International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# Hybrid Deep Learning Model for Prediction of Systemic Lupus Erythematosus

**M. Sobhana[1], Dodda Pragathi[2], Ginjupalli Naga Deepika[3], Yaswant Sai Kavuri[4], Nihitha Vemulapalli[5], M. G. S. S. Venkatesh[6]**

**Submitted**: 04/05/2023 **Revised**: 17/07/2023 **Accepted**: 08/08/2023

**Abstract:** Artificial Intelligence is widely used in health care to classify and predict diseases. Systemic Lupus Erythematosus (SLE) is the most common type of lupus and presents intricate challenges in accurate prediction due to its multifaceted nature. SLE is an inflammatory disease caused by the immune system attacking its tissues. Lupus most likely originates from a synthesis of genetics and environmental challenges. This study used the GEO dataset to develop an accurate and precise model for the prediction of SLE. However, choosing the right features is crucial in training a model. This study aims to enhance the predictive capabilities of SLE using a hybrid approach of Genetic Algorithm (GA) integrated with neural networks. The subset of features used by Artificial Neural Network (ANN) is optimized by feature selection using GA. The proposed model is GA-ANN, and experimental results indicate that the model performed well in comparison to other models, achieving an accuracy of 96.32%.

***Keywords:*** *Artificial Neural Networks, Genetic Algorithm, Systemic Lupus Erythematosus, Prediction.*

## 1. Introduction

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune illness that can affect various organs in the human body [1]. Hormonal changes, specific medications, infections, and exposure to ultraviolet radiation can cause or exacerbate SLE symptoms. Hormones, particularly estrogen, are thought to contribute to the higher prevalence of SLE in women [2]. Fluctuations in hormone levels during the menstrual cycle, pregnancy, and menopause can impact the disease's activity. The heterogeneous nature of SLE poses challenges [3] in early diagnosis and prediction, often leading to delayed interventions and increased disease severity. The intricate interplay of genetic and environmental factors contributes to the complexity of SLE pathogenesis, warranting innovative approaches to improve prediction accuracy.

Recent advancements in computational methods have opened avenues for harnessing complex data patterns to enhance disease prediction models. Among these methods, Artificial Neural Networks (ANNs) [4] have demonstrated remarkable proficiency in learning intricate relationships within datasets. However, the success of ANN models can be influenced by factors such as feature selection, dataset dimensionality, and the architecture's design. Genetic Algorithms (GAs) [5] have emerged as powerful optimization techniques to address these challenges by refining feature selection and enhancing the ANN's performance. GAs simulate natural selection to iteratively evolve the most relevant subset of features, thereby improving model efficiency and accuracy.

Early diagnosis of SLE can significantly reduce the severity of the disease. In this study, our primary objective is to predict whether individuals have SLE or not. To achieve this, we have developed a hybrid deep learning model that integrates GA with ANN, aiming to construct a precise and accurate predictive model for SLE. The paper is organized as follows Section 2 provides an overview of the research work, Section 3 provides a detailed explanation of the proposed methodology and technologies used, Section 4 showcases the outcome of our proposed study, and Section 5 summarizes the key findings and implications of the study.

## 2. Literature Review

Several researchers explored machine learning methods to compare the performance of various prediction algorithms in classifying SLE. AM Jorge [6] proposed that the Hospitalizations for SLE can be predicted using algorithms like Gradient Boosting (GB), Logistic Regression (LR), Decision Trees (DT), and Random Forests (RF). A total of 2,780 SLE patients were represented in the dataset, and 0.26 of them underwent at least one hospitalization throughout

[1] *Department of Computer Science and Engineering. V. R Siddhartha Engineering College, Kanuru-520007, INDIA*
*ORCID ID : 0000-0001-5938-5740*
[2] *Department of Computer Science and Engineering. V. R Siddhartha Engineering College, Kanuru-520007, INDIA*
*ORCID ID : 0009-0003-4879-6298*
[3] *Department of Computer Science and Engineering. V. R Siddhartha Engineering College, Kanuru-520007, INDIA*
*ORCID ID : 0009-0005-0436-0008*
[4] *Department of Computer Science and Engineering. Prasad V. Potluri Siddhartha Institute of Technology, Kanuru-520007, INDIA*
*ORCID ID : 0009-0005-7591-2578*
[5] *Department of Computer Science and Engineering. V. R Siddhartha Engineering College, Kanuru-520007, INDIA*
*ORCID ID : 0009-0007-8626-0012*
[6] *Department of Computer Science and Engineering. V. R Siddhartha Engineering College, Kanuru-520007, INDIA*
*ORCID ID : 0009-0004-6406-000X*
** Corresponding Author Email: sobhana@vrsiddhartha.ac.in*

the research period. With AUCs of 0.751 and 0.772 for two methods (averaging, and progressive), the RF algorithm outperformed all other algorithms in terms of performance.

Utilizing longitudinal data, Y Zhao [7] focused on forecasting lupus hospitalizations. The 925 individuals with longitudinal data were included in a multicenter lupus cohort using an electronic health record (EHR). The first Differential technique introduces new delayed variables between successive time steps to account for the temporal relationships in sequential data. Next, assess the performance of LSTM, an innovative time series-specific deep learning algorithm. The bagging strategy, not only balanced the training data but also offered the benefits of ensemble learning, and was responsible for the Differential approach's higher stability than the LSTM model with an accuracy of 88%, specificity of 79%, and recall of 74%.

Women with SLE were investigated in this study [8,9] for their unfavorable pregnancy outcomes. The authors utilized PROMISSE data. An APO prediction model is developed using logistic regression based on this data. An extension of their approach included machine learning algorithms such as LASSO, RF, SVM-RBF (all AUC=0.77), and Super Learner (SL) to determine if they could improve the prediction and identify additional risk factors. With an AUC of 0.78, the SL method performed the best.

J Stojanowski [10] addressed that an effective method for forecasting the fate of Lupus Nephritis is an Artificial Neural Network. An ANN with a backpropagation algorithm was utilized in the study to forecast the course of Lupus Nephritis. The training set of data was used to train the ANN, and cross-validation methods were used to fine-tune the model’s hyperparameters. With an AUC-ROC of 0.9375 (0.94) and an accuracy of 91.67, a multi-layer perceptron architecture with 40 neurons in the first hidden layer followed by 45 neurons in the second hidden layer appeared to perform the best. Y Zhao [11] proposed a calibrated ensemble (CE) model to predict severe flares in Lupus patients. In this study, they used three models Logistic Regression, RF, and Naive Bayes, and combined the predictions of these models to form a CE model with an accuracy of 79%.

According to a study [12], a Random forests algorithm predominately improves the genetic risk prediction of SLE. Choosing relevant genetic variations, preparing the data, employing an ML algorithm for predicting the risk of SLE, and evaluating the model's performance were the phases of the study. In this study, SLE cases were categorized using three ML algorithms: SVM, RF, and ANN. With a mean AUC of 0.84, RF has outperformed other classifiers.

The work proposed by A. Malarvizhi [13] evaluates the performance of the Ensemble-based technique (EARLNP) using the kidney dataset. This study aimed to detect lymph nodes (LNs) using Health Record Systems (HRSs) with minimal tests, focusing on crucial characteristics. The study highlights specific gravity, albumin, blood glucose, sugar, potassium, packed cell volume, serum creatinine, white blood cell, and red blood cell count as key factors for effective LN prediction, identified through a filter feature selection technique. Employing 10-fold cross-validation, researchers trained, tested, and validated classifiers. The proposed EARLNAP model stood out, achieving 88.8% accuracy post-parameter optimization.

An extremely contagious illness called Covid-19 (Coronavirus) has now spread worldwide. Severe cold, cough, sore throat, body soreness, and other symptoms are among them. Traditional diagnostics were expensive and took a long time to discover COVID-19. As a result, the researchers E. Benmalek [14] developed a practical method for detecting COVID-19 using machine learning classifiers. PCA is used to explain the data collecting, data preparation, and feature selection processes. Machine learning classifiers like SVM and RF successfully distinguished between COVID-19 positive and negative persons based on cough noises, which are known to display unique patterns in COVID-19 patients, with an accuracy of 97.48%.

The case study presented by Yevgeniya Gartshteyn et al., [15] explores the interaction between covid19 and SLE. The authors examined a series of cases involving individuals who had both COVID-19 and preexisting SLE. It might discuss whether having SLE influenced the severity of COVID-19 symptoms, how the patient's SLE medications and immune responses interacted with the viral infection, and any particular challenges faced by individuals with both conditions.

## 3. Method

### 3.1. Proposed Methodology

The methodology diagram of our proposed model is represented in Fig 1. Development of the proposed model starts with the collection of datasets from the GEO website. The data is then pre-processed by replacing null values with mean values, Label Encoder is used to convert categorical data into numerical values, and standardization of data is done using Standard Scaler. Feature selection is done using a Genetic algorithm. Then, the dataset is divided into training and testing data. The Multi-Layer Feedforward Neural Network also known as the Multi-Layer Perceptron (MLP) model is built. The features selected by GA are used to train the MLP model. The trained MLP model is tested with testing data, and model performance is assessed using classification metrics.

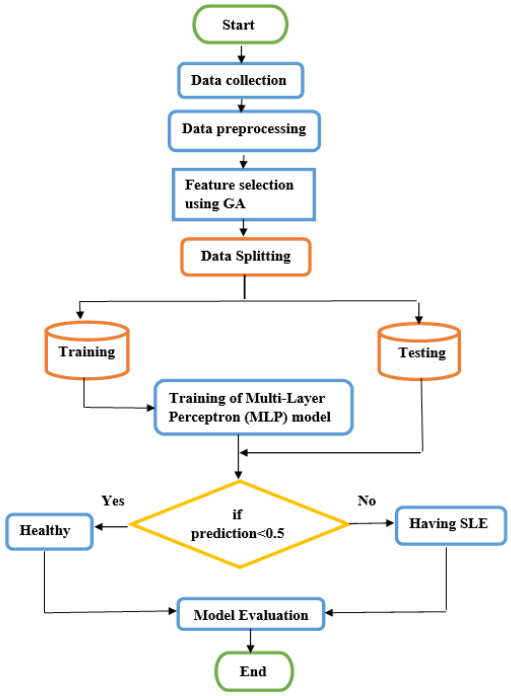


**Fig. 1.** Process flow diagram of Systemic Lupus Erythematosus prediction model.

### 3.2. Data Collection

Our study uses the GEO dataset with accession number GSE65391 to predict SLE. Nicole Baldwin from the University of Baylor Research Institute in the American city of Dallas developed the dataset Longitudinal Transcriptional Pediatric SLE Research including clinical characteristics. With the help of this data, transcriptional correlates at both the cohort and individual levels of SLE disease activity can be found. The researchers obtained 924 SLE samples by profiling the whole blood transcriptomes of 158 individuals over a period of up to 4 years using microarrays. In the end, 996 samples with 88 qualities were examined, of which 924 were SLE samples and 72 were healthy samples.

### 3.3. Data Pre-processing

The Dataset we used in our study GSE65391 includes a huge number of null values. For performing any analysis of the data it should be free from noise. Data cleaning involves handling missing values and identifying the outliers. For handling null values, we replaced null values with mean values. A simple algorithm is used to handle null values. The algorithm first identifies null values in respective columns and then replaces the null value with the mean value of its appropriate column.

Data Encoding is also an important step in data pre-processing. Encoding of data involves converting categorical data into numerical data. In the dataset GSE65391 there are some columns with numerical data and the remaining with categorical data and this uneven data is not suitable for selecting relevant features in the dataset. Also, it will be easy for machine learning models to remember a single integer instead of a complex string label. So, we used a label encoder to convert categorical data into numerical data. Label encoder works by assigning a unique integer for each unique categorical value in the respective category. Label encoder is most suitable for ordinal categorical values. And it is most preferable when there are more categorical groups.

Standardization of data is to convert all the feature values onto a similar scale which will prevent the dominance of one feature over the other and ensure a fair comparison. In the data GSE65391, the feature values are not on a similar scale this will lead to the dominance of the features with high feature values and this data will not select relevant features that cause SLE. So before performing feature selection, we used Standard Scaler to convert all the feature values onto a similar scale. By standardizing the data set the algorithm will treat all features equally thus avoiding biased analysis and comparisons.

### 3.4. Feature Selection Using Genetic Algorithm

The Genetic Algorithm (GA) [15] is the optimization algorithm inspired by the process of natural selection. It works iteratively to obtain the most optimal subset of features. We integrated GA into the feature selection process to further enhance the model's predictive power. GA helped identify the most relevant features that contributed significantly to the prediction of SLE. By selecting an optimal subset of features, we aimed to increase the model's interpretability and generalization ability, which are crucial factors in real-world clinical applications. The subset of features that are selected using GA are listed in Table 1.

**Table 1.** Features selected using Genetic Algorithm

| | |
|---|---|
| Source name | Asa_category |
| Batch | Seizure |
| Visit | Sledai_component_class |
| Mdg | Hgb |
| Fever | Cr |
| Cumulative_time | C4 |
| Thrombocytopenia | Alt |
| Days_last_visit | Treatment |
| Race | Vasculitis |
| Biopsy_history | Disease_activity |
| Lymphocyte_count | Sledaic_lmm2 |
| Metotrexate_category | Cranial_nerve_disorder |

**Algorithm 1:** Feature selection using Genetic Algorithm

**Step 1:** Population ← Generate the initial population with random feature selection.

**Step 2: Begin for loop**

**Step 2.1:** fitness_value ← [Evaluate fitness for each individual in the population using a fitness function (classification accuracy)]

**Step 2.2:** parent 1, parent 2 ← select parents using the selection tournament

**Step 2.3:** offspring1, offspring2 ← single-point crossover between parents

**Step 2.4:** mutated_offspring ← apply bit-flip mutation to offspring1

**Step 2.5:** mutated_offspring2 ← apply bit-flip mutation to offspring2

**Step 2.6:** replace least fitted individuals with the mutated offspring

**End for loop**

**Step 3:** ga_features choose the best individuals from the features with the highest fitness value

### 3.5. Classification

The samples are categorized by assigning them to the SLE and Healthy classes, respectively. If the class label is SLE, then the person has the disease; if the class label is Healthy, then the person is healthy.The SLE classification model is required to categorize. The data is classified using ANN in our suggested method. A Multi-Layer Feed-forward Neural Network (MLFNN), also called Multi-Layer Perceptron (MLP), with two hidden layers is selected for the prediction of SLE.

The optimization algorithm used is 'Adam'. In ANN hyperparameters play an important role such as the Learning rate, the number of hidden layers, the number of neurons in each layer, and the Activation function have to be tuned to increase the performance of classification. The Activation functions used are 'ReLu' [16] and 'Sigmoid' [17]. The output layer employs the sigmoid function, whereas the hidden layers use the ReLu function. Since the Classification of SLE is binary, binary_crossentropy is used as a loss function. Drop-Out layers are also added in between the hidden layers to prevent overfitting. L2 kernel regularizer [18] of rate 0.01 is used in dense layers to improve the generalization of the ANN model by preventing the

Model from preventing.

**Algorithm 2:** GA-ANN model

**Start**

**Step 1:** Build ANN Model ()

MLP with two hidden layers, dropout layers, and l2 regularization.

**Step 2**: Split the data into training and testing parts in the ratio of 80:20

**Step 3:** Train the neural network using the features selected by GA.

**Loop 1:** Until stopping criteria met

**Forward Propagation ()**

$h \leftarrow \text{ReLu}(w_h \times inp)$

Dropout ()

$o \leftarrow \text{Sigmoid}(w_0 \times h)$

Loss = BinaryCrossEntropy(o, true_labels)

**Backward Propagation ()**

$\Delta_0 \leftarrow o$ - true_labels

$\Delta h = f\,1\,(h) \times (\text{weight\_output\_transpose} \times \Delta_0)$

$W_0(\text{new}) \leftarrow$ update weight_output

$W_h(\text{new}) \leftarrow$ update weight_hidden

Apply L2 regularization to updated weights

if (stopping criteria met)

Stop training

**End Loop**

**Step 4: Prediction ()**

feed the testing data into the trained ANN model

y ← trained_model.predict(x_test,y_test)

if (y > = 0.5)

"SLE"

else

"Healthy"

**Stop**

### 3.6. Performance Evaluation

Our study evaluates and compares the performance of GA-ANN with PCA-ANN. Performance metrics like precision, recall, f1-score, and accuracy are used to assess the classification models. The comparison of the performance of PCA-ANN and GA-ANN models is shown in Table 2.

**Table 2.** Comparison of performance of PCA-ANN and GA-ANN models

| *Model* | *Accuracy* | *Precision* | *Recall* | *F1-score* |
|---|---|---|---|---|
| PCA-ANN | 91.63% | 91.94% | 99.63% | 95.63% |
| GA-ANN | 96.32% | 99.80% | 97.81% | 98.89% |

## 4. Results and Discussion

The GA-ANN explores the more influential features of SLE characteristics. In a previous study of PCA-ANN, PCA was used to reduce the dimensionality of the complex dataset while retaining its essential information. As a result, 5 crucial features have been selected from the scree plot. Then, these features are utilized to train and test the ANN. This model achieved an accuracy of 91.63%. To enhance the models' prediction ability, GAs for feature selection optimize the predictive capabilities. The genetic optimization process iteratively evolved a subset of features that were most relevant for accurate SLE prediction. This refined feature set was then used to train and test the ANN. The GA-ANN model demonstrated a significantly improved accuracy of 96.32%. The integration of genetic optimization through GAs facilitated the identification of the most informative features, further enhancing the model's discriminatory power. This impressive accuracy suggests that the GA-ANN model effectively harnessed the synergy between genetic algorithms and neural networks, resulting in highly accurate SLE prediction.

### 4.1. Experimental Results

This section includes the results of our proposed models PCA-ANN and GA-ANN. The Experimental results of both the proposed models are differentiated using training and validation accuracy graphs, and confusion matrices. Experimental Results and their detailed explanation are shown below. Fig 2. depicts the PCA-ANN model's training and validation accuracy for each epoch. This graph shows the model's ability to generalize to new data. The hold-out validation dataset's validation curve provides a primary indication of how well the model generalizes. The number of training iterations is indicated by the epochs on the X-axis. At the beginning of training, typically there is a rapid decrease in the training loss. This is because the ANN is learning to fit the data and reduce the error.

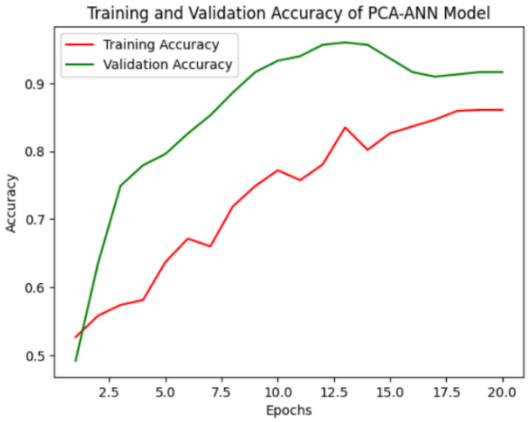


**Fig 2.** Training accuracy and validation accuracy of the PCA-ANN model

Fig 3. shows the GA-ANN model's training and validation accuracy for each epoch. The model's learning efficiency is indicated by the training curve, which was calculated using the training dataset. The testing dataset is used to calculate the validation curve, which evaluates the model's viability. Utilizing new data, validation tests the models' dependability. The number of training iterations is shown by the epoch on the X-axis. This Fig illustrates that the model is able to learn efficiently as well as produce the optimized output with accurate precision.

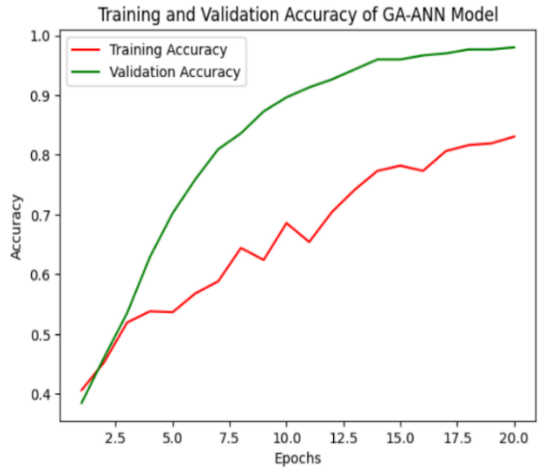


**Fig 3.** Training accuracy and validation accuracy of the GA-ANN model

Fig 4. shows the confusion matrix of the PCA-ANN model when applied to the testing data. It represents the cases where the PCA-ANN model correctly predicted the positive class (e.g., a disease is present) based on the reduced-dimensional input data.

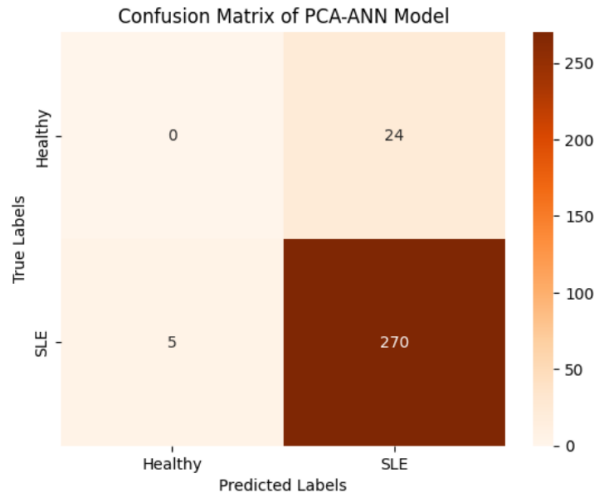


**Fig 4.** Confusion Matrix of PCA-ANN model

Fig 5. represents the confusion matrix of the GA-ANN model it has a high rate of True positives and True Negatives. It is noticeable that the model has the ability to predict both positive samples and healthy samples effectively.

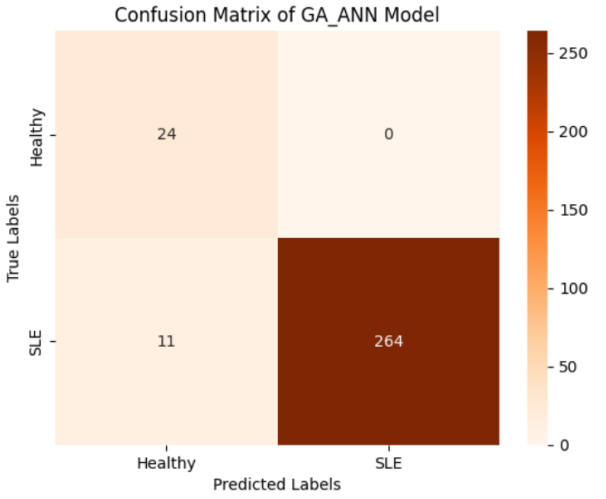


**Fig 5.** Confusion Matrix of GA-ANN model

### 4.2. Discussion

Table 3 describes the comparison of previous related work with our proposed work. So far prediction of SLE is done using various machine learning algorithms. All the previously implemented methods and their performance are noted in Table 3.

**Table 3.** Comparison of performance of PCA-ANN and GA-ANN models

| *Methods* | *Accuracy* | *Specificity* | *Recall* |
|---|---|---|---|
| LSTM [7] | 88% | 79% | 74% |
| Super Learner [8] | - | 77% | 71% |
| ANN [10] 91.67% | 91.67% | | - |
| (LR, RF, NB) [11] | 74% | 71% | 86% |
| RF [12] | - | 68% | 84% |
| PCA-ANN 99.63% | 91.63% | | - |
| GA-ANN 97.81% | 96.32% | | - |

## 5. Conclusion

This research work compares the performance of PCA-ANN with GA-ANN for the prediction of SLE. Based on the experiment's results, the optimal parameters for learning rate and epochs were 0.0001 and 20 respectively for SLE classification. The PCA-ANN model achieved 91.63% accuracy, precision of 91.94%, recall of 99.63%, and F1 score of 95.63%. The GA-ANN outperformed by achieving an accuracy of 96.32% which yielded the most optimized and robust model. Adam's optimizer showcased its efficiency by converging faster during training, enabling the models to achieve higher predictive accuracy within fewer iterations. This is attributed to its adaptive learning rate mechanism, which adjusts the learning rate for each parameter, facilitating smoother and faster convergence. In terms of activation functions, both Sigmoid and Rectified Linear Unit (ReLU) were considered. Sigmoid, with its S-shaped curve, demonstrated suitability for binary classification tasks like SLE prediction. ReLU, on the other hand, being computationally efficient and capable of mitigating the vanishing gradient problem, proved advantageous for deeper networks, potentially enhancing model performance. In conclusion, we believe that the adoption of such advanced predictive models in clinical settings could significantly impact patient care and management, leading to earlier diagnosis and better patient outcomes. However, further research and validation on external datasets are essential to fully establish the model's clinical utility and robustness.

## 6. References and Footnotes

### Acknowledgments

We sincerely appreciate the editors' and reviewers' helpful feedback, which allowed the article to be enhanced to a high degree. The revered Dr. A.V. Ratna Prasad, our Principal, and Dr. D. Rajeswara Rao, Head of the Department of Computer Science and Engineering, are also to be thanked for their assistance with our project.

### Author contributions

**M Sobhana:** Conceptualization, Methodology, Software, Field study **Dodda Pragathi:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Ginjupalli Naga Deepika:** Software Visualization,

Investigation, Writing-Reviewing and Editing. **Yaswant Sai Kavuri:** Reviewing and Editing. **Nihitha Vemulapalli:** Reviewing and Editing. **M G S S Venkatesh:** reviewing and Editing.

**Conflicts of interest**

The authors declare no conflicts of interest.

## References


[1] Rudin, Vanessa, et al. "Digital therapeutic platform for management of systemic lupus erythematosus." 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020. https://doi.org/10.1109/EMBC44109.2020.9175209

[2] Kim, J. W., Kim, H. A., Suh, C. H., & Jung, J. Y. (2022). Sex hormones affect the pathogenesis and clinical characteristics of systemic lupus erythematosus. Frontiers in Medicine, 9, 906475. https://doi.org/10.3389/fmed.2022.906475

[3] Piga, Matteo, and Laurent Arnaud. "The main challenges in systemic lupus erythematosus: where do we stand?." Journal of clinical medicine 10.2 (2021): 243. https://doi.org/10.3390/jcm10020243

[4] Indira, D. N. V. S. L. S., et al. "Improved artificial neural network with state order dataset estimation for brain cancer cell diagnosis." BioMed Research International 2022 (2022). https://doi.org/10.1155/2022/7799812

[5] J. C. T. Arroyo and A. J. P. Delima, "An optimized neural network using genetic algorithm for cardiovascular disease prediction," J. Adv. Inf. Technol., vol. 13, no. 1, 2022. doi: 10.12720/jait.13.1.95-99

[6] Jorge, April M., et al. "Exploration of machine learning methods to predict systemic lupus erythematosus hospitalizations." Lupus 31.11 (2022): 1296-1305. https://doi.org/10.1177/09612033221114805

[7] Zhao, Yijun, Dylan Smith, and April Jorge. "Comparing two machine learning approaches in predicting lupus hospitalization using longitudinal data." Scientific Reports 12.1 (2022): 16424. https://doi.org/10.1038/s41598-022-20845-w

[8] Fazzari, M. J., Guerra, M. M., Salmon, J., & Kim, M. Y. (2022). Adverse pregnancy outcomes in women with systemic lupus erythematosus: can we improve predictions with machine learning?. Lupus Science & Medicine, 9(1), e000769. http://dx.doi.org/10.1136/lupus-2022-000769

[9] Petri, Michelle. "Pregnancy and systemic lupus erythematosus." Best Practice & Research Clinical Obstetrics & Gynaecology 64 (2020): 24-30. https://doi.org/10.1016/j.bpobgyn.2019.09.002

[10] Stojanowski, Jakub, et al. "Artificial neural network-an effective tool for predicting the lupus nephritis outcome." BMC nephrology 23.1 (2022): 1-11. https://doi.org/10.1186/s12882-022-02978-2

[11] Zhao, Yijun, Man Qin, and April Jorge. "A calibrated ensemble algorithm to address data heterogeneity in machine learning: an application to identify severe SLE flares in lupus patients." IEEE Access 10 (2022): 18720-18729. https://doi.org/10.1109/ACCESS.2022.3149477

[12] Ma, W., Lau, Y. L., Yang, W., & Wang, Y. F. (2022). Random forests algorithm boosts genetic risk prediction of systemic lupus erythematosus. Frontiers in Genetics, 13, 902793. https://doi.org/10.3389/fgene.2022.902793

[13] MALARVIZHI, A. "An Ensemble-based technique (EARLNP) for recommendations of Lupus activity in Humans from kidney datasets." Journal of Positive School Psychology (2022): 7264-7273. Available at: Google Scholar

[14] BENMALEK, E., EL MHAMDI, J., JILBAB, A., & JBARI, A. (2022). A COUGH-BASED COVID-19 DETECTION SYSTEM USING PCA AND MACHINE LEARNING CLASSIFIERS. Applied Computer Science, 18(4), 96-115. https://doi.org/10.35784/acs-2022-31

[15] Gartshteyn, Yevgeniya, et al. "COVID-19 and systemic lupus erythematosus: a case series." The Lancet Rheumatology 2.8 (2020): e452-e454. https://doi.org/10.1016/S2665-9913(20)30161-2

[16] Islam, Md Touhidul, Sanjida Reza Rafa, and Md Golam Kibria. "Early prediction of heart disease using PCA and hybrid genetic algorithm with k-means." 2020 23rd International Conference on Computer and Information Technology (ICCIT). IEEE, 2020. https://doi.org/10.1109/ICCIT51783.2020.9392655

[17] Mao, Tong, and Ding-Xuan Zhou. "Rates of approximation by ReLU shallow neural networks." Journal of Complexity (2023): 101784. https://doi.org/10.1016/j.jco.2023.101784

[18] Pratiwi, Heny, et al. "Sigmoid activation function in selecting the best model of artificial neural networks." Journal of Physics: Conference Series. Vol. 1471. No. 1. IOP Publishing, 2020. https://doi.org/10.1088/1742-6596/1471/1/012010

[19] Lewkowycz, Aitor, and Guy Gur-Ari. "On the training dynamics of deep networks with $ L_2 $

regularization." Advances in Neural Information Processing Systems 33 (2020): 4790-4799. Available at: Google Scholar.

[20] Omuya, Erick Odhiambo, George Onyango Okeyo, and Michael Waema Kimwele. "Feature selection for classification using principal component analysis and information gain." Expert Systems with Applications 174 (2021): 114765. https://doi.org/10.1016/j.eswa.2021.114765

[21] Taghizadeh, E., Heydarheydari, S., Saberi, A., JafarpoorNesheli, S., & Rezaeijo, S. M. (2022). Breast cancer prediction with transcriptome profiling using feature selection and machine learning methods. BMC bioinformatics, 23(1), 1-9. https://doi.org/10.1186/s12859-022-04965-8

[22] Usategui, I., Barbado, J., Torres, A. M., Cascón, J., & Mateo, J. (2023). Machine learning, a new tool for the detection of immunodeficiency patterns in systemic lupus erythematosus. Journal of Investigative Medicine, 10815589231171404. https://doi.org/10.1177/10815589231171404

[23] Zhang, Z., Xue, Y., Slowik, A., & Yuan, Z. (2023). Sle-CNN: a novel convolutional neural network for sleep stage classification. Neural Computing and Applications, 1-16. https://doi.org/10.1007/s00521-023-08598-7

[24] Sohail, Ayesha. "Genetic algorithms in the fields of artificial intelligence and data sciences." Annals of Data Science 10.4 (2023): 1007-1018. https://doi.org/10.1007/s40745-021-00354-9

[25] Cui, Jing, et al. "Risk prediction models for incident systemic lupus erythematosus among women in the Nurses' health study cohorts using genetics, family history, and lifestyle and environmental factors." Seminars in Arthritis and Rheumatism. Vol. 58. WB Saunders, 2023. https://doi.org/10.1016/j.semarthrit.2022.152143

[26] Prabhu K, Edison. "IoT-based systemic lupus erythematosus prediction model using hybrid genetic algorithm integrated with ANN." ETRI Journal (2023). https://doi.org/10.4218/etrij.2022-0324

[27] Lazar, Stephanie, and J. Michelle Kahlenberg. "Systemic lupus erythematosus: new diagnostic and therapeutic approaches." Annual review of medicine 74 (2023): 339-352. https://doi.org/10.1146/annurev-med-043021-032611

[28] Roy, Sanjiban Sekhar, et al. "L2 regularized deep convolutional neural networks for fire detection." Journal of Intelligent & Fuzzy Systems 43.2 (2022): 1799-1810. https://doi.org/10.3233/JIFS-219281

[29] Pesqueda-Cendejas, Karen, et al. "Nutritional Approaches to Modulate Cardiovascular Disease Risk in Systemic Lupus Erythematosus: A Literature Review." Nutrients 15.4 (2023): 1036. https://doi.org/10.3390/nu15041036

[30] Usategui, I., Barbado, J., Torres, A. M., Cascón, J., & Mateo, J. (2023). Machine learning, a new tool for the detection of immunodeficiency patterns in systemic lupus erythematosus. Journal of Investigative Medicine, 10815589231171404. https://doi.org/10.1177/10815589231171404

[31] Dhanwanth, B. ., Saravanakumar, R. ., Tamilselvi, T. ., & Revathi, K. . (2023). A Smart Remote Monitoring System for Prenatal Care in Rural Areas. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 30–36. https://doi.org/10.17762/ijritcc.v11i3.6196

[32] Jóhann, Þorvaldsson, Koskinen, P., Meer, P. van der, Steiner, M., & Keller, T. Improving Graduation Rates in Engineering Programs Using Machine Learning. Kuwait Journal of Machine Learning, 1(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/110