# Project Proposal: Enhancing Noninvasive Glucose Monitoring Using Wearable Technology and Machine Learning

## Introduction

Prediabetes affects approximately one in three individuals and has a concerning 10% annual progression rate to type 2 diabetes without timely lifestyle or medical interventions. Effective management of glycemic health is crucial to prevent the onset of type 2 diabetes. Currently, there is no commercially-available, noninvasive method for continuous glucose monitoring that aids in the self-management of prediabetes. This gap highlights the urgent need for innovative solutions to improve the monitoring and management of glycemic health.

In this study, we propose to leverage a dataset comprising 25,000 simultaneous measurements of interstitial glucose and noninvasive smartwatch data . Our goal is to evaluate the feasibility of using these wearable devices in conjunction with food logs recorded over a 10-day period to detect personalized glucose deviations and predict real-time interstitial glucose values.

## Project Scope

Our project aims to replicate and enhance the model presented in a recent Nature publication that predicted interstitial glucose (mg/dL) with an RMSE of 21.2 and a MAPE of 14.3%. The original study utilized eight features and engineered a total of 69 features for their final model. By refining this approach, we seek to improve prediction accuracy and reliability.

**Objectives:**

1. **Model Replication and Enhancement:** Recreate the existing model from the Nature publication, incorporating additional data and feature engineering techniques to improve predictive performance.
2. **Feature Engineering:** Explore and implement advanced feature engineering methods to enhance the model's ability to predict interstitial glucose levels.
3. **Validation and Testing:** Validate the enhanced model using the provided dataset and evaluate its performance against the original metrics.

## Data Overview

The dataset includes a range of files for each patient, with varying sampling periods, and the Dexcom data serves as the target variable for prediction. The dataset files are as follows:

| CSV | Description | Source | Median Sampling Period |
|---|---|---|---|
| **ACC_001** | Tri-axial accelerometry (X-Y-Z) | Empatica E4 wrist-worn device | 0.03125 seconds |
| **BVP_001** | Blood volume pulse | Empatica E4 wrist-worn device | 0.015625 seconds |
| **Dexcom_001** | Interstitial glucose concentration (mg/dL) | Dexcom G6, a continuous glucose monitor system | 300.0 seconds |
| **EDA_001** | Electrodermal activity | Empatica E4 wrist-worn device | 0.25 seconds |
| **HR_001** | Heart Rate | Empatica E4 wrist-worn device | 1.24 seconds |
| **IBI_001** | Interbeat interval | Empatica E4 wrist-worn device | 0.98442 seconds |
| **TEMP_001** | Skin Temperature | Empatica E4 wrist-worn device | 0.25 seconds |
| **food_log** | Log of food intake with timestamps and nutritional information | User input | As needed |
| **demographics_csv** | Sex, HbA1c, Patient ID | User input | One time |

## References


- PhysioNet Dataset: Dataset
- Nature Publication: Article


This project aims to address a critical need in glycemic health management by enhancing noninvasive glucose monitoring methods and providing actionable insights for individuals managing prediabetes. Through advanced modeling and feature engineering, we seek to contribute to more effective and accessible glucose monitoring solutions.