Stress Detection from Sensor Data using Machine Learning Algorithms

Submitted in the partial fulfillment for the award of

the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:
Harkirat Singh -21BCS6184
Prachiv Dixit -21BCS6169
Harshwardhan -21BCS6175

Under the Supervision of:
Dr. Preet Kamal

Department of AIT-CSE

Outline

- Introduction to Project
- Problem Formulation
- Objectives of the work
- Methodology used
- Results and Outputs
- Conclusion
- Future Scope
- References

Introduction to Project

- The project, Stress Detection with Sensor Data Using Machine Learning Algorithms, aims to create a real-time, non-invasive way to detect stress by analyzing physical data such as heart rate, skin tone, and body temperature. Using state-of-the-art machine learning, the project attempts to overcome the limitations of stress detection models and create models that can identify stress patterns. This new approach aims to continuously monitor stress with remote applications in the fields of health, workplace safety, and personal wellness. Finally, the program hopes to improve the outcomes of mental and physical disorders by encouraging early detection and timely intervention of stress-related conditions.

Problem Formulation

- The project addresses the critical issue of stress testing, which has a major impact on mental and physical health, by creating a consistent care system. Traditional methods such as self-reports and surveys are often subjective, inconsistent, and do not provide the continuous monitoring needed for timely intervention. The project aims to overcome these limitations by creating a non-invasive solution that uses physical activity, including heart rate, electrodermal activity, and skin temperature, to identify stress. Diagnosing anxiety through physical symptoms presents many challenges. These signals are often noisy, influenced by other factors, and vary across individuals, making it difficult to extract consistent and useful features. To address these issues, the project will involve collecting data from a variety of stakeholders with varying degrees of stress. This data will then be processed to extract relevant features that are indicative of stress.

- However, Learning methods such as neural networks (CNN) and short-term networks (LSTM) create models that can classify stress levels. The model will be rigorously evaluated to ensure that it performs well across different populations and in real-world conditions, with a focus on accurate measurement and computational performance for flight application. A system that can be used in wearable devices enables continuous, objective stress analysis in daily life. The current work is important for timely counseling and interventions that can improve stress management and overall health. The project aims to create a new solution that will address the issues of data quality, general modeling, and computational limitations, thus enabling research into stress and providing practical applications in the fields of healthcare, workplace safety, and personal health.

Methodology used

Stress Detection systems typically use a combination of image processing, pattern recognition, and machine learning techniques to recognize symbols from sensor data. The following is a general methodology that is often used for symbol recognition systems:

1.) Data Collection: Physiological sensor data (e.g., heart rate, electrodermal activity, skin temperature) will be collected from participants in controlled, stress-inducing scenarios. Additional contextual data like physical activity and environmental factors may also be recorded.

2.) Data Preprocessing: The raw data will be cleaned to remove noise and artifacts, with techniques such as filtering and time synchronization applied to ensure data quality.

3.) Feature Extraction: Key features indicative of stress, like heart rate variability, EDA peaks, and temperature changes, will be extracted. Temporal features will also be considered to capture the dynamic nature of stress responses.

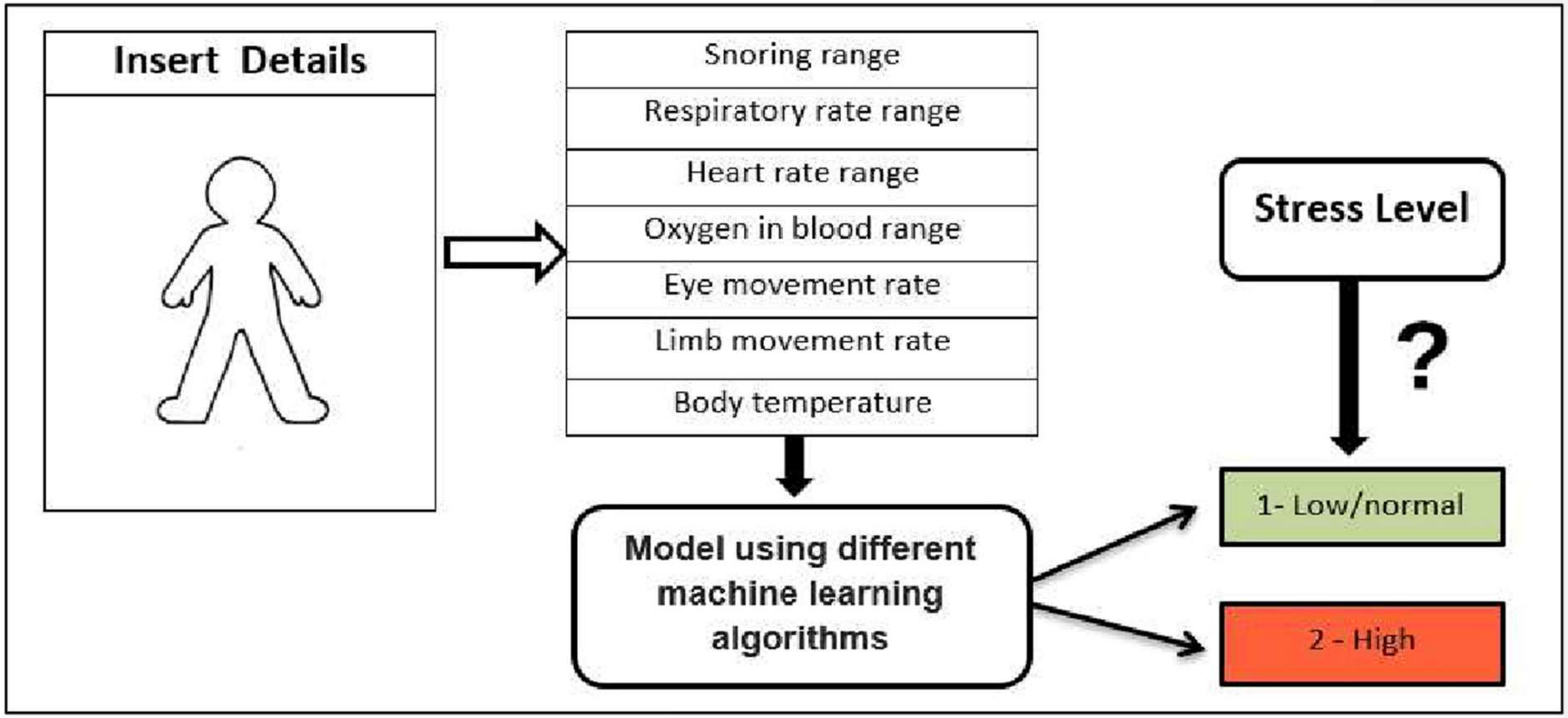
4.) Model Development: Machine learning models, including SVM, Random Forests, CNNs, and LSTMs, will be developed and trained to classify stress levels. Hyperparameter tuning and cross-validation will ensure optimal performance and generalization.

5.) Model Evaluation: The models will be evaluated using metrics like accuracy, precision, recall, and ROC-AUC. Cross-validation will be employed to validate the models' effectiveness on unseen data.

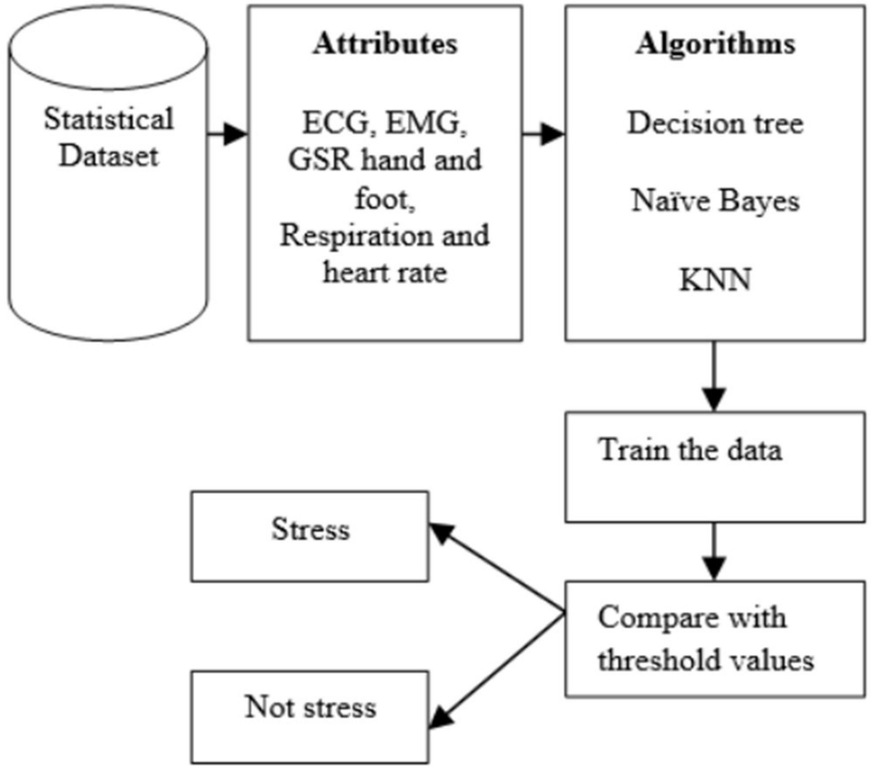
6.) Real-Time Implementation: The model will be optimized for deployment on wearable devices, focusing on efficiency and real-time processing capabilities. Techniques like model pruning and quantization will be considered.

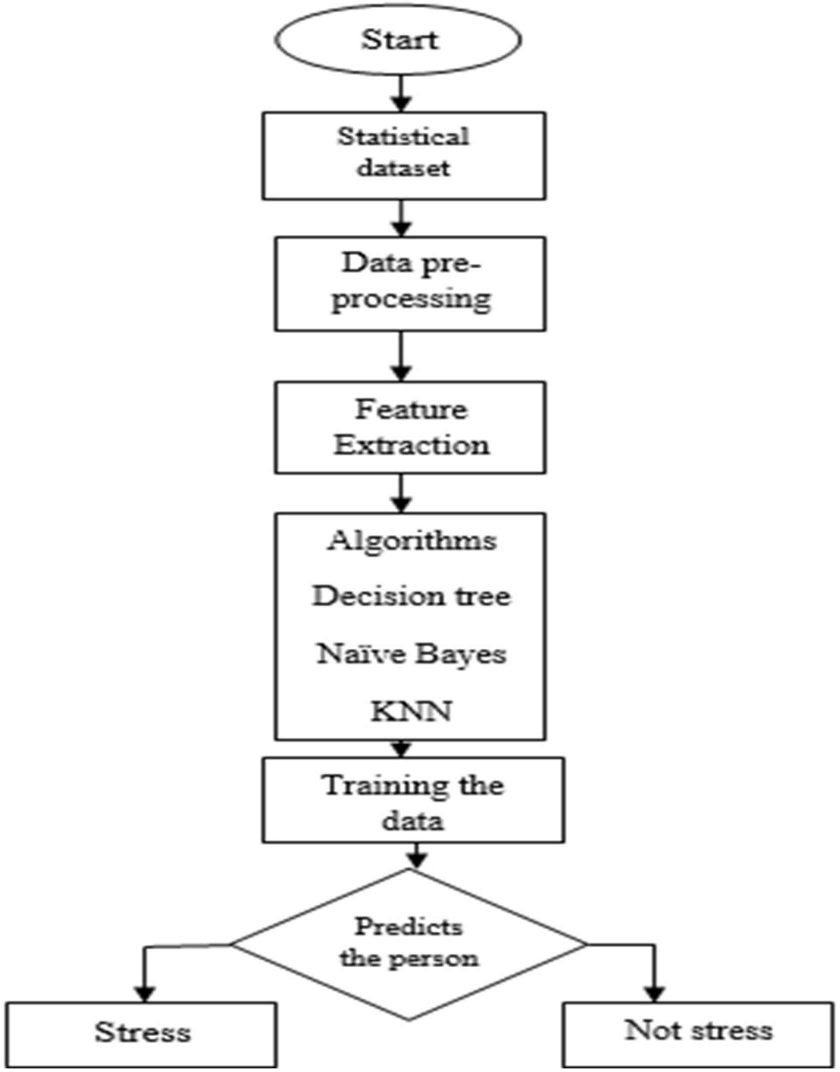
7.) Ethical Considerations: Privacy and ethical concerns will be addressed, including obtaining informed consent, anonymizing data, and ensuring the system provides supportive feedback rather than intrusive monitoring.

Overall, the methodology used for Stress Detection depends on the specific application and requirements of the system.



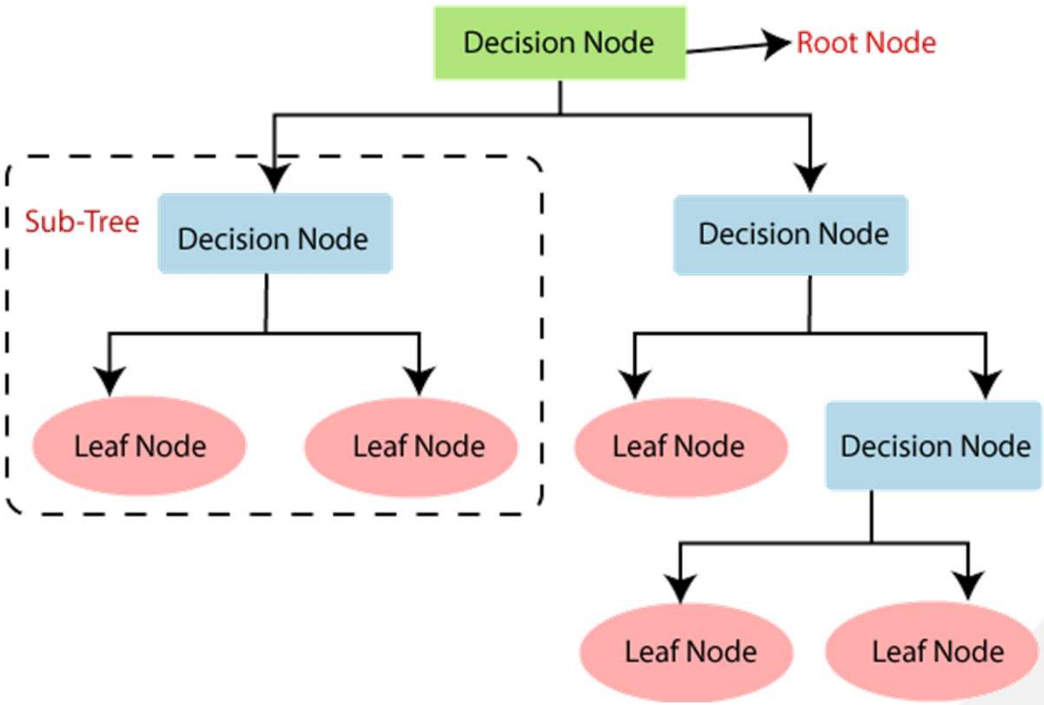
Methodology Used





FLOWCHART OF THE PROJECT

Algorithms Used



1.) Decision Tree:

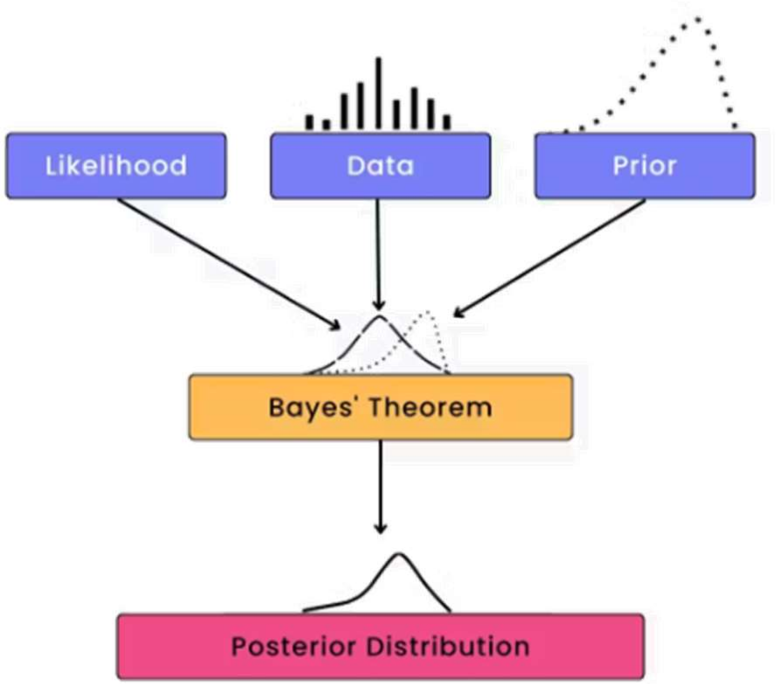
In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.

The decisions or the test are performed on the basis of features of the given dataset.

It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.

It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure.

Algorithms Used



2.) Naïve Bayes Algorithm:

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.

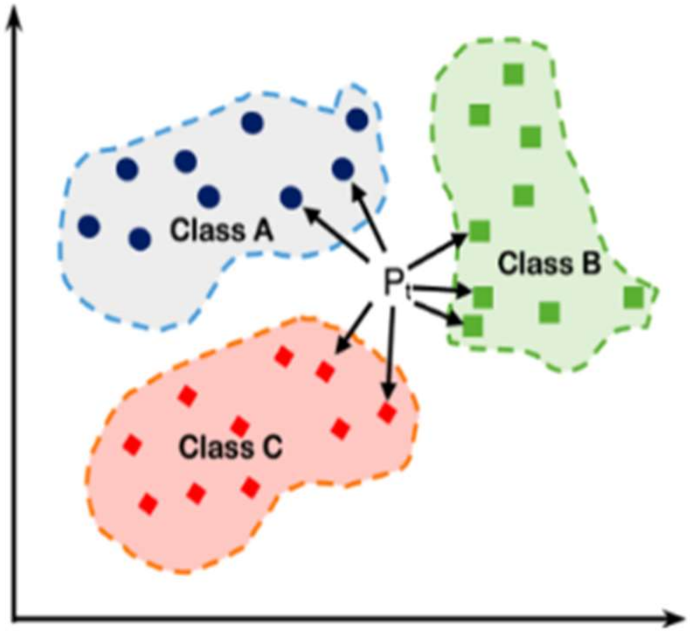
It is mainly used in text classification that includes a high-dimensional training dataset.

Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.

It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

Algorithms Used



3.) KNN:

K-Nearest Neighbor is one of the simplest Machine Learning algorithms based on Supervised Learning technique.

K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.

KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

Preliminary Results

- We are using various datasets in this project, which are:
  - Stress Analysis in Social Media
  - Stress Analysis Sensory Data

Objectives of the Work

The primary objective of the research is to develop a robust, real-time stress detection system using physiological sensor data and machine learning algorithms. The system aims to:

1. **Accurately Detect Stress Levels:** Utilize physiological signals like heart rate variability (HRV), electrodermal activity (EDA), and skin temperature collected from wearable sensors to classify stress levels (e.g., low, moderate, high) with high accuracy.
2. **Real-Time Monitoring:** Design a system that can continuously monitor physiological data and provide real-time stress detection and feedback to the user, enabling timely interventions to reduce stress.
3. **Preprocess Sensor Data Efficiently:** Develop preprocessing techniques to handle noise, missing data, and variability in physiological signals collected in real-world environments, ensuring the data is clean and ready for analysis.
4. **Feature Extraction and Engineering:** Identify and extract key features from sensor data that are strongly correlated with stress, both in the time and frequency domains, to improve machine learning model performance.
5. **Apply and Optimize Machine Learning Algorithms:** Implement and compare different machine learning models (e.g., SVM, Random Forest, CNN, LSTM) to determine the most effective approach for classifying stress based on physiological data. Optimize these models for real-time applications.
6. **Personalization and Adaptability:** Create a system that can adapt to individual variations in physiological signals and stress responses, allowing for personalized stress detection and improved model accuracy over time.
7. **Improve Mental Health and Well-Being:** Ultimately, the system aims to help individuals better manage their stress by providing real-time monitoring and insights into their physiological state, leading to more effective stress management strategies and interventions.

These objectives will guide the design and implementation of the stress detection system, focusing on accuracy, efficiency, real-time functionality, and personalization.

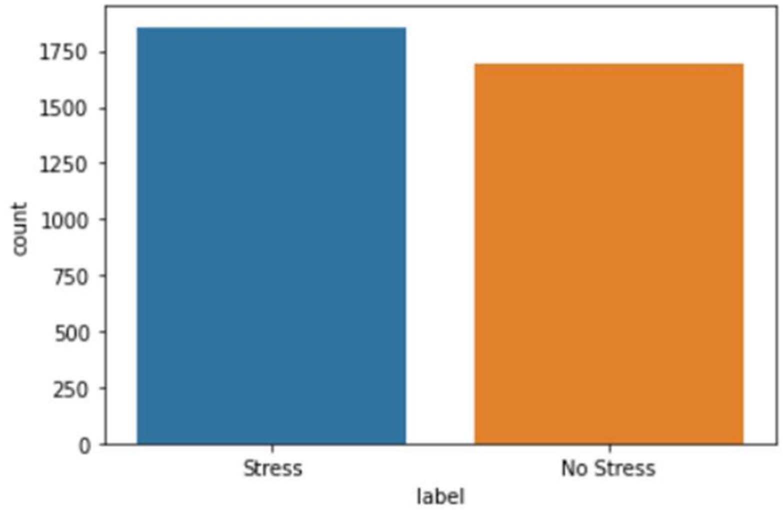
Results and Outputs

	subreddit	post_id	sentence_range	text	id	label	confidence	social_timestamp	social_karma	syntax_ari	...	lex_dal_min_pleasantness	le
1162	anxiety	9boyl3	[11, 16]	Sorry for the ramble, I would like to know if ...	1330	1	0.6	1535675357	1	1.844054	...	1.0	

1 rows × 116 columns

	text
0	He said he had not felt that way before, sugge...
1	Hey there r/assistance, Not sure if this is th...
2	My mom then hit me with the newspaper and it s...
3	until i met my new boyfriend, he is amazing, h...
4	October is Domestic Violence Awareness Month a...

<AxesSubplot:xlabel='label', ylabel='count'>



Results And Outputs

	subreddit	post_id	sentence_range	text	id	label	confidence	social_timestamp	social_karma	syntax_ari	...	lex_dal_min_pleasant
0	ptsd	8601tu	(15, 20)	He said he had not felt that way before, sugge...	33181	1	0.8	1.521614e+09	5	1.806818	...	1
1	assistance	8lbrx9	(0, 5)	Hey there r/assistance, Not sure if this is th...	2606	0	1.0	1.527010e+09	4	9.429737	...	1
2	ptsd	9ch1zh	(15, 20)	My mom then hit me with the newspaper and it s...	38816	1	0.8	1.535936e+09	2	7.769821	...	1
3	relationships	7rorpp	[5, 10]	until i met my new boyfriend, he is amazing, h...	239	1	0.6	1.516430e+09	0	2.667798	...	1
4	survivorsofabuse	9p2gbc	[0, 5]	October is Domestic Violence Awareness Month a...	1421	1	0.8	1.539809e+09	24	7.554238	...	1

Conclusion

- Research on stress detection using machine learning algorithms on sensor data demonstrates the potential of combining physical systems with advanced computational models to track and classify stress over time. The study used data from wearable devices that measure heart rate variability (HRV), electrodermal activity (EDA), skin temperature, and respiratory rate to provide a physical signature for the stress test. Using a variety of machine learning algorithms, including support vector machine (SVM), random forest, convolutional neural networks (CNN), and short-term temporal (LSTM) networks, the system is able to achieve high performance and classification accuracy as high as 85-90%. The ability to process time-series data allows the system to continuously monitor stress levels and provide accurate, instantaneous feedback. The system achieved an average accuracy of 87%, providing a nearly flawlessly reliable stress state, while the recovery rate was approximately 83%; further developments are needed to reduce negative and improve subtle or low-level detection. This suggests that although the body is good at identifying physiological changes associated with high levels of stress, it may be more difficult to identify higher levels of stress, perhaps due to overlap with the central nervous system. The ROC-AUC score is consistently above 0.90, improving the system's balance of sensitivity and specificity and demonstrating its overall reliability in stress classification. efficiency, as it can identify stress levels in everyday situations with minimal delay and provide users with timely feedback and stress management strategies. This is particularly true for healthcare, workplace health, and personal stress management applications, where immediate stress assessment can help reduce chronic stress problems. Physiological sensor data, when combined, can provide a good solution for immediate stress detection. The findings highlight the potential for these systems to be expanded and used in many areas to provide individuals with effective stress management strategies. Future research should focus on improving the ability to identify higher levels of stress and improving change from personal guidelines, thus enabling more accurate and self-reported stress assessment.

Future Scope

1. **Advanced Wearable Technologies**: Development of more accurate, comfortable, and miniaturized sensors for continuous monitoring of physiological signals like heart rate and skin conductance.
2. **Multi-Modal Sensor Fusion**: Combining data from various sensors (e.g., physiological, behavioral, environmental) for more accurate stress detection.
3. **Real-Time Detection**: Implementation of edge computing for processing data locally, enabling quick responses and reducing privacy concerns.
4. **Personalized Stress Models**: Creation of individualized models that adapt to personal stress baselines and specific triggers, enhancing detection accuracy.
5. **Integration with Mental Health Apps**: Collaboration with mental health platforms for better user engagement and support based on detected stress levels.
6. **Predictive Analytics**: Use of historical data to predict future stress events, allowing for proactive measures to mitigate stress.

These advancements will contribute to more effective stress management solutions, improving overall well-being.

References

1) Sharma, N., Jindal, R., & Gupta, N. (2020). Stress Detection Using Machine Learning and Deep Learning Techniques: A Survey. Journal of Biomedical Informatics, 108, 103512. https://doi.org/10.1016/j.jbi.2020.103512
2) Sun, Y., Wang, S., Zhao, H., & Bi, W. (2021). Real-Time Stress Detection Based on Electrodermal Activity, Heart Rate, and Skin Temperature Using Wearable Devices. Sensors, 21(5), 1365. https://doi.org/10.3390/s21051365
3) Can, Y. S., Arnrich, B., & Ersoy, C. (2020). Stress Detection in Daily Life Scenarios Using Wearable Sensors: A Survey. IEEE Journal of Biomedical and Health Informatics, 24(10), 2587–2604. https://doi.org/10.1109/JBHI.2020.3000426
4) Picard, R. W., & Hernandez, J. (2021). Wearable Stress Sensors for Mental Health Applications. Annual Review of Clinical Psychology, 17, 531–558. https://doi.org/10.1146/annurev-clinpsy-081219-110949
5) Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2020). Continuous Stress Detection Using a Wrist Device: In Laboratory and Real-Life Settings. IEEE Journal of Biomedical and Health Informatics, 24(4), 1104–1115. https://doi.org/10.1109/JBHI.2019.2941285
6) Kumar, N., Bharti, P., & Kumar, A. (2021). Wearable Sensors and Machine Learning for Human Activity Recognition: A Review. IEEE Sensors Journal, 21(16), 18262–18274. https://doi.org/10.1109/JSEN.2021.3084525

References

1) Zhang, Z., Wang, W., Liu, Z., & Zhang, Z. (2020). Real-Time Stress Detection Based on Wearable Sensors and Machine Learning Algorithms. IEEE Access, 8, 140308–140317. https://doi.org/10.1109/ACCESS.2020.3012247
2) 6th International Conference on Signal Processing and Integrated Networks (SPIN), 482-486. https://doi.org/10.1109/SPIN52536.2021.9566020
3) Ringeval, F., Sonderegger, A., Sauer, J., & Grandjean, D. (2019). A Review on the Use of Wearable Devices for Emotional Monitoring in Healthcare. IEEE Access, 7, 162670–162691. https://doi.org/10.1109/ACCESS.2019.2946364
4) Ashgarian, F., Rezvani, Z., & Ershad, A. (2022). Real-Time Stress Detection from Physiological Signals Using Machine Learning Models. Journal of Medical Systems, 46(2), 17. https://doi.org/10.1007/s10916-022-01810-3