

## EMOTIONS BASED PATTERNS OF MENTAL DISORDERS ON SOCIAL MEDIA –A CASE OF ANOREXIA AND DEPRESSION

<sup>1</sup>A Sowmya,<sup>2</sup>CH Prudvini,<sup>3</sup>P Avaniketh,<sup>4</sup>M Sai Venkat

<sup>1,2,3</sup>Assistant Professor,<sup>4</sup>Student

Department of CSE

Christu Jyothi Institute of Technology & Science, Colombo Nagar, Telangana

### ABSTRACT :

Millions of people around the world are affected by one or more mental disorders that interfere in their thinking and behavior. A timely detection of these issues is challenging but crucial, since it could open the possibility to offer help to people before the illness gets worse. One alternative to accomplish this is to monitor how people express themselves, that is for example what and how they write, or even a step further, what emotions they express in their social media communications. In this study, we analyze two computational representations that aim to model the presence and changes of the emotions expressed by social media users. In our evaluation we use two recent public data sets for two important mental disorders: Depression and Anorexia. The obtained results suggest that the presence and variability of emotions, captured by the proposed representations, allow to highlight important information about social media users suffering from depression or anorexia. Furthermore, the fusion of both representations can boost the performance, equalling the best reported approach for depression and barely behind the top performer for anorexia by only 1%. Moreover, these representations open the possibility to add some interpretability to the results.

### I. INTRODUCTION :

A mental disorder causes different interferences in the thinking and behavior of the affected person [1]. These interferences could vary from mild to severe, and could result in an inability to live routines in daily life and ordinary demands [2]. Common mental disorders such as depression and anorexia affect millions of people around the world. They may be related to a single incident causing excessive stress on the person or by a series of different stressful events. It is also well known that mental disorders tend to increase in countries experiencing generalized violence or recurrent natural disasters. For example, in 2018 a study of mental disorders in Mexico revealed that 17% of its population has at least one mental disorder and one in four will suffer a mental disorder at least once in their life [3]. In another vein, in the modern world, we take for granted that social life could be experienced either in the physical world or in a virtual world created by social media platforms like Facebook, Twitter, Reddit, or similar platforms. This reality presents some challenges, but also great opportunities which, if properly addressed, could contribute to the understanding of what and how we communicate. In this regard, the goal of this study is to analyze, via the automatic identification of emotional patterns, social media documents with the purpose of detecting the presence of signs of depression or anorexia in the population of that area [4]–[6]. Previous works have addressed the analysis of emotions of social media users by paying attention to their contrast and tone. They have mainly applied this analysis to predict users' age and gender as well as a range of sensitive personal attributes including sexual orientation, religion, political

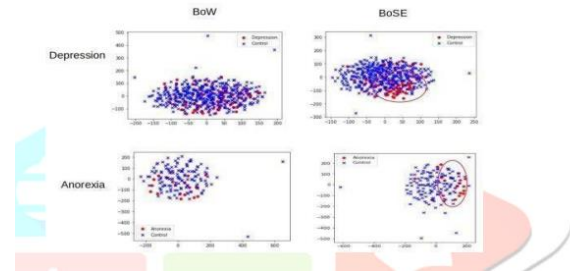
orientation [7], [8], income [9], and personality traits [10], [11]. According to these studies, the analysis of emotions in social media allows capturing important information related to users. This information presents an opportunity for us to extend the use of emotions in the detection of depression and anorexia in social media.

Former studies focused on the detection of depression and anorexia have mainly considered linguistic and sentiment analysis [12]–[14]. Note that the use of sentiments, i.e. polarity, was the preamble for the later use of emotions for the same task [15]. This line of thought exposed the potential of using emotions as features, such as "anger", "surprise" or "joy", instead of linguistic features or general sentiments like positive and negative. In this direction, in our previous work [16], we introduced a novel representation that was built using information extracted from emotions lexicons combined with word embeddings as a way to represent the information contained in users' documents. Then, using a clustering algorithm, we created sub-groups of emotions, that conveniently we named as sub-emotions. These discovered sub-emotions provided a more flexible and fine-grained representation of users and a better performance for the detection of depression. In a few words, the idea behind this representation was to capture the presence of sub-emotions in users' posts. The intuition of our approach is that users suffering from depression would show a distribution of emotions different from healthy users. Motivated by the encouraging results of the representation based on sub-emotions, in this study we give a more complete treatment of the method. In particular, we propose a new

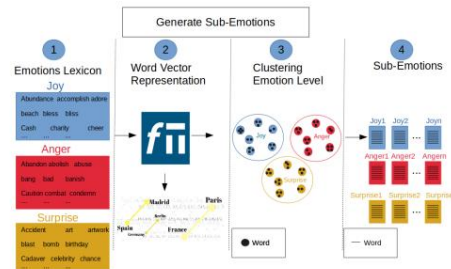
representation that not only captures the presence of sub-emotions, but also models their changes over time. The intuition is to model emotional fluctuations that users with mental disorders could continuously present. This temporal information is later integrated to enrich the original approach. That is, we build a fusion of both representations, that at the end attains very competitive results, practically equal to those of the state-of-the-art approaches. Finally, we envision how these two representations can be applied beyond detecting depression to also detect other important mental disorder such as anorexia. Using this new representation we contrast emotional patterns between the two disorders, possibly finding what could be described as their emotional “silhouette”

**PROBLEM STATEMENT :**

model emotional fluctuations that users with mental disorders could continuously present. This temporal information is later integrated to enrich the original approach. That is, we build a fusion of both representations, that at the end attains very competitive results, practically equal to those of the state-of-the-art approaches. Finally, we envision how these two representations can be applied beyond detecting depression to also detect other important mental disorder such as anorexia. Using this new representation we contrast emotional patterns between the two disorders, possibly finding what could be described as their emotional “silhouette” combined them with traditional approaches to leverage their results. In order to analyze the obtained results, we plotted the users in a plane using both the BoW and the BoSE representations. To generate these visualizations, we used the T-distributed Stochastic Neighbor Embedding (T-SNE) algorithm [5], which is a nonlinear dimensionality reduction technique well-suited for plotting high-dimensional spaces in a low dimensional space. For this analysis we used the vector representation of 3000/1500 features obtained using tf-idf with chi2 distribution for both BoSE and BoW. Figure 2 offers an interesting perspective of the advantage of using BoSE over BoW to allow the classifier to build a better classification function. We even analyzed the boundary cases and found similar distribution in the sub-emotions, this could be due to the similarity in the topics captured by the sub-emotions that users posted and shared.



**MODEL DIAGRAM :**



**II. LITERATURE SURVEY**

**TITLE :**

**The kerberos network authentication service**

**AUTHORS :**

C. Neuman, S. Hartman, K. Raeburn

**ABSTRACT:**

When using authentication based on cryptography, an attacker listening to the network gains no information that would enable it to falsely claim another's identity. Kerberos is the most commonly used example of this type of authentication technology. The authors concentrate on authentication for real-time, interactive services that are offered on computer networks. They use the term real-time loosely to mean that a client process is waiting for a response to a query or command so that it can display the results to the user, or otherwise continue performing its intended function. This class of services includes remote login, file system reads and writes, and information retrieval for applications like Mosaic.

**TITLE :**

**IDFusion: An open architecture for Kerberos based authorization**

**AUTHORS :**

G. Wettstein, J. Grosen, and E. Rodriguez

**ABSTRACT:**

Since its initial development Kerberos has evolved to become the widely accepted system for implementing centralized authentication services. During this time the Lightweight Directory Access Protocol (LDAP) has become the accepted method for the centralized distribution of identity information. Organizations increasingly deploy both infra-structural components in

order to support management of distributed information delivery systems.

uring this evolution no standardized scheme for authorization has emerged. Industry consensus suggests that LDAP is the protocol of choice for storing extended information needed to make authorization decisions. Despite this consensus no standardized scheme has evolved for implementing directory based authorization.

This paper discusses a strategy for using the symmetric key management facilities of Kerberos to implement directory based authorization. The system is architected to provide inherent security in the event of a directory compromise. The system offers the management advantages of role based access systems while providing the option for fine grained authorization control.

The identity based authorization model uses a service oriented approach to managing authorization. As such it is consistent with and supportive of the trend toward services oriented application architectures.

**TITLE :****A nonce-based protocol for multiple authentications****AUTHORS :**

A. Kehne, J. Schonwalder, and H. Langendorfer

**ABSTRACT:**

MIT's Project Athena, is based on the Needham and Schroeder protocol. Timestamps depending on reliable synchronized clocks are used to guarantee the freshness of messages. As an improvement, we present a nonce-based protocol offering the same features as Kerberos. We generate a ticket in an initial message exchange which includes a generalized timestamp. Checking this generalized timestamp is left to the principal who created it. Consequently we do not need synchronized clocks. Our protocol has the property of using a minimal number of messages to establish an authenticated session key.

**III. SYSTEM ANALYSIS****EXISTING SYSTEM :**

Depression is a mental health disorder characterized by persistent loss of interest in activities, which can cause significant difficulties in everyday life [1], [17]. Studies focusing on the automatic detection of this disorder have used crowdsourcing as their main strategy to collect data from users who expressly have reported being diagnosed with clinical depression [18], [19]. Among these studies, the most popular approach considers words and word n-grams as features and employs traditional classification algorithms [13], [20], [21]. The main idea is to capture the most frequent words used by individuals suffering from depression and compare them against the most frequent words used by healthy users. This approach suffers because there is usually a high overlap in the vocabulary of users with and without depression.

Another group of works used a LIWC-based representation [22], aiming to represent users' posts by a set of psychologically meaningful categories like social relationships, thinking styles, or individual differences [18], [23]. These works have allowed a better characterization of the mental disorder conditions, nevertheless, they have only obtained moderately better results than using only the words. Recent works have considered ensemble approaches, which combine word and LIWC based representations with deep neural models such as LSTM and CNN networks [24], [25]. For example, in [25], [26], the combination of these models with features like the frequencies of words, user-level linguistic metadata, and neural word embeddings offered the best-reported result in the eRisk2018 shared task on depression detection [27]. These works show that in social media texts exist useful information to determine if a person suffers from depression, but the results are sometimes hard to interpret. This is an important limitation since these types of tools are naturally aimed to support health professionals and not to take the final decisions. In [28] [29], the authors conduct studies to tackle this problem. They characterize users affected by mental disorders and provide methods for visualizing the data in order to provide useful insights to psychologists.

**DISADVANTAGES OF EXISTING SYSTEM :**

- 1) Less accuracy
- 2)low Efficiency

**PROPOSED SYSTEM :**

To offer a glimpse of the data sets, we demonstrate some examples of posts from the different classes of users. Our goal is to show that users who effect from a mental illness as well as controlusers share their experiences and personal feelings about them, which for both can be positive andnegative, making their detection a great challenge.

Depression :

- 1) After coming home from a road trip with a group of friends to celebrate my birthday.
- 2) Sometimes I can't help but think that they will be so much better off without me, and they knowthat they would be happier without me.

Anorexia

- 1) I'm happy to hear that you're okay with realizing you'll be on anti-depressants for the rest of your life.

- 2) My coach looked over at me then muttered; "It's a shame. If she wasn't so BIG I'd consider herfor the team.

Control

- 1) Nice job; it's not always easy with the clouds. I love the colors of those waters with the glacialmoraine. Beautiful image.

2) It was difficult, I do not expect it to be well-received here, but even if one person find it useful, i will continue.

The steps of proposed system are as follows:

Preprocessing: The texts were normalized by lowercasing all words and eliminating special characters like URLs, emoticons, and #; the stop words were kept. Then, the preprocessed texts were hidden using the created sub-emotions.

Classification:

The main goal is to divide users into one of the two classes (Depressed / Control or Anorexia / Control). The BoSE (Bag of Sub Emotions) approach consists of three main steps: first, a set of fine-grained emotions are tested using unsupervised learning from a lexical resource that contains words associated to different emotions and sentiments, this is achieved using a clustering technique that separates the distribution of each emotion  $e$  in  $K$  sub-groups also known as sub-emotions. Second, the fine-grained emotions are used to represent the documents, each word is masked or substitute by its closer sub-emotion, and each document is represented by a frequency histogram of their sub-emotions. Third, the histogram representation is used to train a classification model that predicts the depression label. On the other hand, the involvement of a dynamic analysis over the sub-emotions, called  $\Delta$ -BoSE, which helps to improve in finding the users having signs of anorexia and depression.

**ADVANTAGES OF PROPOSED SYSTEM :**

- 1) High accuracy
- 2) High efficiency

**SYSTEM REQUIREMENTS**

• **HARDWARE & SOFTWARE REQUIREMENTS:**

➤ **HARDWARE REQUIRMENTS :**

- ◆ System : Pentium IV 2.4 GHz.
- ◆ Hard Disk : 40 GB.
- ◆ Ram : 512 MB.

➤ **SOFTWARE REQUIRMENTS :**

- ◆ Technology : Java 2 Standard Edition, JDBC
- ◆ Web Server : Tomcat 7.0
- ◆ Client Side Technologies : HTML, CSS, JavaScript
- ◆ Server Side Technologies : Servlets, JSP
- ◆ Data Base Server : MySQL
- ◆ Editor : Netbeans 8.1

- ◆ Operating System : Microsoft Windows, Linux or Mac any version

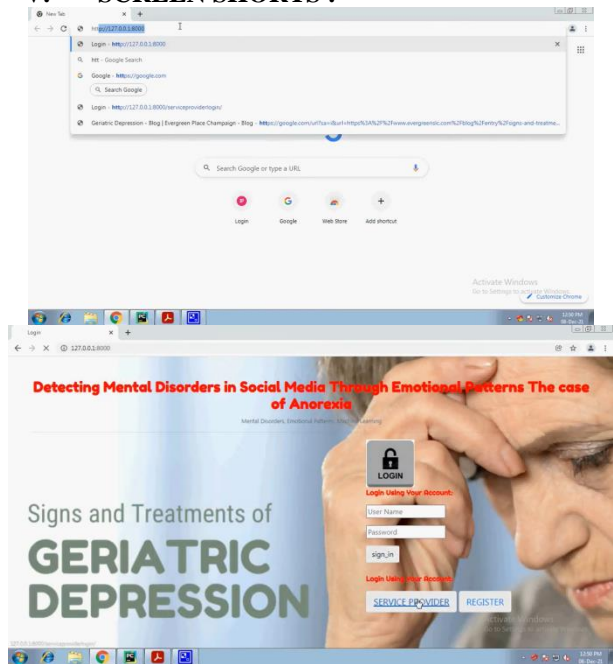
**IV. IMPLEMENTATION:**

**MODULES:**

Upload currency image : use this button to get upload an currency image.

Run template matching currency recognition : use this module to get Run template matching currency recognition.

**V. SCREEN SHORTS :**





results. Finally, the capability to model the emotional behavior of users using their social media data presents an opportunity for future wellness facilitating technologies. This kind of technology can serve as warning systems that provide wide-area analysis and information related to a mental disorder respecting user privacy. This information could include the presence of mental disorders in certain areas, and the authorities could decide to create professional assistance or emotional support, that the users will decide whether to take or not. We believe that it is important to mention when we analyze social media content, we may have concerns regarding individual privacy or certain ethical considerations. These concerns appear due to the usage of information that could be sensitive, given the personal behavior and emotional health of the users. The experiments and usage of this data are for research and analysis only, an

#### REFERENCES

- [1] S. Volkova and Y. Bachrach, "Inferring perceived demographics from user emotional tone and user environment emotional contrast," Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2016.
- [2] D. Ram´irez-Cifuentes and A. Freire, "Upf's participation at the clef erisk 2018: Early risk prediction on the internet," Proceedings of the 9th International Conference of the CLEF Association, CLEF 2018, Avignon, France, 2018.
- [3] H. Schwartz, J. Eichstaedt, M. Kern, G. Park, M. Sap, D. Stillwell, M. Kosinski, and L. Ungar, "Towards assessing changes in degree of depression through facebook," In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2014.
- [4] G. Coopersmith, M. Dredze, and C. Harman, "Quantifying mental health signals in twitter," Workshop on Computational Linguistics and Clinical Psychology, 2014.
- [5] C. Xuetong, D. Martin, W. Thomas, and E. Suzanne, "What about mood swings? identifying depression on twitter with temporal measures of emotions," Companion Proceedings of the The Web Conference 2018, International World Wide Web Conferences Steering Committee, pp. 1653–1660, 2018.
- [6] M. Aragon, A. L´opez-Monroy, L. Gonz´alez-Gurrola, and M. Montes-´y Gomez, "Detecting depression in social media using fine-grained ´emotions," Proceedings of the 2019 Conference of the North American Chapter of the Association for

Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019.

[7] C. Mathers and D. Loncar, "Projections of global mortality and burden of disease from 2002 to 2030," PLOS Medicine, Public Library of Science, 2006.

[8] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," In Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, 2013.

[9] M. De Choudhury, S. Counts, and E. Horvitz, "Social media as a measurement tool of depression in populations.," In Proceedings of the 5th Annual ACM Web Science Conference, 2013.

[10] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh, and H. Ohsaki, "Recognizing depression from twitter activity," In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015.