

# The Blood Glucose Level Prediction Challenge Rules

## Overview

The OhioT1DM Dataset contains 8 weeks worth of data for each of 12 people with type 1 diabetes. The last 10 days' worth of data for each person was held out for testing. The 6 anonymous individuals who contributed data for the second BGLP Challenge were identified as data contributors 540, 544, 552, 567, 584, and 596. These are the contributors for whom results are to be reported. Data for people who contributed to the first BGLP Challenge may be used for pre-training models, if desired. Those individuals were numbered 559, 563, 570, 575, 588, and 591.

**Results should be reported for 30 and 60-minute prediction horizons.**

**Results may be submitted for offline models, online models, or both.** An offline model trains and tunes just one model per patient on the provided training data. At test time, the trained model does not change, and the same model is used to compute predictions for all test points. An online model, on the other hand, can use blood glucose data up to and including data at time  $t-30$  or  $t-60$  to make predictions for each test point  $t$ . Please indicate which type of model(s) you used to obtain your results.

**The set of evaluation points should begin 60 minutes after the start of each 10-day test dataset.** This is important in the offline setting, because the test dataset immediately follows the training dataset, and the initial test points would otherwise be very close chronologically to the training data and thus label-correlated to it. Although not pertinent in the online setting, we will use the same evaluation points in both settings, for consistency.

**Participants should ensure that model predictions are not contaminated by using data from "the future."** For example, this entails that interpolation can not be used to deal with missing data; rather, extrapolation could be used.

**At least one author of each accepted paper must attend the BGLP Challenge in Santiago de Compostela, Spain.** No remote presentations will be allowed.

**To facilitate future work and replication of published results, we require that participants make the code publicly available,** by including a link in the camera ready paper. Participants also need to document the training, tuning, and evaluation procedures that need to be used to replicate the results reported in the camera ready paper. The code will also be linked from the KDH website immediately following the workshop.

**BGLP Challenge participants should submit: (1) Results; (2) System Description Papers; and, for accepted system description papers, (3) Camera Ready Copy.**

## How to Report Results

The first stage of the BGLP Challenge is to report the prediction results for each of the 6 data contributors 540, 544, 552, 567, 584, and 596. The test portion of the dataset contains 2884, 2704, 2352, 2377, 2653, and 2731 test points for these data contributors, respectively. Results should be reported for 30 and 60-minute prediction horizons, as follows:

1. The RMSE for each of the 6 people with diabetes. This will result in 6 RMSE numbers, one for each data contributor, where the RMSE for each person is computed over all of the test points for that person.
2. The mean RMSE over all 6 people, i.e., the average of the 6 RMSE numbers above.
3. The MAE for each of the 6 people.
4. The mean MAE over all 6 people.

Participants using models with performance that depends on random initialization (e.g. neural networks) are encouraged to run multiple evaluations and report the mean and standard deviation for each person's RMSE and MAE.

Please note that participants are encouraged to include the results of any other type of evaluation they deem insightful in their system description papers.

The foregoing results should be reported in a PDF file and uploaded to the BGLP Challenge Results track in [EasyChair](#).

Additionally, the raw data predictions should also be submitted at this time. For each data contributor <dc\_ID>, at each prediction horizon <horizon>, the prediction results should be submitted in a text file <system\_ID> <dc\_ID> <horizon>.txt, where each line contains the timestamp, true blood glucose level (BGL), and predicted BGL, separated by white spaces, in chronological order:

```
<timestamp_1> <true_BGL_1> <predicted_BGL_1>
<timestamp_2> <true_BGL_2> <predicted_BGL_2>
...
```

Specific instructions for uploading these files will be provided as the date approaches.

## Guidelines for Preparing a System Description Paper

The next step for those who have already reported their results is to prepare a system description paper. The guidelines are as follows:

**Results:** Present the originally submitted results. The results of additional evaluations that lend insight may also be presented. Examples include (but are not limited to) results:

- for clinically relevant metrics, such as Clark Error Grid (CEG) analysis or glucose specific RMSE based on CEG (gRMSE);
- for longer (e.g., 90 or 120 minute) prediction horizons;
- on other training vs. test splits of the data;
- on subsets of physiological parameters;
- for predicting hypoglycemia directly through training a classification model instead of regression; or
- for studying the effect of data augmentation techniques, such as pre-training on data from other data contributors, or exploiting incomplete data (e.g., blood glucose levels without meals) from other data contributors.

**Methodology:** Explain in detail the training, tuning, and testing phases of the system used to produce the submitted results. This is especially important for determining whether the submitted results are comparable to the results from other systems.

**Replicability:** Present all details that will allow someone else to replicate your system.

**Analysis:** Focus on results and analysis. Some possibilities are: reporting results on several runs of the system; presenting ablation experiments showing the usefulness of different features and/or techniques; and showing comparisons with baselines.

**Avoid Duplication:** Cite the paper, "The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020." Then you can avoid repeating details of the task and data. However, briefly outlining the task and relevant aspects of the data is a good idea. You may use a "place holder" citation until the official BibTeX citation becomes available for the camera ready copy.

The system description paper should be formatted as per the [Call for Papers](#) and submitted to the BGLP Challenge track in [EasyChair](#).

## Guidelines for Preparing the Camera Ready Copy

The next step for those whose system description papers have been accepted is to prepare the camera ready copy. The guidelines are as follows:

**Results:** Present the originally submitted results. If you have gotten better results since your original submission, you may present them in addition. In this case, be sure to explain how the improvement was obtained.

**Feedback:** Incorporate the feedback of the anonymous reviewers into your camera ready copy.

**Length:** The papers will be published as a CEUR proceedings. CEUR requires papers to be at least 5 pages long. The length specified in our Call for Papers is 4 pages plus 1 page for references. Please ensure that your paper goes onto the 5th page! It will be fine if there is content on the 5th page in addition to the references.

The camera ready copy should be uploaded to [EasyChair](#).