A Data-Driven Approach to Predict Carbohydrate Counting Errors in Diabetes Management

Sahaj Shah
Dartmouth College, Sahaj.S.Shah.21@Dartmouth.edu

Follow this and additional works at: https://digitalcommons.dartmouth.edu/senior_theses

Recommended Citation
https://digitalcommons.dartmouth.edu/senior_theses/226

This Thesis (Undergraduate) is brought to you for free and open access by the Theses and Dissertations at Dartmouth Digital Commons. It has been accepted for inclusion in Dartmouth College Undergraduate Theses by an authorized administrator of Dartmouth Digital Commons. For more information, please contact dartmouthdigitalcommons@groups.dartmouth.edu.
A Data-Driven Approach to Predict Carbohydrate Counting Errors in Diabetes Management

Undergraduate Computer Science Thesis
Sahaj Shah

Advisor: Temiloluwa Prioleau, Ph.D.

In Partial Fulfillment of the Requirements for the Degree of Bachelor of Arts in Computer Science

DARTMOUTH COLLEGE
Hanover, New Hampshire

June 1, 2021
ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Professor Temiloluwa Prioleau, for introducing me to this topic and inspiring me to become an independent researcher. I am forever grateful for her support and encouragement throughout this journey. I thank Abigail Bartolome for her relentless help and insightful discussions about the research in the past year. I would also like to convey my deep gratitude to Professor Sorough Vosoughi for his helpful suggestions and support. Lastly, I would like to dedicate this to my parents, without whom none of this would have been possible.
ABSTRACT

Carbohydrate counting, which refers to estimating the carbohydrate content in meals, is critical for determining mealtime insulin doses and maintaining healthy blood glucose levels in persons with type 1 diabetes (T1D). However, carbohydrate counting errors (i.e., over- or under-estimation of carbohydrate intake) are very common amongst patients and are often a source of poor glycemic control. Fortunately, the prevalence of personal health data from wearable medical devices like continuous glucose monitors (CGMs) and insulin pumps provide unique opportunities for understanding and predicting health management outcomes. In this study, we use adverse glycemic events following meal intakes as a proxy for identifying carbohydrate counting errors, then use supervised machine learning models to predict these carbohydrate counting errors. Our dataset includes an average of 161-days of CGM and insulin pump data from 34 patients with T1D. Using a total of 13 features from both datasets, we observed the highest prediction accuracy of 70.5% with a multilayer perceptron (MLP) classifier compared to a baseline model that only yielded 61% accuracy. This work provides a framework for the development of more data-driven tools that leverage personal health data for decision-support to improve health outcomes for people with T1D.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>Chapter I: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II: Methods</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Data Description</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Exploratory Analysis</td>
<td>5</td>
</tr>
<tr>
<td>2.3 Data Pre-Processing</td>
<td>8</td>
</tr>
<tr>
<td>2.4 Feature Selection</td>
<td>9</td>
</tr>
<tr>
<td>2.5 Model Development</td>
<td>10</td>
</tr>
<tr>
<td>Chapter III: Results</td>
<td>12</td>
</tr>
<tr>
<td>Chapter IV: Discussion</td>
<td>14</td>
</tr>
<tr>
<td>Chapter V: Conclusion and Future Work</td>
<td>16</td>
</tr>
<tr>
<td>Bibliography</td>
<td>17</td>
</tr>
</tbody>
</table>


Chapter 1

INTRODUCTION

Diabetes is a prevalent chronic condition characterized by impaired glucose metabolism, which leads to frequent critical fluctuations in blood glucose levels (Centers for Disease Control and Prevention, 2017; American Diabetes Association, 2017). Thus, the primary goal of a person with diabetes is to maintain their blood glucose levels within the normal or healthy range by managing various factors that affect blood glucose levels such as food intake, physical activity, sleep, and much more (Morton et al., 2020; Brown, 2021). Carbohydrate intake is one of the key factors affecting post-meal glucose response and insulin requirements (Wolever and Bolognesi, 1996; Kang et al., 2013; Bantle et al., 2008). Thus, for a person dependent on insulin as is the case with Type 1 Diabetes (T1D), carbohydrate counting (i.e., estimating the carbohydrate content of meals) is a critical part of diabetes management as this is necessary for determining mealtime insulin doses (Warshaw and Kulkarni, 2011; Kawamura, 2007; Scavone et al., 2010; Rabasa-Lhoret et al., 1999). The recent availability of reliable continuous glucose monitoring (CGM) systems has proven to be a major innovation in diabetes management. Yet, the challenge of carbohydrate counting error persists. Many persons with T1D find carbohydrate counting difficult, despite receiving training about how to estimate the carbohydrate content of different foods and meals. They frequently under- or over-estimate the carbohydrate content of foods consumed, resulting in excess glucose excursions (Freeman and Lyons, 2008). Furthermore, the person’s ability to accurately count carbohydrates is seldom discussed during a visit with a provider, causing it to go unnoticed (Fortin et al., 2017; Meade and Rushton, 2016).

Several studies have found that carbohydrate counting errors often result in glycemic variability, leading to subsequent rises or falls in blood glucose levels. If the amount of carbohydrates in a meal is underestimated this leads to a subsequent rise in blood glucose levels. On the other hand, an over-estimation of the amount of carbohydrates in a meal results in a subsequent fall in blood glucose levels. Smart et al. showed that a variation more than or equal to 20% of the actual carbohydrate content in the meal may result in post-meal hyperglycemia or hypoglycemia, respectively (Smart et al., 2012). Upon the investigation of the extent of carbohydrate counting errors in persons with T1D, studies show that such errors are frequent (63%), with


the majority being underestimated, and associated with higher daily blood glucose variability in adults with T1D (Brazeau et al., 2013). Smart et al. found that larger meals tended to be underestimated while smaller meals (snacks) were more likely to be overestimated. Furthermore, an error of ± 10 grams per meal or snack on a meal size of 60 grams did not significantly deteriorate the post-meal control, while a ±20 grams error had a significant impact on post-meal glycemia (Smart et al., 2012). This suggests that the carbohydrate input entered by the subject and the subsequent response of glucose levels can be used to log carbohydrate counting errors automatically.

Few studies translate the impact of carbohydrate counting errors into assisting patients to make better decisions. Roversi et al. is the only study that proposes a counting error model developed by using multiple linear regression with stepwise variable selection (Roversi et al., 2020). However, the model relies largely on a published dataset from a cross-sectional design study, relying on education level, insulin treatment duration, age, body weight, meal-type, lipid, energy, protein, and fiber content of the patient — data that is not readily available to the patient. A machine-based prediction model of such carbohydrate counting errors in persons with diabetes before post-meal alterations in blood glucose levels would be desirable. However, a primary challenge in predicting carbohydrate counting error lies in the lack of ground truth values, i.e. actual meal composition of meal consumed by an individual. These values are often not available in a real-life setting.

To circumvent the issue of lack of ground truth of meals consumed by an individual, we use meal-related glucose excursions, derived from the individual’s continuous glucose monitoring data, as a proxy for determining if a carbohydrate input is associated with an error. This is informed by several studies that suggest that increased glycemic variability, excessive fluctuations in glucose levels, can be associated with diabetes-related complications (Reiterer, Freckmann, and Re, 2018; Kilpatrick, Rigby, and Atkin, 2006). To address the gap in the literature, we designed a personalized machine-learning-based model to predict carbohydrate counting errors. The primary goal of this study is to develop data-driven solutions that leverage individual patient’s data from wearable devices (i.e. continuous glucose monitors and insulin pumps) to characterize carbohydrate counting errors in free-living adults with T1D using post-meal glucose excursion as a proxy for carbohydrate counting error classification. Furthermore, we aim to improve on state-of-the-art methods for estimating carbohydrate counting errors and predict future occurrences via machine learning.
classification models. Therefore, we define carbohydrate counting error within the context of post-meal blood glucose excursions. To the best of our knowledge, this is the first study that aims to build a classification model to predict carbohydrate counting errors using an individual’s retrospective CGM data. The ability to correctly predict whether a carbohydrate input by the subject was associated with an error will (1) better inform patients with diabetes about their carbohydrate counting habits and recommend a personalized treatment plan for managing glucose level to prevent diabetes-related complications and (2) augment future research on better prediction and management of post-meal hyperglycemia.
Chapter 2

METHODS

2.1 Data Description

All the data used in this study for exploratory analysis and prediction models were contributed by the non-profit organization Tidepool. The data include CGM and insulin data from 34 subjects with type 1 diabetes (age = 39.8 ± 8.73 yrs., time since diagnosis = 18.4 ± 10.58 yrs.) for an average of 161-days per subject. (Tidepool n.d.). CGMs today record 1 sample every 5 minutes, yielding approximately 288 blood glucose samples for a 24-hour period. Therefore, the number of samples is 288 × no. of days/subject. Corresponding insulin pump data for subjects provided details on the amount of carbohydrate input and its associated time. The insulin pump data includes an average of 284 recorded entries of carbohydrate input for each subject. The descriptions of the features and brief statistical summary are shown in Table 2.2.

We define a carbohydrate entry recorded by the subject as being associated with an error if it results in a hyperglycemic or hypoglycemic episode. As outlined by the American Diabetes Association, a hyperglycemic episode is defined as glucose levels above 180 mg/dL, while a hypoglycemic episode is defined as glucose levels below 70 mg/dL (Association, 2020). The selection of time window was informed by literature, where peak post-meal glucose levels are observed to be around 1-1.5 hours after the beginning of a meal (Daenen et al., 2010). This ensures that the hyperglycemic or hypoglycemic episode is associated with a carbohydrate input.

As a result, a time window of 15 minutes prior to carbohydrate input up to 120 minutes after the start of the carbohydrate input was used to effectively capture the subsequent peak in glucose levels. The decision to include 15 minutes prior was to account for instances when the patient might have forgotten to record entry and done so after the start of the meal. Since peak glucose is observed approximately 90 minutes after recording the carbohydrate input, all carbohydrate input entries within that time frame were grouped together, and respective carbohydrate amounts were summed while keeping the latest time recorded. For example, if a subject recorded 20g of carbohydrate input at 10 AM and subsequently another 25g at 10:30 AM, then a combined carbohydrate input of 45g at 10:30 AM was used. Each
carbohydrate input was then compared to glucose values 15 minutes and 120 min after and labeled as an error if a hyper/hypoglycemic episode was observed for more than or equal to 10 minutes. For example, Figure 2.1 illustrates a sample dashboard of glucose values for a particular subject on a given day. Two carbohydrate inputs were recorded, one of 25g at 8:46 am and the second of 30g at 1:02 pm. Both result in post-meal glucose levels above 180 mg/dL lasting after 120 minutes for a period of more than 10 minutes. As a result, they are both associated with an error. Carbohydrate counting errors were thus classified using the subject’s CGM and insulin pump data.

2.2 Exploratory Analysis
To determine which features were important for the prediction of carbohydrate counting errors, we performed an exploratory data analysis. Figure 2.2 demonstrates the distribution of recorded carbohydrate inputs from all 33 subjects in the Tidepool dataset. The average carbohydrate input amount was $48.05 \pm 51.01 \text{g}$.

First, we investigated the extent of carbohydrate counting error in our dataset. On average, 347 out of 615 (56.4%) carbohydrate inputs per subject resulted in an error, with 263 (75.7%) resulting in hyperglycemia (i.e., being underestimated). This was consistent with findings from other studies where patients had a pronounced negative estimation bias, with 63% of 448 analyzed meals being underestimated (Brazeau et al., 2013). Additionally, Deeb et al. found that 165 out of 246 meals (67%) were considered inaccurate when compared to dieticians’ findings. The majority of underestimated carbohydrate counts led to above target glucose levels (Deeb et al.,
Figure 2.2: Histogram showing the distribution of carbohydrate inputs from all 34 subjects in the dataset.

2017).

Table 2.1: Impact of carbohydrate counting errors on diabetes management

<table>
<thead>
<tr>
<th></th>
<th>Mean (\bar{x}) (n=34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of carbohydrate entries per subject</td>
<td>615.91</td>
</tr>
<tr>
<td>Carbohydrate entries associated with errors per subject</td>
<td>347.91</td>
</tr>
<tr>
<td>Number of errors that resulted in hypoglycemia</td>
<td>84.50</td>
</tr>
<tr>
<td>Number of errors that resulted in hyperglycemia</td>
<td>263.41</td>
</tr>
</tbody>
</table>

Then, we investigated whether the carbohydrate counting errors were associated with a particular time of the day. 24 hours of the day were split into four windows, each encompassing six hours. While the overnight window (12am-6am) resulted in the most amount of errors (66.9%), no significant difference was observed among the different time windows (Figure 2.3).

Finally, we investigated whether carbohydrate counting error was associated with the size of carbohydrate input. We observed a small increase in the occurrence of carbohydrate input error from 0 – 40g interval to 60 – 80g interval. However, this effect is not observed in higher carbohydrate intervals as one might expect. This was in contrast to the findings from prior studies, where the error increased two-fold from 80 < carbohydrate amount (g) < 100 to carbohydrate amount (g) > 100 alone
Figure 2.3: (a) Heat map showing carbohydrate counting patterns for 34 individual subjects. Each horizontal strip summarizes the probability of counting errors for a particular subject, associating each hour of the day with the corresponding ratio of carbohydrate inputs that result in an error in appropriately colored boxes. The greater the probability of counting error, the brighter the corresponding box. Black color denotes that no carbohydrate input value was recorded at that hour. (b) Bar plot showing percentage of carbohydrate inputs associated with an error per time interval.
(Roversi et al., 2020). To conclude, based on our dataset, we did not observe strong association between meal size and time of the day with carbohydrate input error (Figure 2.4).

2.3 Data Pre-Processing

To address the issue of missing data evident from data cleaning, we perform the following data pre-processing steps on the dataset:

**Outlier rejection**: Since many classifiers are sensitive to the distribution of data, data points that markedly deviated from others were rejected. Informed by literature (Hasan et al., 2020), outlier rejection is defined as:

\[
    f(x) = \begin{cases} 
    x, & \text{if } Q1 - 1.5 \times IQR \leq x \leq Q3 + 1.5 \times IQR \\
    \text{reject}, & \text{otherwise} 
    \end{cases}
\]

where \(x\) refers to instances of the feature vector, \(x \in IR\). \(Q1, Q3,\) and \(IQR\) represent the first quartile, third quartile, and interquartile range of the features respectively, \(Q1, Q3, IQR \in IR\).

**Handling missing values**: Due to CGM and insulin pump being wearable devices, it is not uncommon to encounter missing data upon feature extraction. After outlier rejection, features with missing values were imputed by using mean values of respective features.
Scaling: Finally, due to different ranges of meal-related features (measured in grams) and glucose-related features (measured in mg/dL), all features were scaled down using a min-max mapping algorithm, linearly transforming each feature into the interval [0,1] while maintaining the original distribution of the dataset using the equation below.

\[ x' = \frac{x - \min(x)}{\max(x) - \min(x)} \]

where \( x \) and \( x' \) refer to instances of the feature vector before and after scaling, \( x, x' \in IR \). Lastly, one subject was removed from the study because they did not have sufficient data in the two classes of interest.

2.4 Feature Selection

Based on results from the exploratory analysis and literature review, a total of 11 features were extracted from the dataset. Table 2.2 describes a full summary of the features used. These features can be classified into three categories:

1. Meal-related features: Meal-related features capture data from an insulin pump. As investigated in exploratory analysis, the size of carbohydrate input and time of the carbohydrate input can play a role in carbohydrate counting errors. Other features such as insulin:carbohydrate ratio, amount of carbohydrate consumed in the last 24 hours of recorded carbohydrate input were also evaluated but did not increase the overall accuracy of the model.

2. Glucose-related features: Since post-meal glucose excursion is used as a proxy for carbohydrate counting errors, glucose-related features can help us capture short-term and long-term patterns in glucose levels before a recorded carbohydrate input. In particular, short-term glycemic features such as ‘Average_glucose_24h’, ‘Std_24h’,

---

**Figure 2.5:** Overall workflow of study and approach to classification model development.
‘Average_glucose_6h’, ‘Std_2h’, and ‘Std_4h’ were used. Finally, snowball effect features, ‘Pos_increments’ and ‘Neg_increments’ were used to capture accruing effects of change in glucose levels over time before a recorded carbohydrate input (Dave et al., 2020). Glucose-related features were transformed into log values due to their non-linear distributions.

(3) Time-related features: Time-related features such as day of the week and time of recorded carbohydrate input were selected to capture the temporal patterns in carbohydrate counting habits of an individual.

2.5 Model Development

We selected eight commonly used machine learning classifiers: Multi Layer Perceptron (MLP), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), multinomial Naive Bayes (NB), Adaptive Boosting (AdaBoost), and an ensemble model consisting of MLP followed by SVM (MLP + SVM). These models were successfully implemented using Python.

A $K$-fold cross validation approach ($K = 5$) was then used on the Tidepool dataset. For a given model, we first partition the dataset into $K$ folds. Using the grid search algorithm, we train and fine-tune different hyperparameters using $K-1$ folds. After selection of the best hyperparameters, the dataset was divided into two parts: 80% for training and validation, and 20% for testing the model. Class distributions were maintained in both cohorts as observed in the dataset. All models were trained for 100 epochs and initialized with a random seed of 42. Overall accuracy (correct classifications per overall classifications), recall (also known as sensitivity; true positives per true positives and false negatives), precision (true positives per true positives and false positives), and f1-scores (weighted average of precision and recall) were used to compare performance across different models. For comparison of performance, a dummy baseline model was created that predicts an error (class = 1) for every carbohydrate input entered. The results from this process are discussed in the next section.
Table 2.2: Summary of features extracted for classification

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Mean ± Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Carbinput’ (g)</td>
<td>Carbohydrate entry recorded by the subject</td>
<td>34.58 ± 2.30</td>
</tr>
<tr>
<td>‘Carb_count’</td>
<td>Frequency of carbohydrate entries recorded by the subject that resulted in an error in the last 24 hours</td>
<td>23.36 ± 11.85</td>
</tr>
<tr>
<td>‘Average_glucose_24h’ (mg/dL)</td>
<td>Mean value of CGM observations in the last 24 hours</td>
<td>142.55 ± 1.27</td>
</tr>
<tr>
<td>‘Average_glucose_6h’ (mg/dL)</td>
<td>Mean value of CGM observations in the last 6 hours of carbohydrate input</td>
<td>140.58 ± 1.41</td>
</tr>
<tr>
<td>‘Std_24h’ (mg/dL)</td>
<td>Standard deviation of CGM observations in the last 24 hours</td>
<td>46.58 ± 18.20</td>
</tr>
<tr>
<td>Std_2h (mg/dL)</td>
<td>Standard deviation of CGM observations in the last 2 hours of carbohydrate input</td>
<td>17.14 ± 12.67</td>
</tr>
<tr>
<td>Std_4h (mg/dL)</td>
<td>Standard deviation of CGM observations in the last 4 hours of carbohydrate input</td>
<td>31.31 ± 19.49</td>
</tr>
<tr>
<td>‘Pos_increments’</td>
<td>Sum of all increments in adjacent CGM observations in last two hours of carbohydrate input</td>
<td>46.29 ± 69.67</td>
</tr>
<tr>
<td>‘Neg_increments’</td>
<td>Sum of all decrements in adjacent CGM observations in past two hours of carbohydrate input</td>
<td>-45.58 ± 66.28</td>
</tr>
<tr>
<td>‘Time’</td>
<td>Time (format hh:mm) recorded for carbohydrate input</td>
<td>NA</td>
</tr>
<tr>
<td>‘Day_of_week’</td>
<td>Day of the week</td>
<td>NA</td>
</tr>
</tbody>
</table>
Chapter 3

RESULTS

Table 3.1 shows a full summary of accuracy, precision, recall, and f1-scores for each classifier. The MLP neural network yields the best performance across all classifiers, resulting in an accuracy of 70.5%. This is an improvement from the baseline model that resulted in an accuracy of 61%. It also led other models in terms of precision and f1-score. The best model consisted of a single layer of 100 hidden units, designed to learn the features and output a binary prediction label. A Rectified Linear Unit (ReLU) activation function was used with a constant learning rate of 0.0001 with Adam optimization. Figure 3.1 outlines subject-level breakdown of accuracy scores for the MLP model and feature importance. Glucose-related features, particularly, average glucose values 24 hours prior to the carbohydrate input as well as standard deviation of glucose values 24 and 6 hour prior to a carbohydrate input contributed the most to the error classification. On the other hand, temporal feature, day of the week at the time of the recorded carbohydrate input, contributed the least. We also repeated the above experiments by combining data across all 33 subjects and training a single classifier (subject-independent) rather than a unique classifier per subject (subject-dependent). We found that the subject-dependent models resulted in a greater accuracy than subject-independent models (68.6%).
Table 3.1: Performance of classifiers (mean ± standard deviation) given by 5-fold validation.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>f1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>0.705 ± 0.08</td>
<td>0.708 ± 0.09</td>
<td>0.829 ± 0.19</td>
<td>0.756 ± 0.14</td>
</tr>
<tr>
<td>LR</td>
<td>0.701 ± 0.09</td>
<td>0.706 ± 0.09</td>
<td>0.824 ± 0.24</td>
<td>0.743 ± 0.18</td>
</tr>
<tr>
<td>RF</td>
<td>0.692 ± 0.09</td>
<td>0.707 ± 0.09</td>
<td>0.787 ± 0.22</td>
<td>0.734 ± 0.16</td>
</tr>
<tr>
<td>SVM</td>
<td>0.688 ± 0.10</td>
<td>0.651 ± 0.19</td>
<td>0.838 ± 0.26</td>
<td>0.729 ± 0.21</td>
</tr>
<tr>
<td>DT</td>
<td>0.614 ± 0.08</td>
<td>0.678 ± 0.13</td>
<td>0.662 ± 0.14</td>
<td>0.668 ± 0.13</td>
</tr>
<tr>
<td>NB</td>
<td>0.641 ± 0.08</td>
<td>0.722 ± 0.11</td>
<td>0.649 ± 0.14</td>
<td>0.681 ± 0.12</td>
</tr>
<tr>
<td>ADA Boost</td>
<td>0.656 ± 0.09</td>
<td>0.693 ± 0.12</td>
<td>0.724 ± 0.16</td>
<td>0.706 ± 0.14</td>
</tr>
<tr>
<td>MLP + SVM</td>
<td>0.695 ± 0.10</td>
<td>0.649 ± 0.21</td>
<td>0.914 ± 0.10</td>
<td>0.740 ± 0.15</td>
</tr>
</tbody>
</table>

Figure 3.1: Plots from the best model. (a) individual breakdown of accuracy scores for each subject. (b) comparison of different feature importances measured using the Gini index.
DISCUSSION

Personalized predictive models were generated using supervised ML algorithms widely applied in supervised learning. Of all the classifiers that were tested, the MLP model resulted in the greatest accuracy of 70.5%, an increase from 61% for the baseline model. The MLP model also led in terms of precision, recall, and f1-score. The highest accuracy observed in Subject 23 is 91% and the lowest of 56% in Subject 25. Upon further observation, this can be attributed to two factors: (1) greater instances of a carbohydrate input being associated with an error in Subject 25 resulting in an unequal class distribution compared to Subject 23, and (2) overall poor association of features with an error in Subject 25 compared to Subject 23. We also found that glucose-related features contributed the most to the model while temporal features contributed the least. This is consistent with how we define post-meal glycemia as a measure for carbohydrate counting error. Upon examination of subjects that resulted in high accuracy from those that resulted in low accuracy, we notice that the model fails to discriminate data points when there is an equal distribution of classes. It could be that the features simply do not provide sufficient information for the classifier or perhaps the way we define carbohydrate counting error is inadequate and requires more nuance.

There are several limitations to this study. Primarily, the concept of using post-meal glucose excursions as a proxy for carbohydrate counting error classification has not been investigated before. Due to the absence of true ground truth values (actual carbohydrate content of the meal consumed), we are unable to go beyond error classification to recommend corrective actions. The study also does not account for carbohydrate inputs that may have been missed or not recorded by the subject. For example, eating a banana every morning for breakfast without entering the carbohydrates may result in a hyperglycemic response but this will not be reflected in our analysis. Thus, the classifier can only predict for carbohydrate inputs that have been recorded by the subject. In addition to amount of carbohydrates in a meal, post-meal glycemic response is affected by other macronutrients including proteins and fats (Bell et al., 2015; Meng et al., 2017). Evidence suggests that meals that are low in carbohydrates but high in dietary fat cause sustained elevations in post-meal glucose levels. However, this model does not account for the effects from different
meal composition. Lastly, we discarded instances of carbohydrate inputs when a subject had glucose levels above 70 mg/dL. We are unable to classify carbohydrate inputs if the person already has hyperglycemia during that instance.
Chapter 5

CONCLUSION AND FUTURE WORK

Prediction of carbohydrate counting errors using CGM and insulin data can provide great benefit to people with T1D with personalized, adaptive, and context-aware coaching of their carbohydrate counting habits. We propose to use meal-related glucose excursions as a marker for determining if a carbohydrate input is associated with an error. The overall extent of such errors and the association of various features such as size and time of the day with an input error was also investigated. We then derived meal-related, glucose-related, and temporal features to build subject-dependent classification models. Our analysis shows that the MLP model resulted in the highest accuracy of 70.5% while decision trees were classified with the lowest accuracy. The main distinguishing factor between subjects with high accuracy and those with low accuracy was class imbalance. To the best of our knowledge, this is the first study that aims to build a classification model to predict carbohydrate counting errors using an individual’s retrospective CGM data.

A number of future directions for this work are possible. More subject-dependent factors can be investigated in order to improve the accuracy of our model. Importantly, further experiments need to be performed to gather the true quantity of known meals and train subject-dependent models. This will allow us to validate our use of post-meal glucose excursions as a proxy for carbohydrate counting errors. Finally, physical activity can have an impact on the postprandial glucose response. However, this study did not take physical activity of the subjects into account while considering the changes in blood glucose levels. We would like to incorporate features based on physical activity and sleep into our model.
BIBLIOGRAPHY


Deeb, Asma et al. (2017). “Accurate carbohydrate counting is an important determinant of postprandial glycemia in children and adolescents with type 1 diabetes on insulin pump therapy”. In: *Journal of diabetes science and technology* 11.4, pp. 753–758.

Fortin, Andréanne et al. (2017). “Practices, perceptions and expectations for carbohydrate counting in patients with type 1 diabetes–Results from an online survey”. In: *Diabetes research and clinical practice* 126, pp. 214–221.


Kang, Xin et al. (2013). “Effects of different proportion of carbohydrate in breakfast on postprandial glucose excursion in normal glucose tolerance and impaired glucose regulation subjects”. In: Diabetes technology & therapeutics 15.7, pp. 569–574.


Smart, CE et al. (2012). “In children using intensive insulin therapy, a 20-g variation in carbohydrate amount significantly impacts on postprandial glycaemia”. In: Diabetic Medicine 29.7, e21–e24.
