JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# ADVANCED STRESS DETECTION WITH DEEP LEARNING TECHNIQUES USING UX EVALUATION

**Dr. Manjula S[1], Nidhishree S[2]**

*Assistant Professor, Department of Computer Science & Engineering, JSS STU, Mysore, India*

*Department of Computer Science & Engineering, JSS STU, Mysore, India*

***Abstract:*** **Physiological measurements have been widely used by researchers and practitioners in order to address the stress detection challenge. So far, various datasets for stress detection have been recorded and are available to the research community for testing and benchmarking. The majority of the stress-related available datasets have been recorded while users were exposed to intense stressors, such as songs, movie clips, major hardware/software failures, image datasets, and gaming scenarios. However, it remains an open research question if such datasets can be used for creating models that will effectively detect stress in different contexts. This paper investigates the performance of the publicly available physiological dataset named WESAD (wearable stress and affect detection) in the context of user experience (UX) evaluation. More specifically, electrodermal activity (EDA) and skin temperature (ST) signals from WESAD were used in order to train three traditional machine learning classifiers and a simple feed forward deep learning artificial neural network combining continues variables and entity embeddings. Regarding the binary classification problem (stress vs. no stress), high accuracy (up to 97.4%), for both training approaches (deep-learning, machine learning), 93.28% was reached by decision tree and 83.85% by logistic regression model was achieved. Regarding the stress detection effectiveness of the created models in another context, such as user experience (UX) evaluation, the results were quite impressive. More specifically, the deep-learning model achieved a rather high agreement when a user-annotated dataset was used for validation.**

***Keyword: stress detection; UX evaluation; electro-dermal activity; deep learning; entity embeddings; machine learning; logistic regression.***

## I. INTRODUCTION

The Internet of Things, human computer interaction, modern computers, artificial intelligence, andother fields have all contributed to a significant increase in interest in the emotional components of packaged goods. User experience (UX) entitles teachers to realise interpretation of the users' interplay events by using equipment and proposals that transcend traditional usability metrics while taking into account physical world living conditions and commercial-off-the-shelf (COTS) instrument panels and sensors. Usability, usefulness, aesthetics, and emotions are just a few of themany components that make up UX. In many circumstances, UX design starts even before the consumer has the product in their hands, effectively anticipating their needs and wants. An in- depth understanding of people feelings while engaging with a system is required when planning and creating for UX. We'll use a range of techniques, including post-questionnaires, interviews, and observation, to experience the emotional components of UX. Otherwise, projections are made for modalities likecountenance, touchscreen patterns and speech tone analysis that may not be inheritable by embedded and robotic instrument panels and sensors.

## II. LITERATURE SURVEY

A critical stage in confirming the value of research contributions is the examination of analysis artefacts. In addition to usability, HCI subfields frequently target fundamental objectives like property, Sustainable HCI (SHCI), HCI for development, or health and safety. It is necessary to build new standards for identifying, debating, and supporting relevant analysis methods for these disciplines because conventional analytical methods are not always decent or applicable. In this study, we prefer to reiterate the purpose and objectives of analysis in HCI and SHCI and elicit 5 essential elements that will provide guidance to various analysis approaches for SHCI research. Our essay is intended to serve as a springboard for discussion of current and emerging SHCI analysis practices. [1] The most significant influences on the quality of camera-based systems for recognising emotionsare variations in head position and light conditions. The methods that provide 2-Dimentional image analysis are particularly sensitive to these issues. Techniques that enforce 3-Dimentional face models are far more promising. Because of its low cost and ease of use, we frequently employ Microsoft Kinect for 3D face modelling in our investigations.[2] Examples include work environments and user experience assessments. In the past, flow was evaluated through questionnaires, preventing its usage in online, timed contexts. In this study, we tend to outline ourconclusions regarding assessing a user's flow state backed by physiological data captured by wearable technology. We frequently carry out research with people who are playing the Tetris game at various levels of difficulty, which causes boredom, stress, and flow. [5] Users' physiological data can be collected utilising sensors to provide important insights that are not possible with just traditional measurements. An indication of both physiological and psychological arousal is electrodermal activity (EDA). There are several uses for measuring arousal. For instance, persistently high arousal levels that occur frequently can be a sign of chronic stress. At the opposite extreme, for instance, consistently low arousal levels in geriatric care can indicate that the patients are not receiving enough movement and attention from the caregivers. Measurement of arousal can reveal when people become enthusiastic and when they are more tranquil in the context of events. In this study, an EDA measurement pilot study that was carried out at a trade show is presented. [8] Despite the recent explosion in popularity of computer games, methods for assessing players' emotional states as they play are far behind. There are few techniques for determining one's emotional state, and even fewer for measuring emotion while playing. The method for continuallymodelling emotion using physiological data is presented in this research. Four physiological inputswere translated into arousal and valence using a fuzzy logic model. Arousal and valence were turned into the five emotional states of boredom, challenge, excitement, aggravation, and joy in asecond fuzzy logic model. The means were also assessed with subjective self-reports, showing thesame tendencies as reported feelings for joy, boredom, and excitement. Modeled emotionsperformed well compared to a manual technique. [10]

# III. PROPOSED WORK

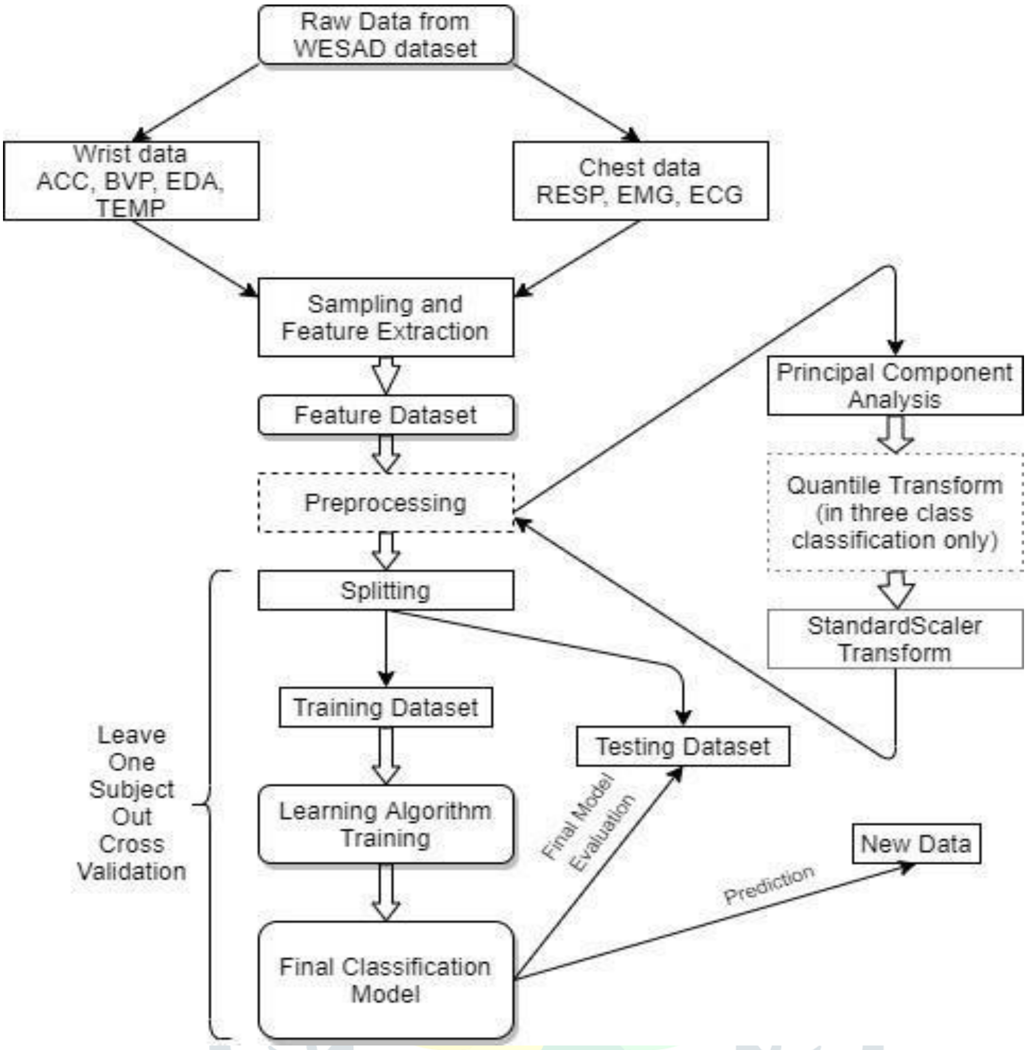
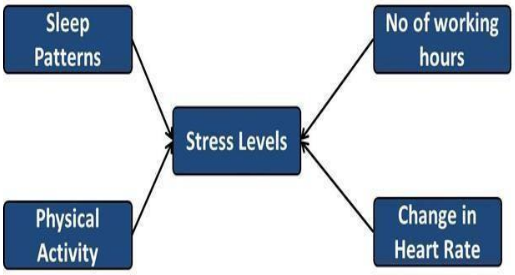

Raw Data from WESAD dataset
Wrist data ACC, BVP, EDA, TEMP
Chest data RESP, EMG, ECG
Sampling and Feature Extraction
Feature Dataset
Preprocessing
Principal Component Analysis
Quantile Transform (in three class classification only)
StandardScaler Transform
Splitting
Training Dataset
Testing Dataset
Learning Algorithm Training
Final Classification Model
Leave One Subject Out Cross Validation
Final Model Evaluation
Prediction
New Data

*Fig 1. Schematic flow representation of Stress Detection Methodology*



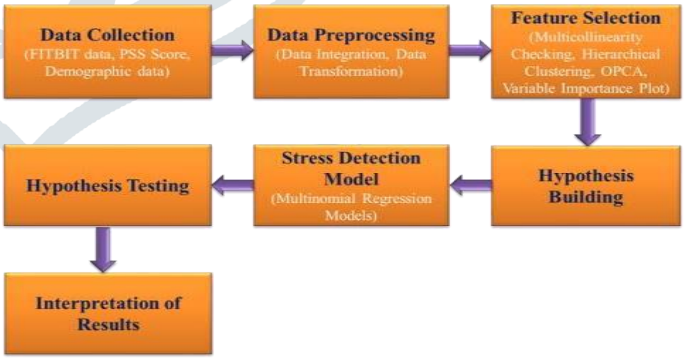

Data Collection (FITBIT data, PSS Score, Demographic data)
Data Preprocessing (Data Integration, Data Transformation)
Feature Selection (Multicollinearity Checking, Hierarchical Clustering, OPCA, Variable Importance Plot)
Hypothesis Building
Stress Detection Model (Multinomial Regression Models)
Hypothesis Testing
Interpretation of Results

***Figure 2.*** *Conceptual Model of STRESS detection.* ***Figure 3.*** *The block diagram*

The conceptual flow diagram of the stress detection methodology is shown in Figure 5.1.

The abstract model of our study on stress is depicted in Figure 5.2. the significance of sleep, exercise, a variety of working hours, and changes in pulse in relation to stress levels. The stress detection approach is divided into several steps, such as knowledge gathering, knowledge pre-processing, feature selection, hypothesis formulation, stress detection model,hypothesis testing, and result interpretation (see Figure 5.3).

## Extraction of the Dataset and Features:

The dataset used for this investigation is WESAD. Attila Reiss, Philip Schmidt, et al. first presented and made this dataset available to the public in 2018 [1]. This multimodal dataset assembles mobility information and physiological characteristics from 15 people using the wrist-and chest-worn sensors Empatica E4 and RespiBAN Professional, respectively. The physiologicalsignals of the subjects were captured while they through various study protocols, including preparation, baseline condition, amusement condition, stress condition, meditation, and recuperation. Reference [1] provides specifics on sensor placement, setup, and the method used tocreate this dataset, including which data are gathered during which subject-specific study protocol.The ACC, RESP, ECG, EDA, EMG, and TEMP were measured using the RespiBAN. At 700 Hz,all signals were captured. The TEMP, EDA, ACC, and BVP were all measured by the E4 using samples at 4 Hz, 4 Hz, 32 Hz, and 64 Hz, respectively. Each subject in the dataset has a folder (SX, where X is the subject ID). The subsequent files are located in each subject folder: • The SXreadme.txt file contains information about the subject (SX), as well as details about the gatheringand grading of data (if applicable). • The SX quest.csv file contains all the data needed to gather ground truth, such as the SX procedure schedule and responses to the self-report questions. • SX respiban.txt: this file contains information from the RespiBAN device, including ECG, EDA, EMG, TEMP (°C), RESP, and other data. Data from the Empatica E4 device, such as ACC, BVP,EDA, and TEMP, are included in the file SX E4 Data.zip. Additional files in this subdirectory include: - ACC.csv: The three data columns represent the three channels of the accelerometer. - BVP.csv: Photoplethysmograph data (PPG). Data is provided in S in the EDA.csv file. - TEMP.csv: Data is given in degrees Celsius. • SX.pkl: contains labels and synchronised data. Allsensor signals were segmented using a sliding window with a 1 second shift. The features that were extracted using various modalities from the WESAD dataset are shown in Table I. These characteristics are a subset of the characteristics listed in [1]. On the raw ACC signal, other statistical aspects were calculated, such as the standard deviation, mean, minimum, and maximumvalue, as well as adding up for all axes (3D) as absolute magnitudes.

## Developing a Deep Learning Model Using Entity Embedding and Continual Variables:

A deep learning methodology for stress detection is projected in this segment. In an extremelyneural network model, the projected model incorporates the continuous and categorical load variables from the dataset. For the illustration of categorical variables, in which every securedutility of the variable is constituted as a numerical vector, often with a low measure, we tendto hold the entity embedding technique. The method stated earlier uses a surface of linear neurons to translate different values to a three-dimensional space. Because of this, the relationship between different values is frequently represented in the distance of the above vectors using a same methodology embedding, which takes into account semantic resemblance in the NLP province (e.g., stated as specific variable, Sunday may be reviewed almost like Saturday than it is to Monday). We include each category and continuous variable in our representation. It specifically contains 21 continuous variables that correspond to the choices removed from the WESAD dataset. The continuous variables consist the mean valueof the SC signals (following ironing as well as normalising them as intended), the median, the standard deviation, and other intended properties. The categorical factors include the user's gender, information on whether or not they smoke, whether they smoked in the hour prior tothe experiment, and whether or not they drank coffee in that hour. The batch normalisation layer for the continuous columns and embedding layers for the individual columns were bothincluded in the proposed neural network model for stress categorization. Following a dropoutlayer, the following representations are sequentially fed into two layers that are completely connected and have 200 and 100 nodes, respectively. A two-hidden network is capable of accurately approximating any smooth mapping and representing any arbitrary decisionboundary with rational activation functions. ReLU was employed as the activation function, as seen in figure below (Figure 7.1). As a result, the embedding layer transforms the categoryvariables before they interact with the continuous input variables. Finally, based on the cross- entropy loss function, the output layer examines the "stress" / "no stress" categories.

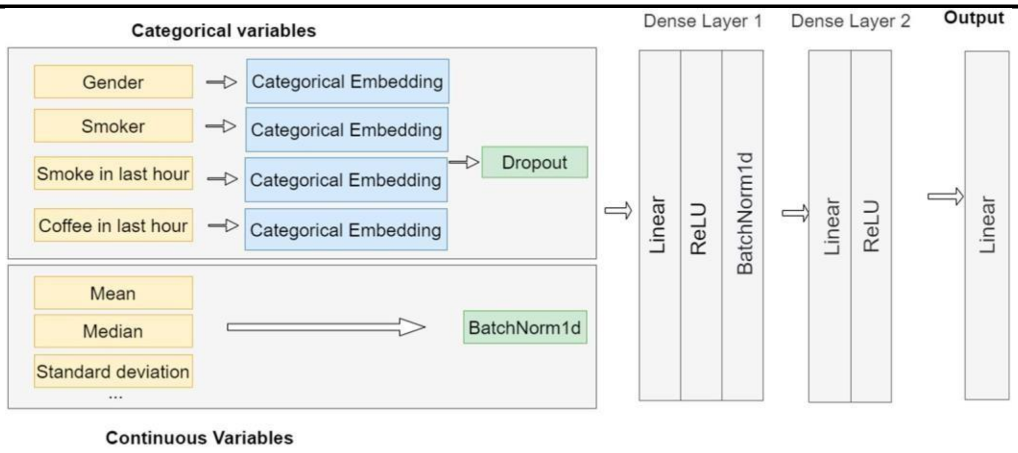


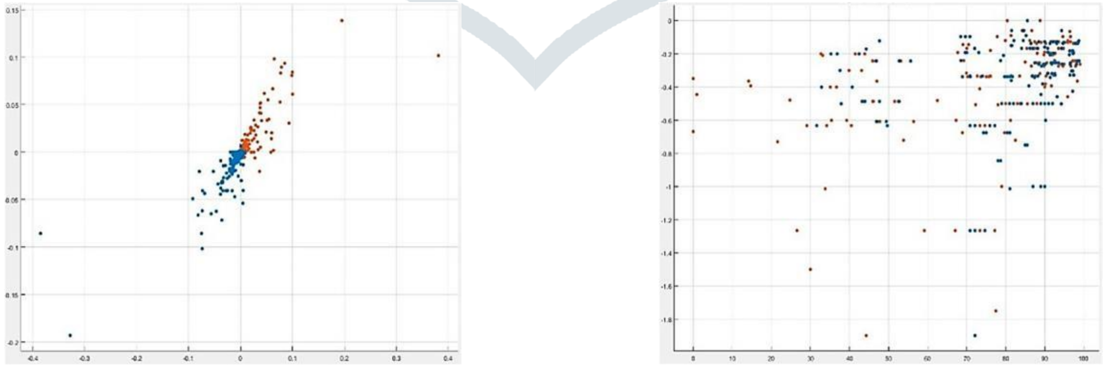
*__Figure 4.__ A visualization of the neural network (deep-learning model) architecture and training process. Continuous and categorical variables (left part) are compressed in order to feed 2 interconnected hidden layers(dense 1 and 2).*

## Training and Classification:

According to a proposal in [40], 21 features were taken from the amplitude of the SC signal for the numeric inputs. As suggested in [26], fifteen features were also derived from amplitude of theST signal. 380 segments were withdrawn from the NS-SCR segments inside each TSST session in order to calculate all of these attributes. The 380 segments of the first characteristic of the SCsignal's amplitude, or the mean value of the signal's first difference, are displayed in Figure 3a, and the corresponding values of the ST signal's amplitude are represented in Figure (b).

It shouldbe observed that 215 of them relate to class non-stress, while 165 of them do not. Figure 3a showsthat the stress and non-stress classes are well separated, which suggests a stronger predictive valuethan in Figure (b) where the feature values are not clearly separated. This is also shown in Table1, where the measurements for skin conductance signal perform better than the analogous metricsfor skin temperature.The retrieved features were then used as input for the deep learning model and the three machinelearning algorithms like C-SVM, L-SVM, and Q-SVM that were designed to distinguish betweenthe two emotional states (stress vs. non-stress).

A 5-fold cross-validation training was used in all classification methods for the machine learning classification. The acquired performance metric for each trained classifier is shown in Table 1 with regard to the binary problem (stress vs. non-stress). High accuracy was attained by all classifiers (at least 91 percent ). The L-SVM classifier produced the best classification outcomes(93.2 percent ). These results show that the classification results utilising our training strategy were better than the 80 percent accuracy reported by [26] when using simply the SC signal.



***(a)*** ***(b)***

***Figure 5.** An instance of skin conductance predictors and skin temperature. The orange color indicates the stress class and blue the non-stress class: (**a**) skin conductance predictors separate classeswell; (**b**) temperature predictors did not separate the classes well.*

| | | C-SVM | L-SVM | Q-SVM |
|---|---|---|---|---|
| Precision | SC | 89.7% | 92.6% | 92.4% |
| ST | | 37.1% | 25.0% | 33.3% |
| Recall | SC | 89,7% | 91.5% | 88.5% |
| ST | | 31.5% | 03.6% | 22.4% |
| Accuracy | SC | 91.1% | 93.2% | 91.8% |
| ST | | 47.1% | 53.4% | 46.8% |
| F1-Score | SC | 89.7% | 92.1% | 90.4% |
| ST | | 34.1% | 06.3% | 26.8% |

***Table 1.*** *Performance for each signal (skin conductance: SC, skin temperature: ST) per classifier. The F1-scoreis also an important metric when there are imbalanced classes as in our case.*

The area under this ROC curve is known as the area under the curve (AUC), and the plot of sensitivity versus 1-Specificity is known as the receiver operating characteristic (ROC) curve (seeFigure 5). Both the ROC and AUC are useful metrics for accuracy. This curve is crucial in assessing how well diagnostic testing can distinguish between people' real states. The AUC can bethought of as the likelihood that a randomly selected stress signal will be regarded or ranked as being more likely to be stress than a randomly selected non-stress signal. Each classifier attained a high AUC (at least 94 percent ). L-SVM classifier obtained the best AUC outcome (98 percent).

Predicted Classes / model

Skin Conductance

| True Class | | C-SVM | | L-SVM | | Q-SVM | |
|---|---|---|---|---|---|---|---|
| | | Stress | No Stress | Stress | No Stress | Stress | No Stress |
| | Stress | 148 | 17 | 151 | 14 | 146 | 19 |
| | No Stress | 17 | 198 | 12 | 203 | 12 | 203 |

Skin Temperature

| True Class | | C-SVM | | L-SVM | | Q-SVM | |
|---|---|---|---|---|---|---|---|
| | Stress | 52 | 113 | 6 | 159 | 37 | 128 |
| | No Stress | 88 | 127 | 18 | 197 | 74 | 141 |

***Figure 6.****. Confusion matrix for each signal per classifier. Figure shows the correctly classified (green rectangles) cases per class. Overall, the training dataset consisted of 380 cases; 165 in the class stress and 215 inthe class non stress. Green parts show the correctly classified cases for each classifier.*

To determine their effect on classification accuracy in deep learning, two versions of the model were examined, one in need of categorical variables and the other lacking. Based on its impact on loss, the learning rate hyper parameter's ideal value was chosen to improve model performance.

The hyper parameter learning rate specifically determines how much gradient will be back propagated. The amount by which we progress towards the minima is then determined by this. When the learning rate is placed too small, the evaluation process takes a too long and only littlemodifies model's weights, which causes the model to converge slowly and with no apparent gain.The optimizer may overshoot the minimum and potentially diverge if the learning rate is too large.

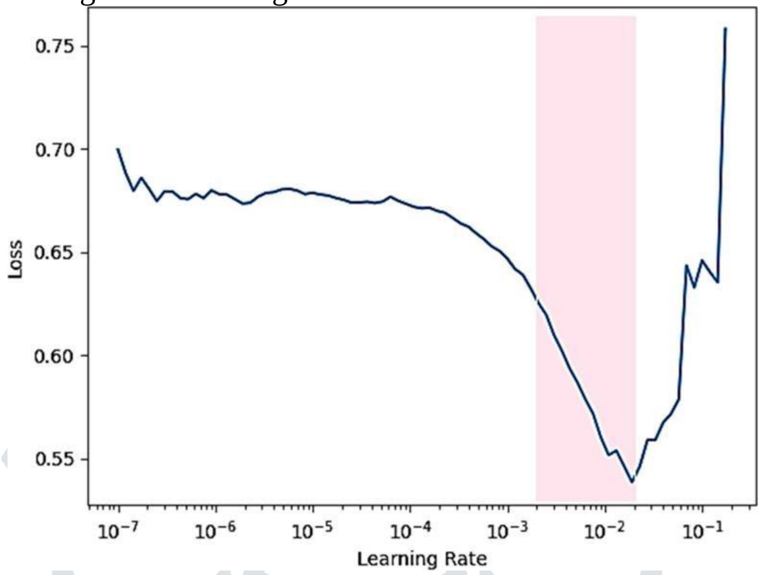
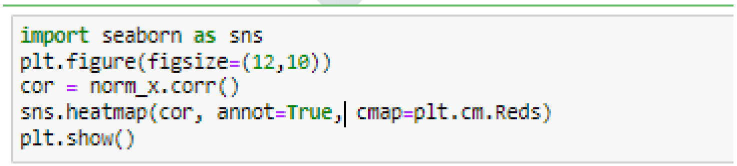


***Figure 7.*** *Learning rate chart. Pink rectangle indicates an area of optimal choices. In our case the ~7 ×$10^{-2}$ learning rate was used*

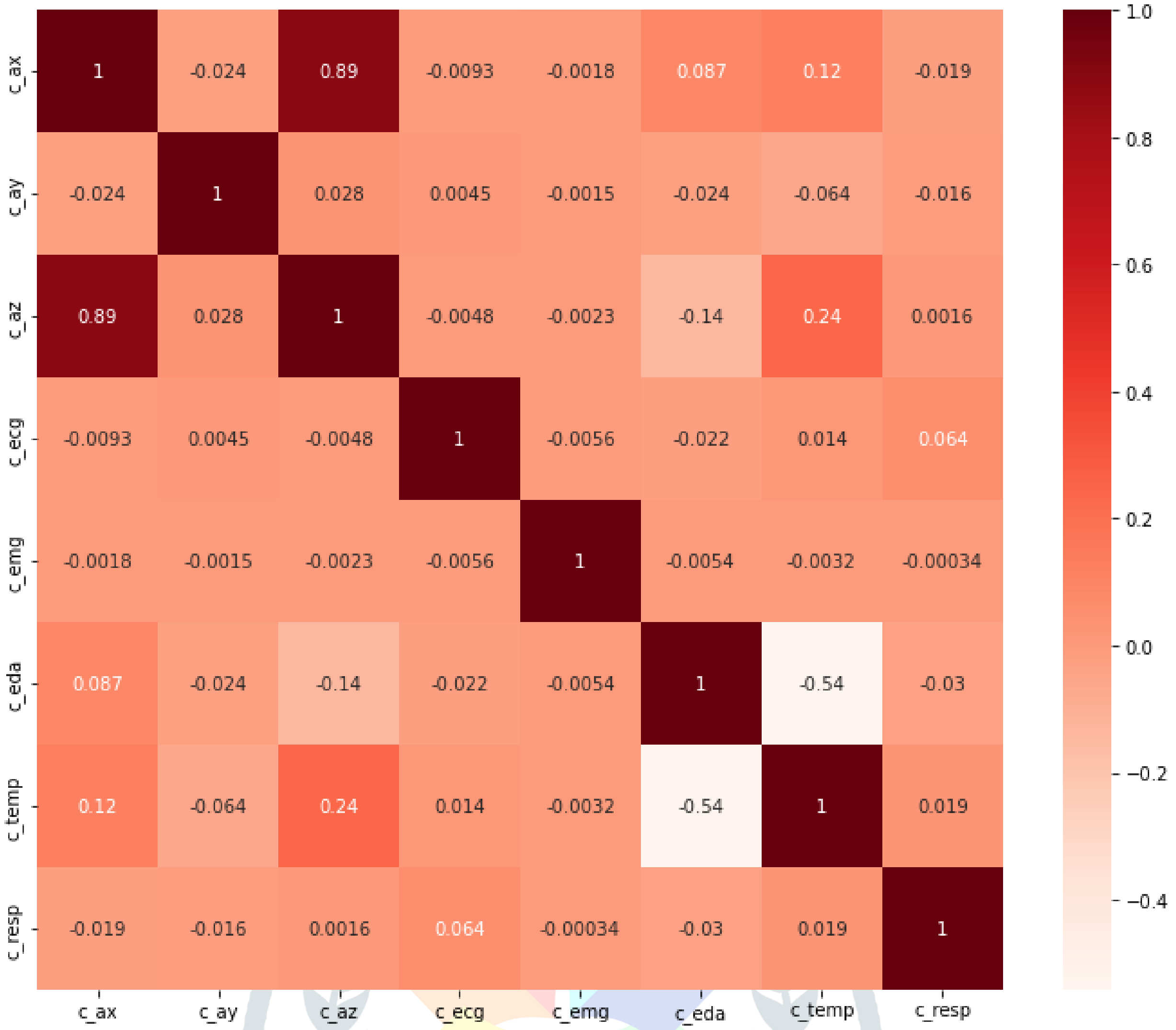
Overfitting is a crucial problem while training a neural network [47]. A neural network model needs to be taught across a number of epochs, however during training, patterns particular to thetest information are found. In other words, the model loses its generalizability when it is overfit to the training data. To avoid overfitting and increase the neural network's potential for generalisation, the model should be trained for the ideal number of epochs. The epoch number atwhich the model starts to overfit is determined by tracking the loss and accuracy on both the training and validation sets. Second table displays the gained performance metric for every trainedclassifier in relation to the binary problem (stress vs. non-stress).

# IV. RESULT

## Normalized Cross Correlation Coefficient:



```
import seaborn as sns
plt.figure(figsize=(12,10))
cor = norm_x.corr()
sns.heatmap(cor, annot=True, cmap=plt.cm.Reds)
plt.show()
```



| | c_ax | c_ay | c_az | c_ecg | c_emg | c_eda | c_temp | c_resp |
|---|---|---|---|---|---|---|---|---|
| c_ax | 1 | -0.024 | 0.89 | -0.0093 | -0.0018 | 0.087 | 0.12 | -0.019 |
| c_ay | -0.024 | 1 | 0.028 | 0.0045 | -0.0015 | -0.024 | -0.064 | -0.016 |
| c_az | 0.89 | 0.028 | 1 | -0.0048 | -0.0023 | -0.14 | 0.24 | 0.0016 |
| c_ecg | -0.0093 | 0.0045 | -0.0048 | 1 | -0.0056 | -0.022 | 0.014 | 0.064 |
| c_emg | -0.0018 | -0.0015 | -0.0023 | -0.0056 | 1 | -0.0054 | -0.0032 | -0.00034 |
| c_eda | 0.087 | -0.024 | -0.14 | -0.022 | -0.0054 | 1 | -0.54 | -0.03 |
| c_temp | 0.12 | -0.064 | 0.24 | 0.014 | -0.0032 | -0.54 | 1 | 0.019 |
| c_resp | -0.019 | -0.016 | 0.0016 | 0.064 | -0.00034 | -0.03 | 0.019 | 1 |

***Figure 8.*** *Normalized Cross Correlation Coefficient*

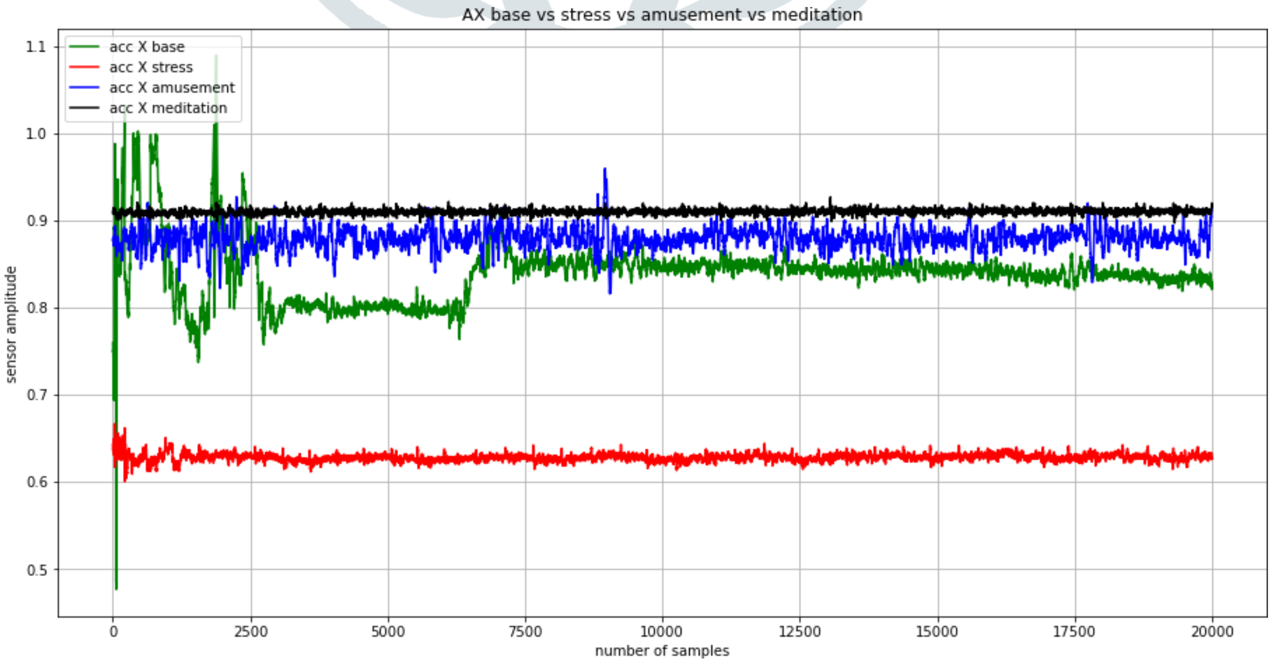


***Figure 9***. *Comparing AX base vs Stress vs Amusement vs Meditation*

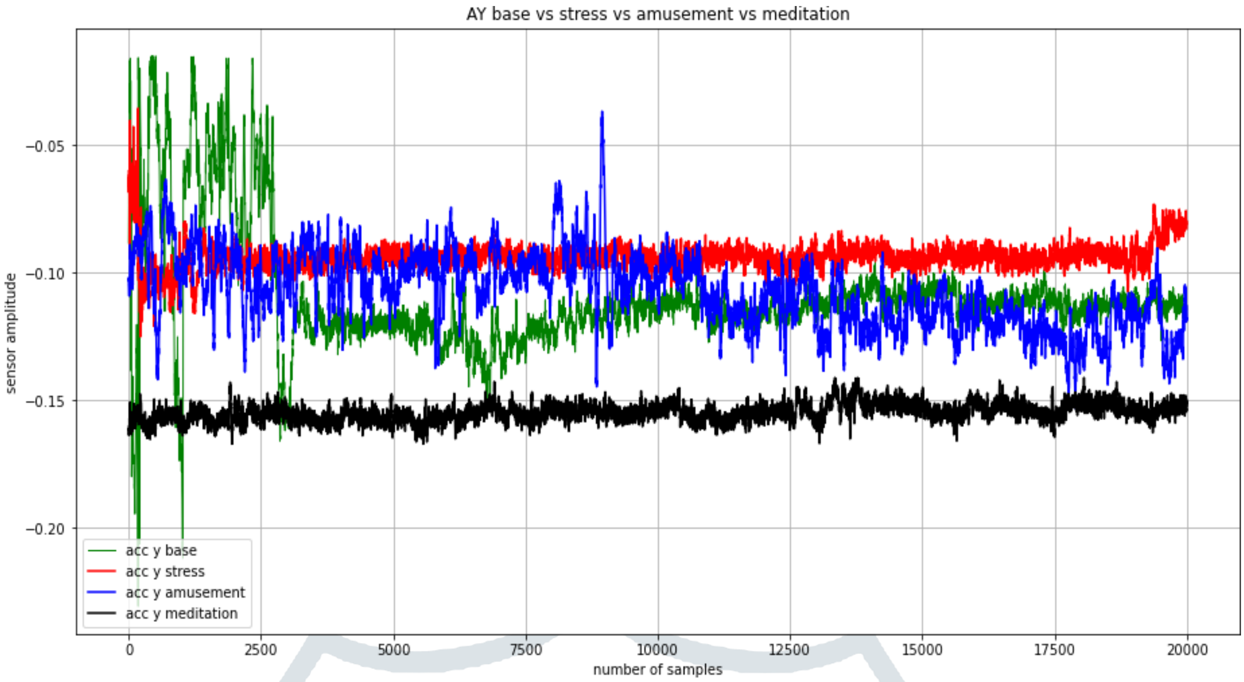


***Figure 10.*** *Comparing AY base vs Stress vs Amusement vs Meditation*

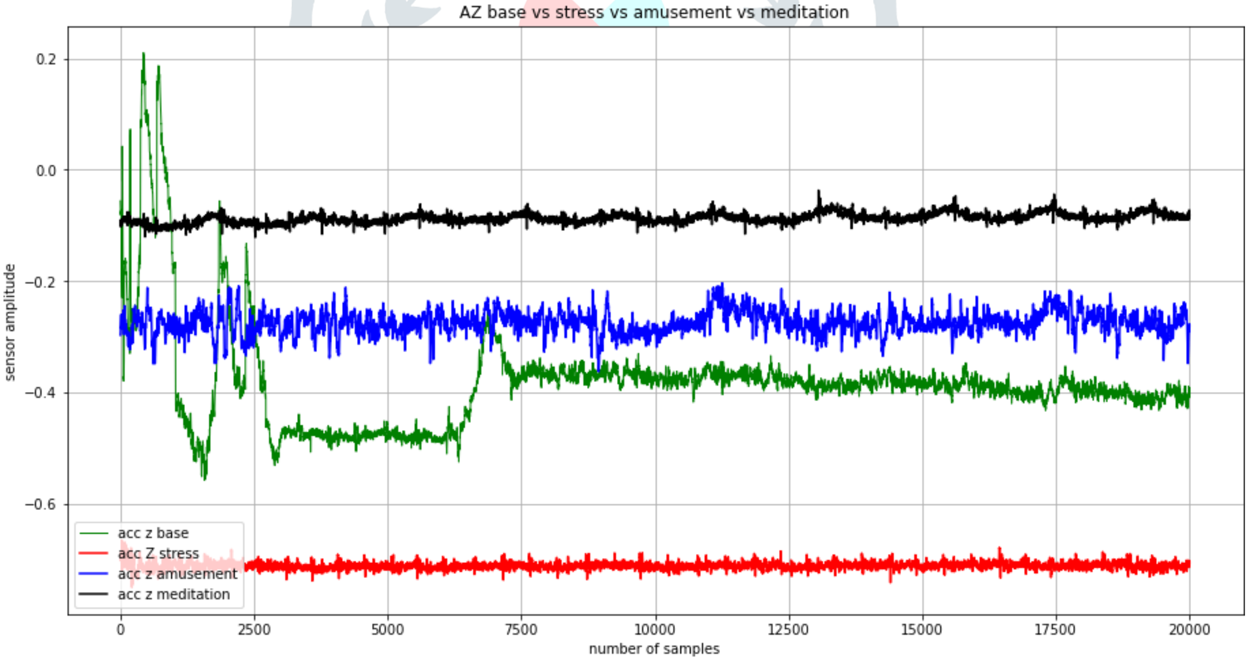


***Figure 11.*** *Comparing AZ base vs Stress vs Amusement vs Meditation*

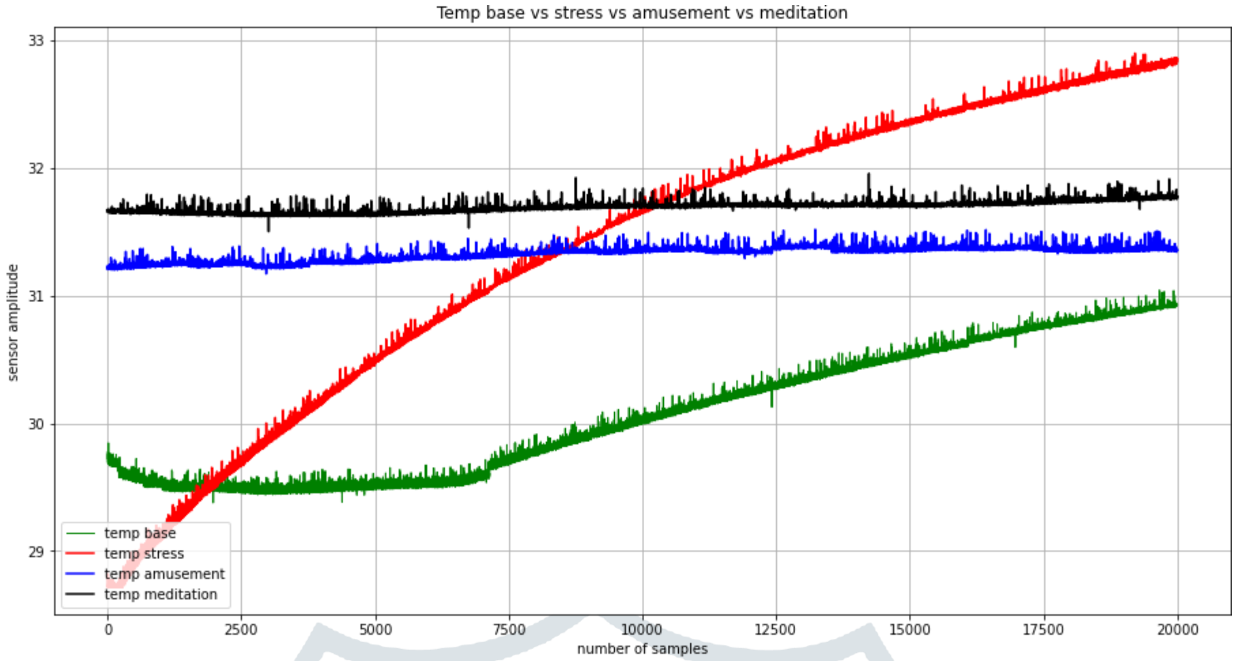


***Figure 12.*** *Comparing Temp base vs Stress vs Amusement vs Meditation*

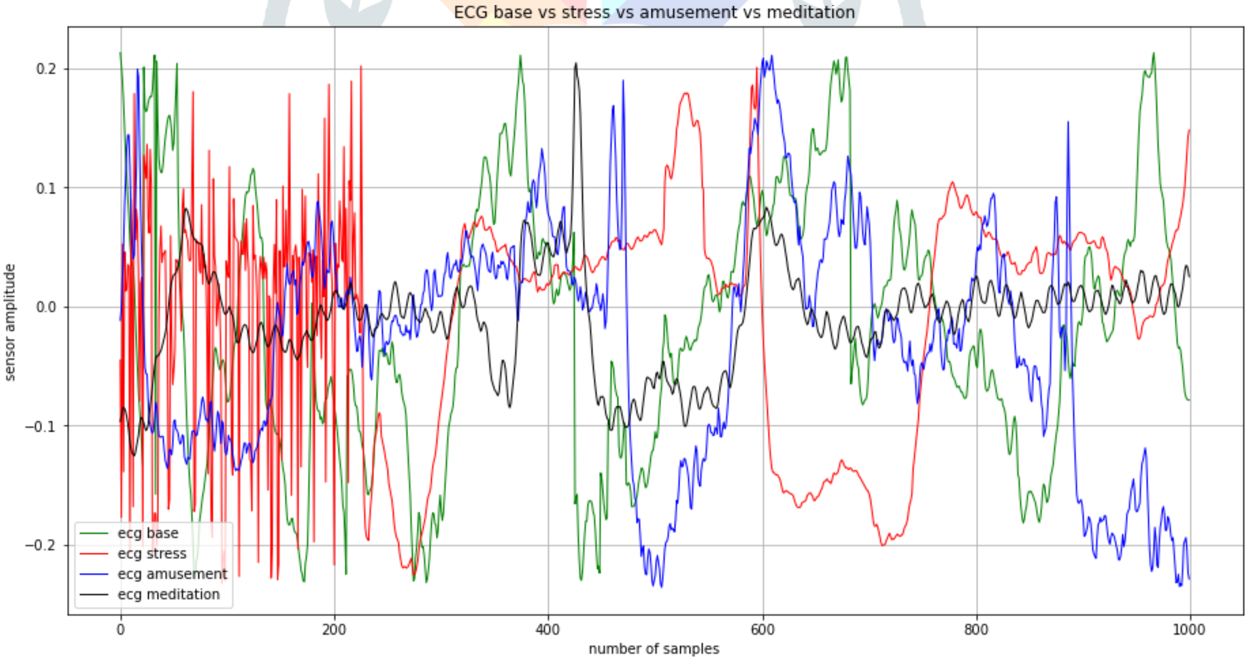


***Figure 13.*** *Comparing ECG base vs Stress vs Amusement vs Meditation*

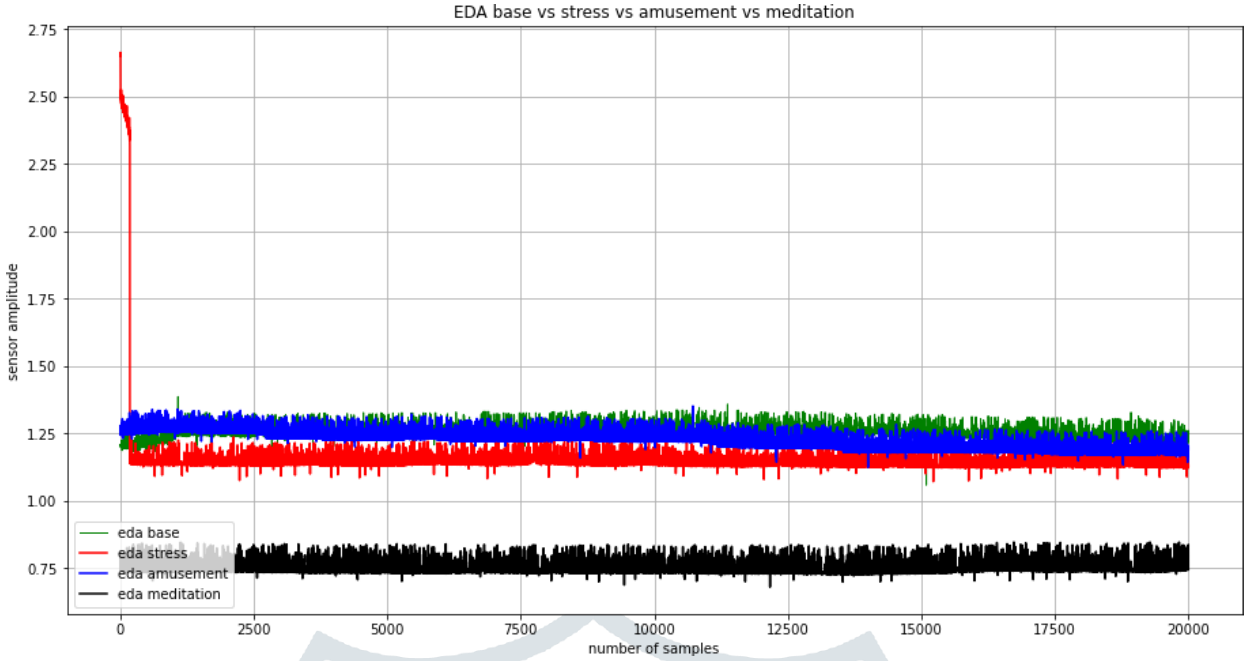


***Figure 14.*** *Comparing EDA base vs Stress vs Amusement vs Meditation*

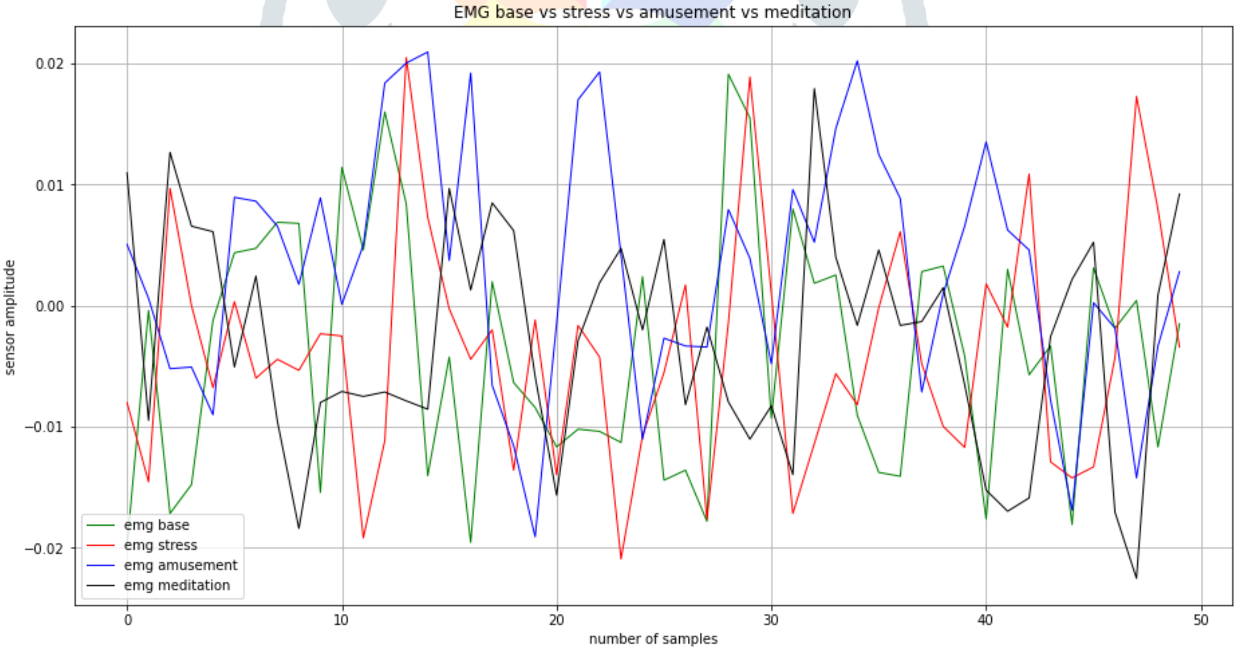


***Figure 15.*** *Comparing EMG base vs Stress vs Amusement vs Meditation*

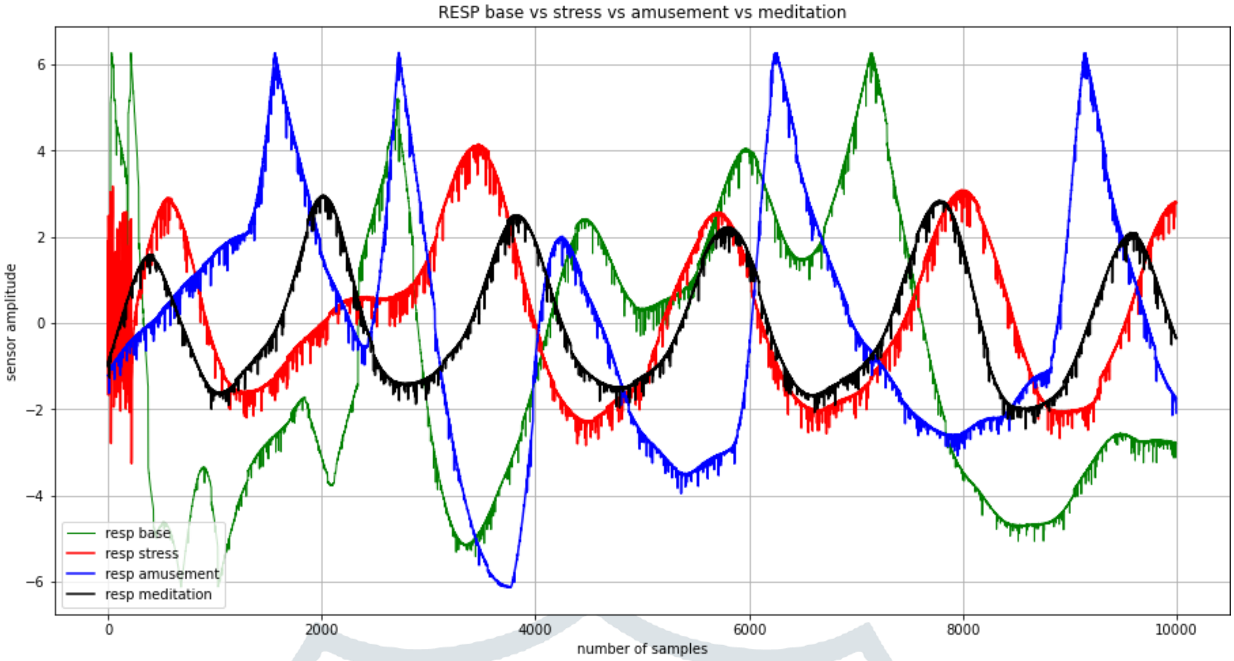


*__Figure 16.__ Comparing RESP base vs Stress vs Amusement vs Meditation*

## V. CONCLUSION AND FUTURE WORK

Numerous physiological datasets that have been collected during stress research are readily accessible to the general public. The majority of them were captured while subjects were subjected to strong stressors that are frequently present in real-life situations. Although these methods can produce stress prediction models with relatively good classification accuracy, it is still debatable whether they can be utilised to successfully capture subtle stress reactions, which are typically anticipated in various contexts, such as UX evaluation studies.

By combining classic machine learning and deep learning techniques, we conduct an extensive analysis of the performance of such a dataset in the context of UX evaluation in this experiment to try to answer the question raised above. To the best of our knowledge, thiswork is the first to employ deep learning to identify stress in a user experience environment.More particular, three well-known machine learning classifiers and a neural network were trained using the WESAD dataset (NN). The NN classifier was able to achieve accuracy of up to 97.4% for the binary classification problem (stress vs. non-stress). Using a decision tree, the accuracy was 93.28 percent, while using a logistic regression model, it was 83.85 percent. By employing the Kappa coefficient in an inter-rater reliability analysis, we evaluated the effectiveness of the stress models. As a result, the ground truth dataset was a pre-existing bio-signals dataset made up of SC segments. The ground truth dataset's SC segments contain user-reported instances of usability problems they encountered when engaging with a web- based platform during a UX study. Overall, the findings of this paper show that careful thought should be given to the usage of existingbio-signal datasets in various scenarios. Although a one-size-fits-all strategy is not advised, this study offers intriguing new information on how generalizable the bio-signals datasets are.

## REFERENCES

[1] Sarsenbayeva, Z.; Marini, G.; van Berkel, N.; Luo, C.; Jiang, W.; Yang, K.; Wadley, G.; Dingler, T.; Kostakos, V.; Goncalves, J. Does Smartphone Use Drive Our Emotions or Vice Versa? A Causal Analysis. In Proceedings of the 2020 CHI Conference on Human Factors inComputing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–15.

[2] Remy, C.; Bates, O.; Dix, A.; Thomas, V.; Hazas, M.; Friday, A.; Huang, E.M. Evaluation Beyond Usability: Validating Sustainable HCI Research. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; ACM: New York, NY, USA, 2018; pp. 216:1–216:14.

[3] Silvennoinen, J.M.; Jokinen, J.P.P. Aesthetic Appeal and Visual Usability in Four Icon DesignEras. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems,San Jose, CA, USA, 7–12 May 2016; Association for Computing Machinery: San Jose, CA, USA, 2016; pp. 4390–4400.

[4] Díaz-Oreiro, I.; López, G.; Quesada, L.; Guerrero, L.A. Standardized Questionnaires for UserExperience Evaluation: A Systematic Literature Review. *Proceedings* **2019**, *31*, 14.

[5] Tarnowski, P.; Kołodziej, M.; Majkowski, A.; Rak, R.J. Emotion Recognition Using Facial Expressions. *Procedia Comput. Sci.* **2017**, *108*, 1175–1184. [CrossRef]

[6] Rathour, N.; Alshamrani, S.S.; Singh, R.; Gehlot, A.; Rashid, M.; Akram, S.V.; AlGhamdi, A.S. IoMT Based Facial Emotion Recognition System Using Deep Convolution Neural Networks. *Electronics* **2021**, *10*, 1289. Mao, Q.; Dong, M.; Huang, Z.; Zhan, Y. Learning Salient Features for Speech Emotion Recognition Using Convolutional Neural Networks. *IEEE Trans. Multimed.* **2014**, *16*, 2203–2213. [CrossRef]

[7] Lazar, J.; Feng, J.H.; Hochheiser, H. *Research Methods in Human-Computer Interaction*; John Wiley & Sons: Hoboken, NJ, USA, 2010; ISBN 978-0-470-72337-1.

[8] Hernandez, J.; Paredes, P.; Roseway, A.; Czerwinski, M. Under Pressure: Sensing Stress of Computer Users. In Proceedings of the SIGCHI Conference on Human Factors in ComputingSystems; ACM: New York, NY, USA, 2014; pp. 51–60.

[9] Boucsein, W. *Electrodermal Activity*, 2nd ed.; Springer: New York, NY, USA; Dordrecht, The Netherlands; Heidelberg, Germany; London, UK, 2012; ISBN 978-1-4614-1125-3.

[10] Quazi, M.T.; Mukhopadhyay, S.C.; Suryadevara, N.K.; Huang, Y.M. Towards the Smart Sensors Based Human Emotion Recognition. In Proceedings of the 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, Austria, 13– 16 May 2012; pp. 2365–2370.

[11] Kaklauskas, A. Web-based Biometric Computer Mouse Advisory System to Analyze a User's Emotions and Work Productivity. In *Biometric and Intelligent Decision Making Support*; Kaklauskas, A., Ed.; Intelligent Systems Reference Library; Springer International Publishing: Cham, Switzerland, 2015; pp. 137–173. ISBN 978-3-319-13659-2.

[12] Koldijk, S.; Sappelli, M.; Verberne, S.; Neerincx, M.A.; Kraaij, W. The SWELL Knowledge Work Dataset for Stress and User Modeling Research. In Proceedings of the 16th InternationalConference on Multimodal Interaction, Istanbul, Turkey, 12–16 November 2014; ACM: NewYork, NY, USA, 2014; pp. 291–298.

[13] Subramanian, R.; Wache, J.; Abadi, M.K.; Vieriu, R.L.; Winkler, S.; Sebe, N. ASCERTAIN: Emotion and Personality Recognition Using Commercial Sensors. *IEEE Trans. Affect. Comput.* **2018**, *9*, 147–160. [CrossRef]

[14] Alberdi, A.; Aztiria, A.; Basarab, A. Towards an Automatic Early Stress Recognition Systemfor Office Environments Based on Multimodal Measurements: A Review. *J. Biomed. Inform.***2016**, *59*, 49–75. [CrossRef] [PubMed]

[15] Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. In Proceedings ofthe 20th ACM International Conference on Multimodal Interaction, Boulder, CO, USA, 16– 20 October 2018; ACM: New York, NY, USA, 2018; pp. 400–408.

[16] Howard, J.; Gugger, S. Fastai: A Layered API for Deep Learning. *Information* **2020**, *11*, 108. [CrossRef]

[17] Liapis, A.; Katsanos, C.; Karousos, N.; Xenos, M.; Orphanoudakis, T. User Experience Evaluation: A Validation Study of a Tool-Based Approach for Automatic Stress Detection Using Physiological Signals. *Int. J. Hum.–Comput. Interact.* **2021**, *37*, 470–483. [CrossRef]

[18] Chow, C.; Gedeon, T. Evaluating Crowdsourced Relevance Assessments Using Self-Reported Traits and Task Speed. In Proceedings of the 29th Australian Conference on Computer- Human Interaction; Association for Computing Machinery: Brisbane, Qld, Australia, 2017; pp. 407–411.