PRINCIPLE OF SENTIMENT ANALYSIS USING CNN AND ONLINE SOCIAL NETWORK DATA, A TECHNOLOGICAL METHOD TO STRESS DETECTION

[1]T Radhika, [2]Ch Sandhya, [3]Punna Mahesh, [4]K Manohar Reddy
[1,2,3]Assistant Professor, [4]UG Student,
[1,2,3,4]Department of Computer Science and Engineering, Brilliant Institute of Engineering and Technology, Hyderabad, India.

Abstract
According to the research, stress is often a human's response to various dangers or wants. When working properly, this reaction may help us stay focused, motivated, and cognitively engaged, but if it becomes out of control, it can be hazardous and result in depression, anxiety, hypertension, and a variety of other life-threatening conditions. Cyberspace is a huge platform for individuals to communicate everything and everything they experience in their daily lives. Then, depending on the posts and status updates the person provides, it may be utilized as a very effective approach to determine that person's degree of stress. This is a suggestion for a website that accepts the subject's Twitter username as an input, scans and analyses the topic's profile using sentiment analysis, and then displays the findings. These findings show the subject's overall stress levels and provide an overview of its mental and emotional condition. Here in this project we are aiming the CNN deep learning architecture for the stress of evolution in online social networks such as Twitter.

Keywords: Social media, factor graph model, social interaction, stress detection, health care

Introduction

Data mining, also known as information exploration, is a computer-assisted method of sifting through and analyzing enormous collections of information in order to discover their underlying meaning. Tools for data mining that forecast consumer preferences and future trends, enabling firms to make strategic decisions under the guidance of experience. Mark e tissues that have previously taken so long to solve may now be answered by data mining technologies. They search through hidden trend databases to find forecasting information that experts would pass over since it contradicts their presumptions. The act of analyzing and synthesizing data from many angles to provide useful information [7] that can be used to increase revenue, decrease expenses, or do both is commonly referred to as data mining. Software for data mining is one of several analytical methods for data analysis. It allows users to examine, categories, and summaries data-related relationships from a variety of perspectives.

Data mining is the process of identifying commonalities [17] or patterns between thousands of fields in large databases. Despite the fact that data mining may yield similar results in the long run, the innovation organization is not. For a long time, organizations have used powerful machines to sift through large amounts of general store scanner data and deconstruct statistical surveying reports. However, nonstop advancements to workstation preparing power, disc storage, and statistical methods, while driving down the cost, are dramatically increasing research accuracy. Even though data mining is still in its infancy ,data mining tools and techniques a real ready being used by companies in a wide range of industries to take advantage of historical data, including retail, finance, healthcare, manufacturing, transport, and aerospace. Data mining serves examiners with comprehend vital information, relationships, trends, patterns[18] [26], exceptions Furthermore anomalies that might overall try unnoticed with filter through warehoused learning Eventually Tom's perusing utilizing example distinguishment advances Furthermore measurable Furthermore scientific systems. Information mining will be used to recognizing examples Furthermore correlations in the information for organizations in place should help make educated business choices. Information mining will encourage with spot bargains trends, create smarter offering campaigns, What's more faultlessly foresee customer devotion the natural provisions for learning mining.

Literature Review

To the detection of psychological is associated with the topic of sentiment analysis and emotion detects.

In recent years, computer-aided identification, analysis and application of emotions, especially on social networks, has attracted great deal of attention [8], [9] Relationships between stress and psychology [11], [16] Personality characteristics could be an important topic to consider. For instance, [1] provides proof that Daily44 Based on behavioral metrics, stress may be accurately identified Service by users on mobile phones. Oriented execution under extreme stress detected numbers social media-based feeling examination investigations would on the tweet level, utilizing text-based semantic Characteristics Also established arrangement methodologies. [23] Recommended an arrangement called temperament lens with perform feeling Investigation on the chinese micro-blog stage Weibo, classifying the feeling classifications under four types, i. E. , angry, gross, happy, and sad.

[9] contemplated the issue for passionate transmission for social networks Furthermore found that resentment need An stronger join the middle of separate clients over joy, demonstrating that negative feelings Might spread a greater amount effectively and comprehensively over the organize. Since stress is often seen as a negative emotion, this inference can allow us to combine the social impact of users who detect stress. However, the set asks mainly leverage the textual content of social networks. Social network data usually consists of sequential and interconnected objects from multiple sources and modalities in practice, essentially rendering it cross-media data.

While the identification of tweet-level emotions represents the immediate emotion reflected in a single tweet, the emotions or psychological stress states of people are typically more lasting, varying over different periods of time. Extensive research in recent years has started to concentrate on user-level identification of emotions on social networks [27] Our recent work [27]suggested detecting social media psychological stress states by learning user-level presentation on sequential tweet series over a certain period of time through a deep convolution network. Motivated by the homophilia theory, [18] integrated social relationships to strengthen Twitter's user-level sentiment analysis. In spite of the fact that a few user-level feeling discovery thinks about have been performed, the part that social connections play in ones states of mental stretch have not been explored however and how able to coordinated such information into push discovery. One of the most significant characteristics of social media sites is social networking. Many researchers are now focused on using knowledge about social interaction to further increase the efficacy of analysis of social media.[12] Studied the links between social interactions and the thoughts and behaviors of users, and discovered that interactioncenteredonTwittercancausesuccessfulcognitions.[19]LeveragedFlickercommentstohelp predict the feelings conveyed by photos shared on Flicker. These works, however, moved essential on the substance of social interactions, e. G. Printed remark content, same time ignoring the systemic subtle elements underlying, for example, such that how clients are interfaced.

Problem Identification

Stress is essentially people's response to various forms of desires or challenges. This reaction will help us to remain focused, energetic and mentally active while working properly, but if it is out of proportion, depression, anxiety, hypertension and a host of threatening disorders will certainly be detrimental. Cyber space is a huge soap box for people to share everything and everything they experience in their day-to-day lives. Subsequently, based on the posts and status changes he/she shares, it can be used as a very useful technique to determine the stress levels of a person. This is a proposal for a website that takes the Twitter username of the topic as associate degree input, scans and analyses the subject's profile through Sentiment Analysis and offers out results. Such results indicate the overall stress levels of the subject and provide a description of his or her mental and emotional status.

Implementation

Difficulties arise in psychological stress recognition. 1) How to extract attributes for the user level from the tweeting sequence and user attribute stack ling the question of lack of modality in tweets

2) We proposed a process in which we gathered messages from Twitter and assigned distinct moods to each detail. The structure and class of each tweet can be used to identify them. It was nostalgic after categorizing all of the tweets for each of the words. It is simple to edit each of the tweets with the help of emotion extraction so that each of the stress rate levels may be identified.

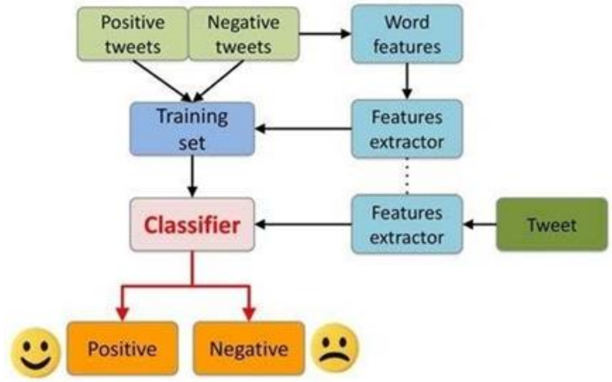

Figure 1:- Represents how all the knowledge that is initially obtained can be analyzed and how all the sentence is extracted using emotions.

Proposed System

Machine learning algorithms can be used to understand the cardinal indication of psychological stress: excessive negative thoughts and a lack of good sentiments. Optical Character Recognition (OCR) for image processing, Natural Language Processing (NLP), and Convolutional Neural Network (CNN) for text content processing are all included in the proposed system. Image processing, which detects and extracts text tweets from photographs, preprocessing, attribute extraction, categorization, and linking to NGOs are all critical components of the proposed system, as shown in Figure 1.

The data was acquired via social media platforms such as Twitter and Facebook. The user's stress level was estimated using a data selection of graphics and text. The photo dataset is extracted and analyzed using OCR, which extracts the text. After that, the text tweet material dataset and image extracted text dataset were used as preprocessing input and feature extraction using NLP. CNN plays a crucial role in categorizing favorable and negative tweet content. Finally, users' disparaging tweets and information are compiled and given to the NGO for advice. The stressed user, stressed user information (i.e. user ID) is extracted and will be given to the CLASSIFER. This is one of the ways in which the suicidal person is detected in the social network, and it is very likely to decrease the Level of depression and outcomes, denoting the non-stress end user and negative (0) as a stress end user

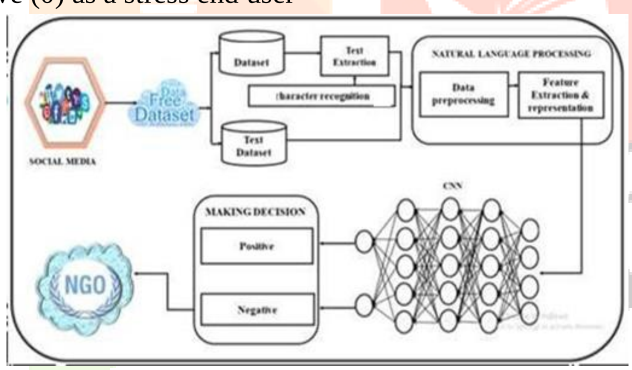

Figure.2. Architecture of Psychological Stress Prediction

Level of depression and outcomes, denoting the non-stress end user and negative (0) as a stress end user. In order to warn CLASSIFER that interfaces with the predicted result can be categorized as positive (1) after the CNN classification of social media drop out layer is finally applied to the network regularization and to avoid over-fitting issues value, which provides the value of likelihood as an output.

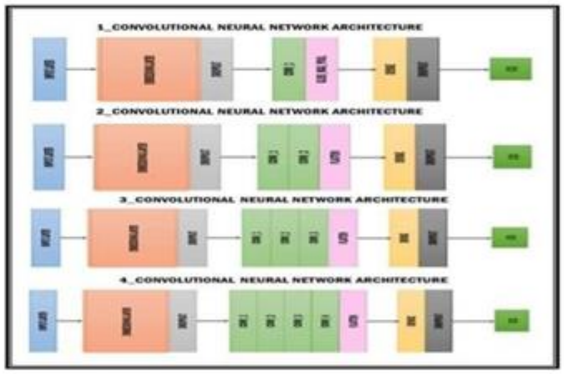

Figure 3. Layered Convolution Neural Networks for Stress Classification

1-CNN is the first layer composed of 1-convculation. The layer and the temporal convolution operation are carried out with the kernel size and 0-padding in this layer. The functionality of relu activation is then applied to the layer and the pooling process is performed. To decrease the dimensionality of the data, global maximum pooling is carried out here. For the fully connected convolution layer that produces the single value, the pooling output is supplied as an input. The sigmoid activation function is performed with this.

Results and Analysis

Figure .4. Scrapping Tweets from the Online Tweets

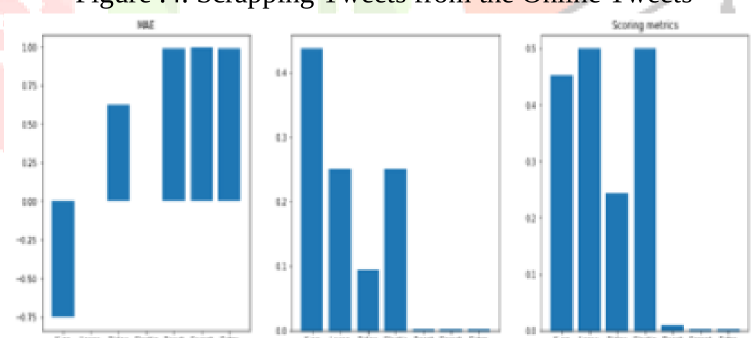

Figure.5.ScoringMetricslikeMean Absolute Error and Scoring Metrics Results Graph

S.No.	Model	MSE	MAE	Score
0	K-nn	0.4	0.45	-75.05
1	Lasso	0.2	0.50	-0.02
2	Ridge	0.1	0.24	62.26
3	Elastic	0.2	0.50	-0.02
4	Boost	0.0	0.01	99.17
5	Forest	0.0	0.00	99.19
6	Extra	0.0	0.00	99.15

Figure .6. Scoring Metrics Results Using Machine Llearning Approach

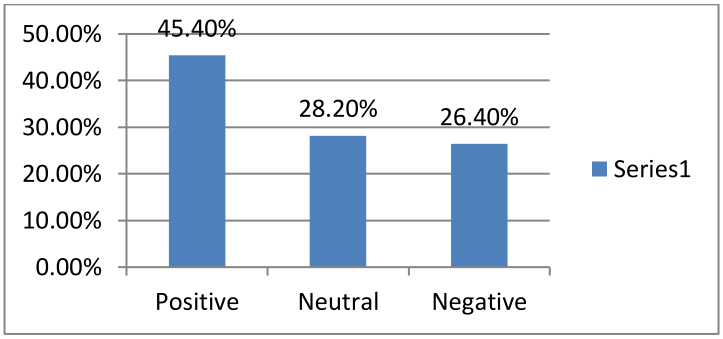

Figure.6. Fraction of Each Tweet for Stress Classification

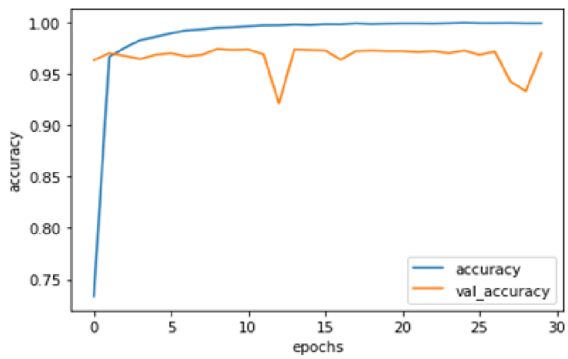

Figure .7. Training and Validation Accuracy Graph

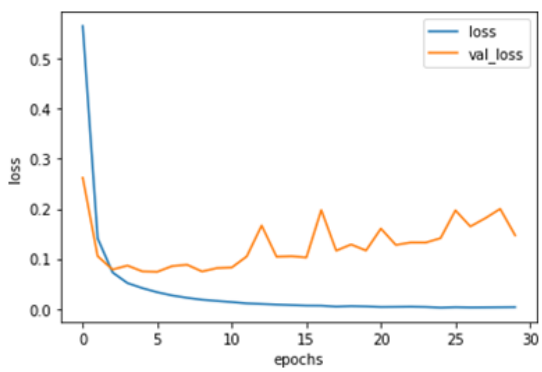

Figure.8.Training and Validation Loss Graph

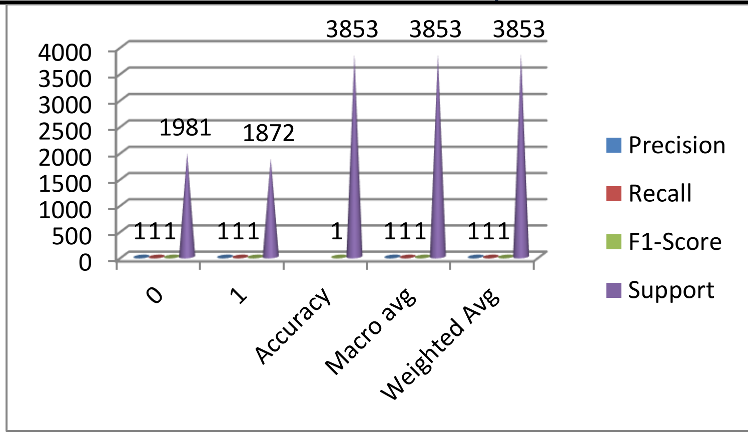

Figure.9. Evolution Metrics Results

Conclusion

It is crucial to assess how stressed a person actually is in today's society when young people and practically the whole population suffer from stress, whether from job pressure, workload, or other household difficulties. Because of this, quick stress diagnosis and prevention are imperative. This project that we developed enables users to investigate the stress problem. Those who don't feel comfortable discussing their problems with others will benefit greatly from this Endeavour. These people may be given a reality check and given the chance to seek out and receive medical care based only on their social contacts. We have utilized the concepts of sentiment analysis and employed both human and machine learning. This technique's primary advantage over earlier methods is its non-invasiveness and quick execution in the detection of stress.

References

1. Ben Verhoeven, Walter Daelemans, and Barbara Plank. Twisty: A multilingual twitter stylo metry corpus for gender and personality profiling. In Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC, pages1632–1637, 2016.
2. Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In ACM International Conference on Multimedia, pages 477–486, 2014.
3. Chris Buckley and Ellen M Voorhees. Retrieval evaluation within complete information. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, pages 25–32, 2004.
4. Xiaojun Chang, Yi Yang, Alexander G Hauptmann, EricPXing, 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.
5. Wanxiang Che, Zhenghua Li, and Ting Liu. Ltp: A Chinese language technology platform. In Proceedings of International Conferenceon Computational Linguistics, pages 13–16,2010.
6. Chihchung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY, 2(3):389–396,2001.
7. Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jurge Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In Proceedings of International Joint Conference on Artificial Intelligence, pages 1237–1242,2011.
8. Sheldon Cohen and Thomas A.W. Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98(2):310–357,1985.
9. Glen Coppersmith, CraigHarman, and MarkDredze. Measuring post traumatic stress disorder in twitter. In Proceedings of the International Conference on Web logs and Social Media, pages 579–582, 2014.
10. Rui Fan, Jichang Zhao, Yan Chen, and KeXu. Anger is more influential than joy: Sentiment correlation in weibo.PLoSONE,2014.
11. Zhanpeng Fang, Xinyu Zhou, Jie Tang, Wei Shao, A.C.M.Fong, Longjun Sun, Ying Ding, Ling Zhou, , and JarderLuo. Modeling paying behavior in game social networks. In Proceedings of the Twenty-Third Conference on Information

and Knowledge Management (CIKM'14), pages 411–420, 2014.

12. Golnoosh Farnadi, Geetha Sitaraman, Shanu Sushmita, Fabio Celli, MichalKosinski, David Still well, 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.
13. Eileen Fischer and A. Rebecca Reuber. Social interaction vianew social media: (how) can interactions on twitter affect effectual thinking and behavior? Journal of Business Venturing,26(1):1–18,2011.
14. JeromeH.Friedman.Greedyfunctionapproximation:A gradient boosting machine. Annals of Statistics, 29(5):1189–1232,1999.
15. RuiGao, BiboHao, He Li, YusongGao, and Tingshao Zhu.Developingsimplifiedchinesepsychologicallinguisticanalysisdictionaryformicroblog.pages359–368,2013.
16. JohannesGettingerandSabineT.Koeszegi.MoreThanWords:TheEffectofEmoticons inElectronicNegotiations.
17. Jennifer Golbeck, Cristina Robles, Michon Edmondson, andKarenTurner.Predictingpersonalityfromtwitter.InPassat/socialcom 2011, Privacy, Security, Risk and Trust,pages149–156,2011.
18. Mark S. Granovetter. The strength of weak ties. AmericanJournalofSociology,1973.
19. QuanGuo, JiaJia, GuangyaoShen, Lei Zhang, LianhongCai, and Zhang Yi. Learning robust uniform features forcross-mediasocialdatabyusingcrossautoencoders.KnowledgeBasedSystem, 102:64–75, 2016.
20. DavidW.Hosmer,StanleyLemeshow,andRodneyX.Sturdivant.Appliedlogisticregression.Wileyseriesinprobabilityandmathematicalstatistics,2013.
21. Sung Ju Hwang. Discriminative object categorization withexternal semanticknowledge.2013.
22. SepandarD.Kamvar.Wefeelfineandsearchingtheemotionalweb.

2. InProceedings ofWSDM,pages117–126,2011.
3. HerbertC.Kelman.Compliance,identification,andinternalization: Three processes of attitude change. Generalinformation,1(1):51–60, 1958.
1. ShigenobuKobayashi.Theaimandmethodofthecolorimagescale.
4. Color research&application,6(2):93–107,1981.

1. NovakPKralj,JSmailovi,BSluban,andIMozeti.Sentimentofemojis.PlosOne, 10(12),2015.
2. Frank R Kschischang, Brendan J Frey, and H-A Loeliger.Factor graphs and the sum-product algorithm. InformationTheory,IEEETransactionson,47(2):498–519,2001.
3. YannLeCun and YoshuaBengio. Convolutional networksfor images, speech, and time series. The handbook of braintheoryandneural networks, 3361,1995.
4. Kathy Lee, AnkitAgrawal, and AlokChoudhary. Real-timedisease surveillance using twitter data: demonstration on fluand cancer. In Proceedings of ACM SIGKDD internationalconference on Knowledge discovery and data mining, pages1474–1477,2013.
5. H. Lin,J.Jia,Q.Guo,Y.Xue,J.Huang,L.Cai,andL.Feng.Psychological stress detection from cross-media microblogdata using deep sparse neural network.In proceedings ofIEEEInternationalConferenceonMultimedia&Expo,2014.