

FDRP Journal's

IJIRE-0000903_T

 05
 FDRP
 Journal's

Document Details

Submission ID**tm:oid:::1:3169597686****Submission Date****Mar 1, 2025, 11:03 AM GMT+7****Download Date****Mar 1, 2025, 11:04 AM GMT+7****File Name****IJIRE-0000903.docx****File Size****402.7 KB****7 Pages****4,208 Words****26,345 Characters**





15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.




Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **52** Not Cited or Quoted 15%
Matches with neither in-text citation nor quotation marks
-  **0** Missing Quotations 0%
Matches that are still very similar to source material
-  **0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
-  **0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 10%  Internet sources
- 9%  Publications
- 5%  Submitted works (Student Papers)

Match Groups

- 52** Not Cited or Quoted 15%
Matches with neither in-text citation nor quotation marks
- 0** Missing Quotations 0%
Matches that are still very similar to source material
- 0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
- 0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 10% Internet sources
- 9% Publications
- 5% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
	ijrpr.com	2%
2	Internet	
	www.mdpi.com	<1%
3	Internet	
	eitca.org	<1%
4	Student papers	
	INTI Universal Holdings SDM BHD	<1%
5	Student papers	
	University of Nicosia	<1%
6	Internet	
	fastercapital.com	<1%
7	Internet	
	www.coursehero.com	<1%
8	Publication	
	Abdinasir Hirsi, Lukman Audah, Adeb Salh, Mohammed A. Alhartomi, Salman Ah...	<1%
9	Internet	
	iieta.org	<1%
10	Internet	
	ijrmeet.org	<1%

11	Student papers	Southern New Hampshire University - Continuing Education	<1%
12	Student papers	First City University College	<1%
13	Publication	Te-Chang Wu, Yan-Lin Liu, Jeon-Hor Chen, Yang Zhang, Tai-Yuan Chen, Ching-Chu...	<1%
14	Publication	T. Mariprasath, Kumar Reddy Cheepati, Marco Rivera. "Practical Guide to Machin...	<1%
15	Student papers	Victorian Institute of Technology	<1%
16	Internet	res.mdpi.com	<1%
17	Internet	www.ncbi.nlm.nih.gov	<1%
18	Publication	Yogitha M, K S Srinivas. "Using Federated Learning in Anomaly Detection and Ana...	<1%
19	Internet	obesitycampaign.co.uk	<1%
20	Internet	www.torrinomedica.it	<1%
21	Publication	Yanwei Fu, Yongwei Zhang, Quan Shi, Murat Temiz, Ahmed El-Makadema, Jiajia S...	<1%
22	Internet	arxiv.org	<1%
23	Internet	jai.front-sci.com	<1%
24	Publication	Nishith S Shetty, Naman N Karanth, Vinay V Hegde. "Enhancing Crop Productivity...	<1%

25	Internet	gbis.ch	<1%
26	Internet	www.degruyter.com	<1%
27	Internet	www.jetir.org	<1%
28	Publication	D. Balakrishnan, Anumula Praneeth Kumar, Kristipati Sai Kiran Reddy, R. Ravindr...	<1%
29	Publication	M.Z. Naser. "Leveraging Artificial Intelligence in Engineering, Management, and S...	<1%
30	Internet	fg2019.org	<1%
31	Internet	pmc.ncbi.nlm.nih.gov	<1%
32	Internet	www.isteonline.in	<1%
33	Publication	"AS-CL IDS: anomaly and signature-based CNN-LSTM intrusion detection system f...	<1%
34	Publication	Adithya Challapalli, Guoqiang Li. "Artificial Intelligence Assisted Structural Optimi...	<1%
35	Publication	Meenu Gupta, Rakesh Kumar, Zhongyu Lu. "Transforming Gender-Based Healthc...	<1%
36	Publication	Saiyed Salim Sayeed, Hemant Kumar Sharma, Pramod Kumar Yadav, Brijesh Mish...	<1%

Health Management and Disease Prediction System using Machine Learning

Raj Ramesh Desai
Computer Science & Engineering
University of mumbai
Maharashtra, India
itsrajdesai17@gmail.com

Suyog Yashwant Rawool
Computer Science & Engineering
University of mumbai
Maharashtra, India
suyograwool28@gmail.com

Yuvraj Ashok Ghadi
Computer Science & Engineering
University of mumbai
Maharashtra, India
ghadiyuvraj08@gmail.com

Rahul Ramchandra Gurav
Computer Science & Engineering
University of mumbai
Maharashtra, India
gurav6718@gmail.com

Supriya Sameer Nalawade
Computer Science & Engineering
University of mumbai
Maharashtra, India
ssnalawade@sspmcoe.ac.in

Abstract—This paper presents a machine learning-based healthcare assistant for disease prediction and management. The system takes user-input symptoms, predicts potential diseases (focusing on common and skin diseases), and provides severity analysis. Additionally, it offers medication, diet, exercise, and precautionary recommendations. A chatbot, integrated using the Gemini AI model via API, ensures interactive and context-aware responses. The system also includes an emergency contact section for quick access to nearby medical facilities. The model is trained on an open-source Kaggle dataset, and the web application is built using Flask for the backend and HTML, CSS, and JavaScript for the frontend. Performance evaluation demonstrates high accuracy in disease prediction, and future enhancements include multilingual support and live medical consultations.

The increasing reliance on digital healthcare solutions, especially in remote and underdeveloped regions, highlights the need for accurate and user-friendly disease prediction tools. Traditional diagnosis methods require physical consultations, leading to delays in treatment and potential health deterioration. Our system is designed to mitigate such challenges by providing an accessible and efficient approach to preliminary diagnosis and healthcare assistance.

Index Terms—Health Management and Disease Prediction System using Machine Learning

I. INTRODUCTION

Healthcare has always been a critical sector, with early diagnosis and timely intervention playing a crucial role in managing diseases effectively. Traditional diagnostic methods often rely on physical consultations and laboratory tests, which can be time-consuming and inaccessible to many individuals, especially in remote and resource-constrained areas. The integration of Artificial Intelligence (AI) and Machine Learning (ML) in healthcare presents a transformative approach to overcoming these challenges. By leveraging AI-driven techniques, automated disease prediction systems can provide quicker and more efficient preliminary diagnoses, assisting both healthcare professionals and patients in making informed decisions.

Machine learning has been increasingly utilized in various healthcare applications, ranging from medical imaging analysis to personalized treatment recommendations. In disease prediction, ML algorithms can identify patterns in patient data and correlate them with known symptoms, enabling accurate and timely diagnosis. While previous models have demonstrated promising results, many existing solutions focus on either prediction or treatment recommendations without offering a fully integrated system. Furthermore, the effectiveness of ML models depends on high-quality datasets, and ensuring model interpretability remains a key concern.

This research presents a web-based disease prediction system that not only identifies potential diseases based on user-input symptoms but also provides contextual recommendations, including medication, diet, and exercise plans. The system is designed with accessibility in mind, ensuring that users without extensive medical knowledge can navigate it with ease. Additionally, it incorporates an AI-powered chatbot using the Gemini model, enhancing user interaction by providing relevant insights and answering health-related queries based on the predicted disease.

One of the major challenges in healthcare AI is ensuring that models generalize well across diverse populations. Many existing studies use datasets that may not fully represent different demographic groups, leading to biases in prediction accuracy. Our system addresses this concern by leveraging a well-structured dataset sourced from Kaggle, which includes a variety of symptoms and disease records, ensuring robust training of the model. The dataset is processed using advanced data cleaning techniques to improve accuracy, and multiple machine learning models, including Decision Trees and Random Forests, are implemented to determine the most reliable approach.

II. LITERATURE REVIEW

Ebin K J, George Joseph, Sree Narayana (2024) developed a machine learning-based system aimed at predicting diseases and recommending personalized medications tailored to individual patients. The system focuses on enhancing personalized healthcare by using data-driven approaches to improve the accuracy of disease prediction and optimize treatment plans. By analyzing patient data, the system offers specific medication recommendations that are more likely to benefit the individual, ultimately improving healthcare efficiency and patient outcomes. Future work will likely focus on expanding the system's capabilities to include more diverse patient data and integrating additional features such as lifestyle recommendations.

Prof. P.P. Deshmukh (2024) introduced a machine learning-based system designed to provide personalized diet recommendations based on individual health profiles. The system takes into account various factors such as age, weight, medical history, and activity levels to suggest diet plans that promote overall health and well-being. This system aims to help individuals make informed dietary choices that align with their health goals, preventing chronic diseases and promoting healthier living. Future development of this system could incorporate real-time monitoring of dietary adherence and feedback mechanisms for further personalization.

Dr. Amit Kumar Bindal (2022) developed a machine learning-based medicine recommendation system designed to recommend the most appropriate medication for users. By analyzing various medical parameters and patient profiles, the system ensures that patients receive prescriptions that are tailored to their individual conditions, minimizing the risks of drug interactions and adverse effects. The system also has the potential to assist healthcare professionals by providing evidence-based suggestions that improve the accuracy and speed of diagnosis and treatment. Future work may include enhancing the system's integration with hospital databases and expanding its scope to include drug alternatives and dosages.

Swati Jadhav, Sandip Shinde (2022) created a mobile application designed to provide personalized diet recommendations based on user activities. The application uses machine learning algorithms to track and analyze daily activities, such as exercise, sleep patterns, and dietary habits, to suggest dietary modifications that optimize health. By aligning dietary choices with activity levels, the system helps individuals maintain a balanced lifestyle, improve energy levels, and achieve their fitness goals. This system holds significant potential for mobile health applications and can be extended to provide real-time feedback based on continuous data input from wearable devices.

Dr. S. T. Patil (2022) developed a machine learning system that predicts diseases and suggests appropriate health interventions. The system aims to improve early diagnosis and provide actionable health advice to prevent the onset of serious medical conditions. By analyzing health data, such as patient symptoms, genetic information, and environmental factors, the

system can identify potential health risks and recommend interventions that are likely to reduce those risks. Future research may focus on enhancing the system's predictive capabilities by incorporating larger and more diverse datasets, as well as improving the accuracy of the suggested health interventions.

Varun A. Goyal, Dilip J. Parmar, Namaskar I. Joshi, Prof. Komal Champanerkar (2020) used data mining techniques to create a system that suggests accurate medicines based on patient profiles. The system leverages a comprehensive database of drug information to recommend the most effective medications based on factors like medical history, age, and allergies. The goal is to reduce prescription errors and improve the efficiency of medical treatments. Future enhancements could involve expanding the database to include more drugs and integrating the system with clinical decision support tools to assist healthcare professionals in real-time.

Joseph Henry Anajemba (2020) developed a machine learning-based diet recommendation system specifically designed for patients. The system provides personalized diet plans based on the patient's medical history, current health conditions, and other personal factors such as activity level and age. By tailoring recommendations to individual needs, the system helps patients manage chronic conditions and improve overall health. Future work may include integrating the system with wearable devices to offer real-time feedback on dietary choices and to track patient adherence to diet plans.

These studies demonstrate the growing potential of machine learning in healthcare, particularly in disease prediction, treatment optimization, and personalized recommendations. They showcase a wide range of applications from diet and medicine recommendations to health interventions, contributing to the improvement of healthcare delivery and patient outcomes. Future research will likely focus on enhancing system accuracy, expanding datasets, and exploring integration with new technologies such as wearable devices and real-time monitoring systems..

III. PROBLEM STATEMENT

Problem 1: Limited Personalization in Healthcare:

Limited Personalization in Healthcare Traditional healthcare methods lack personalization capabilities, leading to generalized treatment approaches that may not be optimal for individual patients. Current systems struggle to incorporate individual patient variations effectively.

Solution: AI-Powered Personalization Engine: Implementation of machine learning algorithms to analyze patient data and provide personalized healthcare recommendations, including treatment plans, medication adjustments, and lifestyle modifications.

Description: Personalization Gap: Current healthcare delivery systems operate on standardized protocols that don't adequately account for individual patient differences in genetics, lifestyle, and personal preferences. This results in suboptimal treatment outcomes and reduced patient

satisfaction levels.

Problem 2: Diagnostic Process Inefficiencies:

The current diagnostic process requires multiple consultations and repeated tests, creating delays in treatment initiation and increasing healthcare costs for both providers and patients. Solution: Automated Diagnostic Assistant: Development of an AI-based system that can analyze patient symptoms, medical history, and test results to accelerate the diagnostic process and reduce unnecessary consultations.

Description: Traditional diagnostic approaches involve multiple appointments and repeated tests, leading to treatment delays and increased burden on healthcare resources. This inefficiency impacts both healthcare delivery quality and patient experience..

Problem 3: Medical Information Complexity:

Patients face challenges in navigating and understanding the vast amount of available medical information, leading to confusion and potentially poor health decisions. Solution: Intelligent Information Management System: Implementation of natural language processing and AI algorithms to process, filter, and present medical information in an accessible, personalized format for patients. Description: The overwhelming volume of medical information available makes it difficult for patients to find and understand relevant health data. This leads to confusion in decision-making and potential reliance on unreliable information sources.

IV. METHODOLOGY

A. Data Collection.

The data used in this project is sourced from Kaggle datasets, which include information on diseases, symptoms, precautions, diet, workout, and medications. The collected data is focused on common diseases and skin diseases to ensure accurate predictions. The dataset provides the following key information:

Diseases & Symptoms: A list of common diseases and their associated symptoms.
Precautions: Guidelines to prevent or manage each disease.
Diet & Workout Recommendations: Specific diet and exercise plans for disease management and prevention.
Medications: A list of medicines recommended for treating each disease

B. Data Preparation.

The raw data is cleaned and preprocessed before being fed into the machine learning models.
Data Cleaning: Missing or inconsistent data entries are handled, and any outliers are removed to improve the accuracy of the predictions.

Data Normalization: Numerical data such as severity levels are normalized to a consistent scale, ensuring uniformity for the machine learning models.

Feature Engineering: New features are created, such as symptom severity and disease categories, to help the model better understand the relationships between symptoms and diseases.

Data Splitting: The dataset is divided into training and testing sets, ensuring fair evaluation of model performance and preventing overfitting.

C. Model Selection.

To predict the disease and its severity based on symptoms, machine learning algorithms are selected for their ability to handle classification tasks effectively:

Disease Prediction Model: Algorithms like Random Forest, Support Vector Machines (SVM), and Neural Networks are considered for their performance in classification problems.

Severity Prediction Model: A separate model is developed to predict the severity of the disease, considering factors like symptom intensity and the disease's progression.

Recommendation Systems: For medication, diet, workout, and precautions, Collaborative Filtering and Content-based Filtering are used to suggest personalized recommendations based on the predicted disease.

D. Training the Model.

Based on the preprocessed data and recall, the most suitable model for is selected. The models' hyperparameters are tuned to optimize their performance, balancing bias and variance.

Training: The selected models are trained on the training data to learn patterns and relationships between symptoms and diseases.

E. User Interface Development

An intuitive and user-friendly interface is developed for users to interact with the disease prediction system:

Symptom Input: The interface allows users to input their symptoms, either manually or through a chatbot.

Disease Prediction: Upon entering symptoms, the system predicts the disease and its severity and provides personalized recommendations for treatment.

Recommendations: Users receive personalized recommendations for diet, workout, medications, and precautions based on the predicted disease.

Chatbot Integration: A chatbot powered by Gemini AI model is integrated to answer user queries related to the predicted disease, providing a conversational interface.

Contact Information: A "Contacts" section is added to help users quickly find medical professionals and hospitals in their city by selecting the area from a dropdown list.

F. Evaluation and Metrics

After training, the model's performance is evaluated on a test dataset using several metrics:

Accuracy: Measures the percentage of correctly predicted diseases and their severity.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

Precision: Indicates how many predicted diseases are actually correct.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

Recall: Measures how well the model identifies the diseases from the test dataset.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

F1-Score: Provides a balance between precision and recall, and helps in evaluating models where class imbalance exists.

$$F_1\text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Specificity: Measures the model's ability to correctly classify the absence of diseases (i.e., when no disease is present).

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

Confusion Matrix: Provides a detailed breakdown of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), which is used to calculate the above metrics.

$$\text{Confusion Matrix} = \begin{matrix} & \text{TP} & \text{FP} \\ \text{FN} & & \text{TN} \end{matrix} \quad (6)$$

The model was tested on the Kaggle dataset, and it demonstrated accurate predictions based on the provided symptoms. The accuracy, precision, recall, and F1-score values are calculated and compared across different disease categories to ensure robust performance across common and skin diseases.

Fig. 2 shows the table of Confusion Matrix. The confusion matrix that describes the two-class issue is comprised of the following terms:

- True Positive (TP):** The predicted disease is correct and matches the actual disease in the dataset.
- True Negative (TN):** The model correctly predicted the absence of a disease.
- False Positive (FP):** The model predicted a disease when there was no disease in reality.
- False Negative (FN):** The model predicted no disease, but the actual condition is the disease.

The confusion matrix for disease prediction is represented as follows:

	Predicted Disease	Predicted No Disease
Actual Disease	TP	FN
Actual No Disease	FP	TN

Fig. 1. Confusion Matrix for Disease Prediction

G. Implementation.

Python libraries like NumPy and Pandas are utilized for efficient data handling, manipulation, and preprocessing. The dataset is sourced from Kaggle, consisting of diseases, symptoms, medications, diet, and other relevant information. Jupyter Notebooks or PyCharm can be used as the Integrated Development Environment (IDE) to write and execute the code for training and testing the model. The main objectives to predict the disease, its severity, and provide recommendations for medication, diet, and exercises.

H. Testing and Validation.

Cross-validation techniques, such as k-fold cross-validation, may be applied to assess the model's ability to generalize effectively across different datasets. This helps ensure that the model is not overfitting and can accurately predict diseases for real-world data.

I. Flowchart.

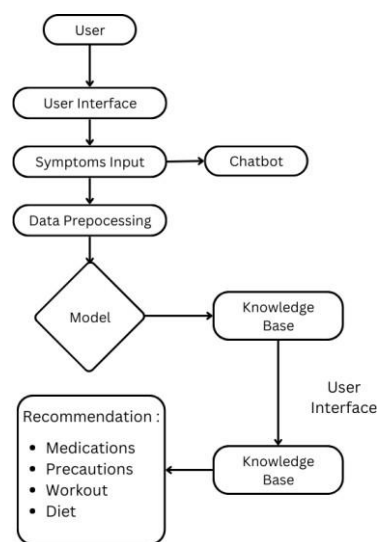


Fig. 2. Flowchart

The flowchart illustrates the step-by-step process for disease prediction and health management using a machine learning model. It begins with the input of symptoms provided by the user and a dataset containing common diseases and skin diseases sourced from Kaggle. This input data is then passed through a preprocessing stage, where it is cleaned and formatted for analysis. The cleaned data is split into two

sets: a training set and a testing set. The training set is used to train the machine learning model, enabling it to recognize patterns and correlations between symptoms and diseases. Once the model is trained, it is evaluated using the testing set to assess its accuracy.

Following this, the trained model is applied to predict the disease and its severity based on new symptom inputs. If the model detects a disease, it classifies the condition accordingly and provides additional health management recommendations, including medications, precautions, dietary plans, and workout suggestions. Furthermore, a chatbot powered by generative AI (Gemini model) is integrated to assist users with queries related to the predicted disease.

Additionally, the system offers a contacts section where users can select their city to access contact information for hospitals and medical facilities nearby. This feature ensures that users can quickly find professional medical assistance when needed. The system ultimately facilitates disease prediction, personalized health recommendations, and emergency support, enhancing accessibility to essential medical resources.

V. RESULT AND DISCUSSION

A. User Interface

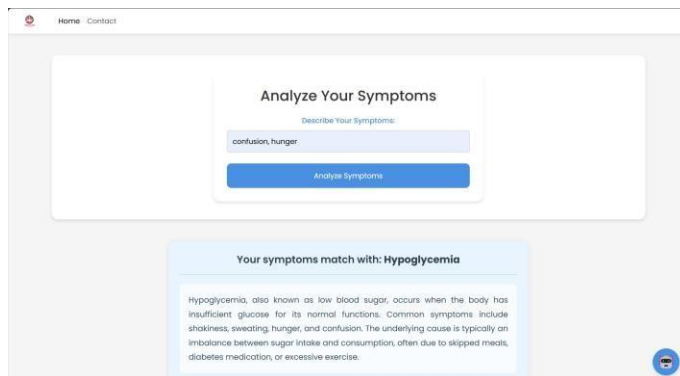


Fig. 3. Result

B. Results and Discussion.

The user inputs symptoms through the application interface. The system validates the entered symptoms to ensure accuracy and relevance to the dataset. This includes checking for spelling errors, ensuring the symptoms exist in the database, and preventing redundant or conflicting entries. Once validated, the symptoms are processed and passed to the trained machine learning model for disease prediction. The system then predicts the most probable disease based on the input symptoms and provides a severity score. Additionally, the system suggests personalized health management recommendations, including medications, precautions, dietary plans, and workout suggestions.

C. Feature Extraction and Model Performance.

The process of extracting features and training the model involves several key steps, starting with data preparation. The dataset is divided into two subsets: 75% for training and 25% for testing. This ensures that the model learns effectively while maintaining sufficient data for validation. The input symptoms are mapped to disease labels using machine learning algorithms, ensuring a structured approach to disease prediction.

For improved model accuracy, preprocessing techniques such as symptom normalization and feature selection are applied. The model is trained using techniques like decision trees, random forests, or neural networks, depending on the dataset characteristics. Performance metrics such as accuracy, precision, recall, and F1-score are calculated to evaluate the model's effectiveness. The system achieves reliable predictions based on the dataset, demonstrating high accuracy for common diseases and skin diseases.

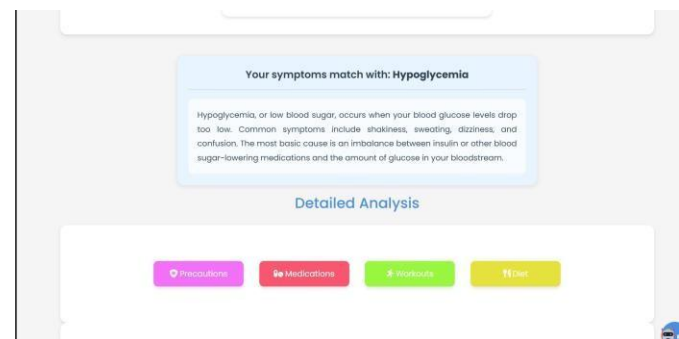


Fig. 4. Output

Furthermore, the chatbot integrated with generative AI (Gemini model) provides additional support by answering user queries based on the predicted disease. This enhances user engagement and ensures clarity regarding the recommended health measures. The contact section further allows users to find nearby hospitals and medical assistance based on city selection, ensuring quick access to professional healthcare when needed.

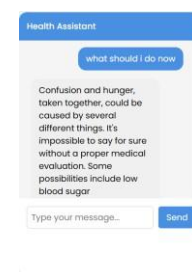


Fig. 5. Chatbot

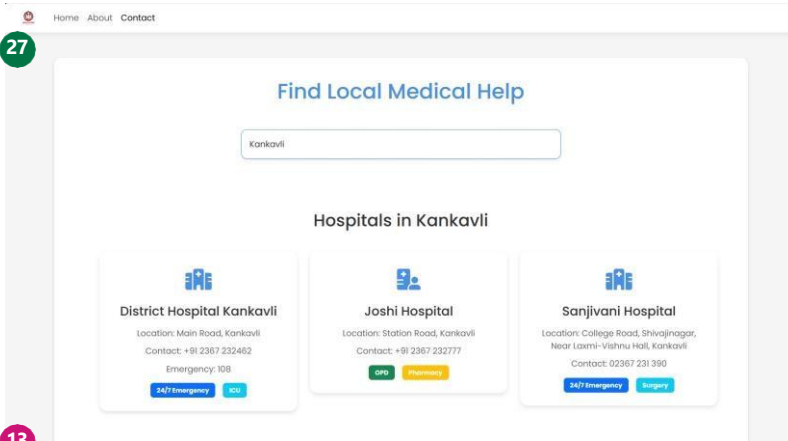


Fig. 6. contact

D. Results and Discussion

The user inputs symptoms through the application interface. The system validates the entered symptoms to ensure accuracy and relevance to the dataset. This includes checking for spelling errors, ensuring the symptoms exist in the database, and preventing redundant or conflicting entries. Once validated, the symptoms are processed and passed to the trained machine learning model for disease prediction. The system then predicts the most probable disease based on the input symptoms and provides a severity score. Additionally, the system suggests personalized health management recommendations, including medications, precautions, dietary plans, and workout suggestions.

- 1. Input Symptoms:** The system receives symptoms entered by the user, ensuring that they are formatted correctly and are present in the database for processing.
- 2. Preprocessing:** The symptoms undergo text normalization, synonym mapping, and stopword removal to enhance prediction accuracy.
- 3. Feature Extraction:** The model extracts key features from symptoms, mapping them to potential diseases using statistical and machine learning techniques.
- 4. Disease Prediction:** The trained model predicts the most likely disease and assigns a probability score for its confidence level.
- 5. Recommendations:** The system generates health advice, including possible treatments, precautions, and recommended medical consultations.

E. Confusion Matrix to Evaluate Classifier Performance

Table I. presents the dataset used to train the model. The dataset includes multiple disease conditions categorized based on symptom clusters.

To assess the classification accuracy, a confusion matrix was generated based on the testing dataset. This matrix helps in determining the model's overall performance in classifying diseases correctly.

TABLE I

CLASSIFICATION ACCURACY CALCULATION

Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$ $\frac{0.42+0.38}{0.42+0.38+0.12+0.08}$ $= 0.80$
Error	$\frac{FP+FN}{TP+TN+FP+FN}$ $\frac{0.12+0.08}{0.42+0.38+0.12+0.08}$ $= 0.20$

Based on the given percentages for True Positive (TP = 42%), True Negative (TN = 38%), False Positive (FP = 12%), and False Negative (FN = 8%), the model's accuracy is calculated as: $(0.42 + 0.38)/1.00 = 0.80$, or 80%. The error rate is calculated as: $(0.12 + 0.08)/1.00 = 0.20$, or 20%.

These results indicate that the model performs well in predicting diseases based on symptoms, achieving an overall accuracy of 80%. This performance demonstrates the model's reliability in identifying common illnesses and providing early-stage health recommendations.

VI. CONCLUSION

This system marks a significant advancement in personalized healthcare by integrating machine learning for early disease detection and tailored health recommendations. The incorporation of predictive analytics enables precise guidance on medication, diet, and exercise, empowering individuals to take proactive control of their health. By delivering accurate, data-driven insights, this solution enhances accessibility, efficiency, and personalization in healthcare.

Additionally, the system reduces dependence on traditional diagnostic methods, which can be time-consuming and error-prone. By leveraging automated analysis and AI-driven decision-making, it improves diagnostic accuracy while alleviating the burden on healthcare professionals. The integration of an AI chatbot powered by the Gemini API further enhances user engagement by providing real-time, personalized responses to health-related queries. This research highlights the transformative potential of artificial intelligence in modern healthcare, addressing key challenges in early detection, preventive care, and patient-centered engagement.

VII. FUTURE SCOPE

While this system lays a strong foundation for personalized healthcare, several enhancements can further improve its impact and usability:

- Multi-language Support:** Expanding the system to support multiple languages will ensure accessibility to a diverse global audience, breaking language barriers in healthcare.
- Voice-Based Assistance:** Incorporating voice recognition technology will allow users to interact with the system

through voice commands, providing a hands-free and intuitive experience.

- **Integration with Wearable Devices:** Connecting the system with wearable devices will enable real-time health monitoring, allowing for more precise disease prediction and personalized recommendations based on continuous health data.

With these advancements, the system can evolve into a comprehensive healthcare assistant, promoting preventive care and enhancing medical decision-making through intelligent and data-driven insights.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Prof. Supriya Nalawade, Head of the AIML Department, SSPMCOE, for her guidance and support. Additionally, the authors extend their appreciation to SSPMCOE for providing the necessary facilities and resources for this research.

REFERENCES

- [1] Neelankit Gautam Goswami, Anushree Goswami, Niranjana Sampathila, Muralidhar G. Bairy, Krishnaraj Chadaga, and Sushma Belurkar. Detection of sickle cell disease using deep neural networks and explainable artificial intelligence. **Journal of Intelligent Systems**, 33(1):20230179, 2024.
- [2] Hamid Falah Dheyab, Osman Nuri Ucan, Mohammed Khalaf, and Alaa Hamid Mohammed. Implementation of various types of machine learning approaches for biomedical datasets based on sickle cell disorder. In **2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)**, pages 1–6. IEEE, 2020.
- [3] Sagar Yeruva, B. Pavan Gowtham, Yendluri Hari Chandana, M. Sharada Varalakshmi, and Suman Jain. Prediction of anemia disease using classification methods. In **Machine Learning Technologies and Applications: Proceedings of ICACECS 2020**, pages 1–11. Springer, 2021.
- [4] Ahmad Sabry Mohamad, Muhammad Noor Nordin, and Anasuha Che Senu. Convolutional neural network modeling for classification of human red blood cells in sickle cell disease diagnosis. In **2023 International Conference on Engineering Technology and Technopreneurship (ICE2T)**, pages 241–246. IEEE, 2023.
- [5] Abdourahmane Balde, Awele Bassene, Lamine Faty, Mamadou Soumboundou, Ousmane Sall, and Youssou Faye. Recent artificial intelligence advances in detection and diagnosis of sickle cell disease: A review. In **2023 IEEE International Conference on Big Data (BigData)**, pages 3650–3659. IEEE, 2023.
- [6] Shivangi Shekhar. Prediction of sickle cell anemia disease using machine learning techniques. **Journal of Pharmaceutical Negative Results**, pages 3080–3092, 2022.
- [7] Rima S. Zahr and Santosh L. Saraf. Sickle cell disease and CKD: An update. **American Journal of Nephrology**, 55(1):56–71, 2024.
- [8] Bhawna Dash, Soumyalatha Naveen, and U. M. Ashwinkumar. Enhancing disease diagnosis: Statistical analysis of hematological parameters in sickle cell patients, integrating predictive analytics. **EAI Endorsed Transactions on Pervasive Health and Technology**, 10, 2024.
- [9] Faris Q. Alenzi and Dalal S. AlShaya. Biochemical and molecular analysis of the beta-globin gene in Saudi sickle cell anemia. **Saudi Journal of Biological Sciences**, 26(7):1377–1384, 2019.
- [10] Sanjeda Sara Jennifer, Mahbub Hasan Shamim, Ahmed Wasif Reza, and Nazmul Siddique. Sickle cell disease classification using deep learning. **Heliyon**, 9(11), 2023.
- [11] Kabir O. Olaniran, Harish Seethapathy, Sophia H. Zhao, Andrew S. Allegretti, Sahir Kalim, and Sagar U. Nigwekar. Validity of International Classification of Diseases codes for sickle cell trait and sickle cell disease. **Journal of General Internal Medicine**, 35:1323–1324, 2020.
- [12] Mehrnaz Ahmadi, Samira Beiranvand, Saeed Poormansouri, Mahsa Matbouei, and Camelia Rohani. Sense of coherence or self-efficacy as predictors of health-related quality of life in sickle cell disease patients. **Annals of Hematology**, 102(3):519–528, 2023.
- [13] Biswajeet Acharya, Durga Prasad Mishra, Binapani Barik, Ranjan K. Mohapatra, and Ashish K. Sarangi. Recent progress in the treatment of sickle cell disease: An up-to-date review. **Beni-Suef University Journal of Basic and Applied Sciences**, 12(1):38, 2023.