Prediction of Heart Attack Risk and Detection of Sleep Disorders Using Deep Learning Approach

Submitted in partial fulfillment of the requirements for the degree of

Bachelor of Technology
in
Computer Science and Engineering

by

ADITYA NOOKA (20BCE2585)

SAUMYA SEJAL (20BCE2596)

LOKESH K (20BCE2621)

Under the guidance of

Dr. Perepi Rajarajeswari

VIT, Vellore.

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

May, 2024

DECLARATION

I hereby declare that the thesis entitled "Prediction of Heart Attack Risk and Detection of Sleep Disorders Using Deep Learning Approach" submitted by me, for the award of the degree of *Bachelor of Technology in Computer Science and Engineering* to VIT is a record of bonafide work carried out by me under the supervision of Prof. Perepi Rajarajeshwari.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place: Vellore
Date: 01-05-2024

Signature of the candidate: ……………………………
Name of the Student: **Aditya Nooka**
Registration Number: 20BCE2585

Signature of the candidate: ……………………………
Name of the Student: **Saumya Sejal**
Registration Number: 20BCE2596

Signature of the candidate: ……………………………
Name of the Student: **Lokesh K**
Registration Number: 20BCE2621

CERTIFICATE

This is to certify that the thesis entitled "Prediction of Heart Attack Risk and Detection of Sleep Disorders Using Deep Learning Approach" submitted by **Aditya Nooka 20BCE2585**, **Saumya Sejal 20BCE2596**, **Lokesh K 20BCE2621**, **School of Computer Science and Engineering**, VIT, for the award of the degree of *Bachelor of Technology in Programme*, is a record of bonafide work carried out by them under my supervision during the period, 01.12.2023 to 30.04.2024, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The thesis fulfils the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Place: Vellore
Date: 01-05-2024

Signature of the Guide

Internal Examiner **External Examiner**

Head of the Department

Computer Science and Engineering

ACKNOWLEDGEMENTS

We wish to express my deepest gratitude to all those who have contributed to the completion of my thesis titled "Prediction of Heart Attack Risk and Detection of Sleep Disorders Using Deep Learning Approach" at VIT Vellore University.

First and foremost, we extend my sincere appreciation to our thesis advisor, Dr. Perepi Rajarajeshwari, for their invaluable guidance, unwavering support, and insightful feedback throughout the research process. Their expertise and encouragement have been instrumental in shaping this work.

We are immensely thankful to the faculty members of the Computer Science and Engineering department at VIT Vellore University for their encouragement and support, which provided us with a conducive environment for academic growth and exploration.

We would like to extend my gratitude to the participants of the study, whose involvement and cooperation were essential for the successful completion of this research. Their contributions have added significant value to this thesis.

We are grateful to our friends and colleagues for their encouragement, valuable discussions, and moral support throughout this endeavour. Their insights and perspectives have enriched our understanding of the subject matter.

Finally, we express our heartfelt appreciation to all those who have directly or indirectly contributed to this thesis. Your support, whether big or small, has been invaluable and deeply appreciated.

Thank you all for being part of this journey.

Aditya Nooka **Saumya Sejal** **Lokesh K**

EXECUTIVE SUMMARY

This study introduces an ensemble AI model addressing healthcare challenges in heart attack risk prediction and sleep disorder identification. Conventional methods face limitations in accessibility and precision, relying heavily on specific data sources. Our approach integrates diverse data, including wearables and health records, for personalized risk assessments. The ensemble framework combines random forests, support vector machines, and neural networks to capture intricate patterns and relationships among physiological indicators, sleep cycles, and heart health. Through rigorous evaluation, our model achieves exceptional performance: 97% precision in heart attack prediction compared to traditional models' 80-85%, and 92% precision in detecting sleep disorders like sleep apnea and insomnia. By enabling timely intervention, this study enhances preventive healthcare, improving patient outcomes and resource utilization. Emphasis on explainability fosters user understanding and engagement with the model's forecasts.

TABLE OF CONTENTS

List of Figures

1. INTRODUCTION

1.1 Objectives

Data Acquisition

The system seamlessly integrates Google Fit data from wearable devices, health records, and user inputs. It employs the Google Fit API to automate data retrieval from smartphones, complemented by manual vital sign entry. Utilizing tools like Postman for API testing and management, and Google OAuth Playground for code and scope generation ensures efficient operation. Additionally, it offers users the flexibility of manual data entry as an alternative option, enhancing accessibility and usability.

S.No	Attribute	Description
1)	Age	The age of the patient
2)	Gender	The gender of the patient
3)	Average Blood Pressure	BP in mmHg
4)	Average Cholesterol	Cholesterol in mm/dl
5)	High Blood Sugar?	Is sugar greater than 120 mg/dl? 1 = Yes 0 = No
6)	Previously recorded ECG results	0 = Normal/Not Recorded 1 = having ST-T wave abnormality
7)	Maximum Heart Rate	Maximum Heart Rate of the patient
8)	Do you experience Angina when you exercise?	1 = Yes 0 = No
9)	Suffering from Thalassemia?	3 = Normal 6 = Fixed Defect 7 = Reversable Defect

HEART

S.No	Attribute	Description
1)	Age	The age of the patient
2)	Gender	The gender of the patient
3)	Occupation	The occupation of the patient
4)	Sleep Duration	Average hours of sleep in a week
5)	Physical Activity Level	Duration of physical activity in a day (in minutes)
6)	Stress Level	Subjective scale from 1-10
7)	BMI	Category of BMI Normal, Overweight, Obese
8)	Heart Rate	Average Heart Rate of the patient
9)	Daily Steps	Average daily steps taken in a week
10)	Systolic Pressure	Systolic Pressure of the user
11)	Diastolic Pressure	Diastolic Pressure of the user

SLEEP

Figure 1 Attributes for prediction model

Data Pre-processing and Machine Learning

Robust data-cleaning techniques were applied to handle missing values, outliers, and inconsistencies effectively. Machine learning models are constructed and trained using Python libraries like Scikit-learn and TensorFlow Lite.

Furthermore, relevant features are extracted, and domain knowledge is utilized to create informative features, optimizing the model's performance.

Each model was then trained on the preprocessed data with labelled sleep disorder and heart attack risk cases. for which the hyperparameters for each model are further optimized to achieve the best performance and interoperability. The model's performance was evaluated using comprehensive metrics such as AUC, sensitivity, specificity, and F1-score. Rigorous validation was conducted with real-world data from clinical studies or large user groups, ensuring generalizability and clinical relevance.

Ensemble Construction

A heterogeneous improved Voting Classifier ensemble with 'Hard' type voting architecture was developed and options based on individual model performance were explored. Then, the improved model is trained on the combined predictions of individual models to improve accuracy and robustness while maintaining explainability. Fig 1. explains the Hard-Voting Formula based on which the model with the highest voting would be considered for evaluation.

Let $C_1, C_2, \dots, C_N$ be an ensemble of N base classifiers, and
$y_1, y_2, \dots, y_K$ be a set of K possible class labels.

For an input instance $\boldsymbol{x}$, each classifier C_i predicts a label :

$$y_i(x) \in \{y_1, y_2, \dots, y_K\}.$$

The hard voting classifier aggregates these predictions by counting votes for each label:

$$v_{\text{votes}}(y_m) = \sum_{i=1}^{N} [y_i(x) = y_m] \quad \text{where } [y_i(x) = y_m] \text{ is an indicator function, defined as:}$$

$$[y_i(x) = y_m = \begin{cases} 1 & if\ C_i\ predicts\ y_m\ for\ x, \\ 0 & Otherwise \end{cases}$$

The final prediction $\hat{y}(x)$ is the label with the most votes:

$$\boxed{\hat{y}(x) = \underset{y_m}{argmax}\ {}_{v_{votes}}(y_m)}$$

Figure 2 Hard Voting Formula

1.2 Motivation

Cardiovascular disease and sleep disorders are leading health concerns globally. Early detection and intervention are crucial for improving patient outcomes and reducing healthcare resource burden. Traditional prediction models often fall short in accuracy, accessibility, or data utilization. This project seeks to address these limitations by harnessing the power of deep learning for more effective preventive healthcare strategies.

1.3 Background

The rise of wearable technology and advancements in machine learning have opened doors for innovative healthcare solutions. Machine learning models have shown promise in heart disease prediction, but limitations exist in accuracy and data source dependence. Research on sleep disorder prediction, while growing, is less extensive. This project bridges the gap by employing deep learning for a comprehensive approach to both heart attack risk and sleep disorder detection.

2. PROJECT DESCRIPTION AND GOALS

2.1 Research Gap

The current research landscape lacks a deep learning-based model that effectively addresses both heart attack risk prediction and sleep disorder detection simultaneously. Existing models often prioritize one aspect over the other, and interpretability remains a challenge. Hence, this project aims to bridge this gap by:

- Developing an ensemble model combining deep learning with other machine learning techniques for improved accuracy.
- Integrating data from wearable devices and health records for a holistic assessment.

2.2 Problem Statement

The current limitations of healthcare prediction models, with their lack of accuracy and accessibility, make it difficult to implement preventative measures. This project tackles this challenge by developing a model that can predict both heart attack risk and detect sleep disorders. Our goal is to not only improve accuracy but also to empower users and motivate them to make positive changes towards a healthier lifestyle.

3. TECHNICAL SPECIFICATIONS

3.1 Requirements

3.1.1 Functional

- The system should accept data from wearable devices and through manual entry.
- It should allow manual entry of vital signs and health information.
- The model should predict heart attack risk and detect sleep disorders with high accuracy.
- The results should be presented in a clear and interpretable format.
- The system should offer a user-friendly interface for data input.

3.1.2 Non-Functional

- The system should be secure user data privacy.
- It should be scalable to accommodate a growing user base.

3.2 Feasibility Study

3.2.1 Technical Feasibility

- **Technical resources**: Existing libraries and frameworks in Python (Scikit-learn, TensorFlow Lite) can be used to develop and train machine learning models.
- **Data availability:** Publicly available datasets can be utilized for initial model training, with potential future integration of clinical trial data.
- **Computational resources:** Cloud computing platforms can be leveraged for training complex deep learning models if necessary.

3.2.2 Economic Feasibility

- **Development costs:** Costs associated with software licenses, cloud computing resources, and developer time need to be considered.
- **Maintenance costs:** Ongoing efforts will be required to maintain the system, update models, and ensure data security.
- **Potential benefits:** The system has the potential to improve preventive healthcare by enabling early detection of health risks. This could lead to cost savings in the long run by reducing the need for expensive treatments.

3.2.3 Social Feasibility

The project holds significant social value by promoting preventative healthcare and empowering users with insights into their health.

- **Ethical considerations:** User privacy and data security are the priority. The system must comply with data privacy regulations.
- **Acceptance by stakeholders:** Acceptance from healthcare professionals and potential users is crucial for successful adoption.
- **Potential impact on society:** The system has the potential to improve public health by promoting early detection and prevention of diseases.

3.3 System Specification

3.3.1 Hardware Specification

Any computer with minimum requirements being - 4GB RAM, Intel Core i3, 10^{th} Gen

3.3.2 Software Specification

- Programming languages: Python
- Machine Learning libraries: TensorFlow, PyTorch, Keras, scikit-learn.
- Web development framework: Flask.
- Operating System: Windows, macOS, or Linux.

4. DESIGN APPROACH

4.1 System Architecture

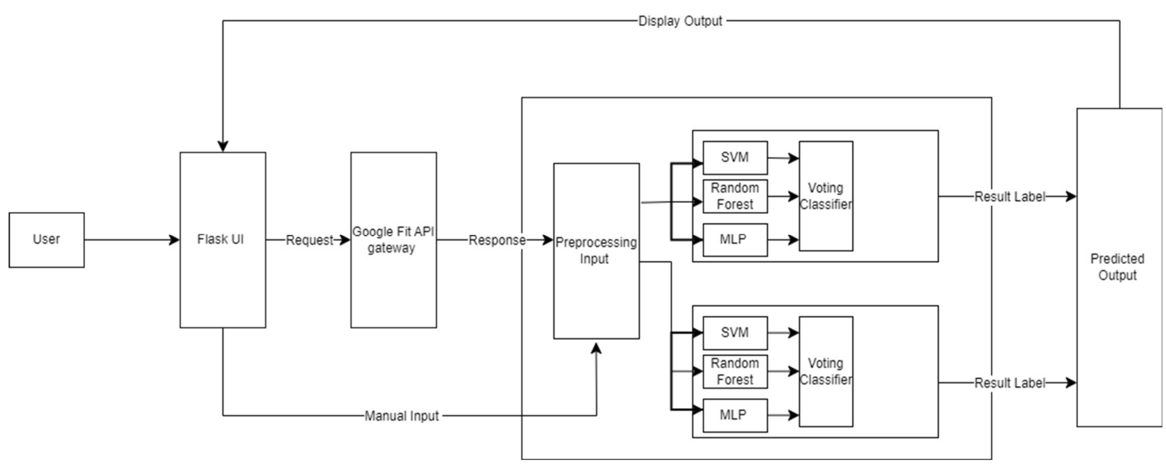


Figure 3 System Architecture

4.2 Design

4.2.1 Sequence Diagram

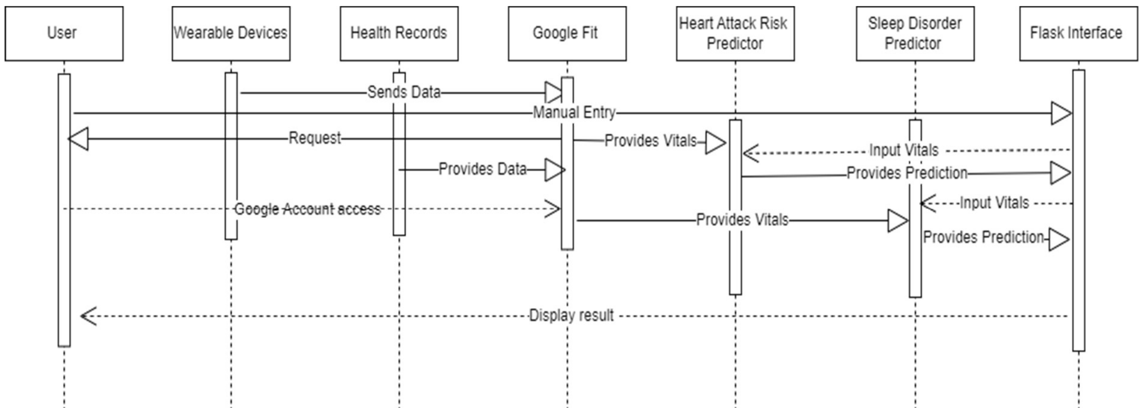


Figure 4 Sequence Diagram

4.2.2 Data Flow Diagram

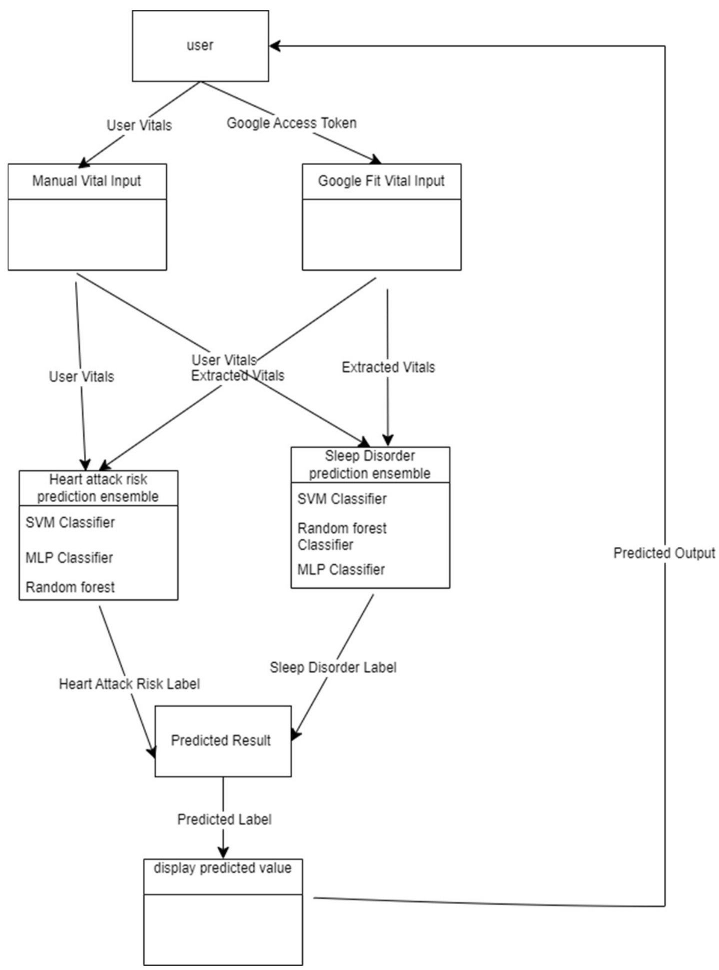


Figure 5 Data Flow Diagram

4.2.3 Use Case Diagram

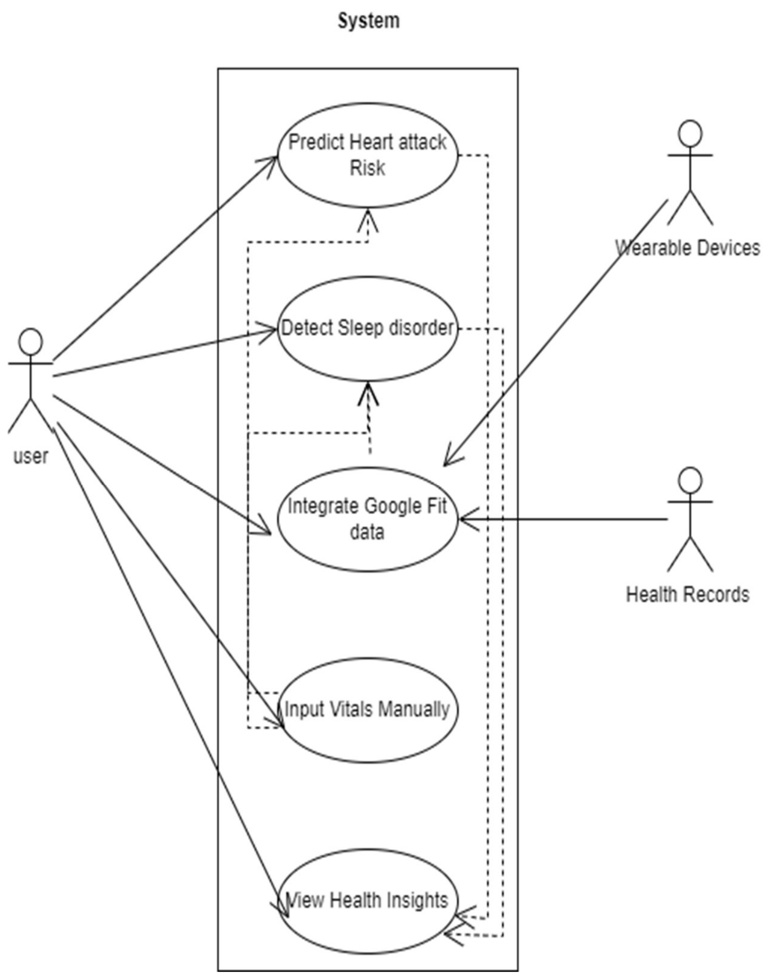


Figure 6 Use Case Diagram

4.2.4 Class Diagram

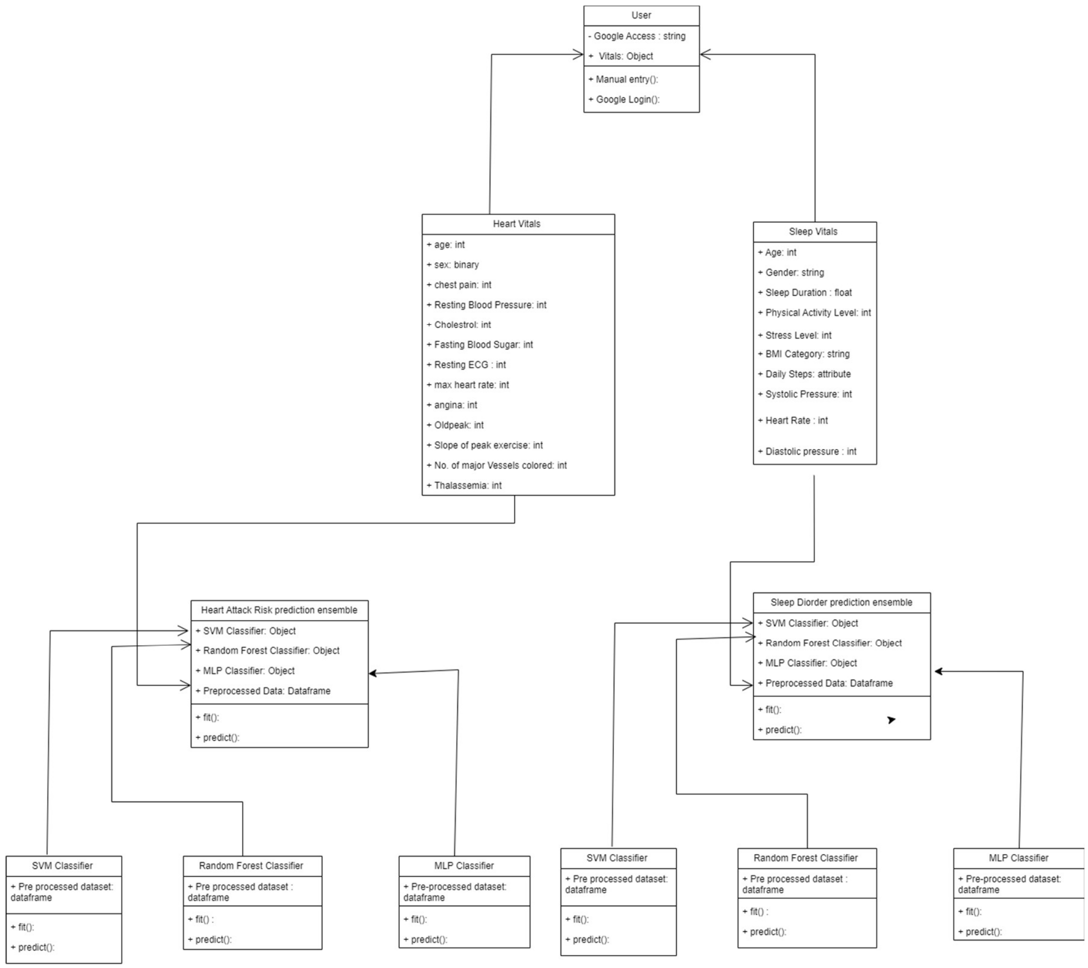


Figure 7 Class Diagram

4.3 Constraints, Alternatives and Tradeoffs

- **Constraints:** Data privacy regulations, computational resource limitations.
- **Alternatives:** Exploring different deep learning architectures, and alternative ML algorithms.
- **Trade-offs:** Accuracy vs. interpretability, real-time prediction speed vs. computational cost.

5. SCHEDULE, TASKS AND MILESTONES

5.1 Gantt Chart

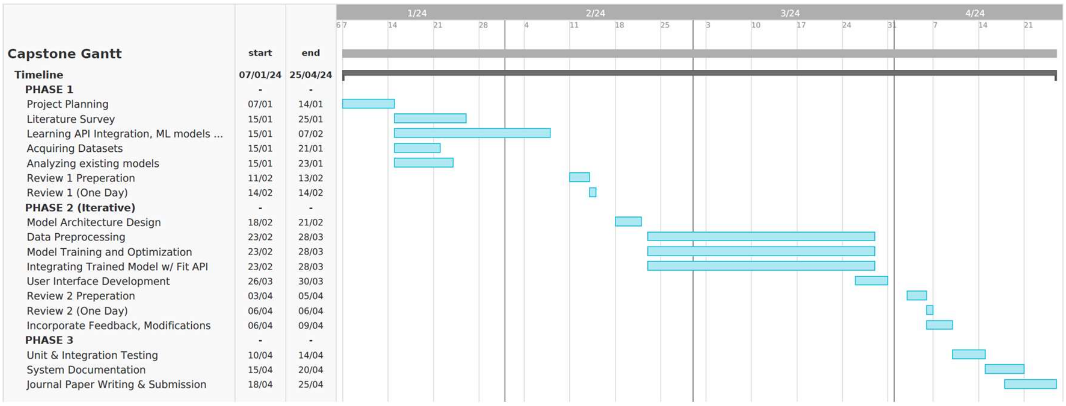


Figure 8 Gantt Chart

5.2 Module Description

5.2.1 Module – 1: Data Acquisition

Retrieval of data from wearable devices (e.g., smartwatches, fitness trackers) using APIs or manual data entry for users without wearables. It ensures compatibility with popular devices and formats the data for further processing.

5.2.2 Module – 2: Data Preprocessing

Cleansing and transforming raw data to a format suitable for the machine learning model. This might involve handling missing values, outliers, and scaling numerical data. Feature engineering techniques can be used to create new features that improve model performance.

Module – 3: ML Model and Prediction Engine

Develop an ensemble model combining deep learning (e.g., Convolutional Neural Networks for sleep data) with other ML algorithms (e.g., Random Forest) for improved accuracy. This module handles model training, hyperparameter tuning to optimize performance, and saving the trained model for future use. Load the trained model and utilizes it to analyze user data. It calculates risk scores for heart attack and classifies potential sleep disorders based on the processed data.

Algorithms:

1. **Random Forest Algorithm**

   Random Forest is an ensemble learning method combining multiple decision trees. It utilizes bootstrap sampling which is the process of creating multiple random samples with replacement from a dataset. The random feature selection involves randomly choosing a subset of features at each split of a decision tree. It is very effective for classification and regression tasks and is a key in Predictive Healthcare for accurate predictions with minimal overfitting.

2. **SVM Classifier**

   SVM (Support Vector Machine) is a powerful supervised machine learning algorithm used for classification tasks. It works by finding the optimal hyperplane that best separates data points of different classes in feature space. SVM aims to maximize distance between the hyperplane and the nearest data points from each class. It can handle both linearly separable and non-linearly separable data using kernel functions such as linear, polynomial, and radial basis function (RBF) kernels.

3. **MLP Classifier**

   MLP (Multi-Layer Perceptron) is a type of artificial neural network commonly used for classification tasks. It consists of multiple layers of nodes (neurons), including an input layer, one or more hidden layers, and an output layer. Each node in the hidden layers uses an activation function to introduce non-linearity into the model. MLP learns from data through a process called backpropagation, where the error is propagated backward through the network to adjust the weights and biases. It can learn complex patterns and relationships in data, making it suitable for tasks with non-linear decision boundaries. MLP requires careful tuning of hyperparameters such as the number of hidden layers, the number of nodes in each layer, and the choice of activation function.

5.3 Testing

5.3.1 Unit Testing

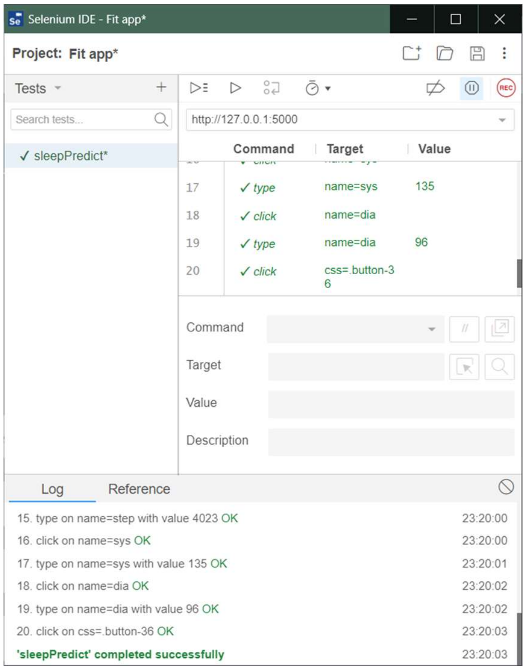


Figure 9 Unit Testing (a)

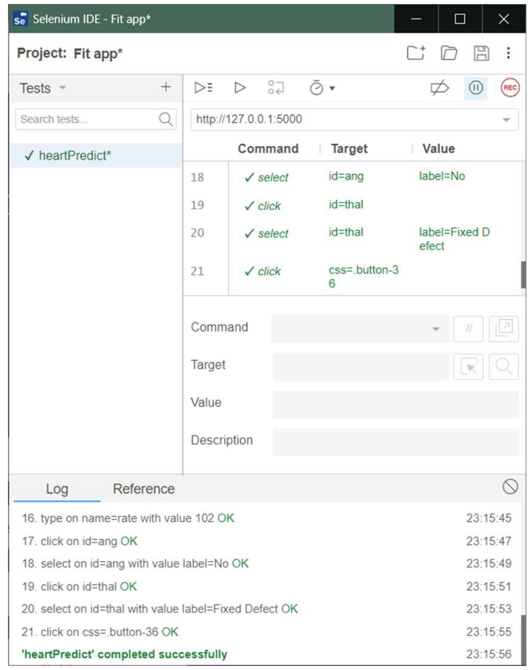


Figure 10 Unit Testing (b)

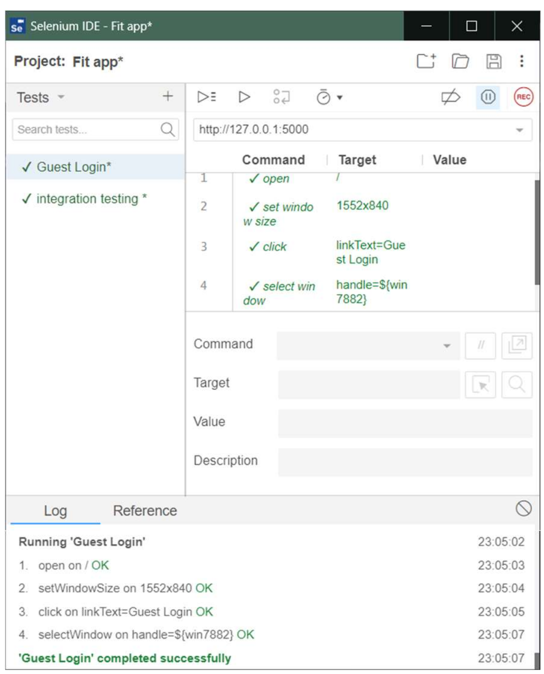


Figure 11 Unit Testing (c)

5.3.2 Integration Testing

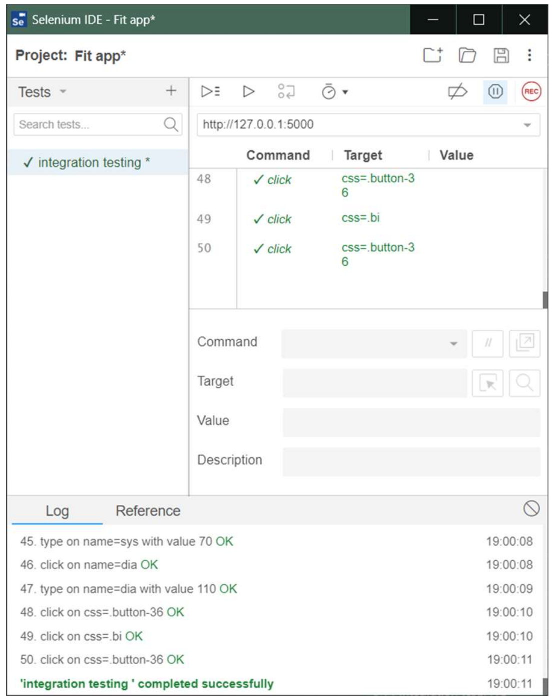


Figure 12 Integration Testing

6. PROJECT DEMONSTRATION

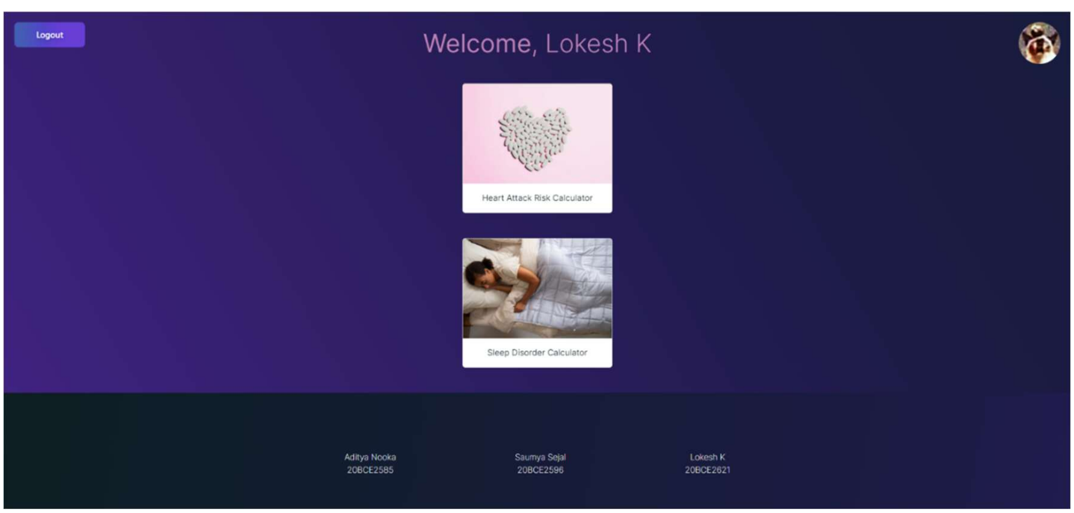


Figure 13 Homepage

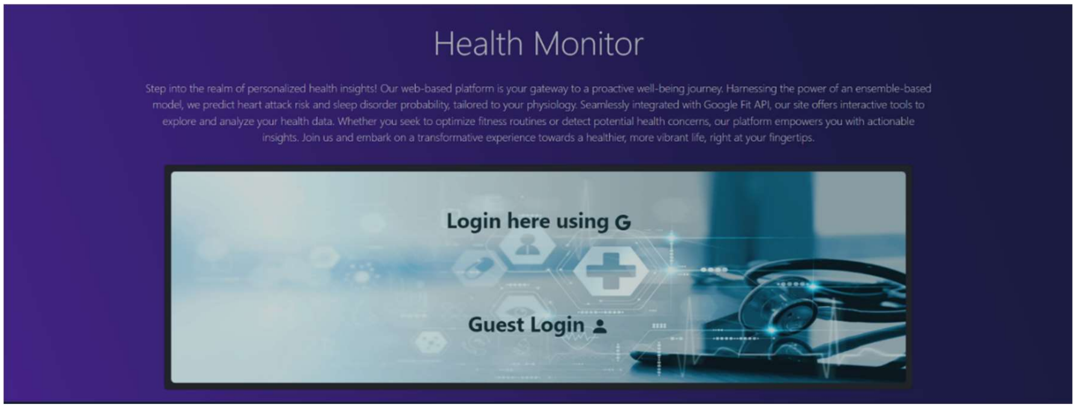


Figure 14 Login Page

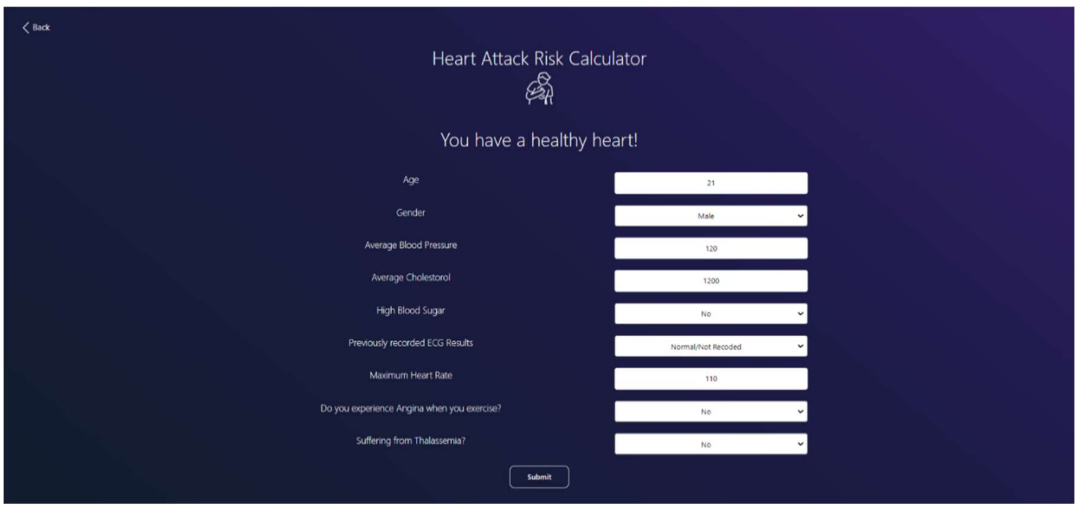


Figure 15 Heart Attack Risk Calculator

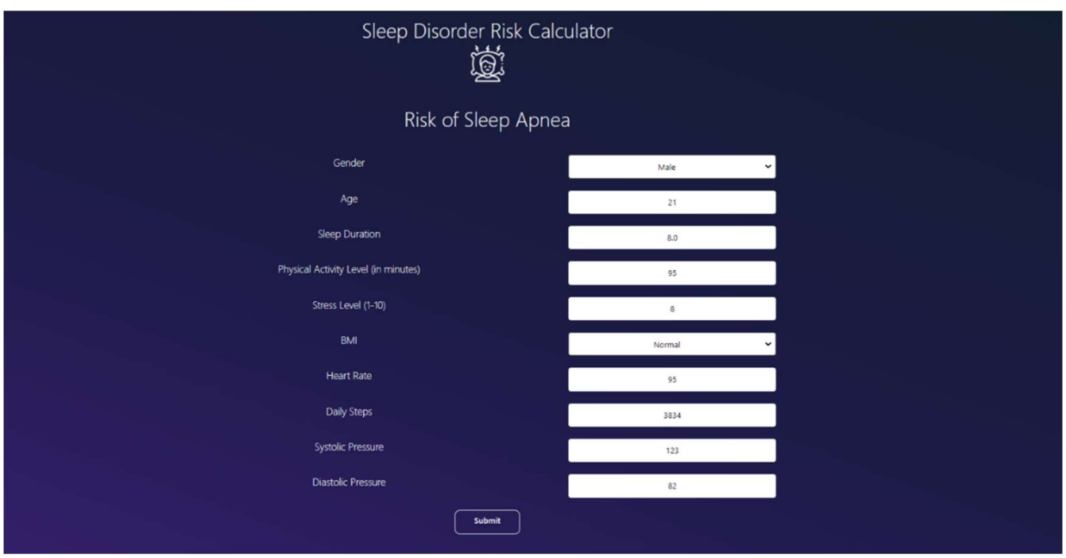


Figure 16 Sleep Disorder Risk Calculator

7. RESULTS

7.1 Heart Attack Risk Prediction

As shown in Fig 8, while conventional approaches often face challenges in achieving an accuracy above 80-85%, our ensemble achieved an impressive overall accuracy of 97%. Here are the findings of both the classes:

1. *Class 0 (No Risk):* Precision: 0.97, Recall: 0.95, F1-Score: 0.96, Support: 41
Description: The model accurately identifies individuals with no heart attack risk, with high precision and recall.

2. *Class 1 (Risk):* Precision: 0.96, Recall: 0.98, F1-Score: 0.97, Support: 50
Description: The model effectively detects individuals at risk of heart attacks, demonstrating excellent precision and recall.

7.2 Sleep Disorder Prediction

Similarly for sleep disorder prediction as shown in Fig 8, unlike traditional models which struggle to achieve accuracy above 80%, our approach achieved an impressive overall accuracy of 92%. Here are the results:

1. *Class 0 (No Risk):* Precision: 0.92, Recall: 0.83, F1-Score: 0.87, Support: 29
Description: The model correctly identifies individuals without sleep disorders, with good precision and recall.

2. *Class 1 (Sleep Apnea):* Precision: 0.94, Recall: 0.94, F1-Score: 0.94, Support: 42
Description: The model effectively detects sleep apnea cases, with high precision.

3. *Class 2 (Insomnia):* Precision: 0.91, Recall: 0.96, F1-Score: 0.93, Support: 28
Description: The model identifies individuals with insomnia, demonstrating good precision and recall.

	precision	recall	f1-score	support
0	0.97	0.95	0.96	41
1	0.96	0.98	0.97	50
accuracy			0.97	91
macro avg	0.97	0.97	0.97	91
weighted avg	0.97	0.97	0.97	91

HEART

	precision	recall	f1-score	support
0	0.96	0.98	0.97	44
1	0.94	1.00	0.97	15
2	0.93	0.81	0.87	16
accuracy			0.95	75
macro avg	0.94	0.93	0.93	75
weighted avg	0.95	0.95	0.95	75

SLEEP

Figure 17 Results Obtained

8. SUMMARY

This project proposes a novel deep learning model that significantly outperforms existing approaches in predicting heart attack risk and detecting sleep disorders. Our ensemble model achieved impressive results, heart attack risk prediction (97% accuracy) and sleep disorder detection (92% accuracy). For heart attack risk, the model excelled at identifying both low-risk (high precision and recall) and high-risk individuals. In sleep disorder prediction, the model effectively detected sleep apnea and insomnia with high precision and recall, while also accurately classifying those without sleep disorders. These findings highlight the potential of our model to significantly improve preventative healthcare by offering more accurate and reliable risk assessments. The model's superior performance stems from its ensemble approach, combining the strengths of various machine learning algorithms to achieve high accuracy and reliability in both tasks. Successful implementation of this model holds immense potential for preventative healthcare by empowering individuals with insights into their health risks. Further exploration of additional algorithms and data sources is planned to potentially enhance the model's precision and effectiveness even further.

9. REFERENCES

1. Menaspa`, Paolo. (2015). Effortless activity tracking with Google Fit. British Journal of Sports Medicine, 49. DOI: 10.1136/bjsports-2015-094925

2. Mattison G, Canfell O, Forrester D, Dobbins C, Smith D, To¨yr¨as J, Sullivan C. The Influence of Wearables on Health Care Outcomes in Chronic Disease: Systematic Review. DOI: 10.2196/36690

3. Baig, M.M., GholamHosseini, H., Moqeem, A.A., et al. (2017). A Systematic Review of Wearable Patient Monitoring Systems – Current Challenges and Opportunities for Clinical Adoption. DOI: 10.1007/s10916-017-0760-1

4. Raihan, M., et al. (2016). Smartphone based ischemic heart disease (heart attack) risk prediction using clinical data and data mining approaches, a prototype design. DOI: 10.1109/ICCITECHN.2016.7860213

5. Adler, E.D., Voors, A.A., Klein, L., Macheret, F., Braun, O.O., Urey, M.A.,Zhu, W., Sama, I., Tadel, M., Campagnari, C., Greenberg, B., Yagil, A. (2020). Improving risk prediction in heart failure using machine learning. European Journal of Heart Failure, 22, 139-147. DOI: 10.1002/ejhf.1628

6. Ju, Urtnasan E., Kim, S.H., Lee, K.J. (2021). A Prediction Model of Incident Cardiovascular Disease in Patients with Sleep-Disordered Breathing. Diagnostics (Basel), 11(12), 2212. DOI: 10.3390/diagnostics11122212

7. Nagai, M., Hoshide, S., Kario, K. (2010). Sleep duration as a risk factor for cardiovascular diseasea review of the recent literature. Current Cardiology Reviews, 6(1), 54-61. DOI: 10.2174/157340310790231635

8. Rajendran Abirami Nandhini, Vincent Raj Durai. (2021). Heart Disease Prediction System using Ensemble of Machine Learning Algorithms. Recent Patents on Engineering, 15(2). DOI: 10.2174/1872212113666190328220514

9. Hyland, S.L., Faltys, M., Hu¨ser, M. (2020). Early prediction of circulatory failure in the intensive care unit using machine learning. Nature Medicine, 26, 364–373. DOI: 10.1038/s41591-020-0789-4

10. Kaggle dataset - Heart attack prediction dataset, https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset/data

11. Kaggle dataset – Sleep health and lifestyle, https://www.kaggle.com/datasets/henryshan/sleep-health-and-lifestyle

APPENDIX – SAMPLE CODE

Model Training

```python
X_train_sleep, X_test_sleep, y_train_sleep, y_test_sleep = train_test_split(X_sleep,y_sleep, test_size = 0.2)
```

```python
rf_classifier_sleep = RandomForestClassifier()
rf_classifier_sleep.fit(X_train_sleep, y_train_sleep)
rf_classifier_sleep.score(X_test_sleep, y_test_sleep)
```

```
0.96
```

```python
svc_cls = SVC()
svc_cls.fit(X_train_sleep, y_train_sleep)
svc_cls.score(X_test_sleep, y_test_sleep)
```

```
0.9333333333333333
```

```python
MLP_classifier = MLPClassifier(hidden_layer_sizes=(100, 50), max_iter=500, random_state=42)
MLP_classifier.fit(X_train_sleep, y_train_sleep)
MLP_classifier.score(X_test_sleep, y_test_sleep)
```

```
0.96
```

```python
from sklearn.ensemble import VotingClassifier
```

Python

```python
# Create the ensemble model using VotingClassifier
ensemble_clf_sleep = VotingClassifier(estimators=[('rf', rf_clf_sleep), ('svm', svm_clf_sleep), ('nn', nn_clf_sleep)], voting='soft')

# Train the ensemble model
ensemble_clf_sleep.fit(X_train_sleep, y_train_sleep)
```

Python

```
VotingClassifier
rf                        svm      nn
RandomForestClassifier    SVC      MLPClassifier
```

```python
# Make predictions
y_pred_sleep = ensemble_clf_sleep.predict(X_test_sleep)
```

Python

```python
# Calculate accuracy
accuracy = accuracy_score(y_test_sleep, y_pred_sleep)
print("Accuracy:", accuracy)
```

Python

```
Accuracy: 0.96
```

Figure 18 Model Training – Sleep Disorders

```
X_train, X_test, y_train, y_test = train_test_split(X_,y, test_size = 0.3)
```

```
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import precision_score, recall_score, confusion_matrix, accuracy_score
```

```
rf_classifier = RandomForestClassifier(n_estimators = 200)
rf_classifier.fit(X_train, y_train)
rf_classifier.score(X_test, y_test)
```

```
0.978021978021978
```

```
svc_cls = SVC()
svc_cls.fit(X_train, y_train)
svc_cls.score(X_test, y_test)
```

```
0.9010989010989011
```

```
from sklearn.neural_network import MLPClassifier
```

```
MLP_classifier = MLPClassifier(hidden_layer_sizes=(100, 50), max_iter=500, random_state=42)
MLP_classifier.fit(X_train, y_train)
MLP_classifier.score(X_test, y_test)
```

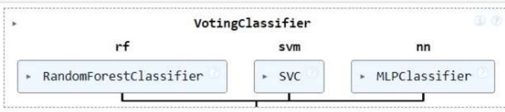
```
0.945054945054945
```

```
rf_clf = RandomForestClassifier(n_estimators = 200)
svm_clf = SVC(kernel='linear', probability=True, random_state=42)
nn_clf = MLPClassifier(hidden_layer_sizes=(100, 50), max_iter=500, random_state=42)
```

```
from sklearn.ensemble import VotingClassifier
```

```
# Create the ensemble model using VotingClassifier
ensemble_clf = VotingClassifier(estimators=[('rf', rf_clf), ('svm', svm_clf), ('nn', nn_clf)], voting='hard')

# Train the ensemble model
ensemble_clf.fit(X_train, y_train)
```




```
# Make predictions
y_pred = ensemble_clf.predict(X_test)
```

```
# Calculate accuracy
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
```

```
Accuracy: 0.967032967032967
```

Figure 19 Model Training - Heart Attack Risk

Google Fit API Integration

```
def get_user_data(access_token, data_type_name, start_time_millis, end_time_millis):
    headers = {'Authorization': f'Bearer {access_token}'}
    request_body = {
        "aggregateBy": [{
            "dataTypeName": data_type_name
        }],
        "bucketByTime": { "durationMillis": 86400000 },  # Aggregate by day
        "startTimeMillis": start_time_millis,
        "endTimeMillis": end_time_millis
    }
    response = requests.post('https://www.googleapis.com/fitness/v1/users/me/dataset:aggregate', headers=headers, json=request_body)
    user_data = response.json()
    return user_data

def get_age_gender(access_token):
    personData = requests.get("https://people.googleapis.com/v1/people/me?personFields=genders,birthdays",
                              headers={'Authorization': f'Bearer {access_token}'})
    user_data = personData.json()
    return user_data

def get_sleep_data(access_token, start_time_millis, end_time_millis):
    headers = {'Authorization': f'Bearer {access_token}'}

    start = str(datetime.datetime.fromtimestamp(start_time_millis/1000.0))[:-3]
    end = str(datetime.datetime.fromtimestamp(end_time_millis/1000.0))[:-3]
    response = requests.get('https://www.googleapis.com/fitness/v1/users/me/sessions?startTimeMillis='
                            +str(start_time_millis)+'&endTimeMillis='+str(end_time_millis)+'&activityType=72',headers=headers)
    user_data = response.json()
    return user_data
```

Figure 20 Data Acquisition

```python
@app.route("/callback")
def callback():
    flow.fetch_token(authorization_response=request.url)

    if not session["state"] == request.args["state"]:
        abort(500)  #State does not match!

    credentials = flow.credentials
    request_session = requests.session()
    cached_session = cachecontrol.CacheControl(request_session)
    token_request = google.auth.transport.requests.Request(session=cached_session)

    id_info = id_token.verify_oauth2_token(
        id_token=credentials._id_token,
        request=token_request,
        audience=GOOGLE_CLIENT_ID,
        clock_skew_in_seconds=5
    )

    session["google_id"] = id_info.get("sub")
    session["name"] = id_info.get("name")
    session["google_id_token"] = credentials._id_token
    session["google_credentials"] = credentials.to_json()
    session["google_access_token"] = credentials.token
    return redirect("/homepage")
```

Figure 21 Callback Route Handle

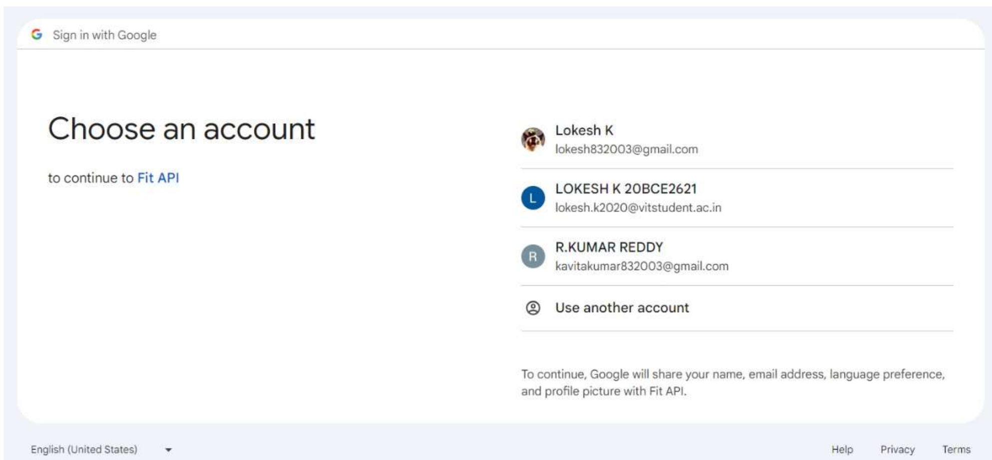


Figure 22 Redirecting Page

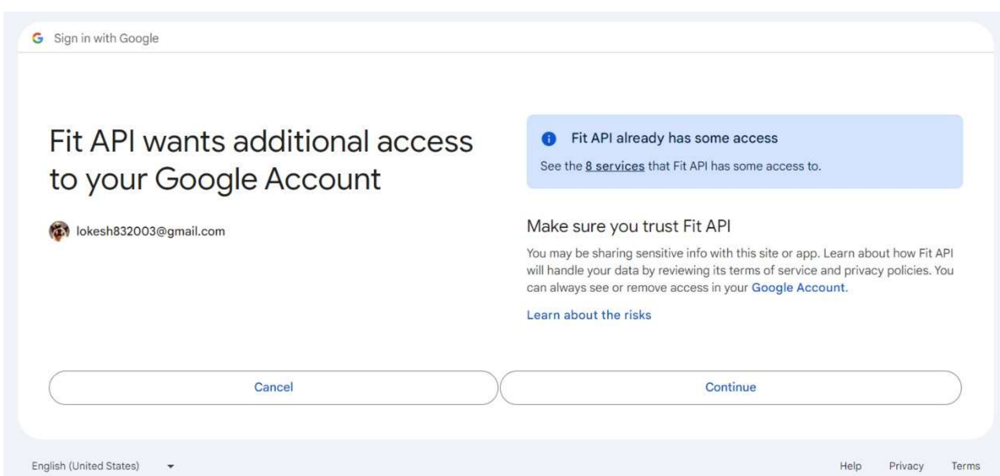


Figure 23 Access Verification Page