SenseFB: Linking Facebook Usage to Mental Health Using Mobile Sensing Data

ABSTRACT
Many of the mental problems like depression and stress remain hidden among college students. Although social media usage has been widely adopted in human mental well-being and personality research, no work has combined online and offline behaviors collectively to understand mental health. The SenseFB links Facebook usage with the activity and mental well-being data of 63 students collected by continuous smartphone and wristband sensing across a 10-week term at XXX College. We first analyze the general patterns of Facebook activities within a XXX term. We then examine different Facebook behaviors under various stress levels using the textual features of messages. Then, we observe a number of Facebook variables, especially the number of interests on a user's profile, significantly correlate with his/her mental well-being scales. Finally, we discuss how a combination of Facebook variables and automatic sensing data from smartphones and wristbands contribute to a better estimation in some aspects of mental wellness.

ACM Classification Keywords
H.4. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Mental health is an important issue among college students. However, much of the depression and stress remain hidden. Although social media has been shown to link with a user's personality or mental health [17, 4, 36, 50], most analysis focuses on the online social media itself, and overlook users' activities in real life – i.e., their offline life. We ask: what is the connection between students' online social media interaction and their offline behaviors around campus? How do students' online behaviors change in different stages in an academic term? If we can track the fluctuations of stress level during the term, can we find any difference in Facebook usage under different levels of mental states such as stress? How predictive it is to use Facebook features to estimate mental well-being states? What are the most important features in this process? Finally, if we combine the features from sensing data, can we improve the accuracy in predicting mental health scores? To find the answers to these questions, we look into the Facebook and mental health issues in a more detailed and comprehensive way by taking inferred behavioral data and ecological momentary assessment [53] (EMA) responses from mobile sensing into account at the same time.

Social networking services like Facebook have become part of the fabric of everyday life around the world. Furthermore, the easy access to social networking, while intertwined with the smartphone is something that made even more of an impression. Provided such easy and immediate approaches for individuals to present themselves to the world, their usage of social media is very likely to reflect the changes in their lives, emotion, mental status and daily activities. However, there is little work combining social media variables and daily behaviors. We consider it is because of two main difficulties. Firstly, to dig into people's daily social media usage, we need a not-so-big but deep and complete Facebook dataset which consists of not only the public profiles but also friend lists, photos, instant messages and the whole timelines. Such a dataset differs from the emphasis on the amount of samples in other studies, and may require full authentications from the users. Secondly, the traditional fully questionnaire-based approaches suffer from problems in tracking participants' daily activities and mental health status. They are more similar to single snapshots rather than a trustable longitudinal record. Thus, we need a novel system which can automatically and continuously infer human behavior.

XXX College is a small self-contained campus. Students here study, live and socialize in a relatively isolated community. We believe college students at XXX are very suitable for this study. On one hand, social networking is crucial for the academic community, as it can exert considerable impact on student motivation, affective learning, and classroom climate [34]. On the other hand, a previous work, StudentLife [61], has paved the way for using students' smartphones to collect sensor data unobtrusively over time.

We get detailed Facebook data from 63 participants. Upon user's agreement, we obtain the Facebook data directly from the copy downloaded by the user himself/herself from Facebook. The user can delete any sensitive information he/she doesn't like to share from the copy. We leverage sensing data derived from our recent CampusLife project, which consists of automatically collected and inferred human behaviors from both smartphones and wristbands.

Sensing data expands the dimensions of information we can gather from social media. For instance, normally we can only know the time, content, sender and receiver from a Facebook message. But simultaneously, the accelerometer can detect physical activity, the GPS can tell the location, the microphone can detect the exposure to human speech, and the EMA can tell the mental states before or after a few hours.

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:
• ACM copyright: ACM holds the copyright on the work. This is the historical approach.
• License: The author(s) retain copyright, but ACM receives an exclusive publication license.
• Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.
This text field is large enough to hold the appropriate release statement assuming it is single spaced.
Every submission will be assigned their own unique DOI string to be included here.

The main contributions of this paper are:

- To the best of our knowledge, we are the first to link Facebook usage with the activity and mental health data collected by continuous smartphone sensing, and collectively use them to estimate mental status.
- We observe trends in the Facebook usage in a term. Some of the behaviors show notable patterns, just like the *XXX term lifecycle* we have seen in the prior StudentLife study. It helps to understand how active students are in different categories of social networking behaviors over an academic term.
- We analyze the linguistic styles of messages under different stress. We find that the emotional tone is more positive, more upbeat, with lower anxiety, sadness or hostility in low stress days. We also pick out a small number of important linguistic features which are less affected by the mental or mood changes of individuals.
- We discover a number of Facebook variables that are associated with people's mental health. Results from our study suggests that the content on Facebook profiles and their Facebook behaviors may reveal information about their mental well-being that is not stated explicitly on their pages. Specifically, the number of interests people include on their pages appears to be an excellent indicator that has statistically significant associations (positive or negative) with all the mental well-being measures (depression, stress, flourishing scale, loneliness and self-esteem).
- We select sets of essential variables to separately predict depression, stress, flourishing, loneliness and self-esteem scales. Finally, we show that the models that predict participant's stress scores achieve better performance when it combines both the Facebook and sensing data.

RELATED WORK

Social media usage among college students attracts interests from numerous researchers [30, 59, 40, 21, 64]. A one-week diary-like measure among 92 undergraduates [40] focuses on college students' social networking activities and experiences, and reveals how much, why and how they use Facebook. Lewis et al. [30] introduces the first publicly available Facebook dataset of college students (in Harvard University) to appeal to scholars interested in the relationship between virtual and real social lives.

Social media researchers take interests in understanding mental health including depression [39, 58], stress and relaxation [56, 13] via online social media usage and interactional patterns. Many other high-level features can be derived from Facebook profiles and manipulations. For example, number of friends, groups, likes, photos, status updates, tags are treated as Facebook profile features and associated with personality traits in [4]. Ryan and Xenos [50] extracts 13 Facebook features, namely status, wall comments, new feed, like, messages, photos, groups, games, fan pages, events, note and chat. A principal component factor analysis with Varimax rotation [23] finds that these 13 features load on four factors: *active social contributions*, *passive engagement*, *news and information* and *real-time social interaction*. [17] involves completeness of profile items as features, i.e. whether or not the user lists the information (e.g. relationship, family, religion) on his page. Current study also explains how personality is manifested through these various features [3, 17, 4, 47]. [36] discovers that people with higher conscientiousness make significantly fewer wall postings. [4] shows the number of Facebook "likes" positively linked with openness, extroversion and neuroticism but negatively associated with conscientiousness and agreeableness.

Besides the features mentioned above, researchers also tend to involve linguistic measurements derived from sentiment analysis and word count to deduce affective states from social media text [17, 28, 13, 43, 28, 17]. Golbeck et al. [17] find a number of weak correlations between linguistic features extracted from users' profiles and their personality scores. Although his work exploratorily applies linguistic analysis methods on Facebook profiles, the persuasiveness of his conclusion is weaken by his rather small (an average of 42.6 words per person) text samples. Kumar et al. [28] analyze linguistic measures of behavior obtained from posts shared on the "SuicideWatch" forum hosted on the popular social media Reddit, and observe that contents following celebrity suicides imply greater negativity, self-attentional focus and less social concerns. Choudhury et al. [13] find the third pronoun uses in Twitter posts achieve high performance in predicting significant emotional and behavioral changes in new mothers postpartum, with an accuracy ranging from 81% to 84%. The all use the Linguistic Inquiry and Word count (LIWC) [43] software to analyze the text and generate linguistic features. The wildly used linguistic attributes in extant literature include affect and emotional expression, cognition, perception, temporal references, social/personal concerns [28, 17], and interpersonal awareness and focus [9, 14],

In recent years, mobile sensing has become more and more popular to infer human behavioral health [2, 46, 10, 15, 61]. Sandstrom et al. [51] indicate that the value of smartphones in research contexts goes beyond self-report assessments. Typically, people have noticed the opportunities for smartphones in clinical care. Ben-Zeev et al. [6] use mobile phones to collect behavioral data and ascertain the location, activity and exposure to conversation of the patients with schizophrenia. Smartphone can also play an essential role in accessing human mental health though embedded EMA [61, 48]. In [5], participants complete daily ratings of depression (Patient Health Questionnaire-9 [32]), stress (Perceived Stress Scale [20]), and loneliness (UCLA Loneliness Scale [49]).

There is little work combining social network analysis with mobile sensing based technology to understand mental well-being. However, we see potential in this research. We hypothesize that sensing data that contains detailed and dynamic offline information is a perfect supplement for the online features, which helps to understand mental health status.

DATA COLLECTION

We collect a 10-week behavioral dataset from 84 participants during the winter (38 students) and the spring (46 students) term in 2016. There are in total 72 undergraduates and 12

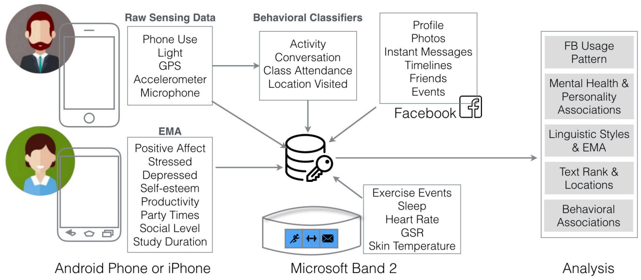


Figure 1: Sensing and analytics system architecture

graduate students. In terms of gender, 41 participants are female and 43 are male. Figure 1 describes the sensing and analytics system architecture. Briefly speaking, our study obtains data from smartphone, waistband, surveys and Facebook. This system provides the basis for us to undertake analysis and evaluate our prediction model.

Sensing System

The CampusLife app automatically infers user activity including stationary, walking, the number of independent conversations and durations, as well as ambient environment such as voice and light. The sensing system is build upon a prior work [62, 61]. CampusLife makes an improvement upon its pioneer StudentLife, providing more abundant and accurate sensing features. In the StudentLife study, the sensing system was implemented only on Android platforms. Although some students used their own primary phones to run the sensing system, over 3/4 students who only had iPhones had to carry additional Android Nexus 4s offered by us to collect sensing data. [61] shows the former group were better data sources than the later, a result that can be expected from less burden of carrying two phones during the study. To solve this problem, an iOS version is implemented in the new research, and all the sensing data are obtained from the primary smartphones of the students. Mobile ecological momentary assessment (EMA) [53] component is integrated in the system to capture students' mental states across the term.

Moreover, students are also asked to wear Microsoft Band 2s. We rely on the bands to obtain users' daily events, calories spent on different activities, and sleep durations. It also brings us additional features like the index of exposure to ultraviolet light, heart rate, Galvanic skin response (GSR) and so on.

Facebook Data Collection

Among the 84 students who complete the CampusLift study, 29 in the winter term and 34 in the spring term are willing to provide their Facebook copies. The languages used in their Facebook files are as follows: 59 in English, 2 in French, 1 in Chinese and 1 in Spanish. All of these 63 students have remained their profiles and timelines in the data folders. 27 students (12 in winter + 15 in spring) grant us permission to access to their messaging data.

A wide range of Facebook variables can be extracted from the data copies we received from our participants. One problem is that not all the files are in English. If the student sets a language other than English as Facebook language preference, the downloaded data also follows that setting. We see French, Chinese and Spanish in our dataset. Thus, to make all use of our dataset, data preprocessing is of necessary.

survey	measure
eight-item patient health questionnaire (PHQ-8) [27]	depression level
PSS [11]	stress level
flourishing Scale [16]	flouring level
UCLA loneliness Scale [49]	loniness level
state self-esteem scale [19]	self-esteem

Table 1: Mental well-being surveys

Profiles include users' basic information (e.g. gender, relationship, email, family), interests and favorites (e.g. books, music, movies, teams, athletes), groups participated in, Apps associated with and so on. A profile consists of a bunch of tag names and following contents. If someone has not included some information, the tag will not appear in the html file. Thus, we firstly scan all the files written in English to get a full list of tags which could be in profile pages. Next, we find all the tag names on profiles in other languages. Finally we carefully map these tag names into their English version and build a dictionary to memorize all the mappings. After these steps, we can easily recognize any tag on profiles by looking up the dictionary. We do not translate the content after a tag, because we only care about whether or not the user has included the information, or how many items are listed. We are not really interested in the user's religion, email or family, but just curious about what fields they are willing to expose and share to the public.

Similarly, we need to do phrase converting when dealing with the timelines. Various information is mixed in the timeline, including social media activities such as adding a new friend, uploading a photo, making a post on the wall or receiving others' comments. Social media activities can be recognized by the special phrase repeated in the document. For instance "A and B are now friends" indicates a behavior of accepting a new friend. We then find its corresponding words in the timelines in other languages. "sont désormais amis" in French and "son amigas" in Spanish are equal to "are now friends" in English. Besides, natural language dates in various formatting and languages need to be structured so that we can easily search, sort, and compare them. We use dateparser package [52] in python which provides modules to easily parse localized dates in almost any string formats commonly found on web pages.

Surveys

During the entry and exit stage of the study, participants accomplish a series of health and psychological baseline surveys as shown in Table 1. The surveys are administered using SurveyMonkey [54]. For every participant, we take average his pre- and post-study scores to get the ground truth of every mental well-being scale.

DATASET

Our dataset generally consists of 4 parts: Facebook features extracted from profile, timeline, messaging and photo album data; behavioral features inferred from automatically collected raw sensing data; responses of pre- and post-study mental well-being and personality surveys; and EMAs.

Facebook Features

We extract features from users' profiles, timelines, messages and albums. In what follows, we describe each feature and the reason to choose it.

From Profile

Personal information. Multiple aspects of personal details can be found from profiles, ranging from birthday, gender, and family to experience like hometown, education, etc. Comparing to the contents in this fields, it is more important to know whether the user is willing to release the information on their profiles. So these features only have two values, namely 0 and 1. Considering the high correlations between these fields, we believe it is not necessary to take all of them into consideration. Instead, we sum up the values of these features as the "completeness" of his or her personal information. The aggregated completeness indicates user's general sharing willingness. Besides, we particularly take the "family" and "relationship" fields into account because they have strong impacts on mental health [1, 57].

Behaviors and activities. On Facebook, people are also allowed to express themselves through activities and favorites. We can get a brief impression of a person from the list of his favorite althetes, teams, books or music and other experiences or activities he may have. We select 8 features out of 23 provided by Facebook. The selected features cover (1) hobbies and favorites("interests", "music", "books", "movies"); (2) social status ("groups"); (3) mobile phone usage ("apps", "games"); and social networking history ("previous Names"). We count how many items are listed in each field, and assign the value to the feature.

From Timeline

Differing from a relatively static image we can get from a profile, a timeline tells longitudinal stories. We track the following behaviors in our study terms: (1) adding a new friend, from a sentence like "A and B are now friends"; (2) uploading a new photo, recognized from phrases such as "added a new photo to the album" or "A with B" (appearing when faces are tagged in the uploaded pictures); (3) giving likes to others' contents, in words "A likes something", (4) listening to a song on Spotify (a famous music streaming service), indicated by "listened to song_name on Spotify"; (5) attaching a post on the wall. Each behavior is combined with a timestamp. A XXX term has ten weeks, namely 70 days. For every study subject in each day, we calculate the frequency he/she takes each type of activity. These features are used to seek the pattern of Facebook behaviors during an academical term.

From Messaging

Messaging activities. 27 out of 63 students give us access to their messages. We count the number of messages in each day of the study term (winter or spring depending on the participant), just like what we do to process the timeline. Compared with the activities disclosed by timelines, messaging is far more a frequent behavior on Facebook among students.

category		feature
summary language variables		analytical thinking, clout, authenticity, emotional tone, word per sentence (wps)
basic linguistic style	functional words	1st person singular,1st person plural, 2nd person, 3rd person singular, 3rd person plural, impersonal pronuns, articles, prepositions, auxiliary verbs,adverbs, verbs, adjectives, comparisons, interrogatives, numbers
	time orientations	past orientation, present orientation, future orientation
cognitive measures	cognitive process	insight, causation, discrepancy, tentative, certainty, differentiation
	perception processes	see, hear, feel
concerns	personal concerns	work, leisure, home, money, religion, death
	social concerns	family, friends, female references, male references
	biological concern	body, health, sexual, ingestion
language formality	informal speech	swear words, netspeak words
	punctuation	periods, commas, colons, semicolons, question marks, exclamation marks, dashes, quotation marks, apostrophes, parentheses, other irregular marks

Table 2: Linguistic features

Linguistic features. We define five categories of, in total 59 linguistic and psychological features in our research: (1) **summary language variables**, (2) **basic linguistic style**, (3) **cognitive measures**, (4) **concerns**, (5) **language formality**. The definitions of features are to a large extent inspired by an existing piece of literature that examine changes in suicide content in social media following celebrity suicides [28]. First, we derive 4 100-point-scale **summary variables** for narrative evaluation: *analytical thinking* [42], *clout* [22], *authenticity* [37], and *emotional tone* [12]. Besides, *Word per sentence* is also counted as a useful measurement. These summary variables help to understand the overall direction of messages. Then, we consider **basic linguistic style** from two aspects: percent of functional words, and time orientations. We choose them because the words structure and tense are an expression of our personality and inner thoughts [44]. Next, we export the **cognitive features**: cognitive processes and perception processes, which directly reflect human perception and thinking. We also pay attention to 3 **psychological concern** categories: personal concerns, social concerns, and biological concern. We believe these highly context-related features record what a participant cares about, and may have some changes in different mental states. At last, we estimate language formality including informal speech and all punctuation. Table 2 shows all the detailed variables in each category, and we use a word-based computerized text analysis software LIWC2015 [41, 55] to extract these linguistic measures.

From Photos

45 among the 63 participants allow us to draw information from their photos. In total, we have on aggregate 48011 pictures. We are interested in the total number of photos each user has uploaded to Facebook, as well as the composition of these photos - more specifically, how many human faces are in the photos. We perform viola-jones face detection [60, 31] using Haar Cascades in OpenCV [7], a famous real-time computer vision library widely used in academic and industry.

If the total number of pictures is too small, the ratio of every category will not be reliable. Thus, we cleaned those data from participants with less than 100 photos before face detection. Although the number of uploaded pictures is often examined in social networking related research, little work has looked into the detailed patterns in these photos. We make an exploratory hypothesis that the ratio of faces in photos might have some relations with human activities and mental health. After all, we consider four features in this category: photo count, the ratio of individual photos (one-face ratio), the ratio of group photos (multi-face ratio) and the percentage of landscape or stills photos (no-face ratio).

Behavioral Sensing Features

Mobile Phone

A number of behavioral sensing features from the raw sensor data and behaviral inferences are collected by the CampusLife app. The features we use in this study include participants' physical movement (sill or walk), ambient acoustic environment (which reflects users' preference of quiet isolated places against noisy busy places), the number and duration of exposure to conversations, and the number of places visited as well as the total distance travelled per day. All of the features above have been shown associated with mental health and wellness in previous studies [62, 29, 8].

Microsoft Band 2

From waistbands, we collected the total calories and time consumed on all the activities (running, exercise, biking, etc.), sleep duration and quality (measured by the number of wake-ups), UV exposure time and the changes in physiological features such as Galvanic skin response (GSR), heart rate and skin temperature.

Survey Instrument Data

Table 1 lists the set of surveys we use to measure behavioral and mental well-being in pre and post stages of study. The eight-item Patient Health Questionnaire depression scale (PHQ-8) [27, 25] is established as a valid diagnostic and severity measure for depressive disorders. Usually, a sum score ≥ 10 represents clinically significant depression [25]. The score of each item reveals how a subject is bothered by a depression-related problem, ranging from 0 (not at all) to 3 (nearly every day). Table 3 shows the number of students that fall into different severe levels of depression. Most of the students experience minimal or mild depression at the beginning of term. However, 10 students suffer from moderate depression and 6 students are considerably depressed in the pre-study measure. That is to say, nearly 26.3% students (with score ≥ 10) are potentially difficult and unpleasant in their lives. We have fewer people finishing the post survey, and approximately 20% students are still with clinically significant depression at the end of term. The Perceived Stress Scale (PSS) [11] is designed to measure the degree to which situations in one's life are considered as stressful. There are 10 items rated on a 5-point scale describing how often a participant thoughts or felts a stressed way, ranging from never (0) to almost always (4). The ratings are summed between 0 and 40, with the positively worded items are reverse scored. A higher score indicates a greater degree of perceived stress. We also harness the Flourishing Scale [16], a brief 8-item summary measure to appraise self-perceived success in important areas like relationships, self-esteem, purpose, and optimism. The possible range of scores is from 8 to 56. A high score represents a person with many psychological resources and strengths. UCLA loneliness scale [49] is a 20-item scale that measures a participant's subjective feelings of loneliness and social isolation, sum to a score between 20 (non-lonely direction) and 80 (lonely direction). Finally, the State self-esteem scale (SSES) [19] is a 20-item scale designed to measure one's self-esteem. The 20 5-point scale items are subdivided into 3 components: performance self-esteem, social self-esteem, and appearance self-esteem. Self-esteem is scored between 20 (lowest) to 100 (highest). Table 4 shows the mean score and standard deviation of each pre- or post-study survey for all the participants.

depression severity	*none - minimal*		*moderate*		*severe*
Score	0-4	5-9	10-14	15-19	20-24
number of students (pre)	24	21	10	4	2
number of students (post)	19	18	5	4	0

Table 3: PHQ-8 depression scale interpretations and outcomes

survey	pre-study			post-study		
outcomes	*Participants*	*mean*	*std*	*Participants*	*mean*	*std*
depression	57/63	6.6	5.3	46/63	6.2	4.3
stress	61/63	20.7	2.5	46/63	21.0	2.7
flourishing	61/63	42.3	8.8	46/63	41.2	7.6
loneliness	61/63	44.6	11.8	46/63	48.1	9.0
self-esteem	61/63	45.3	7.8	46/63	45.6	6.9

Table 4: Statistics of mental well-being surveys

EMA Data

We integrate ecological momentary assessment (EMA) [53] in our app to capture mental health and psychological states. The visual interface of EMA and its detailed implementation can be reached in [61]. EMA component schedules both daily and weekly short surveys at some fixed or random points. Users' responses to EMA surveys are synchronized with our cloud server whenever the smartphones are being charged with WiFi Internet access. EMA in this study is composed of:

- (Daily) Photographic affect meter (PAM) [45] to instantaneously mood of participants. PAM asks the user to select one of 16 pictures which can best reflect his feeling, thus to measure his positive affect at that moment.

- (Daily) Stress EMA includes a simple stress level survey asking to what degree the user is stressed. It comes up simultaneously with a photographic stress meter [18] interface which is similar as PAM but is used to measure stress.
- (Weekly) SSE-3, a simple version (4-item) of SSES to quick estimate people's self-esteem.
- (Weekly) PHQ-4 [26], an ultra-brief screening scale for anxiety and depression.
- (Weekly) Other short surveys including study duration over the week, social level, productivity and party times.

ANALYSIS AND RESULTS

In this section, we start from general patterns of students' social networking behavior. Then we identify the difference in Facebook usage under different levels of mental states. After that, we discuss the relationship between Facebook behaviors and metal well-being. At last, we show how a combination of Facebook features and sensing data contributes to a better estimation of some aspect of mental health.

Overview of Facebook Usage

We discuss students' Facebook usage patterns from two aspects: timing and location. Firstly, we analyze the XXX(College name) Facebook term lifecycle using both number of messages sent each day and activities extracted from timelines. Then, for the location's perspective, we inspect the linguistic styles and keywords from messages in different places.

Facebook Term Lifecycle

Table 5 summarizes the average frequencies of Facebook usage in discovering the Facebook term cycle. For every participant, we start with counting the number of messages, new friends accepted, songs listened, post, photos, and likes on Facebook on each day. We then change the number to ratio dividing by his sum of a activity during the term. Finally, for each social networking activity, an daily average ratio among all valid students is calculated. In this process, I exclude the participant barely take an activity in study term, because their ratios on someday will be especially high and should be viewed as extreme values. As for messages, new friends and songs, I take average among those who have usages in over 5 days in a term. For post, photos and likes, because the averages days of using posts and likes are quite small, I choose individuals with more than 3 usage days. We define these participants as *valid subjects*.

Figure 2(a)(b) shows the average daily Facebook behaviors over the term for all valid subjects. The behaviors in 2(a) illustrate some common trends. They first experience some fluctuation in the beginning weeks, followed by a decline during the mid-term. The usage recovers right after the mid-term, and again slip down in final periods. Messages and posts share similar trends, may be they are all text-related behaviors. However, the behaviors in 2(b) do not follow the trend in 2(a). We speculate students apparently upload more photos in the second half of term. The number of new friends nearly keeps stable and the likes slightly decreases with fluctuations. The reason that trends in 2(a) differs from 2(b) maybe lie in the difference of these behaviors. We can imagine that photos may highly depend on interesting events on campus. And comparing to behaviors in 2(a) (messages, songs and post) which are users' active choices, new friends and likes are more similar to reactions to others' requests or sharing.

activity	*active people in study term*	**active days** *mean*	*std*
messaging	25/27	38.48	24.80
new friends	56/63	15.89	9.53
songs	14/63	18.86	19.62
posts	45/63	4.51	3.31
photos	11/63	3	1.95
likes	44/63	3.66	4.08

Table 5: Overview of students' Facebook use during one term

location	analytic thinking	clout	authenticity	emotional tone	wps
indoor	30.69	54.89	55.96	83.48	6.39
outdoor	27.39	56.01	57.40	91.42	6.07
study	28.74	55.85	57.17	76.50	6.09

Table 6: Summary language variables in different locations

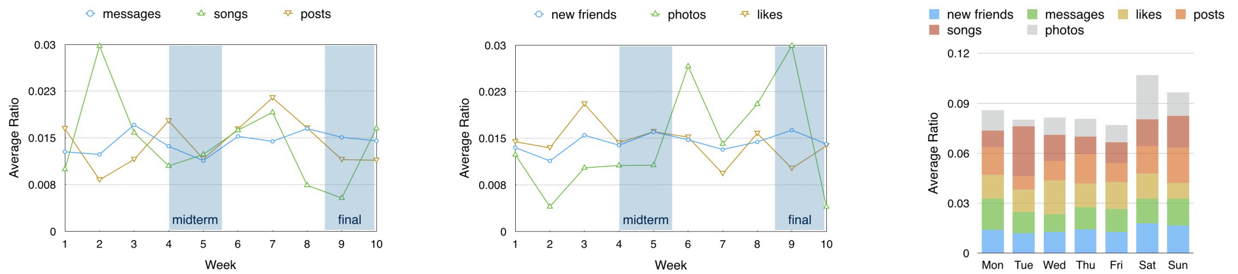
Figure 2(c) shows the average ratio of behaviors on different days in a week. Overall, students are relatively more active on weekends than on a weekday. The number of new friends remain almost unchanged in different days. They send messages mostly on Monday and infrequently on Wednesday. The number of posts peaks on Sunday and the photos uploaded tops on Saturday. On Tuesday, students are inactive in uploading photos and posts but are quite engaged in listening to songs.

Words and locations

It is interesting to combine messages with our sensing data. Our continuous sensing data attach messages with extra features such as the daily ground truth of depression, stress, exposure to conversations, activities, and locations. Here we focus on the locations. We find students send messages mostly in dorms, Greek houses and study places (Figure 3). Table 6 shows the summary language variables of messages sent in different places. We divide the places into three big categories: indoor (dorms and Greek houses), outdoor (food and gym), and study places.

The messages are the most formal, logical thinking in dorms and Greek houses, even though the difference is not so notable. In addition, students send messages with higher positive emotional tone in food places and gym than in study places. The number of words per sentence is slightly higher in dorms than in other places.

We then use TextRank [35] algorithm to extract keyword from messages. TextRank is a graph-based ranking model that decides the importance of a vertex by taking into account global information recursively computed from the entire graph, rather than relying only on local vertex-specific information



(a) Term lifecycle: messages, songs and posts (b) Term lifecycle: new friends, photos and likes (c) Activities on different days in a week

Figure 2: Facebook usage across academic term

location	top-20 keywords
dorms / houses	time, people, stuff, thing, lol, work, class, week, day, house, http, today, friend, term, room, hey, lot, haha, year, night, group
food / gym	time, class, hey, something, chair, kind, year, lot, event, hope, cause, group, haha, way, tomorrow, spring, sound, job, night, hour
study places	people, time, class, lol, thing, day, today, sorry, stuff, way, tomorrow, someone, part, term, house, guy, lot, kind, http, something

Table 7: Top-20 keywords in different locations

linguistic summary varialbes	low stress (avg)	high stress (avg)	*t-test*	*p-value*
Analytic thinking	25.93	28.68	-1.016	0.323
Clout	54.31	52.86	0.449	0.659
Authenticity	63.80	62.48	0.264	0.794
Emotional tone	88.01	76.81	2.189	**0.042**
Word per sentence	11.63	10.62	1.022	0.321

Table 8: Summary language variables in low and high stress days

[35]. Table 12 enumerates the top-20 keywords from TextRank in 3 kinds of locations. Overall, students' messages cover topics about campus life, including daily plans, places, friends and classes (e.g. time, class, week, day, term, house). "Time" ranks in top 2 in all the places. However, words in dorm or study places tend to be more human concerned ("people", "friend", "someone", "guy") than in food places and gym. Schedule-oriented words like "day", "today" and "tomorrow" have higher ranks in study places than in others. Keywords in dorms and study places are quite similar, while there exists some special topics in food and workout places (e.g. "chair", "event", "group", "spring", "job").

Linguistic Features Under Different Mental States

Can we infer the fluctuations of moods through changes in Facebook usage? We find in Table 5 that comparing to other activities, messaging is most frequently used during the term and is therefore the best to be linked with everyday states. We get the ground truth of stress level from the EMA. Students choose one stress level from 1 ("Not at all") to 5 ("Extremely") in stress level EMA. We take average the scores in one day and separate the message set of one students into two categories according to his stress score on that day. If the score is less than or equal to 2, the messages are classified as low-stress set. If the score is above or equal to 4, the messages on that day are tagged as "high-stress". We perform paired samples t-test to check the differences between the two sets.

Table 8 shows the summary of language variables from students' messages, comparing the messages sent in low-stress days to those from high-stress days. As we can see, there are no significant differences in analytic thinking, clout, and authenticity in messages under different stress. It is noteworthy that the emotional tone scores in low-stress days are significantly higher than in high-stress days. A higher emotional tone score indicates a more positive, more upbeat mood, with lower anxiety, sadness or hostility. Thus we know the stress may be detected from emotional tone of Facebook messages. On average, sentences on low-stress days have about 1 more word than what on high-stress days, yet it is far from statistical significance.

Table 8 gives us another inspiration that although the emotional tone is influenced by people's stress, most of individuals' codes of language don't change so much under different environments (e.g. stress). This indicates that those features that keep high correlations in various environments might be used to identify different individuals. To identify these features, we run correlation analysis between the messages in difference stress sets. Table 9 enumerates all the 22 important features which are significantly correlated in low and high stress sets.

To get an insight into these selected features, we use t-Distributed Stochastic Neighbor Embedding (t-SNE) [33] to visualize messages from all participants (Figure 4). One point in the graph is projected from the 22-dimentional linguistic features in Table 9 extracted from a two-week message samples. So for every participant in a term, we have 5 points in the graph. People are identified using different colors. We observe data points are grouped into different clusters in Figure 4(a). If we only consider the significantly highly correlated features ($r > 0.7$, $p < 0.01$), the aggregations in a multi-dimensional space become even more obvious (Figure 4(b)). This observation gives us confidence to say some linguistic measures may serves as a fingerprint to identify an individual. On this basis,

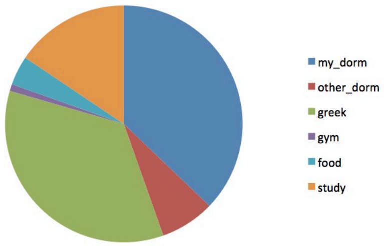
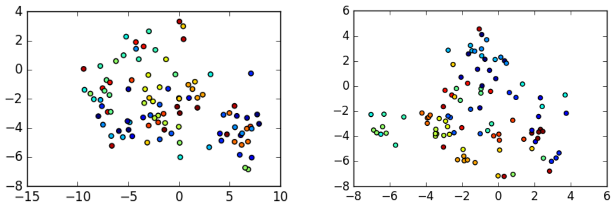


Figure 3: messages at different locations



(a) 22 selected linguistic feaures (b) 7 essential linguistic feaures

Figure 4: Feature visulization using t-SNE

category	linguistic feature
summary language variables	analytical thinking, **clout**
basic linguistic style	2nd person, prepositions, ***auxiliary verbs***, verbs, **comparisons**, **interrogatives**, future orientation
cognitive measures	insight
concerns	**money**, religion, friends, ***sexual***, ***ingestion***,
language formality	***swear words***, **netspeak words**, ***periods***, ***commas***, question marks, ***exclamation marks***, ***apostrophes***
all associations with $p < 0.05$ **bold** if $p < 0.01$ and *italic* if $r > 0.7$	

Table 9: Linguistic features that significantly correlates in different stresses

the emotional tone can indicate stress swings over his base line.

Facebook Usage, Mental Health and EMA

In what follows, we first consider the correlations between Facebook variables and mental well-being scores derived from our pre-post surveys. We also identify a number of significant correlations with EMA responses. The degree of correlation is indicated by the Pearson correlation coefficient r (-1 $\leq$ r $\leq$ 1), where -1 indicates a perfect negative linear relationship between variables, 0 indicates no linear relationship between variables, and 1 indicates a perfect positive linear relationship. A p value is the observed significance level of the test.

Correlation with Mental Health

We find a set of correlations between Facebook variables and mental well-being scores including depression, perceived stress, flourishing, loneliness and self-esteem scales (see Table 10).

PHQ-8 depression scale. We find a number of significant correlations ($p < 0.05$) between number of interests written on profiles, apps linked with Facebook, total number of photos in Facebook album and PHQ-8 depression, as shown in Table 10. We might have thought that people with a lot of interests seem unlikely to be depressed, because it is the loss of interest or pleasure in hobbies and activities often regarded as a sign of depression [38]. However, our research shows in the opposite way. We find the number of interests on profiles positively associated ($r = 0.281, p = 0.028$) with depression. One explanation is that the number of interests on Facebook is not necessarily the perfect duplication of real life. They are more likely a psychological choice - some people may have rich interests, while they do not want to list all of them on Facebook. Also, we see a positive relationship ($r = 0.252, p = 0.047$) between apps registered on Facebook platform and depression. The number of photos uploaded in term has a week positive association ($r = 0.243, p = 0.055$) with depression, while the total number in album is tested to be a significantly stronger indicator of depression ($r = 0.402, p = 0.006$). This suggests that students who prefer to share more pictures on Facebook are more likely to experience depressive symptoms. We often infer those who exposure many photos enjoying their lives but things are very likely just the opposite: they are doing so because they are feeling bad, and are anxious to receive attention from others.

Perceived stress scale (PSS). We observe a small set of correlations between Facebook behaviors and perceived stress. The willingness to put relationship (single/in love with somebody/married) information on Facebook has a provisionally significant ($r = 0.209, p = 0.099$) association with stress. Although relationship appears to be merely one's current status, the decision of whether to write this information may be psychologically meaningful [65]. Because Facebook provides a list of actions that friends have recently undertaken on everyone's homepage, others will rapidly know changes of relationship status. We hypothesize that releasing relationship status reflects a strong desire to elicit interest, contact, cognition and blessing from others. And this behavior may be a sign of stress. The number of apps weekly links with stress ($r = 0.256, p = 0.043$). We also notice that the number of interests again, has a significant positive correlation ($r = 0.256, p = 0.043$) with stress.

Flourishing scale. Releasing family information on Facebook is positively linked with flourishing, close to being statistically significant ($r = 0.241, p = 0.058$). Students who have less interests ($r = -0.390, p = 0.002$) and favorite books ($r = -0.262, p = 0.038$) listed on their pages are more flourishing. There is a positive connection between number of friends and flourishing scale ($r = 0.321, p = 0.015$). This relationship can also be observed between new friends accepted in term and flourishing, in a less significant way ($r = 0.225, p = 0.077$).

Loneliness scale. Number of interests shows a significantly positive association with loneliness ($r = 264, p = 0.037$). We see that the negative correlation between the number of new friends in term and loneliness approaches the borderline of significance ($r = -0.237, p = 0.062$). In addition, the total

mental health scale	*Facebook varialbes*	*r*	*p value*
depression	interest	0.281	0.028
	apps	0.252	0.047
	photos (term)	0.243	0.055
	photos (total)	0.402	0.006
stress	relationship	0.209	0.099
	interests	0.256	0.043
	apps	0.237	0.062
flourishing	family	0.241	0.058
	interests	-0.390	0.002
	books	-0.262	0.038
	friends	0.321	0.015
	new friend in term	0.225	0.077
loneliness	interests	0.264	0.037
	friends	-0.355	0.007
	new friend in term	-0.237	0.062
self-esteem	interests	-0.383	0.002
	music	-0.281	0.026
	books	-0.280	0.026
	movies	-0.266	0.035
	apps	-0.323	0.010

Table 10: Correlations between Facebook variables and mental health scale

Facebook variables	EMA (positively related)	EMA (negatively related)
apps	depression, stress(m)	
books	**depression**, **stress(l)**	
friends	social level	
interests	**depression**, stress(l)	productivity
photo (total)	depression	
no-face ratio	**positive affect**	stress(l), **stress(m)**
multi-face ratio	stress(sm)	**positive affect**
all assosications with $p < 0.05$ **bold** if $p < 0.01$		

Table 11: Correlations between Facebook variables and EMA

number of friends has a stronger negative and more significant association ($r = -0.355, p = 0.007$) with loneliness. It is not hard to imagine that people who have more friends have less chances to suffer from loneliness.

Self-Esteem scale. Students' interests ($r = -0.383, p = 0.002$) and favorite music ($r = -0.281, p = 0.026$), books ($r = -0.280, p = 0.026$) and movies ($r = -0.266, p = 0.035$) are found negatively related to self-esteem. Number of apps also has a negative link to self-esteem ($r = -0.323, p = 0.010$).

Correlation with EMA

Students respond to a number of EMAs which capture their momentarily psychological and behavioral status during the term. The EMAs measure positive affects (PAM), stress (stress level survey & stress meter), depression (PHQ-4), self-esteem (SSE-3), study duration, social level, productivity and party times, as is discussed in the Dataset Section. We examine correlations between Facebook behaviors and the EMA responses from students. Table 11 shows a set of significant correlations we have discovered. Note that we have two different stress EMA, so we label the stress level survey as stress(l) and the picture-presenting stress meter as stress(m). As we can see from Table 11, the number of apps, books, interests and photos in album positively related to depression scales. These evidences strengthen our findings because these Facebook features also correlate with depression from pre-post study surveys. Similarly, the positive associations between apps, books interests and stress are also incline with what we observed in the previous subsection.

The interesting thing here is we observe the percentage of group photos and no-facial (landscape/stills) photos in album can reflect people's positive affect and stress. To be specific, we find a significantly positive association between the ratio of no-face pictures and people's positive affects, as well as a strong negative link between no-face photo ratio and stress. However, the ratio of multi-face pictures acts as an opposite role. The more percentage of multi-face photos include, the more likely a student gets stressed up and less likely to have positive affects. In addition, the number of interests is also negatively related with the productivity, while the number of friends is positively linked with the social level.

Based on the analyses above, we can clearly see that the number of interests is an excellent indicator of personal mental well-being states. It has strong connections with all the mental health scale we have examined in survey and EMA. More precisely, the more number of interests on profile, the more likely a student is suffering from mental health issues like depression, stress, and loneliness. Adversely, a low number of interests on profile may indicate a good mood such as flourishing, self-esteem and positive affects.

Prediction Analysis

In this section, we discuss the performance of predicting mental health status based on 3 different feature sets: Facebook-only, sensing-only, and a combination of both Facebook and sensing features. The ground truth of individual mental health scales comes from the average of his pre- and post-study surveys, including measures for depression, stress, flouring, loneliness and self-esteem, as shown in Table 1.

Since we have far more features comparing to a not-so-large sample size, it is necessary to reduce the dimensions of feature space before prediction. For each mental health measure, we firstly use Random Lasso [63] to rank the importance of features respectively of the 3 different feature sets. Random Lasso works by resampling the train data and computing a Lasso on each resampling. After repeating the process a number of times, the selection results can be aggregated, and the features selected more often are good features. Weaker, but still relevant features will have non-zero scores, because they may be among selected features when stronger features are not in the currently subset. Irrelevant features would have scores zero, for that they would never be selected. After the Random Lasso analysis, we perform a greedy algorithm to select 5 key features in each category, under 3 principals: (1) the feature with higher score in Random Lasso results is of higher priority to be selected; (2) the feature which is considerably correlated ($r \geq 0.4, p < 0.05$) with the selected features should be removed; (3) all the features with zero scores should be discarded. Consequently, we pick the key features as described in Table 12. Note that the selected important features in combined model are not necessarily a sum of the top features in

mental wellness	feature set	top-ranking features	MAE	r	p
depression (0-24)	Facebook	family(-), photos (total)(+), interests(+), friends removed(+), likes(+)	3.178	0.513	0.002
	sensing	UV exposure minutes(+), location visited(-), gsr max(-), heart rate max(+), gsr min(+)	4.792	0.271	0.127
	FB + sensing	gsr min(+), gsr max(-), photos (total)(+), interests(+), family(-)	3.347	0.576	<0.001
stress (0-40)	Facebook	music(+), relationship(+), family(-), photos (total)(+), likes(+)	1.641	0.238	0.181
	sensing	walk duration day(-), gsr min(+), heart rate mean(+), sleep duration(-), calories all activities(+)	1.819	0.059	0.742
	FB + sensing	gsr min(+), walk duration day(-), music(+), likes(+), completeness(+)	1.287	0.600	<0.001
flourishing (8-56)	Facebook	family(+), groups(+), apps(-), posts(+), new friend (term)(+)	4.654	0.673	<0.001
	sensing	location visited(+), ambient sound variation day(+), heart rate median(-)	6.114	0.390	0.024
	FB + sensing	location visited(+), ambient sound variation day(+), family(+), friends removed(-), interests(-)	5.470	0.586	<0.001
loneliness (20-80)	Facebook	family(-), new friend (term)(-), likes(+), groups(-), previous names(+)	6.787	0.554	<0.001
	sensing	location visited(-), ambient sound variation day(-), calories all activities(-), gsr max(-), sleep wake-ups(-)	9.285	0.0370	0.835
	FB + sensing	heart rate max(-), ambient sound variation day(-), family(-), friends(-)	6.589	0.557	<0.001
self-esteem (20-100)	Facebook	photos (total)(-), family(+), friends removed(-), books(-), likes(-)	4.998	0.572	<0.001
	sensing	location visited(+), gsr min(-), heart rate median(-), sleep duration(+), calories all activities(-)	7.654	0.162	0.365
	FB + sensing	location visited(+), heart rate mean(-), interests(-)	5.444	0.443	0.009
(-): negative association, (+) positive association					

Table 12: Key features and prediction performance

separate feature spaces. The reason is that there are many highly correlated features between Facebook and sensing features, although the principles behind these connections still remain obscure. For example, we have discovered a strong positive connection between the number of movies and the ambient sound variance during 12am to 9am ($r = 0.602$, $p< 0.001$).

We apply leave-one-subject-out cross validation [24] and linear regression model to evaluate the performance of each model. In order to make the work properly, each feature is scaled with zero mean and unit variance. We see that both the Facebook-only and combined model can achieve MAE around 3.2 of the predicted depression score (Table 12), indicating that the predictions are within ± 3.2 of the ground truth. However, the predicted depression from the combined model has a stronger correlation with the ground truth with $r = 0.576$ and $p < 0.001$, which further indicates a better capture of the original distribution. As for stress, neither of the Facebook-only and sensing-only models provides acceptable prediction. However, we observe a significant improvement in both accuracy (± 1.287) and correlation ($r = 0.600$, $p < 0.001$) when using the combined model. Facebook features are powerful in flourishing prediction, which achieve strong correlates with ground truth ($r = 0.673$, $p < 0.001$) within ± 4.654 of survey scores. The prediction abilities of Facebook-only and combined features are almost equally prosperous in estimating loneliness (± 6.6) with high correlation ($r = 0.55$, $p < 0.001$). Facebook-only model takes a leading position in predicting self-esteem, with a (± 4.998) bias from the ground truth ($r = 0.572$, $p < 0.001$).

Furthermore, the positive or negative effects of every selected feature on mental well-being scores can be derived from the coefficients in regression models (Table 12). We see in general, people with higher gsr, more interests and photos on Facebook and less willingness to release family information are more likely to suffer from depression and stress. More group, posts, new friends and less apps indicate a flourishing life style. Ambient sound variation during 9am to 6pm serves may imply a distance from loneliness. People with higher self-esteem have fewer photos in Facebook album, fewer books and are less willing to delete a friend or give likes to others' contents.

Overall, we see that the Facebook-only features take advantages in predicting the scores of surveys which reflect positive sides of mental wellness, such as the flourishing scale and self-esteem. Combining with sensing features, the abilities to estimate those negative mental issues like depression, stress and loneliness are improved to varying degrees.

CONCLUSION

SenseFB is the first study to collaboratively use Facebook and sensing data to estimate mental health status. The study give answers to the questions we propose at the beginning of the paper. As the term progresses, the Facebook activities like sending messages, making posts, listening to songs, adding new friends, uploading pictures and giving likes to others shown some interesting patterns. The emotional tones in texts to some extent tell up and downs of students' moods like stress. We prove the Facebook data shows considerable promise in predicting mental health. Provided an extra dataset from mobile sensing, the accuracy and ability to capture differences in scores can be further improved in some of the mental well-being scales. Especially, the predicted stress from features extracted from combined dataset strongly correlates with the ground truth from surveys ($r = 0.6$ and $p < 0.001$) and is on average within ± 1.287 of the true score which can range from 0 to 40.

REFERENCES

1. Substance Abuse. 2010. Mental Health Services Administration.(2009). *Results from the 2008 national survey on drug use and health: National findings* (2010).
2. Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. 2011. Social fMRI: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing* 7, 6 (2011), 643–659.
3. Yair Amichai-Hamburger and Gideon Vinitzky. 2010. Social network use and personality. *Computers in human behavior* 26, 6 (2010), 1289–1295.
4. Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli, and David Stillwell. 2012. Personality and patterns of Facebook usage. In *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, 24–32.
5. Dror Ben-Zeev, Emily A Scherer, Rui Wang, Haiyi Xie, and Andrew T Campbell. 2015a. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric rehabilitation journal* 38, 3 (2015), 218.
6. Dror Ben-Zeev, Rui Wang, Saeed Abdullah, Rachel Brian, Emily A Scherer, Lisa A Mistler, Marta Hauser, John M Kane, Andrew Campbell, and Tanzeem Choudhury. 2015b. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. *Psychiatric Services* (2015).
7. G. Bradski. 2000. *Dr. Dobb's Journal of Software Tools* (2000).
8. Luca Canzian and Mirco Musolesi. 2015. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1293–1304.
9. Stevie Chancellor, Tanushree Mitra, and Munmun De Choudhury. 2016. Recovery Amid Pro-Anorexia: Analysis of Recovery in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2111–2123.
10. Tanzeem Choudhury, Gaetano Borriello, Sunny Consolvo, Dirk Haehnel, Beverly Harrison, Bruce Hemingway, Jeffrey Hightower, Karl Koscher, Anthony LaMarca, James A Landay, and others. 2008. The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Computing* 7, 2 (2008), 32–41.
11. Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A global measure of perceived stress. *Journal of health and social behavior* (1983), 385–396.
12. Michael A Cohn, Matthias R Mehl, and James W Pennebaker. 2004. Linguistic markers of psychological change surrounding September 11, 2001. *Psychological science* 15, 10 (2004), 687–693.
13. Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via Social Media.. In *ICWSM*. 2.
14. Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2098–2110.
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. Ed Diener, Derrick Wirtz, William Tov, Chu Kim-Prieto, Dong-won Choi, Shigehiro Oishi, and Robert Biswas-Diener. 2010. New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research* 97, 2 (2010), 143–156.
17. Jennifer Golbeck, Cristina Robles, and Karen Turner. 2011. Predicting personality with social media. In *CHI'11 extended abstracts on human factors in computing systems*. ACM, 253–262.
18. Sophia Haim, Rui Wang, Sarah E Lord, Lorie Loeb, Xia Zhou, and Andrew T Campbell. 2015. The mobile photographic stress meter (MPSM): a new way to measure stress using images. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 733–742.
19. Todd F Heatherton and Janet Polivy. 1991. Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social psychology* 60, 6 (1991), 895.
20. Paul L Hewitt, Gordon L Flett, and Shawn W Mosher. 1992. The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample. *Journal of Psychopathology and Behavioral Assessment* 14, 3 (1992), 247–257.
21. Reynol Junco, Greg Heiberger, and Eric Loken. 2011. The effect of Twitter on college student engagement and grades. *Journal of computer assisted learning* 27, 2 (2011), 119–132.
22. Ewa Kacewicz, James W Pennebaker, Matthew Davis, Moongee Jeon, and Arthur C Graesser. 2013. Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology* (2013), 0261927X13502654.
23. Henry F Kaiser. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 3 (1958), 187–200.
24. Ron Kohavi and others. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, Vol. 14. 1137–1145.
25. Kurt Kroenke and Robert L Spitzer. 2002. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals* 32, 9 (2002), 509–515.
26. Kurt Kroenke, Robert L Spitzer, Janet BW Williams, and Bernd Löwe. 2009a. An ultra-brief screening scale for anxiety and depression: the PHQ–4. *Psychosomatics* 50, 6 (2009), 613–621.
27. Kurt Kroenke, Tara W Strine, Robert L Spitzer, Janet BW Williams, Joyce T Berry, and Ali H Mokdad. 2009b. The PHQ-8 as a measure of current depression in the general population. *Journal of affective disorders* 114, 1 (2009), 163–173.
28. Mrinal Kumar, Mark Dredze, Glen Coppersmith, and Munmun De Choudhury. 2015. Detecting changes in suicide content manifested in social media following celebrity suicides. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*. ACM, 85–94.
29. Nicholas D Lane, Mashfiqui Mohammod, Mu Lin, Xiaochao Yang, Hong Lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. 2011. Bewell: A smartphone application to

monitor, model and promote wellbeing. In *5th international ICST conference on pervasive computing technologies for healthcare*. 23–26.

30. Kevin Lewis, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. Tastes, ties, and time: A new social network dataset using Facebook. com. *Social networks* 30, 4 (2008), 330–342.
31. Rainer Lienhart and Jochen Maydt. 2002. An extended set of haar-like features for rapid object detection. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, Vol. 1. IEEE, I–900.
32. Bernd Löwe, Jürgen Unützer, Christopher M Callahan, Anthony J Perkins, and Kurt Kroenke. 2004. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical care* 42, 12 (2004), 1194–1201.
33. Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9, Nov (2008), 2579–2605.
34. Joseph P Mazer, Richard E Murphy, and Cheri J Simonds. 2007. I'll see you on "Facebook": The effects of computer-mediated teacher self-disclosure on student motivation, affective learning, and classroom climate. *Communication Education* 56, 1 (2007), 1–17.
35. Rada Mihalcea and Paul Tarau. 2004. TextRank: Bringing order into texts. Association for Computational Linguistics.
36. Kelly Moore and James C McElroy. 2012. The influence of personality on Facebook usage, wall postings, and regret. *Computers in Human Behavior* 28, 1 (2012), 267–274.
37. Matthew L Newman, James W Pennebaker, Diane S Berry, and Jane M Richards. 2003. Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin* 29, 5 (2003), 665–675.
38. National Institute of Mental Health. 2016. Depression. (May 2016). Retrieved September 8, 2016 from https://www.nimh.nih.gov/health/topics/depression/index.shtml.
39. Minsu Park, David W McDonald, and Meeyoung Cha. 2013. Perception Differences between the Depressed and Non-Depressed Users in Twitter.. In *ICWSM*.
40. Tiffany A Pempek, Yevdokiya A Yermolayeva, and Sandra L Calvert. 2009. College students' social networking experiences on Facebook. *Journal of applied developmental psychology* 30, 3 (2009), 227–238.
41. James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. *UT Faculty/Researcher Works* (2015).
42. James W Pennebaker, Cindy K Chung, Joey Frazee, Gary M Lavergne, and David I Beaver. 2014. When small words foretell academic success: The case of college admissions essays. *PloS one* 9, 12 (2014), e115844.
43. James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates* 71 (2001), 2001.
44. James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. 2003. Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology* 54, 1 (2003), 547–577.
45. John P Pollak, Phil Adams, and Geri Gay. 2011. PAM: a photographic affect meter for frequent, in situ measurement of affect. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 725–734.
46. Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury, and Ethan Berke. 2011. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 385–394.
47. Craig Ross, Emily S Orr, Mia Sisic, Jaime M Arseneault, Mary G Simmering, and R Robert Orr. 2009. Personality and motivations associated with Facebook use. *Computers in human behavior* 25, 2 (2009), 578–586.
48. Jason D Runyan, Timothy A Steenbergh, Charles Bainbridge, Douglas A Daugherty, Lorne Oke, and Brian N Fry. 2013. A smartphone ecological momentary assessment/intervention "app" for collecting real-time data and promoting self-awareness. *PLoS One* 8, 8 (2013), e71325.
49. Daniel W Russell. 1996. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of personality assessment* 66, 1 (1996), 20–40.
50. Tracii Ryan and Sophia Xenos. 2011. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage. *Computers in Human Behavior* 27, 5 (2011), 1658–1664.
51. Gillian M Sandstrom, Neal Lathia, Cecilia Mascolo, and Peter J Rentfrow. 2016. Opportunities for Smartphones in Clinical Care: The Future of Mobile Mood Monitoring. *The Journal of clinical psychiatry* 77, 2 (2016), 135–137.
52. Scrapinghub. 2016. dateparser 0.4.0. https://github.com/scrapinghub/dateparser. (2016).
53. Saul Shiffman, Arthur A Stone, and Michael R Hufford. 2008. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* 4 (2008), 1–32.
54. SurveyMonkey. 2016. http://www.surveymonkey.com/. (2016).
55. Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology* 29, 1 (2010), 24–54.
56. Mike Thelwall. 2016. TensiStrength: Stress and relaxation magnitude detection for social media texts. *Information Processing & Management* (2016).
57. Jane Traupmann, Elaine Hatfield, and others. 1981. Love and its effect on mental and physical health. *Aging: Stability and change in the family* (1981), 253–274.
58. Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. 2015. Recognizing depression from twitter activity. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3187–3196.
59. Sebastián Valenzuela, Namsu Park, and Kerk F Kee. 2009. Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation1. *Journal of Computer-Mediated Communication* 14, 4 (2009), 875–901.
60. Paul Viola and Michael Jones. 2001. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 1. IEEE, I–511.

61. Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 3–14.
62. Rui Wang, Min S. H. Aung, Saeed Abdullah, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Martan Hauser, John Kane, Michael Merrill, Emily A. Scherer, and Vincent W. S. Tseng. 2016. CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia. (2016).
63. Sijian Wang, Bin Nan, Saharon Rosset, and Ji Zhu. 2011. Random lasso. *The annals of applied statistics* 5, 1 (2011), 468.
64. Yiran Wang, Melissa Niiya, Gloria Mark, Stephanie M Reich, and Mark Warschauer. 2015. Coming of age (digitally): An ecological view of social media use among college students. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 571–582.
65. Sean Young, Debo Dutta, and Gopal Dommety. 2009. Extrapolating psychological insights from Facebook profiles: A study of religion and relationship status. *CyberPsychology & Behavior* 12, 3 (2009), 347–350.