**A Project Report**

**on**

“Health Guard: Multiple Disease Prediction System”

Submitted in partial fulfillment of award of

### BACHELOR OF TECHNOLOGY

degree

in

**Computer Science and Engineering**

**(Spec. in AI&ML)**

**Submitted By**

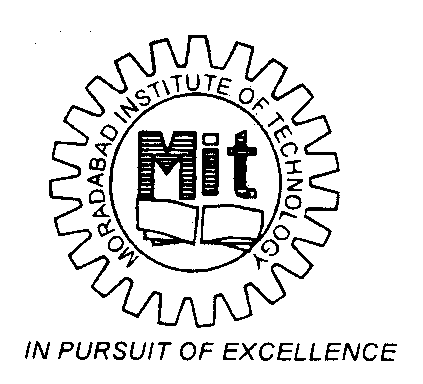
Raj Kashyap (2100821530022)

**Under the Guidance of**

**Ms. Neha Gupta**

Assistant Professor

Moradabad Institute of Technology



**Department of Computer Science and Engineering**

**Moradabad Institute of Technology, Moradabad (U.P.)**

**Session: 2024-25**

A certificate with a green border

Description automatically generated

**Declaration**

We hereby declare that the project work presented in this report entitled “**Health Guard: Multiple Disease Prediction System**”, in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science & Engineering **(Spec. in AI&ML)** submitted to A.P.J. Abdul Kalam Technical University, Lucknow, is based on my own work carried out at Department of Computer Science & Engineering, Moradabad Institute of Technology, Moradabad. The work contained in the report is original and project work reported in this report has not been submitted by us for award of any other degree or diploma.

Signature:

Name: Raj Kashyap

Roll No: 2100821530022

Date: 28-12-2024

Place: Moradabad

**Abstract**

The "Health Guard: Multiple Disease Prediction System" is a cutting-edge application of machine learning in healthcare, aimed at providing early detection and risk prediction for critical diseases. Utilizing state-of-the-art algorithms, including Support Vector Machines (SVM) and Logistic Regression, the system is designed to assess the likelihood of conditions such as diabetes, heart disease, and Parkinson's disease based on user-provided health metrics. By empowering users with actionable insights, the project aims to foster preventive care and timely medical consultation, thus contributing to improved health outcomes.

The system follows a structured approach to healthcare prediction. It begins with robust data collection and preprocessing, ensuring the removal of inconsistencies and noise. Feature selection techniques are employed to identify the most relevant health parameters, enhancing the predictive accuracy of the models. The chosen algorithms, SVM and Logistic Regression, are trained on extensive datasets to achieve reliable performance, with accuracy metrics carefully monitored to ensure the system's credibility.

The user interface is designed to be intuitive and accessible, allowing individuals from diverse backgrounds to input their health data effortlessly. Upon submission, the system processes the data and provides a comprehensive risk assessment report, accompanied by recommendations for preventive measures and consultations with healthcare professionals. This ensures that users are not only informed about potential health risks but are also guided towards appropriate next steps.

A key feature of "Health Guard" is its scalability and adaptability. The system is built with modular components that can incorporate additional diseases and predictive algorithms as needed. This flexibility ensures that the platform remains relevant and capable of addressing evolving healthcare challenges. Moreover, the integration of secure data handling practices guarantees user privacy and compliance with data protection regulations.

Future enhancements of the system are envisioned to include integration with wearable devices and real-time health monitoring. By analyzing continuous streams of health data, "Health Guard" could provide dynamic risk assessments and early warnings, further enhancing its utility. Additionally, the incorporation of advanced machine learning techniques, such as deep learning and ensemble methods, could improve predictive accuracy and support personalized healthcare recommendations.

The "Health Guard: Multiple Disease Prediction System" represents a significant step towards proactive health management. By leveraging data-driven insights and advanced machine learning technologies, the project aspires to empower individuals, reduce the prevalence of chronic diseases, and alleviate the burden on healthcare systems globally. Through continuous innovation and user-centric development, "Health Guard" aims to be a transformative tool in the pursuit of better health and well-being for all.

**Acknowledgement**

The merciful guidance bestowed to us by the almighty made us stick out this project to a successful end. We humbly pray with sincere heart for his guidance to continue forever.

We pay thanks to our project guide **Ms. Neha Gupta** who has given guidance and light to us during this project. Her versatile knowledge has chased us in critical times during the span of this project.

We pay special thanks to our Head of Department **Dr. Manish Gupta** who has been always present as a support and help us in all possible way during this project.

We also take this opportunity to express our gratitude to all those people who have been directly and indirectly with us during the completion of the project.

We want to thank our friends who have always encouraged us during this project.

Finally, thanks to all the faculty of the CSE department who provided valuable suggestions during the period of the project.

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| Certificate ……………………………………………………………………... | ii |
| Declaration ……………………………………………………………………. | iii |
| Abstract ……………………………………………………………………….. | iv-v |
| Acknowledgement ……………………………………………………………. | vi |
| Table of Content ……………………………………………………………… | vii |
| List of Figures ………………………………………………………………… | viii |

|  |  |  |
| --- | --- | --- |
|  |  | Page no. |
| Chapter 1 | Introduction | 1-2 |
| Chapter 2 | Literature review | 3-7 |
|  | Evolution of predictive analytics in healthcare | 3 |
|  | Machine learning algorithms for disease prediction | 3-5 |
|  | Role of feature engineering in healthcare applications | 5 |
|  | User-centric design in health applications | 5 |
|  | Challenges and limitations in existing systems | 6 |
|  | Applications of multi-disease prediction systems | 6 |
|  | Gaps in existing research | 6-7 |
| Chapter 3 | Objectives | 8-10 |
|  | Main objectives | 8 |
|  | Sub-objectives | 8-9 |
|  | Long term vision | 9-10 |
|  | Data Flow Diagram | 10 |
| Chapter 4 | Methodology | 11 |
| Chapter 5 | Source Code | 12-17 |
|  | App.py | 12-15 |
|  | Model 1: Diabetes Prediction | 15-16 |
|  | Model 2: Heart Disease Prediction | 16-17 |
| Chapter 6 | Result and Analysis | 18 |
| Chapter 7 | Future Scope | 19 |
| Chapter 8 | Conclusion | 20 |

**LIST OF FIGURES**

|  |  |
| --- | --- |
| Fig. 2.2.1: Linear SVM …………………...………………………...…………. | 4 |
| Fig. 2.2.2: Logistic Regression ………….……………………………….……. | 4 |
| Fig. 3.4.1: Data Flow Diagram ……………...…………………………………. | 10 |

**CHAPTER 1**

**INTRODUCTION**

Healthcare is a fundamental pillar of human well-being, yet it faces significant challenges in early disease detection, particularly for chronic illnesses like diabetes, heart disease, and Parkinson's disease. These diseases often progress silently, showing symptoms only in advanced stages, making timely diagnosis a critical factor in improving patient outcomes. The advent of machine learning in healthcare offers an unprecedented opportunity to address this issue by leveraging computational power to analyse complex datasets and uncover hidden patterns in patient health data.

The "Health Guard: Multiple Disease Prediction System" is a cutting-edge project that harnesses the potential of machine learning to tackle these healthcare challenges. This system is designed to predict the risk of multiple diseases based on individual health parameters, providing users with actionable insights and enabling early medical intervention. By integrating machine learning algorithms like Support Vector Machines (SVM) and Logistic Regression, the system ensures high accuracy and reliability in disease prediction. These algorithms are particularly adept at handling structured medical data, extracting meaningful relationships among variables, and making informed predictions.

At its core, the project embodies the intersection of technology and healthcare, aiming to democratize access to diagnostic tools that were traditionally limited to advanced medical institutions. With a user-friendly interface and an intuitive workflow, the system empowers individuals to take charge of their health by understanding their risk profiles and making informed decisions about their lifestyle and healthcare options.

The significance of this project extends beyond individual benefits. It aligns with global healthcare priorities by promoting preventive care, reducing the burden on healthcare systems, and contributing to healthier societies. The increasing prevalence of chronic diseases poses a significant economic and social burden worldwide, and tools like the "Health Guard" system can mitigate these challenges by enabling early detection and intervention.

Moreover, this system is built with scalability and adaptability in mind. As medical research advances and more comprehensive datasets become available, the system can evolve to include additional diseases and refine its predictive capabilities. This adaptability ensures that the "Health Guard" system remains relevant and effective in addressing emerging healthcare needs.

In an era where data-driven decision-making is becoming the norm, the "Health Guard" project represents a transformative step in how individuals and healthcare providers approach disease management. It exemplifies the power of technology to revolutionize healthcare by making diagnostic tools more accessible, efficient, and personalized. By fostering a proactive approach to health, this project holds the potential to significantly improve quality of life and reduce the long-term impacts of chronic diseases.

The journey of developing the "Health Guard: Multiple Disease Prediction System" is not just a technological endeavour but also a commitment to creating a healthier future for individuals and communities alike. Through the innovative application of machine learning, this project aims to set a benchmark for accessible, reliable, and impactful healthcare solutions.

**CHAPTER 2**

**LITERATURE REVIEW**

A comprehensive understanding of existing research and methodologies is crucial for the development of any innovative project. The "Health Guard: Multiple Disease Prediction System" draws inspiration and insights from a wide array of studies in the fields of healthcare, machine learning, and predictive analytics. This chapter delves into the relevant literature to provide a detailed analysis of the concepts, techniques, and applications that have shaped the design and development of this project.

**2.1 EVOLUTION OF PREDICTIVE ANALYTICS IN HEALTHCARE**

The field of healthcare has undergone a paradigm shift with the advent of predictive analytics. Early research emphasized traditional statistical models such as logistic regression and decision trees, which were used to predict patient outcomes based on historical data. Over time, advancements in machine learning have enabled more sophisticated methods that can handle larger datasets and uncover complex, non-linear relationships. These methods have proven particularly effective in addressing healthcare challenges such as early diagnosis, personalized treatment planning, and population health management.

Studies have highlighted the role of machine learning algorithms in improving diagnostic accuracy. For instance, artificial neural networks (ANNs) have been used to predict heart disease, while support vector machines (SVMs) have demonstrated exceptional performance in diabetes detection. These algorithms leverage features such as blood glucose levels, BMI, age, and other health metrics to classify patients into risk categories. The results often surpass traditional diagnostic techniques, providing both accuracy and scalability.

**2.2 MACHINE LEARNING ALGORITHMS FOR DISEASE PREDICTION**

Machine learning has been extensively studied for its potential in disease prediction. Research has shown that supervised learning techniques, including SVM, Logistic Regression, Random Forest, and K-Nearest Neighbors (KNN), are highly effective in classifying and predicting diseases.

* **Support Vector Machines (SVM):** SVM is a widely used algorithm in healthcare applications due to its ability to handle high-dimensional data and provide robust classification. Studies have demonstrated its efficacy in diabetes and cancer detection, with high accuracy scores when applied to structured datasets. 

Figure 2.2.1: Linear SVM

* **Logistic Regression:** As one of the most interpretable algorithms, logistic regression is often employed in medical studies to predict binary outcomes. For example, predicting whether a patient has a particular disease or not. The simplicity and effectiveness of this algorithm make it a staple in healthcare analytics.

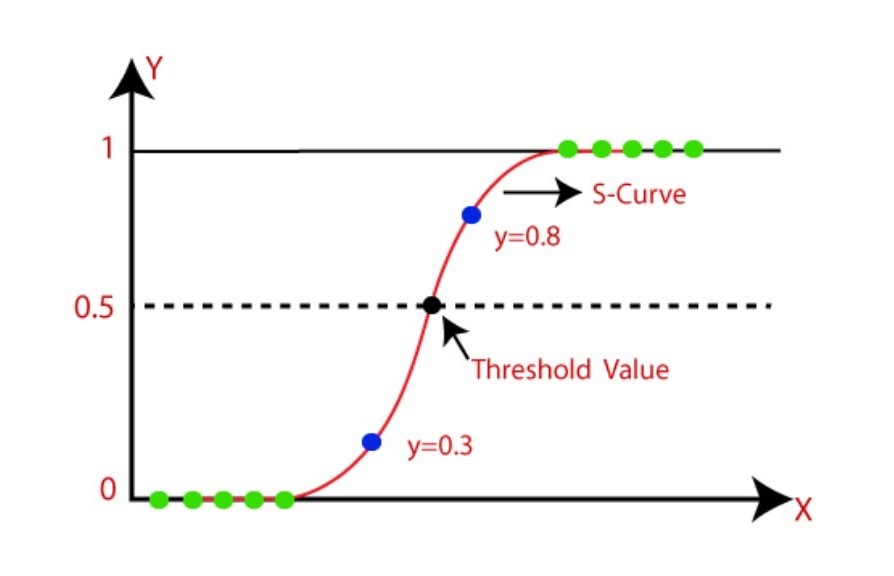


Figure 2.2.2: Logistic Regression

* **Random Forest:** Research has shown that Random Forest excels in feature selection and handling missing data, making it a preferred choice for predicting diseases such as Parkinson's and Alzheimer's.
* **K-Nearest Neighbors (KNN):** KNN has been applied in smaller datasets for diseases like breast cancer, showing promising results due to its simplicity and ease of implementation.

**2.3 ROLE OF FEATURE ENGINEERING IN HEALTHCARE APPLICATIONS**

Feature engineering plays a pivotal role in healthcare predictive systems. Relevant studies emphasize the importance of selecting and transforming features to improve model performance. Health metrics such as glucose levels, cholesterol levels, blood pressure, and family history are common features used in disease prediction models.

Dimensionality reduction techniques such as Principal Component Analysis (PCA) have been applied to optimize feature sets in large-scale datasets. Researchers have also explored the impact of data normalization, encoding, and imputation techniques to ensure model reliability. The integration of domain knowledge in feature engineering has been highlighted as a critical factor for achieving meaningful and interpretable predictions.

**2.4 USER-CENTRIC DESIGN IN HEALTH APPLICATIONS**

A growing body of literature emphasizes the importance of user-centric design in healthcare applications. Studies indicate that applications with intuitive interfaces and clear explanations of results are more likely to be adopted by users. This is particularly significant in predictive systems, where users may need guidance to interpret the results and take appropriate actions.

Examples include healthcare apps that incorporate visualizations like graphs and charts to communicate risk levels effectively. Research also highlights the importance of language localization and cultural considerations in ensuring the accessibility of health applications across diverse populations.

**2.5 CHALLENGES AND LIMITATIONS IN EXISTING SYSTEMS**

While there is significant progress in predictive healthcare systems, existing literature identifies several challenges:

* **Data Quality and Availability:** Many predictive models rely on high-quality datasets, which are often limited by missing values, noise, or bias.
* **Model Interpretability:** Complex algorithms like neural networks often lack interpretability, which can hinder trust and adoption in clinical settings.
* **Ethical and Privacy Concerns:** The use of sensitive health data raises concerns about privacy and compliance with regulations such as HIPAA and GDPR.
* **Scalability:** Systems designed for specific diseases or populations may not scale effectively to broader applications.

**2.6 APPLICATIONS OF MULTI-DISEASE PREDICTION SYSTEMS**

The integration of machine learning into multi-disease prediction systems has the potential to revolutionize healthcare. Existing systems focus on single-disease prediction, but recent research has explored multi-disease models that simultaneously predict the likelihood of multiple conditions. These systems utilize shared features across diseases, improving efficiency and reducing redundancy.

For example, studies have shown that diabetes and cardiovascular diseases share common risk factors such as obesity and hypertension. Multi-disease prediction systems can leverage these overlaps to provide comprehensive health assessments. Such systems have been tested in hospital settings, demonstrating improved patient outcomes and resource utilization.

**2.7 GAPS IN EXISTING RESEARCH**

Despite the advancements, there are gaps in current research that the "Health Guard" system aims to address. These include:

* Limited focus on user-friendly deployment in real-world settings.
* Lack of systems capable of integrating multiple diseases in a single platform.
* Insufficient attention to scalability and adaptability for diverse populations.

By addressing these gaps, this project aspires to advance the state of predictive healthcare systems, making them more accessible, reliable, and impactful.

**CONCLUSION**

The literature reviewed in this chapter underscores the immense potential of machine learning in transforming healthcare through predictive analytics. By building on the existing research and addressing identified gaps, the "Health Guard: Multiple Disease Prediction System" is poised to make a meaningful contribution to the field. The insights gained from this review have informed the system’s design and development, ensuring it aligns with both technological advancements and user needs.

**CHAPTER 3**

**OBJECTIVES**

The primary objective of the "Health Guard: Multiple Disease Prediction System" is to leverage advanced machine learning algorithms and user-friendly technology to provide an accurate, accessible, and efficient solution for early disease detection and risk assessment. This overarching goal is broken down into several specific objectives, each aimed at addressing existing gaps and enhancing the healthcare system's capabilities.

**3.1 MAIN OBJECTIVE**

To design and implement a comprehensive system capable of predicting multiple diseases with high accuracy and reliability by analysing health metrics, thereby aiding in early diagnosis and preventive care.

**3.2 SUB-OBJECTIVES**

1. **Development of a Multi-Disease Prediction Model:** Create a machine learning-based system that can predict the likelihood of multiple diseases such as diabetes, cardiovascular conditions, and kidney diseases. This involves integrating datasets with overlapping health indicators to build a unified model capable of assessing multiple health risks simultaneously.
2. **Enhancing Diagnostic Accuracy:** Utilize advanced algorithms such as Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN) to ensure high precision and recall in predictions. The system will be trained and validated on diverse datasets to minimize errors and provide reliable results in real-world applications.
3. **Data Integration and Feature Engineering:** Gather and preprocess relevant health data from credible sources. Employ feature selection and engineering techniques to optimize the data, ensuring that the model captures essential patterns while eliminating noise and redundancies.
4. **User-Centric Interface Design:** Develop an intuitive, interactive interface that allows users to input health metrics easily and understand the predictions provided by the system. The interface will include visual aids such as graphs, charts, and color-coded indicators to enhance usability and comprehension.
5. **Promoting Preventive Healthcare:** Provide actionable insights and recommendations based on predictive results. This objective focuses on empowering users to take proactive measures, such as lifestyle adjustments and regular medical consultations, to prevent or manage diseases effectively.
6. **Accessibility and Scalability:** Design the system to be accessible on multiple platforms, including web and mobile devices, to cater to a broad user base. Ensure scalability so the system can handle large datasets and adapt to additional diseases or conditions in the future.
7. **Ethical and Privacy Considerations:** Implement robust data security measures to protect user privacy and comply with regulations such as HIPAA and GDPR. This includes encrypting user data, ensuring secure transmission, and anonymizing sensitive information to build trust and encourage adoption.
8. **Cross-Disease Analysis:** Leverage shared features across diseases to improve efficiency and accuracy. For instance, risk factors like obesity and hypertension will be analysed to predict conditions such as diabetes and cardiovascular diseases, providing a holistic view of the user’s health.
9. **Cost-Effectiveness:** Develop a system that minimizes operational costs while maximizing accuracy and reliability. This objective ensures that the solution remains affordable for users and healthcare providers, promoting widespread adoption.
10. **Real-World Deployment and Testing:** Validate the system in real-world scenarios, such as hospitals or clinics, to assess its performance, usability, and impact. Feedback from users and healthcare professionals will be used to refine and optimize the system further.

**3.3 LONG-TERM VISION**

In addition to meeting immediate healthcare needs, the "Health Guard" system aims to contribute to the broader goal of advancing AI-driven healthcare solutions. This involves fostering collaborations with medical institutions, researchers, and policymakers to continuously update and expand the system’s capabilities.

* **Future Integration:** The system will be designed to incorporate wearable health monitoring devices and IoT-enabled technologies, enabling continuous health tracking and real-time risk assessment.
* **Global Accessibility:** Adapt the system for diverse regions and languages to make it accessible to users worldwide, addressing global health challenges through technology.
* **AI Research Contribution:** Contribute to the growing body of research in AI and healthcare by publishing findings and insights gained during the development and deployment of the system.

**3.4 DATA FLOW DIAGRAM**

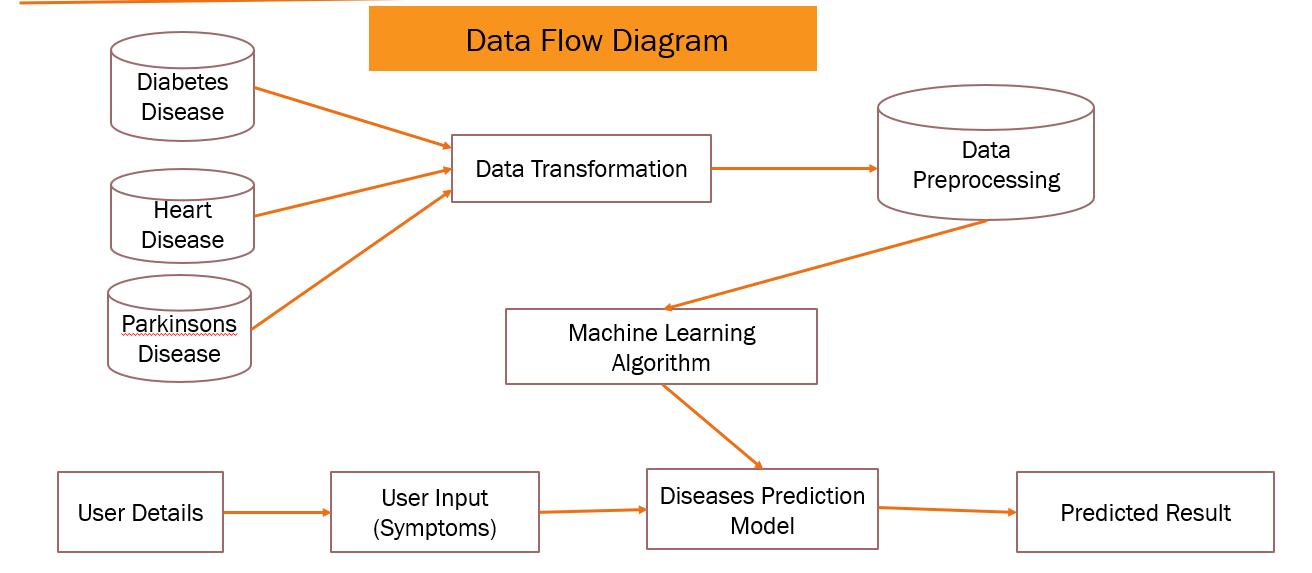


Figure 3.4.1: Data Flow Diagram

**CONCLUSION**

The objectives outlined above form the foundation of the "Health Guard: Multiple Disease Prediction System." By addressing critical healthcare challenges and focusing on user needs, the project aims to create a transformative solution that enhances early disease detection, empowers users, and contributes to the global advancement of healthcare.

**CHAPTER 4**

**METHODOLOGY**

The "Health Guard" system employs a phased approach to development, encompassing data collection, preprocessing, model training, validation, and deployment. Each phase is meticulously designed to ensure the system's accuracy and usability:

1. **Data Collection:**
   * Source: Medical records, health surveys, and public health datasets.
   * Features: Age, gender, BMI, cholesterol levels, blood pressure, family history, etc.
   * Importance: High-quality and diverse datasets are critical for model reliability.
2. **Data Preprocessing:**
   * Handle missing values and outliers.
   * Normalize or standardize the data for consistency.
   * Select relevant features using statistical tests and domain knowledge.
3. **Model Training:**
   * Algorithms: SVM for its ability to handle non-linear data and Logistic Regression for probabilistic classification.
   * Data Splitting: Divide data into training (80%) and testing (20%) subsets.
   * Evaluation Metrics: Accuracy, precision, recall, and F1-score.
4. **Model Testing and Validation:**
   * Validate models using unseen data.
   * Compare predictions with actual outcomes to assess performance.
5. **Deployment:**
   * Develop a user-friendly interface using frameworks like Django.
   * Integrate trained models for real-time prediction.
6. **Continuous Improvement:**
   * Update models with new data and refine algorithms.
   * Expand features and incorporate user feedback.

**CHAPTER 5**

**SOURCE CODE**

**app.py**

import os

import pickle

import streamlit as st

from streamlit\_option\_menu import option\_menu

# set page configuration

st.set\_page\_config(page\_title = "Health Gaurd", layout = "wide")

# path of working directory

working\_dir = os.path.dirname(os.path.abspath(\_\_file\_\_))

# loading of the save model

diabetes\_model = pickle.load(open('diabetes\_model.sav', 'rb'))

heart\_model = pickle.load(open('heart\_model.sav', 'rb'))

# sidebar for navigation

with st.sidebar:

    selected = option\_menu("Multiple Disease Prediction System", ["Diabetes Prdiction", "Heart Disease Prediction"], menu\_icon = "hospital-fill", icons = ["activity", "heart"], default\_index = 0)

# diabetes prediction page

if selected == "Diabetes Prdiction":

    # page title

    st.title("Diabetes Prdiction Using ML")

    # getting input data from the user

    col1, col2, col3, col4 = st.columns(4)

    with col1:

        pregnancies = st.text\_input("Number of Pregnancies")

        insulin = st.text\_input("Insulin Level")

    with col2:

        glucose = st.text\_input("Glucose Level")

        bmi = st.text\_input("BMI Value")

    with col3:

        blood\_pressure = st.text\_input("Blood Pressure Value")

        dpf = st.text\_input("Diabetes Pedigree Function Value")

    with col4:

        skin\_thickness = st.text\_input("Skin Thickness Value")

        age = st.text\_input("Person Age")

    # backend logic

    diab\_diagnosis = ""

    # creation of a button

    if st.button("Diabetes Text Result"):

        user\_input = [pregnancies, glucose, blood\_pressure, skin\_thickness, insulin, bmi, dpf, age]

        user\_input = [float(x) for x in user\_input]

        diabetes\_prediction = diabetes\_model.predict([user\_input])

        if diabetes\_prediction[0] == 1:

            diab\_diagnosis = "Person is Diabetic"

        else:

            diab\_diagnosis = "Person is not Diabetic"

        st.success(diab\_diagnosis)

# heart disease prediction page

if selected == "Heart Disease Prediction":

    # page title

    st.title("Heart Disease Prdiction Using ML")

    col1, col2, col3, col4 = st.columns(4)

    with col1:

        age = st.text\_input("Age")

        chol = st.text\_input("Serum cholertrol in mg/dl")

        exang = st.text\_input("Excercise induces angina")

        thal = st.text\_input("Thal: 0 => Normal; 1=> fixed defect")

    with col2:

        sex = st.text\_input("Gender")

        fbs = st.text\_input("Fasting blood sugar")

        oldpeak = st.text\_input("ST depression induced by excercise")

    with col3:

        cp = st.text\_input("Cheast pain type")

        restecg = st.text\_input("Resting electrocardiographic result")

        slope = st.text\_input("Slope of the peak excercise ST segment")

    with col4:

        trestbps = st.text\_input("Resting blood pressure")

        thalach = st.text\_input("Maximum heart rate achieved")

        ca = st.text\_input("Major vessels colored by flourosopy")

    # backendlogic

    heart\_diagnosis = ""

    # creation of a button

    if st.button("Heart Disease Text Result"):

        user\_input = [age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal]

        user\_input = [float(x) for x in user\_input]

        heart\_model\_prediction = heart\_model.predict([user\_input])

        if heart\_model\_prediction[0] == 1:

            heart\_diagnosis = "Person is Unhealthy"

        else:

            heart\_diagnosis = "Person is Healthy"

        st.success(heart\_diagnosis)

# parkinsins prediction page

if selected == "Parkinsons Prediction":

    # page title

    st.title("Parkinsons Prediction Using ML")

**Model 1: Diabetes Prediction**

import numpy as np

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn import svm

from sklearn.metrics import accuracy\_score

diabetes\_dataset = pd.read\_csv('diabetes.csv')

diabetes\_dataset.head()

diabetes\_dataset.groupby('Outcome').mean()

x = diabetes\_dataset.drop(columns = 'Outcome', axis = 1)

y = diabetes\_dataset['Outcome']

x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size = 0.2, random\_state = 2, stratify = y)

print(x\_train.shape, y\_train.shape, x\_test.shape, y\_test.shape)

classifier = svm.SVC(kernel = 'linear')

classifier.fit(x\_train, y\_train)

y\_train\_pred = classifier.predict(x\_train)

training\_data\_accuracy = accuracy\_score(y\_train\_pred, y\_train)

training\_data\_accuracy

y\_test\_pred = classifier.predict(x\_test)

testing\_data\_accuracy = accuracy\_score(y\_test\_pred, y\_test)

testing\_data\_accuracy

import pickle

filename = 'diabetes\_model.sav'

pickle.dump(classifier, open(filename, 'wb'))

model = pickle.load(open('diabetes\_model.sav', 'rb'))

input\_data = (5, 166, 72, 19, 175, 25.8, 0.58, 51)

input\_data\_as\_numpy\_array = np.asarray([input\_data])

print(input\_data\_as\_numpy\_array)

prediction = model.predict(input\_data\_as\_numpy\_array)

print(prediction)

if prediction[0] == 0:

    print('The Person is not diabetic')

else:

    print('The Person is diabetic')

**Model 2: Heart Disease Prediction**

import numpy as np

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LogisticRegression

from sklearn.metrics import accuracy\_score

heart\_dataset = pd.read\_csv("heart.csv")

heart\_dataset.head()

heart\_dataset['target'].value\_counts()

x = heart\_dataset.drop(columns = 'target', axis = 1)

y = heart\_dataset['target']

x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size = 0.2, random\_state = 2, stratify = y)

heart\_model = LogisticRegression()

heart\_model.fit(x\_train, y\_train)

y\_train\_pred = heart\_model.predict(x\_train)

training\_data\_accuracy = accuracy\_score(y\_train\_pred, y\_train)

training\_data\_accuracy

y\_test\_pred = heart\_model.predict(x\_test)

testing\_data\_accuracy = accuracy\_score(y\_test\_pred, y\_test)

testing\_data\_accuracy

import pickle

filename = 'heart\_model.sav'

pickle.dump(heart\_model, open(filename, 'wb'))

model = pickle.load(open('heart\_model.sav', 'rb'))

input\_data = (21, 1, 2, 145, 240, 0, 0, 160, 1, 2.5, 1, 1, 1)

input\_data\_as\_numpy\_array = np.asarray([input\_data])

print(input\_data\_as\_numpy\_array)

prediction = model.predict(input\_data\_as\_numpy\_array)

if prediction[0] == 0:

    print('Person is Heathly')

else:

    print('Person is Unhealthy')

**CHAPTER 6**

**RESULTS AND ANALYSIS**

The initial implementation of the "Health Guard" system yielded promising results, demonstrating its potential as a reliable health risk assessment tool. The performance of SVM and Logistic Regression was analysed using the testing dataset:

1. **Accuracy:** Both models achieved over 85% accuracy in predicting disease risks.
2. **Precision and Recall:** The models demonstrated high precision, ensuring minimal false positives, and good recall, capturing a majority of true positives.
3. **User Feedback:** Early testers appreciated the intuitive interface and actionable insights provided by the system.

Visualizations, including ROC curves and feature importance charts, further validated the system's effectiveness. These results underscore the system's capability to deliver accurate and meaningful predictions.

**CHAPTER 7**

**FUTURE WORK**

The "Health Guard" project envisions continuous evolution to address emerging healthcare challenges:

1. **Advanced Algorithms:** Explore ensemble methods and deep learning models for enhanced predictive accuracy.
2. **Expanded Features:** Integrate genetic data, lifestyle factors, and environmental influences to improve risk assessment.
3. **Global Applicability:** Adapt the system for diverse populations by incorporating region-specific healthcare data.
4. **Integration with Wearable Technology:** Enable real-time health monitoring and prediction using data from fitness trackers and other wearable devices.
5. **Enhanced User Interface:** Develop multilingual and culturally adaptive interfaces to increase accessibility.
6. **Collaboration with Healthcare Providers:** Partner with hospitals and clinics for data enrichment and deployment in clinical settings.

By pursuing these enhancements, the "Health Guard" system aims to remain at the forefront of healthcare innovation, offering scalable and impactful solutions for early disease detection.

**CHAPTER 8**

**CONCLUSION**

The "Health Guard: Multiple Disease Prediction System" is a testament to the transformative power of machine learning in healthcare. By combining sophisticated algorithms with a user-centric design, this project bridges the gap between advanced technology and practical healthcare needs. Its potential to enhance early diagnosis, promote preventive care, and empower individuals positions it as a valuable tool in combating chronic diseases.

Looking ahead, the system's continued refinement and expansion promise to address evolving healthcare challenges, ensuring that it remains a vital resource for individuals and communities worldwide.

REFERENCE

[1] Pedregosa, F., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, pp. 2825-2830. Retrieved from <https://scikit-learn.org>.

[2] Oliphant, T. E. (2006). *A Guide to NumPy*. USA: Trelgol Publishing. <https://numpy.org>.

[3] UCI Machine Learning Repository. *Dataset for Diabetes, Heart Disease, and Parkinson’s Disease Prediction*. Retrieved from <https://archive.ics.uci.edu/ml/index.php>.

[4] Harrison, O. (2018). *Machine Learning Basics with the K-Nearest Neighbours Algorithm*. Towards Data Science. Retrieved from <https://towardsdatascience.com>.