DEPRESSION DETECTION THROUGH ACTIVITY RECOGNITION: DEEP LEARNING MODELS USING SYNTHESIZED SENSOR DATA

Abdul Kadar Muhammad Masum PhD[1], Md Fokrul Islam Khan[2], Fariha Anjum[3], Sadia Alam[3], Erfanul Hoque Bahadur[3]

[1]Professor, Dept. of Software Engineering, Daffodil International University, Bangladeshakmmasum@yahoo.com
[2]Dept. of Management Information System, International American University, California, USA
Fokrulkhan837@gmail.com
[3]Dept. of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh
Ehb1729@gmail.com

[1]***Abstract*—Despite having a very different kinesthetic sensibility from smartphone sensors, human bodies assess variations and address specific sensor values. Human Activity Recognition (HAR) has been used extensively in various applications; however, despite this, HAR cannot currently be used to correlate activity patterns to identify biomarkers for any disease. This provisional scrutiny leverages HAR to detect depression-symptomatic activities. Data collection was carried out in various ways, combining outdoor and indoor activities, while the smartphone was kept in the slash pocket. The Generative Adversarial Network (GAN) model generates synthesized sensor data to enhance the size of the original dataset. The dataset was improved via preprocessing, including the Butterworth low-pass filter. Since the data was linear, deep learning models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) were used. The evaluation procedure is divided into three sections. To begin with, LSTM outperformed GRU on the combination of actual and generated data, with an accuracy of 96.48%. Furthermore, the selected dataset was then analyzed to identify the impact of the Butterworth low-pass filter, which produced a higher accuracy of 2.01%. Finally, the obtained dataset was compared to two publicly accessible datasets, WISDM and MHEALTH, regarding the two suggested models, with the conjunction of the depression dataset and the LSTM model achieving a greater accuracy of 1.80%.**

***Index Terms*—Biomarkers, Depression Symptomatic Activities,Gated Recurrent Unit (GRU).Long Short-Term Memory (LSTM), Sensor Data**

Introduction

A person's movement within a defined context is the foundation of Human Activity Recognition (HAR), which recognizes a person's physical activity. HAR is noteworthy because of its ubiquitous computing benefits and ability to retrieve meaningful high-level information regarding human activities from raw sensor data. Because they offer consistent tracking assistance, are reasonably priced, and consume little battery power, sensors, particularly those built into electronics (such as smartphones, smartwatches, etc.), are quite popular. Wearable, video-based, wireless, and smartphone sensors are all used by HAR to detect ambulatory activities, postural and bodily motions, and other user activity. The most effective regular activity monitors are smartphones that people use while going about their daily lives.

Previously, methods for identifying human activities included Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and others. Although the techniques have reasonable recognition rates, they focus on routine actions and poorly pick up on complex activities. Additionally, they need extensive efforts in data preparation, human domain expertise, feature extraction and processing, and a particular application. Contrarily, by automating feature extraction and requiring less human involvement, the deep learning subfield of machine learning tackles the shortcomings of manual feature extraction methods. Making precise predictions is made possible by their capacity to collect and analyze attributes from unprocessed sensor data. They have thrived in a number of academic fields recently, including image and audio processing, segmentation, natural language processing, object recognition, and object identification. By creating an end-to-end neural network, deep learning could significantly minimize the labor required to produce features while simultaneously learning a greater number of high-level and significant traits. Deep learning methods for HAR encompass Restricted Boltzmann Machine, Sparse Coding, Autoencoder, Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), Gated Feed Forward Neural Network, and Long Short-Term Memory (LSTM).

This innovative approach focuses on depressed symptoms, including running, sitting, standing, walking, moving downstairs or upstairs, eating, lying down, falling down, cycling, smoking, drinking, and other activities [1]. This research may be broadened to prevent sad or mentally ill persons from committing suicide. The suggested study's whole working methodology entails employing smartphone sensors to capture usual activity patterns and precisely identify all actions indicative of sadness. The GAN or Generative Adversarial Network produced synthetic sensor data. GRU and LSTM, two deep learning algorithms employed as classification models, were given preprocessed sensor data. In order to establish whether the method produced a better classifier, the two methods were contrasted. SVM, random forests, Naive Bayes, and other manually created feature extraction techniques were formerly employed in the majority of works. However, deep learning models rely on feature extraction and the training dataset's volume, sometimes under-fitting and over-fitting. Many also focused on activity detection techniques unsuitable for the intended usage due to privacy issues and the challenge of identifying persons

in a crowd. However, smartphones offer a valuable method for data collecting, except for vision-based approaches and other wearable sensors. Additionally, because both male and female individuals and interior and exterior environments were considered, the data obtained would be in various settings. Also, the employment of the GAN, utilized for rapid generation of new instances from existing ones, was not observed in the relevant studies. Furthermore, the 1 Hz data collection procedure was challenging in and of itself.

Associated Works

Due to the quick advancements in bioinformatics, medical science, and information technology, the challenge of determining human activity from sensor data using deep learning techniques has been extensively researched over the past several years for its various application areas. A Convolutional Neural Network (ConvNet) serving as an automated feature extractor and classifier for identifying human activities has been introduced, exploiting smartphone accelerometer and gyroscope sensor data.in[1, 2]. Even while it surpasses some cutting-edge techniques, achieving an accuracy of 94.79% using only raw sensor data and 95.75% using extra fast Fourier transform data from the HAR data set, it still hangs back to discriminate between stationary activities. Twelve human activities were classified utilizing Deep Belief Network (DBN), preceded by feature extraction and Kernel Principal Component Analysis (KPCA) applied on the smartphone's inertial sensors data[3]. Although it achieves an overall accuracy of 95.85%, surpassing traditional approaches like Artificial Neural Network and Support Vector Machine, the variation in proposed activities was less and addressing such a problem,Barna, Masum [4]focused on infrequent activities with 1Hz data collection frequency. In [5], which used deep-stacked auto-encoders [6] for the feature extraction portion, it is shown how essential smartphone sensors are for disclosing a person's health state. The results revealed 0.15 % and 0.5 % gains for auto feature extraction. With deep learning models being modified and sustaining accuracy of over 95 %, staging several diseases like Dementia, Diabetes, Depression, etc. have been done using HAR[7, 8]. Some researchers worked with categorizing different human activities while the subjects were in a car, on foot, tilting, or in static, with and without agent-based analysis (ABA)[9], employing Deep Recurrent Neural Network (DRNN) and ABA appeared to outperform by a greater accuracy of 4 %.The study byBatool, Khan [10] focuses on developing an ensemble deep learning model for analyzing human activities using wearable sensory data. The use of wearable sensors has become increasingly prevalent in activity recognition applications due to their convenience and effectiveness. The authors propose an ensemble approach that combines multiple deep learning models to enhance the accuracy and robustness of human activity analysis. By leveraging the complementary strengths of different models within the ensemble, they aim to achieve superior performance compared to individual models. This research contributes to the advancement of human activity recognition technology, particularly in real-world scenarios where accurate and reliable detection of activities is essential.Chopra, Zhang [11]propose a novel approach for human action recognition using multi-

stream fusion and hybrid deep neural networks. Human action recognition is a challenging task that requires processing information from multiple sources or modalities, such as video frames or sensor data. The authors introduce a fusion strategy that integrates information from different data streams to enhance the discriminative power of the model. They employ hybrid deep neural networks, which combine various neural network architectures to effectively capture spatio-temporal patterns in human actions. By leveraging multi-stream fusion and hybrid neural networks, this research aims to achieve state-of-the-art performance in recognizing complex human actions from diverse data sources.Asokan, Pathmanaban [12]investigate gait-based human activity recognition using hybrid neural networks. Gait analysis is a promising approach for identifying individuals based on their walking patterns, which can also be applied to activity recognition tasks. The authors propose a hybrid neural network model that integrates gait features with deep learning techniques to recognize various human activities. By incorporating gait-related information into the neural network architecture, the model can effectively differentiate between different activities based on subtle variations in walking patterns. This research contributes to advancing activity recognition technology by exploring the synergy between gait analysis and deep learning methodologies.

Although convolutional neural networks (CNN) continue to dominate this area of research, [13] proposed a multi-layer LSTM model which outperformed not only long-established machine learning models but also the CNN model in terms of less computation. Additionally, the performance of four LSTM networks was investigated for activity recognition purposes in[14], exploiting the smartphone's built-in inertial sensors. Also, that study's recommended 4-layer CNN-LSTM model surpassed previous LSTM networks with an enhanced accuracy rate of 2.24 %. The Gated Recurrent Unit (GRU) is another frequently utilized member of the DL family. The complexity of traditional RNN structures was muted by more sophisticated recurrent neural network topologies, such as LSTM [15, 16] and GRU models [17].Sowmiya and Menaka [18]conduct a comparative analysis of hybrid neural network models for human activity recognition using inertial measurement units (IMUs). IMUs are commonly used for capturing motion-related data in wearable devices, making them suitable for activity recognition applications. The authors investigate different hybrid neural network architectures that combine the strengths of multiple neural network components, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and/or traditional machine learning algorithms. Their study aims to identify the most effective model for accurately determining various human activities based on IMU data. This research contributes valuable insights into the design and selection of neural network architectures optimized for activity recognition tasks. A study in [19] showed that the GRU layers in a multi-input CNN-GRU model exhibited long-period dependencies in time-series data that were well preserved and outplayed its CNN-LSTM counterpart by obtaining accuracies of 96.20 %, 97.21 % and 95.27 % in three publicly available data sets. A bi-directional GRU was proposed in[20] and deployed on the UniMiB SHAR data set. Despite effectively classifying activities, the model hasn't experimented against data

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

augmentation or other conventional models. Research investigating the performance of four deep learning models, LSTM, DBN, bi-directional LSTM and CNN, on publicly available datasets showed that CNN gives the highest performance at lower space and time requirements.

On publicly accessible data sets, research conducting a comparison between the performance of four deep learning models - LSTM, DBN, bi-directional LSTM, and CNNfound that CNN performs best while requiring the least amount of time and space[21]. Conversely, GRU outperforms CNN by greater accuracy when compared based on HHAR, MHEALTH, USC-HAD, PAMAP2 and REALWORLD data sets. According to research by Alawneh, et al. [22], adding more data significantly enhances activity recognition. The findings suggested that the accuracy of GRU and LSTM models may be significantly improved by adding exponential smoothing to the data. However, an automatic hyper-parameter tuning was absent, which could greatly boost the whole process.

Materials and Methods

Given the inverse relationship between physical activity and mental Depression in humans[23], this study strives to conduct HAR by accumulating Smartphone sensor data and employing deep learning to identify 14 different depression symptomatic activities. Additionally, developing enhancements in HAR was focused on setting up a robust system. The study compared two widely recognized time series classification models, LSTM and GRU, formulated based on the Butterworth Low Pass filter as a data preprocessing step, the three versions of the primary datasets, and the proposed and publicly accessible datasets.

Specifying Depression Symptomatic Activities

Depression typically includes signs like a gloomy mood, disturbed sleep, loss of delight or enjoyment in activities, exhaustion, or trouble concentrating. The entire process relied on HAR, and the depressive symptomatic activities were determined first. The risk of death and other cardiovascular diseases rises with the severity of Depression in a graded manner. Furthermore, Depression brings behavioral changes such as smoking, poor diet, and inactivity[24]. Figure 1 displays the 14 distinct depression symptomatic activities taken into account for the study.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

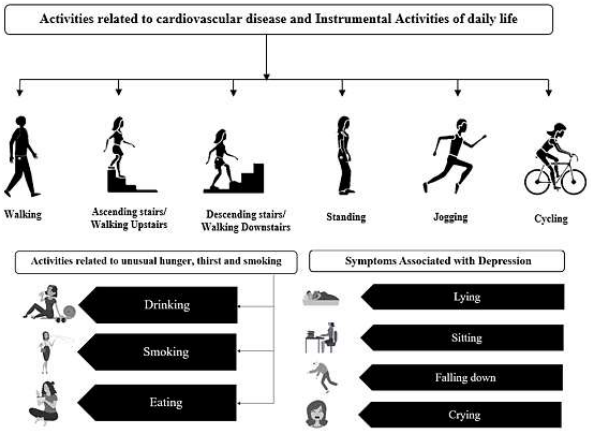

Figure 1: Depression Symptomatic Activities

Walking, sitting, jogging, walking upstairs, lying, walking downstairs, standing, drinking, cycling, eating, smoking, crying and falling down are the selected activities.

Data Collection from Sensors

Since human activities and smartphones are intimately related, the tri-axial accelerometer sensor on smartphones was employed to gather data and store it for categorization purposes. The embedded accelerometer sensor collected data for each activity along three axes during the labeled data collecting process. The 'Accelerometer Meter' application software was used, with data accumulation frequency kept at 1Hz, ensuring less power usage. Smartphones may be used for various tasks, including web surfing, texting, and making phone calls. They can also be placed on a table to charge or for other reasons. Moreover, the 14 activities listed as depressive symptoms and other activities can be carried out when the cell phones are closed. To prevent false positive and false negative data, such activities mentioned above are also collected and labeled as Irrelevant Activities.

No matter their age, height, gender, or weight, the fifteen youthful, active and in good physical shape volunteers participated in the data acquisition process. They individually collected data on every action for roughly nine minutes while holding their smartphones in their front pockets. In addition to building a flawless system, data was gathered in all conceivable positions, including flipped, upside down, and downside up, and both indoor and outdoor environments were considered.

Each volunteer collected data for each activity for about nine minutes through the selected android application. Upon completion of data collection, it was labeled with the activity name and saved in (standard separated value) CSV format. Fig. 2 depicts the whole data collection process.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

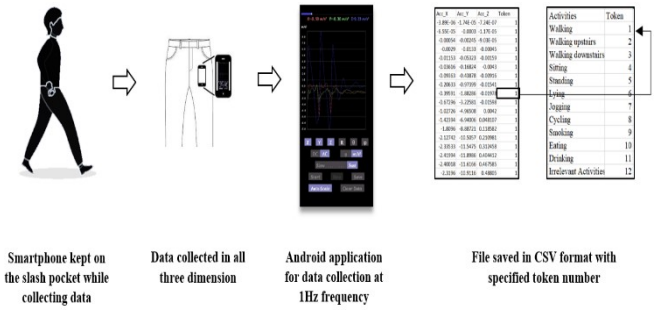

Figure 2 Data collection process

Generative Adversarial Network for Generating Sensor Data

Deep generative models, often referred to as Generative Adversarial Networks, commonly known as GAN, are commonly used in deep learning models that automatically discover and recognize broad patterns or patterns from the inserted data. The model is then becoming prepared to be employed to generate synthetic data. GAN model is comprised of a generator model and a discriminator model. The discriminator model decides whether the instances produced by the generator model are genuine, taken from the domain, or bogus. Figure 3 displays how these models are visualized.

Simultaneously, both models are being trained. The generator and supplies produce a set of samples to the discriminator as input with an actual example from the domain to determine whether they are true or false. The generator becomes adjusted based on how much the created instances deceived the discriminator, and the discriminator becomes upgraded for the subsequent step to increase the capacity to identify between actual and fake samples. If the optimization procedure works according to plan, the created distribution should finally match the actual distribution. The new distribution, sometimes called a zero-sum game, should reproduce the original distribution if everything proceeds as planned during the optimization phase. The generator is altered with significant modifications in model parameters when the discriminator correctly distinguishes actual samples from false ones or when no changes to the model parameters are needed. The model parameters are left alone if the discriminator is lying or the generator is compensated. However, the model's parameters are changed, and the discriminator is made to pay for the change.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

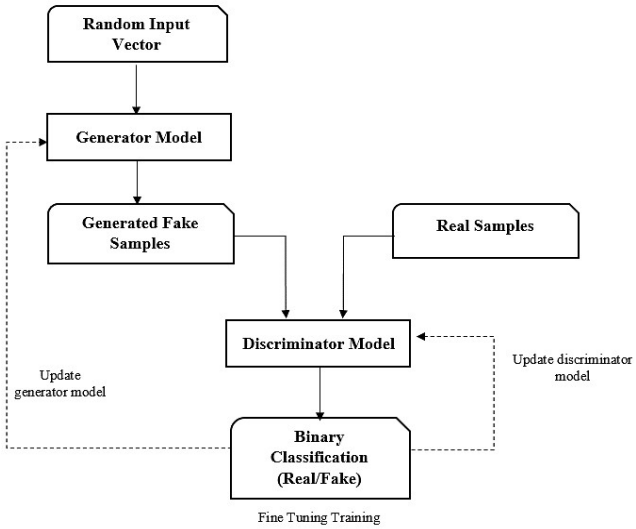

Figure 3: Architectural model of Generative Adversarial Networks

About 105,000 instances of 14 depressive symptomatic activities were accumulated via smartphone sensors and assembled to make up the primary dataset. Noting that GAN can be applied to tabular data, we deployed the Tabgan sampler to produce synthetic sensor data of 229,873 instances for the analysis. To make a proper comparison, the GRU and LSTM models were deployed on the actual sensors' dataset, then on the merged dataset of actual & generated sensors data, and lastly, the dataset with only generated sensors data. The two classification models were used on these three datasets after normalizing the parameters with Keras-tuner.

Data Pre-processing

Although a variety of trustworthy data acquisition procedures are readily available, the availability of appropriate and unnecessary details inside the data, null values or out-of-range, insufficient and disrupted data. Other factors can significantly impact the recovery of useful knowledge from sparsely processed data. Data analysis that hasn't been carefully reviewed for these problems may provide inaccurate findings. One of the most crucial data mining activities is data preprocessing, which is putting data into an appropriate format for mining operations. The independently created Android application has faults and problems and may overburden the execution process, producing inaccurate data.

The initial issue is the existence of null values that are predicted to arise due to incompatible device settings, problematic motion sensors, defective Android programming etc. It was a plus that the dataset had no null values, so no preprocessing steps were needed to account for missing or nonexistent values. Outliers in a dataset contributing to a deviation from the normal data distribution are said to have raucous data. Therefore, the Boxplot procedure was used to identify outliers. Also, the Interquartile range (IQR) method computed the upper and lower limit for a given region, and any observations lower than the lower limit and higher than the upper limit were replaced with the mean of that region. Moreover, since the dataset had characteristics of various ranges, Min-Max Normalization was used to bring all the attributes to the same scale to

help in data mining processes and prevent the efficacy of essential lower-scale features from being diluted. To scale the dataset between a range of zero and one, Min-Max Normalization was used since the dataset contains acceleration characteristics using arbitrary units, such as radians per second or square meters per second. The constant gravitational field interferes with the regular sensor signals, resulting in more considerable variations. The oscillating signal was split apart and transformed into a smooth sinusoidal waveform using the Fourier transformation. The Butterworth low-pass filter transforms a generated frequency into a smooth frequency when used as a low-pass filter. After utilizing the filtering process, we obtained a valuable result for the X-axis of the accelerometer sensor, as shown in Figure 4.

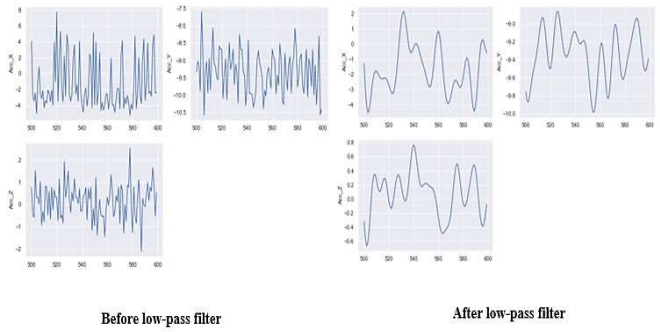

Figure 4: Using a Butterworth low pass filter (lower), asignal from accelerometer sensor data was smoothed out.

To conclude, one hot encoding reduces the limitations of label encoding by encoding the absolute values into whole integers. Each classification or class is given a binary value of 1, and a new column is created with the category's name as the column's name.

Classifier: Long Short-Term Memory

In the deep learning technique, RNNs may forecast the present time output in addition to processing input sequentially. However, RNN networks can only detect input briefly because of the vanishing gradient problem. If not flown deeply, the gradients may disappear using the deep network back propagation approach. Hochreiter and Schmidhuber[15] prove that adding LSTM provided a solution to the long-term dependence issue they were addressing. The hidden state of the recurrent network, which preserves past sequences that remains undissolved, is where memory cells recollect or keep long-period constraints of former inputs, according to the LSTM model.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

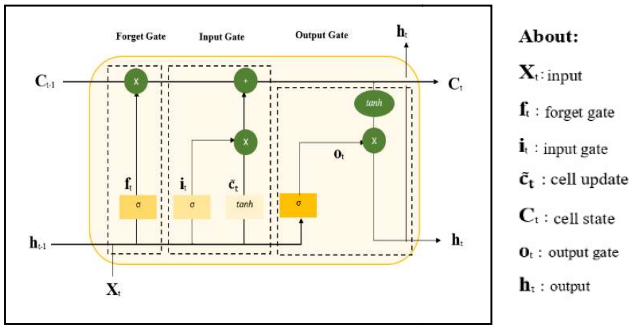

Figure 5: Long Short-Term Memory Cell

Cell and hidden states that transmit the data integrate slight adjustments via a sequence of addition and multiplication processes once the LSTM cells are constructed, with a reasonably large memory cell block for each. An LSTM unit's construction is shown in Fig. 5. Using channel multiplication, the forget gate purges inconsequential data from the network. The logistic function determines the data be dumped and to be preserved by combining the input for that phase (x_t) and the hidden state of the primary cell (h_{t-1}). However, after the forget state is finished, the cell states come to an end as follows:

$$C_t = f_t * C_{t-1} \qquad (1)$$

f_t indicates the forget, and C_{t-1} indicates previous cell states. The following expression refers to a forgetting state:

$$f_t = \frac{1}{(1+e^{-(W_f*[h_{t-1},x_t]+b_f)})} \qquad (2)$$

The b_f represents the bias vector, and the W_f represents the weight matrix. The output vector of the sigmoid function is added to an input gate after being multiplied by the cell state. The extra sigmoid function acts as a sieve that administers the values to be summed to the cell state, and a vector with a hyperbolic tangent function is generated by it, which contains all potential values. The outcome of the sigmoid layer and potential values for hyperbolic tangent function are then supplemented to the cell state. As a result, only necessary sequences, not unnecessary ones, are eligible to contribute to the cell state. Following the conclusion of the input state, the cell state can be expressed in the following way:

$$C_t = f_t * C_{t-1} + i_t * C'_t \qquad (3)$$

Where C'_t carries potential values and it input state retains the sigmoid layer. Such functions are expressed as:

$$i_t = \frac{1}{(1+e^{-(W_i*[h_{t-1},x_t]+b_i)})} \qquad (4)$$

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

$$C'_t = \frac{(1-e^{-2(W_c*[h_{t-1},x_t]+b_c)})}{(1+e^{-2(W_c*[h_{t-1},x_t]+b_c)})} \quad (5)$$

In such scenarios, b_i and b_c are biasing vectors, whereas W_c denotes intermediate cell state and Wi denotes weight matrices of the input cell. A succeeding filtration selects the necessary cell state components for formation. After establishing the cell state via hyperbolic tangent, the sigmoid function and vector are created. Vector output of the hyperbolic tangent function and the output of the regulatory filter are given to the hidden state. The only crucial sequences generated by the product are those. As a result, the output produces the following cell state:

$$C_t = o_t * h_t \quad (6)$$

The o_t represents the regulatory filter, and h_t represents the vector of hyperbolic tangent values. Such functions can be expressed as:

$$o_t = \frac{1}{(1+e^{-(W_c*[h_{t-1},x_t]+b_o)})} \quad (7)$$

$$h_t = \frac{o_t*(1-e^{-2(C_t)})}{(1+e^{-2(C_t)})} \quad (8)$$

The target gate's weight matrix is symbolized as W_c, and the bias vector is symbolized as b_o. The researchers claim that the concealed LSTM model's slight open connection must be adjusted. The spy hole link from internal cells to multiplicative gates was expanded to better understand the detailed information between spike sequences. As a consequence, forget, input, and output were allocated to the following gates:

$$f_t = \frac{1}{(1+e^{-(W_f*[h_{t-1},x_t]+b_f)})} \quad (9)$$

$$i_t = \frac{1}{(1+e^{-(W_i*[h_{t-1},x_t]+b_i)})} \quad (10)$$

$$o_t = \frac{1}{(1+e^{-(W_c*[h_{t-1},x_t]+b_o)})} \quad (11)$$

However, in a study[19], researchers integrated forget and input gates. The integrated gates decide what to preserve and what new information to acknowledge in this paradigm. In the forget gate, the fused model eliminates the input data series while supplementing new data into the state. The following expression can express the resulting cell state from this model:

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

$C_t = f_t * C_{t-1} + i_t * C'_t$□□□□□□□□□□□□□□□□□□□□□(12)

Classifier: Gated Recurrent Unit

The vanishing gradient issue that plagues RNN is also addressed by the GRU network developed by scholars in [24]. To control information sharing, gates are utilized in GRU, but unlike LSTM, it is deprived of a discrete cell state (C_t). There is just a concealed state (h_t). As a result, training or adaptation is quicker with GRUs due to their more straightforward design. The update and reset gates, which are part of GRUs, decide what data should be sent to the output. It is notable because they may be taught to retain data for a very long period without deleting unnecessary data from the forecast. Using the formula, the update gate z_t for time step t is computed.

□□□□□□□□□□□□□□□□□□□□□□$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1})$□□□□□□□□□□□(13)

The $W_z$weight is multiplied by the x_t connected to the network unit. Similarly, the weight of U_z is multiplied by h_{t-1}, extracting data from t-1 units. The weighted total derived from the earlier findings is then given a sigmoid function. By calculating the amount of information from preceding time-steps, to be repeated in the following stage to avoid vanishing gradient issues, update gates help the model. On the other hand, the amount of primary data to be discarded is decided by the reset gate. Its computation utilizes the following function:

$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1})$ (14)

Table 1: A Detailed description of the Architecture of 3-stacked LSTM and GRU Network

	LSTM		GRU	
State	Hyper-parameter	Tuned Value	Hyper-parameter	Tuned Value
Final Composition	LSTM Unit	0.20	GRU Unit	0.10
	Dropout Layer	416	Dropout Layer	305
	LSTM Unit	0.10	GRU Unit	0.00
	Dropout Layer	576	Dropout Layer	465
	LSTM Unit	0.00	GRU Unit	0.20
	Dropout Layer	992	Dropout Layer	781
	Dense Layer	928	Dense Layer	768
	Dropout Layer	0.70	Dropout Layer	0.70
	Dense Layer	896	Dense Layer	890
	Dropout Layer	0.30	Dropout Layer	0.20
Training Keras-Tuner	Loss Function	Categorical CrossEntropy	Loss Function	Categorical CrossEntropy
	Optimizer	Adam	Optimizer	Adam
	Batch-size	256	Batch-size	256
	Learning-Rate	variable	Learning-Rate	variable
	Number of Trials	90	Number of Trials	90

The contemporary memory content makes use of the reset gate to keep track of pertinent information from the past, as shown by the following equation:

$h'_t = tanh(Wx_t + r_t \odot Uh_{t-1})$ (15)

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

First, the product of a weight W and x_t is computed, and then a weight U is multiplied by h_{t-1}. The insignificant information deleted from the initial timesteps is computed using the element-wise product between the reset gate rt and Uh_{t-1}. After compiling the findings, the tanh (non-linear activation) is used. Finally, the network is to compute the h_t vector, storing and transmitting the data from the current unit to the network. The update gate decides what to gather from the present memory content $h'_t$and what to collect from the preceding steps and accomplishes this. The following equation demonstrates the whole process:

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t \ ) \qquad (16)$$

Element-wise multiplication is utilized for the update gates z_t and $h_{(t-1)}$, as well as $(1-z_t)$ and h'_t. Eventually, the outcomes are added together to produce ht. The expression will disappear, which means that the new hidden state will only include a small amount of information from the previous hidden state if z_t is close to 0. The second portion, however, combines, indicating that the candidate state will be the sole data in the concealed state at the present timestamp. Likewise, to first case, if the value of z_t on the second term is zero, the information from the hidden state at timestamp t-1 will be essential for understanding the present hidden state. It spans from 0 to 1 and shows how significant the variable z_t is in the equation. Additionally, r_t is given equal weight in the candidate's hidden state.

Proposed LSTM and GRU Classification Model

The dataset is modified to form a sliding window to produce an installed vector for phase assessment. The transformed dataset held three parameters: input size, time-steps and number of features, where the time-step was set to 20 and the number of features was three referring to Acc_X, Acc_Y, and Acc_Z. The input data is carried through the stacked GRU and LSTM layers unrolled overall time. Afterward, the hyper-parameter optimization for the proposed LSTM and GRU model was done with the aid of Keras-tuner, a package that assists in selecting the best set of hyper-parameters for the TensorFlow algorithm. Employing the Hyperband algorithm of Keras-tuner, the best combination of layer units, dropout ratio in each layer, activation function in each layer and learning rate were obtained.

The proposed LSTM and GRU models were comprised of three LSTM and GRU layers separately, where each layer was assigned to some dropouts, as shown in Table I. The final dense layer consisted of 14 units referring to the 14 activities. ReLU activation function was a component of each layer, with softmax allocated to the final dense layer. The softmax activation function organizes the multiclass properties. The optimizer utilized for the case was Adam. Furthermore, the number of instances supplied to the network at once, which is the batch size, was set to 256 for both models. Table I holds the hyper-parameters for the proposed stacked LSTM network and the hyper-parameters for the proposed stacked GRU network. Fig. 6 illustrates the entire working methodology of the two models.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

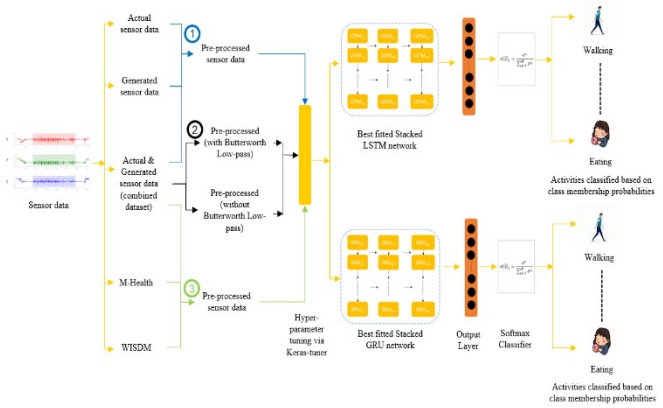

Figure 1 A visual abstract of the model at work

Results

To assess the functioning of the GRU and LSTM networks and address the HAR problem, the trials were modified. Including one dependent variable and three independent variables for 14 depression symptomatic activities. The GAN model increased the dataset, resulting in 229,873 instances of sensor-generated data. The two suggested models, GRU and LSTM, were used to evaluate the three different datasets: actual, generated, and the combination of actual and generated. In order to train the network, The datasets were split up into two sections. The first section had 70% of the data used as an input to the models, while the second section contained 30% of the data used to evaluate how well various datasets performed in the two separate models. Butterworth's low-pass filter was used during the data preprocessing stage to compare the two models. The assessment procedure was compared based on two categories: (a) using three different main dataset variants and (b) using a butter-worth low-pass filter to preprocess data. The bar chart in Figure 7 and Figure 8 presents the outcomes.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

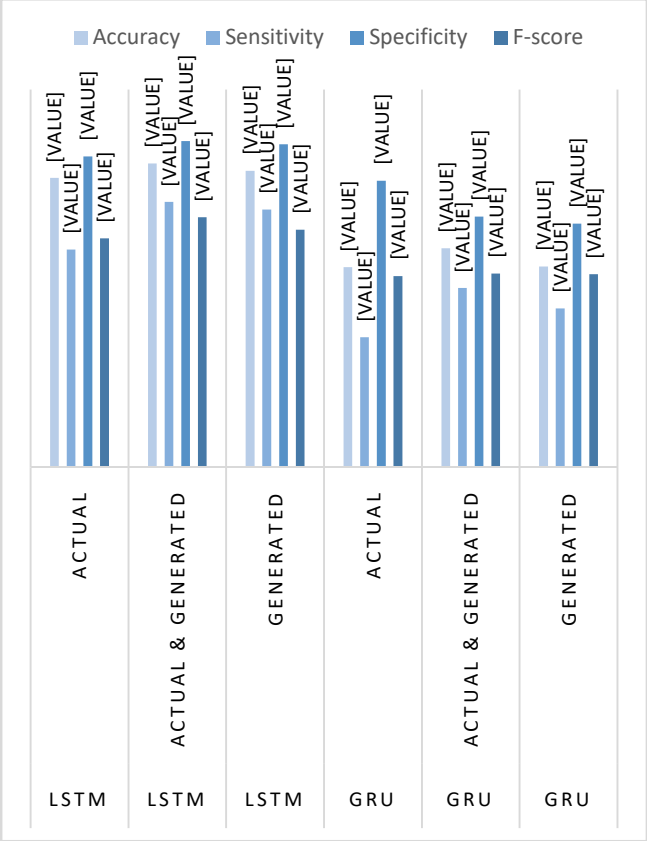

Figure 2 Performance evaluation based on three variations of primary datasets.

Figure 7 shows that the two models, GRU and LSTM, produced the most impressive outcomes from the merged dataset, which included both generated and actual data. Therefore, the merged dataset was only utilized, the Depression dataset, for future comparison. Additionally, the raw data was processed to obtain helpful information, and the Butterworth low-pass filter significantly contributed to that information. The data has been given synthetic coherence with a Butterworth low pass filter. Figure 8 illustrates its potency.

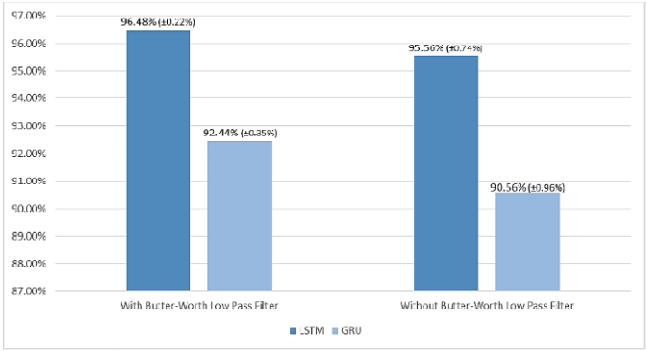

Figure 3 Performance evaluation developedon the applianceof Butterworth low pass filter

GRU and LSTM machine learning methods were used to evaluate their performance on the Depression dataset. With standard deviations of 0.22 % and 0.35 %, respectively, accuracy rates

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

of 92.44 % using the GRU technique and 96.48 % using the LSTM approach were attained. It was, therefore, evident that the LSTM model outperformed the GRU model. After developing a superior performance algorithm, we also employed it on several well-known HAR datasets and a dataset we created by compiling data from diverse activities. In this case, the Depression dataset has an accuracy of 96.48%, while other datasets have lower accuracy, as illustrated in Table II.

Table 2: Performance Assessmentwith LSTM and GRU Using Different Datasets.

Model Name	Dataset Name	Accuracy (%)	Standard Deviation (%)
LSTM	Depression Dataset	96.48	0.22
	M-Health	90.08	0.38
	WISDM	86.03	0.41
GRU	Depression Dataset	92.44	0.35
	M-Health	87.53	0.39
	WISDM	83.09	0.49

Discussion

A significant number of writers were involved in earlier work on HAR using the LSTM. The UCI HAR public dataset achieved an accuracy of 93.79%[25], 92%, 91.7% with their dataset, and so forth. Our suggested system has outscored prior research using the LSTM-Depression dataset combination by performing better than them. A new dataset and practical preprocessing techniques were used to improve performance in this area. Table II displays the accuracy and standard deviation figures for various dataset combinations.

Different results have been generated by the three stacked LSTM models for various dataset combinations. As of Figure 7, it achieved 95.79 % and 92.38 % sensitivity for a dataset created using only raw sensor data, 96.13 % accuracy and 94.28 % sensitivity for the dataset created using generated data, and 96.48 % accuracy and 94.65 % sensitivity for the dataset formed using both actual and generated data. On the other hand, the GRU model exhibits the same traits when applied to the dataset. For the actual dataset, the GRU model obtained 91.54 % and 88.21 % sensitivity; for the dataset including generated data, 91.57 % accuracy and 89.57 % sensitivity; and for the combined dataset, 92.44 % accuracy and 88.21 % sensitivity. Now it is evident that

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

both models performed better for the dataset prepared using a combination of generated and actual data. However, the LSTM model outperforms the GRU model for the same dataset with an incredible accuracy of 96.48%, which is 4.4% higher than the GRU's accuracy. The same goes for the sensitivity, as LSTM performed 4.1% better than GRU.

Data preprocessing is essential in getting effective results from a study. Butterworth low pass filtering is a data processing method that has the power to influence the outcome significantly. After applying the Butterworth low pass filtering to the dataset, a fair comparison has been performed between the effects of the GRU and LSTM models. The accuracy rates of the GRU and LSTM models on the Depression dataset were 95.56% (0.74%) and 90.56% (0.96%), respectively, before the application of Butterworth low pass filtering, which is shown in Figure 8. But following the processing, the LSTM and GRU models' accuracy climbed by 0.92% and 1.88%, correspondingly. Even after the Butterworth low pass filtering, the LSTM model still beat the GRU model in this case by 4.4%.

Subsequently, two recommended models, GRU and LSTM, were trained and evaluated on the combined primary dataset, the Depression dataset and two publicly accessible datasets, WISDM and M-Health, for activity recognition. The proposed LSTM model secured an accuracy of 96.48% (±0.22 %) on Depression dataset, 86.03% (±0.41 %) on the WISDM dataset and 90.08% (±0.38 %) on the M-Health dataset, as shown in Table II. In earlier investigations, LSTM models showed lower performance when deployed on WISDM and M-Health datasets, acquiring accuracies of 84.71% and 78.09%, respectively[25, 26], whereas the proposed LSTM model achieved a greater accuracy of 1.32% and 11.99% respectively. Also, the obtained results show that the recommended LSTM performs best when deployed on the Depression dataset. On the other hand, the GRU model secured an accuracy of 92.44% (±0.35 %) on the Depression dataset, 83.09% (±0.49 %) on the WISDM dataset and 87.53% (±0.39 %) on the M-Health dataset, as shown in Table II. Earlier studies show that the GRU model performs indisposed when deployed on WISDM and M-Health datasets[25]. Conversely, when deployed on WISDM and M-Health datasets, the proposed GRU model achieved higher accuracies than the previous GRU models. The results in Table II show that the recommended GRU model operated with the highest accuracy of 87.53%. As a result, it can be concluded that both the models work finest when deployed on the primary dataset, that is, the Depression dataset. Additionally, out of the two suggested models, LSTM surpassed the GRU model by a greater accuracy of 4.04%, which signifies that LSTM classifies the activities to a greater extent than GRU and earlier LSTM-based activity detection algorithms.

Conclusions

Around the world, 3.8% of people experience Depression, with adult prevalence rates of 5.0% and those over 60 years of age at 5.7%. Globally, 280 million people experience Depression. In its severe cases, Depression may lead to suicide, which results in the deaths of over 700,000 people each year. Hence, it takes time to identify Depression's symptomatic activity. Several

customizable approaches were applied to collect data on the Depression symptomatic activities from volunteers who had no medical conditions, were in good health, and were physically active. After completing the required preprocessing, the GRU and LSTM deep learning models were applied to the datasets, with the recommended LSTM model outperforming the GRU and previously used LSTM-based activity detection algorithms.

Both male and female volunteers accumulate data in different environments, both inside and outside, positioning the smartphone used for data gathering in various settings. Deep learning models perform better when the dataset is more extensive. For the data augmentation, we used GAN. A frequency of 1 Hz also provided exquisite precision. Additionally, deep learning models were utilized to identify the 14 depression symptomatic activities. But first, the classifier handles these particular data types to build a more dependable system. When someone is asleep, their smartphone may not always be with them. This very situation was left unexplored in the study. Therefore, a method for identifying if somebody is nodding off without their smartphone should be designed. This method requires users to keep a connection to the internet to collect and transfer smartphone data to a server. While we focused on smartphones, our research may benefit from integrating portable devices like smartwatches. We want to deal with the smartphone's isolation from the user in further development to avoid creating irrelevant data. Our next objective will be to determine how closely heart rate relates to Depression.

References

1. Ronao, C.A. and S.-B. Cho, *Deep Convolutional Neural Networks for Human Activity Recognition with Smartphone Sensors*, in *Neural Information Processing*. 2015. p. 46-53.
2. Ronao, C.A. and S.-B. Cho, *Human activity recognition with smartphone sensors using deep learning neural networks.* Expert systems with applications, 2016. **59**: p. 235-244.
3. Hassan, M.M., et al., *A robust human activity recognition system using smartphone sensors and deep learning.* Future Generation Computer Systems, 2018. **81**: p. 307-313.
4. Barna, A., et al. *A study on human activity recognition using gyroscope, accelerometer, temperature and humidity data*. in *2019 international conference on electrical, computer and communication engineering (ecce)*. 2019. IEEE.
5. Li, H. and M. Trocan, *Deep learning of smartphone sensor data for personal health assistance.* Microelectronics Journal, 2019. **88**: p. 164-172.
6. Almaslukh, B., J. AlMuhtadi, and A. Artoli, *An effective deep autoencoder approach for online smartphone-based human activity recognition.* Int. J. Comput. Sci. Netw. Secur, 2017. **17**(4): p. 160-165.
7. Bahadur, E.H., et al., *Active sense: early staging of non-insulin dependent diabetes mellitus (NIDDM) hinges upon recognizing daily activity pattern.* Electronics, 2021. **10**(18): p. 2194.

8. Barua, A., et al., *Human activity recognition in prognosis of depression using long short-term memory approach.* International Journal of Advanced Science and Technology, 2020. **29**: p. 4998-5017.
9. Javed, A.R., et al., *A smartphone sensors-based personalized human activity recognition system for sustainable smart cities.* Sustainable Cities and Society, 2021. **71**: p. 102970.
10. Batool, S., M.H. Khan, and M.S. Farid, *An ensemble deep learning model for human activity analysis using wearable sensory data.* Applied Soft Computing, 2024: p. 111599.
11. Chopra, S., L. Zhang, and M. Jiang. *Human Action Recognition Using Multi-Stream Fusion and Hybrid Deep Neural Networks*. in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2023. IEEE.
12. Asokan, R., et al. *Gait Based Human Activity Recognition using Hybrid Neural Networks*. in *2023 12th International Conference on Advanced Computing (ICoAC)*. 2023. IEEE.
13. Yu, T., et al. *A multi-layer parallel lstm network for human activity recognition with smartphone sensors*. in *2018 10th International conference on wireless communications and signal processing (WCSP)*. 2018. IEEE.
14. Mekruksavanich, S. and A. Jitpattanakul, *Lstm networks using smartphone data for sensor-based human activity recognition in smart homes.* Sensors, 2021. **21**(5): p. 1636.
15. Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural computation, 1997. **9**(8): p. 1735-1780.
16. Bahadur, E.H., et al. *LSTM based approach for diabetic symptomatic activity recognition using smartphone sensors*. in *2019 22nd International Conference on Computer and Information Technology (ICCIT)*. 2019. IEEE.
17. Chung, J., et al., *Empirical evaluation of gated recurrent neural networks on sequence modeling.* arXiv preprint arXiv:1412.3555, 2014.
18. Sowmiya, S. and D. Menaka, *Comparative Analysis of Various Hybrid Neural Network Models to Determine Human Activities using Inertial Measurement Units.* International Journal of Intelligent Systems and Applications in Engineering, 2024. **12**(12s): p. 585-599.
19. Dua, N., S.N. Singh, and V.B. Semwal, *Multi-input CNN-GRU based human activity recognition using wearable sensors.* Computing, 2021. **103**(7): p. 1461-1478.
20. Alsarhan, T., et al. *Bidirectional gated recurrent units for human activity recognition using accelerometer data*. in *2019 IEEE SENSORS*. 2019. IEEE.
21. Sansano, E., R. Montoliu, and O. Belmonte Fernandez, *A study of deep neural networks for human activity recognition.* Computational Intelligence, 2020. **36**(3): p. 1113-1139.
22. Alawneh, L., et al., *Enhancing human activity recognition using deep learning and time series augmented data.* Journal of Ambient Intelligence and Humanized Computing, 2021: p. 1-16.

23. Lu, Y. and F.M. Salem. *Simplified gating in long short-term memory (lstm) recurrent neural networks*. in *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*. 2017. IEEE.
24. Cho, K., et al., *On the properties of neural machine translation: Encoder-decoder approaches.* arXiv preprint arXiv:1409.1259, 2014.
25. Shiranthika, C., et al. *Human activity recognition using CNN & LSTM*. in *2020 5th International Conference on Information Technology Research (ICITR)*. 2020. IEEE.
26. O'Halloran, J. and E. Curry. *A Comparison of Deep Learning Models in Human Activity Recognition and Behavioural Prediction on the MHEALTH Dataset*. in *AICS*. 2019.