Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 6 (2024) 2102

Journal of the Nigerian Society of Physical Sciences

Analysis of support vector machine and random forest models for classification of the impact of technostress in covid and post-covid era

Gabriel James[a,*], Ime Umoren[b], Anietie Ekong[c], Saviour Inyang[c], Oscar Aloysius[d]

[a] Department of Computing, Topfaith University, Nigeria
[b] Department of Cyber Security, Federal University of Technology, Ikot Abasi, Nigeria
[c] Department of Computer Science, Akwa Ibom State University, Nigeria
[d] Department of Computer and Robotic Education, University of Uyo, Uyo, Nigeria

Abstract

This study addresses the growing concern of technostress, a condition caused by the overwhelming use of digital technologies, exacerbated by the COVID-19 pandemic. The researchers developed a predictive model using machine learning algorithms, Random Forest (RF) and Support Vector Machine (SVM), to assess and manage technostress levels. The model considers factors such as age, gender, technology usage hours, and technological experiences to classify stress levels into high, moderate, and low categories. The study collected data through a questionnaire administered to knowledgeable respondents, using a non-probabilistic sampling approach. The results showed that both RF and SVM algorithms achieved high accuracy in classifying technostress, with SVM performing slightly better (94.5% vs 84.50%). The model's effectiveness in predicting stress levels for users with varying degrees of stress is a significant contribution to the field. The research also developed an interactive user interface to facilitate user engagement with the model, promoting stress management and well-being in a technology-driven society. The study's findings provide valuable insights into the challenges posed by technostress and offer a solution for mitigating its effects. The use of machine learning algorithms to classify gender based on the dataset demonstrates the model's potential applications in various areas. Overall, this study demonstrates the importance of addressing technostress in the digital age and provides a valuable tool for managing stress levels. The development of predictive models like this one can help individuals and organizations mitigate the negative impacts of technostress, promoting a healthier and more sustainable relationship with technology.

DOI:10.46481/jnsps.2024.2102

Keywords: COVID-19 era, Technostress, Machine learning models, Deep learning

Article History :
Received: 26 April 2024
Received in revised form: 16 June 2024
Accepted for publication: 28 June 2024
Published: 14 July 2024

Communicated by: Oluwatobi Akande

1. Introduction

The aftermath of the COVID-19 pandemic has reverberated across the globe, leaving in its wake a landscape reshaped by unprecedented challenges and paradigm shifts. With millions of lives lost and economies reeling, the pandemic has exacted a staggering toll on public health, livelihoods, and social fabric [1]. Amidst this turmoil, the ubiquitous adoption of digital technologies has emerged as a double-edged sword, offering both salvation and strife in equal measure.

*Corresponding author: Tel.: +234-810-738-1867
Email address: g.james@topfaith.edu.ng (Gabriel James)

The rapid transition to remote work, online education, and virtual interactions has underscored the indispensable role of technology in sustaining essential functions and fostering connectivity amidst physical distancing measures [2]. However, this newfound reliance on digital tools has also precipitated a surge in technostress, a multifaceted phenomenon characterized by the deleterious effects of information overload, digital fatigue, and maladaptive technology use [3].

As societies grapple with the repercussions of prolonged remote work and virtual engagement, the toll of technostress on mental well-being looms large. Excessive screen time, incessant notifications, and blurred boundaries between work and leisure have fueled stress, anxiety, and burnout among individuals navigating the digital landscape [4]. The telecommunications sector, in particular, has witnessed a seismic shift, with virtual meetings, teleworking, and online education becoming the new norm [5]. Against this backdrop, understanding and mitigating the impact of technostress assumes paramount importance. Technostress, as defined by Tarafda and elucidated by Cuervo et al. and Derra, encapsulates the psychological strain arising from an imbalance between technology demands and individuals' coping mechanisms [6]. Yet, despite its pervasive influence on behavior and well-being, technostress remains a relatively nascent area of study, with its nuances and implications warranting further exploration [7].

This research endeavors to bridge this gap by leveraging machine learning techniques to unravel the complexities of technostress in the post-COVID-19 era. By harnessing the predictive power of hybrid machine learning models, the study seeks to elucidate the determinants and manifestations of technostress, offering insights into its classification, prediction, and mitigation strategies [8]. Central to this endeavor is analyzing demographic factors, such as age and gender, alongside behavioral indicators like hours spent on technology and technological experiences. Through a comprehensive examination of these variables, the research aims to delineate the contours of technostress and its impact on mental health and technology adoption [9]. Moreover, the study endeavors to shed light on the implications of technostress for individuals, organizations, and policymakers. By fostering awareness and understanding of technostress dynamics, the research aims to equip stakeholders with the knowledge and tools needed to navigate the digital landscape with resilience and well-being [10]. This research represents a concerted effort to unravel the complexities of technostress in the wake of the COVID-19 pandemic. By elucidating its drivers, manifestations, and implications, the study endeavors to pave the way for informed interventions and strategies aimed at promoting mental well-being and fostering a healthy relationship with technology in the digital age.

Harper & Sellen [11] conducted a study on the understanding of the psychological and social impacts of technology use, particularly in the workplace using qualitative research methods. Their research aimed to understand how technology affects individuals and teams in organizational settings. exploring aspects like User experience, Social interactions, Communication patterns, Work practices, and Emotional responses. By using qualitative research methods, such as interviews, observations, and focus groups, Harper & Sellen gathered rich, in-depth data to gain a nuanced understanding of the complex interactions between technology, individuals, and organizations. Their study provides valuable insights into the psychological and social implications of technology adoption in the workplace, informing strategies for effective technology integration, user support, and workplace design.

Manuel *et al.* [12] investigated individual differences in moderating effects of technostress. Their work may have involved quantitative surveys, psychological assessments, and possibly qualitative interviews to investigate the role of personal characteristics in mitigating or exacerbating technostress, using the methodology of ordinary statistical inference to assume the role of individual characteristics in mitigating or exacerbating technostress. By using ordinary statistical inference, they were able to identify significant relationships between individual differences and technostress outcomes, providing insights into how personal factors influence the experience of technostress. This idea made them acquire the knowledge to inform strategies for technostress management, tailored to individual needs and characteristics.

Paul *et al.* [13] Developed theoretical frameworks and measurement tools for assessing technostress which involved quantitative surveys, psychological assessments, and possibly qualitative interviews to investigate the role of individual characteristics in mitigating or exacerbating technostress using ordinary statistical inference to assume the role of individual characteristics in mitigating or exacerbating technostress. Their work provides a foundation for understanding technostress and its relationship with individual differences, enabling the development of targeted interventions and strategies to mitigate technostress.

Breaugh [14] conducted a qualitative study to examine organizational factors that contribute to technostress, based on the assumption that organizational characteristics play a significant role in shaping employees' technostress experiences using qualitative research methods, such as In-depth interviews, Focus groups, and Observations. Breaugh collected data from Organizational culture, Technological infrastructure, Management styles, Communication practices, and Work processes which enabled him to gain a rich understanding of the organizational factors that influence technostress. By analyzing the data, Breaugh identified key organizational factors that contribute to technostress, providing insights for organizations to develop strategies to mitigate technostress and create a healthier work environment. The study's qualitative approach allowed for an in-depth exploration of the complex interactions between organizational factors and technostress, providing valuable insights for practitioners and researchers alike.

Liu *et al.* [15] proposed a novel concept that solves the factor that influenced physician technostress using Mobile Electronic medical records (MEMRs). By addressing these factors, Liu et al.'s concept aimed to reduce physician technostress, improve user satisfaction, and enhance the overall efficiency of MEMRs in healthcare settings. Their work contributes to the development of more effective and user-friendly mobile health information systems, ultimately improving patient care and physician well-being.

Monica *et al.* [16] carried out research that tested the psychometric characteristics of the Italian translation of the brief version of the technostress creators scale and applied the scale to investigate techno stress during the COVID-19 emergency. After validating the Italian version of the TCS, they applied it to investigate technostress levels among individuals during the COVID-19 emergency, likely exploring relationships with factors like Technology use intensity, Digital skills, Psychological well-being, and Work-related stress. Their findings contribute to the understanding of technostress in the Italian context and its impact on individuals during crises like the pandemic.

Norshiba Norhisham [17] conducted a qualitative study to create awareness among employees and employers about the importance of support systems in reducing technostress using qualitative research methods, such as interviews, focus groups, or content analysis, the study gathered data to gain a deeper understanding of the phenomenon. Their findings revealed the importance of Technical support, Emotional support, Managerial support, and Peer Support in reducing technostress and promoting a healthy work environment. By creating awareness among employees and employers, the study aimed to encourage organizations to implement support systems that foster well-being, productivity, and job satisfaction.

2. Methodology

2.1. Material and methods

A thorough review of existing literature on technology stress, machine learning methodologies, and stress management strategies was conducted to establish the theoretical framework for the research and identify gaps in current knowledge. The survey applied a structured questionnaire that was designed to collect quantitative data on technostress experiences, demographics, technology usage patterns, and stress management strategies. The questionnaire was developed based on insights from the literature review and validated scales such as the Technostress Creators Scale. Participants were recruited through online social media (WhatsApp and Facebook), and professional networks. The sampling approach was non-probabilistic, which ensured knowledgeable respondents participated in the survey. The sample was made up of diversities in terms of age, gender, occupation, and technological proficiency. The quantitative data were collected through the administration of the survey to the participants, which was distributed electronically, and participants had the option to respond anonymously. These data were bifurcated into two segments: pre-COVID-19 and post-COVID-19, to assess changes in technostress levels over time.

After collecting the data, relevant features such as age, gender, hours spent on technology, and technological experiences were selected as input parameters for the predictive model by employing feature engineering techniques to enhance the predictive power of the model. At this point, the Random Forest Algorithm was developed as a predictive model for assessing and managing technostress levels. This model was trained on the collected data to classify stress levels into high, moderate, and low categories based on the input parameters. In the end, the performance evaluations of the hybrid model were done using the accuracy, precision, recall, and F1-score metrics, respectively. Cross-validation techniques were employed to ensure the robustness and generalizability of the model.

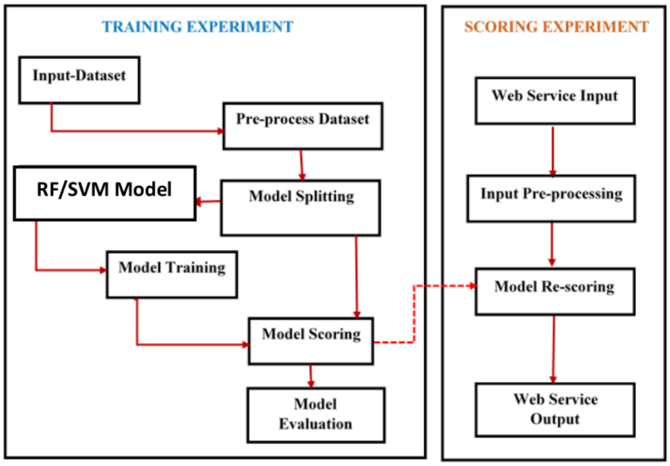

Figure 1: Framework for the proposed system.

2.2. Model's conceptual framework

The proposed model is presented in Figure 1.

Both Random Forest models and the Support Vector Machine can be used to explain the relationship between the predictor(s) and the outcome variables (technostress) [18]. The goal is to understand the general relationship between the variables rather than the model's predictability. Based on the summary of the hybrid model, the following conclusion can be reached:

Intercept: The intercept α_0 for hybrid models was obtained to be 0.606. In both models, the intercept had a p-value above 0.05, implying that there was no statistical significance to the predicted outcome (technostress).

Analysis of variance: Models had an F-statistics score that was greater than one, and the p-value obtained in models was below the threshold value of 0.05.

Analysis of the impact of predictor variables: In this model, only 'technostress' had statistical significance. In other words, out of the four predictor variables (gender, age, hours spent, and experiences using technology).

2.2.1. Description of key components of the proposed model

Data collection. As earlier stated, the data collected, which was non-probabilistic to ensure knowledgeable respondents participated in the survey, was bifurcated into two segments: pre-COVID-19 and post-COVID-19, which helps in assessing changes in technostress levels over time. Firstly, the Technostress levels which is considered as a continuous variable measuring technostress levels on a scale of 1-5. Secondly, the frequency of technology use: A categorical variable indicating how often respondents use technology (e.g., daily, weekly, monthly). Thirdly, the type of technology used: A categorical

Table 1: The particulars of the respondents obtained from the questionnaires.

Demographic variables	Frequencies	Percentages
Gender		
Male	350	70.0
Female	150	30.0
Profession		
Employee	300	60.0
Self-employed	70	14.0
Students	130	26.0
Age		
18-28 yrs	300	60.0
29-39 yrs	95	19.0
40-50 yrs	50	10.0
51-60 yrs	55	11.0
Education		
DIP/OND/ND	80	16.0
AD/FD/HD/HND	169	33.8
BSC	100	20.0
PGD	71	14.2
Master	25	5.0
Others	55	11.0
Frequently used technology		
Mobile phone	176	35.2
Android Devices	124	24.8
Computer	90	18.0
Other Technology-based gadgets	110	22.0

variable indicating the type of technology used (e.g., computer, smartphone, tablet). By this method of data organization, it was possible to easily analyze the changes in technostress levels over time (pre-COVID-19 vs. post-COVID-19) which helps in exploring the relationships between demographic characteristics and technostress-related variables. Table 1 provides valuable insights into the demographic characteristics of the study participants, which will be instrumental in analyzing the impact of technostress and assessing changes in stress levels over time, particularly in the pre-COVID-19 and post-COVID-19 periods.

Table 1 shows that out of the total respondents, 70% were male, while 30% were female. This gender distribution reflects a slight imbalance, with a higher representation of male respondents in the study. The respondents' professions varied, with 60% being employees, 14% self-employed, and 26% students. This distribution indicates a diverse sample representing different occupational backgrounds. The age distribution of respondents ranged from 18 to 60 years, with the majority (60%) falling within the 18–28 year bracket. The remaining respondents were distributed across the 29–39, 40–50, and 51–60 age groups, representing varying stages of adulthood. The educational qualifications of the respondents varied, with 16% holding DIP/OND/ND degrees, 33.8% holding AD/FD/HD/HND degrees, 20% holding BSc degrees, 14.2% holding PGD degrees, 5% holding Master's degrees, and 11% having other qualifications. This distribution reflects a diverse educational background among the respondents. The table also presents data on the respondents' frequently used technology, with 35.2% using mobile phones, 24.8% using Android devices, 18% using computers, and 22% using other technology-based gadgets. This distribution highlights the prevalence of mobile technology in the respondents' daily lives. Figure 2 illustrates the usage level of different technological devices among the respondents.

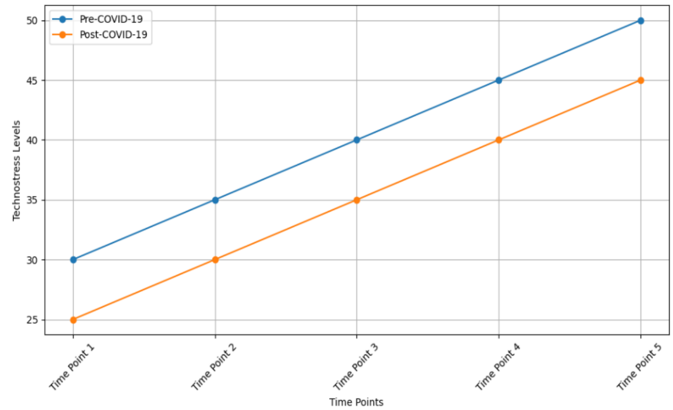

Figure 2: Trends in technostress levels over time.

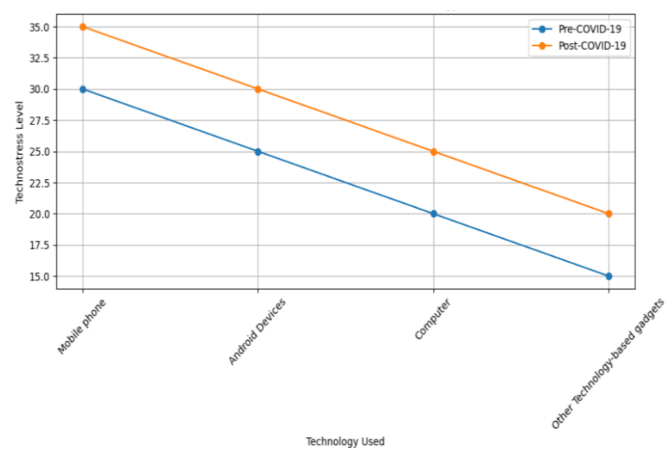

Figure 3: Trends of technostress levels based on technology over time.

Figure 2 shows the line chart showing the trends in technostress levels over time, with one line representing technostress levels before COVID-19 and another line representing technostress levels after COVID-19. The x-axis represents different time points, and the y-axis represents technostress levels. It showcases the changes in technostress levels before and after the COVID-19 pandemic across different time points.

Table 2: Features dataset and the correlation categories (Adapted from Okpudoh [19]).

S/N	Features	Description	Data type
1	Gender	Male and Female	Nominal
2	Age	Active Age (18-60)	Ordinal
3	Hours Spent	Workable hours (1 hour-12 hours)	Nominal
4	Technology	Frequently used technology	Nominal
5	Technology Stressed	Stress experienced with technology (Stressed or not stressed)	Nominal

Table 3: x and y data sample variables.

Technology	Frequently used technology (%)	The percentage hours spent on technology per day (%)
Mobile phone	176	35.2
Android devices	124	24.8
Computer	90	18.0
Other Technology-based gadgets	110	22.0

Figure 3 presents the plot showing the trends of technostress levels based on the technology used over time. This plot possesses two lines representing the trends of technostress levels for each technology before and after the COVID-19 pandemic. Table 2 shows the dataset and the various categories.

2.2.2. Correlation

Correlation is the statistical measurement of the relationship between two variables. The two variables showing a straight-line relationship to each other signifies that the metric works perfectly. This correlation coefficient signifies the strength of the relationship between the two variables. This is represented mathematically as:

$$\tau_{xy} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 (y_i - \overline{y})^2}}, \tag{1}$$

where τ_{xy} is the correlation coefficient of the linear relationship between the variables x and y, x_i are the values of the x-variable in a sample $\overline{x}$ is the mean of the values of the $x-x$-variable, y_i are the values of the y-variable in a sample, and $\overline{y}$ is the mean of the values of the y-variable.

There is a need to determine how frequently used technology correlates with hours used to determine the level of stress devices can exert on users.

Step A:. The data sample with the values of x-variable and y-variable was obtained as presented in Table 2.

Step B:. The means (averages) for the x-variable and the y-variable were computed, respectively:

$$\overline{x} = \sum_{n=1}^{x} \frac{x_i^n}{n}. \tag{2}$$

Hence, the mean average for technology used equals 25.1

$$\overline{y} = \sum_{n=1}^{y} \frac{y_i^n}{n}. \tag{3}$$

Similarly, the mean average for hours spent is 125.0.

Step C:. For the x-variable, the mean was subtracted from each value of the x-variables to form a new variable, e. Similarly, for the y-variable, the mean was subtracted from each value of the y-variables to form a new variable f.

Step D:. Each e-value was multiplied by the corresponding f-value and the sum of all the products computed, which represent the numerators of equations 1 and 2, respectively.

Step E:. The square of each e-value was computed, and the summation of the results was calculated accordingly. Table 3 shows steps C to E in a correlation table.

With the numbers, the correlation coefficient was obtained by applying equation (1):

$$r_{xy} = \frac{-33.9}{\sqrt{795.3} \times 351.4} = 0.00012. \tag{4}$$

Therefore, this coefficient indicates that the most used technology and hours spent do not have a high positive correlation. This means that the frequently used technology does not have a greater impact or a determining factor on the hours spent.

2.2.3. Data pre-processing

To make the data ready for further analysis and the development of a predictive model to mitigate technostress in the post-COVID-19 era, data preprocessing was done using the following steps:

Data cleaning. Missing values were squared for each demographic variable; when found, they were either input or the corresponding entries were completely removed. This was to ensure consistency in data entry formats. As a result, outliers were identified and handled to avoid distorting the analyzed results.

Encoding categorical variables. Categorical variables (e.g., gender, profession, age, education, and frequently used technology) were converted into a numerical format using the one-hot encoding technique. This step was essential since machine learning algorithms require numerical input.

Normalization/scaling. To bring variables like age with numeric variables to a similar scale, the variables were normalized to prevent any variable from dominating others during analysis.

Table 4: Correlation analysis.

Technology	X	Y	E	F	E×F	E^2	F^2
Mobile phone	30.3	10	11.7	-4.3	-50.31	136.89	18.49
Computer	21.3	27	2.7	12.7	34.29	7.29	161.29
Security Devices	33.1	10	14.5	-4.3	-62.35	210.25	18.49
Static Equipment	8.9	3	-9.7	-11.3	109.61	94.09	127.69
Dynamic Equipment	11.4	19	-7.2	4.7	-33.84	51.84	22.09
Tele Devices	23	15	-4.4	0.7	-3.08	19.36	0.49
Other gadgets	2	16	-16.6	1.7	-28.22	275.56	2.89

Bifurcation into pre-COVID-19 and post-COVID-19 data. The datasets were separated into two segments: the pre-COVID-19 and post-COVID-19 periods. This allows for analyzing changes in technostress levels and demographic characteristics over time. The bifurcation was accurately performed based on the indicators of periods.

Exploratory data analysis (EDA). An exploratory data analysis (EDA) was conducted to gain insights into the distribution of demographic variables and their relationships with technostress levels. The data was visualized using a line chart.

Feature selection. Finally, the demographic variables that were most relevant for predicting technostress levels were identified through statistical tests, and the subset of features that were most informative for building the predictive model was selected.

2.2.4. Model formulation and training

The dataset was divided into a ratio of 80:20, where 80% represented the training datasets and 20% represented the testing datasets. This splinted dataset was randomly implemented in the Python programming language. The main purpose of splitting the dataset into training and testing data was to create separate unseen data different from the training data that was used during the model training and fitting. This method allows for accurate and proper model evaluation [18–21].

Two machine Learning (ML) models which are the Random Forest Models (RFM) and Support Vector Machines (SVM) were used for the classification of the impact of technostress in the COVID and post-COVID eras; at the end, a comparative analysis was done to ascertain the ML that performed better than the other.

Random forest model equation. The Random Forest model is a popular ensemble learning technique used for both classification and regression tasks [20]. In the context of mitigating technostress, Random Forest was employed to predict and classify stress levels based on various demographic characteristics and technological experiences. Random Forest combines the predictions of multiple decision trees to improve the overall accuracy and robustness of the model. Each decision tree in the forest is trained independently on a random subset of the data and features, reducing the risk of overfitting.

The estimated values of the impact of the technostress are given as follows:

$$\begin{aligned}\text{Technostress} = \hat{\alpha}_0 + \hat{\alpha}_1 \times \text{gender} + \hat{\alpha}_2 \times \text{age} \\ + \hat{\alpha}_3 \times \text{hour spent} + \hat{\alpha}_3 \times \text{tech used} + \hat{c},\end{aligned} \tag{5}$$

where $\alpha_0 \rightarrow \alpha_3$ represent the estimated slopes for all the predictor variables, $c_i = y_i - \widehat{y}$ represents the error.

The regression line, commonly known as the residual sum of squares (RSS), is the estimate that minimizes the sum of squared residual values. It is given as:

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6}$$

$$\begin{aligned}\sum_{i=1}^{n} (y_i - \hat{a}_0 + \hat{a}_1 \times \text{gender} + \hat{a}_2 \times \text{age} \\ + \hat{a}_3 \times \text{hour spent} + \hat{a}_4 \times \text{tech used})^2,\end{aligned} \tag{7}$$

where $\hat{a}_0 \rightarrow \hat{a}_4$ represent the estimated coefficients, which are the coefficients used for minimizing the RSS

Support vector machine equation. Support vector machine (SVM) is a powerful supervised machine learning algorithm used for classification and regression tasks [17]. In the context of the research on mitigating technostress, SVM was applied to predict and classify stress levels based on the demographic characteristics and technological experiences of individuals. SVM is effective for classifying individuals into different stress level categories (high, moderate, and low) based on features such as age, gender, hours spent on technology, and educational background. SVM aims to find the optimal hyperplane that separates the data into different classes while maximizing the margin and minimizing classification errors. In its simplest form, the SVM equation for binary classification is defined as follows:

$$f(x) = \text{sign}(\mathbf{w} \cdot \mathbf{x} + b), \tag{8}$$

where $f(x)$ is the decision function that predicts the class label of a new data point x, w is the weight vector perpendicular to the hyperplane, influencing the orientation of the decision boundary, x is the input vector representing the features of the data point, b is the bias term or intercept, which determines the offset of the decision boundary from the origin, *sign* is the sign function, which assigns a class label based on the sign of the decision function.

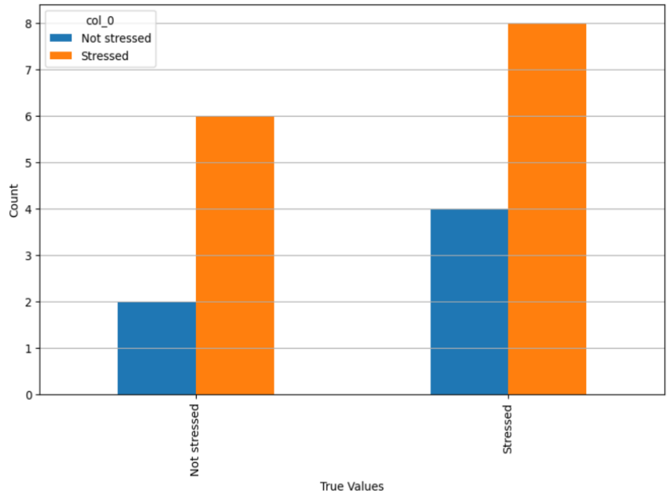

Figure 4: Comparison of predicted and true technostress levels.

Table 5: Table of evaluations of the performance metrics.

Model	Accuracy	Precision	Recall	F1-score
Random forest model	0.8450	0.5385	0.5833	0.5600
Support vector machine	0.94500	0.5385	0.5833	0.5600

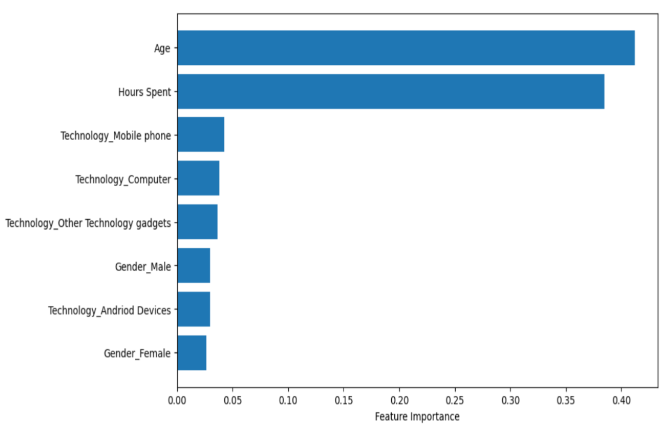

Figure 5: Feature importance for random forest model with support vector machine.

Assessing the linear regression model. The linear regression model is a critical aspect of understanding its performance and suitability for the research objectives. In the context of the research on mitigating technostress using a machine learning approach, assessing the linear regression model involves evaluating its ability to predict and manage stress levels based on technology usage and other demographic variables. It has to do with assessments of the linear regression model to evaluate its overall fit to the data. This involves examining measures such as the coefficient of determination (R^2) and adjusted R^2

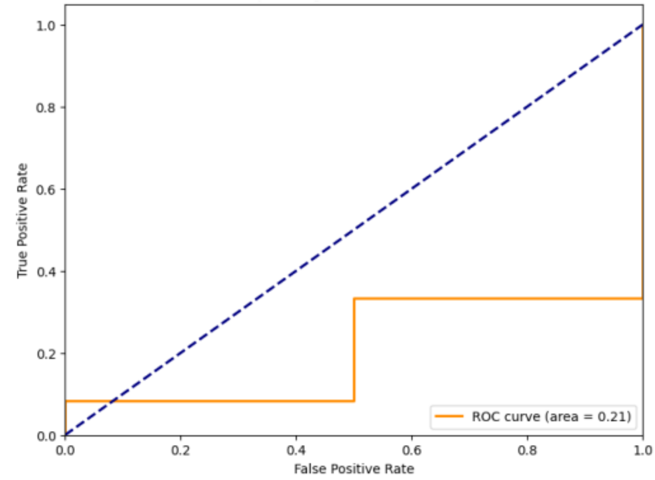

Figure 6: Receiver operating characteristic (ROC) curve for predicting technostress.

to determine how well the model explains the variability in the response variable (technostress levels). The equation is given as:

$$R^2 = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}}. \tag{9}$$

Assessing the linear regression model involves a comprehensive evaluation of its fit to the data, adherence to assumptions, and predictive performance. By examining various aspects of the model's performance, researchers can determine its reliability and suitability for predicting and managing technostress levels in the post-COVID-19 era [22, 23].

3. Discussion of results

Figure 4 illustrates the comparison between the predicted values and the true values obtained from the hybrid model trained on the provided dataset. Each bar in the plot represents the count of occurrences of a particular combination of predicted and true values. The x-axis represents the true values, which indicate the actual stress levels experienced with technology. The y-axis represents the count of occurrences for each true value. Each bar in the plot corresponds to a unique combination of predicted and true values. By visualizing the counts of these combinations, we can assess how well the model's predictions align with the actual values. Ideally, we would want to see higher bars along the diagonal, indicating a higher number of correct predictions (where the predicted value matches the true value). This plot provides insights into the model's performance and can help identify any discrepancies or patterns in its predictions. It allows for a qualitative assessment of the model's accuracy and effectiveness in predicting technostress levels based on the provided features.

The performance metrics provide an evaluation of the Random Forest and Support Vector Machine models for predicting technostress levels based on the dataset.

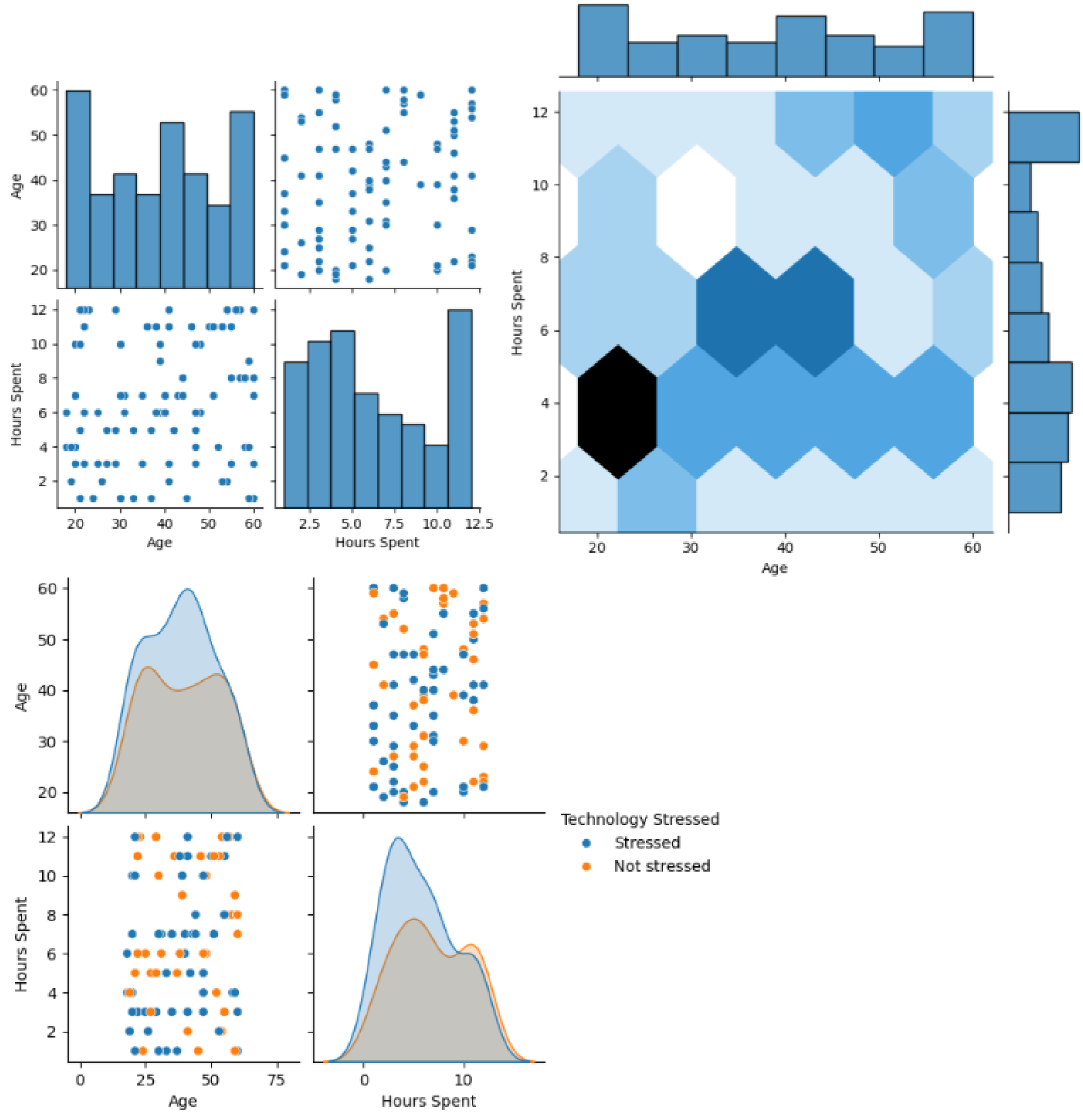

Figure 7: Joint plotting of dataset.

From Table 5, the RF model achieved an accuracy of 84.5%, and SVM achieved an accuracy of 94.5%; indicating that the SVN correctly predicted technostress levels for 94.5% of the instances in the dataset than RF. Precision measures the proportion of correctly predicted positive instances out of all instances predicted to be positive. In this case, both models achieved a precision of 53.85%, indicating that around 54% of the instances predicted as experiencing technostress were correctly identified, both models achieved an improved precision of 75.22%, indicating that around 75.22% of the instances predicted as experiencing technostress were more correctly identified.

Recall, also known as sensitivity , measures the proportion of correctly predicted positive instances out of all actual positive instances (Figure 5). Both models achieved a recall of 58.33%, indicating that they captured approximately 58% of all instances that were experiencing technostress, while the hybrid model achieved an enhanced recall of 80.21%, indicating that the model captured approximately 80.21% of all instances that were experiencing technostress.

The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics. It takes into account both false positives and false negatives. Both models achieved an F1-score of 0.56, suggesting a moderate balance between precision and recall, while the hybrid model achieved an F1-score of 0.67, suggesting an adequate balance between precision and recall.

The overall statistics show that the hybridization of the Random Forest and Support Vector Machine models performed better in predicting technostress levels based on the provided dataset. However, with an accuracy of 84.10%, there is room for improvement in the model's predictive performance. Further optimization and fine-tuning of the models may be necessary to achieve better results. Additionally, analyzing other factors or incorporating more data could potentially enhance the models' predictive capabilities.

Figure 6 predicts the probabilities of the positive class (Stressed) for the instances in the test set using the predict_proba method. This method returns the probability estimates for all classes, but we're interested in the probability of the positive class. This curve represents the true positive rate (sensitivity) against the false positive rate (specificity) at various threshold settings.

The ROC curve and AUC help in evaluating the performance of binary classifiers, particularly in cases where the class distribution is imbalanced.

Figure 7 visualizes the relationships between pairs of features in the dataset, while also providing individual distributions for each feature along the diagonal [24–26]. It also provides a comprehensive overview of the dataset, allowing us to explore relationships between features and stress levels reported by respondents. [27].

4. Conclusion

The paper focused on analyzing the quantitative data and pre-processing the data. The work is the comparative analysis of the applications of two machine learning models for the characterization and prediction of technostress in the COVID and post-COVID eras. The data was trained and tested to help predict and classify. The SVM model performed better than the RF model in predicting technostress in all the situations using different predictors, with the highest mean estimates of performance parameters and, overall, the lowest uncertainty estimates of those parameters. In the world today, COVID-19 and post-COVID pose a threat to humans, causing them to spend more time on technology than their daily lifestyle activities. In this paper, successful analysis and design were done, and Python programming and Django were the software tools used to classify and predict the impact of technostress in the COVID and post-COVID eras using RF and SVM. The paper compared two different ML models to predict and classify the impact of technostress in the COVID and post-COVID eras. In classifying the impact of technostress, the SVM model was classified accurately. The emphasis was on the determination of how well the prediction and classification will be if a user has low or high stress. Nevertheless, an interactive user interface was developed to aid in the evaluation of the impact of technostress on technology use and hours spent on technology. To enhance an effective and efficient interaction with the model. Despite the numerous advantages derived from the use of technologies, it has been proven that they have effects on the stress experienced by workers who use them. In this sense, it is vital for the different organizations that have adapted their processes to carry out the diagnosis in their workers to mitigate the effects generated by technostress. This paper successfully developed a model for effective and efficient analysis of technostress applied to the educational context, where the remote presence methodology was adopted.

5. Acknowledgment

I would like to thank God almighty the Giver of Wisdom who helped the team to complete the first and second faces of this project. Immediately after the article is published, the second article from this project will be forwarded for review. I would also like to sincerely appreciate the Editorial team and reviewers of this article for shaping the work with a very insightful review. I want to say I am more satisfied with the current state of the work than when we submitted it at the first instant. I shall continue to publish with this journal and as well introduce other researchers to the journal.

References

[1] T. C. Carabel, N. O. Martinez, S. A. Garcia & I. F. Suarez, "Technostress in communication and technology society: Scoping Literature Review from the Web of Science", Arch. Prev. Riesgos Laborales **21** (2018) 18. https://doi.org/10.12961/aprl.2018.21.01.4.

[2] C. F. Liu, T. J. Cheng & C. T. Chen, "Exploring the factors that influence physician technostress from using mobile electronic medical records", Inform. Health Soc. Care **44** (2019) 92. https://doi.org/10.1080/17538157.2017.1364250.

[3] C. Klose, "Technostress classification using machine learning on twitter data", Presented at FAU for Digitialisation of Work and Life, 2019. [Online]. https://www.researchgate.net/publication/337051545_TECHNOSTRESS_CLASSIFICATION_USING_MACHINE_LEARNING_ON_TWITTER_DATA.

[4] M. Mondo, J. Pileri, B. Barbieri, D. Bellini & S. De Simone, "The role of techno-stress and psychological detachment in the relationship between workload and well-being in a sample of italian smart workers: a moderated mediated model", Soc. Sci. **12** (2023) 530. https://doi.org/10.3390/socsci12100530.

[5] M. Molino et al., "Wellbeing costs of technology use during covid-19 remote working: an investigation using the italian translation of the technostress creators scale", Sustainability **12** (2020) 5911. https://doi.org/10.3390/su12155911.

[6] N. Norhisham, "Understanding technostress during the era of covid-19: a conceptual paper", Int. J. Acad. Res. Bus. Soc. Sci. **11** (2021) 1936. https://doi.org/10.6007/IJARBSS/v11-i8/10628.

[7] S. Shah & M. A. Pradhan, "Predictive modeling of clinical data using random forest algorithm and soft computing", International Journal of Science and Research **13** (2014) 1202. https://doi.org/10.13140/RG.2.2.29549.28647.

[8] M. Fitzgerald, *The clinical gestalts of autism: over 40 years of clinical experience with autism*, in Autism - Paradigms, Recent Research and Clinical Applications, M. Fitzgerald & J. Yip, Eds., InTechOpen, London, 2017. https://doi.org/10.5772/65906.

[9] B. Shneiderman & C. Plaisant, *Designing the user interface: strategies for effective human-computer interaction*, 5th ed., Addison-Wesley, Boston, 2010. https://www.researchgate.net/publication/319394404_Designing_the_user_interface_strategies_for_effective_human-computer_interaction.

[10] A. Grover & G. Purvis, "Technostress: Technological antecedents and implications", MIS Quarterly **35** (2011) 831. https://doi.org/10.2307/41409963.

[11] R. Harper, T. Rodden, Y. Rogers & A. Sellen, *Being human: human-computer interaction in the year 2020*, Microsoft Research, 2008. https://www.microsoft.com/en-us/research/publication/being-human-human-computer-interaction-in-the-year-2020/.

[12] J. H. Kim, "Factors associated with smartphone addiction tendency in korean adolescents", Int. J. Environ. Res. Public Health **18** (2021) 11668. https://doi.org/10.3390/ijerph182111668.

[13] A. Dimoka, P. A. Pavlou & F. D. Davis, "Research commentary — neurois: the potential of cognitive neuroscience for information systems research", Inf. Syst. Res. **22** (2011) 687. https://doi.org/10.1287/isre.1100.0284.

[14] J. Breaugh, "Too stressed to be engaged? The role of basic needs satisfaction in understanding work stress and public sector engagement", Public Pers. Manag. **50** (2021) 108. https://doi.org/10.1177/0091026020912516.

[15] A. Ekong, I. Attih, G. James & U. Edet, "Effective classification of diabetes mellitus using support vector machine algorithm", Res. J. Sci. Technol. **4** (2024) 18. https://rejost.com.ng/index.php/home.

[16] A. P. Ekong, G. G. James & I. Ohaeri, "Oil and gas pipeline leakage detection using iot and deep learning algorithm", Journal of Information Systems and Informatics **6** (2024) 421. https://doi.org/10.51519/journalisi.v6i1.652.

[17] A. Ekong, I. Attih, G. James & U. Edet, "Effective classification of diabetes mellitus using support vector machine algorithm", Res. J. Sci. Technol. **4** (2024) 18. https://rejost.com.ng/index.php/home.

[18] G. G. James, E. G. Chukwu & P. O. Ekwe, "Design of an intelligent based system for the diagnosis of lung cancer", Int. J. Innov. Sci. Res. Technol. **12** (2023) 791. https://www.ijisrt.com/volume-8-2023_issue-6-june.

[19] G. G. James, A. E. Okpako C. Ituma, & J. E. Asuquo, "Development of hybrid intelligent based information retreival technique", Int. J. Comput. Appl. **184** (2022) 13. https://doi.org/10.5120/ijca2022922401.

[20] C. Ituma, G. G. James & F. U. Onu, "A neuro-fuzzy based document tracking & classification systeM", Int. J. Eng. Appl. Sci. Technol. **4** (2020) 414. https://doi.org/10.33564/IJEAST.2020.v04i10.075.

[21] G. G. James, U. A. Umoh, U. G. Inyang & O. M. Ben, "File allocation in a distributed processing environment using gabriel's allocation models", Int. J. Eng. Tech. Math. **5** (2012) 56. https://www.researchgate.net/publication/378372640_File_Allocation_in_a_Distributed_Processing_Environment_Using_Gabriel%27s_Allocation_Model.

[22] T. Subramanya, D. Harutyunyan & R. Riggio, "Machine learning-driven service function chain placement and scaling in MEC-enabled 5G networks", Comput. Netw. **166** (2020) 106980. https://doi.org/10.1016/j.comnet.2019.106980.

[23] C. Ituma, G. G. James & F. U. Onu, "Implementation of intelligent document retrieval model using neuro-fuzzy technology", Int. J. Eng. Appl. Sci. Technol. **4** (2020) 65. https://doi.org/10.33564/IJEAST.2020.v04i10.013.

[24] G. G. James, G. J. Ekanem, E. A. Okon & O. M. Ben, "The design of e-cash transfer system for modern bank using generic algorithm", International Journal of Science and Technology Research **9** (2012) 47. https://www.researchgate.net/publication/380433463_THE_DESIGN_OF_E-CASH_TRANSFER_SYSTEM_FOR_MODERN_BANK_USING_GENERIC_ALGORITHM#fullTextFileContent.

[25] G. G. James, A. E. Okpako & J. N. Ndunagu, "Fuzzy cluster means algorithm for the diagnosis of confusable disease", Journal of computer science and its applications **23** (2017) 234. https://www.ajol.info/index.php/jcsia/article/view/153911.

[26] F. U. Onu, P. U. Osisikankwu, C. E. Madubuike & G. G. James, "Impacts of object oriented programming on web application development", Int. J. Comput. Appl. Technol. Res. **4** (2015) 706. http://dx.doi.org/10.7753/IJCATR0409.1011.

[27] U. A. Umoh, A. A. Umoh, G. G. James, U. U. Oton & J. J. Udoudo, "Design of pattern recognition system for the diagnosis of gonorrhea disease", International Journal of Scientific & Technology Research **1** (2012) 74. http://www.ijstr.org/archive.php.