STRESS DETECTION AND SENTIMENT PREDICTION: A SURVEY

Dr. G V Garje
HOD of Computer Engineering and IT
PVG's College of Engineering and Technology
Pune, Maharashtra, India

Apoorva Inamdar, Harsha Mahajan,
Apeksha Bhansali, Saif Ali Khan
Department of Computer Engineering and IT
PVG's College of Engineering and Technology
Pune, Maharashtra, India

Abstract— **Self-expression is a vital usage of social media, in the form of daily life updates, experience description, information and comments sharing, etc. Social media is slowly replacing the practice of face-to-face communication. In a way, virtual communication is taking over the real-life communication. Active participation of Teen-age users on social media is one of the reasons behind this changing dynamics. Twitter is a popular micro blogging site amongst youth today. The humongous volume of tweeting content, along with various tweeting, retweeting and comments statistics, presents an opportunity for human sentiment analysis and behavioral study. Twitter users' behavior and their mood transition is an in demand study domain with gigantic scope and many unexplored aspects. One such area is Stress Detection. There are some studies which have implemented stress identifying and prediction techniques. This paper presents an analytical survey of such recent published studies.**

Keywords—Prediction, Stress, Twitter, Detection

Abbreviations: SVM- Support Vector Machine, ARIMA-Auto Regressive Integrated Moving Average Model, NB- Naïve Bayes, GPC-Gaussian Process Classifier, LIWC- Linguistic Inquiry and Word Count

I. INTRODUCTION

Twitter is a very popular social media platform in today's digital era. Every second, 600 tweets are approximately published on Twitter [1]. Twitter implements limited character policy, in which a user can tweet up to 142 characters. It enables users to convey their message in a short version. This has encouraged the use of short forms to fit into the provided character limit. These short forms involve use of alphabets, digits and exclamations. This has been quite handy for today's generation. Twitter also has the hashtag feature wherein any phrase or word written after a hashtag is considered as the topic of the tweet. Twitter trends are decided by the highest frequency of similar hashtag phrases all around Twitter. The most discussed topics become trends for the day. Twitter also offers regional trends, mainly country-wise.

Since the boom in the E-commerce section, social media platforms are crucial in customer feedback. Sentiment analysis of a product is a major field of study in this culture. But, along with product sentiment and opinions, human emotions can also be analyzed by tweet contents. Twitter offers a quick self-expression format, which is unique because of features like retweeting, hashtag trending, limited characters etc. Twitter gives a chance for quick updates to everyone. The Twitter behavior pattern of a person can be very useful to study human emotions and changing form of self-expression, along with all its effects on other aspects of social life.

For such study, Teen age Twitter users prove to be a good subject. Tweets from such users are mostly spontaneous, unfiltered thoughts expressed without least possible regard for its consequences. This can help in identifying correct emotions of a person at that point. Teen age users are also constantly active, which makes a large chunk of continuous data to analyze. Influences upon Teenage users can also be determined.

In case of Teen agers, one of the prominent problems in recent years has been stress. Stress can be caused by a number of factors such as peer pressure, parental pressure, exam, etc. But, it has been observed that social media is also becoming one of the contributing factors for the teenage stress. Less popularity on internet, being trapped in social engineering traps are leading to problems such as early stress, nervous breakdown, depression, suicidal tendencies among children. When it comes to such problems, prevention is better than cure. Analyzing the online activity of the user, his/her stress level can be determined, even predicted.

This topic is being explored by many to cope with the negative side effect of social media. A survey of this study clarifies the effort taken on the subject, various analysis techniques and possible solutions. This survey reviews a part of work in this domain. The paper is divided into two sections: Systems for Stress Detection and systems for Sentiment

Prediction. It also includes tabular analysis, summarizing the sections. The paper includes analysis of recent IEEE papers.

II. EXISTING SYSTEMS FOR STRESS DETECTION

Paper titled "The adolescent brain and age-related behavioral manifestations", authored by L. P. Spear, published in 2000, describes adolescent behavior and reasons behind it in detail. It particularly emphasizes on the influence of social media for behavioral analysis. The focus of this paper is adolescent behavior review. The paper explores relationship between adolescent brain and behavior function. It differentiates between adolescence and puberty, reviews adolescence in most species at different stages and presents conclusions. This paper does not implement any machine learning algorithm, but presents an analysis which can be used as groundwork for future behavior pattern analysis.

Paper titled "Detecting Adolescent Psychological Pressures from Micro-Blog" authored by Yuan yuan Xue, Qi Li , Li Jin, Ling Feng, David A. Clifton,Gari D. Clifford, was published in 2013. The paper states that the traditional face-to-face psychological diagnosis and treatment cannot meet the demand of relieving teenagers' stress completely due to its lack of timeliness and diversity. Herein, a microblog platform is envisioned to sense psychological pressures through teenagers' tweets, and assist teenagers to release their stress through micro-blog. It implements five classifiers for tweet analysis, namely Naive Bayes, Support Vector Machines, Artificial Neural Network, Random Forest, and Gaussian Process Classifier for pressure detection. The paper presents a well-researched scenario about Tweeting habits, subtle classification of emotion polarities, and standard classification methods. The paper also gives single-tweet based pressure detection. This model is implemented on China Sina Microblog.

Paper titled "Helping Teenagers Relieve Psychological Pressures: A Micro-blog Based System" is an extension to the above mentioned paper. Written by Qi Li, Yuan yuan Xue, Jia Jia, Ling Feng, it presents a system, tHelper for sensing and easing teenagers' psychological pressures in study, communication, affection, or self-recognition through micro-blog. The system adopts Gaussian Process to classify a teenager's pressure based on a number of features extracted from users' tweets. The system collaborates with parents to notify them about their children's' stress level, using Mobile SMS. It also sends positive stories to identified "stressed" children (Registered with system) to reduce their stress. This system is just a proposed concept and not yet implemented on considerable research level. GUI designs for tHelper is under experiments.

The 2014 IEEE invited paper, titled "A Mathematical Approach to Gauging Influence by Identifying Shifts in the Emotions of Social Media Users" is penned by Les Servi and Sara Beth Elson for IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS. It presents a new definition of influence that is tailored uniquely for online contexts and an associated methodology for gauging influence. The paper depicts implementation of LIWC 2007 program for tweet analysis and presents a mathematical model to plot influence in emotion shifts. It has a strong mathematical model. It does not particularly focus on "stress" factor, but covers emotion polarities, such as positive, negative and neutral. The database used for analysis is a third party database available on the internet. The suggested model can also be used to draw emotion trends of Twitter user.

The paper titled "User-Level Psychological Stress Detection From Social Media Using Deep Neural Network" is written by Huijie Lin, Jia Jia, Quan Guo, Yuanyuan Xue, Qi Li, Jie Huang, Lianhong Cai, Ling Feng, from Department of Computer Science and Technology, Tsinghua University. It was published in 2014. The paper employs real online micro-blog data to investigate the correlations between users' stress and their tweeting content. It also defines two types of stress-related attributes: - Low-level content attributes from a single tweet, including text, images and social interactions; and User-scope statistical attributes through their weekly micro-blog postings, mapping information of tweeting time, tweeting types and linguistic styles. A deep Neural Network model is designed & proposed based on this study to analyze users' stress pattern. It studies all users, not just teenagers. The experimental implementations are done on China Sina Weibo database, along with Twitter Database. The model proposed is very complex. It does not implement the widely implemented emotion polarities.

Table -1 Existing Stress Detection Systems

Title of the paper	New Systems Introduced	Use of Machine Learning Algorithm	Database used
The adolescent brain and age-related behavioural manifestations,2000	None	None	N.A.
Detecting Adolescent Psychological Pressures from Micro-Blog,2013	Single tweet based pressure detection	NB, SVM, Artificial Neural Network and GPC	Twitter
Helping Teenagers Relieve Psychological Pressures: A Micro-blog Based System,2014	tHelper	Gaussian Process	Twitter
A Mathematical Approach to Gauging Influence by Identifying	None	None	Twitter

Shifts in the Emotions of Social Media Users,2014			
Title of the paper	**New Systems Introduced**	**Use of Machine Learning Algorithm**	**Database used**
User-Level Psychological Stress Detection From Social Media Using Deep Neural Network,2014	None, Studied system implemented with 78.57% accuracy	Deep Neural Network	China Sina Weibo, Twitter

	82.6% accuracy		
Title of the paper	**New Systems Introduced**	**Use of Machine Learning Algorithm**	**Database used**
Twitter User Behavior Understanding with Mood Transition Prediction,2012	None	SVM	Stanford SNAP database
Real Time Sentiment Analysis of Twitter Data Using Hadoop,2014	None, Study model implemented with 72.27% accuracy	Gaussian Process	Twitter

III. EXISTING SYSTEMS FOR SENTIMENT PREDICTION

A system of Teenagers' Future Stress Prediction using Machine Learning Algorithm implements a 2015 IEEE paper titled "Predicting Teenager's Future Stress Level from Micro-blog". The paper is authored by Yiping Li, Jing Huang, Hao Wang, Ling Feng. The paper proposes a model based on ARIMA for detection and prediction of Teenagers' stress. The paper addresses the challenges such as data incompleteness and multi-faceted prediction using machine learning and multi-variant time series prediction techniques.

The paper titled "Twitter User Behavior Understanding with Mood Transition Prediction" was published by Aditya Mogadala, Vasudeva Varma in the 21st ACM International Conference on Information and Knowledge Management (CIKM-2012). It uses history of mood information to predict future moods. It implements regression analysis techniques on Tweets. It defines mood labels and describes mood swings based on it. The paper appears vague about experimental setups. Stanford SNAP database is used for implementation of this paper. It gives generalized results for moods.

The 2014 paper titled "Real Time Sentiment Analysis of Twitter Data Using Hadoop" is written by Sunil B. Mane, Yashwant Sawant, Saif Kazi, Vaibhav Shinde, from College of Engineering, Pune. It implements a sentiment analysis method using Hadoop. It focuses on analysis speed, rather than analysis accuracy. It converts unstructured Twitter data into structured data. It implements Map Reduce Algorithm.

Table -2 Existing Sentiment Prediction Systems

Title of the paper	**New Systems Introduced**	**Use of Machine Learning Algorithm**	**Database used**
Predicting Teenager's Future Stress Level from Micro-blog,2015	Teenagers' Future Stress Prediction, with	Gaussian Process Regression, ARIMA	China Sina, Twitter

IV. CONCLUSION

From this survey, it can be stated that, while there are many systems for stress detection, very few systems particularly emphasize on stress prediction. There are also a few systems exploring both stress detection and prediction. Also, the systems above work on word analysis, rather than content analysis. Accuracy measures and parameters for systems are rather undisclosed in most papers for copyright reasons. The sentiment polarities explored are clichéd, conventional and inadequate. Still, the ongoing work presents a great hope for human sentiment analysis and prediction.

V. REFERENCE

[1] Yiping Li, Jing Huang, Hao Wang, Ling Feng 'Predicting Teens' Future Stress Level from Microblog', 2015

[2] Sitaram Asur, Bernardo A. Huberman 'Predicting the future with social media', 2015

[3] H. Lin, J. Jia, Q. Guo, Y. Xue, Q. Li, J. Huang, L. Cai, and L. Feng, "User-level psychological stress detection from social media using deep neural network," in Proc. of MM, 2014.

[4] Q. Li, Y. Xue, J. Jia, and L. Feng, "Helping teenagers relieve psychological pressures: A microblog based system," in EDBT, 2014.

[5] L.P. Spear, "The adolescent brain and age-related behavioural manifestations",2000

[6] Yuanyuan Xue, Qi Li , Li Jin, Ling Feng, David A. Clifton,Gari D. Clifford ,"Detecting Adolescent Psychological Pressures from Micro-Blog", 2013.

[7] Les Servi and Sara Beth Elson, "A Mathematical Approach to Gauging Influence by Identifying Shifts in the Emotions of Social Media Users", 2014

[8] Aditya Mogadala, "Twitter User Behavior Understanding with Mood Transition Prediction" ,20