Cardiovascular Disease Decision Support System

Huating Sun, Michelle Menachery, Wei Yu

Problem Description

The project is centered on developing a Decision Support System aimed at improving the early detection and prevention of cardiovascular diseases (CVD). Cardiovascular diseases, including disorders like coronary heart disease and stroke, are a leading cause of mortality worldwide, responsible for approximately 17.9 million deaths annually.

Why is it a problem?

Despite advancements in healthcare, the challenge in managing CVD lies in its complex, multifactorial nature, influenced by diverse risk factors such as high blood pressure, cholesterol levels, and lifestyle choices. Current strategies often fail to utilize patient-specific data effectively to predict and manage these diseases, leading to suboptimal prevention and treatment strategies.

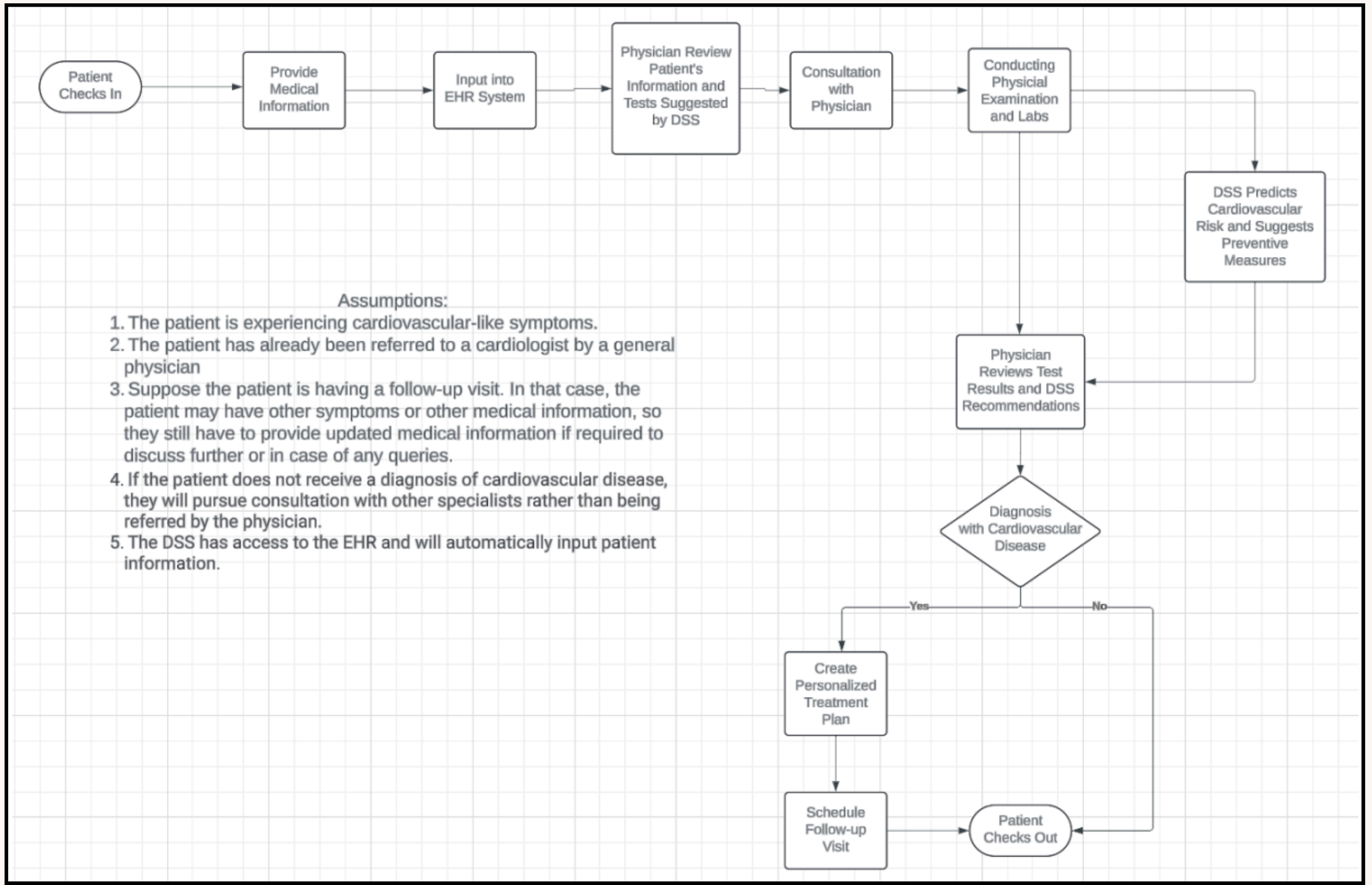
Stakeholders

- Patients at risk of cardiovascular diseases
- Healthcare providers such as physicians and nurses
- Healthcare organizations
- IT professionals responsible for the development and maintenance of the DSS.

Role of IT in Decision Support

IT plays a pivotal role in enhancing the decision-making process through the Decision Support System which will integrate ML algorithms with patient-specific data to predict cardiovascular risk more accurately and offer personalized recommendations. Such a system not only aids healthcare providers in making informed decisions but also empowers patients by providing them with actionable insights into their health risks.

Process Model



Key Activities

- Patient Check-In
- Collection of Medical Information
- Data Entry into EHR System
- Review of Patient Information
- Physician Consultation
- Physical Examinations and Additional Labs
- Review of Test Results and DSS Assessments
- Diagnosis
- Creation of a Personalized Treatment Plan
- Scheduling a Follow-up Visit
- Patient Check-Out

Major components - IT Solution

01. **Relational Database:** To store detailed patient data relevant to cardiovascular diseases.

02. **Modeling Component:** Algorithms that predict CVD based on diverse patient data.

03. **User Interface**: Facilitating easy access to medical recommendations and patient data management for healthcare professionals.

This system will leverage the Cardiovascular Disease dataset from Kaggle, which includes comprehensive medical data, to train the predictive algorithms. The objective is to provide a tool that not only aids in diagnosis but also in the proactive management of cardiovascular health through timely interventions based on guideline-based reminders and predictive analytics.

Data Source

Kaggle Cardiovascular Disease Dataset

The dataset is a comprehensive collection of medical data relevant to cardiovascular health, widely used for predictive modeling and medical research.

Data Elements Extracted

- Age
- Gender
- Height and Weight
- Blood Pressure (Systolic and Diastolic)
- Cholesterol Levels
- Glucose Levels
- Smoking Status
- Alcohol Consumption
- Physical Activity

Relational Database

The database consists of 5 tables:

users: Contains users information

patient: patient information

physician: physician information

physicianpatientlink: patient assignment information

medicaltest: patient’s medical test results information

USERS

	User ID [PK] integer	username character varying (50)	password character varying (50)	User Type character
1	1	goodwinkenneth	NcwOwDck	patient
2	2	xmorris	4Du1V2yQ	patient
3	3	qyoung	dF4nlNmT	patient
4	4	charleshendrix	yusl8sQI	patient
5	5	halltyler	tzVuIrlV	patient

PATIENTS

	User ID [PK] integer	First Name character	Last Name character	age integer	gender integer	Contact Information character varying (100)	height integer	weight numeric	Smoke History integer	Alcohol Consumption integer
1	1	Megan ...	Clark ...	60	1	arnoldjustin@example.org	156	64	0	0
2	2	Sheila ...	Holloway ...	60	1	swaters@example.org	170	85	0	0
3	3	Julia ...	Hale ...	64	1	stephanie76@example.org	151	90	0	0
4	4	Linda ...	Horton ...	54	1	casey70@example.net	159	97	0	0
5	5	Jennifer ...	Yoder ...	50	1	steven19@example.org	164	68	0	0

PHYSICIANS

	User ID [PK] integer	First Name character	Last Name character	Contact Information character varying (100)
1	1001	Mary ...	Cole	monica28@example.net
2	1002	Barry	Christensen	kscott@example.com

MEDICALTEST

	Test ID [PK] integer	Physician ID integer	Patient ID integer	Diastolic Blood Pressure integer	Systolic Blood Pressure integer	cholesterol integer	glucose integer	Cardiovascular Disease integer
1	1	1002	1	80	140	2	1	1
2	2	1001	2	90	160	1	1	1
3	3	1001	3	80	130	1	1	1
4	4	1001	4	80	120	1	1	1
5	5	1002	5	80	120	1	1	0

PHYSICIANPATIENTLINK

	Physician ID [PK] integer	Patient ID [PK] integer
1	1001	2
2	1001	3
3	1001	4
4	1001	8
5	1001	9

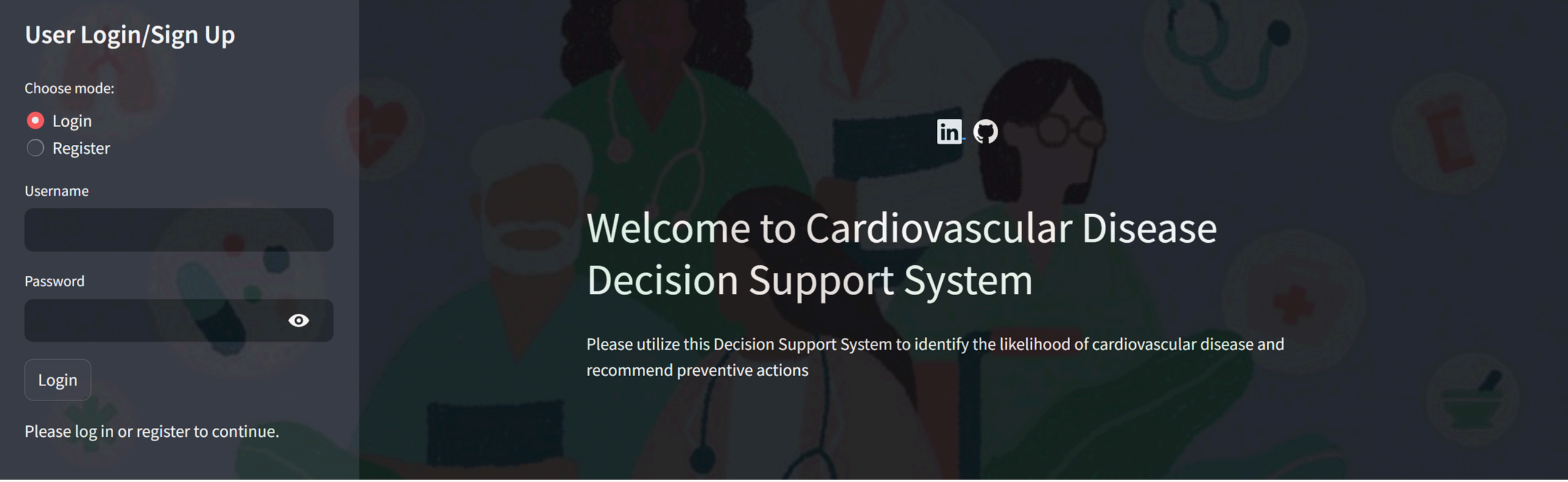
Modeling Component

The EDA performed for this project started with descriptive statistics by using univariate analysis using histograms to identify skewness and outliers in age, weight, and blood pressure. Bivariate analysis with pair plots and correlation matrices then helped understand relationships between these variables and cardiovascular disease outcomes.

For the modeling component of the project we tested several ML algorithms to analyze patient data and predict cardiovascular disease risk. The models were:

1. Logistic Regression
2. Decision Tree
3. Random Forest
4. Support Vector Machine
5. Gradient Boosting

Gradient Boosting was selected as the primary model for the project as it achieved the highest accuracy and ROC-AUC score, indicating its effectiveness at correctly classifying instances and distinguishing between the presence and absence of cardiovascular disease.



User Interface

- Streamlit: a free and open-source framework to developed data web application
- psycopg2: Python library to connect to PostgreSQL database and perform queries

What if Scenarios

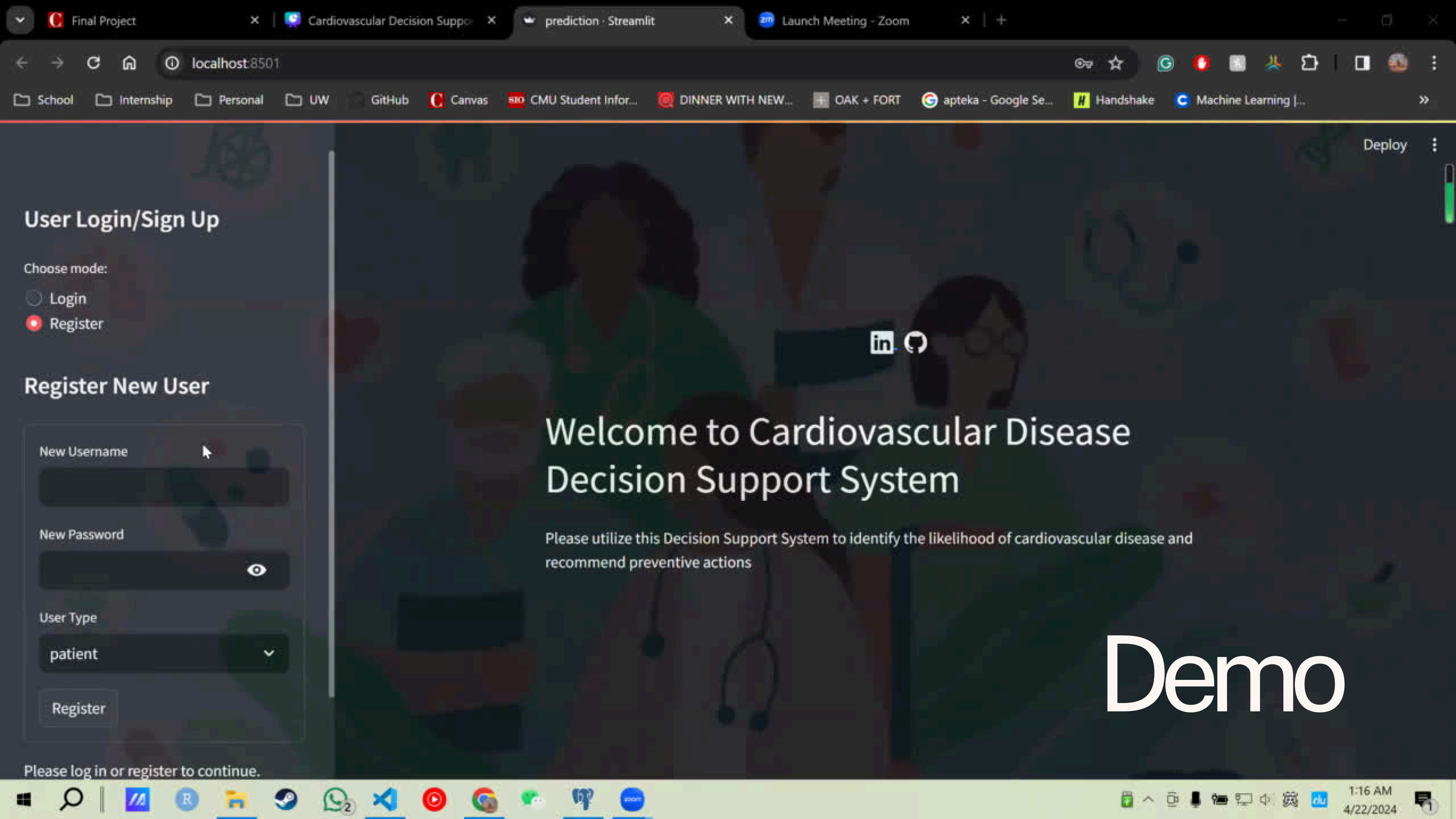
Scenario 1

Patient: Sheila Holloway
Age: 60
Gender: Male
Height: 170 cm
Weight: 85 kg
Smoke History: Yes
Alcohol Consumption: Yes
Exercise Regularly: Yes
Diastolic Blood Pressure: 90
Systolic Blood Pressure: 160
Cholesterol Level: Normal
Glucose Level: Normal
Initial Prediction of CVD: Yes

Scenario 2

Patient: Sheila Holloway
Age: 60
Gender: Male
Height: 170 cm
Weight: 60 kg
Smoke History: No
Alcohol Consumption: No
Exercise Regularly: Yes
Diastolic Blood Pressure: 80
Systolic Blood Pressure: 120
Cholesterol Level: Normal
Glucose Level: Normal
Prediction of CVD: No

- Changing factors such as BMI, Lifestyle Choices, Reduced Blood Pressure, and Normal Glucose and Cholesterol Levels will help patients avoid the risk of cardiovascular disease.
- Emphasize the need to continue monitoring these changing factors to ensure the risk of cardiovascular disease remains low.
- Doctors can focus on changing these factors to help the patients who are at risk.


Final Project
Cardiovascular Decision Suppo
prediction · Streamlit
Launch Meeting - Zoom
localhost:8501
School
Internship
Personal
UW
GitHub
Canvas
CMU Student Infor...
DINNER WITH NEW...
OAK + FORT
apteka - Google Se...
Handshake
Machine Learning |...
Deploy
User Login/Sign Up
Choose mode:
Login
Register
Register New User
New Username
New Password
User Type
patient
Register
Please log in or register to continue.
Welcome to Cardiovascular Disease Decision Support System
Please utilize this Decision Support System to identify the likelihood of cardiovascular disease and recommend preventive actions
Demo
1:16 AM
4/22/2024

Top Relevant Standards

HIPAA

It is important because there are a great amount of data involved in this project and this standard ensure data security by establishing information security standards and safeguards to prevent patient data from being compromised.

RxNorm

As a normalized drug naming system 6, RxNorm is used for both patients' medication history & current medications and medication prescription in our project. This standard is particularly important because it standardizes medications and reconciles conflicts.

UMLS

It is important because it helps improve "understanding" of computer programs for the biomedical meaning and helps in information retriving and integration for users.

CAHPS

This is important because we have follow-up visit with the patients and we need to assess their experiences. This standard ensures the assessment is reliable and consistent.

ISO/IEEE 11073

This standard is important because it facilitates the communication between internal medical devices and external computer systems which are both needed in our project since we need the classification by machine learning.

Conclusion

The implementation of this DSS in clinical settings can enhance the precision of cardiovascular disease diagnosis and the personalization of treatment plans, leading to improved patient outcomes.

Future Use Cases

- The architecture and algorithms used in this DSS can be adapted to manage other chronic diseases such as diabetes or respiratory conditions, broadening the scope of the system's applicability.
- Future enhancements could include real-time data analysis from wearable health devices, providing continuous updates to patient profiles and allowing more dynamic risk assessments.
- Integrating the DSS with patient-facing applications could empower individuals to monitor their health proactively and features could include personalized health tips & medication reminders.
- As ML and data analytics evolve, further development could incorporate more sophisticated models and ensemble techniques to improve the accuracy and reliability of predictions.
- Collaborating with public health organizations to integrate with national health databases could enhance epidemiological studies and improve the overall quality of healthcare services.

References:

1. Secureframe. (n.d.). What is HIPAA Compliance and Why is it Important? Secureframe. https://secureframe.com/hub/hipaa/what-is-hipaa-compliance#:~:text=Because%20it%20establishes%20information%20security,for%20keeping%20sensitive%20data%20safe.
2. Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med. 1993 Aug;32(4):281-91. doi: 10.1055/s-0038-1634945. PMID: 8412823; PMCID: PMC6693515.
3. Consumer Assessment of Healthcare Providers & Systems (CAHPS) | CMS. (n.d.). https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems#:~:text=CAHPS%20surveys%20follow%20scientific%20principles,be%20compared%20across%20healthcare%20settings.
4. ISO/IEEE 11073-20601:2022(en) Health informatics — Device interoperability — Part 20601: Personal health device communication — Application profile — Optimized exchange protocol. (n.d.). In ISO. https://www.iso.org/obp/ui/en/#iso:std:iso-ieee:11073:-20601:ed-3:v1:en

Thank you very much!