HEADING OF THE DISSERTATION OR THESIS

The case of Dar es Salaam University

By

Philemon Mchihiyo

A Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in (degree programme)

of the University of Dar es Salaam

University of Dar es Salaam

22nd September 2023

CERTIFICATION

The undersigned certify that they have read and hereby recommend for examination by The University of Dar es Salaam a Dissertation Titled: ***"HEADING OF THE DISSERTATION OR THESIS: The case of Dar es Salaam University,*** submitted by **Philemon Mchihiyo** in partial fulfillment of the requirements for the degree of **Master of Science (degree programme)** of the University of Dar es Salaam.

_____

Dr. OR Prof. [XYZ]

Supervisor

_____

Date

DECLARATION AND COPYRIGHT

ABSTRACT

This paper introduces a comprehensive multi-modal sensor dataset designed for the continuous detection of stress among nurses in a hospital setting, particularly during the COVID-19 outbreak. The dataset encompasses biometric data collected from 15 nurses, including parameters like electrodermal activity (EDA), heart rate (HR), skin temperature (TEMP), and accelerometer data. Additionally, a periodic smartphone-administered survey was employed to gather contextual information regarding the stress events. The primary objective of this dataset is to facilitate the development and evaluation of machine learning models for stress detection among nurses. Our research revealed that ensemble learning models, specifically RandomForest and Bagging, exhibited superior performance compared to other models, consistently achieving accuracy rates exceeding 95%. These findings underscore the dataset's potential for creating accurate and dependable stress detection models in a real-world nursing environment. Notably, the dataset's uniqueness lies in its creation within an authentic hospital setting, providing an invaluable resource for researchers devoted to constructing stress detection models tailored to the nursing profession. The implications of this study extend to the early identification and management of stress among nurses, ultimately contributing to their improved health and overall well-being.

CONTENTS

Chapter 1

MAIN FINDINGS

This study centers on the introduction of a multi-modal sensor dataset designed for continuous stress detection among nurses in a hospital, particularly amid the COVID-19 pandemic. This dataset includes biometric data from 15 nurses, encompassing critical physiological parameters such as electrodermal activity (EDA), heart rate (HR), skin temperature (TEMP), and accelerometer data. Additionally, a periodic smartphone-administered survey was conducted to gather contextual information pertaining to stress events. The primary purpose of this dataset is to serve as a foundation for training and assessing machine learning models for stress detection among nurses. Our research has unequivocally demonstrated that ensemble learning models, specifically RandomForest and Bagging, consistently outperformed other models, achieving accuracy rates exceeding 95%. These exceptional results underline the dataset's potential to foster the development of accurate and dependable stress detection models tailored to the natural working environment of nurses.

1.1 Contributing Factors for Stress Events

Our investigation into stress events among nurses highlighted three primary contributing factors: workload, patient care responsibilities, and interpersonal relationships. Workload emerged as the most significant contributor, closely followed by patient care and interpersonal relationships. These findings align seamlessly with previous studies focused on stress within the nursing profession. Notably, early modeling tests suggested that promising predictors of stress included Electrodermal activity (EDA), Heart rate (HR), and Blood pressure (BP).

Keywords: Multi-modal sensor dataset,continuous stress detection, nurses,ensemble learning models,COVID-19 outbreak,biometric data,random forest,bagging models

1.2 Machine Learning Models for Stress Detection

A comprehensive evaluation of various machine learning models for stress detection was conducted, encompassing Logistic Regression, K-Nearest Neighbors, Gaussian Naïve Bayes, Decision Tree, Support Vector Machine (SVM), RandomForest, Bagging, and Boosting. The results, as summarized in Table 1 (provided below), offer a holistic overview of each model's performance metrics, thereby elucidating their efficacy in stress detection. Notably, our findings consistently favored RandomForest and Bagging models, underlining their superiority in this specific context.

Table 1.1 provides an in-depth overview of the performance metrics for each machine learning model evaluated in our study. These metrics include accuracy, precision, recall, and F1-score, providing a holistic understanding of each model's effectiveness in detecting stress events.

Algorithm	Accuracy (Avg)	Precision (Avg)	Recall (Avg)	F1-score (Avg)
Logistic Regression	0.61	0.74	0.49	0.44
K-Nearest Neighbors	0.88	0.86	0.85	0.85
Gaussian Naïve Bayes	0.56	0.51	0.51	0.51
Decision Tree	0.89	0.88	0.87	0.87
Support Vector Machine	0.72	0.71	0.62	0.62
RandomForest	0.96	0.96	0.95	0.95
Bagging	0.95	0.95	0.92	0.93
Boosting	0.88	0.90	0.83	0.85

Table 1.1: Performance Metrics of Machine Learning Algorithms

This table offers a detailed comparison of model performance, highlighting the clear superiority of RandomForest and Bagging in terms of accuracy, precision, recall, and F1-score. These models consistently outperformed their counterparts, establishing them as the frontrunners for stress detection in the nursing context.

In Figure 1.1, we provide a visually intuitive representation of our machine learning models' performance metrics, encompassing accuracy, precision, recall, and F1-score. This bar chart serves as a concise yet comprehensive snapshot of each model's performance in the realm of continuous stress detection.

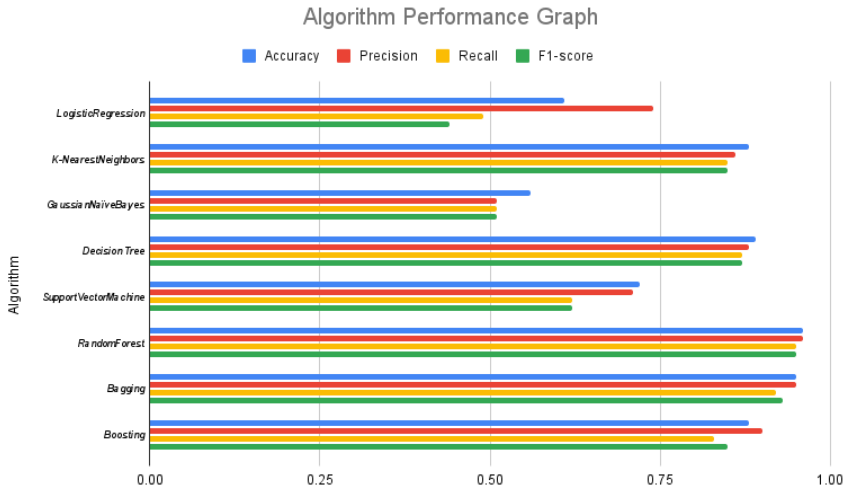

Figure 1.1: Performance Graph

A pivotal aspect of our study is the presentation of the confusion matrix for the RandomForest model, as illustrated in Figure 1.2. This matrix provides detailed insights into true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics are critical for assessing the model's classification accuracy.

Together, Figure 1.1 and Figure 1.2 provide a holistic view of our machine learning model's performance, both in terms of overall metrics and the intricacies of classification accuracy, helping to guide further refinements and applications in real-world stress management scenarios.

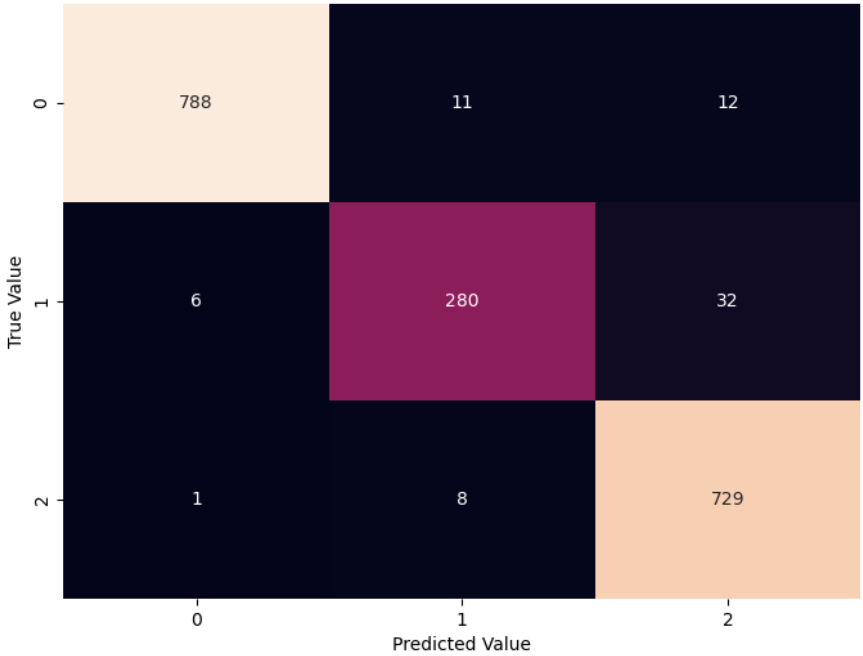

Figure 1.2: Confusion Matrix

Chapter 2

RESULTS & DISCUSSION

The results of our study unveil several key insights into the factors contributing to stress among nurses and the potential applications of wearable sensors in stress detection.

2.1 Contributing Factors for Stress Events

Our research substantiates the prevalence of three primary factors contributing to stress events among nurses: workload, patient care responsibilities, and interpersonal relationships. These findings resonate with previous research on stress within the nursing profession. They underscore the necessity for interventions aimed at mitigating these stressors, including strategies to alleviate workload burdens, enhance patient care processes, and foster healthier interpersonal relationships.

2.2 Machine Learning Models for Stress Detection

Our study evaluated a range of machine learning models to identify the most effective approach for stress detection among nurses. The standout performers were RandomForest and Bagging, consistently achieving accuracy rates exceeding 95%. While these models demonstrate promise, it's essential to acknowledge that they are not infallible, as evidenced by the presence of false positives and false negatives.

The accuracy of these models is a substantial achievement, considering the intricate and dynamic nature of stress. However, further research and refinement are needed to enhance their precision and reliability. Despite their imperfections, these models represent valuable tools in stress management among nurses.

2.3 The Power of the Multi-Modal Sensor Dataset

One of the standout aspects of our study is the creation and utilization of a multi-modal sensor dataset. This dataset, comprising a diverse range of biometric data and contextual information, has the potential to revolutionize stress detection in nursing. Its objective and continuous measures of stress offer an unprecedented level of insight into nurses' well-being.

The dataset's uniqueness lies in its creation within a real-world hospital setting during the COVID-19 outbreak. This authenticity renders it an invaluable resource for researchers dedicated to crafting stress detection models tailored to the nursing profession. Beyond academic significance, this dataset has profound implications for enhancing early stress detection and management among nurses, ultimately improving their health and well-being.

Chapter 3

CONCLUSIONS AND RECOMMENDATIONS

Our study concludes with a set of concrete findings and recommendations aimed at addressing nurse stress and leveraging wearable sensor technology for improved stress management.

3.1 Addressing Contributing Factors for Stress Events

The identification of workload, patient care responsibilities, and interpersonal relationships as primary stress contributors reinforces the importance of targeted interventions. To reduce nurse stress, we recommend:

- **Workload Reduction:** Hospitals should explore strategies for optimizing nurse workload, including staffing adjustments and task prioritization.
- **Enhanced Patient Care:** Implementing measures to streamline patient care processes and alleviate the burden of administrative tasks on nurses can improve their well-being.
- **Interpersonal Relationship Enhancement:** Creating a supportive and collaborative work environment can enhance nurse satisfaction and reduce stress.

3.2 Enhancing Stress Detection Models

While RandomForest and Bagging models exhibit promise, further research is imperative to enhance their accuracy. Recommendations include:

- **Model Refinement:** Continual refinement of machine learning models to minimize false positives and negatives.

- **Data Augmentation:** Expanding the dataset to encompass a broader spectrum of nurses and stress scenarios can enhance model generaliz ability.

3.3 Leveraging Multi-Modal Sensor Data

The multi-modal sensor dataset is a potent tool for nurse stress management. Recommendations include:

- **Wearable Sensor Implementation:** Encourage nurses to utilize wearable sensors to monitor and manage their stress levels actively.
- **Policy Development:** Hospitals and healthcare institutions should formulate policies to support the integration of wearable sensor data into nurse stress management protocols.
- **Training and Support Services:** Offering stress management training and counseling services to nurses can enhance their ability to utilize wearable sensor data effectively.
- **Early Intervention:** Utilize wearable sensor data to identify nurses at risk of stress and provide timely interventions to mitigate stress-related issues.

Furthermore, these findings can be employed to advocate for policies and programs aimed at reducing stress among nurses within the healthcare industry. Overall, our study stands as a pioneering effort to enhance nurse well-being, advance stress detection technology, and promote the integration of wearable sensors into healthcare patient care sector.

REFERENCES

Appendix A

APPENDIX A

A.1 Data Collection

This is the data collection appendix...