ISSN: 0975-766X
CODEN: IJPTFI
Available Online through *Research Article*
www.ijptonline.com

DECISION SUPPORT SYSTEM FOR SYSTEMIC LUPUS ERYTHEMATOSUS USING MACHINE LEARNING TECHNIQUES

S. Samundeswari[1,2*], V. Ramalingam[1], B. Latha[2], S. Palanivel[1]
[1]Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar - 608002, Tamil Nadu, India;
[2]Department of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai – 600044, Tamil Nadu, India;
Email: samu_siva@yahoo.co.in

Received on: 03-02-2017 *Accepted on: 12-03-2017*

Abstract

Systemic Lupus Erythematosus(SLE) is an autoimmune disease in which the human immune system becomes hyperactive and attacks normal, healthy tissues. Diagnosing the lupus is difficult and the doctor may take long time to diagnose this complex disease accurately. The research design made use of clinical dataset collected from various hospitals. Clinical data of 600 patients (300 SLE and 300 normal) are taken into the experimental study. This research paper used machine learning techniques, that can offer a support to the physician decision. Classification algorithms are used to classify the lupus data set into SLE or Normal. Three classifiers, Multilayer Perceptron, Radial Basis Functions and Support Vector Machines are used to classify and design the decision support system.

Experimental results show that the SVM with polynomial kernel model provides higher accuracy than other models for identifying SLE.

Keywords: Systemic Lupus Erythematosus (SLE), Classification, Machine Learning techniques, Multilayer Perceptron(MLP), Radial Basis Functions(RBF), Support Vector Machines(SVM).

1. Introduction

Systemic Lupus Erythematosus (SLE) is the systemic autoimmune disease in which body tissues are attacked by its own immune system. SLE can affect multiple organs in the body like heart, lungs, liver, kidneys and central nervous system. When compared to men, the disease appears nine times commonly in women, especially it is more common in women of child-bearing ages. Unfortunately, the cause of lupus is unpredictable with periods of illness and wellness. This variation leads to diagnosis of SLE very challenging. As Lupus has no single diagnostic tool, it can be identified through clinical and laboratory criteria. The American College of Rheumatology (ACR) developed the

criteria for classifying patients with SLE[1] which is shown in Table 1.

Table 1. The American College of Rheumatology Revised Criteria for Classification of Systemic Lupus Erythematosus

Criterion	Definition
1. Malar Rash	Fixed erythema, flat or raised, over the malar eminences, tending to spare the nasolabial folds
2. Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
3. Photosensitivity	Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation
4. Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by physician
5. Nonerosive Arthritis	Involving 2 or more peripheral joints, characterized by tenderness, swelling, or effusion
6. Pleuritis or Pericarditis	1.Pleuritis--convincing history of pleuritic pain or rubbing heard by a physician or evidence of pleural effusion OR 2.Pericarditis--documented by electrocardigram or rub or evidence of pericardial effusion
7. Renal Disorder	1. Persistent proteinuria > 0.5 grams per day or > than 3+ if quantitation not performed OR Cellular casts--may be red cell, hemoglobin, granular, tubular, or mixed
8. Neurologic Disorder	1.Seizures--in the absence of offending drugs or known metabolic derangements; e.g., uremia, ketoacidosis, or electrolyte imbalance OR 2.Psychosis--in the absence of offending drugs or known metabolic derangements, e.g. uremia, ketoacidosis, or electrolyte imbalance
9. Hematologic Disorder	1.Hemolytic anemia--with reticulocytosis OR 2.Leukopenia--< 4,000/mm3on ≥ 2 occasions OR 3.Lyphopenia--< 1,500/ mm3on ≥ 2 occasions OR 4.Thrombocytopenia--<100,000/ mm3in the absence of offending

	drugs
10. Immunologic Disorder	1.Anti-DNA: antibody to native DNA in abnormal titer OR 2.Anti-Sm: presence of antibody to Sm nuclear antigen OR 3.Positive finding of antiphospholipid antibodies on: 1. an abnormal serum level of IgG or IgM anticardiolipin antibodies, 2. a positive test result for lupus anticoagulant using a standard method, 3. a false-positive test result for at least 6 months confirmed by Treponema pallidum immobilization or fluorescent treponemal antibody absorption test
11. Positive Antinuclear Antibody	An abnormal titer of antinuclear antibody by immunofluorescence or an equivalent assay at any point in time and in the absence of drugs

For the purpose of identifying patients in clinical studies, a patient shall be diagnosed to have systemic lupus erythematosus if any 4 or more of the 11 criteria are present. Clinical features in individual patients varies from mild joint disease to severe life threatening diseases. The clinical manifestations of disease is sensitive and also specific for the diagnosis of disease activity. The disease could be identified through a combination of laboratory criteria. Antinuclear Antibody (ANA) titer is the primary laboratory test to predict SLE. Generally, symptoms vary from person to person, with joint pain, swelling, fever, fatigue, sensitivity to sunlight, skin rash, chest pain while taking deep breath, hair loss.

Severity reduction and flares prevention are included in the SLE treatment. It can also include corticosteroids and anti-malarial drugs. Diagnosis of Systemic Lupus Erythematosus should be accurate as it is important for the treatment that can reduce morbidity and mortality. In this proposed method, the features are extracted with the help of the rheumatologists. Machine learning plays an important role in the medical field for the automation of the diseases. Once the features are extracted, a classifier can be trained to classify the patients as SLE or healthy.

2. Related Work

Various kinds of approaches have been proposed to improve the accuracy level for the prediction of SLE. Diego F. Hernandez Ramirez and Javier Cabiedes, [2] presented a paper for Immunological techniques that allow the detection of multiple autoantibodies at the same time with small volume of samples. Immunological techniques like Indirect Immunofluorescence (IFI), Enzyme Linked Immunosorbent Assay (ELISA), Electro Immuno Transference (EIT) that confirm the presence of antibodies in patient samples. R Rajimehr, [3] proposed a new method for predicting lupus

nephritis in patients with SLE was compared with a logistic regression model and clinical diagnosis.

Melisa R. Arbuckle, [4] developed a method to investigate onset and progression of SLE autoantibody development before the clinical diagnosis. The serologic and clinical findings, suggest that there are at least three phases in the development of SLE autoantibodies. Normal phase includes asymptomatic persons with no SLE autoantibodies. In the second phase, benign autoimmunity, there is a laboratory finding but without clinical manifestations. The third phase, pathogenic autoimmunity, is marked by the presence of more antibodies namely anti-dsDNA, anti-sm, and anti- nuclear ribonucleoprotein by the onset of signs and symptoms leading to clinical presentation and diagnosis.

Tobias Schmidt-Wilcke, [5] proposed a method to identify regional changes in white matter integrity which can be detected in SLE patients without neuropsychiatric symptoms (non-NPSLE patients). DTI and TBSS **(**tract based spatial statistics) methods are used to investigate SLE patients from healthy one. Steven R. Binder, [6] introduced a new method that combines a multiplex immunoassay using k nearest neighbour (kNN) algorithm for computer assisted pattern recognition.

This algorithm was developed and tested using "leave one out" strategy.

RenuSaigal, Amitkansal, [7] proposed a method to identify the clinical pattern and disease course in patients with SLE at Western India. All patients satisfying the revised American College of Rheumatology criteria for SLE were included in the study over a one year period. Antinuclear antibody (ANA) and anti-dsDNA was positive in 98.3% and 65% patients respectively.

Mohamad Hasan Bahari, [8] introduced a method for recognizing pathogenic antibodies in SLE. dsDNA antibodies have been present in the pathogenesis of autoimmune diseases.

2.1. Outline of the work

This work deals with the classification of SLE using classification algorithms Multilayer perceptron, Radial Basis Functions, Support Vector Machine and their performances are compared. The proposed method is evaluated using clinical dataset.

After consulting rheumatologist, 12 features are selected from the case sheet. The obtained features are used to classify the patients into SLE or non-SLE. The results are then compared using different machine learning techniques.

The remaining of this paper is organised as follows. Section 3 describes about methodology. Experimental results are discussed in section 4. Finally, Section V concludes the paper followed by references.

3. Methodology

The proposed methodology is to design and develop a decision support system for SLE using classification algorithms. The performance of the classifiers are compared based on the accuracy. The general overview of the proposed approach is illustrated in Fig. 1.

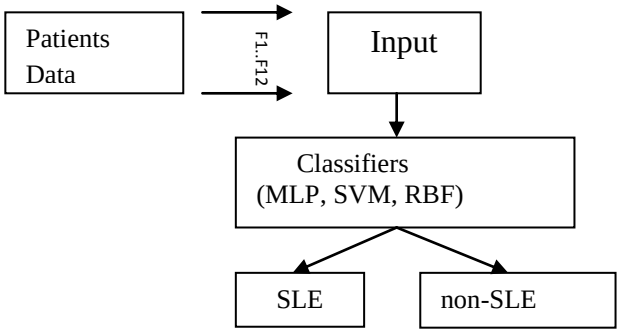

Fig. 1. Overview of proposed method for SLE Classification.

3.1. Feature Extraction

The Dataset is collected from various hospitals containing records of 600 patients of age group 15 to 35. They were classifed into two groups : 300 were SLE patients and 300 were Non-SLE(normal)patients. Features plays a vital role in analysis part which gives a meaningful data. 12 features are extracted from the records on consultation with Rheumatologist. Table 2 shows the features used for classification.

Table 2. Features used for classification.

Features	Description	Type
Age	Patient's age	Numeric
Hb	Haemoglobin	Numeric
Urine pcr	Urine – Proteincreatine	Numeric
ESR	Erthrocyte Sedimentation rate	Numeric
C3	Complement component 3	Numeric
C4	Complement component 4	Numeric
RBC Cast	Checking the kidney function	Numeric
platelet	Component of Blood	Numeric
ANA	Antinuclear antibody	String
Anti ds DNA	Anti double stranded antibody	String
Anti smith	Anti smith antibody	Numeric
WBC Cast	White Blood Cell count	String

3.2. SVM Classifier

Support Vector Machines (SVM) are set of supervised learning methods which is used for classification, outliers detection and regression. SVM constructs a linear model to estimate the decision function using non-linear class boundaries based on support vectors. SVM determines a linear separating hyperplane with the maximal margin in the higher dimensional space. Fig. 2 shows the architecture diagram of the SVM. Support Vector Machine (SVM) can be used for classifying the obtained data. It constructs a set of hyper planes in a high dimensional space. Let us done a feature vector by $x = (x_1, x_2, \ldots x_n)$ and its class label by y such that $y = \{+1, -1\}$. Therefore, let us take a problem of separating the set of n training samples which belongs to two classes,

$$(x_i, y_i), x_i \in R^n, y = \{-1, +1\}, i = 1,2, \ldots, n \quad (1)$$

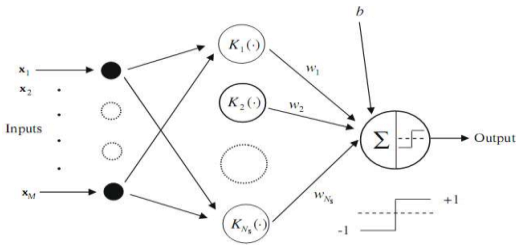

Fig. 2. Architecture of SVM.

3.3. Multilayer Perceptron

A Multi Layer Perceptron(MLP) is an artificial neural network that maps input data onto a set of appropriate outputs. An MLP consists of one or more layers of nodes with each layer are fully connected to the next one. It consists of three layers namely input layer, hidden layer and output layer. The training and testing data are given to input layer and processed by the hidden and output layers. Żak[10] and Hassoun [11] proposed the analysis and study of multilayer perceptrons. MLP utilizes supervised learning technique called backpropagation for training the network. The Fig. 3 illustrates the architecture of a multilayer network .

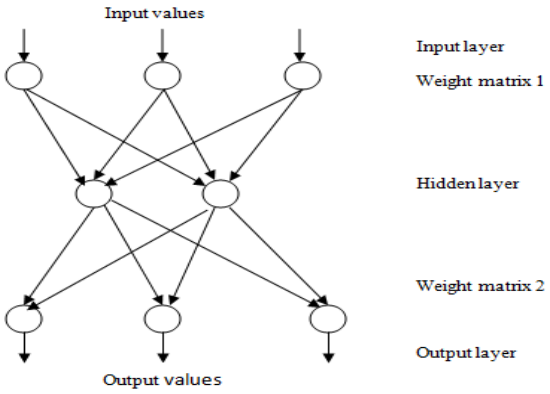

Fig. 3. Architecture of MLP

3.4. Radial Basis Function

A Radial Basis Function nework is an artificial network consists of layer of neurons. RBF is based on supervised learning. Its structure is similar to structure of MLP. RBF differs from MLP as RBF contains a hidden layer that contains called RBF units. The location of the function's centre and its deviation or width are the two parameters of each RBF. The hidden unit finds the distance between the center of its RBF and input data vector. RBF has only one hidden layer. It uses Gaussian transfer functions.

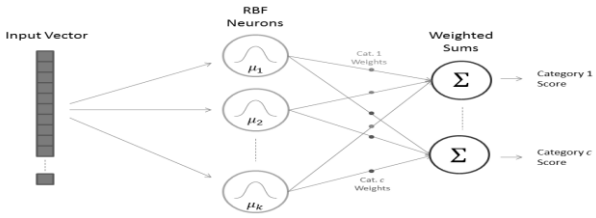

Fig. 4. Schematic diagram of RBF.

4. Results and Discussion

4.1. Accuracy Measures

For our experimentation, we use three classifiers namely Multilayer Perceptron, Radial Basis Function and SVM networks for classification. The tests were performed using 10 folds cross validation. The various performance measures which are used to assess the classifiers performance are precision, recall and accuracy(correct classification). These are given by

$$\text{Precision} = TP/(TP+FP) \quad (2)$$

$$\text{Recall} = TP/(TP+FN) \quad (3)$$

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (4)$$

TP stands for true positive, **FP** for false positive, **FN** for false negative, and **TN** for true negative counts.

4.2. Classifier performances

Table 3 represents the classifiers with corresponding Precision and recall values. From Fig. 5, it was observed that SVM (with Polynomial Kernel) outperformed the other two classifiers MLP and RBF.

Table 3. Classifiers with corresponding Precision and Recall values.

Classifier	Precision		Recall	
	SLE	Non-	SLE	Non-

		SLE		SLE
MLP	0.993	0.997	0.997	0.993
RBF	0.99	0.99	0.99	0.99
SVM	0.993	1	1	0.993

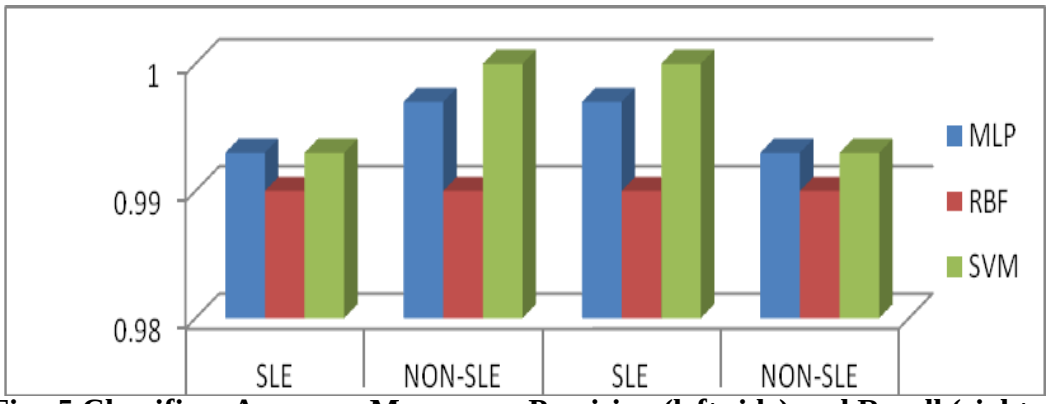

Fig. 5 Classifiers Accuracy Measures - Precision (left side) and Recall (right side)

The clinical dataset was given as input to the model which uses all classifications discussed in section III. The classifiers with corresponding Error rate and Accuracy values are shown in Table 4.

Table 4. Error Rate and Accuracy Values.

Classifier	Error Rate	Accuracy value (%)
MLP	0.0066	99.5
RBF	0.0157	99
SVM	0.0033	99.7

From Fig. 6, one can observe that the SVM provides lower error rate when compared to other TWO classification algorithms.

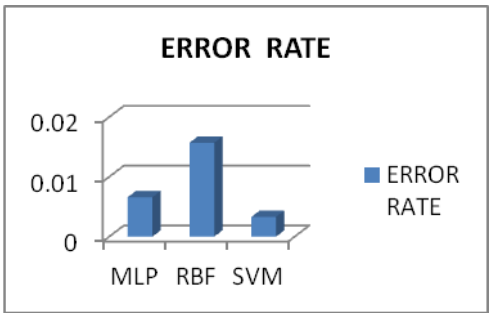

Fig. 6 Classifier Error Rate.

The Accuracy values of all classifier algorithms are shown in Fig. 7. It is observed that SVM classifier has achieved a better performance with 99.7 percent of accuracy whereas MLP has given 99.5 percent and RBF has given 99 percent

of accuracy.

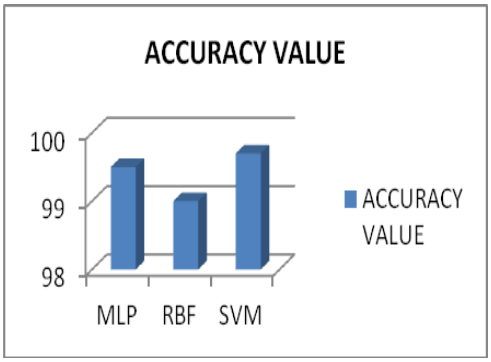

Fig. 7. Classifier Accuracy Value.

5. Conclusion

A decision support system for early diagnosis of Systemic Lupus Erythematosus is proposed. Lupus affects all the major parts of the body such as lungs, heart, kidneys and also central nervous system. It is important to diagnose the lupus at early stage to ensure speedy treatment as it mimics other diseases. The diagnosing methodology uses different classifier algorithms for the classification of SLE from a group of patients. SVM shows the maximum performance than other classifiers MLP and RBF. SVM yields maximum of 99.7% as accuracy. If precision, recall, accuracy and error rate are considered with combined together, one can conclude that SVM with polynomial kernel is the only better choice than other classifiers, namely, MLP and RBF. Further this decision is supported by Rheumatologist also.

6. References

1. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erthematosus, Arthritis Rheum. 1982; 25:1271-7.
2. Diego F. Hernandez Ramirez and Javier Cabiedes, “Immunological techniques that support the diagnosis of autoimmune diseases”, Reumatol Clin. vol. 6, no. 3, pp. 173-177, 2010.
3. R Rajimehr, S Farisu, L Montaser Kousari, A Bidari , C Lucas, S Yousefian and F Bahrami, “Prediction of lupus nephritis in patients using artificial neural networks”, Lupus vol. 11, pp. 485-492, 2002.
4. Melisa R. Arbuckle, M.D., Ph.D., Micah T. McClain, Ph.D., Mark V. Rubertone, M.D., “Development of Autoantibodies before the Clinical Onset of Systemic Lupus Erythematosus”, N Engl J Med; vol. 349, 2003.
5. Tobias Schmidt-Wilcke, Patricia Cagnoli, Page Wang, Thomas Schultz, “White matter integrity in patients with systemic lupus erythematosus”, NeuroImage: Clinical 2014.

6. Steven R. Binder, Mark C. Genoves, Joan T. Merrill, "Computer – Assisted Pattern Recognition of Autoantibody Results", Division of Immunology and Rheumatology, Stanford University, Stanford, California 94305.
7. RenuSaigal, Amitkansal, Hari ram Maharia, Manish Juneja, "Clinical profile of systemic lupus erythematosus patients at a tertiary care centre in Western India", vol. 1, pp. 27-32, 2011.
8. Mohamad Hasan Bahari, Asad Azemi and Morteza khadime, "Recognizing pathogenic antibodies in SLE using GRNN", ICIC International, vol. 8, no. 4, April 2012.
9. Eman M. Elserougy, Iman I. El-Gazzar, Iman M. Fikry, "Study of SLE Patients correlation with disease activity and damage indices", The Egyptian Rheumatologist vol. 35, pp. 207-215, 2013.
10. Zak S.H., "Systems and Control", NY: Oxford Uninversity Press, 2003.
11. Hassoun M.H, "Fundamentals of Artificial Neural Networks", Cambridge, MA: MIT press, 1999.
12. William H.Robinson, Derek Fong, Josef S.Smolen and Paul J.Utz, "Autoantigen microarrays for multiplex characterization of autoantibody responses ", Nature Medicine, vol. 8, March 2002.
13. George Bertsias, Ricard Cervera, Dimitrios T Boumpas, "SLE: Pathogenesis and Clinical features", EULAR Textbook on Rheumatic Diseases.
14. La Revue de and Sante de la, "Eastern Mediterreanean Health Journal", vol. 16, no. 8, 2010.
15. J. C. Burges Christoper, "A tutorial on support vector machine for pattern recognition", vol. 52, 1998.
16. V. Vapinik, "Statistical Learning theory", Wiley, New York, 1998.
17. Ronan Collobert and Samy Bengio, "SVMTorch: Support Vector Machines for large scale regression problems", Journal of Machine Learning Research, vol. 1, no. 1, 2001.
18. Powell, M. J. D, " Radial basis functions for multivariable interpolation : A review" In Algorithms for Approximation, J. C. Mason and M. G. Cox, Eds. Oxford University Press, Oxford, UK, 1987, pp. 143–167 .
19. Mrinal Kanti Dhar, Quazi M.Hasibul Haque, Md. Tanjimuddin, "Facial Recognition by Radial Basis Function Network ", International Journal of Computer Applications, vol. 78. No. 3, Sept 2013.