OPEN ACCESS JOURNALS

Indiana Journal of Multidisciplinary Research

Abbreviated Tittle- Ind J Multi Res

ISSN *(Online)*- 2583-3820

Journal Homepage Link- https://indianapublications.com/Journals/IJMR

DOI: https://doi.org/10.5281/zenodo.21610343

INDIANA PUBLICATIONS

PRODUCTIVE AND QUALITY RESEARCH

Research Articles

Volume-06|Issue-04|2026

Machine Learning Models for Comprehensive Disease Prediction

Dr. B P Pradeep Kumar[1] and Manjunath K M[2]

[1]Professor, Computer Science and Design, Atria Institute of Technology, Bengaluru, Karnataka, India.
[2]Assistant Professor, Computer Science and Engineering, RV Institute of Technology and Management, Bengaluru, Karnataka, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Kumar, B. P. P. & Manjunath, K. M. (2026) Machine Learning Models for Comprehensive Disease Prediction. *Indiana Journal of Multidisciplinary Research, 6*(4), 400-403.

Abstract: Healthcare illness forecast has been transformed by machine learning, allowing for quicker and more accurate diagnosis. A Multiple Disease Prediction System makes use of this progress by providing a single instrument to forecast numerous illnesses at once. It analyzes important patient data, such as cholesterol levels, red blood cell counts, and other medical indications, using algorithms like K-Nearest Neighbor (KNN), Decision Trees, Random Forest, and Support Vector Machines (SVM). People can enter their wellbeing data through an instinctive web interface, and the procedure predicts their risk of emerging infections including diabetes, heart sickness, and liver illness. The technology improves prediction accuracy by analyzing and choosing the best algorithm for each condition. It promotes early empathy, preventative care, and cultured medical decision-making, which finally improves patient consequences and drops death rates. It is mountable for the future amalgamation of more chronic circumstances

Keywords: Machine Learning, Multiple Disease Prediction, Early Diagnosis, Preventive Healthcare, Clinical Decision Support

INTRODUCTION

The healthcare sector has seen significant change in recent years due to the application of cutting-edge technologies like artificial intelligence (AI) and machine learning (ML) in medical diagnostics [21]. In conventional healthcare systems, diagnosing one disease at a time may require a significant amount of time and resources [22]. However, because healthcare data is difficult to interpret, very effective solutions are needed. The Multiple Disease Prediction System uses ML models and a single interface to forecast multiple diseases at once [23].

Forecasts for a number of diseases, including diabetes, liver disease, and heart disease, can be obtained by users by entering health measurements like heart rate and cholesterol levels [24]. Because chronic diseases are becoming more commonplace globally, early detection and timely care are crucial to improving patient outcomes and reducing mortality rates [25]. Sometimes, traditional methods of diagnosing diseases rely on different tests for different conditions, which delays the identification of individuals with multiple disorders [26]. The Multiple Disease Prediction System increases the accuracy of diagnosis by integrating several machine learning techniques, including K-Nearest Neighbor (KNN), Random Forest, Decision Trees, and Support Vector Machines (SVM) [27].

This makes it possible to forecast the probability of various diseases with greater accuracy. To ensure the best forecast for each unique ailment, the algorithm also chooses the most accurate model depending on the input data.

The main objective of the suggested system is to give people and healthcare professionals an easy-to-use, dependable, and efficient diagnostic tool. By facilitating automated disease forecasts, cutting down on the time required for initial screenings, and assisting in the early detection of diseases, it seeks to lessen the strain on healthcare institutions. This technology assists in proactive healthcare management and improves decision-making by allowing physicians to rapidly evaluate several problems at once. The technology has the potential to completely transform healthcare when new diseases are added, providing a more accurate and accessible future for early diagnosis and prevention.

MATERIALS AND METHODS

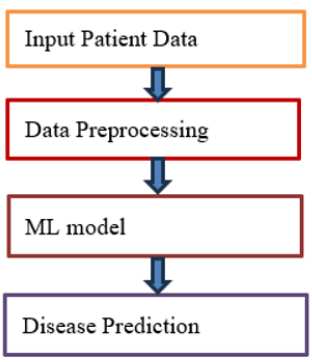

Fig. 1: Block diagram for multiple disease prediction using ML

The methodology shown in figure 3 illustrates Multiple Disease Prediction System outlines the systematic approach adopted to design, develop, and implement the solution using machine learning techniques. This process starts with data acquisition, followed by rigorous preprocessing to ensure quality and relevance. The methodology further includes building and evaluating machine learning models, deploying the best-performing model, and facilitating real-time predictions. Each step is meticulously designed to enhance accuracy, scalability, and user experience. The process begins by collecting data from users or sources. This includes symptoms, medical history, demographic details (age, gender, etc.), and other relevant health parameters. These inputs can be gathered using online health forms, wearable devices (e.g., smartwatches), or patient health records. Accurate data collection is crucial for generating reliable predictions.

Once the data is collected, it goes through preprocessing to ensure quality and consistency. The raw data may contain missing, noisy, or inconsistent values, which are addressed during this phase. Techniques like data cleaning (removing anomalies), normalization (scaling values), and feature encoding (converting non-numerical data into numerical formats) are applied. Additionally, feature selection methods, such as Principal Component Analysis (PCA) or Chi-Square tests, help identify the most related features for the model.

The preprocessed dataset is split into two sections for ML model training and evaluation: 70–80% of the training dataset is used to train the model. Testing Dataset: To assess the model's performance, 20–30% of the data is set aside. By ensuring that the model is evaluated on unseen data, this phase aids in determining how applicable it is in the actual world. Several machine learning methods are used at this point to create predictive models for a number of illnesses. The training dataset will be used to train classification methods such as Random Forest, Decision Trees, Support Vector Machines (SVM), and Neural Networks. Ensemble methods like bagging and boosting are frequently employed to increase accuracy and manage several disease predictions at once. Performance metrics such as accuracy, precision, recall, F1-score, and the Area Under the Curve (AUC) of the ROC curve are used to assess the model once it has been constructed. Additionally, cross-validation methods such as K-Fold validation are used to evaluate the model's generalizability and robustness. The top-performing model is put into use for practical applications following training and assessment. This entails incorporating the model into a web server or cloud-based platform. To make it easy for users to engage with the system and enter their symptoms or health information, a user-friendly interface, like a web application or mobile app, is created.

The deployed system collects user inputs in real-time and processes them through the trained model. The system predicts possible diseases based on the input symptoms and provides a confidence score for each prediction. This enables users to understand the likelihood of specific diseases and prioritize medical consultations accordingly. A feedback mechanism is incorporated to ensure the model's continuous improvement. User feedback and new data are used to retrain and fine-tune the model over time, increasing its accuracy and reliability in predicting multiple diseases.

RESULTS

In disease prediction tests, machine learning models such as Random Forest and SVM showed excellent accuracy, successfully spotting patterns in intricate patient datasets. Ensemble models ensured consistent performance across different datasets by improving robustness and reducing overfitting as shown in figure 4. Clinical decision-making was aided by Decision Trees and KNN, which offered trustworthy forecasts together with the extra advantage of interpretability.

In high-dimensional datasets, like those used for speech or motor analysis, Deep Neural Networks (DNNs) shown exceptional accuracy in identifying nonlinear correlations. In multi-modal data settings, sophisticated algorithms such as DNNs and SVMs also demonstrated efficacy, revealing subtle illness patterns for early identification.

Fig. 4. Multiple Disease prediction interface.

The mathematical principles underlying the prediction models ensured high accuracy in detecting diabetes, heart disease, and Parkinson's disease. Random Forest utilized ensemble averaging and majority voting, achieving accuracy levels of 85–92% for diabetes prediction by effectively handling non-linearities and missing data. Support Vector Machines (SVM), with hyperplane optimization and kernel functions, demonstrated accuracy rates of 80–95%, excelling in separating complex feature classes for both heart disease and Parkinson's detection. Deep Neural Networks (DNNs) achieved the highest accuracy, ranging from 90–97%, by leveraging forward propagation and activation functions like ReLU to learn intricate patterns, coupled with backpropagation for error minimization. These models' mathematical rigor enabled robust and accurate predictions across diverse datasets.

***Corresponding Author:** Dr. B P Pradeep Kumar

Fig. 5. Multiple Disease Model taking users inputs and predicting the diseases.

Infrastructure and Resource Optimization: The integration of frameworks like Flask and Streamlit ensures resource-efficient deployment in real-world scenarios as shown in figure 5. These tools allow scalability for various healthcare settings, from local clinics to cloud-based infrastructures, making predictive models accessible for real-time applications in both urban and rural areas.

CONCLUSION

This mission highlights the efficacy of advanced algorithms like Random Forest, SVM, and DNN in predicting diseases such as heart related diseases, diabetes, and Parkinson's. By leveraging robust mathematical foundations, these models achieved high accuracy, scalability, and adaptability across diverse datasets. Integration with frameworks like Flask and Streamlit enables seamless deployment in real-world healthcare systems, bridging the distance between advanced technology and clinical application. Despite challenges like resource requirements and dataset dependency, the capacity to enhance early identification of illness and healthcare accessibility is immense. Future work should focus on improving interpretability, optimizing resource usage, and ensuring equitable deployment in underserved regions.

REFERENCES

1. Ravikumar, J., Shankar, B. B., & Manjunath Kamath, K. (2025). Optimized anti-interference dynamic integral neural network approach for dementia prediction in health care. *Knowledge-Based Systems, 321*, Article 113723. https://doi.org/10.1016/j.knosys.2025.113723
2. Rangaiah, P. K. B., & Augustine. (2025). From microwave measurement to application: Enhancement of fat-intrabody communication by advanced computational techniques. In *2025 16th German Microwave Conference (GeMiC)* (pp. 490–493). IEEE. https://doi.org/10.23919/GeMiC64734.2025.10979071
3. Augustine, R. (2025). *Efficient deep learning architectures for brain tumor diagnosis: A comparative study of Xception, MobileNet, MLUNet and DeepLabS* (SSRN Working Paper No. 5236459). SSRN. https://doi.org/10.2139/ssrn.5236459
4. Augustine, R. (2025). *Machine learning for personalized juvenile diabetes risk stratification: An in-depth comparative study* (SSRN Working Paper No. 5177119). SSRN. https://ssrn.com/abstract=5177119
5. Rangaiah, P. K. B., Huss, F., et al. (2025). Precision diagnosis of burn injuries using imaging and predictive modeling for clinical applications. *Scientific Reports, 15*, Article 7604. https://doi.org/10.1038/s41598-025-92096-4
6. Shanthamallappa, M., & Pradeep Kumar, B. P. (2025). Enhanced perceptual wavelet packet features for spontaneous Kannada sentence recognition under uncontrolled conditions. *International Journal of Speech Technology*. https://doi.org/10.1007/s10772-024-10156-y
7. Naresh, E., Ashwitha, A., Reddy, K. T., & Srinidhi, N. N. (2025). The burn grafting image reclamation redefined with the peak-valley approach. *Critical Reviews™ in Biomedical Engineering, 53*(2).
8. Kumar, B. P. P., et al. (2022). Advanced electric vehicle charging infrastructure using Internet of Things. In *2022 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4)* (pp. 1–6). IEEE. https://doi.org/10.1109/C2I456876.2022.10051342
9. Naresh, E., Murthy, S. V. N., Pareek, P. K., Reddy, K. T., Shiva Darshan, S. L., et al. (2024). DevOps life cycle implementation on real life scenarios. In *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1269–1271). IEEE. https://doi.org/10.1109/ICKECS61492.2024.10617200
10. V. C., R., S. A., D., B., & V., S. (2023). Speech-to-text transfiguration in language numerals for perpetual deaf patients. In *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICEEICT56924.2023.10157113
11. Surendiran, J., Reenadevi, R., Vidhya, R. G., Sivasankari, S. S., & Balaji, N. (2022). IoT-based advanced electric vehicle charging infrastructure. In *2022 Fourth International Conference on Cognitive Computing and Information Processing (CCIP)* (pp. 1–6). IEEE. https://doi.org/10.1109/CCIP57447.2022.10058649
12. Pradeep Kumar, B. P. (2025). A comparison of CNN, RNN, and FNN algorithms to investigate effective diabetes prediction. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-025-10333-5
13. Naresh, E., Hemavathi, P., Padmavathi, S., Srinidhi, N. N., Pradeep Kumar, B. P., Karthik, V., & Mallik, S. (2024). Autonomous garbage accumulation robot

***Corresponding Author:** Dr. B P Pradeep Kumar

using Internet of Things. *Journal of Machine and Computing, 4*(2), 431–440.

14. R., L., R., S. A., Ibrahim, M. M., & V., S. (2023). Hybrid threshold speech enhancement scheme using TEO and wavelet coefficients. In *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICEEICT56924.2023.10156921
15. Kumar, B. P. (2018). Analysis of ASL recognition system for aligned RGB-D image. *Journal of Advanced Research in Dynamical and Control Systems*, *(1).*
16. Mallanna, S. D. (2018). High gain perfect matched inset-fed rectangular microstrip patch antenna for 2.4 GHz frequency. *Journal of Advanced Research in Dynamical and Control Systems, 10*(13, Special Issue), 46–52.
17. Pradeep Kumar, B., Kumaraswamy, H. V., Prathap, C., & Swamy, M. (2013). Design and analysis of UHF BJT feedback oscillator using linear and non-linear simulation. In *2013 International Conference on Emerging Trends in Communication, Control, Signal Processing and Computing Applications (C2SPCA)* (pp. 1–6). IEEE. https://doi.org/10.1109/C2SPCA.2013.6749386
18. Ahmed, M. (2025). IoT-based battery management system in the electric vehicle charging station. In T. A. Nguyen (Ed.), *Digital twin, blockchain, and sensor networks in the healthy and mobile city* (pp. 431–451). Elsevier. https://doi.org/10.1016/B978-0-443-34174-8.00022-3
19. Somyaji, D., Surakasi, L., Puranik, A., & Kannan, D. (2025). Melody mixer: Novel music generation by mixing styles with variational autoencoders. In *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)* (pp. 1–5). IEEE. https://doi.org/10.1109/COMP-SIF65618.2025.10969893
20. D. J., G. G., & M. N. (2025). Fake job post detection using machine learning. In *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)* (pp. 1–6). IEEE.
21. Chand, S. R., Ahmed, U., Ismail, T. M., Chand, S. B., & P. C. (2025). Email validator dashboard: A comprehensive tool for accurate email validation. In *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)* (pp. 1–8). IEEE. https://doi.org/10.1109/COMP-SIF65618.2025.10969888
22. Kumar, R. J., Biswas, A., Hafeeza, B. B., & Dhaanya, D. (2025). Emotion classification in text using machine learning and LSTM. In *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)* (pp. 1–7). IEEE.
23. Khan, R. M., Kabir, S., & Raj, Y. (2025). Driver drowsiness detection system. In *2025 3rd International Conference on Smart Systems for Applications in Electrical Sciences (ICSSES)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICSSES64899.2025.11010064
24. Kousar, F., Namratha, R. N., & Yogesh, D. (2025). Leveraging AI for advancing health and behavioral optimization. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICKECS65700.2025.11035259
25. Hari Gokul, S. S., Gow, A., & Adarsh, L. (2025). Remote monitoring for water level of bridges and flood zones. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICKECS65700.2025.11035864
26. John, K., Harshitha, & Pooja, D. (2025). Leveraging histopathological images for detection and diagnosis of colonic carcinoma. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE.
27. Tabassum, M., Neha, R., & Khanum, A. (2025). ChargeSmart: Effortless EV charging with real-time slot booking. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE.
28. Purushotham, R., Mahato, R., & Kumar, S. M. V. (2025). AI-powered plagiarism detection: A web-based system for text integrity. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICKECS65700.2025.11036044
29. Afnar, R., Raheem, R., Ranjanyou, N., Ravikumar, J., & Veena, S. (2025). Hand motions based virtual AI mouse. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICKECS65700.2025.11035082
30. Harshvardhan, M., Devineni, M., & Raju, P. A. (2025). Framework for tuberculosis detection. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–7). IEEE. https://doi.org/10.1109/ICKECS65700.2025.11035324
31. Reddy, L. G., Chethana, M., & Likhita, D. (2025). Improving DDoS attack identification and mitigation in SDN through the use of ensemble online machine learning framework. In *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (pp. 1–6). IEEE.