Detection of employee stress using machine learning

[1]V.Venkatesh, [2]Shaik.Mahaboob pasha, [3]K.Krishnaveni,[4] G.Ganesh,[5] Mr.B.Thirupathi,

[1,2,3,4,]U.G.Scholor, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.
[5] Research Guide, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

ABSTRACT

With the help of modern decision trees and support vector machines, we can make stress detection more precise. Procedures and Tools for the Study: This work primarily focuses on two areas: support vector machines and novel decision trees. Each group consists of 10 people, and the study's parameters specify a beta of 0.2 and an alpha of 0.05. For the dataset, G-Power, the SPSS program suggested a significance level of 80%. Different algorithms provide different results. New decision trees have an accuracy rate of 89%, whereas support vector machines only manage 84.54%. The dispersion of this data is 0.093. With a p-value less than 0.05—far lower than the significance threshold of 0.020—stress detection in IT personnel is possible. Novel Decision Tree Performs Better Than Support, According to the Results. Vector Machine (84.54% accuracy) by 5.14 percentage points, yielding an 89% output.

Keywords: Included in these parts are the following, IT, MH, ML, stress detection, NEW decision tree, and SVM.

INTRODUCTION

The physiological response of the body to any external stimulus, sometimes referred to as stress, is an emotionally charged and aesthetically arresting condition. Due to the fact that long-haul pressure has detrimental consequences that may cause a variety of issues, such as increased risk of cardiovascular disease, headaches, and poor sleep quality [1]. Factors such as constant risk of harm, short deadlines, difficult job, or even boring housework may all contribute to an unpleasant work environment. By discreetly wearing pressure-detecting devices that continually monitor levels of anxiety, employees in the information technology area may begin pressure-lowering treatment with minimum disturbance to their workdays. Differentiating between stress and depression is currently only possible for those working in the medical field and physiological research. Paying close attention to form is a common strategy for stress assessment [2]. Because this approach is dependent on the respondent's words alone, they may be reluctant to admit that they are conventional or anxious. By installing automated stress detectors, we can improve community health and reduce the likelihood of mental health issues emerging.

Consequently, a method for automatically identifying human stress levels by analyzing physiological data must be developed that is grounded on solid science.The potential for stress detection to be a massive societal benefit that boosts people's happiness has led several writers to highlight its significance. As new goods and technologies are being introduced into the information technology industry, public relations is being revamped. An approach to stress detection based on facial expressions has previously been developed [3]. The second study tracked participants' levels of professional stress and psychological well-being using a mix of visual cues and physiological signals. Without a doubt, these examinations are both invasive and unpleasant [4]. A stress index compares the results of each sensor and serves as a stress detection threshold.

Twelve articles in Science Direct and twenty-three in IEEE Xplore discuss this research. Formulate a Novel Decision Tree using Structured Data [5]. The goal of using machine learning for stress detection was accomplished. In order to construct categorization modeling approaches, product application evaluations were analyzed [6]. The first study focused on stress detection using ML techniques. Automated stress detection might be improved by considering a person's age, level of professional stress, mental health, family concerns, and emotional state [7]. Using these machine learning categorization techniques, differentiation is feasible [6].

In light of the method and the results of improved accuracy, the purpose of the article's last discussion is to establish the validity of all one thousand datasets [8]. Modern standard feature extraction methods evolved out of a need for fast analysis [9]. Using a variety of machine learning approaches, this work aims to enhance stress detection.

MATERIALS AND METHODS

The Saveetha School of Engineering in Chennai is home to the Machine Learning Lab, where researchers do their work. The institute is also known as the Saveetha Institute of Medical and Technical Sciences.Machine learning approaches include novel decision trees and support vector machines. The classifier is where you may find these approaches, according to [10]. A novel decision tree was employed by one group of participants in this study, while support vector machines were used by the other. As a result of the To reach parity, ten iterations are needed using An 80.0 percent or higher G-power, a p-value below 0.05, and a 95% CI. Among its many recent updates is the stress dataset. In order to get this data, we use the open-source software application Kaggle.

Using python, Google Collaboratory analyzed a dataset consisting of 1,008 rows and 14,002 columns. The fourteen columns cover a wide range of topics, including variables, such as chronological age, gender, chest pain, s.n.o., cholesterol, fasting blood sugar levels, ECG at rest, exercise-induced angina, and peak heart rate during activity. Always remember that the testing environment may include several software and hardware versions.This laptop has aEquipped with a 64-bit OS and 16 GB of RAM, an x64-based central processing unit, and a hard drive. The

latest build of the program is compatible with Windows 11 and was co-coded in Python. Once the software finishes running, the accuracy number will be shown. Google Collaborative is

solely responsible for the completion of the python code for our research. We got 10 different accuracy numbers when we ran the final file through 10 iterations and saved it. We use SPSS to calculate the mean accuracy once we have collected 10 sets of accuracy data.

Novel Decision Tree Algorithm

Novel One kind of supervised learning approach that doesn't need parameters is decision trees. has found usage in regression and classification applications. There is a single root node, two intermediate nodes, and a single terminal node in its hierarchical structure.

Algorithm

Input : Training dataset S.

Output: A collection of datasets used for testing

Step 1: discover the whole inventory for each category

Step 2 :Make advantage of the given information to build a fresh decision tree.

Step 3 :In order to run tests, data collection is required.

Step 4 : Repeat the steps of 1 & 2 .

Step 5 :Applying the method, rank each set according to its significance.

Support VectorMachine Algorithm

Support vector machines are often used in stress detection for future prediction. Its effect on stress prediction is substantial. So, it's possible for the software to tell whether someone is nervous.

Algorithm

Input :Data for the Kth training set.

Output: A dataset has to be put together before any testing can be done.

Step 1 :Picking K random data points at random from the training set is the first step.

Step 2 :Make a New Decision Tree Using the Chosen Data Points

Step 3 :To tell the program how many new decision trees to create, type N.

Step 4 : Repeat steps 1 and 2 again.

Step 5 :Sort the additional data points by popularity once the Novel Decision Tree has predicted them all.

Statistical Analysis

This SPSS program is used to compute the independent t-test. We assess the iterations after 10 and make predictions about the future performance of the collaborative code based on them. In order to compare the two groups' statistically significant results, a t-test was used. This test is based on independent samples. The components for stress detection Knowledge of experts and, on the other side, generic information make up the independent variables.

RESULTS

These two machine learning techniques are SVM and Novel Decision Tree. Table 1 displays the numerical accuracy values of the innovative decision tree as an alternative to StVM iterations. The Novel Decision Tree has an average accuracy of 89% and a standard deviation of 2.8, as shown in Table 2. The precision and standard deviation of support vector machines are 84.82% and 3.02, respectively. The Novel Decision Tree consistently outperformed the Support Vector Machine in all tests. Assuming a mean of 9.6 and a standard deviation of 0.93.Whereas support vector machines Novel Decision Tree that was suggested performed better in our tests. Our new decision trees and support vector machines performed well in the independent samples T-test, as shown in Table 3. Using 0.020% level, the findings are considered statistically significant. Figure 1 is a bar graph comparing and contrasting the usual accuracy of the SVM algorithm with that of the Novel Decision Tree approach. The innovative decision tree achieved an average accuracy of 84.54% by using support vector machines.

DISCUSSION

Great results with low standard deviation (84.54%) were achieved by both techniques, however the Research demonstrated an astounding 89% success rate in stress detection by using the state-of-the-art Decision Tree method. By using SVM method was also excellent.
A novel decision tree for application stress detection is described in this paper, which has an 89% success rate [11]. Data from 2021 reveals that the revolutionary decision treeTreeTree is 89% accurate. Another approach that has been either used or compared to evaluate stress detection is support vector machines; in the previous research, it achieved an accuracy range of 80%, and in the present one, it reached 84.54% [12]. Stress datasets have found utility in mental health and prediction models due to their predictive power.

The task at hand and the dataset at hand dictate the accuracy level that any one of many methods can achieve [13]. While both old and new methods of analyzing a given article use metrics like support vector machine (SVM) to evaluate stress detection, Novel Decision Tree seems to be more precise [14].The decreasing accuracy number is due in part to the study's restrictions, most notably the dataset's duration. Research has slowed down since there aren't enough training and testing accuracy datasets that work with smaller datasets. Developing computationally efficient and highly accurate algorithms for stress detection categorization is the next step in our research.

CONCLUSION

In order to detect application stress, this study used novel decision trees and support vector machines. Taking into account the SVM 84.54% accuracy rating, the Novel Decision Tree's 89% is much higher. Compared to support vector machines, novel decision trees seem to be better at stress detection. The main motivation for its invention was to find a mechanism to identify stress.

On the testing dataset for the research, we found that Novel Decision Tree had better classification accuracy than SVM classifier. Results from evaluations utilizing training data showed that the SVM classifier performed better than competing approaches. If you'd like, you may apply the same method to further instances of the sampled dataset. Thanks to the data visualization and the newly found characteristics, the dataset became simpler to study while also improving the classification accuracy. You may evaluate the performance of each algorithm according to its accuracy criteria by looking at the proportion of correct approaches.

DECLARATIONS

Conflicts of Interest

Nothing in this piece implies that the author has any kind of conflict of interest.

Authors Contribution

Not only did Author SP write the article, but she also gathered and analyzed the facts. Author BBB conceived of the concept, checked the facts, and gave the text a critical eye.

Acknowledgements

Thanks to Saveetha's generous support, the engineering department at what was formerly Saveetha University (now Saveetha Institute of Medical and Technical Sciences) was able to finish the project.

Funding

Thanks to the following organizations for providing financial support that enabled us to complete the study.

1. HKM Info Systems Pvt.Ltd.
2. Saveetha School of Engineering.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha University.

REFERENCES

[1] M. B. Bin Heyat*et al.*, "Role of Oxidative Stress and Inflammation in Insomnia Sleep Disorder and Cardiovascular Diseases: Herbal Antioxidants and Anti-inflammatory Coupled with Insomnia Detection Using Machine Learning," *Curr. Pharm. Des.*, Dec. 2022, doi: 10.2174/1381612829666221201161636.

[2] M. M. Hussain, M. MahammadHussain, and V. Karthick, "Efficient Search in Cloud Storage with Reduced Computational Cost using Token Generation Method over Crypto Hash Algorithm," *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*. 2022. doi: 10.1109/icosec54921.2022.9952137.

[3] A. M. Elsherbini*et al.*, "Decoding Diabetes Biomarkers and Related Molecular Mechanisms by Using Machine Learning, Text Mining, and Gene Expression Analysis," *Int. J. Environ. Res. Public Health*, vol. 19, no. 21, Oct. 2022, doi: 10.3390/ijerph192113890.

[4] D. Swain, P. K. Pattnaik, and P. K. Gupta, *Machine Learning and Information Processing: Proceedings of ICMLIP 2019*. Springer Nature, 2020.

[5] N. S. Chandel, Y. A. Rajwade, K. Dubey, A. K. Chandel, A. Subeesh, and M. K. Tiwari, "Water Stress Identification of Winter Wheat Crop with State-of-the-Art AI Techniques and High-Resolution Thermal-RGB Imagery," *Plants*, vol. 11, no. 23, Dec. 2022, doi: 10.3390/plants11233344.

[6] F. Ritsert, M. Elgendi, V. Galli, and C. Menon, "Heart and Breathing Rate Variations as Biomarkers for Anxiety Detection," *Bioengineering (Basel)*, vol. 9, no. 11, Nov. 2022, doi: 10.3390/bioengineering9110711.

[7] K. V. Bhavitha and S. J. J. Thangaraj, "Novel detection of accurate spam content using logistic regression algorithm compared with Gaussian algorithm," in *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, IEEE, Feb. 2022. doi: 10.1109/icbats54253.2022.9759003.

[8] Z. Tan *et al.*, "PollenDetect: An Open-Source Pollen Viability Status Recognition System Based on Deep Learning Neural Networks," *Int. J. Mol. Sci.*, vol. 23, no. 21, Nov. 2022, doi: 10.3390/ijms232113469.

[9] J. Zhang, S. Sami, and R. Meiser-Stedman, "Acute stress and PTSD among trauma-exposed children and adolescents: Computational prediction and interpretation," *J. Anxiety Disord.*, vol. 92, p. 102642, Dec. 2022.

[10] W.-T. Kang *et al.*, "Qualitative and Quantitative Detection of Multiple Sexually Transmitted Infection Pathogens Reveals Distinct Associations with Cervicitis and Vaginitis," *MicrobiolSpectr*, p. e0196622, Oct. 2022.

TABLES AND FIGURES

Table 1 Use Support Vector Machine (SVM) and Novel Decision Tree (NDT) on a 10-item sample. The correct NDTC and SVM parameters were found.

ITERATION	NDTC	SVM
1	89.00	84.54
2	88.73	84.12
3	88.42	83.84
4	87.97	83.57
5	87.73	83.43
6	87.56	82.86
7	87.22	82.62
8	86.94	82.11
9	86.62	81.91
10	86.12	81.51

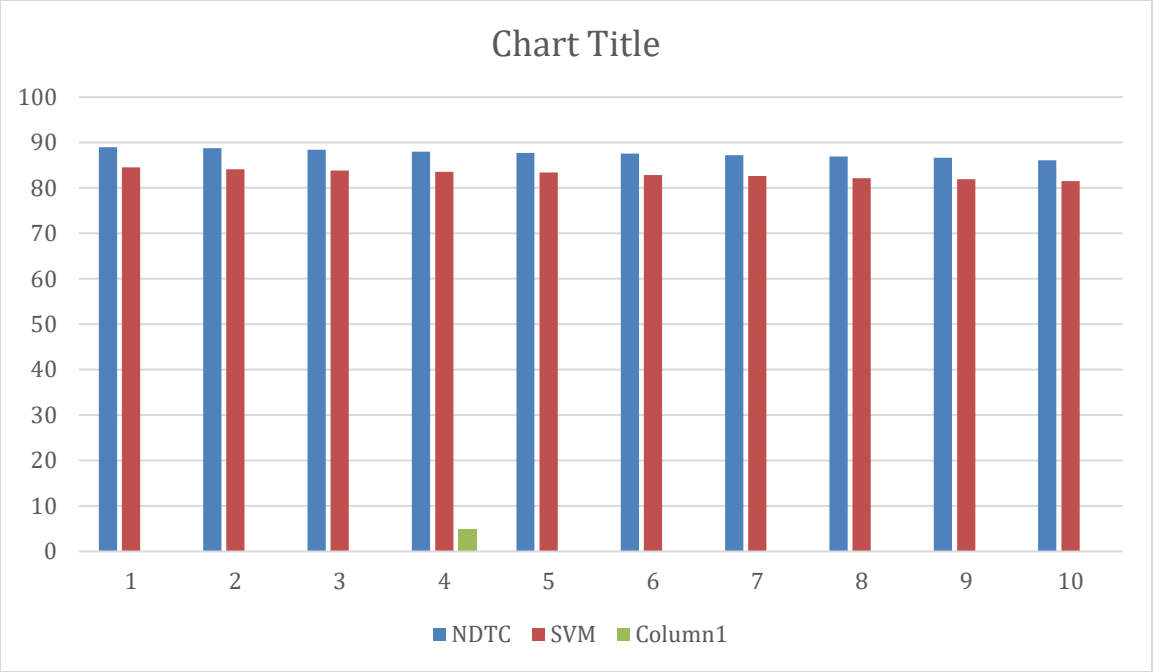

Fig 1. Finding out how Novel Decision Tree (NDT) and Support Vector Machine (SVM) stack up against one another might be challenging.

Table 2. Support Vector Machines are less accurate than Novel Decision Trees, which have a standard deviation of 1.21 and a performance level of 84.54%.

	Computer programme	**N**	**Mean**	**Typical Deviation**	**Typical Error Mean**
Accuracy	NDTC	10	89.0001	.93011	.29413
	SVM	10	84.5400	1.21579	.38447

Table 3. Using a randomly chosen sample, determine the margin of error and conduct a significance test. We considered the result statistically significant when, after computing the 95% confidence intervals, the p-value was less than 0.05.

Checking for Equality of Variances with Leven's Test				**A t-test to look for mean equality**			**Radius from the Midpoint with 95% Reliability**	
		F	**Sig**	**Important (2-tailed)**	**Average Distinction**	**Variance in Standard Error**	**Lower**	**Better still**
Accuracy	Presumption of Equal Variance	.232	.020	.001	4.01300	.48407	2.99600	5.03000
	Equal Variance not assumed			.001	4.01300	.48407	2.99099	5.03501

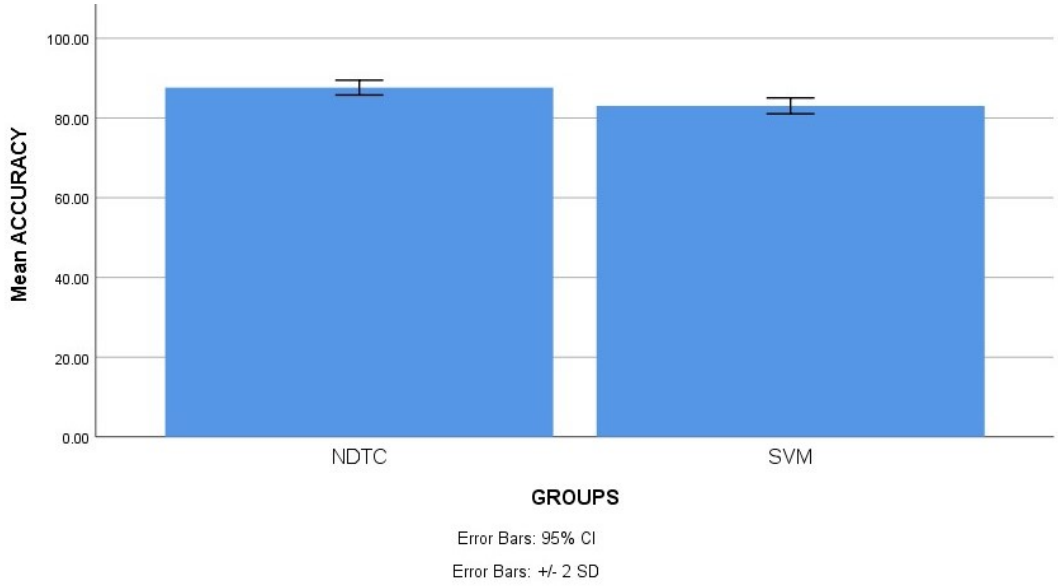

Fig. 2 Two techniques' mean gains, as calculated using SPSS, are shown in the bar graph. A t-test and an independent samples bar chart may be created when all the average data is entered into the SPSS program. We may see the categories on the one hand, and the average accuracy on the other. The average precision is given by the 95% CI, which is plus or minus 2 standard deviations.