer perceptrons are incapable of learning some linearly separable patterns. A multi-layered network can overcome its limitation of having only one single layer by adding more layers (effectively more internal layers) to learn different abstract features in each layer by finding patterns in the layer below it. Therefore, the ability to find patterns between the current and prior layers successively will create internal representations. The motivation behind the backpropagation lies in the methodology to train multi-layered neural networks in such a way as to make it learn the convenient internal representations for its granted learning of any random mapping between input and output. It needs a training set created by a dataset of the desired output in consideration of many inputs. It is most beneficial for feed-forward type

networks. It also requires a differentiable activation function for its artificial neurons. Fundamentally, the steps of the backpropagation algorithm can be summarized as follows [13]:

1. Introduce a training pattern to the network.
2. Compare the network's output with the desired one and calculate the error for each output neuron.
3. Calculate each neuron's output with its local error.
4. Adjust each neuron's weight to minimize the local error.
5. Assign "blame" to neurons for the local error at the former level.
6. Repeat the process starting with step 3 using each neuron's "blame" as its error on the ones at the former level.

2.2 Linear Vector Quantization (LVQ)

LVQ was developed by Kohonen in 1984 [14]. Its basic tenet is to represent an n-dimensional vector by lower dimensional vectors. Learning in LVQ is to determine which output vector will match up with an input vector. Vector set in the output layer is called reference vector. The objective is to match up the input vector with the reference vectors. Generally, it is used in the solution of classification problems. Only one of the output vectors takes the value of 1 and the others take 0. The reference vector with the value of 1 determines the class of input vector. The distance between the inputs and the outputs is found by Euclidian method. It runs according to "winner takes all" principle, and not the all weights but the weights of the winning reference vector only is updated in such a way that it gets closer to the input vector if both belong to the same class, and is moved away from it if they belong to different classes [15, 16].

LVQ is a supervised classifier and is used to adjust the bounds between classes to maximize classification. It has three layers like multi-layer nets. There is no information processing in the input layer. However, data entry to LVQ net is performed through the input layer. Distance that will provide the closest vector to the input vector is computed in hidden or Kohonen layer (a.k.a. intermediate layer). Each weight vector between the input and the intermediate layer is called reference vector. Assigning or classifying the output layer to one of the reference vectors is fulfilled in output layer. The weights between the hidden and the output layers are constants and their value are equal to 1. An LVQ net includes a single layer of nodes where each node shows a class. The weight adjustment procedure on this single layer's nodes is applied to adapt each node to the input vectors occurring with the corresponding class. LVQ is trained as based on the following steps:

1. Assign very small numbers between 0 and 1 to weight vectors initially.
2. Pick up a training vector.
3. Discover the class (node) of minimum distance with this training vector.
4. Update the node's weights in accordance with the following expression:
   If the target and the output class are the same,

$$\underline{w}_{j_{new}} = \underline{w}_{j_{old}} + \alpha(\underline{x} - \underline{w}_{j_{old}}) \tag{1}$$

else

$$\underline{w}_{j_{new}} = \underline{w}_{j_{old}} - \alpha(\underline{x} - \underline{w}_{j_{old}}) . \tag{2}$$

5. Reduce learning rate α.
6. Go to step 2 and repeat until the termination criterion is met.

LVQ usually performs efficiently if the initial weight vectors are close to their final values. That is why some heuristic procedures are used for weight initialization, like Kohonen's SOM or any other vector quantization procedure, like *k-means*.

2.3 Self Organizing Map (SOM)

SOM is an unsupervised learning type of artificial neural network that transforms its input space of training samples with arbitrary dimension into a low-dimensional discrete map by preserving topological properties during its training session [17, 18]. Self-organizing maps, not the other artificial neural networks, use only a neighborhood function in order not to lose the topological properties of the input space [18]. It was first introduced by Teuvo Kohonen, and therefore is also known as Kohonen map [19]. It accomplishes a competitive learning. Euclidian distances of all weight vectors to the training samples are computed for a given training sample to the network. Best identical neuron with the input is called the best matching unit (BMU). The BMU weights and the neurons in its instantaneous neighborhood are adjusted in connection with the input vector. The amount of the change is expected to become smaller over time and with distance from the BMU. A neuron with weight vector $W_v(t)$ is updated by the following equation:

$$W_v(t+1) = W_v(t) + \Theta(v,t)\alpha(t)(D(t) - W_v(t)),$$ (3)

where α(t) is a monotonically decreasing learning parameter, Θ(v, t) is the neighborhood function that depends on the distance between the BMU and neuron v, and $D(t)$ is the input vector. The neighborhood function generally utilizes a Gaussian function. Nevertheless, it can simply be used as one for all neurons in the BMU neighborhood and zero otherwise. The neighborhood function concurs with time without regard to the functional form [20]. When the neighborhood is large initially, the self-organizing takes place in the overall scope. When the neighborhood is shrunken to just a couple of neurons the weights converge to their local estimates.

The neuron whose weight vector stands closest to the input vector at the end of mapping is called *winning* neuron. This is clearly resolved by computing the Euclidean distance between input and weight vectors.

2.4. Support Vector Machines (SVMs)

Another methodology used for classification purposes is known to be Support Vector Machines [21, 22]. In this method, the classification is aimed to be done through the aid of a linear or a nonlinear function. This method is based on the estimation of the most suitable function to separate the data from each other. A hyperplane is a concept in geometry. It is a generalization of the plane into a different number of dimensions. A H_0 hyperplane in terms of the points on itself can be expressed as

$$H_0 : W^T X + b = 0.$$ (4)

If the space between the two sides of H_0 hyperplane, say, H_1 and H_2, is required to be maximized, the following objective function needs to solved:

$$\text{Minimize} \quad \frac{w^T w}{2}$$ (5)

constrained by $y_i(w^T x_i + b) \geq 1$ $(i = 1,2..n)$ where W represents the weight vector, X represents the training set, y represents the classes, b is a constant number, and finally, n the number of attributes. The problem of the nonseperability of the data can be resolved through the addition of nonnegative and error-indicative ξ_i loose variables into the optimization model. The cases where the data can not be separated in linear or linear classifiers are not complex enough sometimes, then the nonlinear classifiers can be used instead of the linear ones to map the data

into a richer feature space including nonlinear ones by constructing a hyperplane in that space. In these cases, ϕ functions can be utilized to analyze by carrying the data to the spaces with larger dimensions. ϕ functions are used for choosing a good mapping which turns out to be the improved results for the problem. Encoding prior knowledge and getting the right complexity of function class will serve the understanding of choosing the good mapping [22].

2.5. Random Forests (RFs)

Random forest is a collection of CART-like decision trees that follows certain rules. There is no pruning in RF as opposed to classical decision tree since the ensemble and bootstrapping schemes help random forest overcome overfitting issues [23-25]. Each tree is constructed using the following algorithm:

Step 1. Form a training set by N sample cases for this tree in which the number of training sample cases is chosen at random but with replacement from original data (i.e., take a bootstrap sample). This sample will be the training set for growing the tree.

Step 2. Specify a number m of the M input variables to determine the decision at a node of the tree. The number m should be less than the number M such that at each node, m variables are selected at random out of the M. Calculate the best split based on these m variables in the training set to split the node. The value of m is held constant during the forest growing.

Step 3. Each tree is fully grown to the largest extent possible and not pruned.
Forest error rate is dependent on two things: *correlation* and *strength*. Correlation is directly proportional to the forest error rate. Increasing the correlation between any two trees in the forest will increase the forest error rate. Strength is attributed to each individual tree in the classifier.

Strength of the individual trees is inversely proportional to the forest error rate. That is, increasing the strength of the individual trees will decrease the forest error rate

The value of m affects both the correlation and the strength. Therefore, reducing it reduces both and increasing it increases both. m lies on somewhere in between as its optimal range. A value of m in the range can quickly be found by the oob (out-of-bag) error estimation rate. Random forests is somewhat sensitive to this adjustable parameter only.

2.6 Naïve Bayes Classifier

A Naïve Bayes classifier is a simple probabilistic feature model based on an application of Bayes' theorem along with an assumption of strong independence on a feature of a class. Here, the more explanatory term for the probabilistic model is "independent feature model" and the prefix *naïve* is driven by the word "strong". Independence means that all the features of a class contribute independently to the probability of a particular object that is classified. Simply, the presence of one feature of the class has no relationship, on no account whatsoever, with the presence of any other feature, given the class variable. Although it is simplified considerably by its assumptions and naïve design, these types of classifiers have been proven to perform reasonably better in many complicated real-world environments [26-28].

Naïve Bayes models use the maximum likelihood method in many practical applications for estimating the particular values of the mean and variance as the model parameters to make the observed results the most likely, given the model. Generally, the model parameters are picked out by the maximum likelihood method to maximize the likelihood function between the model picked and the data observed [27, 28].

Naïve Bayes has the advantage of requiring only a small percent of training data to estimate the parameters when it is used as classifier, and is inclined to train very efficiently for some types of probability models in a supervised learning paradigm. Additionally, each class requires only the

calculation of the variances of the variables, not the entire covariance matrix since the variables are assumed independent [27, 28].

2.7 Bayes Nets

Bayes Nets or Bayesian networks are graphical representation for probabilistic relationships among a set of random variables. In contrast to undirected graphical models, which provide a simple definition of independence, Bayesian networks, also called directed graphical models, presents a more complicated concept of independence. However, their advantages are evident and undeniable ones. Bayes nets can be used for causal, or "top down", reasoning. For example, an arc from A to B can be treated causal as it is taking into consideration the directionality of the arcs by indicating that A "causes" B. This can ease the way on how the graph is built [29]. Furthermore, Bayes nets can also be used to encode the deterministic relationships, and are easier to learn as the model fits the data well [30, 31].

3. EXPERIMENTAL ANALYSIS

The differential diagnosis, which is a medical term defined as a systematic method of diagnosing a disorder that lacks unique symptoms or signs [32], creates a substantial problem in dermatology in diagnosing skin disease (erythemato-squamous). In this type of disease, the clinical features of erythema and scaling are generally overlapped. The severity of erythema and scaling, the presence of itching and koebner phenomenon, the formation of papules matter all critically on every account. A biopsy is usually applied as the last option to explicitly diagnose the disease of the clinical features detailed in [33]. Even these types of diseases can harbor many histopathological features. Indicating the features of another disease initially, having the characteristic features thereafter is a common cause that complicates the diagnosis [33]. Additionally, some skin samples may not show the typical histopathological features of the disease [33]. In this work, the same dataset for classes and features (both clinical and histopathological) used in [33] have also been employed.

The networks have been trained by using the training data set with 33 features. Furthermore, 40 inputs have been used for training and 20 inputs for testing for each class. 66.66 percent of the entire data of 360 has been split up for training and 33.33 percent for testing. The input layer has been composed of 33 nodes; each was a representative of 33 features from the data entry form. The output layer has been composed of 6 nodes; each was a representative different class of diagnostic diseases such as C_1: Psoriasis, C_2: Seborrheic dermatitis, C_3: Lichen planus, C_4: Pityriasis rosea, C_5: Chronic dermatitis, and C_6: Pityriasis rubra pilaris [33].

4. PERFORMANCE EVALUATION BY CONFUSION MATRICES

In this paper, the objective was initially driven to compare supervised algorithms with unsupervised ones for the given dataset above. Therefore, BP and LVQ were chosen randomly to represent the supervised algorithms and SOM the unsupervised one. The error rates shown in Figure 1 were discussed in detail in [34]. In summary, SOM yielded the highest error rate and failed to separate one or two classes for some lower iteration numbers. Furthermore, LVQ learned the earliest due possibly to the most appropriate weight vector initialization; hence it achieved the maximum success rate the earliest, too. Although LVQ achieved a relatively fast learning, BP generated the minimum error rate. The results have been obtained by using MATLAB [35].
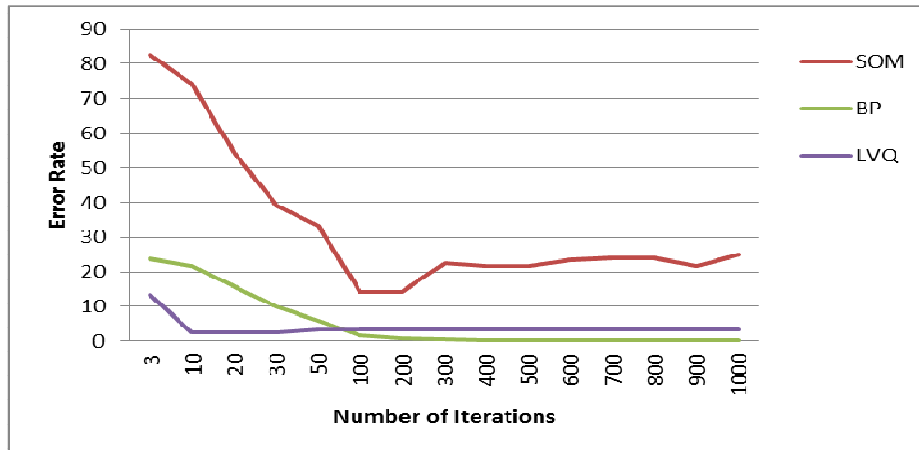
FIGURE 1: Error rates for BP, LVQ, and SOM algorithms [34].

Figure 2 shows the time performance of the supervised and unsupervised algorithms [34].
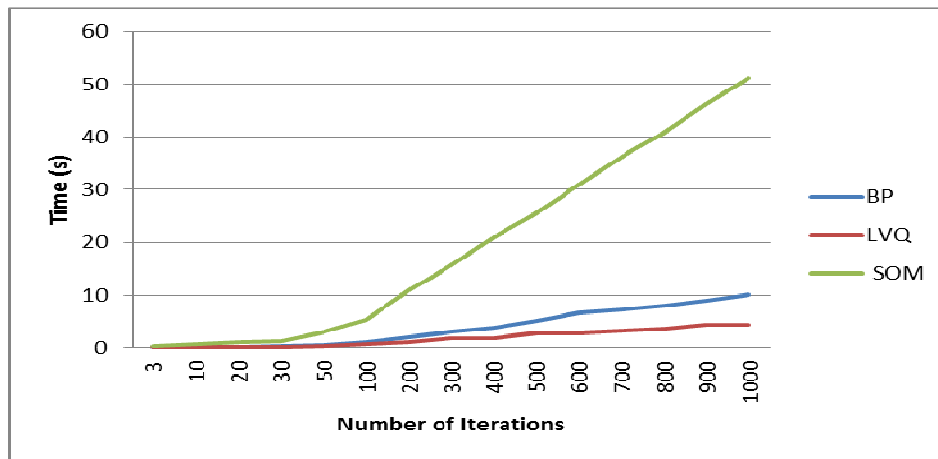

FIGURE 2: Time performance of BP, LVQ, and SOM algorithms [34].

However, a more comprehensive performance evaluation which also covers the emerging algorithms as presented in this paper has been a prime motive to extend the work beyond these three. Therefore, the need to extend the performance evaluation has been satisfied greatly by the aid of confusion matrices. A confusion matrix is a specific table used to evaluate the performance of a supervised or an unsupervised algorithm. Each column and row of the matrix represents the instances in a predicted and an actual class, respectively. Since it is easy to see if the system is confusing two classes, most possibly, in the form of mislabeling one as another, therefore the name is given and used as the confusion matrix. The confusion matrix is often called the contingency table or the error matrix outside the artificial intelligence community. The confusion matrices below have been obtained for the algorithms, Naïve Bayes, Backpropagation, Random forest, Support vector machines, Linear vector quantization, Self-organizing maps, and finally Bayes theorem to diagnose the skin diseases.

NaïveBayes

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
0	18	0	0	0	2	**2**
0	0	20	0	0	0	**3**
0	0	0	20	0	0	**4**
0	0	0	0	20	0	**5**
0	0	0	0	0	20	**6**

BP

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
0	18	0	2	0	0	**2**
0	0	20	0	0	0	**3**
0	0	0	20	0	0	**4**
0	0	0	1	19	0	**5**
0	0	0	0	0	20	**6**

RF

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
1	17	0	2	0	0	**2**
0	0	20	0	0	0	**3**
0	0	0	20	0	0	**4**
0	0	0	0	20	0	**5**
0	0	0	0	0	20	**6**

SVM

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
0	18	0	2	0	0	**2**
0	0	20	0	0	0	**3**
0	0	0	20	0	0	**4**
0	0	0	0	20	0	**5**
0	0	0	0	0	20	**6**

LVQ

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
0	17	0	3	0	0	**2**
0	0	20	0	0	0	**3**
0	1	0	19	0	0	**4**
0	0	0	0	20	0	**5**
0	0	0	0	0	20	**6**

SOM

1	2	3	4	5	6	
0	0	0	0	0	20	**1**
0	8	0	12	0	0	**2**
20	0	0	0	0	0	**3**
0	19	0	1	0	0	**4**
0	0	20	0	0	0	**5**
0	0	0	0	20	0	**6**

BayesNET

1	2	3	4	5	6	
20	0	0	0	0	0	**1**
0	19	0	1	0	0	**2**
0	0	20	0	0	0	**3**
0	0	0	20	0	0	**4**
0	0	0	0	20	0	**5**
0	0	0	0	0	20	**6**

FIGURE 3: Confusion Matrices of the Neural Classifiers.

The basis of our problem underlies the classification of skin diseases which shows the close symptoms to each other. Once the confusion matrices are studied in detail, it can be observed by the employed methods that the diseases the 1st, 3rd, 5th, and the 6th can disassociate themselves successfully. On the other hand, misclassification over the entire classes has occurred in the 2nd and the 4th classes, i.e., in the Seborrheic dermatitis, a chronic skin disease characterized by the inflammation of the scalp or the face along with the appearance of scaly and flaky substance over the surface of the skin, and Pityriasis rosea, a harmless, common skin condition that causes a rash, although the methods differ greatly from each other. This can particularly be seen prominently in the LVQ and SOM algorithms. Hence, the errors arising can be interpreted as the result of these two diseases that gives more close symptoms as compared to the others, rather than the methods used.

5. COMPARATIVE EVALUATION

As mentioned previously, the objective of this study aims to classify the skin conditions and diseases which, in particular, shows the close symptoms to each other. Furthermore, we have deliberated on diagnosing skin conditions and diseases that needed to be treated at their early stages before getting out of control towards skin cancer.

Any comparative evaluation to the similar cases can only be meaningful. However, it is rather difficult to observe any case study similar to the ones used in our research. In general, a huge variety of applied methods and differing skin conditions and diseases as the application area can complicate further the success of such comparative evaluation. We are confident that, in terms of Artificial Intelligence and Machine Learning perspective, our algorithms not only cover the conventional ones proven to be successful and reliable when applied to several medical diagnostic tasks, but also include some recently developed methods that seemed promising for overall medical diagnosis and treatment [36-39]. For example, we have obtained the best result when we have applied the Bayesian network classifier. In addition, naïve Bayes has yielded very good result, too. Our conclusion has also been drawn by some other experiments and studies in the literature [36-39]. Therefore, it shows that physicians prefer explanations as based on the Bayesian classifiers.

6. CONCLUSION

ANNs have been proven to be very efficient tools for pattern recognition and they can be successfully used in dermatological applications. They can increase the detectability of dangerous diseases and lower the mortality rate of patients. Both supervised and unsupervised based neural network algorithms can readily be used for implementing an application similar to the considered one here, and utilized in the field of medicine to determine the skin diseases. Medical doctor has to decide if there is a need to analyze some part of the skin. Therefore, it is up to the medical doctor what to do with the results after running the program. The outcome obtained from the program can be very helpful. Although computers cannot replace the dermatologist, they can make their work easier and more effective. Proposed study might be also very useful for general practitioners, who do not have wide knowledge about dermatology.
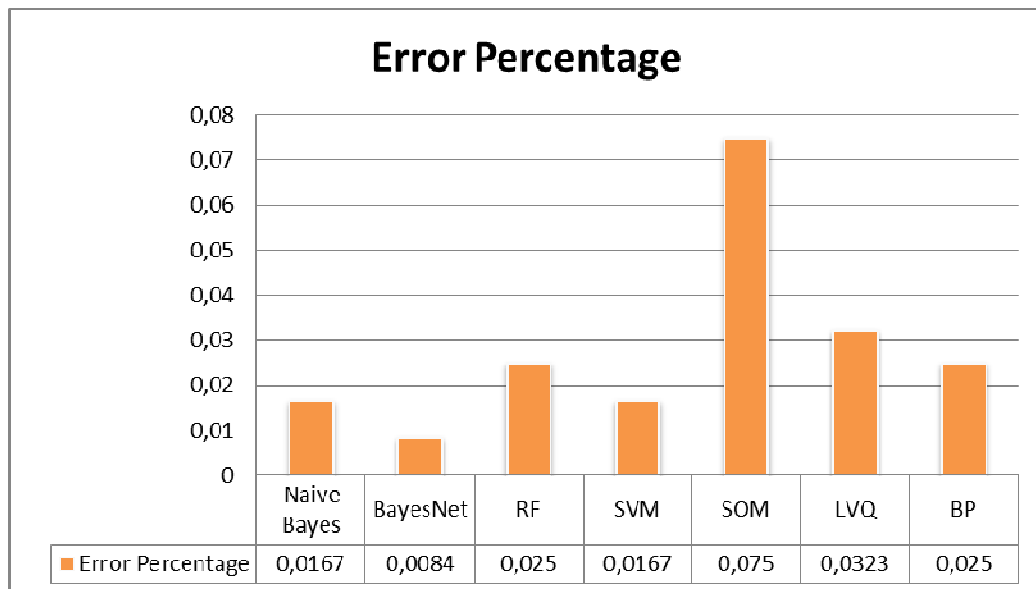

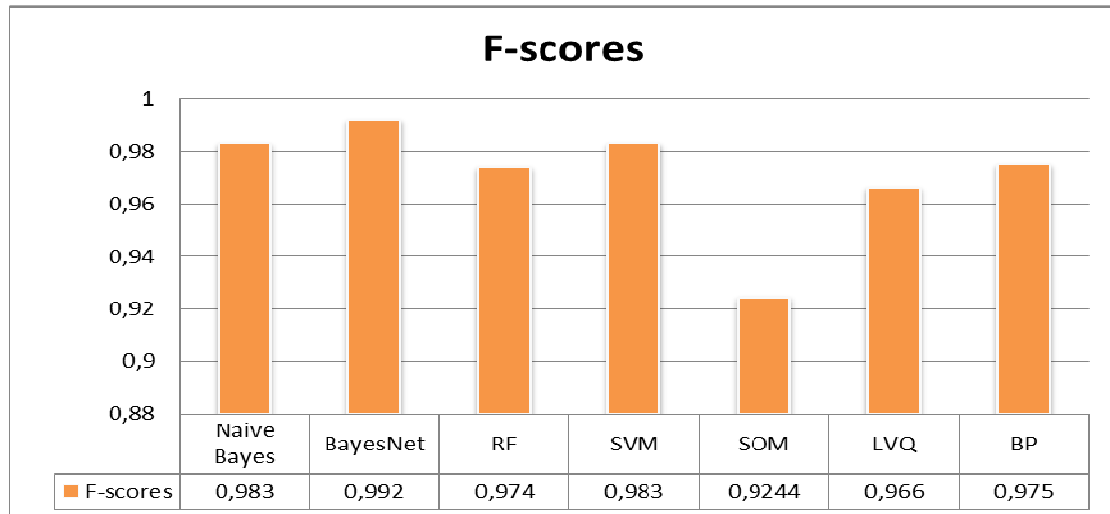**FIGURE 4:** Error Percentages of the Neural Classifiers.

FIGURE 5: F-scores of the Neural Classifiers.

The results have been summarized in the form of error percentages and F-scores as can be seen from Figures 4 and 5. All the methodologies considered in this paper have produced satisfactory results except for the SOM algorithm that has generated a relatively high percentage of error (Please see Figures 4 and 5). Despite its lower success rate in this study, SOM is successful to make or draw a distinction amongst classes, i.e., it is capable of classifying a difficult dataset as an unsupervised learning method. Other experiments and/or studies in the literature tend to encourage us to use this method for a wide range of analyses, even from such domain as medical data. The best result has been obtained from Bayesian network classifier since it makes its classification with an assumption of dependency among the features of its own data. One can easily conclude that there exists correlations among the symptoms of a disease. Therefore, the results obtained seem to be consistent if one considers the area of application. However, naive Bayes has surprisingly produced a considerably good result in this application although, ironically, it assumes no dependency among the features. Generally, ANNs which shows a lower performance rate produces a higher accuracy rate once compared to SVMs. However, in this study, SVM has produced a better result although the difference between them has remained small; RF has produced very close result to BP is the another observation.

7. REFERENCES

[1] M. Antkowiak, Artificial Neural Networks vs. Support Vector Machines for Skin Diseases Recognition, Msc. Thesis, Umea University, Sweden, 2006.

[2] http://www.skincancer.org/Skin-Cancer-Facts/, accessed May 27, 2010.

[3] A. Frank and A. Asuncion, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science, 2010.

[4] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", 3rd edition, Prentice Hall, p.578, 2009.

[5] C.M. Bishop, "Neural networks and their applications", Review of Scientific Instruments, 65(6), pp. 1803–1832, 1994.

[6] C.M. Bishop, "Neural Networks for Pattern Recognition", Clarendon Press, Oxford, UK, 1995.

[7] B.D. Ripley, "Pattern Recognition and Neural Networks", Cambridge University Press, Cambridge, UK, 1996.

[8] C. Stergiou and D. Siganos, "Neural networks", SURPRISE 96 Journal, 4, 1996.

[9] R. Tadeusiewicz, "Sieci neuronowe, (Neural Networks)", Akademicka Oficyna Wydawnicza, Warszawa, 1993.

[10] M. Antkowiak, "Recognition of skin diseases using artificial neural networks", In Proceeding of USCCS'05, pp. 313–325, 2005.

[11] A.E. Bryson and Y. Ho, "Applied optimal control: optimization, estimation, and control", Blaisdell Publishing Company or Xerox College Publishing, pp. 481, 1969.

[12] D.E. Rumelhart, G.E. Hinton, and R.I. Williams, "Learning representations by back-propagating errors", Nature, 323, pp. 533-536, 1986.

[13] E. Sivasankar, R.S. Rajesh, and S.R.Venkateswaran, "Diagnosing Appendicitis Using Backpropagation Neural Network and Bayesian Based Classifier", International Journal of Computer Theory and Engineering, Vol. 1, No. 4, pp. 1793-8201, 2009.

[14] T. Kohonen, "Self-Organizing Maps", Springer, Berlin, 1997.

[15] T. Kohonen, G. Barna, and R. Chrisley, "Statistical pattern recognition with neural networks: Benchmarking studies", In Proceedings of the International Conference on Neural Networks (ICNN), Vol. I, pp. 61-68, 1988.

[16] T. Kohonen, Learning vector quantization. In: M.A. Arbib, editor, The Handbook of Brain Theory and Neural Networks, pages 537–540. MIT Press, Cambridge, MA, 1995.

[17] http://www.nd.com/models/sofm.htm, Self-Organizing Feature Maps, accessed April 5, 2011.

[18] http://en.wikipedia.org/wiki/Self-organizing_map, Self-organizing map, accessed April 5, 2011.

[19] T. Kohonen and T. Honkela, "Kohonen network", Scholarpedia 2007, http://www.scholarpedia.org/article/Kohonen_network, accessed May 25, 2010.

[20] S. Haykin, "9. Self-organizing maps", Neural networks - A comprehensive foundation (2nd ed.), Prentice-Hall, 1999.

[21] N. Cristianini, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods", Cambridge University Press, 2000.

[22] B. Schlkopf and A.J. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond", MIT Press, 2001.

[23] L. Breiman, "Random Forests". Machine Learning 45 (1): 5–32, 2001.

[24] T.K. Ho, "Random Decision Forest". Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282, 1995.

[25] T.K. Ho, "The Random Subspace Method for Constructing Decision Forests". IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (8): 832–844, 1998.

[26] M. Kantardzic. Data Mining - Concepts, Models, Methods, and Algorithms, IEEE Press, Wiley-Interscience, 2003.

[27] S. Theodoridis and K. Koutroumbas. Pattern Recognition. San Diego: Academic Press, 2006.

[28] T.M. Mitchell. Machine Learning. McGraw-Hill, 1997.

[29] J. Pearl, Causality: Models, Reasoning, and Inference, Cambridge University Press, 2000.

[30] R.E. Neapolitan. Learning Bayesian networks. Prentice Hall, 2004.

[31] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using Bayesian Networks to Analyze Expression Data", Journal of Computational Biology (Larchmont, New York: Mary Ann Liebert, Inc.) 7 (3/4): 601–620, 2000.

[32] http://www.thefreedictionary.com/differential+diagnosis, accessed October 7, 2013.

[33] G. Demiröz, H. A. Güvenir, and N. İlter, "Learning differential diagnosis of eryhemato-squamous diseases using voting feature intervals", Artificial Intelligence in Medicine, 13, pp. 147-165, 1998.

[34] İ. Ö. Bucak, Y. Sahin, and A. Serbetci, "Supervised vs. Unsupervised neural classifiers for diagnosing skin diseases", The Proceedings of Second International Conference on Informatics (ICI'2011), Çanakkale-Turkiye, pp. 73-80, 27-29 April 2011.

[35] MATLAB R2008b, Natick, Massachusetts: The MathWorks Inc., 2008.

[36] K.K Manjusha, K. Sankaranarayanan, and P. Seena, "Prediction of Different Dermatological Conditions Using naïve Bayesian Classification", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, pp. 864-868, 2014.

[37] R. Blanco, I. Inza, M. Merino, J. Quiroga, and P. Larranaga, "Feature selection in Bayesian classifiers for the prognosis of survival of cirrhotic patients treated with TIPS", Journal of Biomedical Informatics, Volume 38, Issue 5, pp. 376-388, 2005.

[38] S. Shirali-Shahreza, M.E. Mousavi, "A New Bayesian Classifier for Skin Detection", 3rd International Conference on Innovation Computing and Control, pp. 172-175, 2008.

[39] A. Masood, A.A. Al-Jumaily, "Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms", International Journal of Biomedical Imaging, Volume 2013, 22 pages, 2013.

INSTRUCTIONS TO CONTRIBUTORS

The main aim of International Journal of Artificial Intelligence and Expert Systems (IJAE) is to provide a platform to AI & Expert Systems (ES) scientists and professionals to share their research and report new advances in the field of AI and ES. IJAE is a refereed journal producing well-written original research articles and studies, high quality papers as well as state-of-the-art surveys related to AI and ES. By establishing an effective channel of communication between theoretical researchers and practitioners, IJAE provides necessary support to practitioners in the design and development of intelligent and expert systems, and the difficulties faced by the practitioners in using the theoretical results provide feedback to the theoreticians to revalidate their models. IJAE thus meets the demand of both theoretical and applied researchers in artificial intelligence, soft computing and expert systems.

IJAE is a broad journal covering all branches of Artificial Intelligence and Expert Systems and its application in the topics including but not limited to technology & computing, fuzzy logic, expert systems, neural networks, reasoning and evolution, automatic control, mechatronics, robotics, web intelligence applications, heuristic and AI planning strategies and tools, computational theories of learning, intelligent system architectures.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJAE.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with Volume 6, 2015, IJAE will appear with more focused issues related to artificial intelligence and expert systems studies. Besides normal publications, IJAE intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS
The realm of International Journal of Artificial Intelligence and Expert Systems (IJAE) extends, but not limited, to the following:

- AI for Web Intelligence Applications
- AI Parallel Processing Tools
- AI Tools for Computer Vision and Speech Understand
- Application in VLSI Algorithms and Mobile Communication
- Case-based reasoning
- Derivative-free Optimization Algorithms
- Evolutionary and Swarm Algorithms
- Expert Systems Components
- Fuzzy Sets and logic
- Hybridization of Intelligent Models/algorithms

- AI in Bioinformatics
- AI Tools for CAD and VLSI Analysis/Design/Testing
- AI Tools for Multimedia
- Automated Reasoning
- Data and Web Mining
- Emotional Intelligence
- Expert System Development Stages
- Expert-System Development Lifecycle
- Heuristic and AI Planning Strategies and Tools
- Image Understanding

- Inference
- Intelligent Planning
- Intelligent System Architectures
- Knowledge-Based Systems
- Logic Programming
- Multi-agent Systems
- Neural Networks for AI
- Parallel and Distributed Realization of Intelligence
- Reasoning and Evolution of Knowledge Bases
- Rule-Based Systems
- Uncertainty

- Integrated/Hybrid AI Approaches
- Intelligent Search
- Knowledge Acquisition
- Knowledge-Based/Expert Systems
- Machine learning
- Neural Computing
- Object-Oriented Programming for AI
- Problem solving Methods

- Rough Sets
- Self-Healing and Autonomous Systems
- Visual/linguistic Perception

CALL FOR PAPERS

Volume: 6 - **Issue: 1**

i. Submission Deadline : November 30, 2014   **ii. Author Notification:** December 31, 2014

iii. Issue Publication: January 2015

CONTACT INFORMATION