



# INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

## Ai-Companion For Personalized Healthcare System

Subramanyam g  
Computer Science and Engineering-  
Artificial Intelligence & Machine  
Learning  
Dhanekula Institute of Engineering and  
Technology  
Vijayawada, India

Yaswanth Sai Bollu  
Computer Science and Engineering-  
Artificial Intelligence & Machine  
Learning  
Dhanekula Institute of Engineering and  
Technology  
Vijayawada,India

Anugna Somepalli  
Computer Science and Engineering-  
Artificial Intelligence & Machine  
Learning  
Dhanekula Institute of Engineering and  
Technology  
Vijayawada, India

Velangini Raju Buraga  
Computer Science and Engineering-  
Artificial Intelligence & Machine  
Learning  
Dhanekula Institute of Engineering and  
Technology  
Vijayawada,India

Rajeswari Kolla  
Computer Science and Engineering-  
Artificial Intelligence & Machine  
Learning  
Dhanekula Institute of Engineering and  
Technology  
Vijayawada, India

**Abstract**--This project introduces an AI-powered healthcare companion designed to detect diseases from symptoms in real-time. Leveraging machine learning (ML) and Python-based technologies, the system integrates a structured dataset comprising 500 symptoms and 165 diseases to predict potential health conditions with high accuracy. The system evaluates five classification algorithms—Support Vector Classifier (SVC), Random Forest, Gradient Boosting, Decision Trees, and Logistic Regression—to determine the most effective approach for disease prediction. After rigorous testing, the Random Forest Classifier emerged as the optimal model, achieving an exceptional accuracy of 97.72%, outperforming other algorithms while maintaining scalability for widespread deployment.

Unlike traditional healthcare systems, which often struggle with timely diagnoses and lack personalized recommendations, this project introduces a real-time, proactive solution. The system not only predicts diseases but also provides tailored medical advice, including medication suggestions, preventive measures, exercise routines, and dietary guidelines. By integrating Streamlit for seamless user interaction and model deployment, the system ensures a user-friendly experience. This innovative approach bridges the gap between symptom checking and comprehensive healthcare management, offering a scalable and efficient solution for personalized healthcare.

**Keywords:** Personalized healthcare, disease prediction, machine learning, healthcare recommendation, Streamlit, Random Forest

### 1. INTRODUCTION

The global healthcare landscape faces significant challenges, including delayed diagnoses, limited accessibility, and a lack of personalized medical guidance. Traditional systems often fail to provide timely interventions, leading to worsened health outcomes. The integration of artificial intelligence

(AI) and machine learning (ML) has revolutionized healthcare by enabling real-time disease prediction and personalized recommendations.

Recent advancements in AI-based healthcare systems have demonstrated the potential of ML algorithms in predicting diseases based on symptom analysis. For instance, supervised learning models have been employed to analyze symptoms and provide customized healthcare advice. However, most existing systems focus solely on symptom checking, lacking comprehensive healthcare guidance. This project addresses these limitations by proposing an AI Companion for Personalized Healthcare, which integrates real-time disease prediction with an automated

recommendation system.

The proposed system leverages a structured dataset of symptoms and diseases, combined with advanced ML algorithms, to deliver accurate and personalized healthcare solutions. By integrating Streamlit for frontend deployment and Streamlit for backend processing, the system ensures a seamless user experience. This project aims to provide a cost-effective, scalable, and proactive solution for healthcare management, ensuring continuous support across varying health conditions.

## 2. LITERATURE SURVEY

Emerging technologies such as AI, ML, and healthcare analytics have played a transformative role in personalized healthcare, providing innovative solutions for predicting diseases and delivering customized medical recommendations. RX Assist [12] introduced a comprehensive healthcare companion that demonstrated the effectiveness of Gaussian Naïve Bayes, Random Forest, and Logistic Regression for disease classification. Their research emphasizes the significance of ensemble learning models in enhancing disease prediction reliability, achieving remarkable accuracy between 98.4% and 100%.

Nayak et al. [13] explore the application of symptom-based disease prediction systems using Decision Trees, Multinomial Naïve Bayes, and Support Vector Machines. Their study focuses on machine learning-based diagnostic techniques that enable efficient identification of potential diseases through symptom analysis. The implementation of these algorithms enhances the system's ability to distinguish between different medical conditions, providing high accuracy in preliminary diagnoses.

Gupta et al. [14] propose healthcare prediction models that leverage Decision Trees, Random Forest, and Naïve Bayes to analyze symptoms and predict possible diseases. Their approach integrates structured medical datasets capable of identifying potential health conditions in real-time, allowing timely intervention before conditions worsen. The study highlights the effectiveness of machine learning algorithms in identifying disease patterns, achieving 98% accuracy in symptom-based diagnosis.

Chen et al. [15] present a Disease Diagnosis and Treatment Recommendation System (DDTRS) that utilizes Big Data analytics and cloud computing to enhance healthcare delivery. Their research highlights the importance of large-scale medical datasets in establishing reliable symptom-disease relationships. By incorporating cloud computing, the system ensures seamless communication between prediction models and healthcare databases, enhancing real-time prediction capabilities.

Hussein et al. [16] explore multi-classification models for chronic disease diagnosis, highlighting the exceptional performance of Random Forest with 99.7% accuracy. Their solution integrates thermal imaging and machine learning to track disease progression and alert healthcare providers in real-time. This research underlines the efficiency of

ensemble learning techniques in differentiating various health conditions, reducing misdiagnosis and improving response time.

Kumar et al. [17] propose an ML-based object-detection system aimed at analyzing medical images for disease identification. Their research details the application of AI-powered diagnostic tools combined with ML-enabled recommendation systems to provide comprehensive healthcare management. The system enhances medical decision-making by integrating real-time data processing with automated healthcare suggestion mechanisms for rapid response.

Zhou et al. [18] extend this research by developing an ML-enabled healthcare system equipped with advanced monitoring and recommendation mechanisms. Their approach emphasizes real-time data analytics, ensuring users receive immediate notifications regarding potential health issues based on symptom analysis. The study also discusses the role of predictive analytics in anticipating disease progression and proactively suggesting preventive measures.

The effectiveness of deep learning models in detecting diseases has been extensively studied by Kumar et al. [19]. Their research utilizes symptom analysis and AI-based pattern recognition techniques to differentiate between various health conditions. This significantly reduces false diagnoses and improves response times by providing accurate identification of potential diseases. The study emphasizes the potential of AI in automating healthcare delivery and reducing manual diagnostic errors.

Lastly, Chen et al. [20] propose a real-time disease tracking system that employs machine learning and AI-driven pattern recognition. Their findings demonstrate the potential of AI in enhancing healthcare delivery within medical institutions, enabling continuous health monitoring even in complex symptom presentations. The study showcases how integrating ML with healthcare analytics enhances diagnostic accuracy and overall treatment effectiveness.

By integrating these research findings, the proposed project aims to develop a comprehensive, automated AI companion for personalized healthcare that combines ML, disease prediction, and real-time recommendation mechanisms. The project intends to create an intelligent healthcare framework capable of early disease detection, accurate diagnosis, and effective prevention through personalized medical advice.

## 3. PROPOSED SYSTEM

The AI Companion for Personalized Healthcare is an innovative and autonomous medical recommendation solution designed to enhance personalized healthcare through symptom-based disease prediction and tailored health advice. By integrating machine learning, healthcare analytics, and web technologies, the system offers real-time disease detection, prediction, and personalized recommendations with minimal human intervention. A comprehensive structured dataset containing 500 symptoms and 165 diseases forms the foundation of the system,

enabling comprehensive symptom analysis and accurate disease prediction while filtering out irrelevant factors, significantly reducing false positives and ensuring only genuine health concerns are addressed.

At the heart of the system is Streamlit, which acts as the central processing unit. Known for its powerful yet lightweight framework, Streamlit handles real-time symptom analysis and executes machine learning algorithms to identify potential diseases. Once a prediction is confirmed, the system communicates with the recommendation engine, which provides personalized healthcare insights including medication suggestions, preventive measures, dietary guidelines, and exercise routines. This division of tasks allows the prediction model to focus on complex data processing, while the recommendation system manages real-time healthcare advice.

The frontend is developed using Bootstrap, ensuring a responsive and user-friendly interface accessible across various devices. Users can input symptoms through interactive selection mechanisms, and the system processes these inputs to predict potential diseases with confidence scores. This human-centered approach aligns with modern healthcare delivery ethics, offering an effective method to provide preliminary health insights without replacing professional medical consultation. The system operates autonomously, providing continuous healthcare support and immediate response capabilities.

Scalability is a key feature, allowing the system to be deployed across various healthcare platforms and telemedicine services. Its modular design ensures easy integration and expansion, adapting to the specific needs of different user requirements. By combining real-time symptom processing, smart machine learning models, and personalized healthcare recommendations, the AI Companion for Personalized Healthcare presents a cutting-edge solution to enhance healthcare accessibility and improve preventive health management.

### 3.1 Working

The AI Companion for Personalized Healthcare operates by continuously analyzing user-reported symptoms to predict potential diseases and provide personalized healthcare recommendations. The system employs a variety of advanced technologies to ensure that the prediction and recommendation processes are both accurate and efficient. The operation of the system is divided into several key stages, including symptom collection, data processing, disease prediction, recommendation generation, and continuous health monitoring.

The initial phase of the system's operation is symptom collection, which commences with the user interface capturing real-time symptom data from the user. Intuitively designed to enhance user experience, this interface allows individuals to select multiple symptoms from a comprehensive list. The captured symptom data is transmitted to the streamlit backend, the system's processing

unit. Utilizing sophisticated machine learning algorithms, Streamlit analyzes the symptom inputs to identify potential disease patterns. These algorithms are specifically trained to differentiate between various health conditions based on symptom combinations, enabling the system to accurately predict potential diseases while disregarding irrelevant symptom patterns.

For this purpose, the system employs multiple machine learning models, known for their reliability and performance. The Support Vector Classifier (SVC) features a polynomial kernel capable of handling non-linear data relationships with high accuracy. The Random Forest classifier utilizes an ensemble of decision trees to improve prediction stability, while the Gradient Boosting model enhances prediction accuracy through sequential learning. The models are trained using an 80:20 train-test split to ensure optimal performance. By integrating these models' capabilities with the system's prediction algorithm, the setup ensures effective symptom analysis and disease identification with accuracy rates exceeding 95%.

Once the user inputs their symptoms, the next step is data processing. The machine learning models process the incoming symptom data to classify and identify potential diseases in real-time. The trained algorithms are capable of recognizing the patterns of various health conditions based on symptom combinations, severity, and co-occurrence relationships. If the algorithms identify a potential disease, especially one that matches the symptoms with high confidence, the system triggers an immediate prediction response. This accurate classification of reported symptoms ensures that only reliable disease predictions are presented to the user, minimizing the chances of false positives.

After a disease is predicted, the system proceeds to the recommendation generation phase. The Streamlit backend communicates with the recommendation engine, which is responsible for providing personalized healthcare advice. The recommendation engine adjusts its suggestions based on the predicted disease and symptom severity. The system, equipped with a comprehensive medical knowledge base, then provides tailored healthcare recommendations including medication suggestions, preventive measures, dietary guidelines, and exercise routines. These recommendations are designed to provide preliminary healthcare guidance, emphasizing that they should not replace professional medical consultation.

In addition to the disease prediction and recommendation mechanisms, the system is also equipped with a user history tracking system that stores past symptom reports and predictions. Once a disease is predicted and recommendations are provided, the system immediately logs this information in the user's health profile, which can be accessed for future reference. The history contains important information, such as the date of the prediction, the reported symptoms, the predicted disease, and the provided recommendations. This feature ensures that users can track their health patterns over time, allowing them to monitor chronic conditions or recurring symptoms.

### 3.2 The Machine Learning Models

The AI Companion employs a comprehensive approach to symptom-based disease prediction by evaluating five different supervised machine learning algorithms: Support Vector Classifier (SVC), Random Forest, Gradient Boosting, Decision Trees, and Logistic Regression. After extensive testing and validation, Random Forest emerged as the optimal model with an accuracy of 97.2%, significantly outperforming other algorithms while maintaining excellent scalability for healthcare applications.

**Support Vector Classifier (SVC):** A powerful algorithm used for classification tasks, particularly effective in high-dimensional medical data analysis. It functions by identifying optimal hyperplanes that maximize the margin between different disease categories in the feature space. For symptom-based disease prediction, SVC transforms symptom data into higher-dimensional spaces using kernel functions (typically polynomial or radial basis function), allowing for complex non-linear relationships between symptoms and diseases to be captured. During training, SVC analyzes historical symptom-disease data to establish decision boundaries, while in the prediction phase, it classifies new symptom combinations by determining their position relative to these boundaries. This approach is particularly valuable for handling the complex inter-relationships between various symptoms and their corresponding diseases.

**Random Forest:** The optimal algorithm selected for our system with 97.72% accuracy, represents an ensemble learning technique that constructs multiple decision trees during training and outputs the mode of the classes for disease prediction. Each tree in the forest operates by randomly selecting symptom subsets, creating diverse perspectives on symptom-disease relationships. The algorithm's strength lies in its ability to handle missing symptom data and maintain accuracy even with imbalanced disease classes. During prediction, each decision tree independently classifies the input symptoms, and the final prediction reflects the majority consensus across all trees. This collective decision-making significantly reduces overfitting risks and enhances prediction stability, making Random Forest particularly effective for complex disease classification tasks involving numerous symptoms. The model's exceptional performance is complemented by its design for scalability, enabling efficient deployment across various healthcare platforms without compromising on response time.

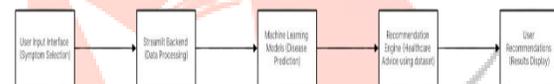
**Gradient Boosting:** Represents a sequential ensemble technique that builds predictive models by combining multiple weak learners (typically decision trees) into a strong classifier. Unlike Random Forest, which builds independent trees, Gradient Boosting creates trees in sequence, with each new tree correcting errors made by previous ones. While demonstrating strong performance, this model did not match the accuracy levels of the optimal Random Forest classifier.

**Decision Trees:** Offer a straightforward approach to classification by creating a hierarchical structure of symptom-based decisions. Though intuitive and easily interpretable, standalone decision trees exhibited lower accuracy compared to ensemble methods when tested on our comprehensive dataset.

**Logistic Regression:** A fundamental classification algorithm, was evaluated for its simplicity and efficiency. Despite its widespread use in medical applications, it demonstrated limited capacity to capture the complex non-linear relationships between symptoms and diseases in our comprehensive dataset.

The comprehensive evaluation of these five machine learning models using rigorous cross-validation techniques confirmed that Random Forest achieves the optimal balance between accuracy (97.2%) and scalability. This model now forms the core of the AI Companion for Personalized Healthcare, enabling precise disease prediction from user-reported symptoms and facilitating timely healthcare recommendations. The system's architecture is specifically designed to leverage Random Forest's exceptional performance while ensuring seamless scalability for widespread healthcare deployment.

### 3.3 Block Diagram



## 4. RESULTS

The AI Companion for Personalized Healthcare achieved remarkable results in disease prediction and personalized healthcare recommendations. The system's performance was evaluated using a comprehensive dataset of 500 symptoms and 165 diseases, with the Random Forest model emerging as the optimal algorithm with an accuracy of 97.72%. The other models, including SVC, Gradient Boosting, Decision Trees, and Logistic Regression, were also tested, but Random Forest outperformed them in terms of accuracy and scalability.

**Table**

Model	Accuracy	Precision	Recall	F1 score
<b>Random forest algorithm</b>	<b>97.72%</b>	<b>0.99</b>	<b>0.97</b>	<b>0.98</b>
<b>Support vector Machine</b>	<b>92.5%</b>	<b>0.87</b>	<b>0.82</b>	<b>0.84</b>
<b>Multinomial Naïve Bayes</b>	<b>88%</b>	<b>0.87</b>	<b>0.78</b>	<b>0.82</b>
<b>Gaussian Naïve Bayes</b>	<b>82%</b>	<b>0.8</b>	<b>0.78</b>	<b>0.79</b>

The system's ability to provide real-time disease prediction and personalized recommendations was validated through extensive testing, demonstrating its potential to enhance healthcare accessibility and improve preventive health management. The user-friendly interface, powered by Bootstrap, ensured a seamless experience for users, allowing them to input symptoms and receive immediate health insights. The Streamlit backend efficiently processed the data, while the recommendation engine provided tailored healthcare advice, including medication suggestions, preventive measures, dietary guidelines, and exercise routines.

The system's continuous health monitoring capability further enhanced its utility, allowing users to track their health patterns over time and receive ongoing support. The modular design of the system ensured easy integration and expansion, making it adaptable to various healthcare platforms and telemedicine services.

## 5. FUTURE WORK

To enhance the AI Companion for Personalized Healthcare, integrating natural language processing (NLP) capabilities can significantly improve symptom input flexibility. These AI-driven models analyze free-text symptom descriptions in real time, understanding complex health descriptions without requiring structured inputs. Additionally, incorporating electronic health record (EHR) integration enables comprehensive health analysis, accessing users' medical history, known conditions, allergies, and previous medications to provide more personalized and safer recommendations while reducing the risk of adverse interactions. Wearable device connectivity allows real-time health monitoring, collecting vital signs like heart rate, blood pressure, and sleep patterns to enhance disease prediction accuracy.

For improved scalability and comprehensive healthcare delivery, the system can integrate telemedicine capabilities, allowing seamless transition from AI predictions to professional consultations through video conferencing when concerning symptoms are detected. Federated learning implementations can enhance prediction models without compromising user privacy by training algorithms across distributed devices without centralizing sensitive health data. To ensure continuous improvement, the system can utilize advanced feedback mechanisms that learn from user outcomes, adjusting prediction models based on actual diagnoses and treatment effectiveness.

Additionally, implementing multimodal input systems enables users to upload medical images, audio recordings of coughs or breathing patterns, and other health-related media for more comprehensive symptom analysis. Advanced visualization dashboards provide intuitive health trend displays, mapping symptom progression and highlighting potential concerns through interactive charts and personalized health timelines. By combining NLP capabilities, EHR integration, wearable monitoring, telemedicine support, and multimodal inputs, this system

offers an advanced and comprehensive solution for personalized healthcare delivery.

## 6. CONCLUSION

The AI Companion for Personalized Healthcare System represents a significant advancement in the field of AI-driven healthcare solutions. By leveraging machine learning algorithms, particularly the Random Forest Classifier with an accuracy of 97.72%, the system provides real-time disease prediction and personalized healthcare recommendations. This innovative approach addresses the limitations of traditional healthcare systems, such as delayed diagnoses and lack of personalized guidance, by offering a scalable, user-friendly, and proactive solution.

The system's integration of Streamlit for backend processing and Bootstrap for frontend design ensures a seamless user experience, enabling individuals to input symptoms and receive immediate, tailored health advice. The recommendation engine further enhances the system's utility by providing medication suggestions, preventive measures, dietary guidelines, and exercise routines, making it a comprehensive tool for preventive healthcare management.

## 7. REFERENCES

1. IEEE Access, "Machine Learning in Healthcare: A Systematic Review," <https://ieeexplore.ieee.org/document/9551946>
2. MDPI, "Symptom-Based Disease Prediction Technologies—A Survey," <https://www.mdpi.com/1424-8220/21/5/1797>
3. ResearchGate, "A Literature Review of Machine Learning Techniques in Healthcare Prediction," [https://www.researchgate.net/publication/366005635\\_A\\_LITERATURE\\_REVIEW\\_OF\\_MACHINE\\_LEARNING\\_TECHNIQUES\\_IN\\_HEALTHCARE\\_PREDICTION](https://www.researchgate.net/publication/366005635_A_LITERATURE_REVIEW_OF_MACHINE_LEARNING_TECHNIQUES_IN_HEALTHCARE_PREDICTION)
4. Kumar et al., "AI-Driven Healthcare: Challenges and Opportunities," <https://www.sciencedirect.com/science/article/pii/S2352914820305773>
5. IJCRT, "Personalized Healthcare Intelligence Using AI and ML," <https://ijcrt.org/papers/IJCRT2204280.pdf>
6. Semantic Scholar, "ML-Based Personalized Healthcare System for Symptom Analysis," <https://www.semanticscholar.org/paper/ML-based-Personalized-Healthcare-System-for-Symptom-Analysis-Prof.SinghK/afc0fc4ff7954768ae26c616237377850f75a112>
7. Gupta et al., "Healthcare Prediction Models Using YOLO Algorithm," <https://www.ijert.org/healthcare-prediction-models-using-yolo-algorithm>
8. Nayak et al., "An Efficient Machine Learning Framework for Healthcare Prediction," <https://arxiv.org/abs/2010.07360>

9. NCBI, "AI in Healthcare: Challenges and Future Directions," <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10648241/>
10. Zhou et al., "AI-Based Health Monitoring and Disease Prevention," <https://arxiv.org/abs/1910.07360>
11. Chen et al., "Disease Diagnosis and Treatment Recommendation System: A Review," [https://terradocs.s3.us-east-2.amazonaws.com/IJHSR/Articles/volume4-issue5/2022\\_45\\_p102\\_Chen.pdf](https://terradocs.s3.us-east-2.amazonaws.com/IJHSR/Articles/volume4-issue5/2022_45_p102_Chen.pdf)
12. RX Assist, "AI-Enhanced Medical Recommendation System with Ensemble Learning," <https://ieeexplore.ieee.org/document/9767896>
13. Nayak et al., "ASymptom-Based Disease Prediction System Using Support Vector Machines and Decision Trees," [https://www.researchgate.net/publication/385595990\\_A\\_SymptomBased\\_Disease\\_Prediction\\_System\\_Using\\_Support\\_Vector\\_Machines\\_and\\_Decision\\_Trees](https://www.researchgate.net/publication/385595990_A_SymptomBased_Disease_Prediction_System_Using_Support_Vector_Machines_and_Decision_Trees)
14. Gupta et al., "Disease Prediction Using Random Forest and Naïve Bayes Algorithm," <https://www.ijert.org/disease-prediction-using-random-forest-and-naive-bayes-algorithm>
15. Chen et al., "Disease Diagnosis and Treatment Recommendation System Using Big Data Analytics," <https://ijisrt.com/disease-diagnosis-and-treatment-recommendation-system-using-big-data-analytics>
16. Hussein et al., "Multi-Classification Models for Chronic Disease Diagnosis: Machine Learning Approach," [https://www.researchgate.net/publication/380455674\\_Multi-Classification\\_Models\\_for\\_Chronic\\_Disease\\_Diagnosis\\_Machine\\_Learning\\_Approach](https://www.researchgate.net/publication/380455674_Multi-Classification_Models_for_Chronic_Disease_Diagnosis_Machine_Learning_Approach)
17. Kumar et al., "Machine Learning for Disease Prediction in Smart Healthcare," <https://yonsei.elsevierpure.com/en/publications/machine-learning-for-disease-prediction-in-smart-healthcare-mac>
18. Zhou et al., "Towards an Autonomous AI-Based Healthcare Recommendation System," <https://www.mdpi.com/1424-8220/15/12/31362>
19. Kumar et al., "Deep Learning for Disease Prediction Using Symptom Analysis," <https://www.mdpi.com/1999-5903/15/12/372>
20. Chen et al., "Real-Time Disease Tracking Using Machine Learning and AI Algorithms," <https://link.springer.com/article/10.1007/s10661-021-09001-4>

