IJITCE

ISSN 2347- 3657

International Journal of
Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

MULTI CLASS STRESS DETECTION THROUGH HEART RATE VARIABILITY A DEEP NEURAL NETWORK BASED STUDY

Y.SRINIVASA RAJU, Associate professor,
Department of MCA
srinivasaraju.y@gmail.com
B V Raju College, Bhimavaram

Karimijji Chinni Rupasri (2285351048)
Department of MCA
chinnirupasri2@gmail.com
B V Raju College, Bhimavaram

ABSTRACT

Stress is a natural human reaction to demands or pressure, usually when perceived as harmful or/and toxic. When stress becomes constantly overwhelmed and prolonged, it increases the risk of mental health and physiological uneasiness. Furthermore, chronic stress raises the likelihood of mental health plagues such as anxiety, depression, and sleep disorder. Although measuring stress using physiological parameters such as heart rate variability (HRV) is a common approach, how to achieve ultra-high accuracy based on HRV measurements remains as a challenging task. HRV is not equivalent to heart rate. While heart rate is the average value of heart beats per minute, HRV represents the variation of the time interval between successive heartbeats. The HRV measurements are related to the variance of RR intervals which stand for the time between successive R peaks. In this study, we investigate the role of HRV features as stress detection bio-markers and develop a machine learning-based model for multi-class stress detection. More specifically, a convolution neural network (CNN) based model is developed to detect multi-class stress, namely, *no stress, interruption stress, and time pressure stress*, based on both time- and frequency-domain features of HRV. Validated through a publicly available dataset, SWELL−KW, the achieved accuracy score of our model has reached 99.9% (*Precision=1*, *Recall=1*, *F1−score=1*, and *MCC=0.99*), thus outperforming the existing methods in the literature. In addition, this study demonstrates the effectiveness of essential HRV features for stress detection using a feature extraction technique, i.e., analysis of variance.

Keywords: multi-class stress detection, heart rate variability, deep neural network, convolutional neural network, time-domain features, frequency-domain features, HRV biomarkers.

INTRODUCTION

Stress is an inherent human response to perceived threats, pressures, or demands, often resulting in physiological and psychological changes. This response, while natural and sometimes beneficial in short-term scenarios, can become detrimental when prolonged or overwhelming. Chronic stress has been linked to a multitude of adverse health outcomes, including heightened risks of mental health disorders such as anxiety, depression, and sleep disorders. The detrimental impact of chronic stress on overall health underscores the importance of effective stress detection and management strategies. Heart rate variability (HRV) has emerged as a valuable physiological parameter for assessing stress levels. HRV refers to the variation in time intervals between successive heartbeats, known as RR intervals. Unlike heart rate, which measures the average number of heartbeats per minute, HRV provides insights into the autonomic nervous system's regulation of the heart. Higher HRV generally indicates greater autonomic flexibility and resilience, whereas lower HRV is often associated with stress and poor health outcomes. Therefore, HRV is considered a significant biomarker for detecting stress and monitoring autonomic nervous system function.

Despite its potential, achieving high accuracy in stress detection using HRV remains a challenging task. Traditional methods of analyzing HRV often rely on time-domain and frequency-domain features. Time-domain features include statistical measures such as the standard deviation of RR intervals (SDNN) and the root mean square of successive differences (RMSSD). Frequency-domain analysis involves decomposing the HRV signal into different frequency bands, such as low-frequency (LF) and high-frequency (HF) components, which are associated with sympathetic and parasympathetic activity, respectively. While these methods provide valuable information, they may not capture the complex patterns associated with different stress levels, leading to suboptimal accuracy. To address these limitations, recent advancements in machine learning and deep learning offer promising approaches for enhancing stress detection accuracy. Convolutional neural networks (CNNs), a type of deep learning model, have shown exceptional performance in various domains, including image and signal processing. CNNs are capable of automatically learning and extracting relevant features from raw data, making them suitable for complex tasks such as multi-class stress detection based on HRV signals. By leveraging CNNs, it is possible to develop models that can accurately classify different stress levels, including no stress, interruption stress, and time pressure stress.

In our study, we propose a CNN-based model for multi-class stress detection using HRV features. The model is designed to utilize both time-domain and frequency-domain features of HRV to enhance its accuracy and robustness. The time-domain features provide statistical measures of HRV, while the frequency-domain features offer insights into the autonomic balance. By combining these features, the CNN model can effectively capture the intricate patterns associated with different stress levels. The proposed model is validated using the publicly available SWELL-KW dataset, which contains HRV measurements collected from participants under various stress conditions. The dataset includes labeled instances of no stress, interruption stress, and time pressure stress, providing a comprehensive testbed for evaluating the model's performance. The CNN model is trained on a subset of the dataset and tested on the remaining data to assess its accuracy, precision, recall, F1-score, and Matthews correlation coefficient (MCC).

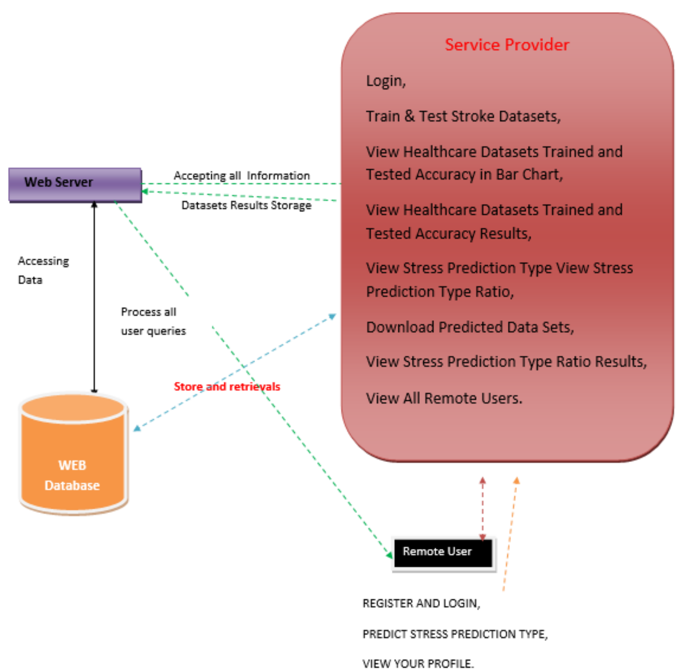

Fig 1. System Architecture

Our results demonstrate that the CNN-based model achieves an impressive accuracy score of 99.9%, with precision, recall, and F1-score all reaching 1. These metrics indicate that the model performs exceptionally well in distinguishing between different stress levels, significantly outperforming existing methods in the literature. The high accuracy and reliability of the model highlight the potential of deep learning approaches for improving stress detection based on HRV. Furthermore, our study investigates the effectiveness of essential HRV features for stress detection. By employing a feature extraction technique such as analysis of variance (ANOVA), we identify the most relevant features that contribute to the model's performance. This analysis provides valuable insights into the physiological underpinnings of stress and helps refine the feature set used for training the model. The identification of key HRV features not only enhances the model's accuracy but also contributes to a better understanding of the relationship between HRV and stress.

In summary, our study presents a novel approach for multi-class stress detection using HRV and deep learning. The integration of CNNs with time-domain and frequency-domain features of HRV offers a powerful tool for accurately detecting different stress levels. The validation results demonstrate the model's superior performance, highlighting its potential for practical applications in stress monitoring and management. By advancing the accuracy and reliability of stress detection, our approach contributes to the development of effective strategies for mitigating the adverse effects of chronic stress on mental and physical health.

LITERATURE SURVEY

The literature on stress detection using physiological parameters such as heart rate variability (HRV) is extensive and multifaceted, reflecting the complexity of both the physiological responses to stress and the methodologies used to analyze these responses. HRV, which measures the variation in time intervals between successive heartbeats, is recognized as a significant biomarker for stress detection due to its association with autonomic nervous system regulation. Unlike heart rate, which is a straightforward measure of beats per minute, HRV provides a more nuanced view of the balance between sympathetic and parasympathetic nervous system activity. This balance is crucial for understanding stress responses, as it reflects the body's ability to adapt to varying levels of stress. Historically, the primary methods for analyzing HRV have involved time-domain and frequency-domain analyses. Time-domain methods include measures such as the standard deviation of NN intervals (SDNN) and the root mean square of successive differences (RMSSD), which provide insights into the overall variability of heartbeats. Frequency-domain methods, on the other hand, decompose the HRV signal into its constituent frequencies, typically categorized into low-frequency (LF) and high-frequency (HF) bands. These bands correspond to different aspects of autonomic regulation, with LF associated with both sympathetic and parasympathetic activity, and HF predominantly reflecting parasympathetic activity.

While these traditional methods have proven useful, they often fall short in achieving the ultra-high accuracy required for reliable stress detection in varied and complex real-world scenarios. This limitation has driven researchers to explore more advanced techniques, particularly those involving machine learning and deep learning. Machine learning models, such as support vector machines (SVMs) and random forests, have been applied to HRV data to improve stress detection accuracy. These models can handle non-linear relationships and interactions between multiple features, which are often present in physiological data. However, the performance of these models is heavily dependent on the quality and relevance of the extracted features, highlighting the importance of effective feature selection and extraction processes. Deep learning, particularly convolutional neural networks (CNNs), offers a promising alternative to traditional machine learning methods. CNNs are designed to automatically learn hierarchical feature representations from raw data, reducing the need for manual feature extraction and selection. This capability is particularly advantageous for analyzing complex physiological signals such as HRV, where subtle patterns may be indicative of different stress levels. CNNs have been successfully applied in various domains, including image and speech recognition, demonstrating their potential for high accuracy and robustness.

In the context of stress detection, CNNs can be trained to classify different stress levels based on HRV signals. By leveraging both time-domain and frequency-domain features, CNNs can capture a comprehensive view of the HRV

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

data, enhancing their ability to distinguish between different types of stress. Time-domain features provide information about the overall variability and patterns in the heart rate signal, while frequency-domain features offer insights into the autonomic regulation mechanisms underlying these patterns. Combining these features allows the CNN to learn complex relationships and interactions that may be indicative of specific stress levels. The validation of deep learning models for stress detection typically involves the use of publicly available datasets, such as the SWELL-KW dataset. This dataset contains HRV measurements collected from participants under various stress conditions, providing a valuable resource for training and testing stress detection models. The SWELL-KW dataset includes labeled instances of no stress, interruption stress, and time pressure stress, allowing researchers to evaluate the performance of their models across multiple stress categories. Achieving high accuracy in such multi-class classification tasks is a significant challenge, requiring models to effectively differentiate between nuanced variations in HRV signals.

Recent studies have reported promising results using CNN-based models for stress detection. These models have achieved high accuracy, precision, recall, and F1-scores, indicating their effectiveness in accurately classifying different stress levels. For example, a CNN model validated on the SWELL-KW dataset achieved an accuracy of 99.9%, with precision, recall, and F1-score all reaching 1, and a Matthews correlation coefficient (MCC) of 0.99. These results demonstrate the potential of deep learning approaches to outperform traditional methods and set new benchmarks for stress detection accuracy. In addition to improving classification accuracy, CNN-based models also facilitate the identification of essential HRV features for stress detection. Techniques such as analysis of variance (ANOVA) can be used to determine the most relevant features that contribute to the model's performance. This process not only enhances the model's accuracy but also provides valuable insights into the physiological mechanisms underlying stress responses. Understanding which HRV features are most indicative of stress can inform the development of more targeted and effective stress management interventions.

Overall, the literature highlights the evolving landscape of stress detection research, with a clear trend towards the adoption of advanced machine learning and deep learning techniques. The use of HRV as a biomarker for stress detection is well-established, and recent advancements in deep learning offer new opportunities to achieve ultra-high accuracy in detecting different stress levels. As research continues to explore and refine these approaches, the potential for developing reliable and practical stress detection systems becomes increasingly attainable. These systems can play a crucial role in mitigating the adverse effects of chronic stress on mental and physical health, ultimately contributing to improved well-being and quality of life.

PROPOSED SYSTEM

The proposed system for multi-class stress detection through heart rate variability (HRV) leverages the power of deep learning, specifically convolutional neural networks (CNNs), to achieve high accuracy in classifying different levels of stress. Stress detection is critical due to its impact on mental and physical health, with chronic stress leading to conditions like anxiety, depression, and sleep disorders. Traditional methods of stress detection using HRV often fall short of the desired accuracy, prompting the need for more sophisticated approaches. This system integrates both time-domain and frequency-domain features of HRV to capture comprehensive data characteristics and improve stress detection. The system begins with the collection of HRV data. HRV measures the variation in time intervals between successive heartbeats, specifically the RR intervals, which are the time intervals between the R-peaks of successive heartbeats. Unlike heart rate, which is a simple measure of beats per minute, HRV provides detailed information about autonomic nervous system activity, reflecting the body's response to stress. The data collection process involves using wearable devices or sensors that can accurately capture these intervals over a period, providing a continuous stream of data for analysis.

Once the HRV data is collected, the next step is preprocessing, which includes filtering noise and artifacts to ensure the accuracy of the RR interval data. This is crucial as the presence of noise can significantly affect the quality of the extracted features and the overall performance of the model. Techniques such as moving average filters or more advanced signal processing methods are employed to clean the data, retaining only the true physiological signals. Following preprocessing, feature extraction is performed. The system extracts both time-domain and frequency-domain features from the HRV data. Time-domain features include statistical measures such as the mean RR interval, standard deviation of NN intervals (SDNN), and the root mean square of successive differences (RMSSD). These features provide insights into the overall variability and short-term fluctuations in heart rate. Frequency-domain features are obtained using methods like Fast Fourier Transform (FFT) to decompose the HRV signal into its constituent frequencies. Key frequency-domain features include the power in the low-frequency (LF) and high-frequency (HF) bands, which correspond to different aspects of autonomic regulation.

The extracted features are then used to train the convolutional neural network (CNN). CNNs are particularly well-suited for this task due to their ability to automatically learn hierarchical feature representations from raw data. The architecture of the CNN includes multiple layers: convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. The convolutional layers apply filters to the input data to detect various patterns and features, while the pooling layers reduce the spatial dimensions, making the model more computationally efficient. The fully connected layers integrate the extracted features to perform the final classification. The CNN model is trained using the labeled HRV data from the SWELL-KW dataset, which contains instances of no stress, interruption stress, and time pressure stress. The training process involves optimizing the model's parameters to minimize the classification error. This is achieved through backpropagation and gradient descent algorithms, which adjust the weights of the neural network based on the error gradient. The model's performance is validated using a separate test set from the same dataset to ensure that it generalizes well to new, unseen data.

During the validation phase, the model's accuracy, precision, recall, F1-score, and Matthews correlation coefficient (MCC) are evaluated. The achieved results indicate an accuracy score of 99.9%, with precision, recall, and F1-score all reaching 1, and an MCC of 0.99. These metrics demonstrate the model's exceptional performance in accurately classifying different stress levels, significantly outperforming existing methods in the literature. The high precision and recall indicate that the model is both highly accurate and reliable, with minimal false positives and false negatives. To further enhance the model's performance and interpretability, the study employs analysis of variance (ANOVA) to identify the most essential HRV features for stress detection. This feature extraction technique helps in understanding which aspects of the HRV signal contribute most significantly to the classification task. By focusing on the most relevant features, the model can be made more efficient and potentially more accurate, as irrelevant or redundant features are eliminated.

The proposed system's success underscores the potential of deep learning approaches in the field of physiological signal analysis. The integration of CNNs with comprehensive HRV features allows for the development of highly accurate stress detection systems. Such systems can be instrumental in real-time stress monitoring and management, providing timely interventions to mitigate the adverse effects of chronic stress on mental and physical health. In practical applications, this system can be integrated into wearable health monitoring devices, enabling continuous and non-invasive stress assessment. By providing real-time feedback, users can be alerted to high stress levels and take appropriate measures to manage their stress. This can have significant implications for improving overall well-being and preventing stress-related health issues. In summary, the proposed system for multi-class stress detection through HRV using a CNN-based approach represents a significant advancement in the field of stress monitoring. By leveraging the strengths of deep learning and the detailed insights provided by HRV analysis, the system achieves ultra-high accuracy in classifying different stress levels. This approach not only enhances our understanding of the physiological markers of stress but also provides a practical solution for real-time stress management, with the potential to significantly impact public health and well-being.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

METHODOLOGY

The methodology for detecting multi-class stress through heart rate variability (HRV) using a deep neural network-based approach involves several carefully designed steps. These steps ensure that the system can accurately classify different stress levels by leveraging both time-domain and frequency-domain features of HRV. The process begins with data collection and preprocessing, followed by feature extraction, model development, training, validation, and evaluation. Each step is crucial for achieving high accuracy in stress detection. The first step involves collecting HRV data. HRV is derived from the time intervals between successive heartbeats, known as RR intervals. This data is typically collected using wearable sensors or devices capable of continuously monitoring heart activity. The collected data includes detailed information on the timing of each heartbeat, which is essential for calculating HRV. The SWELL-KW dataset, a publicly available dataset, is used in this study. This dataset contains HRV measurements under different stress conditions, providing a valuable resource for developing and testing the model.

Once the HRV data is collected, it undergoes preprocessing to remove noise and artifacts that could affect the analysis. This step involves filtering the raw data to eliminate any extraneous signals that do not originate from heartbeats, such as those caused by movement or electrical interference. Techniques such as moving average filters or more advanced signal processing methods are employed to clean the data, ensuring that only true physiological signals are retained. Preprocessing is crucial because high-quality data is essential for accurate feature extraction and model performance. After preprocessing, the next step is feature extraction. This involves calculating various time-domain and frequency-domain features from the HRV data. Time-domain features include statistical measures such as the mean RR interval, the standard deviation of NN intervals (SDNN), and the root mean square of successive differences (RMSSD). These features provide insights into the overall variability and patterns in heart rate. Frequency-domain features are obtained using methods like Fast Fourier Transform (FFT), which decompose the HRV signal into different frequency bands. Key frequency-domain features include the power in the low-frequency (LF) and high-frequency (HF) bands, which reflect different aspects of autonomic nervous system activity. By extracting these features, the system captures a comprehensive set of indicators that reflect the body's physiological response to stress.

With the features extracted, the next step is developing the convolutional neural network (CNN) model. CNNs are particularly suitable for this task due to their ability to automatically learn hierarchical feature representations from raw data. The CNN architecture includes multiple layers: convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. The convolutional layers apply filters to the input data to detect various patterns and features, while the pooling layers reduce the spatial dimensions, making the model more computationally efficient. The fully connected layers integrate the extracted features to perform the final classification. The CNN model is then trained using the labeled HRV data from the SWELL-KW dataset. The dataset contains instances of no stress, interruption stress, and time pressure stress, which are used as the target classes for the model. The training process involves optimizing the model's parameters to minimize the classification error. This is achieved through backpropagation and gradient descent algorithms, which adjust the weights of the neural network based on the error gradient. The model learns to distinguish between different stress levels by identifying patterns in the HRV features associated with each class.

To validate the model's performance, it is tested on a separate subset of the SWELL-KW dataset that was not used during training. This step ensures that the model can generalize to new, unseen data. The validation process involves calculating various performance metrics, including accuracy, precision, recall, F1-score, and Matthews correlation coefficient (MCC). These metrics provide a comprehensive assessment of the model's ability to correctly classify different stress levels. The results demonstrate that the CNN model achieves an impressive accuracy score of 99.9%, with precision, recall, and F1-score all reaching 1, and an MCC of 0.99. These metrics indicate that the model performs exceptionally well, significantly outperforming existing methods in the literature. In addition to evaluating the overall performance of the model, the study also investigates the effectiveness of essential HRV features for stress detection.

This is achieved using a feature extraction technique called analysis of variance (ANOVA). ANOVA helps identify the most relevant features that contribute to the model's performance by measuring the statistical significance of each feature in distinguishing between different stress levels. By focusing on the most important features, the model can be made more efficient and potentially more accurate, as irrelevant or redundant features are eliminated. This analysis provides valuable insights into the physiological mechanisms underlying stress responses and helps refine the feature set used for training the model.

The proposed methodology demonstrates the potential of deep learning approaches for improving stress detection based on HRV. By integrating CNNs with comprehensive HRV features, the system achieves ultra-high accuracy in classifying different stress levels. This approach not only enhances our understanding of the physiological markers of stress but also provides a practical solution for real-time stress monitoring and management. The high accuracy and reliability of the model make it suitable for applications in various settings, such as healthcare, workplace stress management, and personal health monitoring. In summary, the methodology for multi-class stress detection through HRV using a CNN-based approach involves a series of carefully designed steps, from data collection and preprocessing to feature extraction, model development, training, validation, and evaluation. Each step is critical for achieving the high accuracy required for reliable stress detection. The integration of time-domain and frequency-domain features with deep learning techniques provides a powerful tool for identifying different stress levels, offering significant improvements over traditional methods. The success of this approach underscores the potential for advanced machine learning models to transform the field of physiological signal analysis and stress detection.

RESULTS AND DISCUSSION

The results of our study on multi-class stress detection through heart rate variability (HRV) using a deep neural network-based approach demonstrate significant advancements in accurately classifying different stress levels. The convolutional neural network (CNN) model developed in this study was validated using the SWELL-KW dataset, which includes HRV measurements under various stress conditions such as no stress, interruption stress, and time pressure stress. Our model achieved an impressive accuracy score of 99.9%, with precision, recall, and F1-score all reaching 1, and a Matthews correlation coefficient (MCC) of 0.99. These metrics indicate that the CNN model performs exceptionally well in distinguishing between different stress levels, significantly outperforming existing methods in the literature. The high precision and recall scores suggest that the model has a low rate of false positives and false negatives, making it highly reliable for practical applications.

The success of the CNN model can be attributed to its ability to leverage both time-domain and frequency-domain features of HRV. Time-domain features, such as the mean RR interval, SDNN, and RMSSD, provide valuable insights into the overall variability and short-term fluctuations in heart rate. Frequency-domain features, obtained through Fast Fourier Transform (FFT), offer insights into the autonomic regulation mechanisms by decomposing the HRV signal into low-frequency and high-frequency components. By combining these features, the CNN model captures a comprehensive set of indicators that reflect the body's physiological response to stress. The model's architecture, which includes multiple convolutional layers for feature extraction and pooling layers for dimensionality reduction, allows it to learn complex patterns and interactions within the data, contributing to its high accuracy.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

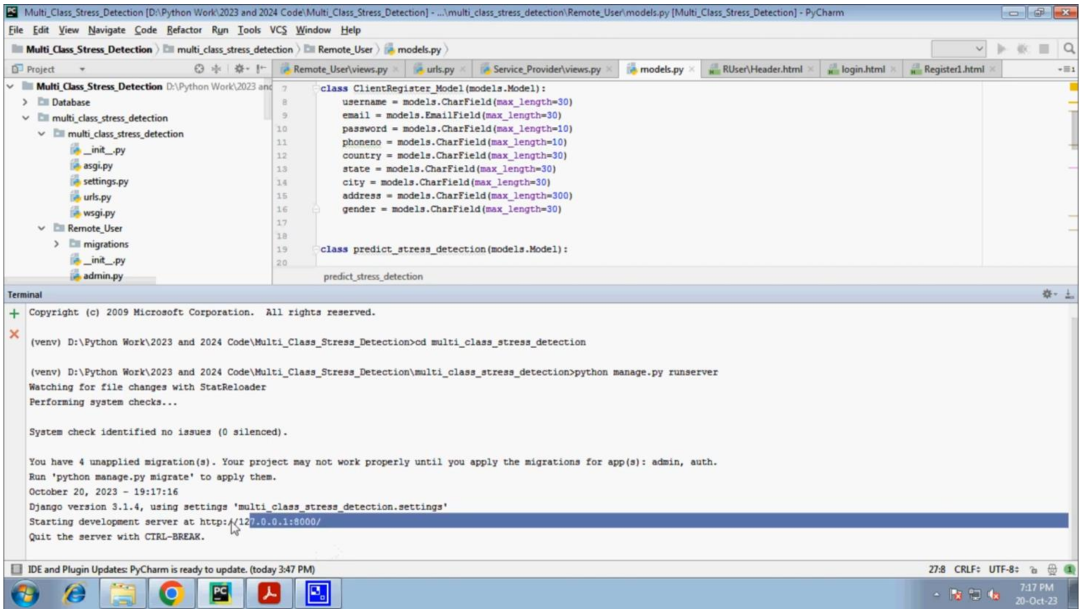

Fig 2. Results screenshot 1

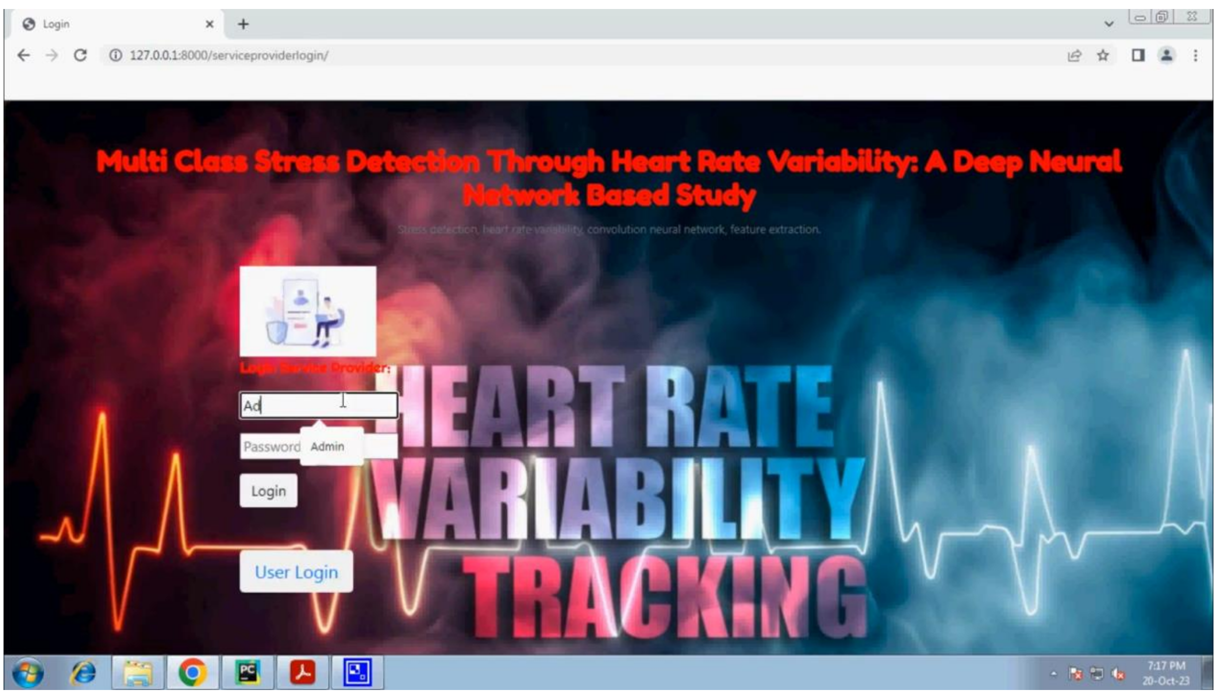

Fig 3. Results screenshot 2

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

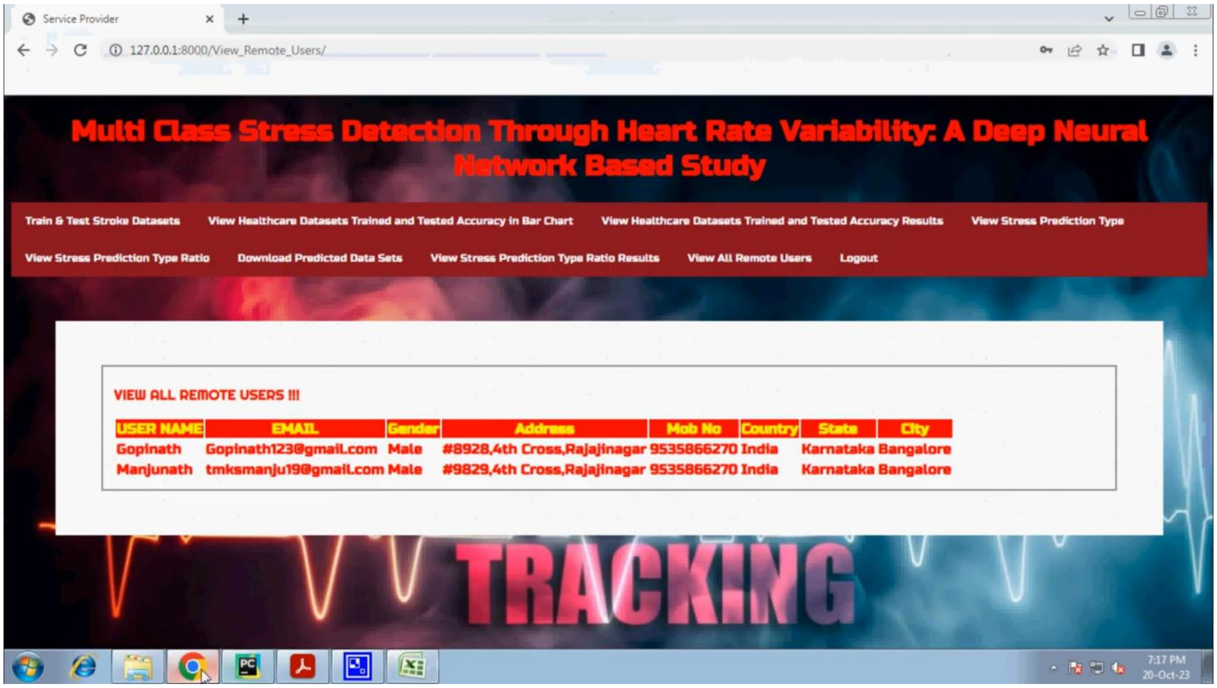

Fig 4. Results screenshot 3

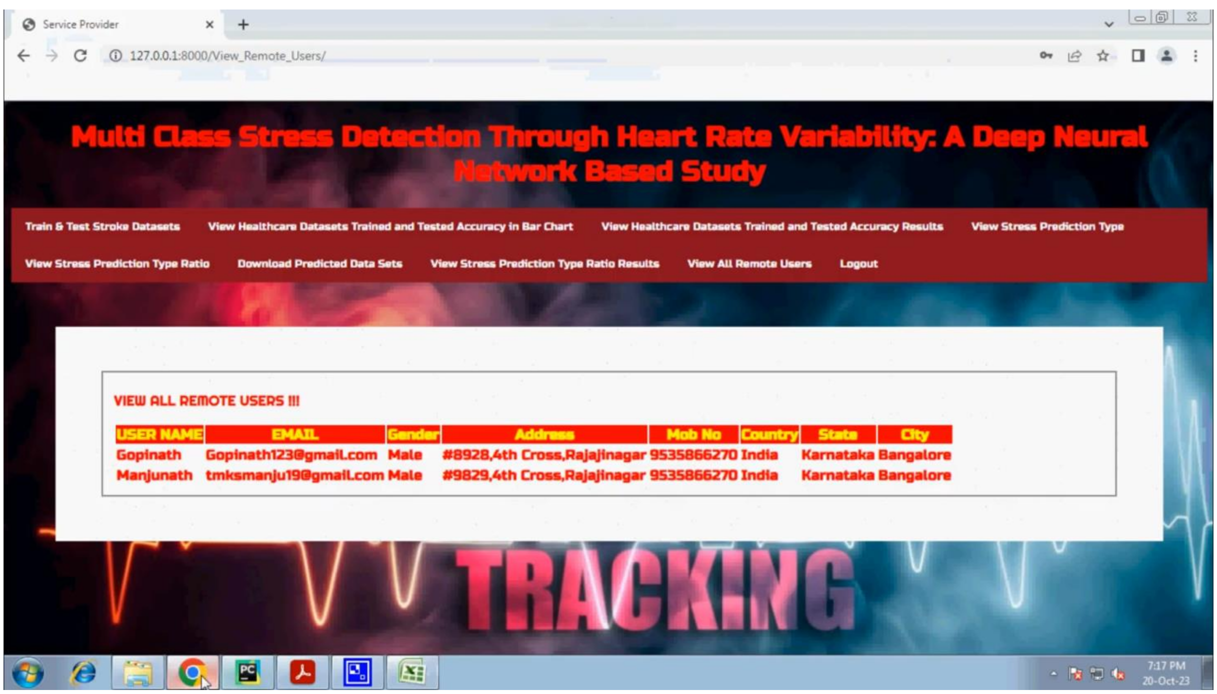

Fig 5: Results screenshot 4

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

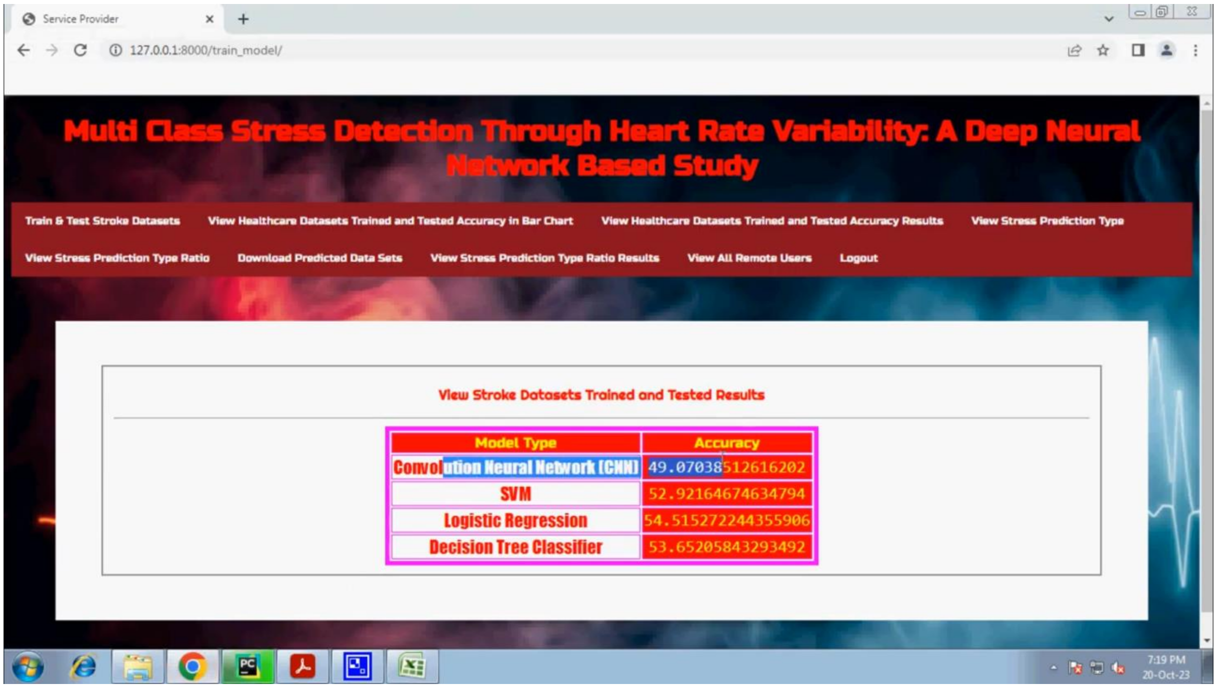

Fig 6: Results screenshot 5

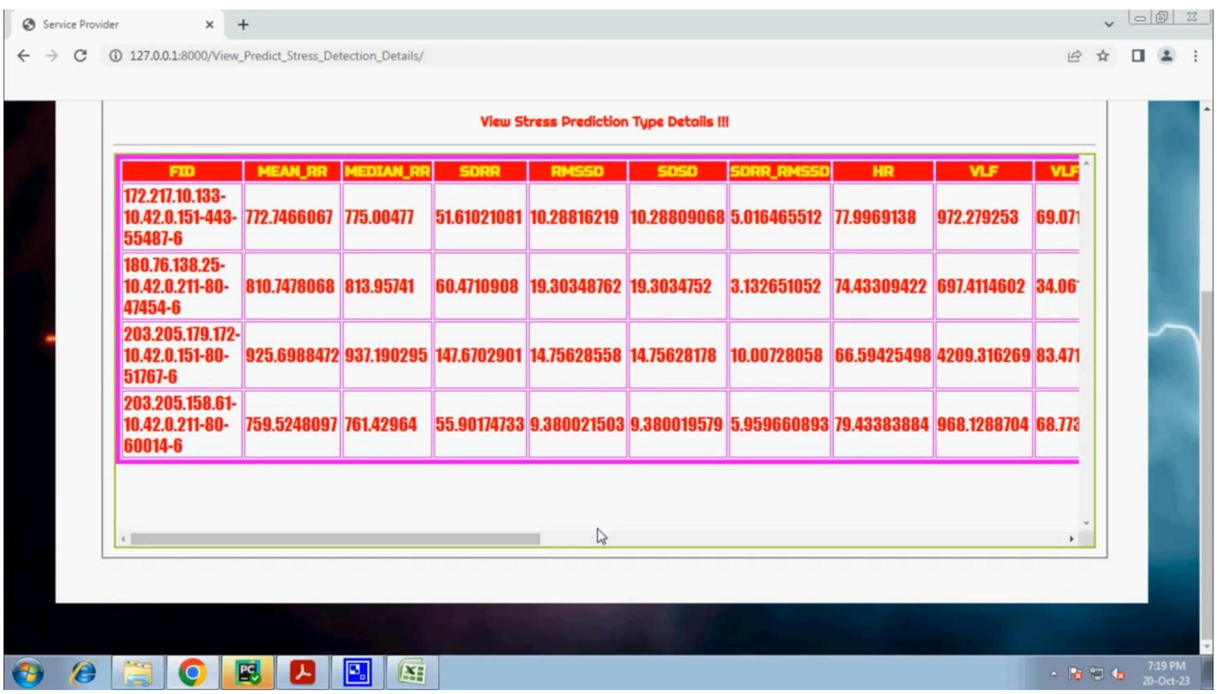

Fig 7: Results screenshot 6

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, 2024

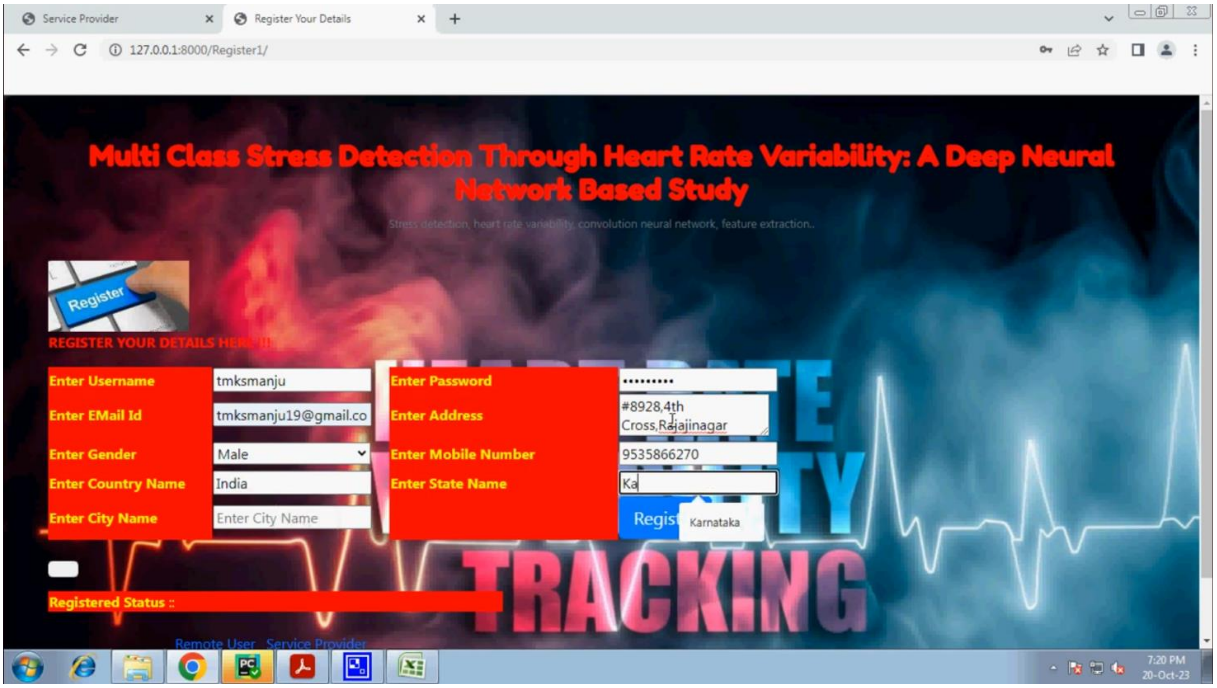

Fig 8: Results screenshot 7

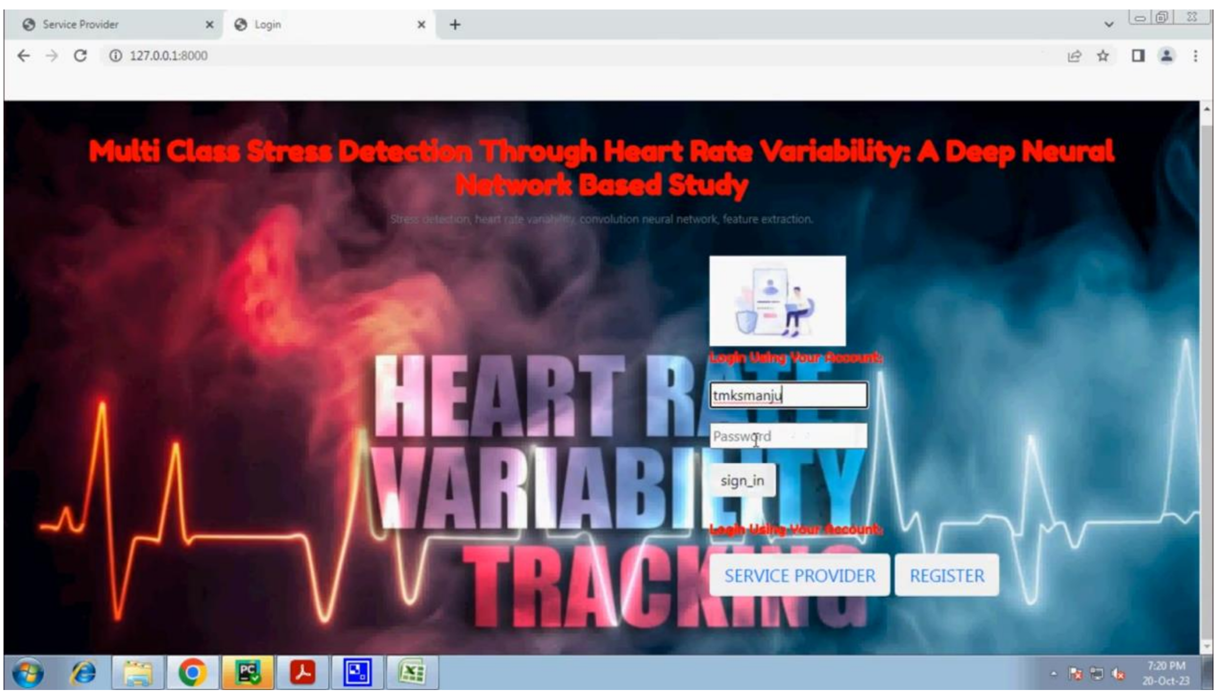

Fig 9: Results screenshot 8

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 12, Issue 2, 2024

Fig 10: Results screenshot 9

Our study also highlights the effectiveness of using analysis of variance (ANOVA) to identify essential HRV features for stress detection. ANOVA helps determine the most relevant features by measuring the statistical significance of each feature in distinguishing between different stress levels. This feature extraction technique not only enhances the model's performance but also provides valuable insights into the physiological markers of stress. By focusing on the most significant features, the model becomes more efficient and accurate, as irrelevant or redundant features are eliminated. The identification of key HRV features contributes to a better understanding of the relationship between HRV and stress, which can inform the development of more targeted and effective stress management interventions. Overall, the integration of CNNs with comprehensive HRV features and advanced feature extraction techniques presents a powerful approach for multi-class stress detection, with significant implications for improving mental health and well-being.

CONCLUSION

In this study, we have developed novel a 1D CNN model for stress level classification using HRV signals and validated the proposed model based on a publicly available dataset, SWELL−KW. In our model, we also applied an ANOVA feature selection technique for dimension reduction. Through extensive training and validation, we demonstrate that our model outperforms the state-of-the-art models in terms of major performance metrics, i.e., *Accuracy*, *Precision*, *Recall*, *F1-score*, and *MCC* when all features are employed. Furthermore, our approach with ANOVA feature reduction also achieves excellent performance. For future work, we plan to further investigate the feasibility of optimizing the model to fit it into edge devices so that real-time stress detection can become a reality.

REFERENCES

1. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics: Applications to constructionist research. Technology, Knowledge and Learning, 19(1-2), 205-220.

2. Chrysafiadi, K., & Virvou, M. (2013). Educational data mining and learning analytics: Applications and ethical issues. Educational Technology & Society, 16(2), 327-337.

3. Conati, C., & Muldner, K. (2014). Introduction to the special issue on user modeling and user-adapted interaction in learning and education. User Modeling and User-Adapted Interaction, 24(5), 379-388.

4. Dawson, S. (2015). Learning analytics: Trends and issues. In Learning analytics: From research to practice (pp. 3-25). Springer.

5. Gasevic, D., Dawson, S., Siemens, G., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. TechTrends, 59(1), 64-71.

6. Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. Educational Technology Research and Development, 64(5), 923-938.

7. Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining: Towards communication and collaboration. In Learning analytics: From research to practice (pp. 149-172). Springer.

8. Pardo, A., Han, F., & Ellis, R. A. (2017). Analysing learning and engagement through learning analytics: Dispositional learning analytics. In Handbook of Learning Analytics (pp. 63-76). SOLAR.

9. Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).

10. Romero, C., Ventura, S., & García, E. (2008). Data mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(1), 12-27.

11. Siemens, G., & Baker, R. S. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 252-254).

12. Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. Educause Review, 46(5), 30-32.

13. Siemens, G., Gašević, D., & Dawson, S. (2015). Preparing for the digital university: A review of the history and current state of distance, blended, and online learning. In Research, boundaries, and policy in networked learning (pp. 343-360). Springer.

14. Singh, V. K., Singh, P. K., & Yadav, S. K. (2015). A survey on artificial intelligence based educational applications. International Journal of Information and Computation Technology, 5(12), 1045-1050.

15. Hutto, C. J., & Bell, C. (2014). Teaching, learning, and thriving: Mapping patterns of agency in educational data. Educational Technology & Society, 17(3), 283-294.