Multimodal Time-Series Activity Forecasting for Adaptive Lifestyle Intervention Design

Abdullah Mamun[1,2], Krista S. Leonard[1], Matthew P. Buman[1], Hassan Ghasemzadeh[1]

Abstract*—Physical activity is a cornerstone of chronic conditions and one of the most critical factors in reducing the risks of cardiovascular diseases, the leading cause of death in the United States. App-based lifestyle interventions have been utilized to promote physical activity in people with or at risk for chronic conditions. However, these mHealth tools have remained largely static and do not adapt to the changing behavior of the user. In a step toward designing adaptive interventions, we propose *BeWell24Plus*, a framework for monitoring activity and user engagement and developing computational models for outcome prediction and intervention design. In particular, we focus on devising algorithms that combine data about physical activity and engagement with the app to predict future physical activity performance. Knowing in advance how active a person is going to be in the next day can help with designing adaptive interventions that help individuals achieve their physical activity goals. Our technique combines the recent history of a person's physical activity with app engagement metrics such as when, how often, and for how long the app was used to forecast the near future's activity. We formulate the problem of multimodal activity forecasting and propose an LSTM-based realization of our proposed model architecture, which estimates physical activity outcomes in advance by examining the history of app usage and physical activity of the user. We demonstrate the effectiveness of our forecasting approach using data collected with 58 prediabetic people in a 9-month user study. We show that our multimodal forecasting approach outperforms single-modality forecasting by 2.2% to 11.1% in mean-absolute-error.

***Index Terms*—machine learning, physical activity, wearables, time-series forecasting**

I. Introduction

Mobile health (mHealth) interventions delivered through mobile apps have the potential to promote physical activity and reduce sedentary time as a means of reducing the risk of chronic diseases (e.g., cardiovascular disease, diabetes, and some cancers). There is some evidence suggesting a modest effect of these interventions in promoting healthy behaviors whereas some show no or limited effects [1]. One approach to improve the effectiveness of mHealth interventions is to offer real-time, personalized, and adaptive interventions. A potential direction for adaptive intervention design is to forecast a person's physical activity performance and offer interventions tailored toward the person's needs. Such a forecasting approach must not only take into account the person's past physical activity performance but also integrate information about the user's involvement with the mobile app that delivers intervention components (e.g., reminders, motivational feedback, goal settings). There are several ways researchers have defined user engagement with apps, or app usage, including qualitative interviews, system usage data (e.g., frequency of use, duration of use, number of log-ins, pages viewed, etc.), and sensor data [2]. However, to the best of our knowledge, the user engagement data have not been used to forecast a person's behavioral health outcomes.

Prior research on time-series forecasting focused primarily on using current data to predict future data of the same type. Examples include weather forecasting, stock-market prediction, traffic forecast, and energy requirements prediction utilizing a variety of machine learning models [3] [4]. The utility of neural network based models such as LSTM and ARIMA on stock market datasets has been studied previously [5]. Prior research also developed a regression tree to predict the time when a particular activity will take place in a smart home setting [6]. Nonetheless, none of these prior studies examined how engagement in an mHealth app can be combined with sensor-based metrics such as physical activity performance to predict future performance levels.

In this work, our main focus is on developing machine learning algorithms that combine physical activity and app engagement data to predict a person's future physical activity performance (e.g., next-day step counts). To this end, our approach examines how engagement with an mHealth app affects a person's physical activity and whether we can forecast physical activity with higher accuracy when we consider the app usage history. Understanding engagement with smartphone app intervention and its associations with physical activity aids in the development of personalized app-based treatments that maximize the efficacy of mHealth interventions.

Our contributions in this work are as follows: (i) we propose a framework for activity and user engagement monitoring and adaptive intervention design; (ii) we present a formal definition of the time-series activity forecasting; (iii) we propose an overall architecture for designing a machine learning algorithm for multimodal activity forecasting; and (iv) we show a realization of our proposed architecture based on an LSTM model; and (v) we demonstrate the effectiveness of our forecasting approach using data collected with 58 prediabetic people. Our results show that a suitable multimodal forecasting solution's mean-

This work was supported in part by the National Science Foundation, under grants CNS-2210133, CNS-2227002, IIS-1954372, and IIS-1852163, and also by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under the award number R18DK109516. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

[1]Authors are with the College of Health Solutions at Arizona State University, Phoenix, AZ 85054, USA email: {a.mamun, krista.leonard, mbuman, hassan.ghasemzadeh}@asu.edu

[2]Also affiliated with the School of Computing and Augmented Intelligence at Arizona State University

absolute-error can be 2.2% to 11.2% less than that of single-modality forecasting.

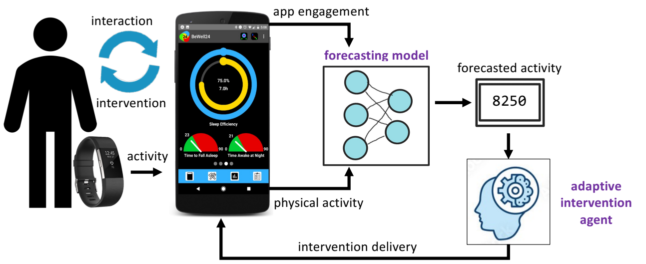


Fig. 1: The overall architecture of *BeWell24Plus*: the system combines lifestyle education, physical activity tracking, and machine learning for real-time health monitoring and adaptive physical activity intervention. The user interacts with the educational smartphone app while their physical activity is monitored continuously. The machine learning model uses data collected from the app and the fitness tracker to forecast the daily number of steps.

II. System Architecture & Machine Learning

We discuss the overall architecture of *BeWell24Plus* (shown in Fig. 1), formulation of the forecasting problem, and machine learning approaches for time-series activity forecasting.

A. Overview

BeWell24Plus builds on our work in designing *BeWell24* [7], a multi-component smartphone app intervention that targets behavior change in the 24h-spectrum (i.e., sleep, sedentary time, physical activity). As shown in Fig. 1, BeWell24Plus gathers data about the physical activity using an activity tracker (e.g., Fitbit). It also collects data about user engagement in various components of the lifestyle intervention app. The longitudinal data are then used to train a machine learning algorithm to forecast the user's activity performance (e.g., forecasting model). Our focus in this paper is on designing this forecasting module. Nonetheless, the forecasting model can be used to provide adaptive interventions through an intervention agent (e.g., another machine learning model) that adjusts the intervention components and parameters in the app based on the person's physical activity performance, the predicted performance, and other person-specific data.

B. Multimodal Forecasting

As discussed in Section I, it is possible to forecast a variable using the past values of the same variable with deep learning. We follow a similar approach, but in our case, we consider additional information to train a model and make an inference. In our case, the variable we want to forecast is an activity pattern, such as, the number of steps in a day, and the additional information contains the app usage history and other activity measures, such as time spent in sedentary activity, low-intensity physical activity (LPA), and moderate to vigorous physical activity (MVPA) in the previous days. Therefore, there are two broad categories of input data that we

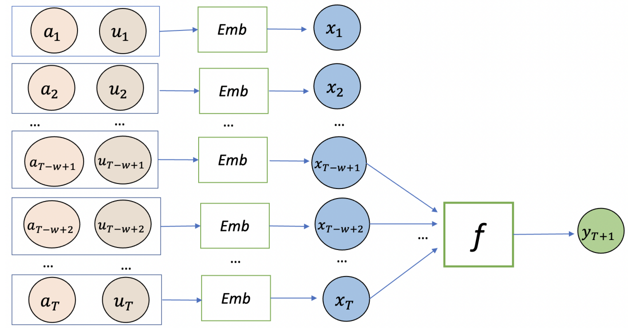


Fig. 2: Block diagram of our forecasting with past engagement and activity metrics. Here, for any t, a_t is the physical activity measure, u_t is the app usage measure, and x_t is the embedding of those two. Also, w is the window size, T is the current time step, and $T+1$ is the next time step. We want to forecast y_{T+1}, which is a physical activity outcome in the next time step.

use in our forecasting, physical activity data and engagement data.

Assume we are given k activity features, $a = (a^{(1)}, a^{(2)}, ..., a^{(k)})$, and p user engagement features, $u = (u^{(1)}, u^{(2)}, \ldots, u^{(p)})$, for a user. Let $\{a_1, a_2, ..., a_T\}$ be the values of a for time steps (e.g., days) 1 to T. Similarly, let $\{u_1, u_2, ..., u_T\}$ be the values of u during time 1 to T. Let w be the window length, i.e., number of previous days we are considering to make our forecasting. Our goal is to forecast the value of some physical activity outcome, y_{T+1} (e.g., one of the features in a).

Depending on what input features are used, various prediction models can be constructed. Using only activity features, the forecasting model will learn function f_a given by

$$y_{T+1} = f_a(a_{T-w+1}, a_{T-w+2}, ..., a_T) \tag{1}$$

Using only app engagement features, the forecasting model will learn function f_u as shown in

$$y_{T+1} = f_u(u_{T-w+1}, u_{T-w+2}, ..., u_T) \tag{2}$$

To design a forecasting model that uses both activity metrics and engagement features, we combine a_t and u_t to create x_t for every past value of t using an embedding function, Emb.

$$x_t = Emb(a_t, u_t), \;\; 1 \leq t \leq T \tag{3}$$

Then we forecast the activity outcome, y_{T+1}, according to the following equation.

$$y_{T+1} = f(x_{T-w+1}, x_{T-w+2}, ..., x_T) \tag{4}$$

This proposed forecasting approach incorporated into our BeWell24Plus framework is visualized in Fig. 2. To realize this forecasting approach, we propose to use an LSTM (Long Short-Term Memory) network. LSTM networks have shown promising results on time series classification tasks. These models capture long-distance dependencies from sequential data through the integration of memory cells and RNNs [8]. The forecasting models that can be obtained by training f_a and f_u in (1) and (2) will be used as two baseline methods in our analyses.

III. Validation Approach

A. User Study

We collected data from 58 people with prediabetic conditions in a 9-month clinical study. The participants were instructed to wear a Fitbit device for as much as they could throughout the day. For each user, we filter the data by removing the days with a daily wear time of less than 10 hours. We also exclude the users who have data for less than 10 days. After these steps, our dataset contains data from 54 participants (40 male, 10 female). Their average age is 56.13 ± 10.55, mean BMI is 34.91 ± 6.98. For our first set of experiments, we keep only the top 25% most active users, based on how many times they opened the app throughout the total intervention period. That leaves us with 14 super users in total. The concept of super users was inspired by [9]. However, all 54 users were involved in our second set of experiments.

B. Data Preprocessing

Initially, we have minute-level app usage and activity data for each user. For every minute, we have a binary value indicating whether a user accessed the app in that minute. Additionally, we have the number of steps and duration of sedentary activity, LPA, and MVPA in seconds in that minute. We convert the minute-level data to daily-level values by aggregating the minute-level features. A day is defined as the period from 12:00 AM to 11:59 PM according to the local time zone of the user.

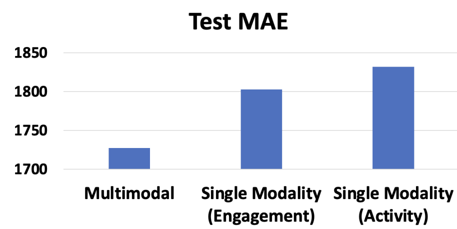
The final feature set is composed of engagement measures and activity measures. The engagement features have app usage metrics and a piece of additional information, 'day of the week'. Among the app usage features, there are total *minutes used* and *times opened* for a day. The same features have been created on an hourly basis as well, which implicitly embeds the 'time of the day' information. People usually have different schedules for weekends and weekdays. Therefore, we believe that the day of the week may influence the activity levels. The 'day of the week' feature is implemented by creating 7 one-hot encoded columns, one for each day of the week. For our experiments, 30 was chosen as the window size, which means, the machine learning model would take last 30 days' engagement metrics and physical activity metrics as input and forecast the next-day physical activity outcomes.

C. Forecasting Model

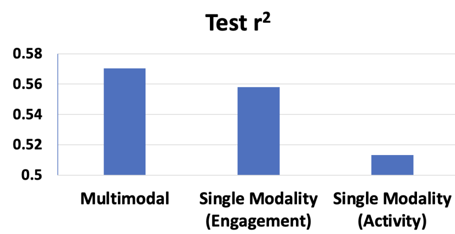
An LSTM layer with 100 neurons, followed by a dense layer with 50 neurons and the output neurons, was implemented for the forecasting. The model was trained for 20 epochs with the Adam optimizer for validating with each user. As there are 14 super users in total and each of them has to be on the validation set for 20 epochs, the total number of epochs the model is trained for is 14 x 20 = 280 for activity forecasting on super users. We divide our experiments into two categories, regression, and classification. For the regression experiments, mean-absolute-error (MAE) is reported, and for the classification problem, accuracy and f-score are reported for each threshold for the daily number of steps to walk. We vary the threshold from 5000 to 10000 with an increment of 1000. We use the leave-one-subject-out validation method in all our experiments.

D. Baseline Models

Our two baselines are the the performance of the models trained with only past activity data, and the same of the models trained with only engagement data. The first baseline is similar to traditional forecasting models that use past data to predict future data of the same type (i.e., here physical activity). The second baseline is inherently different because it uses data of different modalities for input and output.



(a) mean-absolute-error



(b) r^2 value

Fig. 3: Comparing the performance of the regression-based LSTM forecaster for different modalities on super users.

IV. Results and Discussion

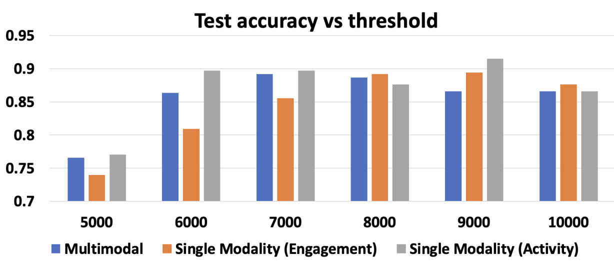
A. Activity Forecasting on Super Users

Fig. 3 shows the mean-absolute-error (MAE) and r^2 values of activity forecasting for multiple modalities (i.e., all features) and single modality (i.e., either engagement features or activity features) inputs using data from super users. When both physical activity and engagement features are used, the forecasting model performs better on the regression problem where the machine learning model estimates the number of steps that the user will walk the next day. As shown in Fig. 3a, the forecasting approach achieves MAE values of 1727, 1803, and 1832 for multimodal, engagement-based single-modal, and activity-based single-modal prediction algorithms, respectively. These results suggest that multimodal forecasting reduces the mean-absolute-error of next-day step counts prediction by 4.2%–5.7%. Similarly, as shown in Fig. 3b, the value of r^2 is 0.57,

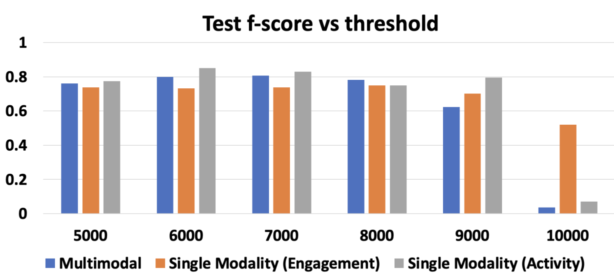
TABLE I: Performance of the LSTM model for activity forecasting on all users.

Metric	*Multimodal*	*Single Modality (Engagement)*	*Single Modality (Activity)*
MAE	**2081**	2372	2090
r^2	0.350	0.267	**0.351**

0.56, and 0.51 for multimodal, engagement-only, and activity-only forecasting cases, respectively. This suggests a 2.2% improvement in r^2 value from the model with engagement-only input features and an 11.1% increase in the r^2 value from the model with activity-only features.



(a) accuracy



(b) f-score

Fig. 4: Performance of the classification-based LSTM forecasting model for different modalilties on super users. The threshold for classification is varied from 5000 to 10000 steps.

Fig. 4 shows accuracy and f-score results for the multimodal model, single modality model with engagement features, and single modality model with activity features. As Fig. 4b shows, for the classification task of predicting a next-day step count of 8000 steps, the multimodal forecasting model achieves the highest f-score (i.e, 0.782), which is a 4.3% improvement in the f-score of the two single modality model (i.e., 0.750). We, however, note that for other step count thresholds (e.g., predicting a next-day step count of 6000 steps), the results are not consistent with these findings, which suggests further research in this area.

B. Activity Forecasting on All Users

Table I presents the result for forecasting on all 54 users. Similar to Fig. 3a, we notice a performance improvement in the multimodal approach. The multimodal forecasting reduces the mean-absolute-error (MAE) by 0.4%–12.27%. Although the r^2 coefficient of the multimodal approach was 0.28% lower than that of the activity-only approach, it is noteworthy that r^2 coefficients can be adversely affected by outlier data points.

Finally, we note that overall, the forecasting on super users was more effective than forecasting on all users for our dataset. One possible reason for that observation can be the existence of users who had very low engagement overall.

V. Conclusions and Future Work

Time-series forecasting has been an active research area with applications in domains including weather, traffic, and the stock market. We formulated the problem of time-series forecasting for activity prediction in the mHealth domain. We presented an approach for multimodal forecasting that combines behavioral data with user engagement with an intervention app to provide insight into the relationship between app usage and physical activity performance. Our study covers fitness tracking, lifestyle intervention apps, and time-series forecasting with neural networks. From the experiments, we found that it is possible to forecast physical activity outcomes with past app engagement and physical activity patterns. Our findings imply the existence of a relationship between app adherence and the physical activity of a person. The proposed forecasting model can be used in a closed-loop system to provide continuous, real-time, and personalized lifestyle interventions in mHealth. Our future work will focus on realizing such a closed-loop system in active intervention studies.

References

[1] R. Gal, A. M. May, E. J. van Overmeeren, M. Simons, and E. M. Monninkhof, "The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis," *Sports Medicine - Open*, vol. 4, no. 1, Sep. 2018.

[2] C. E. Short, A. DeSmet, C. Woods, S. L. Williams, C. Maher, A. Middelweerd, A. M. Müller, P. A. Wark, C. Vandelanotte, L. Poppe, M. D. Hingle, and R. Crutzen, "Measuring engagement in eHealth and mHealth behavior change interventions: Viewpoint of methodologies," *Journal of Medical Internet Research*, vol. 20, no. 11, p. e292, Nov. 2018.

[3] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long- and short-term temporal patterns with deep neural networks." ACM, Jun. 2018.

[4] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, p. 20200209, Feb. 2021.

[5] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of arima and lstm in forecasting time series," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*. IEEE, 2018, pp. 1394–1401.

[6] B. Minor and D. J. Cook, "Forecasting occurrences of activities," *Pervasive and mobile computing*, vol. 38, pp. 77–91, 2017.

[7] M. P. Buman, D. R. Epstein, M. Gutierrez, C. Herb, K. Hollingshead, J. L. Huberty, E. B. Hekler, S. Vega-López, P. Ohri-Vachaspati, A. C. Hekler, and C. M. Baldwin, "BeWell24: development and process evaluation of a smartphone "app" to improve sleep, sedentary, and active behaviors in US veterans with increased metabolic risk," *Translational Behavioral Medicine*, vol. 6, no. 3, pp. 438–448, Nov. 2015.

[8] P. Alinia, S. Parvaneh, S. I. Mirzadeh, A. Arefeen, and H. Ghasemzadeh, "Boosting lying posture classification with transfer learning," in *The 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2022)*, 2022.

[9] S. Edney, J. C. Ryan, T. Olds, C. Monroe, F. Fraysse, C. Vandelanotte, R. Plotnikoff, R. Curtis, and C. Maher, "User engagement and attrition in an app-based physical activity intervention: Secondary analysis of a randomized controlled trial," *Journal of Medical Internet Research*, vol. 21, no. 11, p. e14645, Nov. 2019.