Research Proposal

Background

Diabetes mellitus is a metabolic disorder affecting more than 24 million people in the United States (1). This disorder ultimately manifests as chronic hyperglycemia and a number of potential comorbidities, including depression, cardiovascular disease, and cancer (2–5). Continuous glucose monitoring technology provides real-time measurement of glucose and is a powerful tool to help manage glucose levels and improve outcomes for people with diabetes (6–8). Individuals wearing continuous glucose monitoring technology can be alerted if they are approaching hyperglycemic (blood glucose $>$180 mg/dL) or hypoglycemic (blood glucose $<$ 70 mg/dL) thresholds or if the rate of change in blood glucose levels is fast (8).

Machine learning methods have been used with continuous glucose monitoring devices to produce models that can forecast potential future adverse glycemic events and warn before those events occurs (9–13). With the advent of wearable activity tracker technology, there is an opportunity to improve the predictions of these models by incorporating data on activity levels(14,15). However, the majority of research to date has been done on continuous glucose monitoring in people with Type 1 Diabetes, in part because accurate forecasting of blood glucose levels requires a large amount of data and most of the publicly available datasets have been compiled from people with Type 1 Diabetes (16,17). Additionally, these datasets often contain a relatively small number of individuals, which makes it difficult to train models that can generalize to a larger population (18,19).

Blood glucose dynamics vary within individuals with both types of diabetes and make it difficult to predict blood glucose levels for some individuals (20). This heterogeneity is mediated by both modifiable and unmodifiable factors, such as behavioral health (*e.g.* smoking, diet, exercise), genetics, clinical characteristics and environmental exposures (21). We aim to produce more accurate forecasts of blood glucose levels in individuals with Type 2 Diabetes by coupling wearable activity monitor data with blood glucose levels. In this study, we will also evaluate patient-level factors that may limit the predictive accuracy of these models.

Study Design

This study will use data from individuals in the AI-READI dataset, including individuals with and without Type 2 Diabetes. Individuals without diabetes will be included as a comparison group to provide a baseline of what model predictive accuracy looks like for individuals who do not have to externally control their blood glucose levels.

We will be forecasting blood glucose levels using Recurrent Neural Nets (RNN) and evaluating how predictive accuracy decays as the forecasting horizon decreases, with the goal of accurately predicting adverse glycemic events in the future. To help improve the accuracy of these forecasts we will include both time-varying and non-time-varying covariates in the models. The time-varying covariates will include data captured by the wearable activity trackers: heart rate, physical activity, respiratory rate, stress, and sleep data. We will also include data from the home environment sensor: light levels, particulate matter, temperature, humidity, volatile organic compound, nitrogen oxide. Other non-time-varying covariates will include physical attributes (age, weight, and height), behavioral attributes (responses to the dietary questionnaire, substance use questionnaire, and the diabetes survey), general health (responses from the general health questionnaire).

To evaluate model performance, we will evaluate model accuracy over progressively longer forecasting horizons starting at 15 minutes and increasing up to 8 hours. We will forecast both binary outcomes (hyperglycemic events and hypoglycemic events) as well as continuous outcomes (blood glucose concentrations). Model performance will be evaluated with standard evaluation metrics for both continuous and binary outcomes. Finally, we will evaluate model performance among different groups of individuals within the data as a way to generate additional hypotheses about what factors may contribute to the difficult in forecasting blood glucose levels. These factors will include disease severity, sex, and social determinants of health, such as education and health insurance coverage.

References

1. Surveillance - United States Diabetes Surveillance System. https://gis.cdc.gov/grasp/diabetes/diabetesatlas-surveillance.html

2. Zelenko Z, Gallagher EJ. Diabetes and cancer. *Endocrinol Metab Clin North Am.* 2014;43(1):167-185. doi:10.1016/j.ecl.2013.09.008

3. Wang M, Yang Y, Liao Z. Diabetes and cancer: Epidemiological and biological links. *World Journal of Diabetes.* 2020;11(6):227-238. doi:10.4239/wjd.v11.i6.227

4. Liu L, Simon B, Shi J, Mallhi AK, Eisen HJ. Impact of diabetes mellitus on risk of cardiovascular disease and all-cause mortality: Evidence on health outcomes and antidiabetic treatment in united states adults. *World Journal of Diabetes*. 2016;7(18):449-461. doi:10.4239/wjd.v7.i18.449

5. Shaya FT, Chirikov VV, Howard D, et al. Effect of social networks intervention in type 2 diabetes: A partial randomised study. doi:10.1136/jech

6. Dupre ME, Silberberg M, Willis JM, Feinglos MN. Education, glucose control, and mortality risks among u.s. Older adults with diabetes. *Diabetes Res Clin Pract*. 2015;107(3):392-399. doi:10.1016/j.diabres.2014.12.013

7. V ENaveen, A J, V E, Jerith GG, Reeves SJ, Solomon I. 2023 international conference on self sustainable artificial intelligence systems (ICSSAS). In:; 2023:1248-1253. doi:10.1109/ICSSAS57918.2023.10331808

8. DeSalvo D, Buckingham B. Continuous glucose monitoring: current use and future directions. *Curr Diab Rep*. 2013;13(5):657-662. doi:10.1007/s11892-013-0398-4

9. Ali H, Niazi IK, White D, Akhter MN, Madanian S. Comparison of Machine Learning Models for Predicting Interstitial Glucose Using Smart Watch and Food Log. *Electronics*. 2024;13(16):3192. doi:10.3390/electronics13163192

10. Arefeen A, Ghasemzadeh H. 2023 IEEE EMBS international conference on biomedical and health informatics (BHI). In:; 2023:1-5. doi:10.1109/BHI58575.2023.10313443

11. Cichosz SL, Jensen MH, Hejlesen O. Short-term prediction of future continuous glucose monitoring readings in type 1 diabetes: Development and validation of a neural network regression model. *Int J Med Inform*. 2021;151:104472. doi:10.1016/j.ijmedinf.2021.104472

12. Dave D, Vyas K, Branan K, et al. Detection of Hypoglycemia and Hyperglycemia Using Noninvasive Wearable Sensors: Electrocardiograms and Accelerometry. *J Diabetes Sci Technol*. 2024;18(2):351-362. doi:10.1177/19322968221116393

13. Duckworth C, Guy MJ, Kumaran A, et al. Explainable machine learning for real-time hypoglycemia and hyperglycemia prediction and personalized control recommendations. *J Diabetes Sci Technol*. 2022;18(1):113-123. doi:10.1177/19322968221103561

14. Campanella S, Palma L. Sensor-Based Monitoring of Physical Activity for Glucose Management in Diabetic Patients: A Review. In: Valle M, Gastaldo P, Limiti E, eds. Springer Nature Switzerland; 2025:177-188. doi:10.1007/978-3-031-71518-1_20

15. Ahmed A, Aziz S, Abd-alrazaq A, Farooq F, Househ M, Sheikh J. The Effectiveness of Wearable Devices Using Artificial Intelligence for Blood Glucose Level Forecasting or Prediction: Systematic Review. *J Med Internet Res*. 2023;25(1):e40259. doi:10.2196/40259

16. Sun Y, Kosmas P. Integrating Bayesian Approaches and Expert Knowledge for Forecasting Continuous Glucose Monitoring Values in Type 2 Diabetes Mellitus. doi:10.48550/arXiv.2409.07315

17. Sergazinov R, Chun E, Rogovchenko V, Fernandes N, Kasman N, Gaynanova I. GlucoBench: Curated list of continuous glucose monitoring datasets with prediction benchmarks. doi:10.48550/arXiv.2410.05780

18. Fiorini S, Martini C, Malpassi D, et al. 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). In:; 2017:1680-1683. doi:10.1109/EMBC.2017.8037164

19. Sergazinov R, Chun E, Rogovchenko V, Fernandes N, Kasman N, Gaynanova I. GlucoBench: Curated list of continuous glucose monitoring datasets with prediction benchmarks. doi:10.48550/arXiv.2410.05780

20. Cichosz SL, Hejlesen O, Jensen MH. Identification of individuals with diabetes who are eligible for continuous glucose monitoring forecasting. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*. 2024;18(2):102972. doi:10.1016/j.dsx.2024.102972

21. Redondo MJ, Hagopian WA, Oram R, et al. The clinical consequences of heterogeneity within and between different diabetes types. *Diabetologia*. 2020;63(10):2040-2048. doi:10.1007/s00125-020-05211-7