24-25J-109

PATIENT-CENTERED MOBILE APPLICATION FOR COMPREHENSIVE DIABETES SELF-MANAGEMENT AND OPTIMIZATION

OUR MEMBERS

Supervisor
Dr.Junius Anjana

Co-Supervisor
Ms.Gaya Thamali Dassanayake

Gunawardhana D.H.M.G

Kajeevan J

De Silva L.K.N

Wimansa
P.P.S.D.

OUR EXTERNAL SUPERVISORS

Dr. RameshKumar Thevraj
Senior Registrar in
Endocrinology
National Hospital Sri Lanka

DR.SOMESWARAPILLLAI PATHANTHAN
GRADE I MEDICAL OFFICER, NATIONAL
HOSPITAL, SRILANKA.

INTRODUCTION

- Integrated mobile app for comprehensive management.
- Real-time data integration from devices.
- Feedback system to track progress and adjust recommendations.
- User-friendly interface to enhance patient engagement and education.

RESEARCH PROBLEM

Insulin Optimization

How can advanced machine learning algorithms be utilized to adjust insulin dosage effectively, minimizing the risks of hypo- and hyperglycemia?

Side Effect Management

How can a system effectively monitor and manage side effects like hypoglycemia, weight gain, and cardiovascular issues?

Simplified Food Tracking

How can automated systems ensure accurate and user-friendly tracking of food intake for diabetes patients?

GUNAWARDHANA D.H.M.G. | IT21289798

RESEARCH GAP

Limited Personalization

Data Overload Challenges

Side Effect Management

Inadequate Food Tracking Tools

Glycemic Variability Insights

GUNAWARDHANA D.H.M.G. | IT21289798

Project Timeline: Gantt Chart

Commercializtion Plan

1. Target Market & Value Proposition
    - Targeting Type 1 & Type 2 diabetes patients, healthcare providers, and clinics.
    - Real-time blood glucose prediction, insulin optimization, and personalized care using AI.
2. Product Development
    - Initial mobile app with cloud-based AI models.
    - Future integration with wearables and medical devices.
3. Marketing Strategy
    - Educational campaigns, social media influencers, and healthcare conferences.
    - Targeted advertising to patients and providers.
4. Scaling & Expansion
    - Strategic partnerships with tech and pharmaceutical companies.
5. Customer Support & Engagement
    - 24/7 support and community building.
6. Financial Projections

- Initial Funding: Seek investment through venture capital, crowdfunding, or partnerships with healthcare organizations to fund product development and marketing.

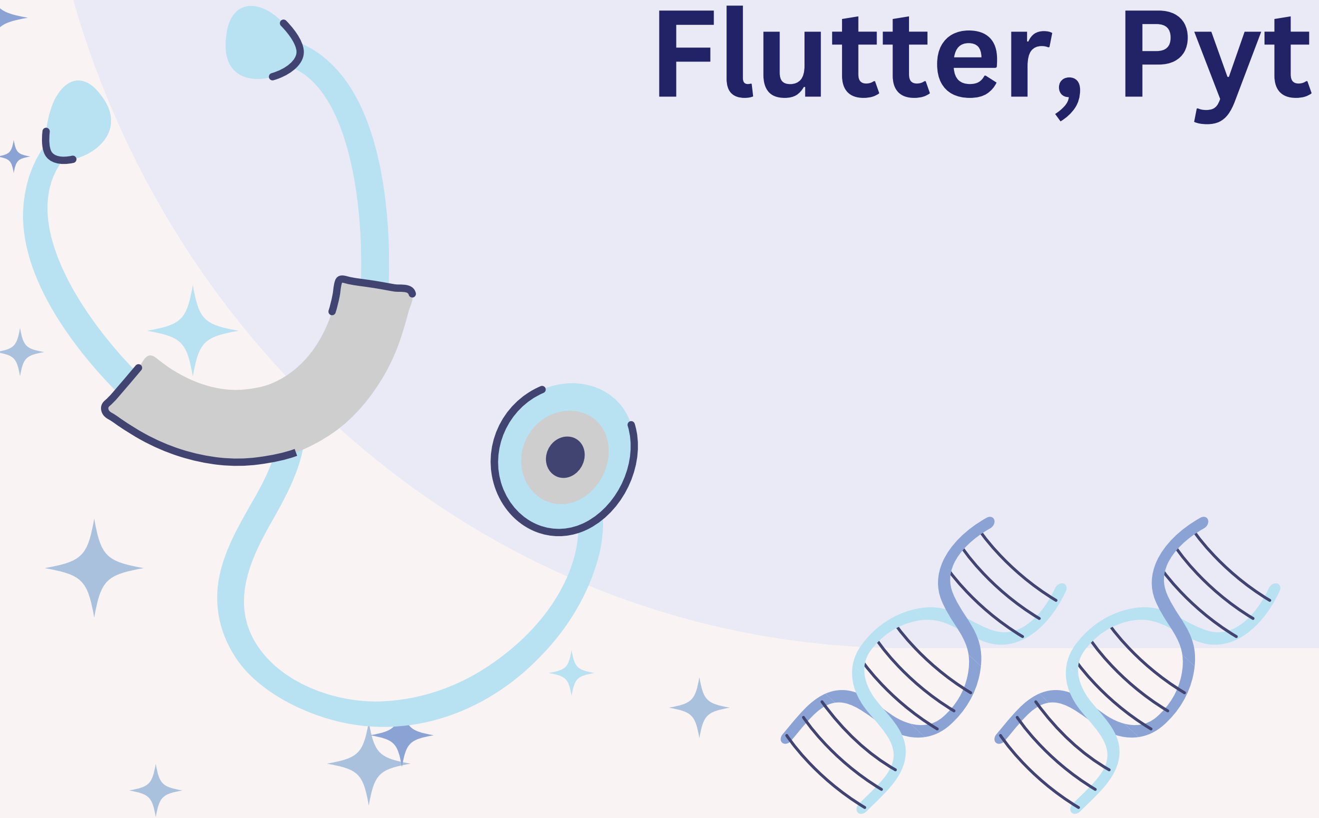
IT21289798| GUNAWARDHANA D.H.M.G.

SPECIALIZATION - DATA SCIENCE

PERSONALIZED INSULIN AND MEDICATION PREDICTION WITH PATIENT ENGAGEMENT

INTRODUCTION

This innovative system leverages AI, time-series analysis, and real-time data processing to predict blood glucose levels and optimize insulin dosing, revolutionizing diabetes care. Built with Flutter, Python, and Cloud Computing.



GUNAWARDHANA D.H.M.G. | IT21289798

RESEARCH PROBLEM

Managing diabetes requires precise blood glucose predictions and insulin dosing. Current methods lack personalization and real-time accuracy, leading to suboptimal care.

RESEARCH SOLUTION

Develop a machine learning-powered system for accurate glucose prediction and insulin optimization using time-series analysis and real-time data, enabling better health outcomes and decision-making.

GUNAWARDHANA D.H.M.G. | IT21289798

RESEARCH GAP

Patient-Centered Design

Limited tools that provide personalized insights tailored to individual patient profiles and conditions.

Data Integration

Lack of systems that effectively combine diverse data sources (e.g., glucose levels, insulin doses, carbohydrate intake, and lifestyle factors).

Decision Support

Insufficient integration of actionable recommendations for healthcare providers and patients to improve treatment decisions.

Technological Integration

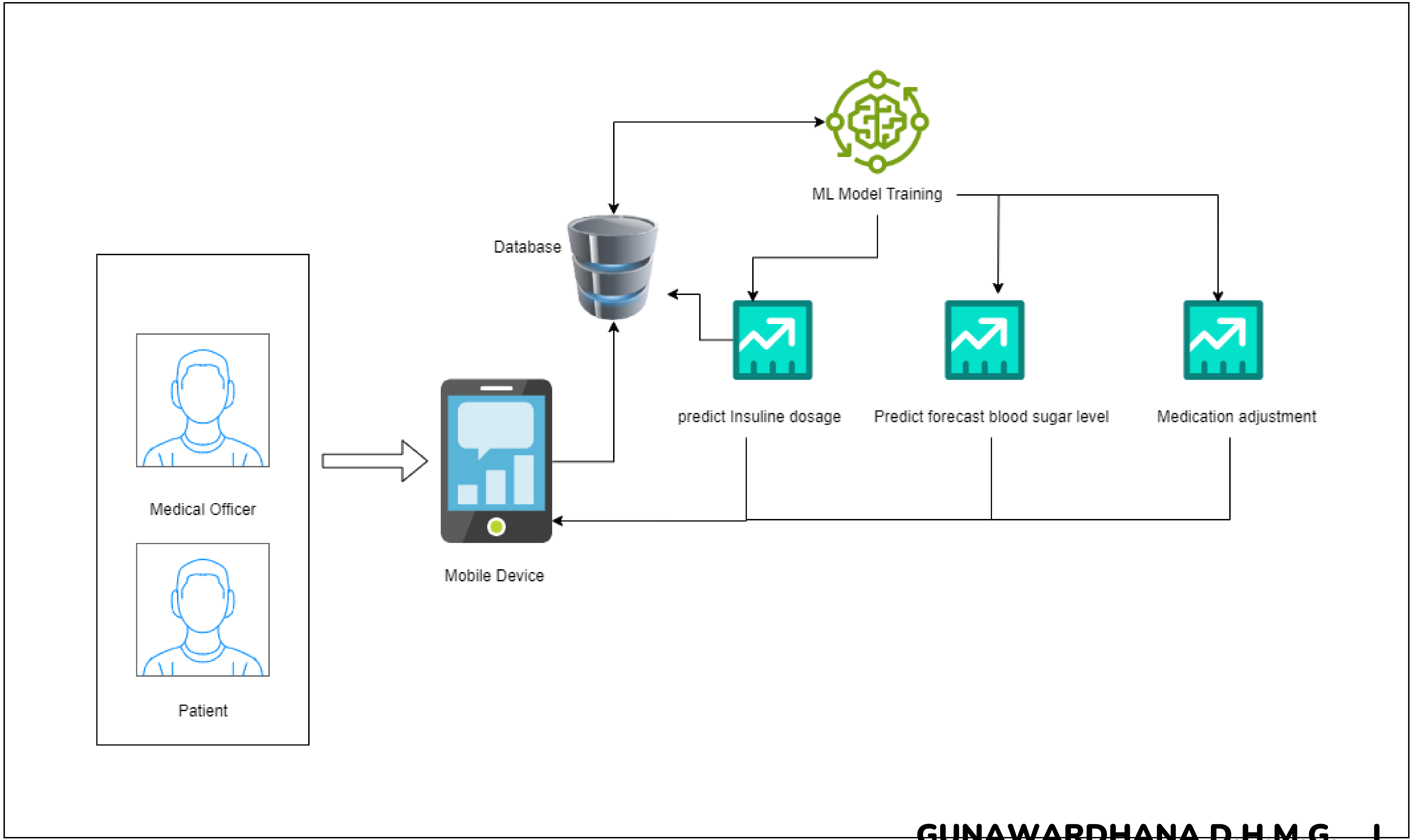
Underutilization of advanced AI, time-series analysis, and cloud computing for seamless and scalable diabetes care solutions.

DATA SET

- Size: 269,768 rows and 104 columns
- Key Features:
- Timestamps: Track blood glucose and insulin changes over time.
- Blood Glucose Levels: Real-time glucose readings.
- Insulin Doses: Administered insulin amounts.
- Carbohydrate Intake: Records of carb consumption affecting glucose levels.
- Lag Variables: Previous values (e.g., glucose_lag_1) for predictive analysis.
- Patient Info: Details like weight and insulin type for personalized care.

GUNAWARDHANA D.H.M.G. | IT21289798

SYSTEM DIAGRAM



FUTURE WORKS

Fine Tune the Model

Intergrate Frontend and Backend

Develop Patient Profile with his data

REFERENCES

- [1] C. R. Marling and R. C. Bunescu, "The OhioT1DM Dataset for Blood Glucose Level Prediction," Proceedings of the 3rd International Workshop on Knowledge Discovery in Healthcare Data, Stockholm, Sweden, 2018.
- [2] D. R. Zisser, J. E. Jovanovic, and M. D. Laffel, "Clinical Evaluation of a Continuous Glucose Monitoring System for Type 1 Diabetes," Journal of Diabetes Science and Technology, vol. 3, no. 4, pp. 1-8, 2009.
- [3] K. C. Santarcangelo, P. J. Johnson, and C. D. Garcia, "A Machine Learning Approach for Predicting Blood Glucose in Diabetes Management," Journal of Healthcare Engineering, vol. 2019, Article ID 2909471, 2019.
- [4] R. D. P. Haug, M. F. Gubbi, and S. P. George, "Optimizing Insulin Dosing: A Data-Driven Approach," IEEE Transactions on Biomedical Engineering, vol. 68, no. 3, pp. 1-10, Mar. 2021.
- [5] S. R. Basak, S. Ghosh, and A. K. Bhattacharya, "Predictive Models for Diabetes Management: A Review," Journal of Medical Systems, vol. 42, no. 6, pp. 23-45, 2018.
- [6] A. G. Ganaie and M. A. Ziarati, "Real-Time Blood Glucose Monitoring and Prediction using Machine Learning Algorithms," IEEE Access, vol. 8, pp. 150345-150353, 2020.
- [7] P. M. Santos and R. M. Lima, "Cloud-Based Healthcare System for Diabetes Management with AI Integration," Journal of Health Informatics, vol. 28, no. 2, pp. 58-70, 2022.
- [8] L. M. Greenfield, "Integrating Time-Series Data for Predictive Analytics in Healthcare," International Journal of Data Science and Analytics, vol. 5, no. 3, pp. 211-219, 2021.

IT21286032| KAJEEVAN J

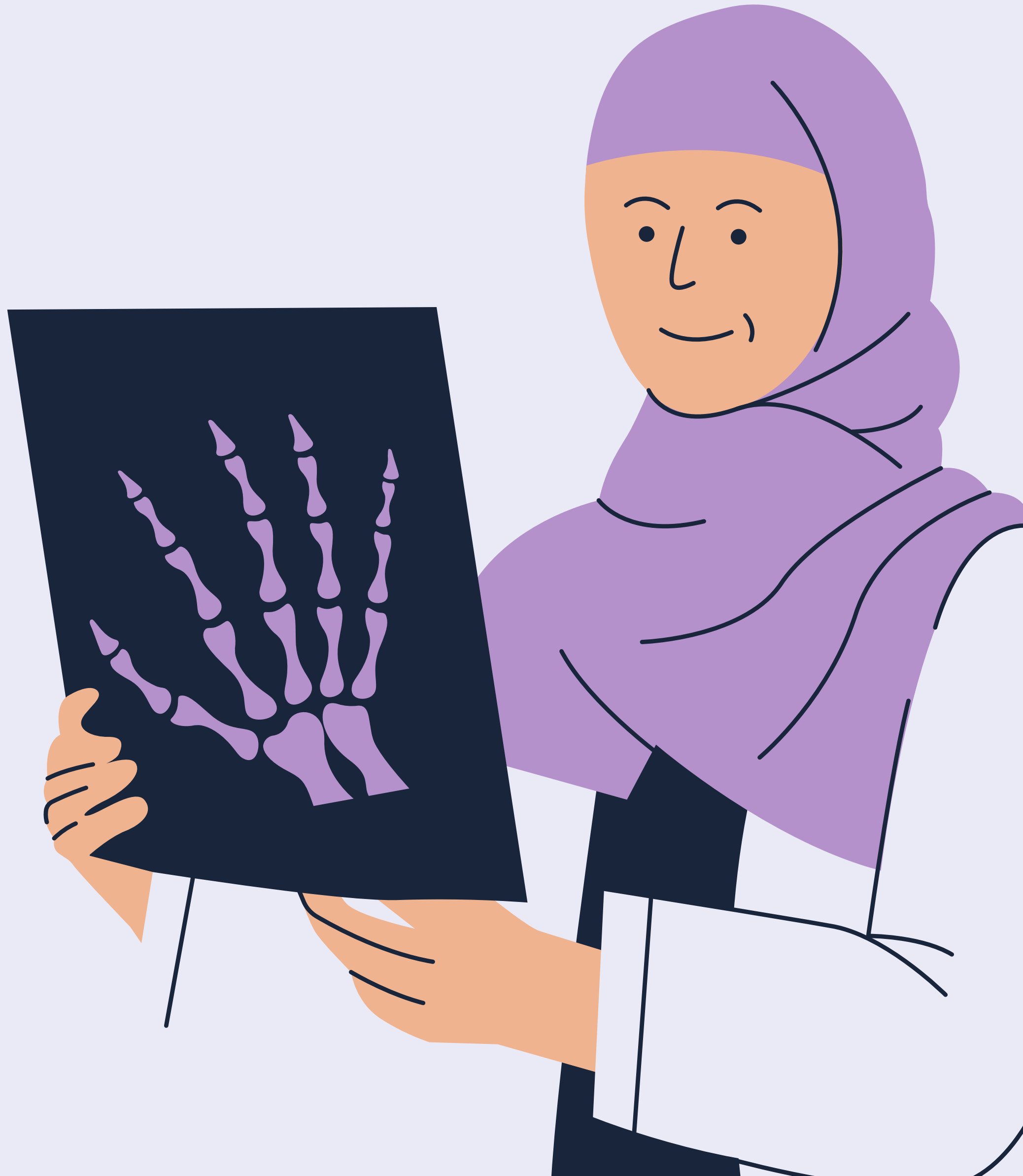
SPECIALIZATION - DATA SCIENCE

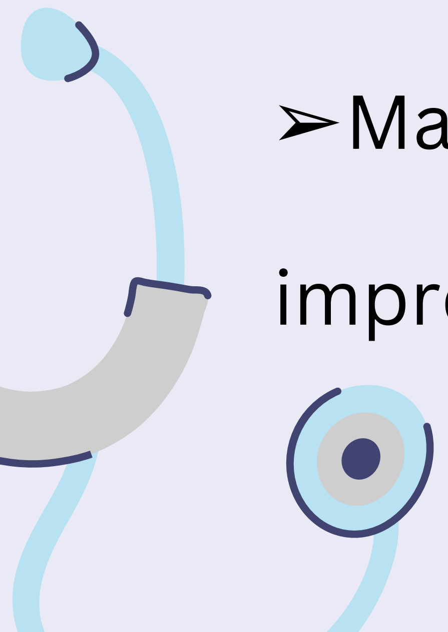
REAL-TIME GLYCEMIC EVENT PREDICTION AND MANAGEMENT TOOL.

Introduction

Diabetes is a growing health challenge in Sri Lanka, with rising cases affecting both individuals and healthcare systems. Effective management requires continuous monitoring of blood glucose levels to prevent glycemic events. This research aims to address gaps by developing a real-time, personalized solution for better diabetes care

- ➢Rising diabetes prevalence in Sri Lanka, impacting public health.
- ➢Existing healthcare solutions lack real-time, personalized diabetes management.
- ➢Machine learning and predictive models show potential for improving glycemic event prediction and management.





Research Question

How can CGM data and machine learning predict and manage glycemic events in real-time for diabetes patients in Sri Lanka?

•What are the benefits of personalized recommendations and real-time alerts in improving diabetes management?

Objectives

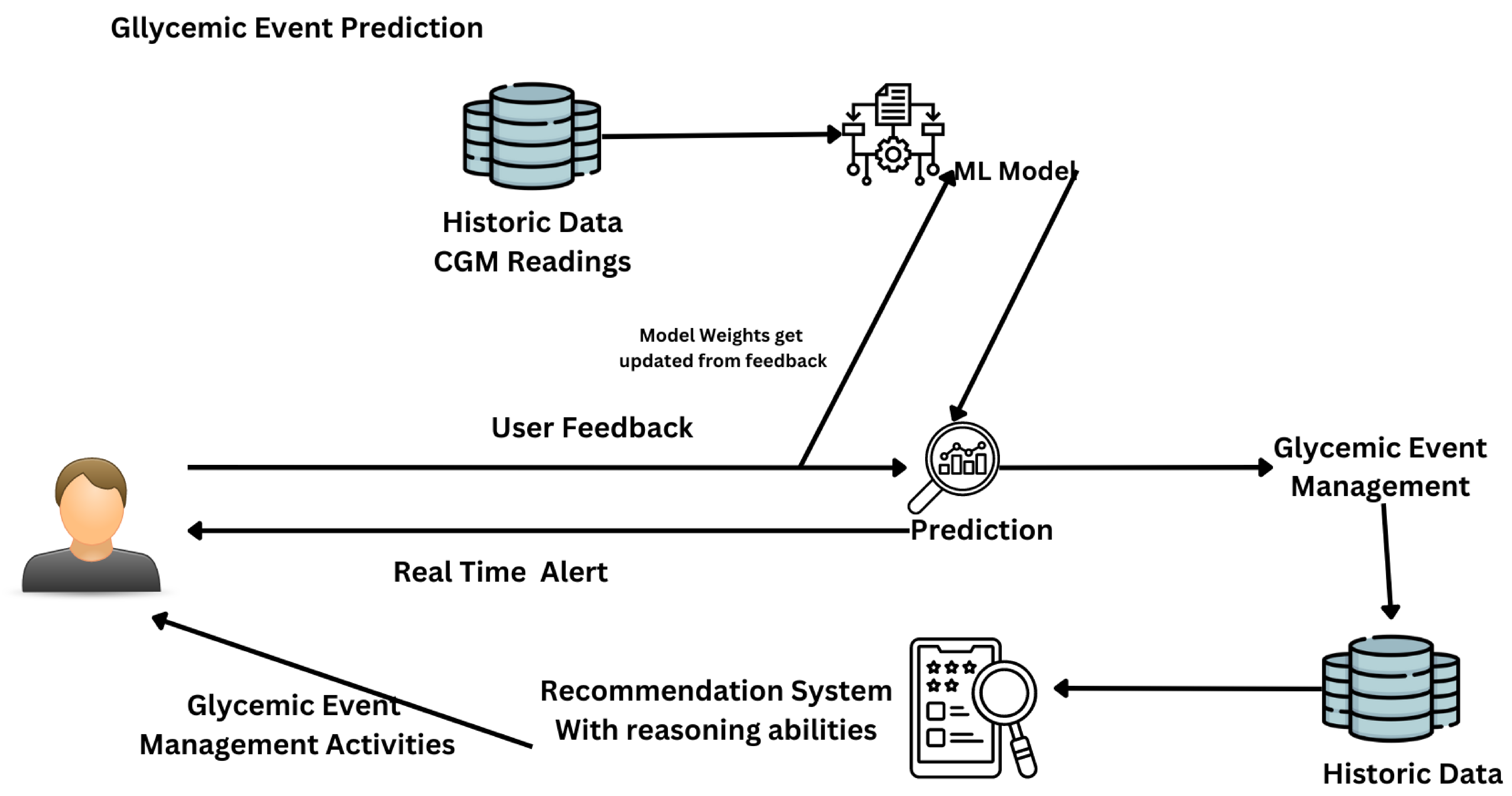
Specific Objectives

- Develop Predictive Algorithms & Management System for Glycemic Events

Sub Objectives

- Integrate CGM data with insulin, activity, and diet.
- Develop predictive models for accurate event forecasting.
- Create a recommendation system for real-time management.

Component Diagram



What I Have Finished:

- **Time Series Machine Learning Model**

Built an GRU-based predictive model to forecast glycemic events using historical CGM data.

- **UI Implementation**

What I Need to Finish:

- **Integrate Real-Time Data:**

Integrate CGM and health band data for continuous monitoring and real-time updates in the mobile app.

- **Personalized Adaptive Recommendation System:**

Build a system for personalized management strategies (insulin adjustments, meal recommendations, alerts) when glycemic events occur.

Technologies

- Backend: Python with Flask APIs
- Frontend: React Native
- Machine Learning: Keras with TensorFlow (GRU for time series prediction), Q-learning for recommendation system
- Data Integration: CGM devices, fitness bands (via APIs)
- Deployment: Kubernetes (for scheduling data extraction jobs)

kubernetes

Proof Of Completion

Mendeley Data

HUPA-UCM Diabetes Dataset

Published: 25 April 2024 | Version 1 | DOI: 10.17632/3hbcscwz44.1
Contributors: J. Ignacio Hidalgo, Jorge Alvarado, Marta Botella, Aranzazu Aramendi, J. Manuel Velasco, Oscar Garnica

Description

This dataset provides a collection of Continuous Glucose Monitoring (CGM) data, insulin dose administration, meal ingestion counted in carbohydrate grams, steps, calories burned, heart rate, and sleep quality and quantity assessment acquired from 25 people with type 1 diabetes mellitus (T1DM). CGM data was acquired by FreeStyle Libre 2 CGMs, and Fitbit Ionic smartwatches were used to obtain steps, calories, heart rate, and sleep data for at least 14 days. This dataset could be utilized to obtain glucose prediction models, hypoglycemia and hyperglycemia prediction models, and research on the relationships among sleep, CGM values, and the rest of the mentioned variables. This dataset could be used directly from the preprocessed version or customized from raw data.

Download All 77.7 MB

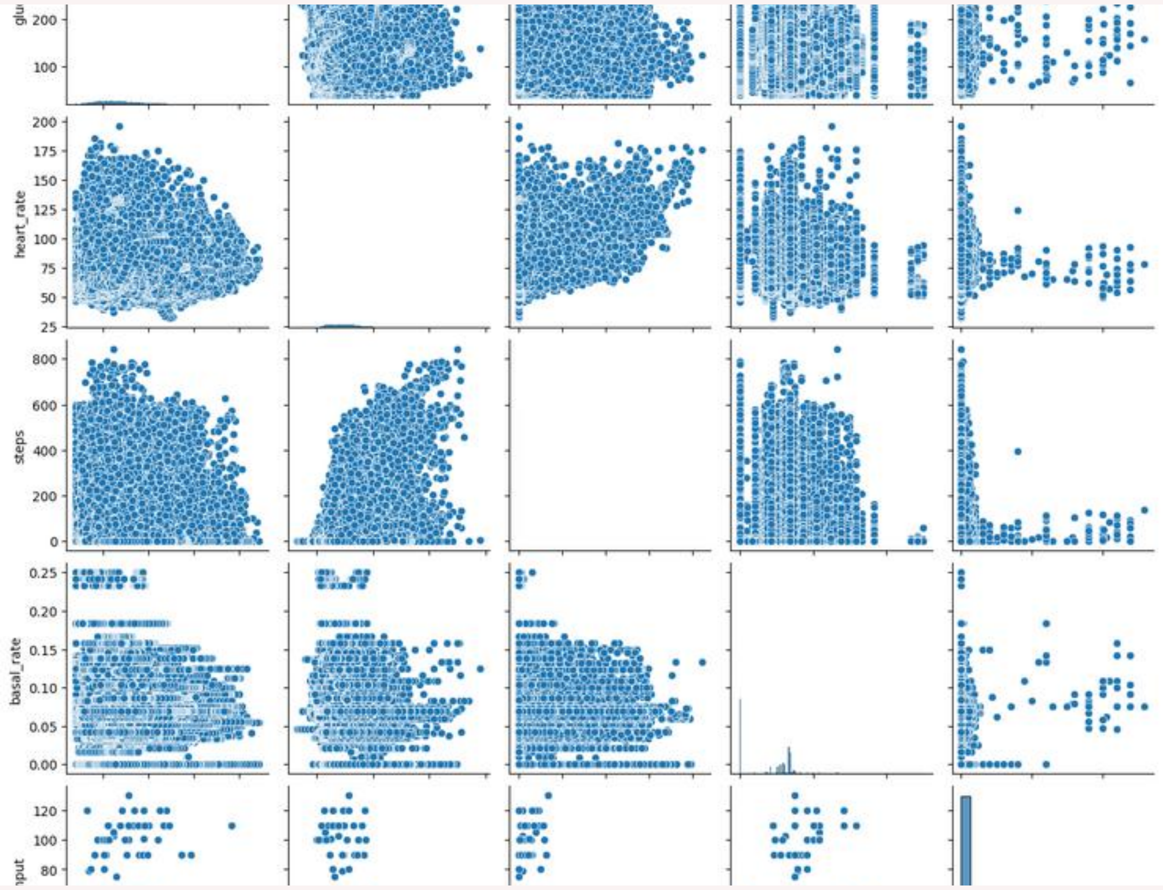
Files

Root > Preprocessed

HUPA0016P.csv	375 KB
HUPA0017P.csv	302 KB

First few rows after merging datasets:

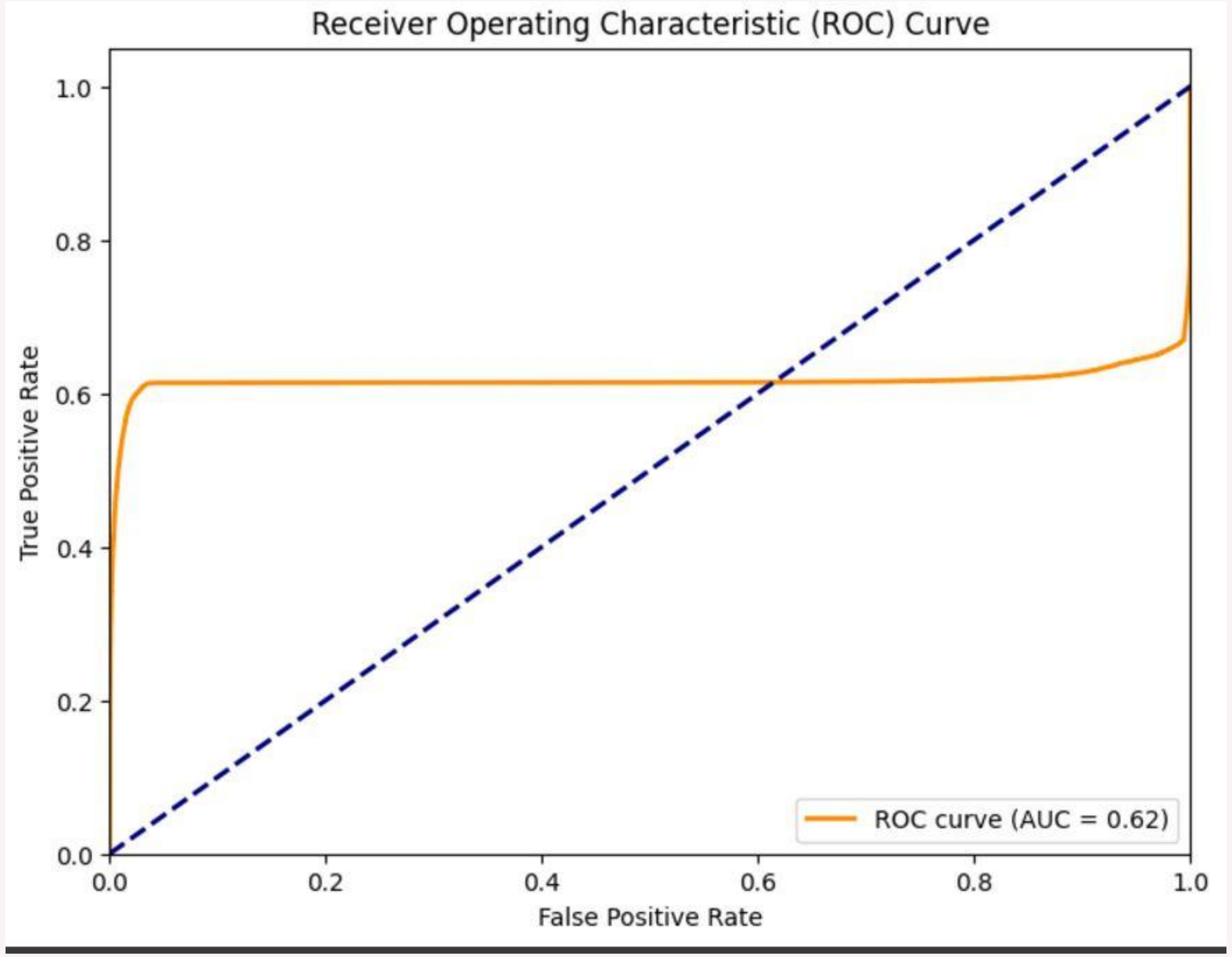
	time	glucose	calories	heart_rate	steps	basal_rate	bolus_volume_delivered	carb_input
0	2018-06-13T18:40:00	332.0	6.3595	82.322835	34.0	0.091667	0.0	0.0
1	2018-06-13T18:45:00	326.0	7.7280	83.740157	0.0	0.091667	0.0	0.0
2	2018-06-13T18:50:00	330.0	4.7495	80.525180	0.0	0.091667	0.0	0.0
3	2018-06-13T18:55:00	324.0	6.3595	89.129032	20.0	0.091667	0.0	0.0
4	2018-06-13T19:00:00	306.0	5.1520	92.495652	0.0	0.075000	0.0	0.0



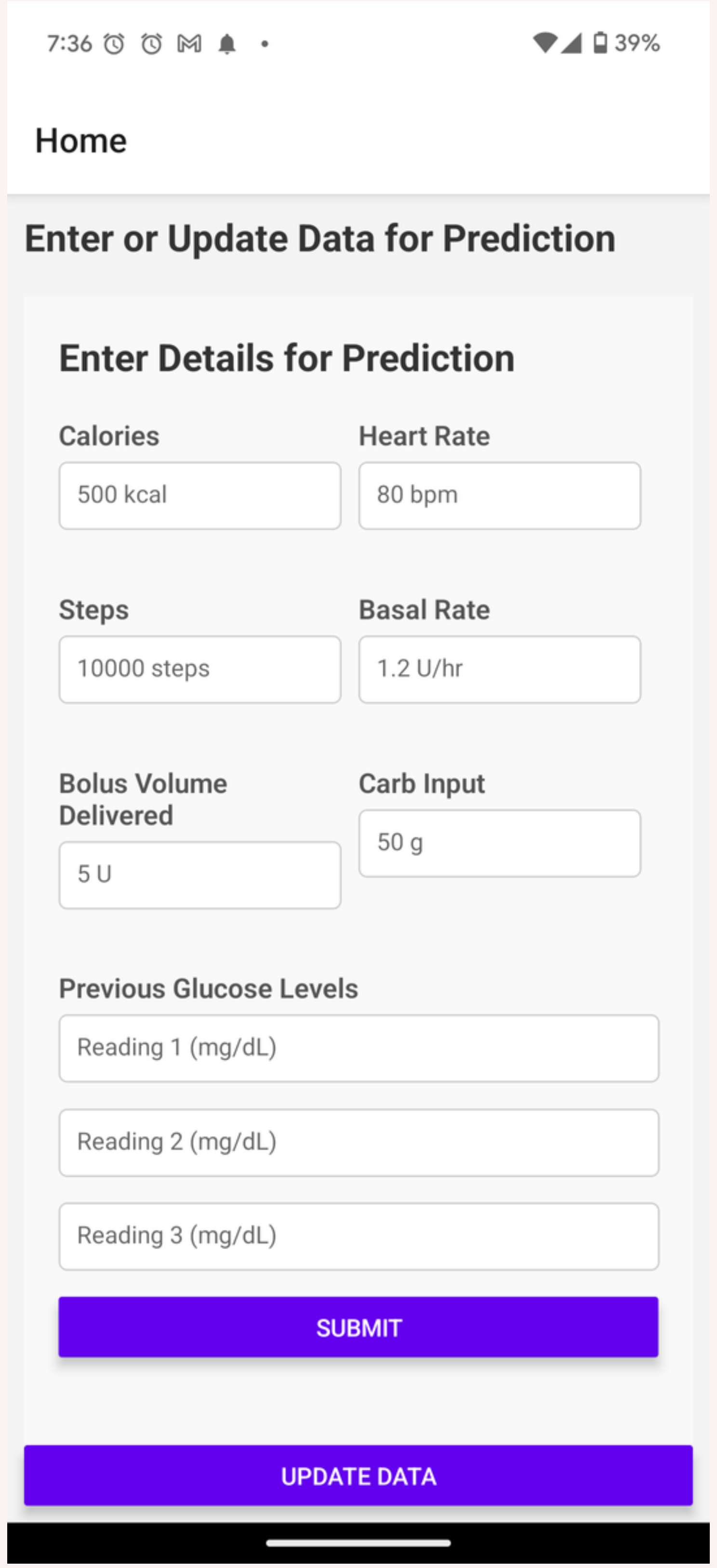

```
Epoch 1/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 296s 35ms/step - event_output_accuracy: 0.8537 - event_output_loss: 0.3325 - loss: 0.4706 - time_output_loss: 0.1
Epoch 2/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 320s 35ms/step - event_output_accuracy: 0.9797 - event_output_loss: 0.0554 - loss: 0.1139 - time_output_loss: 0.0
Epoch 3/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 299s 36ms/step - event_output_accuracy: 0.9811 - event_output_loss: 0.0503 - loss: 0.1030 - time_output_loss: 0.0
Epoch 4/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 319s 35ms/step - event_output_accuracy: 0.9819 - event_output_loss: 0.0461 - loss: 0.0956 - time_output_loss: 0.0
Epoch 5/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 321s 35ms/step - event_output_accuracy: 0.9830 - event_output_loss: 0.0425 - loss: 0.0893 - time_output_loss: 0.0
Epoch 6/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 319s 35ms/step - event_output_accuracy: 0.9833 - event_output_loss: 0.0410 - loss: 0.0857 - time_output_loss: 0.0
Epoch 7/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 321s 35ms/step - event_output_accuracy: 0.9840 - event_output_loss: 0.0382 - loss: 0.0814 - time_output_loss: 0.0
Epoch 8/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 320s 34ms/step - event_output_accuracy: 0.9847 - event_output_loss: 0.0359 - loss: 0.0766 - time_output_loss: 0.0
Epoch 9/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 323s 35ms/step - event_output_accuracy: 0.9854 - event_output_loss: 0.0343 - loss: 0.0733 - time_output_loss: 0.0
Epoch 10/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 326s 35ms/step - event_output_accuracy: 0.9871 - event_output_loss: 0.0307 - loss: 0.0686 - time_output_loss: 0.0
Epoch 11/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 316s 34ms/step - event_output_accuracy: 0.9882 - event_output_loss: 0.0277 - loss: 0.0642 - time_output_loss: 0.0
Epoch 12/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 324s 35ms/step - event_output_accuracy: 0.9892 - event_output_loss: 0.0256 - loss: 0.0614 - time_output_loss: 0.0
Epoch 13/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 321s 34ms/step - event_output_accuracy: 0.9894 - event_output_loss: 0.0252 - loss: 0.0605 - time_output_loss: 0.0
Epoch 14/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 323s 35ms/step - event_output_accuracy: 0.9907 - event_output_loss: 0.0227 - loss: 0.0571 - time_output_loss: 0.0
Epoch 15/50
8390/8390 ━━━━━━━━━━━━━━━━━━━━ 320s 34ms/step - event_output_accuracy: 0.9907 - event_output_loss: 0.0217 - loss: 0.0557 - time_output_loss: 0.0
```

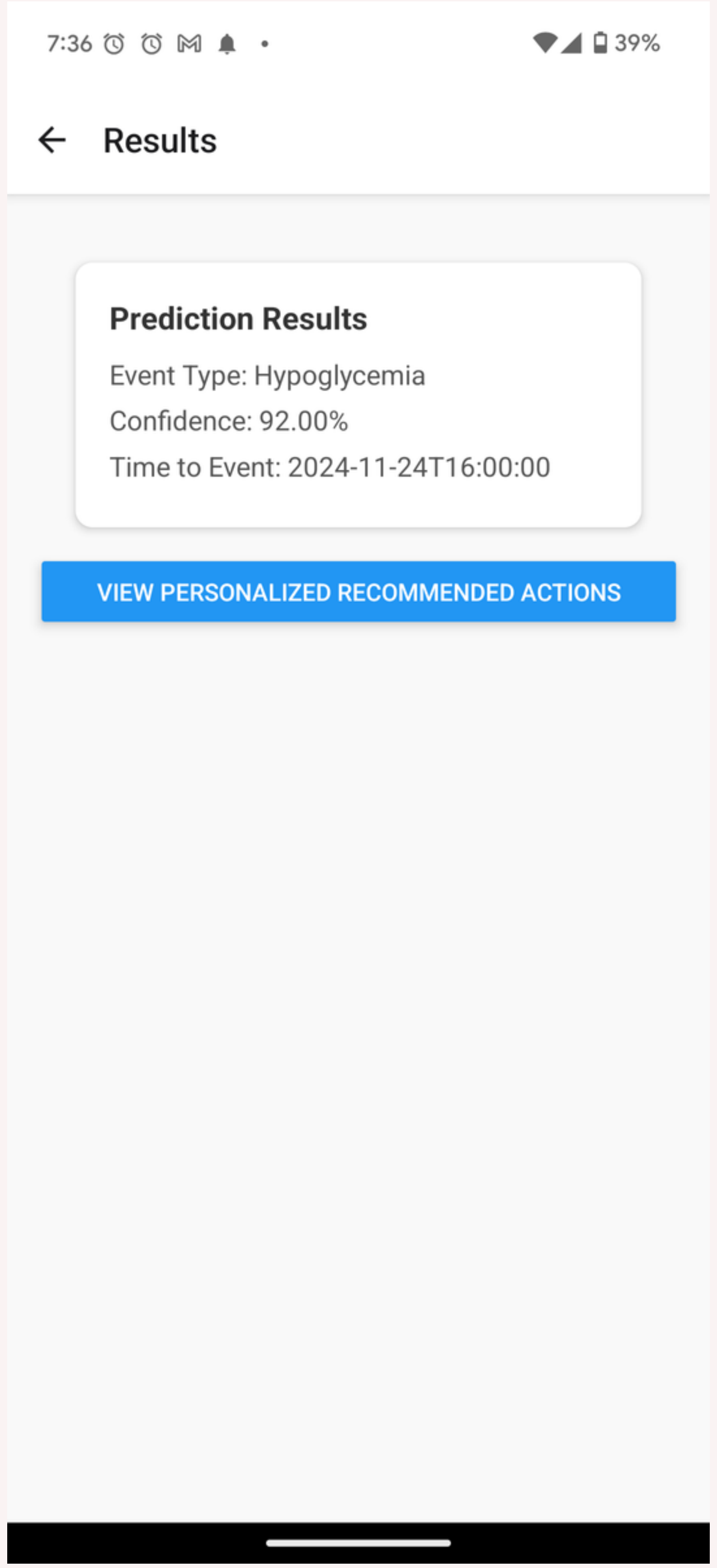
Proof Of Completion

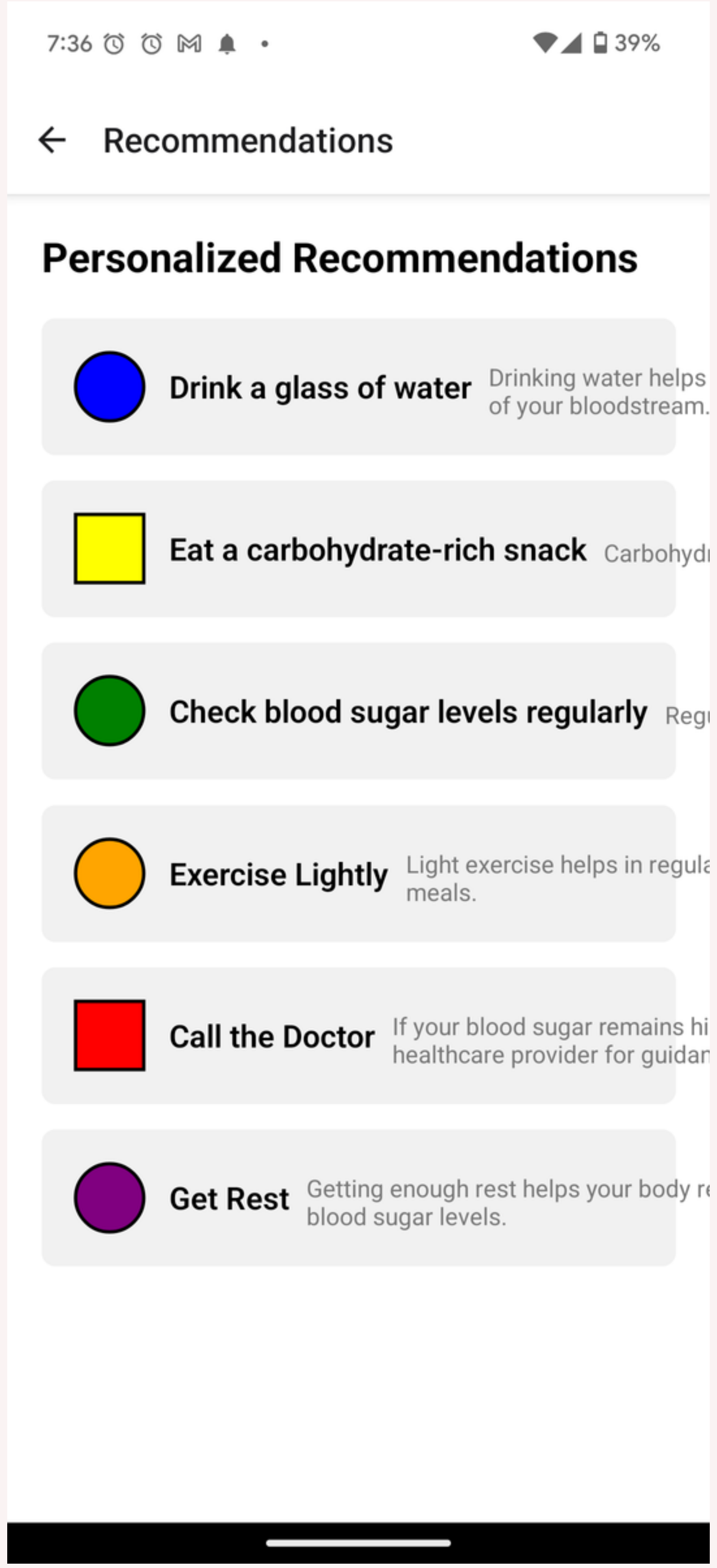
```
Predicted Event Type (Hypoglycemia/Hyperglycemia/Normal):
Glucose: 117.66666666666669 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 125.0 - Normal Glycemic Event, Time to Event: 0.87 minutes
Glucose: 125.0 - Normal Glycemic Event, Time to Event: 0.87 minutes
Glucose: 125.0 - Normal Glycemic Event, Time to Event: 0.87 minutes
Glucose: 125.0 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 136.66666666666666 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 148.33333333333334 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 160.0 - Normal Glycemic Event, Time to Event: 0.86 minutes
Glucose: 160.0 - Normal Glycemic Event, Time to Event: 0.87 minutes
Glucose: 160.0 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 160.0 - Normal Glycemic Event, Time to Event: 0.89 minutes
Glucose: 168.66666666666666 - Normal Glycemic Event, Time to Event: 0.88 minutes
Glucose: 177.33333333333334 - Normal Glycemic Event, Time to Event: 0.87 minutes
Glucose: 186.0 - Hyperglycemia, Time to Event: 0.88 minutes
Glucose: 191.33333333333331 - Hyperglycemia, Time to Event: 0.90 minutes
Glucose: 196.66666666666663 - Hyperglycemia, Time to Event: 0.89 minutes
Glucose: 202.0 - Hyperglycemia, Time to Event: 0.86 minutes
Glucose: 201.0 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 200.0 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 199.0 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 195.33333333333331 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 191.66666666666663 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 188.0 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 183.33333333333331 - Hyperglycemia, Time to Event: 0.84 minutes
Glucose: 178.66666666666666 - Normal Glycemic Event, Time to Event: 0.84 minutes
Glucose: 174.0 - Normal Glycemic Event, Time to Event: 0.84 minutes
```

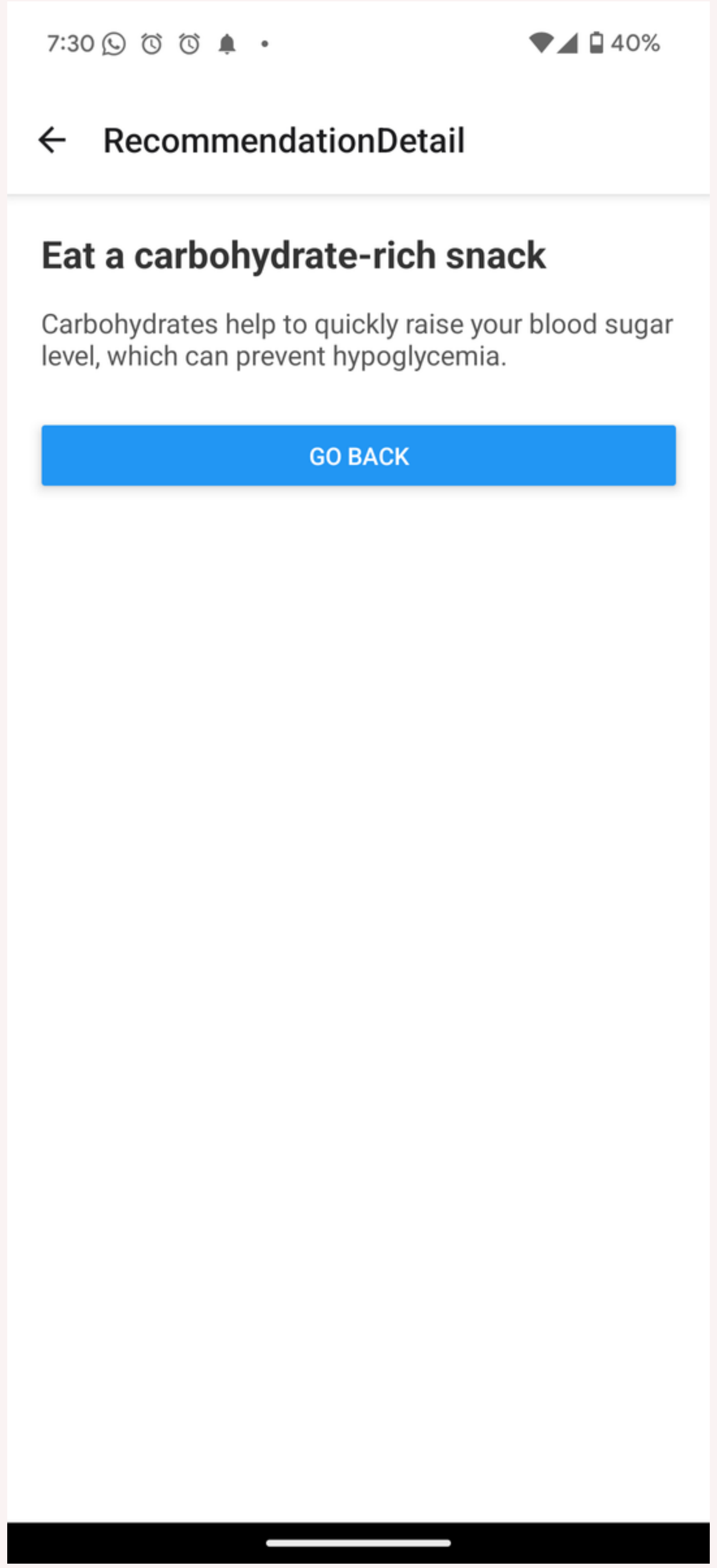


Proof Of Completion









References

- M. Gadaleta, A. Facchinetti, E. Grisan, and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," IEEE Journal of Biomedical and Health Informatics, vol. PP, no. 99, pp. 1-1, April 2018, doi: 10.1109/JBHI.2018.2823763. [1]
- N. Nuwarapaksha, H.M.R.K.G. Nandasena, and N. Nanayakkara, "Prevalence of Long-term Complications among Type 2 Diabetes Mellitus Patients at National Hospital - Kandy, Sri Lanka," Diabetes Research and Clinical Practice, vol. 186S, p. 109535, January 2022. [2]
- Deldjoo, Y., He, Z., & McAuley, J., "A Review of Modern Recommender Systems Using Generative Models (Gen-RecSys)," Polytechnic University of Bari, Bari, Italy; University of California, La Jolla, USA.[3]
- Lin, Y., Liu, Y., Lin, F., Miao, C., "A Survey on Reinforcement Learning for Recommender Systems," September 2021.[4]
- Syafrudin, M., Alfian, G., Fitriyani, N. L., Anshari, M., "Future Glycemic Events Prediction Model Based On Artificial Neural Network," 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), November 2022, [5]
- Crilly, P., Uddin, N., "A mixed method study investigating the use of the Sugar Smart app to help community pharmacists reduce public sugar intake," Royal Pharmaceutical Society (RPS) Winter Summit 2017, London, December 2017.[6]
- Zhu, Y., Yi, J., Xie, J., Chen, Z., "Deep Causal Reasoning for Recommendations," arXiv preprint arXiv:2201.01593, 6 Jan 2022 (v1), last revised 21 Nov 2022 (v2).[7]

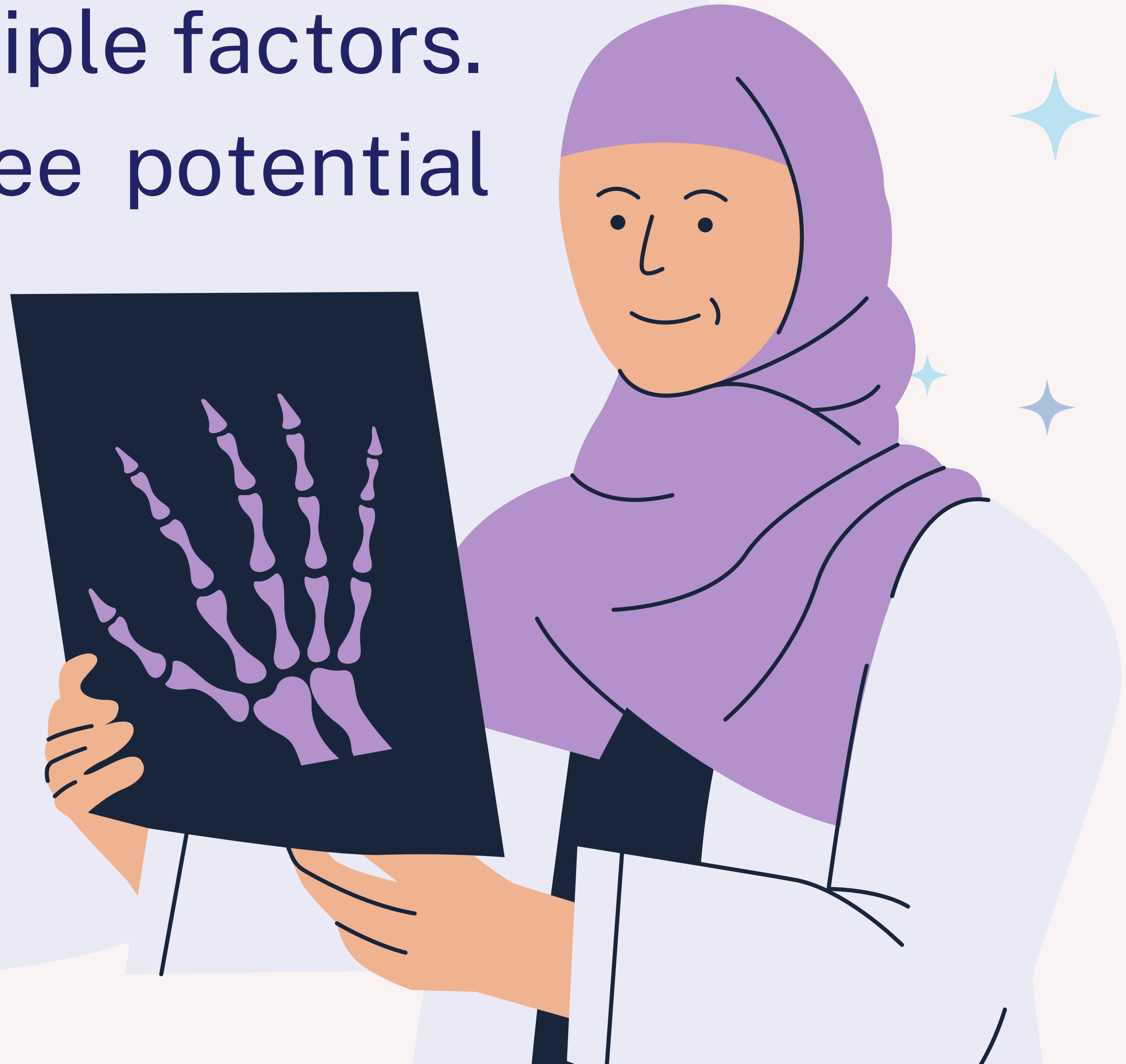
IT21287718| WIMANSA P.P.H.S.D.

SPECIALIZATION - DATA SCIENCE

SIDE EFFECTS PREDICTION AND ALERT MECHANISM FOR DIABETES TREATMENT

INTRODUCTION

- Diabetes treatment often causes unpredictable side effects.
- Side effects vary across individuals due to multiple factors.
- Goal -- Develop a predictive system to foresee potential side effects.



WIMANSA P.P.H.S.D. | IT21287718

RESEARCH PROBLEMS

- **Unpredictability** - Side effects from diabetes treatments vary across patients.

- **Lack of Personalization** - Current systems fail to account for individual daily symptom logs or ongoing trends.

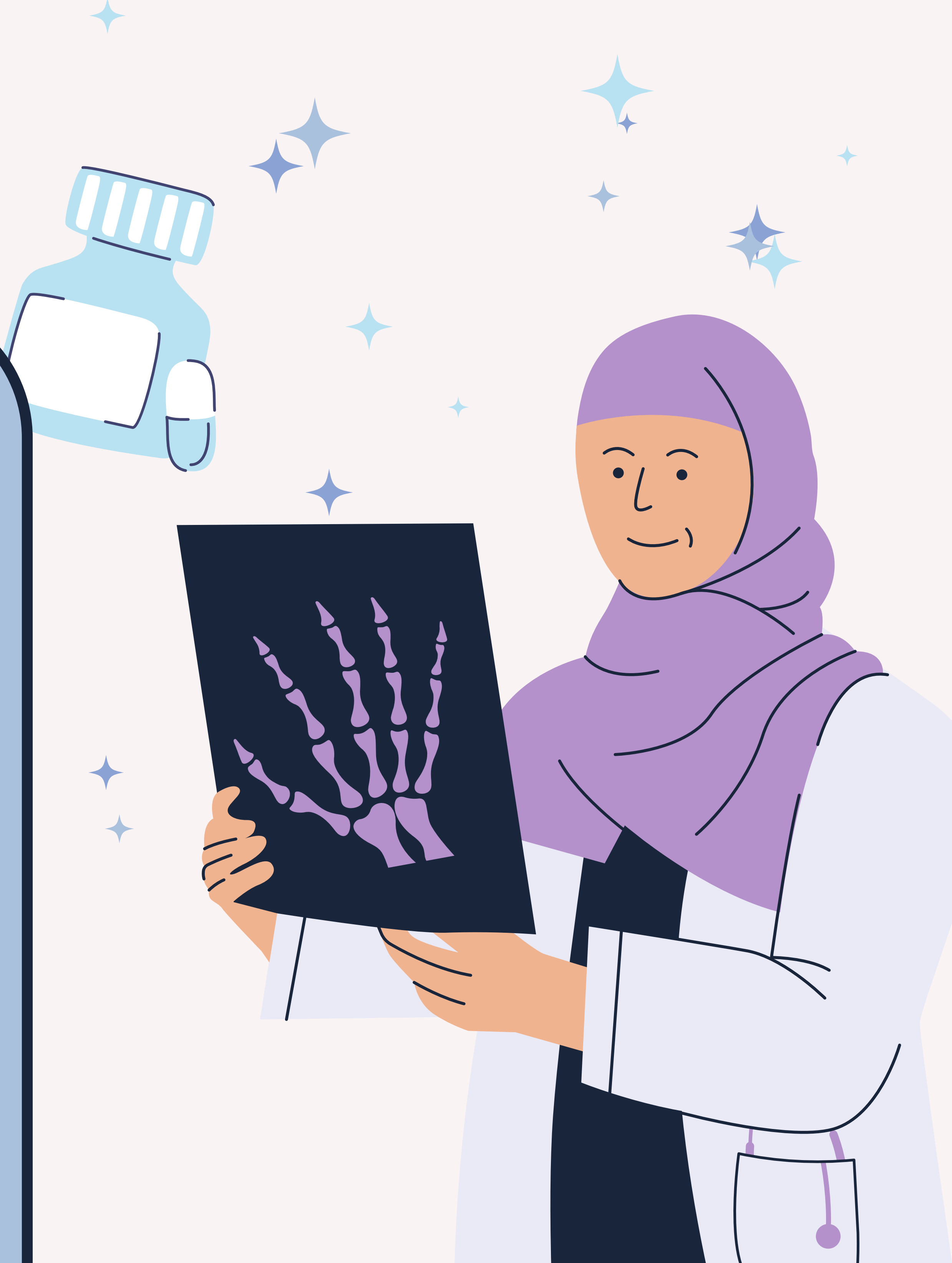
- **Impact** - Poor management of side effects leads to reduced treatment adherence and poorer outcomes.

SOLUTIONS

- **Predictive System -** Develop a machine learning model to forecast side effects.
- **Personalization -** Incorporate daily symptom logs, health metrics, and treatment dosages.
- **Proactive Care -** Alert patients and healthcare providers about potential side effects in advance.

OBJECTIVES

- *Predict potential side effects of diabetic treatments for individuals.*
- *Alert patients and healthcare providers about potential side effects in advance.*



WIMANSA P.P.H.S.D. | IT21287718

DATA SET

Here's what the dataset columns represent:

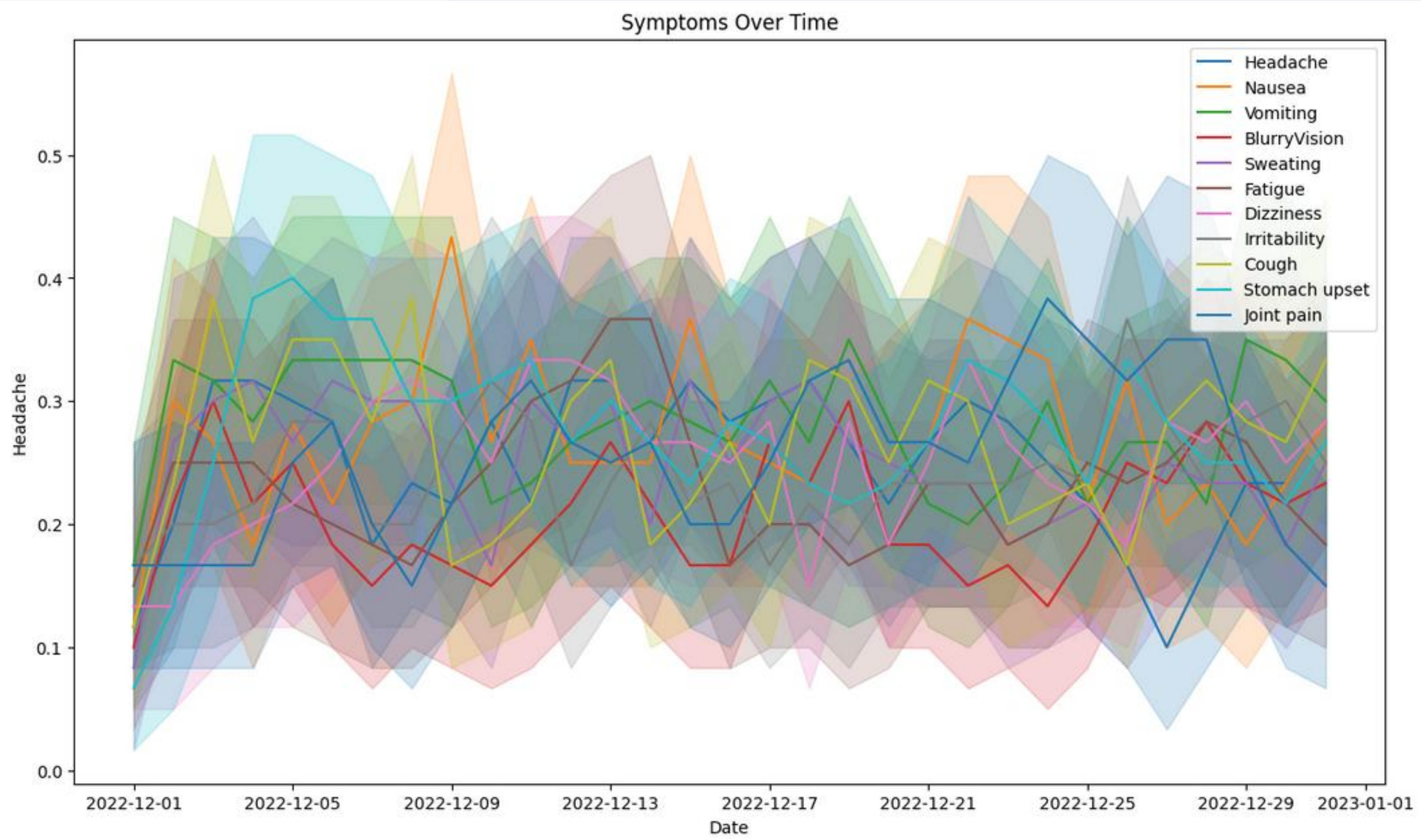
- PatientID: Unique identifier for each patient.
- Date: The date of the record.
- Glucose, BMI, Insulin: Quantitative health metrics.
- Symptoms (e.g., Headache, Nausea, Vomiting, etc.): Binary (0/1) indicators for whether a symptom was experienced.
- SleepHours: Hours of sleep recorded.
- DietQuality: Categorical column indicating the quality of diet (e.g., "Good," "Average").

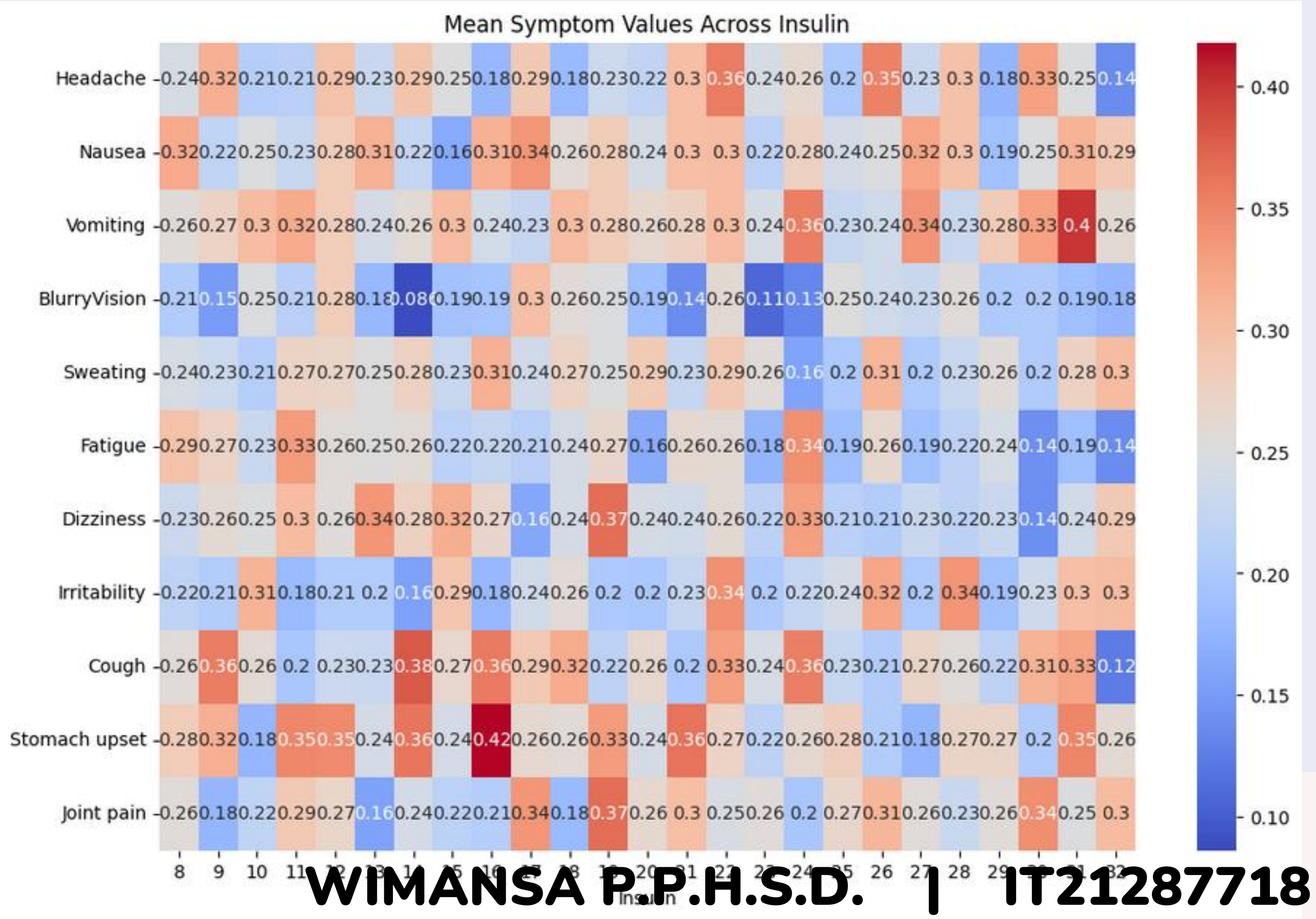
```
data.head()
```

	PatientID	Date	Glucose	BMI	Insulin	Headache	Nausea	Vomiting	BlurryVision	Sweating	Fatigue	Dizziness	Irritability	Cough	Stomach upset	Joint pain	SleepHours	DietQuality
0	1	2022-12-01	118	34	29	0	0	0	0	0	0	0	1	1	1	0	7	Average
1	1	2022-12-02	118	34	32	0	0	0	0	0	0	0	1	1	0	0	8	Good
2	1	2022-12-03	121	34	24	0	0	1	0	0	0	0	0	1	0	0	6	Good
3	1	2022-12-04	80	34	18	0	0	0	0	0	0	0	0	0	1	0	9	Good
4	1	2022-12-05	143	34	15	0	0	0	0	0	0	0	1	1	1	0	8	Average

COMPLETED TASKS

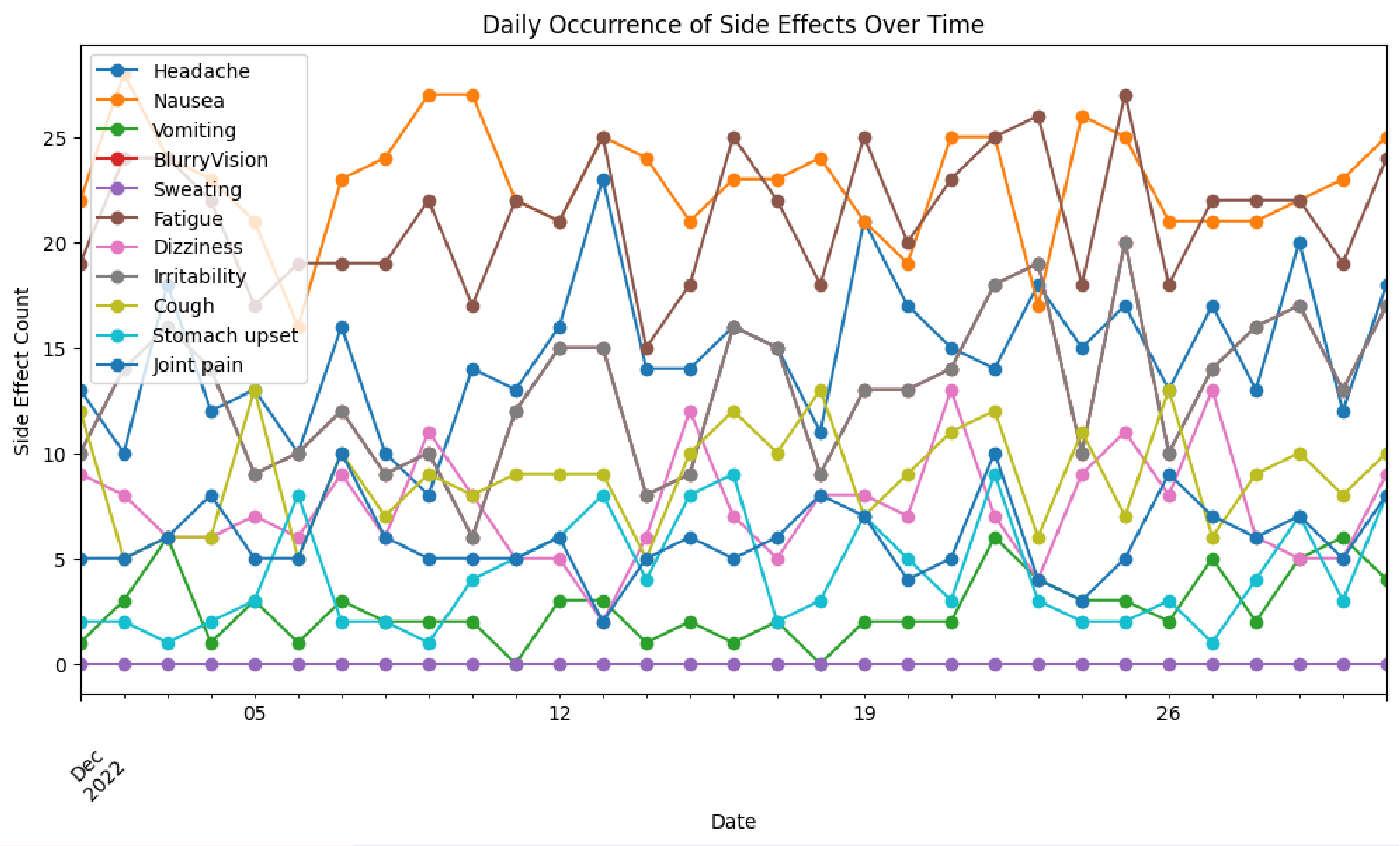
- Cleaned and preprocessed patient data.
- Conducted exploratory data analysis to uncover trends and patterns.
- Developed a machine learning model to predict side effects.

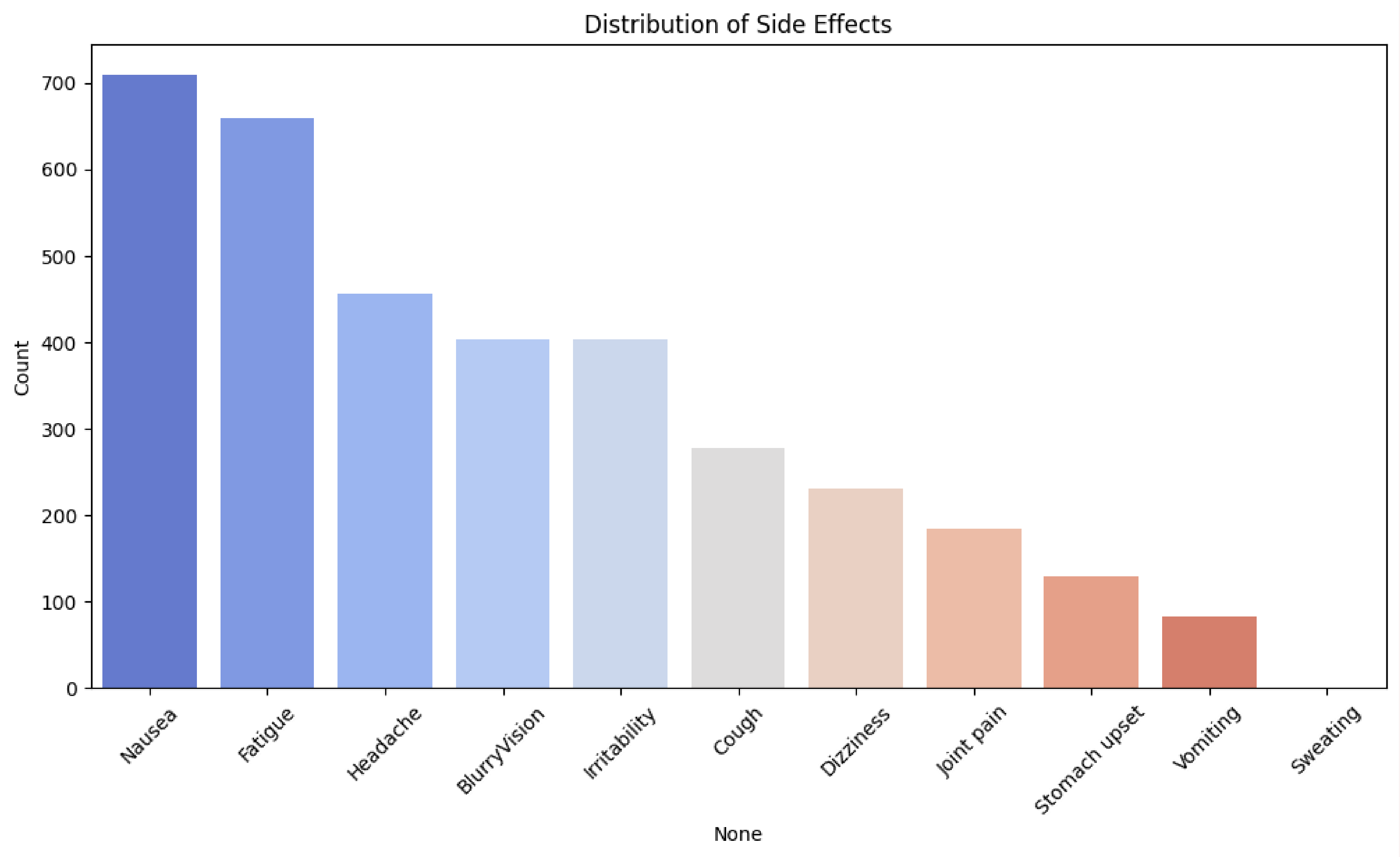




WIMANSA P.P.H.S.D. | IT21287718

DATA ANALYSIS





TRAINING MODELS

Model Used - LSTM (Long Short-Term Memory).

Why LSTM?

- **Captures sequential trends in daily symptom logs.**
- **Handles time-dependent relationships effectively.**

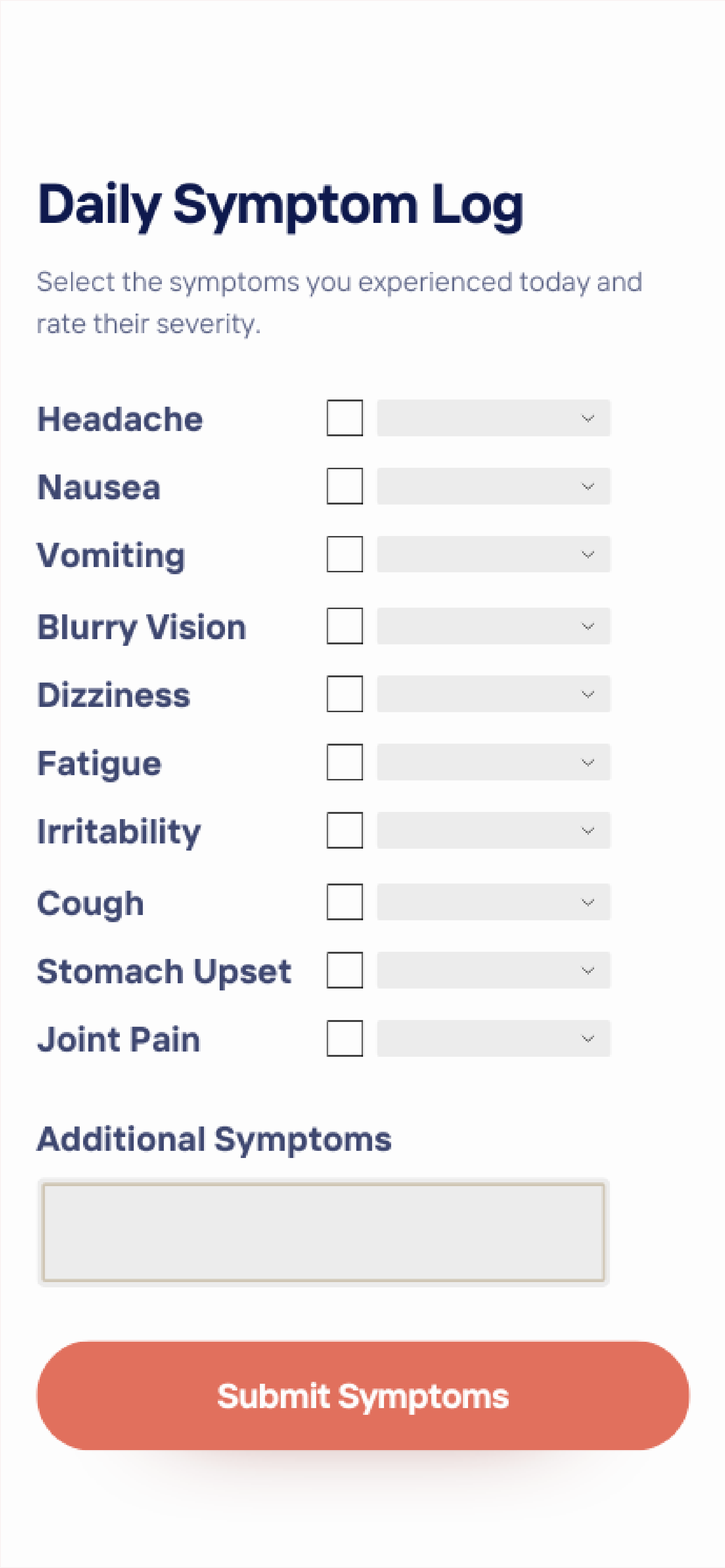
```
# Train the model
from tensorflow.keras.layers import Reshape

model = Sequential([
    LSTM(64, input_shape=(X_train.shape[1], X_train.shape[2]), return_sequences=True),
    Dropout(0.2),
    LSTM(32, return_sequences=False),
    Dropout(0.2),
    Dense(y_train.shape[2], activation='sigmoid'),
    Reshape((1, y_train.shape[2]))
])

model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)

history = model.fit(
    X_train, y_train,
    validation_split=0.2,
    epochs=50,
    batch_size=32,
    callbacks=[early_stopping],
    verbose=1
)
```

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 23ms/step
Accuracy for Headache: 75.00%
Accuracy for Nausea: 66.67%
Accuracy for Vomiting: 75.00%
Accuracy for BlurryVision: 91.67%
Accuracy for Sweating: 66.67%
Accuracy for Fatigue: 66.67%
Accuracy for Dizziness: 50.00%
Accuracy for Irritability: 75.00%
Accuracy for Cough: 58.33%
Accuracy for Stomach upset: 58.33%
Accuracy for Joint pain: 75.00%
```

FUTURE WORKS

- Collect more diverse data for generalizability.
- Refine models with hyperparameter tuning.
- Develop a mobile app for real-time symptom logging and feedback.
- Integrate alert systems for severe side effects.

FUTURE WORKS



WIMANSA P.P.H.S.D. | IT21287718

REFERENCES

M. S. Abdul Rahman, S. M. Rahimifard, and M. M. Al-Shami, "Prediction of DiabetesRelated Complications Using Machine Learning Algorithms: A Comprehensive Review," in 2021 IEEE International Conference on Artificial Intelligence and Machine Learning (AIML), 2021, pp. 1-6.

N. Jothi and A. A. N. Islam, "A Framework for Predicting Adverse Drug Reactions in Diabetic Patients Using Machine Learning Techniques," in 2021 IEEE International Conference on Biomedical Engineering and Sciences (IECBES), 2021, pp. 212-217

R. Al-Mahdi, M. Elshamli, and A. S. Mokhtar, "A Machine Learning Approach for RealTime Prediction of Side Effects in Diabetes Treatment," in 2022 IEEE International Conference on Healthcare Informatics (ICHI), 2022, pp. 45-50.

J. Wu, J. Zhang, and Z. Li, "Mobile Health Application for Diabetes Management: RealTime Prediction of Blood Glucose Levels and Side Effects," in 2021 IEEE International Conference on Health Informatics (ICHI), 2021, pp. 74-80.

IT21283062| DE SILVA L.K.N.

SPECIALIZATION - DATA SCIENCE

NUTRITIONAL GUIDANCE SYSTEM WITH PERSONALIZED MEAL PLANS

INTRODUCTION

Managing diabetes requires a careful balance of nutrition and lifestyle adjustments. This research proposes a machine learning-based personalized meal recommendation system designed for diabetic patients. The system uses patient profiles, including health metrics, dietary preferences, and allergies, to generate meal plans tailored to their diabetes type and other variables.

RESEARCH PROBLEMS

Lack of real-time data utilization for dynamically adjusting meal plans

Inaccurate or non-personalized predictive models for glucose response to meals.

Generic recommendations with insufficient deep personalization for individual users.

OBJECTIVES

To create a personalized nutritional guidance system
that uses dietary, medical,
and activity data
to provide optimized meal plans for better diabetes management.

DATA COLLECTION

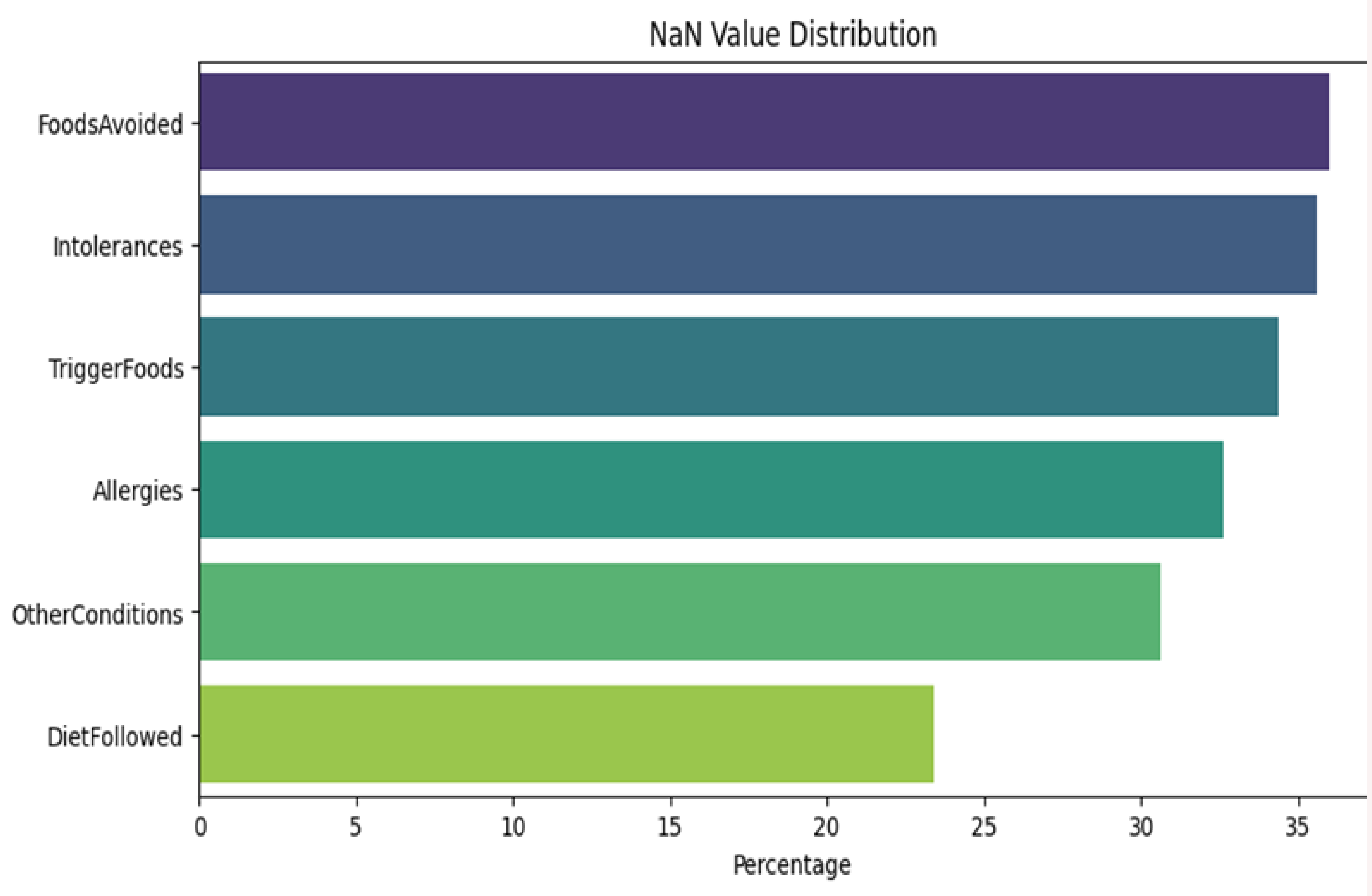
```python
datapath = 'diabetes_user_profiles_with_mealID.csv'
df = pd.read_csv(datapath)
df
```

Python

	RecordID	Name	Age	Gender	Height	Weight	Location	Occupation	DiabetesType	DiagnosedYearsAgo	...	DietFollowed	TriggerFoods	Allergies	Int
0	1	Nimal Fernando	64	Male	168	80	Jaffna	Homemaker	Gestational	2	...	Vegetarian	NaN	Dairy	
1	2	Kumari Rathnayake	55	Female	178	96	Kandy	Office worker	Type 1	14	...	Low-carb	Sugary snacks	Nuts	
2	3	Anjali Perera	57	Female	180	55	Jaffna	Homemaker	Type 1	10	...	NaN	Sugary snacks	Dairy	
3	4	Amara Wijesinghe	48	Female	156	56	Colombo	Office worker	Type 1	4	...	Vegetarian	Sugary snacks	Dairy	
4	5	Kumari Rathnayake	70	Female	175	50	Colombo	Homemaker	Type 2	5	...	Vegetarian	White bread	NaN	

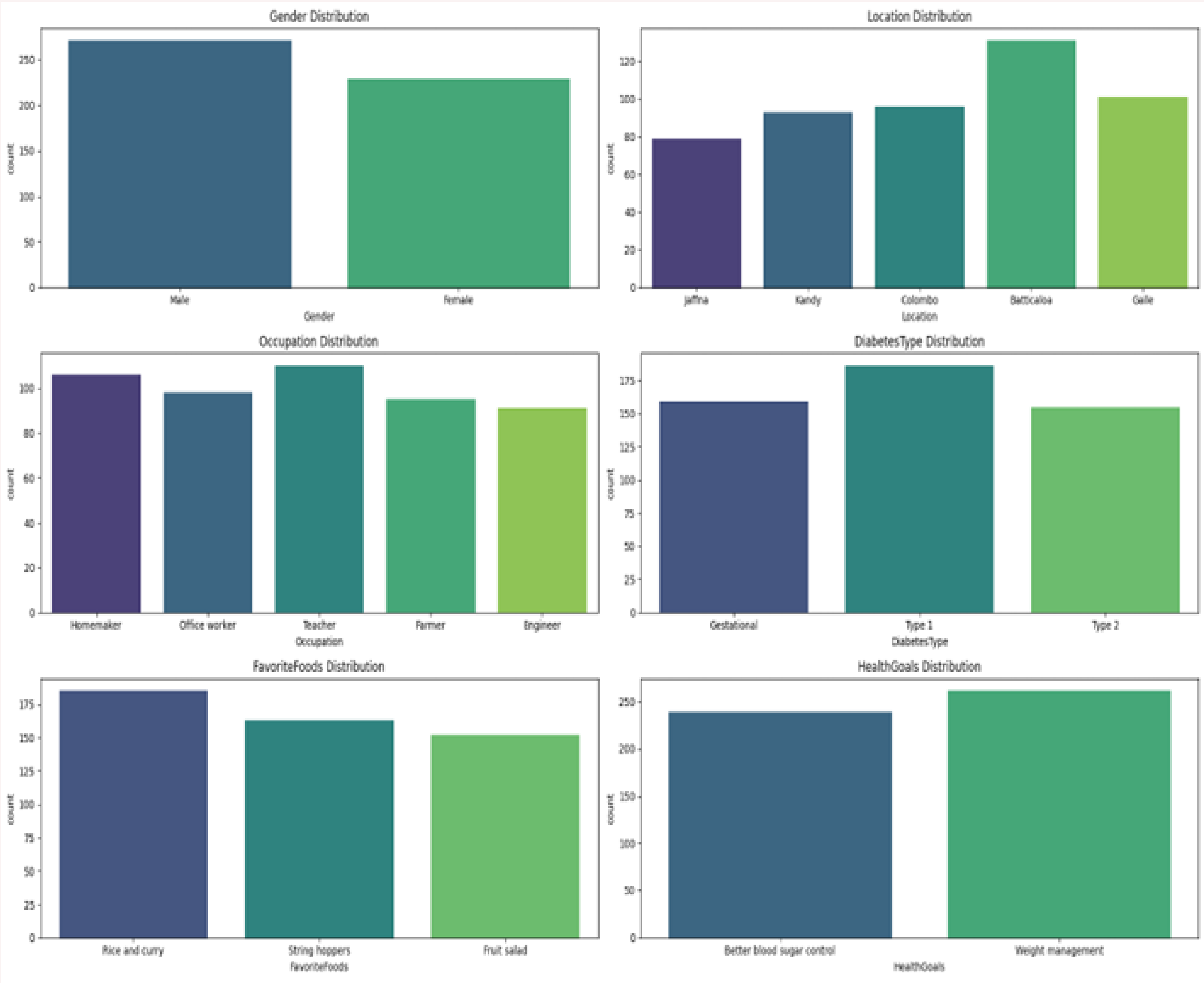
MealID	MealName	MealDetails	CalorieCount	AllergyStatus	Preferences	Type
1	Traditional Sri Lankan Breakfast	String hoppers, pol sambol, dhal curry	350	Contains coconut	Vegetarian	Breakfast
1	Rice and Curry Lunch	White rice, fish curry, mallung, papadam	600	Contains fish, gluten	Non-Vegetarian	Lunch
1	Light Dinner	Vegetable soup, brown bread	250	Contains gluten	Vegetarian	Dinner
2	Coconut Milk Rice Breakfast	Kiribath, lunu miris	300	Contains coconut	Vegetarian	Breakfast
2	Seafood Curry Lunch	White rice, prawn curry, gotukola salad	650	Contains shellfish	Non-Vegetarian	Lunch
2	String Hopper Dinner	String hoppers, chicken curry	450	Contains gluten	Non-Vegetarian	Dinner
3	Healthy Breakfast	Oats porridge, banana, nuts	320	Contains nuts	Vegetarian	Breakfast
3	Chicken Fried Rice Lunch	Brown rice, chicken, vegetables	700	Contains egg, soy	Non-Vegetarian	Lunch
3	Vegetable Roti Dinner	Whole wheat roti, mixed vegetables	400	Contains gluten	Vegetarian	Dinner
4	Spicy Breakfast	Hoppers, lunu miris, dhal curry	300	Contains coconut	Vegetarian	Breakfast
4	Fish Curry Lunch	Red rice, fish curry, beetroot salad	550	Contains fish	Non-Vegetarian	Lunch

DATA ANALYSIS

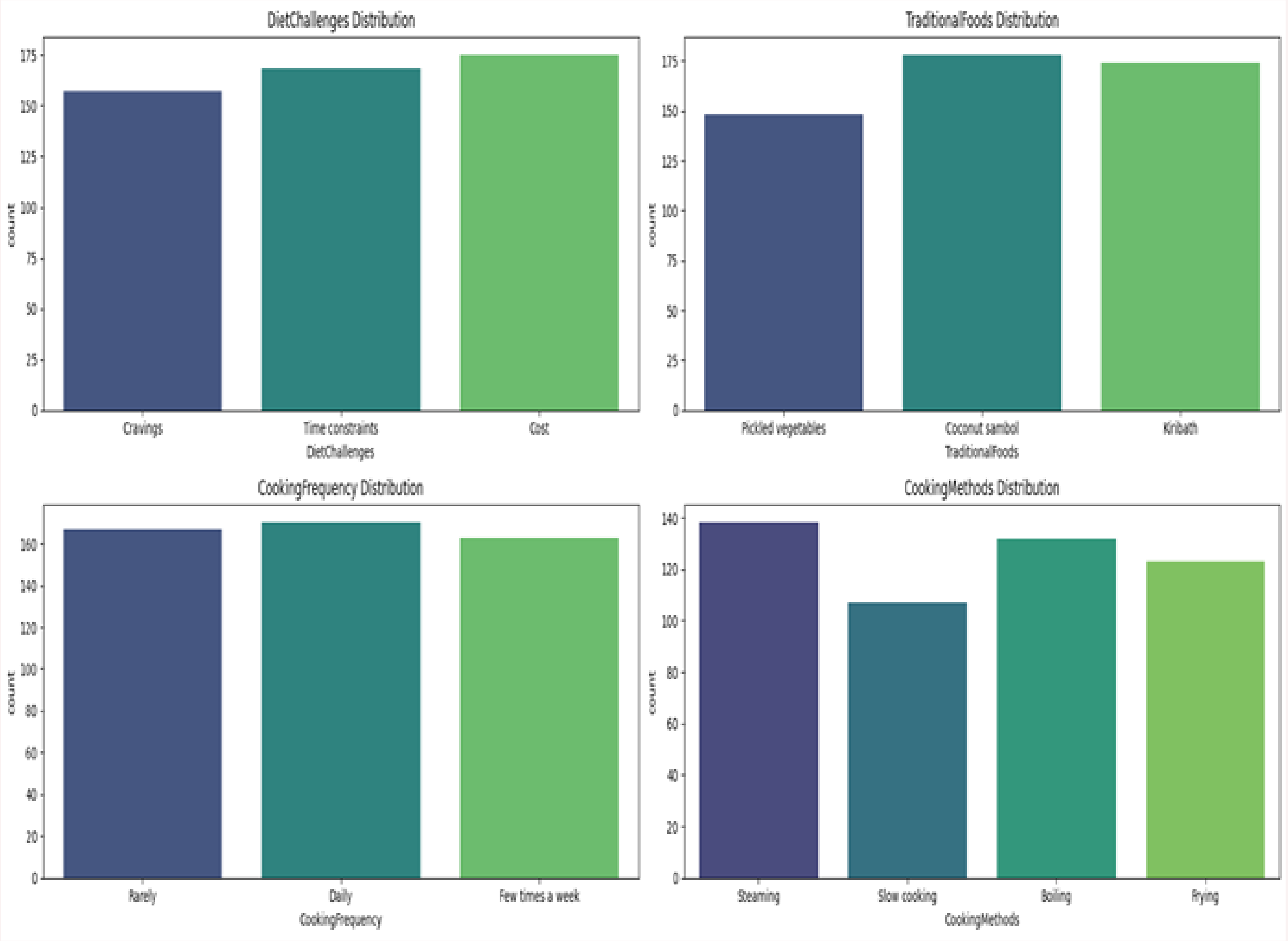



```python
### NaN Value Distribution - Identifies Missing Data
nan_dist = df.isna().sum()
nan_dist = nan_dist[nan_dist > 0]
nan_dist = nan_dist.sort_values(ascending=False)
nan_dist = nan_dist / df.shape[0] * 100
nan_dist = nan_dist.round(2)

# plot
plt.figure(figsize=(10, 5))
sns.barplot(x=nan_dist.values, y=nan_dist.index, palette='viridis')
plt.title('NaN Value Distribution')
plt.xlabel('Percentage')
plt.ylabel('Columns')
plt.show()
```

DATA ANALYSIS

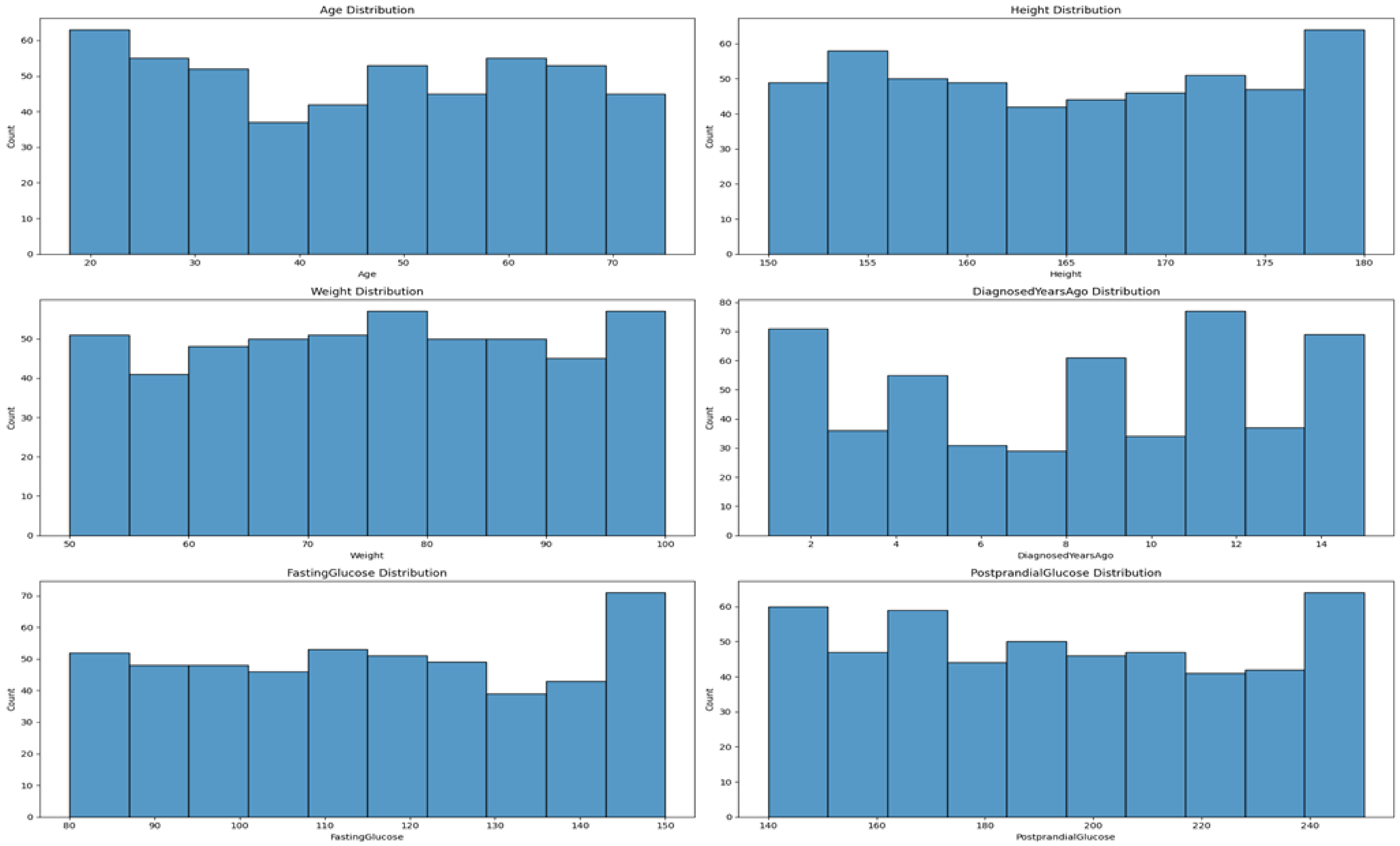


visualize the distribution of each ***categorical column*** *in the Dataset using count plots.*



DE SILVA L.K.N. | IT21283062

DATA ANALYSIS



Visualize the distribution of each ***numerical column*** *in the Dataset using* ***histograms.***

DE SILVA L.K.N. | IT21283062

TRAINING MODELS

CatBoostClassifier model

- Ensemble Learning - boosting framework
- Classification Tasks
- Categorical Data

```
cat = CatBoostClassifier(
                        iterations=200,
                        learning_rate=0.1,
                        loss_function='MultiClass',
                        depth=6
                        )
cat.fit(
        X, Y,
        eval_set=(
                  X_test,
                  Y_test
                  ),
        verbose=100
        )
```

[54] ✓ 0.5s

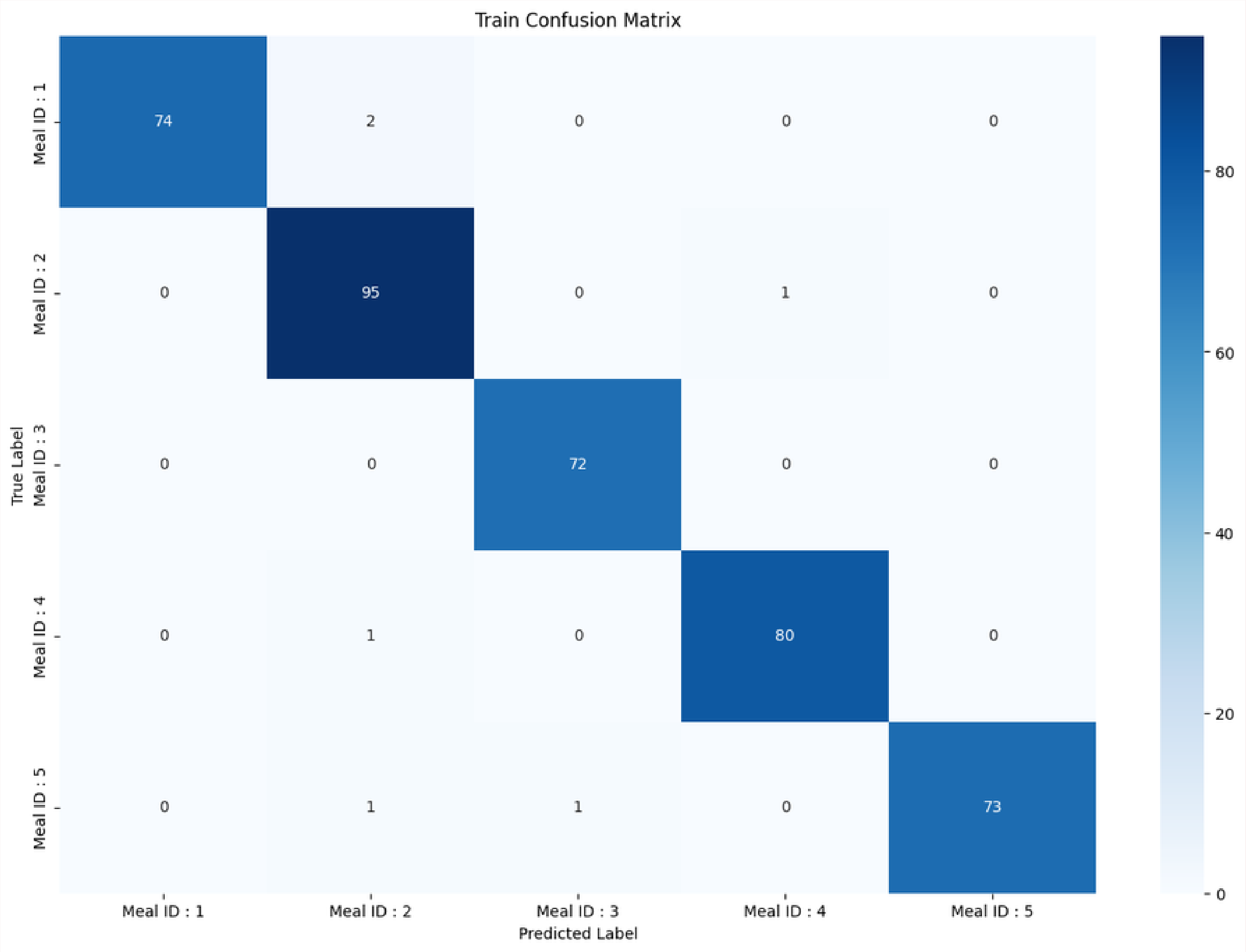
```
0:      learn: 1.5930468        test: 1.5937598 best: 1.5937598 (0)     total: 5.02ms   remaining: 1000ms
100:    learn: 0.9514731        test: 0.9540896 best: 0.9540896 (100)   total: 197ms    remaining: 193ms
199:    learn: 0.6058105        test: 0.6028302 best: 0.6028302 (199)   total: 416ms    remaining: 0us

bestTest = 0.6028302123
bestIteration = 199

<catboost.core.CatBoostClassifier at 0x18f7604a920>
```

DE SILVA L.K.N. | IT21283062

EVALUATING THE MODEL

```
..   -------------------- Train CLS REPORT ----------------------
                precision    recall  f1-score   support

 Meal ID : 1       1.00      0.97      0.99        76
 Meal ID : 2       0.96      0.99      0.97        96
 Meal ID : 3       0.99      1.00      0.99        72
 Meal ID : 4       0.99      0.99      0.99        81
 Meal ID : 5       1.00      0.97      0.99        75

    accuracy                           0.98       400
   macro avg       0.99      0.98      0.99       400
weighted avg       0.99      0.98      0.99       400

-------------------- Test CLS REPORT ---------------------
                precision    recall  f1-score   support

 Meal ID : 1       1.00      1.00      1.00        22
 Meal ID : 2       0.95      1.00      0.98        20
 Meal ID : 3       1.00      1.00      1.00        26
 Meal ID : 4       1.00      1.00      1.00        16
 Meal ID : 5       1.00      0.94      0.97        16

    accuracy                           0.99       100
   macro avg       0.99      0.99      0.99       100
weighted avg       0.99      0.99      0.99       100
```

- classification report
- confusion matrix



DE SILVA L.K.N. | IT21283062

COMPLETED TASKS

- Data Cleaning and Preprocessing
- Data Analysis
- Created a model to generate and recommend meal plan

FUTURE WORKS

- Collect Realtime data for generalizability.
- Improve the model accuracy for better performance.
- Develop a mobile app for real-time symptom logging and feedback.
- develop feedback system for future improvement.

DE SILVA L.K.N. | IT21283062

REFERENCES

- "Mobile health applications for managing diabetes: A systematic review" by El-Gayar, O., Timsina, P., Nawar, N., & Eid, W. (2013). Journal of Diabetes Science and Technology.
- "Data preprocessing techniques for classification without noise in real-world" by García, S., Luengo, J., & Herrera, F. (2015). Knowledge and Information Systems.
- "Collaborative filtering for dietary recommendations" by Harvey, M., Ludwig, B., & Elsweiler, D. (2013). Proceedings of the 7th ACM conference on Recommender systems
- "Adaptive learning algorithms for personalized diet recommendations" by Ueta, T., Koizumi, Y., & Shirota, S. (2019). IEEE Access
- "Clinical trials and diabetes treatment: A guide to conducting clinical research" by Gale, E. A. M., & Collier, A. (2006). Clinical Medicine & Research