# Characterizing Physical Activity in a Health Social Network

Javid Ebrahimi
University of Oregon
javid@cs.uoregon.edu

NhatHai Phan
University of Oregon
haiphan@cs.uoregon.edu

Dejing Dou
University of Oregon
dou@cs.uoregon.edu

Brigitte Piniewski
PeaceHealth Laboratories
bpiniewski@peacehealthlabs.org

David Kil
HealthMantic, Inc
david.kil@healthmantic.com

## ABSTRACT

New horizons are emerging within healthcare delivery, education, intervention provision, and tracking. We study a health social network that has tracked physical activities, biomarkers, and posts the participants have shared, throughout a one-year program. The program was aimed at helping people to adopt healthy behaviors and to lose weight. In this paper, we focus on users' posts that relate to physical activities. Prior papers characterize health based solely on users' information disclosed through natural language or questionnaires. The drawback of these works is their lack of medical records or health-related information to validate their findings. By contrast, with our direct access to users' physical and medical data, we investigate the implication of users' posts at both individual and group levels. We are able to validate our hypotheses about the effects of certain social network activities, by contextualizing them in the specific users' actual medical progress and documented levels of exercise. Our findings show that activity self-disclosure posts are good indicators of one's real-world physical activity, which makes them good resources for monitoring the participants. In addition, using a physical activity propagation model, we show how these posts can influence the physical activity behavior at the network level. Further, posts exhibit distinctive affective, biological, and linguistic style markers. We observe that these characteristics can be used in a predictive capacity, to detect positive activity signals with $\sim 88\%$ accuracy, which can be utilized for an unobtrusive monitoring solution.

## Keywords

health social network; topic modeling; physical activity propagation; classification

## 1. INTRODUCTION

Regular engagement in physical activity reduces the risk of developing cardiovascular disease, diabetes, obesity, some cancers, and other chronic conditions [1]. The estimated annual healthcare cost of obesity-related illness is a staggering $190.2 billion, nearly 21% of annual medical spending in the United States [9]. Public health standards recommend that adults participate in at least 30 minutes of moderate-intensity physical activity on five or more days a week [16]. However, less than 50% of the adult population meets these standards in many industrialized countries [3]. Thus, monitoring physical activity, characterizing healthy behaviors, and spreading them to a large population are important.

Recent technological advancements have opened new horizons for healthcare delivery. The Internet is identified as an important conduit for physical activity behavior suggestions and interventions, engaged through newsletters, emails, patient portals, etc. [20]. Several studies have shown the positive impact of these internet-based interventions [29]. In addition, numerous mobile apps have provided new opportunities to support healthy behaviors, through lifestyle monitoring and through tracking and recording of physical activities. Studies have found that 19% of smartphone owners have at least one health app on their phone; the most popular of which are exercise, diet and weight apps [23]. Another important development is the growth of online social media. One-third of consumers use social media for health-related matters, including sharing information about symptoms, drugs, and medical devices [23].

Utilizing these technologies, our recent study, named YesiWell, was conducted in 2010-2011 as a collaboration between PeaceHealth Laboratories, SK Telecom Americas, and University of Oregon. The goal of the study was to help overweight participants adopt healthy lifestyles and lose weight. We recorded daily physical activities, social activities (e.g., text messages), biomarkers (e.g., weight), and biometric measures (e.g., triglyceride level) for a group of 254 individuals who formed a social network. Physical activities were measured by a mobile device carried by each user, called h-pod. All users enrolled in an online social network application allowing them to make friends with and communicate to each other. Biomarkers and biometric measures were recorded via monthly medical tests performed at our labs.

One of the ways people communicate with each other on our social network is through public posting. People ask questions, support each other, express their feelings, and disclose their physical activities. We found that activity disclosure posts are good indicators of individual behavior and group-level behavior. Specifically, we distinguish posts that contain positive signals for physical activity (e.g., walking, running) from those that contain negative signals (e.g., sickness). We address the following three research questions:



*DH '16, April 11-13, 2016, Montreal, QC, Canada*
 ISBN 978-1-4503-4224-7/16/04…$15.00
DOI: http://dx.doi.org/10.1145/2896338.2896349

**RQ1:** Can we quantify the individual-level and group-level behaviors related to these posts?

**RQ2:** What are the linguistic and psycholinguistic attributes of positive and negative posts?

**RQ3:** Can these attributes be used to predict whether a future post is positive or negative?

To study the individual-level behavior, we show that positive posts can point to an increase of physical activity compared to (the average of) previous and future days. And likewise, a negative post can indicate a decrease in the activity compared to (the average of) previous and future days. To study the group-level behavior caused by the posts, we build on our previous research [25] that demonstrated how social influence can contribute to physical activity propagation. Findings of the current paper show that there is a significant difference between positive and negative posts in how they influence other members of the network. Specifically, a person's positive post has a significantly higher probability of propagating physical activity to other people in the network than does a negative post.

To address the second research question, we characterize the content of the posts. We use different techniques and lexicons. Using the psycholinguistic lexicon, LIWC[1], we find that the posts can be differentiated significantly by affective and biological processes. After running topic modeling [5], we also observe that there exist significant differences in topic distribution between the two groups of posts. In addition, making use of the MPQA[2] lexicon, we find that positive posts express stronger subjective words.

To answer the last research question, we combine the linguistic, and the psycholinguistic characteristics. We are able to predict whether a post is positive or negative with $\sim 88\%$ accuracy. We do this via supervised learning with a linear-kernel Support Vector Machine (SVM), and we perform 10-fold cross-validation. This model could be used for similar programs run in the future; we can monitor users' posting behavior, and we can deliver, therefrom, appropriate interventions to improve their levels of physical activity. Monitoring through text analysis can compensate for intervals in which pedometers are not utilized or are not accurate [7].

Our work has a fundamental distinction from similar studies that characterize health in social networks. In previous works, detection of the target population is contingent on self-reporting or affiliation (e.g., use of a particular hashtag). Our work is the first that studies the diction and social attributes of a population in conjunction with their actual health-related data.

# 2. RELATED WORK

## 2.1 Physical activity intervention

Website-delivered physical activity interventions have the potential to overcome many of the obstacles associated with traditional in-person exercise counseling or group-based physical activity programs. An Internet user can seek advice at any time, any place, and often at a lower cost compared with other delivery modalities [8]. In [29], fifteen website-delivered intervention programs that used the Internet or e-mail were evaluated. Improvement in physical activities was reported in eight. Better outcomes were identified when interventions had more than five contacts with participants and when the time to follow-up was short ($\leq$3 months; 60% positive outcomes), compared to medium-term (3-6 months, 50%) and long-term ($\geq$6 months, 40%). Phan et al. [25] introduced a hierarchical approach to analyze physical activity propagation through social communications at the community level. Their findings show that: (a) social networks have great potential to propagate physical activities via social communications; and (b) physical activity-based influence behavior has a strong correlation to health outcome measures such as BMI. We later showed that the influence of a post is related with the distribution of the topics in the content of the post [26]. Moreover, in [24] a deep learning model for human behavior prediction in health social networks was proposed. By incorporating all the human behavior determinants which are self-motivation, implicit and explicit social influences, and environmental events, the model can predict the future activity levels of users more accurately and more stably than conventional methods.

## 2.2 Natural language and health

The abundance of personalized data on social media has sparked researchers to study it for healthcare applications. A common topic in this line of research is how computational techniques may be applied to the data that people share on online social media to understand their health behaviors.

Self-reported diagnoses have been used to examine mental health conditions, such as depression, PTSD, bipolar disorder, and seasonal affective disorder [13, 12], and physical health conditions, such as asthma and diabetes [30], and flu [27]. Kumar et al. [19] studies the Werther effect, which describes the increased rate of completed or attempted suicides following the coverage of a celebrity's suicide in the media. Additional research investigates the use of online platforms, in discourse around pro eating disorder websites [6]. Similarly, De Choudhury [10] characterizes pro-anorexia and pro-recovery communities on Tumblr.

There are works that have studied users over a period of time. For instance, factors impacting abstinence and relapse during cessation attempts by smokers were identified in [22], after analyzing two years of their data. Similarly, the impact of pregnancy on new mothers was studied in [14].

Online health communities have been studied to identify the key features in interactions between the people who seek help and those who can influence and support them. Two types of support are exchanged in these communities: informational, and emotional. Biyani et al., [4] found that the sets of posts by the most prolific users had a higher percentage of emotional support than those of less prolific users. Zhao et al., [31] showed that the sentiment expressed in emotional support can be used to detect influential users in a health community.

# 3. DATA

Our dataset consists of a social network of 254 participants, in which 78% of the participants were female, and the median age was 51. We have access to their weekly physical activity data, in terms of the number of steps they have taken, and monthly biometric and biomarker metrics over ten months, from October 2010 to August 2011. The physical activity data, collected by a special electronic de-

[1]http://www.liwc.net/

[2]http://mpqa.cs.pitt.edu/

| |
|---|
| Taking a mini horse jogging sure helps get those steps in, I'll have to do that more often! Tomorrow I start Zumba and I am sure I'll get my steps in there. |
| Ran a new route yesterday: 2.8 miles in 30:48. Need to break out the GPS again to try another route to get up to 3 miles. |
| Alas, out with an infection... Antibiotics and rest ordered. I'm readjusting my goals. |
| Had cataract surgery on Monday. didn't get much exercise in that day. |

Table 1: Examples of positive and negative activity posts.

| | |
|---|---|
| Number of activity disclosure posts | 265 |
| Number of users | 63 |
| Number of positive posts | 199 |
| Number of negative posts | 66 |
| Number of users with only positive posts | 28 |
| Number of users with only negative posts | 11 |
| Number of users with both types of posts | 24 |
| Average length of a post in words | 11 |

Table 2: Data statistics

vice (h-pod) worn by each user, includes such information as the number of steps taken. The social network application provides friendship and posting capabilities. We only study the activity disclosure-related posts, examples of which are shown in Table 1. As we locate disclosure-related posts, we include original posts and comments upon posts, and we value them equally.

Out of the 1601 posts, we collected 265 posts related to activity disclosure. Two human annotators labeled these posts independently, and a third human annotator resolved any mismatches. For each of the 265 posts, we removed the stop words and the words that had appeared fewer than two times in the dataset which leads to a small vocabulary of 529 words. The total number of users authoring these posts was 63, with similar demographic distribution as the original data. Statistics summarizing this data are presented in Table 2.

# 4. RQ1: BEHAVIOR IMPLICATIONS

## 4.1 Individual-level behavior

We show that positive posts can point to an increase of physical activity compared to the average of previous and future days. And likewise, a negative post can indicate a decrease in the activity compared to (the average of) previous and future days.

We take the average number of steps taken by the participant in $2w+1$ days between days $d-w$ and $d+w$ where $d$ is the date of the posting and $w$ is the window size (set to three). We call this measure $avg(S_w)$. We also take the average of the number of steps taken in $2w$ days from $d+w+1$ to $d+w+w$ and from $d-w-w$ to $d-w-1$. We call this measure $avg(S_o)$. We posit that with a positive post, we are more likely to have $avg(S_w) > avg(S_o)$, whereas with a negative post, we are more likely to have $avg(S_w) < avg(S_o)$. On aggregate, positive and negative posts reflect a change in the activity more than randomly selected dates on which no such posting exists.

In order to study this, we use two random sets of control dates. In one (random1), we only consider the users with self-disclosure posts and select some random dates that do

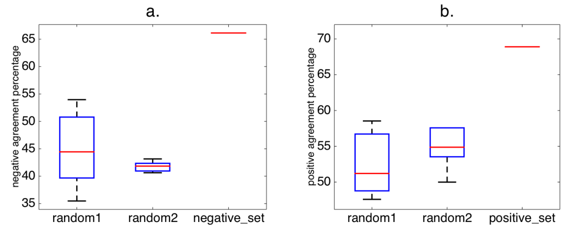


Figure 1: Comparing the dates with self-disclosure posts and two random dates. a) Percentage of dates where $avg(S_w) < avg(S_o)$ holds for our positive set and the two random sets. b) Percentage of dates where $avg(S_w) > avg(S_o)$ holds for our negative set and the two random sets.

not overlap with our dates selected for study. In the second case (random2), we use random dates from random users. To avoid selection biases, the second group is selected such that the population has similar age and gender distribution as the users who had self-disclosure posts.

Figures 1.a and 1.b compare the three types of dates: dates with self-disclosure posts, random1, and random2. For the random groups, we pick a number of instances equal to the number of instances in the first group and report the results of 20 runs, with 95% confidence. In both positive and negative cases, dates selected based on the polarity of the post reflect the activity behavior of the user much better than randomly picked dates. In the positive case, for 68.9% of the posts, $avg(S_w) > avg(S_o)$, whereas in the negative case, in 66.1% of the times $avg(S_w) < avg(S_o)$ holds. In other words, people report their activity/inactivity and express positive/negative sentiment when there exists a relative increase/decrease in their level of activity.

## 4.2 Group-level behavior

The next question we address is whether these posts can result in different behavior at the group-level. In other words, do these posts impact physical activities of other participants? To investigate this, we must equip ourselves with methods that can model influence in a social network. To model physical activity propagation, Phan et al. [25] introduced the Community-level Physical Activity Propagation (CPP) model, which was inspired by the ideas of the Independent Cascade (IC) model [17], and the Community-level Social Influence (CSI) model [21]. Similar to our Topic aware Community-level Physical Activity Propagation (TaCPP) model [26], we take the content of the posts into account. However, to study the physical activity of the network better, the current work is based on the labels of activity-disclosure posts and not the topics of all the posts available.

The building block for activity propagation modeling is the definition of a trace [25]. Assume that at time $t$, user $v$ makes a posting on which user $u$ comments; given a $\Delta t$, $v$ is considered to *activate* $u$ at time $t$ if the total number of (walking & running) steps of $u$ in $[t, t+\Delta t]$ is larger than or equal to the total number of steps of $u$ in the past period $[t-\Delta t, t]$. Given a chain of users $\alpha = \{U_1, \ldots, U_n\}$ such that $U_i$ is a set of users, $U_1 \cap U_2 \cap \ldots \cap U_n = \emptyset$; $\alpha$ is called a single trace if $\forall i \in [1, n-1], \forall u \in U_{i+1}$ is activated by some user $u' \in U_i$ such that $t_\alpha(u) \in [t_\alpha(u'), t_\alpha(u') + w]$ where $t_\alpha(u)$ is the *activation time* of $u$ in $\alpha$. In this work, $U_1$ can be a user instead of a set of users, and the chain of users is given by the people who have posted or commented on a post.

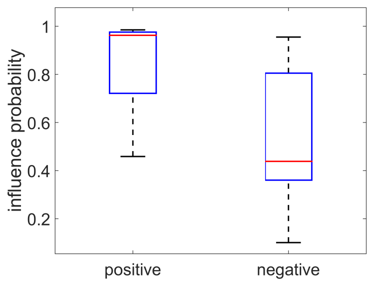


Figure 2: Activity propagation based on the type of posts.

Let $I_{\alpha,u}$ be the set of users that potentially influence $u$'s activation in the trace $\alpha$:

$$I^{+}_{\alpha,u} = \{v|(v,u) \in E,\ \textit{iff } u \in U_i \textit{ then } v \in U_{i-1}\}$$

Similarly, we define the set of users who clearly failed in activating $u$ in the trace $\alpha$:

$$I^{-}_{\alpha,u} = \{v|(v,u) \in E,\ \textit{iff } v \in U_{i-1} \textit{ then } u \notin U_i\}$$

Let $p : V \times V \rightarrow [0,1]$ denote a function that maps every pair of nodes to a probability. By the i.i.d assumption, the log likelihood of the traces in $D$ given $p$ can be defined as:

$$\log \mathcal{L}(D|p) = \sum_{\alpha \in D} \log \mathcal{L}_\alpha(p)$$

Each $v \in I^{+}_{\alpha,u}$, $v$ succeeds in activating $u$ on the considered trace $\alpha$ with probability $p(v,u)$ and fails with probability $1 - p(v,u)$. Following [21, 26], we define $\gamma_{\alpha,v\rightarrow u,l}$ as users' responsibility which represents the probability that in trace $\alpha$, the activation of $u$ was due to the success of the activation trial performed by $v$ with a post with label $l$. By using $\gamma_{\alpha,v\rightarrow u,l}$, we can define the likelihood of the observed propagation as:

$$\mathcal{L}_\alpha(p) = \prod_{u \in V}\Big[1 - \prod_{v \in I^{+}_{\alpha,u}} (1 - p(v,u)^{\frac{\sum_{l \in L} m_{l,v\rightarrow u}\gamma_{\alpha,v\rightarrow u,l}}{Z(\alpha,v\rightarrow u)}})\Big] \times \Big[\prod_{v \in I^{-}_{\alpha,u}} (1 - p(v,u)^{1-\frac{\sum_{l \in L} m_{l,v\rightarrow u}\gamma_{\alpha,v\rightarrow u,l}}{Z(\alpha,v\rightarrow u)}})\Big] \quad (1)$$

where $m_{l,v\rightarrow u}$ is the number of posts with label $l$ posted by $v$, and commented on by $u$ in the propagation trace $\alpha$, and $Z(\alpha, v \rightarrow u)$ is a normalization function.

Using Expectation Maximization, we can derive $p$ and $\gamma$. For the sake of brevity, we just give the influence probability: the probability that a participant $u$ will improve their level of physical activity after commenting on a post with the label $l$, posted by user $v$.

$$p(v \rightarrow u|l) = \sum_{\alpha \in D} \frac{m_{l,v\rightarrow u}\gamma_{\alpha,v\rightarrow u,l}}{\sum_{z \in I^{+}_{\alpha,u} \cup I^{-}_{\alpha,u}} \sum_{l \in L} m_{l,z\rightarrow u}\, \gamma_{\alpha,z\rightarrow u,l}}$$

Out of the 265 posts in our dataset, 44 positive and 34 negative posts were found in propagation traces. We compare the influence probability of these posts in Figure 2 with 95% confidence. Based on Mann-Whitney U tests, positive posts wield much more significant influence than the negative ones ($p < 0.0001$, $z = 5.8$).

| Topic-1 | Topic-2 | Topic-3 |
|---|---|---|
| day | **miles** | back |
| today | time | days |
| work | **run** | walk |
| great | today | **sick** |
| good | h-pod | week |
| time | **mile** | work |
| **step** | day | day |
| **2** | **walk** | **frustrate** |
| long | yesterday | feel |
| **3** | days | hard |
| **level** | **jog** | **bad** |
| **run** | morning | **shoulder** |
| food | good | weight |
| tired | **goal** | **therapy** |
| **team** | pace | game |
| **happy** | **trail** | **antibiotics** |
| **join** | Friday | vacation |
| **treadmill** | **uphill** | normal |
| night | **:)** | miss |
| **motivate** | **gym** | **cold** |

Table 3: Top words for topics extracted from the posts.

# 5. RQ2: CHARACTERISTICS

To characterize the language in the posts, we apply a variety of language-based feature extractors. We perform topic modeling [5] to find the topic distribution of each user in the dataset. We also use higher-level features extracted by LIWC categories [11] and a subjectivity lexicon, MPQA corpus [28]. LIWC,which has received much attention in the literature, provides affective, social, biological, linguistic style, and other measures. We use the subjectivity lexicon compiled from the MPQA to get weak and strong subjective words. This corpus also includes the polarity of the words.

## 5.1 Topical processes

We perform Latent Dirichlet Allocation (LDA) and extract three topics from the posts. LDA finds topic distributions for each word, and each post in the corpus, for a given number of topics. We set the number of topics to three. Table 3 summarizes the top 20 words for each topic, with boldface used for the words that form a coherent topic. Based on Mann-Whitney U tests, positive posts contain more of the first two topics (positive z), and contain less of the third topic (negative z), with significant differences across topics two and three. See Table 4 for details.

Clearly, the third topic contains more negative signals for physical activity (e.g., therapy, antibiotics, cold, etc.). It also includes more negative sentiment compared with the other two topics (e.g., frustrate, bad). Topic two contains more physical activity-related words (e.g., run, jog, mile, etc.). The first topic has some information about the program and the social network application that participants used. Words such as *team* and *join* refer to the competition feature of the application. For the competitions, people joined teams and competed with each other, in terms of their collective levels of physical activity. In addition the word *level*, and numbers *2*, *3* refer to the progress that participants make, advancing to a higher level of the program according to the number of steps they have walked or ran.

## 5.2 Psycholinguistic processes

Table 5 shows the differences in how lexical density, affective processes, biological processes, relativity, and personal concerns are manifested in positive and negative posts. The

| | POS | NEG | z | p |
|---|---|---|---|---|
| Topic1 | 0.33 | 0.26 | 1.74 | - |
| Topic2 | 0.46 | 0.18 | 4.03 | *** |
| Topic3 | 0.21 | 0.56 | -5.46 | *** |

* $p \leq .01$; ** $p \leq .001$; *** $p \leq .0001$

Table 4: Comparison of positive and negative posts along topics based on Mann-Whitney U tests.

| | POS | NEG | z | p |
|---|---|---|---|---|
| **Lexical density** | | | | |
| future tense | 0.0004 | 0.0027 | -2.92 | * |
| adverbs | 0.006 | 0.0120 | -2.54 | * |
| **Biological processes** | | | | |
| health | 0.0145 | 0.0305 | -3.8 | ** |
| **Relativity** | | | | |
| motion | 0.0638 | 0.0341 | 4.38 | *** |
| **Personal concerns** | | | | |
| work | 0.0355 | 0.0472 | -2.13 | - |
| home | 0.0165 | 0.0334 | -4.44 | *** |
| **Affective processes** | | | | |
| anxiety | 0.0037 | 0.0166 | -4.66 | *** |
| sadness | 0.0098 | 0.0261 | -4.70 | *** |
| weak-subjective | 0.50 | 0.66 | -3.73 | ** |
| strong-subjective | 0.50 | 0.34 | 3.79 | ** |
| negative-polarity | 0.26 | 0.5 | -5.17 | *** |
| positive-polarity | 0.45 | 0.32 | 2.14 | - |

Holm-Bonferroni correction:
* $p \leq .01$; ** $p \leq .001$; *** $p \leq .0001$

Table 5: Comparison of positive and negative posts along attributes of psycholinguistic processes. The categories that are shown are significantly different ($p < 0.05$) based on Mann-Whitney U tests.

table contains only the categories that were significantly different ($p < 0.05$).

In positive posts, users exhibit higher positive polarity ($z = 2.14, p < 0.05$) and stronger subjectivity ($z = 3.79, p < 0.001$). This is due to the satisfaction, and joy gained after successful physical activity, and progress in the program. This can also be seen in the higher value ($z = 4.38$, $p < 0.0001$) in the motion category.

In negative posts users express higher negative polarity ($z = 5.17$, $p < 0.0001$), anxiety ($z = 4.66$, $p < 0.0001$), weak subjectivity ($z = 3.73$, $p < 0.001$), and sadness ($z = 4.66$, $p < 0.0001$). This could be attributed to their sense of underachievement, and negativity towards the outcome of the program. They also express more personal concerns with regard to work ($z = 2.13$, $p < 0.05$), and home ($z = 4.44$, $p < 0.0001$), which are factors that hinder their physical activity. This can also be seen in the higher values for future tense ($z = 2.92$, $p < 0.01$), and adverbs ($z = 2.54$, $p < 0.01$), as they mention possible physical activities at a later time. Physical health factors such as sickness are mentioned more in negative posts, as can be seen in the biological processes' health category ($z = 3.8$, $p < 0.001$).

# 6. RQ3: PREDICTION

To distinguish between positive and negative posts, we used a Support Vector Machine (SVM) with linear kernel with 10-fold cross-validation for evaluation. Table 6 shows

| Feature set | Accuracy |
|---|---|
| unigram | 83.10 |
| unigram+LDA | 86.09 |
| unigram+LDA+MPQA | 87.22 |
| unigram+LDA+MPQA+LIWC | 87.94 |

Table 6: Classification of positive and negative posts.

| |
|---|
| Oh yay, so nice to be back to feeling well and racking up the steps. |
| Still feel a bit weak and tired but glad to be back to getting physical! Sweat Out the rest of those germs!!! |
| I Didn't get sick, just neasea...It is the end of the third day and I am able to at least walk but not at my usual pace. |

Table 7: Three examples of false negatives.

how the accuracy of the model increases by enriching the feature set. The dataset is imbalanced, in which the number of positive instances is three times larger than the negative instances. As a result, 64% of the errors are false positives. In addition, many of the false negatives belong to the posts that denote restarting one's physical activity after a break . Such posts are inherently difficult to classify, a couple of which are shown as examples in Table 7.

For the sake of completeness, we sampled 265 posts from the posts that were not related to activity; these posts include technical questions about the software or the h-pod, diet-related information, etc. We also combined the positive and negative posts to form an activity-related class. We then performed a binary classification to distinguish between the activity-related posts and the rest. The accuracy with unigram features, linear-kernel SVM, and 10-fold cross-validation was 85.20%.

# 7. CLINICAL IMPLICATIONS

We study the relationship between the participants' posts and their health outcomes. Particularly, we want to see if there is a difference between the health outcomes of participants with positive and those with negative posts. To address this question, we first demonstrate the relationship between participants' self-disclosure through posts with their actual physical activities. Then we make use of a health-related score, derived from the health information to which we have access. We compare two groups of people with activity disclosure posts: people with only positive posts (28 people), and people with only negative posts (11 people), excluding the users with both types of posts.

Figure 3 reports the average number of steps taken by these two groups of people, over the course of the program, with a 95% confidence interval. Users who had positive posts (pos) performed better than those who had negative posts (neg). They took 23794 additional steps on average, which represents approximately 12 miles of walking or jogging. This shows a relationship between the language of the participants and their real world physical activity.

Wellness Score [18] is a weekly composite score of one's health based on physical activity, biomarkers, and biometric measures. The Wellness Scores are then recomputed via a Markov Chain Monte Carlo sampler such that the new scores mimic percentile ranking. For instance, a Wellness Score of 90 means among the top 10%.

Figure 4 shows the average weekly Wellness Score of the group with only positive posts, and that of the group with

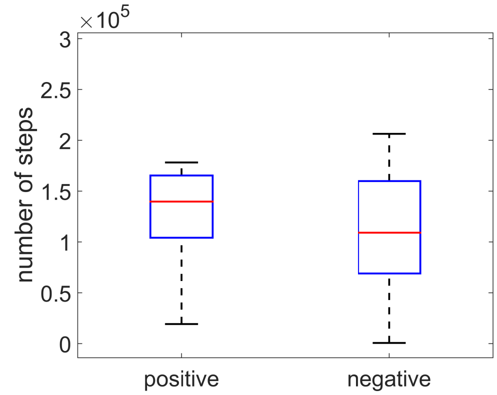


Figure 3: Comparing the number of steps taken by the two groups of people.

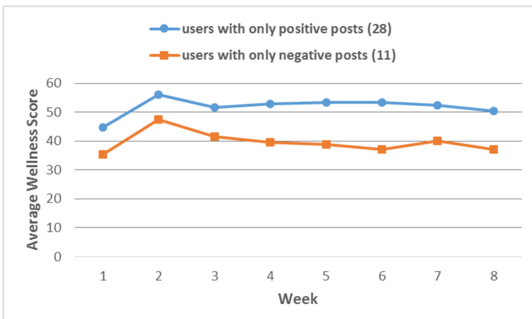


Figure 4: Weekly Wellness Score for the group of users with only positive or only negative.

only negative posts, over a two-month period. There is a consistent pattern wherein the former performs better than the latter, which shows the importance of intervention for the users with negative posts.

Investigating the effectivness of the program for each group is not the focus of this paper. Figures 3-4 just show there is a correlation between people's posts that are publicly shared, and their physical activity and health. However, we can give some pointers as to why such differences may, in fact, be due to bias. First of all, since the participants may be (or become) real-world friends and even do physical activity in groups, they can take some of their conversation to the social network. This increases the correlation between the positive posts and corresponding real-world activities. In addition, users with positive posts might have higher self-motivation, and exercise more in general.

Our predictive model could be used for significant clinical benefits. We can monitor users' posting behavior, and thereby, we can become able to tailor appropriate interventions to improve their levels of physical activity. By including affective and psycholinguistic factors in our analysis, we not only gain higher accuracy, but we also may be able to help identify surmountable obstacles to progress in programs of clinical benefit administered via, or complemented with, social networks. For example, is the nature of a given post with positive signals for physical activity one of simple self-disclosure, encouragement of others, or competition with others, and how might those differently affect fellow participants in a social network?

# 8. CONCLUSIONS AND FUTURE WORK

Two thirds of the US population are now overweight or obese [2]. This incurs significant health risks and financial costs to society. Traditionally, support groups and other social reinforcement approaches have been popular and effective in dealing with unhealthy behaviors including overweight. Of the factors associated with sustained weight loss, one of the most important is continued intervention with frequent social contacts. The YesiWell study, conducted in 2010-2011, was designed to help overweight participants adopt healthy lifestyles and lose weight through interaction with other participants, facilitated with mobile and social network technologies.

In this work, we only focused on the activity self-disclosure posts that participants had publicly shared throughout the program. These posts can contain either positive or negative signals for activity. Ignoring technical glitches in recording the number of steps, we found that these posts are good indicators of one's real-world physical activity. We found several significant differences between the positive and negative posts in terms of their linguistic, affective, and biological processes. Interestingly, there is a significant difference between positive and negative posts in how they influence other members of the network. Concretely, positive posts can lead other participants to increase their physical activity considerably more than the negative posts.

We propose the following two future directions:

- The current work is limited by the number of posts, and the number of people in the study. However, we can measure physical activities that people self-report on social media through the language. Employing affiliation techniques [15], we can find people who are using physical activity apps, and who are actively discussing them on social media. This would help us conduct a similar study on a large scale, and to be more inclusive demographically. Even without the access to people's health-related information, e.g., actual lab results and other biomarker measures, we can extract useful information about the physical well-being of communities from the social media. Using the network structure, we can also study the impact of social influence for physical activity.

- Delineating use cases by tone and focus could also identify opportunities to moderate content and/or to tailor interventions, such as to note if participants need any emotional, logistic, or informational supports. Is the nature of a given post with negative signals of physical activity sad, angry, matter-of-fact, etc., and how might those differently affect fellow participants? If they are sad, that could elicit messages of emotional support. If they are angry about, e.g., some piece of health monitoring equipment not functioning correctly, that could elicit instructions from tech support. If they are simply commenting on not knowing how to progress further, sending health education information might be appropriate; and so forth.

# 9. ACKNOWLEDGMENTS


This work is supported by the NIH grant R01GM103309. We acknowledge Ellen Klowden for her contributions to the manuscript. We also thank anonymous reviewers for their constructive comments, which helped improve the paper.

## 10. REFERENCES


[1] Physical activity and health: A report of the Surgeon General. *U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion*, 1996.

[2] Overweight and obesity - epidemiology forecast to 2022. *EpiCast Report*, 2013.

[3] A. Bauman, T. Armstrong, J. Davies, N. Owen, W. Brown, B. Bellew, and P. Vita. Trends in physical activity participation and the impact of integrated campaigns among Australian adults, 1997-99. *Australian and New Zealand Journal of Public Health*, 27(1):76–79, 2003.

[4] P. Biyani, C. Caragea, P. Mitra, and J. Yen. Identifying emotional and informational support in online health communities. In *Proceedings of COLING*, pages 827–836, 2014.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[6] D. L. Borzekowski, S. Schenk, J. L. Wilson, and R. Peebles. e-ana and e-mia: A content analysis of pro-eating disorder web sites. *American Journal of Public Health*, 8(100):1526–1534, 2010.

[7] M. A. Case, H. A. Burwick, K. G. Volpp, and M. S. Patel. Accuracy of Smartphone Applications and Wearable Devices for Tracking Physical Activity Data. *JAMA*, 313(6):625+, 2015.

[8] J. Cawley and C. Meyerhoefer. Internet interventions: in review, in use, and into the future. *Professional Psychology: Research and Practice*, 34(5):527–534, 2003.

[9] J. Cawley and C. Meyerhoefer. The medical care costs of obesity: An instrumental variables approach. *Australian and New Zealand Journal of Public Health*, 31(1):219–230, 2012.

[10] M. D. Choudhury. Anorexia on tumblr: A characterization study. In *Proceedings of DH*, pages 43–50, 2015.

[11] C. Chung and J. Pennebaker. The psychological functions of function words. *Social Communication*, pages 343–359, 2007.

[12] G. Coppersmith, M. Dredze, and C. Harman. Quantifying mental health signals in twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60, 2014.

[13] G. Coppersmith, M. Dredze, C. Harman, and K. Hollingshead. From ADHD to SAD: Analyzing the language of mental health on twitter through self-reported diagnoses. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 1–10, 2015.

[14] M. De Choudhury, S. Counts, and E. Horvitz. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of CHI*, pages 3267–3276, 2013.

[15] V. L. Dos Reis and A. Culotta. Using matched samples to estimate the effects of exercise on mental health via Twitter. In *Proceedings of AAAI*, pages 182–188, 2015.

[16] W. Haskell, I. Lee, R. Pate, K. Powell, S. Blair, B. Franklin, C. Macera, G. Heath, P. Thompson, and A. Bauman. Physical activity and public health: Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *JAMA*, 273(5):402–407, 1995.

[17] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of KDD*, pages 137–146, 2003.

[18] D. Kil, F. Shin, B. Piniewski, J. Hahn, and K. Chan. Impacts of social health data on predicting weight loss and engagement. In *O'Reilly Strata Rx Conference, San Francisco, CA*, 2012.

[19] M. Kumar, M. Dredze, G. Coppersmith, and M. De Choudhury. Detecting changes in suicide content manifested in social media following celebrity suicides. In *Proceedings of HT*, pages 85–94, 2015.

[20] A. Marshall, E. Eakin, E. Leslie, and N. Owen. Exploring the feasibility and acceptability of using internet technology to promote physical activity within a defined community. *Health promotion journal of Australia,*, 2005(16):82–84, 2005.

[21] Y. Mehmood, N. Barbieri, F. Bonchi, and A. Ukkonen. CSI: Community-level social influence analysis. In *Proceedings of ECML-PKDD'13*, pages 48–63, 2013.

[22] E. L. Murnane and S. Counts. Unraveling abstinence and relapse: Smoking cessation reflected in social media. In *Proceedings of CHI*, pages 1345–1354, 2014.

[23] M. Ottenhoff. Infographic: Rising use of social and mobile in healthcare. *The Spark Report*, 2012.

[24] N. Phan, D. Dou, B. Piniewski, and D. Kil. Social restricted boltzmann machine: Human behavior prediction in health social networks. In *Proceedings of ASONAM*, pages 424–431, 2015.

[25] N. Phan, D. Dou, X. Xiao, B. Piniewski, and D. Kil. Analysis of physical activity propagation in a health social network. In *Proceedings of CIKM*, pages 1329–1338, 2014.

[26] N. Phan, J. Ebrahimi, D. Kil, B. Piniewski, and D. Dou. Topic-aware physical activity propagation in a health social network. *IEEE Intelligent Systems*, 31(1):5–14, 2016.

[27] A. Sadilek, H. Kautz, and V. Silenzio. Predicting disease transmission from geo-tagged micro-blog data. In *Proceedings of AAAI*, pages 136–142, 2012.

[28] V. Stoyanov, C. Cardie, and J. Wiebe. Multi-perspective question answering using the OPQA corpus. In *Proceedings of EMNLP*, pages 923–930, 2005.

[29] C. Vandelanotte, K. M. Spathonis, E. G. Eakin, and N. Owen. Website-delivered physical activity interventions: A review of the literature. *American Journal of Preventive Medicine*, 33(1):54 – 64, 2007.

[30] C. Weeg, H. A. Schwartz, S. Hill, M. R. Merchant, C. Arango, and L. Ungar. Using Twitter to measure public discussion of diseases: A case study. *JMIR Public Health Surveillance*, 1(1), 2015.

[31] K. Zhao, G. E. Greer, B. Qiu, P. Mitra, K. Portier, and J. Yen. Finding influential users of an online health community: A new metric based on sentiment influence. *JAMIA*, 21(e2), 2012.