Data Mining Project - Cell Glucose Monitoring (CGM) Prediction Models

This project applies data mining and machine learning techniques on real-time glucose levels obtained from Cell Glucose Monitoring (CGM) sensors of 5 different subjects to identify patterns, recognize if the CGM data is recorded while taking the meal and segment the data based on carbohydrate intake.

Tasks implemented:

Project -1 Feature Engineering: Data Preprocessing techniques were applied to the time-series data and four different types of time-series features (statistical, frequency based etc) were extracted by observing patterns in CGM sensor data to form a feature matrix. Standardization and Dimensionality Reduction techniques like Principal Component Analysis (PCA) are applied to the feature matrix.

Project-2 Classification : Trained a model using KNN (K-Nearest Neighbours) and SVM (Support Vector Machines) to classify if a given time-series CGM data indicates the subject is having a meal (1 - Meal) or not having a meal (0 - No Meal) during the respective time frames. Used k-fold cross validation technique to train the model and reported Accuracy, F1 score, Precision and Recall of the machine.

Project -3 Clustering : Trained a unsupervised learning model to cluster the 'Meal Data' (CGM sensor data collected while having a meal) of 5 different subjects based on the amount of carbohydrates in each meal using KMeans and DBSCAN algorithms.

Programming Language : Python
Packages Used : scipy, pandas, matplotlib, numpy, sklearn