2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Machine Learning-based Meal Detection Using Continuous Glucose Monitoring on Healthy Participants: An Objective Measure of Participant Compliance to Protocol

Victor Palacios[1], Diane Myung-kyung Woodbridge[2] and Jean L. Fry[3]

***Abstract*— Meal timing affects metabolic responses to diet, but participant compliance in time-restricted feeding and other diet studies is challenging to monitor and is a major concern for research rigor and reproducibility. To facilitate automated validation of participant self-reports of meal timing, the present study focuses on the creation of a meal detection algorithm using continuous glucose monitoring (CGM), physiological monitors and machine learning. While most CGM-related studies focus on participants who are diabetic, this study is the first to apply machine learning to meal detection using CGM in metabolically healthy adults. Furthermore, the results demonstrate a high area under the receiver operating characteristic curve (AUC-ROC) and precision-recall curve (AUC-PR). A cold-start simulation using a random forest algorithm yields .891 and .803 for AUC-ROC and AUC-PR respectively on 110-minutes data, and a non-cold start simulation using a gradient boosted tree model yields over .996 (AUC-ROC) and .964 (AUC-PR). Here it is demonstrated that CGM and physiological monitoring data is a viable tool for practitioners and scientists to objectively validate self-reports of meal consumption in healthy participants.**

***Index Terms*— Machine Learning, Meal Detection, Patient Compliance, Continuous Glucose Monitoring**

I. Introduction

In nutrition science research, a lack of automated tools to validate compliance with study protocols in the free world is widely recognized as a major barrier to progress in the field [1]. Though meal timing and fasting intervals have recently emerged as an important determinant of metabolic health, it is not yet known how long fasting periods must be, how these timing patterns affect different populations, or whether the benefits of limiting eating windows extend beyond their effects on total energy intake. Though federal agencies are currently funding several human studies to answer some of these important questions, these research efforts are hampered by the inability to easily and objectively evaluate compliance with diet study protocols. Currently, investigators conducting these studies rely on participant-initiated prompts, for example time timestamps in digital food logs [2], to signal the beginning and end of eating periods, which limits the ability to determine a diet protocol's efficacy.

[1]Victor Palacios is a student in the MS in Data Science Program, University of San Francisco. vpalacios@usfca.edu

[2]Diane Myung-kyung Woodbridge, Ph.D. is an Assistant Professor at the MS in Data Science Program, University of San Francisco. dwoodbridge@usfca.edu

[3]Jean L. Fry, Ph.D., R.D.N. is an Assistant Professor in the Department of Athletic Training and Clinical Nutrition, University of Kentucky. jean.fry@uky.edu

Continuous glucose monitoring (CGM) has been used to objectively identify meals in people with diabetes [3][4][5][6]. In this study, we utilize CGM and physiological monitoring data to identify eating occasions in metabolically healthy adults whose post-meal blood glucose excursions deviate less from baseline when compared with people having diabetes or impaired glucose tolerance. Our data was collected in a highly controlled inpatient setting allowing us to know exact meal times. Here we use tree-based models, including random forests and gradient boosted trees, to determine how glucose, heart rate, physical activity, core temperature, skin temperature, and respiration rate may be utilized to develop an automated tool for meal detection in healthy participants.

II. Data Collection

A. Participant Characteristics

Nine healthy, young (24.3 ± 4.6 years), lean (BMI 24.4 ± 2.9 and body fat $16.6 \pm 5.8\%$), non-smoking male volunteers participated in this study. Participants had normal fasting blood glucose concentrations (<100 mg·dL^{-1}). The data presented here were collected on post-resistance exercise nutrient timing study with two experimental conditions where participants received immediate post-exercise nutrition (IPEN) or 3-hour post-exercise nutrition (3h-PEN). The trial was registered on clinicaltrials.gov and complete enrollment and trial details may be viewed on the website (NCT01674049). The study protocol was approved by the institutional review board of the George Washington University Medical Center and all participants gave written and oral informed consent prior to participating.

B. Baseline Measures and Familiarization

After passing an eligibility screen and providing consent, participants completed a baseline assessment and familiarization session. We collected fasting blood glucose measured with an Accu-Check Advantage (Roche Diagnostics, Indianapolis, IL). Standing height was measured by using a wall-mounted stadiometer and participants were weighed in undergarments on a digital scale. Body fat percentage was determined using air displacement plethysmography (BodPod®, Life Measurement, Inc, Concord, CA).

After completing baseline testing, participants were familiarized to resistance exercise (RE) equipment and the exercise protocol to be used in the calorimeter, which was a 40-minute circuit using resistance bands (Tower 200,

Pacoima, CA). Participants were given meal plans to follow for two days preceding their experimental sessions. Energy content of the meal plans were assigned by estimating basal metabolic rate with the Harris-Benedict equation [7]. The resulting value was multiplied by a factor of 1.5 to estimate total daily energy expenditure (TEE) for a healthy, untrained young person.

C. Experimental Procedures

Participants spent approximately 48-hours in a whole-room calorimeter at the Beltsville Human Nutrition Research Center (BHNRC) within the US Department of Agriculture (USDA). The first 24-hours served as a control period and participants exercised on the second day. We continuously measured respiratory gas exchange (O_2 consumption, CO_2 output), core body temperature (BT), and heart rate (HR) throughout the 48-hour protocol.

1) Continuous Glucose Monitoring: Before entering the calorimeter, the iPRO^{TM} Professional Continuous Glucose Monitoring System (Medtronic MiniMed, Inc, Northridge, CA) was placed on the patient's abdomen and remained for the duration of the experiment. The data stored in the monitor were transferred and converted to blood glucose concentrations with the CGM system's software. A venous catheter was placed in the antecubital vein to allow for blood draws throughout the experimental session.

2) Ambulatory Monitoring: Participants wore the Equivital Sensor Electronics Module (Equivital I: Hidalgo Ltd., Cambridge, UK) for the entire 48-hour period. This device measures heart rate (HR), heat flux, and core body temperature, in combination with a heat-sensitive transmitter (pill) that was ingested immediately before entering the calorimeter. Participants were asked not to nap, but they were allowed to choose their own sleep times overnight.

3) Diet: Participants consumed a standardized modestly hypercaloric diet providing equal amounts of energy on both days in the metabolic chamber. Energy needs were determined using the TEE estimate (as above for meal plans) plus a 450 kcal post-exercise supplement (low-fat chocolate milk) that was administered on both the control and exercise days. The supplement was administered either immediately after a resistance exercise bout or 3 hours after the exercise bout. Since this was a cross-over study, participants completed both sessions. Total daily energy intake had an approximate macronutrient distribution of 52% carbohydrate, 32% fat, and 16% protein. Based on prior estimates of approximately 6 kcal/min of increased energy expenditure in response to in-chamber RE [8], we expected the 450 kcal supplement would provide an energy surplus of approximately 450 kcal on the control day and 210 kcal on the exercise day. Participants consumed meals at approximately 08:00, 11:20, 16:00 (supplement) and 18:00 on both days, and participants were asked to consume all food provided. Exact meal times were recorded for all meals and supplements ingested.

4) Resistance Exercise Bout: Participants performed 40 minutes of circuit-style RE in the calorimeter at approximately 12:20 p.m. (one hour after eating lunch). Participants completed as many circuit sets as possible in 40-min, while taking 90-second breaks between each circuit exercise set. Exercise and rest periods between each circuit set were monitored at all times by an exercise physiologist.

TABLE I: Outlier data definition

Readings	Lower Limit	Upper Limit
Heart Rate	40	210
Respiration Rate	7	45
Core Temperature	36	40
Skin Temperature	33	42

Algorithm 1 Outlier removal function

1: **for** Rows with five sensor readings **do**
2: Detect values outside of normal ranges.
3: Replace outlier start and end points with the nearest realistic value.
4: After replacement, take the nearest realistic points and average between them.
5: **end for**
6: **return** Data without outliers caused by sensor noises

III. Data Preprocessing

Since the calorimeter samples every minute, the Equivital physiological monitoring and CGM data were resampled to match the calorimeter data using cubic spline interpolation [9].

The algorithm for data removing outlier values not within the normal range of a human is given in Algorithm 1. Data to be labeled as outlier data were determined by a review of the literature [10][11][12][13] and a clinical nutritionist and exercise physiologist as given in Table I. For instance, heart rate readings of [1, -1000, 40, 50, -10, 60, 70, 20, 10000] become [40, 40, 40, 50, 55, 60, 70, 70, 70] after data cleansing.

After outlier removal, Algorithm 2 extended from our previous work [14] creates a window (w) that includes previous k sensor readings and extends the current reading to incorporate the readings themselves, the difference between the first and last readings, and statistical values within w. This results in adding $5k + 40$ extra variables.

IV. Machine Learning

1) Training and Test Set Splits: The present study applied a cold-start approach and non-cold start approach. The cold-start approach splits the data into a training set using 70% (all but two random participants) of the original data and the remaining 30% as the test set to simulate the case where a doctor receives data from a brand new patient. The non-cold start approach splits the data into a training set using 80% of samples and a test set with the remaining 20% of samples to simulate a doctor receiving from new data from a patient they have prior data for. In both cases, tree models (random forest machine learning model [15] and gradient boosting machines [16]) were trained on the training set and made predictions on the test set. These models were chosen based

Algorithm 2 Window function

for Each readings of a participant sorted by timestamp **do**
  for Each sensor, s **do**
    Create a window (w) including the previous k readings of s.
    Add $\delta = d_k - d_0$, where d_k is the current sensor reading and d_0 is the sensor reading of the first reading within w.
    Add $\delta_{percent} = \frac{\delta}{d_0}$.
    Add δ^2 and $\delta^2_{percent}$.
    Add mean, standard deviation, minimum and maximum values in w.
  end for
  return Extended readings that includes readings of k consequent minutes and statistical values
end for

on previous studies which demonstrated robust results with tree-based models for CGM studies [17].

2) Variable Selection: For the machine learning model, we used the data which consisted of six primary variables: glucose, heart rate, physical activity, core temperature, skin temperature, and respiration rate. Physical activity was not found to aid in prediction and was thus removed as a variable. Conversely, new lag variables that capture value differences across time were created as well as other statistics such as the minimum, maximum, and standard deviation which did aid in prediction. As well, unlike other studies using age, gender, body mass index, hemoglobin, etc. [17] the focus of this study was on rapidly-changing physiological variables which allow the machine learning model to better generalize its predictions.

3) Library: The present study utilized H2O, a distributed, scalable machine learning library [18]. That library includes automated machine learning (AutoML) which automates processes of the machine learning cycle such as hyperparameter tuning. In addition, the time needed to train several random forests or gradient boosted trees each with varying hyperparameters was faster with H2O by distributed in-memory data processing.

V. Results

We evaluate the model performance based on 1) the area under the receiver operating characteristic curve (AUC-ROC) and 2) the area under the precision-recall curve (AUC-PR). In this study, one goal is to find the minimal window size (k) which yields high AUC-ROC and AUC-PR values. Since this is an imbalanced problem where one class (eating) does not occur as frequently as the other class (not eating), the primary goal would be finding the optimal window size with a higher AUC-PR value. As well, because of this imbalance accuracy is not an appropriate metric to report.

For the cold-start case, two participants were randomly selected as hold-out data which the random forest model was never exposed to when generating its algorithm for prediction. This offers a chance to validate how well the created model performs on unseen data. With $k = 110$ minutes, the AUC-ROC value was .891 and the AUC-PR value was .803. For non-cold-start cases, with $k = 20$ minutes, AUC-ROC was .996 and AUC-PR was .964 (Figure 1).

Beyond single digit metrics, Figure 2 also shows how a doctor, nutritionist or researcher may use the results of their own similarly constructed model to find hot-spots indicative of eating. As it relates to the overall eating window, which is of interest to those studying chrononutrition, our methodology typically predicted a window that was about one hour longer than the actual window, and the timing was delayed one hour later than actual eating times.

VI. Conclusion

The present study aimed to create a meal detection model for healthy participants using continuous glucose monitoring, physiological monitors and machine learning. The results demonstrate that a random forest and gradient boosted machine learning model can detect when a participant has eaten with high AUC-PR and AUC-ROC scores without food type or physical activity data.

Future studies should minimize high inflation in the performance of evaluation metrics by minimizing the use of invasive procedures and using an anxiety screening survey during the protocol [19]. In this work, finger sticks and blood draws from a catheter took place immediately before all the meals, which may have induced anxiety in patients and affected glucose and heart rate levels. Prior evidence indicates that different CGM devices show variable glucose responses to the same meals in healthy people [20], so data collected with different CGM devices may need to be standardized for consistency.

Our novel approach enables scientists to passively approximate participants' meal times and feeding windows with low participant burden utilizing readily available CGM technology. Future work refining the model and most useful data inputs may be a viable way for scientists in the field of chrononutrition to collect objective meal timing data in the free world, which would enhance rigor and reproducibility of clinical trials in nutrition sciences.

Acknowledgment

The US Department of Agriculture Beltsville Human Nutrition Research Center provided food, facilities and equipment.

References

[1] NIH Nutrition Research Task Force, "2020–2030 strategic plan for nih nutrition research–a report of the nih nutrition research task force," *https://dpcpsi.nih.gov/sites/default/files/2020NutritionStrategicPlan_508.pdf*, 2020.

[2] M. T. Stratton, G. M. Tinsley, M. G. Alesi, G. M. Hester, A. A. Olmos, P. R. Serafini, A. S. Modjeski, G. T. Mangine, K. King, S. N. Savage *et al.*, "Four weeks of time-restricted feeding combined with resistance training does not differentially influence measures of body composition, muscle performance, resting energy expenditure, and blood biomarkers," *Nutrients*, vol. 12, no. 4, p. 1126, 2020.

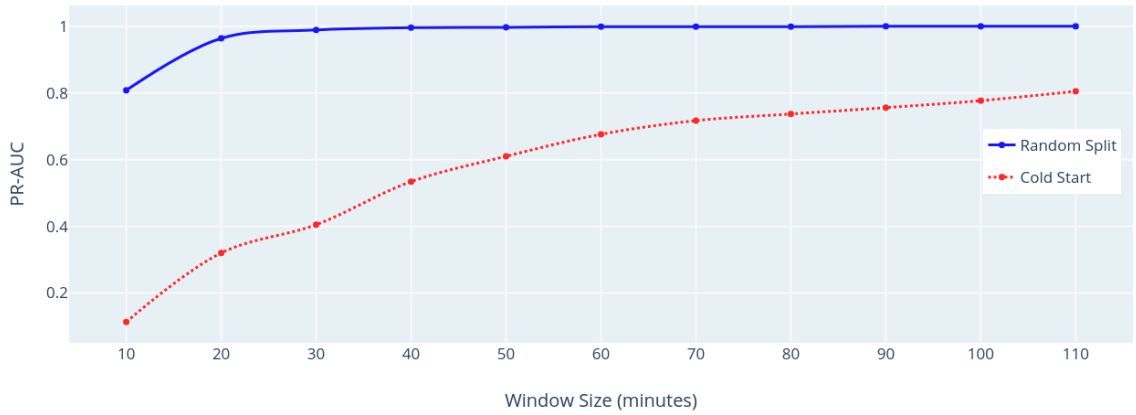

Fig. 1: Comparison between AUC-PR of random-split and cold-start cases

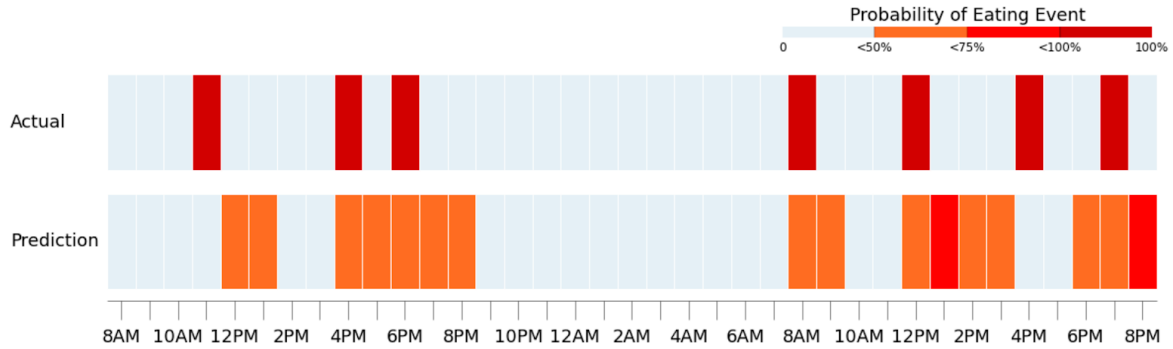

Fig. 2: Actual vs Predicted Eating Events: Algorithm's confidence of eating event by hour for a single participant

[3] S. Chen, J. Weimer, M. R. Rickels, A. Peleckis, and I. Lee, "Towards a model-based meal detector for type 1 diabetics," *6th Workshop on Medical Cyber-Physical Systems (MedicalCPS 2015)*, 2015.

[4] C. M. Ramkissoon, P. Herrero, J. Bondia, and J. Vehi, "Unannounced meals in the artificial pancreas: detection using continuous glucose monitoring," *Sensors*, vol. 18, no. 3, p. 884, 2018.

[5] L. Meneghetti, A. Facchinetti, and S. Del Favero, "Model-based detection and classification of insulin pump faults and missed meal announcements in artificial pancreas systems for type 1 diabetes therapy," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 1, pp. 170–180, 2020.

[6] C. M. Ramkissoon, P. Herrero, J. Bondia, and J. Vehi, "Meal detection in the artificial pancreas: implications during exercise," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5462–5467, 2017.

[7] J. A. Harris and F. G. Benedict, "A biometric study of human basal metabolism," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 4, no. 12, p. 370, 1918.

[8] E. L. Melanson, T. A. Sharp, H. M. Seagle, W. T. Donahoo, G. K. Grunwald, J. C. Peters, J. T. Hamilton, and J. O. Hill, "Resistance and aerobic exercise have similar effects on 24-h nutrient oxidation," *Medicine & Science in Sports & Exercise*, vol. 34, no. 11, pp. 1793–1800, 2002.

[9] S. McKinley and M. Levine, "Cubic spline interpolation," *College of the Redwoods*, vol. 45, no. 1, pp. 1049–1060, 1998.

[10] L. Chhabra, N. Goel, L. Prajapat, D. H. Spodick, and S. Goyal, "Mouse heart rate in a human: diagnostic mystery of an extreme tachyarrhythmia," *Indian pacing and electrophysiology journal*, vol. 12, no. 1, pp. 32–35, 2012.

[11] I. Lin, L. Tai, and S.-Y. Fan, "Breathing at a rate of 5.5 breaths per minute with equal inhalation-to-exhalation ratio increases heart rate variability," *International Journal of Psychophysiology*, vol. 91, no. 3, pp. 206–211, 2014.

[12] M. Torii, M. Yamasaki, T. Sasaki, and H. Nakayama, "Fall in skin temperature of exercising man." *British journal of sports medicine*, vol. 26, no. 1, pp. 29–32, 1992.

[13] E. Breathe (Sheffield, "Your lungs and exercise," *European Respiratory Society*, vol. 12 (1), pp. 97–100, 2016.

[14] D. Fozoonmayeh, H. V. Le, E. Wittfoth, C. Geng, N. Ha, J. Wang, M. Vasilenko, Y. Ahn, and D. M.-k. Woodbridge, "A scalable smartwatch-based medication intake detection system using distributed machine learning," *Journal of medical systems*, vol. 44, no. 4, pp. 1–14, 2020.

[15] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[16] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers in neurorobotics*, vol. 7, p. 21, 2013.

[17] D. Dave, D. J. DeSalvo, B. Haridas, S. McKay, A. Shenoy, C. J. Koh, M. Lawley, and M. Erraguntla, "Feature-based machine learning model for real-time hypoglycemia prediction," *Journal of Diabetes Science and Technology*, p. 1932296820922622, 2020.

[18] E. LeDell and S. Poirier, "H2o automl: Scalable automatic machine learning," in *7th ICML workshop on automated machine learning*, 2020.

[19] R. K. Samala, H.-P. Chan, L. Hadjiiski, and S. Koneru, "Hazards of data leakage in machine learning: a study on classification of breast cancer using deep neural networks," in *Medical Imaging 2020: Computer-Aided Diagnosis*, vol. 11314. International Society for Optics and Photonics, 2020, p. 1131416.

[20] R. Howard, J. Guo, and K. D. Hall, "Imprecision nutrition? different simultaneous continuous glucose monitors provide discordant meal rankings for incremental postprandial glucose in subjects without diabetes," *The American Journal of Clinical Nutrition*, vol. 112, no. 4, pp. 1114–1119, 2020.