International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 10, Issue 7, July-2025, http://ijmec.com/, ISSN: 2456-4265

Improving Medical Diagnosis Systems using Efficient Machine Learning System

Ketankumar Chaturbhai Patel
Research Scholar, Dept. of. Computer Science & Engineering, University of Technology, Jaipur, Rajasthan, India.
Prof. (Dr.) Satish Narayanrao Gujar
Dept. of. Computer Science & Engineering, University of Technology, Jaipur, Rajasthan, India.

ABSTRACT

This paper proposes a unified and mathematically grounded hybrid framework for disease prediction by integrating machine learning, applied mathematics, and fuzzy logic. The study focuses on predictive modeling of infectious and non-communicable diseases, including COVID-19, lung cancer, swine flu, and dengue, using structured clinical, epidemiological, and climatic datasets. Data preprocessing techniques such as cleaning, normalization, and train–test splitting are employed to ensure reliability and consistency. Multiple supervised learning algorithms Support Vector Machine, Random Forest, k- Nearest Neighbors, Decision Tree, and Artificial Neural Network are implemented and comparatively evaluated using accuracy, precision, recall, F1-score, mean squared error, and confusion matrices. To address uncertainty in epidemiological parameters, fuzzy mathematical modeling and bifurcation analysis are incorporated, enabling uncertainty-aware interpretation of disease dynamics through the fuzzy basic reproduction number. Experimental results indicate that Random Forest achieves the most stable and reliable performance, while SVM and ANN show competitive outcomes. Overall, the proposed hybrid analytical–ML framework enhances interpretability, robustness, and scalability, making it suitable for epidemiological studies and clinical decision support systems.

Keywords: *Machine Learning, Fuzzy Logic, Disease Prediction, Epidemiological Modeling, Clinical Decision Support.*

I. Introduction

The rapid expansion of digital technologies has transformed modern healthcare, making Machine Learning (ML) one of the most influential tools in medical science. ML enables computer systems to analyze large and complex medical datasets, detect patterns, and generate reliable diagnostic predictions with minimal human intervention. With the growth of electronic health records, medical imaging, laboratory reports, genomic data, and wearable sensor outputs, traditional diagnostic approaches often struggle to manage data complexity and scale. ML overcomes these limitations by identifying hidden relationships that may not be evident to clinicians, thus supporting early disease detection, risk assessment, and accurate prognosis. Beyond diagnosis, ML contributes significantly to treatment planning and personalized medicine by predicting treatment responses and optimizing clinical decisions. However, the increasing use of ML also raises issues related to data quality, interpretability, ethics, and clinical trust, requiring careful evaluation to ensure safe and effective implementation in healthcare.

Emergence of Machine Learning in Healthcare

Machine Learning marks a shift from experience-based to evidence-based medicine. With the digitization of healthcare data—from EHRs and diagnostic images to laboratory and genomic records—ML has become essential for processing high-dimensional information and generating clinically meaningful insights. Early ML applications relied on simple rule-based models, but advancements in computational power and deep learning have enabled sophisticated applications such as disease prediction, medical image analysis, and clinical decision-support systems. ML now enhances diagnostic accuracy, reduces human error, and facilitates preventive healthcare, making it a core component of modern medical systems.

Growth of Medical Data and Need for Intelligent Analysis

The volume of medical data has increased exponentially due to digital recordkeeping, imaging technologies, genomic sequencing, and wearable devices. Although this data has tremendous potential, its complexity makes manual interpretation slow and error-prone. Traditional rule-based analysis often fails to detect subtle patterns, leading to delayed or incorrect diagnoses. ML provides intelligent, adaptive analysis capable of processing large datasets efficiently, uncovering hidden trends, and generating accurate clinical insights. As healthcare data continues to expand, intelligent automated systems are essential for improving diagnosis, treatment decisions, and patient outcomes.

Role of Machine Learning in Medical Diagnosis

Machine Learning plays a vital role in enhancing the accuracy, speed, and efficiency of medical diagnosis. Conventional methods rely on manual interpretation of medical data, which can be time- consuming and susceptible to errors. ML techniques analyze medical images—such as X-rays, CT scans, and MRIs to detect tumors, fractures, and other abnormalities. ML models also process electronic health records and laboratory data to predict disease risks and identify high-priority patients. Natural language processing further supports diagnosis by extracting meaningful information from unstructured clinical notes. By enabling early detection, reducing diagnostic errors, and supporting evidence-based decisions, ML has become a crucial tool in modern diagnostic practice.

Applications of Machine Learning in Medical Diagnosis and Treatment

Machine Learning (ML) has become a transformative tool across multiple areas of medical diagnosis and treatment. In radiology, ML—especially deep learning techniques like CNNs—enhances the detection of lung nodules, tumors, fractures, and organ abnormalities from X-rays, CT scans, MRI, and ultrasound images, improving accuracy and reducing human error. In pathology, ML models analyze digital biopsy slides to classify tissues, detect cancer, and grade tumors with high precision. In neurology, ML processes MRI, fMRI, EEG, and EMG data to identify early signs of disorders such as Alzheimer's disease, epilepsy, and neurodegenerative conditions. For infectious diseases, ML integrates clinical, laboratory, and epidemiological data to detect outbreaks, predict severity, and support triage decisions, as seen during COVID-19. In cardiology, ML improves ECG and echocardiogram interpretation, enabling early identification of arrhythmias and heart failure. In oncology, ML assists personalized treatment planning by predicting therapy responses based on tumor and genetic profiles. ML also supports mental health diagnosis by analyzing speech, facial expressions, and behavioral patterns for early detection of disorders. In chronic disease management, ML models predict disease flare-ups using wearable and environmental data. Additionally, ML accelerates drug discovery and ICU decision-making by predicting complications such as sepsis and organ failure

.

2. Background Study on Machine Learning Applications in Medical Diagnosis

Author(s) & Year	Disease / Application Area	Methodology / Approach	Key Findings / Conclusions
Hai Minh et al. (2025)	Colorectal cancer diagnosis	Machine learning classification using medical datasets	The study reported that machine learning provided an effective and practical computational approach for colorectal cancer diagnosis by improving classification accuracy and supporting clinical decision-making.
Wang et al. (2025)	Cardiovascular disease (Point-of-care systems)	ML-assisted point- of-care diagnostic models	The authors stated that machine learning- assisted point-of-care systems significantly enhanced cardiovascular healthcare by enabling rapid detection, risk stratification, and real-time decision support.
Jdey et al. (2024)	Malaria diagnosis	Deep learning and machine learning techniques	The study noted that ML and DL methods improved malaria detection accuracy, while also highlighting implementation challenges and the need for interdisciplinary research.
Almakhzoumi et al. (2024)	Malaria detection	Machine learning with imaging datasets	The authors demonstrated that ML techniques enabled automated recognition of infected cells, supporting faster and more reliable diagnostic workflows.
Yan et al. (2023)	General medical diagnostics	Multimodal machine learning	The study explained that integrating multiple data sources through multimodal ML enhanced prediction, classification, and clinical decision support across medical domains.

Kokabi et al. (2023)	Cancer diagnosis	Biosensors integrated with ML	The authors concluded that combining biosensors with ML enabled early cancer diagnosis through real-time monitoring and accurate interpretation of biological signals.
Jameela et al. (2022)	Malaria detection	Deep learning and transfer learning	The study observed that DL and transfer learning significantly improved detection accuracy by extracting discriminative features from microscopic images.
Li et al. (2022)	Lung cancer	Machine learning on imaging, genomic, and clinical data	The authors reported that ML supported diagnosis, treatment planning, and prognosis prediction using integrated computational models.
Bhavsar et al. (2021)	General medical diagnosis	Machine learning-based decision systems	The study concluded that ML enhanced diagnostic accuracy, automated clinical tasks, and supported decision-making across diverse diseases.
Battineni et al. (2020)	Chronic disease diagnosis	ML predictive models	The authors stated that ML improved long-term disease management by analyzing patient histories and identifying risk patterns.
Yue et al. (2018)	Breast cancer	Machine learning classification algorithms	The study indicated that ML improved diagnosis and prognosis by enhancing classification accuracy and identifying malignant patterns.
Safdar et al. (2018)	Heart disease	ML-based decision support systems	The authors explained that ML strengthened heart disease diagnosis by integrating clinical data with predictive analytics.
Sumathi & Poorna (2016)	Mental health in children	Machine learning classification techniques	The study reported accurate prediction of mental health problems through analysis of behavioural and physiological indicators.
Prasad et al. (2016)	Thyroid disease	Hybrid rough set theory and ML models	The authors stated that hybrid ML architectures improved diagnosis by refining feature selection and increasing precision.
Barua et al. (2015)	Driver stress detection	Supervised ML using physiological sensors	The study found that ML algorithms effectively diagnosed stress levels by identifying physiological response patterns.
Dilsizian & Siegel (2014)	Cardiac imaging	Artificial intelligence and big data analytics	The authors emphasized that AI-driven analytics advanced personalized cardiac diagnosis through automated image interpretation and improved clinical decisions.

3. Research Methodology

This section outlines a systematic and scientifically structured research methodology adopted for the predictive modeling of infectious diseases, namely COVID-19, Lung Cancer, Swine Flu, and Dengue, using machine learning and fuzzy mathematical techniques. The proposed framework integrates multiple stages, including data acquisition, preprocessing, mathematical formulation, supervised learning model development, and rigorous performance evaluation. Emphasis is placed on transforming raw clinical and climatic data into normalized and analyzable forms to ensure consistency and accuracy in model training. It is proposed to use the trained models (including support vector machine, random forest, nearest neighbor, and decision tree algorithms which are developed based on mathematics) for comparison. Popular metrics like accuracy, fine-tuning, recall, F1 metric, mean squared error and confusion matrix are employed to estimate the quality of prediction and generalize. Also, fuzzy logic is used in combination with branching process to account for the presence of uncertainty on epidemiological parameters such as incidence and recovery rates. This fuzzy logic/machine learning hybrid methodology is interpretable, enables nonlinearity modeling and offers robustness to uncertainty. So, our proposed method yields the faithful, interpretable, and scalable data representation for disease prediction that is readily applicable to epidemiological studies or clinical decision support systems. Recent fast developments in the area of machine learning and artificial intelligence techniques enormously modified medical diagnosis and treatment planning. The medical field's growing availability of clinical, epidemiological, and environmental data has opened the door for intelligent diagnostic systems that can aid healthcare workers in making complex decision. The successful applications of these technologies, however, demand well-defined research targets in terms of clinical and mathematical rigor. The objectives of the present study are formulated in response to the limitations identified in existing literature, particularly the over-reliance on traditional diagnostic approaches, limited interpretability of machine learning models, and the lack of generalized frameworks applicable across multiple disease conditions. These challenges highlight the need for a systematic investigation into the role of machine learning techniques, supported by applied mathematical and fuzzy modelling, in improving diagnostic accuracy, robustness, and clinical usability.

3.1 Data Collection and Preprocessing

Let the dataset be represented as

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$$

where, $x_i = (x_{i1}, x_{i2}, \ldots, x_{id}) \in \mathbb{R}^d$ denotes the feature vector (clinical or climatic variables), $y_i \in \{0,1\}$ represents the class label (absence/presence of disease). Preprocessing steps include:

Data cleaning (handling missing and noisy values),

Normalization using Min–Max scaling: $x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$

Train–test split using an 80:20 ratio: $\mathcal{D} = \mathcal{D}_{train} \cup \mathcal{D}_{test}$

3.2 Mathematical Modeling of Machine Learning Algorithms

Support Vector Machine (SVM)

SVM aims to find an optimal hyperplane:

$$\mathbf{w} \cdot \mathbf{x} + b = 0$$

that maximizes the margin between two classes by solving:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2} \| \mathbf{w} \|^2 + C \sum_{i=1}^{N} \xi_i$$

subject to:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

For nonlinear data, the **RBF kernel** is used:

$$K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|^2)$$

Random Forest (RF)

Random Forest is an ensemble of decision trees: $\hat{y} = \text{mode}\{T_1(x), T_2(x), \dots, T_M(x)\}$
where each tree $T_m$is trained on a bootstrap sample. The impurity at a node is measured using Gini

Index: $G = 1 - \sum_{k=1}^{K} p_k^2.$

k-Nearest Neighbours (k- NN)

The classification is based on distance:

$$d(x_i, x_j) = \sqrt{\sum_{l=1}^{d} (x_{il} - x_{jl})^2}$$

The predicted class is

$$\hat{y} = \text{mode}\{y_j \mid x_j \in \mathcal{N}_k(x)\}$$

Decision Tree (DT)

Decision Trees recursively partition data by maximizing **Information Gain**:

$$IG(S, A) = H(S) - \sum_{v \in A} \frac{|S_v|}{|S|} H(S_v)$$

where entropy:

$$H(S) = -\sum_{i=1}^{c} p_i \log_2 p_i$$

Artificial Neural Network (ANN)

An ANN neuron output is given by:

$$y = f(\textstyle\sum_{i=1}^{n} w_i x_i + b)$$

where $f(\cdot)$is the activation function (sigmoid or ReLU).
The loss function used is Mean Squared Error (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

Weights are updated using gradient descent:

$$w^{(t+1)} = w^{(t)} - \eta \frac{\partial L}{\partial w}$$

3.3.6 Performance Evaluation Metrics

Model performance is evaluated using:

Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Fuzzy Mathematical Modeling and Bifurcation Analysis

To model uncertainty, epidemiological parameters are represented as **triangular fuzzy numbers**

$$\tilde{\beta} = (\beta_l, \beta_m, \beta_u)$$

The fuzzy basic reproduction number is defined as:

$$\tilde{R}_0 = \frac{\tilde{\beta}}{\tilde{\gamma}}$$

where $\tilde{\gamma}$ denotes the recovery rate.

If

$$\tilde{R}_0 > 1$$

the system exhibits endemic equilibrium; otherwise, a disease-free equilibrium exists. Fuzzy bifurcation analysis helps study system stability under parameter uncertainty, enhancing epidemiological interpretation pasted.

Hybrid Analytical–ML Framework

Feature importance obtained from Random Forest is integrated into analytical estimation:

$$\hat{C} = \sum_{i=1}^{d} w_i\, x_i$$

where w_i denotes feature importance weight. This hybrid framework enables interpretable forecasting with improved robustness.

4. Simulation and Result Analysis

This section presents a comprehensive implementation and relative evaluation of multiple machine learning algorithms used for COVID-19 predictive modelling, supported by mathematical formulation, experimental metrics, and analytical extensions. The discussion integrates classical ML models with fuzzy logic–based bifurcation analysis to strengthen epidemiological interpretation.

Algorithmic Performance Analysis

Five models Random Forest, k-Nearest Neighbours, Decision Tree, Artificial Neural Network (ANN), and Support Vector Machine (SVM) were implemented and evaluated using standard classification metrics.

- Random Forest achieved the highest overall presentation with an accuracy of 60.50% and balanced precision, recall, and F1-score (0.55, 0.54, 0.54). The confusion matrix shows stronger identification of Class-1 cases, indicating robustness of ensemble learning in handling feature interactions.
- k-NN and Decision Tree exhibited comparatively lower accuracies (53.50% and 51.00%), struggling particularly with Class-2 prediction, likely due to class imbalance and limited feature separability.
- ANN delivered moderate accuracy (58.50%) but showed signs of overfitting, as reflected by low training MSE and higher validation/test errors.
- SVM (RBF kernel) performed competitively (59.00% accuracy), benefitting from nonlinear kernel mapping, though misclassification of Class-2 remained notable.

Fuzzy Bifurcation and Epidemiological Insight

The fuzzy bifurcation analysis incorporated uncertainty in epidemiological parameters (infection, recovery, and death rates). The fuzzy basic reproduction number $\tilde{R}_0 = (1.24, 1.67, 2.32)$ remained greater than 1 across all fuzzy intervals, indicating persistent endemic behaviour. This confirms that under current parameter uncertainty, the system does not transition to a disease-free equilibrium.

Analytical Estimation Framework

Using Random Forest feature importance weights, an analytical estimation predicted 1,800 confirmed cases for a sample scenario. This hybrid ML-analytical approach enables rapid forecasting while maintaining interpretability.

Random Forest Regression

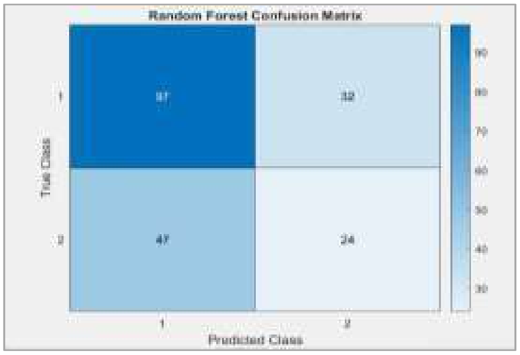

Fig. 1: Random Forest Confusion Matrix

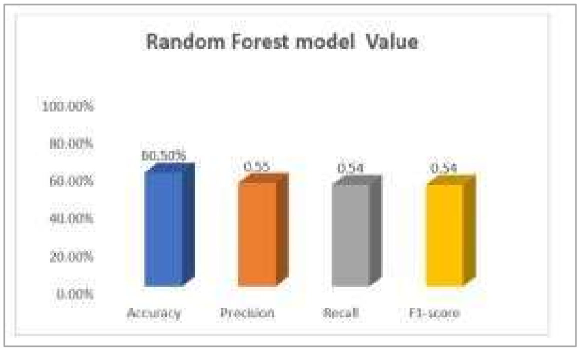

Fig. 2: Random Forest Model Value

The Random Forest model demonstrates moderate predictive presentation in the given classification task. The general accuracy of 60.50% can be interpreted that the model is able to correctly categories a little bit higher than half of all cases, it implies acceptable but not very robust performance. As for the precision, we see that it is 0.55, which means that about 55% of the instances guessed as positive are correct which clearly indicates a medium reliability level on predicting positives. The recall of 0.54 also tells that the model can identify around 54 % of all real positive cases, however this again shows there are still a significant number which are not identified true cases. Also, the low F1-score (0.54), as this is balanced between precision and recall, reinforces that moderate effectivity.

The confusion matrix reveals that the model performs better in identifying **Class 1** compared to **Class 2**, with a higher number of true positives for Class 1. However, misclassification of Class 2 remains significant, indicating scope for improvement through feature enhancement, class balancing, or hyperparameter tuning.

k-Nearest Neighbours (k-NN)

Fig. 3: k-NN Confusion Matrix

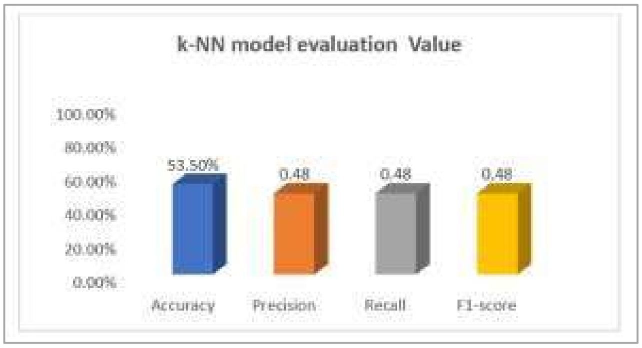

Fig. 4: k-NN Model Evaluation Value

The k-Nearest Neighbours (k-NN) model exhibits **relatively low predictive presentation** in the given classification task. This is translated as the model classifying instances correctly just over half of the time, and hence only a marginal improvement on random guessing. The model performs poorly, as indicated by the precision and recall values of 0.48 as well as F1-score values 2 ≈ below (Table III). The precision of 0.48 means that less than half of the examples we predicted are positive (were correct), and the recall of 0.48 shows us that our model misses over half the actual positives! The non symmetrical feature of false positives opposed to false negatives is evidenced as well by the F1-score. As can be seen from the confusion matrix, model had less accuracy in predicting class 2 compared to class 1 and it frequently made wrong predictions as Class 2 irrespective of correct one. This might be attributed to their similarity of features, data scarcity and class imbalance. In general, these results suggest that k-NN need to be optimized, feature scaled or distance metrics other than the Euclidean would provide a better performance on classification.

Decision Tree Regression

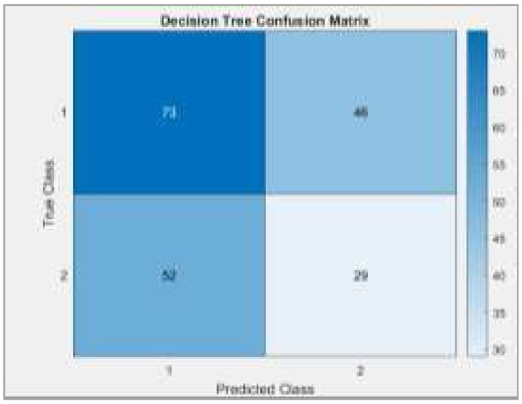

Fig. 5: DT Confusion Matrix

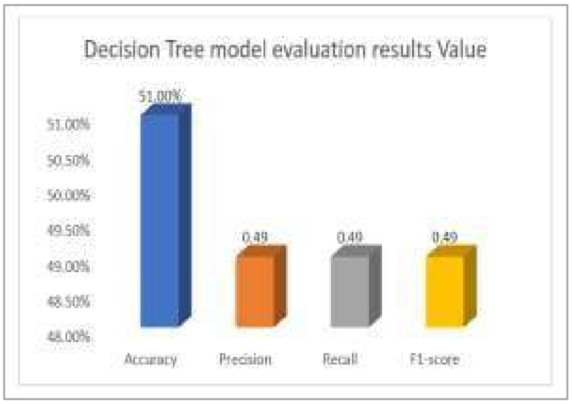

Fig. 6: Decision Tree Model Evaluation Results Value

The Decision Tree model shows **limited predictive capability** for the given classification task. The global performance, accuracy = 51.00%, of the CR model shows that the CR classification correctly identifies a little better than half of all cases (only slightly better than chance). The Precision, Recall, and the F1-score of 0.49 are additional indicators of a weak and unstable performance on this dataset. 0.49 precision means that about half of our positive predictions are wrong and 0.49 recall says that we miss a lot of positive cases. The balanced and low F1-score indicates that the model does not handle both false positives and false negatives well.

The confusion matrix shows serious misclassification in two classes. While the model is able to distinguish Class 1 a bit better than Class 2, it still predicts many of the samples belonging to both classes incorrectly. This indicates the Decision Tree tends to overfit or is not able to capture complex relations. In general, the results suggest pruning or hyperparameter tuning or adopting an ensemble system in order to achieve good performance.

Artificial Neural Network (ANN)

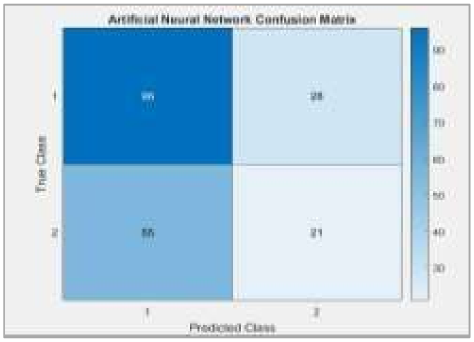

Fig. 7: ANN Confusion Matrix

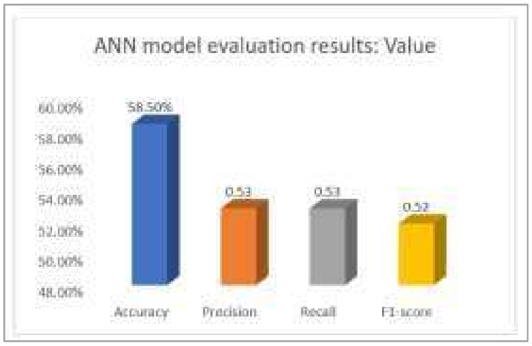

Fig. 8: ANN Model Evaluation Results: Value

The Artificial Neural Network (ANN) model demonstrates moderate predictive performance in the classification task. The performance score of 58.50% demonstrates the model's ability to correctly predict just over a half of the instances, and is an improvement compare to compeller models as k-NN Ans Decision Tree. The precision and recall of 0.53 indicates that the test performs with moderate accuracy, in a sufficiently optimal way so as to positively identify positive cases while maintaining an acceptable degree of correctness in predictions. The F1-score of 0.52, which is a harmonic average of the precision and recall, also supports the moderate performance of this model.

The confusion matrix indicates that ANN achieves high true positive rate in recognising Class 1, while the misclassification of Class 2 is still significant. This imbalance suggests that the model is biased towards Class 1 patterns and fails to generalize equally in both classes. While ANN can fit non-linear relationships from data, the findings indicate what might be overfit or lack of proper feature representation. Further performance enhancements could have been realized through architectural tuning, regularization, or a larger training set.

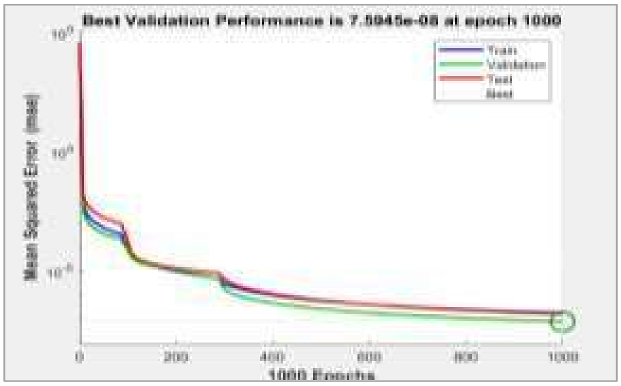

Fig. 9: Mean Squared Error (MSE)

The Mean Squared Error (MSE) plot illustrates the learning behaviour of the Artificial Neural Network across training, validation, and test datasets over 1,000 epochs. MSE decrease rapidly in the first epochs, showing that learning is good and prediction errors are quickly decreasing. Towards the end of training, all three curves eventually converge, indicating that learning is stable and generalizes well. A validation performance of 7.5945 ×10^{-8} at epoch 1000 indicates good convergence with low error. The convergence of training, validation and test curves close to each other shows that the model is not overfitting heavily, and suggests strong predictive reliability on unobserved data.

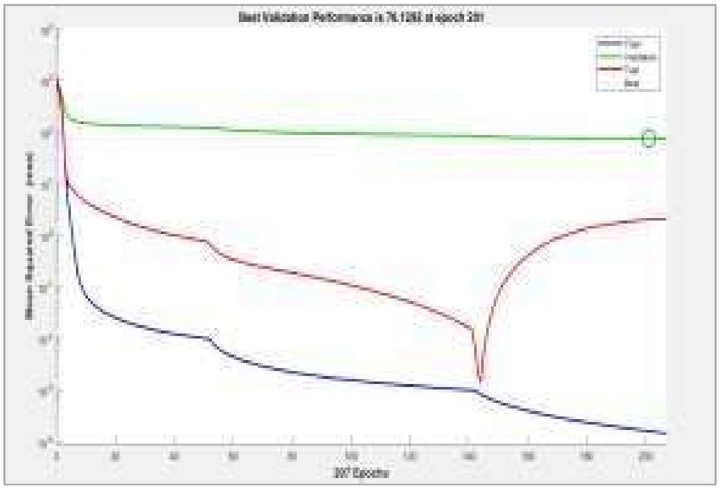

Fig. 10: Mean Squared Error (MSE) Changes Over 207 Epochs

The Mean Squared Error (MSE) plot over 207 epochs illustrates the learning and generalization behaviour of the model across training, validation, and test datasets. The training MSE reduces so quickly that the learning is efficient and fits well to the data. Nevertheless, the validation means squared error is still high and does not decrease much. The test means squared error (MSE) variation is quite prominent, and it clearly has a sudden minimization around epoch 140, then grows back quickly evidence of over-fit to unseen observations. The peak of the validation performance 76.1292 at epoch 201 indicates that while training keeps growing, it overfits showing that either regularization or better model tuning are needed.

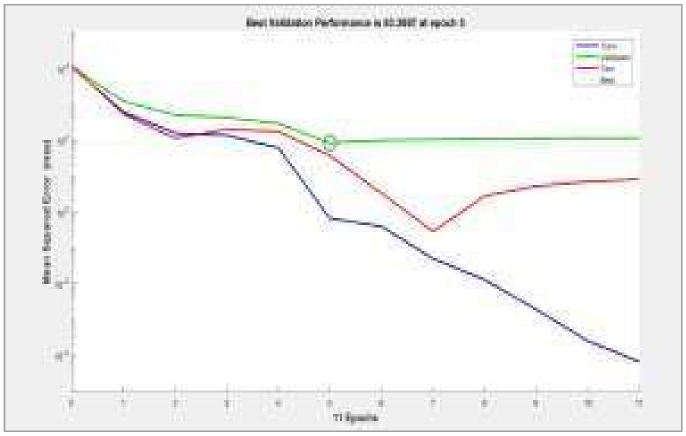

Fig. 11: Mean Squared Error (MSE) for Training

The Mean Squared Error (MSE) plot over 11 epochs illustrates the learning behaviour of the model during early training. The training MSE reduces rapidly, which means the model fast captures patterns in the Web logs. The validation MSE reaches its minimum of 83.3087 in epoch 5, and from this point on it stays about the same or rises slightly indicating the beginning of overfitting. In the meantime, the test MSE decreases initially and increases then to indicate limited generalization to other unseen samples. Such a mismatch between training and validation/test errors were reported to originate from the overtraining of NN beyond early epochs, and are suggestive to the significance of regularization and early stopping.

Support Vector Machine (SVM) Regression

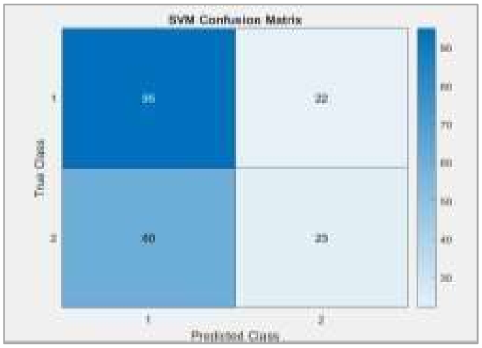

Fig 12: SVM Confusion Matrix

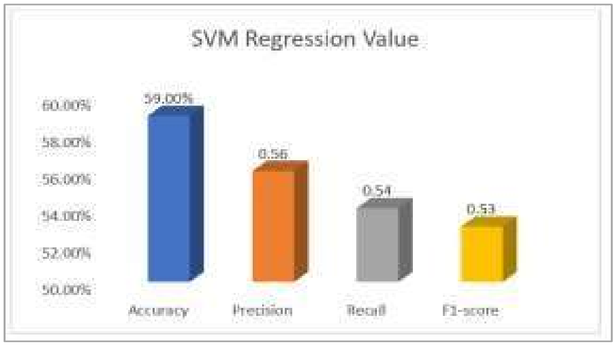

Fig. 13: SVM Regression Value

The Support Vector Machine (SVM) model demonstrates moderate classification performance. An overall accuracy of 59.00% implies that the model can classify just under 60% of all instances correctly. The precision of 0.56 means 56% of the observations predicted as positive were correct. On the other hand, recall of 0.54 means that if we have 1 positive sample in test set our model can identify ~54% of them , which is just a bit higher than average. The F1-score of 0.53, which is an average of the two former measures, validates such performance and highlights that our model performs moderately. The confusion matrix indicates that SVM can predict Class 1 objects better than Class 2 objects, and misclassified the Class 2 samples more often. Altogether, the SVM seems to work quite well, however additional optimization or better feature-set might be required to boost up its classification robustness.

5. Conclusion and Future Scope

The present study demonstrated that integrating machine learning with applied mathematical and fuzzy logic–based modeling provides an effective and interpretable framework for disease prediction across multiple medical conditions. The comparative analysis of supervised learning algorithms revealed that ensemble-based approaches, particularly Random Forest, offer greater stability and robustness in handling complex and uncertain medical datasets. The incorporation of fuzzy mathematical modeling and bifurcation analysis enhanced epidemiological interpretation by explicitly accounting for uncertainty in key parameters such as infection and recovery rates. This hybrid analytical–ML framework addressed major limitations of conventional diagnostic systems, including limited interpretability and sensitivity to data variability, thereby improving clinical usability and decision support potential. Future research may focus on extending the framework to larger, real-time, and multi-center datasets to improve generalizability. The integration of deep learning architectures, temporal models, and real-time sensor data can further enhance predictive accuracy. Additionally, incorporating explainable AI techniques and advanced uncertainty quantification methods may strengthen transparency, trust, and adoption in clinical and public health applications.

References

1. Hai Minh, N., Quy, T. Q., Tam, N. D., Tuan, T. M., & Son, L. H. (2025). A practical approach for colorectal cancer diagnosis based on machine learning. *PLoS One*, *20*(4), e0321009.
2. Wang, K., Tan, B., Wang, X., Qiu, S., Zhang, Q., Wang, S., ... & Yu, Y. (2025). Machine learning-assisted point-of-care diagnostics for cardiovascular healthcare. *Bioengineering & Translational Medicine*, e70002.
3. Jdey, I., Hcini, G., & Ltifi, H. (2024). Deep learning and machine learning for Malaria detection: overview, challenges and future directions. *International Journal of Information Technology & Decision Making*, *23*(05), 1745-1776.
4. Almakhzoumi, A., Bonny, T., & Al-Shabi, M. (2024, June). Malaria detection using machine learning. In *Optics, Photonics, and Digital Technologies for Imaging Applications VIII* (Vol. 12998, pp. 259-268). SPIE.

5. Yan, K., Li, T., Marques, J. A. L., Gao, J., & Fong, S. J. (2023). A review on multimodal machine learning in medical diagnostics. *Math. Biosci. Eng*, *20*(5), 8708-8726.
6. Kokabi, M., Tahir, M. N., Singh, D., & Javanmard, M. (2023). Advancing healthcare: synergizing biosensors and machine learning for early cancer diagnosis. *Biosensors*, *13*(9), 884.
7. Jameela, T., Athota, K., Singh, N., Gunjan, V. K., & Kahali, S. (2022). Deep learning and transfer learning for malaria detection. *Computational intelligence and neuroscience*, *2022*(1), 2221728.
8. Li, Y., Wu, X., Yang, P., Jiang, G., & Luo, Y. (2022). Machine learning for lung cancer diagnosis, treatment, and prognosis. *Genomics, proteomics & bioinformatics*, *20*(5), 850- 866.
9. Bhavsar, K. A., Abugabah, A., Singla, J., AlZubi, A. A., & Bashir, A. K. (2021). A comprehensive review on medical diagnosis using machine learning. *Computers, Materials and Continua*, *67*(2), 1997.
10. Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). Applications of machine learning predictive models in the chronic disease diagnosis. *Journal of personalized medicine*, *10*(2), 21.
11. Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine learning with applications in breast cancer diagnosis and prognosis. *Designs*, *2*(2), 13.
12. Safdar, S., Zafar, S., Zafar, N., & Khan, N. F. (2018). Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. *Artificial Intelligence Review*, *50*(4), 597-623.
13. Sumathi, M. R., & Poorna, B. (2016). Prediction of mental health problems among children using machine learning techniques. *International Journal of Advanced Computer Science and Applications*, *7*(1).
14. Prasad, V., Rao, T. S., & Babu, M. S. P. (2016). Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms. *Soft Computing*, *20*(3), 1179-1189.
15. Barua, S., Begum, S., & Ahmed, M. U. (2015). Supervised machine learning algorithms to diagnose stress for vehicle drivers based on physiological sensor signals. In *pHealth 2015* (pp. 241-248). IOS Press.
16. Dilsizian, S. E., & Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current cardiology reports*, *16*(1), 441.