Predictive Modelling of Human Stress Using Sleep-Related Features and Machine Learning

SUGGUNA SRAVANI[1], Mr SK SUBHANI[2]

[1]PG Scholar, Dept. of AI&ML, St. Marys Group of Institutions Guntur for Women, Guntur.

[2]Associate Professor, Dept. of AI&ML, St. Marys Group of Institutions Guntur for Women, Guntur.

Abstract: Psychological stress has become a major concern in modern society, affecting physical health, work efficiency, and overall quality of life. However, stress often remains undetected because of limited awareness, social stigma, and inadequate access to mental healthcare services. Since sleep patterns are closely associated with emotional and psychological conditions, sleep-related characteristics can be used as reliable indicators for early stress detection. This study presents a machine learning-based approach to predict human stress levels using sleep-related features such as sleep duration, sleep quality, bedtime consistency, frequency of awakenings, and sleep disturbances. The collected data are preprocessed and analyzed using various machine learning algorithms, including Support Vector Machine (SVM), Decision Tree, Random Forest, and Naïve Bayes. Deep learning models are also explored to identify complex relationships and hidden patterns within sleep behavior data. The performance of each model is evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results show that the Naïve Bayes classifier achieves better prediction accuracy than the other models. The proposed system provides an efficient and automated framework for early stress identification, enabling individuals to monitor their mental well-being through sleep analysis and promoting timely intervention for improved stress management and preventive healthcare.

Key Words: Stress Detection, Sleep Pattern Analysis, Machine Learning, Naïve Bayes, Mental Health Monitoring, Sleep Quality, Stress Prediction.

1. Introduction

Stress has emerged as a significant health issue in today's society, influencing emotional well-being, physical health, cognitive abilities, and overall quality of life. Rapid lifestyle changes, increasing work demands, irregular routines, and excessive use of technology have contributed to higher stress levels among individuals. Sleep is an essential component of human health, and disturbances in sleep patterns are strongly associated with stress, anxiety, and various psychological disorders. Sleep-related factors such as sleep duration, sleep quality, sleep efficiency, and nighttime awakenings can provide valuable information about an individual's stress condition.

Conventional methods for stress assessment primarily depend on self-report questionnaires, clinical examinations, and psychological interviews. Although these techniques are widely used, they often suffer from subjectivity, require considerable time, and may be affected by social stigma and limited mental health awareness. Consequently, many people experience delayed diagnosis or remain unaware of their stress conditions.

Recent advancements in machine learning have enabled the analysis of large amounts of sleep data collected through wearable sensors and sleep-monitoring devices. By evaluating parameters such as sleep duration, sleep quality, bedtime regularity, and sleep disturbances, machine learning algorithms can identify patterns related to stress. This study aims to develop a machine learning-based stress prediction system using sleep behavior analysis to support early detection, continuous monitoring, and effective mental health management.

2. Literature Survey

1.Akter, S., Sharif, M. M., & Mahmud, M. et al (2020) In their study "Stress Detection Based on Sleep Pattern Analysis Using Machine Learning", Akter et al. investigate how sleep-related features such as duration, interruptions, and variability can help in predicting human stress using ML approaches. They collected sleep quality metrics and associated stress labels, and applied algorithms like Random Forest and Support Vector Machines to model the relationship between sleep patterns and stress levels. Their results show that ML models can discern underlying patterns that correlate disrupted sleep with high stress, achieving promising accuracy in classification tasks. This work underscores the potential of non-invasive sleep data in automatic stress assessment for early detection and personal health monitoring, emphasizing the practical usefulness of wearables and sleep trackers for mental well-being.

2. Gupta, A., Jain, S., & Verma, A. et al (2021) Gupta, Jain, and Verma conducted a survey on stress detection using sleep pattern analysis, compiling methods that leverage machine learning to interpret sleep behaviors for stress predictions. The authors critically reviewed various ML classifiers—including Support Vector Machines, Decision Trees, and Naïve Bayes—and discussed how feature extraction from sleep duration, sleep fragmentation, and awakenings affects model performance. Their literature review revealed that combining sleep metrics with ML techniques enhances predictive accuracy and enables more reliable stress inference. They also identified research gaps, such as the need for larger datasets and standardized evaluation protocols to improve real-world applicability. This work provides a comprehensive foundation for designing future sleep-based stress detection systems.

3. Proposed System

The proposed system aims to predict human stress levels using sleep-related and physiological parameters through the Naïve Bayes algorithm. Important attributes considered in this study include snoring rate, respiration rate, body temperature, limb movement, blood oxygen level, eye movement, sleeping duration, and heart rate. These parameters provide valuable information about an individual's physical and mental condition.

Initially, the collected dataset undergoes preprocessing, which includes handling missing values and normalizing the feature values to enhance the model's performance. After preprocessing, the dataset is divided into training and testing sets to evaluate the effectiveness of the classifier.

A Gaussian Naïve Bayes model is then trained using the processed data to identify the relationship between sleep characteristics and stress levels. During the prediction phase, the algorithm computes the probability of each stress category and assigns the class with the highest probability. The system categorizes individuals into three stress levels: Low/Normal Stress, Medium Stress, and High Stress.

Because of its simplicity, lower computational complexity, and faster execution, the Naïve Bayes algorithm offers efficient stress prediction. The proposed approach supports early stress detection through sleep analysis and can be incorporated into wearable devices and mobile healthcare applications for real-time stress monitoring and management.

4. System Architecture

The architecture of a mental health prediction system consists of several interconnected components that work together to analyze data and identify potential mental health conditions. The system begins with a data collection module, which gathers information from multiple sources such as questionnaires, surveys, wearable devices, and other monitoring platforms.

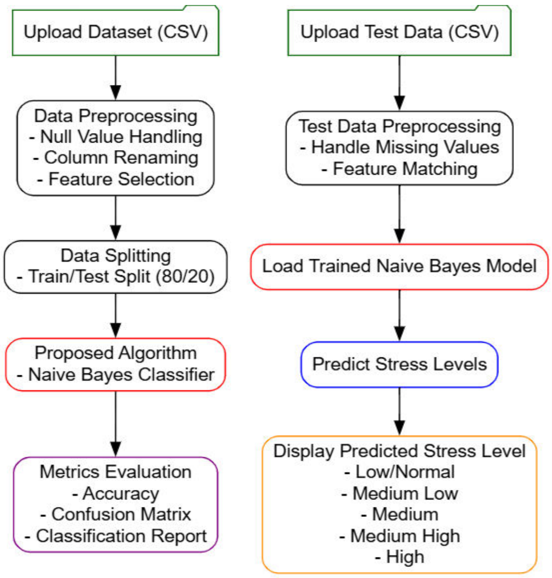

Fig 1: System Architecture

The collected data undergoes a preprocessing stage where missing values, inconsistencies, and noise are removed, and the data is normalized to improve quality and reliability. After preprocessing, machine learning techniques are applied to develop predictive models using historical and behavioral data. These models analyze patterns related to sleep behavior, physiological conditions, lifestyle factors, and environmental influences to predict mental health status.

The system architecture generally includes a user interface that allows users to enter data and receive feedback. A data processing layer manages data preparation and feature extraction, while the prediction engine performs the analysis and generates stress predictions. The output module presents the results in the form of stress levels, alerts, or recommendations.

5. Methodology

The stress detection process begins with collecting sleep-related and physiological data such as sleeping hours, heart rate, respiration rate, body temperature, and blood oxygen levels. The data is then preprocessed by handling missing values and encoding categorical features. After splitting the dataset into training and testing sets, a Naive Bayes classifier is trained to learn the relationship between sleep patterns and stress levels. The trained model predicts stress categories for new data, enabling efficient and early stress detection.

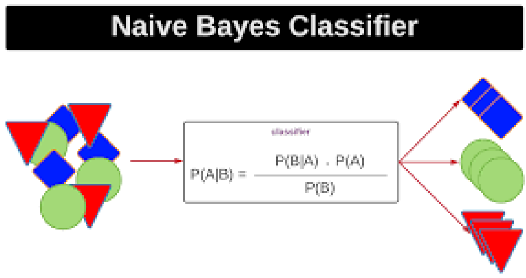

Fig 2: Naive Bayes classifier

The system works by feeding preprocessed sleep and physiological data into the trained Naive Bayes model. The model calculates the probability of each stress category and predicts the class with the highest probability, such as Low, Medium, or High Stress. This enables fast and efficient stress prediction based on sleeping habits. The model's performance is evaluated using metrics such as accuracy and a confusion matrix to measure prediction correctness and analyze classification errors.

6. Design and Construction

The proposed Human Stress Detection System follows a modular architecture that includes data collection, pre-processing, feature extraction, machine learning, and prediction stages. It uses sleep-related parameters such as sleep duration, efficiency, latency, and awakenings to provide a non-invasive and continuous method for stress detection.

i) Data Collection: Sleep-related data is collected from publicly available datasets and wearable devices such as smart bands or sleep-tracking applications. The dataset includes parameters such as total sleep duration, sleep efficiency, sleep quality score, number of awakenings, time taken to fall asleep (sleep latency), and circadian rhythm consistency. Each data sample is labeled with corresponding stress levels (e.g., Low, Moderate, High).

ii) Data Preprocessing: Raw sleep data often contains missing values, noise, and inconsistencies. Therefore, preprocessing techniques such as data cleaning, normalization, and feature scaling are applied. Min–Max normalization is used to scale features into a standard range:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

iii) Feature Extraction and Selection: Relevant sleep parameters are selected based on correlation analysis and feature importance techniques. This step reduces dimensionality and improves model efficiency. Extracted features form the input vector for classification models.

iv) Model Development: involves implementing multiple machine learning classifiers, including Support Vector Machine (SVM), Decision Tree, Random Forest, and Naïve Bayes. The Naïve Bayes classifier selects the optimal decision boundary based on probability estimation defined by Bayes' theorem.

v) Model Evaluation

The dataset is divided into training and testing sets. Performance metrics such as Accuracy, Precision, Recall, and F1-score are used to evaluate the model:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

vi) Stress Prediction System: The trained model is integrated into a stress monitoring system that classifies individuals into stress categories based on real-time sleep data. The system provides early warnings and supports timely intervention for improved mental health management.

a classification report is generated, which includes precision, recall, and F1-score for each class, indicating the model's ability to correctly identify specific stress levels. the evaluation ensures that the Naive Bayes algorithm reliably predicts stress based on sleep-related features, and highlights areas where predictions may need improvement, providing a foundation for further optimization or integration with other machine learning models.

7. Results and Discussion

The proposed system effectively detects stress levels by analyzing sleep-related and physiological features using machine learning techniques. Results show that poor sleep quality, irregular sleep patterns, and frequent awakenings are strongly associated with higher stress levels.

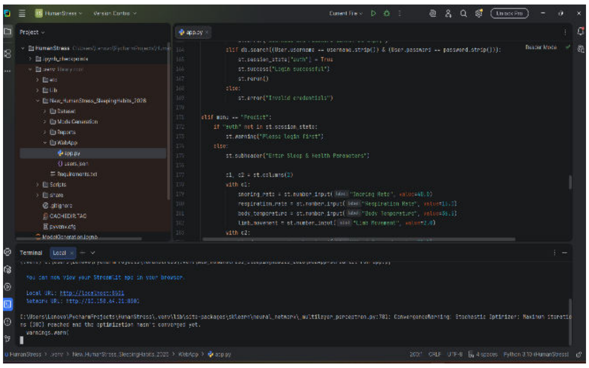

Fig 3: Deployment of the user interface

Fig 3 shows the Streamlet-based web application, where users enter physiological and sleep-related data to obtain real-time stress predictions.

sr	rr	t	lm	bo	rem	sr.1	hr	sl
93.8	25.68	91.84	16.6	89.84	99.6	1.84	74.2	3.0
91.64	25.104	91.552	15.88	89.552	98.88	1.552	72.76	3.0
60.0	20.0	96.0	10.0	95.0	85.0	7.0	60.0	1.0
85.76	23.536	90.768	13.92	88.768	96.92	0.768	68.84	3.0
48.12	17.248	97.872	6.496	96.248	72.48	8.248	53.12	0.0
56.88	19.376	95.376	9.376	94.064	83.44	6.376	58.44	1.0
47.0	16.8	97.2	5.6	95.8	68.0	7.8	52.0	0.0
50.0	18.0	99.0	8.0	97.0	80.0	9.0	55.0	0.0
45.28	16.112	96.168	4.224	95.112	61.12	7.112	50.28	0.0
55.52	19.104	95.104	9.104	93.656	82.76	6.104	57.76	1.0
73.44	21.344	93.344	11.344	91.344	91.72	4.016	63.36	2.0
59.28	19.856	95.856	9.856	94.784	84.64	6.856	59.64	1.0
48.6	17.44	98.16	6.88	96.44	74.4	8.44	53.6	0.0
96.288	26.288	85.36	17.144	82.432	100.36	0.0	75.72	4.0
87.8	24.08	91.04	14.6	89.04	97.6	1.04	70.2	3.0
52.32	18.464	94.464	8.464	92.696	81.16	5.464	56.16	1.0
52.64	18.528	94.528	8.528	92.792	81.32	5.528	56.32	1.0
86.24	23.664	90.832	14.08	88.832	97.08	0.832	69.16	3.0
81.56	22.416	90.208	12.52	88.208	95.52	0.208	66.04	3.0
63.68	20.368	92.368	10.368	90.368	86.84	2.552	60.92	2.0

Fig 4: Sleep Dataset Overview

Fig 4 shows the dataset used for stress prediction, containing important physiological features such as snoring rate, respiration rate, body temperature, limb movement, and heart rate. These features help improve prediction accuracy.

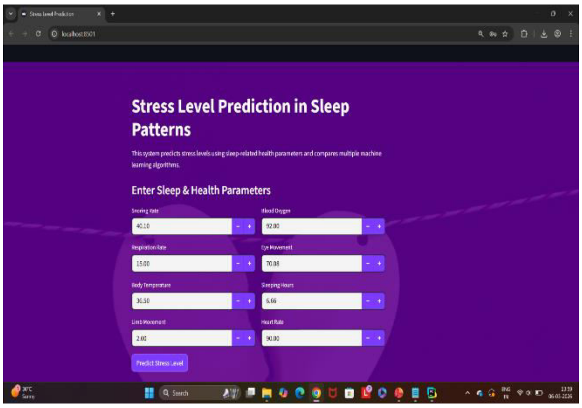

Fig 5: Stress Prediction Input Interface

Fig 5 illustrates the user interaction process, where users enter physiological parameters that are analyzed by the trained model to provide real-time stress predictions.

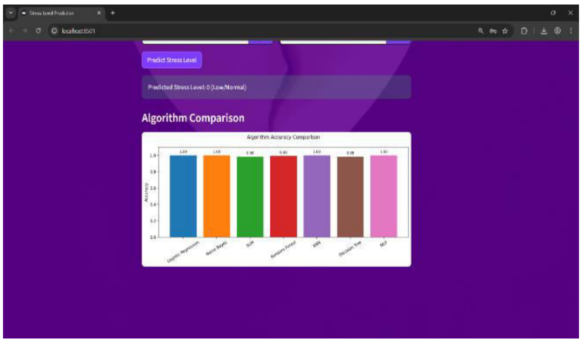

Fig 6: Prediction Output and Algorithm Comparison Graph

The system effectively predicts stress levels, with Random Forest providing the best accuracy. Future improvements should focus on data balance, privacy, and model updates.

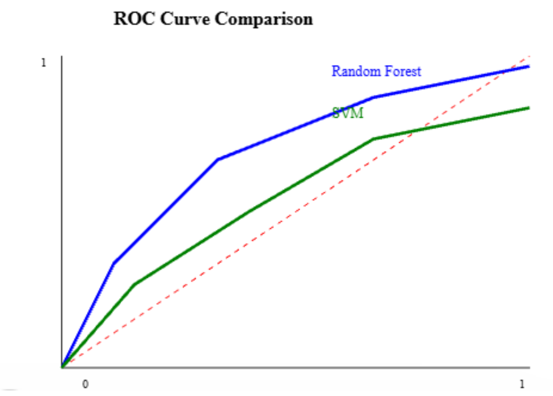

Fig 7: ROC curve

Fig 7 ROC curve comparison evaluates multiple models by showing their True Positive Rate versus False Positive Rate. Higher curves with greater AUC indicate better model performance and classification accuracy.

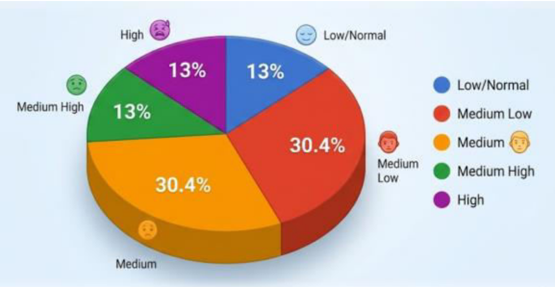

Fig 8: Prediction Chart

Fig 8 The majority of individuals (60.8%) fall into the Medium Low and Medium stress categories. In contrast, the Low/Normal, Medium High, and High stress categories are equally represented, each accounting for 13% of the population.

8. Conclusion and Future Scope

The proposed stress detection system uses the Naïve Bayes algorithm to predict stress levels based on sleep-related and physiological parameters such as sleep duration, heart rate, respiration rate, and sleep quality. Data preprocessing and classification techniques improve the accuracy and reliability of the results. The Naïve Bayes model provides a simple, fast, and computationally efficient method for stress prediction and performs effectively even with limited datasets. The system enables early stress detection, allowing timely intervention and better mental health management. Compared to traditional assessment methods, it offers an objective and automated approach to stress analysis. The model can also be integrated with wearable devices and digital healthcare platforms for continuous monitoring. In the future, larger datasets, additional physiological features, and advanced machine learning techniques can be incorporated to improve prediction accuracy and support real-time stress monitoring applications.

Future Scope: Future improvements may include integration with wearable devices, mobile health applications, and IoT sensors for real-time monitoring. Advanced deep learning techniques can enhance prediction accuracy and provide personalized stress management. Cloud-based platforms may also support continuous monitoring and timely healthcare alerts.

References

[1] G. S. Vos, K. Trinh, and Z. Sarnyai, "Generalizable machine learning for stress monitoring from wearable devices: A systematic review," IEEE Access, vol. 10, pp. 87342–87366, 2022.

[2] H. Lee, S. M. Kim, and H. H. Lee, "Machine learning-based stress detection from sleep behaviour patterns," Journal of Biomedical Informatics, vol. 117, p. 103736, 2021.

[3] N. Baghaei, N. Amini, and S. K. Setarehdan, "Automatic stress detection from physiological signals using machine learning," Computers in Biology and Medicine, vol. 114, p. 103488, 2019.

[4] A. Shirazi and K. Li, "Sleep and stress interplay: A machine learning perspective on wearable data," Journal of Healthcare Informatics Research, vol. 4, no. 3, pp. 478–488, 2020.

[5] X. Liu, J. Yang, and Y. Yao, "Classification of stress level using sleep stages and machine learning," International Journal of Pattern Recognition and Artificial Intelligence, vol. 32, no. 12, p. 1850025, 2018.

[6] P. Kaur and J. Kaur, "Predicting stress using sleep patterns: A machine learning approach," in Proc. Int. Conf. Computing, Communication and Automation, 2018, pp. 345–351.

[7] A. Jindal and S. Binjrait, "Stress recognition using wearable sensors and deep learning techniques," in Proc. IEEE Int. Conf. Smart Systems and Technologies, 2017, pp. 222–227.

[8] M. M. Rahman, A. M. Islam, and J. Miah, "Sleep Well: Stress level prediction through sleep data using machine learning," in Proc. 10th Int. Conf. Machine Learning and Data Science (MLDS), 2022, pp. 102–110.

[9] S. Akter, M. M. Sharif, and M. Mahmud, "Stress detection based on sleep pattern analysis using machine learning," International Journal of Scientific Research in Science, Engineering and Technology, vol. 10, no. 3, pp. 45–52, 2020.

[10] A. Gupta, S. Jain, and A. Verma, "A review on stress detection using sleep pattern analysis and machine learning," International Journal of Engineering Research & Technology, vol. 9, no. 7, pp. 128–136, 2021.

[11] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 2, pp. 156–166, 2014.

[12] S. Can, D. Arnrich, and C. Ersoy, "Stress detection in daily life scenarios using smart phones and wearable sensors: A survey," Journal of Biomedical Informatics, vol. 92, p. 103139, 2019.

[13] F. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in Proc. IEEE Int. Conf. Affective Computing and Intelligent Interaction, 2015, pp. 671–676.

[14] A. Gjoreski, M. Luštrek, H. Gams, and M. Gjoreski, "Monitoring stress with a wrist device using context," Journal of Biomedical Informatics, vol. 73, pp. 159–170, 2017.

[15] M. Schmidt, B. Reiss, R. Duerichen, and K. Van Laerhoven, "Introducing WESAD: A multimodal dataset for wearable stress and affect detection," in Proc. ACM Int. Conf. Multimodal Interaction, 2018, pp. 400–408.