SPECTRUM OF ENGINEERING SCIENCES
ISSN (E) 3007-3138 (P) 3007-312X VOLUME 4, ISSUE 6, 2026

GLUCOTWIN: AN ARTIFICIAL INTELLIGENCE–DRIVEN INSULIN DOSAGE PREDICTION SYSTEM

Sanam Nayab[1], Sadia Zahra[2], Hina Mahjabeen[3], Babar Majeed[4], Muhammad Sajid Maqbool[5]

[1]Department of Software Engineering, National University of Modern Languages (NUML), Islamabad, Pakistan
[2,3]Department of Computer Science, National University of Modern Languages (NUML), Islamabad, Pakistan
[4]Department of Electrical Engineering, University of Lahore, Islamabad Campus
[5]Department of Computer Science, NFC Institute of Engineering and Technology, Multan

[1]sanam.nayab@numl.edu.pk, [2]sadia.zahra@numl.edu.pk, [3]hina.mahjabeen@numl.edu.pk,
[4]Babarmajeed120@gmail.com, [*5]sajid.maqbool@nfciet.edu.pk

DOI: https://doi.org/10.5281/zenodo.21125455

Keywords
GLUCOTWIN, Diabetes, Machine Learning, Insulin Prediction, Mobile App, Healthcare, Flutter, Firebase, Child Monitoring

Article History
Received: 25 April 2026
Accepted: 04 June 2026
Published: 21 June 2026

Copyright @Author
Corresponding Author: *
Muhammad Sajid Maqbool
Sanam Nayab

Abstract
Diabetes management in children requires continuous monitoring and accurate insulin dosage adjustments based on multiple physiological and dietary factors. Incorrect insulin administration can lead to serious health complications, particularly when caregivers lack the necessary medical expertise to make informed decisions. This study presents GLUCOTWIN, an AI-driven healthcare application that assists with insulin dosage prediction for children aged 3–12 years by integrating machine learning into a user-friendly mobile platform. The proposed system utilizes health-related parameters, including blood glucose level, blood pressure, body temperature, Body Mass Index (BMI), meal timing, and carbohydrate intake, to predict personalized insulin dosage recommendations. A role-based architecture was developed comprising Admin, Guardian, and Doctor modules, enabling secure data management, medical supervision, and real-time decision support. The mobile application was developed using Flutter, while Firebase was employed as the backend database for storing patient records, prediction history, and user information. Multiple machine learning regression algorithms were trained and evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 Score. Among the evaluated models, the Gradient Boosting Regressor achieved the highest predictive performance with an MAE of 0.2921, RMSE of 0.3736, and R^2 score of 0.9570, demonstrating high accuracy and reliability in insulin prediction. The trained model was successfully deployed and integrated into the application to provide real-time insulin dosage recommendations, which can be reviewed and validated by healthcare professionals. The results indicate that GLUCOTWIN has the potential to enhance diabetes management, support caregivers in making informed decisions, and contribute to improved healthcare outcomes for paediatric diabetes patients.

1. Introduction:

GLUCOTWIN was developed because diabetes management in children aged 3 to 12 years is a serious and sensitive healthcare problem that needs careful daily attention. Diabetes can cause high blood sugar and, if it is not managed properly, it can lead to harmful effects on different parts of the body, including the eyes, kidneys, nerves, heart, and blood vessels, which shows why proper care is necessary from an early stage. For children, this challenge becomes even more important because daily care usually depends on guardians, and those guardians may not always have enough medical knowledge to decide the correct insulin amount in changing conditions [1], [2]. In real situations, insulin requirement is not fixed, because it can be affected by several health and lifestyle factors such as blood sugar condition, blood pressure, body temperature, BMI, and carbohydrate intake. This makes manual decision-making difficult, and a wrong insulin dose can place a child at risk. Because of this, there was a strong need for a system that could support guardians with a more accurate, fast, and organized method for handling child diabetes care. GLUCOTWIN was created to answer this need by providing a smart application that collects child health details, sends them to a trained machine learning model, and returns an insulin prediction based on learned patterns from the training dataset. The use of machine learning in diabetes care has been widely studied for prediction, decision support, and better treatment planning, which shows that intelligent systems can play an important role in improving diabetes management [3], [4]. This made the development of GLUCOTWIN highly relevant, because it applies that same idea to a practical child-focused system where prediction is connected with real application use. The application was built in Flutter, which supports the development of mobile applications from a single codebase, and it was linked with Firebase Realtime Database, which allows data to be stored and synchronized in real time across connected clients [5], [6]. This technical structure made it possible to create a role-based healthcare application with three main actors: admin, guardian, and doctor. In this system, doctors register and receive approval before getting access, guardians add children and enter complete details, and all important testing values are processed through the integrated prediction workflow. The system also supports carbohydrate estimation from food details, which is important because food intake directly affects diabetes control and insulin-related decisions [2], [4]. Another major reason this application became necessary is that prediction alone is not always enough in healthcare. For this reason, GLUCOTWIN also includes doctor review, so that predicted insulin values can be checked by a medical professional before final suggestions are trusted for decision support. This combination of intelligent prediction, doctor verification, guardian support, and admin monitoring makes the system safer and more reliable than a basic calculator or simple record-keeping app. GLUCOTWIN is therefore important not only as a software project, but as a healthcare support solution that addresses a real problem faced by families of diabetic children. It helps reduce confusion, supports more informed insulin-related decisions, improves data organization, and creates a structured connection between guardians, doctors, and child health records. For these reasons, the development of GLUCOTWIN was necessary, practical, and meaningful in both healthcare and technology contexts [1], [3], [5], [6].

The problem addressed in this project is the lack of an intelligent and structured system for supporting insulin dosage decisions in diabetic children aged 3 to 12 years. Determining the correct insulin requirement is a complex task because it depends on several changing factors, including blood pressure, body temperature, BMI, blood sugar-related condition, and carbohydrate intake. Guardians often face difficulty in understanding these factors and making correct decisions without medical guidance, which increases the risk of improper insulin administration and possible harm to the child's health. Furthermore, the absence of a unified platform for prediction, medical review, child record management, and administrative monitoring reduces the efficiency and reliability of diabetes care. Therefore, there is a need for a smart application that can process child health data, predict insulin dosage through machine learning, and provide doctor-supported recommendations in a structured environment. GLUCOTWIN was developed to address this need by offering an integrated system that improves decision-making, reduces confusion, and supports safer diabetes management for children.

The main aim of GLUCOTWIN is to develop a smart and easy-to-use healthcare application that helps manage diabetes in children aged 3 to 12 years by predicting the required insulin dose based on important health factors and supporting guardians with doctor-reviewed guidance. The first objective is to create a mobile application that can be used easily by admin, guardians, and doctors through a role-based system. The second objective is to collect important child health information, such as blood pressure, temperature, BMI, and carbohydrate intake, and use this data for insulin prediction. The third objective is to train a machine learning model on a proper dataset and select the model with the best accuracy for prediction. The fourth objective is to deploy the trained model through Python and connect it with the application for real-time prediction. The fifth objective is to integrate Firebase database with the application so that user data, test records, and history can be stored and managed properly. The sixth objective is to allow guardians to add children, enter details, and get insulin prediction in a simple and organized way. The seventh objective is to involve doctors in the system so that predicted results can be reviewed and proper medical suggestions can be given. The eighth objective is to provide admin control for approving doctors, monitoring system activity, and managing overall records. The final objective is to reduce confusion for guardians and improve safe diabetes management for children through a reliable and intelligent digital system. The proposed solution is to develop GLUCOTWIN as a smart healthcare application that helps manage diabetes in children aged 3 to 12 years in a more safe, accurate, and organized way. This system solves the problem by combining a mobile application, machine learning model, real-time database, and doctor support in one platform. In this solution, guardians can create an account, add children, and enter important health details such as blood pressure, temperature, BMI, blood sugar condition, and food-related information. An intelligent agent is used to estimate carbohydrate values from food details, and all these values are then sent to the trained machine learning model. The model, which was trained on a proper dataset and selected based on best accuracy, predicts the required insulin dose for the child. To make the system more reliable, the predicted result is also reviewed by a doctor, who can provide medical suggestions after checking the values. The admin manages the overall system by approving doctor accounts, monitoring test activities, and maintaining records and history. The application is developed in Flutter for a user-friendly interface, while Firebase is used for storing and managing data in real time, and the trained model is deployed through Python on a local machine. In this way, the proposed solution provides a complete and practical system that reduces confusion for guardians, supports doctors in reviewing cases, improves decision-making, and helps provide better diabetes care for children.

2. Literature Review

Diabetes management in children is a very sensitive and difficult healthcare issue because children need regular monitoring, proper food control, correct insulin support, and continuous care from guardians and healthcare professionals. In recent years, researchers have given more attention to digital healthcare systems because these systems can support monitoring, prediction, and decision-making in a more structured way. Current literature shows that machine learning, mobile health applications, and digital support tools are becoming more useful in diabetes care, especially for children and adolescents [7]–[20]. This literature is important for GLUCOTWIN because the project is based on the same idea of using intelligent technology and mobile application support to improve diabetes management for young children. Machine learning has become an important area in diabetes research because it can identify patterns in medical data and support prediction-based decision-making. In recent years 2023, 2024, 2025 systematic review and research on machine learning studies in pediatric diabetes reported that machine learning has strong potential in child diabetes care, but many published studies still have weak reporting quality and limited transparency, which can make clinical use more difficult [7][11][12][13][14] . This finding is important because it shows that machine learning is useful, but it must be applied in a careful and practical way. For a system like GLUCOTWIN, this means that prediction should not only be accurate, but also understandable and connected to real healthcare use. Another recent review published in 2025 focused on prediction models for diabetes in children and adolescents [15]. This review explained that machine learning and deep learning models are becoming more useful for identifying risk factors, improving early detection, and supporting disease prediction in younger age groups [8]. The same study also highlighted that stronger datasets and better data repositories are still needed for wider use in real healthcare environments [8]. This is directly related to GLUCOTWIN because the application also depends on trained data and health-related variables to support insulin-related decisions. The literature therefore supports the idea that child-focused prediction systems are necessary, especially when they are built around real clinical needs.

Recent research also shows that mobile health applications are playing a major role in diabetes care. A 2024 systematic review on mHealth coaching technologies for children and adolescents with type 1 diabetes found that digital tools are increasingly being used to support diabetes management in younger populations, and these tools often involve not only patients, but also parents and healthcare professionals [9][16] [17]. This point is very important because effective diabetes care for children usually requires shared responsibility. GLUCOTWIN also follows this direction by involving guardians, doctors, and admin in one connected system. The same review also showed that digital diabetes technologies are most useful when they support real management needs instead of only working as simple data storage tools [9][17]. This is highly relevant because GLUCOTWIN is not designed only to save records. It is designed to collect child health values, support carbohydrate-related input, and provide insulin prediction in a practical way. This makes the application more useful in daily life, especially for guardians who may feel confused when they need to decide insulin requirements under changing health conditions. A healthcare application becomes more useful when its design is practical, connected, and easy to use. A 2025 systematic review and meta-analysis on mobile health interventions for diabetes management reported that applications with effective technological functionality and proper system architecture can significantly improve diabetes self-management and reduce HbA1c levels in many cases [10] [18] [19]. The study explained that app functions, intervention models, and technical structure all play an important role in the success of digital diabetes systems [10]. This finding strongly supports the design of GLUCOTWIN because the project combines machine learning prediction, role-based access, real-time data handling, and mobile user interaction in one platform. This literature also shows that good healthcare systems need more than one strong feature. A useful system should combine prediction, structured data flow, and easy interaction for users. In GLUCOTWIN, this appears through the use of a machine learning model for insulin prediction, Flutter for mobile development, and Firebase for data management. The reviewed literature supports this kind of integrated approach because it shows that strong technical design is necessary for a healthcare application to be practically useful [10]. The literature clearly shows that machine learning has strong potential in paediatric diabetes care, and mobile health systems are becoming more important in supporting daily diabetes management [7]–[10]. However, the literature also shows several gaps. Many machine learning studies in paediatric diabetes still have weak reporting quality and limited transparency, which reduces trust and makes practical implementation more difficult [7][18][19]. In addition, recent reviews show that better datasets and stronger data repositories are still needed for wider use of prediction models in children and adolescents [8]. Mobile health studies also show that applications are most effective when they provide practical support and strong technical functionality, but many systems still focus on limited features instead of offering a complete solution [9], [10][17][18][19][20]. This creates a need for a more integrated system that can combine multiple important functions in one platform. A major gap in the current literature is the limited availability of child-focused systems that bring together health data input, insulin-related prediction, guardian support, doctor review, and organized record management in one mobile application [7]–[20]. GLUCOTWIN was developed to address this gap by providing a structured and intelligent solution for diabetic children aged 3 to 12 years. It combines data collection, machine learning prediction, guardian use, doctor supervision, and admin monitoring in a single system, which makes it more practical and complete than many isolated tools discussed in the literature.

3. Proposed System

GLUCOTWIN is a role-based healthcare application designed to help manage diabetes in children by using machine learning, mobile technology, and real-time data management in one connected system. The overall system includes three main actors: Admin, Guardian, and Doctor. Each actor has a separate role and specific access in the application. The system is developed using Flutter for the mobile front end, Firebase for storing and managing data, and a machine learning model deployed through Python for insulin prediction. The guardian enters child details and health-related values such as blood pressure, temperature, BMI, and food information. The system also estimates carbohydrate values from food details and sends all required inputs to the machine learning model. After processing the data, the model returns the predicted insulin value. This result is then available for doctor review, and the doctor can provide suggestions based on the prediction and child condition. The admin monitors the whole system, approves doctor accounts, and manages records and activities. In this way, the system works as a complete and structured healthcare platform that connects users, health data, intelligent prediction, and medical supervision in one application. Figure 1 shows system architecture of our proposed system.

IEER
Institute for Excellence in Education & Research

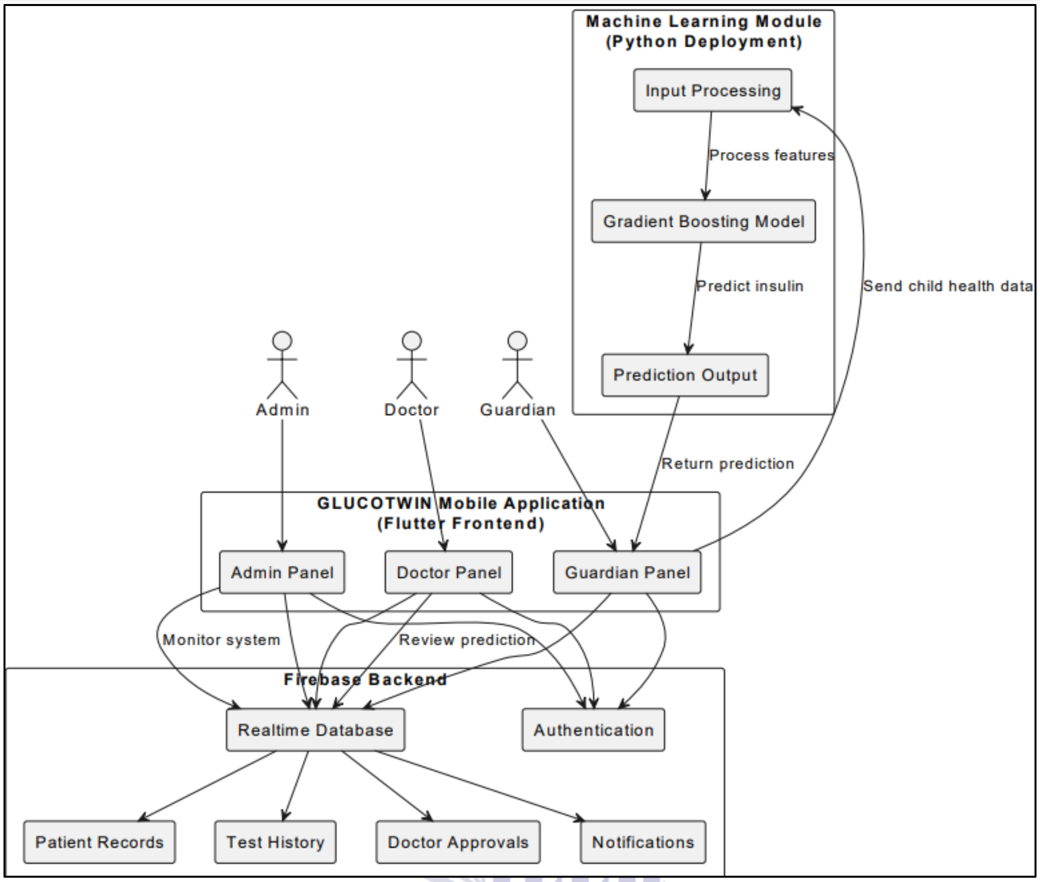

Figure 1: System Architecture

4. Results

The results of our proposed system are mentioned below:

4.1 Dataset and Model Training

The dataset file named **Glocotwin.xlsx** was first uploaded to Google Drive. After that, Google Drive was mounted in Google Colab so that the file could be accessed directly from the notebook. Once the connection was established, the dataset was loaded using the Pandas library. This step allowed the data to be read into a dataframe for further analysis and training. According to the notebook output, the dataset contained **3000 rows and 11 columns**, which means it had 3000 records and 11 attributes. These attributes were **PatientID, Gender, Age, SystolicBP_mmHg, DiastolicBP_mmHg, BodyTemp_C, Sugar_mg_dL, BMI, MealTiming, Carbohydrates_g, and InsulinUnits_Simulated**. This shows that the dataset included both health-related input features and the insulin output value that the model needed to learn.

4.2 Understanding the Dataset

After loading the dataset, the next step was to understand its structure and identify which columns should be used for training. The target variable in this project was **InsulinUnits_Simulated**, because this is the value the model needed to predict. The **PatientID** column was not useful for prediction, so it was removed from the training process. This is a correct step because patient ID only identifies the record and does not help the model learn medical patterns. The remaining input features included both numeric and categorical data. The categorical columns were **Gender** and **MealTiming**, while the numeric columns were **Age, SystolicBP_mmHg, DiastolicBP_mmHg, BodyTemp_C, Sugar_mg_dL, BMI, and Carbohydrates_g**. This separation was important because machine learning models cannot directly work with text values, so categorical data needed special preprocessing.

Based on the dataset file, the age values ranged up to **15 years**, while the application focus is on children aged **3 to 12 years**. The data also included realistic health-related ranges such as body temperature, blood pressure, sugar level, BMI, carbohydrate intake, and simulated insulin units. The median insulin value in the dataset was around **3.2**, while the maximum value reached **9.8**. These values gave a useful range for the model to learn how insulin changes according to health conditions.

```
from google.colab import drive
drive.mount('/content/drive')

import pandas as pd

file_path = "/content/drive/MyDrive/Glocotwin.xlsx"
df = pd.read_excel(file_path)

print("Shape:", df.shape)
print(df.columns.tolist())
df.head()
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
Shape: (3000, 11)
['PatientID', 'Gender', 'Age', 'SystolicBP_mmHg', 'DiastolicBP_mmHg', 'BodyTemp_C', 'Sugar_mg_dL', 'BMI', 'MealTiming', 'Carbohydrates_g', 'InsulinU
```

	PatientID	Gender	Age	SystolicBP_mmHg	DiastolicBP_mmHg	BodyTemp_C	Sugar_mg_dL	BMI	MealTiming	Carbohydrates_g	InsulinUnits_Simulated
0	P00001	Male	8	90	62	36.3	122	17.4	AfterMeal	78	4.9
1	P00002	Female	15	104	57	37.1	179	22.4	AfterMeal	100	8.1
2	P00003	Male	10	97	47	36.7	84	18.5	BeforeMeal	13	0.7
3	P00004	Male	7	75	53	36.9	152	12.5	AfterMeal	81	5.6
4	P00005	Male	4	78	41	36.8	118	19.8	BeforeMeal	59	4.4

Figure 2: Dataset Upload

4.

3 Data Preprocessing

Before training the models, preprocessing was applied to prepare the data properly. First, the input data **X** and target data **y** were separated. The input data contained all the feature columns, while the target data contained only **InsulinUnits_Simulated**. After this, the dataset was split into two parts: **80% for training and 20% for testing** using train_test_split with random_state=42. This means the model was trained on most of the data and tested on unseen data to check how well it performs in real prediction situations.

A preprocessing pipeline was then prepared using ColumnTransformer. For numeric columns, **StandardScaler** was used. This step standardizes numeric values so that all features remain on a similar scale, which helps some machine learning models perform better. For categorical columns, **OneHotEncoder** was used. This converted text values such as Male, Female, BeforeMeal, and AfterMeal into numeric format so that the models could use them correctly. This preprocessing step was very important because the dataset contained mixed types of data and the models required a consistent numeric format.

```python
target = "InsulinUnits_Simulated"

# Drop obvious ID columns if present
possible_id_cols = ["PatientID", "Patient_ID", "PatientCode", "PatientCodeID", "Patient_Id", "ID", "Patient"]
df = df.drop(columns=[c for c in possible_id_cols if c in df.columns], errors="ignore")

# Auto-drop ID-like object columns (95%+ unique)
obj_cols = df.select_dtypes(include=["object"]).columns.tolist()
id_like = [c for c in obj_cols if df[c].nunique() / len(df) > 0.95]

print("Object cols:", obj_cols)
print("Dropping ID-like cols:", id_like)

df = df.drop(columns=id_like, errors="ignore")

# Split X/y
X = df.drop(columns=[target])
y = df[target]

print("Remaining object columns in X:", X.select_dtypes(include=["object"]).columns.tolist())
```

```
Object cols: ['Gender', 'MealTiming']
Dropping ID-like cols: []
Remaining object columns in X: ['Gender', 'MealTiming']
```

Figure 3: Data Preprocessing

4.3.1 Model Training

After preprocessing, multiple machine learning regression models were trained and compared. The notebook used the following models:

- **Linear Regression**
- **Ridge Regression**
- **Random Forest Regressor**
- **Gradient Boosting Regressor**
- **Support Vector Regressor (SVR)**

Each model was trained using the same training data and preprocessing pipeline. Some models also used **GridSearchCV** for hyperparameter tuning, which means different parameter combinations were tested to find the best settings for performance. This is a good training practice because it improves the chances of selecting the most accurate and stable model.

The purpose of trying different models was to compare their results and choose the one that predicts insulin values most accurately. Since the problem is numerical prediction, regression models were the correct choice.

```python
for name, (model, param_grid) in models.items():
    pipe = Pipeline(steps=[("preprocess", preprocess), ("model", model)])

    if param_grid:
        search = GridSearchCV(
            pipe,
            param_grid=param_grid,
            scoring="neg_mean_absolute_error",
            cv=5,
            n_jobs=-1
        )
        search.fit(X_train, y_train)
        best_model = search.best_estimator_
        best_params = search.best_params_
    else:
        best_model = pipe.fit(X_train, y_train)
        best_params = {}

    preds = best_model.predict(X_test)

    mae = mean_absolute_error(y_test, preds)
    rmse = root_mean_squared_error(y_test, preds)
    r2 = r2_score(y_test, preds)

    results.append((name, mae, rmse, r2, best_params))
    best_estimators[name] = best_model

res_df = pd.DataFrame(results, columns=["Model", "MAE", "RMSE", "R2", "BestParams"]).sort_values("MAE")
res_df
```

	Model	MAE	RMSE	R2	BestParams
3	GradientBoosting	0.292139	0.373556	0.956989	{'model__learning_rate': 0.05, 'model__max_dep...
4	SVR	0.294526	0.380667	0.955336	{'model__C': 1, 'model__epsilon': 0.1, 'model_...
0	LinearRegression	0.305494	0.390387	0.953026	{}
1	Ridge	0.305595	0.390320	0.953042	{'model__alpha': 10.0}
2	RandomForest	0.320295	0.407090	0.948920	{'model__max_depth': 16, 'model__min_samples_s...

Figure 4: Model Training

4.3.2 Training Results

According to the notebook results, the performance of the trained models was as follows: These results show that all models performed well, because every R^2 score was above **0.94**, which means each model explained most of the variation in insulin values. However, the best performance came from **Gradient Boosting**, because it had the **lowest MAE**, the **lowest RMSE**, and the **highest R^2 score** among all tested models. This means Gradient Boosting produced the most accurate predictions with the smallest average error.

Model	MAE	RMSE	R^2 Score
Gradient Boosting	0.292139	0.373556	0.956989
SVR	0.294526	0.380667	0.955336
Linear Regression	0.305494	0.390387	0.953026
Ridge Regression	0.305595	0.390320	0.953042
Random Forest	0.320295	0.407090	0.948920

4.3.3 Best Model Selection

After comparing all models, **Gradient Boosting Regressor** was selected as the final model. This was the correct choice because it performed better than all other models in all three evaluation metrics. The notebook clearly showed:

BEST MODEL = GradientBoosting

This means the final insulin prediction system in GLUCOTWIN is based on the Gradient Boosting model. The result also suggests that the relationship between health variables and insulin value is not purely linear, and a boosting-based model was better able to learn the patterns from the data.

```
best_name = eval_df.iloc[0]["Model"]
best_model = best_estimators[best_name]
best_preds = pred_store[best_name]

print("BEST MODEL =", best_name)
print(eval_df.iloc[0].to_dict())

# Show few samples
preview = X_test.iloc[:10].copy()
preview["y_true"] = y_test.iloc[:10].values
preview["y_pred"] = best_model.predict(X_test.iloc[:10])
preview
```

```
BEST MODEL = GradientBoosting
{'Model': 'GradientBoosting', 'MAE': 0.29213943271338644, 'RMSE': 0.3735562607191907, 'R2': 0.956989098798979}
```

	Gender	Age	SystolicBP_mmHg	DiastolicBP_mmHg	BodyTemp_C	Sugar_mg_dL	BMI	MealTiming	Carbohydrates_g	y_true	y_pred
1801	Female	6	101	40	36.9	187	19.3	AfterMeal	82	5.8	6.193504
1190	Male	11	94	54	37.0	174	18.3	AfterMeal	47	3.4	3.835881
1817	Male	2	79	49	36.9	152	15.7	AfterMeal	71	4.1	4.614749
251	Male	14	102	68	36.9	182	23.7	AfterMeal	44	4.1	4.236942
2505	Male	5	87	56	36.6	92	18.6	BeforeMeal	14	0.7	0.896690
1117	Male	15	104	57	37.0	130	20.4	BeforeMeal	47	3.8	3.671651
1411	Female	6	88	51	37.0	130	21.5	AfterMeal	69	4.6	4.582414
2113	Female	11	99	63	36.5	88	19.1	BeforeMeal	17	1.1	1.104359
408	Male	2	85	52	36.7	133	15.8	AfterMeal	56	3.2	3.500435
2579	Female	5	87	50	36.9	101	14.8	BeforeMeal	28	1.7	1.588209

Figure 5: Model Testing

IIEER
Institute for Excellence in Education & Research

4.3.4 Actual vs Predicated

The graph shows the comparison between the actual insulin values and the predicted insulin values using the Gradient Boosting model. In this plot, the x-axis represents the actual values, while the y-axis represents the predicted values generated by the model. Most of the data points are very close to the straight diagonal line, which means that the predicted values are very close to the actual values. This indicates that the model is performing very well and making accurate predictions. The points are evenly distributed along the line, showing that the model is consistent across different value ranges. There are only a few points that are slightly away from the line, which means the prediction error is small and acceptable. Overall, this graph clearly shows that the Gradient Boosting model has high accuracy and is reliable for predicting insulin values in the system.

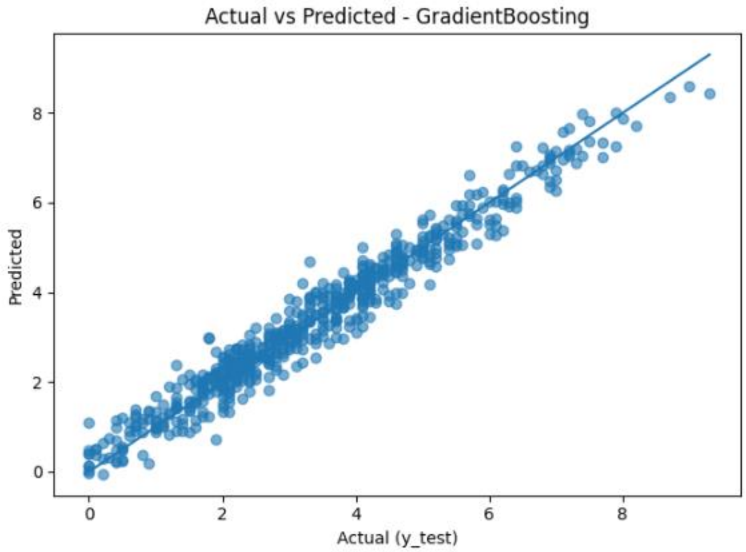

Figure 6: Actual vs Predicated

4.3.5 Cross-validated predictions of the Gradient Boosting model

The graph shows the cross-validated predictions of the Gradient Boosting model on the training data. In this plot, the x-axis represents the actual insulin values from the training dataset, while the y-axis represents the predicted values generated by the model during cross-validation. Most of the data points are closely aligned along the diagonal line, which means that the predicted values are very similar to the actual values. This indicates that the model has learned the patterns in the training data very well. The points are evenly distributed across the entire range, showing that the model performs consistently for both low and high insulin values. There are only a few small deviations from the line, which means the prediction error is low. Overall, this graph shows that the Gradient Boosting model is stable, accurate, and reliable, even when evaluated using cross-validation, which strengthens confidence in its performance for real-world predictions.

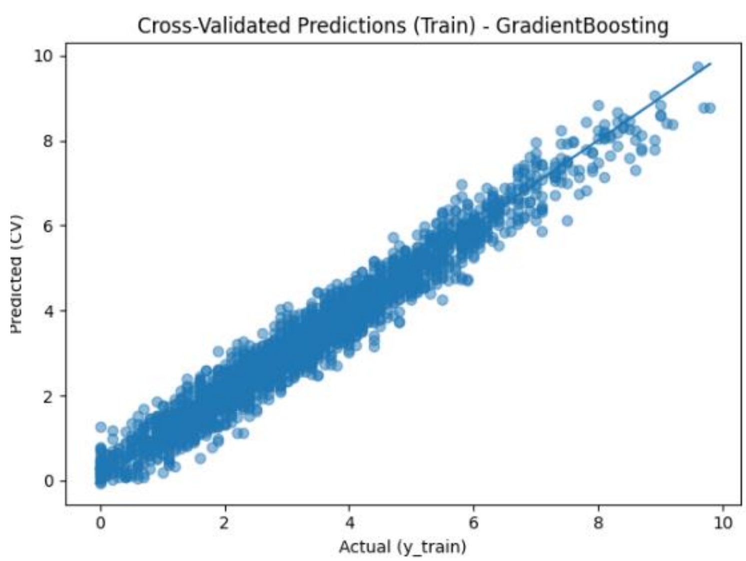

Figure 7: Cross-Validated predictions

4.4

Model Saving and Deployment

After selecting Gradient Boosting as the best model, it was saved to Google Drive as **best_insulin_model.pkl** using the joblib library. A metadata file was also created to save the best model name and evaluation metrics. This was an important step because it allowed the trained model to be reused later without retraining it again.

The saved model was then used in a Python script that loads the file, accepts new input values, and returns an insulin prediction. This shows that the training phase was completed successfully and the model was ready to be connected with the GLUCOTWIN application for real-time use.

Dataset training process was completed successfully. The dataset was first loaded from Google Drive, then cleaned and prepared for training. The target variable was selected as **InsulinUnits_Simulated**, while unnecessary identification data was removed. The data was split into training and testing sets, and both numeric and categorical columns were preprocessed properly. Multiple regression models were trained and evaluated using MAE, RMSE, and R^2 score. Among all tested models, **Gradient Boosting Regressor** gave the best results with **MAE = 0.292139**, **RMSE = 0.373556**, and **R^2 = 0.956989**. These results show that the model was highly accurate and suitable for insulin prediction. After evaluation, the model was saved and successfully used for new predictions, which made it ready for integration into the GLUCOTWIN healthcare application.

4.5 User Interface

The user interface of the application is designed to be simple, clear, and easy to use for all users, especially guardians who may not have technical knowledge. The interface includes different screens for admin, guardian, and doctor, each with specific functions and access. The goal of the user interface design is to make the system user-friendly so that users can easily navigate, enter data, view results, and understand the information without confusion.

The meal details and testing screens in the GLUCOTWIN application are designed to guide the guardian step by step in entering health and food-related information in a simple way. In the meal details screen, the guardian can add food items by selecting categories and specifying quantity, which helps in calculating total carbohydrates. After adding items, the system calculates the total carbohydrate value automatically, as shown in the screen where white rice results in 45.0 grams of carbohydrates. This value is important because it is used as input for insulin prediction. Once the meal information is complete, the guardian proceeds to the sugar test screen, where important health values such as glucose level, body temperature, blood pressure, and meal timing are entered. These inputs are then submitted to the system for analysis. After submission, a confirmation message appears showing that the AI analysis has been completed successfully, and the user is informed to wait for the doctor's review before taking any action. Finally, the test results screen displays the predicted insulin level along with all entered health details in a clear and organized format. This complete flow shows how the system collects data, processes it using the machine learning model, and presents results in a user-friendly way, helping guardians make better decisions for child health

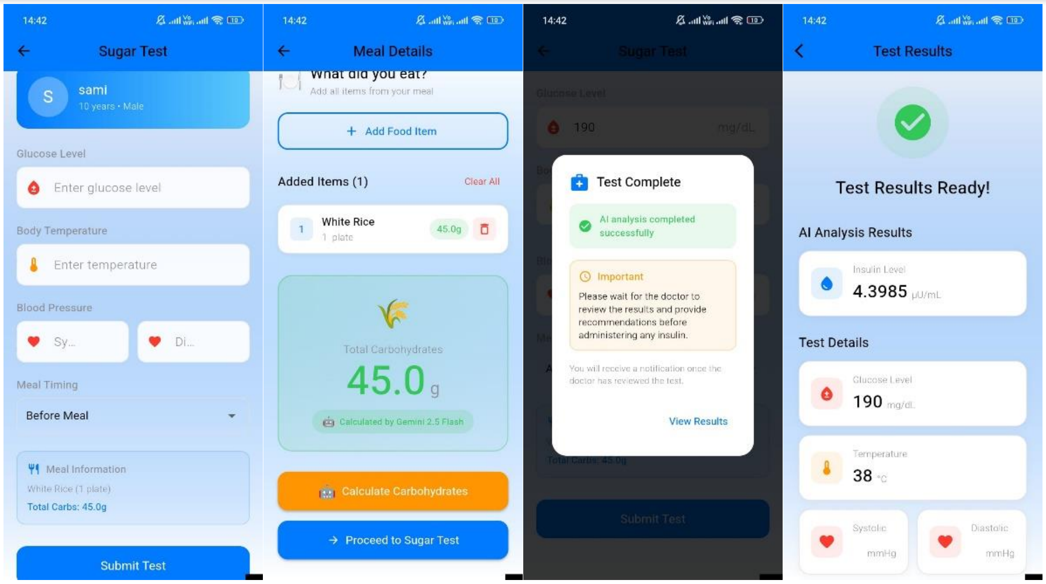

Figure 8: Meal and Insulin Results

The doctor screens in the GLUCOTWIN application are designed to help medical professionals review patient data, provide recommendations, and manage treatment in a simple and organized way. The doctor dashboard shows an overview of total tests, pending cases, completed reviews, and total patients, which helps the doctor quickly understand their workload. The system also separates tests into categories such as pending review and completed cases, making it easy to manage patient records. Each patient card displays important information such as glucose level, insulin prediction, carbohydrate intake, and blood pressure, allowing the doctor to quickly assess the patient's condition.

The patient detail screen provides complete information including AI prediction results, patient details, and test measurements. The doctor can clearly see whether the condition is low risk and review all health values such as glucose, temperature, blood pressure, meal timing, and carbohydrate intake. Based on this information, the doctor can write a clinical recommendation, prescribe medication or insulin dose, and add additional notes if needed. The system also allows the doctor to mark when a dose has been given, which helps in tracking treatment progress. Overall, these screens are designed to support doctors in making informed decisions, reviewing AI-based predictions, and providing safe and proper medical guidance for child diabetes management.

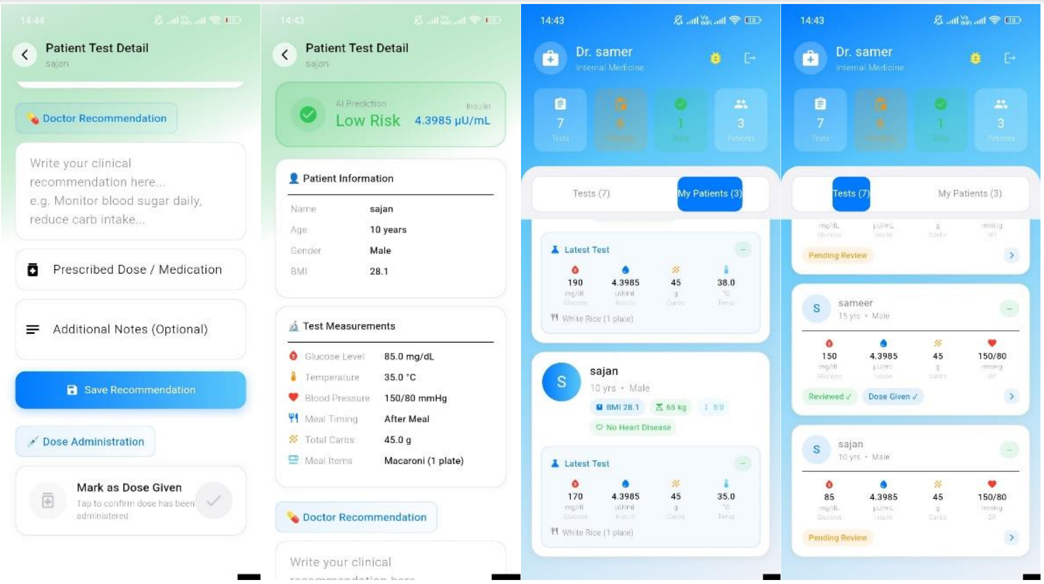

Figure 9: Doctor Medical Recommendations

5. Conclusion and Future Work:

This project presented GLUCOTWIN as a smart healthcare application developed to support diabetes management for children aged 3 to 12 years. The main purpose of this system was to help guardians who often face difficulty in deciding the correct insulin requirement when a child's health condition changes. Since insulin dosage depends on several important factors such as glucose level, body temperature, blood pressure, BMI, meal timing, and carbohydrate intake, manual decision-making can become confusing and risky. To solve this problem, GLUCOTWIN was designed as a complete system that combines mobile application development, machine learning, real-time database support, and doctor involvement in one platform. The project successfully showed how a trained machine learning model can be used in a practical healthcare application. A proper dataset was used for training different machine learning models, and after comparing their results, the best-performing model was selected for insulin prediction. The selected model was then deployed using Python and connected with the mobile application so that real-time input values could be processed easily. This prediction process made the system intelligent and useful because it allowed health data to be converted into meaningful insulin-related results in a short time. The project also demonstrated the importance of role-based system design. GLUCOTWIN included three main users: admin, guardian, and doctor. Each role was designed with different responsibilities to make the system more secure, organized, and easy to manage. Guardians were able to add children, enter meal and health information, calculate carbohydrates, perform sugar tests, and view prediction results. Doctors were able to review patient tests, check AI-generated insulin values, and provide medical recommendations. Admin managed doctor approvals, patient assignments, and system monitoring. This structure helped ensure that the system was not only technically complete, but also practical for real use. Another important achievement of this project was the development of a simple and user-friendly interface. The screens were designed in a clean and understandable way so that users could interact with the application without confusion. From the splash screen and onboarding pages to login, patient management, sugar testing, doctor review, and admin monitoring, every part of the system was designed to support clear and smooth use. This is very important because healthcare applications must not only work properly, but must also be easy for real users to understand and use. Overall, GLUCOTWIN provides a meaningful solution for a real healthcare problem.

It helps reduce confusion for guardians, supports doctors with organized patient data and prediction results, and gives admin full control over system activities. The project shows that machine learning can be applied in a practical and helpful way when it is combined with proper application design and medical review. In conclusion, GLUCOTWIN is not only a technical project, but also a useful healthcare support system that can improve diabetes management for children in a safer, smarter, and more organized way.

In future work, the model can be deployed on a cloud-based server or a web API so that the system can provide faster, more scalable, and more stable real-time predictions. Cloud deployment would also make the system more practical for larger numbers of users because the prediction service could be accessed from anywhere without depending on one local machine. This would improve system performance and make GLUCOTWIN more suitable for real-world deployment on a larger scale.

REFERENCES

World Health Organization, "Diabetes," Nov. 2024.

Centers for Disease Control and Prevention, "National Diabetes Statistics Report," 2023.

American Diabetes Association, "Standards of Medical Care in Diabetes—2024," Diabetes Care, vol. 47, 2024.

Z. Zrubka et al., "The Reporting Quality of Machine Learning Studies on Pediatric Diabetes Mellitus: Systematic Review," J. Med. Internet Res., vol. 26, 2024.

L. Cveticanin and M. Arsenovic, "Prediction Models for Diabetes in Children and Adolescents," Appl. Sci., vol. 15, 2025.

P. Trnka et al., "mHealth Coaching Technologies for Children with Diabetes," JMIR Pediatrics, vol. 7, 2024.

Yu, X., Wang, Y., Liu, Z., & Jung, E. (2025). Technological functionality and system architecture of mobile health interventions for diabetes management: a systematic review and meta-analysis of randomized controlled trials. Frontiers in public health, 13, 1549568.

Kalshetty, S., Kulkarni, N. S., Salunkhe, M. S., & Kulkarni, R. (2026). Comparative Analysis of Machine Learning Models for Rheumatoid Arthritis Prediction.

Ahmad, F., Alasskar, A., Samui, P., & Asteris, P. G. (2025). Machine learning-based graphical user interface for predicting high-performance concrete compressive strength: comparative analysis of gradient boosting machine, random forest, and deep neural network models. Frontiers of Structural and Civil Engineering, 19(7), 1075-1090.

Gerber, B. S., Biggers, A., Tilton, J. J., Smith Marsh, D. E., Lane, R., Mihailescu, D., ... & Sharp, L. K. (2023). Mobile health intervention in patients with type 2 diabetes: a randomized clinical trial. JAMA Network Open, 6(9), e2333629.

Usman, T. M., Saheed, Y. K., Nsang, A., Ajibesin, A., & Rakshit, S. (2023). A systematic literature review of machine learning based risk prediction models for diabetic retinopathy progression. Artificial intelligence in medicine, 143, 102617.

Cveticanin, L., & Arsenovic, M. (2025). Prediction models for diabetes in children and adolescents: A review. Applied Sciences, 15(6), 2906.

Kumar, A. Implementation of Machine Learning Techniques for the Detection of the Pediatric Diseases.

Theodore Armand, T. P., Nfor, K. A., Kim, J. I., & Kim, H. C. (2024). Applications of artificial intelligence, machine learning, and deep learning in nutrition: a systematic review. Nutrients, 16(7), 1073.

IEER Excellence in Education & Research

Cveticanin, L., & Arsenovic, M. (2025). Prediction models for diabetes in children and adolescents: A review. Applied Sciences, 15(6), 2906.

Shan, R., Sarkar, S., & Martin, S. S. (2019). Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. Diabetologia, 62(6), 877-887.

Årsand, E., Frøisland, D. H., Skrøvseth, S. O., Chomutare, T., Tatara, N., Hartvigsen, G., & Tufano, J. T. (2012). Mobile health applications to assist patients with diabetes: lessons learned and design implications. Journal of diabetes science and technology, 6(5), 1197-1206.

Yu, X., Wang, Y., Liu, Z., & Jung, E. (2025). Technological functionality and system architecture of mobile health interventions for diabetes management: a systematic review and meta-analysis of randomized controlled trials. Frontiers in public health, 13, 1549568.

Versluis, A., Boels, A. M., Huijden, M. C., Mijnsbergen, M. D., Rutten, G. E., & Vos, R. C. (2025). Diabetes self-management education and support delivered by mobile health (mHealth) interventions for adults with type 2 diabetes—A systematic review and meta-analysis. Diabetic Medicine, 42(5), e70002.

Kim, S. K., Park, S. Y., Hwang, H. R., Moon, S. H., & Park, J. W. (2025). Effectiveness of mobile health intervention in medication adherence: a systematic review and meta-analysis. Journal of medical systems, 49(1), 13.

Chen, Y., Chen, Z., Zhao, L., Li, S., Ying, Z., Yu, P., ... & Li, X. (2023). Real-time artificial intelligence assisted insulin dosage titration system for glucose control in type 2 diabetic patients: a proof-of-concept study. Current Medicine, 2(1), 2.

Maqbool, M. S., Hanif, I., Iqbal, S., Basit, A., & Shabbir, A. (2023). Optimized feature extraction and cross-lingual text reuse detection using ensemble machine learning models. Journal of Computing & Biomedical Informatics, 5(01), 26-40.

Abid, K., Aslam, N., Fuzail, M., Maqbool, M. S., & Sajid, K. (2023). An efficient deep learning approach for prediction of student performance using neural network. VFAST Transactions on Software Engineering, 11(4), 67-79.

Kanwal, F., Abid, M. K., Maqbool, M. S., Aslam, N., & Fuzail, M. (2023). Optimized classification of cardiovascular disease using machine learning paradigms. VFAST Transactions on Software Engineering, 11(2), 140-148.

Basit, A., Hanif, I., Maqbool, M. S., Qayyum, W., Hasnain, M. A., & Nazeer, R. (2023). Cross-lingual information retrieval in a hybrid query model for optimality. Journal of Computing & Biomedical Informatics, 5(01), 130-141.

Fazal, U., Khan, M., Maqbool, M. S., Bibi, H., & Nazeer, R. (2023). Sentiment analysis of omicron tweets by using machine learning models. VFAST Transactions on Software Engineering, 11(1), 67-75.

Hasnain, M. A., Ali, S., Malik, H., Irfan, M., & Maqbool, M. S. (2023). Deep learning-based classification of dental disease using x-rays. Journal of Computing & Biomedical Informatics, 5(01), 82-95.

Hasnain, M. A., Ali, Z., Maqbool, M. S., & Aziz, M. (2024). X-ray image analysis for dental disease: A deep learning approach using efficientnets. VFAST Transactions on Software Engineering, 12(3), 147-165.

Maqbool, M. S., Zahra, S. R., Ismail, S., Nadeem, M., Fatima, N., & Ahmad, J. (2026). A CNN-BASED FRAMEWORK FOR EFFICIENT DETECTION OF EYE DISEASE IN FUNDUS IMAGES. Spectrum of Engineering Sciences, 4(4), 1157-1169.

Rafiqee, M. M., Qaiser, Z. H., Fuzail, M., Aslam, N., & Maqbool, M. S. (2023). Implementation of efficient deep fake detection technique on videos dataset using deep learning method. Journal of Computing & Biomedical Informatics, 5(01), 345-357.

Zainab, F., Nazim, F., Kashaf, M., Aslam, N., & Maqbool, M. S. (2026). Predictive analytics for customer churn in subscription-based businesses using machine learning. Spectrum of Engineering Sciences, 4(4), 596-618.

Tanveer, K., Aslam, N., Saleem, H., Maqbool, M. S., Akhter, M., & Rashid, M. H. (2026). INTEGRATING BLOCKCHAIN AND MACHINE LEARNING FOR ROBUST IOT DATA INTEGRITY IN CRITICAL INFRASTRUCTURES. Spectrum of Engineering Sciences, 4(3), 1875-1886.

Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). Understanding Urban Expansion Through Multi-Temporal Satellite Data Analysis. Kashf Journal of Multidisciplinary Research, 2(09), 252-273.

Malik, F., Fuzail, M., Aslam, N., Sarwar, R., Abid, K., Maqbool, M. S., & Yousaf, A. (2024). A hybrid machine learning model to predict sentiment analysis on X. Journal of Computing & Biomedical Informatics, 6(02), 64-79.

Maqbool, M. S., Nazeer, R., Basit, A., & Zahra, K. (2023). Automated detection and localization of fungal infections on cotton leaves using YOLO-based object detection model. Machines and Algorithms, 2(2), 121-136.

Hasnain, M. A., Ali, Z., Rehman, K. U., Ehtsham, M., & Maqbool, M. S. (2024). Idd-net: A deep learning approach for early detection of dental diseases using x-ray imaging. Journal of Computing & Biomedical Informatics, 7(02).

Aslam, N., Maqbool, M. S., Nadeem, M., & Saleem, H. (2026). DEEPFAKESHIELD: ENHANCED VIDEO AUTHENTICITY DETECTION VIA CONVOLUTIONAL VISION TRANSFORMER. Spectrum of Engineering Sciences, 4(3), 1650-1665.

Manzoor, I., Maqbool, M. S., Shahzad, F., Nadeem, M., Zulfiqar, A., & Naqvi, S. Q. (2026). A FRAMEWORK FOR HATE SPEECH IDENTIFICATION USING OPTIMIZED TEXT FEATURES AND NATURAL LANGUAGE PROCESSING ON TWITTER DATASET. Spectrum of Engineering Sciences, 4(6), 1389-1405.

Maqbool, M. S., Fatima, N., Nazeer, R., Aslam, N., Abbas, F., Sumra, U., & Nadeem, M. (2025). A HYBRID DATASET-BASED ENSEMBLE STRATEGY FOR EFFICIENT BREAST CANCER DETECTION. Kashf Journal of Multidisciplinary Research, 2(12), 39-57.

Aslam, M., Maqbool, M. S., Aoun, M., Aslam, N., Razzaq, A. M., & Ali, S. (2026). HIGH-PERFORMANCE AND EFFICIENT BRAIN TUMOR SEGMENTATION FOR ENHANCED CLINICAL ANALYSIS. Spectrum of Engineering Sciences, 4(3), 195-210.

Israr, S., Maqbool, M. S., Hanif, I., Nadeem, M., Basit, A., & Batool, A. A. (2026). A ROBUST DEEP LEARNING FRAMEWORK FOR PREDICTING ACADEMIC SUCCESS USING ADVANCED NEURAL ARCHITECTURES. Spectrum of Engineering Sciences, 4(5), 1006-1022.

Ahmed, T. (2025). Glycemic Control and Insulin Management: The Role of AI in Glycemic Index Control and Insulin Regulation. In AI-Based Nutritional Intervention in Polycystic Ovary Syndrome (PCOS) (pp. 185-211). Singapore: Springer Nature Singapore.

Alhijailan, H., Alyahya, R., Alhawass, R., Alyousef, W., & Alyahya, K. (2025, September). InsulinSync: A Type 1 Diabetes management application with AI-Decision Support System for Insulin Dosing and Computer Vision-Based Nutrition Prediction. In 2025 16th Student Research Conference on Applied Computing (SRC) (pp. 1-6). IEEE.

Arefeen, A., Khamesian, S., Grando, M. A., Thompson, B., & Ghasemzadeh, H. (2025). Glytwin: Digital twin for glucose control in type 1 diabetes through optimal behavioral modifications using patient-centric counterfactuals. arXiv preprint arXiv:2504.09846.

Ahmed, T. (2025). Glycemic Control and Insulin Management: The Role of AI in Glycemic Index Control and Insulin Regulation. In AI-Based Nutritional Intervention in Polycystic Ovary Syndrome (PCOS) (pp. 185-211). Singapore: Springer Nature Singapore.

Pellizzari, E., Cappon, G., Nicolis, G., Sparacino, G., & Facchinetti, A. (2025). Development of Insulin Bolus Calculators in Type 1 Diabetes using A Framework Based on Real-world Data, Digital Twins and Machine Learning. IEEE Transactions on Biomedical Engineering.

Tominaga, H., Hamaguchi, M., Hamaguchi, Y., Yashiki, R., Yamaguchi, A., Arai, T., ... & Fukui, M. (2025). Prediction of Postprandial Blood Glucose Variability Using Machine Learning in Frequent Insulin Injection Therapy with a Simplified Carbohydrate Counting Model. Nutrients, 17(24), 3832.

Wilson, S. M. (2022). GLUCAGON: AI-Based Insulin Dosage Prediction Application.

IEER
Institute for Excellence in Education & Research