International Journal of Future Generation Communication and Networking
Vol. 13, No. 3, (2020), pp. 1701–1706

Human Stress Detection Based on Social Media Interaction Using Machine Learning

Pragati M Ramteke, Nisha Balani, Dr. Pradnya Borkar

Computer Science & Engineering, Jhulelal Institute of Technology, Nagpur, Maharashtra.

Abstract

Now a day's most of the people's faces stress that will lead to psychological problem. Hence it is vital to detect causes of stress before it turns to a big health problem. Generally BP, heart attack or some death are occurred because of taking excessive stress. Detecting stress of human we use social media data such as post on Facebook page, twits on twitter etc. because people share their feelings on social media hence it becomes easy to get social data to find stress based on their behavior. Since regular methods are very time consuming and costly. We utilize tweets of user posted on media from twitter dataset. The states of user stress are classified using Convolutional Neural Network (CNN) algorithm and are divided as stressed or non-stressed user.

Index Terms: *Stress detection, social media, Convolutional Neural Network, healthcare, data mining, social interaction.*

I. INTRODUCTION

The social networks are developed to share images, videos and reviews between the friends. The users can connect from anywhere and anytime. The social media data values are analyzed to calculate the user's behavior and stress levels. The relationship of user interactions and stress are analyzed for the stress detection process. Stress-related textual, visual and social attributes are extracted for the stress recognition process. Proposed a novel hybrid model which combines factor graph model and Convolutional Neural Network (CNN) to pull tweet content and social interaction data for stress detection.

Psychological stress is becoming a major problem in people's health nowadays more peoples are feeling stressed. Stress is non-clinical and common in every one's life, excessive stress can be vulnerable to people's physical and mental health and excessive stress is considered to be a major cause of suicide. Thus, it is crucial to detect stress before it turns into severe health issue.

The social media is grown to change people's life, as well as research in healthcare and wellness with the development of social networks like Twitter most of the peoples share their daily events and moods, and interact with friends through the social media.

As these social media data timely reflect users' real-life states and emotions in a timely manner, it offers new opportunities for representing, measuring, modeling, and mining user's behavior patterns through the large-scale social networks, and such social data can detect its theoretical basis in psychology study.

Most of opinion analysis makes utilization of emoticons as labels to decrease dependency in machine learning techniques for sentiment classification. Twitter hashtags and smileys to increase sentiment learning. Improved target-based Twitter opinion or sentiment classification by taking target-based features and related tweets taken into account. Used machine learning techniques such as CNN to classify social data for stress detection. A CNN classifier is an algorithm that uses Neural Network to classify objects. CNN classifiers will assume strong, independence between attributes present in the data points. CNN classifiers are widely utilized for machine learning because they are simple in implementation.

The section I explains the Introduction ofhuman stress detection using classification method such as CNN. Section II presents the literature review of existing systems and Section III present proposed system implementation details Section IV presents experimental analysis, results and discussion of proposed system. Section V concludes our proposed system. While at the end list of references paper are presented.

II. LITERATURE REVIEW

The aim of Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland [1] to investigate the automatic recognize daily stress of people's from three sets of data: people activity detected on their smartphones, weather conditions and personality traits. Proposed multifactorial statistical model for detecting stress in independent individual by using activity performed on mobile.

Huijie Lin et al. [2] present user-level psychological stress detection from social media utilizing deep neural network. Here author employs real online microblog data to detect the correlations among users' stress and their tweet content. It also states two kinds 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 kind and linguistic styles.

Xiaojun Chang, Yi Yang, Alexander G Hauptmann, Eric P Xing, and Yao-Liang Yu [3] objective is to detect complex events in unconstrained Internet videos. Since most of previous system utilizes labeled training data, they take a more difficult zero-shot setting where training data is not supplied. They applied concept of classifiers on all test videos to obtain multiple score vectors. These vectors are converted into pairwise comparison matrices and the nuclear norm rank aggregation framework is adopted to seek consensus. To address the challenges author proposed an efficient, highly scalable algorithm that is an order of magnitude has more speed than existing alternatives.

Li-fang Zhang et al. [4] proposed titled- Occupational stress and teaching methods in Chinese academics. Author suggested that, controlling the self-rating abilities of the participants, the favorable conceptual changes in teaching approach and their role insufficiency predicated that the conceptual change in teaching strategy is negative.

Qi Li, Yuan yuan Xue, Jia Jia, Ling Feng [5] proposed tHelper for sensing and easing teenagers' psychological pressures in study, communication, affection, or self-recognition through micro-blog. The system use Gaussian Process to classify a teenager's stress depends on a number of features extracted from user's tweets. The system relate with parents to advise them about their children's' stress level, utilizing Mobile SMS. It also submit positive stories to detected stressed children for reducing their stress.

Aditya Mogadala [6] utilizes history of mood data to predict future moods. It develop regression analysis techniques on Tweets. It states mood labels and present mood swings depends on it. Stanford SNAP database is utilized for development of this system. It returns generalized results for moods.

Yuan yuan Xue, Qi Li , Li Jin, Ling Feng, David A. Clifton, Gari D. Clifford [7] states that the general face-to-face psychological detection and treatment cannot get the demand of relieving teenagers' stress completely because of its lack of timeliness and diversity. A microblog platform is anticipated to sense psychological pressures via teenagers' tweets, and assist teenagers how to release their stress via micro-blog. It utilized five classifiers for tweet analysis, Naive Bayes, Support Vector Machines, Artificial Neural Network, Random Forest, and Gaussian Process Classifier for stress or pressure detection.

III. SYSTEM ARCHITECTURE

A. System Architecture

Here Twitter dataset is taken as an input to detect human stress then preprocessing is done on that dataset to remove stemming and stop words. Then POS tagging applied on files with removed stemming and stop words. After that linguistic features are extracted based on that generated training file and then we apply CNN classification to classify training file and to find stress. Then in tweet behavior we choose tweet of user to identify its behavior or stress level. We generally use classes to define stress level that is class 0, class 1 and class 2 for positive, negative and neutral stress.

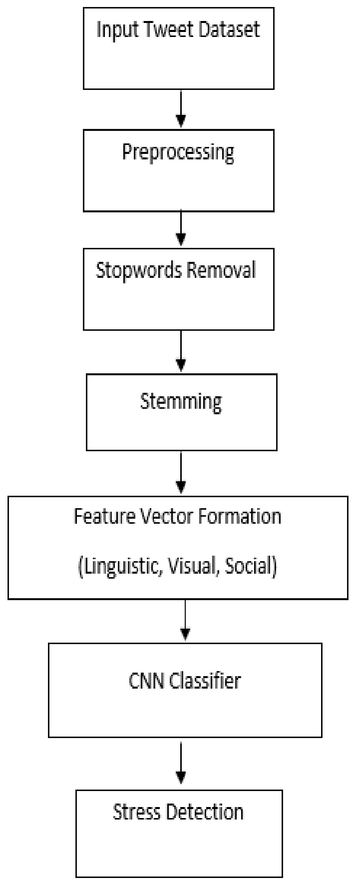

Fig 1. System Architecture

B. Algorithm

Step 1:Convolution
A convolution is a joined integration of two methods that demonstrates to you how one methods changes the other.

ISSN: 2233-7857 IJFGCN

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t-\tau)\, d\tau$$
$$= \int_{-\infty}^{\infty} f(t-\tau) g(\tau)\, d\tau.$$

Step 2: Apply the ReLu (Rectified Linear Unit)

In this progression we apply the rectifier function to increment non-linearity in the CNN. Dataset are made of various items that are not linear to one another. Without applying this function, the information grouping will be treated as a linear issue while it is really a non-straight one.

Step 3: Pooling

Spatial invariance is a concept where the location of an object in a dataset doesn't affect the ability of the neural network to detect its specific features. Pooling enables the CNN to detect features in various data. There are different types of pooling, for example, max pooling and min pooling. Max pooling works by placing a matrix of 2x2 on the feature map and picking the largest value in that box. The 2x2 matrix is moved from left to right through the entire feature map picking the largest value in each pass.These values then form a new matrix called a pooled feature map. Max pooling works to preserve the main features while also reducing the size of the data. This helps reduce overfitting, which would occur if the CNN is given too much information, especially if that information is not relevant in classifying the data.

Step 4: Flattening

Once the pooled featured map is obtained, the next step is to flatten it. Flatteninginvolves transforming the entire pooled feature map matrix into a single column which is then fed to the neural network for processing.

Step 5: Full connection

After flattening, the flattened feature map is passed through a neural network. This step is made up of the input layer, the fully connected layer, and the output layer. The fully connected layer is similar to the hidden layer in ANNs but in this case it's fully connected. The output layer is where we get the predicted classes. The information is passed through the network and the error of prediction is calculated. The error is then backpropagated through the system to improve the prediction.The final figures produced by the neural network don't usually add up to one. However, it is important that these figures are brought down to numbers between zero and one, which represent the probability of each class. This is the role of the Softmax function.

$$\sigma : \mathbb{R}^K \to (0,1)^K$$
$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

IV. Result And Discussions

A. Experimental Setup

I. All the experimental cases are implemented in Java in congestion with Netbeans tools and MySql as backend, algorithms and strategies, and the competing classification approach along with various feature extraction technique, and run in environment with System having configuration of Intel Core i5-6200U, 2.30 GHz Windows 10 (64 bit) machine with 8GB of RAM

ISSN: 2233-7857 IJFGCN

II. Dataset Description: The data was taken from twitter dataset, twitter dataset taken from kaggle and is in the form of no, userid, time and tweet.

B. Analytical Result

This section presents the tweet behaviour of user. Fig 6.1 behaviour graph of tweets. Here we utilize pie chart to indicate tweet behaviour. Class 0, class 1 and class 2 indicate user behaviour as positive behaviour, negative behaviour and neutral behaviour.

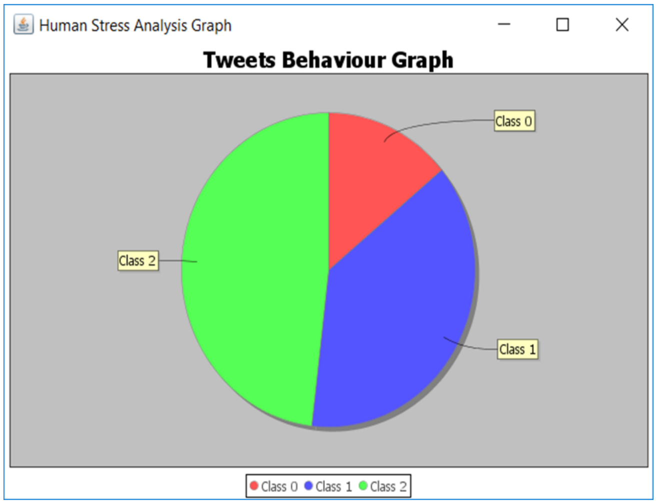

Figure 2. Result Analysis

Conclusion

Human stress is growing now a days and that leads to major health issue. So need to reduce stress for that detection of stress is vital. Here we propose a model to detect human stress based on social data because peoples generally prefer to share their feelings on social media via tweets or post. We utilize twitter dataset to detect user behaviour and applied CNN classifier to classify the training data. Because data on social media is vastly increasing by using traditional method cannot possible to classify large data files so we utilize classifier to classify training file. Result of user tweet shows as class 0, class1 and class 2. Class 0 indicate positive stress level, Class 1 indicate negative stress level and class 2 indicate neutral stress level.

REFERENCES

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

[2] Huijie Lin, Jia Jia, Jiezhon Qiu, Yongfeng Zhang, Lexing Xie, Jie Tang, Ling Feng, and Tat-Seng Chua, "Detecting Stress Based on Social Interactions inSocial Networks", IEEE Transactions on Knowledge and Data Engineering, 2017.

[3] Xiaojun Chang, Yi Yang, Alexander G Hauptmann, Eric P Xing, and Yao-Liang Yu. Semantic concept discovery for large-scale zero-shot event detection. In Proceedings of International Joint Conference on Artificial Intelligence, pages 2234–2240, 2015.

[4] Li-fang Zhang, -Occupational stress teaching approach among Chinese academics, Educational Psychology, Volume 29, Issue 2, March 2009, pages 203 – 219.

[5] Q. Li, Y. Xue, J. Jia, and L. Feng, “Helping teenagers relieve psychological pressures: A microblog based system,” in EDBT, 2014.
[6] Aditya Mogadala, “Twitter User Behavior Understanding with Mood Transition Prediction” 20
[7] Yuanyuan Xue, Qi Li , Li Jin, Ling Feng, David A. Clifton,Gari D. Clifford ,“Detecting Adolescent Psychological Pressures from Micro-Blog”, 2013.
[4] Wanxiang Che, Zhenghua Li, and Ting Liu. Ltp: A Chinese language technology platform. In Proceedings of International Conference on Computational Linguistics, pages 13–16, 2010.
[5] Chih chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY, 2(3):389–396, 2001.
[6] 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.
[7] Sheldon Cohen and Thomas A. W. Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98(2):310–357, 1985.
[8] Glen Coppersmith, Craig Harman, and Mark Dredze. Measuring post-traumatic stress disorder in twitter. In Proceedings of the International Conference on Weblogs and Social Media, pages 579–582, 2014.
[9] Rui Fan, Jichang Zhao, Yan Chen, and Ke Xu. Anger is more influential than joy: Sentiment correlation in weibo. PLoS ONE, 2014.
[10] Zhanpeng Fang, Xinyu Zhou, Jie Tang, Wei Shao, A.C.M. Fong, Longjun Sun, Ying Ding, Ling Zhou, , and Jarder Luo. Modeling paying behavior in game social networks. In In Proceedings of the Twenty-Third Conference on Information and Knowledge Management (CIKM’14), pages 411–420, 2014.
[11] Golnoosh Farnadi, Geetha Sitaraman, Shanu Sushmita, Fabio Celli, Michal Kosinski, David Stillwell, Sergio Davalos, Marie Francine Moens, and Martine De Cock. Computational personality recognition in social media. User Modeling and User- Adapted Interaction, pages 1–34, 2016.
[12] Eileen Fischer and A. Rebecca Reuber. Social interaction via new social media: (how) can interactions on twitter affect effectual thinking and behavior? Journal of Business Venturing, 26(1):1–18, 2011.
[13] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5):1189–1232, 1999.
[14] Rui Gao, Bibo Hao, He Li, Yusong Gao, and Tingshao Zhu. Developing simplified chinese psychological linguistic analysis dictionary for microblog. pages 359–368, 2013.
[15] Johannes Gettinger and Sabine T. Koeszegi. More Than Words: The Effect of Emoticons in Electronic Negotiations.