Social Media Discussions Predict Mental Health Consultations on College Campuses

Koustuv Saha[1*+]**, Asra Yousuf**[2]**, Ryan L. Boyd**[3–5]**, James W. Pennebaker**[6]**, and Munmun De Choudhury**[2]

[1]Microsoft Research, Montréal, QC, Canada
[2]School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, USA
[3]Department of Psychology, Lancaster University, Lancaster, UK
[4]Security Lancaster, Lancaster University, Lancaster, UK
[5]Data Science Institute, Lancaster University, Lancaster, UK
[6]Department of Psychology, University of Texas at Austin, Austin, TX, USA

[*]Corresponding Author:
Koustuv Saha
Microsoft Research Lab - Montréal
6795 Rue Marconi, Suite 400
Montréal, Québec, H2S 3J9
Canada
koustuvsaha@microsoft.com
[+]The first author did the majority of this work while at Georgia Institute of Technology

ABSTRACT

The mental health of college students is a growing concern, and gauging the mental health needs of college students is difficult to assess in real-time and in scale. While social media has shown potential as a viable "passive sensor" of mental health, the construct validity and in-practice reliability of such computational assessments remain largely unexplored. Towards this goal, we study how assessing the mental health of college students using social media data correspond with ground-truth data of on-campus mental health consultations. For a large U.S. public university, we obtained ground-truth data of on-campus mental health consultations between 2011–2016, and collected 66,000 posts from the university's Reddit community. We adopted machine learning and natural language methodologies to measure symptomatic mental health expressions of depression, anxiety, stress, suicidal ideation, and psychosis on the social media data. Seasonal auto-regressive integrated moving average (SARIMA) models of forecasting on-campus mental health consultations showed that incorporating social media data led to predictions with r=0.86 and SMAPE=13.30, outperforming models without social media data by 41%. Our language analyses revealed that social media discussions during high mental health consultations months consisted of discussions on academics and career, whereas months of low mental health consultations saliently show expressions of positive affect, collective identity, and socialization. This study reveals that social media data can improve our understanding of college students' mental health, particularly their mental health treatment needs.

Keywords: social media; language; college students; mental health; counseling services; Reddit

1 Introduction

Mental health on college campuses is a matter of growing concern as an increasing number of college students show rising levels of anxiety, depression, and suicidal ideation. According to the 2019 National College Health Assessment[1] 16.7% students felt too depressed to function in the last two weeks from when the survey was conducted while 8.6% seriously considered suicide or tried to harm themselves in the past 12 months. Another decade-spanning study found that the percentage of students diagnosed with mental illness rose up from 22% in 2007 to 36% in 2017 even though the rate of treatment increased from 19% to 34%[2]. Mental health services on college campuses, including on-campus counseling centers and psychiatric clinics, therefore continuously struggle to address the increasing demands of mental health consultations in a timely fashion[3]. A research study conducted by Penn State's Center for Collegiate Mental Health, for instance, reported a 30-40% increase in the on-campus counseling consultations between 2009–2015, despite an only 5% increase in enrollment[4]. In short, these services often lack in resources, staff, and preparedness, leading to long waiting lists and selective/infrequent consultations of many[5].

This underscores an urgent need to meet the rising demand for mental health services with adequate and accessible resources.

However, campus mental health services do not currently have adequate means to assess the evolving nature of demand or needs. While periodic surveys of students' mental health provides some barometer of mental health incidence, in terms of medication use, daily lifestyle, suicidal thoughts, depression symptoms, as well as potentially contributing academic, environmental, personal, and social factors[6], they are accurate only in snapshots, and are prone to retrospective and susceptible to biases[7]. Since it is practically and financially unsustainable to administer surveys in real-time, when administered after a while from actual incidences, such data may not capture the time-sensitive nature of mental health needs and demands. Surveys are also difficult to scale and are limited to smaller study groups. With an increasing gap in the supply of mental health resources and their growing demand, college campuses need to find alternative means to gauge and forecast the demand of counselling services in order to cater to everyone who needs them.

To overcome such limitations, researchers and practitioners have started exploring passive sources of data, which provide dense and longitudinal behavior of individuals at scale[8]. Given the ubiquity and widespread use of social media, especially among the college student demographic, social media data has also been leveraged as a "passive sensor" that can act as a complementary source of unobtrusive, real-time, and naturalistic data to infer wellbeing[9]. Social media data is low-cost, large-scale, non-intrusive to collect, and has the potential to comprehensively reveal naturalistic patterns of mood, behavior, cognition, psychological states and social milieu, both in real-time and across longitudinal time for individuals and collectives[10]. Social media language consists of an individual's personal and social discourse about day-to-day concerns, and effectively reflects their health and psychosocial wellbeing in a variety of states and contexts[11;12;13;14]. Linguistic cues and social interactions on social media platforms have therefore, enabled researchers to study psychopathologies including depression, anxiety, stress, and loneliness[15;16;17;18;19]. Closely related to this research, Guntuku et al. showed how social media language distinctly associates with healthcare utilization, in terms of emergency and inpatient hospital visits[20].

Social media plays a pivotal role in college students' lives and is reflective of their behavior and psychosocial wellbeing[21;22;23;24;25]. Social media helps the college students to draw social capital benefits[26], to satisfy psychosocial needs[22;27], and seek solidarity and support during both normalcy and crisis[23]. Recently, researchers have studied college student wellbeing through campus-specific discussions on Reddit, which function as online analogs of offline and geographically situated college communities[23;24;28]. Bagroy et al. measured campus-specific Mental Wellbeing Index (MWI), and found seasonal trends of mental health expressions which were higher during academic terms as compared to holidays[24], and Saha et al. measured the efficacy of counseling recommendations following student deaths on college campuses[28].

Although considerable research reveals the potential of social media data, its predictive ability in correspondence with *ground-truth* mental health data in colleges is yet to be validated. If validated, social media data can be appropriated to help meet the varying demands of college students' mental healthcare at various points of the academic year both in normalcy and crisis, such as the ongoing COVID-19 pandemic. This study targets this previously unexplored problem to examine if social media expressions of college students reflect mental health service utilization by college students. We conduct our study on a Reddit dataset of the online community of a large U.S. public university, which includes 66,020 posts by 18,401 unique users. We employ transfer learning classifiers to identify the language indicative of symptomatic mental health expressions of depression, anxiety, stress, suicidal ideation, and psychosis. We build seasonal auto-regressive time series models using the monthly number of mental health consultations (ground-truth) in the same university campus. We also examine the social media language using unsupervised language modeling and psycholinguistic characterization to explain the predictive ability of this data in terms of how it captures mental health concerns and challenges faced by college students. Our study bears positive implications on resource planning and management of on-campus mental health services. The ability to forecast demand of mental health services can enable on-campus health centers to plan better and cater to the students' needs during critical times.

2 Results

2.1 Inferring Symptomatic Mental Health Expressions on Social Media

We employed transfer learning classifiers to identify language indicative of depression, anxiety, stress, suicidal ideation, and psychosis on the university's subreddit posts. Out of the 66,020 total posts in 2011–2016, 23.49% were related to depression, 21.62% to anxiety, 42.23% to stress, 14.19% to suicidal ideation, and 31.94% to psychosis. We aggregated the monthly averages of each type, and Figure 1 shows the temporal evolution of these posts. Interestingly, greatest mental health expressions occur in April, July, and November, which also coincide with the periods before examinations for the university in consideration, as well as that for most U.S. colleges that follow three-semester cycle in an academic term (spring, summer, and fall semesters). On the other hand, the mental health expressions are observed to be lowest in December, February, and October, which coincide with winter, spring, and fall holidays respectively.

2.2 Associating Social Media Expressions and On-campus Mental Health Consultations

Next, we examined if inferring the symptomatic mental health expressions bears relevance to the ground-truth data of on-campus mental health consultations. Figure 1 also shows some form of trend and seasonality in the occurrence of symptomatic mental health expressions. A Dicky-Fuller test revealed that these time series are not stationary (p>0.05). Therefore, for each time series, we conducted trend and seasonality decomposition, and applied moving window based trend and seasonality removal to obtain transformed residual time series that passed the stationarity test (p<0.05).

We conducted similar time series decomposition on our ground-truth data. Then, we obtained the cross-correlation between the residual time series of social media mental health expressions and ground-truth data of mental health visits. We built linear regression models at various lags by controlling for base-rates of the previous month's number of mental health consultations and the prevalence of mental health expressions on social media. A lag of n months indicates a comparison where the social media data is shifted by n months behind the ground-truth data. A higher standardized coefficient would explain a greater predictive ability of the social media expressions towards the ground-truth data. Next, Figure 2 plots and reports the cross-correlations for each of the mental health symptomatic expressions. Except for suicidal ideation, we find a positive correlation coefficient for all other measures. We find the greatest coefficient for stress (mean e=0.38), followed by depression (mean e=0.26), psychosis (mean e=0.23), and anxiety (mean e=0.20); all with statistical significance (p<0.05). These correlations suggest that social media data can be potentially effective to predict on-campus mental health consultations.

2.3 Predicting On-campus Mental Health Consultations

Now, we predict on-campus mental health consultations using seasonal auto-regressive integrated moving average (SARIMA) time series modeling. Table 1 shows the predictive performance of the two models, M_0 and M_1 — M_0 is the *baseline model* which uses only the time series of on-campus mental health consultation data, and M_1 combines the time series of on-campus mental health consultation data and mental health expressions captured from the college subreddit..

We find that M_1 shows 13.16% better correlation and 41.25% lower error than M_0. A dependent overlapping correlation between the two model predictions shows a statistical significance (t=-2.07, p<0.01). Figure 3 shows the model predictions in comparison to the actual values. Drawing on permutation test approaches [29,30], we permuted (randomized) the predictions of mental health consultations. 1,000 such permutations of randomized predictions show a Pearson's r=0.09 and SMAPE=32.40 at average, and a probability of 0 of better performance than either of M_0 or M_1. This rejects the null hypothesis that any prediction improvement is by chance. Overall, our results reveal that combining baseline model (M_0) with social media based inferences of symptomatic mental health outcomes (in M_1) is an effective means to predict on-campus mental health consultations.

2.4 Examining How Social Media Language Explains Mental Health Consultations

Finally, we illuminate the characteristics of social media language that corresponds with our ground-truth. For this, we separated the social media data of the months that showed high mental health consultations (Hi-MHC) and those that showed low mental health consultations (Lo-MHC) on a median split. We conducted two types of language analysis, which we describe below.

2.4.1 Analyzing Linguistic Cues

First, we employed an unsupervised language modeling technique called Sparse Additive Generative Model (SAGE). Table 2 shows the most salient keywords distinctly used in Hi-MHC and Lo-MHC months. We find that the Hi-MHC months show greater prevalence of keywords related to academics and examination, such as, "finals", "hours semester", "summer classes", "textbooks", etc, and keywords related to disciplines such as, "cs majors", "geology", and "psychology", e.g., "I need urgent help. I'm about to get kicked out of my CS major. I need a 2.65 entry level GPA to advance. I made an A- in 312 and a C- in data structures, so my CS gpa at 2.66." Hi-MHC also show keywords related to "commencement" and "graduation", which could associate with the stress during post-college transition period of students [31], for instance, "When my met my advisor to apply for graduation he told me that I needed a BF certificate to count as my minor, I wish I knew this before."

In contrast, Lo-MHC months show a greater prevalence of keywords related to events, such as "parties", "football", "events", "hangout", and "social", such as "The parties were pretty lame and we were bored at one. My friends and I stole some beers and broke into a pool only to get nearly arrested!" Likewise, Lo-MHC also show keywords related to people and friends. Other forms of social gatherings such as "game" and "football", and accommodations such as "frat" and "dorm" occur saliently in the Lo-MHC months, e.g., "I'm a freshman, currently pledging a frat! I like partying, programming, drinking, playing, lifting weights, and mindlessly scrolling social media for hours!."

2.4.2 Psycholinguistic Characterization

We next discuss the results from our psycholinguistic characterization. First, we extracted the normalized occurrences of the 50 psycholinguistic categories as per LIWC [32]. Then, for each category, we conducted an independent sample t-test between the occurrences in Hi-MHC and Lo-MHC months followed by a Benjamini-Hochberg-Yekutieli False Discovery Rate (FDR) correction. We present the results in Table 3.

Affective and Cognitive Attributes. Affective and cognitive attributes are indicative of an individual's disclosure and expressiveness in social media language. Among affective attributes, we find that Lo-MHC months show greater prevalence of affective categories, including *anger*, *negative affect*, and *swear*. While all of these categories bear a negative connotation, their greater occurrence reflects greater expressiveness, which is known to be a positive wellbeing indicator[33]. This might associate with people venting out more often about their campus life, such as in, "Now I have even more reason to not live here next year. Fuck this place!" Among cognitive attributes, we find that the Hi-MHC months show a greater prevalence of *tentativeness* and *discrepancies*, which indicate an individual's insecurity and low degree of immediacy about the situation[32;34]. In contrast, the Lo-MHC months show a greater prevalence of *certainty*, *percept*, *hear*, and *see*, e.g., "If you want to save your bandwidth, go to a computer lab, and watch youtube / listen to grooveshark / watch netflix all day long." The greater use of these category of language has been associated with an individual's better cognitive functioning and mental health[35].

Linguistic Style Attributes. We first examine pronoun usage; pronouns are markers of social attention and connectedness[36]. We find that Hi-MHC months show a greater prevalence of *first person singular* and *second person pronouns* — these could be indicative of heightened self-attentional focus, first-hand accounts of personal events, narrative, and conversational language[35], for example, "I added psychology in my second year. I have learned that this is a very rigorous path to take, a huge commitment, and that you may need to take an extra year to complete. I must have taken at least 15 semester hours for every semester I spent here, peaking at 21 hrs last semester.", where an individual describes the challenges of their college journey. In contrast, Lo-MHC months show a greater prevalence of first person plural pronoun which associates with narrating as a collective identity[23;37], such as in "There are plenty of ways to socialize here, as we have several student organizations.". We also see a greater use of several function words including *preposition*, *conjunction*, *relative*, and *inclusive* in the Hi-MHC months, which are known to relate with personal narrative writing style[37].

Personal and Social Concerns. Among personal and social attributes, we find that Hi-MHC show a greater usage of *work* and *achievement* keywords, which could associate with discussions on career and self-actualization, and a greater use of *money* may associate with students discussing financial concerns, e.g., "Is it possible to consolidate jobs, save money and improve level of service? What would be an implementation to achieve this?". Hi-MHC also show a greater use of keywords related to *home*, which could include challenges with roommate, e.g., "I got stuck on a top floor between a bad roommate and an old, tiny room, such a terrible year." In contrast, Lo-MHC months show a greater use of social words, such as in "I have a lot of free time and realized I really don't have a lot of friends. I've always been a social person, but it's been hard to make friends at this time of the year, since classes and clubs and everything are ending. How do you recommend I meet some new peeps?".

3 Discussion

3.1 Principal Findings
This study showed that social media interactions of college students can help predict ground-truth data of on-campus mental health consultations. We adopted machine learning approaches to infer mental health expressions on a university's Reddit community, and then incorporated the model outcomes in time series forecasting models of the normalized number of on-campus mental health consultations. First, we found that (online) mental health expressions of college students correlate with (offline) mental health service utilization on college campus. Second, we found that the SARIMA model of forecasting on-campus mental health consultations accounting for social media data could predict the ground truth within 10.65% of error, which was also 38% lower error than models that did not include social media data. Finally, we conducted a deeper dive into the language of social media posts by comparing the data of months with high and low mental health visits using psycholinguistic characterization and an unsupervised language modeling called SAGE. We found that the months of high mental health visits tend to show a greater prevalence of words related to academics, academic examinations, career, and psycholinguistic attributes indicative of worse mental wellbeing, whereas the months of lower mental health visits show a greater prevalence of words related to social, partying, leisure, and psycholinguistic attributes indicative of better mental health. Together, social media data bears the ability to capture the language and social interaction of college students, and therefore can function as a "verbal sensor" to assess mental health needs and demands of college students.

3.2 Methodological and Practical Implications
This work establishes the construct validity of computational assessments of mental health from social media data. This data can therefore serve as an unobtrusive and passive lens to gauge offline critical measures that are otherwise challenging to predict, including other forms of mental health service utilization, such as uptake of peer support interventions, should that type of data be accessible or easily gathered. Our study also demonstrated the face validity of this data, where it revealed discussions and concerns related to local, contextual, and contemporary events of interest, for example, during certain political event on gun laws in U.S., a student posted, "It's fucking nonsensical to carry a pistol around campus despite a handgun license!" Likewise,

following a student death on the campus, students felt stressed and anxious about the event, e.g., "It is so depressing! Seems like he jumped out wanting to die." This construct and face validity showcases promise that, since our machine learning models were built on considerable amounts of social media data, we believe they would be applicable across time periods of various lengths, capturing the ebbs and flows of a typical academic year, such as expectedly stressful periods as well as those when students typically recuperate and rejuvenate. Nevertheless, we note that due to the underlying sensitive nature of the mental health consultation data and practical challenges in gathering and gaining access to it, we had to rely on data from a single university. Consequently, we cannot claim generalizability at this stage. Still, this paper provides a first feasibility study of validity of social media data that can be extended in future research, spanning different universities, contexts, and datasets.

Next, this work provides empirical evidence that can help to move toward constructing practical applications of on-campus mental health assessments using passive and unobtrusive data sources. Recent research has highlighted the extent to which stakeholders — including campus stakeholders and more generally, clinicians — value the potential of these technologies, such as in the form of proactive mental health assessment tools[38;39]. This work established the construct and face validity of these assessments, as described above, and therefore, can guide building tools and dashboards that proactively assess the mental health of college students from online social chatter. Although not ready for real-world use immediately, we foresee two applications that our work could inspire.

The first application sounds ways to assess campus pulse or campus morale — timely, contextual information regarding the mental wellbeing of students. These can be in the form of interfaces, visualizations, and systems[38] which help college stakeholders, including administrators, policymakers, and wellbeing councils, to gauge the needs of the students and accordingly ensure that adequate resources are available and measures are taken to meet the demands of mental health related services. Because our approach can yield assessments over time, the models can further be used to capture the ebbs and highs of mental wellbeing as well as its temporally-varying and evolving characteristics, such as during a typical academic year. Since our methods were predictive of mental health consultations on campus, these assessments could also be used to understand the impacts of academic events like examinations, regulations and policy decisions in campus life.

A second application of this work could center around facilitating improved preparedness in campus in case of an emergency or crisis, and assessing mental resilience of the student body in response to adverse events that affect mental wellbeing of student (e.g., shooting incident on campus[23], an infectious disease outbreak like COVID-19[40], and so on). Speaking more specifically, such preparedness may mean managing/increasing the allocation of resources in the student health clinics on campus in the form of available clinicians or consulting hours, amplifying avenues for seeking alternative sources of mental health help, such as peer support, peer counseling, or crisis rehabilitation services, or even organizing awareness and educational initiatives/campaigns that encourage students to seek help and care more proactively. In essence, with our models, decision-making and resource allocation around college student mental health could be made more evidence-based.

3.3 Limitations and Future Directions

Our study has limitations, some of which also suggest novel and impactful future directions. We cannot claim clinical validity to our assessments, and building upon prior work[41] is a direction to evaluate in the future. The findings of our study is limited to one college campus and a single form of ground truth data (on-campus mental health consultations). However, our computational approaches can be translated and adapted on other college campuses and for other wellbeing measures. We note that social media data suffer from limitations of sparsity and self-selection, i.e., this data only allows us to observe those who use and choose to post on social media. Therefore, the utility of these approaches are bounded by how active and generally engaged the social media discussion board and students of a college campus are, although we expect our methods to be applicable to comparably sized institutions and with similar demographic makeup as the one studied in this paper. Prior work noted that Reddit communities with at least 500 subscribers are somewhat representative of the campus population[23;24]. Future research can thus expand the models developed here to varied university settings with active social media presence, such as a rural or suburban institution, or a liberal arts college, to test generalizability and robustness of the construct validity findings explored in this research. Further, as in the case of any large data source, Reddit data is not immune to noise. Despite moderation strategies, this data can contain discussions not related to personal and campus lives of students (e.g., advertisements, promotions, etc.), and members who do not belong to the college communities — these need to be accounted for when considering practical implementations of computational and data-centric assessments. Future work can also validate mental wellbeing assessments from other social media streams that allow longitudinal posting, instantaneous interactions, and private socializations such as Facebook, Twitter, or Snapchat, which can provide complementary information about individual and collective mental health on college campuses. Finally, replicating and reproducing the validity results from this work to other types of mental health service utilization data would bolster confidence in the application of our methods to real world settings.

4 Methods

4.1 Ground-truth Data of On-Campus Mental Health Service Utilization

This research builds upon health center data stemming from a large public university in the southern U.S. with an enrollment of over 50,000 students. Our ground-truth dataset comprises the count of monthly health center visits by students at the same university. The visits are classified into two types: visits related to mental health issues, and those unrelated to mental health issues. This data spans a period of 84 months: September 2009 to August 2016.

For the purposes of our study, we normalized the monthly measure of mental health consultations as the percentage of enrolled students who sought mental health service in the same month. Such a normalization facilitates two goals – 1) minimalization of confounding outliers and distortion due to total number of enrolled students; and 2) preservation of the privacy of the university and the students of the university whose data is being studied.

4.2 Social Media Data

We focus on the social media data pertaining to the college students from the same university. For this purpose, we use data from Reddit. Reddit is a popular social media platform among the age group between 18-29 years: Pew Research found that 65% of Reddit users are young adults[42]. This age group aligns with the typical college student demographic. Reddit facilitates focused conversations through "subreddits" that comprise of members interested in a specific topic. Many colleges have a dedicated subreddit community, which provides a common portal for the students on a campus to share and discuss about a variety of issues related to their personal, social, and academic life[23;24]. Reddit is suitable data source for the study as it allows us to isolate posts from students from a particular college campus. Reddit, by design, facilitates candid disclosures by allowing pseudonymous and throwaway accounts, and community-driven moderation to maintain authenticity of members and civility and relevance of discussions[43;44;45]. In the case of college subreddits, the members often need to provide proof of their authenticity status to the moderators before participating in the discussions. While the subreddits also remain open to the alumni and staff, typically, only active students engage the most in ongoing discussion threads. Prior work has also leveraged Reddit data to study college students[23;24;28;46]. We obtained the data from the subreddit corresponding to the same university under study, using the BigQuery API which hosts Reddit data archives[23;24]. This archive included 66,020 posts by 18,401 unique users averaging at 33 posts per day between May 2011 and August 2016. The rest of the paper studies this period as this overlaps with our ground-truth data availability as well.

4.3 Modeling Approach

Our primary objective concerned examining if the online college community data is reflective of on-campus mental health service consultations. We identified the language indicative of symptomatic mental health outcomes from these social media posts. Then, we conducted time-series modeling to predict the mental health visits. We evaluated if including information gathered from social media data improved the predictions.

Measuring Symptomatic Mental Health Expressions on Social Media Data. We quantified mental health related expression in Reddit posts using machine learning classifiers identifying the language indicative of symptomatic mental health expressions of depression, anxiety, stress, suicidal ideation, and psychosis. We adopted the approach presented in Saha et al.[17]. Essentially, these classifiers are built using transfer learning methodologies, i.e., transferring a classifier trained on a different labeled dataset. These classifiers are *n*-gram (*n*=1,2,3) based binary SVM models where the positive class of the training datasets come from appropriate subreddits, i.e., *r/depression* for depression, *r/anxiety* for anxiety, *r/stress* for stress, *r/SuicideWatch* for suicidal ideation, and *r/psychosis* for psychosis, and the negative class of training data comes from non-mental health content on Reddit — a collated sample of 20M posts, gathered from 20 subreddits from Reddit's home page such as *r/AskReddit*, *r/aww*, *r/movies*, etc. These classifiers perform at a high accuracy of approximately 0.90 on test data[17]. We used the classifiers to label each post in our Reddit dataset with binary (0 or 1) labels of each symptomatic mental health expression.

Predicting Mental Health Service Utilization. To predict monthly mental health consultations, we adopted a time series modeling approach. We used seasonal auto regressive integrated moving average techniques (SARIMA) — a standard time series forecasting method based on past behavior accounting for seasonality[47]. SARIMA incorporates seasonality in addition to auto regressive integrated moving average techniques (ARIMA)[48], and is suitable in time series with seasonality (e.g., in our case there is known seasonality in academic cycles). We adopted a *k*-fold (*k*=10) cross-validation approach to predict and evaluate our modeling approaches. We first set aside the data from the first year of our dataset (2011) as the default training set so that the models could learn from the same baseline historical data. Then, we obtained various combinations of 10-folds, i.e., 90% of the remaining data was used to build a model that predicted the monthly mental health consultations on the remaining 10% data, and we iterated on various combinations to predict the entire dataset. As our work primarily targets the efficacy of social media data in understanding mental health service utilization, we built two kinds of models, as listed below.

- Model M_0 is trained using only the time series of on-campus mental health consultation data. This model can be considered to be the one used in most in-practice purposes, or as our *baseline model*.

- Model M_1 is trained using the time series of on-campus mental health consultation data, in conjunction with the time series of monthly aggregated mental health discussions in social media. For this, we calculated the monthly average of posts relating to depression, anxiety, stress, and suicidal ideation as identified by our classifiers.

We used the above trained models to separately predict the number of monthly consultations in test data. We pooled all the predictions together to compare against the actual values and computed the Pearson correlation coefficient (r), where higher values directly associate with better performance. We also measured the prediction error between the actual and predicted data as mean absolute error (MAE) and symmetric mean absolute percent error (SMAPE). MAE calculates the arithmetic average of the absolute errors ($|y_i - x_i|$) where y_i and x_i are the predicted and actual values respectively, and SMAPE calculates percentage of relative errors ($|y_i - x_i|/[(|y_i| + |x_i|)/2]$), and is bounded between 0–100. For both error measures, lower values indicate lower error and better predictive performance.

While comparing M_0 and M_1, if M_1 shows comparatively better predictive performance than M_0, we would conclude that using social media data contributes to better predict on-campus monthly mental health consultations. To measure statistical significance in prediction differences between M_0 and M_1, we conducted t-tests using the dependent overlapping correlation method, which controls for comparing against a common variable of interest (here, the ground-truth number of monthly on-campus mental health consultations)[49].

4.4 Analyzing the Social Media Language of Mental Health

Finally, we interpreted how social media language associates with on-campus mental health consultations. We obtained the months of high and low number of mental health visits — we adopted a median split on the normalized number of visits in a month. Then, we examined the distinction of these periods as per social media language. This examination would help establish the face validity of the social media language in correspondence to the ground-truth. We conducted two analyses:

First, we adopted an unsupervised language modeling technique called the Sparse Additive Generative Model (SAGE)[50]. Given two documents, SAGE finds the keywords that distinguish the documents by comparing the parameters of two logistically parameterized multinomial models using a self-tuned regularization parameter that controls the tradeoff between frequent and rare terms. We aimed to obtain keywords that would relate with the key concerns faced by college students that lead to heightened mental health concerns.

Second, we conducted a psycholinguistic analysis. For this, we used the well-validated psycholinguistic lexicon, Linguistic Inquiry and Word Count (LIWC)[32]. LIWC characterizes social media language in 50 psycholinguistic attributes ranging across affect, cognition and perception, interpersonal focus, temporal references, lexical density and awareness, and personal and social concerns. This analysis would help to contextualize the social media language of college students in the literature on mental health and therefore explain the predictive ability of social media language.

Data Availability

The datasets analyzed during the current study are not publicly available to protect the privacy of the university campus, health center, and students whose data was used in this study. The public Reddit data can be made available upon request, subject to an appropriate data use agreement, if applicable.

Acknowledgements

Ethics Declaration

Because this work used publicly accessible Reddit data without any interaction with individuals whose data is analyzed, it did not qualify for ethics board approval. However, utmost care was taken to secure the privacy of the individuals in the dataset. This paper used de-identified data for analyses and presents paraphrased quotes to minimize traceability.

Contributors

KS and MDC designed the research; KS conceptualized and developed the analytic techniques; KS and AY gathered and analyzed the data; KS, AY, RB, JP, and MDC interpreted the results; KS, AY, and MDC drafted the paper; and RB and JP read, edited, and provided feedback on the paper.

Competing Interests

None declared.

References

1. American College Health Association et al. American college health association-national college health assessment spring 2019 reference group data report (abridged): the american college health association. 2019.

2. Sarah Ketchen Lipson, Emily G Lattie, and Daniel Eisenberg. Increased rates of mental health service utilization by us college students: 10-year population-level trends (2007–2017). *Psychiatric services*, 70(1):60–63, 2019.

3. Emily G Lattie, Sarah Ketchen Lipson, and Daniel Eisenberg. Technology and college student mental health: challenges and opportunities. *Frontiers in psychiatry*, 10:246, 2019.

4. (CCMH) Center for Collegiate Mental Health (CCMH). 2018 annual report. 2018.

5. Robert P Gallagher. National survey of college counseling centers 2014. 2015.

6. Nuran Bayram and Nazan Bilgel. The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. *Social psychiatry and psychiatric epidemiology*, 43(8):667–672, 2008.

7. Roger Tourangeau, Lance J Rips, and Kenneth Rasinski. *The psychology of survey response*. Cambridge University Press, 2000.

8. Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Ubicomp*.

9. Koustuv Saha, Larry Chan, Kaya De Barbaro, Gregory D Abowd, and Munmun De Choudhury. Inferring mood instability on social media by leveraging ecological momentary assessments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):95, 2017.

10. Scott A Golder and Michael W Macy. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051):1878–1881, 2011.

11. Kokil Jaidka, Salvatore Giorgi, H Andrew Schwartz, Margaret L Kern, Lyle H Ungar, and Johannes C Eichstaedt. Estimating geographic subjective well-being from twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences*, 117(19):10165–10171, 2020.

12. Johannes C Eichstaedt and Aaron C Weidman. Tracking fluctuations in psychological states using social media language: A case study of weekly emotion. *European Journal of Personality*, 34(5):845–858, 2020.

13. Ryan L Boyd and James W Pennebaker. A way with words: Using language for psychological science in the modern era. *Consumer psychology in a social media world*, pages 222–236, 2015.

14. Ryan Boyd, Steven Wilson, James Pennebaker, Michal Kosinski, David Stillwell, and Rada Mihalcea. Values in words: Using language to evaluate and understand personal values. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, 2015.

15. Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*, 2013.

16. Glen Coppersmith, Craig Harman, and Mark Dredze. Measuring post traumatic stress disorder in twitter. In *International AAAI Conference on Web and Social Media*, 2014.

17. Koustuv Saha, Benjamin Sugar, John Torous, Bruno Abrahao, Emre Kıcıman, and Munmun De Choudhury. A social media study on the effects of psychiatric medication use. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01):440–451, Jul. 2019.

18. Sharath Chandra Guntuku, Rachelle Schneider, Arthur Pelullo, Jami Young, Vivien Wong, Lyle Ungar, Daniel Polsky, Kevin G Volpp, and Raina Merchant. Studying expressions of loneliness in individuals using twitter: an observational study. *BMJ open*, 9(11), 2019.

19. H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, et al. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9):e73791, 2013.

20. Sharath Chandra Guntuku, H Andrew Schwartz, Adarsh Kashyap, Jessica S Gaulton, Daniel C Stokes, David A Asch, Lyle H Ungar, and Raina M Merchant. Variability in language used on social media prior to hospital visits. *Scientific reports*, 10(1):1–9, 2020.

21. Megan A Moreno, Lauren A Jelenchick, Katie G Egan, Elizabeth Cox, Henry Young, Kerry E Gannon, and Tara Becker. Feeling bad on facebook: depression disclosures by college students on a social networking site. *Depress. Anxiety*, 2011.

22. Adriana M Manago, Tamara Taylor, and Patricia M Greenfield. Me and my 400 friends: the anatomy of college students' facebook networks, their communication patterns, and well-being. *Developmental psychology*, 2012.

23. Koustuv Saha and Munmun De Choudhury. Modeling stress with social media around incidents of gun violence on college campuses. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW), December 2017.

24. Shrey Bagroy, Ponnurangam Kumaraguru, and Munmun De Choudhury. A social media based index of mental well-being in college campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017.

25. Yiran Wang, Melissa Niiya, Gloria Mark, Stephanie M Reich, and Mark Warschauer. 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*, pages 571–582. ACM, 2015.

26. Nicole B Ellison, Charles Steinfield, and Cliff Lampe. The benefits of facebook "friends": Social capital and college studentsúse of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

27. Zheng Wang, John M Tchernev, and Tyler Solloway. A dynamic longitudinal examination of social media use, needs, and gratifications among college students. *Computers in human behavior*, 28(5):1829–1839, 2012.

28. Koustuv Saha, Ingmar Weber, and Munmun De Choudhury. A social media based examination of the effects of counseling recommendations after student deaths on college campuses. In *Twelfth International AAAI Conference on Web and Social Media*, 2018.

29. Koustuv Saha, Manikanta D Reddy, Vedant Das Swain, Julie M Gregg, Ted Grover, Suwen Lin, Gonzalo J Martinez, Stephen M Mattingly, et al. Imputing Missing Social Media Data Stream in Multisensor Studies of Human Behavior. In *Proceedings of International Conference on Affective Computing and Intelligent Interaction (ACII 2019)*, 2019.

30. Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 7–15. ACM, 2008.

31. Jude P Mikal, Ronald E Rice, Audrey Abeyta, and Jenica DeVilbiss. Transition, stress and computer-mediated social support. *Computers in Human Behavior*, 29(5):A40–A53, 2013.

32. Yla R Tausczik and James W Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54, 2010.

33. James W Pennebaker and Cindy K Chung. Expressive writing, emotional upheavals, and health. *Handbook of health psychology*, pages 263–284, 2007.

34. James W Pennebaker and Laura A King. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296, 1999.

35. James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1):547–577, 2003.

36. Ryan L. Boyd and H. Andrew Schwartz. Natural language analysis and the psychology of verbal behavior: The past, present, and future states of the field. *Journal of Language and Social Psychology*, 40(1):21–41, January 2021. ISSN 0261-927X. doi: 10.1177/0261927X20967028.

37. Michael A Cohn, Matthias R Mehl, and James W Pennebaker. Linguistic markers of psychological change surrounding september 11, 2001. *Psychological science*, 15(10):687–693, 2004.

38. Dong Whi Yoo and Munmun De Choudhury. Designing dashboard for campus stakeholders to support college student mental health. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 61–70, 2019.

39. Dong Whi Yoo, Michael L Birnbaum, Anna R Van Meter, Asra F Ali, Elizabeth Arenare, Gregory D Abowd, and Munmun De Choudhury. Designing a clinician-facing tool for using insights from patients' social media activity: Iterative co-design approach. *JMIR Mental Health*, 7(8):e16969, 2020.

40. Koustuv Saha, John Torous, Eric D Caine, and Munmun De Choudhury. Psychosocial effects of the covid-19 pandemic: Large-scale quasi-experimental study on social media. *Journal of medical internet research*, 22(11):e22600, 2020.

41. Michael L Birnbaum, Sindhu Kiranmai Ernala, Asra F Rizvi, Munmun De Choudhury, and John M Kane. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *Journal of medical Internet research*, 19(8), 2017.

42. Pew. `pewinternet.org/fact-sheet/social-media`, 2018.

43. Munmun De Choudhury and Sushovan De. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Eighth international AAAI conference on weblogs and social media*, 2014.

44. Nazanin Andalibi, Oliver L Haimson, Munmun De Choudhury, and Andrea Forte. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proc. CHI*, 2016.

45. Jessica Vitak. The Impact of Context Collapse and Privacy on Social Network Site Disclosures. *Journal of Broadcasting & Electronic Media*, 56, 2012. ISSN 0883-8151.

46. Koustuv Saha, Eshwar Chandrasekharan, and Munmun De Choudhury. Prevalence and psychological effects of hateful speech in online college communities. In *Proceedings of the 10th ACM Conference on Web Science*, pages 255–264, 2019.

47. Flávio Fonseca Nobre, Ana Beatriz Soares Monteiro, Paulo Roberto Telles, and G David Williamson. Dynamic linear model and sarima: a comparison of their forecasting performance in epidemiology. *Statistics in medicine*, 20(20):3051–3069, 2001.

48. Brian K Nelson. Time series analysis using autoregressive integrated moving average (arima) models. *Academic emergency medicine*, 5(7):739–744, 1998.

49. Olive Jean Dunn and Virginia Clark. Comparison of tests of the equality of dependent correlation coefficients. *Journal of the American Statistical Association*, 66(336):904–908, 1971.

50. Jacob Eisenstein, Amr Ahmed, and Eric P Xing. Sparse additive generative models of text. In *ICML*, 2011.
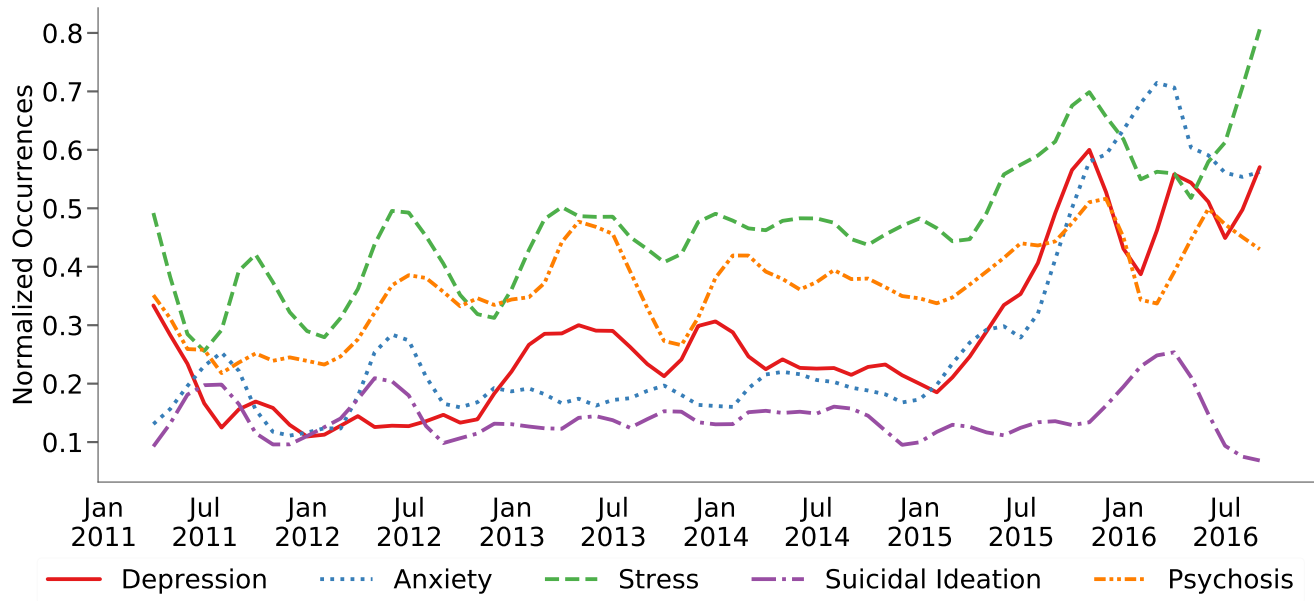
Figure 1. Temporal evolution of the normalized prevalence of social media expressions indicative of mental health symptomatic outcomes on the college subreddit.

(a) Depression **(b)** Anxiety **(c)** Stress

(d) Suicidal Ideation **(e)** Psychosis
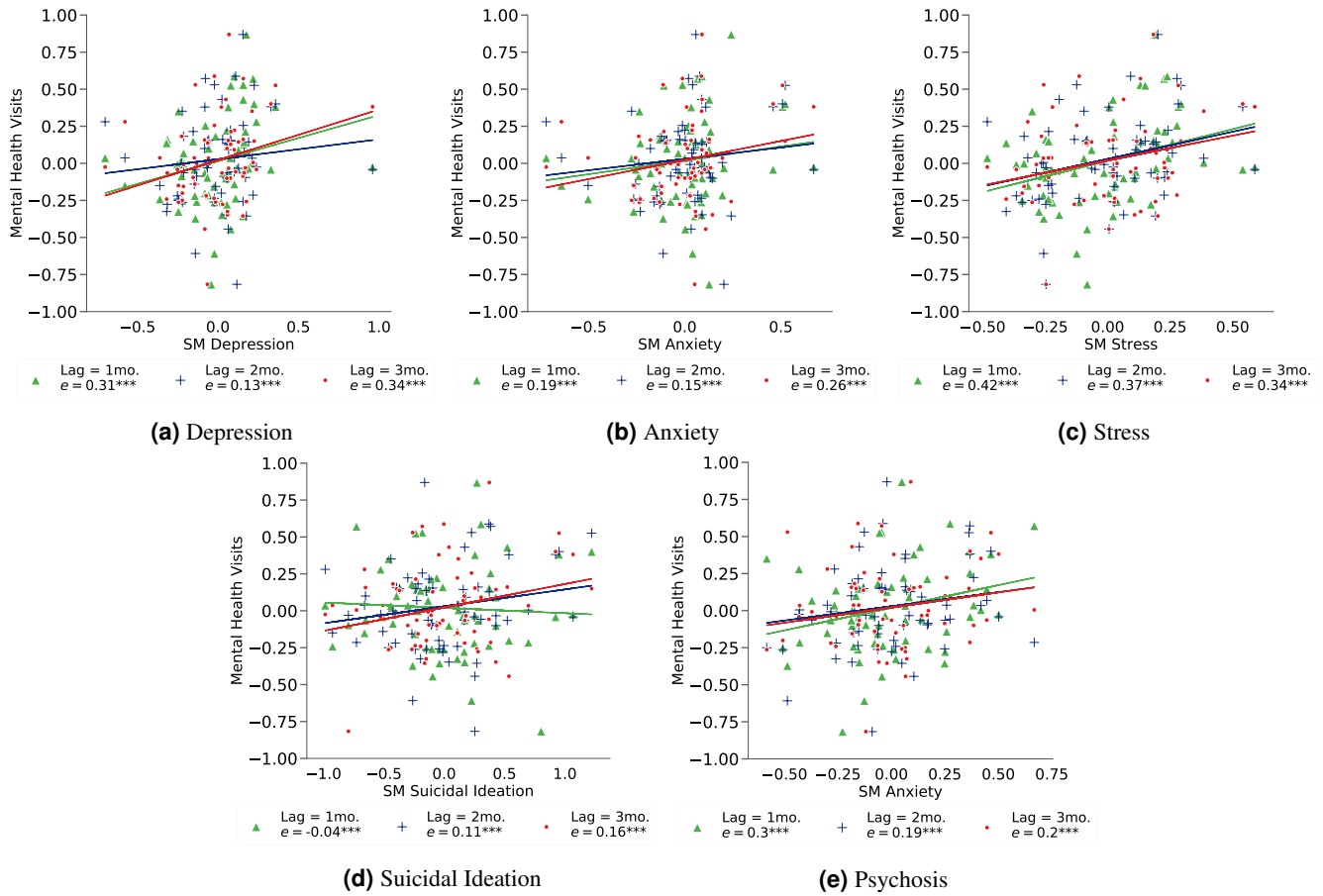
Figure 2. Lagged Cross-correlation between the trend-seasonality removed residuals of mental health visits and mental health symptomatic expressions on social media (SM).

Table 1. Seasonal Auto Regressive Integrated Moving Average (SARIMA) models of predicting normalized number of mental health consultations in the test dataset (year 2016 data) (* $p<0.01$, ** $p<0.001$, *** $p<0.0001$).

Model	Pearson's r	MAE	SMAPE
Model M_0	0.76***	2.51	22.64
Model M_1	0.86***	1.96	13.30
Dependent Overlapping Correlation: t=-2.07**			
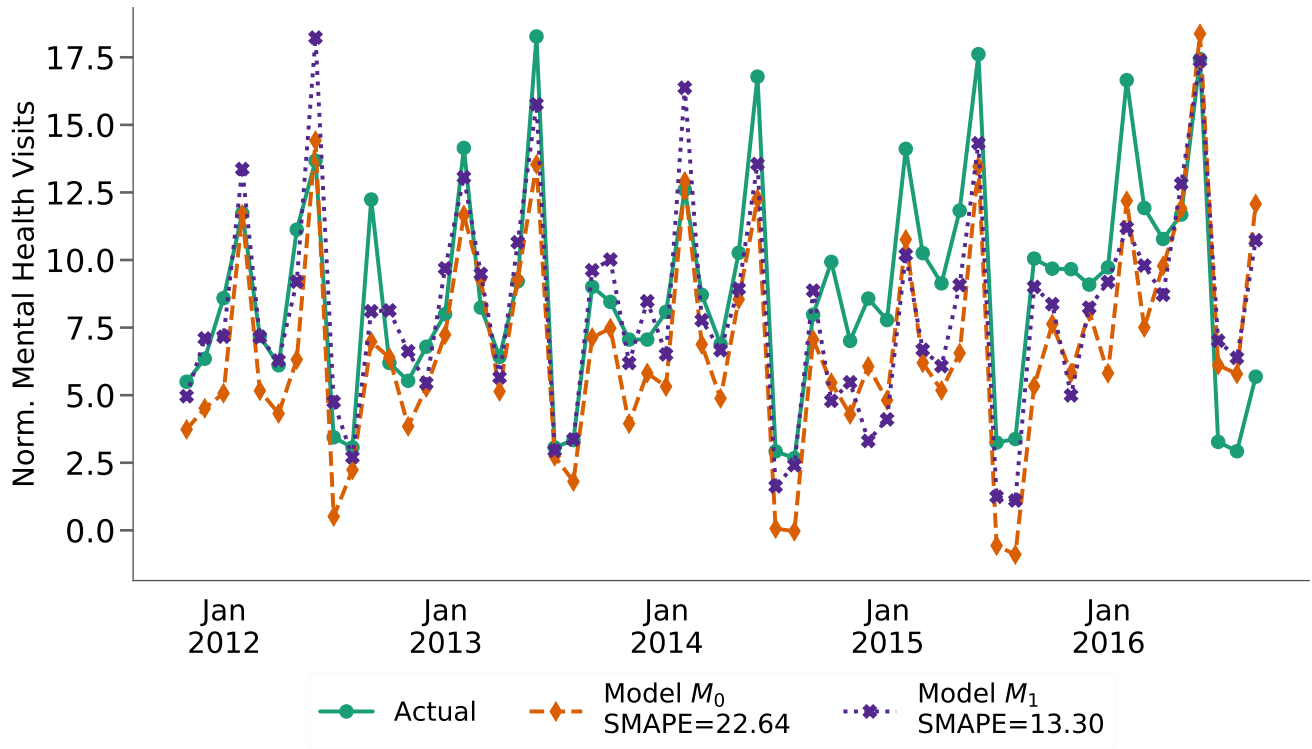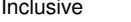
Figure 3. SARIMA models to predict the number of mental health service utilization.

Table 2. Top salient *n*-grams (*n*=1,2,3) distinguishing months of high and low mental health visits as per SAGE[50]. Bar lengths and color indicate magnitude and sign of SAGE value; **pink** bars (positive SAGE) indicate a saliency in Hi-MHC months whereas **green** bars (negative SAGE) indicate a saliency in Lo-MHC months.

High Mental Health Visit Months (Hi-MHC)			
n-gram	SAGE	*n*-gram	SAGE
commencement	2.51	grade	0.65
sword	2.02	gender	0.63
smokers	1.48	mom	0.63
startup	1.14	entry level	0.61
swim	1.08	summer classes	0.61
mass	1.04	equivalent	0.59
april	0.98	drug	0.59
counselors	0.91	dance	0.57
blackboard	0.89	completed	0.56
smell	0.88	summer	0.55
graduation	0.87	comfortable	0.55
shooting	0.85	laws	0.53
finals	0.85	books	0.52
psychology	0.83	swimming	0.52
engineering physics	0.82	professional	0.52
driver	0.77	geology	0.51
permit	0.77	metro	0.51
midland	0.75	final	0.50
cs majors	0.74	textbooks	0.49
lower division	0.74	labs	0.48
february	0.73	aerospace	0.48
appeal	0.71	exchange	0.48
elective	0.69	gay	0.47
busy	0.68	astronomy	0.47
hours semester	0.68	textbook	0.47
phy	0.65	residency	0.47

Low Mental Health Visit Months (Lo-MHC)			
n-gram	SAGE	*n*-gram	SAGE
software	-0.78	bring	-0.38
parties	-0.63	laptop	-0.38
greek	-0.57	buildings	-0.37
football	-0.57	project	-0.37
internet	-0.57	hangout	-0.37
season	-0.57	funny	-0.36
make friends	-0.56	union	-0.36
game	-0.56	weird	-0.36
choice	-0.54	ticket	-0.36
downtown	-0.53	freshmen	-0.35
weekend	-0.52	agree	-0.34
quiet	-0.47	games	-0.34
phone	-0.47	floor	-0.31
act	-0.46	living campus	-0.31
hell	-0.46	hall	-0.30
events	-0.46	guy	-0.29
frat	-0.44	score	-0.29
god	-0.43	lol	-0.27
tech	-0.42	world	-0.27
car	-0.42	smart	-0.27
reasons	-0.39	tickets	-0.26
action	-0.39	opinion	-0.26
teach	-0.39	culture	-0.25
dorms	-0.39	reason	-0.25
bikes	-0.39	friends	-0.25
party	-0.38	social	-0.19

Table 3. Comparing psycholinguistic attributes across clusters, with mean normalized percentage occurrence in Hi-MHC and Lo-MHC months and *t*-test statistic. Statistical significance reported after Benjamini-Hochberg- Yekutieli False Discovery Rate correction (*** $p<.001$, ** $.001<p<.01$, * $.01<p<.05$). Bar length and color indicates magnitude and sign of *t*-statistics. **Pink** bars (positive *t*) indicate greater prevalence in Hi-MHC months and **green** bars (negative *t*) indicate a a greater prevalence in Lo-MHC months.

Category	Hi-MHC Mean%	Lo-MHC Mean%	t-stat.	p
Affect				
Negative Affect	0.44	0.50		-3.01 **
Swear	0.28	0.40		-5.00 ***
Anger	0.46	0.65		-6.38 ***
Cognition and Perception				
Tentativeness	2.84	2.74		2.60 ***
Discrepancies	1.48	1.43		1.29 *
Hear	0.38	0.43		-1.62 *
Certainty	1.03	1.10		-2.04 *
Percept	1.35	1.45		-2.64 *
See	0.61	0.67		-2.71 *
Interpersonal Focus				
2nd P.	1.65	1.57		1.78 *
1st P. Singular	3.08	3.00		1.05 *
Indef. Pronoun	4.40	4.56		-2.77 *
1st P. Plural	0.29	0.33		-3.06 **
3rd P.	0.28	0.38		-5.60 ***

Category	Hi-MHC Mean%	Lo-MHC Mean%	t-stat.	p
Lexical Density and Awareness				
Preposition	9.44	9.13		4.69 **
Relative	10.21	9.95		2.98 **
Conjunction	4.18	4.06		2.64 *
Inclusive	2.65	2.58		2.18 *
Quantifier	1.73	1.67		1.32 *
Personal and Social Concerns				
Work	4.59	3.92		10.26 ***
Achievement	1.26	1.15		3.45 ***
Money	0.87	0.80		2.56 *
Home	0.58	0.52		2.27 *
Social	6.15	6.46		-1.06 **
Religion	0.08	0.11		-1.62 *
Bio	0.97	1.04		-1.92 *
Humans	0.52	0.58		-2.72 *
Sexual	0.23	0.30		-3.63 **
Body	0.23	0.30		-4.58 ***