# Research on the Relationship Between Perceived Social Support and Exercise Behavior of User in Social Network

**LICAI LEI** [ID] **AND ENYU ZENG** [ID]

Business School, Xiangtan University, Xiangtan 411105, China

Corresponding author: Enyu Zeng (zey1024@foxmail.com)

**ABSTRACT** A growing number of information technology systems and services have occurred to change users' attitudes or behaviors or both in the rapid development of mobile social media platforms. It is a new topic in the field of health communication whether the digitization and socialization of individual exercise behavior can stimulate health behavior. WeRun is a typical platform for the digitization and socialization of individual exercise. Based on 689 WeRun users' questionnaires, this study first repairs the missing and abnormal data by BP neural network. Then, the decision tree is used to evaluate the relationship between the perceived social support and exercise behavior under different intervention conditions, and detects the heterogeneous intervention effects for different pre-intervention profiles. In addition, this study further discusses the performances of social support features of persuasion technology in the WeRun. The data-driven method used in this study is beneficial to reducing self-selection bias and evaluate the intervention effect. The decision tree does not require decision-makers to have much expertise or to make parameter hypotheses while evaluating the intervention effect, and the results are more direct and intelligible. The results show that the decision tree can detect the heterogeneous intervention effect. In some cases, there is not a perfectly positive correlation between the degree of perceived social support and the number of average daily steps, and the relationship with friends has a great impact on the user's perceived social support. In addition, it also reveals the relationship between social comparison and perceived social support, and their interaction on exercise behavior. Finally, this study provides practical suggestions for the design and operation of e-health social network platform. The platforms are supposed to take corresponding persuasive strategies according to the various characteristics of users, so as to improve the continuous attention and participation of users.

**INDEX TERMS** Exercise behavior, perceived social support, persuasion technology, C4.5 decision tree, intervention, BP neural network.

## I. INTRODUCTION

Along with the increasing richness of material life, people are attaching more and more importance on personal health. To have a healthy body, people could start by taking a walk, which the World Health Organization (WHO) says is the best exercise in the world. Nowadays, there are a growing number of wearable devices for users to choose to record their exercise data with the Internet in real-time, such as smartphones, smart bands, etc. All those devices can monitor and manage users' exercise states and physical health, and promote the visualization and socialization of health data.

Numerous social media has become a kind of widespread tool for changing people's lifestyles [1], and the current escalating consumption structure is driving huge changes in consumer demand and preferences. Consumers have changed their original state of accepting services passively, and have more involved in activities that create value with businesses and other consumers, resulting in more positive consumer experiences [2]. One of the most important form is that consumers can get more consumer information and good emotional experiences through some behaviors, such as

---

The associate editor coordinating the review of this manuscript and approving it for publication was Honggang Wang [ID].

communication, interaction, sharing and mutual assistance, and those also serve the purpose of improving self and people's social welfare [3]. In this process, consumers can continuously develop their potential, seek self-expression and self-realization, and gain a higher level of inner happiness.

As a representative of mobile social media, WeChat (WeiXin in Chinese) is an all-inclusive platform of information acquisition, image display, service promotion and social communication [4], [5]. "The 2018 WeChat Annual Data Report" shows that the number of monthly active users of WeChat exceeded 1.082 billion all around the world. Obviously, WeChat has become the fastest-growing and largest-scale new media as well as one of the most important social platforms. What's more, it is free and easy to use. These are the major functions and features of WeChat:

- Messaging: text, voice, stickers, photos, group chat, video conference, location sharing...
- Moments: friends' updates via text, images, music, videos, articles, and allows likes and comments, just like Facebook and Twitter.
- Official Accounts: enterprise accounts and subscription accounts offer services to and update feeds for subscribers.
- WeChat Pay: a digital wallet allowing mobile payment for goods or services, money transaction and digital red packets between friends. (More information about WeChat can be obtained from https://www.wechat.com)

WeRun, similar to the database of a pedometer, is an Official Account of WeChat that can be paired with wearable fitness devices selected. As of September 2017, the number of daily active users of WeRun reached 115 million. WeRun is a health management system based on persuasion technology and it uses leaderboard as an incentive to promote user exercise. WeRun records the steps the user takes a day with the user's permission and the user will receive a daily message informing him(her) how close (s)he is to reaching his goal. In addition, the user will be able to know exactly how many steps (s)he has taken every day and that of his friends on WeChat who follow the official account. The user can also give the thumbs-up to friends' steps data as a form of social support to encourage individuals to get a better ranking on the leaderboard. What's more, users can view their own exercise records over the past period, share their real-time steps and rankings to friends, and even publish them on Moments. As an added incentive, if one user is number one on the leaderboard at the end of the day, (s)he is able to set the cover photo on all his(her) friends' leaderboards.

Nowadays, an increasing number of users are using WeRun to share exercise data and manage health, and interact with friends through the leaderboard to obtain certain image display and social support. The effects of user's exercise behaviors interacting and promoting each other in social networks are increasingly concerned by researchers and practitioners. Gradually, it became one of the hottest research topics in the field of electronic health service management. So, in WeRun, is the application of persuasion technology effective in health

management? Does the leaderboard have an incentive effect on users' exercise behavior? Does social support interact with social comparison? What are the differences in users' exercise behavior when they have different degrees of perceived social support? Will exercise behavior increase positively as social support increases in any case? This study is mainly based on these issues.

This study is based on a data-driven approach that provides a more direct and understandable result. We focus on exploring the heterogeneous relationship between users' perceived social support and exercise behavior in social networks, and further explores the performances of social support features of persuasion technology in the WeRun. In this paper, based on the issues and background, we first review the available literature related to this topic; then we introduce the data sources used in this study. Next, we introduce the research methods and report the analysis results after establishing the model. The paper concludes with research contributions and implications, as well as limitations.

## II. THEORETICAL BACKGROUND
### A. SOCIAL NETWORKS
Social networks are different from other forms of networks. Social networks refer to the relatively stable relationship system among individual members of society because of interaction. Network transitivity is an important property of social networks [6]. The existence of social networks enables individuals in the network to share knowledge better. Individuals in social networks do not exist independently, and their behaviors are often influenced by the people and society they are exposed to, that is, the existence of social networks affects individuals in the network in many ways [7], [8].

The influence of social networks on individual cognition may lead to individual behavioral changes. In early research, first of all, the researchers have studied how the interaction between users' behaviors on social networks affects healthy behaviors, pointing out that personal decisions and behaviors affect peer decisions and behaviors. In terms of physical health, Christakis and Fowler [9] show that obesity can spread through social relationships: a pair of twins, if one becomes obese, the other is 40 percent more likely to become obese. Similarly, mental health is also influenced by social networks. Fowler and Christakis [10] find that happiness in social networks spreads dynamically and it depends on the happiness of others associated with them, and that those who are surrounded by many happy people are more likely to be happy in the future.

The current development of social network applications embeds everyone tightly into an interactive social network [11]. Now, social networks include not only relationships that exist in reality, but also those on the Internet. Ghose and Han [12] proposed an empirical framework for user behavior based on mobile Internet space, arguing that social networks have a strong positive impact on user behavior. For example, in a study about exercise behavior, Aral and Nicolaides [13] showed that the user's exercise
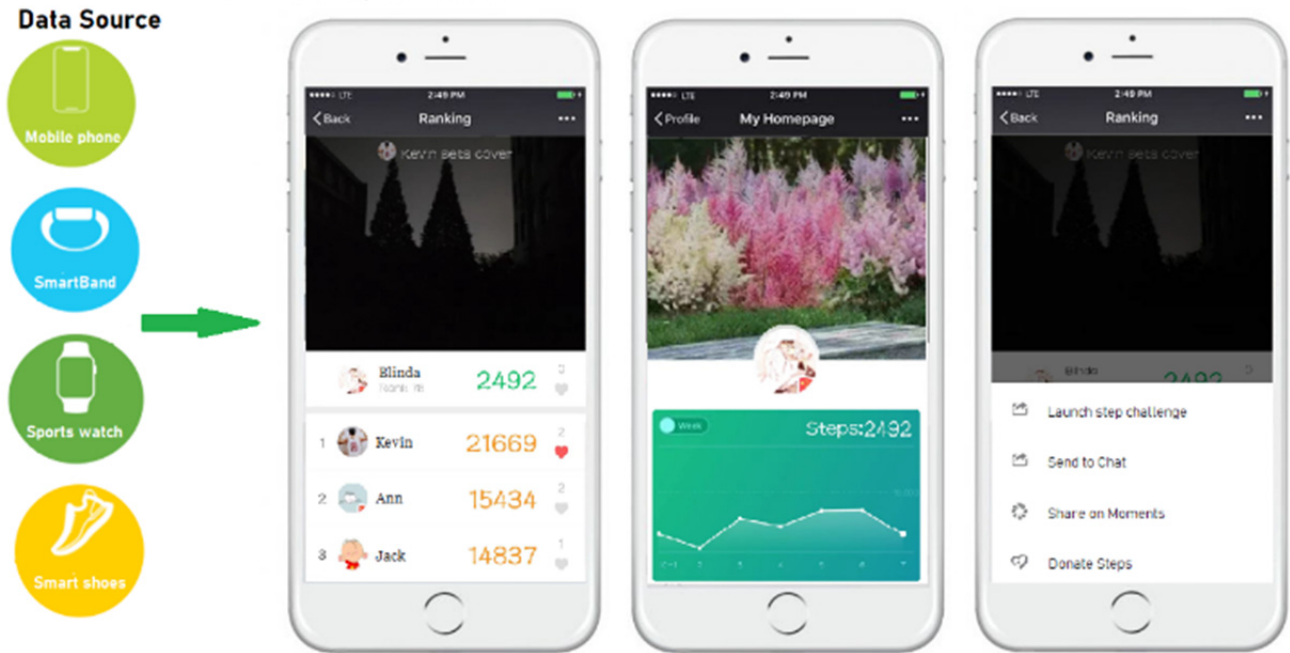
**FIGURE 1.** Some pages of WeRun.

behavior is socially contagious in global social networks. The increase of a user's steps may help increase the steps of his(her) peers over a period, and this contagion is also accompanied by differences in personal characteristics.

WeChat, as a representative of China's social applications, has a heterogeneous distribution of its user networks. Users with obvious core positions in the WeChat network have strong interpersonal skills and will have a certain influence on the behavior of others and the acquisition of information resources [14]. As one of the functions of WeChat, WeRun integrates social function and health management. The existence of leaderboards can motivate individuals' pursuit of health through external incentives, and transform personal health management behaviors by self-monitoring into group participation behaviors. Ordinarily, the leaderboard can cause competition among users, and the existence of competition will affect subsequent actions taken by users [15], [16].

### B. SOCIAL SUPPORT

Social support is the product of relying on social networks. It refers to a general term used to describe all kinds of support besides the individual, which is a social behavior accompanied by the existence of vulnerable groups. Cohen and Wills [17] believe that social support means providing individuals with psychological support and material resources in order to help them effectively cope with stress. Social support comes in many forms, Barrera [18] divides social support into six forms: material assistance, behavioral support, intimate interaction, guidance, feedback, and positive social interaction. In social networks, users' online participation behaviors can promote information exchange, image display and emotional support. Emotional support can more enhance users'

continuous participation and attention to online healthy communities than information exchange. Perceived social support is the cognition of social support, which can alleviate psychological stress and improve social adaptability [19], [20]. According to the direct effect model of perceived social support, perceived social support can directly promote people's posttraumatic adaptation by enhancing their healthy behavior [21], [22].

For one thing, many scholars have studied the relationship between social support and subjective feelings. Many studies have shown that social support is significantly related to subjective well-being, and positively correlated with individual life satisfaction and positive emotions [23], [24]. For another, many scholars have studied the impact of social support on health status. For example, in terms of weight control, people who receive social support have a higher rate of completing weight-loss treatment and maintaining weight loss than those who are treated alone [25]. Moreover, different levels of social support can cause significant differences in weight-loss effects [26]. Social support can motivate individual behavior by promoting self-regulation, especially self-efficacy [27].

With the rapid development of interactive technology based on the Internet, the number of virtual communities has increased dramatically, and the platforms for people to exchange social support are more reliant on online interactive communities [28]. Internet technology has greatly improved the timeliness and convenience of interaction, which expands the scope of social networks. It can make online mutual assistance as a kind of social behavior that can occur at any time. Yao *et al.* [29] found that patients can get more emotional support and information support through online social networks, which is conducive to promoting their physical and mental

health. In terms of the impact of social support on exercise behavior, social support behaviors such as comment, giving thumbs-up and forwarding will have a positive short-term impact on the user's exercise behavior in the social network relying on the Internet [30]. Similarly, in WeRun, users can give thumbs-up to each other's steps data. This behavior of giving thumbs-up means expressing a kind of concern and appreciation to others, which is also a sign of social support behavior.

### C. PERSUASION TECHNOLOGY

In daily life, people's behavior is tempted and influenced by various external factors. The role of computer technology in changing people's behaviors and attitudes has become more apparent with the development and rapid popularization of computer technology and communication technology, especially network and personal intelligent terminals have penetrated all aspects of our daily life [31]. Persuasion technology refers to a computer interaction system that changes the attitudes and behaviors of users. ''Persuasion'' is the process of behavior change, and ''technology'' refers to computer technology [32]. It is worth mentioning that although some means of coercion and deception can temporarily achieve the purpose of changing users' behaviors and attitudes, this is not the original intention of persuasion. The persuasion in persuasion technology refers to an attempt to change behavior and attitude, which emphasizes voluntary and spontaneous change. Fogg behavior model (FBM) proposes behaviors change needs three elements—motivation, abilities, and trigger [33]. The generation of target behavior must be provided with three indispensable conditions: sufficient motivation, sufficient ability and effective trigger. In practical applications, one's motivation and ability indicate the probability of behavior, while the trigger requires the use of persuasion strategies to generate and promote the behavior. Persuasion strategy is the specific method to realize persuasion technology. Computer technology can play the role of tools, media and social man in persuasion technology. Generally, different roles adopt different persuasion strategies.

Persuasion technology plays an important role in areas that influences users to change certain behaviors or attitudes. For instance, a system uses persuasion technology to convey environmental protection information to residents, thereby affecting people's consumption patterns and improving energy-saving awareness of people, especially young people [34]. It also can be used in ubiquitous computing systems that improve hand-washing behavior at the sink [35]. In social interaction activities, systems such as Facebook and Mixi use persuasion technology to make closer communication between family and friends, thereby establishing good social relationships [36]. In the specific application of self-health management, persuasion technology mainly focuses on motivating users to strengthen physical exercise, facilitating users to adjust and develop healthy eating habits, and helping obese patients to lose weight and self-management of chronic diseases [37]–[39]. As a health

management platform with social function, one of WeRun's important systematic feature is social support. The system features explain how the suggested design principles can and should be transformed into software requirements and further implemented as actual system features. The design principles in the social support feature describe how to design the system so that it motivates users by leveraging social influence, which is a key feature discussed in this paper. Based on the design principles of social support category developed by Oinas-Kukkonen *et al.* [40], we summarize the implementations of the WeRun under different principles, as shown in Table 1.

In summary, the behavior of users in social networks is contagious. Changes in one user's behavior will affect other users' behaviors in the social network. The existence of social support can enhance certain behaviors in the social network. The application of persuasion technology in WeRun can motivate users to better manage their own health. In the previous research, however, there are still many flaws and shortcomings in the methods and perspectives. Firstly, when evaluating the effect of social support interventions, it is best to make the characteristics (except the social support) of the control group be as similar as possible to the control group, which can better explain the relationship between social support and exercise Behavior. However, due to the costs, ethic and other considerations, strictly controlled experiments are rarely seen in social science research. Secondly, most of the existing studies do not consider the selection bias of users and the heterogeneous intervention effects. The common PSM supposes the intervention effect is homogeneous, which is difficult to explore the heterogeneous intervention effects of different subgroups. However, because of the existence of differences it is more effective to formulate policies for different groups of population than policies that only consider the average treatment effect. Finally, social support is the mutual behavior among users. In WeRun, users can give social support to others while accepting support from other friends. However, some studies neglected this kind of reciprocity and the subjective feelings of users on social support. Considering these shortcomings, we established a data-driven method. The decision tree does not need to make parameter hypotheses when evaluating intervention effects. Moreover, it provides a more direct and easier-to-understand result for communication in research and practice. Lastly, the results provide practical guidance for the operation of the e-health social platform.

## III. DATA SET
### A. DATA SOURCE

This study designed a questionnaire with many measurement scales and then used an online survey to collect data. Before the questionnaire was formally issued, 40 students were asked to conduct a pre-interview of the questionnaire. Then, after the questionnaire was collected, the revised opinions were collected and preliminarily analyzed. Finally, the questionnaire was formally issued after further revision based on

**TABLE 1.** Implementations of the WeRun under social support feature.

| Principle | Example requirement | Example implementation |
|---|---|---|
| Social learning | System should provide means to observe other users who are performing their target behaviors and to see the outcomes of their behavior. | Users of the WeRun leaderboard can share steps data. |
| Social comparison | System should provide means for comparing performance with the performance of other users. | The WeRun leaderboard allows users to update and rank their steps in real-time. |
| Normative influence | System should provide means for gathering together people who have the same goal and get them to feel norms. | WeRun offers ranking of user's friend in real-time; share user's steps to other friends, and even publish them on Moments. |
| Social facilitation | System should provide means for discerning other users who are performing the behavior. | Real-time update of steps data and ranking. |
| Cooperation | System should provide means for co-operation | None. |
| Competition | System should provide means for competing with other users. | Steps leaderboard; select friend to launch step challenge. |
| Recognition | System should provide public recognition for users who perform their target behavior. | Number one on the leaderboard can set the cover photo on leaderboards; give thumbs-up to the number of friend's steps; get thumbs-up through share the data into Moments. |

the opinions and preliminary analysis results. This questionnaire uses the sample service of the Questionnaire Star (WenJuanXing in Chinese, http://www.wjx.cn) to formally collect the questionnaire data. The main way is to publish links and invite Chinese WeChat users to fill out the questionnaire through various social network channels.

The questionnaire is divided into three parts: the questionnaire participant's basic information, the measurement of exercise attitude and the measurement of WeRun participation. The user's basic information includes the user's personal characteristics and main factors that affect the number of steps: gender, age, willingness to invest in exercise and the main source of the steps, etc. It is generally believed that high-income people will pay more attention to physical health, and education is closely related to income. So, education and income [41], [42] are also worth to be attached importance to. In order to be applicable to all users, the measurement scale of user's exercise attitude was revised based on Mao's [43] "Exercise Attitude Scale." Both the Theory of Reasoned Action (TRA) and the Theory of Planned Behavior (TPB) proposed by Fishbein and Ajzen points out that attitudes affect the behavioral intentions of individuals, and eventually affects individual behaviors. Based on these two theories, Mao established the Attitude-Behavior Causal Model, and then designed the "Exercise Attitude Scale." The scale is composed of eight subscales, including the user's behavioral habit, behavioral attitude, target attitude, behavioral cognition, behavioral intention, emotional experience, behavioral control, and subjective criteria. The questionnaire used the 5-point Likert scale to measure the degree of the user's exercise attitude. There were 66 attitude measurement questions including both positive and negative ones. Then, the total score was used to measure the user's exercise attitude. The measurement of WeRun participation, the third part, is used to measure the user's performance in the whole process of participating WeRun when under the guidance of persuasion technology [42].

In order to ensure the participants in this survey are target subjects, we set up multiple filtering questions in the questionnaire to ensure that the participants have real experience of using the pedometer software. From the release of the questionnaire in December 2017 to the end in late June 2018, a total of 969 answers were collected, among which 689 valid questionnaires were retained after excluding some invalid samples(such as too short response time, not using pedometer software, variables with abnormal values and with an unserious attitude to respond), the valid rate is 71.104%. The basic information of sample demographics is shown in Table 2.

Social support includes the exchange of resources between the conscious individuals, who are both the giver and the recipient [44]. In the WeRun relying on the Internet, the user can not only be a giver who gives a thumbs-up to friends' steps data, but also be a recipient who accepts a thumbs-up from friends. Because of the reciprocity in the exchange of social support, the user's subjective perception of social support (that is, whether users can truly feel the social support from others and willing to provide social support) is also a question worth considering. The perceived social support variables in this study were measured by the mean of the possibility of users giving thumbs-up to friends and the perceived degree of being given the thumbs-up by friends, both of those two variables were obtained by 5-point Likert scale. The calculation formula of perceived social support is

$$
\begin{aligned}
SociSup = (&\textit{the possibility of users giving thumbs} \\
&-\textit{up to friends} \\
+&\textit{the perceived degree of being given the thumbs} \\
&-\textit{up by friends})/2
\end{aligned}
$$

## B. RELIABILITY AND VALIDITY
The reliability and validity of the exercise attitude scale were tested. The KMO value of the sample data is 0.973, which was obtained by using principal component analysis and indicates

**TABLE 2.** Statistical characteristics of sample variable.

| Variable | Variable Description | Scale | Mean | Standard Deviation | Percentage |
|---|---|---|---|---|---|
| Gender | Gender of user | 1=male | | | 48.476% |
| | | 0=female | | | 51.524% |
| Age | Age range of user | 0=under 18 | | | 2.322% |
| | | 1=18~24 | | | 56.023% |
| | | 2=25~34 | 1.594 | 0.880 | 25.835% |
| | | 3=35~44 | | | 11.611% |
| | | 4=45~59 | | | 4.209% |
| | | 5=60 or over 60 | | | 0% |
| Edu | Level of education of user | 0=high school and below | | | 12.627% |
| | | 1 =Bachelor | 1.025 | 0.556 | 73.875% |
| | | 2=Master | | | 11.901% |
| | | 3=Doctor | | | 1.597% |
| IncoMon | Monthly disposable income | 0=under 1000 RMB | | | 19.594% |
| | | 1=1000~2499 | | | 37.446% |
| | | 2=2500~4999 | 1.583 | 1.271 | 19.448% |
| | | 3=5000~7999 | | | 14.514% |
| | | 4=8000~11999 | | | 6.531% |
| | | 5=12000 RMB or above | | | 2.467% |
| InveMon | The willingness to invest money in exercise | 1 = willing | | | 93.033% |
| | | 0=unwilling | | | 6.967% |
| StepSour | The main source of steps | 1 = sports and fitness | | | 31.930% |
| | | 0 = daily walking | | | 68.070% |
| AttituSco | Total score of exercise attitude | Continuous variable (range 66~330) | 245.409 | 40.628 | |
| CheckRank | Possibility of viewing the leaderboard every day | 5-point Likert scale | 3.975 | 1.085 | |
| CompaFro | Possibility of comparison with top-ranked friends | 5-point Likert scale | 3.475 | 1.227 | |
| CompaClo | Possibility of comparison with close friends | 5-point Likert scale | 3.515 | 1.277 | |
| CatchUp | The idea of trying to catch up with your friends | 1=have such an idea | | | 52.104% |
| | | 0=not have | | | 47.896% |
| SociSup | The degree of perceived social support | 5-point Likert scale | 3.499 | 1.207 | |

**TABLE 3.** Tests results of reliability and convergent validity.

| Dimension | CR | Cronbach's α | AVE |
|---|---|---|---|
| Behavioral attitude | 0.901 | 0.890 | 0.536 |
| Target attitude | 0.945 | 0.866 | 0.632 |
| Behavioral cognition | 0.920 | 0.929 | 0.623 |
| Behavioral habit | 0.960 | 0.959 | 0.706 |
| Behavioral intention | 0.938 | 0.948 | 0.657 |
| Emotional experience | 0.932 | 0.960 | 0.580 |
| Behavioral control | 0.950 | 0.916 | 0.702 |
| Subjective criteria | 0.902 | 0.885 | 0.651 |

that the scale is suitable for factor analysis. The reliability of the scale is measured by calculating each measurement dimension's combined reliability (CR), Cronbach's alpha coefficient and average variance extracted (AVE), and calculating the factor loading of each measured item. The specific results are shown in Table 3 and Appendix. Table 3 shows that the CR and the α coefficient value of all measurement dimensions are both greater than 0.700, indicating adequate reliability. Moreover, the AVE of each measurement dimension is above 0.500, and the factor loading of each item in Appendix is greater than 0.500 and most of them are

above 0.700, indicating that the sample data has a good convergent validity. The analysis results show that the reliability and validity of the measurement scale meet the study's requirement.

## IV. DATA REPAIRING

Among the 689 questionnaires, some of them were filled out by users who were unable to assess their average number of daily steps, and some were filled with abnormal data on the steps. These questionnaires were listed as samples of missing dependent variable values, a total of 175. In order not to waste samples and make the decision tree model have more analysis samples, this study repaired the missing and abnormal values. In the preliminary exploratory analysis, we found that there is a certain non-linear relationship between the average daily steps and other independent variables. Among the commonly used prediction models, the traditional econometric regression models need to make premise assumptions about the complex non-linear relationships between independent variables, while BP neural networks have strong non-linear mapping, high self-learning and adaptive capabilities. Therefore, the BP neural network model is used to predict and correct the sample of missing dependent variable values in this study.
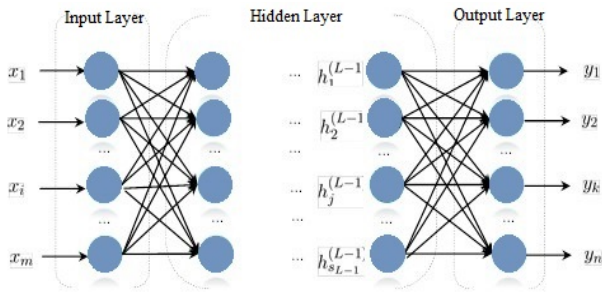
**FIGURE 2.** Multi-layer neural network framework.

## A. BP NEURAL NETWORK ALGORITHMS

The neural network is a broadly parallel interconnected network composed of simple adaptive units whose organization can simulate the interaction of biological nervous systems with real-world objects. BP neural network is a kind of multilayer feedforward neural network based on the backpropagation algorithm. Its main feature is that the signal forward propagation and error backward propagation. By continuously adjusting the network weight value, the final output of the network is as close as possible to the expected output, so as to achieve the purpose of training [45].

### 1) MULTI-LAYER NEURAL NETWORK FRAMEWORK

Fig. 2 shows a typical multi-layer neural network consisting of many nodes and interconnections between them. Usually, a multi-layer neural network consists of L-Layer neurons, in which the first layer is called the input layer, the last layer (the lth layer) is called the output layer, and the other layers are called the hidden layer (the 2nd–l-1th layer). The number of layers of the network and the number of nodes in each layer are set according to specific conditions. each of these nodes represents a specific output function called activation function. The connection between every two nodes represents a weighting value for the connection signal, called weight, which is equivalent to the memory of the artificial neural network. The output of the network varies depending on the connection method, the weight value, and the activation function of the network.

### 2) FORWARD TRANSMISSION PROCESS OF BP NEURAL NETWORK

Initialize all the weights and thresholds in the neural network and set them to a smaller random number, based on which the forward propagation of the input is performed. Each node accepts input signals from the upper-level nodes, each signal is passed through a weighted connection. The nodes add these signals together to get a total input value, and then compare the total input value with the threshold value of the node. Then, through an "activation function" process to get the final output, this output will be passed as a layer of the input of the next layer.

First, calculate the total input of the node. The output of each previous-layer node connected to the node is multiplied

by the corresponding link weight and then accumulated to form the total input of the node. i is a node in layer l, the total input $I_i$ of this node is calculated as:

$$I_i = \sum_{j=1}^{S_{l-1}} \omega_{ji} O_j + \theta_i \tag{1}$$

In (1), $S_{l-1}$ is the number of nodes in the l-1th layer, $\omega_{ji}$ is the connection weight between the node j of the l-1th layer and the node I of the lth layer, $O_j$ is the output of the node j in the l-1th layer, and $\theta_i$ is the bias of the node I of the lth layer.

Then the activation function $f(I) = 1/(1+e^{-I})$ is applied to the total input $I_i$ to form the output $O_i$ of the node, and the calculation formula is:

$$O_i = 1/(1+e^{-I_i}) \tag{2}$$

### 3) BACK PROPAGATION OF ERRORS AND UPDATE OF PARAMETERS

The basic idea of back propagation is to constantly adjust the network parameters by calculating the error between the output layer and the expected value so that the error becomes smaller and the direction is gradually fed back from the output layer to the input layer. In order to find the minimum error parameter, the neural network needs to find the minimum value of the error according to the negative gradient direction, so as to obtain the final model parameters. Firstly, the error formula is defined as:

$$Err = \frac{1}{2} \sum_{i=1}^{S_l} (T_i - O_i)^2 \tag{3}$$

In (3), $S_l$ is the number of nodes in the lth layer, and $T_i$ is the actual value of node I in the lth layer.

Each iteration updates the weights and bias according to the following formula:

$$\omega_{ji} \xleftarrow{\text{update}} \omega_{ji} - \alpha \frac{\partial Err}{\partial \omega_{ji}} \tag{4}$$

$$\theta_i \xleftarrow{\text{update}} \theta_i - \alpha \frac{\partial Err}{\partial \theta_i} \tag{5}$$

In (4) and (5), $\alpha$ is the learning rate, and its value range is (0,1).

## B. MODELING PROCESS

Firstly, pre-processing the entire sample. Before the neural network is established, the data is normalized. Normalization can prevent the hidden nodes from being prematurely saturated and avoid the impact of larger values on the output [46]. This study uses the Min-Max normalization principle to convert the data to the value in the range [0,1]. The formula is as follows:

$$x^* = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

In (6), $x_{max}$ is the maximum value of the variable x in the sample, $x_{min}$ is the minimum value of the variable x in the sample, x is the original value, and $x^*$ is the normalized value.

Secondly, dividing the data set. There are 514 samples with known values of dependent variables in the whole sample. Then, 70% of these samples were randomly selected and set to the training set, the remaining 30% of these samples as a test set to test the performance of the constructed neural network. The sample classification is shown in Fig.3.
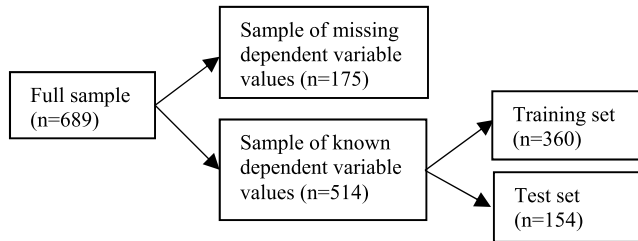


**FIGURE 3.** Sample classification.

Finally, constructing a BP neural network model. The input layer of the model is the 12 variables in Table 2, and the output layer is only 1 variable, which is the number of user's daily average steps. The structure of the network directly affects its approximation and generalization ability, so we adopt the exploratory method to confirm the number of layers and nodes in the hidden layer. The mathematical theory shows that the three-layer neural network has a good predictive ability. Due to the complexity of the neural network, we only explore the performance of the three-layer and four-layer neural networks, that is, one or two hidden layers.

We specify that the next layer has fewer nodes than the previous layer in the forward direction. After determining the neural network parameter values, the training set is used to construct the network, and then the test set is put into the constructed network. Because the neural network constructed in this study is to predict the value of the dependent variable, which is a continuous variable, we use the mean squared error (MSE) of the test set to measure the performance of different neural network models. Taking the network with two hidden layers as an example, when the hidden layer of the first layer is 10 nodes and the second layer takes different values, the MSE values of the test set of different neural network models are shown in Table 4. Fig. 4 shows the difference between the predicted values and the actual values of the sample of known dependent variable values under the different network models.

**TABLE 4.** MSE values of the test set of different neural network models.

| The number of nodes | (12,10,1,1) | (12,10,2,1) | (12,10,3,1) |
|---|---|---|---|
| MSE | 0.023 | 0.023 | 0.021 |
| The number of nodes | (12,10,4,1) | (12,10,5,1) | (12,10,6,1) |
| MSE | 0.0275 | 0.025 | 0.027 |
| The number of nodes | (12,10,7,1) | (12,10,8,1) | (12,10,9,1) |
| MSE | 0.020 | 0.025 | 0.026 |

Generally, there is a contradiction between the predictive ability and the training ability of the BP neural network.
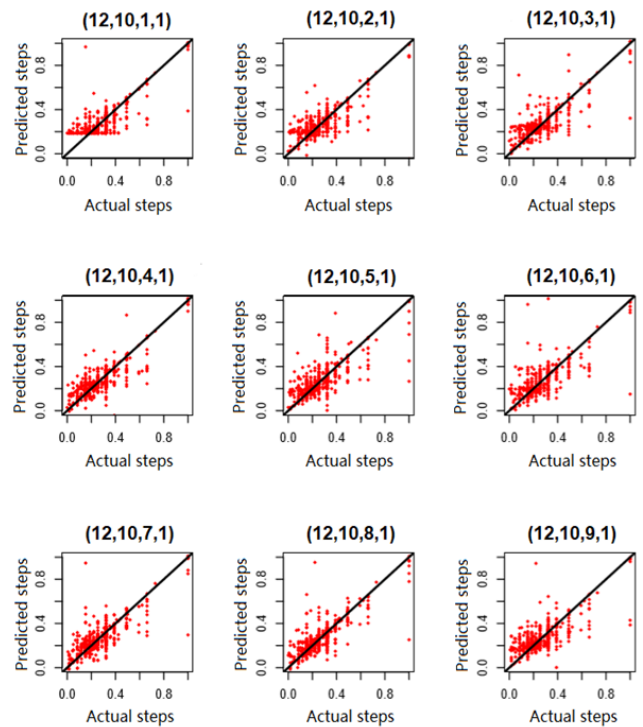


**FIGURE 4.** The difference between the predicted steps and the actual steps.

In order to reduce the possibility of "overfitting", the number of repetitions for the neural network's training is 3 in this study. After continuous exploration of the number of layers and nodes in the hidden layer, we eventually construct a four-layer neural network, which in the first hidden layer is 10 nodes and the second hidden layer is 7 nodes. That is, the number of nodes is (12, 10, 7, 1).

## C. RESULT ANALYSIS

In order to illustrate the advantages of the constructed neural network model, the BP neural network model is compared with commonly used prediction models, including K-nearest neighbors regression (KNN), support vector regression (SVR) and linear regression (LR). In addition, in order to make the predictor variables in the constructed linear regression model explain the response variables better, we use the all-subsets regression to select the final predictor from the 12 candidate variables. That adjusted $R^2$ is used to evaluate the performance of the linear regression model under different subsets, the results are shown in Fig. 5. In Fig. 5, the ordinate presents the adjusted $R^2$ value which gradually increases from the bottom to the top. The adjusted $R^2$ value of the top is the largest, indicating that the model has the best explanatory ability than others. At this time, the variables of the model include gender, education level, the main source of steps, total score of exercise attitude, possibility of viewing the leaderboard every day and possibility of comparison with top-ranked friends. The final constructed linear regression
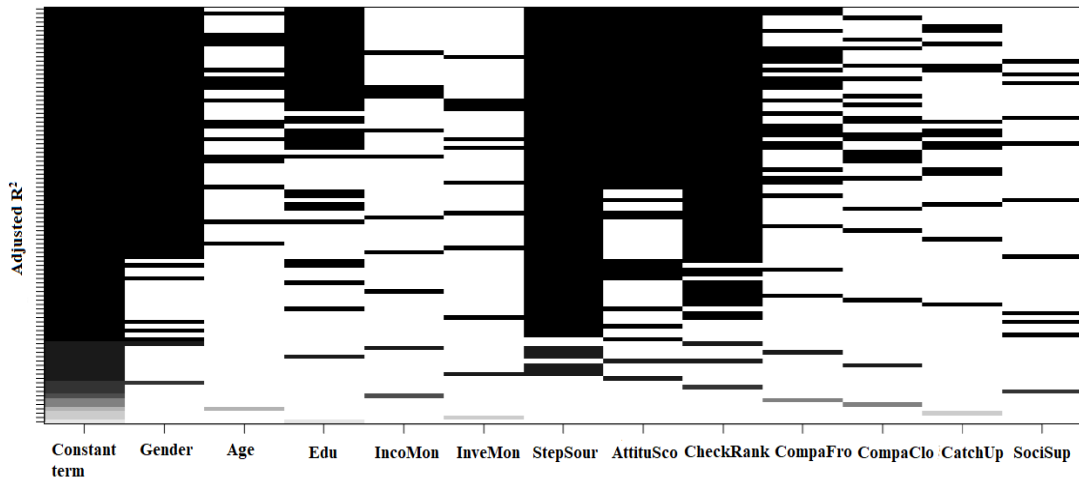
**FIGURE 5.** Regression model with different subset.

model is:

$$
\begin{aligned}
Averaged&ailynumberofsteps_i \\
&= \beta_0 + \beta_1 \times Gender_i + \beta_2 \times Edu_i \\
&\quad + \beta_3 \times StepSour_i + \beta_4 \times AttituSco_i \\
&\quad + \beta_5 \times CheckRank_i + \beta_6 \times CompaFro_i \\
&\quad + \varepsilon
\end{aligned}
\tag{7}
$$

In (7), i represents user i, $\beta_0$ is a constant term, $\beta_1 \sim \beta_6$ are regression coefficients, $\varepsilon$ is a random error term.

The generalization ability of the network model is closely related to the typicality of the training set. The models established on the existing training set may not perform well on the new dataset. Therefore, the method of 10-fold cross-validation is used to evaluate the generalization ability of the model. The main idea of 10-fold cross-validation is to divide the sample into 10 sub-samples, and then take 9 sub-samples in turn as the training set and the remaining 1 sub-sample as the test set. This method generates 10 predictive models, and the predicted results and evaluation indexes of 10 test sets are recorded. Finally, the average of the indexes is calculated. Since the samples of the test set do not involve the selection of the model's parameters, the model can obtain a more accurate estimation of the new data.

10-fold cross-validation was performed on the BP neural network, KNN, SVR, and LR respectively. Then, the average values of MSE, MAE, and MAE% of the test sets were used to evaluate the performance of each model. The evaluation index values of the 4 models are shown in Table 5, and the difference between the predicted steps and the actual steps is shown in Fig. 6.

The index values of each model in Table 5 show that the BP neural network model's MSE is 0.017, the MAE is 0.095, and the MAE% is 0.359, which are smaller than the evaluation index values of the other 3 models. Besides, Fig.6 shows that the relationship between predicted and actual values in the BP neural network model is more concentrated on the

**TABLE 5.** The evaluation index values of the 4 models.

| | MSE | MAE | MAE% |
|---|---|---|---|
| KNN | 0.028 | 0.118 | 0.446 |
| SVR | 0.022 | 0.101 | 0.381 |
| LR | 0.022 | 0.102 | 0.387 |
| BP neural network | 0.017 | 0.095 | 0.359 |

Note: In K-nearest neighbors regression, where K (the number of neighbors) is 3. In support vector regression, where the type of kernel function is RBF and the regularization parameter is 0.1.
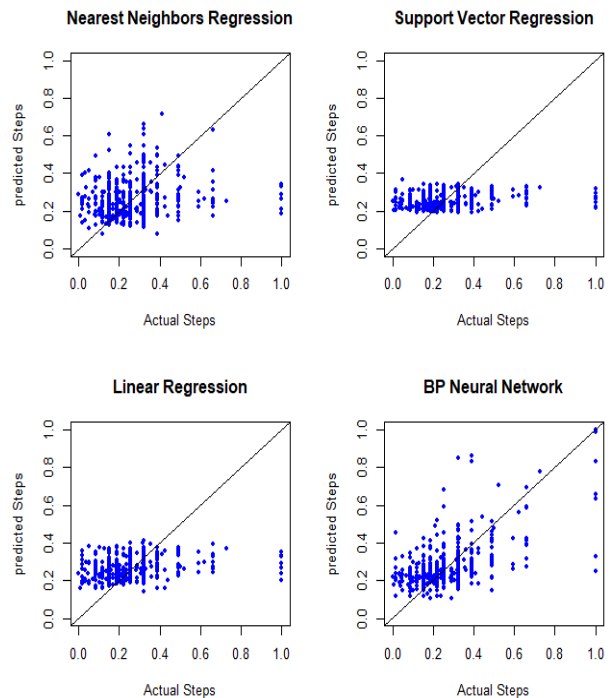


**FIGURE 6.** The difference between the predicted steps and the actual steps.

diagonal line. In sum, this shows that the BP neural network model constructed in this paper is more ideal for predicting the whole sample.
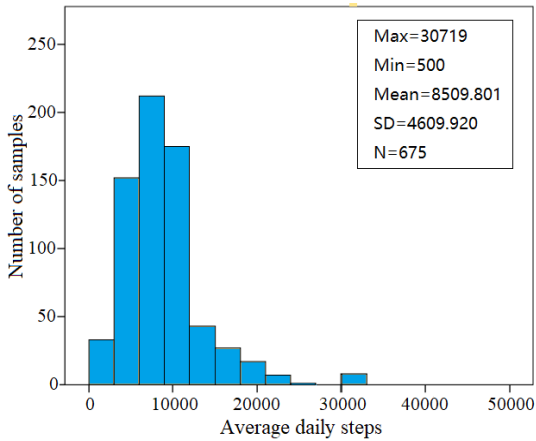
**FIGURE 7.** Bar graph of repaired average daily steps.

Finally, we used the neural network model constructed to repair the 175 samples with missing and abnormal values. In order to overcome the influence of outliers, we truncated the steps data at the end of 1% tail after repairing the samples, and the final sample size after processing was 675. Fig. 7 shows the average daily steps distribution of the processed sample.

## V. INTERVENTION EFFECT ASSESSMENT
### A. C4.5 DECISION TREE ALGORITHMS
The decision tree is a classic data mining method, which aims to predict the value of the target/output variable by establishing an adaptive tree model based on the values of several input variables. Decision trees are simple, intuitive and highly interpretable, which in data mining can be divided into two categories: classification trees and regression trees. The classification tree is for the situation where the target variable is discrete, that is, the final goal is to predict the category of each sample, such as to determine whether a certain product will be purchased according to the customer characteristics. A regression tree is applicable to the continuous target, such as predicting someone's monthly income. The decision tree model is highly automated and the results are easy to understand. It does not require to make parameter hypotheses during modeling and is well used for forecast and classification [47]–[49].

Let $D$ be the set of current data samples and $|D|$ be the number of sample sets. The category attribute has $K$ different values, corresponding to $K$ different category sets $C_k$. $k \in \{1, 2, 3, \cdots, K\}$, $|C_k|$ is the number of samples in the category set $C_k$, then the formula of information entropy is:

$$H(D) = -\sum_{k=1}^{K} p_k \log_2 p_k \qquad (8)$$

In (8), $p_i = |C_i|/|D|$, which represents the probability that any one data object belongs to the category set $C_i$.

The information entropy formula required to divide the current sample set using attribute A is:

$$H(D \mid A) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} H(D_i) \qquad (9)$$

In (9), $D_i$ represents a subset which is divided from the data set $D$ based on the attribute $A$, $|D_i|/|D|$ represents the weight given to each subset, and $n$ is the total number of categories of the attribute $A$, that is, how many $D_i$ there are.

Therefore, the *Information Gain* obtained by using attribute A to partition the current branch node into its corresponding subsets is:

$$Gain(D, A) = H(D) - H(D \mid A) \qquad (10)$$

The ID3 algorithm is based on information gain. The core of the algorithm is to apply the information gain criterion to select the classification attribute on each sub-node of the decision tree and to recursively construct the decision tree. The method is as follows: calculating the information gain of all possible attributes from the root node; selecting the attribute with the largest information gain as the attribute of the node, and establishing the child node by different values of this attribute; then recursively calling the above method on the child nodes to construct the decision tree's branch until the information gain of all attribute is very small or no attribute can be selected; finally, a decision tree is formed. However, the disadvantage of ID3 algorithm is that it tends to select the attribute with a large number of value and cannot handle continuous attributes. C4.5 is an improved algorithm of ID3 which can process non-discrete data and use an information gain called gain ratio to compensate for the defects of ID3 [50]. The gain ratio punishes attributes with more values by introducing an item called "split information", which is used to measure the breadth and uniformity of the attribute in data splitting. The formula for the gain ratio is:

$$GainRatio(D, A) = \frac{Gain(D, A)}{SplitInformation(D, A)} \qquad (11)$$

$$SplitInformation(D, A) = -\sum_{i=1}^{n} \frac{|D_i|}{|D|} \log_2 \frac{|D_i|}{|D|} \qquad (12)$$

In (11) and (12), the $GainRatio(D, A)$ is the gain ratio based on attribute $A$, and the $SplitInformation(D, A)$ is the balance of the information gain for attribute $A$.

At each branch node, the decision tree C4.5 algorithm calculates the information gain ratio of each attribute, and selects the attribute with the highest information gain ratio as the attribute of the subset partitioning on the node. The method above is recursively called on the child nodes, and then stops the construction of the decision tree until the information gain ratio falls below a certain threshold.

### B. MODELING PROCESS
Decision trees are commonly used in business analysis and data mining applications. As a transparent algorithm, it is easy to interpret in data mining and does not have to make
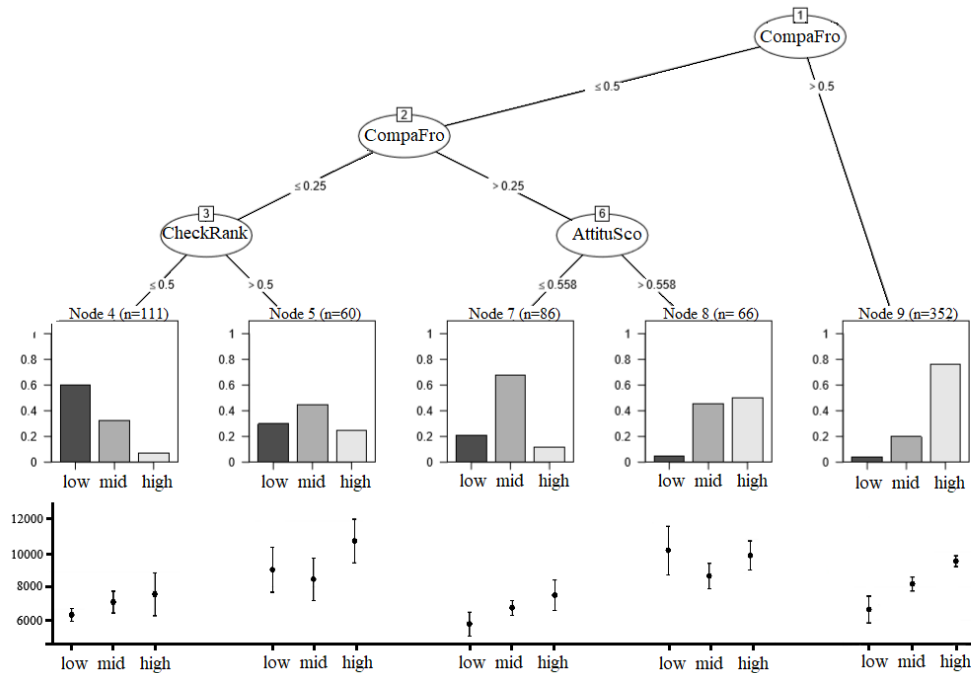
**FIGURE 8.** Results of decision tree.

parameter assumptions about the relationship between results and input variables. The decision tree divides the data into multiple regions, each region is a terminal node. Observations are homogeneous in their predictor values and outcome probabilities or distributions are within a particular terminal node. This feature enables decision trees to effectively predict the outcome values of individual observations and to select important predictor variables [51]. This study is based on the tree-based method for assessing the effects of self-selective interventions proposed by Yahav *et al.* [52] to evaluate the differences in average daily steps between users under different intervention conditions, i.e., under varying degrees of perceived social support.

This study constructs a decision tree based on the C4.5 algorithm to assess the intervention effect, including four main steps, as follows:

1. *Classify users according to the degree of perceived social support.* The perceived social support is a discrete variable measured by the scale and its value is between 0 and 1 after normalization. We define users with a perceived social support value of {0, 0.125, 0.25} as users with low perceived social support, with a value of {0.375, 0.5, 0.625} as users with medium perceived social support, and with a value of {0.75, 0.875, 1} as users with high perceived social support.

2. *Generate a decision tree model.* A classification tree is constructed. The 11 variables in Table 2 except perceived social support are input variables (predictors), and the perceived social support is used as the outcome variable of the intervention condition.

3. *Identify imbalance variables.* The generated tree shows the splitting variables and their values. The split and combination of the variables are unbalanced, and each terminal node displays the distribution of the intervention conditions for its observations.

4. *Measure the intervention effect of each terminal node.* Each terminal node is treated as a separate subsample, and all observations within each terminal node have the same input variable value and intervention probability. Within each terminal node, the effect of the intervention is evaluated by comparing the average daily steps values of users with different degrees of perceived social support.

If the parameters of the decision tree are not set, the resulting decision tree will be very detailed and huge, and each input variable will be considered in detail. Although this tree can classify all observations with 100% accuracy, the observation size of the root node is very small because of overfitting, which is meaningless for our comparative analysis. Therefore, we set the minimum size of observations in the terminal node to 50 and the confidence threshold to 0.25. The resulting tree is shown in Fig. 8.

### C. RESULT OF INTERVENTION EFFECT
In Fig. 8, the generated tree found that the possibility of comparison with top-ranked friends (CompaFro) has the greatest effect on perceived social support, and it largely determines the relationship between the users' perceived social support and their exercise behavior. Node 9 indicates that users with high CompaFro values are also most likely to have a high degree of perceived social support. CompaFro is

the first unbalanced variable detected by the tree, and also the largest contributor of all input variables to tree generation. So, according to the value of CompaFro, this study divided the sample into 2 subsets, A and B. Through one-way analysis of variance, we found that in the whole sample and two subgroups, the steps of low perceived social support and medium perceived social support were not significantly different. Therefore, then we combined the subsets of low perceived social support and medium perceived social support into one set (non-high perceived social support). The average daily steps of different sets are shown in Table 6 and Fig. 9.
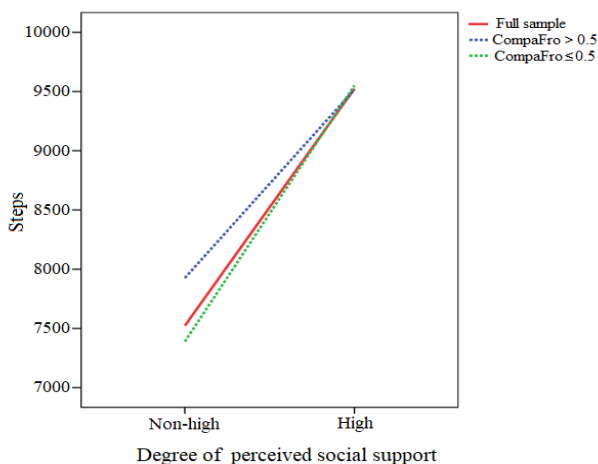


**FIGURE 9.** The effects of perceived social support on average daily steps.

**TABLE 6.** The average daily steps of different subsets.

|  | Full sample (n=675) | CompaFro>0.5 (A, n=352) | CompaFro≤0.5 (B, n=323) |
|---|---|---|---|
| High perceived social support | 9,523.784 | 9,517.082 | 9,550.894 |
| Non-high perceived social support | 7,522.503 | 7,925.824 | 7,389.109 |
| Difference | **2,001.281** (t=5.773) | **1,591.258** (t=2.691) | **2,161.785** (t=3.704) |

The results in Table 6 and Fig. 9 show that users with different perceived social support degrees have significant differences in the number of steps. In the whole sample, the daily steps of the users with high perceived social support is 2,001.281 more than that of other users. The first imbalance variable (CompaFro) detected by the tree divides the whole sample into 2 subsets: A and B. In subset B (CompaFro≤0.5), the positive correlation between perceived social support and the number of steps is more significant–the daily steps of users with high perceived social support are 2161.785 more than those without high perceived social support. In contrast,

in subset A (CompaFro>0.5), this positive impact is relatively low, with only 1,591.258 steps increased. This means that when users lack a strong sense of competition (i.e., low possibility of comparison with top-ranked friends), perceived social support can better motivate exercise behavior. When the user has a strong sense of competition (i.e., high possibility of comparison with top-ranked friends), the incentive effect would be correspondingly reduced. At the same time, when the user has non-high perceived social support, the users in subset A walk more than those in subset B. This means that when users lack social support, a strong sense of competition can better motivate exercise behavior. When the user with high perceived social support, the difference in steps between subset A and subset B is very small. This means that when users can truly feel the social support from others and are willing to provide social support, the difference in steps caused by the sense of competition is very small.

The decision tree depicted in Fig. 8 also shows the proportional distribution of different perceived social support and its average daily steps on each terminal node. In some cases, there is not a perfectly positive correlation between the degree of perceived social support and the number of average daily steps. Decision trees can detect different effects (heterogeneous effects) and show them intuitively. For example, when the value of CompaFro is more than 0.5 (node 9), the perceived social support is positively correlated with their average daily. The characteristic of the users in node 5 is that they are more likely to view the leaderboard every day, but they hardly compare with their top-ranked friends. Besides, in node 5, the average daily steps of users in different degrees of perceived social support are generally higher than those of other nodes. In addition, there is no positive correlation between perceived social support and the number of average daily steps (the same as node 8), it is difficult to detect such nuances in common linear regression models. Comparing users in node 7 with users in node 8, the users with a higher score of exercise attitude (node 8) have more average daily steps than those with a lower score of exercise attitude (node 7). That is, users with a more active attitude are more likely to walk more. In addition, in node 8, users' attitude towards exercise is positive, and they will consciously exercise even if the perceived social support is low. By comparing the samples of node 4 and node 5, it can be seen that users with a high probability of viewing the leaderboard every day (CheckRank) are more likely to walk than those with a low probability of viewing the leaderboard, indicating that there is an incentive relationship between the leaderboard and users' exercise behavior.

## VI. DISCUSSION

Research into exercise activity correlates is an evolving field showing that the impact factor of exercise activity is complex and varies [42]. In the process of sharing knowledge in social networks, users' behaviors are often affected by the people

and society they are in contact with, and these behaviors are contagious. Exercise behavior in social networks also with the feature of contagious [13]. The effects of user's exercise behaviors interacting and promoting each other in social networks are one of the hot research topics in the field of electronic health service management. Social support is an important factor affecting user behavior and an important system feature of WeRun. On the basis of the previous research background and analysis, this paper discusses the important factors affecting exercise behavior and the relationship between perceived social support and exercise behavior. The study mainly focuses on heterogeneous effects in interventions. Finally, based on the research results, we provide useful persuasion strategies for motivating users' exercise behavior.

### A. STUDY RESULTS

In this study, personal characteristics and daily steps data of users are obtained through questionnaire survey, and a decision tree model was constructed to analyze the relationship between perceived social support and daily steps of users. The results showed that:

(1) The possibility of comparison with top-ranked friends (CompaFro) has the greatest effect on perceived social support. Moreover, most users who like to compare their steps with friends would have higher perceived social support (in most cases, they give thumbs-up to their friends and care about the impact of thumbs-up from others). On the contrary, most users with low values of CompaFro and CheckRank would have lower perceived social support.

(2) In full sample, users with high perceived social support generally walk more every day. In the sub-sample, when the user lacks a strong sense of competition (i.e., the value of CompaFro is low), perceived social support can better motivate exercise behavior. When the user has a strong sense of competition, the incentive effect would be correspondingly reduced. Moreover, it also shows the relationship between CompaFro and perceived social support, and their interaction on exercise behavior. The user's sense of social comparison and competition stimulate his(her) exercise behavior even when the user's perceived social support is low. Overall, the social support feature of persuasion technology has a certain incentive effect on the user's exercise behavior.

(3) In some cases, the degree of perceived social support is not perfectly positively correlated with the number of steps. Users with high perceived social support do not always walk more than those with non-high perceived social support. Decision trees can detect different effects (heterogeneous effects) between perceived social support and average daily steps and present them intuitively. Most users with a positive exercise attitude would consciously exercise even if their perceived social support is low. The leaderboard has a certain incentive effect on the user's exercise behavior, users who view the leaderboard more were more likely to walk more.

### B. THEORETICAL CONTRIBUTIONS

From a theoretical perspective, this study expands the research on the influence mechanism of user's exercise behavior in social networks, and further explores its heterogeneous effects. Social networks allow individuals to share knowledge better, and Internet-based WeRun provides a way for them to do so. Individuals in social networks do not exist independently, and their behaviors are often influenced by the people and environment they come into contact with. That is to say, the existence of social networks will affect the individuals of the network in many aspects [7], [8]. This study suggests that the existence of the leaderboard can further promote users' pursuit of personal health by promoting comparison among users and enhancing their sense of competition, and at the same time changes the behavior of personal health management from self-monitoring to a competitive behavior of group participation. The existence of competition relationships will affect the subsequent behavior of users [15], [16]. Increasing the frequency of users viewing the leaderboard can strengthen the competitive relationship among users. In social networks, users' online participation can promote information exchange, image display and emotional support. Emotional support is more effective than information exchange in increasing users' continued engagement and concern with the online health community. Giving thumbs-up this kind of social support behavior will have a positive short-term impact on the user's exercise behavior [30], and the results of this study indicate that social support can compensate for the lack of motivation caused by the lack of competitive consciousness. WeRun's design, based on persuasion technology, allows people to pay more attention to their exercises and encourages people to act voluntarily. Finally, this study also extends the research on the feature of social support's applicability in real-life design and use.

### C. MANAGERIAL IMPLICATIONS

This study also bears important managerial implications.

(1) Persuasion technology can stimulate users' exercise behavior through certain strategies, helping them to conduct more targeted self-health management. For example, the possibility of comparing with friends and checking leaderboard has a great impact on perceived social support, so the system can increase the push frequency of the leaderboard to increase the possibility of users' checking.

(2) The decision tree detects the heterogeneity of intervention effect, which indicates that there are individual differences in the user's exercise behavior. The e-health service platform should design differentiated services according to the characteristics of users to encourage them to exercise more. For example, real-time information interaction in social networks allows users to get feedback to review and motivate themselves during the exercise. The service provider can give the top-ranked users in the leaderboard some virtual rewards to motivate their exercise behavior, such as virtual medals

or points. In addition, achievement rewards need to be set according to the user's characteristic differences.

(3) Establish an effective incentive mechanism, improve social support and participation mechanism, and highlight the incentive role of friends. For instance, add the function of group competition, because peer cooperation and encouragement can motivate lazier users to walk more [53]. At the same time, by expanding and improving the information sharing and social functions of e-health platform (such as likes, comments and adoption), users can feel the impact of social network.

(4) Internet-based social support is closely related to the user's exercise behavior. Policymakers should make more targeted incentive policies according to the social network effect of the Internet, persuasion technology and regular pattern of user's exercise behavior.

### D. LIMITATIONS

The data-driven data mining method is helpful for decision-makers to formulate more accurate strategies. However, this study still has some limitations, which need to be solved in the future. Firstly, the sample is limited and static, which can not determine the overall social phenomena and the changes of users' future behavior. In addition, the data come from the users' self-report, which is greatly influenced by the user's subjectivity. There may be some subjective bias, so the follow-up research can use the real-time data of the user's walking steps for analysis. Thirdly, the breadth and depth of the factors affecting users' perceived social support are insufficient. If the information of users' physical and psychological qualities, living environment [54], friends network structure [55] and other aspects are considered in data mining analysis, more accurate results will be obtained.

### VII. CONCLUSION

The decision tree algorithm used in this study is a data-driven algorithm, which automatically identifies variables in the process of generating. The results obtained are easy to understand and use, and can be well applied to high-dimensional data. Besides, the decision tree algorithm does not need to make parameter assumptions about the relationship between results and input variables, and it not only has strong adaptability, but also gets relatively robust results. It can play an important role in the research of design, including random large sample experiments [56], [57]. At the same time, decision tree can capture heterogeneous effects, which are more likely to exist in big data. The model constructed in this study captures the heterogeneous effect between perceived social support and exercise behavior. It also clarifies the factors that affect exercise behavior and the interrelation between social support and social comparison. The tree can accurately identify the differences of individual characteristics, and better evaluate the intervention effect. The results of the tree show the incentives relationship between perceived social support and exercise behavior. Moreover, the "leaderboard"

based on persuasion technology in WeRun is an extremely effective way to help users form healthy behavior. In a health management platform with a social attribute, leaderboards and providing social support are good ways to motivate users to exercise.

### REFERENCES

[1] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," *Bus. Horizons*, vol. 53, no. 1, pp. 59–68, Jan. 2010, doi: 10.1016/j.bushor.2009.09.003.

[2] L. Xia and R. Suri, "Trading effort for money: Consumers' cocreation motivation and the pricing of service options," *J. Service Res.*, vol. 17, no. 2, pp. 229–242, May 2014, doi: 10.1177/1094670513507338.

[3] L. Anderson and A. L. Ostrom, "Transformative service research: Advancing our knowledge about service and well-being," *J. Service Res.*, vol. 18, no. 3, pp. 243–249, Aug. 2015, doi: 10.1177/1094670515591316.

[4] Z. Wen, X. Geng, and Y. Ye, "Does the use of WeChat lead to subjective well-being?: The effect of use intensity and motivations," *Cyberpsychol., Behav., Social Netw.*, vol. 19, no. 10, pp. 587–592, Oct. 2016, doi: 10.1089/cyber.2016.0154.

[5] G. Wang, W. Zhang, and R. Zeng, "WeChat use intensity and social support: The moderating effect of motivators for WeChat use," *Comput. Hum. Behav.*, vol. 91, pp. 244–251, Feb. 2019, doi: 10.1016/j.chb.2018.10.010.

[6] M. E. J. Newman and J. Park, "Why social networks are different from other types of networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 68, no. 3, Sep. 2003, Art. no. 036122, doi: 10.1103/physreve.68.036122.

[7] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 24, pp. 8788–8790, Jun. 2014, doi: 10.1073/pnas.1320040111.

[8] N. A. Christakis and J. H. Fowler, "Social contagion theory: Examining dynamic social networks and human behavior," *Statist. Med.*, vol. 32, no. 4, pp. 556–577, Feb. 2013, doi: 10.1002/sim.5408.

[9] N. A. Christakis and J. H. Fowler, "The spread of obesity in a large social network over 32 years," *New England J. Med.*, vol. 357, no. 4, pp. 370–379, Jul. 2007, doi: 10.1056/NEJMsa066082.

[10] J. H. Fowler and N. A. Christakis, "Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the framingham heart study," *BMJ*, vol. 337, no. dec04 2, pp. a2338–a2338, Dec. 2008, doi: 10.1136/bmj.a2338.

[11] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, Feb. 2009, doi: 10.1126/science.1165821.

[12] A. Ghose and S. P. Han, "An empirical analysis of user content generation and usage behavior on the mobile Internet," *Manage. Sci.*, vol. 57, no. 9, pp. 1671–1691, Sep. 2011, doi: 10.1287/mnsc.1110.1350.

[13] S. Aral and C. Nicolaides, "Exercise contagion in a global social network," *Nature Commun.*, vol. 8, no. 1, Apr. 2017, Art. no. 14753, doi: 10.1038/ncomms14753.

[14] Y. Liu and H. Zhang, "Empirical analysis of WeChat customer relationship based on social network theory," (in Chinese), *Inf. Documentation Services*, vol. 35, no. 4, pp. 56–61, Jul. 2014. [Online]. Available: http://www.cnki.com.cn/Article/CJFDTotal-QBZL201404011.htm

[15] K. Murayama and A. J. Elliot, "The competition–performance relation: A meta-analytic review and test of the opposing processes model of competition and performance," *Psychol. Bull.*, vol. 138, no. 6, pp. 1035–1070, 2012, doi: 10.1037/a0028324.

[16] K. D. Brownell, R. Y. Cohen, A. J. Stunkard, M. R. Felix, and N. B. Cooley, "Weight loss competitions at the work site: Impact on weight, morale and cost-effectiveness.," *Amer. J. Public Health*, vol. 74, no. 11, pp. 1283–1285, Nov. 1984, doi: 10.2105/ajph.74.11.1283.

[17] S. Cohen and T. A. Wills, "Stress, social support, and the buffering hypothesis.," *Psychol. Bull.*, vol. 98, no. 2, pp. 310–357, 1985, doi: 10.1037/0033-2909.98.2.310.

[18] M. Barrera, "Distinctions between social support concepts, measures, and models," *Amer. J. Commun. Psychol.*, vol. 14, no. 4, pp. 413–445, Aug. 1986, doi: 10.1007/BF00922627.

[19] R. Schwarzer and N. Knoll, "Functional roles of social support within the stress and coping process: A theoretical and empirical overview," *Int. J. Psychol.*, vol. 42, no. 4, pp. 243–252, Aug. 2007, doi: 10.1080/00207590701396641.

[20] J. Huang, X. Wang, W. Li, and Y. An, "The relationship between conscientiousness and posttraumatic stress disorder among young chinese firefighters: The mediating effect of perceived social support," *Psychiatry Res.*, vol. 273, pp. 450–455, Mar. 2019, doi: 10.1016/j.psychres.2019.01.053.

[21] M. Ali, N. Farooq, M. A. Bhatti, and C. Kuroiwa, "Assessment of prevalence and determinants of posttraumatic stress disorder in survivors of Earth quake in Pakistan using davidson trauma scale," *J. Affect. Disorders*, vol. 136, no. 3, pp. 238–243, Feb. 2012, doi: 10.1016/j.jad.2011.12.023.

[22] Y. Neria, A. Nandi, and S. Galea, "Post-traumatic stress disorder following disasters: A systematic review," *Psychol. Med.*, vol. 38, no. 4, pp. 467–480, Apr. 2008, doi: 10.1017/S0033291707001353.

[23] P. S. Chu, D. A. Saucier, and E. Hafner, "Meta-analysis of the relationships between social support and well-being in children and adolescents," *J. Social Clin. Psychol.*, vol. 29, no. 6, pp. 624–645, Jun. 2010, doi: 10.1521/jscp.2010.29.6.624.

[24] S. Shahyad, M. Ali Besharat, M. Asadi, A. Shir Alipour, and M. Miri, "The relation of attachment and perceived social support with life satisfaction: Structural equation model," *Procedia Soc. Behav. Sci.*, vol. 15, pp. 952–956, 2011, doi: 10.1016/j.sbspro.2011.03.219.

[25] R. R. Wing and R. W. Jeffery, "Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance.," *J. Consulting Clin. Psychol.*, vol. 67, no. 1, pp. 132–138, 1999, doi: 10.1037/0022-006X.67.1.132.

[26] P. W. Ballantine and R. J. Stephenson, "Help me, I'm fat! Social support in online weight loss networks," *J. Consum. Behav.*, vol. 10, no. 6, pp. 332–337, Nov. 2011, doi: 10.1002/cb.374.

[27] P. Rackow, U. Scholz, and R. Hornung, "Received social support and exercising: An intervention study to test the enabling hypothesis," *Brit. J. Health Psychol.*, vol. 20, no. 4, pp. 763–776, Nov. 2015, doi: 10.1111/bjhp.12139.

[28] J. W. Turner, J. A. Grube, and J. Meyers, "Developing an optimal match within online communities: An exploration of CMC support communities and traditional support," *J. Commun.*, vol. 51, no. 2, pp. 231–251, Jun. 2001, doi: 10.1111/j.1460-2466.2001.tb02879.x.

[29] T. Yao, Q. Zheng, and X. Fan, "The impact of online social support on Patients' quality of life and the moderating role of social exclusion," *J. Service Res.*, vol. 18, no. 3, pp. 369–383, Aug. 2015, doi: 10.1177/1094670515583271.

[30] L. Wang, X. Guo, T. Wu, L. Lv, and Z. Zhang, "Short-term effects of social encouragement on exercise behavior: Insights from China's wanbu network," *Public Health*, vol. 148, pp. 25–29, Jul. 2017, doi: 10.1016/j.puhe.2017.03.004.

[31] M. M. CASSELL CHRISTINE JACKS, "Health communication on the Internet: An effective channel for health behavior change?" *J. Health Commun.*, vol. 3, no. 1, pp. 71–79, Jan. 1998, doi: 10.1080/108107398127517.

[32] B. J. Fogg, *Persuasive Technology? Using Computers to Change what we Think and Do*. San Francisco, CA, USA: Morgan Kaufmann, 2003.

[33] B. Fogg, "A behavior model for persuasive design," in *Proc. 4th Int. Conf. Persuasive Technol.*, Claremont, CA, USA, Apr. 2009, pp. 1–7, doi: 10.1145/1541948.1541999.

[34] M. Bang, C. Torstensson, and C. Katzeff, "The PowerHouse: A persuasive computer game designed to raise awareness of domestic energy consumption," *Proc. Int. Conf. Persuasive Technol.* Berlin, Germany: Springer, May 2006, pp. 123–132, doi: 10.1007/11755494_18.

[35] L. Bonanni, E. Arroyo, C.-H. Lee, and T. Selker, "Exploring feedback and persuasive techniques at the sink," *Interactions*, vol. 12, no. 4, p. 25–28, Jul. 2005, doi: 10.1145/1070960.1070980.

[36] B. J. Fogg and D. Iizawa, "Online persuasion in Facebook and Mixi: A cross-cultural comparison," in *Proc. Int. Conf. Persuasive Technol.* Berlin, Germany: Springer, 2008, pp. 35–46, doi: 10.1007/978-3-540-68504-3_4.

[37] R. Hurling, M. Catt, M. De Boni, B. W. Fairley, T. Hurst, P. Murray, A. Richardson, and J. S. Sodhi, "Using Internet and mobile phone technology to deliver an automated physical activity program: Randomized controlled trial," *J. Med. Internet Res.*, vol. 9, no. 2, p. e7, Apr. 2007, doi: 10.2196/jmir.9.2.e7.

[38] K. Patrick, S. J. Marshall, E. P. Davila, J. K. Kolodziejczyk, J. H. Fowler, K. J. Calfas, J. S. Huang, C. L. Rock, W. G. Griswold, A. Gupta, G. Merchant, G. J. Norman, F. Raab, M. C. Donohue, B. J. Fogg, and T. N. Robinson, "Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART)," *Contemp. Clin. Trials*, vol. 37, no. 1, pp. 10–18, Jan. 2014, doi: 10.1016/j.cct.2013.11.001.

[39] R. Looije, M. A. Neerincx, and F. Cnossen, "Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors," *Int. J. Hum.-Comput. Stud.*, vol. 68, no. 6, pp. 386–397, Jun. 2010, doi: 10.1016/j.ijhcs.2009.08.007.

[40] H. Oinas-Kukkonen and M. Harjumaa, "A systematic framework for designing and evaluating persuasive systems," in *Proc. Int. Conf. Persuasive Technol.* Berlin, Germany: Springer, Jun. 2008, pp. 164–176, doi: 10.1007/978-3-540-68504-3_15.

[41] S. G. Trost, "Correlates of adults' participation in physical activity: Review and update," *Med. Sci. Sports Exercise*, vol. 34, no. 12, pp. 1996–2001, 2002, doi: 10.1097/00005768-200212000-00020.

[42] A. E. Bauman, "Correlates of physical activity: Why are some people physically active and others not?" *The Lancet*, vol. 380, no. 9838, pp. 258–271, Jul. 2012, doi: 10.1016/S0140-6736(12)60735-1.

[43] R. J. Mao, "The establishment and test of nine factors model of exercise attitude-behavior for young students," (in Chinese), M.S. thesis, Beijing Sport Univ., Beijing, China, 2003. [Online]. Available: http://cdmd.cnki.com.cn/Article/CDMD-10043-2003107433.htm

[44] J. Shakespeare-Finch and P. L. Obst, "The development of the 2-Way social support scale: A measure of giving and receiving emotional and instrumental support," *J. Personality Assessment*, vol. 93, no. 5, pp. 483–490, Sep. 2011, doi: 10.1080/00223891.2011.594124.

[45] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986, doi: 10.1038/323533a0.

[46] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, 2000, doi: 10.1016/S0167-7012(00)00201-3.

[47] G. Parker, S. McCraw, and D. Hadzi-Pavlovic, "The utility of a classificatory decision tree approach to assist clinical differentiation of melancholic and non-melancholic depression," *J. Affect. Disorders*, vol. 180, pp. 148–153, Jul. 2015, doi: 10.1016/j.jad.2015.03.057.

[48] S. M. Saghebian, M. T. Sattari, R. Mirabbasi, and M. Pal, "Ground water quality classification by decision tree method in ardebil region, iran," *Arabian J. Geosci.*, vol. 7, no. 11, pp. 4767–4777, Nov. 2014, doi: 10.1007/s12517-013-1042-y.

[49] N. Archak, A. Ghose, and P. G. Ipeirotis, "Deriving the pricing power of product features by mining consumer reviews," *Manage. Sci.*, vol. 57, no. 8, pp. 1485–1509, Aug. 2011, doi: 10.1287/mnsc.1110.1370.

[50] J. R. Quinlan, *C4.5 Programs for Machine Learning*. San Mateo, CA, USA: Morgan Kaufmann, 1993. [Online]. Available: http://vlebooks.com/vleweb/product/openreader?id=none&isbn=9780080500584&uid=none

[51] G. Shmueli, N. R. Patel, and P. C. Bruce, *Data Mining for Business Analytics: Concepts, Techniques and Applications in XLMiner*, 3rd ed. Hoboken, NJ, USA: Wiley, 2016.

[52] I. Yahav, G. Shmueli, and D. Mani, "A Tree-Based Approach for Addressing Self-Selection in Impact Studies with Big Data," *Mis Quart.*, vol. 40, no. 4, pp. 819–848, 2013, doi: 10.2139/ssrn.2276770.

[53] T. Toscos, A. Faber, S. An, and M. P. Gandhi, "Chick clique: Persuasive technology to motivate teenage girls to exercise," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, 2006, pp. 1873–1878, doi: 10.1145/1125451.1125805.

[54] T. Althoff, R. Sosiä, J. L. Hicks, A. C. King, S. L. Delp, and J. Leskovec, "Large-scale physical activity data reveal worldwide activity inequality," *Nature*, vol. 547, no. 7663, pp. 336–339, Jul. 2017, doi: 10.1038/nature23018.

[55] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, Sep. 2010, doi: 10.1126/science.1185231.

[56] S. Aral and D. Walker, "Creating social contagion through viral product design: A randomized trial of peer influence in networks," *Manage. Sci.*, vol. 57, no. 9, pp. 1623–1639, Sep. 2011, doi: 10.1287/mnsc.1110.1421.

[57] R. Bapna and A. Umyarov, "Do your online friends make you pay? A randomized field experiment on peer influence in online social networks," *Manage. Sci.*, vol. 61, no. 8, pp. 1902–1920, Apr. 2015, doi: 10.1287/mnsc.2014.2081 .

**ENYU ZENG** received the B.E. degree in computer science and technology from the China University of Petroleum, Qingdao, Shandong, China, in 2016. She is currently pursuing the master's degree in management with Xiangtan University, Xiangtan, Hunan, China. Her research interests include E-health and data mining.

● ● ●

**LICAI LEI** received the B.S. degree in electronic commerce from the University of Electronic Science and Technology, Chengdu, Sichuan, China, in 2007, and the Ph.D. degree in management engineering from Nanjing University, Nanjing, Jiangsu, China, in 2012. Since 2012, she has been an Assistant Professor with the School of Business, Xiangtan University, China. Her research interests include E-commerce operation mode, behavioral decision theory, and its application.