Research Article

ISSN 2641-4317

International Journal of Psychiatry Research

Diagnosis of Mental Health Issues In Social Forums Using Semantic Biomarkers, Markovian Models and Artificial Intelligence

Nithin Parthasarathy*

Northwood High School, Irvine, CA.

***Correspondence:**
Nithin Parthasarathy, Northwood High School, Irvine, CA.

Received: 21 October 2019; **Accepted**: 21 November 2019

Citation: Nithin Parthasarathy Diagnosis of Mental Health Issues In Social Forums Using Semantic Biomarkers, Markovian Models and Artificial Intelligence. Int J Psychiatr Res. 2019; 2(6): 1-9.

Keywords

Mental health, Behavior, Biomarkers, Bipolar disorder.

Introduction

At every stage of life, our mental health is an integral part of our overall well-being. In the emotional roller coaster of life, what constitutes good mental health is somewhat fuzzy. Our behavior and indirectly our actions are indeed dependent on our mental well-being underlying its importance. Figure 1 depicts statistics on the prevalence of mental illness in the United States. Note that approximately 1 out of 5 adults is affected by mental illness which corresponds to a significant portion of the population!

- Approximately 1 in 5 adults in the U.S.—43.8 million, or 18.5%—experiences mental illness in a given year.[1]
- Approximately 1 in 25 adults in the U.S.—9.8 million, or 4.0%—experiences a serious mental illness in a given year that substantially interferes with or limits one or more major life activities.[2]
- Approximately 1 in 5 youth aged 13–18 (21.4%) experiences a severe mental disorder at some point during their life. For children aged 8–15, the estimate is 13%.[3]
- 1.1% of adults in the U.S. live with schizophrenia.[4]
- 2.6% of adults in the U.S. live with bipolar disorder.[5]
- 6.9% of adults in the U.S.—16 million—had at least one major depressive episode in the past year.[6]
- 18.1% of adults in the U.S. experienced an anxiety disorder such as posttraumatic stress disorder, obsessive-compulsive disorder and specific phobias.[7]

Figure 1: Mental health statistics illustrating the prevalence of mental issues in the U.S population (from https://www.nami.org/Learn-More/Mental-Health-By-the-Numbers).

From a global perspective, it is estimated that 15% of the world population suffer some form of Mental health issue [1]. Motivated by the dramatic impact of an automated and continuous diagnosis tool, this work employs semantic biomarkers (or language) as a strong indicator of mental health. Unlike previous work, semantic equivalence of afflicted mental health are deduced from clinical definitions of behavioral traits for issues such as depression. These are called "Behavioral Semantic Biomarkers" (BSB) and are based on semantic word similarity that happen in depressed as well as normal conversation. Next, another set of markers called "Associative Semantic Biomarkers" (ASB) are defined which are based on semantic similarity or closeness of verbs and adjectives in tweets to normal and abnormal behavioral traits deduced using Artificial Intelligence (AI) algorithms.

It is shown that an optimally weighted combination of BSB's and ASB's used with dynamic programming techniques and applied to a Markovian model of mental state not only accurately diagnoses mental illness such as depression, but also provides an instantaneous deduction of the mental state triggering alerts for extreme conditions! Needless to say, this serves as an automated, cheap and accurate diagnosis tool for evaluating mental health conditions!

Obtaining feeds from social media such as Twitter, both a "target" and "reference" dataset are assembled. First, tweets from users who have reported being diagnosed with specific illness such as depression is obtained. Such an event is typically recorded by a tweet indicating "I was diagnosed with XYZ". Simultaneously, a number of "normal" tweets by individuals are obtained making sure that there is no reported instance of "I was diagnosed with XYZ" in this reference data set. Thus, a "target" and a "reference" dataset (which may also alternatively be called a "depressed" and "normal" dataset while evaluating for identifying depression) is obtained.

The concepts of this work have been partially motivated by the field of astronomy [2] where "Pulsars" or certain stars in deep space which spin so accurately that their use has been proposed as a marker to obtain relative distance. Further, stars have been used in the ancient world as successful navigation guides as has GPS [3] in the modern world. In a similar spirit, a new and novel technique to use semantic distance between tweets and "anchor" words as a guide of a mental health marker has been introduced in this work!

Results obtained from the technique of semantic similarity merged with behavioral traits provide a new direction in mental health study. While extensive results are presented for depression related mental conditions, the technique has broad reaching implications and easily extend to evaluating Schizophrenia, Bipolar Disorder, Seasonal Affective Disorder (SAD) and Autism.

Relevent work

Mental illness has been an active field of study in the field of medicine and psychology. Such studies have typically been behavioral observations/evaluation of patients followed by medical and psychological prescriptions. The earliest work on diagnosis of depression and other related ailments in social media was reported by DeChoudhary [4]. The authors solicited responses from users on specific questions which they augmented with their social profiles to aid in diagnosis. Coppersmith [5] on the other hand approached the issue wherein they obtained social dialogue of users in forums such as Twitter without requiring them to actively interact. With these they conducted a linguistic word count enquiry to ascertain and differentiate users with mental illness from normal users. Differences in language use in social media for users with mental health problems was also reported by both DeChoudhary [4] and Coppersmith [6]. Other papers [7-9] have focused on using Artificial Intelligence (AI) models to the user messages in social forums by converting them to word embeddings [10-12].

In contrast to these papers, rather than treating the entire postings/tweets of users as a black box fed in to an AI program, this work develops a novel "Anchor" word based semantic biomarkers which is then augmented by a linguistic symptomatic approach (which is based on published psychological/behavioral case studies) as an accurate and automated diagnostic tool. The technique developed in this work allows behavioral and psychological attributes to be easily translated and mapped to a probabilistic model. Further, as the mental condition of the user at any instant is modelled as a Markov chain and the most likely state is then decoded allowing evaluation of user's mental states at any given time allowing an intuitive feel on their mental condition.

Dataset Study

The dataset is mined from Twitter by using the public domain scraping tool Twint [10]. Tweets from depressed users (for example) are mined by searching for the phrase "I was diagnosed with depression" in the range from 2008 to 2014. Once the users Twitter ID's are obtained, the entire tweet history of the users are obtained in the time period. The user tweets are first converted to lower case (making the message case insensitive). Next, stopwords which do not convey any specific information are removed (stopwords are words like "and", "the" etc.). If desired, further processing can also be performed such as "Stemming" which changes words into their root form (for example, the word "running" would be changed into its root form of "run").

These cleaned tweets then form our evaluation dataset. Note that similar datasets can be obtained for other mental illness like Schizophrenia, PTSD etc. Next, to obtain the normal dataset, tweets from random twitter ID's in the same time frame are downloaded with care taken to check that the tweets do not contain any indication of diagnosis of depression. This will form the "normal" dataset. 850 users tweets are obtained in the "depression" dataset and 2500 user tweets form the "normal" dataset.

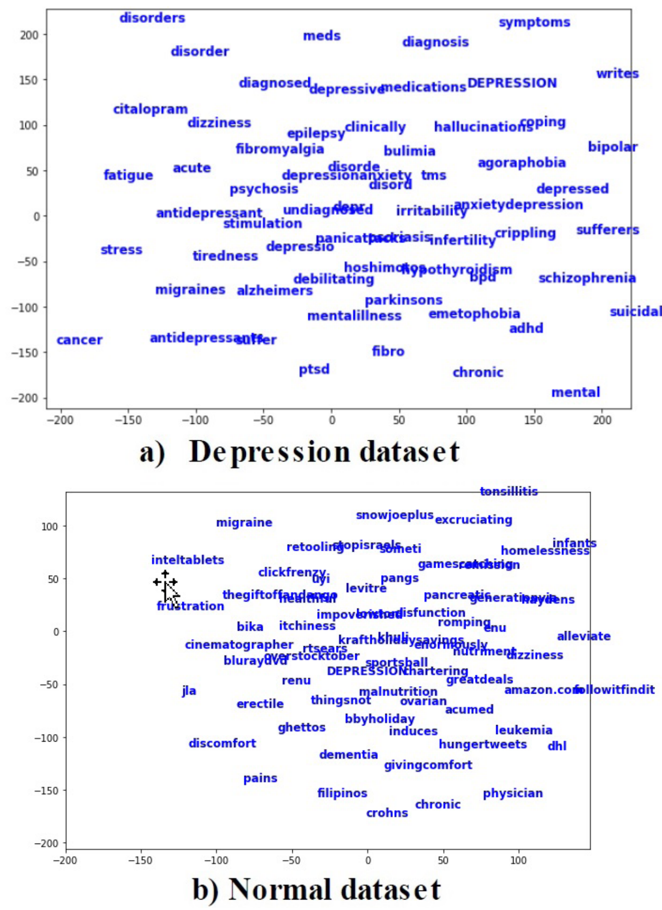

Figure 2: TSNE plot of words close to the word "DEPRESSION" in the two datasets.

Figures 2 shows a comparison plot of words close to depression in the two databases being evaluated. This was obtained by using Word2Vec [11,12] which is a Neural Network producing word embedding's. To visualize these embedding's in the form of a two-dimensional plot where the distance of the closest words to a word of interest (in this case, the word "depression"), the dimensionality reduction technique of TSNE [13] is used. Notice in Figure 2 how close the depicted words are to strong negative emotions as well as medication in the depression dataset.

Figure 3 shows a WordCloud [14] depiction of most frequent adjectives in the "depressed" and "random" dataset. Notice that the depressed dataset carries more extreme emotional words as well as negative adjectives. How can one translate these emotions into effectively separating the normal and depression datasets? The next few sections will elaborate further on this topic.

Figure 3: WordCloud of most frequent adjectives in depressed and normal dataset.

Modelling and Diagnosis of Mental Illness

Markov Chain Models of Mental State Evolution

Earlier sections dealt with the characteristics of the dataset that is used in this work. In this section, models are developed for mental health diagnosis. Emotional state extracted from each tweet of an individual is modelled as a Markov chain. Each time instant in the Markov chain corresponds to the information processed from a user interaction such as a tweet and can potentially be extended to any other user interaction (such as images or videos).

There are three parts to the modelling. First, extracting the relevant information from a given tweet. Second, computing the appropriate state transition probabilities (or transition weights) based on how positive or negative the tweet is. Next, given the Markov state evolution, a final determination has to be made on whether the person is depressed or normal (or any other affliction) as well as the severity of the affliction.

Figure 4 illustrates a depiction of a Markov chain with corresponding transition probabilities. These transitions corresponds to the weights which denote likelihood of moving from say a "depressed" state at time *t* either a new or same state at time *t*+1. Hence, with every processed tweet, the state of the Markov chain is advanced. At any given time *t*, the complete mental state evolution of the individual is then captured by the Markov chain from time 0 until time *t*.

Figure 5 which depicts a path taken by a set of tweets from a "normal" individual beginning at time 0 until time t and another path taken by a "depressed" individual. Of all the possible transition weights, the most probabilistically likely path is retained while less likely paths have been dropped. The exact technique of dropping the least likely path and retaining the most likely path will be described next.

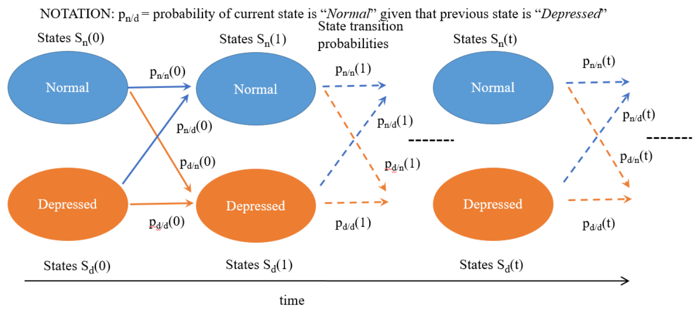

Figure 4: Mental state evolution: Markov state diagram of evolution over time.

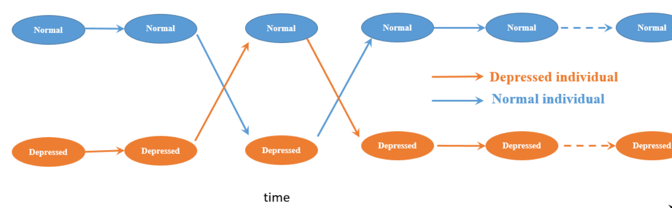

Figure 5: Example of a potential tweet state evolution for a normal and a depressed individual. In this illustration, the most likely paths for the "Depressed" state and "Normal" state are depicted while other possible paths have been dropped.

Deducing the entire state sequence of an evolving Markov chain (also known as a "trellis") has exponential complexity. For example, when the state space comprises of two variables, and the overall length of the trellis is N, the total number of paths to be evaluated is 2N. Fortunately, dynamic programming [15] techniques can be used to reduce the exponential complexity to linear complexity without any loss in evaluation accuracy. In particular, the Viterbi algorithm [16] which is well known in the field of speech processing and digital communication provides a technique to evaluate the most likely candidate path for the Normal and Depressed state.

As shown in Figure 6, dynamic programming generates a path *S(0)*, *S(1)* … *S(n)* through the state space (Normal, Depressed) given the observations (tweets) *T1*, *T2*, …. *Tn* in a manner which maximizes the joint probability of the observation sequence and the state sequence. This is a recursive algorithm and at every time instant, the following weights update is calculated for each state.

$S_n(k) = max\ (S_n(k-1) + p_{n/n}(k)\ ,\ S_n(k-1) + p_{n/d}(k)) \rightarrow$ ***Equation 1***

$S_d(k) = max\ (S_d(k\text{-}1) + p_{d/n}(k)\ ,\ S_d(k\text{-}1) + p_{d/d}(k)) \rightarrow$ ***Equation 2***

One can alternatively use $S_n(k) = max\ (S_n(k\text{-}1)^*\ p_{n/n}(k),\ S_n(k\text{-}1)^*\ p_{n/d}(k))$ as the update metric, but since pn/n(k) in this work is not a strict probability (it ranges from -1 to 1) measure, but a relative branch weight the update is as defined in Equation 1 and 2. To obtain the most likely sequence for the depressed/normal state, a backwards trace common in Viterbi decoding literature [16] is adopted. Final indication of whether the individual is depressed/normal is obtained

Final Diagnosis at time k = *"Depressed" if* $S'_d(k) > S'_n(k)$ *"Normal" otherwise* → ***Equation 3***

where ***S'd*** and ***S'n*** represent the sum of branches in the most likely depressed and normal sequence respectively. Note that the value of ***S'd – S'n*** indicates the magnitude/severity of the affliction

Computing the Transition Weights

How does one then construct the transition probabilities? The approach taken in this work is two-fold. First, behavior persistent in depressed individuals is used to evaluate mental state and construct temporal state transition probabilities. Such linguistically based behavioral markers of mental state are called "Behavioral Semantic Biomarkers" (BSB). Second, words in tweets which are semantically similar to certain emotional target words called "Anchors" (which are characteristics of "depressed" and "normal" individuals as well as their emotional states) are termed "Associative Semantic Biomarkers" (ASB) and used to create a second set of transition probabilities. The complete state transition probability is then obtained by the weighted addition of the ASB's and BSB's.

$p_{n/n}(k) = w_1{}^*p\_ASB_{n/n}(k) + w_2{}^*p\_BSB_{n/n}(k)$ *for the* k^{th} *message or tweet* → ***Equation 4***

where $p_{n/n}$ represents the conditional probability of the current state being normal given the previous state was normal and w_1 and w_2 are weights with $w_1+w_2=1$. Similarly,

$p_{d/d}(k) = w_1{}^*p\_ASB_{d/d}(k) + w_2{}^*p\_BSB_{d/d}(k)$ *at time k* → ***Equation 5***

where $p_{d/d}$ represents the probability of the current state being depressed given the previous state was depressed

Explicitly determining the conditional probability measures *p_ASB* and *p_BSB* in Equation 5 will be covered in the next subsections.

"Behavioral semantic biomarkers" (BSB) computation

Depressed individuals are known to have some linguistic characteristics which can be converted into appropriate state transition probabilities. In this section, we will explore the behavioral characteristics individually. Note that as new behavioral patterns get determined and established, the model can get further refined. Symptoms described in these sections follow the conditions described in current established depression research [17,18].

Behavioral Symptom 1: Excessive use of personal pronouns [18]

Depressed individuals have been behaviorally observed to use an excess of personal (self-referring) pronouns such as "I" and "me" (complete list is provided in the appendix). This is also verified in the datasets used here by computing the frequency of self-referring words in the normal and depressed dataset. More specifically, the probability of an *I_PRONOUN* computed in the normal dataset is 0.07 while that in a depressed dataset is 0.093.

The following steps are then used to map this probability into the Markov model. If a tweet has a self-referring pronoun (also called as an *I_PRONOUN*), it is given a branch weight according to the class ("depressed" or "normal") to which it belongs.

$p\_i\_pronoun_{n/d} = p\_i\_pronoun_{n/n} = 0.07$ *if i_pronoun is present, 0 otherwise*

$p\_i\_pronoun_{d/d} = p\_i\_pronoun_{d/n} = 0.093$ *if i_pronoun is present, 0 otherwise*

In the above equations, the notation *p_i_pronounn/d* denotes the probability of a *I_PRONOUN* when the current state is "normal" given that the previous state was "depressed". Note that the current conditional probabilities are previous state independent and assumed to be stationary (does not change with time) and hence *p_i_pronounn* = $p\_i\_pronoun_{n/d}$ at any time instant.

Behavioral Symptom 2: Excessive use of words expressing absolute emotion [18]

Now, depressed individuals are also observed to use more words expressing absolute emotions such as 'absolutely', 'all', 'always', 'complete', 'completely' etc which is verified in the data set used here. The full list of such words is provided in the Appendix. Then similar to the earlier subsection, based on the probabilities computed from the dataset, we have

$p\_absolutist_{n/d} = p\_i\_absolutist_{n/n} = 0.016$ *if absolutist word is present, 0 otherwise*

$p\_absolutist_{d/d} = p\_i\_absolutist_{d/n} = 0.018$ *if absolutist word is present, 0 otherwise*

Behavioral Symptom 2: Excessive use of words expressing absolute emotion [18]

Now, depressed individuals are also observed to use more words expressing absolute emotions such as 'absolutely', 'all', 'always', 'complete', 'completely' etc which is verified in the data set used here. The full list of such words is provided in the Appendix. Then similar to the earlier subsection, based on the probabilities computed from the dataset, we have

$p\_absolutist_{n/d} = p\_i\_absolutist_{n/n} = 0.016$ *if absolutist word is present, 0 otherwise*

$p\_absolutist_{d/d} = p\_i\_absolutist_{d/n} = 0.018$ *if absolutist word is present, 0 otherwise*

Behavioral Symptom 3: Use of positive and negative emotion words [17]

Occurrence of certain words which are based on emotion or medication are used as linguistic anchors to distinguish between normal and depressed individuals. Human behavior involves 'primary' emotions and Paul Ekman [19,20] described the six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) which Robert Plutchik [21] later expanded and grouped into four opposite sets (joy-sadness, anger-fear, trust-distrust, surprise-anticipation). While these emotions are characteristics prevalent in every individual, there are some emotions which tend to dominate in depressed individuals such as fear and sadness in Ekman's classifications.

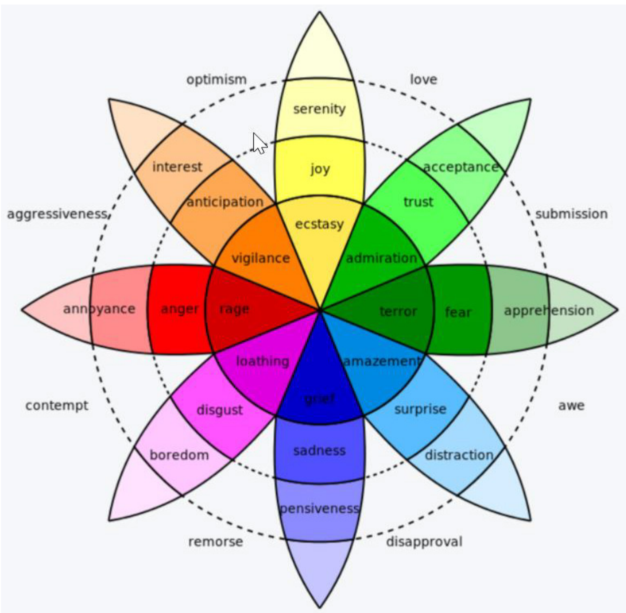

Figure 7: Emotions categorized by Plutchik [21] on a wheel with positive and negative emotions on opposite sides*.

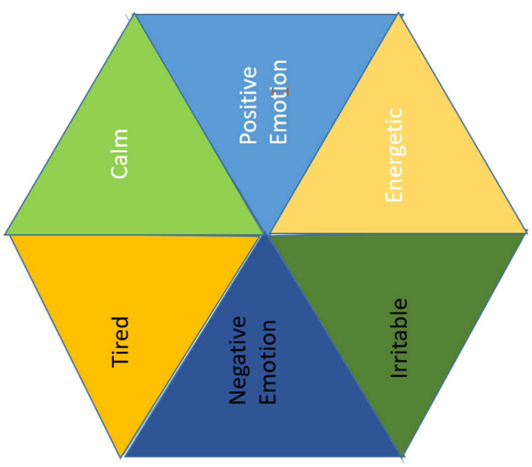

Figure 8: Relevant emotions used in computing the Markov model BSB weights.

To take advantage of such emotional markers (listed in Appendix), those relevant to the dataset used in this work are incorporated as transition probabilities.

$p\_pos\_emotion_{n/d}$, $p\_pos\_emotion_{n/n}$ = *1 if any of the tweet words are in positive emotion list*

$p\_neg\_emotion_{d/n}$, $p\_neg\_emotion_{d/d}$ = *1 if any of the tweet words are in negative emotion list*

Once again note that conditional probabilities are independent of the previous state and are stationary in that their values do not change over time. Once again, note that all these listed behavioral traits follow established medical symptoms [17,18].

Behavioral Symptom 4: Frequent use of negative phrases (limited here to "bigrams" or two consecutive words only)

$p\_bigram_{d/n}$, $p\_bigram_{d/d}$ = *1 if word is in bigram list*

Behavioral Symptom 5: Irritable - Get restless or more cranky than usual

$p\_irritable_{d/n}$, $p\_irritable_{d/d}$ = *1 if any of the tweet words are in the irritable list*

$p\_calm_{n/d}$, $p\_calm_{n/n}$ = *1 if any of the tweet words are in the calm list*

Behavioral Symptom 6: Less energetic, feel extremely tired or think more slowly. Daily routines and tasks may seem too hard to manage

$p\_tired_{d/n}$, $p\_tired_{d/d}$ = *1 if any of the tweet words are in the tired list*

$p\_energetic_{n/d}$, $p\_energetic_{n/n}$ = *1 if any of the tweet words are in the energetic list*

Behavioral Symptom 7: Discussions on the use of medication particularly related to mental health

$p\_medication_{d/n}$, $p\_medication_{d/d}$ = *1 if word is in depressed list*

Behavioral Symptom 8: Accompanied by co-occurrence of other related mental problems (also known as comorbidity)

$p\_comorbid_{d/n}$, $p\_comorbid_{d/d}$ = *1 if word is in comorbid list*

Behavioral Symptom 9: Accompanied by suicidal tendencies

$p\_suicidal_{d/n}$, $p\_suicidal_{d/d}$ = *1 if word is in depressed list*

"Associative semantic biomarkers" (ASB) probability computation

While many of the conditions described in the earlier section can be captured by appropriate linguistic markers, it is not exhaustive. The important point is that while the linguistic aides in the previous section do indeed help, it only captures some known and clear conditions and there are many traits which are fuzzy and hard to evaluate. Drawing inspiration from the field of astronomy where stars are used as navigational guides, certain words called "Anchors" which are indicative of the emotions in the particular data set are selected as "navigational" guides. These Anchors (specified in Appendix A) are chosen from words that are semantically representative of the dataset or the mental condition being evaluated. For every tweet, semantic closeness or similarity is evaluated to these "Anchors" in both the depressed and normal dataset. Note that only action descriptor words such as Verbs and Adjectives in the tweets are used in the evaluation.

Semantic closeness is derived by the "cosine similarity" (statistical terminology which computes "closeness" or distance between two words and is defined in Appendix B) which is a word distance measure ranging from -1 to 1 with a value of 1 indicating very semantically similar words and a small value indicating dissimilar words. To do so, first word embeddings are obtained using Word2Vec [11-12] (Glove [25] is another approach that is also viable) from which this metric can then be easily computed.

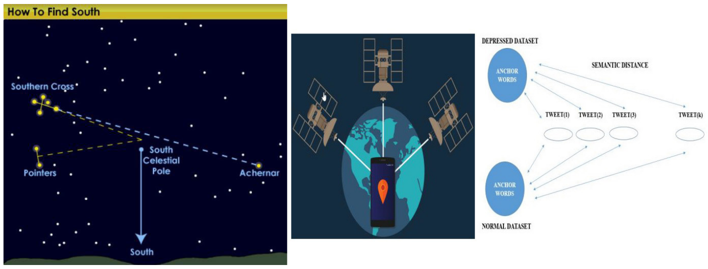

Figure 9: (a) Using Anchor stars to find direction (b) GPS for location identification (c) Using semantic similarity to "Anchor" words in respective datasets to identify closeness.

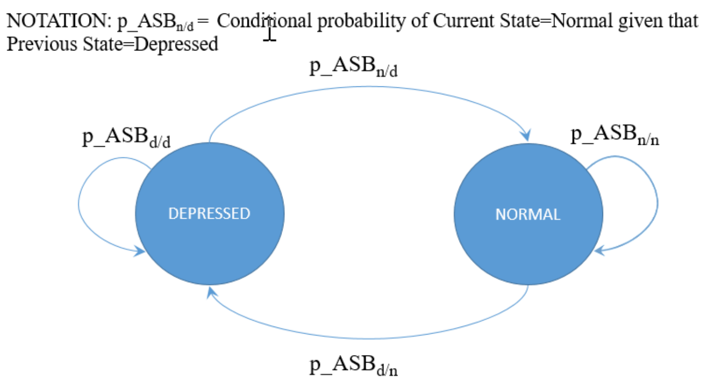

Figure 10: Markov model of "Associative semantic biomarkers" (ASB) transitions between depressed and normal states over time.

p_ASB$_{d/n}$ = *cosine distance between tweets and ASB_depressed_anchor_list in the normal data set*
p_ASB$_{d/d}$ = *cosine distance between tweets and ASB_depressed_anchor_list in the depressed data set*
p_ASB$_{n/n}$ = *cosine distance between tweets and ASB_normal_anchor_list in the normal data set*
p_ASB$_{n/d}$ = *cosine distance between tweet and ASB_normal_anchor_list in the depressed data set*

Note that the motivation here is to automatically compute the distance between every tweet to the depressed and normal dataset resulting in a self-guided approach to evaluating the mental state in tweets. The "Anchor" reference list for the depressed and normal dataset is provided in the Appendix.

Final transition probabilities

Now that all the component transition probabilities have been described, final expressions are obtained. Recall that $p_{d/d}(k)$ is a weighted sum of $p\_ASB_{d/d}(k)$ and $p\_BSB_{d/d}(k)$. Therefore, the most general form of constructing $p_{d/d}(k)$ is by a weighted sum of the individual probabilities. Hence,

$$p_{d/d}(k)= w_1*p\_i\_pronoun_{d/d}(k) + w_2*p\_absolutist_{d/d}(k)+ w_3*p\_neg\_emotion_{d/d}(k)+ w_4*p\_bigram_{d/d}(k) + w_5*p\_irritable_{d/d}(k) + w_6*p\_neg\_tired_{d/d}(k)+ w_7*p\_medication_{d/d}(k)+ w_8*p\_comorbid_{d/d}(k)+ w_9*p\_suicidal_{d/d}(k)+ w_{10}*p\_ASB_{d/d}(k) \rightarrow \textbf{Equation 6}$$

where $w_1+ w_2+ w_3+ w_4+ w_5+ w_6+ w_7+ w_8+ w_9+ w_{10}+ w_{11} = 1$

$$p_{d/n}(k)= w_1*p\_i\_pronoun_{d/n}(k) + w_2*p\_absolutist_{d/n}(k)+ w_3*p\_neg\_emotion_{d/n}(k)+ w_4*p\_bigram_{d/n}(k) + w_5*p\_irritable_{d/n}(k) + w_6*p\_tired_{d/n}(k)+ w_7*p\_medication_{d/n}(k)+ w_8*p\_comorbid_{d/n}(k)+ w_9*p\_suicidal_{d/n}(k)+ w_{10}*p\_ASB_{d/n}(k) \rightarrow \textbf{Equation 7}$$

$$p_{n/n}(k)= q_1*p\_i\_pronoun_{n/n}(k) + q_2*p\_absolutist_{n/n}(k)+ q_3*p\_pos\_emotion_{n/n}(k)+ q_4*p\_calm_{n/n}(k) + q_5*p\_energetic_{n/n}(k)+ q_6*p\_ASB_{n/n}(k) \rightarrow \textbf{Equation 8}$$

$$p_{n/d}(k)= q_1*p\_i\_pronoun_{n/d}(k) + q_2*p\_absolutist_{n/d}(k)+ q_3*p\_pos\_emotion_{n/d}(k)+ q_4*p\_calm_{n/d}(k) + q_5*p\_energetic_{n/d}(k)+ q_6*p\_ASB_{n/d}(k) \rightarrow \textbf{Equation 9}$$

where $q_1+ q_2+ q_3+ q_4+ q_5+ q_6 = 1$

If *w's* and *q's* were known, then the transition probabilities are easy to compute. The optimal individual weight computation is described in the next section.

Optimization of weights

In the above Equations 6-9, there are a number of weights that would require to be optimized. Based on the proportional value of the weights, the constituent probability takes on lesser or greater importance in the overall sum. To optimize, first the number of weights are reduced to keep the problem manageable in terms of computational complexity. Note that there is no reason to place a restriction on linear addition of weights. One could very easily train a Neural Net to provide an transition probability $p_{d/d}(k)$ which is not linearly dependent on its inputs. By reducing the number of weights to 4, the following simplification is obtained.

$$p_{d/d}(k)= w_1*p\_i\_pronoun_{d/d}(k) + w_2*p\_absolutist_{d/d}(k)+ w_3*p\_neg\_emotion_{d/d}(k)+ w_3*p\_bigram_{d/d}(k) + w_3*p\_irritable_{d/d}(k) + w_3*p\_neg\_tired_{d/d}(k)+ w_3*p\_medication_{d/d}(k)+ w_3*p\_comorbid_{d/d}(k)+ w_3*p\_suicidal_{d/d}(k)+ w_4*p\_ASB_{d/d}(k) \rightarrow \textbf{Equation 10}$$

where $w_1+ w_2+ w_3+ w_4 = 1$

Similarly,

$$p_{n/n}(k)= w_1*p\_i\_pronoun_{n/n}(k) + w_2*p\_absolutist_{n/n}(k)+ w_3*p\_pos\_emotion_{n/n}(k)+ w_3*p\_calm_{n/n}(k) + w_3*p\_energetic_{n/n}(k)+ w_4*p\_ASB_{n/n}(k) \rightarrow \textbf{Equation 11}$$

where $w_1+ w_2+ w_3+ w_4 = 1$

Weight optimization of w's is performed by a brute force search (though perhaps linear programming techniques can be used). To keep the complexity manageable, coarse discrete step sizes are used in the optimization. More specifically, the optimization identifying the value of the w's is defined as

$$\underset{w_1, w_2, w_3, w_4}{\textbf{MAX}} \ (\textit{Evaluation Precision of the Depressed data set})$$
$$\text{subject to the constraint } w_1+w_2+w_3+w_4 = 1 \quad \rightarrow \textbf{\textit{Equation 12}}$$

An Alternate Bayesian Classifier

Suppose *Mw* = Markov chain Mental state (set of Normal and Depressed state), ***Tw*** =Tweet vector comprising of all the tweets, from standard Baye's formulae

$P(Mw/Tw) = P(Tw/Mw)*P(Mw)/ P(Tw) \rightarrow$ ***Equation 13***

where as usual, *P(.)* represents the probability function. Then $P(Tw/Mw)=\Pi P(\prod_{j=1}^{n} P/Twj/Mw)$ assuming independence of tweets. Due to this assumption, this classifier is also called as "Naïve Bayes" [22].

Note that P(*Tw*) need not be explicitly computed as it is common in both the normal and depressed evaluation and hence can be dropped. The decision rule for choosing the depressed state "*Dw*" over the normal state *Nw*) is then *P(Dw/Tw) > P(Nw/Tw)*. By using Equation 13, this becomes (after dropping the common *P(Tw)* in the denominator)

$(\prod_{j=1}^{n} P(Tw_j/Dw)*P(Dw)) > (\prod_{j=1}^{n} P(Tw_j/Nw)* P(Nw)) \rightarrow$ ***Equation 14***

where all the probabilities are obtained from the training set. Now simplifying the above calculation using logarithms thereby converting the multiplications to additions (as the probabilities are small to begin with and hence multiplication is best avoided) results in

$\Sigma_{j=1}^{n} log(P(Tw_j/Dw))+log(P(Dw))\Sigma_{j=1}^{n} log(P(Tw_j/Nw))+ log(P(Nw)) \rightarrow$ ***Equation 15***

Finally, the classifier decision is then obtained as follows: Choose DepressedState (*Dw*) over NormalState (*Nw*) if left hand side of Equation 15 is greater than the right hand side else choose NormalState (*Nw*). Interestingly, notice that the Baye's classifier in Equation 15 corresponds to the top and bottom direct path in the trellis shown in Figure 8.

Results and Discussion

Metrics for evaluation of accuracy

To evaluate the accuracy of classifiers and evaluate the results, relevant metrics need to be first identified. The frequently used metrics in document retrieval are Precision, Recall and F1 score23. As an example, the "Precision" metric for the depressed dataset evaluation shows that among the predicted depressed individuals, how many were actually correctly diagnosed. "Recall" metric shows the other side of predictions. For example, it shows that of the total depressed individuals, how many were correctly predicted. "F1 score" captures both of these metrics into a single metric in the form of their harmonic mean.

Precision = True positive/(True positive + False positive)
Recall = True positive/(True positive + False negative)
F1 = 2*Precision*Recall / (Precision + Recall)

Table 1: Description of accuracy metrics.

Performance results

Results evaluated (as per the metrics of Table 1) on the "depressed" and "normal" are shown in Table 2. First, note that the dynamic programming technique is quite accurate. First, the weights w's were varied to determine whether the ASB's and BSB's are more relatively important in terms of the best performance. What is interesting is the specific weights which provided the best precision performance.

Compare the optimal weights for maximizing the precision shown in Table 2 with the performance obtained with a random weight as shown in Table 3. The increase in precision with optimal weights is not all that dramatic indicating that precision is not very sensitive to w values beyond a point. *p_ASB's* described earlier serves as an excellent marker for depression as indicated by the precision scores. While *p_BSB's* also serve as a very good marker in Table 3, (as indicated in the row where only *w3* is set to 1), using all the weights together does not increase the precision dramatically underlying the fact that the *p_ASB's*, *p_BSB's* have reasonable correlation between them. (whereby if one of them takes a certain value, most likely, so has the other). Note that the optimization of the *w's* here is performed as a linear addition and an AI algorithm would be able to account for non-linear techniques and this will be the subject of future work. Further, as more behavioral constructs are added into the model, the accuracy is only expected to further improve.

	Precision	Recall	F1	Weights w1, w2, w3, w4
Depressed	0.68	0.82	0.74	0, 0, 0, 1
Normal	0.65	0.47	0.55	0, 0, 0, 1

Table 2: Precision/Recall and Optimal weights for dynamic programming approach.

	Precision	Recall	F1	Weights w1, w2, w3, w4
Depressed	0.67	0.84	0.75	0, 0, 1, 0
Normal	0.67	0.43	0.53	0, 0, 1, 0
Depressed	0.62	0.84	0.72	0, 1/3,1/3.1/3
Normal	0.67	0.39	0.49	0, 1/3,1/3.1/3

Table 3: Precision/Recall for two random weight settings.

	Precision	Recall	F1	Weights w1, w2, w3, w4
Depressed	0.65	0.84	0.73	0, 0, 0, 1
Normal	0.67	0.42	0.51	0, 0, 0, 1

Table 4: Precision/Recall and Optimal weights for Bayesian classifier.

Once again demonstrate the viability of the model construction as well as the classifier in prediction depression. As more behavioral constructs are added into the model, the accuracy is only expected to further improve. Statisitical signifcance is studied using the McNemar [24] test on paired data. The p value the test yields is much lesser than 0.01! This means that the null hypothesis (random occurrence) is rejected. In a sense, this is not surprising given the large amount of data that is used for testing.

Conclusion and Future Directions

A new approach to identifying mental illness in social forums was developed in this work. Inspired by navigational guides that are used both in Astronomy and in GPS, symptomatic techniques were developed using semantic word distance of tweets from "Anchors" or guide words. Coupled with Linguistic word count, this creates a powerful foundation for developing a probabilistic model using semantic biomarkers to evaluate presence as well as severity of mental illness. A Markovian model tracking mental state evolution was then constructed. Subsequently, dynamic programming techniques were used to optimally decode the most likely state sequence over time and accurately decipher the mental state. This is the first known application of dynamic programming techniques and Markovian models to decipher the mental state and diagnose mental ailments.

Results demonstrate that the model can predict the onset of depression with reasonable accuracy while being open to include more behavioral conditions to enhance performance. For example, more emotional states can be incorporated in the model as probabilistic constructs based on recent psychiatric and medical research on mental afflictions. The Markovian model presented in this work has memory that relates the current state with the previous state. It would be worthwhile to compare the results of this model with AI based algorithms with both long and short term memory to find the best fit for mental ailment prediction. This will be the subject of future work.

While the primary mental illness evaluated in this work is "depression", the techniques extend in a straightforward manner to other mental afflictions too and will also be the study of future work. Results demonstrate that the Markov model can not only detect "depression" accurately but can also predict it well before self-reported diagnoses of the condition. When path metrics involving semantic biomarkers in mental health evolution exceed a threshold, "Alarm" conditions can be triggered thus providing a simple, cheap and accurate diagnosis tool for remedial action. Note that facial emotions, speech and other multimedia data can also be used to augment the semantic data and will be the subject of future study. While the work in this paper is based on social forums such as Twitter, any other social forum of platform such as Reddit, inter-personal communication can be equally applied serving as continuous real-time monitor against suicidal tendencies.

In summary, a novel scheme of using AI for "social good" by automatically detecting semantic bio-markers and deducing mental health condition has been described opening a number of promising topics for future study.

References

1. https://ourworldindata.org/mental-health
2. https://www.space.com/5849-navigating-stars.html
3. GPS Technology Tutorial & Basics. https://www.radio-electronics.com/info/satellite/gps/gps- technology-basics-tutorial.php.
4. Munmun De Choudhury, Michael Gamon, Scott Counts, et al. Predicting Depression via Social Media. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. 2013.
5. Glen Coppersmith, Mark Dredze, Craig Harman. Quantifying Mental Health Signals in Twitter. Associational for Computational Linguistics Workshop of Computational Linguistics and Clinical Psychology. 2014.
6. Margaret Mitchell, Kristy Hollingshead, Glen Coppersmith. Quantifying the language of schizophrenia in social media. Proceedings of the 2nd workshop on Computational linguistics and clinical psychology: From linguistic signal to clinical reality. 2015.
7. Gatis Mikelsons, Matthew Smith, Abhinav Mehrotra, et al. Towards Deep Learning Models for Psychological State Prediction using Smartphone Data: Challenges and Opportunities. 31st Conference on Neural Information Processing Systems, Long Beach, CA. 2017.
8. George Gkotsis, Anika Oellrich, Sumitra Velupillai, et al. Characterization of mental health conditions in social media using Informed learning. Scientific Reports. 2017; 7: 45141.
9. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: Our words, our selves. Annual review of psychology. 2003; 54: 547-577.
10. Twint A. Twitter scraping tool. https://github.com/twintproject/twint
11. Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space. International Conference on Learning Representations. 2013.
12. Gensim: Topic Modelling for humans. http://radimrehurek.com/gensim/
13. Laurens van der Maaten, Geoffrey Hinton. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008; 9: 2579-2605.
14. https://en.wikipedia.org/wiki/Tag_cloud
15. Daniel Jurafsky, James Martin. Speech and Language Processing. 2nd Edition, Chapter 9 is on Hidden Markov Models, Prentice Hall. 2009.
16. Andrew Viterbi. The Viterbi Algorithm Demystified. https://viterbischool.usc.edu/news/2017/03/viterbi-algorithm-demystified/
17. https://www.webmd.com/depression/guide/depression-symptoms-causes.
18. People with depression use language differently and how to spot it. http://theconversation.com/people-with-depression-use-language-differently-heres-how-to-spot- it-90877.
19. https://www.psychologytoday.com/us/blog/hide-and-seek/201601/what-are-basic-emotions
20. https://www.paulekman.com/wp-content/uploads/2013/07/Basic-Emotions.pdf

21. Plutchik, Robert. Emotion: Theory, research, and experience: Vol. 1. Theories of emotion. New York: Academic. 1980.
22. Naïve Bayes: Predicting Movie Review Sentiment. https://www.dataquest.io/blog/naive- bayes-tutorial.
23. What are precision, recall and F1? https://simplyml.com/what-are-precision-recall-and- f1/.
24. McNemar's test using SPSS statistics. ttps://statistics.laerd.com/spsstutorials/mcnemars- test-using-spss-statistics.php.
25. Pennington J, Socerh R, Manning C. Glove: Global vectors for word representation. Empirical methods in natural language processing (EMNLP). 1532-1543.

APPENDIX

APPENDIX A: Definition of the various word lists

1. absolutist_list = ['absolutely', 'all', 'always', 'complete', 'completely', 'constant', 'constantly', 'definitely', 'entire', 'ever', 'every', 'everyone', 'everything', 'full', 'must', 'never', 'nothing', 'totally', 'whole']
2. i_pronoun_list = ['i', 'my', 'me', 'i've', 'I'm', 'myself']
3. pos_emotion_list = ['succeed', 'upbeat' 'confident', 'friendly', 'initiative', 'enthusiastic', 'determined', 'confident', 'optimistic', 'pleased', 'keen', 'eager', 'amazing', 'wonderful', 'happy', 'excellent', 'excited', 'delighted', 'thrilled' ,'outstanding']
4. neg_emotion_list = ['miserable', 'hopeless', 'pessimistic', 'die', 'panic', 'panicking', 'crying', 'desperate', 'attacks', 'suffer', 'overwhelming', 'worthless', 'depressed', 'overwhelmed', 'pain', 'misery', 'anxious', 'lonely', 'suffering', 'hopeless', 'sadness', 'unhappy', 'sad', 'help', 'suffering', 'hurts', 'awful']
5. irritable_list = ['irritated', 'restless', 'cranky', 'irritability']
6. calm_list = ['calm', 'relaxed', 'relaxing', 'relaxation', 'chill']
7. energetic_list = ['energetic', 'lively', 'dynamic', 'active', 'exercise', 'workout']
8. tired_list = ['tired', 'exhausted', 'exhaustion', 'weary', 'drained', 'fatigued', 'fatigue', 'tiredness', 'insomnia']
9. medication_list = ['anti-depressant', 'antidepressant', 'citalopram', 'saphris', 'klonopin', 'tramadol', 'resulti', 'paxil', 'neurontin' 'fluoxetine', 'depressant', 'trazodone', 'effexor', 'klonopin', 'cymbalta', 'metformin', 'cephalexin', 'zoloft', 'antipsychotics', 'finasteride', 'ativan']
10. comorbid_list = ['depression', 'schizophrenia', 'ptsd', 'bipolar', 'ocd', 'psychosis', 'hoshimotos', 'ocd', 'disorder', 'addiction', 'psychiatric', 'counseling', 'therapy', 'phobia', 'psychologist', 'agoraphobia', 'asperger', 'aspergers', 'chronic', 'severe', 'psychologists', 'shrink']
11. suicidal_list = ['suicidal', 'suicide']
12. bigram_list = ['hate myself', 'kill myself', 'life sucks', 'feel down', 'really down', 'social anxiety', 'need help']
13. ASB_depressed_anchor_list = ['pleased', 'excited', 'delighted', 'thrilled']
14. ASB_normal_anchor_list = ['panic', 'overwhelmed', 'anxious', 'lonely']

APPENDIX B: Cosine similarity of two n-dimensional vectors A and B is defined as:

$$\frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}=\frac{\sum_{i=1}^{n}A_iB_i}{\sqrt{\sum_{i=1}^{n}A_i^2}\sqrt{\sum_{i=1}^{n}B_i^2}}$$

which simply is the cosine of the angle between A and B.