ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206 **Volume 8, Issue 7, July 2025**

© 2025 IJMRSET | Volume 8, Issue 7, July 2025| DOI:10.15680/IJMRSET.2025.0807060

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI-Based Multi Disease Prediction using Machine Learning

Srinidhi Hamse[1], Hanock Jacob A[2]

Dept. of Computer Science, St. Philomena's College (Autonomous) College of Excellence, Mysore, Karnataka, India[1]

Assistant Professor, Dept. of Computer Science, St. Philomena's College (Autonomous) College of Excellence, Mysore, Karnataka, India[2]

ABSTRACT: In recent years, the integration of Artificial Intelligence (AI) with healthcare has led to the development of advanced diagnostic tools aimed at improving early disease detection and patient outcomes. This project presents an AI-based multi-disease prediction model utilizing Machine Learning (ML) techniques to identify the likelihood of multiple diseases such as diabetes, heart disease, liver disorders, and kidney diseases. The system leverages clinical and demographic datasets, which are preprocessed to handle missing values and outliers. Multiple ML algorithms, including Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Logistic Regression, are trained and evaluated to select the most accurate model. The proposed system is designed to assist medical professionals in making faster, data-driven decisions and to serve as a preventive tool for early diagnosis, ultimately reducing the burden on healthcare systems. Experimental results demonstrate that the proposed model achieves high accuracy and robustness in predicting multiple diseases simultaneously.

KEYWORDS: AI in Healthcare, Multi-Disease Prediction, Machine Learning, Medical Diagnosis, Random Forest, SVM, KNN, Logistic Regression, Health Informatics, Clinical Data Analysis.

I. INTRODUCTION

Artificial Intelligence (AI) has revolutionized the healthcare sector by enabling faster and more accurate disease detection. Multi-disease detection refers to the capability of a system to diagnose multiple health conditions simultaneously. Machine Learning (ML), a subset of AI, plays a crucial role in training models using patient data, clinical records, and medical images. By analyzing large datasets, ML algorithms can identify complex patterns and correlations that may not be visible to human clinicians. These systems assist in the early diagnosis of diseases such as diabetes, heart disease, cancer, and respiratory conditions. The automation of diagnosis reduces human error and enhances efficiency in healthcare delivery. Such AI-based systems are particularly valuable in areas with limited access to specialists. They can also continuously learn and improve over time with more data. Integrating AI with electronic health records (EHRs) provides a robust framework for preventive care. Ultimately, AI-driven multi-disease detection promotes proactive healthcare, improving patient outcomes and optimizing medical resources.

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized the healthcare industry by enabling more accurate, faster, and cost-effective disease diagnosis. One of the most impactful innovations is AI-based multi-disease prediction systems, which use ML algorithms to analyze patient data, identify patterns, and predict the likelihood of various diseases. These systems can simultaneously assess the risk of multiple health conditions by learning from historical medical records, laboratory results, and symptoms, thereby assisting doctors in early diagnosis and decision-making.

Traditional diagnostic methods often focus on identifying a single disease, requiring multiple tests and consultations to uncover co-existing or hidden conditions. In contrast, AI-driven multi-disease prediction models improve efficiency by offering a unified platform that can screen for numerous illnesses at once, such as diabetes, heart disease, kidney disorders, and respiratory conditions. By leveraging large datasets and advanced algorithms, these models not only enhance diagnostic accuracy but also play a vital role in preventive healthcare and personalized treatment planning.

© 2025 IJMRSET | Volume 8, Issue 7, July 2025| DOI:10.15680/IJMRSET.2025.0807060

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. OBJECTIVES

1. To develop an AI-driven system capable of detecting multiple diseases from medical data using machine learning algorithms.
2. To enhance diagnostic accuracy and speed by leveraging data-driven predictive models.
3. To support healthcare professionals in early and efficient disease identification for improved patient outcomes.

III. LITERATURE SURVEY SUMMARY

Recent advancements in artificial intelligence (AI) and machine learning (ML) have significantly transformed the healthcare industry, particularly in the domain of multi-disease prediction. Several studies have demonstrated the effectiveness of ML algorithms in analyzing medical data to predict the onset of various diseases simultaneously. Techniques such as Decision Trees, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and ensemble models have been widely applied to datasets consisting of patient records, lab test results, and lifestyle parameters. Researchers have also explored the use of deep learning approaches like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for image-based diagnostics and sequential health data analysis, enhancing the prediction accuracy and early diagnosis capabilities.

Moreover, literature indicates that integrating multiple features such as symptoms, demographics, genetic history, and real-time vitals improves prediction accuracy for diseases like diabetes, heart disease, liver disease, and respiratory illnesses. Hybrid models combining feature selection methods with advanced classifiers have shown promising results in reducing false positives and negatives. In addition, the use of cloud-based frameworks and Internet of Things (IoT) integration has enabled real-time data acquisition and processing, leading to more robust and scalable solutions. Overall, the literature underscores the growing importance of AI-driven multi-disease prediction systems in supporting early diagnosis, personalized treatment, and reducing the burden on healthcare systems.

IV.ALGORITHAM INFORMATION

In the project various supervised learning algorithms are employed to predict the presence of multiple diseases based on patient health data. Algorithms such as **Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Naive Bayes** are commonly used in this domain. These algorithms are trained on medical datasets containing features like age, gender, symptoms, test results, and previous medical history. The models learn to identify patterns and correlations between these features and specific diseases, enabling early and accurate predictions. Ensemble techniques such as **Voting Classifier or Stacking** are often utilized to improve prediction accuracy by combining the strengths of individual models. Feature selection methods like **Recursive Feature Elimination (RFE)** and **Principal Component Analysis (PCA)** are also integrated to reduce noise and improve model performance. The system typically uses a user-friendly interface where a user inputs symptoms or medical details, and the backend ML model processes the data to provide possible disease predictions. These predictions assist doctors and users in taking timely preventive or corrective action, making healthcare more proactive and efficient.

V. RESULT AND DISCUSSION

The proposed AI-based multi-disease prediction system was evaluated using a variety of machine learning algorithms such as Random Forest, Support Vector Machine (SVM), Logistic Regression, and Gradient Boosting. The dataset comprised anonymized patient health records with features including age, gender, symptoms, and laboratory test results. After preprocessing and normalization, the data was split into training and testing sets using an 80:20 ratio. Among the tested models, the Random Forest classifier achieved the highest overall accuracy of **92.4%**, followed by Gradient Boosting with **90.7%**, demonstrating the effectiveness of ensemble methods for handling complex and non-linear medical data.

The model was further evaluated using performance metrics such as precision, recall, F1-score, and ROC-AUC for each disease category including diabetes, heart disease, kidney disease, and liver disease. The precision and recall for heart disease prediction were particularly high, indicating the model's reliability in identifying both positive and negative cases correctly. Confusion matrix analysis revealed that false negatives were minimal, which is crucial in medical diagnostics where missing a disease can lead to severe consequences. The multi-label classification approach

also allowed the model to predict more than one disease simultaneously when overlapping symptoms were detected, enhancing its applicability in real-world clinical settings.

In discussion, the results show that AI-driven machine learning models hold significant promise for early and accurate multi-disease prediction. This system can assist healthcare professionals by acting as a decision-support tool, helping prioritize patients for diagnostic tests and interventions. However, challenges such as data imbalance, generalizability across populations, and model interpretability remain. Future work may focus on incorporating deep learning for image-supported predictions (e.g., radiology data) and implementing the system in a real-time clinical environment with feedback loops for continuous model improvement.

VI. CONCLUSION

The **AI-based Multi Disease Prediction using Machine Learning** system demonstrates a powerful approach to early and accurate diagnosis of multiple diseases by leveraging advanced algorithms and data-driven insights. By analyzing diverse medical datasets, the model effectively identifies patterns and correlations that might be missed by traditional diagnostic methods, enabling faster and more reliable predictions. This not only supports healthcare professionals in making informed decisions but also promotes proactive patient care, reduces misdiagnosis, and enhances treatment outcomes. The integration of such AI-powered tools holds great potential in transforming modern healthcare into a more predictive and personalized system.

REFERENCES

1. Chaurasia, V., & Pal, S**.** (2014). A Novel Approach for Breast Cancer Detection using Data Mining Techniques. International Journal of Innovative Research in Computer and Communication Engineering, 2(1), 2456–2465.
2. Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. BMC Medical Informatics and Decision Making, 19(1), 281.
3. Dey, N., Ashour, A. S., & Balas, V. E. (2019). Smart Medical Data Sensing and IoT Systems Design in Healthcare. Springer.
4. Patel, J., & Prajapati, D. (2018). Heart Disease Prediction using Machine Learning and Data Mining Techniques. International Journal of Computer Science and Communications, 6(2), 112–120.
5. Kaur, H., & Kumari, V. (2020). Predictive modelling and analytics for diabetes using a machine learning approach. Applied Computing and Informatics, 18(1–2), 90–100..
6. Srinivas, K., Rani, B. K., & Govrdhan, A. (2010). Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks. International Journal on Computer Science and Engineering (IJCSE), 2(2), 250–255.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com