



Designing A Machine Learning System For Detecting The Stress In IT Professionals

P.Bhaskar¹, G.Eswaramma²

¹ Professor in the Department of CSE/MCA at QIS College of Engineering and Technology(Autonomous), Vengamukkapalem, Ongole-523272, Prakasam Dt., AP

² MCA Student in the Department of MCA at QIS College of Engineering and Technology(Autonomous), Vengamukkapalem, Ongole-523272, Prakasam Dt., AP.

ABSTRACT_ Stress disorders are a very common occurrence among employees who work in corporate sectors. We can observe an increase in stress among employees as a result of people's changing jobs and lifestyles. Even though many corporate sectors offer a range of mental health programs and work to lessen stress-related disorders in the workplace, the problem is far from being resolved. In our paper, we'll use two machine learning techniques to identify the stress levels of employees working in corporate sectors and to try to focus on the problems that cause those levels of stress. Once the preprocessing and cleaning of the data is complete, we will apply two machine learning techniques (SVM and Random Forest). Our trained model's accuracy was easily read and evaluated. Sex, family history, and ease of access to health benefits at work are found to be the three main factors that lead to stress disorders by using these two machine learning techniques. With these outcomes, corporate industries are now able to reduce stress and create a very welcoming workplace for their employees.

1.INTRODUCTION

Unwinding gives wellbeing while stress influences your health[1]. Situation Stress is increasing the fastest right now. This is why, despite prosperity, people are unhappy. Stress is a compressed inclination. The stress can be mental, physical, or even emotional. Emotional, physical, or even mental pressure can occur whenever someone is under pressure. One feels like they would fail or not succeed when under pressure[2]. Stress can be found anywhere. It is found in every

single character. Stress can be brought on by any thought or situation. However, nature is distinct. There are many different types of stress. It's mental, physical, and financial. Stress can result from any thought or circumstance[2]. Under stress, one's denial gradually rises in multiples. Personality shifts dramatically in these circumstances. Stress is also caused by the inferior complex and superior complex. An upset man can barely do anything effectively. Stress can be beneficial or, more often than not, harmful. Valuable pressure makes you work, it keeps



you dynamic, occupied and spurred. Negative stress, on the other hand, makes you dull, inactive, afraid, and lonely. So far, there are two types of stress: short-term stress and long-term stress, also known as chronic stress[2]. Situational short-term stress disappears as soon as the situation changes. A problem that lasts a long time is stress. Because of this, it is dangerous. Chronic stress can sometimes be mostly inherited, or it can be linked to genes. It can end to any person's life. One dies from excessive stress. Tolerating pressure is the initial step to lead tranquil blissful life[2]. Before stress becomes acute, it is critical to identify and manage it[3]. Simple, cost-effective remedies are available. Since a long time ago, studies have been conducted to identify people who are stressed. There is a lot of literature on stress detection. There are scientific and traditional methods for identifying stressed individuals. A) Survey: A large questionnaire is provided by the psychiatrist, and the results are used to determine whether or not the patient is experiencing stress. Because frequently the responses are not factual, this method has its own limitations and drawbacks. The questionnaire's questions can sometimes be inappropriate. B) The sensor measuring method is the alternative method. The limit of this technique is, the time has come

consuming and a piece costly. Social media is another, more recent, approach to stress detection[1]. Social media makes it possible to identify stress. A person's posts on social media, how they respond to a particular issue, and the things they like and dislike on social media all contribute to the detection of stress. A psychiatrist can find people who are under pressure, crazy, or mad after a typical subject by reading the constant posts on social media. Additionally, this is a sign of stress. The state of one's mind and thinking can be analyzed through the use of social networking sites[5]. Twitter and Facebook have a wide number of clients. Individuals share their own contemplations, thoughts, belief system, state of mind and so on. Whether a writer is normal or abnormal is evident in their comments on Twitter or Facebook. Furthermore, unusual is doubtlessly under pressure. The individual's comment reveals his personality. The property of tweeter is short, casual and restricted characters. 58 million tweets are sent out every day on Twitter, which is beyond comprehension. Tweets draw out one's own internal, character center, feelings and tension. Tweets describe what they are made of and how they look. One's tweet can without much of a stretch demonstrate whether one is loose or like freedom or under tension. This is early recognition of



stress which in the drawn out will be useful to stay away from the extreme issue.

2. LITERATURE SURVEY

2.1 Measuring Post Traumatic Stress Disorder in Twitter. Glen Coppersmith, Mark Dredze, and Craig Harman. 2014.

Traditional mental health studies rely on information primarily collected

through personal contact with a health care professional.

Recent work has shown the utility of social adequate treatment. This position paper highlights some recent attempts examining the potential for leveraging social media postings as a new type of lens in media data for studying depression, but there have been limited evaluations of

and populations. Information gleaned from social media bears potential to complement traditional survey techniques in its ability to provide finer grained measurements of other mental health

conditions. We consider post traumatic stress disorder (PTSD), a serious condition that affects millions worldwide, with especially high rates in military veterans.

We also present a novel method to obtain a PTSD classifier for social media using simple searches of available Twitter data, a significant reduction in training data cost

compared to previous work. We demonstrate its utility by examining differences in language use between PTSD and random individuals, building classifiers to separate these two groups and by detecting elevated rates of PTSD at and around U.S. military bases using our classifiers. Introduction Mental health conditions affect a significant percentage of the U.S. adult population each year, including depression (6.7%), eating disorders like anorexia and bulimia (1.6%), bipolar disorder (2.6%) and post traumatic stress disorder (PTSD) (3.5%).

2.2 Role of Social Media in Tackling Challenges in Mental Health. Munmun De Choudhury. 2013.

Mental illness is a serious and widespread health challenge in our society today. Tens of millions of people each year suffer from depression and only a fraction receives

understanding mental illness in individuals behavior over time while radically

expanding population sample sizes. We for identifying the onset of depressive disorders in individuals, for use by healthcare agencies; or on behalf of individuals, enabling those suffering from mental illness to be more proactive about their mental health.



3. PROPOSED SYSTEM

In this paper author is describing concept to detect employee stress by using machine learning algorithms such as SVM and Random Forest Algorithms. To detect stress author is using social media dataset such as tweets where employee can share their views and by analysing this views we can identify whether employee is in relax or stress mood but by analysing this views manually may take lot of human efforts so author using machine learning algorithms and the experiment with this algorithms show stress detection accuracy more than 90%.

To implement this project we are using twitter tweets dataset and this dataset is pre-process with NLTK (natural language tool kit) to remove stop words and special characters. Processed dataset will be splitted to train and test part where application used 80% dataset for machine learning model training and 20% dataset is used to evaluate or test machine learning prediction accuracy.

Random Forest: This algorithms first select random values and then start looking for class which is close to that random values and assign that class to that values and conclude highlighting how this research this

direction may be useful in developing tools

process continues till all classes assign with closer random values.

SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes. In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. In particular, it is commonly used in support vector machine classification. As a simple example, for a classification task with only two features (like the image above), you can think of a hyperplane as a line that linearly separates and classifies a set of data.

Intuitively, the further from the hyperplane our data points lie, the more confident we are that they have been correctly classified. We therefore want our data points to be as far away from the hyperplane as possible, while still being on the correct side of it.

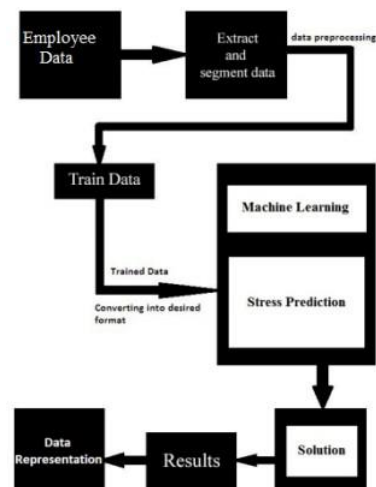


Fig 1:Architecture

3.1 IMPLEMENTATION

Gathering the datasets: We gather all the r data from the kaggle website and upload to the proposed model

Generate Train & Test Model: We have to preprocess the gathered data and then we have to split the data into two parts training data with 80% and test data with 20%

Run Algorithms: For prediction apply the machine learning models on the dataset by splitting the datasets in to 70 to 80 % of training with these models and 30 to 20 % of testing for predicting

Input data: In this module we will give test data as our input based on we will get output

Predict output: in this module we will show output we got stress detection accuracymore than 90%.

4. RESULTS AND DISCUSSION

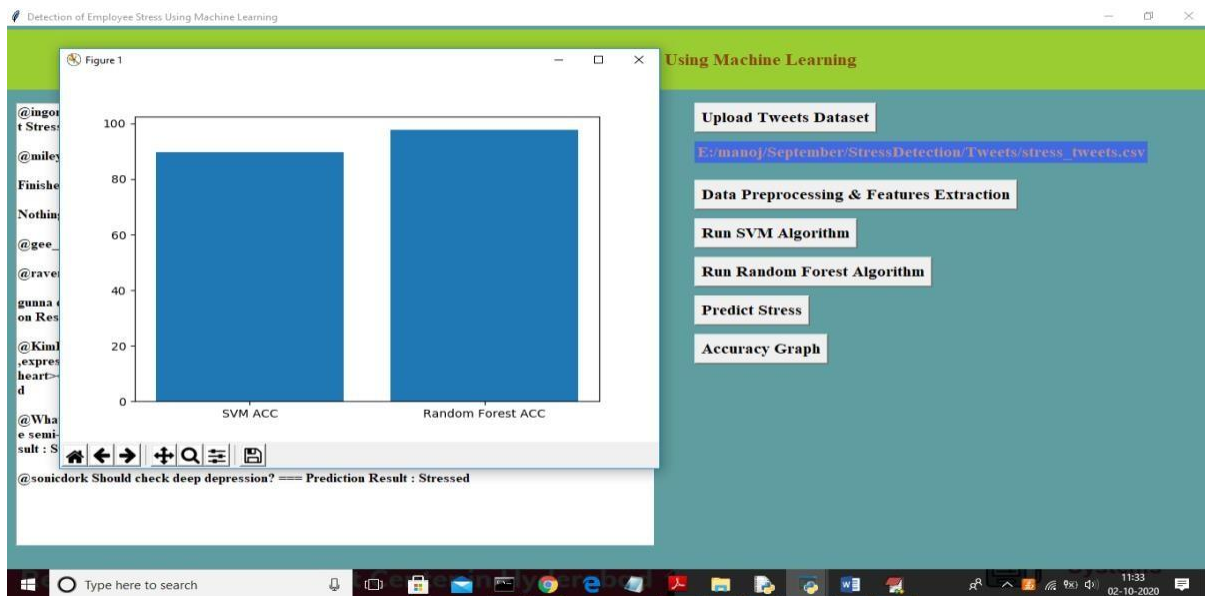


Fig 2: In above x-axis represents algorithm name and y-axis represents accuracy of those algorithms and from above graph we can say random forest is better than SVM

Sno	Algorithm Name	Accuracy
1	SVM	88
2	Random Forest	99

5. CONCLUSION

When compared to the other factors that affect whether an employee can experience conceptual health-related issues, gender, the family history of the illness, and whether a particular employer provides the benefits of health for their employees were of greater significance. In spite of the fact that their jobs were not tech-related, our study revealed that those who work for tech companies are more

likely to experience stress. Business organizations could effectively use these perceptions to develop HR strategies that working employees would find more appealing. A 75% correctness rate demonstrates that using two machine learning techniques (SVM and Random forest) to predict stress and conceptual health conditions yields results that are deserving of further investigation and thus fulfill the purpose of this paper.



FUTURE SCOPE

Since this paper is associated with the social problem which is one of the enormously growing field, the scope is pretty high and it helps the society in a way which can identify the victims of stress which is one of the most commonly identified disorder among the adolescents.

REFERENCES

[1] Detecting and characterizing Mental Health Related Self-Disclosure in Social Media. Sairam Balani and Munmun De Choudhury. 2015. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems -CHI EA '15, pages 1373–1378.

[2] Measuring Post Traumatic Stress Disorder in Twitter. Glen Coppersmith, Mark Dredze, and Craig Harman. 2014.

[3] Role of Social Media in Tackling Challenges in Mental Health. Munmun De Choudhury. 2013.

[4] Bhattacharyya, R., & Basu, S. (2018). India Inc looks to deal with rising stress in employees. Retrieved from „The Economic Times“

[5] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning*

research, 12(Oct), 2825-2

[6] OSMI Mental Health in Tech Survey Dataset, 2017 from Kaggle.

[7] Van den Broeck, J., Cunningham, S. A., Eeckels, R., & Herbst, K. (2005). Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS medicine*, 2(10), e267.

[8] Relationship between Job Stress and Self-Rated Health among Japanese Full Time Occupational Physicians Takashi Shimizu and Shoji Nagata 2007 Academic Papers in Japanese 2007.

[9] Tomar, D., & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology*, 5(5), 241-266.

[10] Gender and Stress. (n.d.). Retrieved from APA press release 2010

[11] Julie Aitken Harris, Robert Saltstone and Maryann Fraboni. (2000) An Evaluation of the Job Stress Questionnaire with a Sample of Entrepreneurs” 2000 JSQ scale Entrepreneurs.

[12] “Demographic and Workplace Characteristics which add to the Prediction of Stress and Job Satisfaction within the Police Workplace” , Jeremy D. Davey, Patricia L. Obst, and Mary C. Sheehan 2015 IEEE 14th International



Authors Profiles



Dr.P.Bhaskar Professor in the Department of CSE/MCA at QIS College of Engineering and Technology.(Autonomous),

Vengamukkapalem, Prakasam (Dt.) He is having 20 Years of Teaching

Experience & 13 years of Research Experience and Published more than 20 Research Publications & his area of interests is Artificial Intelligence and Image processing& Biometric Systems.



G.Eswaramma PG Scholar in the department of MCA Qis College of Engineering and Technology(autonomous), Vengamukkapalem, Prakasam (DT).

She has completed B.Sc in Computer Science from Krishnaveni Degree College Areas of Interests Machine Learning.