

Methods in Predictive Techniques for Mental Health Status on Social Networks using Machine Learning

Nalini. L

Research Scholar

College of Computer Science and Information Science, Srinivas University, Mangalore, India
Assistant Professor, Department of Computer Science, BMS College for Women, Bangalore, India
nalinabdvt@gmail.com

Abstract: *Purpose:* Social networks have been developed as a great point for its users to communicate with their interested friends and share their opinions, photos, and videos reflecting their moods, feelings and sentiments. This creates an opportunity to analyze social network data for user's feelings and sentiments to investigate their moods and attitudes when they are communicating via these online tools.

Methods: Although diagnosis of mental health using social networks data has picked an established position globally, there are several dimensions that are yet to be detected. In this study, we aim to perform psychological analysis on Facebook data collected from an online public source. To investigate the mental health, we propose machine learning technique as an efficient and scalable method.

Results: With the implementation of the proposed method using a set of various psycholinguistic features, we can significantly improve the accuracy and classification error rate. In addition, we can achieve expected results in different experiments such as Decision Tree (DT) which gives the highest accuracy in ML approaches to find the mental health.

Conclusions: Machine learning techniques identify high quality solutions of mental health problems among Facebook users.

Keywords: Health, Relationship, Classifier, Exploration, Community

I. INTRODUCTION

The proliferations of internet and communication technologies, especially the online social networks have rejuvenated how people interact and communicate with each other electronically. The applications such as Facebook, Twitter, Instagram and alike not only host the written and multimedia contents but also offer their users to express their feelings, emotions and sentiments about a topic, subject, or an issue online. On one hand, this is great for users of social networking site to contribute and respond to any topic online; on the other hand, it creates opportunities for people working in the health sector to get insight of what might be happening at mental state of someone who reacted to a topic in a specific manner openly and freely. To provide such insight, machine learning techniques could potentially offer some unique features that can assist in examining the unique patterns hidden in online communication and process them to reveal the mental state (such as 'happiness', 'sadness', 'anger', 'anxiety', Mental health) among social networks' users. Moreover, there is growing body of literature addressing the role of social networks on the structure of social relationships such as breakup relationship, mental illness ('Mental health', 'anxiety', 'bipolar' etc.), smoking and drinking relapse, sexual harassment and for suicide ideation.

II. LITERATURE SURVEY

In this study, we aim to analyze Facebook data to detect any factors that may reflect the mental health of relevant Facebook's users. Various machine learning techniques are employed for such purpose. Considering the key objective of this study, the following are subsequent research challenges that will be addressed.

Define what mental health are and what are the common factors contributing toward mental health. What are the factors to look for mental health detection in Facebook comments? How to extract these factors from Facebook comments? What is the relationship between these factors and attitudes toward mental health? When is the most influential time to communicate within depressive indicative Facebook user? What are the most influential machine learning techniques for

detection of mental health in Facebook comments?

In the context of above-mentioned challenges, we analyze mental health from Facebook users' data. As users express their feeling as a post or comments in the Facebook platform, sometimes their posts and comments refer to as emotional state such as 'joy', 'sadness', 'fear', 'anger', or 'surprise'. We analyze various features of Facebook comments by collecting data through an effective method of machine learning classification techniques and to make overall judgments regarding their various parts. In this study, we used publically available Facebook data (from bipolar, Mental health and anxiety Facebook page) containing users' comments.

In this study, we examine various linguistic cues which help to detect emotion cause events: the position of cause event and experience relative to the emotion keyword: emotional process like positive emotion (e.g. 'happy', 'love', 'nice'), negative emotion (e.g. 'worthless', 'loser', 'hurt', 'ugly', 'nasty'), sadness (e.g. 'worry', 'crying', 'grief', 'sad'), anger (e.g. 'stop', 'shit', 'hate', 'kill', 'annoyed') and anxiety (e.g. 'worried', 'fearful'). A temporal process like present focus (e.g. 'today', 'is', 'now'), past focus (e.g. 'ago', 'did', 'talked') and future focus (e.g. 'shall', 'may', 'will', 'soon'). Linguistic words like articles (e.g. 'a', 'an', 'the'), prepositions (e.g. 'for', 'in', 'of', 'to', 'with', 'above'), auxiliary verbs (e.g. 'do', 'have', 'am', 'will'), conjunctions (e.g. 'and', 'but', 'whereas'), personal pronoun (e.g. 'I', 'them', 'her', 'him'), impersonal pronouns (e.g. 'it', 'it's', 'those'), verbs (e.g. 'go', 'good') and negation (e.g. 'deny', 'dishonest', 'no', 'not', 'never').

III. STATEMENT OF THE RESEARCH PROBLEM AND OBJECTIVES

The main contributions of this paper are listed as follows: First, we synthesize the literature on various emotion detection techniques to detect mental health. Second, we designated four features for our specific research problem and elaborate on the lesson learned from using each type. Third, our experiments are carried out on datasets of Facebook user comments. Fourth, we suggest machine learning techniques to utilize all factors and maintain robustness. We also identify that a Decision Tree classifier outperforms other classifiers (a SVM, KNN and Ensemble) for our dataset. Finally, our work also shows the importance of mental health detection for mental disorder.

IV. MATERIALS AND METHODS

In this study, we will first focus on four types of factors such as emotional process, temporal process, linguistic style and all (emotional, temporal, linguistic style) features together for the detection and processing of depressive data received as Facebook posts. We then apply supervised machine learning approaches to study each factor types independently. The classification techniques such as 'decision tree', 'k-Nearest Neighbor', 'Support Vector Machine', and 'ensemble' are deemed suitable for each type. Below are the methods at high level to be applied for finding the mental health using the posts/comments in social networks.

- Data set exploration
- Data set preparation
- Building ground truth dataset
- Feature extraction
- Measuring depressive behavior
- These higher-level categories are also divided into sub- categories such as
- Psychological process— affective process, social process, cognitive process, perceptual process, biological process, drives, time orientations, relativity, personal concerns
- Linguistic process— word count, word/sentence, pronoun, personal pronoun, articles, prepositions, auxiliary verbs, adverbs, conjunctions, Negations
- Others grammar— verbs, adjectives, comparisons, interrogatives, number, quantifiers.
- Biological processes— sexual, body, ingestion, and health
- Affective processes— anxiety, anger, sadness, positive emotion, negative emotion
- Time orientations— present, past, future
- Social processes— family, friends, male, female
- Perceptual processes— see, hear, and feel.

V. EXPECTED OUTCOME

In this research we will be exhibiting the capability of using Facebook as a tool for measuring and detecting major mental health among its users which we feel will be useful in improving the mental health of our community or society.

- **Timeline:** Following activities will be done to complete the research and conclude on the thesis,
- **Data Set Exploration:** We will work on Facebook users' comments for psychological behavioural exploration and detection. We will collect data from the social network. Preparation social network data, in particular Facebook user's comments is one of the primary challenges which bear information
- On whether or not they could contain psychological content. We will be doing qualitative data analysis of the captured social media content.
- **Data Set Preparation:** After collecting the raw data from Facebook, it was analysed by using a suitable software that will heart of our software engine which will do the text analysis strategy and can process text on a line by line. Our primary dataset contains information that represents the linguistic style (articles, prepositions, auxiliary verbs, conjunctions, personal pronoun, impersonal pronouns, verbs, negation etc.)
- **Building Ground Truth Dataset:** In this this section we will process the information and employ it to construct our dataset with ground truth label information (on whether the comments is psychological indicative). The Facebook data containing users' comments will be divided into two sets (a) for the positive (YES) class (psychological indicative comments) and (b) for the negative (NO) class (non-psychological indicative comments).
- **Feature Extraction:** To describe and demonstrate amongst psychological and non-psychological effecting posts, we extract the different features in view of psycholinguistic measurements from the user's post.
- **Measuring Psychological Behaviour:** Finally, with all the information and data sets we will measure the Mental health on human mind.

REFERENCES

- [1]. Scott J. Social network analysis. Thousand Oaks: Sage; 2017.
- [2]. Mikal J, Hurst S, Conway M. Investigating patient attitudes towards the use of social media data to augment depression diagnosis and treatment: a qualitative study. In: Proceedings of the fourth workshop on computational linguistics and clinical psychology—from linguistic signal to clinical reality. 2017.
- [3]. Conway M, O'Connor D. social media, big data, and mental health: current advances and ethical implications. *Curr Opin Psychol.* 2016; 9:77–82.
- [4]. Ofek N, et al. Sentiment analysis in transcribed utterances. In: Pacific Asia conference on knowledge discovery and data mining. 2015. Cham: Springer.
- [5]. Holleran SE. The early detection of depression from social networking sites. Tucson: The University of Arizona; 2010.
- [6]. Greenberg LS. Emotion-focused therapy of depression. *Per Centered ExpPsychother.* 2017;16(1):106–17.
- [7]. Haberler G. Prosperity and depression: a theoretical analysis of cyclical movements. London: Routledge; 2017.
- [8]. Pennebaker JW, Francis ME, Booth RJ. Linguistic inquiry and word count: LIWC 2001, vol. 71. Mahway: Lawrence Erlbaum Associates; 2001.