Erdos Institute Summer 2024

June 1, 2024

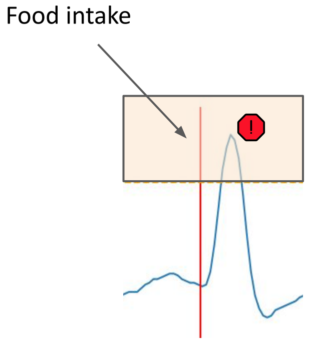
Predicting Glucose Levels using Smartwatch Data

Noah Gillespie, Margaret Swerdloff, Oladimeji Olaluwoye,
S. C. Park, Daniel Visscher

Overview:

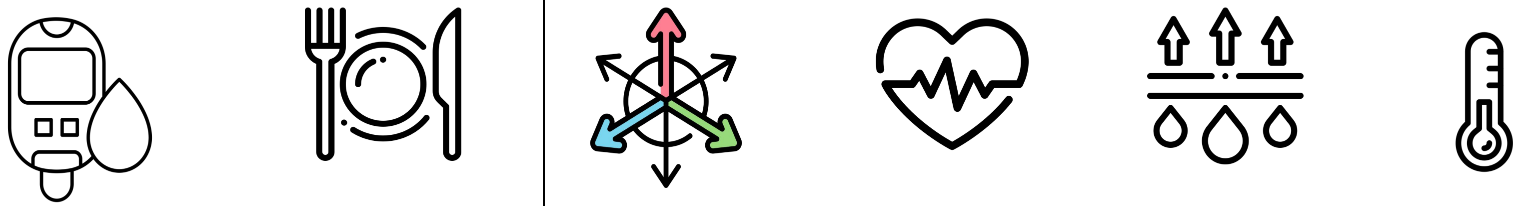
1. **High glucose levels** have negative long-term health effects.

2. **Predicting adverse events** such as high glucose levels can help people with prediabetes **protect their health**.



Dataset

Smartwatch



Glucose | Food Log | Movement | Heart Rate Measures | Sweat | Skin Temp

Duke BIG IDEAS LAB

1. Cho, P., Kim, J., Bent, B., & Dunn, J. (2023). BIG IDEAs Lab Glycemic Variability and Wearable Device Data (version 1.1.2). *PhysioNet*. https://doi.org/10.13026/zthx-5212
2. Bent, B., Cho, P.J., Henriquez, M. et al. Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. npj Digit. Med. 4, 89 (2021). https://doi.org/10.1038/s41746-021-00465-w

Data Prep: Accelerometer

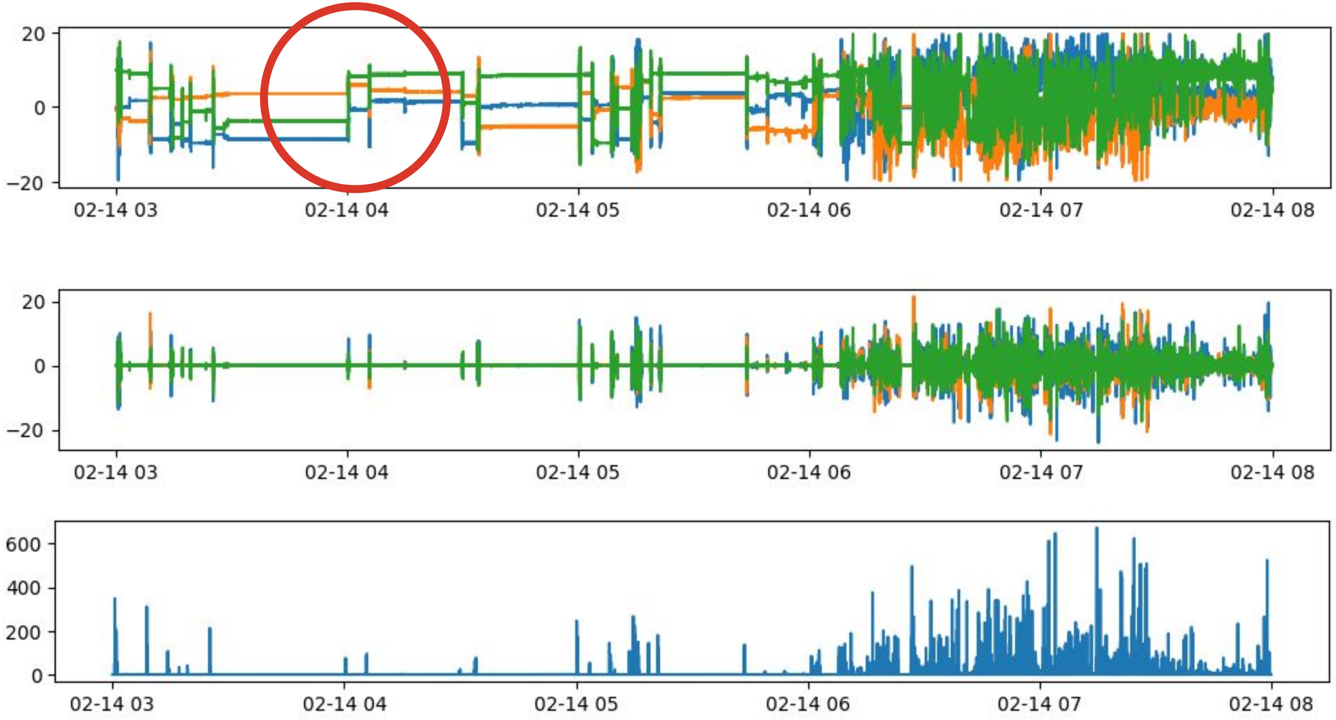
Gravity + Activity (Vector)

High-Pass Filter

Activity (Vector)

Magnitude

Activity (Scalar)



Data Augmentation: Glycemic Index (GI)

- GI quantifies how much a unit of carbohydrates in the given food raises the blood sugar.
- High GI means carbohydrates in that food are quickly absorbed → rapid rise in blood glucose.

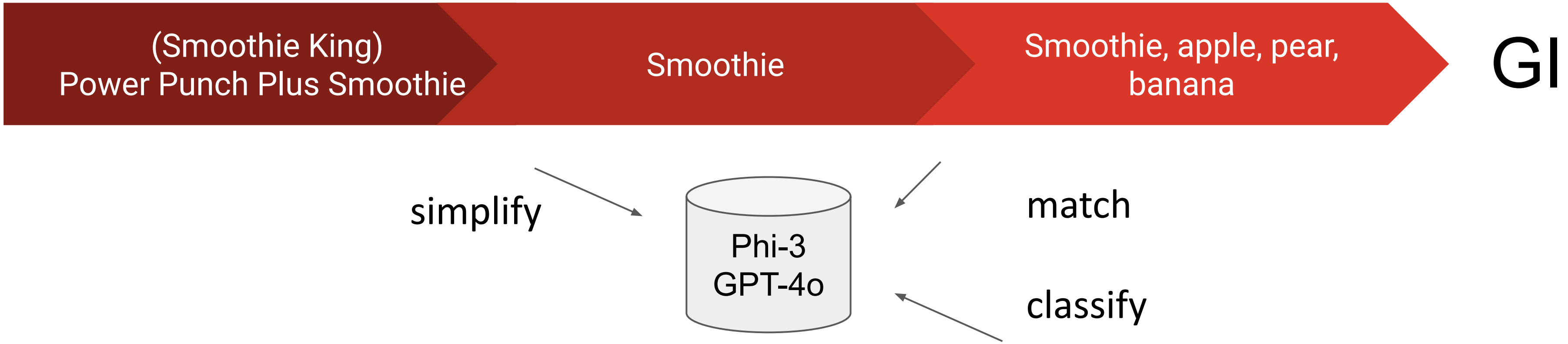
THE UNIVERSITY OF SYDNEY

Our dataset has a food log (e.g., “mashed potato”, “chicken pot pie”)
→ scrape and incorporate external GI database (University of Sydney)

Image and data source: https://glycemicindex.com/

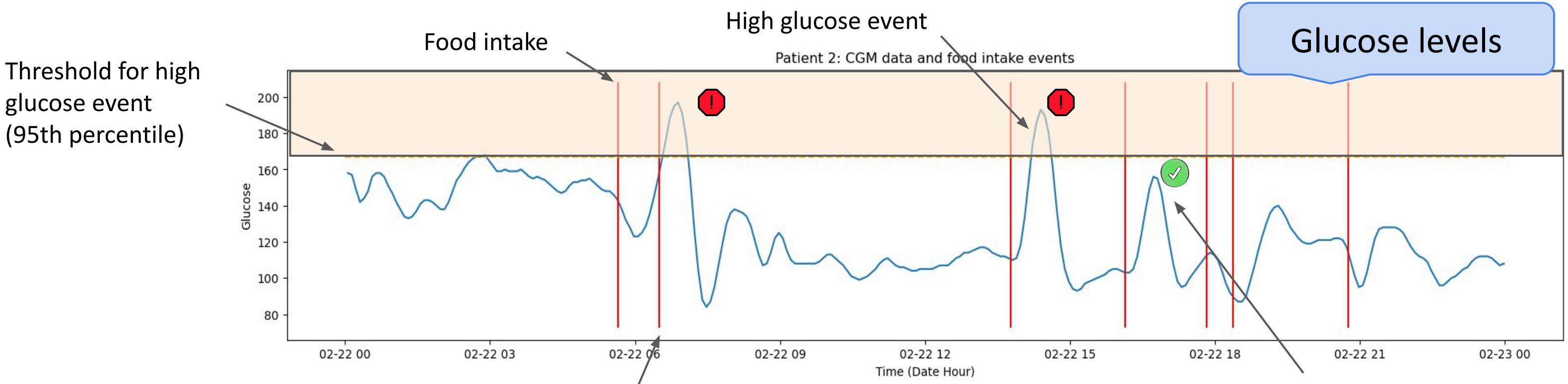
Matching GI Data with LMs

We used SLM/LLM (Phi-3/GPT-4o) to match food description to the GI database:



For foods not found, we used it to classify:

High glucose events

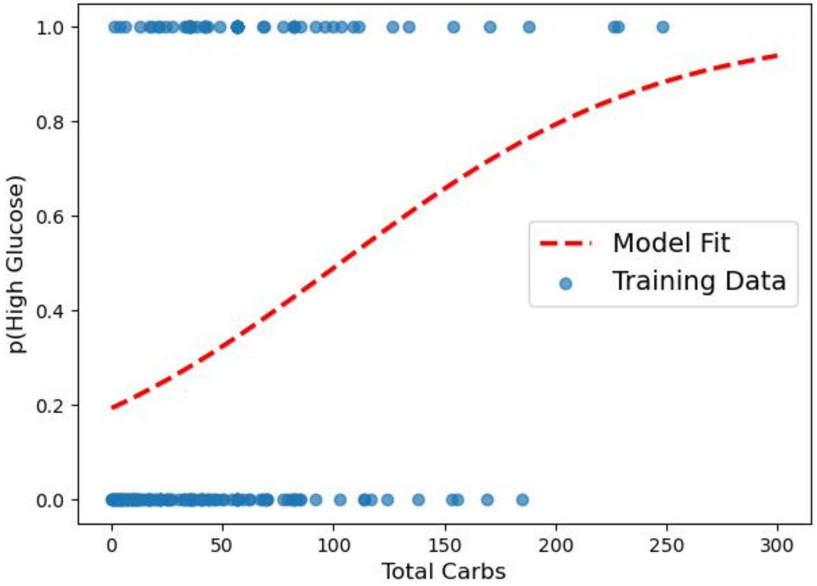


Food data: Calories, Carbohydrates, Sugar, Protein, Fat, Name of Food
Computed: Glycemic Index

Using carbohydrates to predict high glucose level

Total Carbohydrates → Model → High Glucose?

N = 13



Discarded meal data:
- Intervals too close
- Missing glucose data
- Unreasonably large intake

Sensitivity: .3279
Specificity: .8374

Glycemic index improves prediction

Total Carbohydrates

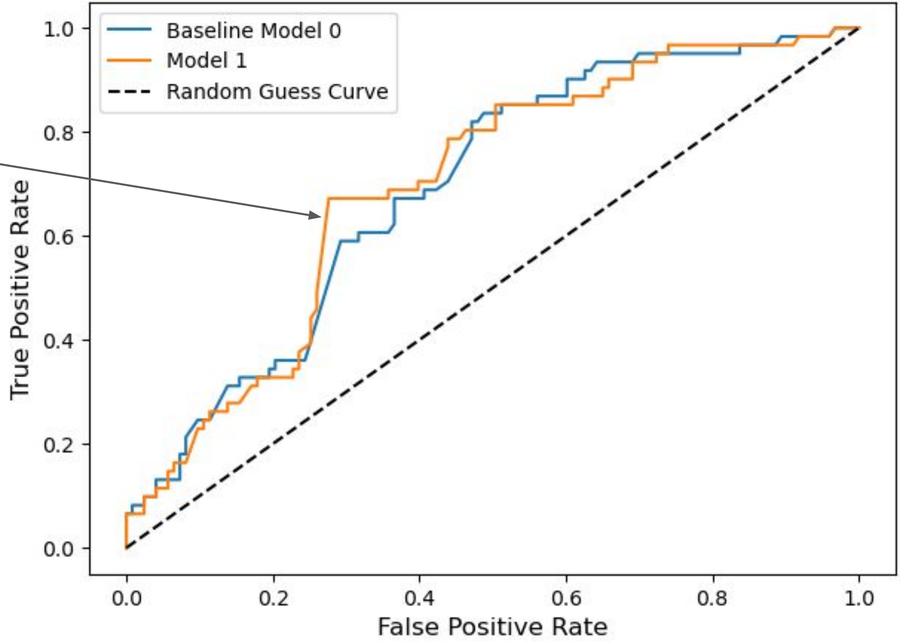
Glycemic Index

Model

High Glucose?

N = 13

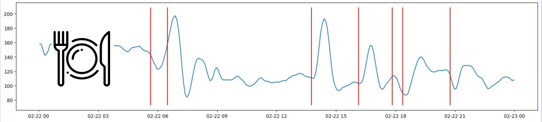
Increase in accuracy for predicting high glucose



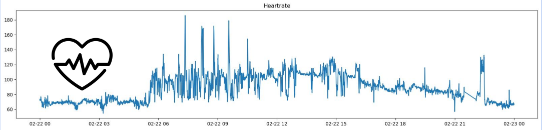
Sensitivity: .6721
Specificity: .7236

Predicting glucose levels over time

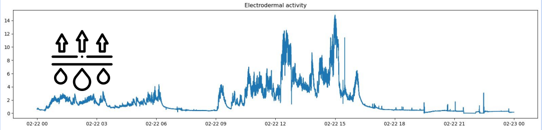
Inputs: Time-Dependent



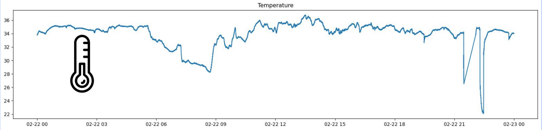
Food data



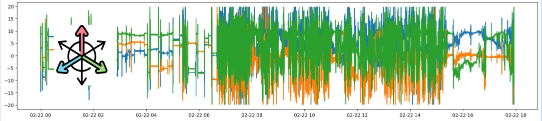
Heart rate



Skin moisture



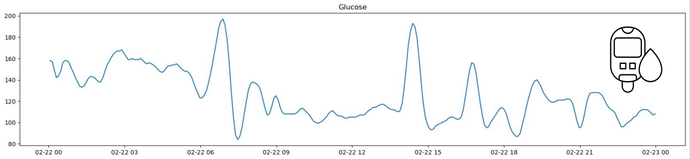
Temperature



3D Movement

Our model

Output: Time Series Prediction



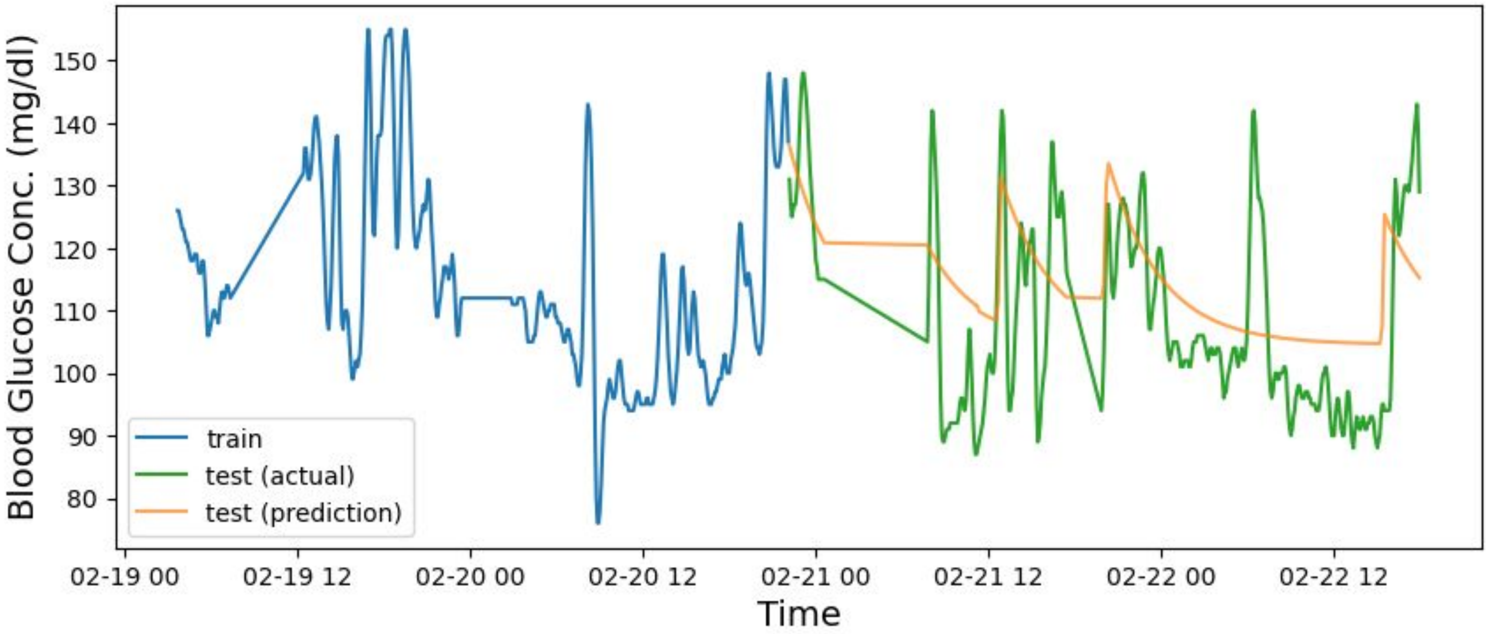
Glucose

Autoregressive Distributed Lag (ADL)

- ADL functions like an autoregressive (AR) model, but incorporates information from other time series into its predictions
- Our ADL model uses the smartwatch and food consumption data to predict a glucose reading
- Model was fit to each participant
- Hyperparameter optimization was used to find the best model for each participant
    - Hyperparameters optimized: Predictor set, Glucose lags, Predictor lags

ADL: GI Only

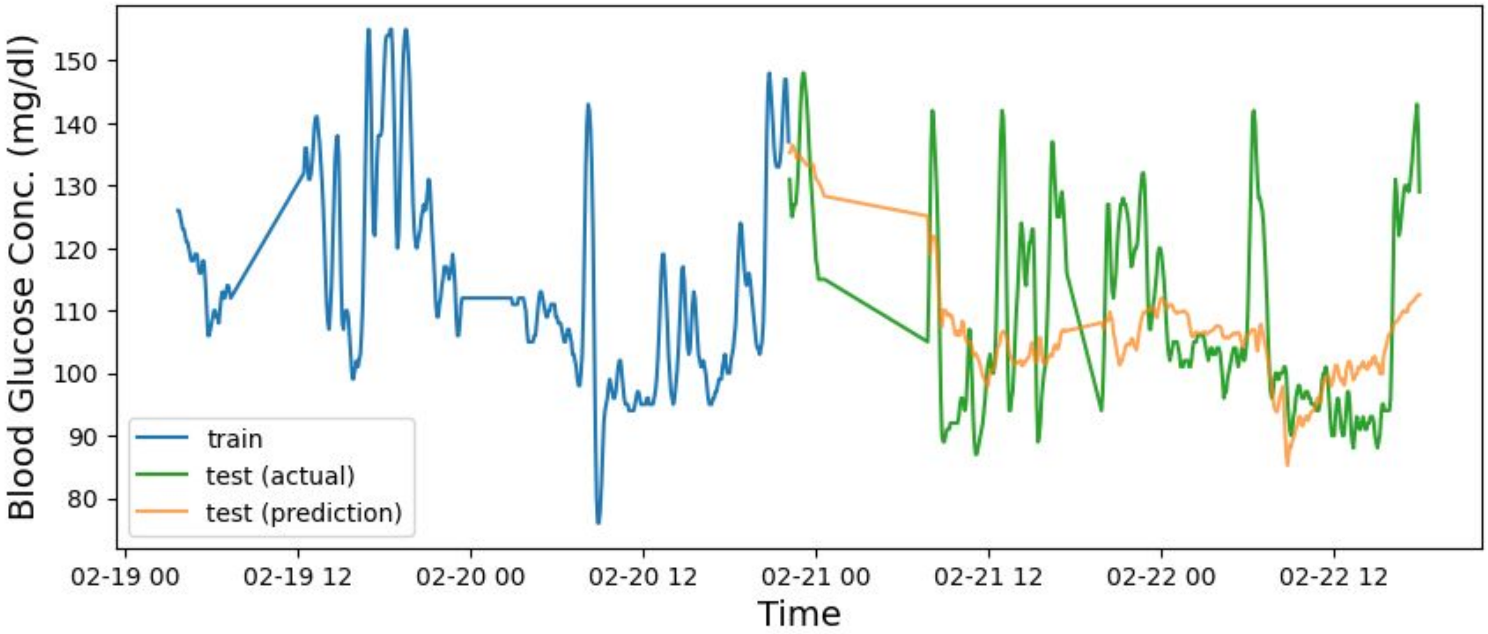
CV RMSE: 15.4 (mg/dL)



Participant 1, Glucose Lag 1, Predictor Lag 3

ADL: Smartwatch Only

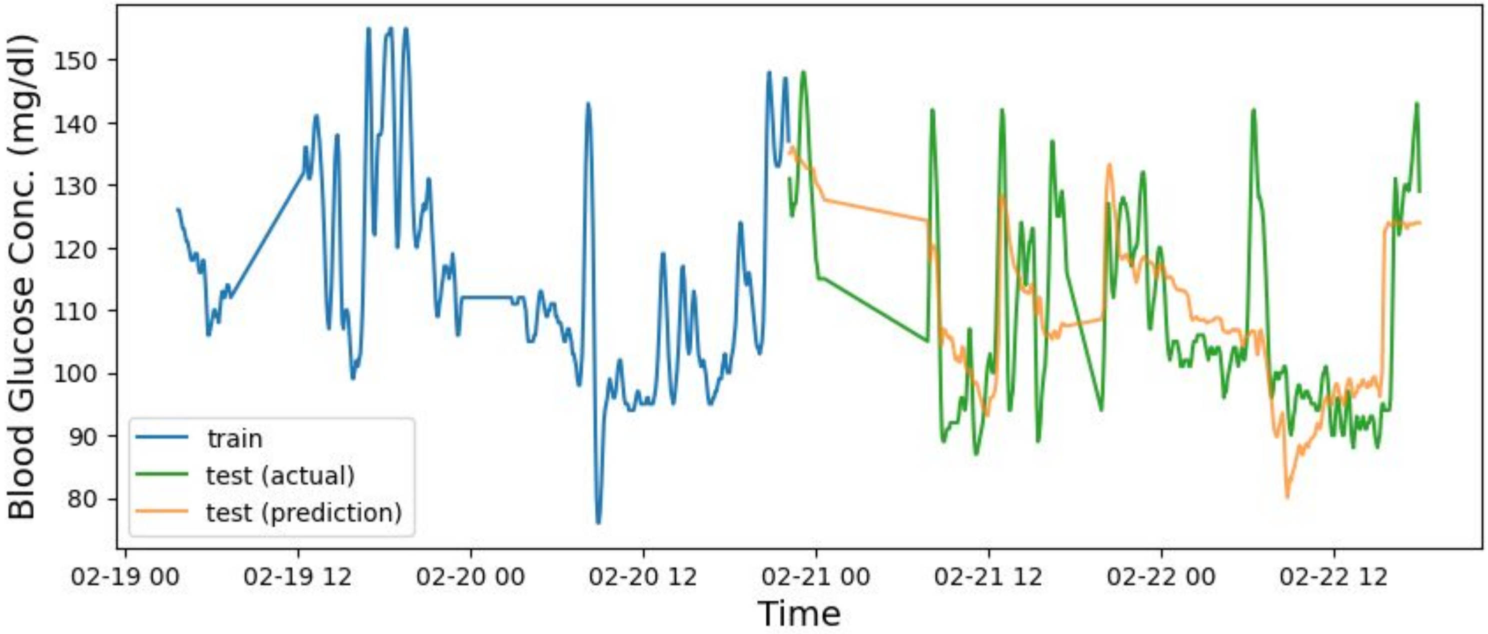
CV RMSE: 15.5 (mg/dL)



Participant 1, Glucose Lag 1, Predictor Lag 3

ADL: GI + Smartwatch

CV RMSE: 14.6 (mg/dL)



Participant 1, Glucose Lag 1, Predictor Lag 3

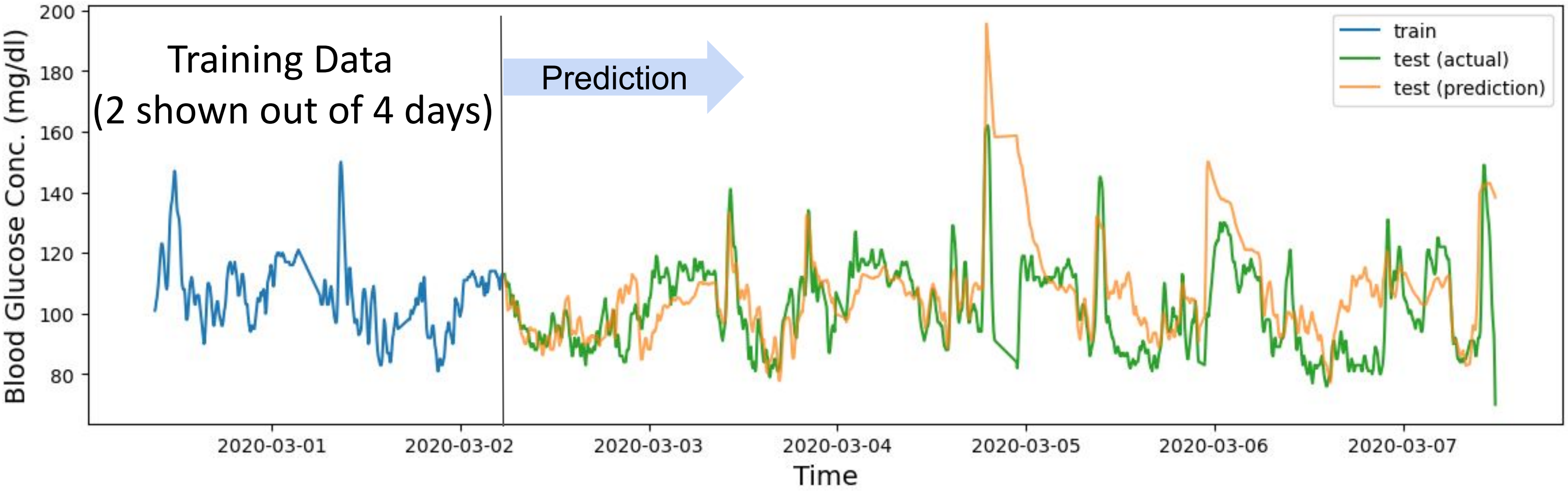
Results

- Good fit for participants with reliably logged biometric data
- Participants with highly varying glucose data had worse fits
- Glycemic index was present in 8 of 13 best models

Participant #	Best Model RMSE Glucose Conc.(mg/dL)
1	14.19
2	19.09
3	19.39
4	18.75
5	10.62
6	22.76
8	16.54
9	22.87
10	23.20
11	19.98
12	16.34
14	20.75
16	16.82

Long Term Prediction

After training on 4 days of data, this model is capable of predicting the glucose readings for 5 days, given only smartwatch and food data for the same period.



Participant 5, Glucose Lag 1, Predictor Lag 6

Future Work

- Developing a mobile application that provides real-time alerts and personalized dietary recommendations.
- Improve data cleanup and alignment process.
- Incorporating additional data such as inter-beat interval.
- Investigating the feasibility of predicting the time to reach/recover from the peak glucose level after a meal.

Acknowledgements

- Gleb Zhelezov
- Steven Gubkin
- Alec Clott
- Roman Holowinsky

The Erdos Institute

Image source: https://www.erdosinstitute.org