Human Stress Prediction Using Activity Data

Namratha Renjal
School of Computing
Dublin City University
Dublin, Ireland
namratha.renjal2@mail.dcu.ie

Vaishnavi Manjunatha
School of Computing
Dublin City University
Dublin, Ireland
vaishnavi.manjunatha2@mail.dcu.ie

***Abstract*—Predicting stress is important because it facilitates early intervention, prevents serious mental health problems, and enhances general well-being. It also increases productivity by making it possible to promptly handle pressures in both personal and professional contexts. This paper proposes a two-fold approach to enhance stress prediction. The first step involves Human Activity Recognition, which aims to reduce the user's effort in labeling activities by automatically predicting them. The second step focuses on predicting stress by analyzing the daily activity label data obtained from HAR. The publicly-available ETRI dataset comprises 570 days of experimental sessions, amounting to approximately 7,350 hours of data from 22 subjects. The primary goal of the study is to contribute to stress prediction using activity data, enabling early intervention and adjustments in daily routines to mitigate stress effectively. We conducted a comparative analysis of various machine learning algorithms, including LSTM, Bi-LSTM, Decision Tree, and SVM, for Human Activity Recognition and stress prediction. Our findings revealed that Bi-LSTM achieved the highest accuracy of 92.18% for HAR. For stress prediction, Bi-LSTM outperformed the other algorithms, delivering an accuracy of 72.5% with activities performed during the second half of the day.**

***Index Terms*—Stress Prediction, Activity labels, Human Activity Recognition**

I. Introduction

In today's fast-paced world, stress has become a common experience for many individuals. Whether it's due to work pressures, financial concerns, relationship issues, or other life challenges, stress can have a significant impact on both our physical and mental well-being. Understanding what stress is and learning effective strategies to manage it plays a major role in maintaining a healthy and balanced life. [1], [2] While some levels of stress can be beneficial, motivating us to take action and adapt to challenging situations, excessive or chronic stress can have detrimental effects on our health [3]. Prolonged exposure to stress can lead to a range of physical and mental health problems, including high blood pressure, anxiety, depression, and weakened immune function [4].

Stress prediction aids stress management by offering early awareness of impending stressors, allowing individuals to intervene before stress becomes overwhelming. By analysing patterns in behaviour and physiology, predictive models can be used to provide personalised recommendations tailored to individual stress triggers and responses [5]. Continuous monitoring and prediction of stress contribute to resilience building by fostering self-awareness and facilitating the development of adaptive coping strategies. Ultimately, stress prediction may help empower individuals to take proactive steps towards self-care, leading to improved overall well-being and a greater sense of control over stress levels [6].

It has been generally observed that activities like working, commuting and making financial decisions can be stressful. On the other hand, activities like watching TV, engaging in a hobby or playing with your pet can help alleviate stress. Identifying such activities can help us strike a balance and manage stress in a healthy manner. Sleep plays a crucial role in regulating stress levels and promoting emotional resilience [7], better sleep quality is also observed to undo the detrimental effects of stress [8].

Our proposed project aims to provide insights into the relationship between daily activities and stress levels logged by users. Early prediction of stress allows individuals to take preventive measures to manage and reduce stress before it escalates into more serious mental health issues. We will be using the activity performed and the details associated with the activities, with sleep and work being the major activities, to make inferences of the stress and fatigue levels of the user. The stress levels that will be used for training was collected from a survey that users were asked to fill at the end of the day during the period in which they participated in the study [9].

II. Literature Review

Frustration, nervousness, and anxiety are the most common symptoms of stress (50%) in younger people [10]. It creates a negative impact on human lives and damages the performance of each individual [11]. Early prediction of stress and the level of stress can help to reduce its impact and different serious health issues related to this mental state.

Previous studies have primarily focused on utilising biological and physiological sensor data for predicting stress without any contextual meaning added to it. We believe attaching activity context would add much more insight to physiological sensor data while predicting stress. However, these sensors lack inherent activity awareness, necessitating manual labelling efforts to associate sensor data with specific activities. To streamline this process, our approach aims to autonomously recognise activities from sensor data, thereby eliminating the need for manual labelling. Subsequently, we aim to leverage the recognised activities to predict the corresponding level of induced stress.

We begin our work by examining the features that have been predominantly used for stress prediction, followed by reviewing previous research across a diverse array of daily activity recognition.

A. Human Activity Recognition

Activity recognition has garnered significant attention in recent years, with notable advancements in various research endeavours. The study in [12] uses a non-parametric approach to analyze human behaviors more effectively than traditional methods, automatically learning model parameters from sensor data. Additionally, wearable cameras, which are now widely used, can autonomously record data for extended periods without user intervention [13]. The experiments demonstrated in this study shows that the CORAL adaptation method allows models to be successfully transferred to different datasets from various wearable cameras, yielding competitive results with minimal labeled data. [14] proposes a deep learning system that recognizes activities from wearable sensor data with a high Accuracy rate of around 99% and provides explainable AI for understanding behavior in various settings. There have been a lot of studies on HAR using image and video data. Deep learning approaches, particularly convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have been effective in classifying human activities from image and video data [15] [16]. Most prior research has utilized physiological sensors to predict physical activity labels, while camera photo or video data has been employed for identifying behavioral activity labels. Building on previous research in user-activity analysis, we aim to predict stress levels based on daily activities and various motion and physiological sensor data. This requires activity recognition as the first step to eliminate the need for manual labeling.

B. Stress Prediction

Datasets on stress, unlike typical datasets, are characterised by extensive data volumes sourced from wearable devices, smartphones, electronic health records and daily activity data, offering a multifaceted perspective on stress. Machine learning models are suitable for predicting stress in individuals due to their ability to analyse large, diverse datasets and identify complex patterns indicative of stress. Common features used in various papers for predicting stress with machine learning encompass a wide range of physiological, behavioural, and environmental indicators. In [17] psychological features like mood experiences, cognitive and behavioural experiences, psychiatric and physical symptoms were used to predict happiness. Meanwhile, [18] used bio-metric data, diet, historical mood data, sleep data and environment data to detect and predict the current mood of the user. The use of biological sensor data like heart rate and galvanic skin resistance and results of perceived stress scale have proven effective to predict a positive or negative change in stress levels [19]. The physiological data and other features like results from personality tests, sleep quality and music related data are used in another study [20]. [21] discusses the effect of time and location on prediction of activity, they found a strong correlation between the location and activity.

The use of logistic regression in [20] to predict mood and ensemble methods with weak classifiers to detect music mood and style. Being a supervised machine learning method, logistic regression predicts the likelihood of an outcome, typically for binary classification tasks by estimating probabilities based on input features. It specialises in delivering binary outcomes (Yes/No, 0/1) allowing for straightforward decisions between two possibilities.The study [17] made use of SVM to predict happiness levels of the study subject, they achieved an accuracy of around 80%, building on this. The methodology employed a nonlinear SVM with a Gaussian kernel to predict happiness levels from various features, implemented using the ”e1071” package in R. Features were preprocessed by centering, normalising, and selecting based on significant correlations with happiness and effect sizes. Fourteen features were selected as inputs for the SVM. Training involved using a radial basis function kernel, with hyperparameters (C,γ) tuned via grid search for optimal classification accuracy. Testing involved evaluating the SVM's prediction accuracy using lifelog data from a separate test dataset. The study [18] made use of C4.5 to boost the dataset to improve the accuracy of the SVM algorithm in predicting the mood of a user.

Another study [19] compared the performance of three different supervised learning algorithms like Linear Discriminant Analysis, quadratic discriminant analysis and decision tree in predicting changes in stress levels. Decision tree gave an accuracy of 70%. Validation was conducted using k-fold crossvalidation (k = 10).

The study [22] uses machine learning algorithms for activity recognition and stress detection based on data collected from a wristband. For activity recognition, the study employs machine learning models such as Artificial Neural Network (ANN), Decision Tree (DT), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Logistic Regression (LR). These models are trained and evaluated using 3-axis acceleration signals recorded on the wrist to distinguish daily motion activities with different levels of physical intensities. They achieved approximately 97% accuracy in recognizing physical activities using Decision trees and over 80% accuracy in detecting stress using SVM. It is observed that daily stress detection should not solely rely on physiological signals, as models based only on these signals struggle to effectively differentiate between high-intensity activities and stress conditions.

As can be seen, the machine learning methodologies usually used for activity and stress recognition tends to be SVM as it excels in adjusting to feature-based classification methods, particularly when dealing with complex relationships between features. It defines its functions by identifying a boundary in the feature space that effectively separates different classes during training [23].

The existing body of research on stress prediction predominantly focuses on sensor-based approaches, leveraging physiological signals such as heart rate, skin conductance, and cortisol levels to infer stress. They rely on input data that requires manual effort from users, such as surveys, mood logs, or dietary information that they need to fill out. While these studies have laid a crucial foundation, there remains a gap in exploring the potential influence of using activities performed by users. The impact the activities have on stress prediction has been a topic which has not been explored in the technical aspect. It has however been backed up by many philosophical studies. Our study aims to bridge this gap by evaluating the accuracy of stress prediction through diverse activity data of individuals and employing various machine learning algorithms to identify the most effective approach. We hypothesize that activity data can serve as a robust indicator of stress when analyzed with sophisticated machine learning techniques. Our research is guided by two key questions:

1) Can the accuracy of stress prediction be enhanced by incorporating activity duration, activity frequency, or both for users?

2) Can the accuracy of stress prediction be improved by utilizing data collected closer to the time when the ground truth was captured, as opposed to using data from the entire day?

By addressing these questions, our study seeks to advance the field of stress prediction, providing insights that could lead to more precise and timely stress monitoring solutions.

III. Methodology

In this study, we adopt a two-fold approach. First, we focus on Human Activity Recognition using motion and physiological sensor data. We apply various machine learning models to determine which one yields the highest accuracy in identifying user activities. This step reduces the labeling effort required from users, facilitating the subsequent phase: stress prediction. Our primary objective is to predict stress levels and identify which activities contribute to stress.

A. Dataset Description

The dataset comprises lifelog data which consolidates 7350 hours of data from 22 subjects.

We have utilized the ETRI Lifelog dataset [11] as a foundational resource. This dataset offers a comprehensive insight into human behavior, comprising 570 days of experimental sessions and approximately 7,350 hours of data from 22 subjects. The dataset encompasses a rich array of physiological metrics, including Blood Volume Pulse (BVP), Electrodermal Activity (EDA), Heart Rate (HR), and skin temperature, recorded via a wrist-worn sensor (Empatica E4). Moreover,

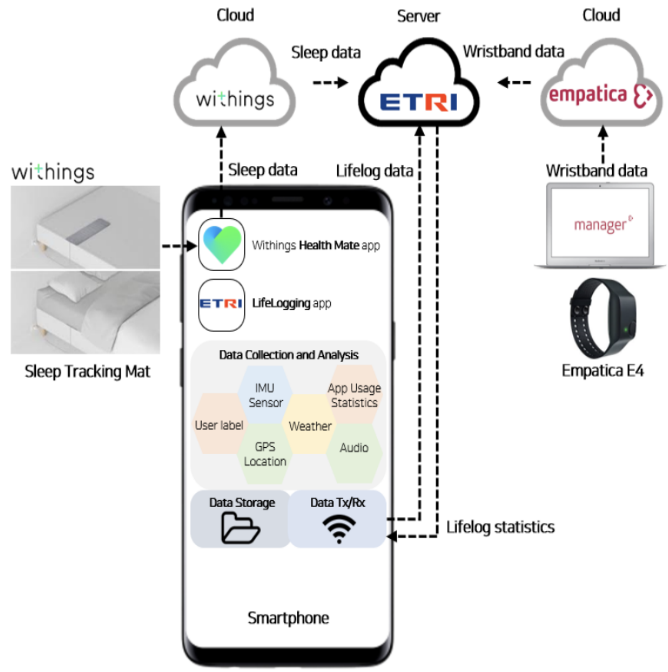

Fig. 1. Sensor data collection in detail

it incorporates multivariate behavioral data such as Inertial Measurement Unit (IMU) data from both mobile phones and the E4 sensor, along with Global Positioning System (GPS) data. The dataset includes 440,830 processed labels, representing 10,732 unique categories that 16 daily activities (including different modes of transportation) and contextual details such as semantic places and social states. After grouping and merging similar activities, we have taken the following final list of activities which include sleep, personal care, work, study, household, caregiving for household members, recreation and media, entertainment, outdoor activities, hobby, free time (recreation, etc.), shop, travel, meal and socializing. The different sensor data collected in our ETRI dataset is depicted in Figure 1. The sensors and their description are mentioned in Table I.

B. Data Pre-processing

For the first step of HAR, we need to deal with sensor data.The pre-processing steps for the sensor data such as noise removal, feature extraction, and normalization were typically performed to enhance the quality of the data. During this stage, data selection techniques were applied to filter out irrelevant or redundant features, thereby reducing computational complexity and improving model performance.

- Consolidation of Sensor Data: Integration of diverse sensor data into a unified file was imperative for streamlined processing. The dataset comprises sensor reading for each minute of the day within a folder, alongside a separate labels file containing activity-related data. To streamline the data and facilitate analysis, we integrated both sensor

TABLE I
SENSORS AND DESCRIPTION

Sensor Code	Description
mAcc and e4Acc	Triaxial acceleration (m/s^2) Mobile (30 Hz) and E4 (32 Hz) accelerometers.
mGps	Location (lat/long) and accuracy (meters) GPS updates every 5 seconds.
mGyr	Triaxial rotation (rad/s, Degrees) Gyroscope at 30 Hz.
mMag	Triaxial geomagnetic field (in μT) Magnetometer at 30 Hz.
e4Bvp	Blood volume pressure (in nano Watt) 4 photoplethysmography (PPG) sensor (64 Hz)
e4Eda	Electrodermal activity (skin conductance in μS) E4 EDA sensor (4 Hz)
e4Hr	Average heart rate values (in bps) (1 Hz)
e4Temp	Peripheral skin temperature (in Celsius degrees) E4 infrared thermopile (4 Hz)

readings and activity data into a single consolidated file, ensuring comprehensive access to all relevant information within one unified dataset.

- Frequency Normalization: Given variations in sensor frequencies, adjustments such as upsampling or downsampling were performed to standardize the consolidated file, ensuring compatibility across all data streams.
- Dimensionality Reduction: Due to the large size of the dataset, we employed dimensionality reduction techniques, we plotted a correlation matrix to examine the correlation between features. By analyzing the matrix, we identified highly correlated features that could be eliminated to reduce redundancy. This method helped us streamline the dataset by removing unnecessary features while preserving relevant information. The refined dataset enhances the efficiency of our model.

For stress prediction, the data required specific formatting to calculate and store the duration and frequency of each activity based on timestamps. Additionally, we integrated this activity data with user's emotion and stress scores collected from the user survey, ensuring a comprehensive dataset for accurate stress prediction analysis.

C. Data Augmentation

We noticed some class imbalance in the activity labels while performing our exploratory data analysis. To address this issue we applied **SMOTE** (Synthetic Minority Oversampling Technique) which in-turn helped us reduce overfitting by providing more samples. It also helped us improve the model performance by improving the accuracy. SMOTE addresses the class imbalance problem by generating synthetic samples for the minority class, creating new instances along the lines connecting minority class samples and their nearest neighbors.This technique effectively increases the presence of minority class examples, which helps in shifting the decision boundary towards the minority class, potentially improving the classification performance.The method has been shown to improve classification metrics like Gmean, indicating better balance in classifier performance between minority and majority classes after applying SMOTE [24]. It has been utilized in this study during both the Human Activity Recognition phase and the Stress Prediction phase.

D. Machine Learning Techniques

*1) **Human Activity Recognition**:* To perform human activity recognition, we conducted a comparative analysis using various machine learning techniques, including decision tree, Long-short term memory(LSTM), and bi-directional LSTM models, on the pre-processed dataset. For training the models, we utilized sensor data paired with activity labels collected from user surveys. The predicted activity labels aim to eliminate the need for manual labeling by users, facilitating the next crucial phase of our research: stress prediction.

Decision tree: It recursively splits the dataset into smaller subsets based on the most significant attribute, creating a tree-like structure to predict the target variable.Decision trees are suitable for sensor data due to their ability to handle non-linear relationships and interactions between features effectively. They can discern complex patterns in sensor readings and provide interpretable results, aiding in meaningful insights and actionable decisions.

*2) **Stress Prediction**:* The predicted activity labels from human activity recognition were utilized to determine the duration and frequency of each activity performed by users daily.The duration and the frequency of the activity was extracted from the data and it was then pivoted and rearranged to bring it into a usable form. We considered various feature variations to evaluate which would yield the best accuracy in predicting stress:

- Duration of the activity
- Frequency of the activity
- Both frequency and duration of the activity
- Activities in the first half and second half of the day

The training set will include the necessary activity data alongside the stress score data obtained from user surveys. We combined emotion and stress scores reported by each user to achieve the highest accuracy. The stress score is measured on a scale of 10 and is categorized into three levels: low stress, moderate stress, and high stress. We used Support Vector Machine(SVM) as a baseline model for stress prediction. Additionally, we applied LSTM and bi-directional LSTM models to the dataset, as previous research indicated promising results with these approaches. The models were trained for 10 epochs with a batch size of 16 and a dropout

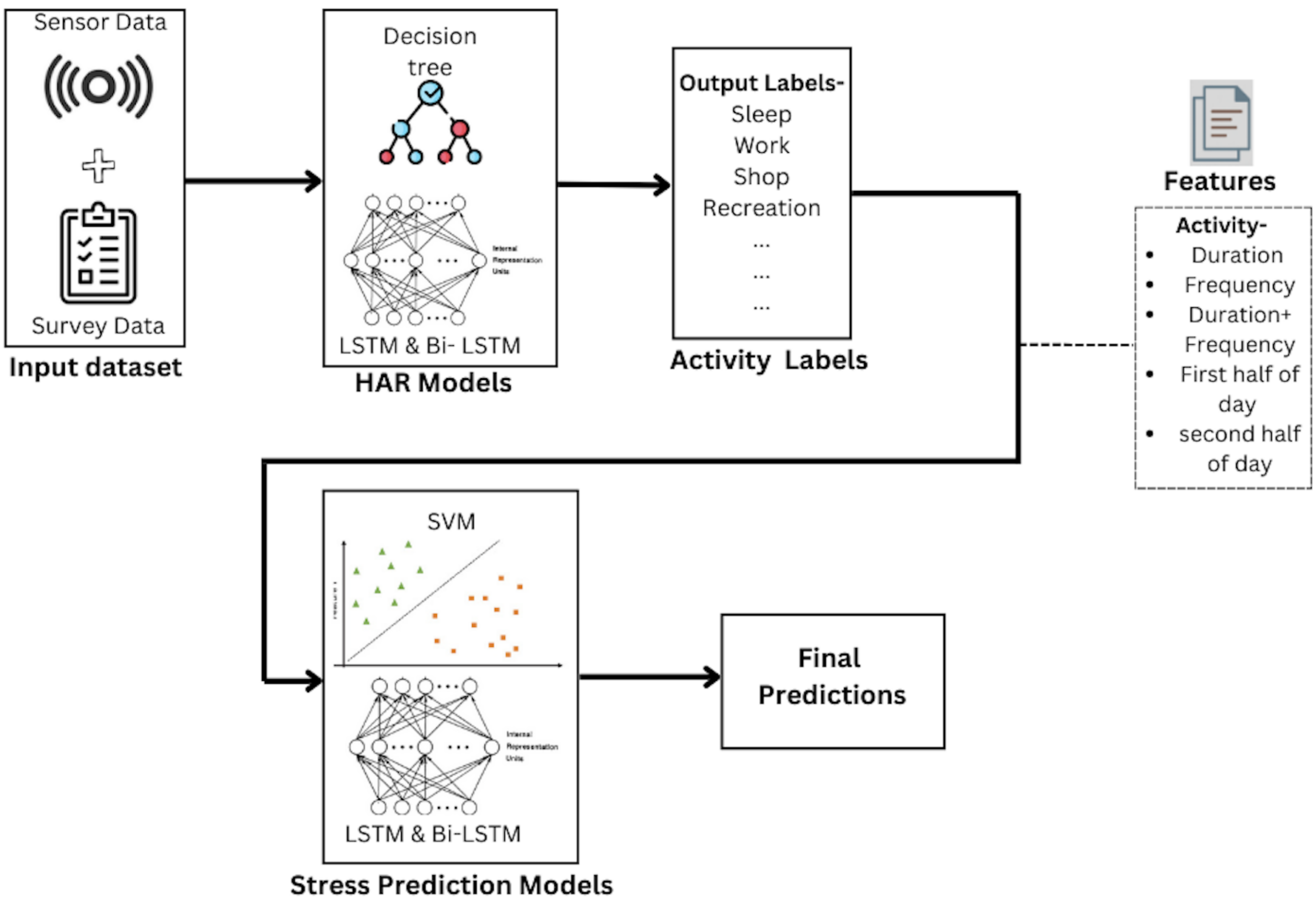

Fig. 2. Representation of the flow the study

rate of 0.5.

SVM : They are machine learning model known for their robustness and efficacy in handling high-dimensional data. In stress prediction tasks, SVMs are employed to classify physiological signals or behavioral data into stress and non-stress categories. By finding the optimal hyperplane that separates the data into these classes, SVMs can accurately predict stress levels. Their ability to manage non-linear relationships through kernel functions makes them particularly suited for the complex patterns often present in stress-related data. Furthermore, SVMs are effective in scenarios with limited data, which is common in personalized stress monitoring applications, providing reliable baseline performance for stress prediction models.

LSTM : It is a specialized type of RNN developed to address the shortcomings of traditional RNNs in capturing long-term dependencies in sequential data. It incorporates memory cells and gating mechanisms, which allconvol to selectively retain and forget information over time. With its internal memory state, LSTM can store information for extended periods, enabling it to capture dependencies that span multiple time steps.

Bi-LSTM: Unlike traditional RNNs, which process input sequences in a single direction (either forward or backward), Bi-LSTM processes sequences simultaneously in both directions. It comprises two LSTM layers: one that processes the sequence in the forward direction and another that processes it in the backward direction. Each layer maintains its own hidden states and memory cells, allowing the model to capture information from both past and future contexts.
The comprehensive workflow of our entire approach is detailed in Figure 2.

IV. Results

This section delves into a detailed discussion of the evaluation metrics and results derived from the comparative analysis of various machine learning algorithms. We will begin by examining the evaluation metrics for human activity recognition and stress prediction.

A. Evaluation metrics

We have used the evaluation metrics Precision and Accuracy used in machine learning classification tasks.

Precision:It measures the proportion of correctly predicted positive instances (true positives) out of all instances predicted as positive by the model.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

Accuracy: It measures the proportion of correctly predicted instances (both true positives and true negatives) out of all instances evaluated.
Formula:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

Recall: It is a performance metric used in classification problems to evaluate the effectiveness of a model in identifying all relevant instances in a dataset.
Formula:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

- True Positives (TP): Instances correctly predicted as positive.
- False Positives (FP): Instances incorrectly predicted as positive.
- True Negatives (TN): Instances correctly predicted as negative.
- Total Number of Instances: The total number of instances in the dataset.

B. Model Performance

The performance of our model was majorly improved by SMOTE which addressed the class imbalance by generating synthetic samples for the minority classes in case of . It did this by interpolating between existing minority class samples, creating new, similar instances. This process enhances the representation of the minority class without merely duplicating existing samples, thereby improving model performance by ensuring it doesn't become biased towards the majority class. This technique effectively balances the dataset, contributing to the creation of more robust and accurate predictive models.For time series classification, a novel approach combining shapelets transform with SMOTE preserves important variable relationships while augmenting data [25]. In recruitment recommendation systems, SMOTE-based augmentation has shown promising results in overcoming dataset imbalances and improving classification accuracy, particularly with the MultiLayer Perceptron algorithm [26].

In the context of Human Activity Recognition, we observed notable differences in the performance of the various models tested. The decision tree algorithm underperformed, achieving an accuracy of only 69.22%. In contrast, the LSTM model demonstrated significantly improved results. However, the bi-directional LSTM (bi-LSTM) outperformed both, delivering the highest accuracy of 92.18%. This indicates that bi-LSTM is the most effective model for HAR in our experiments, capturing both past and future context in the data more efficiently than the other models.The results are shown in table II.

TABLE II
PERFORMANCE OF HAR MODELS

Model	Accuracy	Precision	Recall
Decision Tree(with GridSearchCV)	69.22%	72%	69%
LSTM	85.74%	90.04%	82.81%
Bi-LSTM	**92.18%**	**93.72%**	**90.92%**

Moving on to the stress prediction phase, we experimented with different feature variations to improve accuracy. Among the three models tested—SVM, LSTM, and Bi-LSTM—the Bi-LSTM consistently delivered superior results across all feature variations. Notably, it performed exceptionally well when analyzing activities performed during the second half of the day, achieving an precision of 75.94%. Detailed accuracy and precision metrics for each model and feature variation are provided in the accompanying table III.

C. Comparative Analysis

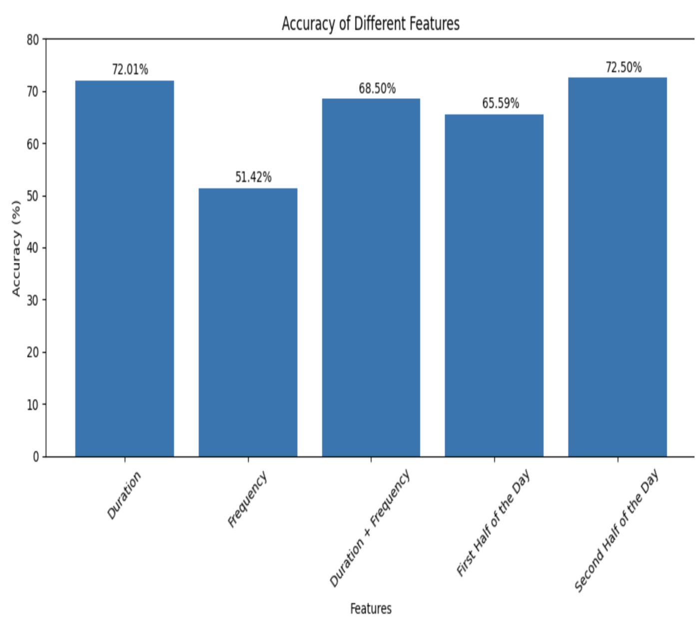

Fig. 3. Accuracy of different features using Bi-LSTM model

In our comparative analysis of various feature sets and machine learning models, several notable findings emerged. Initially, we experimented with the duration and frequency of the activities performed by users. Our results indicated that using the duration as a feature yielded a significantly higher accuracy of 72.01%, compared to using the frequency of activities, which resulted in an accuracy of only 51.42%.

TABLE III
COMPARISON OF MODEL PERFORMANCE FOR STRESS PREDICTION

Features	Evaluation metrics	SVM	LSTM	Bi-LSTM
Activity duration	Accuracy	70.40%	69.79%	**72.01**%
	Precision	71.54%	70.75%	72.23%
	Recall	70.40%	58.36%	63.80%
Activity duration + Frequency	Accuracy	68.16%	62.91%	68.50%
	Precision	69.88%	66.13%	72.42%
	Recall	68.16%	50.07%	60.22%
Activity frequency	Accuracy	50.65%	51.57%	51.42%
	Precision	49.82%	57.17%	56.60%
	Recall	50.65%	24.55%	27.24%
Activities in the first half of the day	Accuracy	43.49%	61.84%	65.59%
	Precision	51.20%	68.64%	68.63%
	Recall	43.49%	51.91%	61.10%
Activities in the second half of the day	Accuracy	56.02%	65.96%	**72.50%**
	Precision	69.69%	70.65%	**75.94**%
	Recall	56.02%	59.71%	68.16%

Interestingly, when both duration and frequency were combined, the accuracy achieved was 68.5%. This suggests that while the combination improves over frequency alone, duration remains the most impactful feature for our Bi-LSTM model. Overall, we determined that Bi-LSTM was the most effective model for both HAR and stress prediction with our chosen dataset.A graph of the comparative graph with accuracy racies of different features is shown in Figure 3

To further understand the temporal influence on activity patterns, we divided the activities into two distinct periods: the first half of the day and the second half. Our analysis revealed that activities performed in the first half of the day resulted in an accuracy of 65.59% and a recall score of 68.63%. However, the second half of the day produced superior results, with the highest overall accuracy and precision, recorded at 72.5% and 75.94%, respectively. These findings suggest a significant temporal variation, indicating that activities performed later in the day have a stronger correlation with the user's emotional states, possibly due to their proximity to the time when emotions are recorded.

V. DISCUSSION

The dual-phase strategy in our project, incorporating Human Activity Recognition followed by stress prediction, eliminates the necessity for manual activity labeling. HAR autonomously identifies activities, which can subsequently be utilized for stress prediction.

*1) **Data Augmentation**:* Data augmentation was essential for our experiment because stress and emotion levels were recorded only once a day. Despite compiling activity and stress data from multiple days, the dataset remained insufficient for effective modeling and accurate results. Augmentation significantly enhanced the dataset, proving invaluable in overcoming these limitations and improving model performance. We employed the Synthetic Minority Over-sampling Technique for data augmentation, an effective approach to mitigate class imbalance in machine learning [27]. SMOTE generates synthetic examples for underrepresented classes, such as recreation and shopping, thereby balancing the dataset. This prevents the model from becoming biased towards the majority class, such as work, while leaving the majority class data untouched. In our case, minimal augmentation of minority classes proved sufficient for achieving the desired balance for HAR. The ability of SMOTE to create diverse yet realistic synthetic data contributed to the development of more robust and accurate models. Without augmentation, a significant portion of our stress data was being misclassified as moderate stress. Given the nature of our data, SMOTE has proven to be the most effective technique for generating realistic synthetic samples.

*2) **Data close to the emotion capture time**:* To refine our stress prediction model, we experimented with multiple machine learning techniques and evaluated five distinct feature sets to determine the most accurate predictors. These features included activity duration, frequency, a combination of duration and frequency, and activities segmented into the first and second halves of the day. Our analysis revealed that activities performed in the second half of the day were the most predictive of stress levels, yielding a maximum precision of 75.94% and an accuracy of 72.5% with Bi-directional LSTM. With just 570 days of data we are were able achieve a significant accuracy. This result diminishes the importance of considering the activities performed the entire day in relation to stress prediction and, suggesting that recent activities, especially those performed within the last 4 to 5 hours before prediction, substantially impact stress levels.This approach is efficient, requiring less data while delivering better results, thereby saving time and effort. The Figure 4 shows a graph of the comparative precision value achieved while considering the all the activities performed in a day (Duration+ Activity), first half and the second half of the day.

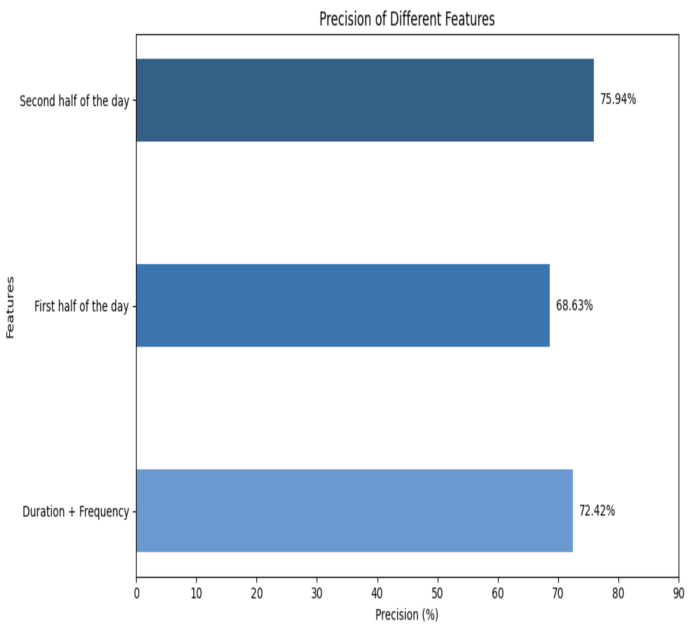

Fig. 4. Precision of different features using Bi-LSTM model

*3) **Improvements in HAR**:* Previous studies have predominantly concentrated on physical activity labels such as walking, running, and lying down. In contrast, our study extends the focus to a broader set of behavioral and contextual activity labels, including work, travel, household chores, personal care, and recreation. Despite the increased complexity and variety of these activity labels, our human activity recognition models demonstrated robust performance. By leveraging sensor readings as input, our models effectively classified these diverse activities, showcasing the capability to accurately recognize a wide range of behaviors and contexts. This advancement highlights the versatility and efficacy of our approach in handling more comprehensive activity sets.

In conclusion, our investigation into various feature sets revealed that activities in the latter half of the day are highly predictive of stress, achieving a notable accuracy of 72.5% with Bi-directional LSTM and an accuracy of 92.18% for HAR. Precise stress prediction can assist in identifying those who are susceptible to stress-related illnesses before they worsen and become more serious health problems, including depression or anxiety.

VI. Future scope

The current study has laid a strong foundation for predicting stress using machine learning models based on human activity recognition. However, there are several promising avenues for future research that could enhance the effectiveness and applicability of this work. One potential direction is the development of specialized models tailored for individual users. Personalized models could provide more accurate insights into the specific activities that contribute most significantly to each user's stress levels and identify activities that help alleviate stress. By understanding these individual variations, interventions can be more precisely targeted, allowing users to modify their behavior to better manage their stress. Furthermore, future research could benefit from incorporating additional contextual data. This might include environmental factors such as weather, location, or social interactions, as well as physiological signals like heart rate variability or sleep patterns. Combining these contextual data points with activity data could provide a more holistic view of the factors influencing stress, leading to more accurate predictions. Advancements in deep learning architectures also hold significant promise for improving stress prediction models. Exploring the use of more sophisticated neural network designs, such as attention mechanisms, transformers, or hybrid models that combine convolutional and recurrent neural networks, could enhance the model's ability to capture complex patterns in the data.

Additionally, longitudinal studies observing users over extended periods could provide valuable insights into the long-term effects of activity modifications on stress levels. This would help in understanding the sustainability of stress management interventions and their impact on overall well-being. By building specialized user-specific models, integrating richer contextual data, and leveraging advanced deep learning techniques, future research can significantly advance the field of stress prediction. These enhancements will contribute to the development of more effective and personalized stress management strategies, ultimately promoting better mental health and quality of life.

VII. Conclusion

In this study, we investigated the effectiveness of various machine learning algorithms for predicting human stress using activity data obtained from Human Activity Recognition. Our approach involved a two-step process: first, recognizing activities from sensor data to minimize manual labeling efforts, and second, predicting stress levels based on these recognized activities combined with user-reported stress scores.

Our experimental results demonstrated that the bi-directional Long Short-Term Memory (bi-LSTM) model achieved the highest accuracy in HAR, significantly outperforming other models such as decision trees and standard LSTM networks. Specifically, the bi-LSTM model attained an accuracy of 92.18%, highlighting its superior capability in capturing both forward and backward dependencies in sequential data.

For stress prediction, we utilized Support Vector Machines (SVM) as a baseline model and explored the application of LSTM and bi-LSTM models which gave us an accuracy of 72.5%. The models were trained using a comprehensive dataset consisting of activity data and user-reported stress scores, categorized into low, moderate, and high stress levels. Our findings indicated that incorporating both frequency and duration of activities, as well as considering the activities in the second half of the day, significantly enhanced the predictive performance of our models. While our findings are promising, they are based on a limited dataset in terms of duration and the number of participants. We hypothesize that the accuracy of stress prediction could be further enhanced with a larger and more diverse dataset, encompassing more days of activity data and a broader user base. This potential improvement highlights the need for extensive longitudinal studies to better understand the complex relationship between daily activities and stress. Our study opens new avenues for future research, suggesting that the inclusion of a wider array of activities and more comprehensive datasets could lead to more accurate and reliable stress prediction models. The integration of HAR and advanced machine learning techniques for stress prediction provides a promising avenue for early detection and intervention, ultimately contributing to improved mental well-being and productivity.

REFERENCES

[1] J. Viveros and D. G. Schramm, "Why stress management strategies work," *Families and Communities*, 2018.

[2] M. S. Abdugapparovna, "Improving a healthy lifestyle among university youth," *IJODKOR O'QITUVCHI*, vol. 3, no. 36, pp. 5–10, 2024.

[3] J. P. Herman, M. Maroun, and G. Richter-Levin, "Good stress, bad stress and very bad stress," pp. 267–268, 2015.

[4] M. Wahrendorf, G. Sembajwe, M. Zins, L. Berkman, M. Goldberg, and J. Siegrist, "Long-term effects of psychosocial work stress in midlife on health functioning after labor market exit—results from the gazel study," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 67, no. 4, pp. 471–480, 2012.

[5] R. Kocielnik and N. Sidorova, "Personalized stress management: enabling stress monitoring with lifelogexplorer," *KI-Künstliche Intelligenz*, vol. 29, pp. 115–122, 2015.

[6] K. Lee, H. Cho, K. Toshnazarov, N. Narziev, S. Y. Rhim, K. Han, Y. Noh, and H. Hong, "Toward future-centric personal informatics: Expecting stressful events and preparing personalized interventions in stress management," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.

[7] J. M. Blaxton, C. S. Bergeman, B. R. Whitehead, M. E. Braun, and J. D. Payne, "Relationships among nightly sleep quality, daily stress, and daily affect," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 72, no. 3, pp. 363–372, 2017.

[8] B. L. Fredrickson, "The broaden–and–build theory of positive emotions," *Philosophical transactions of the royal society of London. Series B: Biological Sciences*, vol. 359, no. 1449, pp. 1367–1377, 2004.

[9] S. Chung, C. Y. Jeong, J. M. Lim, J. Lim, K. J. Noh, G. Kim, and H. Jeong, "Real-world multimodal lifelog dataset for human behavior study," *ETRI Journal*, vol. 44, no. 3, pp. 426–437, 2022.

[10] A. Kene and S. Thakare, "Mental stress level prediction and classification based on machine learning," in *2021 Smart Technologies, Communication and Robotics (STCR)*. IEEE, 2021, pp. 1–7.

[11] S. Cohen, "Aftereffects of stress on human performance and social behavior: a review of research and theory." *Psychological bulletin*, vol. 88, no. 1, p. 82, 1980.

[12] F.-T. Sun, Y.-T. Yeh, H.-T. Cheng, C. Kuo, and M. Griss, "Nonparametric discovery of human routines from sensor data," in *2014 IEEE international conference on pervasive computing and communications*. IEEE, 2014, pp. 11–19.

[13] A. Cartas, P. Radeva, and M. Dimiccoli, "Activities of daily living monitoring via a wearable camera: Toward real-world applications," *IEEE access*, vol. 8, pp. 77 344–77 363, 2020.

[14] M. Z. Uddin and A. Soylu, "Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning," *Scientific Reports*, vol. 11, no. 1, p. 16455, 2021.

[15] A. P. Ismail, M. A. B. Azahar, N. M. Tahir, K. Daud, and N. M. Kasim, "Human action recognition (har) using image processing on deep learning," in *2023 IEEE 13th International Conference on Control System, Computing and Engineering (ICCSCE)*. IEEE, 2023, pp. 179–183.

[16] M. Meratwal, N. Spicher, and T. M. Deserno, "Multi-camera and multi-person indoor activity recognition for continuous health monitoring using long short term memory," in *Medical imaging 2022: imaging informatics for healthcare, research, and applications*, vol. 12037. SPIE, 2022, pp. 64–71.

[17] T. Yamamoto, J. Yoshimoto, E. Murillo-Rodriguez, and S. Machado, "Prediction of daily happiness using supervised learning of multimodal lifelog data," *Revista Psicologia e Saúde*, vol. 11, no. 2, p. 145–152, jul. 2019. [Online]. Available: https://pssaucdb.emnuvens.com.br/pssa/article/view/823

[18] P. Soleimaninejadian, M. Zhang, Y. Liu, and S. Ma, "Mood detection and prediction based on user daily activities," in *2018 First Asian Conference on Affective Computing and Intelligent Interaction*. IEEE, 2018.

[19] C. Dobbins, S. Fairclough, P. Lisboa, and F. F. G. Navarro, "A lifelogging platform towards detecting negative emotions in everyday life using wearable devices," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops*. IEEE, 2018, pp. 306–311.

[20] P. Soleimaninejadian, Y. Wang, H. Tong, Z. Feng, M. Zhang, Y. Liu, and S. Ma, "Thir2 at the ntcir-13 lifelog-2 task: Bridging technology and psychology through the lifelog personality, mood and sleep quality." in *NTCIR*, 2017.

[21] S.-L. Chua, L. K. Foo, and H. W. Guesgen, "Predicting activities of daily living with spatio-temporal information," *Future internet*, vol. 12, no. 12, p. 214, 2020.

[22] J. C. Y. Wong, J. Wang, E. Y. Fu, H. V. Leong, and G. Ngai, "Activity recognition and stress detection via wristband," in *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia*, 2019, pp. 102–106.

[23] M. Sevil, M. Rashid, M. R. Askari, Z. Maloney, I. Hajizadeh, and A. Cinar, "Detection and characterization of physical activity and psychological stress from wristband data," *Signals*, vol. 1, no. 2, pp. 188–208, 2020.

[24] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (smote) for handling class imbalance," *Information Sciences*, vol. 505, pp. 32–64, 2019.

[25] P. Li, S. F. Boubrahimi, and S. M. Hamdi, "Shapelets-based data augmentation for time series classification," in *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2021, pp. 1373–1378.

[26] A. Nurhopipah, Y. Ceasar, and A. Priadana, "Improving machine learning accuracy using data augmentation in recruitment recommendation process," in *2021 3rd east indonesia conference on computer and information technology (EIConCIT)*. IEEE, 2021, pp. 203–208.

[27] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary," *Journal of artificial intelligence research*, vol. 61, pp. 863–905, 2018.