Understanding and Predicting Weight Loss with Mobile Social Networking Data

Zhiwei Wang
Data Science and Engineering Lab
Michigan State University
wangzh65@msu.edu

Tyler Derr
Data Science and Engineering Lab
Michigan State University
derrtyle@msu.edu

Dawei Yin
Data Science Lab
JD.com
Yindawei@acm.org

Jiliang Tang
Data Science and Engineering Lab
Michigan State University
tangjili@msu.edu

ABSTRACT

It has become increasingly popular to use mobile social networking applications for weight loss and management. Users can not only create profiles and maintain their records but also perform a variety of social activities that shatter the barrier to share or seek information. Due to the open and connected nature, these applications produce massive data that consists of rich weight-related information which offers immense opportunities for us to enable advanced research on weight loss. In this paper, we conduct the initial investigation to understand weight loss with a large-scale mobile social networking dataset with near 10 million users. In particular, we study individual and social factors related to weight loss and reveal a number of interesting findings that help us build a meaningful model to predict weight loss automatically. The experimental results demonstrate the effectiveness of the proposed model and the significance of social factors in weight loss.

KEYWORDS

Weight Loss, Social Network Analysis, Mobile Applications

ACM Reference format:
Zhiwei Wang, Tyler Derr, Dawei Yin, and Jiliang Tang. 2017. Understanding and Predicting Weight Loss with Mobile Social Networking Data. In *Proceedings of ACM Conference, Washington, DC, USA, July 2017 (Conference'17),* 10 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

It is informed by a recent report from the American Heart Association that over the last 30 years there has been a startling increase in overweight and obesity in the United Stated across all sections of age, sex, race/ethnicity, and socioeconomic status, which has led to an estimated total of 154.7 million adults (≥ 20 years of age) [17]. This is quite alarming because overweight and obesity are leading risk factors for cardiovascular disease and stroke [20, 32], which are the first and fifth [21, 30] leading causes of death in the United States. Given the importance of maintaining a healthy weight, increasing attention has been attracted from multiple disciplines on understanding and helping weight loss [15, 27, 39].

With the advance in mobile devices, a growing number of self-help mobile applications have been developed for weight management [1, 10]. Some enable their users to report weight-related data such as weight, sleep and physical activities, and receive feedback on input [28]; while others help track calorie balance [39] and promote physical activities [12]. Due to the ease of self-monitoring and flexible access, these mobile applications can enhance the delivery of interventions that could benefit weight loss [15, 28]. The majority of such applications focus on individual-level weight management. On the other hand, social networking systems such as Facebook and Twitter have greatly enabled people to participate in online activities and shattered the barrier for online users to create and share information at any place at any time. Increasing efforts have been made to incorporate social networking elements into weight management that encourages the emergence of weight-related mobile social networking applications [18]. With such applications, in addition to self-monitoring, users can perform a number of social activities that allow them to share and receive information from other users. This process produces rich weight-related data that offers unprecedented opportunities for large-scale weight loss studies.

In this paper, we perform the initial investigation on understanding weight loss with large-scale data from weight-related mobile social networking applications. Specifically, we aim to answer two questions - (1) what factors are related to weight loss? and (2) how to use these factors to help predict weight loss? We have mainly dedicated our efforts to understanding individual and social factors related to weight loss to answer the first question. Individual factors are from users' attributes such as age, gender and body mass index (BMI[1]); while social factors are extracted from various types of social activities such as following, mentioning and commenting. Our major findings about the two types of factors are summarized as follows:

Conference'17, Washington, DC, USA

DOI: 10.1145/nnnnnnn.nnnnnnn

[1] BMI is a value derived from the mass (weight) and height of an individual from which we can categorize the person as underweight, normal weight, overweight, or obese.

- For individual factors, (1) with the increase of age, the likelihood of successfully losing weight also increases; (2) females are more likely to lose weight than males; and (3) the probability of weight loss firstly increases and then decreases with the growth of BMI.
- For social factors, (1) users active in mobile applications are likely to lose weight; (2) peer pressure exists in weight loss; (3) the gender of your "friends" matters; and (4) social correlations exist in weight loss.

The significance of such understandings is two-fold. First, findings revealed from our study can in turn help build better weight management mobile applications. Second, social factors can be greatly complimentary to existing behavioral and psychological factors [41] since social factors can be relevant and influential in weight loss [23]. To answer the second question, we build a meaningful framework based on aforementioned understandings that can automatically predict weight loss. Empirical results on a large-scale mobile social networking dataset with near 10million users demonstrate (a) the effectiveness of the proposed framework and (b) the importance of social factors in weight loss.

The rest of this paper is organized as follows. In Section 2, we introduce the dataset and present our analysis results. We first provide a formal definition for the weight loss prediction problem, followed by our proposed prediction models in Section 3. Details of our experiments along with their results are given in Section 4. In Section 5, we discuss related work briefly. Finally, conclusions are given in Section 6 along with some future research directions.

2 DATA ANALYSIS

In this section, we conduct preliminary analysis with mobile social networking data aiming to answer the question – what are the important factors that can indicate weight loss. Such understandings can pave us ways to develop meaningful models for weight loss prediction. Before presenting the analysis results, we first introduce the dataset we used in this study.

2.1 Data

We collected a large-scale dataset with near 10 million users from one of the most popular mobile social apps for weight loss in China, BOOHEE[2]. People can register an account from BOOHEE and then they can create their profiles, record the weight loss process, and share information and interact with others. For each user, we gathered information of age, gender and body mass index (BMI) from her/his profiles. In BOOHEE, users can mainly perform three types of social activities – (1) they can follow users they are interested in; (2) they can share their weight loss progress with an option to mention others in the post; and (3) they can comment on others' posts. For these three activities, we can build three corresponding networks – following, mentioning and commenting networks. Note that although we also crawled post and comment content, we only use these three networks in this work and would like to exploit content information for weight loss as one future work. Some statistics of the dataset are shown in Table 1.

Table 1: Statistics of the dataset.

Description	BOOHEE
# of users	9,967,290
Avg of age	23.98
Ratio of female	0.62
Avg of BMI	23.74
# following links	127,425,354
# mentions	26,502,741
# comments	25,960,524

Next we introduce notations and definitions used in the work. Let $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ be the set of n users. We use a_i and f_i to denote u_i's age and gender, respectively. Meanwhile, b_i^f and b_i^l are the first and latest reported BMIs of u_i, separately. In this work, if $\frac{b_i^f - b_i^l}{b_i^f} > \mu$, we consider that u_i loses weight successfully where $\mu \in (0, 1)$ is the significance level. We adopt $\mathbf{F} \in \mathbb{R}^{n \times n}$ to denote the following network where $\mathbf{F}_{ij} = 1$ if there is a link from u_i to u_j and $\mathbf{F}_{ij} = 0$, otherwise. Similarly, we use $\mathbf{M} \in \mathbb{R}^{n \times n}$ and $\mathbf{C} \in \mathbb{R}^{n \times n}$ to represent the mention and commenting networks, respectively, where $\mathbf{M}_{ij}$ (or $\mathbf{C}_{ij}$) is the mentioning (or commenting) frequency from u_i to u_j. In the following subsections, we will investigate factors indicating weight loss with respect to user attributes and social activities.

2.2 Individual Factors

In this subsection, we study individual factors from user attributes that can indicate the weight loss. Specifically, we analyze the impacts of age, gender and BMI on weight loss.

I1: Does Age Matter? People of different ages may have distinct motivations to lose weight. For example, young people are more likely to lose weight to improve their physical appearance, while senior people are more likely to focus on losing weight for the purpose of health longevity. Meanwhile, people of different ages will have different self-discipline ability, which may affect a lot of their weight loss progress. To analyze the correlation between age and weight loss, we first divide users into different groups by their ages and then calculate the ratios of users who successfully lost weights for each group. The weight loss successful ratios for different age groups are shown in Figure 1. We make two observations. First, for each age group, there is a considerable portion of people losing weights successfully. Not matter in which age group, people all have motivations to lose weights. Second, with the increase of age, the successfulness of losing weight tends to increase and people with age between 35 to 45 have the highest successful ratio.

I2: Gender Difference. It is consistently reported by social sciences that (1) women are more likely to perceive themselves as overweight than men; and (2) compared to men, women are more likely to attempt to lose weight [24]. Hence, gender could have impact on weight loss. We compute the ratios of women and men who successfully lost weight and the results are demonstrated in Figure 1. We notice that women have a much higher successful ratio than men. The observation supports that women are more likely to lose weight than men.

I3: BMI Patterns. BMI is obtained from mass and height of users. More specifically, $BMI = \frac{\text{mass}_{kg}}{\text{height}^2_m} = \frac{\text{mass}_{lb}}{\text{height}^2_{in}} \times 703$. BMI has been

[2]www.boohee.com

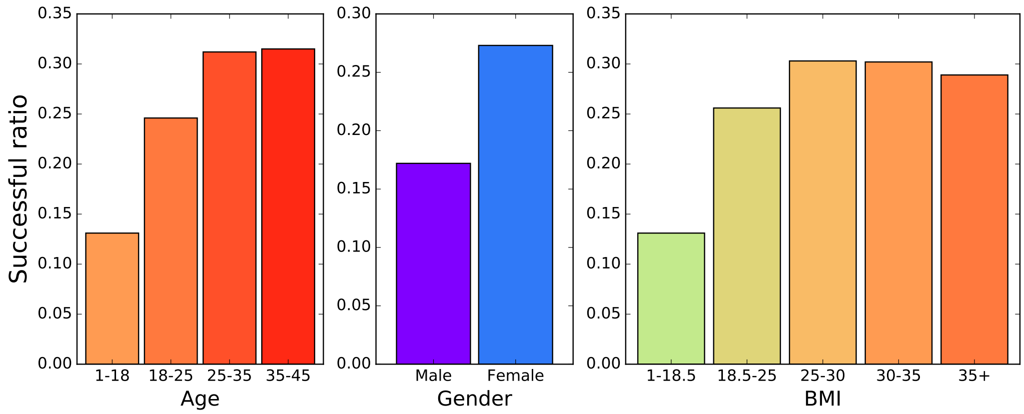

Figure 1: Gender, age, and BMI differences in weight loss.

widely adopted to assess how much one's weight departs from what is normal or desirable for her/his height. According to BMI, users can be roughly categorized into the following five groups – underweight (<18.5), normal (18.5-25), overweight (25-30), moderately obese (30-35) and severely obese (>35). For each group, we calculate the ratio of users who are considered to lose weight successfully and the results are demonstrated in Figure 1. We note that in general, with the increase of BMI, the successful ratio first increases and then decreases and among the five groups of users, overweight users are more likely to lose weight. From underweight to overweight, the successful ratio consistently increases. This observation is desirable. For underweight and normal users, they are likely to use the mobile social system to maintain their current weight or to even healthily gain weight; while overweight users have stronger motivations to lose weight. From moderately obese to severely obese, the ratio decreases that is consistently with findings from psychological studies – it becomes increasingly hard to lose weight for obese users with the growth of BMIs [37].

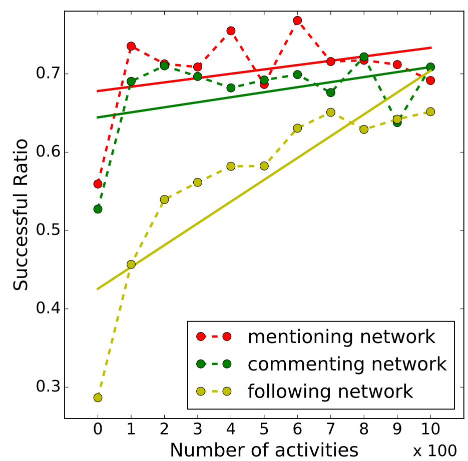

Figure 2: Number of activities vs. weight loss. Note that x-axis denotes the number of activities.

2.3 Social Factors

Online social activities have been used to help researchers understand a number of health and social issues such as depression [14], and autism [29]. In this subsection, we aim to understand weight loss from the social activity perspective. In particular, we want to investigate social factors from following, mentioning and commenting activities that can affect weight loss.

S1: Are More Active Users More Likely to Lose Weight? We show how the successful ratios change in terms of the number of followings, the number of mentions and the number of comments in Figure 2, respectively. The intuition is if you are more active in the mobile social loss weight system, you could have stronger desire to lose weight and as a consequence, it can be more likely for you to lose weight successfully. From the figure, in general, we consistently observe that the successful ratio tends to increase with the growth of the number of following, mentioning and commenting activities. Hence, it is evident from the figure that more active users are more likely to lose weight.

S2: Peer Pressure in Weight Loss We tend to worry about what people who are close to us think about our behaviors and appearance. Such peer pressure has been proven to exist in a variety of social activities such as drug use [11], education [35] and consuming behaviors [2, 16]. Intuitively, if a remarkable portion of "friends" (people we have connections to in the following, mentioning and commenting networks) of overweight people are not overweight, they should have pressure on their weight that can motivate them to lose weight. To verify the correctness of the intuition, we demonstrate how the likelihood of overweight people to lose weight changes with respect to the ratios of non-overweight "friends" in mentioning, or commenting networks respectively in Figure 3. It can be observed when the ratio of the non-overweight "friends" grows, people become increasingly likely to lose weight.

S3: Does the Gender of your "Friends" Matter? Since our weight is a major factor affecting the physical attraction, one intuition is that due to the cultural norm of heterosexual dating we tend to care more about what our friends of the opposite sex think about

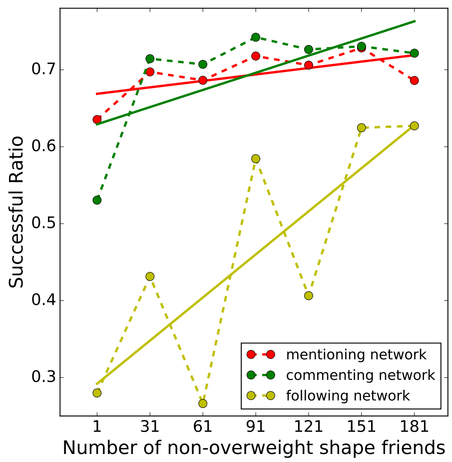

Figure 3: Peer pressure in weight loss. Each line corresponds to a social network.

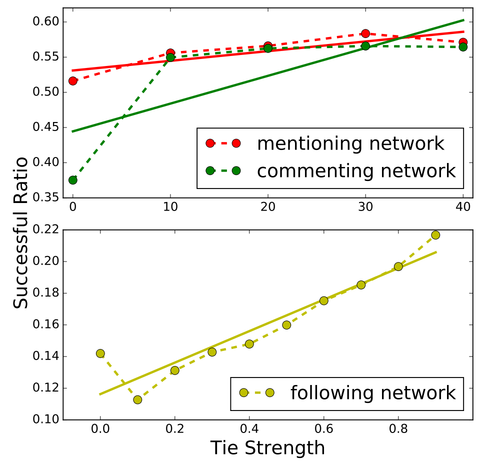

Figure 5: Impact of tie strength. Note that x-axis denotes tie strength and y-axis means the likelihood of the same weight loss status.

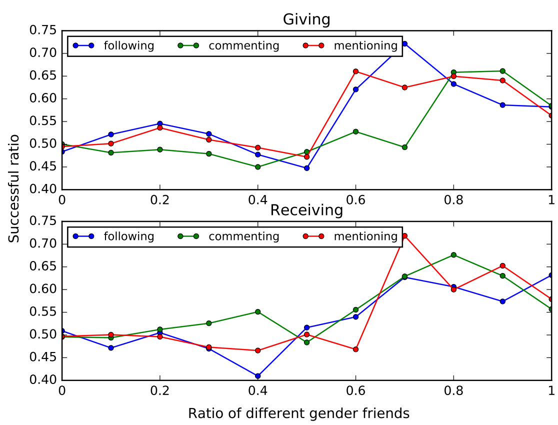

Figure 4: The Impact of your "friends" of the opposite gender.

our appearance. We first show how the weight loss successful ratio changes with the ratio of our "friends" with the opposite sex in Figure 4 with following, mentioning and commenting networks, respectively. It can be seen that the successful ratio tend to increase with the growth of the ratio of "friends" of the opposite gender. This aligns with one of the conclusions from a study investigating the effects of gender on eating behaviors in adolescence [3]. We also check how the successful ratio changes with the ratio of female (male) "friends" whom followings (mentions or comments) are given to (received from) and we have similar observations as those in Figure 4. One possible reason is that female are likely to have more male friends [36]. In other words, female are likely to give followings (mentions or comments) to users with the opposite gender (male).

S4: Social Correlations in Weight Loss Social correlations suggested by homophily and social influence have been widely observed in a variety of social networks [38]. We ask if social correlations exist in networks related to weight loss. To answer this question, we investigate if a linked pair of users in following (mentioning or commenting) networks have higher probability to have the same weight loss status (i.e., either both successfully lose weight or neither) than a randomly chosen pair. We follow a similar procedure as [38] to validate social correlations in weight loss that provides a positive answer to the previous question – linked users in following (mentioning or commenting) networks are likely to share the same weight loss status. The connection strength should affect the likelihood. For example, if user u_i mentions user u_j more frequently than u_k, intuitively u_i is likely to be more similar to u_j than u_k in terms of their weight loss statuses. In this study, we directly use the mentioning (or commenting) frequencies to indicate connection strengths for mentioning and commenting networks. User attribute similarities are good indicators of tie strengths [42]; hence in the following network, if $\mathbf{F}_{ij} = 1$, we use the attribute similarities $\mathbf{S}_{ij}$ to denote its tie strength that is calculated as the cosine similarity in this work. Note that $\mathbf{S} \in \mathbb{R}^{n \times n}$ is the matrix representation of the following network with tie strengths. We show the impact of connection strengths on the likelihood in Figure 5. It can be observed that with the growth of connection strengths, the likelihood increases.

3 WEIGHT LOSS PREDICTION

In reality, the frequencies to report weights could vary dramatically for different users, hence at one given time stamp, we may only know the weight loss of a small portion of users. Predicting weight loss for those unknown users can help us provide useful and timely interventions and recommendations. Therefore, we study the problem of weight loss prediction in this section. Our findings in the previous section pave us a way to build an interpretable model to predict weight loss automatically. Next, we first formally define the problem of weight loss prediction and then detail how to utilize our findings for weight loss prediction.

3.1 Problem Statement

Assume that $\mathcal{U}^l = \{u_1, u_2, \ldots, u_K\}$ $(K < n)$ is a subset of users in $\mathcal{U}$. For each user $u_i \in \mathcal{U}^l$, we know how much weight this user has lost, and that is denoted as y_i. In this paper, we define the problem of weight loss prediction as:

Given a set of n users $\mathcal{U}$, their attributes of age $\{a_i\}_{i=1}^n$, gender $\{f_i\}_{i=1}^n$, height $\{h_i\}_{i=1}^n$, the first recorded weight $\{w_i\}_{i=1}^n$, the first recorded BMIs $\{b_i\}_{i=1}^n$, the number of recorded weights $\{r_i\}_{i=1}^n$, number of days between the first and the last recorded weights $\{d_i\}_{i=1}^n$, the following $\mathbf{F}$, mentioning $\mathbf{M}$ and commenting $\mathbf{C}$ networks, and a subset of users $\{\mathcal{U}^l, \{y_i\}_{i=1}^K\}$ with known weight loss, we aim to learn a function f that can predict how much weight the remaining users in $\mathcal{U}$ will lose.

3.2 The Proposed Framework

Our basic idea is to first extract a set of features to represent each user u_i. Then based on the representations, we boil down the weight loss prediction problem into a regression problem where we can choose any regression models such as linear regression and Lasso for prediction. Therefore, the key problem of the proposed framework is how to extract features to denote users from social information in addition to user attributes such as age, gender, and the first reported BMI. Next, we introduce two ways for feature extraction based on our previous findings, i.e., feature engineering and user embedding with multiple networks.

3.2.1 Feature Engineering. We use feature engineering to extract features according to our findings from social information. These features are extracted from users' following, mentioning and commenting activities as follows:

- From findings in *S1*, we extract the following features: # of followings (mentions or comments) giving (or receiving);
- We extract features according to findings in *S2* as: ratios (numbers) of underweight (normal or overweight) users whom followings (mentions or comments) are given to (received from); and
- According to observations in *S3*, we extract features like ratios (numbers) of female (male or opposite gender) whom followings (mentions or comments) are given to (received from).

3.2.2 User Embedding with Multiple Networks. Through our analysis on following, commenting and mentioning networks, we find that users, who are linked in the networks, are likely to share similar weight loss statuses. Assume that $\mathbf{u}_i \in \mathbb{R}^{m_E}$ and $\mathbf{u}_j \in \mathbb{R}^{m_E}$ are latent representations of two linked users (u_i, u_j), where m_E is the number of latent dimensions. The finding suggests that their latent representations $\mathbf{u}_i$ and $\mathbf{u}_j$ should be similar. This observation allows us to develop an algorithm that can learn the user representations $\{\mathbf{u}_i\}_{i=1}^n$ automatically from the following, commenting and mentioning networks. Next we will first use the following network $\mathbf{F}$ as an example to illustrate how to learn representations of users from each network, and then we discuss how to learn unified representations of users by combing three networks.

If u_i and u_j are linked in the following network (or $\mathbf{F}_{ij} = 1$), we should force their latent representations to be close by minimizing their distance $\|\mathbf{u}_i - \mathbf{u}_j\|_2^2$. This process is equivalent to performing the spectral embedding on $\mathbf{F}$ [7]. Let $\mathbf{D} \in \mathbb{R}^{n\times n}$ be a diagonal matrix where $\mathbf{D}_{ii} = \sum_{j=1}^n \mathbf{F}_{ij}$. We define the normalized graph Laplacian matrix as $\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{F}\mathbf{D}^{-1/2}$; and use $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]^\top \in \mathbb{R}^{n\times m}$ to denote the representation matrix of all users. Then $\mathbf{U}$ can be learnt from the following network by solving the following optimization problem as:

$$\max_{\mathbf{U}} Tr(\mathbf{U}^T\mathbf{L}\mathbf{U}), \ s.t.\ \mathbf{U}^T\mathbf{U} = \mathbf{I} \tag{1}$$

where $Tr(\mathbf{A})$ is the trace of the matrix of $\mathbf{A}$. Next, we discuss how to learn a unified representation matrix $\mathbf{U}^* \in \mathbb{R}^{n\times m_E}$ from three networks $\mathbf{M}, \mathbf{C}, \mathbf{F}$.

Let $\mathbf{S}^{(v)}$ denote the network v ($v \in \{\mathbf{M}, \mathbf{C}, \mathbf{F}\}$) and its normalized graph Laplacian matrix is $\mathbf{L}^{(v)} = \mathbf{D}^{(v)^{-1/2}}\mathbf{S}^{(v)}\mathbf{D}^{(v)^{-1/2}}$. We use $\mathbf{U}^{(v)} \in \mathbb{R}^{n\times m_E}$ to denote the user latent representation matrix from the network v. Since the three user representations from three networks denote the same set of users, intuitively, they should be close to the unified representation matrix $\mathbf{U}^*$. We use the disagreement measurement proposed in [22] to calculate the distance between $\mathbf{U}^{(v)}$ and $\mathbf{U}^*$ as:

$$D(\mathbf{U}^{(v)}, \mathbf{U}^*) = -Tr(\mathbf{U}^{(v)}\mathbf{U}^{(v)^T}\mathbf{U}^{(*)}\mathbf{U}^{(*)^T}) \tag{2}$$

Combining Eq. (1) and Eq. (2), we can reach the optimization formulation to obtain $\mathbf{U}^{(v)}$ ($v \in \{\mathbf{M}, \mathbf{C}, \mathbf{F}\}$) and the unified representation matrix of users $\mathbf{U}^*$ as follows:

$$\begin{aligned} \max_{\mathbf{U}^{(v)}, \mathbf{U}^*} & \sum_{v\in\{M,C,F\}} Tr(\mathbf{U}^{(v)^T}\mathbf{L}^{(v)}\mathbf{U}^{(v)}) \\ & + \sum_{v\in\{M,C,F\}} \lambda_v Tr(\mathbf{U}^{(v)}\mathbf{U}^{(v)^T}\mathbf{U}^{(*)}\mathbf{U}^{(*)^T}) \\ & s.t.\ \mathbf{U}^{(v)^T}\mathbf{U}^{(v)} = \mathbf{I}\ \forall v \in \{M,C,F\}, \mathbf{U}^{(*)^T}\mathbf{U}^{(*)} = \mathbf{I} \end{aligned} \tag{3}$$

where λ_v is introduced to control the contribution from the network v in the unified representation of $\mathbf{U}^*$.

The unified representation of users $\mathbf{U}^{(*)}$ is general, which may be not optimal for the problem of weight loss prediction. In the studied problem, we assume that the weight loss of some users are given, which can be used to guide the learning process of $\mathbf{U}^{(*)}$ for weight loss prediction. We propose the following optimization formulation to include the information of labeled users as:

$$\begin{aligned} \max_{\mathbf{w}, \mathbf{U}^{(v)}, \mathbf{U}^*} & \sum_{v\in\{M,C,F\}} tr(\mathbf{U}^{(v)^T}\mathbf{L}^{(v)}\mathbf{U}^{(v)}) \\ & + \sum_{v\in\{M,C,F\}} \lambda_v tr(\mathbf{U}^{(v)}\mathbf{U}^{(v)^T}\mathbf{U}^{(*)}\mathbf{U}^{(*)^T}) \\ & - \alpha||\mathbf{S} \odot (\mathbf{U}^{(*)}\mathbf{w} - \mathbf{y})||_F^2 - \beta||\mathbf{w}||_2^2 \\ s.t.\quad & \mathbf{U}^{(v)^T}\mathbf{U}^{(v)} = \mathbf{I}\ \forall v \in \{M,C,F\}, \mathbf{U}^{(*)^T}\mathbf{U}^{(*)} = \mathbf{I} \end{aligned} \tag{4}$$

where $\mathbf{S} = [s_1, s_2, \cdots, s_n]^T \in \mathbb{R}^{n\times 1}$ is the indicator vector that $s_i = 1$ if $u_i \in \mathbf{U}_l$, otherwise $s_i = 0$. $\mathbf{y} = [y_1, y_2, \cdots, y_n]^T$and y_i is the weight loss for user u_i. The term $||\mathbf{S} \odot (\mathbf{U}^{(*)}\mathbf{w} - \mathbf{y})||_F^2$ forces that the unified representation $\mathbf{U}^{(*)}$ can fit the label information

$\mathbf{y}$ via a mapping vector $\mathbf{w} \in \mathbb{R}^{m_E}$. α controls the contribution of the label information in the learning process. The term of $||\mathbf{w}||_2^2$ is used to avoid overfitting.

We use Alternating Direction Method of Multiplier (ADMM) [9] to optimize Eq. (4). Adding the auxiliary variable $\mathbf{U}^{(*)} = \mathbf{Z}$, the original objective function can be rewritten as:

$$
\begin{aligned}
\min_{\mathbf{w},\mathbf{U}^{(v)},\mathbf{U}^{(*)}} \; & - \sum_{v \in \{M,C,F\}} tr(\mathbf{U}^{(v)^T} \mathbf{L}^{(v)} \mathbf{U}^{(v)}) \\
& - \sum_{v \in \{M,C,F\}} \lambda_v tr(\mathbf{U}^{(v)} \mathbf{U}^{(v)^T} \mathbf{Z}\mathbf{Z}^T) \\
& + \alpha ||\mathbf{S} \odot (\mathbf{Z}\mathbf{w} - \mathbf{y})||_F^2 + \beta ||\mathbf{w}||_2^2 \\
s.t. \quad & \mathbf{U}^{(v)^T} \mathbf{U}^{(v)} = \mathbf{I} \; \forall v \in \{M,C,F\}, \mathbf{U}^{(*)^T} \mathbf{U}^{(*)} = \mathbf{I}, \mathbf{U}^{(*)} = \mathbf{Z}
\end{aligned}
\tag{5}
$$

By adding the Lagrange multiplier $\mathbf{A}$ and the Lagrangian parameter μ, the Eq.(5) can be solved by using the following ADMM format:

$$
\begin{aligned}
\min_{\mathbf{w},\mathbf{A},\mathbf{Z},\mathbf{U}^{(v)},\mathbf{U}^{(*)}} \; & - \sum_{v \in \{M,C,F\}} tr(\mathbf{U}^{(v)^T} \mathbf{L}^{(v)} \mathbf{U}^{(v)}) \\
& - \sum_{v \in \{M,C,F\}} \lambda_v tr(\mathbf{U}^{(v)} \mathbf{U}^{(v)^T} \mathbf{Z}\mathbf{Z}^T) \\
& + \alpha ||\mathbf{S} \odot (\mathbf{Z}\mathbf{w} - \mathbf{y})||_F^2 + \beta ||\mathbf{w}||_2^2 \\
& + \frac{\mu}{2} ||\mathbf{Z} - \mathbf{U}^{(*)}||_F^2 + tr(\mathbf{A}^T (\mathbf{Z} - \mathbf{U}^{(*)})) \\
s.t. \quad & \mathbf{U}^{(v)^T} \mathbf{U}^{(v)} = \mathbf{I} \; \forall v \in \{M,C,F\}, \mathbf{U}^{(*)^T} \mathbf{U}^{(*)} = \mathbf{I}
\end{aligned}
\tag{6}
$$

Update $\mathbf{U}^{(*)}$: To update $\mathbf{U}^{(*)}$, we fix other terms that are irrelevant and drop the constant ones. Eq.(6) becomes:

$$
\begin{aligned}
\min_{\mathbf{U}^*} \quad & \frac{\mu}{2} ||\mathbf{Z} - \mathbf{U}^{(*)}||_F^2 + tr(\mathbf{A}^T (\mathbf{Z} - \mathbf{U}^{(*)})) \\
s.t. \quad & \mathbf{U}^{(*)^T} \mathbf{U}^{(*)} = \mathbf{I}
\end{aligned}
\tag{7}
$$

By expanding the expression and dropping the variables that are independent of $\mathbf{U}^*$, we further convert Eq.(6) to the following:

$$
\min_{\mathbf{U}^*} \quad \frac{\mu}{2} ||\mathbf{U}^{(*)} - \mathbf{H}||_F^2, \; s.t. \quad \mathbf{U}^{(*)^T} \mathbf{U}^{(*)} = \mathbf{I} \tag{8}
$$

where $\mathbf{H} = \mathbf{Z} + \frac{1}{\mu}\mathbf{A}$. Then the optimization can be solved by the following lemma [40][19]:

lemma 1: *Given the objective in Eq.(8), the optimal* $\mathbf{U}^{(*)}$ *is defined as:*

$$
\mathbf{U}^* = \mathbf{P}\mathbf{Q}^T \tag{9}
$$

where $\mathbf{P}, \mathbf{Q}$ *are the left and right singular vectors of the economic singular value decomposition of* $\mathbf{H}$

Update Z: Similar to previous procedure, after dropping the terms that are irrelevant to $\mathbf{Z}$, we convert Eq.(6) to the following problem:

$$
\begin{aligned}
\min_{\mathbf{Z}} & - \sum_{v \in \{M,C,F\}} \lambda_v tr(\mathbf{U}^{(v)} \mathbf{U}^{(v)^T} \mathbf{Z}\mathbf{Z}^T) + \alpha ||\mathbf{S} \odot (\mathbf{Z}\mathbf{w} - \mathbf{y})||_F^2 \\
& + \frac{\mu}{2} ||\mathbf{Z} - \mathbf{U}^{(*)}||_F^2 + tr(\mathbf{A}^T (\mathbf{Z} - \mathbf{U}^{(*)}))
\end{aligned}
\tag{10}
$$

Let $\mathbf{M} = \sum_{v \in \{M,C,F\}} \lambda_v \mathbf{U}^{(v)} \mathbf{U}^{(v)^T}$, we rewrite the Eq.(10) as:

$$
\begin{aligned}
\min_{\mathbf{Z}} \quad & \alpha ||\mathbf{S} \odot (\mathbf{Z}\mathbf{w} - \mathbf{y})||_F^2 - tr(\mathbf{Z}^T \mathbf{M}\mathbf{Z}) \\
& + \frac{\mu}{2} ||\mathbf{Z} - \mathbf{U}^{(*)}||_F^2 + tr(\mathbf{A}^T (\mathbf{Z} - \mathbf{U}^{(*)}))
\end{aligned}
\tag{11}
$$

Again, by expanding Eq.(11) and dropping the terms that are independent of $\mathbf{Z}$, we have

$$
\begin{aligned}
\min_{\mathbf{Z}} \quad & tr(\alpha \mathbf{S}^T \odot (\mathbf{w}^T \mathbf{Z}^T \mathbf{Z}\mathbf{w} - 2\mathbf{y}^T \mathbf{Z}\mathbf{w}) - \mathbf{Z}^T \mathbf{M}\mathbf{Z} \\
& + \frac{\mu}{2} (\mathbf{Z}^T \mathbf{Z} + 2\mathbf{Z}^T \mathbf{U}^{(*)}) + \mathbf{A}^T \mathbf{Z})
\end{aligned}
\tag{12}
$$

The Eq.(12) can be solved by using gradient decent method as follows:

$$
\begin{aligned}
& \nabla_{\mathbf{Z}} = \alpha((\mathbf{S} \odot \mathbf{Z}\mathbf{w})\mathbf{w}^T - \mathbf{S} \odot \mathbf{y}\mathbf{w}^T) - \mathbf{M}\mathbf{Z} + \frac{\mu}{2}(\mathbf{Z} + \mathbf{U}^{(*)}) + \frac{1}{2}\mathbf{A} \\
& \mathbf{Z} = \mathbf{Z} - \gamma \nabla_{\mathbf{Z}}
\end{aligned}
\tag{13}
$$

Algorithm 1: User Embedding with Multiple Networks

Input: $\mathbf{L}^{(v)} \in \mathbf{R}^{n \times n} (v \in \{\mathbf{M}, \mathbf{C}, \mathbf{F}\})$, $\mathbf{y} \in \mathbf{R}^{n \times 1}$, $\mathbf{S} \in \mathbf{R}^{n \times 1}$, k, λ_v, γ
Output: $\mathbf{U}^{(*)}$
Initialize: $\mu = 1$
Initialize $\mathbf{U}^{(v)}$,$\mathbf{U}^{(*)}$, $\mathbf{Z}$ and $\mathbf{A}$ and $\mathbf{w}$ randomly
repeat
  Calculate $\mathbf{H} = \mathbf{Z} + \frac{1}{\mu}\mathbf{A}$
  Update $\mathbf{U}^{(*)}$ by applying Lemma 8
  Calculate $\mathbf{M} = \sum_{v \in \{M,C,F\}} \lambda_v \mathbf{U}^{(v)} \mathbf{U}^{(v)^T}$
  Calculate $\nabla_{\mathbf{Z}} = \alpha((\mathbf{S} \odot \mathbf{Z}\mathbf{w})\mathbf{w}^T - \mathbf{S} \odot \mathbf{Y}\mathbf{w}^T) - \mathbf{M}\mathbf{Z} + \frac{\mu}{2}(\mathbf{Z} + \mathbf{U}^{(*)}) + \frac{1}{2}\mathbf{A}$
  Update $\mathbf{Z}$ by using gradient decent as Eq.(13)
  Calculate the largest k eigenvectors of $(\mathbf{L}^{(v)} + \lambda_v \mathbf{Z}\mathbf{Z}^T)$
  Update $\mathbf{U}^{(v)}$ by combining the k eigenvectors
  Update $\mathbf{w}$ by gradient decent as Eq. (17)
  Update $\mathbf{A}$ using the Eq. (18)
until *convergence*

Update $\mathbf{U}^{(v)}$: We fix the other terms except $\mathbf{U}^{(v)}$ and drop the irrelevant terms, then Eq.(6) becomes:

$$
\begin{aligned}
\max_{\mathbf{U}^{(v)}} \quad & tr(\mathbf{U}^{(v)^T} \mathbf{L}^{(v)} \mathbf{U}^{(v)}) + \lambda_v tr(\mathbf{U}^{(v)} \mathbf{U}^{(v)^T} \mathbf{Z}\mathbf{Z}^T) \\
s.t. \quad & \mathbf{U}^{(v)^T} \mathbf{U}^{(v)} = \mathbf{I}
\end{aligned}
\tag{14}
$$

By utilizing the trace properties, Eq.(14) can be rewrote as the following form:

$$
\max_{\mathbf{U}^{(v)}} \quad tr(\mathbf{U}^{(v)^T} (\mathbf{L}^{(v)} + \lambda_v \mathbf{Z}\mathbf{Z}^T) \mathbf{U}^{(v)^T}), \; s.t. \; \mathbf{U}^{(v)^T} \mathbf{U}^{(v)} = \mathbf{I} \tag{15}
$$

Eq.(15) is the standard spectral cluster objective function regarding the $(\mathbf{L}^{(v)} + \lambda_v \mathbf{Z}\mathbf{Z}^T)$ as the Laplacian matrix. The solution can be given by the k largest eigenvectors of $(\mathbf{L}^{(v)} + \lambda_v \mathbf{Z}\mathbf{Z}^T)$ [7].

Update w: By fixing other terms and dropping those who are independent of $\mathbf{w}$, we get

$$\min_{\mathbf{w}} \alpha||\mathbf{S} \odot (\mathbf{Z}\mathbf{w} - \mathbf{Y})||_F^2 + \beta||\mathbf{w}||_2^2 \tag{16}$$

We use the gradient decent method to solve Eq.(16)

$$\begin{aligned} \nabla_{\mathbf{w}} &= \mathbf{Z}^T(\mathbf{S} \odot \mathbf{Z}\mathbf{w}) - \mathbf{Z}^T(\mathbf{S} \odot \mathbf{Y}) + \frac{\beta}{\alpha}\mathbf{w} \\ \mathbf{w} &= \mathbf{w} - \gamma\nabla_{\mathbf{w}} \end{aligned} \tag{17}$$

Update A: Finally, we update the Lagrange multiplier $\mathbf{A}$ as follows:

$$\mathbf{A} = \mathbf{A} + \mu(\mathbf{Z} - \mathbf{U}^*) \tag{18}$$

The detailed algorithm to optimize Eq.(5) is shown in Algorithm 1. Here we briefly introduce the algorithm. In line 2, we randomly initialize each matrix. From line 5 to line 12, we use ADMM to alternatively update $\mathbf{U}^{(v)}$,$\mathbf{U}^{(*)}$, $\mathbf{Z}$, $\mathbf{w}$ and $\mathbf{A}$ until convergence. Specifically, in line 5, we construct the $\mathbf{H}$ matrix, which is used to update $\mathbf{U}^{(*)}$according to Lemma 8. From line 7 to line 8, we update $\mathbf{Z}$ using the gradient decent method. The matrix $\mathbf{U}^{(v)}$ is updated in lines 9 and 10 by computing the eigenvectors of the corresponding revised Laplacian matrix. At last we update $\mathbf{w}$ and $\mathbf{A}$ according to Eq. (17) and Eq. (18), repectively.

Time complexity analysis: The computation cost for updating $\mathbf{U}^{(*)}$ involves computing $\mathbf{H} = \mathbf{Z} + \frac{1}{\mu}\mathbf{A}$ and **SVD** on $\mathbf{H}$. They are at cost of $O(nm_E)$ and $O(n(m_E)^2)$, respectively. For updating $\mathbf{Z}$, calculating $\nabla_{\mathbf{Z}}$ will cost $O(n^2 m_E)$ in each gradient decent iteration. Assuming c_z iterations are needed for the gradient decent method, the total time complexity will be $O(c_z n^2 m_E)$. Similarly, to update $\mathbf{w}$, $\nabla_{\mathbf{w}}$ will be calculated at cost of $O(nm_E)$. Assuming c_w iterations are needed, then time complexity for updating $\mathbf{w}$ will be $O(c_w nm_E)$. For updating $\mathbf{U}^{(v)}$, the first cost is the computing $\mathbf{L}^{(v)} + \lambda_v \mathbf{Z}\mathbf{Z}^T$, whose complexity is $O(n^2 m_E)$. Then to get the largest m_E eigenvectors, $O(n^2 m_E)$ is needed. Thus, the overall time complexity is $O(n(m_E)^2 + n^2 m_E)$ for each iteration in the algorithm.

3.3 Predicting Weight Loss

In addition to user attributes, we use (1) feature engineering to manually extract a set of m_F features according to our findings to represent users from social information, denoted as $\mathbf{U}_i^F \in \mathbb{R}^{m_F}$ for u_i; and (2) user embedding to learn the representations of users automatically from social information, indicated as $\mathbf{U}^E \in \mathbb{R}^{m_E}$ for u_i where m_E is the number of latent dimensions. For each type of features, we can apply a regression model for prediction. However, these two types of features may contain complementary information to each other. Hence, a better strategy is to combine them. Next we will use combining $\mathbf{U}_i^F$ and $\mathbf{U}^E$ as an example to illustrate our strategy.

We assume that the weight loss of u_i can be computed as: $\mathbf{y}_i = (1-\theta)\mathbf{U}_i^F\mathbf{w}^\mathbf{F} + \theta\mathbf{U}_\mathbf{i}^\mathbf{E}\mathbf{w}^\mathbf{E}$, where $\mathbf{w}^\mathbf{F}$ and $\mathbf{w}^\mathbf{E}$ are feature coefficients for $\mathbf{U}_i^F$ and $\mathbf{U}_i^E$, respectively. θ is the parameter controlling the relative importance of the two categories of features. Larger value of θ indicates more importance on $\mathbf{U}^{(E)}$. Then $\mathbf{w}^\mathbf{F}$ and $\mathbf{w}^\mathbf{E}$ can be learned by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w}^\mathbf{F},\mathbf{w}^\mathbf{E}} \sum_{u_i \in \mathcal{U}^l} &||(1-\theta)\mathbf{U}_i^F\mathbf{w}^\mathbf{F} + \theta\mathbf{U}_\mathbf{i}^\mathbf{E}\mathbf{w}^\mathbf{E} - \mathbf{y_i}||_2^2 \\ &+ \phi^F||\mathbf{w}^\mathbf{F}||_1 + \phi^\mathbf{E}||\mathbf{w}^\mathbf{E}||_1 \end{aligned} \tag{19}$$

we apply ℓ_1-norm on $\mathbf{w}^\mathbf{F}$ and $\mathbf{w}^\mathbf{E}$ to embed feature selection into the regression model, which can help eliminate irrelevant features into the learning process. Eq.(19) can be solved by updating $\mathbf{w}^\mathbf{E}$ and $\mathbf{w}^\mathbf{F}$ alternatively. Each subproblem is a standard lasso regression problem as follows:

$$\begin{aligned} &\min_{\mathbf{w}^\mathbf{E}} ||\theta\mathbf{U}^E\mathbf{w}^\mathbf{E} - (\mathbf{y} - (\mathbf{1}-\theta)\mathbf{U}^\mathbf{F}\mathbf{w}^\mathbf{F})||_2^2 + \phi^\mathbf{E}||\mathbf{w}^\mathbf{E}||_1 \\ &\min_{\mathbf{w}^\mathbf{F}} ||(1-\theta)\mathbf{U}^F\mathbf{w}^\mathbf{F} - (\mathbf{y} - \theta\mathbf{U}^\mathbf{E}\mathbf{w}^\mathbf{E})||_2^2 + \phi^\mathbf{F}||\mathbf{w}^\mathbf{F}||_1 \end{aligned} \tag{20}$$

4 EXPERIMENTS

In this section, we conduct experiments to verify (a) whether social information can help us predict weight loss better and (b) what are the important factors in weight loss prediction. We first introduce experimental settings, then assess the performance of weight loss prediction and finally study parameter sensitivity and the importance of factors.

4.1 Experimental Settings

We construct a dataset for this evaluation from that used in the data analysis section by filtering these users with less than 10 links in the following networks and less than 30 interactions in the commenting and mentioning networks. We divide the dataset into two parts – $\mathcal{A}$ containing 30% of the dataset and $\mathcal{B}$ containing the remaining 70%. In reality, at a given time stamp, the percent of users who reported their weights is very small. Therefore, we randomly choose x% of $\mathcal{A}$ as labeled for training and the remaining $1-x$% of $\mathcal{A}$ is used as unlabeled data for feature extraction. While the part $\mathcal{B}$ is fixed for testing. We vary x as $\{5, 15, 25, 35, 45\}$ in the evaluation. For each x, we repeat the experiments 10 times and report the average performance. We choose root-mean-square error (RMSE) as the measurement for the model performance. Lower RMSE indicates better prediction performance.

4.2 Performance of Weight Loss Prediction

To answer the question (a), we systematically investigate how different features affect the weight loss prediction performance. For one type of features, we choose the lasso regression for prediction; while we choose the strategy we discussed in the last section when combining multiple types of features. The comparison settings are below:

- **wLoss+A:** we only use the user attributes as features for prediction;
- **wLoss+M, wLoss+C, wLoss+S**: The features consist of the embedding obtained from a single network information. In particular, wLoss+M, wLoss+C, and wLoss+S denote embedding features from mentioning, commenting and following networks, separately;
- **wLoss+F, wLoss+E**: The features are extracted from multiple networks. wLoss+F and wLoss+E denote features from

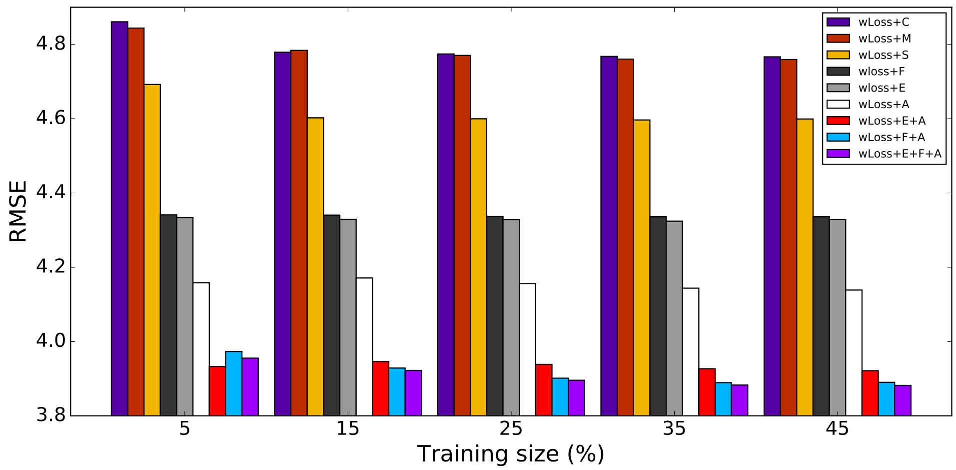

Figure 6: Performance comparison for weight loss prediction.

multiple networks via feature engineering and user embedding, respectively;

- **wLoss+F+A**: It combines features from user attributes and social information via feature engineering;
- **wLoss+E+A**: It combines features from user attributes and social information via user embedding; and
- **wLoss+E+F+A**: It combines features from user attributes, and social information via feature engineering and user embedding.

Note that the parameters of all methods are selected as these achieving the best performance on the validation set. More details about the parameter selection of the proposed model are discussed in the following subsections. Note that, we also can choose other regression models instead of lasso regression and we will leave that investigation as one future work. The results are shown in Figure 6.

It can be observed:

- **wLoss+E** obtains much better performance than **wLoss+M, wLoss+C, wLoss+S**. These results suggest that embedding from multiple networks is better than that from each individual in weight loss prediction;
- **wLoss+E+A** and **wLoss+F+A** achieves better performance than **wLoss+A**. These results suggest that social information contains complementary information to user attributes for weight loss prediction.
- Most of the time, **wLoss+E+F+A** gains performance improvement than both **wLoss+E+A** and **wLoss+F+A**. This observation supports that feature engineering and user embedding provide complementary information to each other. Further experiments will be conducted to investigate the impact of feature engineering and user embedding on weight loss prediction in the following subsection.

We perform t-test on all comparisons that suggests the improvement is significant. Based on above observations, we can draw an answer to question (a) that social factors play an important role in weight loss prediction.

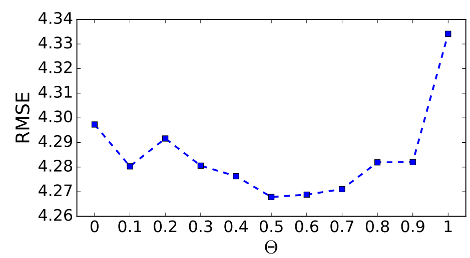

Figure 7: Performance variance with respect to θ.

4.3 Feature Engineering vs User Embedding

In this work, we proposed two ways to extract features from social information. One way is to manually extract features according to the findings from our preliminary data analysis via feature engineering. The other way is to learn the user representations automatically via user embedding. In the last subsections, when we combine them together, we can achieve better performance than each individual, which suggests that they contain complementary information. In this subsection, we further study the contributions of these two ways to the weight loss prediction performance.

Recall the strategy we combine two sets of features as $\mathbf{y}_i = (1-\theta)\mathbf{U}_i^F\mathbf{w}^\mathbf{F} + \theta\mathbf{U}_\mathbf{i}^\mathbf{E}\mathbf{w}^\mathbf{E}$, where θ balances the contributions from feature engineering and user embedding. We give more weights to user embedding when we choose larger values of θ. Hence, we study the impact of features from feature engineering and use embedding on weight loss prediction via showing how the weight loss prediction performance changes with θ. We vary θ from 0 to 1 with one step of 0.1. The results with 5% of training data is demonstrated in Figure 7. Note that we do not show the results with other settings since we have similar observations.

In general, with the increase of θ from 0, the performance first increases that further supports the importance of features from user embedding; at a certain region, the performance is quite stable that eases the difficulty to choose θ in practice; and when we further increase θ, the performance reduces. These results suggest that (1) both features from feature engineering and user embedding are useful for weight loss prediction and (2) they contain complementary information and an appropriate combination of them can improve the weight loss prediction performance.

4.4 Feature Analysis

We explicitly extract two groups of features via feature engineering. One group is from user attributes and the other is from social information. We concatenate these two groups of features into one set and perform lasso regression to learn the coefficient weights of features, which can indicate the importance of features. Therefore, we can study the importance of features via these coefficient weights. We rank features by the absolute values of the t-statistic for model parameters and the top 8 ranked features are shown below:

(1) BMI;
(2) Number of mentions received from other users;
(3) Initial weight;
(4) Number of comments giving to normal weight users;
(5) Number of followings from the opposite gender users;
(6) Ratio of followings from the overweight users;
(7) Number of comments received from other users;
(8) Number of followings giving to overweight users;

Among the top 8 features, 6 of them are from social information mainly related to peer pressure, and gender of “friends”. These results not only suggest the importance of social information but also indicate the necessity to study social factors with mobile social networking data.

5 RELATED WORK

In this section, we first give a brief overview on existing weight loss and management research. Since our model is related to collective classification on networks, we will also briefly review existing collective classification methods.

5.1 Weight Loss

An increasing number of weight-related mobile applications have been built [1, 10] for self-monitoring [28], tracking calorie balance [39] and promoting physical activities [12]. Studies have shown that these mobile applications can enhance the delivery of interventions that could benefit weight loss and controlling [15, 28]. Meanwhile there are online forums designed to allow users to discuss topics about weight loss. Some studied weight loss from the perspective of medical and psychological domains [5, 13]; some performed content analysis to understand the behavior of users who want to lose weight [27]; while others mine health application data to find more and less successful weight loss subgroups [34]. We conducted the initial study to understand individual and social factors in weight loss with a large-scale mobile social networking data.

5.2 Collective Classification on Networks

Many real-world problems can be modeled as the node classification problem in networks, which aims to predict labels of unlabeled nodes in a network by giving the network with some labeled users [33]. According to [8], these algorithms can be mainly divided into local classifier methods and random walk based methods. Local classifier based methods use local neighborhood information to learn local classifiers. Iterative classification methods (or ICA) [31] and its variants [25] construct feature vectors for nodes from the information known about them and their neighborhood. These feature vectors are then used along with labeled users to build a local classifier. Weighted-vote relational neighbor (wvRN) takes a weighted average of the class probabilities in the neighborhood for classification [26]. Random walk based methods propagate the labels by performing random walks on the network such as label propagation [4, 44], graph regularization [43] and adsorption [6]. Our proposed model is closely related to these in [26, 43, 44].

6 CONCLUSION

In this paper, we study the individual and social factors related to weight loss. First, we performed an initial investigation to understand weight loss with large-scale data from the weight related mobile social networking application BOOHEE, which is one of the most popular applications of its kind in China. Our aim was twofold: firstly to discover what factors are related to weight loss, and then to harness these factors into a weight loss prediction model. In answering what factors are related to weight loss we focused on extracting two types of factors, individual and social, and these factors were later used to construct a meaningful model to predict weight loss automatically. Empirical results on this large-scale data with near 10 million users have demonstrated not only the effectiveness of our framework, but also just as important, the significance social factors are in weight loss.

There are several directions that could be taken next for further investigation with large-scale mobile social networking application data. First, we plan to focus our attention more on the social factors, as the individual factors, such as behavioral and psychological, have been previously studied [41], and our results have shown that social factors do indeed play a critical role in determining one’s success in weight loss. Another avenue that could be quite interesting to explore is the inclusion of time into our prediction model. Third, content is also available in mobile social networking data and we would like to incorporate content analysis into our weight loss understanding and prediction. Finally, it would be interesting to use our findings to provide timely interventions and recommendations in weight management and control.

ACKNOWLEGEMENTS

This material is based upon work supported by, or in part by, the National Science Foundation (NSF) under grant number XXXX and XXXXX also would like to thank XXX and XXXX for their helpful discussions.

REFERENCES

[1] Sasan Adibi. 2015. *Mobile health: a technology road map.* Vol. 5. Springer.
[2] Nancy D Albers-Miller. 1999. Consumer misbehavior: why people buy illicit goods. *Journal of consumer Marketing* 16, 3 (1999), 273–287.
[3] Rheanna N Ata, Alison Bryant Ludden, and Megan M Lally. 2007. The effects of gender and family, friend, and media influences on eating behaviors and body image during adolescence. *Journal of Youth and Adolescence* 36, 8 (2007), 1024–1037.

[4] Arik Azran. 2007. The rendezvous algorithm: Multiclass semi-supervised learning with markov random walks. In *Proceedings of the 24th international conference on Machine learning*. ACM, 49–56.
[5] Paul W Ballantine and Rachel J Stephenson. 2011. Help me, I'm fat! Social support in online weight loss networks. *Journal of Consumer Behaviour* 10, 6 (2011), 332–337.
[6] Shumeet Baluja, Rohan Seth, D Sivakumar, Yushi Jing, Jay Yagnik, Shankar Kumar, Deepak Ravichandran, and Mohamed Aly. 2008. Video suggestion and discovery for youtube: taking random walks through the view graph. In *Proceedings of the 17th international conference on World Wide Web*. ACM, 895–904.
[7] Mikhail Belkin and Partha Niyogi. 2001. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *NIPS*, Vol. 14. 585–591.
[8] Smriti Bhagat, Graham Cormode, and S Muthukrishnan. 2011. Node classification in social networks. In *Social network data analytics*. Springer, 115–148.
[9] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning* 3, 1 (2011), 1–122.
[10] Michelle Clare Carter, Victoria Jane Burley, Camilla Nykjaer, and Janet Elizabeth Cade. 2013. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *Journal of medical Internet research* 15, 4 (2013), e32.
[11] Niall Coggans and Susan McKellar. 1994. Drug use amongst peers: peer pressure or peer preference? *Drugs: education, prevention and policy* 1, 1 (1994), 15–26.
[12] Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, and others. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1797–1806.
[13] Anita Das and Arild Faxvaag. 2014. What influences patient participation in an online forum for weight loss surgery? A qualitative case study. *Interactive journal of medical research* 3, 1 (2014), e4.
[14] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via Social Media.. In *ICWSM*. 2.
[15] Laura Dennison, Leanne Morrison, Gemma Conway, and Lucy Yardley. 2013. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *Journal of medical Internet research* 15, 4 (2013), e86.
[16] Richard Elliott and Clare Leonard. 2004. Peer pressure and poverty: Exploring fashion brands and consumption symbolism among children of the âĂŸBritish poorâĂŹ. *Journal of Consumer Behaviour* 3, 4 (2004), 347–359.
[17] Alan S Go, Dariush Mozaffarian, Véronique L Roger, Emelia J Benjamin, Jarett D Berry, Michael J Blaha, Shifan Dai, Earl S Ford, Caroline S Fox, Sheila Franco, and others. 2013. AHA statistical update. *Circulation* 127 (2013), e62–e245.
[18] Mowafa Househ. 2012. Mobile Social Networking Health (MSNet-Health): beyond the mHealth frontier. *Studies in health technology and informatics* 180 (2012), 808.
[19] Jin Huang, Feiping Nie, Heng Huang, and Chris Ding. 2014. Robust manifold nonnegative matrix factorization. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 8, 3 (2014), 11.
[20] Samuel Klein, Lora E Burke, George A Bray, Steven Blair, David B Allison, Xavier Pi-Sunyer, Yuling Hong, and Robert H Eckel. 2004. Clinical implications of obesity with specific focus on cardiovascular disease a statement for professionals from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism: Endorsed by the American College of Cardiology Foundation. *Circulation* 110, 18 (2004), 2952–2967.
[21] Kenneth D Kochanek, Sherry L Murphy, Jiaquan Xu, and Elizabeth Arias. 2013. Mortality in the United States, 2013. (2013).
[22] Abhishek Kumar, Piyush Rai, and Hal Daume. 2011. Co-regularized multi-view spectral clustering. In *Advances in neural information processing systems*. 1413–1421.
[23] Tricia M Leahey, Jessica Gokee LaRose, Joseph L Fava, and Rena R Wing. 2011. Social influences are associated with BMI and weight loss intentions in young adults. *Obesity* 19, 6 (2011), 1157–1162.
[24] Stephenie C Lemon, Milagros C Rosal, Jane Zapka, Amy Borg, and Victoria Andersen. 2009. Contributions of weight perceptions to weight loss attempts: differences by body mass index and gender. *Body image* 6, 2 (2009), 90–96.
[25] Qing Lu and Lise Getoor. 2003. Link-based classification. In *ICML*, Vol. 3. 496–503.
[26] Sofus A Macskassy and Foster Provost. 2003. *A simple relational classifier*. Technical Report. DTIC Document.
[27] Lydia Manikonda, Heather Pon-Barry, Subbarao Kambhampati, Eric Hekler, and David W McDonald. 2016. Venting Weight: Analyzing the Discourse of an Online Weight Loss Forum. In *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*.
[28] Elina Mattila, Raimo Lappalainen, Juha Pärkkä, Jukka Salminen, and Ilkka Korhonen. 2010. Use of a mobile phone diary for observing weight management and related behaviours. *Journal of Telemedicine and Telecare* 16, 5 (2010), 260–264.
[29] Micah O Mazurek, Paul T Shattuck, Mary Wagner, and Benjamin P Cooper. 2012. Prevalence and correlates of screen-based media use among youths with autism spectrum disorders. *Journal of autism and developmental disorders* 42, 8 (2012), 1757–1767.
[30] D Mozaffarian, EJ Benjamin, AS Go, DK Arnett, MJ Blaha, M Cushman, S de Ferranti, JP Després, HJ Fullerton, VJ Howard, and others. 2015. Heart disease and stroke statistics–2015 update: a report from the American Heart Association. *Circulation* 131, 4 (2015), e29–322.
[31] Jennifer Neville and David Jensen. 2000. Iterative classification in relational data. In *Proc. AAAI-2000 Workshop on Learning Statistical Models from Relational Data*. 13–20.
[32] Paul Poirier, Thomas D Giles, George A Bray, Yuling Hong, Judith S Stern, F Xavier Pi-Sunyer, and Robert H Eckel. 2006. Obesity and cardiovascular disease pathophysiology, evaluation, and effect of weight loss. *Arteriosclerosis, thrombosis, and vascular biology* 26, 5 (2006), 968–976.
[33] Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Galligher, and Tina Eliassi-Rad. 2008. Collective classification in network data. *AI magazine* 29, 3 (2008), 93.
[34] Katrina J Serrano, Mandi Yu, Kisha I Coa, Linda M Collins, and Audie A Atienza. 2016. Mining health app data to find more and less successful weight loss subgroups. *Journal of medical Internet research* 18, 6 (2016).
[35] Michael Shiner. 1999. Defining peer education. *Journal of Adolescence* 22, 4 (1999), 555–566.
[36] William E Snell. 1989. Willingness to self-disclose to female and male friends as a function of social anxiety and gender. *Personality and Social Psychology Bulletin* 15, 1 (1989), 113–125.
[37] Kimber L Stanhope, Jean Marc Schwarz, Nancy L Keim, Steven C Griffen, Andrew A Bremer, James L Graham, Bonnie Hatcher, Chad L Cox, Artem Dyachenko, Wei Zhang, and others. 2009. Consuming fructose-sweetened, not glucose-sweetened, beverages increases visceral adiposity and lipids and decreases insulin sensitivity in overweight/obese humans. *The Journal of clinical investigation* 119, 5 (2009), 1322–1334.
[38] Jiliang Tang, Huiji Gao, Xia Hu, and Huan Liu. 2013. Exploiting homophily effect for trust prediction. In *Proceedings of the sixth ACM international conference on Web search and data mining*. ACM, 53–62.
[39] Christopher C Tsai, Gunny Lee, Fred Raab, Gregory J Norman, Timothy Sohn, William G Griswold, and Kevin Patrick. 2007. Usability and feasibility of PmEB: a mobile phone application for monitoring real time caloric balance. *Mobile networks and applications* 12, 2-3 (2007), 173–184.
[40] Suhang Wang, Jiliang Tang, and Huan Liu. 2015. Embedded Unsupervised Feature Selection.. In *AAAI*. Citeseer, 470–476.
[41] Rena R Wing, George Papandonatos, Joseph L Fava, Amy A Gorin, Suzanne Phelan, Jeanne McCaffery, and Deborah F Tate. 2008. Maintaining large weight losses: the role of behavioral and psychological factors. *Journal of consulting and clinical psychology* 76, 6 (2008), 1015.
[42] Rongjing Xiang, Jennifer Neville, and Monica Rogati. 2010. Modeling relationship strength in online social networks. In *Proceedings of the 19th international conference on World wide web*. ACM, 981–990.
[43] Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf. 2004. Learning with local and global consistency. *Advances in neural information processing systems* 16, 16 (2004), 321–328.
[44] Xiaojin Zhu, Zoubin Ghahramani, John Lafferty, and others. 2003. Semi-supervised learning using gaussian fields and harmonic functions. In *ICML*, Vol. 3. 912–919.