Personality Traits and Mental Disorder Prediction Through Twitter Analysis: Applications of Social Media Analytics in Digital Psychiatry

Dr. Karin Tamm[1]*
[1]University of Tartu Faculty of Medicine, Department of Psychiatry and Digital Mental Health Research, Tartu, Estonia

ABSTRACT

With the rise in popularity of social networking, various types of social network mental disorders (SNMDs), such as Cyber-Relationship Addiction, Information Overload, and Net Compulsion, has shown up on the surface. Symptoms of these mental disorders are usually observed passively, resulting in delayed clinical intervention. From the studies it was observed that the core reason for some of this addictive behavior was depression. In this paper we propose a machine learning framework, enhanced with sentimental analysis that exploits features extracted from social network data, to accurately identify potential cases of SNMDs, depression and personality of the user. The sentimental analysis feature extraction process uses multi-class classification which considers basic human emotions (happiness, sadness, surprise, disgust) and neutral as emotion classes. Twitter API is used for extracting the tweets and profile details. A combination of Naive Bayes classifier and ML-KNN classifier is used for classification to yield better results.

KEYWORDS: Social network mental disorder; Feature Extraction; Primary Features; Secondary Features.

1. INTRODUCTION

With the rise in growth and trending of social networking and its apps, online social networks have become an integral part of many people's day to day lives and Twitter is the prominent among them. The number of users joining the social media is increasing rapidly day by day. Most research conducted on social media mining focuses on exploring the knowledge from the social media data for improving people's life. The practice of social media data mining collects and processes the unstructured information (such as posts, comments, tweets, images) shared on social networks like Twitter. While social media has seemingly revolutionized their user's capability in expanding social contacts in the cyber word, they have actually reduced the face-to-face interpersonal interactions in the offline real world. Some social network mental disorders (SNMDs), such as Cyber-Relationship Addiction, Information Overload and Net Compulsion, have recently surfaced. SNMDs are web-oriented and tend to happen to users who usually choose online social media to interact with others. People suffering from SNMDs usually show less interest for offline interactions, and compensate it through cyber-relationships. As the noticeable physical risk factors are less, the patients act passively in seeking medical or psychological services. Therefore, by the time patients seek clinical services; their conditions would have become chronic. The delay in early diagnosis may seriously ruin the individual's social functioning. In short, there is an alarming need to have some mechanism to detect potential SNMD users on online social network at an early stage.

We propose a machine learning framework to predict personality of the user and whether he/she is suffering from any of the social network mental disorders .This framework works based on multiclass sentiment analysis and multilabel classification. The multiclass sentimental feature analysis considers basic human emotions like happiness, sadness, surprise, disgust; and neutral as emotion classes .This adds up to the accuracy of prediction.

2. PREVIOUS WORKS

Mining social media data for mental illness detection has become an area of research. The traditional way of predicting mental illness is usually using questionnaires.

S.C.Guntuku et al. [1] reviewed using social media to predict mental illness. Symptoms associated with mental illness are observable on social medias like Twitter, Facebook, and web forums. Mentally ill users have been identified using screening surveys, their public sharing of a diagnosis on Twitter, or by their membership in an online forum, and they were distinguishable from control users by patterns in their language and online activity.
M.D.Choudhury et al. [2] proposed a system that used the potential of using Twitter as a tool for measuring and predicting major depression in individuals. Initially crowdsourcing was used to collect labels on depression. They proposed a variety of social media measures such as language, emotion, style, egonetwork, and user engagement to characterize depressive behavior. The findings showed that patients with depression show lowered social activity, higher negative emotion, greater self-attention focus, elevated relational and medicinal concerns, and heightened expression of religious thoughts. They have used SVM classifier that can predict, ahead of the reported onset of depression of an individual i.e; his or her likelihood to depression. The classifier yielded results with 70% classification accuracy.

B.Y.Pratama et al. [3] Personality is a fundamental basis of human behavior. Personality affects the interaction and preferences of an individual. People are required to take a personality test to find out their personality. Social media is a place where users express themselves to the world. Posts made by users of social media can be analyzed to obtain their personal information. This research work has used text classification to predict personality based on text written by Twitter users. Two languages English and Indonesian are used. Classification methods implemented are Naive Bayes, K-Nearest Neighbors and Support Vector Machine. Testing results showed Naive Bayes slightly outperformed the other methods. Personality can influence a person's choice in various things. User's personality successfully predicted from text written on Twitter.

A.Z.Riyadh et al. [4] have proposed a system for emotion analysis of tweets using only the core text. Tweets are usually short, more ambiguous and contain a huge amount of noisy data; sometimes it is difficult to understand the user's opinion. Most of the researches in this topic have been focused on binary (positive and negative) and 3-way (positive, negative and neutral) classifications. In this system, we have focused on emotion classification of tweets as multi-class classification. The basic human emotions (happiness, sadness, surprise, disgust) and neutral were chosen as emotion classes. A technique for emotion analysis of tweets using unigram model and unigram model with POS tags for feature extraction were applied. It has used Bag of Words Model. Naive Bayes classifier is used for classification of emotions. The emoticons, URLs, targets, punctuation, stop-words were removed to simplify and make the classification more accurate. According to the experimental results, this approach improved the performance of multi-class classification of twitter data.

M.L.Zhang et al. [5] Multi-label learning originated from the research on text categorization problem, where each document may belong to several predefined topics simultaneously. In multi-label learning, the training set is composed of instances each associated with a set of labels, and the task is to predict the label sets of unseen instances through analyzing training instances with known label sets. In this paper, a multi-label (ML-KNN) is presented, which is derived from the traditional k-Nearest Neighbor (kNN) algorithm. In detail, for each unseen instance, its k nearest neighbors in the training set is firstly identified. After that, based on statistical information gained from the label sets of these neighboring instances, i.e. the number of neighboring instances belonging to each possible class, maximum a posteriori principle is utilized to determine the label set for the unseen instance. In text categorization, each document may belong to several topics, such as government and health in functional genomics, each gene may be associated with a set of functional classes, such as metabolism, transcription and protein synthesis.

H.H.Shuai et al. [7] have proposed a system for detecting social network mental disorders (SNMDs), such as Cyber-Relationship Addiction, Information Overload, and Net Compulsion which exploits features extracted from social network data. Mining online social behavior provides an opportunity to actively identify SNMDs at an early stage and it is challenging to detect SNMDs because the mental status cannot be directly observed from

online social activity logs. It uses multi-source learning from datasets of facebook and instagram in social network mental disorder detection and proposes a new SNMD-based Tensor Model (STM) to improve the accuracy.

The disadvantage of this system is that users may behave differently on different OSNs i.e if the user is not equally active in all OSNs, resulting in inaccurate SNMD detection. Also it has not explored the sentiment as feature which can contribute to the accuracy.

3. PROBLEM DEFINITION

While using social network has many benefits; its excessive use if not controlled may lead to mental disorder categorized as social network mental disorders (SNMDs). SNMDs tend to happen to users who usually choose online social media to interact with others and for expressing themselves. Specifically, we are focusing on three types of SNMDs:
i) Cyber-Relationship Addiction, which shows addictive behavior for connecting and building relationships online. This includes getting hooked up in the cyber social world and get involved in online messaging to the point, where the person prefer virtual social relationships and online friends than real-life ones with friends and families in the offline world.
ii) Net Compulsion, which shows obsessive interest for online social gaming or gambling; usually resulting in monetary and job related issues.
iii) Information Overload, which shows a compulsive habit of surfing of user status and newsfeeds in untimely manner; i.e night surfing during wee hours leading to lower work productivity and deterioration in social interactions with families and friends offline.
The personality traits of a person have deep influence over the chances of developing SNMD condition. So in our analysis we are predicting both personality and social network mental disorder.

4. Proposed System

Twitter API
For our analysis we ought to get the profile details like age, description, followee, followers and favorites as well as other details like the tweets tweeted by the user, retweet count, replies given by the user to others etc. The details are accessed via Twitter API by giving twitter profile handle as input.

Preprocessing
Before extracting the primary features, preprocessing of the tweets is performed. The following are the steps performed as part of preprocessing:s

- URL links usually doesn’t carry any polarity or meaning .Hence they are removed from each tweet as part of preprocessing.
- Stop words like “a”, “an”, “the” etc are removed as they too doesn’t carry polarity or meaning.
- Perform tokenizing i.e change sentence into a collection of a single word.
- Replace short forms to original format (eg: gdnt to good night).
- Stemming is performed i.e. each word in the tweet is mapped to its root word. This is done by using Porter Stemmer.
- Convert the emoticons to text. This not only reduces the ambiguity while considering emoticons for analysis but also increases accuracy.
- Perform POS tagging (part of speech tagging) in order to avoid ambiguity while interpreting between nouns and verbs.
- Calculate tf- idf score for each word

Primary Features
Primary features refer to the features that can be extracted from the profile details like age, gender, description, followee, followers etc. The details are extracted via Twitter API by giving twitter profile handle as input.

Sentiment Analysis
On extracting the tweets of the user multi-class sentiment analysis is performed over it. The sentimental analysis feature extraction process performs multi-class classification over the tweets which considers basic human emotions like happiness, sadness, surprise, disgust ;etc and neutral as emotion classes.

Secondary Features
The secondary features refer to the collection of features that can be extracted from primary features and the features extracted through sentiment analysis. Features that can be extracted from primary features include openness, conscientiousness, extraversion, agreeableness, neurotism and the features extracted through sentiment analysis performed on the tweets are happiness, sadness, surprise, disgust and neutral.

Feature Extraction
The collected dataset is used for feature extraction and it will be used for training the classifiers.We are using two classifiers Naïve Bayes and ML-KNN for deriving the results.Naive Bayes classification algorithm works

based on the Bayes theorem.The primary features classification is done by using Naïve Bayes classifier as it purely focuses on text classification.During our study it was observed that Naive Bayes classifier showed the best performance for multi-class classification for sentimental analysis.The same is used for deriving secondary features. The primary features,the secondary features and the features extracted through sentimental analysis collectively forms the input for ML-KNN algorithm. Since there will be patients who are suffering from multiple disoders we prefered ML-KNN (multi label K-nearest neighbours) algorithm to represent the result.This algorithm uses a distance function (euclidean distance) between training data to test data,and the ‘k’ number of nearest neighbours are determined and finally arrives at the classified result.
The architecture diagram is shown in Figure 1.

5. EXPERIMENTS AND RESULTS

Experimental Set-up
We have collected our own datasets of Twitter messages with profile details. Twitter’s application programming interface (Twitter API) is used for programmatically accessing tweets based on query term. The Twitter API has a parameter for specifying which language tweets has to be retrieved and this has been set to English in our case. For our system we have collected tweets and profile details by either giving the twitter handle or words or hash tags as query term. We have manually labeled them. The algorithms are implemented in C#.

We are considering the last 100 tweets of each user for analysis. The test data is collected as random samples from training set.

Results
The collected data belonging to each class is divided into three equal sized folds, out of which two folds were used for training and one fold for testing. We tested with three classifiers Logistic regression, SVM and Naïve Bayes at the stage of sentiment analysis; as part of secondary feature extraction.

It was observed that Naïve Bayes outperformed the other two classifiers for sentiment analysis.SVM performs well for sentiment analysis with longer documents, but does not perform well with shorter texts. Since tweets are limited to limited characters Naïve Bayes forms the perfect fit. Both Logistic regression and Naïve Bayes are linear classifiers. It was observed that for small datasets Logistic regression outperformed Naïve Bayes but for larger datasets Naive Bayes performed better.

The experimental results are shown in Table I. From the analysis it was observed that using Naïve Bayes gave better accuracy.

TABLE I. COMPARISON WITH DIFFERENT TECHNIQUES

Classification Technique	***Accuracy***
Logistic Regression	89.9%
SVM	90.1%
Naïve Bayes	93.2%

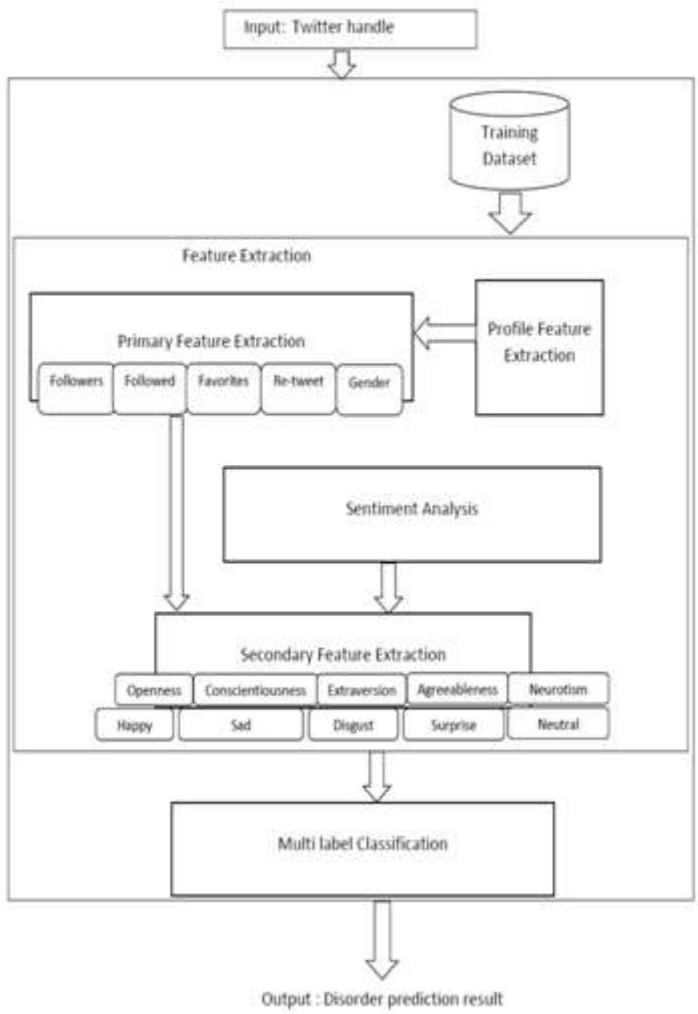

Figure 1. Architecture Diagram

6. CONCLUSION

In this paper, we make an attempt to use machine learning framework to predict the personality, depression and the potential cases of social network mental disorders like Cyber-Relationship Addiction, Information Overload, and Net Compulsion in Twitter users. We have used the profile details and tweets of the users for our analysis. Our system uses multi-class sentimental analysis which enhances identification of sentiment associated with each sentence of the tweets. These data can be complementing to the work of healthcare professionals.

As part of the future work, we can add more features for analysis like considering location how the tendency for developing social network mental disorder varies with urban and countryside people. Also we can extend our analysis to measure how much a person is addicted to apps in personal devices. Thereby, we can track and detect internet user's addiction from turning to a disorder at an early stage.

7. ACKNOWLEDGMENT

We would also like to show our gratitude to Prof.Sabeena M.V (Principal), Mr.Gireesh.T.K (Head of department), Mrs.Shahana.T(Project Coordinator) and other faculties of AWH Engineering College for sharing their wisdom with us during the course of this research. We are also immensely grateful to the Department of Computer Science & Engineering for providing us with facilities to complete our work.

REFERENCES

[1] Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar,and Johannes C Eichstaedt. "Detecting depression and mental illness onsocial media: an integrative review.Current Opinion in Behavioral Sciences",18:43–49, 2017.

[2] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz."Predicting depression via social media". InSeventh international AAAI con-ference on weblogs and social media, 2013.

[3] Bayu Yudha Pratama and Riyanarto Sarno. “Personality classification basedon twitter text using naive bayes, knn and svm”. In2015 InternationalConference on Data and Software Engineering (ICoDSE), pages 170–174.IEEE, 2015.

[4] Abu Zonayed Riyadh, Nasif Alvi, and Kamrul Hasan Talukder. “Exploring human emotion via twitter”. In2017 20th International Conference ofComputer and Information Technology (ICCIT), pages 1–5. IEEE, 2017.

[5] Min-Ling Zhang and Zhi-Hua Zhou. “ Ml-knn: A lazy learning approach to multi-label learning”.Pattern recognition, 40(7):2038–2048, 2007.

[6] Maryam Hasan, Elke Rundensteiner, and Emmanuel Agu. “Emotex: Detect-ing emotions in twitter messages”. 2014.

[7] Hong-Han Shuai, Chih-Ya Shen, De-Nian Yang, Yi-Feng Carol Lan, Wang-Chien Lee, S Yu Philip, and Ming-Syan Chen. “A comprehensive study on social network mental disorders detection via online social media mining”.IEEE Transactions on Knowledge and Data Engineering, 30(7):1212–1225, 2018.

CITE AN ARTICLE

It will get done by IJESRT Team