New Approach for Knowledge-Based Recommendation System that includes Sentiment Analysis and Machine Learning

Shradha Panchal, Shital Y. Gaikwad

ME 2nd Year Student, Department of Computer Science and Engineering, Matoshri Pratishthan Group of Institutions, Nanded, Maharastra, India

Asst. Professor, Department of Computer Science and Engineering, Matoshri Pratishthan Group of Institutions, Nanded, Maharastra, India

ABSTRACT: Online social networks provide relevant information on users' opinions and posts on various topics. So applications, such as monitoring and detection systems can collect and analyse this data. This paper presents a knowledge-based system, which includes an emotional health monitoring system to detect users with possible psychological disorders specially depression and stress. Symptoms Of these psychological disorder are usually observed passively. In this situation, author argue that online social behaviour extraction offers an opportunity to actively identifypsychological disorder at an early stage. It is difficult to identify the disorder because the psychological factors considered in standard diagnostic criteria questionnaire cannot be observed by the registers of online social activities. Our approach, New and innovative for the practice ofpsychological disorder detection, it does so do not trust the self-disclosure of those psychological factors through the questionnaires. Instead, propose a machine learning technique that is detection of psychological disorder in social networks which exploits the features extracted from social network data for identify with precision possible cases of disorder detection. We perform an analysis of the characteristics and we also apply machine learning in large-scale data sets and analyse features of the two types of psychological disorders.

KEYWORDS: Sentiment analysis, knowledge personalization and customization, detection system, social networks, machine learning.

I. INTRODUCTION

The number of active users of online social networks (OSN) it has grown considerably. This maximum the number of users, in OSN, is mainly due to the increase of number of mobile devices, such as smartphones and tablets, connected to the Internet Currently, the OSN has become a rich man and universal means to express opinions, feelings and reflect each user's bad habits or wellness practices. In recent years, the analysis of messages published in OSN has been used by many applications in the field of health informatics The feelings and emotions expressed in the messages published in OSN, provides clues on various aspects of user behaviour; for example, sentences containing words with negative meaning it can indicate sadness, stress or dissatisfaction.

In our system, we find that user's depression and stress is closely related to that of his/her friends in social media, and we employ a large-scale dataset from real-world social platforms to systematically study the correlation of users' psychological disorder states and social interactions. We first define a set of psychological disorder-related textual, visual, and social attributes from various aspects. Fast pace of life, progressively and more individuals are feeling stressed. Though disorder itself is non-clinical and common in our life, excessive and chronic disorder can be rather harmful to people's physical and mental health. Users' social interactions on social networks contain useful cues for stress and depression detection. Social psychological studies have made two interesting observations. The first is mood contagion: a bad mood can be transferred from one person to another during social interaction. The second Social Interaction: people are known to social interaction of user. The advancement of social networks like Twitter, Facebook and Sina Weibo2, an ever increasing number of people will share their every day events and moods, and interact with friends through the social networks. We

can classify using machine learning framework. Due to leverage both Facebook post content attributes and social interactions to enhance stress and depression detection.

II. RELATED WORK

Literature survey is the most important step in any kind of research. Before start developing we need to study the previous papers of our domain which we are working and on the basis of study we can predict or generate the drawback and start working with the reference of previous papers.

In this section, we briefly review the related work on health monitoring system and their different techniques.

In this paper [1], we propose an integrated social media content analysis platform that leverages three levels of features, i.e., user-generated content, social graph connections, and user profile activities, to analyse and detect anomalous behaviours that deviate significantly from the norm in large-scale social networks. Several types of analyses have been conducted for a better understanding of the different user behaviours in the detection of highly adaptive malicious users. We attempted a novel approach regarding the process of data extraction and classification to contextualize large-scale networks in a proper manner. We also collected a significant number of user profiles from Twitter and YouTube, along with around 13 million channel activities. Extensive evaluations were conducted on real-world datasets of user activities for both social networks. The evaluation results show the effectiveness and utility of the proposed approach.

This paper [2], presents a music recommendation system based on a sentiment intensity metric, named enhanced Sentiment Metric (eSM) that is the association of a lexicon-based sentiment metric with a correction factor based on the user's profile. This correction factor is discovered by means of subjective tests, conducted in a laboratory environment. Based on the experimental results, the correction factor is formulated and used to adjust the final sentiment intensity. The users' sentiments are extracted from sentences posted on social networks and the music recommendation system is performed through a framework of low complexity for mobile devices, which suggests songs based on the current user's sentiment intensity.

This paper [3], presents a monitoring solution for users with potential psychological disturbances specially stress. The phrase extracted from social networks are filtered by mood and scored using a sentiment analysis that considered such as gender and age.

In this paper [4], we find that users stress state is closely related to that of his/her friends in social media, and we employ a large-scale dataset from real-world social platforms to systematically study the correlation of users' stress states and social interactions. We first define a set of stress-related textual, visual, and social attributes from various aspects, and then propose a novel hybrid model - a factor graph model combined with Convolutional Neural Network to leverage tweet content and social interaction information for stress detection

This paper [5] , presents Mental scatters frequently happen in mixes, for instance a patient with a nervousness issue can likewise create gloom. This attending psychological wellness condition gives consideration regarding our work in the characterization of online networks with an enthusiasm for sorrow. For this, we followed an expansive collection of 620,000. Distributions made by 80,000 clients in 247 online networks. We have the psycho-etymological topics and attributes communicated in. The productions, utilizing them as contribution for our model. Following a vehicle Technique of learning, we have defined a joint displaying. System for existing together characterizations identified with psychological wellness online network of these qualities. At last, we perform experimental approval of the model in the informational index drawn where our model surpasses the most recent vanguard essential lines.

In this paper [6], we will probably construct prescient models that misuse them Language and standards of conduct, utilized particularly in the social circle .normal, to decide whether a client experiences two instances of mental issue. These prescient models are conceivable utilizing another information gathering process, begat as an intuitive. Publicly supporting, which encourages you gather all the more rapidly and dependably. Persistent informational index. Our examinations recommend that mining explicit phonetic models and qualities of social association of Reliable patient informational indexes can contribute altogether to encourage investigation and recognition of mental issue.

PC programs [7] It ought not be in the business to choose which questions are deserving of study. Despite the fact that

Hessians that are not non-feasible are at times indications of confused and negligible inquiries insufficient models, or estimators, likewise happen every now and again when data about the amounts of intrigue exist in the information through the likelihood work. The creators clarify the issue in detail and present two starter proposition on how manage non-invertible hessians without changing the inquiry.

In this work [8], we incorporate both the extraction of noteworthy issues. Furthermore, sifting messages by means of Twitter. We build up a transmission calculation for a recurrence grouping of the archive Tables; our calculation permits ongoing observing of the initial 10. Points of about 25% of all Twitter posts, while naturally filtering of uproarious and good for nothing subjects. We apply our own proposed transmission calculation for the stream of Japanese and Twitter effectively show that, contrasted with other non-negative online Matrix factorization techniques, our structure monitors this present reality. Occasions with high exactness as far as perplexity and Eliminates immaterial points.

In this work [9], we examine the development of exercises among clients. In the interpersonal organization of Facebook to catch this idea. We find that joins in the movement organize will in general go back and forth. Rapidly after some time, and the quality of the bonds shows a diminish the diminishing propensity of the movement as an association of an informal community hundreds of years. For instance, just 30% of Facebook client sets communicate. Continually starting with multi month then onto the next. It is intriguing to take note of that we additionally locate this, regardless of whether the associations of the action organize numerous properties of diagram hypothesis change quickly after some time. The system of exercises stays unaltered.

This paper [10] presents two contributions. First, it proposes the method SentiHealth to detect the mood of cancer patients that are also users of communities ofpatients in online social networks. Second, it presents an instantiated tool from themethod, called SentiHealth-Cancer (SHC-pt), dedicated to automatically analyse postsin communities of cancer patients, based on SentiHealth. This context-tailored tooloutperformed other general-purpose sentiment analysis tools at least in the cancer context. This suggests that the SentiHealth method could be instantiated as otherdisease-based tools during future works, for instance SentiHealth-HIV, SentiHealth-Strokeand SentiHealth-Sclerosis.

III. **EXISTING APPROACH**

A lot of work has been done in this field thanks to its extensive use and applications. This section mentions some of the approaches that have been implemented to achieve the same purpose. These works are mainly differentiated from the algorithm for psychological disorders detection systems.

Existing jobs have shown that social networks can be used for health care and, in particular, for the detection of mental disorders. There are some limitations in Facebook content based on the detection of mental disorders. Users do not always express their stressful states directly in the Facebook post. Although no disturbance of the publication itself is revealed, from the interactive follow-up comments made by the user and his friends, we can discover that the user is really stressed at work. Therefore, it is sufficient to trust the contents of a user's Facebook post for the detection of mental disorders. Users with high psychological mental disorders may exhibit little activity on social networks.

IV. **PROPOSED APPROACH**

In the proposed systemic approach, we formulate the task as a classification problem to detect two types of detection of psychological disorders in social networks using the machine learning framework:

  i.   Stress
  ii.   Depression

An innovative solution to monitor and detect potentialusers with emotional disorders, according to the classificationof sentences with depressed or stressed content.A machine learning is used for representation at the character level and for the recognition of the extent of the disturbance.
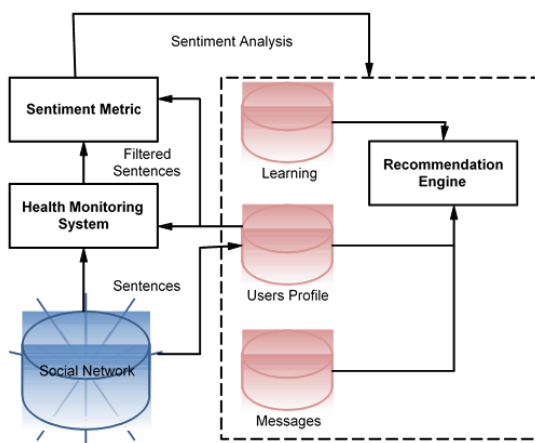
System Diagram:

Fig 1. System Architecture

PROPOSED ALGORITHM

Naive Bayes
Steps:

- Given training dataset D which consists of documents belonging to different class say Class A and Class B
- Calculate the prior probability of class A=number of objects of class A/total number of objects
    - Calculate the prior probability of class B=number of objects of class B/total number of objects
- Find NI, the total no of frequency of each class
    - Na=the total no of frequency of class A
    - Nb=the total no of frequency of class B
- Find conditional probability of keyword occurrence given a class:
    - P (value 1/Class A) =count/ni (A)
    - P (value 1/Class B) =count/ni (B)
    - P (value 2/Class A) =count/ni (A)
    - P (value 2/Class B) =count/ni (B)
    - …………………………………..
    - …………………………………..
    - …………………………………..
    - P (value n/Class B) =count/ni (B)
- Avoid zero frequency problems by applying uniform distribution
- Classify Document C based on the probability p(C/W)
- Find P (A/W) =P (A)*P (value 1/Class A)* P (value 2/Class A)……. P(value n /Class A)
- Find P (B/W) =P (B)*P (value 1/Class B)* P(value 2/Class B)……. P(value n /Class B)
- Assign document to class that has higher probability.

V. EXPERIMENTAL RESULT

In experimental results, we evaluate the proposed system with real time social networking posts dataset. A user study with no of peoples is conducted to evaluate the accuracy of system and analyze the detected mental disorder type's i.e. net compulsion users, cyber relationship users, information overload users using OSN features and personal features.
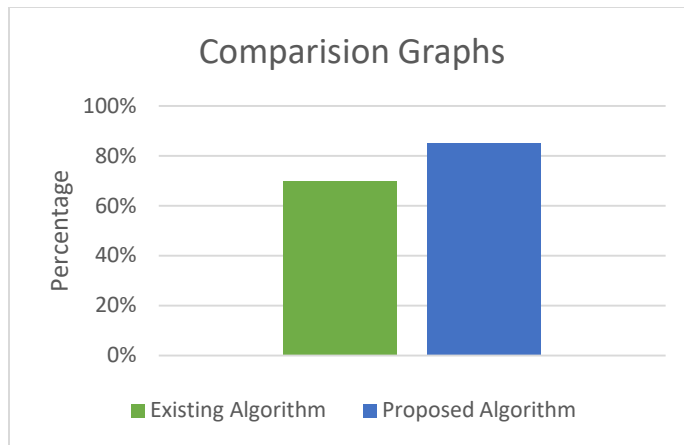
A. *COMPARISON GRAPH*

Fig2. Graph

B. *COMPARISON TABLE:*

Sr.No	Existing Algorithm (TSVM)	Proposed Algorithm(NB)
1	65%	86%

Table 1.comparative result

VI.CONCLUSION

In this paper, automatically identifying potential online users with depression and stress is threatening people's health. It is not trivial to detect the disorder in time for proactive care. Therefore, we present a framework to detect the states of stress and depression of users from the monthly data of users' social networks, exploiting the content of Facebook publications and the social interactions of users. Using the data of the social networks of the real world as a basis, we study the correlation between the states of psychological disorder of users and their social interaction behavior, we recommend the user for health precautions to send by mail for user interaction

REFERENCES

[1] M. Al-Qurishi, M. S. Hossain, M. Alrubaian, S. M. M. Rahman, and A. Alamri, "Leveraging analysis of user behavior to identify malicious activities in large-scale social networks," IEEE Transactions on Industrial Informatics, vol. 14, no. 2, pp. 799–813, Feb 2018.

[2] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Music recommendation system based on user's sentiments extracted from social networks," IEEE Transactions on Consumer Electronics, vol. 61, no. 3, pp. 359–367, Oct 2015.

[3] R. Rosa, D. Rodr, G. Schwartz, I. de Campos Ribeiro, G. Bressan et al., "Monitoring system for potential users with depression using sentiment analysis," in 2016 IEEE International Conference on Consumer Electronics (ICCE). Sao Paulo, Brazil: IEEE, Jan 2016, pp. 381–382.

[4] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T. S. Chua, "Detecting stress based on social interactions in social networks," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 9, pp. 1820–1833, Sept 2017.

[5] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh. A framework for classifying online mental health-related communities with an interest in depression. IEEE Journal of Biomedical and HealthInformatics, 2016.

[6] Chun-Hao Chang, Elvis Saravia, Yi-Shin Chen "Subconscious Crowdsourcing: A Feasible Data Collection Mechanism for Mental Disorder Detection on Social Media" 2016.

[7] J. Gill and G. King. What to do when you're Hessian is not invertible: Alternatives to model specification in nonlinear estimation. Sociological Methods and Research, 2004.

[8] Kohei Hayashi† Takanori Maehara§ Masashi Toyoda Ken-ichi Kawarabayashi "Real-Time Top-R Topic Detection on Twitter with Topic Hijack Filtering",2015

[9] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi. On the evolution of user interaction in Facebook. WOSN, 2009.

[10] R. Rodrigues, R. das Dores, C. Camilo-Junior, and C. Rosa,"Sentihealth-cancer: A sentiment analysis tool to help detecting moodof patients in online social networks," International Journal of MedicalInformatics, vol. 1, no. 85, pp. 80–95, 2016.