



# **NAAN MUDHALVAN PROJECT (IBM)**

**IBM AI 101 ARTIFICIAL INTELLIGENCE –GROUP 1**

**PROJECT:**

**TEAM -3 AI-BASED DIABETES PREDICTION SYSTEM**

**PREPARED BY**

Sunil. R

email: sunilprime03@gmail.com

**NM ID:** au110321106053

ECE- IIIrd YEAR

from the department of

**ELECTRONICS AND COMMUNICATION  
ENGINEERING**

---

## **Background:**

Diabetes is a chronic medical condition affecting millions of people worldwide. Early detection and timely intervention are crucial in managing and preventing the complications associated with diabetes. AI-based predictive models have the potential to assist healthcare professionals in identifying individuals at risk of developing diabetes or those who may already have undiagnosed diabetes.

## **Problem Statement:**

Develop an AI-Based Diabetes Prediction System that can accurately predict the risk of diabetes in individuals based on their health data and lifestyle factors. The system should aim to achieve the following objectives:

1. **Data Collection:** Create a comprehensive database of health-related data, including but not limited to age, gender, family history of diabetes, body mass index (BMI), blood pressure, blood sugar levels, cholesterol levels, diet patterns, physical activity, and other relevant factors.
2. **Data Preprocessing:** Clean and preprocess the collected data, handling missing values and outliers, and ensuring data quality for model training.
3. **Feature Selection:** Identify the most relevant features for diabetes prediction from the dataset. Consider using techniques such as feature ranking, feature importance, or domain knowledge.
4. **Model Development:** Develop machine learning or deep learning models capable of predicting diabetes risk. Explore various algorithms, such as logistic regression, decision trees, random forests, support vector machines, neural networks, or ensemble

methods. The model should take the selected features as input and provide a binary classification output (diabetic or non-diabetic).

5. **Model Evaluation:** Assess the performance of the developed model(s) using appropriate evaluation metrics, such as accuracy, precision, recall, F1-score, and AUC-ROC. Perform cross-validation to ensure the model's generalizability.

6. **User Interface:** Create a user-friendly interface that allows users to input their health and lifestyle data and receive a diabetes risk prediction. The interface should also provide explanations for the prediction and suggest actionable recommendations for individuals at risk.

7. **Privacy and Security:** Implement robust data privacy and security measures to protect the sensitive health data of users.

8. **Scalability and Deployment:** Ensure that the system is scalable to accommodate a growing user base. Deploy the system on a reliable and accessible platform, such as a web application or a mobile app.

9. **Continuous Improvement:** Implement mechanisms for collecting user feedback and improving the model's accuracy over time by incorporating new data and research findings.

10. **Regulatory Compliance:** Ensure that the system complies with relevant healthcare regulations and data protection laws, such as HIPAA (in the United States) or GDPR (in Europe).

**Deliverables:**

- A functional AI-Based Diabetes Prediction System.
- Documentation detailing the data collection process, preprocessing steps, feature selection, model development, and evaluation metrics.
- User guides and documentation for healthcare professionals and end-users.
- A secure and scalable deployment of the system.
- Compliance with privacy and security standards.
- Ongoing maintenance and improvement plan.

**Note:** This problem statement outlines the development of an AI-based diabetes prediction system; however, it's essential to work closely with healthcare experts and comply with ethical and legal guidelines throughout the project. Additionally, consider seeking necessary approvals and permissions when dealing with healthcare data.