International Journal of Public Health Science (IJPHS)
Vol.13, No.2, June 2024, pp. 944~954
ISSN: 2252-8806, DOI: 10.11591/ijphs.v13i2.23577

Evaluation of stress based on multiple distinct modalities using machine learning techniques

Samarendra Narayana Pradhan[1], Reddy Shiva Shankar[1,2], Shekharesh Barik[1], Bhabodeepika Mohanty[1], Venkata Rama Maheswara Rao[1,3]
[1]Department of Computer Science and Engineering, Faculty of CSE, Dhaneswar Rath Institute of Engineering and Management Studies (DRIEMS), Tangi, Odisha, India
[2]Faculty of CSE, Sagi Rama Krishnam Raju Engineering College (A), Bhimavaram, Andhra Pradesh, India
[3]Faculty of CSE, Shri Vishnu Engineering College for Women (A), Bhimavaram, Andhra Pradesh, India

Article Info

Article history:

Received Jun 22, 2023
Revised Oct 23, 2023
Accepted Nov 2, 2023

Keywords:

Artificial intelligence
Auto encoders
Convolutional neural networks
Deep learning
Machine learning
Principal component analysis

ABSTRACT

Nowadays, one of the most time-consuming and complex study subjects is predicting working professionals' stress levels. It is thus crucial to estimate active professionals' stress levels to aid their professional development. Several machine learning (ML) and deep learning (DL) methods have been created in earlier articles for this goal. But they also have drawbacks, such as increased design complexity, a high rate of misclassification, a high incidence of mistakes, and reduced efficiency. Considering these issues, the objective of this study is to make a prognosis about the stress levels experienced by working professionals by using a cutting-edge deep learning model known as the convolutional neural networks (CNN). In this paper, we offer a model that combines CNN-based classification with dataset preprocessing, feature extraction, and optimum feature selection using principal component analysis (PCA). When the raw data is preprocessed, duplicate characteristics are eliminated, and missing values are filled.

Corresponding Author:

Samarendra Narayana Pradhan
Department of Computer Science and Engineering, Dhaneswar Rath Institute of Engineering and Management Studies (DRIEMS)
Tangi, Cuttack, Odisha, India
Email: sunilcst263@gmail.com

1. INTRODUCTION

Stress affects more than 70% of the population. Long-term stress causes reduced immunity, cancer, cardiovascular disease, depression, diabetes, and drug addiction [1]. Stress damages mental and physical health. Developing trustworthy technologies to identify human stress quickly is crucial. These technologies might detect stress continuously. Hence, stress may be reduced by regulating daily activities, and healthcare practitioners can better manage stress-related disorders. Researchers have developed many stress-detecting methods. The sympathetic nervous system releases adrenaline and cortisol when threatened [2]. So, this condition may significantly damage a stressed person's daily life and health [3]. Hence, stress and other factors may cause fatal car accidents [4]. According to the World Health Organization (WHO), 1.35 million road traffic crashes kill 5–29-year-olds [5]. So, it is important to create strategies to immediately spot stress in drivers so they may avoid automobile accidents and injuries, which can greatly impact the driver's and accident victims' lives. These artificial intelligence (AI) models may assess stress in a variety of contexts, such as the workplace [6], while driving [7], in the classroom [8], and during crises [9].

AI models may help self-regulate stress, and extreme stress might alert human resources (HR), management, or teachers to change the atmosphere and workload. AI models may help with affective computing in other soft skills. Stress may cause depression, addiction, and cardiovascular diseases. Emotional tension now harms mental and physical health. Psychological stress is mitigated by ecological momentary assessment (EMA) [10].

Nevertheless, the two systems require real-time psychological stress monitoring, which is the main problem. The sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) are linked to psychological stress, which is fortunate PNS. Stress enhances SNS activity by reducing the inter-beat (RR) interval, increasing the low-frequency energy of the heart rate variability (HRV), increasing the respiration rate, decreasing the high-frequency energy, and other mechanisms [11]. The reverse happens when stress stimulates the PNS. The RR interval and other physiological signs may detect stress. RR interval information from the electrocardiogram (ECG) signal has been utilized to detect psychological stress. The ECG data collected over five minutes may be used to identify stress. Sadly, breathing signals necessitate a bandage, which is painful. Real-time monitoring needs less than one minute.

A kind of stress that positively influences an individual is referred to as eustress. A person feels stress when they expect a thrilling occurrence in their region. The pulse rate rises, yet there is no fear or danger. Whether pursuing a promotion or giving birth, people are more prone to suffer this kind of stress. Eustress may improve mental power and productivity. By pushing people beyond their comfort zone, eustress helps them grow. Stress causes anxiety and worry, which may cause misery. It might last a short time or longer. Stress may cause performance deterioration and mental fog. Chronic or severe diseases can cause suffering for the brain and body; the body cannot handle depression, and other health difficulties result. External and internal stimuli may produce pain. Fear, concern, bad thoughts, perfectionism, overscheduling, poor future planning, high work expectations, job insecurity, inability to assert, and other factors may cause pain. Personal concerns such as grief, illness, accident, financial difficulties, job loss, insomnia, and legal issuesmay contribute to an already difficult situation. Hence, stress should be diagnosed early since it may majorly impact people's lives.

Stress may trigger the "fight or flight" response. For survival, strain helped individuals respond quickly to life-threatening or unpleasant events. When threatened or challenged, the body engages in its self-defense systems. These tools aid or accelerate evasion. This produces shortness of breath, faster pulse, higher blood pressure, muscle tension, and alertness. In difficult circumstances, his strength, endurance, alertness, and response speed help him determine whether to fight or run.

Figure 1 depicts the conceptual framework for our study on stress. This article examines the stress-reducing effects of sleep, physical exercise, working hours, and heart rate. The stress levels experienced by individuals working in various professions are now the subject of investigation by a number of scholars. With Fitbit's wireless activity tracker, we're examining professionals' stress levels (low, medium, high). Currently, we have the following research questions (R) and hypotheses (A):

Q: How much stress do these factors cause professionals?
A1: Is there a correlation between stress and sleep?
A2: Is there a connection between their physical activity and stress levels?
A3: Is there a correlation between working hours and stress?
A4: Is there a connection between any variations in heart rate and their stress levels?

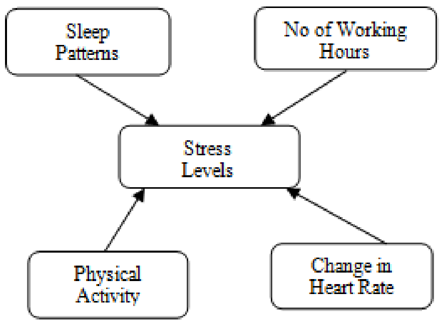

Figure 1. The flow model of stress

2. RELATED WORK

Reddy *et al.* [12] stated that it is not uncommon for individuals in the workforce to experience stress-related mental health disorders. Machine learning, a type of artificial intelligence, lets machines learn from their prior experiences. Flesia *et al.* [13] stated that the study sample showed a higher prevalence of perceived stress levels when compared to standard values in Italy. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a coronavirus strain that can cause a spectrum of respiratory symptoms, from acute cold-like to severe acute respiratory syndrome. Coping, a widely studied dispositional trait, is crucial in modulating responses to stressful events. While education does not appear to impact psychological health significantly, the SARS outbreak revealed that higher education can affect mental health. Therefore, it is essential to sustain the utilization of psychological services to help prevent the emergence of chronic outcomes. Furthermore, ensuring that those in need have access to timely psychological services is critical.

Mathangi [14] probed that increasing stress levels among employees can lead to low self-esteem. They collected data from 143 individuals and found that workload had the most significant impact on job productivity; job security and shift work come next in the sequence. The standardized coefficient for workload was 0.369, which is much higher than that for job security (0.221) and shift work (0.351). By taking appropriate measures to reduce job stress levels, management can boost employee morale and productivity. According to Izard's theory, a critical safety scenario might cause emotions, according to Kania *et al.* [15]. This approach specifies four systems that activate emotions: the neural system, activated by brain electrical stimulation. The best classifier accuracy was 81.63% using k-nearest neighbors (kNN). While kNN with principal component analysis (PCA) had the maximum sensitivity, it had a low specificity, suggesting that the classifier was less accurate than kNN.

S.S. Reddy *et al.* [16] have worked on diabetes mellitus (DM) patients and their correlated ailments [17] like renal fault [18]. Through this research, they have identified that people's stress may affect DM. So many people have been readmitted to the hospital [19]. Jaques *et al.* [20] discussed the advantages of using the multimodal autoencoder (MMAE) for detecting mood, health, and stress, especially in the presence of missing data sources. With the increasing need for long-range and the challenge of incomplete data sources in multimodal data collection, the MMAE offers a solution to this issue. Unlike other techniques, such as PCA, the MMAE can accurately reconstruct data from a missing modality. This is demonstrated by the authors, who show that the MMAE's encoding can facilitate robust mood prediction, even in the case of significant data loss-up to 75% of the data sources. Data was gathered through physiological sensors, a smartphone app, and questionnaires to predict future mood and diagnose mental illness, sadness, and anxiety. The study's ultimate objective is to establish a real-world system that can assist people in modifying their attitudes and spotting mental diseases early.

Rosa *et al.* [21] contemplated that the proposed system aims to detect depression and stress among users and send warning messages to authorized individuals. The system utilizes both a convolutional neural network (CNN) and a bi-directional long short-term memory (BLSTM)-recurrent neural network (RNN) for detecting depressive and stressful sentences with an accuracy of 0.89 and 0.90, respectively. The system includes a sentiment metric and ontologies to improve user satisfaction. As a result of incorporating these features, the system received a satisfaction rating of 94%, compared to 69% for a system lacking such features. Furthermore, based on the monitoring results, the system is programmed to send messages to users with psychological disorders intended to be uplifting, calming, relaxing, or motivational. The proposed solution is memory-efficient and has low processing and energy consumption, making it suitable for mobile electronic devices.

Ghosal *et al.* [22] developed a deep learning (DL) based approach for plant stress identification and classification, focusing on explain ability. Their model uses high-resolution feature maps to identify and segregate visible symptoms precisely, contributing to stress prediction and allowing for more transparent and interpretable results. DL has proven to be a powerful tool in various scientific fields, including medical imaging and neural response modeling, due to its high accuracy levels. Aceto *et al.* [23] proposed that the proliferation of handheld devices has resulted in a massive increase in cellular traffic passing through the home and corporate networks and the Internet. This has made it more challenging to design accurate classifiers, especially with the growing use of encrypted protocols like TLS. Effective deep packet inspection techniques have become less suitable as a result. The effectiveness of DL classifiers is extensively evaluated based on three real human mobile datasets. The research highlights design guidelines, challenges, and potential pitfalls in using these classifiers. The study also proposes a practical strategy for designing mobile traffic classifiers that can handle encrypted traffic using automatically extracted features that reflect complex traffic patterns.

Jaber *et al.* [24] aimed to improve the user-friendliness and practicality of AI-based stress prediction by creating an explanatory report modeled after standard blood test reports. The report presents a stress prediction model that incorporates physiological signals, which can aid patients and practitioners in comprehending the rationale behind the prediction. The authors suggest a novel approach for developing explainable AI in stress prediction that can be customized to suit the preferences of individual users. With the help of this report, users can identify the most influential biological features and any health abnormalities that may contribute to stress prediction, making it easier to take appropriate action.

Liang *et al.* [25] analyzed that DL approaches have become increasingly important in complex evaluation, maintenance, and safety. They utilized a unique DL approach that can quickly and accurately estimate stress distributions, which can substitute for the time-consuming and expensive finite element analysis (FEA) in stress analysis of the human thoracic aorta. Their approach uses a shape encoding module that resembles a bidirectional neural network. However, unlike other studies, shape decoding was not required since they didn't need to reconstruct the shape. A separate neural network must be developed and trained if such a mapping is required.

Jaques *et al.* [26] aimed to predict an individual's mood for the following day and personalized the approach for each person. They utilized multitask learning (MTL) and domain adaptation (DA) methods to train a personalized model that can be created for each individual while also using data from the larger population. They collected data continuously from wrist-worn physiological sensors, a mobile application, and daily surveys to monitor participants. This information and weather data were utilized to construct customized deep neural network (DNN) and Gaussian process (GP) models using multitask learning and domain adaptation. To personalize ML models, they developed systematic techniques by creating personalized models for each individual while still capitalizing on similar people's data. They trained DNNs and GPs to forecast an individual's mood, health, and Stress intensity for the following day based on current physiological, behavioral, and weather data.

Sumathi *et al.* [27] inquired that, nowadays, faculty responsibilities go beyond teaching courses to students. They also provide guidance, counseling, and mentoring to prepare them for future employment or entrepreneurship opportunities. Female faculty face the added challenge of balancing family responsibilities with work obligations. The management plays a critical role in ensuring that faculty members are free from stress and gratified with their work. The stress levels among the faculty members, it is recommended that the management avoid any indications of preferential treatment or discrimination during the promotion and salary review process. All faculty members should be treated equally. A recent study has identified that stress among faculty members can be attributed to several factors, such as unanticipated changes to their class schedules, inadequate promotion policies, overwhelming documentation requirements, and instances of partiality or discrimination.

Liapis *et al.* [28] contemplated that there are multiple datasets accessible to researchers for stress detection, which have been recorded and can be used for benchmarking and testing purposes. These datasets usually entail subjecting users to high-stress scenarios, including gaming scenarios, image datasets, software and hardware failures, songs, and movie clips. To gain an in-depth understanding of users' interaction experiences beyond traditional usability metrics, researchers and practitioners use user experience (UX) techniques. These methods involve post-questionnaires, observation, and interviews to measure the emotional aspects of UX. This research used traditional ML and DL techniques to evaluate the performance of an existing bio signal dataset in the context of UX assessment. The study results indicate that using available bio signal datasets in diverse contexts requires careful consideration. Rachakonda *et al.* [29] declared sleep essential in managing stress and maintaining optimal health and well-being. This data is securely transferred to an Ethereum private blockchain to monitor sleep quality and detect stresses that may be straining the body.

Richter *et al.* [30] created a machine learning method to identify anxiety and depression by recognizing a distinct pattern of biassed emotional responses. The algorithm was tested on a dataset of participants from both groups, with 80% used for training and 20% for validation. The results showed 71% sensitivity, 70% specificity, and 68% classification accuracy for the high-anxiety group and 74% for the high-depression group in a two-group model. Lin *et al.* [31] analyzed that identifying and tracking stress can greatly impact well-being. Traditional methods of categorizing emotional states rely on machine learning algorithms that calculate features from various sensor inputs. This stress detection technique has the best accuracy of 97.8% compared to other approaches.

3. METHOD

DL is a subfield of symbolic learning, and it is characterized by the use of layered models to acquire representations of data in an all-the-way-through method. These methods have already had a sizeable impact up to this point, and it is anticipated that they will continue to revolutionize how complex data analysis tasks are tackled in various fields. Within the realm of DL, in addition to the time-honored supervised and unsupervised learning methodologies, many alternative learning paradigms have emerged as potentially valid. It is considered that enormous quantities of data (on the scale of hundreds of thousands) are needed to correctly train the high number of parameters in a deep model and prevent overfitting. In certain situations, labeling (supervision) and sufficiently thorough data collection may be prohibitively expensive or impossible. It is usual practice to use several tactics and methods for data augmentation.

3.1. Objectives

This study intends to improve the accuracy and effectiveness of a system that predicts stress using a cutting-edge classification strategy that employs deep learning. The stress dataset, accessible on Kaggle, is a public dataset for this investigation. The collection includes information on the stress levels of a wide range of workers. A subset of characteristics is picked from the normalized dataset to enhance prediction further.

3.2. Fully-connected

Networks (or layers of a network) are said to be fully connected has been illustrated in Figure 2, if and only if every node in one layer can directly connect to every node in the next layer, as shown in Figure 2(a). About the information coming into it, each neuron performs the function of a summation node. A non-linear activation function will be applied to the output at some point in the process. The fully connected layer is one of the simplest in a network and is typically used for the final stage of classification or regression.

3.2.1. Convolutional neural networks

When trained on sampled data, CNNs are more effective and more accessible to train than traditional deep feed-forward networks. This is because of the limitations placed on the hypothesis space, which impose order and simplify the set of variables to be considered. Because of the constraints imposed by the mandated structure, the resulting features are rotationally and spatially invariant. This is made feasible by using pooling layers, shared weights, and local connections. Color photographs are needed to process the matrices or tensors, and CNNs are built, as shown in Figure 2(b).

3.2.2. Recurrent neural networks (RNNs)

It is a significant subset of the DL family and is primarily designed to handle sequential input, as shown in Figure 2(c). A simple RNN is, in fact, not nearly as strong as other types of RNNs, and it is only sometimes utilized in contemporary works. However, recurrent hidden units such as long short-term memory (LSTM) and gate recurrent unit (GRU) may perform very well. These modules are built from many independent data pathways that temporarily store and release data to solve the vanishing gradient issue.

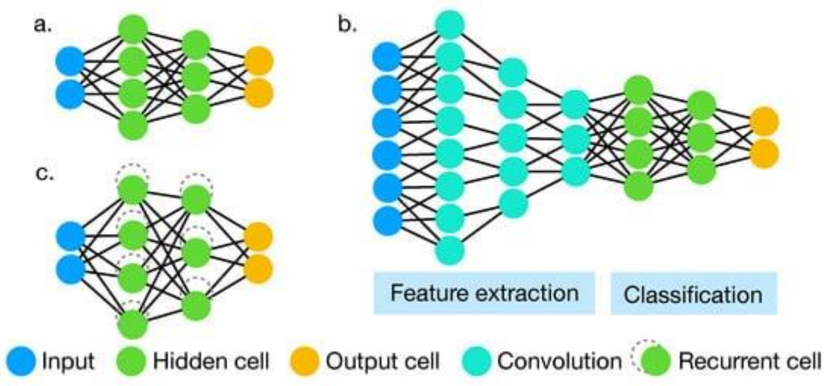

Figure 2. Network architectures (a) fully-connected (b) convolutional neural network (c) recurrent neural network

3.2.3. Autoencoders

One layer of inputs, one layer of hidden units, one layer of reconstruction units, and an activation function comprise an auto encoder (AE) Figure 3. The latent vector is generated by first projecting the input into the hidden layer during training. The analogous network is referred to as the encoder in this context. The encoder's output is then decoded to a layer the same size as the input as shown in Figure 3(a). AEs' strength rests in their unsupervised training, which mandates a meaningful compact representation at their very heart.

3.2.4. Deep belief networks (DBNs)

It is a mixture of simple, unsupervised networks like restricted Boltzmann machines (RBM) or autoencoders, where each hidden layer of a sub-network is the visible layer for the next layer Figure 3(b). When it comes time to fine-tune, a feed-forward network is added if required. The input characteristics are represented as activity in the hidden layer and are the sole information used by AE during reconstruction. Stack-trained encoders reduce abstract semantic information loss and boost model capacity SAE. The SAE work flow was shown in Figure 4.

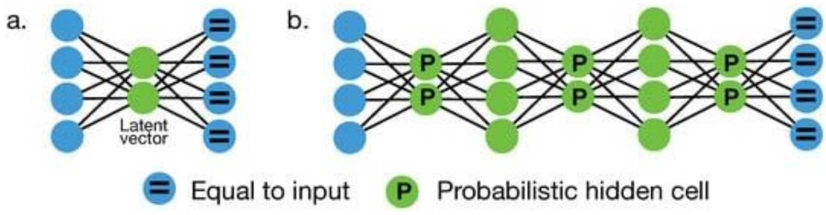

Figure 3. Network architectures, (a) autoencoders (b) deep beliefnetworks **(**DBNs)

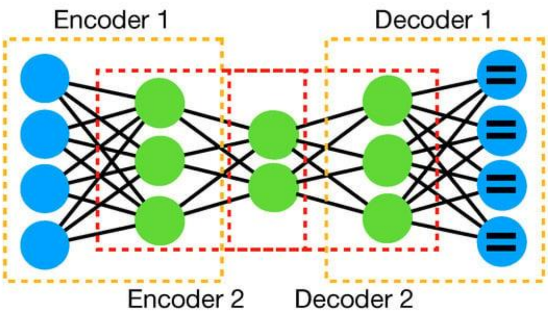

Figure 4. Network architecture of a stacked autoencoder

3.2.5. Generative adversarial networks

One interesting new method for building and training generative models is using adversarial networks, or generative adversarial networks (GANs). Two adversarial neural networks, a generator G and a discriminator D Figure 5, are trained in tandem in this system. Using random inputs, the generator is trained to produce samples from a specified distribution, while D looks for patterns indicating whether the data is genuine. Training occurs in a two-player min-max game setting until the produced data can be reliably differentiated from the actual ones. D may be employed as a well-trained feature extractor after an appropriate training technique and then applied to a particular issue by including a final block that utilizes the required output (for example, a fully connected layer for classification).

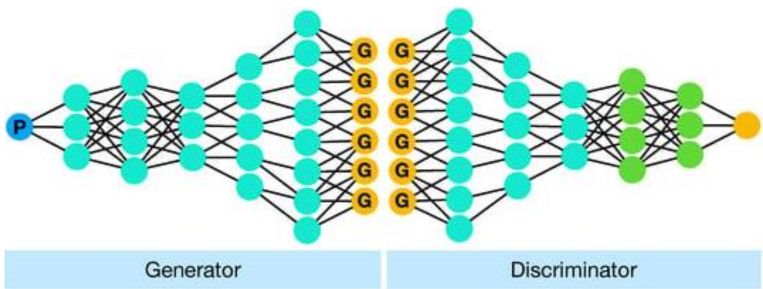

Figure 5. Architecture of GANs

3.3. Proposed work

By minimizing dataset attributes while keeping the most important data, PCA, a dimensionality reduction approach, may be applied. In contrast, CNNs are a popular deep neural network for image categorization. This subsection will focus on integrating a CNN into a Python PCA. This Python code implements the principal component analysis approach for a convolutional neural network with 10 hidden layers:

- Load the dataset and preprocess the data, if necessary.
- Perform PCA on the dataset to reduce the number of features. The number of principal components to retain can be specified using a parameter such as the number of components (n_components) or the explained variance threshold (explained_variance_ratio_).

- Split the dataset into training and testing sets.
- CNN architecture with 10 hidden layers. Hidden layers may have numerous convolutional and pooling layers followed by a dense layer with a given number of neurons and an activation function.
- Compile the CNN with a loss function, an optimizer, and a metric for evaluation, such as accuracy.
- Train the CNN on the training data using a specified number of epochs and batch size.
- Evaluate the trained CNN on the testing data using the specified evaluation metric.
- Optionally, use techniques such as regularization, dropout, or batch normalization to improve the performance of the CNN.

The mathematical equation for the PCA with 10 hidden layers CNN 1D model can be represented as:

$$y = W_2 h_{10} + b_2 \tag{1}$$

$$h_1 = \sigma(w_1 x + b_1) \tag{2}$$

$$h_2 = \sigma(w_2 h_1 + b_2) \tag{3}$$

$$h_{10} = \sigma(w_{10} h_9 + b_{10}) \tag{4}$$

Bring the data to a normal distribution: The data matrix should be normalized so each characteristic has a mean of zero and a variance of one. Perform the computations necessary to create the covariance matrix: A covariance matrix may be calculated after the data has been normalized. Check the eigenvalues and eigenvectors: Find the covariance matrix's eigenvectors and eigenvalues. Data variance along each principal component is represented by its associated eigenvalue, whereas the eigenvectors comprise the principal components. Ordering the eigenvectors: Neutralize the matrix by putting the eigenvectors in descending order of the eigenvalues. The top k eigenvectors should be chosen. Matrix W is constructed from the top k eigenvectors in terms of eigenvalues.

Data should be projected onto the new subspace. To generate a matrix Y representing the data projected onto the new subspace, we multiply the original data matrix by the matrix W. To learn the 1D CNN's 10 hidden layers, create a 1D CNN with 10 hidden layers and train it on the new Y matrix by using the equations from 1 to 4. Analyze the model: Use the trained CNN 1D to classify input from a different dataset after projecting it into the same subspace as the original dataset using the matrix W.

3.3.1. Goals of PCA and FA

To simplify an otherwise overwhelming number of variables (say, p) down to a manageable number (say, m) of components (or factors) for further analysis, a PCA or FA may be performed. We could utilize FA to check test results to determine whether the objects fall under these categories. The researchers assigned each individual a score of 100 based on how well they performed across all five criteria. The data from each participant was factored down from 19 variables to 5 components. A stepwise multiple regression was conducted with these five parameters as predictors (of the single criteria).

FA may reveal and summarize intercorrelations. Exploratory FA arranges variables closely associated with factors, presumably because the same factor impacts them. One may alternatively operationalize (measure) the underlying dimension by a linear combination of the factors that contributed most to the component. Confirmatory factor analysis tests a hypothesis about an event's fundamental aspects. We could utilize FA to check test results to determine whether the objects fall under these categories.

4. RESULTS AND DISCUSSION

The final output is the concatenated one, containing the label indicating whether or not professionals in the working world are stressed. In the process of developing tests, psychometricians often make use of FA. You could want to create a test to determine whether someone would be successful as a teacher, for instance. qTo quantify these factors, you might create several indicators in writing. You give the examination to many individuals and then evaluate the outcomes. I hope that a significant number of the items will cluster into factors that reflect the dimensions you planned to assess. The following is a summary of the most important advantages that come as a direct consequence of carrying out this work.

Accuracy, precision, recall, F1-score, error rate, and area under the curve (AUC) of receiver operating characteristic (ROC) are evaluated as per Figure 6. Evaluation metrics are specified in Table 1. Table 1 shows seven metrics for artificial neural network (ANN), logistic regression (LR), Naïve Bayes

(NB), support vector machine (SVM), and decision tree (DT) proposed work CNN techniques. In this table, ANN, LR, NB, SVM, DT, and proposed work have 78%, 77%, 81%, 88%, 90%, and 96.78% accuracy.

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\_Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Error\_Rate = 1 - Accuracy$$

Figure 6. Mathematical equations for evaluation metrics

Table 1. Evaluation metrics for various models

Models	Accuracy	Precision	Recall	F1-score	RMSE
ANN	78	84	94	86	58
LR.	77	83	94	87	62
NB	81	87	93	86.34	61
SVM	88	93	95	84	59.38
DT	90	93	95	96.48	64
Proposed work	96.78	97.28	97	97.36	31.69

From the table above, the proposed work got 97.28% highest in precision. Likewise in recall and F1-score, the proposed work got 97% and 97.36% the highest value among the existing models. The proposed work also got the highest accuracy, 96.78%, compared with existing models. Accuracy, precision, recall, F1-score, and error rate are developed using the confusion matrix. All metrics, accuracy, precision, recall, F1-score, and RMSE got the highest values for the proposed work. Figure 7 shows model accuracy from these values, while Figure 8 compares the six techniques' precision, recall, and F1-score. Figure 7 shows that the proposed work produced the highest values of the six Models. Figure 9 shows the RMSE graph.

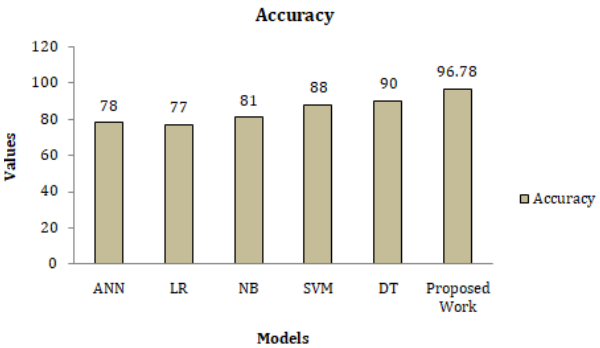

Figure 7. Graph for accuracy

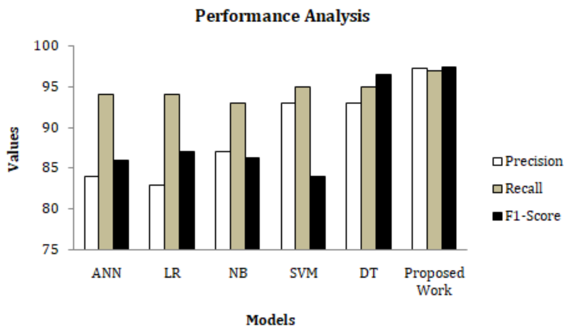

Figure 8. Graph for various evaluation metrics

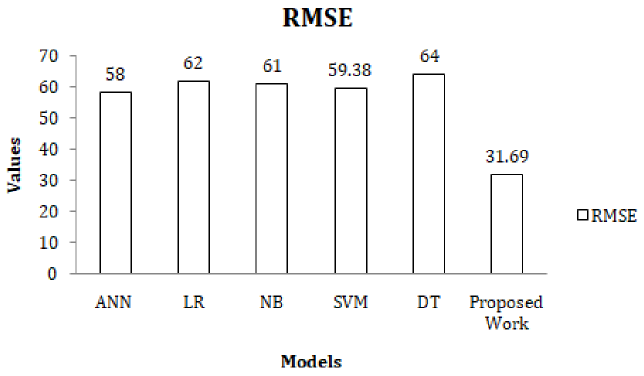

Figure 9. Graph for RMSE

On Stress Detection so many authors worked on various categories by using ML models. From that driver distraction [4], hazard anticipation using real-time in cardiology patients [6], working employees in various organizations [12], performance of the employer after feeling the stress [14], Students assessments on mood disorders [10], there is chance to get diabetes when any person feels stress and they used to get some correlated diabetes ailments [16], predicted the stress on COVID-19 patients [13], after doing the more work without taking the rest, then they feel stress and they may go for depression [30], we are regularly observing the news and reports on social media [31]. Observed that, till now various ML models were used on various datasets but the PCA was not applied to the features. Finally, the proposed system was compared with all existing models and observed that the features in dataset was applied an PCA and classified the dataset with proposed CNN model.

5. CONCLUSION

In this research, we use deep-learning-based classification to predict working professionals' stress levels more accurately. Determine which qualities are best for classification to improve predictions. This article covers preprocessing, feature extraction, selection, and classification. Before analysis, the dataset is normalized by removing unnecessary data and filling in missing values. To enhance categorization, a collection of features is extracted and optimized. The work showed how stress and its stages can be forecast using practical algorithms. The suggested algorithms, ANN, LR, NB, SVM, DT, and proposed work CNN, were examined. CNN had greater accuracy, precision, recall, F1-score, and error rate, with 96.78%, 97.28%, 97%, 97.36%, and 31.69%. Results from CNN can predict stress among people.

REFERENCES

[1] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on Psychological Stress Detection Using Biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2022, doi: 10.1109/TAFFC.2019.2927337.

[2] S. A. Useche, V. G. Ortiz, and B. E. Cendales, "Stress-related psychosocial factors at work, fatigue, and risky driving behavior in bus rapid transport (BRT) drivers," *Accident Analysis and Prevention*, vol. 104, pp. 106–114, 2017, doi: 10.1016/j.aap.2017.04.023.

[3] S. Greene, H. Thapliyal, and A. Caban-Holt, "A survey of affective computing for stress detection: evaluating technologies in stress detection for better health.," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 44–56, 2016, doi: 10.1109/MCE.2016.2590178.

[4] R. Ghandour, B. Neji, A. M. El-Rifaie, and Z. Al Barakeh, "Driver distraction and stress detection systems: a review," *International Journal of Engineering and Applied Sciences (IJEAS)*, vol. 7, no. 4, 2020, doi: 10.31873/ijeas.7.04.10.

[5] S. A. Khowaja, A. G. Prabono, F. Setiawan, B. N. Yahya, and S. L. Lee, "Toward soft real-time stress detection using wrist-worn devices for human workspaces," *Soft Computing*, vol. 25, no. 4, pp. 2793–2820, 2021, doi: 10.1007/s00500-020-05338-0.

[6] L. Kerautret, S. Dabic, and J. Navarro, "Detecting driver stress and hazard anticipation using real-time cardiac measurement: A simulator study," *Brain and Behavior*, vol. 12, no. 2, 2022, doi: 10.1002/brb3.2424.

[7] A. H. Celdrán, J. A. Ruipérez-Valiente, F. J. G. Clemente, M. J. Rodríguez-Triana, S. K. Shankar, and G. M. Pérez, "A scalable architecture for the dynamic deployment of multimodal learning analytics applications in smart classrooms," *Sensors (Switzerland)*, vol. 20, no. 10, 2020, doi: 10.3390/s20102923.

[8] U. Pluntke, S. Gerke, A. Sridhar, J. Weiss, and B. Michel, "Evaluation and classification of physical and psychological stress in firefighters using heart rate variability," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2207–2212, 2019, doi: 10.1109/EMBC.2019.8856596.

[9] A. D. S. Sierra, C. S. Avila, J. G. Casanova, and G. B. Del Pozo, "A stress-detection system based on physiological signals and fuzzy logic," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4857–4865, 2011.

[10] U. W. Ebner-Priemer and T. J. Trull, "Ecological momentary assessment of mood disorders and mood dysregulation," *Psychological Assessment*, vol. 21, no. 4, pp. 463–475, 2009, doi: 10.1037/a0017075.

[11] K. Hovsepian, M. Al'absi, E. Ertin, T. Kamarck, M. Nakajima, and S. Kumar, "Stress: Towards a gold standard for continuous stress assessment in the mobile environment," *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 493–504, 2015, doi: 10.1145/2750858.2807526.
[12] U. S. Reddy, A. V. Thota, and A. Dharun, "Machine learning techniques for stress prediction in working employees," *2018 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2018*, 2018, doi: 10.1109/ICCIC.2018.8782395.
[13] L. Flesia*et al.*, "Predicting perceived stress related to the covid-19 outbreak through stable psychological traits and machine learning models," *Journal of Clinical Medicine*, vol. 9, no. 10, pp. 1–17, 2020, doi: 10.3390/jcm9103350.
[14] V. Mathangi, "Impact of job stress on employees' job performance in Aavin, Coinbatore," *Journal of Organisation& Human Behaviour*, vol. 6, no. 3, pp. 21–29, 2017.
[15] P. Romaniszyn-Kania *et al.*, "Affective state during physiotherapy and its analysis using machine learning methods," *Sensors*, vol. 21, no. 14, 2021, doi: 10.3390/s21144853.
[16] S. S. Reddy, N. Sethi, and R. Rajender, "Mining of multiple ailments correlated to diabetes mellitus," *Evolutionary Intelligence*, vol. 14, no. 2, pp. 733–740, 2021, doi: 10.1007/s12065-020-00432-6.
[17] S. S. Reddy, N. Sethi, and R. Rajender, "A review of data mining schemes for prediction of diabetes mellitus and correlated ailments," in *2019 5th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Sep. 2019, pp. 1–5, doi: 10.1109/ICCUBEA47591.2019.9128880.
[18] S. S. Reddy, N. Sethi, and R. Rajender, "Diabetes correlated renal fault prediction through deep learning," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 6, no. 24, pp. 1–17, 2020, doi: 10.4108/eai.11-11-2020.166958.
[19] S. S. Reddy, N. Sethi, and R. Rajender, "Evaluation of deep belief network to predict hospital readmission of diabetic patients," *Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA 2020*, pp. 5–9, 2020, doi: 10.1109/ICIRCA48905.2020.9182800.
[20] N. Jaques, S. Taylor, A. Sano, and R. Picard, "Multimodal autoencoder: A deep learning approach to filling in missing sensor data and enabling better mood prediction," *2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, vol. 2018-January, pp. 202–208, 2017, doi: 10.1109/ACII.2017.8273601.
[21] R. L. Rosa, G. M. Schwartz, W. V. Ruggiero, and D. Z. Rodriguez, "A knowledge-based recommendation system that includes sentiment analysis and deep learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2124–2135, 2019, doi: 10.1109/TII.2018.2867174.
[22] S. Ghosal, D. Blystone, A. K. Singh, B. G. Subramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," *Proceedings of the National Academy of Sciences*, vol. 115, no. 18, pp. 4613–4618, 2018.
[23] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges," *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 445–458, 2019, doi: 10.1109/TNSM.2019.2899085.
[24] D. Jaber, H. Hajj, F. Maalouf, and W. El-Hajj, "Medically-oriented design for explainable AI for stress prediction from physiological measurements," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, 2022, doi: 10.1186/s12911-022-01772-2.
[25] L. Liang, M. Liu, C. Martin, and W. Sun, "A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis," *Journal of the Royal Society Interface*, vol. 15, no. 138, 2018, doi: 10.1098/rsif.2017.0844.
[26] N. Jaques, S. Taylor, A. Sano, and R. Picard, "Predicting tomorrow's mood, health, and stress level using personalized multitask learning and domain adaptation ognjen (Oggi) rudovic," *Journal of Machine Learning Research*, vol. 66, pp. 17–33, 2017.
[27] V. Sumathi, R. Velmurugan, J. Sudarvel, and P. Sathiyabama, "Intelligent classification of women working in ICT based education," *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, pp. 1711–1715, 2021, doi: 10.1109/ICACCS51430.2021.9441803.
[28] A. Liapis, E. Faliagka, C. P. Antonopoulos, G. Keramidas, and N. Voros, "Advancing stress detection methodology with deep learning techniques targeting ux evaluation in aal scenarios: Applying embeddings for categorical variables," *Electronics (Switzerland)*, vol. 10, no. 13, 2021, doi: 10.3390/electronics10131550.
[29] L. Rachakonda, A. K. Bapatla, S. P. Mohanty, and E. Kougianos, "SaYoPillow: blockchain-integrated privacy-assured IoMT framework for stress management considering sleeping habits," *IEEE Transactions on Consumer Electronics*, vol. 67, no. 1, pp. 20–29, 2021, doi: 10.1109/TCE.2020.3043683.
[30] T. Richter, B. Fishbain, A. Markus, G. Richter-Levin, and H. Okon-Singer, "Using machine learning-based analysis for behavioral differentiation between anxiety and depression," *Scientific Reports*, vol. 10, no. 1, 2020, doi: 10.1038/s41598-020-72289-9.
[31] H. Lin *et al.*, "Detecting stress based on social interactions in social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1820–1833, 2017, doi: 10.1109/TKDE.2017.2686382.

BIOGRAPHIES OF AUTHORS

Samarendra Narayan Pradhan is a Data scientist, having more than 15+ years. of experience in wireless network planning, optimization, Maximizing ROI: Budgeting and CAPEX Optimization in Wireless Networks, Customer Experience, Sales and Marketing Analytics in the Telecom Industry, Leveraging Deep Learning for Customer Churn Prediction., Natural Language Processing Models for Customer Feedback Analysis. He is a MTech postgraduate in Computer Science and Engineering from DRIEMS (Dhaneswar Rath Institute of Engineering and Management Studies). He can be contacted at email: luckysamar@yahoo.com.

Reddy Shiva Shankar is an Assistant Professor at the Department of Computer Science and Engineering in Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhrapradesh, INDIA. He is pursuing PhD degree in Computer Science and Engineering with a specialization in Medical Mining and Machine Learning. His research areas are Image Processing, Medical Mining, Machine Learning, Deep Learning and Pattern Recognition. He published 30+ papers in International Journals and Conferences. S.S. Reddy has filed 05 patents. His research interests include Image Processing, Medical Mining, Machine Learning, Deep Learning and Pattern Recognition. He can be contacted at email: shiva.csesrkr@gmail.com.

Shekharesh Barik is an accomplished author who has several books under their belt. With a wealth of experience, they have 13 published papers in both national and international journals and conferences and have been teaching for a staggering 19 years. Their research interests are diverse and include Machine Learning, IoT, and Computer Graphics, and they are highly passionate about sharing their knowledge and experience to help others grow and develop in their respective fields. He can be contacted at email: shekharesh@gmail.com.

Bhabodeepika Mohanty is a part-time guest faculty with Bhubanananda Orissa School of Engineering in Cuttack. Before that, she was aMTech postgraduate in Computer Science and Engineering from DRIEMS (Dhaneswar Rath Institute of Engineering and Management Studies). She can be contacted with email: mohantybhabodeepika@gmail.com.

Venkata Rama Maheswara Rao is a leading Researcher & Academician in Computer Science & Engineering and holds Ph.D. degree. He is currently working as a Professor in the Dept. of Computer Science & Engineering at Shri Vishnu Engineering College for Women (A), Andhra Pradesh, India. He is actively involved and successfully implemented three projects funded by DST. He has 45 research papers, 17 of which are Scopus-indexed and 7 of which are Web of Science-indexed. He has 23 years of experience that include 6 years of Industry experience, 19 years of Teaching experience and 15 years of Research experience. His Research interests include Data Mining, Web Mining, Cloud Computing, Big Data Analytics, Data Science, Artificial Intelligence and Machine Learning. He can be contacted at email: mahesh_vvr@yahoo.com.