ECRONICON OPEN ACCESS

EC EMERGENCY MEDICINE AND CRITICAL CARE

Research Article

Mental Anxiety of Employees due to Lockdown in Recent Situations

Shawni Dutta[1] and Samir Kumar Bandyopadhyay[2]*

[1]Lecturer, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India

[2]Academic Advisor, The Bhawanipur Education Society College, Kolkata, India

***Corresponding Author**: Samir Kumar Bandyopadhyay, Academic Advisor, The Bhawanipur Education Society College, Kolkata, India.

Received: May 23, 2020; **Published:** June 08, 2020

Abstract

Any company or organization is strongly dependent on employees work dedication. Dedication to work may be hampered if they loss mental peace while working. Work anxiety caused by stress in workplace and it causes poor outcomes both for employees and organizations. This situation needs attention by the organization to improve the working condition as well throughput of employees work. Discussion related to mental illness may create suspicious situation among the co-workers against the patient in the workplace. This paper attempts to predict consequences in the workplace observed due to mental anxiety. For this purpose, neural network classifier followed by 10-fold cross validation is proposed as an automated tool that recommends about the concerns perceived in the workplace. This proposed method is evaluated as well as compared with six classifiers such as Support Vector Machine, k-Nearest Neighbor, naïve bayes, Decision Tree, Adaboost and Gradient Boost classifiers. Experimental analysis concludes that proposed method outperforms well over aforementioned classifiers in terms of performance measure metrics.

Keywords: *Employee Mental Health, Negative Consequences; Co-Workers; Organization; Neural Network; k-fold Cross-Validation*

Introduction

Mental health of employees is one of the important factors that need to be considered along with physical health. Employees' professional dedication is significantly dependent on their mental health. Hence an organization must put their emphasis on employees' mental health. However, depression is often taken lightly than physical illness. Applying a leave for mental illness may create negative consequences for other co-workers and form trouble for the organization.

Relations with co-workers, superior employees, benefits provided by the organization, discussions on mental illness, seriousness and conflict between mental and physical illness etc. are useful for identifying a person with severe mental trauma. Applying a leave for mental illness, may create problem for the organization side as well if there are some resource constraints like huge workload, less number of employees and many more. In fact, these may generate dissatisfaction among other employees as well. Hence, these outcomes are also required to consider while sanctioning a leave. Data mining and knowledge discovery approaches are explored in order to forecast such consequences beforehand. In this paper, observable consequences for applying leave due to mental illness are considered as a crucial problem. By analyzing mental health issues along with several related factors are utilized for predicting further consequences.

The system proposed in this paper automatically captures the aforementioned interfering factors while deciding the consequences. Handling the negative consequences with care may help the organization environment to obtain stability. The proposed system is basically a classifier model that intended to predict whether there will be any negative consequences in the workplace or not. A neural network based framework followed by 10-fold cross validation procedure is implemented. After implementing the model, evaluation process takes place. This paper also implements other set of classifier models and a comparative study is drawn between these set of classifiers and proposed method. Other classifier set include Support Vector Machine (SVM) [1], K-nearest Neighbor (K-NN) [2], Naïve bayes classifier [3], Decision tree (DT) classifier [4], and ensemble classifiers such as Adaboost classifier [5], Gradient Boosting Classifier [6].

Related work

It is difficult to identify the cause of metal stress but it grows very fast in recent environment both in in the place of work as well in other places also. Smets., *et al*. [7] investigates the use of six machine learning techniques for the detection of mental illness. A binary classification problem was considered with classes corresponding to rest and stress periods. It is also reported that Bayesian networks and SVM provided reasonable accuracy with 84.6% and 82.7% respectively. Karthikeyan., *et al*. [8] in studied stress assessment based on the electrocardiography (ECG) and heart rate variability (HRV) signals. They successfully classified stressful and not stressful events based on HRV features with classification accuracy of 79.2%.

Strauss., *et al*. [9] in explored the application of machine learning algorithms like cluster analysis, K-nearest neighbours (KNN), decision trees and support vector machines (SVM) for clinical forms analysis of mental health. Their study concluded that SVM had the best performance with a precision of 64.6%. Kessler., *et al*. [10] utilised machine learning model by taking inputs from survey. They predicted major depressive disorder (MDD) keeping its persistence with good accuracy.

Human mental stress is measured using multimodal bio-sensors and obtained dataset are classified using SVM and fuzzy logic in [11]. Considering related health parameters, decision-tree model and random forest algorithm classifies mental stress level. A predictive model is finally proposed in [11] that decides wellness contents by using Expectation Maximization (EM).

Proposed Methodology

A multi-step procedure is followed to build the proposed model and the other classifiers also. The target of this method is to early detection of possible consequences raised due to mental illness issues. The required steps are explained as follows.

Dataset collection and preprocessing

A technical survey data related to Employee mental health issues are collected from kaggle [12]. The dataset contains 1259 number of sample records along with the following parameters which are summarised in table.

Feature Name	Description
Timestamp	Record Time
Age	Respondent's Age
Gender	Respondent's gender
Country	Country of Employee
State	State of residence
Self_employed	Self-Employed or not
Family_history	Is there any family history of mental illness
Treatment	Sought Treatment for a mental health condition
Work_Interfere	Whether mental health condition interfere work.
No_employees	Number of employees of company of organization.
Remote_work	Remote working of an employee at least 50% of the time
Tech_company	Is employer primarily a tech company/organization?
Benefits	Mental health benefits provided by employer
Care_options	Knowledge of mental health care options provided by employer.
Wellness_program	Discussion regarding mental health as part of an employee wellness program by employer
Seek_help	Does employer provide resources to learn more about mental health issues and how to seek help?
Anonymity	anonymity protection if anyone choose to take advantage of mental health or substance abuse treatment resources
Leave	Easiness to take leave for mental health issues.
Mental health consequence	Observable negative consequences on discussion of mental health issue with employer
Phys health consequence	Observable negative consequences on discussion of physical health issue with employer

Coworkers	Discussion of mental health issue with coworkers
Supervisor	Discussion of mental health issue with direct supervisor(s)
Mental Health interview	Discussion of mental health issue with a potential employer in an interview.
Phys health interview	Discussion of physical health issue with a potential employer in an interview.
Mental vs physical	Importance of mental health and physical health
Obs_consequence	Observable negative consequences for co-workers with mental health conditions in workplace
Comments	Additional Notes or Comments

Collected data are preprocessed and a multistep procedure is followed for obtaining a balanced dataset. Pre-processing techniques include missing values handling such as NaNs, scaling of attribute values of Gender. Gender and Age are the attributes that contain irrelevant values which are cleaned with relevant values. The attribute Timestamp and Comments are eliminated from the dataset. In order to fit the data into classifier, non-numeric data is transformed into numeric data. This will be followed by scaling values of every feature with large set of data points. Feature scaling will assist the classifier to work using normalized data with an enhanced efficiency. The attribute 'obs_consequence' is fixed as the target class that assumes values either Yes or No.

Methodology

In this framework, deep learning based methodology is implemented while analyzing the concerns caused due to emotional state. Deep learning is often regarded as a subfield of machine learning techniques that identifies and learns complicated linear as well as non-linear relationships between input datasets and output [13]. Human brain like operations is simulated in order to accompany complex problem solving approach. For this purpose Neural network architecture is proposed in this paper that accepts several issues related to mental state of an employee and finally predicts the outcome of negative consequences with mental health conditions in workplace. Neural network proposed in this paper is comprised of several neurons. Each of these neurons will accept necessary parameters and apply some activation functions in order to produce outputs. Activation functions [14] are useful to perform diverse computations and produce outputs within a certain range. In other words, activation function is a step that maps input signal into output signal.

After configuring this neural model, training process is executed. The training process goes through one cycle known as an epoch where the dataset is partitioned into smaller sections. An iterative process is executed through a couple of batch size that considers subsections of training dataset for completing epoch execution.

Implementation

While designing this model it is necessary to tune hyper-parameters in order to achieve maximized efficiency. This section describes specification of the model along with its hyper-parameters. This model consists of three Dense layers with 32,16,1 number of nodes respectively. In this context, sigmoid and relu [14] are two popular activation functions those are applied in each of these specified layer. The first two layers apply relu as activation function and the final layer applies sigmoid activation function.

Finally, these aforementioned layers are compiled using adam solver through 30 epochs and with a batch size of 10. Adjustment of the hyper-parameters assists the model to obtain best predictive result. The neural network receives a total of 1,377 parameters and trains those parameters in order to obtain prediction. The summarization of the model is given in figure 1.

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 32)	832
dense_2 (Dense)	(None, 16)	528
dense_3 (Dense)	(None, 1)	17

Figure 1: *Structure of neural network.*

Implementation of this model produces a feed-forward neural network. This implementation is followed by 10-fold cross-validation method [15] for estimating the skill of the model. It is a resampling methodology where the dataset is partitioned into 10 groups and in each iteration one group is considered as the test data and the remaining nine folds are considered as training data. The above-mentioned model is fitted into the training dataset and it is evaluated against the test dataset. Later evaluation scores for each of these iterations are accumulated and mean score is calculated.

This neural network structure along with 10-fold cross validation procedure is applied on technical survey dataset. Implementation of this model is evaluated and compared with other benchmark classifiers such as SVM, Naïve Bayes Classifier, K-NN, Decision Tree, Adaboost and Gradient Boost Classifier. These three classifier models are considered as baseline for comparing the proposed method.

Other classifiers used

Classification is a supervised machine learning technique that analyses specified set of features and identifies data as belonging to a particular class. Different classification algorithms such as support vector machines, decision trees, K-nearest neighbour classifier, Adaboost, Gradient Boost are used to predict the target class. For these classifier models the pre-processed and transformed data are partitioned into training and testing dataset with the ratio of 7:3. Training dataset is fitted to the classifier and later predictions are obtained using testing dataset. Brief description of the classifiers are provided as follows-

Support Vector machine (SVMs) [1] belongs to the category of linear classifiers. It identifies different classes by separating samples with the help of decision boundary known as hyperplane. Both linear as well as non-linear data can be classified with the help of SVMs [1].

K nearest neighbour [2] is often considered as lazy learner which considers instances during classification process. It is known as lazy learners because during training phase it just stores training samples. This identifies objects based on closest proximity of training examples in the feature space. The classifier considers k number of objects as the nearest object while determining the class. The main challenge of this classification technique relies on choosing the appropriate value of k.

The Naive Bayes classifier [3] is a supervised classification tool that exploits the concept of Bayes Theorem [3] of Conditional Probability. The decision made by this classifier is quite effective in practice even if its probability estimates are inaccurate. This classifier obtains a very promising result in the following scenario- when the features are independent or features are completely functionally dependent.

The accuracy of this classifier is not related to feature dependencies rather than it is the amount of information loss of the class due to the independence assumption is needed to predict the accuracy.

A Decision Tree (DT) [4] is a classifier that exemplifies the use of tree-like structure. It gains knowledge on classification. Each target class is denoted as a leaf node of DT and non-leaf nodes of DT are used as a decision node that indicates certain test. The outcomes of those tests are identified by either of the branches of that decision node. Classification results are obtained by starting from the beginning at the root this tree are going through it until a leaf node is reached [4].

For improving the accuracy of classification, several unstable learners are accommodated into a single learner using an efficient technique known as Boosting. Classification algorithms are applied to the reweighted versions of the training data and the weighted majority vote of the sequence of classifiers are chosen. AdaBoost [5] is a good example of boosting technique that produces improved output even when the performance of the weak learners is inadequate. Gradient boosting [6] algorithm is another boosting technique based classifier that exploits the concept of decision tree. It also minimizes the prediction loss.

The above specified classifiers are implemented by considering and adjusting appropriate hyper-parameters for obtaining the maximised performance. The SVM classifier utilizes 'rbf' kernel and regularization parameter C = 1. The K-NN classifier gives a promising result for the value k = 5 considering all the evaluating metric. For naïve bayes classifier, multinomial naïve bayes classifier is employed. The decision tree classifier implemented in this paper uses Gini index while choosing objects from dataset. The nodes of the decision tree are expanded until all leaves are pure or until all leaves contain less than minimum number of samples. In this case, minimum number of samples is assigned a value as 2. On the other hand, ensemble classifiers, such as, AdaBoost and Gradient Boost classifiers are built based on 500 numbers of estimators on which the boosting is terminated.

Performance evaluation metrics

While evaluating performance skill of a model, it is necessary to emphasise on some metrics to justify the evaluation. For this purpose, following metrics are taken into consideration in order to identify the best relevant problem-solving approach. Accuracy [16] is a metric that detects the ratio of true predictions over the total number of instances considered. However, the accuracy may not be enough metric for evaluating model's performance since it does not consider wrong predicted cases. Hence, for addressing the above specified problem, precision and recall is necessary to calculate.

Precision [16] identifies the ratio of correct positive results over the number of positive results predicted by the classifier. Recall [16] denotes the number of correct positive results divided by the number of all relevant samples. F1-Score or F-measure [16] is a parameter that is concerned for both recall and precision and it is calculated as the harmonic mean of precision and recall [16]. Mean Squared Error (MSE) [16] is another evaluating metric that measures absolute differences between the prediction and actual observation of the test samples. A model that exhibits lower value of MSE and higher values of accuracy, F1-Score indicate a better performing model.

Experimental Results

The proposed model is evaluated in terms of performance measure metrics. The results are summarized along with specified baseline classifiers such as Naïve Bayes, SVM, K-NN, Decision Trees, Adaboost and Gradient Boost. The summarized results are shown in table 1. Figure 2 provides overall comparison of all the specified classifier models. This indicates that proposed method provides better performance with respect to other classifiers. Accuracy and MSE obtained in each fold with respect to training and testing dataset is shown in figure 3.

Classifiers Used	Accuracy	MSE	F1-Score
Proposed Method			
Neural Network Classifier with Cross-Validation	85.14%	0.15	0.85
Baseline Classifiers			
Multinomial Naïve Bayes	73.08%	0.27	0.73
SVM	83.89%	0.16	0.84
K-NN	82.69%	0.17	0.83
Decision Tree	78.12%	0.78	0.22
Adaboost Classifier	80.77%	0.81	0.19
Gradient Boost Classifier	83.17%	0.83	0.17

Table 1: *Performance summarization of specified classifiers.*

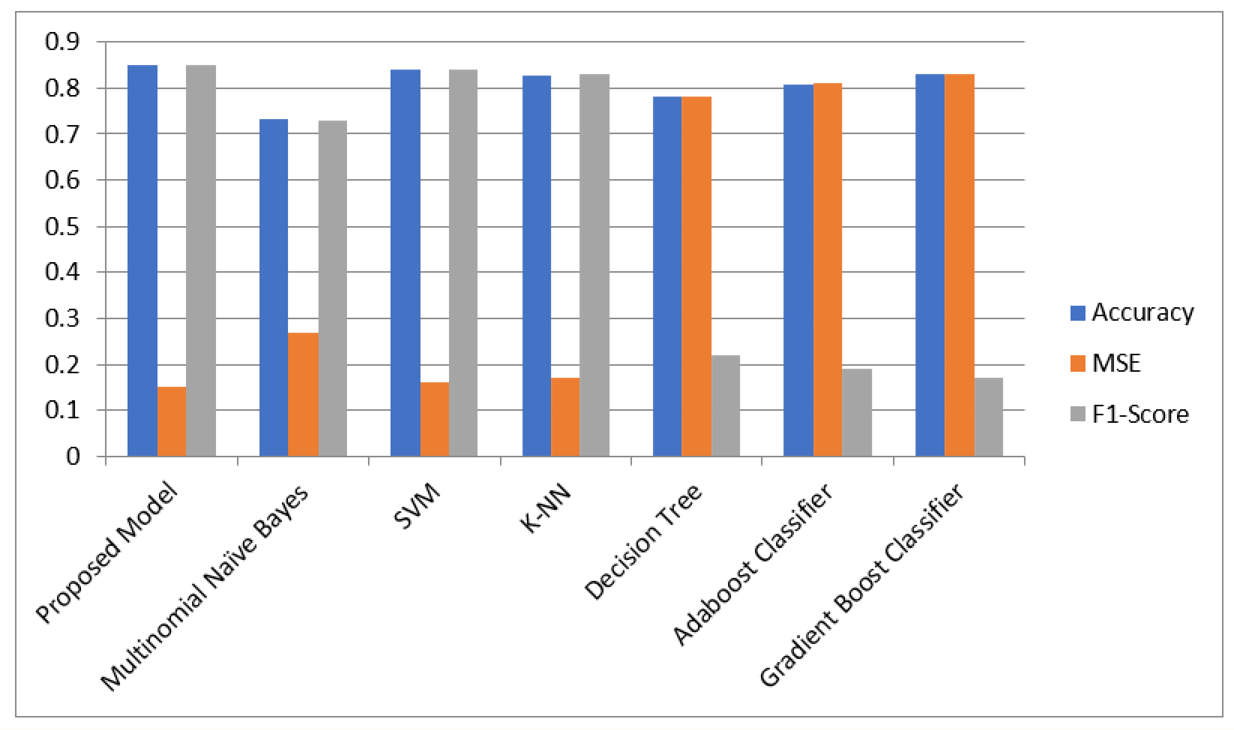

Figure 2: *Overall performance comparison of classifier models.*

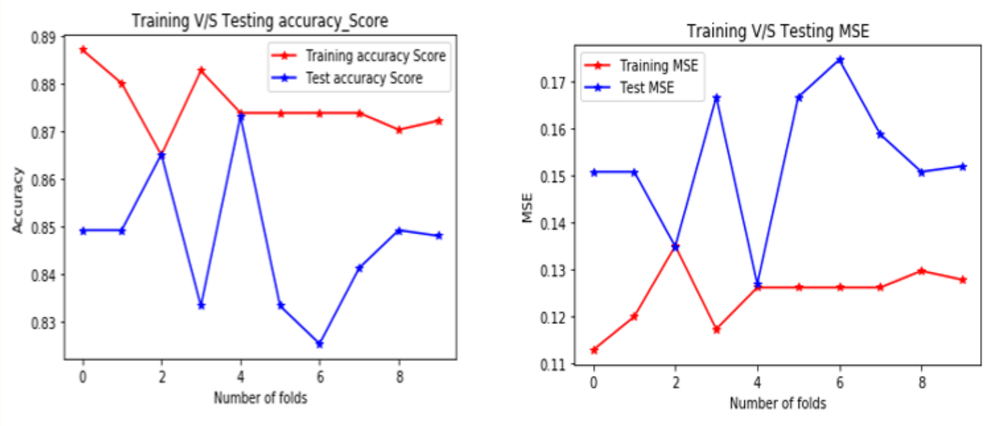

Figure 3: *Training and testing accuracy and MSE shown in each iteration of cross-validation.*

Conclusion

In this paper, identification of negative consequences raised in the workplace due to mental trouble faced by an employee is focused. While analyzing mental health of an employee, parameters in terms of relations with co-workers as well as organization point of view are also taken into consideration. A neural network based framework accompanied by 10-fold cross-validation method is proposed that automatically captures interfering factors to predict the possible outcomes. Designing the model is accompanied by fine-tuning the parameters. Experimental study concludes that proposed method often performs with an accuracy of 85.14% in terms of prediction over its peer classifiers.

Bibliography

1. H Wang., *et al.* "Data classification using support vector machine". Proceedings of International Conference on Tools with Artificial Intelligence ICTAI 1 (2010): 3-6.
2. P Cunningham and SJ Delany. "K -Nearest Neighbour Classifiers". Multiple Classifier Systems (2007): 1-17.
3. I Rish. "An Empirical Study of the Naïve Bayes Classifier An empirical study of the naive Bayes classifier" (2001): 41-46.
4. H Sharma and S Kumar. "A Survey on Decision Tree Algorithms of Classification in Data Mining". *International Journal of Science and Research* 5.4 (2016): 2094-2097.
5. J Friedman., *et al*. "Additive logistic regression: a statistical view of boosting". *Annals of Statistics* 28.2 (2000): 337-407.
6. A Natekin and A Knoll. "Gradient boosting machines, a tutorial". *Frontiers in Neurorobotics* 7 (2013): 21.
7. A Díaz-García., *et al*. "Pervasive Computing Paradigms for Mental Health". *Pervasive Computing Paradigms for Mental Health* 604 (2019): 147-156.
8. P Karthikeyan., *et al*. "Analysis of stroop colorword test-based human stress detection using electrocardiography and heart rate variability signals". *Arabian Journal for Science and Engineering* 39.3 (2012): 1835-1847.
9. J Strauss., *et al*. "Machine learning methods for clinical forms analysis in mental health". *Studies in Health Technology and Informatics* 192.1-2 (2013): 1024.
10. KRC., *et al*. "Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports". *Molecular Psychiatry* 21.10 (2016): 1366-1371.
11. Y Jung and YI Yoon. "Multi-level assessment model for wellness service based on human mental stress level". *Multimedia Tools and Applications* 76.9 (2017): 11305-11317.
12. Open Sourcing Mental Illness, LTD. Mental Health in Tech Survey, Version 3 (2016).
13. J Liu., *et al*. "Performance analysis and characterization of training deep learning models on mobile device". Proceedings International Conference on Parallel and Distributed Systems- ICPADS (2019): 506-515.
14. C Nwankpa., *et al*. "Activation Functions: Comparison of trends in Practice and Research for Deep Learning" (2018): 1-20.
15. RH Kirschen., *et al*. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection". *American Journal of Orthodontics and Dentofacial Orthopedics* 118.4 (2000): 456-461.
16. HM and SMN. "A Review on Evaluation Metrics for Data Classification Evaluations". *International Journal of Data Mining and Knowledge Management Process* 5.2 (2015): 1-11.

Volume 4 Issue 7 July 2020