

Detection of Employee Stress Using Machine Learning

¹Anitha.K, ²Kiran Kumar, ³Kishora GV, ⁴Balaji S, ⁵Likith Reddy NA

¹Asst.Professor, ^{1,2,3,4,5}School of C&IT, REVA University, Bengaluru, India
¹anitha.k@reva.edu.in, ²kirankumar7256@gmail.com, ³gvkishor61998@gmail.com,
⁴balajisreenivas07@gmail.com, ⁵likithreddyna@gmail.com

Article Info

Volume 83

Page Number: 5023-5026

Publication Issue:

May - June 2020

Abstract

Disorders of stress are very casual thing among the employees who are working in corporate sectors. As with changing work of people and their living lifestyle, we can see the increment of stress in the working employees. Even many corporate sectors are providing variety of schemes related to mental health and trying to reduce the disorders of stress in the working environment, the disorder is very far from stopping. In our paper, we are going to make use of two techniques of machines to determine the amount of stress the employee is having who is working in corporate sectors and try to narrow down the issues that identify the stress levels. We are going to apply two techniques of machine learning (i.e. SVM and Random Forest) when the data preprocessing and the cleaning of data is once finished. The correctness of our trained model was clearly read and analyzed. By using these two techniques of machine learning, the main features that result in disorders of stress are found to be as sex, background of family and ease of benefits of health in the working place of employee. With these results, corporate industries can now narrow down the stress and can establish a very friendly working place for the corporate sectors employees.

Article History

Article Received: 19 November 2019

Revised: 27 January 2020

Accepted: 24 February 2020

Publication: 16 May 2020

Keywords: Stress detection, Healthcare Management, Machine Learning algorithms

1. Introduction

Disorders of stress which are related to mental health are not rare for the employees working in corporate sectors. Some analysis done earlier have created some concern on the very same. Based on the work done by Association of Industry, Assocham, we come to know that above 42% of the professional working employees in the corporate private sectors of India are suffering from stress or common disorders of anxiety because of late night working hours and also due to fixed timings. This part of singles are growing as mentioned in the Economic Times of 2018 article which is dependent on the survey that was managed by the Optum[4]. There is a survey that considers the replies of nearly eight lakh working employees who are working from more than seventy huge companies, with each single company having its employees more than 4,500 working professionals. The

workplace which is free from stress must be given at most importance for higher productivity and happy living for the working employees. There are many steps which we can take to help the employees come up with the disorder of stress for well-being of the mental health like assistance for counselling, guidance given for the career, sessions for management of stress, and creating an awareness of health identification of working employees who will need such kind of help will definitely improve the rates of such kind of measures for becoming victorious. We try to make this happen by using our machine learning techniques to overcome with a model that predicts the rate of the stress that is accomplished. This approach is not only going to help company HR managers to know better about their working professionals, it will also help in taking proper precautions to reduce the chances of stress in their working employees.

2. Literature Survey

Here, literature review on different techniques given by various researchers is being presented.

[1] Self leak is a major component for promoting better well being in singles with severe conceptual disorders. As social networking applications has enlargingly embraced in conceptual health associated conversations, we are going to examine how some modern policies may allow the truthful and informal policies that express the thoughts, encounters and faith. Clearly, we are seeking to determine stages of self leak which are cleared in shafts which have been sent on variety of conceptual health accomplished conventions on social news websites. We evolved and arranged for a cause that is dependent on gratified factors. The divider is now capable of distinguishing a social news website post with a low, high, or no own Revelation with an accuracy of 78% . Using this divider for common conceptual health leaks on social news website, we will obtain that the mass of such kind of is now distinguished by major self-leaks, and the group that replies typically to shafts that reveal more or less. We finally conclude with possibility of tackling our initiated detection of self-leaks algorithm in mental treatment through social networking websites.

[2] Conventional conceptual health learnings depend on the facts that are firstly gathered through private contact with an executive who is in charge for the health care. Modern work has show cased the usage of data available in the social media for learning depression, but there were finite assessments for the remaining conceptual health conditions. We are going to examine PTSD (Post Traumatic Stress Disorder), which is a crucial condition that effects millions of people throughout the world, with mainly having high rates in the old retired military soldiers. We would also showcase a book method for obtaining the PTSD distinguisher for social networking apps using casual searches that are accessible for the data available in twitter, a important depletion in the training data will cost in contrast to the past work that has been done. We would like to explain its usage by inspecting dissimilarities in speech that are used between random individuals and PTSD, constructing distinguishers to differentiate these two categories and by detecting raising rates for PTSD at and over military bases in U.S. which use our distinguishers.

[3] conceptual illness is a major and universal health provocative in current Civilization. Every year more than million no of people under go depression but only few among them are capable of recovery. This document focus on some present strives by analysing the possible for buying social networking apps posts as a new way of grasping conceptual illness in community and also in every single person. The data gathered from social media convey possible way to the compliment for usual survey methods are capable of serving efficient value of performance above time interval completely growing community of sample measurements. We deduce focusing how this work will help in useful in improving

tools for recognizing the initial stage of depression in a single person, for health care organization or as a representative of a person authorizing those who are facing conceptual illness to be more resourceful regarding conceptual illness.

3. Proposed System

We have few systems which uses basic algorithms .they attempt to detect the problems through tweets but yielding less efficiency. existing system uses Naive Baye's algorithms, Support vector Machine, Gaussian Classifiers etc, but the disadvantage being they need lot of Data for training which apparently takes more execution time and also yields less efficient outputs. In this paper, we propose an efficient system which is uses Machine learning algorithms SVM and Random Forest it give the best accuracy compare to Existing system. In this system we use the Employee Twitter data for predicting the employee is stress are not.

4. Methodology

Support Vector Machine (SVM)

Support-vector machine builds a hyperplane or group of hyperplanes in a high- or boundless-dimensional space, that could be used for regression, classification or other function such as outliers identification. Instinctively, a great segregation is fulfilled by the hyperplane which has a large span to the nearest trained-data point which is of any kind of a class (commonly known as a Useful margin), familiarly, the smaller the generalisation fault for the classifier ,the larger the edge of the classifier. Fig 4.1, shows the data points containing the hyperplane between them.

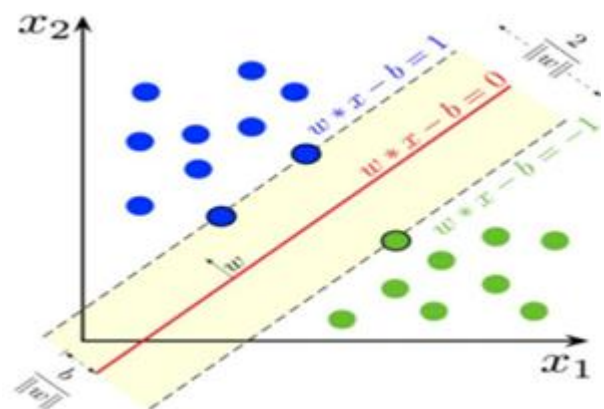


Figure 4.1: Hyperplane containing data points showcasing SVM.

Random Forest

Random forests which are also known as random decision forests are a set of learning approach for regression, classification and other functions that performs by constructing a lot of decision trees during training period

and turns out the class which is a way for all the classes i.e. categorisation or mean forecast i.e. regression of the single trees. Random forests rectify and fixes the decision tree's routine of over fitting for their training group. Fig 4.2, shows the decision tree for random forest algorithm containing instances and leaf nodes.

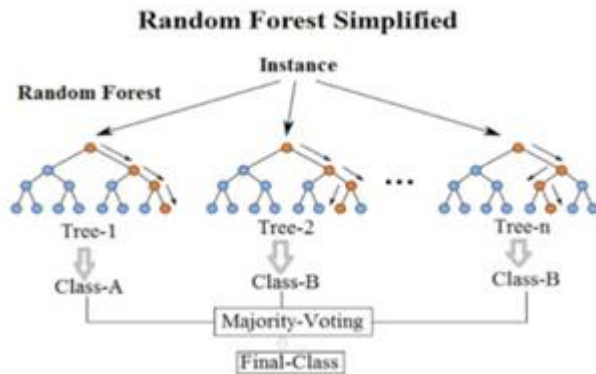


Figure 4.2: Random Forest Decision Tree

5. Architecture Diagram

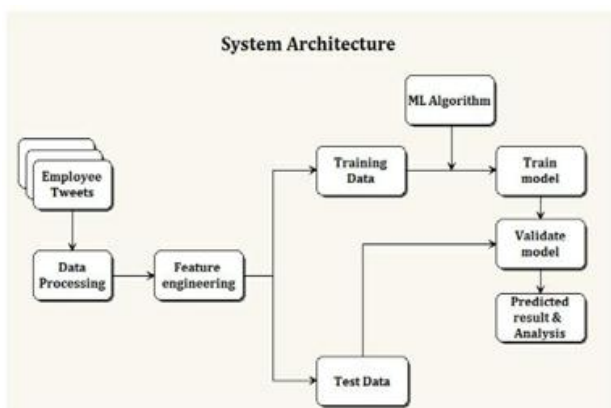


Figure 5.1: System Architecture for predicting the stress in an employee using employee tweets

As the Fig 5.1 depicts, we are going to take the employee tweets from their Twitter data. After the employee tweets is collected we perform data preprocessing to remove the unnecessary special characters from the data. This can be done by using the Natural Language Preprocessing (NLP). Once NLP processing is done we are going to perform feature engineering. By performing feature engineering we can obtain the accurate structure of our training data and also obtain the good model. Here we are splitting our data into training and test data. Now, we are going to apply the machine learning algorithm on the training data for training our model to obtain the result. After the trained model is once obtained validate the trained model with the test model which was obtained in the feature engineering process. Once the validation is done we are now having the predicted and analysed result for the data

i.e. employee tweets that is collected from the employee which shows how much stress the employee is facing.

6. Expected Outcome and Results

All the analysed models are Executed using python with the help of scikit-learn[5] to check whether the individual requires a therapy or not. The tabular column for the visualized results are noted. The diagram illustrates the characteristic significance of the fourteen aspects considered, which is attained by decision tree. The bar chart explains that gender is extremely affected on stress and conceptual health between selected variables. It can be considered by the reality that female has more mental stress compared to male as shown in the Fig.6.1. It will help in discovery of 2010 survey proposed by American Psychological Association on the interconnection between the stress and gender.

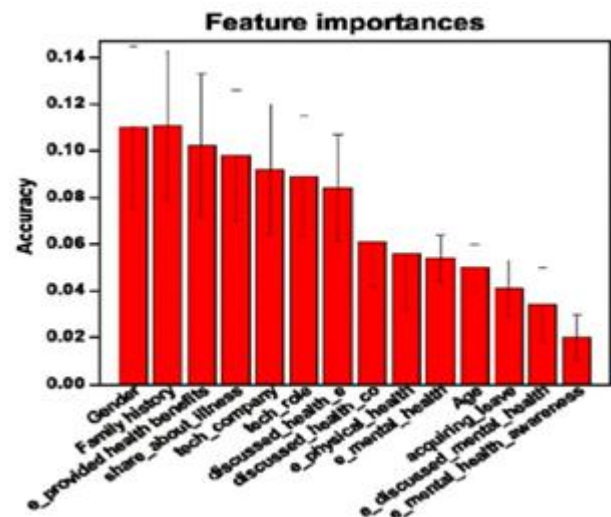


Figure 6.1: Significance of various aspects that are being considered

7. Conclusion

Gender, also the family background which has the illness, and considering whether a single employer provides the conceptual benefits of health for their employees was having more significance compared to the other factors for determining whether an employee can obtain conceptual health associated issues. From our study, we were able to find that the people who are working in the tech companies are at more risk of obtaining stress, even though their job role was not based on tech. These perceptions could be successfully used by business companies to make more desirable HR strategies for the working employees. A 75% correctness shows that the application of two Machine Learning techniques (i.e. SVM and Random forest) for predicting the stress and conceptual health conditions provides worthy results and could be searched further, and thus meets the aim of this paper.

8. Future Scope of Work

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. SairamBalani 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 Conference on Cognitive Informatics & Cognitive Computing (ICCIIC). 2015.
- [13] Mario Salai, Istv an Vass anyi, and Istv an Kosa, “Stress Detection using low cost Heart rate sensors”, Journal of Healthcare Engineering, pp.1-13,Hindawi Publishing corporation , 2016
- [14] Shwetha, S, Sahil, A, Anant Kumar J, (2017) Predictive analysis using classification techniques in healthcare domain, International Journal of Linguistics & Computing Research, ISSN: 2456-8848, Vol. I, Issue.I, June-2017.
- [15] O.M.Mozos et al, “Stress detection using wearable physiological and sociometric sensors”. International Journal of Neural Systems,vol 27,issue 2, 2017.