

Early Disease Prediction using Healthcare Datasets

Shantanu Sunil Patankar

Department of Science and Technology, G.H.Raisoni Skill Tech University, Nagpur, India

ABSTRACT

The digital healthcare data is growing fast and this has created new chances for using data analytics and machine learning to make better medical decisions. Finding diseases early is very important to reduce the number of deaths lower the cost of treatment and improve the health of patients.. The old ways of diagnosing diseases often need people to analyze data manually and sometimes they fail to find diseases early because the symptoms are similar and the health conditions are complex.

This research paper is about a system that uses data to predict diseases. This system uses healthcare data that has information about the patients, their medical history, symptoms and the results of their tests. The system first cleans up the data then it analyzes the data to understand it and finally it uses machine learning to find patterns and predict the likelihood of diseases early.

The system tries out predictive models like Logistic Regression, Decision Tree, Random Forest and Support Vector Machine. It checks how well these models work by looking at things like how accurate they're how precise they are how well they remember things and their F1-score. The results show that machine learning models can really help analyze healthcare data and assist doctors in making decisions. The system shows how important data analytics is in making healthcare better and in supporting early diagnosis of digital healthcare data. The use of healthcare data in this system is a key factor, in its success.

Keywords: Early Disease Prediction, Healthcare Analytics, Machine Learning, Data Mining, Predictive Modeling, Exploratory Data Analysis, Healthcare Datasets, Data Analytics.



This is an open-access article under the [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

I. Introduction

Within Healthcare organizations create a lot of data from things like health records, lab reports and patient monitoring devices. Usual diagnostic methods often rely on checks and doctors experience. This can delay disease detection. Many illnesses progress quietly without symptoms. This makes it hard to predict them early.

Data analytics and machine learning are tools. They help analyze healthcare datasets to find patterns linked to disease occurrence. We can use models on information such as age, medical history, symptoms and test results. This helps predict the likelihood of diseases before they get severe. Healthcare data is used to make these predictions. This research is, about creating a healthcare data analytics system. The system predicts diseases early.

The system includes:

- **Data preprocessing**
- **Feature selection**

Machine learning algorithms to generate insights

The goal is to support healthcare professionals. We want to help them make decisions and improve healthcare strategies with healthcare data. Healthcare data analytics plays a role in this. It helps healthcare professionals make decisions. They can use data to improve healthcare strategies.

Proposed AI-Based Solution

The new system uses machine learning to look at healthcare information and figure out who might get sick. The system takes information and does a few things to it like getting it ready finding important details teaching the system and then making predictions. The system that makes predictions finds connections, between how healthy someone's whether they get sick. When it looks at these connections the system can tell how likely it is that a new patient will get sick.

The system that uses intelligence wants to help doctors and nurses by giving them information that is based on real data. This information can help them find out if someone is sick on and stop them from getting very sick.

Motivation

The main reason for this research is that data-driven healthcare systems are becoming more and more important. Finding diseases early can really help lower the cost of treatment and make patients more likely to survive. A lot of healthcare data isn't being used as much as it could be because there aren't good systems, in place to analyze it. When we use machine learning on healthcare data we can find patterns that weren't obvious and get information that can help us make predictions. This project is trying to connect healthcare data with systems that can help make decisions. We want to show how using data can help predict diseases and plan ways to prevent them.

II. Related Work

Machine learning is really useful for figuring out if someone has a disease. People have used things like Decision Trees and Random Forests and Support Vector Machines to look at information.

These things are like tools that help us understand the data. When we use them on the kind of healthcare data they can be really good at predicting what is going on. Some people have even used machine learning to try to figure out if someone has heart disease or diabetes or cancer. They use things like what the doctor sees and what the tests say. Machine learning is pretty cool because it can really help us make sense of all the healthcare information. We have to make sure we are using the right data and looking at the right things. If we do that then machine learning can be really good at helping us predict what is going on with someones health. Machine learning is about using computers to help us understand things and it can be really helpful, in healthcare.

III. Research Methodology

People have done a lot of research on using machine learning to predict diseases. They used tools like Decision Trees and Random Forests to look at medical information.

Machine learning is really good at predicting diseases when it has the information to work with. Some people looked into using machine learning to predict diseases like heart disease and diabetes. To make a system that can predict diseases you have to do a things. First you collect information from sources like Kaggle. Then you clean up the information to make sure it is correct and complete. After that you use tools to understand what the information means.

Next you use machine learning to make a model that can predict diseases. You test the model to see how well it works. Finally you pick the model to use for predicting disease risks.

For example machine learning can be used to predict heart disease, diabetes and cancer by looking at things like blood pressure and blood sugar levels. These studies show that machine learning can be really helpful in making healthcare better.. To get good results you have to make sure the information

is correct and you use the right tools to evaluate the models. Machine learning is a tool, for disease prediction and it can help doctors make better decisions.

Phases of Disease Project Development

The Early Disease Prediction System is made using Healthcare Datasets. We do this in an organized way. This helps us get the information and make good predictions. Each step is important. It helps turn healthcare data into useful information. This information helps doctors find out what is wrong with people, on. The main steps to make the Early Disease Prediction System are explained below.

1. **Data Collection** : We get this healthcare data from places like Kaggle or healthcare research databases. The healthcare data we collect is usually stored in files like CSV files or databases. Collecting healthcare data is very important because it affects how accurate our disease prediction model is. We need to make sure the healthcare data is good so our disease prediction model can make predictions.

2. **Data Preprocessing** : The healthcare data we collect can be messy with a lot of information. So we need to organize the healthcare data to make it usable. This means we handle missing values remove duplicates and correct errors in the healthcare data. We also make sure the healthcare data is consistent and convert the healthcare data into numbers. This step makes sure our healthcare data is clean and ready for machine learning analysis. We need healthcare data to make good predictions with our disease prediction model.

3. **Exploratory Data Analysis** : We use things like histograms and charts to examine the healthcare data. Exploratory Data Analysis helps us find features that can help us predict diseases with our disease prediction model and understand how the healthcare data behaves overall. By analyzing the healthcare data we can see what the healthcare data tells us about diseases.

4. **Feature Selection** : Here we choose the features that're important for predicting diseases with our disease prediction model. We use things like correlation analysis and statistical tests to choose the features from the healthcare data. By removing features that're not important we make the healthcare data less complex and our machine learning models more efficient and accurate at predicting diseases. We focus on the features in the healthcare data.

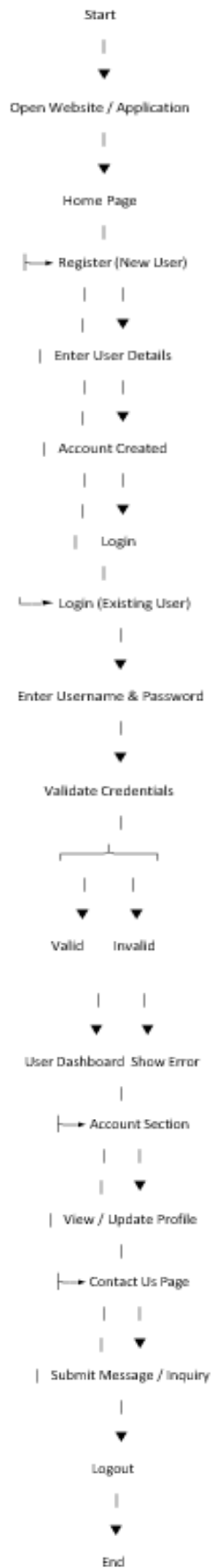
5. **Model Development** : Now that our healthcare data is ready we develop machine learning models to predict diseases. We use algorithms like Logistic Regression, Decision Tree and Support Vector Machine on the healthcare data. We split our healthcare data into two parts: one for training our disease prediction model and one for testing our disease prediction model. The training healthcare data teaches our disease prediction model to recognize patterns of diseases while the testing healthcare data checks how well our disease prediction model works.

6. Model Evaluation :

After we train our disease prediction models we evaluate how well they work to see how accurate and reliable they are at predicting diseases. We use things like accuracy, precision, recall and F1-score to see how effective our disease prediction models are. This step helps us compare algorithms and choose the one for our disease prediction model. We want to choose the disease prediction model.

7. **Visualization** : In this step we use our chosen disease prediction model to predict the risk of diseases for the healthcare data. Our disease prediction model looks at the input healthcare data. Estimates how likely it is that a disease will occur. We show the results using graphs and charts. These visuals help healthcare professionals understand the results and make decisions about diagnosis and preventive care using our disease prediction model. By following these steps we make sure our disease prediction system is developed efficiently using the healthcare data. This approach helps us turn healthcare data into insights that support early disease detection and better healthcare decision-making, with our disease prediction model.

➤ **System Flowchart (Text Representation)**



IV. Project Development

1. Requirement Analysis Phase

When we start working on the Project Development we need to figure out what the system needs to work. We are looking for healthcare datasets that have information about patients, such as how old they're if they are male or female what symptoms they have and what their clinical test results show. To do this Project Development we need to use the Python programming language, libraries that help with machine learning and tools that help us visualize the data. The healthcare datasets are very important for this Project Development.

2. System Design Phase

Now we design the system to work with the healthcare data and give us predictions. The Project Development system has parts, including one for getting the data ready one for picking the right features one for training the models and one for showing us the results. The system is designed to handle healthcare data and the Project Development system is what we are focusing on. We want to make sure the Project Development system works with the healthcare data.

3. Implementation Phase

In this phase we actually write the code for the disease prediction system using Python for the Project Development. We use libraries like Pandas and NumPy to work with the data and libraries like Matplotlib and Seaborn to visualize the data for the Project Development. We also use Scikit-learn to develop the models for the Project Development. The Implementation Phase is a part of the Project Development, where we put everything together.

4. Testing Phase

We test the Project Development system to make sure it is working correctly and giving us results. We use data that the models have not seen before to test how well they can predict things for the Project Development. This is a part of the Project Development to make sure the system is reliable and accurate. The Testing Phase is very important, for the Project Development system because we need to know that it works well.

Data Preprocessing and Analysis

- 1. Collecting Information :** We get healthcare datasets from places where people share information. These datasets have lots of details like what's wrong with the patients, who they are and what the tests say about them.
- 2. Data Cleaning :** We clean up the data to get rid of things that're missing things that are the same and things that do not make sense. This makes the data better and more trustworthy.
- 3. Data Transformation :** We change the data so it is ready for the machines to look at it.

➤ We work with words as part of Data Transformation:

We take things like what's wrong with the patient or what the doctor says and we turn them into numbers using special codes.

➤ We also make sure everything is for Data Transformation:

We make sure all the numbers are on the same level so they all count the same when we try to predict something.

- 4. Data Analysis : Descriptive Analysis :** We look at the data and calculate things like the average the middle number and how spread out the numbers are. This helps us understand what the data is like.

➤ **Correlation Analysis** : We look for connections, between how healthy someone's what happens to them when they are sick.

➤ **Feature Selection** : We pick the things that help us predict what will happen to someone when they are sick. We use tools to figure out what those things are.

➤ **Visualization** : We use pictures, charts and special maps to see what is going on in the data and how things are connected.

✓ **Bar charts are one way to do this.**

✓ **Word clouds are another way.**

✓ **Graphs also help to show the data patterns.**

Outcome

The result of going through all the information and making sense of it is a set of data that's clean and easy to understand. This clean data is really useful, for teaching computers to predict when someone will get a disease. The computers can learn from the data and get better at predicting diseases. The data is structured in a way that makes it easy to use for this purpose. We can use the structured data to train computers to predict diseases.

The outcome of this project is a website that helps users access services easily.

➤ **It has a home page that helps users find what they need.**

➤ **There is an account section where users can manage their information.**

➤ **Users can also contact administrators through the contact us page.**

This project helps users create and manage their accounts safely. In the account section users can update their details. This makes sure their information is always correct and current. The account section makes the website more user-friendly and functional. The contact us page allows users to send messages give feedback or report problems to administrators. This helps administrators and website users communicate better. As a result problems, with the website get solved faster. Website users are happier. The website combines web technologies, database management and user interface design well. The project meets its goal of creating a platform. The website platform makes it simple for website users to interact with it. It also keeps website user data safe and easy to access.

V. Proposed Algorithm

The system we are talking about uses some classification algorithms like:

➤ **Logistic Regression**

➤ **Decision Tree**

➤ **Random Forest**

➤ **Support Vector Machine**

✓ **Logistic Regression**

In healthcare logistic regression helps us understand how likely a person is to get a disease based on things like their age, blood pressure and medical history. It is different from linear regression, which tries to predict numbers that can be anything. Logistic regression uses a function to turn its predictions into probabilities that are between 0 and 1. If the probability is than 0.5 the system says the person has the disease. If it is than 0.5 it says the person is healthy. Logistic Regression is used a lot in research because it is easy to understand and it gives us clear ideas about how different health things affect whether or not someone gets a disease. For example it can show us how things like smoking being overweight or having high blood pressure can increase the chance of getting heart

disease or diabetes. In this project we use regression as a basic model to predict if someone will get a disease. We then compare how well it works with machine learning algorithms.

✓ **Decision Tree**

The model looks like a tree, where each part of the tree's a test on a feature each branch is what happens after the test and each leaf is the final prediction. In healthcare a decision tree can look at things about a person, like their age, weight, blood sugar levels and symptoms to figure out how likely they are to get a disease. The model might first look at the persons age their cholesterol levels and then other health things before making a prediction. One good thing about decision trees is that they are easy to understand. Doctors can see how the model made its decision because it is an logical process. However decision trees can sometimes be too good at fitting the data they are trained on which means they might not work well with new data. To fix this we use advanced methods like Random Forest.

✓ **Random Forest**

Random Forest is an advanced machine learning algorithm that is part of a group called ensemble learning methods. It works by making decision trees during training and combining their predictions to get a more accurate answer. Of using just one decision tree the random forest algorithm makes many trees using different parts of the data and features. Each tree makes its prediction and the final prediction is what most of the trees agree on. This makes the predictions more accurate and less likely to be too good for the training data. Random Forest is especially good with data that has many features and non-linear relationships. In healthcare random forest models are often used to find the important health factors that affect whether someone gets a disease. The algorithm can also give us scores that show which health things have the impact on disease risk. Because it is so accurate and reliable Random Forest is often one of the algorithms for predicting diseases.

✓ **Support Vector Machine (SVM)**

Support Vector Machine (SVM) is a machine learning algorithm that is mainly used for classification problems. The goal of SVM is to find the line or plane that separates the data into different groups. When it comes to predicting diseases SVM looks at a persons health data. Tries to find the best line that separates people who have a disease from those who do not. The algorithm focuses on the data points that're closest to the line, which are called support vectors and these points are crucial in determining the best way to separate the data. SVM is particularly good with data that has features. It can also handle -linear relationships by using special functions that transform the data into higher-dimensional spaces where it is easier to separate the data. Although SVM can be slower, than models it often gives very accurate predictions, especially when the data has complex patterns. In this project we use SVM to analyze healthcare data and compare it to machine learning models like logistic regression, decision trees and random forest.

Random Forest usually does a better job of predicting things because it uses a lot of different models to make a decision. This is what we call learning and it is used in Random Forest. Random Forest is often very good, at predicting things because of this.

Proposed Solution

This system checks things like age and medical history to find patterns that show when a disease might happen. The main goal of the Early Disease Prediction Using Healthcare Datasets project is to help doctors make decisions by giving them useful information. We do this by looking at healthcare data and using techniques to find patterns. The Early Disease Prediction Using Healthcare Datasets project uses these patterns to find health risks before they become problems. We follow some steps to make this system. First we get healthcare data from sources. This data has things like age, gender and test results. Then we make sure the data is clean and good to use. After that we look at the data to see how different things are connected to diseases. We use pictures like charts and graphs to see these

connections. This helps us find the things that help us predict diseases. Next we pick the things that help us predict diseases. This makes the system better and faster. Then we use computer programs like Logistic Regression and Decision Tree to make models that can predict diseases. We teach these models using healthcare data. Then we test them to see how well they work. We pick the model and use it to predict diseases. The model can look at data and say what chance they have of getting a disease. We show the results in a way that's easy to understand. The Early Disease Prediction Using Healthcare Datasets project is good for healthcare because it helps us find diseases early. It also helps doctors plan ahead and make patients healthier.

Overall the Early Disease Prediction Using Healthcare Datasets project is a system that helps doctors find health risks early. It does this by looking at healthcare data and using computer programs to predict what might happen to a patient. The Early Disease Prediction Using Healthcare Datasets project is really helpful, for doctors and patients.

The main parts of a Disease Prediction System are

1. Data Collection

We need to find healthcare datasets with lots of information and medical records. These datasets are crucial for the Disease Prediction System. We have to get data for the Disease Prediction System to work well.

2. Data Preprocessing

This step is where we clean up the data for the Disease Prediction System. We fix errors. Make sure all features are in the same format. This helps the Disease Prediction System give results.

3. Model Training

Now we use the cleaned data to train machine learning algorithms for the Disease Prediction System. The algorithms learn from the data. This is a part of the Disease Prediction System.

4. Disease Prediction

We use the trained model to check if a new patient's at risk of getting a disease, with the Disease Prediction System. The Disease Prediction System looks at the patients information. The Disease Prediction System then tells us if they are likely to get sick. The Disease Prediction System makes this all possible.

Steps of Data Preprocessing

- **Data Collection**
- **Data Cleaning**
- **Data Transformation**
- **Data Integration**
- **Feature Selection**

Steps of Data Preprocessing : Data preprocessing is really important when we are building a disease prediction system. This is because the raw healthcare data we get is often not complete not consistent and not in a format.

1. Data Collection

We start with collecting data. This is the step where we gather healthcare datasets from sources we can trust. These datasets have a lot of information such as details, medical history, symptoms, laboratory test results and disease diagnosis records. The data may come from places like Kaggle or the UCI Machine Learning Repository. We usually store the collected data in CSV files Excel

spreadsheets or databases. The quality of the collected data is very important because it directly affects how accurate the disease prediction model is.

2. Data Cleaning

This involves handling missing values removing duplicates correcting data formats and eliminating attributes that're not relevant. This process makes the dataset more reliable and accurate for analysis. We need to do this to make sure that our disease prediction system is accurate.

3. Data Transformation

Data transformation is the step where we convert the dataset into a format that machine learning algorithms can use. Many datasets have variables like gender or disease type that must be converted into numbers. We use techniques like label encoding and one-hot encoding to convert data into numbers. We may also standardize features to make sure that everything is on the scale. This helps ensure that all variables contribute equally to the disease prediction model. Data transformation is a step in data preprocessing.

4. Data Integration

Data integration is the step where we combine data from sources into one unified dataset. In healthcare analytics patient information may come from systems like hospital records or laboratory reports. During this step we merge these data sources into a dataset while making sure that the data is consistent and that we do not have any information. Data integration helps create a dataset that improves the disease prediction model. This is a step in data preprocessing.

5. Feature Selection

Feature selection is the step where we identify the variables that're relevant to disease prediction. The healthcare datasets we have may have a lot of attributes. Not all of them are useful for building a prediction model. We use methods like correlation analysis to determine which features have a relationship with disease outcomes. By removing features that're not relevant the disease prediction model becomes more efficient, faster and more accurate. Feature selection also reduces complexity. Improves machine learning algorithms. This is a step in data preprocessing.

Overall these data preprocessing steps ensure that the dataset is clean organized and optimized for machine learning analysis. This improves the accuracy and reliability of the disease prediction system. Data preprocessing is crucial, for a disease prediction model. Data preprocessing helps make the disease prediction system more accurate.

Vi. RESULT EVOLUTION AND ANALYSIS

The performance of machine learning models is looked at using things like:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

When we compare machine learning models we can see which one is the best for predicting diseases on. We use ways to show the results, like bar charts and heatmaps so they are easy to understand. Machine learning models are really important, for this kind of thing. We use machine learning models to get the results.

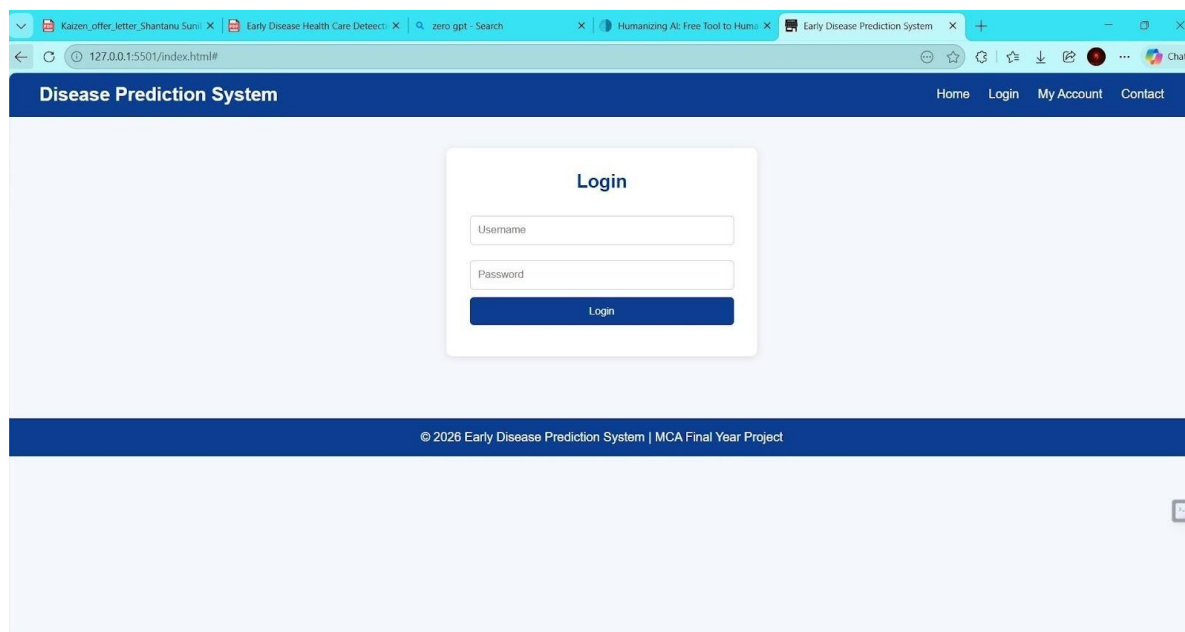


Figure 1: User Login Interface of the System

The picture shows the **Login Interface** of the Early Disease Prediction using Healthcare Datasets or Web Application. This is where people who have signed up can get into the system by putting in their email address and password. The login page looks really nice with a theme and a special box in the middle where you can log in. This page has spaces where you can put in your email and password and a button that says Login. When you click this button the system checks your information to make sure it is correct. If everything is okay you can then use the Early Disease Prediction using Healthcare Datasets and all its features, such as checking if newss real looking at it closely and using the dashboard.

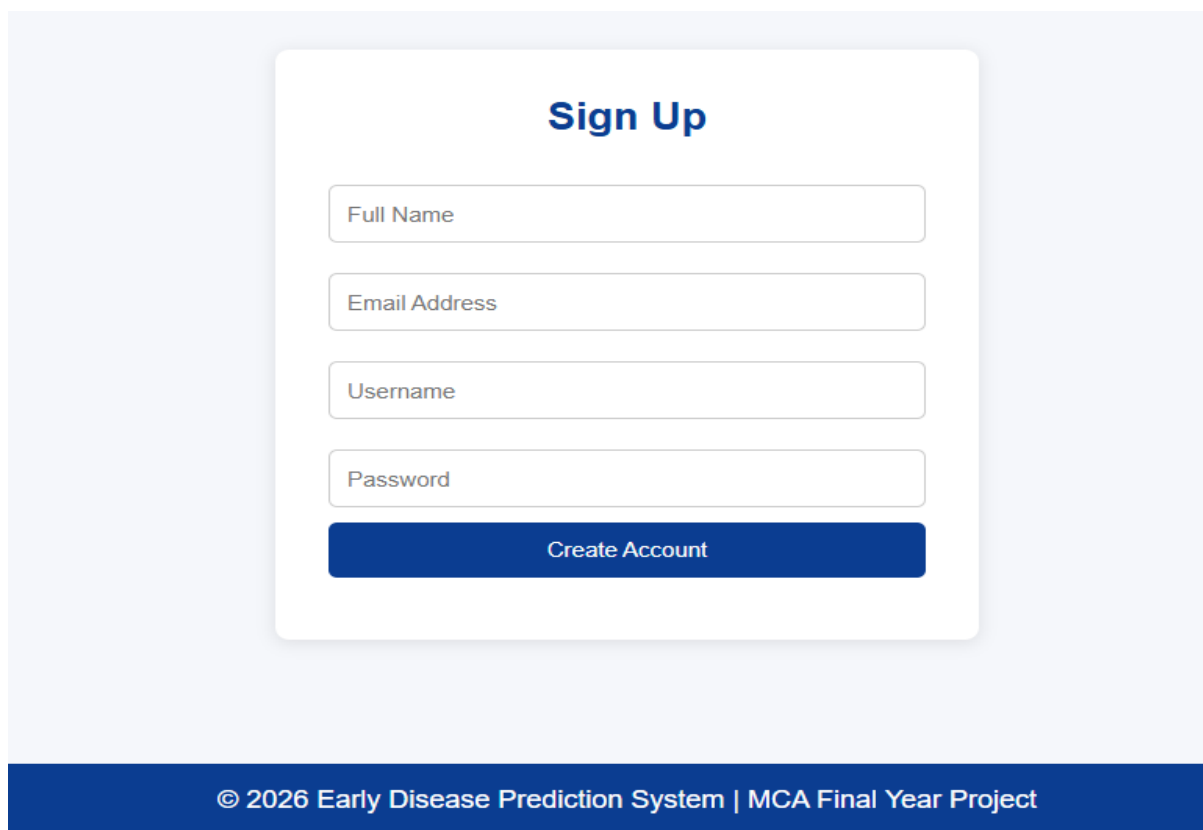


Figure 2: User Registration (Sign Up) Interface

The picture above shows how new users **sign up** to use the system. This is where people can create an account and get access to the system. The design of the system is easy to use. It looks clean. It has a background that makes the form easy to see on the system. When users are done filling in all their details they click the Create Account button to Early Disease Prediction using Healthcare Datasets, complete the sign up process for the system. Their details are then added to the system. After users have signed up for the system they can log in to the system using their email address and password for the system. The sign up part of the system is important for keeping user information safe and secure for all users of the system. It also makes sure that only the right people can use the system. The system keeps track of who's using the system and what they are doing on the system. Only people with an account can use the system. Only people with an account, for the system can use the system.

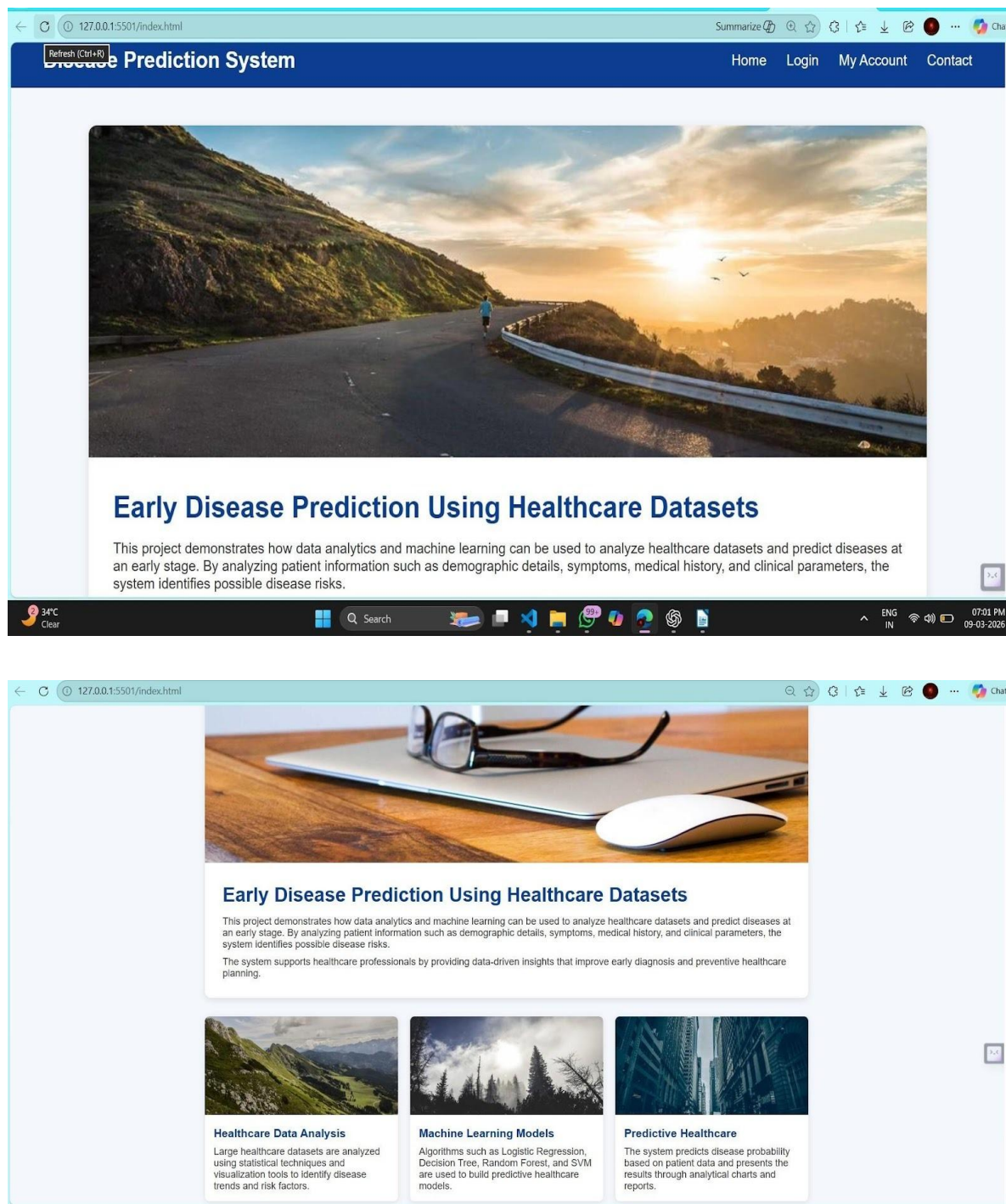


Figure 3 : Main Interface (HOME PAGE)

The **Home Page is like the door of the system**. It shows users what the platform can do.

➤ It is made to be easy to use so users can find their way

The home page has a menu some info about the system and links to parts. In our project the home page explains what the system is for. It helps users find things like registration and login.

➤ The layout is simple and easy to use.

Users can quickly see how to use the system. The home page may also show news or featured stuff.

➤ This helps keep users interested.

The home page helps users get started and find their way It makes the application easy to use. The Home Page is very important, for the system. The home page guides users to use the system. The home page presents the system.

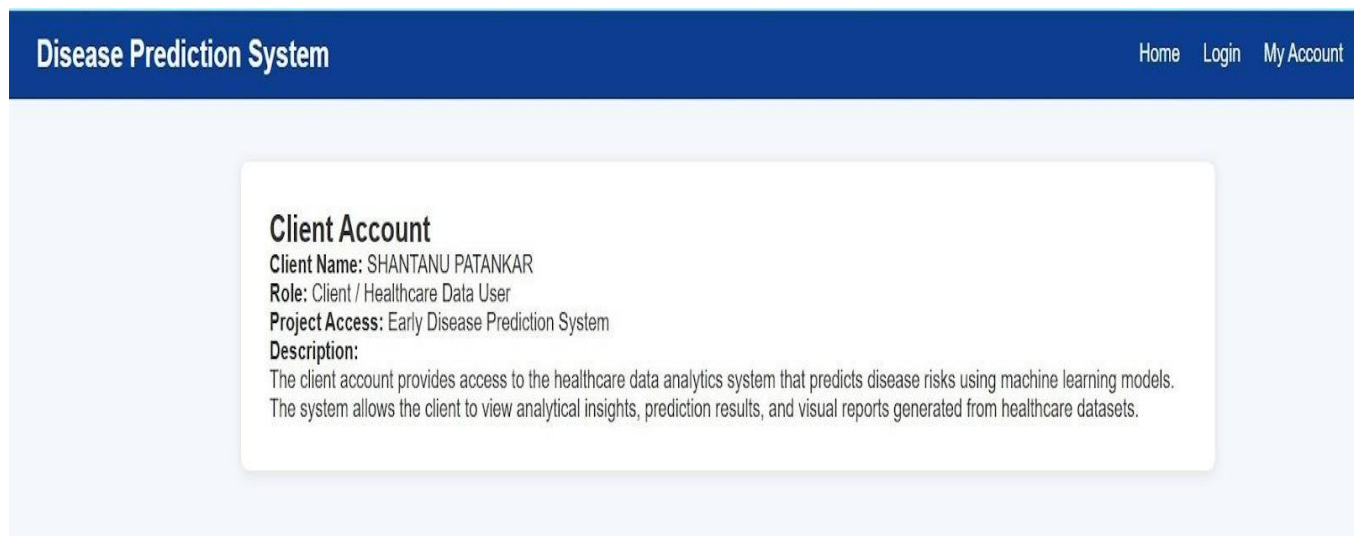
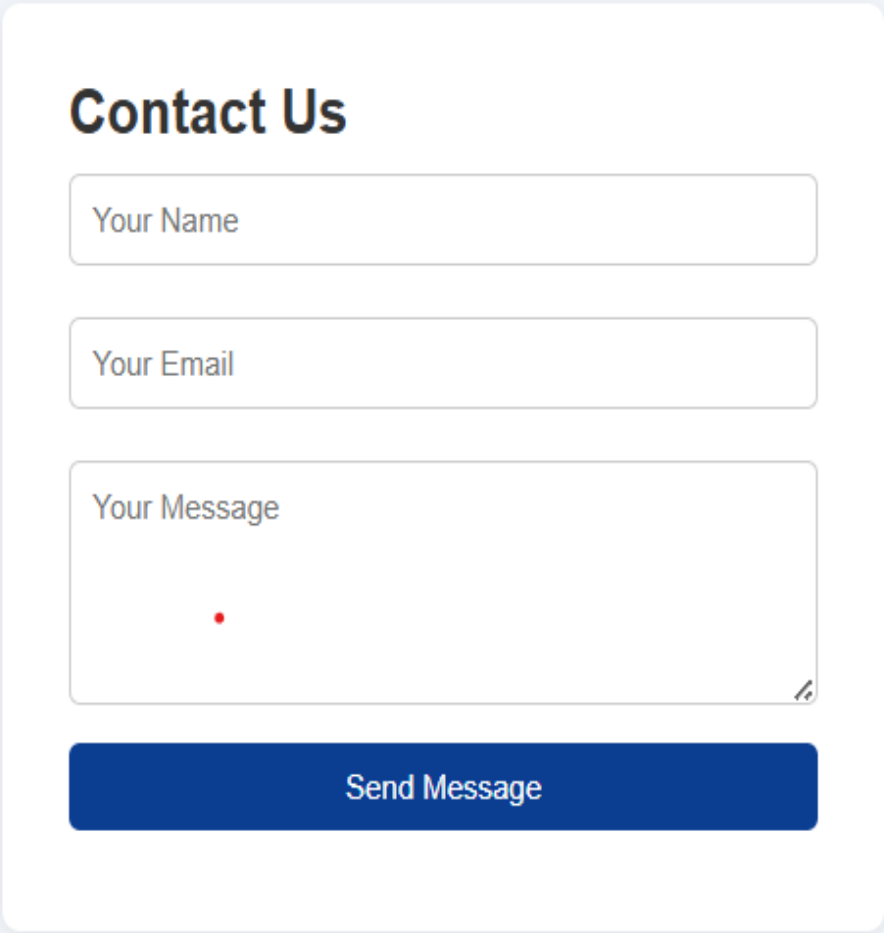


Figure 4: Account Section Interface

The **Account Section** is a part of the system. It lets users take care of their information and account settings. This is a place where users can look at, update and control the details of their profile. The Account Section is where users can see things like their username, email address and contact details. They can also look at information that is related to their profile. The Account Section lets users update their information when they need to. This way their account will always have the information. The Account Section also lets users do things like change their passwords take care of security settings and look at what's happening with their account. The people who made the Account Section thought about how users would interact with it. They made sure that only the right users can get into their account information. This keeps user data safe. Makes sure the system works correctly. The Account Section is good, for users because it makes it easy for them to manage their account. At the time it keeps their information safe and private. The Account Section is a part of the system because it helps users take care of their Account Section and it keeps their data secure.



The image shows a 'Contact Us' form with a white background and rounded corners, set against a light blue gradient. The form contains three input fields: 'Your Name', 'Your Email', and 'Your Message'. The 'Your Message' field is a larger text area with a small red dot and a cursor icon at the bottom right. Below the fields is a blue 'Send Message' button. At the bottom of the page, there is a dark blue footer bar with white text.

Contact Us

Your Name

Your Email

Your Message

Send Message

© 2026 Early Disease Prediction System | MCA Final Year Project

Fig : 5 : Contact Us Page Description

The **Contact Us Page** is an essential component of the system that allows users to communicate with the website administrators or support team. This page provides users with a simple and convenient way to send inquiries, feedback, or report issues related to the platform. Typically, the contact page includes a form where users can enter their details such as name, email address, subject, and message. Once the form is submitted, the information is sent to the system administrators for review and response. This helps maintain effective communication between users and the system management. The design of the contact page focuses on simplicity and accessibility, ensuring that users can easily provide their feedback or request assistance. Additionally, the page may include

other contact information such as email addresses, phone numbers, or social media links to offer multiple communication options.

Overall, the Contact Us page enhances user engagement and support by providing a reliable channel for communication and problem resolution.

Conclusion

This project shows that we can make a web system that people can use easily. The web system has a home page, an account section and a contact us page. These parts work together so people can use the system without any problems.

The home page is where people start when they use the web system. From here people can get to all the parts of the system. The account section is where people can change their information update their profile and control their account. This makes it easy for people to use the system. It also keeps their information safe. The contact us page is how people can talk to the people who run the system. They can ask questions give feedback or tell us about problems they are having.

To make this project we used web technologies and design principles. We wanted to make a system that people can use on any device. That is easy for everyone to use. We also used a database to store peoples information. This keeps everything easy to find. We also made sure the system is safe by using things like passwords and checks to make sure everything is correct.

The project is. This shows that planning, design and development are important when making web applications. This project is a start and we can use it to make the system even better, in the future. We can add things and make it work even better. The web system is a tool and it will help people. The project is a success. It shows what we can do when we work hard and plan carefully.

This study shows that data analysis and machine learning work well in predicting diseases early using healthcare data. The system we propose looks at information to find patterns and risk factors that can lead to disease. By using prediction models our system gives forecasts that can help doctors make smart choices. We found that getting data ready choosing the features and testing models are crucial, for building trustworthy prediction systems. Overall our project proves that using data can greatly improve healthcare analysis and help patients get results through early diagnosis and prevention strategies. Data analytics helps doctors. Machine learning helps doctors.

References:

1. Tom M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill Education, 1997.
2. Jiawei Han, Micheline Kamber, and Jian Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2012.
3. S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Machine Learning: A Review of Classification Techniques," *Artificial Intelligence Review*.
4. Ziad Obermeyer and Ezekiel J. Emanuel, "Predicting the Future — Big Data and Clinical Medicine," *The New England Journal of Medicine*.
5. Andre Esteva *et al.*, "A Guide to Deep Learning in Healthcare," *Nature Medicine*, 2019.
6. F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*.
7. Kaggle, "Healthcare Data Repository," Available: Kaggle Datasets.
8. UCI Machine Learning Repository, "Medical Datasets," Available: UCI Machine Learning Repository.
9. Python Software Foundation, "Python Documentation," Available: Python.org.

10. World Health Organization, “Digital Health Analytics,” Available: WHO Website.
11. Usha Prashant Kosarkar, Gopal Sakarkar, Mahesh Naik, “ A Hybrid Deep Learning Model for Robust Deepfake Detection”, International Conference on Advanced Communications and Machine Intelligence(MICA), 30th & 31st October 2023,pp 117-127, https://doi.org/10.1007/978-981-97-6222-4_9