Blood Glucose Prediction in Wearable Devices *

Andrew Allman

Mandar Kapadia

Dept. of Computer Science
University of Massachusetts
Lowell, MA 01854, USA
Andrew_Allman@student.uml.edu
Mandar Kapadia
Dept. of Computer Science
University of Massachusetts
Lowell, MA 01854, USA
Mandar_Kapadia@student.uml.edu

***Abstract*— About 86 million people with pre-diabetes 90% of them are undiagnosed[1]. Pre-diabetes can lead to health consequences with elevated fasting glucose levels. To help solve this problem our project will develop a predictive artificial intelligence using machine learning to help predict Blood Glucose levels using wearable devices.**

Keywords— pre-diabetes, blood glucose, artificial intelligence, machine learning, predictive model, LSTM(long short-term memory networks),RNN(Recurrent neural network)

I. INTRODUCTION

Wearable devices currently provide useful monitoring to individuals managing diabetes - the standard being Continuous Glucose Monitoring (CGM). Contemporary research has extended this aim into broad collection of bio-markers, attributable to estimating risk of developing pre-diabetes.

As noted in Dataset, continuous measurements of relevant bio-markers were captured over the course of 8-10 days, along with a CGM wearable device, used for capturing risk index of patients.

The aim of model is to explore paths for classification of at-risk patients, based on collected bio-marker information - and further, identify significant relationships between non-obvious features and risk-index.

II. DATASET

The selected dataset for this project is "BIG IDEAs Lab Glycemic Variability and Wearable Device Data" from physionet.org, which can be found with the link:"https://physionet.org/content/big-ideas-glycemic-wearable/1.1.2/"

Form this dataset, we use "tri-axial accelerometry"," blood volume pulse", "blood volume pulse", "Gender","skin temperature" and other column will be used in to prediction the correct blood glucose levels.

III. METHODOLOGIES

Preparation and development were done in an ordinary fashion. Complete data records were reduced to five total participants, and within all patients, all but one biomarker (HR) was considered in the creation of each participant's composite dataset. Instead of considering all data points therein, we captured the averages and standard deviations of all considered biomarkers between blood glucose readings, capturing statistically meaningful relationships, while reducing the storage requirements and eventual technological demand of the data. This correspondence between data averages and standard deviations and classification of blood-glucose levels became our feature representations and classification tags. To be covered more in respective length, our data preparation and model development existed in the exploratory for some time.

This exploration led us, initially to classification of data using Naive Bayes, and Logistic Regression, to the state-of-the-art Keras LSTM model (for establishing correlation between biomarker data and blood-glucose classification / baseline scores), then to an unsuccessful attempt to develop our own LSTM model, and finally our current state Recurrent Neural Net (RNN) of lesser sophistication. Early considerations for the right model for classification in time-series health data, informed the decision to attempt implementation of a Long Short Term Memory model. One of the reasons this model seemed most fit for this task, is its ability to consider long term connections between data points through time. And, with our group having little familiarity with the biological elements of the project (and whether classification could be done as effectively with snapshot information as with long-term consideration between datapoints), an LSTM was pursued in the case that the assertion was incorrect.

A. EXPERIMENTAL EVALUATION

Naive-Bayes Classification

Our first attempt at classification of the data was performed using Naive-Bayes, a simple classification

method that calculates the probability of a given class given a set of features. Due to the relatively high spread of accuracies, we received in trialing this method (and eventually poor performance in some of the other models) our group found one of the overall highest degrees of success in classifying all three classes with this method.

Logistic Regression

Our second evaluation was performed using Logistic Regression, a method that would seem more fitting to attributing class predictions given a set of independent variables. However, our model's performance was without value in this case.

Keras Sequential

For the sake of proving positive correlation, feasibility of classification, and establishing baseline classification performance, our group implemented a network using Keras sequential model connecting a high-performance, pre-built LSTM layer, connected to a Dense output layer for classification. Our results were very good for all classifications measured.

Recurrent Neural Network

As mentioned previously, our attempts to implement a functional LSTM of our own did not surpass development. Instead, a less sophisticated RNN, involving activation, forward and backward propagation, and evaluation was created in its stead. Initially, performance seemed quite poor, with results similar to Logistic Regression that yielded high overall classification accuracy, with little accuracy for marginal examples (Low/High classification). However, after modifying the training data from a representative set to a more balanced data set, and adjusting the hyperparameter to a unintuitive setup, we started to see fair performance that fell short of only the Kera pre-built model on classifying testing samples.

B. ALGORITHMS / IMPLEMENTATION

Data-Preparation

Data preparation was performed in a manner of identifying the times associated with each blood-glucose reading, and collecting the averages and standard deviations of biomarkers recorded between two sequential timestamps.

Data-Loading

Datasets were parsed and split into feature vectors, and tag values. The feature vectors were created from parsing rows of the CSV datasets, splitting. The tags were generated by simple comparison of blood-glucose values against threshold values of either Low (<70mg/dL), Healthy (between 70 and 140 mg/dL) and High (>140mg/dL).

RNN

RNN forward and backward propagation was performed in rather basic stature, with 4 variables pertaining to weights and biases of the hidden layers, additional variables for cache maintenance, chain rule derivation for updating previous layers, and softmax activation for generating class probabilities (forward output). Evaluation captured cross-entropy loss, and accuracy in overall classification, high-glucose classification, and low-glucose classification. Training was performed with a variety of different hyper-parameter settings, with the best results identified with:

Learning Rate: 1e-3
Hidden Units: 5000
Batch Size: 8

using a more balanced dataset (>50% healthy samples, >35% high glucose samples, <15% low glucose samples).

C. RESULTS AND DISCCUSION

Naive-Bayes

As discussed earlier, performance with Naive-Bayes was fairly good due to relatively high-spread of accuracies. Over the course of 25 iterations of testing randomly shuffled samples, we saw an average classification of:

- Overall: 73%
- High: 47%
- Low: 44%

It can't be discounted that, with a higher overall accuracy with the Naive-Bayes, we likely would have seen a lower accuracy predicting high or low.

Logistic Regression

With logistic regression classification, again we saw high-overall accuracy of classification that was likely due to an over-tendency to classify a sample as healthy. Over twenty-five iterations with randomly shuffled samples we saw:

- Overall: 90%
- High: <1%
- Low: 7%

Keras Sequential

The Keras model provided the most reassurance that strong correlation exists between these biomarkers and Blood-Glucose classifications. Training for one thousand epochs, and validating against new samples yielded accuracies of:

- Overall: 97%
- High: 87%
- Low: 71%

RNN

With confirmation of strong correlation between our feature data and tag values, expectation was again to see positive classification of all three ranges of Blood-Glucose. This was (initially) not the case – the results from our deep learning model (on a representative dataset) looked most similar to the Logistic Regression classifier, in that with high overall accuracies, we neared classification performance for non-healthy scores.

- Overall: 89%
- High: <1%
- Low: <1%

Regardless of hyperparameter tunings, it seemed our model couldn't maintain an ability to accurately classify healthy scores in tandem to non-healthy scores, and improvement in either of the two major camps of accuracies saw a near reciprocal decline in the other.

After adjusting the samples, to a more balanced dataset, and expanding the number of hidden units, and increasing the learning rate. We found the fair classification accuracies, with the highest scores shown below:

- Overall: 83%
- High: 65%
- Low: 50%

CONCLUSION AND LIMITATIONS

Based on correlation analysis between our trialed biomarker data, and associated blood-glucose readings, we are confident in the usefulness of further research in the prediction of blood-glucose classification through the use of bio-marker data. The impediment to performance of our model, was largely in implementation, and in part development inexperience. It was demonstrable that given proper implementation, an LSTM network could find connections between the features and tags - and, it seems that simple methods such as Naive-Bayes could identify rough correlation as well.

With additional time, it would have been useful to continue development of the LSTM model we had been building. It also would have been useful to expand our dataset beyond 5 patients, providing more samples that resembled biomarkers associated with high-blood glucose. Although we found fair performance after changing our dataset composition, was lack of availability of anomalous blood-glucose measures was a constraint in performance.

ACKNOWLEDGEMENTS

Writing Code for Naive Bayes, Linear Regression, RNN, and LSTM models has been fairly difficult and complicated. To help us understand more codes we used various online resources like the medium geeks for geeks and references 7-11. Along with those references we also used the internet in a Chat GPT to understand general themes and understanding about each model. Learning from these online resources has helped. We understand so much more about the models we were implementing.

References.

1. Bent, B., Cho, P. J., Henriquez, M., Wittmann, A., Thacker, C., Feinglos, M., Crowley, M. J., & Dunn, J. P. (2021, June 2). Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. Nature News. Retrieved June 14, 2022, from https://www.nature.com/articles/s41746-021-00465-w#article-info

2. Bent, B., Cho, P. J., Wittmann, A., Thacker, C., Muppidi, S., Snyder, M., Crowley, M. J., Feinglos, M., & Dunn, J. P. (2021, June 1). Non-invasive wearables for remote monitoring of HbA1c and glucose variability: Proof of concept. BMJ Open Diabetes Research & Care. Retrieved June 14, 2022, from https://drc.bmj.com/content/9/1/e002027I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

3. Cho PJ, Bent B, Wittmann A, Merwin RM, Thacker CR, Feinglos MN, et al. 73-LB: Expanding the Definition of Intraday Glucose Variability. Diabetes. 2020 Jun 1;69(Supplement_1):73-LB.

4. DBDP. (n.d.). Retrieved September 1, 2022, from https://dbdp.org/

5. Digital Health Data Repository (DHDR). (2022). DBDP. https://github.com/DigitalBiomarkerDiscoveryPipeline/Digital_Health_Data_Repository (Original work published 2021)

6. A. Turin, "Logistic regression from scratch," Medium, https://towardsdatascience.com/logistic-regression-from-scratch-69db4f587e17 (accessed Dec. 23, 2023).

7.[1] "Deriving LSTM gradient for backpropagation," Deriving LSTM Gradient for Backpropagation - Agustinus Kristiadi, https://agustinus.kristia.de/techblog/2016/08/12/lstm-backprop/ (accessed Dec. 23, 2023).

8.[1] "A new way to chat with Friends & Communities," Discord, https://discord.com/channels/@me/1171544308143370335/1188050000534581298 (accessed Dec. 23, 2023).

9.[1] J. Weisberg, "Building a neural network from scratch: Part 2," Building a Neural Network from Scratch: Part 2, https://jonathanweisberg.org/post/A%20Neural%20Network%20from%20Scratch%20-%20Part%202/ (accessed Dec. 23, 2023).

10. Turin, A. (2020) *Logistic regression from scratch*, *Medium*. Available at: https://towardsdatascience.com/logistic-regression-from-scratch-69db4f587e17 (Accessed: 23 December 2023).

11.[1] J. Brownlee, “Naive Bayes classifier from scratch in Python,” MachineLearningMastery.com, https://machinelearningmastery.com/naive-bayes-classifier-scratch-python/ (accessed Dec. 23, 2023).