NON-INVASIVE DIABETES DETECTION USING MACHINE LEARNING

Project Reference No.: *48S_BE_4737*

College : Sri Krishna Institute of Technology, Bengaluru
Branch : Electronics and Communication
Guide : Dr. J. Divya Lakshmi
Student(s): Ms. Amreen Firdose A
Mr. Manu R
Ms. Rachana R k
Ms. Varshini S

Keywords:

Diabetes, PPG, MAX30100, ESP32, Machine Learning, ThingSpeak, Non- Invasive.

Introduction:

Diabetes is a chronic metabolic disorder that affects millions worldwide and can lead to serious health complications if undiagnosed or poorly managed. Traditional diagnostic methods, such as fasting blood glucose tests and oral glucose tolerance tests, are invasive and inconvenient. This project introduces an innovative, non-invasive approach to diabetes detection utilizing Photoplethysmography (PPG) signals captured using the MAX30100 sensor. By integrating machine learning algorithms with real-time hardware and cloud-based monitoring, this system offers an efficient, user-friendly, and scalable alternative to conventional testing. PPG signals, along with critical patient metrics such as age, BMI, blood pressure, glucose levels, and more, are used to predict diabetes status using Logistic Regression and XGBoost models. The ESP32 microcontroller collects sensor data and transmits it to the ThingSpeak platform for analysis and visualization, while a web interface allows healthcare professionals and patients to access prediction results in real time.

Objectives:

- Design a non-invasive diabetes detection system using the MAX30100 sensor and ESP32 microcontroller to monitor PPG signals and transmit data to ThingSpeak for real-time cloud analysis.

- Apply machine learning models—Logistic Regression and XGBoost—on combined sensor data and patient-specific parameters (e.g., BMI, age, glucose level) to classify individuals as diabetic or non-diabetic with high accuracy.
- Develop an intuitive web-based GUI to enable real-time data entry and visualization of diabetes prediction results, aiding in early diagnosis and personalized healthcare interventions.

Methodology:

The project methodology begins with setting up hardware components: the MAX30100 sensor captures PPG signals, which are read by the ESP32 microcontroller and transmitted via Wi-Fi to the ThingSpeak cloud platform. An LCD display provides immediate feedback to the user. Data is preprocessed for consistency by applying normalization and cleaning techniques. Patient data including BMI, glucose levels, blood pressure, age, insulin, and skin thickness is included in the dataset. Feature selection and extraction processes are applied to refine the input for the machine learning model. The machine learning phase utilizes Logistic Regression and XGBoost for binary classification. Logistic Regression provides interpretable outputs, while XGBoost offers better accuracy for complex patterns. Models are trained and validated on the Pima Indian Diabetes dataset using Python, and evaluation metrics like accuracy, recall, and F1-score are calculated. The models are serialized using pickle and deployed to provide real-time predictions through a Flask web interface. ThingSpeak integration enables cloud-based visualization and remote monitoring. A flow diagram illustrates the entire pipeline from data collection to prediction and result visualization.

Result and Conclusion:

In conclusion, the developed system successfully predicts diabetes status using non-invasive data collected via MAX30100. Logistic Regression and XGBoost models were trained and evaluated using standard metrics, achieving high levels of accuracy and recall. Real-time heart rate and glucose trends are visualized on ThingSpeak, and prediction results are accessible via a web dashboard. The LCD display provides on-device feedback. XGBoost demonstrated slightly superior performance due to its ability to model complex relationships. Overall, the solution provides a practical, low-cost, and scalable tool for early diabetes detection, promoting proactive healthcare and patient engagement. The results validate the feasibility and effectiveness of non-invasive diabetes prediction systems when enhanced with machine learning and IoT integration.

Project Outcome & Industry Relevance:

• Developed a working prototype that combines sensor hardware, machine learning, and cloud computing for non-invasive diabetes detection.
• Demonstrated the potential of machine learning in healthcare diagnostics, particularly for early identification and management of chronic diseases.
• Created an end-to-end IoT solution with real-time monitoring capabilities and user-friendly visualization through web and hardware interfaces.
• Addresses a significant healthcare need by offering an affordable, non-invasive alternative to traditional glucose testing.
• The solution aligns with current industry trends in digital health, wearable tech, and remote patient monitoring, making it suitable for integration in medical devices and health-tech startups.
• Offers scalability and potential for commercialization in preventive healthcare systems and telemedicine applications.

Working Model vs. Simulation/Study:

The project includes both a working model and a simulation study. The working model uses the MAX30100 sensor and ESP32 to capture real-time PPG signals, displaying heart rate and SpO_2 on an LCD and sending data to ThingSpeak for cloud storage. A Flask-based GUI shows glucose predictions to the user. Meanwhile, the simulation uses the Pima Indian Diabetes Dataset to train Logistic Regression and XGBoost models using Python. These models are evaluated based on accuracy, precision, recall, and F1-score. The simulation validates the ML models, while the working model brings the system to life for real-time health monitoring.

Project Outcomes & Learnings:

• Successfully implemented a real-time, non-invasive diabetes detection system using PPG sensor data and machine learning algorithms.

• Gained in-depth knowledge of integrating hardware (MAX30100, ESP32), cloud services (ThingSpeak), and software components (Flask, Python ML libraries).
• Developed skills in preprocessing biomedical data, applying classification models, and evaluating model performance using metrics like accuracy, precision, recall, and F1-score.
• Understood the importance of user interface design in healthcare applications by building a web-based GUI for intuitive data input and result display.
• Learned how IoT and ML can be leveraged to solve real-world healthcare challenges with low-cost, accessible solutions.

Future Scope:

The future scope of this project includes:

1. Integration with mobile applications for broader accessibility.
2. Incorporating additional non-invasive sensors for comprehensive health analysis.
3. Expanding dataset diversity to improve model generalization.
4. Real-time alert system for patients and healthcare providers.
5. Implementing advanced deep learning models for improved accuracy.
6. Bluetooth-enabled data sharing and offline storage capabilities.
7. Testing the system in clinical settings for real-world validation.
8. Adding encryption layers for enhanced data privacy and security.
9. Exploring predictive capabilities for other chronic diseases using the same platform.