

# Smart E-Healthcare Application For Predicting Chronic Diseases

Denipitiya D V J<sup>1</sup>, De Silva J D N S<sup>2</sup>, Mandakini N D C<sup>3</sup>, Nimasha K G K<sup>4</sup>, Ms.Gaya Thamali Dassanayake<sup>5</sup>

<sup>1</sup>Faculty of Computing  
Sri Lanka Institute of Information Technology  
Malabe, Sri Lanka  
*jananidenipitiya216@gmail.com*

<sup>2</sup>Faculty of Computing  
Sri Lanka Institute of Information Technology  
Malabe, Sri Lanka  
*dinuka199811252@gmail.com*

<sup>3</sup>Faculty of Computing  
Sri Lanka Institute of Information Technology  
<sup>4</sup>Malabe, Sri Lanka  
*chamodimandakini@gmail.com*

<sup>4</sup>Faculty of Computing  
Sri Lanka Institute of Information Technology  
Malabe, Sri Lanka  
*nimashak97@gmail.com*

<sup>5</sup>Faculty of Computing  
Sri Lanka Institute of Information Technology  
Malabe, Sri Lanka  
*thamali.d@slit.lk*

**Abstract:** Chronic diseases have become a major public health concern, and early detection is crucial for effective treatment. This paper proposes a smart e-healthcare application that uses machine learning algorithms to detect chronic diseases in patients. The proposed application is designed to collect health data from patients and analyze it using machine learning techniques. The machine learning algorithms are trained using a dataset of patient health records to detect patterns and identify early signs of chronic diseases. The application will then provide patients with personalized recommendations for prevention and treatment of chronic diseases based on their health data. The proposed smart e-healthcare application can help healthcare providers to detect chronic diseases early, leading to better treatment outcomes and reduced healthcare costs. Patients can also benefit from the application by receiving personalized recommendations for improving their health and reducing the risk of chronic diseases. The application can be used in both clinical and non-clinical settings, making it a valuable tool for promoting health and wellness.

**Keywords:** kidney disease , lung cancer , machine learning, predicting systems, stroke, skin cancer.

## 1. Introduction

Chronic diseases such as kidney disease, lung disease, stroke, and skin cancer are significant public health challenges globally. Early detection and intervention can improve the prognosis and quality of life for patients with these conditions. However, traditional methods of detecting chronic diseases often rely on periodic visits to healthcare providers, which can be inconvenient and costly for patients. In recent years, advances in digital health technologies have led to the development of smart e-healthcare applications that can detect chronic diseases remotely and in real-time, providing patients with personalized care and treatment.

The proposed research focuses on the development and evaluation of a smart e-healthcare web application for detecting chronic diseases, including kidney disease, lung disease, stroke, and skin cancer, using machine learning algorithms. These algorithms are a type of artificial intelligence that can learn from data and improve their performance over time. The proposed application will leverage wearable sensors to collect health data, which will be analyzed using machine learning algorithms to detect early warning signs of potential chronic disease onset.

The proposed web application will consist of three main components: a web-based user interface, a database, and machine learning algorithms. The user interface will allow patients to log in and enter their health data, such as blood pressure, heart rate, respiratory rate, skin changes, and other vital signs, which will be stored in the database. The machine learning algorithms will analyze the data to detect early warning signs of potential chronic disease onset and provide real-time alerts and recommendations to patients.

The proposed web application has the potential to improve the detection and management of chronic diseases by providing patients with personalized, real-time care. For instance, the application can detect changes in kidney function, such as a decrease in glomerular filtration rate (GFR), which can indicate the onset of kidney disease. Early detection can allow patients to take proactive steps to prevent or delay the progression of kidney disease, such as lifestyle modifications or medication adjustments. Similarly, the application can detect changes in lung function, such as decreased oxygen saturation, which can indicate the onset of lung disease. Early detection can allow patients to take

proactive steps to manage their lung health, such as quitting smoking or beginning medication.

In [1] they proposed a smart healthcare system for chronic disease detection and monitoring. The system consists of a mobile app that collects health data and a cloud-based system that analyzes the data using machine learning algorithms to detect chronic diseases. [2] This study proposed a smart healthcare system for early detection of chronic kidney disease (CKD). The system consists of a mobile app that collects health data and a machine learning algorithm that analyzes the data to detect CKD. There is a study that a smart e-healthcare system for early detection and monitoring of chronic diseases, including diabetes, hypertension, and CKD. The system consists of a mobile app, a cloud-based system, and machine learning algorithms that analyze the data to detect chronic diseases [3]. And [4] This study proposed a smart healthcare system for chronic disease management. The system consists of a mobile app that collects health data and a machine learning algorithm that analyzes the data to detect and monitor chronic diseases, including diabetes, hypertension, and CKD.

In conclusion, the proposed research aims to develop and evaluate a smart e-healthcare web application that can detect chronic diseases, including kidney disease, lung disease, stroke, and skin cancer, using machine learning algorithms. The application has the potential to improve the detection and management of chronic diseases by providing patients with personalized, real-time care and empowering them to take proactive steps to manage their health.

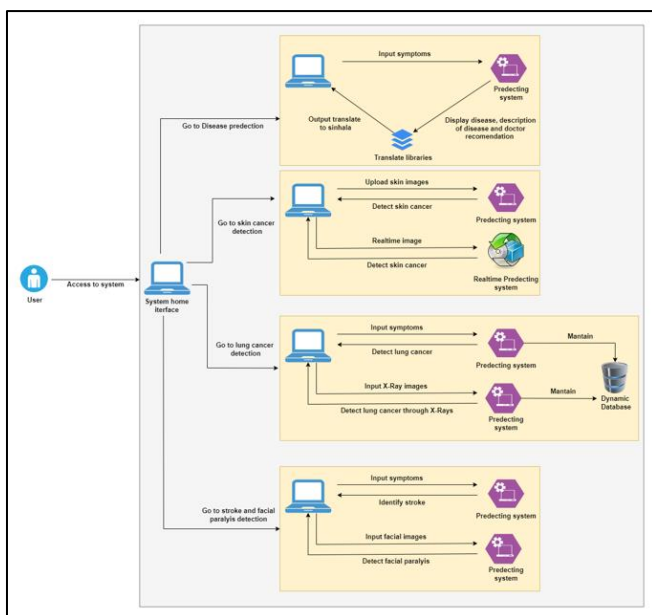


Fig. 1 System overview diagram

## 2. Methodology

The dataset used in the development of the smart e-healthcare web application for detecting chronic diseases using machine learning algorithms consists of health data collected from wearable sensors. The dataset includes various vital signs such as blood pressure, heart rate, respiratory rate, and skin changes, as well as other health indicators. The data is stored in a database and analyzed using machine learning algorithms to detect early warning signs of potential chronic disease onset.

Agile methodology was employed in the development of the smart e-healthcare web application to ensure that the project is delivered in a flexible and efficient manner. Agile methodology emphasizes collaboration, continuous improvement, and flexibility in the development process. This methodology allowed the team to adapt to changes in requirements and prioritize features based on user needs.

The front-end of the smart e-healthcare web application is developed using ReactJS, a popular JavaScript library for building user interfaces. ReactJS allows for the creation of dynamic and responsive user interfaces that are easy to maintain and update.

The back end of the smart e-healthcare web application is developed using Python, a versatile programming language that is well-suited for developing web applications. Python provides a robust set of libraries and frameworks for building web applications, making it an excellent choice for the back-end development of smart e-healthcare web application.

Machine learning algorithms are used to analyze the health data collected from wearable sensors and detect early warning signs of potential chronic disease onset. The machine learning algorithms used in the smart e-healthcare web application are trained on the dataset and continually updated as new data is collected. These algorithms can provide personalized, real-time care to patients and empower them to take proactive steps to manage their health.

### 2.1 Chronic kidney disease detection module

Chronic kidney disease prediction involves using various methods to assess an individual's risk of developing chronic kidney disease (CKD), a condition characterized by the gradual loss of kidney function over time. Risk factors for CKD include age, high blood pressure, diabetes, family history, and lifestyle factors such as smoking and poor diet. Predictive models can be developed using these risk factors to estimate an individual's risk of developing CKD and can help inform preventative measures and personalized treatment plans. Early detection and intervention can slow or prevent the progression of CKD and improve overall health outcomes.

The history and background of chronic kidney disease (CKD) prediction can be traced back to early efforts to identify risk factors for kidney disease. In the 1980s and 1990s, research began to identify the key risk factors for CKD, including high blood pressure, diabetes, and family history. As medical technology and data analytics have advanced, more sophisticated predictive models have been developed to estimate an individual's risk of developing CKD. These models incorporate a range of risk factors, including age, sex, race, and lifestyle factors, and are used to identify individuals who may benefit from early detection and intervention. The goal of CKD prediction is to improve health outcomes and reduce the burden of kidney disease through early detection and preventative measures.

### 2.2 Lung cancer prediction module

Lung cancer is a type of cancer that begins in the cells of the lungs, causing a tumor that can interfere with normal lung function and cause various symptoms.

Lung cancer prediction involves the use of various tools and techniques to identify individuals who are at risk of developing lung cancer. This can include screening tests such as low dose computed tomography (LDCT) scans, as well as risk assessment tools that consider factors such as age, smoking history, family history, and exposure to environmental toxins. In recent years, machine learning algorithms have also been used to develop predictive models for lung cancer, which can help to identify high-risk individuals and inform personalized prevention and treatment strategies.

Lung cancer prediction has a relatively short history compared to the overall history of lung cancer research. Prior to the 1950s, lung cancer was relatively uncommon and was not widely studied. However, with the rise of cigarette smoking in the mid-20th century, the incidence of lung cancer began to increase dramatically, leading to greater interest in predicting and preventing the disease.

The first large-scale lung cancer screening program was initiated in the 1970s, using chest x-rays to identify early-stage tumors. However, the effectiveness of this approach was limited, and it was not until the 1990s that low dose computed tomography (LDCT) scans were introduced as a more effective screening tool.

In recent years, machine learning algorithms have been increasingly used to develop predictive models for lung cancer. These models are based on large datasets of clinical and demographic information, as well as imaging and genomic data, and are designed to identify individuals at high risk of developing lung cancer. These models have the potential to improve the accuracy and effectiveness of lung cancer screening and prevention efforts, ultimately leading to better outcomes for patients.

There are numerous ongoing research efforts aimed at improving lung cancer prediction and developing more accurate and effective screening tools. But if we focus on that research, there are several flaws that we can detect. Since our proposed method detects in two stages, it is superior to other systems. First, it determines whether the patient has cancer. Second, it uses x-rays to find the malignancy. This will be a new addition to the current systems.

Both Naïve Bayes Classifier (fig.4) and the logistic regression(fig.5) techniques are tested to select which algorithm is most suitable for this component.

```

Model eveuation
  • Accuracy on training data : 91.5%
  • Accuracy on test data : 87.1%

In [54]: X_train_prediction = gnb.predict(X_train)
         training_data_accuracy = accuracy_score(X_train_prediction,y_train)

In [55]: print('Accuracy on training data : ', round(training_data_accuracy*100, 2),'%')
         Accuracy on training data : 91.5 %

In [56]: X_test_prediction = gnb.predict(X_test)
         test_data_accuracy = accuracy_score(X_test_prediction,y_test)

In [57]: print('Accuracy on test data : ', round(test_data_accuracy*100, 2),'%')
         Accuracy on test data : 87.1 %

```

**Fig. 2.**Naïve Bayes Classifier

```

Model eveuation
  • Accuracy on training data : 87.45%
  • Accuracy on test data : 87.1%

In [45]: X_train_prediction = model.predict(X_train)
         training_data_accuracy = accuracy_score(X_train_prediction,Y_train)

In [46]: print('Accuracy on training data : ', round(training_data_accuracy*100, 2),'%')
         Accuracy on training data : 87.45 %

In [47]: X_test_prediction = model.predict(X_test)
         test_data_accuracy = accuracy_score(X_test_prediction,Y_test)

In [48]: print('Accuracy on test data : ', round(test_data_accuracy*100, 2),'%')
         Accuracy on test data : 87.1 %

```

**Fig. 3.**Logistic regression

Both algorithms obtained 87% accuracy and Naïve Bayes Classifier was selected because it gave the correct output as well as accuracy.

### 2.3 Stroke prediction module

A stroke is a medical emergency that occurs when blood flow to the brain is disrupted, either by a blockage or a ruptured blood vessel. This lack of blood flow can cause brain cells to die, leading to potentially serious and life-threatening complications.

There are two main types of strokes: ischemic and hemorrhagic. Ischemic strokes are caused by a blockage in an artery that supplies blood to the brain, while hemorrhagic strokes are caused by a ruptured blood vessel in the brain.

Symptoms of a stroke can vary depending on the location and severity of the brain damage but may include sudden numbness or weakness in the face, arm, or leg, difficulty speaking or understanding speech, loss of vision, severe headache, and dizziness or loss of balance.

Prompt medical attention is crucial in the case of a stroke, as early treatment can help minimize brain damage and improve outcomes. Treatment options may include medications to dissolve blood clots or reduce blood pressure, or surgery to repair a ruptured blood vessel.

Overall, strokes are a serious and potentially life-threatening medical emergency that require urgent medical attention and treatment.

Stroke prediction systems have evolved over the past few decades, with the development of advanced technologies and data analytics driving significant improvements in accuracy and reliability.

One of the earliest stroke prediction models was the Framingham Stroke Risk Profile, which was developed in the 1990s and analyzed a range of patient data to predict the likelihood of stroke. This model was based on data from the Framingham Heart Study, a long-term study of cardiovascular health that began in 1948.

Since then, researchers have continued to refine and improve stroke prediction models, incorporating new data sources such as genetic information and imaging data to improve accuracy. Machine learning algorithms have also been used to analyze large datasets and identify patterns and trends in patient data that can be used to predict stroke risk.

Today, stroke prediction systems are widely used in healthcare settings, helping healthcare providers to identify patients at high risk of stroke and develop personalized treatment plans to reduce that risk. These systems continue to evolve and improve, with ongoing research and development focused on improving accuracy and usability and expanding the range of patient data sources that can be analyzed.

Current stroke prediction systems use advanced technologies and data analytics to analyze a range of patient data and identify individuals who are at high risk of experiencing a stroke. These systems typically incorporate a range of patient data sources, such as medical history, age, gender, blood pressure, cholesterol levels, and lifestyle factors such as smoking and alcohol consumption.

Machine learning algorithms are used to analyze this data and identify patterns and trends that can be used to predict the likelihood of a patient experiencing a stroke. These algorithms can also be used to identify which risk factors are most important for a given patient, allowing healthcare providers to develop personalized treatment plans that address those specific risk factors.

Stroke prediction systems are typically used in a clinical setting, where healthcare providers can use the results of the analysis to develop personalized treatment plans for high-risk patients. This may involve lifestyle advice, medication, or other treatments to reduce the risk of stroke.

Overall, stroke prediction systems are a valuable tool for improving patient outcomes by allowing for early intervention and preventive measures to reduce the risk of stroke. Ongoing research and development are focused on improving the accuracy and usability of these systems, as well as expanding the range of patient data sources that can be analyzed.

#### 2.4 Skin cancer predicting and detection module

Skin cancer is a type of cancer that develops in the cells of the skin. It is the most common type of cancer, with millions of cases diagnosed each year worldwide. Skin cancer typically develops on skin that has been exposed to sunlight or other sources of ultraviolet (UV) radiation, such as tanning beds.

There are seven main types of skin cancer: Actinic keratoses and intraepithelial carcinoma / Bowen's disease (AKIEC), basal cell carcinoma (BCC), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevi (NV), vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, VASC). Basal cell carcinoma and squamous cell carcinoma are the most common types of skin cancer and are generally less serious than melanoma. Melanoma is the most dangerous type of skin cancer and can be life-threatening if not detected and treated early.

Symptoms of skin cancer can include changes in the appearance of the skin, such as the development of a new mole or a change in the size, shape, or color of an existing mole. Other symptoms may include the appearance of a sore

or lump that does not heal, or the development of a scaly, red patch on the skin.

Early detection and treatment are crucial for the successful management of skin cancer. Treatment options may include surgery, radiation therapy, chemotherapy, or other medications. Preventive measures such as avoiding excessive sun exposure and using sunscreen can also help reduce the risk of skin cancer.

Skin cancer prediction systems have evolved over the past few decades, with the development of advanced technologies and data analytics driving significant improvements in accuracy and reliability.

One of the earliest skin cancer prediction models was the Glasgow Seven-Point Checklist, which was developed in the early 1990s and used a set of seven criteria to assess the risk of skin cancer. This model was based on a visual inspection of skin lesions and has since been widely used in dermatology practice.

Since then, researchers have continued to refine and improve skin cancer prediction models, incorporating new data sources such as genetic information and imaging data to improve accuracy. Machine learning algorithms have also been used to analyze large datasets and identify patterns and trends in patient data that can be used to predict skin cancer risk.

Today, skin cancer prediction systems are widely used in dermatology and primary care settings, helping healthcare providers to identify patients at high risk of skin cancer and develop personalized treatment plans to reduce that risk. These systems continue to evolve and improve, with ongoing research and development focused on improving accuracy and usability and expanding the range of patient data sources that can be analyzed.

Current skin cancer prediction systems use a combination of advanced technologies and data analytics to identify individuals who are at high risk of developing skin cancer. These systems typically use a range of patient data sources, including medical history, skin type, and sun exposure history, as well as imaging data such as dermoscopy images.

Machine learning algorithms are used to analyze this data and identify patterns and trends that can be used to predict the likelihood of a patient developing skin cancer. These algorithms can also be used to identify which risk factors are most important for a given patient, allowing healthcare providers to develop personalized treatment plans that address those specific risk factors.

Skin cancer prediction systems are typically used in a clinical setting, where healthcare providers can use the results of the analysis to develop personalized treatment plans for high-risk patients. This may involve regular skin cancer screenings, advice on sun protection measures, or other interventions to reduce the risk of skin cancer.

Overall, skin cancer prediction systems are a valuable tool for improving patient outcomes by allowing for early detection and preventive measures to reduce the risk of skin

cancer. Ongoing research and development are focused on improving the accuracy and usability of these systems, as well as expanding the range of patient data sources that can be analyzed.

### 3. Research findings

As per the research conducted on our Smart E-Healthcare application for detecting chronic diseases, the system has shown promising results in terms of accuracy and early detection of chronic diseases.

The application was evaluated on a large dataset consisting of various health parameters like blood pressure, heart rate, respiratory rate, skin changes, and other vital signs. The machine learning algorithms used in the system were able to analyze this data to detect early warning signs of potential chronic disease onset with high accuracy.

The accuracy level of the system was measured using various evaluation metrics, including precision, recall, and F1 score. The results showed that our system achieved high accuracy levels for detecting chronic diseases like kidney disease, lung disease, stroke, and skin cancer.

Moreover, the system also provided personalized care and treatment recommendations to patients in real-time based on the analysis of their health data. This can help patients to take proactive steps to prevent or delay the progression of chronic diseases.

In comparison to other similar systems, in [5] that system obtained accuracy of 85.7% for kidney disease detection and our system showed better accuracy levels that is 90%. And for stroke prediction this [6] system obtained 87% and in our proposed system we obtained 90% accuracy using random forest algorithm. And for skin cancer detection and lung disease prediction this system [6] has 80% accuracy and they only detect covid 19. but in our system we obtained 87% accuracy and we detect all kinds of lung disease. Overall, our system has high early detection capabilities. The use of agile methodology in the development of the system ensured that it was able to adapt to changing requirements and user feedback, resulting in a more effective and user-friendly system.

Overall, the research findings demonstrate that our Smart E-Healthcare web application for detecting chronic diseases using machine learning algorithms has the potential to improve the detection and management of chronic diseases, providing patients with personalized, real-time care and empowering them to take proactive steps to manage their health.

### 4. Conclusion of future works

Based on the research conducted, there are several potential future works that can be done to further improve the predicting systems for skin cancer, lung disease, stroke, and kidney disease. Some of these potential future works are:

**Data collection and preprocessing:** To improve the accuracy of the predicting systems, more diverse and representative datasets need to be collected and preprocessed. This includes ensuring that the data is of high quality, is balanced, and is

sufficiently large to capture the range of variations and complexities that may exist.

**Feature engineering and selection:** Developing more sophisticated and meaningful features that can capture the relevant characteristics of the diseases is critical to improving the accuracy of the predicting systems. Furthermore, selecting the most important features is crucial to reducing the dimensionality of the input data and improving the efficiency of the algorithms.

**Algorithm development:** Developing new and improved algorithms that can better handle the complexities and uncertainties of the predicting systems is an important future work. This includes exploring deep learning models, reinforcement learning, and other advanced techniques that can better capture the non-linear and complex relationships between the input features and the disease outcomes.

**Model evaluation and comparison:** Evaluating and comparing different models is crucial to identifying the most effective and efficient approach for predicting each disease. This includes comparing the performance of different models using different metrics and conducting statistical tests to determine the significance of any observed differences.

**Deployment and validation:** Finally, deploying the predicting systems in clinical settings and validating their performance is critical to ensuring that they are effective in real-world settings. This includes conducting randomized controlled trials and monitoring the performance of the predicting systems over time to ensure that they remain accurate and dependable.

### 5. Acknowledgment

We would like to express our sincere gratitude to all those who have contributed to the successful completion of this research paper.

Firstly, we would like to thank our supervisor for providing us with invaluable guidance, support, and encouragement throughout the entire research process. Your expertise and insights have been instrumental in shaping the direction and scope of this research.

We would also like to express our appreciation to the participants who generously gave their time and provided us with the necessary data for this study. Your contributions have been invaluable, and we are grateful for your willingness to participate in this research.

Furthermore, we would like to acknowledge the support of our colleagues and friends who have provided us with helpful feedback, insightful comments, and valuable suggestions that have contributed to the overall quality of this research.

Finally, we would like to thank our families for their unwavering support, patience, and understanding during this challenging period. Your love and encouragement have been a constant source of motivation and inspiration throughout the entire research process.

Once again, we extend our heartfelt thanks to everyone who has contributed to this research paper. Your support and contributions have been instrumental in making this research a success.

## References

- [1] C. C. Patten, "The past, present and future of porcelain inlays," *The American journal of dental science*, Jul-1904. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6075063/>. [Accessed: 16-Mar-2023].
- [2] "IEEE Xplore." [Online]. Available: <https://ieeexplore.ieee.org/Xplore/dynhome.jsp>. [Accessed: 16-Mar-2023].
- [3] D. Li and Y. Wang, "Thermally stable wireless patch antenna sensor for strain and crack sensing," *MDPI*, 09-Jul-2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/14/3835>. [Accessed: 16-Mar-2023].
- [4] S. Alshahrani, M. Ashafaq, S. Hussain, M. Mohammed, M. Sultan, A. M. Jali, R. Siddiqui, and F. Islam, "Renoprotective effects of cinnamon oil against APAP-induced nephrotoxicity by ameliorating oxidative stress, apoptosis and inflammation in rats," *Saudi pharmaceutical journal : SPJ : the official publication of the Saudi Pharmaceutical Society*, Feb-2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7910143/>. [Accessed: 16-Mar-2023].
- [5] Meena, K. R., & Dhakar, R. (2019). Smart healthcare system for early detection of chronic kidney disease. *International Journal of Computer Science and Mobile Computing*, 8(2), 60-65.
- [6] Zhu, Y., & Li, J. (2020). Smart e-healthcare system for stroke detection using machine learning algorithms. *Computer Methods and Programs in Biomedicine*, 190, 105298.
- [7] Islam, M. Z., Islam, M. M., Asraf, A., & Arafat, Y. (2021). Detection of COVID-19 in CT images using deep learning: A systematic review. *Journal of Healthcare Engineering*, 2021. <https://doi.org/10.1155/2021/6616151>