International Journal of Health Sciences

Available online at www.sciencescholar.us
Vol. x No. x, Month 2021, pages: xx-xx
e-ISSN: 2550-696X, p-ISSN: 2550-6978
https://doi.org/10.53730/ijhs

Mental Stress Level Prediction Of Staff Community

Dr. P. Keerthika [1], Dr. P. Suresh [2], Mr. K. Devendran [3], Ms. S. Vaishnavi [4], Ms. C. Shanmathi [5], Mr. V. Surendar [6]

Manuscript submitted: … , Manuscript revised: … , Accepted for publication: …

Corresponding Author [7a]

Keywords

Mental Health;
Stress;
80 Questionnaires;
Prediction;
Machine and Deep Learning;

Abstract

Stress Based - Mental Health of an individual is very important. It determines the behavior of an individual, emotional activity of an individual. It plays a major role in how an individual thinks and reacts to a situation and how an individual handles it to overcome that. An individual is mentally ill when they are imbalanced in the Brain Chemistry. Mental Health can affect an individuals Physical health, relationship with the others and daily living of an individual. There are many factors that cause Mental illness for example work pressure and stress also leads to Mental illness. An individual can be observed with serious changes in Sleeping and Eating habits and also long lasting loneliness and anxiety. Many Researches are being done to examine the psychiatric behavior of an individual and have predicted the psychiatric behavior of an individual using diverse machine learning supervised and unsupervised algorithms.

The project is to apply both the machine learning technique and deep learning technique to predict the stress level of the staff community based on the 80 questionnaires answered by the individuals. The dataset includes 80 features that comprises questions related to mental illness (i.e) Guilt, Sadness, Loneliness, Work Related questions etc,. From the Random Forest Algorithm Machine Learning Approach an accuracy of 45% was only achieved. To increase the accuracy of the prediction, Convolutional Neural Network (CNN) is being applied to break down the individual based on stress level. By using CNN an accuracy of 96% is being achieved.

[1] Affiliation, City, Country/Territory
[2] Affiliation, City, Country/Territory
[3] Affiliation, City, Country/Territory
[4] Affiliation, City, Country/Territory
[5] Affiliation, City, Country/Territory
[6] Affiliation, City, Country/Territory

e-ISSN: 2550-696X p-ISSN: 2550-6978

Contents

1 Introduction

Stress based mental health has become a biggest crisis for all the age group of persons without any indifference between children and adults. Stress is also a mental health issue of all age groups of citizens, according to the World Health Organization. *(Jorn Bakker et al., 2012)* says stress is an important symptom of mental health. Stress affects all aspects of human life like emotions, thinking and behavior. Stressful mental health disorder is not unusual under the work class. Several of the past studies have the same concern. According to the research of industrial bands, Assam's Sasam has long been working with long and narrow deadlines, depression and general anxiety disorders for narrow deadlines, beyond the proportion of work of the private sector in India. *(J. Bakker et al.,2011)* says today, individuals experience extreme actual issues and mental pressure because of an assortment of inside and outside parts. Individuals change their feelings depending upon the circumstance of spots like the work place, the study place or the place they live and on the circumstance of time and climate. Despite the fact that downturn is for the most part found in individuals in their age of 30 years or around 40 years encountering melancholy, it is in many cases recognized in adolescents because of scholastic pressure and relational relationship and in older people.

Stress is high in programming callings in view of their inclination of the work they do, the project target that is to be completed, accomplishments that have been done, the work time and over responsibility. *(Alban Maxhuni et al.,2016)* says IT experts are feeling the squeeze to convey benefits proficiently and must be financially savvy. Sicknesses are either actuated, maintained or exacerbated by mental stress. Stress score assists us with screening who might be inclined to pressure related actual ailment and individuals with a score more than 6 are in danger of disease and mind ought to be taken at the earliest to ease their pressure. *(Ghaderi. A et al., 2015)* article says sound workers mean better execution by representatives that thus produce a solid local area. Yearly pressure scoring must be done, and workers having a score more than 6 ought to be engaged with dynamic antistress the board.

Nowadays, everyone is facing the stress issue. *(Prakruthi Manjunath et al., 2021)* says that the current situation and educational system has made the both students and staff community much harder which leads to mental stress. The staff community gets more stressed because of the number of working hours, number of subjects they handle, the working organization, the working environment, the students they come across and so on. So it is considered very important to diagnose the staff community for the better well being of the students and a better output from them. Early diagnosing and helping them to overcome makes the education for the students easier to handle.

Convolutional brain network(CNN) is utilized to handle this information. Inside Deep Learning, a Convolutional Neural Network or CNN is a sort of counterfeit brain organization, which is generally used to prepare machines/PCs with the goal that they can gain as a matter of fact, group and perceive

information/pictures very much like a human mind does. The term “Convolution” in CNN indicates the numerical capacity of convolution which is an extraordinary sort of direct activity wherein two capacities are duplicated to deliver a third capacity which communicates how the state of one capacity is altered by the other.
(A. Kene and S. Thakare 2021) says staff are diagnosed by various questionnaires that are asked related to mental stress level and they are categorized as normal, slightly high stress and high stress. Based on the result they can be figured out to sort the issue they face for the betterment of the health and also for the betterment of the students. This is done with the help of Deep Learning Convolutional Neural Network algorithm to achieve a higher accuracy and there won't be any misunderstanding between the staff and the health department helping to sort stress level. The modal records the accuracy.

2 Materials and Methods

Moment's world pressure, global pressure has come from a number of people around the world. Dovetailing to such a situation, mortal physiology grounded medical individual systems are needed other than others. *(Gray-Toft et al.,1985)* says mortal physiological exploration plays an important part in the discovery of mortal internal stress. There's also a possibility of circumstance of stress discovery grounded on face feelings. *(Widanti, N et al., 2015)* says we present a thorough collection of stress, DI-FF measures, stress, operation, and anticipated task issues in the field of sanctuary, and a thorough collection of recognition vaccination systems used to prognosticate. An analysis of well- known point selection styles is also presented. *(Zarski, J. J. et al., 1986)* says research is also being conducted on the relationship between mortal facial feelings and cerebral stress. A variety of exploration in this area will be concentrated and will be more useful for unborn exploration in this area.
(Selvaraj N et al., 2015) says the pressure sensor classifies stressed-out people from normal people by detecting physiological signals through applicable detectors similar to electrocardiogram (ECG) and electrocutaneous response (GSR). These signals are preprocessed to prize bodacious features that characterize the stress situations of working people. Examine support vector machines (SVMs) and KNearest Neighbors (KNNs) to classify these uprooted point sets. The results show a point vector with the stylish features that have a strong impact on stress discovery. assays are formed to arbitrate the stylish point set for maximum bracket delicacy. The proposed system pertains to the SWELLKW standard dataset to give the conventional results.
(C. Setz et al., 2010) says continuous exposure to pressure is harmful to one's internal and physical well-being, yet the best way to combat it is to condemn it. We propose a technique for continuous detection of stressful occurrences using data from a commercial wrist flimflam in this article. *(E. Tromp et al., 2011)* says the framework is comprised of three machinelearning parts: a lab pressure sensor that distinguishes transient pressure each two sparkles; an effort recognizer that consistently perceives the stoner's effort and along these lines gives climate information; and a setting based pressure sensor that takes advantage of the connection between the research facility pressure sensor and the stoner's current circumstance to pursue a last choice each 20 sparkles stretch. The framework was assessed both in the lab and, in actuality. For a two-class task, the delicacy on 55 days of genuine information was 92. This arrangement is at present connected into inside wellbeing and government assistance exercises on cell phones.
(Ingegerd Hildingsson et al., 2014) says stress has grown to be an critical fitness problem, however current pressure detectors are inconvenient in long-time period real-existence use due to the fact customers both ought to put on devoted gadgets or use up remarkable interplay efforts in device version to specifics of every person. Adaptation is vital due to the fact people extensively range of their belief of pressure and pressure responses, however standard version employs supervised studying techniques and consequently calls for pretty big units of labeled records (i.e. records on whether or not every reporting length changed into disturbing or not) from each user. *(David Taylo et al., 2006)* says to cope with those problems, we recommend a unique unsupervised pressure detector, primarily based totally on the usage of a telecel smartphone because the most effective tool and the usage of discrete hidden Markov models (HMM) with most posterior marginal

e-ISSN: 2550-696X p-ISSN: 2550-6978

(MPM) selections for evaluation of telecel smartphone records. Our detector calls for neither extra hardware nor records labeling and consequently is absolutely unobtrusive and appropriate for lifelong use. Its accuracy evaluated the usage of real-existence datasets: within the first case, version changed primarily based totally on very short (some days) telecel smartphone interplay histories of every individual, and within the 2d case—on longer histories. In those tests, the proposed HMM-MPM accomplished fifty nine and 70% accuracies, respectively, that's similar with effects of absolutely supervised techniques, suggested with the aid of using different works.

(K. Glanz et al., 2008) says through the eccentricities of VCR coded facials, this study makes a worldview for distinguishing and evaluating pressure/nervousness feeling state. Thorough experimental protocols are configured to induce systematic variations in the affected state (neutral, relaxation, emphasis / anxiety) through various external and internal stress factors. Investigation centers chiefly around non-power and semi-dynamic facial highlights to appraise more expressive articulations. The qualities under ID incorporate occasions, verbal impacts, head moving boundaries, and pulse assessed by camera-based rayret high grafies. The component determination strategy kept on choosing a characterization technique for distinguishing pressure/tension and impartial circumstances, in view of the unwinding state in any trial stage, and picked the most vigorous capacity. What's more, a positioning change utilizing self-revealing was proposed to inspect the relationship between facial boundaries and members' apparent pressure/nervousness levels.

(Palanisamy K et al., 2013) says the outcomes show that specific facial boosts obtained from eye action, mouth action, head development, and camera-based heart movement accomplish great exactness and are reasonable as discriminants for stress and uneasiness. *(Houtan Jebelli et al., 2019)* says facial feeling acknowledgment is far reaching today and is an appealing region for feeling registering, particularly PC vision with medical care applications. Looks change in different cases relying upon time and individuality. Looks assume the main part in consequently getting a handle on feelings through a PC, and furthermore help the human-machine interface. Individuals might vary marginally from facial looks, yet it is as yet quite difficult for PCs. *(S. Sriramprakash et al., 2017)* says the distributed work proposes a viable dispensed with innovation. By utilizing PCA (head part investigation), you can remove all significant data in a casing where the human face is recognized.The expression of the face is known to promote emotions. Use PCA to reduce calculation dimensions. In this process, the face is extracted and the characteristics are extracted, and the dimensions of the PCA will be reduced, and then the classification of emotions with Euclidean distance measurement criteria is reduced, and then time dynamics (time of continuous face emotion detection system 2017 Turn dynamics). *(G. Rigas et al., 2006)* says emotional reduction frame redundancy. The arrangement of preparing pictures that characterize face rooms determines restrictive vectors. Pack 8 pressure and edge points with PCA. PCA makes use of an information base with a few edges for the purpose of planning. The edges that are left over are used to test recommendations. We worked on outlines for resentment, disdain, bliss, nonpartisanship, and shock, among other emotions. I tried using IndianFaceDatabase. 30 edges are used for preparing the framework and 50 casings are used for testing from this database. We achieved a 91.26 percent identification rate by trial.

(J. Wijsman et al., 2011) says Pressure is a collected psychophysiological province of the human body emerging in light of a tiring occasion or a requesting condition. Ecological elements that flash pressure are called stressors. If there should be an occurrence of delayed openness to different stressors influencing contemporaneously, an individual's interior and actual wellbeing can be adversely impacted which can additionally prompt routine medical problems. To assist with pushing subsidiary issues, it's important to depict them in the nascent stages which are conceivable exclusively by relentless observation of pressure. Wearable predisposition guarantees continuous and constant information assortment, which helps specifically pressure observing. *(J. Zhai et al., 2006)* says an extensive survey has been introduced, which centers around pressure revelation utilizing wearable indicators and applied machine education ways. *(Maryam Memar et al., 2021)* researches the pressure disclosure approaches upheld in concurrence with the touchy inclination comparable as wearable finders, Electrocardiogram (ECG), Electroencephalography (EEG), and Photoplethysmography (PPG), and furthermore relying upon brilliant environmental factors like during driving, considering, and working. The stressors, ways, results, benefits, constraints, and issues for each study are focused on and expected to give a way for unborn investigation studies. Additionally, a multimodal stress revelation framework utilizing a wearable identifier grounded profound education design has been proposed toward the end.

In the U. S. Reddy et al. have applied different calculations to track down the most dependable one and analyzed the connection between different boundaries in the dataset. In M. P. Dooshima et al. have utilized segment, natural, mental and ecological elements for expectation. Different emotional wellness specialists were counseled to approve the acquired boundaries. In *(Srividya et al., 2018).* have involved a poll to get values for various characteristics that can be useful for expectation of psychological well-being. The thought process of this paper was to dissect various calculations and foresee the most reliable one. Different order calculations, for example, Decision Tree, Naïve Bayes as well as SVM were utilized in this paper. The marks from the information gathered were utilized to register a MOS. The above calculations were then applied to view as the most reliable one. The paper reasoned that Support Vector Machine, K-Closest Neighbor and Random Forest are the most reliable calculations with comparable exactness results. In D.Filip and C. Jesus. have utilized Neural Networks to foresee the mental states of people, for example, discouragement, PTSD, nervousness and so on. They additionally concentrated on the impact of blackout or wounds on sportspersons. In S. G. Alonso et al. have directed a broad audit of various calculations utilized for psychological wellness expectation. Various methods, for example, Association Rule Mining and Randomization were contemplated and their expectations were noted for our undertaking. This paper additionally inspected different calculations, for example, SVM, Decision tree, KNN, ANN, Naïve Bayes.

(M. Gjoreski et al., 2016) says relapse examination, which is many times utilized in insights, numerically assesses direct relationships between information with the utilization of relapse. As factors, free factors (cause) and ward factors (result) are utilized. Relapse investigation is recognized into basic relapse examination and various relapse examination as indicated by the quantity of free factors. In basic relapse investigation, there is one autonomous variable and one ward variable, and the circumstances and logical results connection between the two factors is dissected. In different relapse examinations, there are one ward variable and something like two free factors. The investigation technique is utilized to find how no less than two autonomous factors influence one ward variable. Also, the relapse examination strategy is utilized to investigate the circumstances and logical results connection among reliant and autonomous factors.

There are various types of frameworks that presently exist. The greater part of them utilize various strategies to foresee dysfunctional behavior. A portion of the ongoing frameworks incorporate a web-based review which predicts regardless of whether the client has psychological maladjustment. These overviews are sickness explicit for example an alternate overview for sorrow, an alternate one for stress, etc. Every one of these overviews are accessible on the web and anybody can fill them *(E. Vildjiounaite et al., 2018)*. There are a couple of frameworks which use chatbots to foresee dysfunctional behavior by scrutinizing a client and afterward examining their reaction. A few frameworks additionally use Image Processing to screen the looks of clients, dissect their way of behaving to a specific inquiry to help in more precise expectation of psychological instability. The vast majority of these studies incorporate the conduct and actual parts of an individual yet no business related issues. Thus an extremely insignificant measure of exploration has been accomplished on business related psychological instability *(G. Giannakakis et al., 2017)*. The vast majority of these frameworks are focused on the overall parts of psychological sickness. They incorporate the most widely recognized boundaries to assess the outcome. Frameworks that pay attention to the representatives and business related psychological sickness are not accessible on a more noteworthy scale.

Choice tree calculation is a directed learning calculation. Tackling both relapse and arrangement problems can be utilized. *(A. R. Surve et al., 2019)* says it utilizes the tree portrayal to take care of the issue in which each leaf hub relates to a class name and traits are addressed on the inner hub of the tree. The choice is done on Entropy or Gini esteem said by *(S. Gedam et al., 2021)*. Arbitrary woods calculation, similar to its name infers, comprises an enormous number of individual choice trees that work together. *(J. Healey et al., 2005)* says every individual tree in the irregular woods lets out a class expectation and the class with the most votes turns into our model's forecast, says *(Stanley Leonard Tito et al., 2019)*.

(P. S. Pandey et al., 2017) says the research basically targeted the IT sector community. Both clustering and classification was done to foresee the psychological wellness of a person. Grouping algorithms like various leveled bunching and K-Medoids, K-Means Clustering and classification algorithms like Logistic regression, Decision tree, Support vector machines and K-nearest neighbors were utilized to accomplish the most elevated exactness of forecast. The dataset consist of 656 samples, which consist of 20 questionnaires answered by the

e-ISSN: 2550-696X p-ISSN: 2550-6978

IT community. Questionnaires include answers in the form of ratings from 1 to 5 in which 1 denotes least and 5 denotes the highest.
The dataset is clustered into 3 as 0-30, 31-60 and 61-100. 0-30 indicates mentally distressed, 31-60 indicates neutral and 61-100 indicates happy individual. The presentation measurements, for example, accuracy, recall, Precision and F1 score were assessed. (Table 1) Gives the accuracy of the classification model used in the existing system.

Table 1
Accuracy for different algorithms

CLASSIFIER	**ACCURACY**
Logistic Regression	84
Naïve Bayes	73
Support Vector Machine	89
Decision Tree	81
K Nearest Neighbor	89
Ensemble (Bagging)	90
Random Forest	90

In this exploration, machine learning and deep learning strategies have been executed to anticipate the pressure of the staff local area. The various questionnaires answered by various individuals are analyzed to predict the stress-based mental health of each individual. The questions related to stress and mental health were framed, and answers to the questions were collected for the prediction. The data was normalized using PCA. After normalizing the data, the machine learning algorithms Random Forest and Deep Learning Algorithm CNN were used to detect an individual's stress level. Nearly 1500 samples of data have been taken to train the model collected from the IT community. Once after the training process, testing is done on the dataset gathered from the replies of the staff community. Performance metrics such as accuracy, recall, F1 Score, precision, sensitivity and specificity are included in the results.

Random Forest

Random Forest algorithm is an ensemble algorithm. Ensemble methods are used to increase the predictability in the models by using various models to achieve a dependable model. Bagging, Boosting and Stacking are the commonly used methods. An ensemble is a cluster of elements that are considered as a whole, not individually. These techniques help to increase the wellness / simplification of the model. It is a supervised learning algorithm. It builds a forest of decision trees using a bagging method. This combines many decision trees into a single tree to get more accurate predictions. (Figure 1) gives a clear picture about the working of Random Forest Algorithm.

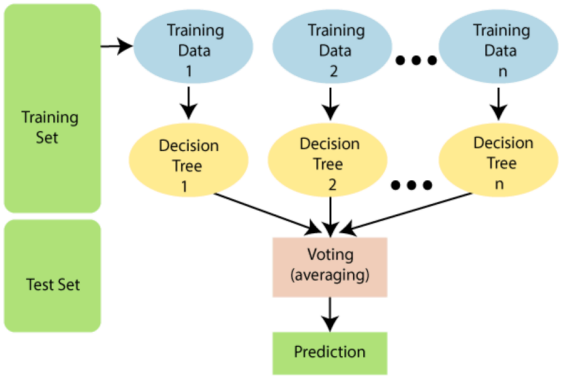


Figure 1. Random Forest

The steps to be followed in Random Forest Algorithm :

- Select K Data Points from the selected Dataset
- Decision trees are build based on K Data Points
- Select N from the Decision trees
- Continue the Step 1 and Step 2
- Using the data points, the test data is classified based on the higher votes

Cnn (Convolutional Neural Networks)

The Neural networks mainly contain three layers. They are the Input Layer, the Hidden Layers and an Output Layer typically. CNN works as stated on the inspiration of the Neurons of the Human Brain. As Neurons take in input and pass the information throughout the body, Convolutional Neural Networks take in input and process them through many hidden layers and send the desired output. Convolutional Neural Networks take an input array. Convolutional Layer is the First layer of Convolutional Neural Networks. There are many layers in Convolutional Neural Networks that are used for feature extraction. (Figure 2) gives a clear picture about the layers in CNN.

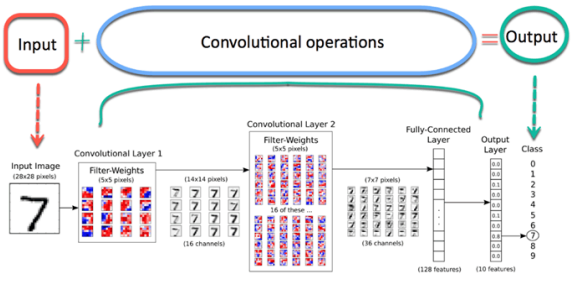


Figure 2. CNN

The steps to be followed in Convolutional Neural Network Algorithm :

- Upload the dataset and perform Scaling feature on the applied input
- The Input Layer
- Convolutional Layer has 14 filters
- Pooling Layer used to reduce the dimensionality of the dataset
- Second Convolutional Layer and Pooling Layer has 32 filters

e-ISSN: 2550-696X p-ISSN: 2550-6978

- Dense Layer connects 1764 neurons, have to define a fully-connected layer
- Final step in the process of prediction

Dataset

The dataset contains Mental and stress level related questions. It contains 80 questions related to predicting the mental health of an individual. Contains 1859 individual tech data for training the model collected from OSMI Mental Health inTech Survey, which contains data for 4 years from 2016 to 2020. For Testing, the response from the staff is used. Daily Routine of an individual, Stress related questions and Disorder related. Disorder dataset contains social anxiety, depression, Anxiety disorder, Mood disorder, Psychotic disorder, Eating Habits, Personality disorders, Addictive disorder, Sleeping Habits & Feelings etc,.

Methodology

Machine Learning Algorithms and Deep Learning Algorithms are applied to achieve the highest prediction accuracy. Random Forest Algorithm is a Machine Learning algorithm that is widely being used for regression and classification. Random Forest Algorithm builds Decision Trees stated on the K number of data points selected from the dataset. Since it handles large numbers of datasets more efficiently, Random Forest is being chosen to predict the stress of an individual. Random Forest Algorithm gives higher accuracy with 20 features in the existing model.
On building the model and testing, the accuracy of the model is 45%. On Increasing the number of features in the dataset and on increasing the number of samples the Random Forest Algorithm's accuracy depreciates. To increase the accuracy of the Prediction, Convolutional Neural Network is used.
Convolutional Neural Network is a Deep Learning Algorithm which can easily classify differences between one object and another. CNN can reduce the complexity in the data and can predict data more accurately. CNN processes the data through many layers. This takes in the input and processes the data in many hidden layers and provides the output. On building the model and testing the model, the accuracy of the model is 96%. The Model is validated using the Accuracy, Specificity, F1-Score, Precision and recall.

Data Preprocessing

The data obtained from the dataset is pretty preprocessed. This information is dissected, and the dataset is produced from this crude information. To utilize the dataset with the python capacities, the information should be standardized. Standardization is a procedure generally applied as a feature of information groundwork for AI. The target of standardization is to change the upsides of numeric segments in the dataset to utilize a standard scale without having contrasts in the scopes of values or losing data. The feature and the label are split and made ready for further processing with the model.

Training and Testing the Model

The model is developed and trained in the python programming environment using various libraries like Pandas, Numpy and matplot. In this research, for training the model, IT sector dataset is used. For testing the model educational sector dataset is used. Training dataset is utilized to fit the model, and Validation Dataset is utilized to give a fair assessment of a model fit on the preparation dataset while tuning model hyperparameters.

Validating the model

Validation is the process in which testing the dataset is done based on the training model. The testing informational index is a different part of the equivalent dataset from which the preparation set is derived. The motivation behind model approval is to check the exactness and execution of the model in light of the past information for which we as of now have actuals. The presentation of the proposed framework has been assessed utilizing Accuracy.

Workflow

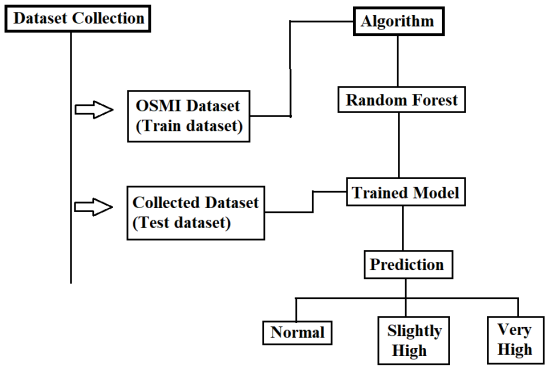


Figure 3. Workflow for Random forest

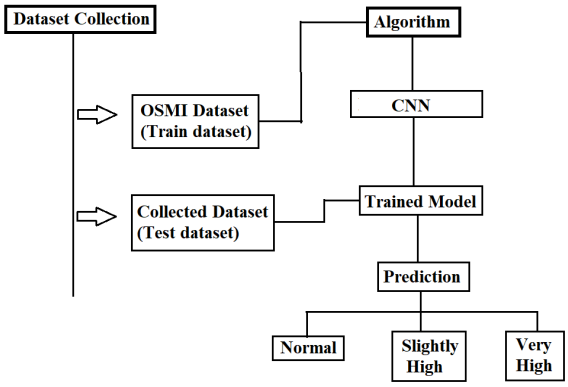


Figure 4. Workflow for CNN

(Figure 3) and (Figure 4) depicts the workflow of the research. The dataset consists of 1859 samples and it is gathered from the OSMI Mental Health in Tech Survey which contains data collected over 4 years. The data in the dataset are 0,1 and 2 where 0 represents NO, 1 represents YES and 2 represents MAY BE.
The dataset gathered is used as the training dataset for random forest and CNN algorithms. Once the models are trained the .sav file is generated. To test the model the dataset collected from the staff community is used. The model is testing from the .sav file generated by the training dataset. The model is validated using the Accuracy, F1-Score, Specificity, Precision and Recall.

3 Results and Discussions

Performance Metrics

Model approval is alluded to as the interaction where a prepared model is assessed with a testing informational index. The testing informational index is a different part of similar informational collection from

e-ISSN: 2550-696X p-ISSN: 2550-6978

which the preparation set is derived. The reason for model approval is to check the accuracy and execution of the model in light of the past information for which we as of now have actuals.

Metrics considered for evaluation:
1. Accuracy
2. Precision
3. F1 Score
4. Specificity
5. Sensitivity

Confusion Matrix: A greatly improved method for assessing the presentation of a classifier is to check out at the disarray framework. The overall thought is to count the times occurrences of class A are named class B. For instance, to know the times the classifier befuddled pictures of 5s with 3s, you would thoroughly search in the fifth line and third segment of the Confusion Matrix.

Precision: Precision is a metric that measures the quantity of right certain expectations made. Precision, subsequently works out the exactness for the minority class. It is determined as the proportion of accurately anticipated positive models divided by the total number of positive models that were anticipated. (Formula 1) portrays precision as the proportion of genuine true positives and the addition of genuine true positives and misleading false positives.

$$Precision = \frac{TP}{(TP+FP)} \quad (1)$$

Sensitivity: Sensitivity is a metric that evaluates the quantity of right certain expectations made from all positive predictions that might have been made. Dissimilar precision that main remarks on the correct positive expectations out of every single positive forecast, Sensitivity gives a sign of missed positive expectations. (Formula 2) portrays the sensitivity as the proportion of genuine true positives and the addition of genuine true positive and genuine false negative.

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

F1 Score: The F1-score, is a proportion of a model's exactness on a dataset. The F1-score is an approach to consolidating the precision and recall of the model, and it is characterized as the consonant mean of the model's precision and recall. Hence, this score considers both misleading false positives and bogus false negatives. Naturally it isn't as straightforward as Accuracy, yet F1 is typically more valuable than exactness, particularly assuming that you have an unevenly distributed class. (Formula 3) depicts the F1 score as two times the proportion of duplication of precision and recall to the addition of precision and recall.

$$F1\ Score = 2 * (\frac{(precision*recall)}{(precision+recall)} \quad (3)$$

Specificity: Specificity is characterized as the proportion of real negatives, which got anticipated as the genuine true negative. This suggests that there will be one more extent of real regret, which got anticipated as sure and could be named as false positives. This extent could likewise be known as a misleading positive rate. (Formula 4) portrays the sum of specificity and misleading positive rate would be 1 all the time.

$$Specificity = \frac{TN}{(TN+FP)} \quad (4)$$

Accuracy: Accuracy is a measurement for assessing arrangement models. Casually, accuracy is the small part of expectations that the model was correct. Officially, (Formula 5) accuracy is the quantity of right expectations per complete number of expectations. The Accuracy acquired is

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (5)$$

Performance Evaluation

(Figure 5) describes the Accuracy of the model by having 80 features and 1859 data for Random Forest and CNN. (Figure 6) describes the Specificity for Random Forest and Convolutional Neural Networks. (Figure 7) describes the Sensitivity for Random Forest and CNN.

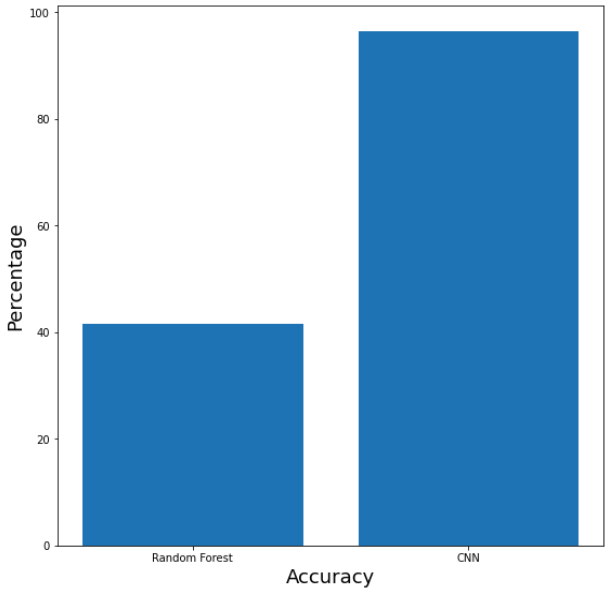


Figure 5. Accuracy

In this research, an IT sector dataset is taken for training the model and a staff community dataset is taken to test the model. First of all, the Random Forest algorithm gives 45 %. So we go for the CNN algorithm and it gives an accuracy of 96 %.

e-ISSN: 2550-696X p-ISSN: 2550-6978

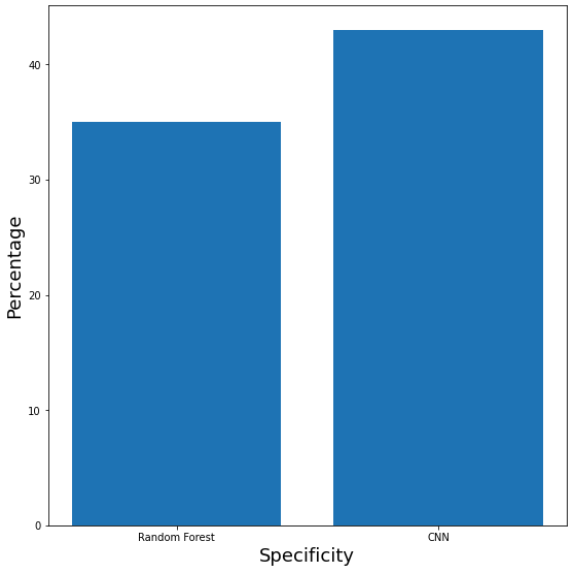


Figure 6. Specificity

In this research, an IT sector dataset is taken for training the model and a staff community dataset is taken to test the model. First of all, the Random Forest algorithm gives 35 %. So we go for the CNN algorithm and it gives the Specificity of 43 %.

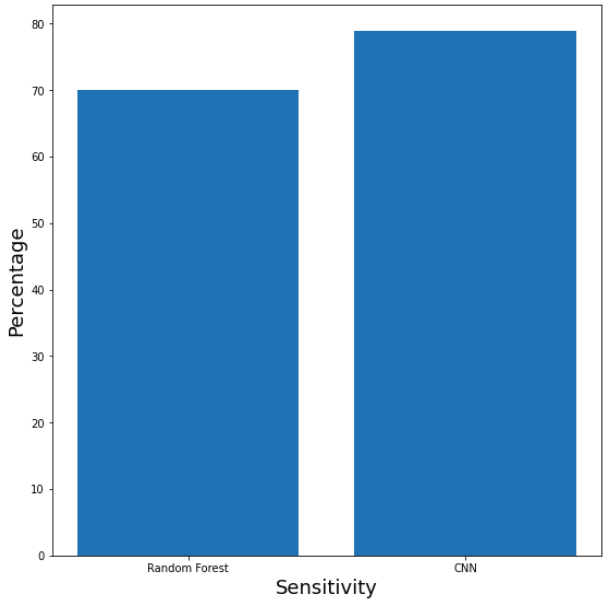


Figure 7. Sensitivity

In this research, an IT sector dataset is taken for training the model and a staff community dataset is taken to test the model. First of all, the Random Forest algorithm gives 50 %. So we go for the CNN algorithm and it gives a Sensitivity of 59 %.

4 Conclusion

The goal of this research is to predict the mental stress level of the staff community based on the survey of questionnaires answered by the individual. Mental stress is being the most common problem for each and every individual in this busy and stressful world. Many machine learning researches are being carried out to forecast the mental health of the IT community. Random Forest is the most commonly used classifier algorithm as it takes less training time and gives more accuracy. It takes data in the dataset and classifies the data, with Random Forest an accuracy of 45% is achieved.

To increase the accuracy of the prediction Deep Learning Convolutional Learning Algorithm is being implemented. With CNN, the accuracy of the model has been increased to 96%. The output includes classifying the data into NORMAL, SLIGHTLY HIGH STRESS AND HIGH STRESS. The future work of this project is to develop a web based application which will be useful for the staff community to self evaluate themselves on the stress level. And have planned to develop an additional module, assessment using RNN. The analysis using RNN module will act as a voice enabled chat bot and ask some questions to the users. It is intelligent to change the questions from the user's mental state and their previous answers. Voice enabled chat bot will be yet more efficient for the user.

References

Jorn Bakker, Leszek Holenderski, Rafal Kocielnik, Mykola Pechenizkiy, Natalia Sidorova "Stess@Work: from measuring stress to its understanding, prediction and handling with personalized coaching", Pages : 673–678, https://doi.org/10.1145/2110363.2110439, 2012.

J. Bakker, M. Pechenizkiy, and N. Sidorova. "What's your current stress level? detection of stress patterns from gsr sensor data". In Proc. of 2nd HaCDAIS Workshop @ ICDM 2011. IEEE Press, 2011.

Alban Maxhuni, Pablo Hernandez-Leal, L. Enrique Sucar, Venet Osmani, Eduardo F. Morales, "Oscar Mayora Stress modeling and prediction in presence of scarce data", Pages 344-356, Volume 63, https://doi.org/10.1016/j.jbi.2016.08.023, 2016.

Ghaderi, A., Frounchi, J., Farnam, A.. "Machine learning-based signal processing using physiological signals for stress detection". 22nd Iranian Conference on Biomedical Engineering (ICBME), p. 93–98, 2015.

Prakruthi Manjunath, Twinkle S, Pola Shreya, Vismaya Ashok, Dr. Shabana Sultana, "Predictive Analysis of Student Stress Level using Machine Learning", INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCCDS – 2021 (Volume 09 – Issue 12), 2021.

A. Kene and S. Thakare, "Mental Stress Level Prediction and Classification based on Machine Learning," 2021 Smart Technologies, Communication and Robotics (STCR), pp. 1-7, doi: 10.1109/STCR51658.2021.9588803, 2021,.

Gray-Toft, P A, and J G Anderson. "Organizational stress in the hospital: development of a model for diagnosis and prediction." Health services research vol. 19,6 Pt 1: 753-74, 1985.

Widanti, N., Sumanto, B., Rosa, P., Miftahudin, M.F.. "Stress level detection using heart rate, blood pressure, and gsr and stress therapy by utilizing infrared". Industrial Instrumentation and Control (ICIC), 2015 International Conference on. IEEE; 2015, p. 275–279, 2015.

Zarski, J. J., Bubenzer, D. L., & West, J. D. "Social interest, stress, and the prediction of health status". Journal of Counseling & Development, 64(6), 386–389. https://doi.org/10.1002/j.1556-6676.1986.tb01143.x, 1986.

Selvaraj, N.. "Psychological acute stress measurement using a wireless adhesive biosensor". 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015, p. 3137–3140, 2015.

C. Setz, B. Arnrich, J. Schumm, R. L. Marca, G. Trster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device." IEEE Trans. Inform. Technol. Biomed., vol. 14, no. 2, pp. 410–417, Mar 2010.

E. Tromp and M. Pechenizkiy. "Senticorr: Multilingual sentiment analysis of personal correspondence". In Proc. of IEEE ICDM 2011 Workshops. IEEE Press, 2011.

e-ISSN: 2550-696X p-ISSN: 2550-6978

Ingegerd Hildingsson, Helen Haines, Margareta Johansson, Christine Rubertsson, Jennifer Fenwick," Childbirth fear in Swedish fathers is associated with parental stress as well as poor physical and mental health", Midwifery, 10.1016/j.midw.2013.12.012, 30, 2, (248-254), 2014.

David Taylo, Jan-Herman Kuiper, "The prediction of stress fractures using a 'stressed volume' concept", https://doi.org/10.1016/S0736-0266(01)00009-2, 2006.

K. Glanz and M. Schwartz. "Stress, coping, and health behavior. Health behavior and health education", Theory, research, and practice, pages 211–236, 2008.

Palanisamy, K., Murugappan, M., Yaacob, S.. "Multiple physiological signal-based human stress identification using non-linear classifiers". Elektronika ir elektrotechnika; 19(7):80–85, 2013

Houtan Jebelli, Byungjoo Choi, SangHyun Lee, "Application of Wearable Biosensors to Construction Sites. I: Assessing Workers' Stress", Journal of Construction Engineering and Management, vol.145, no.12, pp.04019079, 2019.

S. Sriramprakash, V. D. Prasanna, and O. V. R. Murthy, "Stress Detection in Working People," Procedia Comput. Sci., vol. 115, pp. 359–366, 2017, doi: 10.1016/j.procs.2017.09.090.

G. Rigas, C. D. Katsis, P. Bougia and D. I. Fotiadis, "A reasoning-based framework for car driver's stress prediction," 16th Mediterranean Conference on Control and Automation, pp. 627-632, doi: 10.1109/MED.2008.4602162, 2006.

J. Wijsman, B. Grundlehner, H. Liu, H. Hermens and J. Penders, "Towards mental stress detection using wearable physiological sensors," 2011 Annual International Conference of the IEEE Engineer- ing in Medicine and Biology Society, pp. 1798-1801, doi: 10.1109/IEMBS.2011.60905, 2011.

J. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables." in IEEE Engineering in Medicine and Biology Society, vol. 1, pp. 1355–1358, 2006

Maryam Memar, Amin Mokaribolhassan, Amir Aminzadeh Ghavifekr, "Review on Machine Learning Frameworks in Drivers' Physiological Signal Analysis to Detect Stress", 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA), pp.1-6, 2021.

Srividya, M., Mohanavalli, S. & Bhalaji, N." Behavioral Modeling for Mental Health using Machine Learning Algorithms". J Med Syst 42, 88. https://doi.org/10.1007/s10916-018-0934-5, 2018.

M. Gjoreski, H. Gjoreski, M. Lu s˘trek, and M. Gams, "Continuous stress detection using a wrist device - in laboratory and real life," UbiComp 2016 Adjun. - Proc. 2016 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput., pp. 1185–1193, doi: 10.1145/2968219.2968306, 2016.

E. Vildjiounaite et al., "Unobtrusive stress detection on the basis of smartphone usage data," Pers. Ubiquitous Comput., vol. 22, no. 4, pp. 671–688, doi: 10.1007/s00779-017-1108-z, 2018.

G. Giannakakis et al., "Stress and anxiety detection using facial cues from videos," Biomed. Signal Process. Control, vol. 31, pp. 89–101, doi: 10.1016/j.bspc.2016.06.020, 2017.

A. R. Surve, V. R. Ghorpade, and A. S. Patthe, "Continuous Facial Emotion Recognition System Using PCA for Ambient Living", vol. 839. Springer Singapore, 2019.

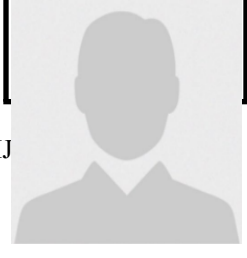
S. Gedam and S. Paul, "A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques," in IEEE Access, vol. 9, pp. 84045-84066, 2021, doi: 10.1109/ACCESS.2021.3085502.

J. Healey and R. Picard, "Detecting stress during real-world driving tasks using physiological sensors," IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 2, pp. 156–166, June 2005.

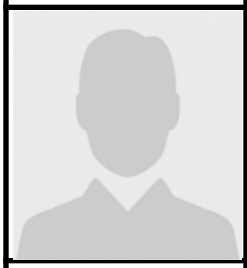
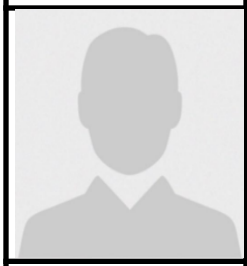
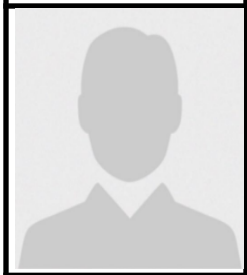
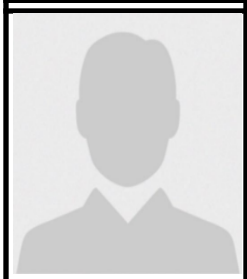
Stanley Leonard Tito, "A Simplified Outpatient Health Monitoring System in Resource Constrained IoT Infrastructure" , International Journal of Engineering Research and Advanced Technology (ijerat): Vol. 5 No. 12: December-2019

P. S. Pandey, "Machine Learning and IoT for prediction and detection of stress," Proc. 2017 17th Int. Conf. Comput. Sci. Its Appl. ICCSA 2017, doi: 10.1109/ICCSA.2017.8000018, 2017.

Biography of Authors



Include in the manuscript a short biography of each author. This section begins on their own page included the photo of the author as well as the short biography about 25 to 100 words.

	Email: firstauthor@gmail.com
	Include in the manuscript a short biography of each author. This section begins on their own page included the photo of the author as well as the short biography about 25 to 100 words. *Email: secondauthor@gmail.com*
	Include in the manuscript a short biography of each author. This section begins on their own page included the photo of the author as well as the short biography about 25 to 100 words. *Email: thirdauthor@gmail.com*
	Include in the manuscript a short biography of each author. This section begins on their own page included the photo of the author as well as the short biography about 25 to 100 words. *Email: fourthauthor@gmail.com*
	Include in the manuscript a short biography of each author. This section begins on their own page included the photo of the author as well as the short biography about 25 to 100 words. *Email: fifthauthor@gmail.com*