Mining Online Social Networking Data for Detecting Mental Disorders

Surabhi Mohite and Satish Gujar

March 30, 2021

Mining Online Social Networking Data for Detecting Mental Disorders

Ms. Surabhi S. Mohite
Student in Department of Computer Engineering,
TSSM's BSCOER, Narhe, Pune
surabhimohite20@gmail.com

Dr. Satish N. Gujar
Professor in Department of Computer Engineering
Coordinator,TSSM's BSCOER,Narhe Pune
satishgujar@gmail.com

***Abstract*—The growth of social network communication has resulted in risky use. Recently, there has been an increase in the number of social network mental disorders (SNMD), such as reliance on cybernetic relationships, data overload, and network constriction. Currently, these mental illnesses' manifestations are passively detected, resulting in late clinical intervention. In this paper, the authors argue that mining online social activity provides a way to systematically classify the SNMD at an early level. Since the mental state cannot be detected directly from reports of online social interactions, it is difficult to identify SNMD.Our new and groundbreaking approach to SNMD detection is not focused on self-disclosure of mental factors by psychological questionnaires. Instead, we suggest a machine learning algorithm called SNMD (Detection of Mental Illnesses in Social Networks), which uses features derived from social network data to reliably classify possible SNMD cases. We also propose a new SNMD-based tensor model (STM) to increase accuracy and use multiple sources learning in SNMD. We boost STM's efficiency with performance guarantees to increase its scalability.A user analysis with a large number of network users is used to test our framework. We evaluate the features of the three types of mental disorders using feature analysis and SNMD in large-scale data sets.**

***Index Terms*—Decision Tree classifier, social network, mental illness identification, feature extraction**

I. INTRODUCTION

Nowadays, mental illness is posing a serious threat to people's welfare. With the fast pace of life, an increasing number of people are experiencing mental distress. It is difficult to diagnose a user's mental illness early enough to protect them. Because of the popularity of web-based social networking, people are accustomed to sharing their daily activities and communicating with friends through web-based networking media platforms, making it possible to use online social network data to identify mental disorders.We discovered that a user's disorder state is closely linked to that of his or her social media friends in our scheme, and we used a large-scale dataset from real-world social networks to systematically research the association between users' disorder states and social interactions. We begin by defining a collection of textual, visual, and social characteristics associated with mental disorders from different perspectives. Because of the fast pace of life, an increasing number of people are becoming overwhelmed. Though mental disorders are non-clinical and normal in our lives, they can be detrimental to people's physical and mental health if they are excessive and persistent.The social experiences of users on social media provide valuable clues for stress detection.

Two fascinating findings have emerged from social psychological research. The first is mood contagion, which occurs when one person's bad mood is passed on to another during social contact. The second type of social interaction is when people are aware of a user's social interactions. With the rise of social media platforms such as Twitter, Facebook, and Sina Weibo, an ever-increasing number of people will use these platforms to share their daily activities and moods, as well as connect with friends. Using a machine learning system, we can classify.Due to a combination of Facebook post content attributes and social experiences, mental illness identification has improved. After determining the severity of the disorder, the system can recommend a hospital for further care, which we can display on a chart, and the system can also advise the consumer to take precautions to prevent disorder.

II. LITERATURE SURVEY

"In this paper[1], we introduce MaxMin-CNN, a new deep CNN architecture that improves the encoding of both positive and negative filter detections in the net [1]."
Advantages:1. We suggest modifying CNN's standard convolutional square with the aim of exchanging further data layer after layer while maintaining some invariance within the framework. 2. Our main idea is to take advantage of both positive and negative high scores in convolution charts. Until pooling, this behaviour is acquired by altering the standard enactment job venture.
Disadvantages: More time is needed for this. It is a lengthy procedure.

We investigate[2] a method for automatically detecting stress in cross-media microblog data.
Advantages include: 1.A three-tiered system for detecting stress in cross-media microblog data. The system is very feasible and efficient for stress detection because it combines a Deep Sparse Neural Network with different features from cross-media microblog data. 2.With this context, the proposed approach will help to identify psychological stress in social networks automatically.
Disadvantages: In order to enhance the detection accuracy,

we aim to look into social associations in psychological stress.

We're [3] curious about our clients' identities. Identity seems to be relevant to a wide range of collaborations.
Advantages:1. We care about our clients' identities. Identity has been shown to be relevant to a wide range of collaborations; it has been shown to be useful in predicting job satisfaction, relationship success, and even inclination. 2. We're curious about our clients' identities. Identity has been shown to be applicable to a wide range of communications; it has been shown to be useful in predicting job satisfaction, professional and sentimental relationship success, and even preference for different interfaces.
Advantages: we can start to address more complex questions about how to provide users with trustworthy, socially important, and well-presented content.

Daily stress recognition from mobile phone data, weather conditions, and individual traits has been studied[4].
Advantages:1. Daily stress may be reliably perceived based on behavioural metrics obtained from the client's mobile phone action as well as additional markers such as weather conditions (information relating to short-term properties of the condition) and identification attributes.
Disadvantages:1. In workplaces where stress has become a serious issue impacting productivity, causing occupational problems, and causing health diseases. 2. Our framework could be expanded and used to identify stress-related disputes and stress contagion early, as well as to help healthy workloads. Disadvantages:1. In workplaces where stress has become a serious issue impacting productivity, causing occupational problems, and causing health diseases. 2. Our framework could be expanded and used to identify stress-related disputes and stress contagion early, as well as to help healthy workloads.

This is used[5] to research a subject. Using cross autoencoders, we can learn robust uniform features for cross-media social data.
1. To solve learning models to handle cross-modality associations in cross-media social elements. 2. We propose CAE to learn uniform modality-invariant features, and AT and PT phases to train the CAE by leveraging large crossmedia data samples.
Disadvantages: Using cross autoencoders to learn robust standardised functionality for cross-media socialdata takes longer.

We may conduct[6] research into when an individual is in good spirits and searching the emotional Web.
Advantages include: 1. On the use of We Feel Fine to propose a new type of visualisation known as Experiential Data Visualization, which focuses on interactive item-level data interaction. 2. The social science implications of such visualisations for crowdsourcing qualitative analysis.
Disadvantages:Repeated information in relevant answers forces the consumer to sift through a large number of responses in order to find the information they need.

To research [7] on using a global learning approach to bridge the language gap between health seekers and healthcare expertise.
Advantages include: 1. a medical terminology assignment scheme to close the gap between health seekers' language and their understanding of healthcare. Local mining and global learning are the two elements of the scheme. 2. Extensive tests on a real-world dataset show that our scheme is capable of producing promising results as compared to existing coding methods.
Disadvantages: Using suggested medical terminologies, we will explore how to flexibly organise unstructured medical content into user needs-aware ontology.

The use of everyday terms to define pictures is studied using picture [8] tags and world information learning tag relations from visual semantic sources.
Advantages: The proposed tagging algorithm generalises to unseen tags and improves further when tag-relation features obtained from ICR are incorporated.
Disadvantages:Multi-word terms and out-of-vocabulary words; advanced NLP techniques for learning word relations from free-form text; evaluation of latent concept relation suggestion, and predicting the type of relations.

We used this to investigate a novel issue of emotion prediction in social networks.
Advantages include: 1. MoodCast is a tool for modelling and predicting emotion dynamics in social networks. 2. The proposed method can accurately model each user's emotional state, and its prediction accuracy outperforms other baseline methods for emotion prediction.
Disadvantages: Due to the small number of participants, it is used to.

This is related to research on the influence maximisation problem, which aims to identify a small subset of nodes (users) in a social network that can optimise influence spread.
1. A Pairwise Factor Graph (PFG) model to formalise the problem in a probabilistic model, which we extend by including time details, resulting in the Dynamic Factor Graph (DFG) mode. 2. The proposed method is capable of detecting complex social factors.
Disadvantages:1. Our algorithm can be parallelized in future work to speed it up even further."

III. Proposed Methodology

We formulate the challenge as a classification issue in the proposed system approach to detect three forms of social network mental disorders using a machine learning framework: i) Stress and non-Stress.
ii) Cyber-Relationship Addiction, which shows addictive

behavior for building online relationships.
iii) Net Compulsion, which shows compulsive behavior for online social gaming or gambling.
iv) Over load of data, which is related to uncontrollable surfing.

A. Architecture

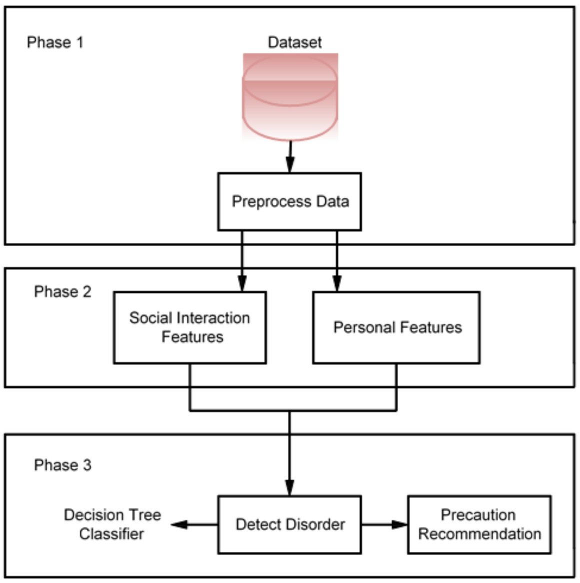

Fig. 1. System Architecture

B. Algorithms

Decision Tree Machine Learning Algorithm:

The decision tree algorithm is a type of supervised learning algorithm. They can be used to solve problems involving regression and classification.
The decision tree solves the problem by using the tree's representation, where each leaf node corresponds to a class mark and the attributes are expressed in the tree's inner node.

- At the beginning, we consider the whole training set as the root.
- Feature values are preferred to be categorical. If the values are continuous then they are discretized prior to building the model.
- On the basis of attribute values records are distributed recursively.
- We use statistical methods for ordering attributes as root or the internal node.

C. Hardware and Software Requirements

Hardware Requirements:
1. Processor - Pentium –III
2. RAM - 2 GB(min)
3. Hard Disk - 20 GB
4. Key Board - Standard Windows Keyboard
5. Mouse - Two or Three Button Mouse
6. Monitor - SVGA

Software Requirements:
1. Operating System - Windows
2. Application Server - Apache Tomcat
3. Coding Language - Java 1.8
4. Scripts - JavaScript.
5. Server side Script - Java Server Pages.
6. Database - My SQL 5.0
7. IDE - Eclipse

D. Mathematical Model

Input:
Step 1: Upload training dataset.
Step 2: Social Networking Posts set is the set of input attributes.
Step 3: Mental Disorder is the set of output attributes.
Step 4: sample is a set of training data.
Function Iterative Dichotomiser returns a decision tree
1. Create root node for the tree
2. If (all inputs are positive, return leaf node positive)
If Else (if all inputs are negative, return leaf node negative)
Else (Some inputs are positive and some inputs are negative, check condition (Positive¿negative——Positive¡negative), then return result)
3. Calculate the entropy of current state H(S)
4. For each attribute, calculate the entropy with respect to the attribute 'X' denoted by H(S,X)
5. Select the attribute which has maximum value of IG(S,X)
6. Remove the attribute that offers highest value from the set of attributes
7. Repeat until we run out of all attributes or the decision tree has all leaf nodes.

Output:
Dataset value will be retrieved.

IV. Result and Discussion

For performance assessment, an experimental evaluation is carried out to equate the proposed system to the current system. The simulation platform is based on the Java system (version jdk 8) and runs on Windows. The programme does not need any special hardware to run; it can be run on any standard computer.

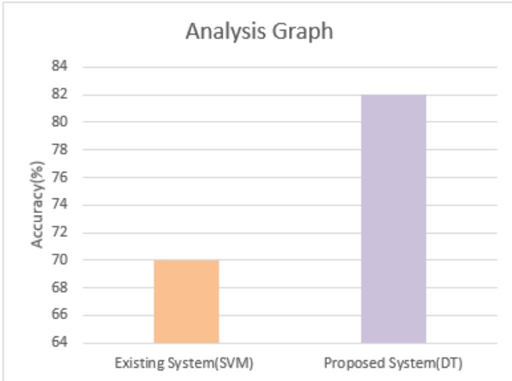

Fig. 2. Graph

Sr. No.	Existing System	Proposed System
1	70%	82%

Table 1:Comparative Result

V. Conclusion

As a result, identify possible online users with SNMDs in this article. Individuals' well-being is being jeopardised by psychological mental disorders. It is insignificant to recognise Mental Disorder in a timely manner for proactive consideration. As a result, we demonstrated a mechanism for identifying clients' Mental Disorder states from month-to-month online networking data, using Facebook post content as well as clients' social associations. We found the link between users' Mental Disorder states and their social collaboration activities using genuine internet-based life knowledge as the assumption and recommended the user to a health counsellor or specialist. On a map, we show the medical facilities for further care and the shortest route from the current area patient to the emergency clinic.

References

[1] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella,and J ¨ urgen Schmidhuber. "Flexible, high performance convolutional neural networks for image classification." In Proceedings of International Joint Conference on Artificial Intelligence, pages 1237–1242, 2011.

[2] H. Lin, J. Jia, Q. Guo, Y. Xue, J. Huang, L. Cai, and L. Feng. "Psychological stress detection from cross-media microblog data using deep sparse neural network. " In proceedings of IEEE International Conference on Multimedia Expo, 2014.

[3] Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. "Predicting personality from twitter." In Passat/socialcom 2011, Privacy, Security, Risk and Trust, pages 149–156, 2011

[4] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. "Daily stress recognition from mobile phone data, weather conditions and individual traits. " In ACM International Conference on Multimedia, pages 477–486, 2014.

[5] Quan Guo, Jia Jia, Guangyao Shen, Lei Zhang, Lianhong Cai, and Zhang Yi." Learning robust uniform features for cross-media social data by using cross autoencoders." Knowledge Based System, 102:64– 75, 2016.

[6] Sepandar D. Kamvar. "We feel fine and searching the emotional web." In Proceedings of WSDM, pages 117–126, 2011.

[7] Liqiang Nie, Yi-Liang Zhao, Mohammad Akbari, Jialie Shen, and Tat-Seng Chua." Bridging the vocabulary gap between health seekers and healthcare knowledge." Knowledge and Data Engineering, IEEE Transactions on, 27(2):396–409, 2015.

[8] Lexing Xie and Xuming He. "Picture tags and world knowledge: learning tag relations from visual semantic sources. "In ACM Multimedia Conference, pages 967–976, 2013.

[9] Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao. "Moodcast: Emotion prediction via dynamic continuous factor graph model. "2013 IEEE 13th International Conference on Data Mining, pages 1193–1198, 2010.

[10] Chi Wang, Jie Tang, Jimeng Sun, and Jiawei Han." Dynamic social influence analysis through time-dependent factor graphs." Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on, pages 239 – 246, 2011.