The Journal of Open Source Software

GluPredKit: A Python Package for Blood Glucose Prediction and Evaluation

Miriam Kopperstad Wolff [1], **Sam Royston** [2], **and Rune Volden** [1]

1 Norwegian University of Science and Technology, Norway **2** Replica Health, United States of America

DOI: 10.21105/joss.06904

Software

- Review
- Repository
- Archive

Editor: Julia Romanowska

Reviewers:

- @panagiotisanagnostou
- @JonnyTran

Submitted: 06 June 2024
Published: 26 September 2024

Summary

Managing blood glucose levels is crucial for individuals with diabetes. Historically, non-linear physiological modeling of glucose dynamics laid the groundwork for automated insulin delivery. Blood glucose prediction can be used as decision support for patients or as a component in an automated insulin delivery control strategy. Today, machine learning and deep neural networks offer new pathways for improvement, and the literature is vast on proposed models. Yet, comparing these advanced models is challenging. Differences in the datasets used for testing and how results are evaluated can make comparisons from existing studies unreliable (Jacobs et al., 2023). Additionally, many research studies do not share their code, making it hard to build upon previous work. `GluPredKit` addresses these issues by standardizing the pipeline steps needed for any blood glucose prediction research (see Figure 1). This includes the collection, organization, and preparation of data, as well as the ability to easily compare different models and measure their effectiveness. Additionally, the software incorporates state-of-the-art components, including the ability to integrate and standardize data from various sources, utilize existing prediction models, and apply established evaluation metrics. It also features automated generation of detailed model evaluation reports, guided by the consensus on blood glucose model evaluation (Jacobs et al., 2023).

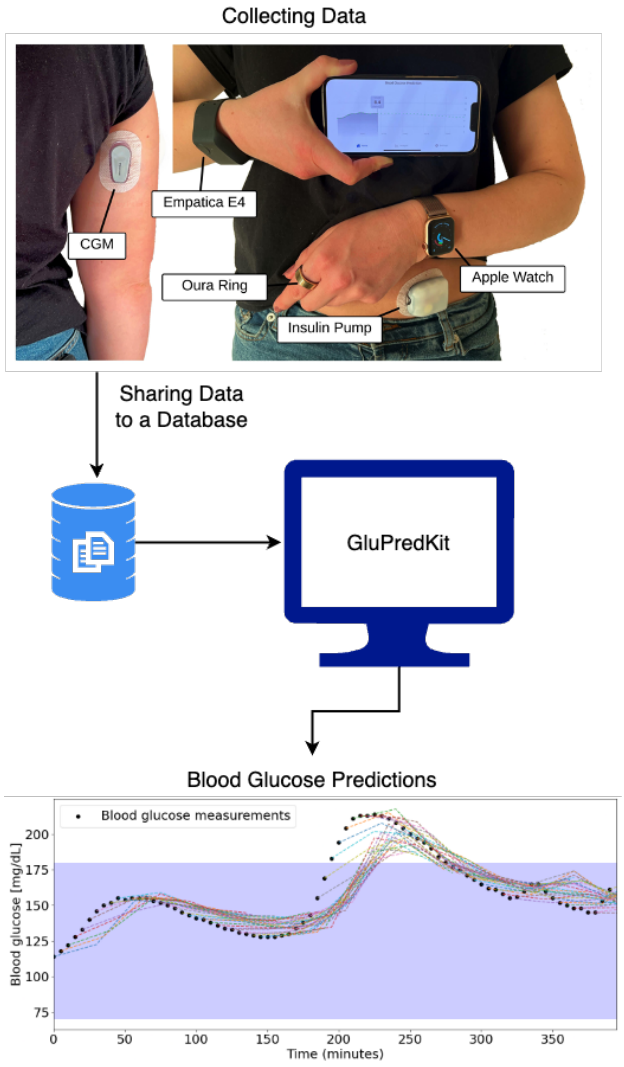

Figure 1: High-level visualization of the GluPredKit ecosystem.

Statement of need

`GluPredKit` is a Python package designed to streamline the process of blood glucose prediction, accessible via both a Command Line Interface (CLI) and as a library on PyPi. There is a need for standardized evaluation guidelines to leverage the potential of artificial intelligence in enhancing glycemic control for diabetes management (Jacobs et al., 2023). Research indicates that modern deep learning models can provide superior predictions compared to traditional complex non-linear physiological models (Cappon, Prendin, et al., 2023).

`GluPredKit` addresses this need by facilitating the evaluation of individual or multiple models. Its modular design and standardized approaches facilitate community contributions, enabling researchers to integrate new models, metrics, or data sources while maintaining compatibility with existing components in a plug-and-play framework. The software includes state-of-the-art features such as integration with common data sources, ready-to-use white- and black-box models, and evaluation metrics.

Despite the vast literature on proposed blood glucose prediction algorithms and benchmarking studies (Xie & Wang, 2020), (Meijner & Persson, 2017), many do not provide open-source code or are not designed for scalability and integration with other models or data sources.

The GLYFE study represents the closest existing package to GluPredKit, as it benchmarks several models and shares its source code, allowing the addition of new components (Bois et al., 2022). However, `GluPredKit` differs from GLYFE in being more flexible in accommodating different dataset input features and hypothetical scenarios, and visualizations such as plots and predicted trajectories in addition to evaluation metrics. Researchers often need to generate these visualizations and metrics repeatedly for each experiment and research publications. Furthermore, `GluPredKit` is designed not only as a standalone package but also as a dependency that can be integrated into other software systems, unlike GLYFE's GUI-centric approach.

To ensure broad usability and scalability, `GluPredKit` consolidates prominent work in the field into a single repository with a scalable architecture that supports future community contributions. Its usability has been validated through user tests (Wolff et al., 2024). Additionally, Oh et al. utilized the platform in a master's thesis, incorporating existing models from the literature and validating them against reported results (Oh, 2024). Integrated test datasets include the Ohio dataset (Marling & Bunescu, 2020), Tidepool API, and Nightscout API. The software implements physiological models based on Uva Padova, using implementations from ReplayBG (Cappon, Vettoretti, et al., 2023) and PyLoopKit, both of which are open-source, in contrast to proprietary models in commercial systems. Moreover, off-the-shelf models such as Ridge Regressor, Random Forest, and LSTM have been implemented, based on common blood glucose prediction model approaches in benchmarks (Xie & Wang, 2020), (Cappon, Prendin, et al., 2023). The complete and evolving list of components is documented in the `GluPredKit` documentation.

`GluPredKit` Workflow

The `GluPredKit` workflow is typically used through the CLI or as a dependency in external projects via PyPi. The first step involves parsing the input data to prepare it for processing. Users then configure the settings tailored to their specific needs before moving on to model training and testing phases. After testing, the user can generate an Excel sheet and a PDF report for a standardized evaluation report based on evaluation consensus guided by Jacobs et al. (Jacobs et al., 2023).

The software consists of four key modules: data source parsers, preprocessors, prediction models, and evaluation metrics. Detailed instructions and standardized code interfaces are provided in the repository's documentation, guiding contributors on how to add and integrate their modules.

Acknowledgements

This project is supported by the Department of ICT and Natural Sciences at the Norwegian University of Science and Technology. We acknowledge contributions from Martin Steinert, Hans Georg Schaathun, and Anders Lyngvi Fougner during the genesis of this project.

References

Bois, M. D., Yacoubi, M. A. E., & Ammi, M. (2022). GLYFE: Review and benchmark of personalized glucose predictive models in type 1 diabetes. *Medical & Biological Engineering & Computing*, *60*(1), 1–17. https://doi.org/10.1007/s11517-021-02437-4

Cappon, G., Prendin, F., Facchinetti, A., Sparacino, G., & Del Favero, S. (2023). Individualized models for glucose prediction in type 1 diabetes: Comparing black-box approaches to a physiological white-box one. *IEEE Transactions on Biomedical Engineering*, *70*(11), 3105–3115. https://doi.org/10.1109/TBME.2023.3276193

Cappon, G., Vettoretti, M., Sparacino, G., Favero, S. D., & Facchinetti, A. (2023). ReplayBG: A digital twin-based methodology to identify a personalized model from type 1 diabetes data and simulate glucose concentrations to assess alternative therapies. *IEEE Transactions on Biomedical Engineering*, *70*(11), 3227–3238. https://doi.org/10.1109/TBME.2023.3286856

Jacobs, P. G., Herrero, P., Facchinetti, A., Vehi, J., Kovatchev, B., Breton, M., Cinar, A., Nikita, K., Doyle, F., Bondia, J., Battelino, T., Castle, J. R., Zarkogianni, K., Narayan, R., & Mosquera-Lopez, C. (2023). Artificial intelligence and machine learning for improving glycemic control in diabetes: Best practices, pitfalls and opportunities. *IEEE Reviews in Biomedical Engineering*, *PP*. https://doi.org/10.1109/RBME.2023.3331297

Marling, C., & Bunescu, R. (2020). The OhioT1DM dataset for blood glucose level prediction: Update 2020. *CEUR Workshop Proceedings*, *2675*, 71–74. http://ceur-ws.org/Vol-2675/paper11.pdf

Meijner, C., & Persson, S. (2017). *Blood glucose prediction for type 1 diabetes using machine learning: Long short-term memory based models for blood glucose prediction* [Master's thesis]. Chalmers University of Technology; University of Gothenburg.

Oh, D. (2024). *Improving blood glucose prediction for people with T1DM during physical activity using machine learning on participant collected data* [Master's thesis, UiT The Arctic University of Norway]. https://hdl.handle.net/10037/33803

Wolff, M. K., Royston, S., Fougner, A. L., Schaathun, H. G., Steinert, M., & Volden, R. (2024). *GluPredKit: Development and user evaluation of a standardization software for blood glucose prediction*. https://arxiv.org/abs/2406.08915

Xie, J., & Wang, Q. (2020). Benchmarking machine learning algorithms on blood glucose prediction for type i diabetes in comparison with classical time-series models. *IEEE Transactions on Biomedical Engineering*, *67*(11), 3101–3124. https://doi.org/10.1109/TBME.2020.2975959