International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Recent Advances in Computer Applications and Information Technology
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Vitalsphere – AI-Powered Smart Health Diagnostic System

Satyam Tiwari, Shreyashi Ratnaparkhi

G H Raisoni University, Amravati, Maharashtra, India

Abstract

The Medisen: Disease Prediction Website is a unique and intelligent health care solution that has been created to enhance the accessibility and availability of health care services by predicting possible diseases based on the symptoms provided by the user. The rapid growth in technology and the demand for digital health care solutions have led to the creation of this health care solution, which promises to provide a quick, dependable, and easily accessible solution for the identification of diseases. The solution has employed advanced machine learning algorithms like K-Nearest Neighbors (KNN), Decision Tree, and Random Forest to analyze the symptoms provided by the user and compare them with a comprehensive list of diseases in its database. The primary aim of the Medisen platform is to help users determine possible diseases in their early stages without the need for immediate in-person consultations. This is especially important for minor symptoms, which, though they may cause concern, do not necessarily require immediate medical attention. This helps users make informed decisions about whether they should seek immediate professional medical help. This not only saves time and resources but also helps in the early awareness of one's health and how to manage it.

Aside from the prediction of diseases, the platform has an integrated feature that allows users to connect with specialized doctors who deal with the disease in question. This bridges the gap from the prediction of diseases to receiving proper medical treatment from professionals. This helps users receive proper guidance on how to proceed with their healthcare. The platform recommends specialists who can help users receive proper diagnosis and treatment for their diseases. In order for the prediction model to be effective, a comparative analysis was done on various machine learning algorithms. The results show that the accuracy and precision of the model are higher when using the Random Forest algorithm compared to using the KNN and Decision Tree algorithms. Therefore, the Random Forest algorithm was chosen for use in the system for disease prediction.

The Medisen platform is created in a way that is simple and user-friendly for people to use, even for those who do not have a lot of knowledge in technology. All a user needs to do is enter their symptoms, and then the system can process the data and give possible predictions for diseases. Not only can this platform be useful for regular people, but it can also be useful for medical professionals. In conclusion, the Medisen Disease Prediction Website is an essential tool in the development of advanced digital healthcare services. This is particularly because it enables early disease detection, facilitates the accessibility of healthcare services, and reduces the load on healthcare institutions. By incorporating artificial intelligence with healthcare services, the system enables quicker decision-making and encourages individuals to adopt preventive measures for maintaining their health. Therefore, this website presents the possibility of artificial intelligence in the development of advanced healthcare systems.

KEYWORDS*: Symptom-Based Diagnosis, Machine Learning in Healthcare, Random Forest Algorithm, K-Nearest Neighbors (KNN), Decision Tree, Medical Decision Support System, Early Disease Detection, Healthcare Accessibility, Expert Doctor Recommendation, User-Friendly Healthcare Platform, Preventive Healthcare, Digital Health.*

1. Introduction

The way we get help is changing very quickly due to new digital healthcare technologies[3]. Now we can get information about our health and some medical help without having to visit a doctor's office There are many problems with healthcare systems and not enough doctors in some places[1]. This makes it very important to find diseases early and do something about them away. Many people do not visit a doctor when they have problems with their health or symptoms that are not clear because they do not have much time or it is too expensive or they live too far away, from a hospital. Digital healthcare technologies and digital healthcare technologies are helping to make things better. If people wait too long to get medical help, they could become very ill with diseases that could have been prevented with digital healthcare technologies. Machine learning is very good at helping us with problems. It can examine a lot of information and determine what it means. We can use this to guess what disease a person might have based on their symptoms. We examine what happened to people in the past who had the symptoms.

This helps us understand if there is any link between the symptoms and the disease. Therefore, individuals can try to determine what might be wrong with them before visiting a doctor. Machine learning is very efficient in analyzing a lot of data and determining what is significant[2]. K- Nearest Neighbors, Decision Tree, and Random Forest are some examples of machine learning that are widely used in the medical field because they are efficient in organizing information into groups. They can also analyze a lot of information simultaneously. Machine learning algorithms, such as these, are used to help us understand information and make sound decisions[14]. The Vitalsphere Disease. Prediction Website is a website that assists individuals in determining what might be wrong with them. It analyzes the symptoms that people give to the website using computer algorithms. The website tries to figure out what disease they might have. The website sees what people say is wrong with them. Then it checks that against a big list of diseases. Then it tells people what they might have away. The people who made the website wanted to make sure it was right so they tried a few computer programs to see which one was the

best. They found out that the Random Forest program was the one, for figuring out what disease someone might have.

The Vitalsphere Disease Prediction Website uses this program to help people understand what is going on with their health[13]. Vitalsphere does more than just tell you what is wrong with you. It also helps you find a doctor who knows a lot about your health problem. This means that after Vitalsphere figures out what might be wrong with you it can connect you with a healthcare professional who really understands that problem. This helps people get the medical care[4]. The Vitalsphere platform is user-friendly, for those who are not very tech-savvy. Vitalsphere is a tool, for patients and medical professionals. It is designed to be simple and easy to understand so that anyone can use it. Vitalsphere is here to help people get the care they need. By encouraging early disease detection, improving the accessibility of healthcare, and alleviating the pressure on the conventional healthcare system, Vitalsphere helps to democratize the healthcare system[5]. The project showcases the application of machine learning in medical decision-making and illustrates how intelligent web-based platforms can improve the efficiency of healthcare. The showcases the homepage of a website named "Vitalsheper," which emphasizes health care and medical consultation. The header includes the message, "Your Health Is Our Priority," along with a prominent search bar inviting users to input their symptoms. The search bar is central, asking "What symptoms are you experiencing?" to encourage interaction. Below the header, there is a disclaimer message in red, informing visitors that the services provided are intended to complement professional medical advice, urging users to Consult healthcare.

Fig 1 : Predicted Disease and Doctor Recommendations

In the Figure 1: displays the output of a symptom-based prediction model, where the primary predicted disease is shown along with four related diseases. The model ranks these conditions based on the input symptoms, providing a primary diagnosis and alternative possibilities. In addition to the predicted diseases, this figure also presents doctors' information, guiding users towards appropriate healthcare professionals for further consultation.

Fig 1.1 : Other Related Diseases with Recommended Doctor

This figure 1.1 displays two additional diseases related to the primary prediction, along with corresponding doctor recommendations. Similarly, the system will show a total of four related diseases with their respective doctor details, ensuring users have access to comprehensive healthcare information.

2. Literature Review

et al., 2023). The capability to make real-time The application of machine learning (ML) in the healthcare sector has been extensively researched for its applications in disease prediction, providing better diagnostic efficiency, accuracy, and real-time healthcare delivery. Various studies

have emphasized the importance of predictive tools capable of analyzing symptoms and medical information to provide early disease detection, especially in regions where there is a lack of access to healthcare experts. Gharghan et al. (2023) [1] discussed the effect of artificial intelligence (AI) and machine learning algorithms in the healthcare sector, focusing on their ability to provide real-time diagnosis and suggestions based on the input information. The study by Gharghan et al. (2023) emphasizes the increasing use of AI to fill the gap in access to healthcare, which corresponds to Vitalsheper's objective of facilitating early disease detection through accurate symptom predictions (Gharghan predictions about diseases is especially useful in underserved regions where there is a lack of access to healthcare experts. Khan, M., & Al-Mustafa, M. (2021) [2]. talked about a disease prediction system based on symptom input, highlighting the efficiency of machine learning algorithms like Decision Trees and Random Forests in predicting diseases with a high level of accuracy. They also highlighted the need to incorporate healthcare professionals into such systems to improve user engagement after prediction. This is in line with Vitalsheper's strategy of not only predicting diseases but also providing users with information about doctors to consult, thus filling the gap between disease prediction and consultation (Khan et al., 2021). On the same note, Sharma, R., Sharma, S., & Gupta, R. (2022) [3] analyzed the existing digital health solutions and pointed out that most symptom checkers lack post-diagnosis support services, including integration with healthcare providers. This lack of engagement with patients resulted in a need for more inclusive systems that can help patients connect with healthcare providers for further consultations. Vitalsheper solves this problem by including a doctor referral service, which allows patients to easily connect with healthcare providers according to their disease predictions, providing a more comprehensive healthcare experience (Sharma et al., 2022). Patel, V., & Patel, S. (2023) [4] presented a disease prediction system using machine learning that not only predicts diseases but also has a doctor referral system. Their proposed system used algorithms like KNN and Random Forest to predict possible

diseases, followed by giving users contact information of healthcare professionals. This combined system aligns with Vitalsheper's dual interest in both predictive analysis and smooth doctor consultation, thereby improving the overall experience of users in the healthcare domain (Patel & Patel, 2023). Zahra, M., & Karim, A. (2022) [5] also stressed the importance of machine learning algorithms like KNN and Random Forest in developing predictive systems that help in disease prediction and healthcare management. They also stressed the need for combining user-friendly aspects in healthcare platforms to improve patient engagement and accuracy in diagnosis. This study aligns with Vitalsheper's design, which focuses on being a user-friendly and scalable platform for disease prediction and doctor referral systems (Zahra & Karim, 2022).

3. Methodology

The Vitalsheper: Disease Prediction Website Methodology explains the systematic way of approaching the development and implementation of a reliable and user-friendly system for predicting diseases using symptoms inputted by the user [2]. This section explains the major elements involved in the development stage, which include machine learning algorithms for predicting diseases, the frontend and backend technologies, and the incorporation of other features like doctor details and consultation with a healthcare professional [7]. The methodology is centered on making the system user-friendly and easy to navigate for users who lack extensive medical knowledge while still being able to provide accurate and valuable health information. By incorporating machine learning algorithms, a user-friendly interface, and smooth interaction between the frontend and backend technologies, the system is intended to enable users to make well-informed decisions regarding their healthcare and have easy access to medical professionals [1]. Tree, and red for Random Forest. This graph offers a simple way of understanding the performance of each model in making predictions Based on the dataset utilized.

3.1. K-Nearest Neighbours

The K-Nearest Neighbours (KNN) algorithm is a simple and efficient algorithm used for predicting . The basic working of the KNN algorithm involves matching the input symptoms with the already known symptom-disease pairs in the database [2]. The algorithm involves calculating the distance between the input symptom vector and the database points. After calculating the distances, the KNN algorithm selects the 'K' closest data points, also known as neighbours. Based on the majority class (disease) of the neighbours, the KNN algorithm predicts the disease. In the context of disease prediction, the 'K' closest diseases are selected, and the disease with the highest frequency among the neighbours is returned as the most likely disease. The K-Nearest Neighbours (KNN) algorithm is a simple classification algorithm. The algorithm starts with the collection of labelled data points and the choice of the number of neighbours (K). The algorithm selects the K nearest neighbours and the majority class among them. Finally, the majority class is assigned as the predicted output for the new data point, and the algorithm terminates [11].

In the context of disease prediction, 'K' number of closest diseases are selected, and the disease with the highest frequency of occurrence among these 'K' nearest neighbours is predicted as the final output. This means that if the majority of the nearest neighbours belong to a particular disease, the system would predict the new input data as belonging to that particular disease[5]. This ensures a higher level of accuracy in the prediction, as it is not solely dependent on a single comparison. This way, a more accurate prediction can be made, which is meaningful to the user of the system.

The K-Nearest Neighbors (KNN) is a simple classification algorithm that does not require any complex model training. The algorithm begins with a set of labeled data points, along with the number of neighbors, denoted by 'K'. The algorithm begins with the availability of labelled data points and the number of neighbours (K) to be considered [3]. The value of K is of significant importance in the prediction process, as it affects the number of data points considered for prediction purposes.

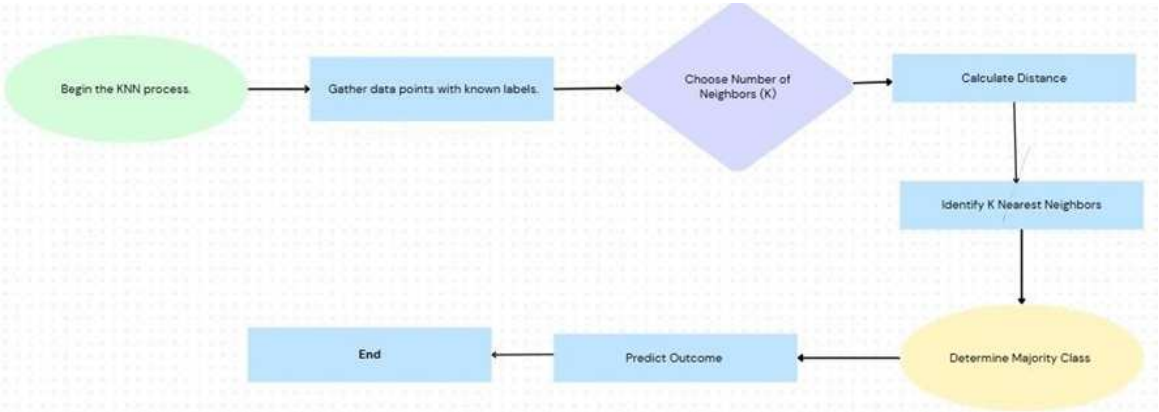

Fig 3.1 : k- Nearest Neigbour (KNN) Algorithm Workflow

3.2. Decision Tree:

The Decision Tree algorithm is a supervised learning classification technique that constructs a tree-like model of decisions given the symptoms [15]. The algorithm breaks down the dataset into smaller subsets by making decisions at each node, based on the symptom that yields the greatest amount of information gain or splitting. The decision tree begins at the root node with the most significant symptom and then expands into multiple possible outcomes based on the characteristics of the symptoms. Each path of the tree corresponds to a decision point, which leads to another symptom or the final prediction (disease). The algorithm terminates when a specified stopping criterion is satisfied, such as reaching the maximum depth or when splitting stops yielding information gain. This decision tree model is simple to understand, as it provides a clear illustration of the decision-making process for predicting a disease given the symptoms [5].

The Vitalsheper disease prediction platform is an effective solution for the growing need for accessible, timely, and accurate healthcare solutions. The platform, which uses machine learning to predict potential diseases based on user-input symptoms, gives users the opportunity to make proactive health choices. The platform's ability to make predictions in real-time and with a high level of accuracy is particularly important in areas where there is a lack of accessible healthcare. One of the most important aspects of the disease prediction platform is its ability to integrate with doctor information, enabling users to quickly connect with medical professionals based on the predicted disease. This aspect of the platform includes the doctor's name, specialization, and contact information, which can be used by users to make proactive health choices.

The model employs three machine learning algorithms: K-Nearest Neighbors (KNN), Random Forest, and Decision Tree. KNN has an accuracy of 84.73%, with smooth decision boundaries but high computational complexity and noise sensitivity [11]. Random Forest, with the highest accuracy of 87.78%, performs well in disease prediction but is prone to overfitting when dealing with noisy data. Decision Tree, with an accuracy of 81.45%, performs well with mixed data types and has low preprocessing requirements but may be noise-sensitive or vary significantly with small data changes. The analysis of the algorithms shows their strengths and weaknesses, with Random Forest being chosen as the main algorithm for disease prediction due to its high accuracy.

Looking forward, there are numerous opportunities that exist in order to further enhance the abilities of Vitalsheper. By increasing the data set to include a broader spectrum of symptoms and diseases, it will be able to make more precise predictions and will be able to assist a broader spectrum of individuals [2]. By utilizing more sophisticated machine learning algorithms and enhancing the existing models, it will be able to offer a broader spectrum of diagnostic assistance and will be able to accommodate a broader spectrum of diseases.

4. Result

The Disease Prediction Website is very effective at what it does. The website provides people with a means of determining what their disease might be based on the symptoms they are experiencing [5]. The website employs computer algorithms such as K Nearest Neighbors and Decision Tree and Random Forest to analyze the symptoms that people provide to it. One of the things about the website is that it has a list of doctors. When people find out what disease they might have, they can look at a list of doctors who specialize in that disease [2]. They can see the doctor's name. What they specialize in, and how to get in touch with them, and where their office is, and when they are working. This is a great way for people to get in touch with a doctor.

The website contains a table that compares the efficiency of K Nearest Neighbors, Decision Tree, and Random Forest in predicting diseases. The table indicates that Random Forest is the most efficient in predicting diseases; therefore, the website uses this program [3]. The people who created the website used ReactJS for the part that people see and Flask for the part that people do not see which makes everything work together smoothly [10]. The Disease Prediction Website is a tool, for people s health because it helps them figure out what disease they might have and learn about health and get in touch with doctors.

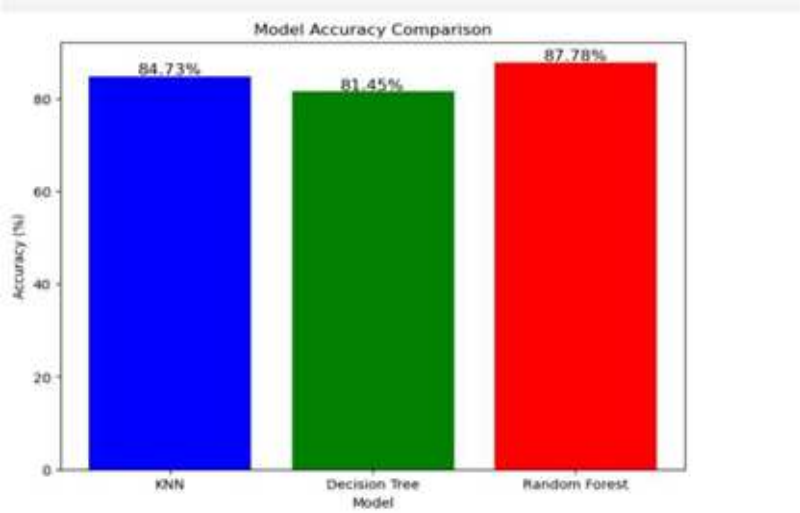

Fig 4 : Model Accuracy Comparison Using Bar Graph

5. Conclusion

Medisen disease prediction platform has been able to address the growing need for accessible, timely, and accurate healthcare solutions [1]. The platform, which uses machine learning to predict potential diseases from various symptoms, gives individuals the power to make informed and proactive decisions regarding their health. The real-time prediction of potential diseases, along with its high accuracy, has been beneficial to individuals living in areas where access to healthcare services is limited or delayed. Another feature of the Medisen disease prediction platform is its inclusion of information regarding doctors, which enables individuals to access relevant and timely information from various doctors depending on the disease they have been predicted to have. The feature includes the doctor's name, specialization, and contact information [4].

The platform uses three algorithms, including K-Nearest Neighbors (KNN), Random Forest, and Decision Tree. The accuracy of the KNN algorithm is 84.73%, which has smoother decision surfaces but is computationally expensive and sensitive to noise. The Random Forest has the highest accuracy of 87.78%, which makes it suitable for disease prediction. However, this algorithm has a high risk of overfitting in the presence of noise. The accuracy of the Decision Tree algorithm is 81.45%, which makes it suitable for use in mixed data types and requires minimal preprocessing [12]. However, this algorithm has a high risk of being sensitive to small variations in data. The comparison of the algorithms shows their advantages and disadvantages, and Random Forest has been chosen for disease prediction on account of its high accuracy.

	Algorithm Used	Advantages	Limitations	Accuracy
1	KNN	Smoother decision surface, less data dependency	KNN is computational ly expensive, sensitive to irrelevant features and noise	84.73%
2	Random Forest	Good accuracy for predicting disease	It is prone to overfitting on noisy data.	87.78%
3	Decision Tree	Handles mixed data types, require minimal preprocessing	It is sensitive to small data changes.	81.45%

Reference

[1] S. Gharghan, M. Khasahmadi, and M. Molla, "AI in Healthcare: A Review on the Role of Machine Learning Models in Disease Prediction," E3S Web of Conferences, vol. 67, p. 01051, 2023.

[2] M. Khan and M. Al-Mustafa, "Symptom-Based Disease Prediction System Using Machine Learning," 2021.

[3] R. Sharma, S. Sharma, and R. Gupta, "A Review of Digital Healthcare Platforms for Disease Prediction," International Journal of Computer Applications, vol. 9, no. 3, pp. 45–52, 2022.

[4] V. Patel and S. Patel, "Integrated Disease Prediction and Doctor Consultation System Using Machine Learning," International Research Journal of Engineering and Technology, vol. 10, no. 1, pp. 58–64, 2023.

[5] M. Zahra and A. Karim, "Machine Learning Models for Healthcare: Disease Prediction and Diagnosis," International Journal of Electronics and Communication Engineering, vol. 15, no. 2, pp. 99–108, 2022.

[6] React Team, "React Documentation," Meta Platforms. [Online]. Available: https://reactjs.org/docs/getting-started.html

[7] Pallets Projects, "Flask Documentation." [Online]. Available: https://flask.palletsprojects.com/en/latest/

[8] Axios Authors, "Axios Documentation." [Online]. Available: https://axios-http.com/docs/intro

[9] Flask-CORS Authors, "Flask-CORS Documentation." [Online]. Available: https://flask-cors.readthedocs.io/en/latest/

[10] M. Grinberg, "How to Create a React + Flask Project." [Online]. Available: https://blog.miguelgrinberg.com/post/how-to-create-a-react-flask-project

[11] Scikit-learn Developers, "K-Nearest Neighbors (KNN) Algorithm." [Online]. Available: https://scikit-learn.org/stable/modules/neighbors.html

[12] Scikit-learn Developers, "Decision Tree Algorithm." [Online]. Available: https://scikit-learn.org/stable/modules/tree.html

[13] Scikit-learn Developers, "Random Forest Algorithm." [Online]. Available: https://scikit-learn.org/stable/modules/ensemble.html#forest

[14] T. Hastie, R. Tibshirani, and J. Friedman, Introduction to Machine Learning.

[15] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830.