Analysis of Machine Learning Models for Wearable Devices Centric Human Activity Recognition

Conference Paper · June 2024

DOI: 10.1109/ICAAIC60222.2024.10575575

Proceedings of the 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC-2024)
IEEE Xplore Part Number: CFP24BC3-ART; ISBN: 979-8-3503-7519-0

2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) | 979-8-3503-7519-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICAAIC60222.2024.10575575

Analysis of Machine Learning Models for Wearable Devices Centric Human Activity Recognition

Namia Jalal Zara
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
namia.zara@northsouth.edu

Rashiqur Rahman Rifat
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
rashiqur.rifat@northsouth.edu

Jarin Tasnim
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
jarin.tasnim10@northsouth.edu

Md. Baker
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
md.baker@northsouth.edu

Riasat Khan
Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
riasat.khan@northsouth.edu

***Abstract*—This work presents a comprehensive study on applying machine learning models for human activity recognition. The increasing number of wearable devices and the growing interest in health monitoring and intelligent environments have fueled the demand for robust and accurate systems that automatically recognize human activities. This study explores and compares the performance of different machine learning (ML) models in human activity recognition. This study intends to identify the most effective ML model to classify and predict human activities accurately. In this study, various models are evaluated, including ensemble-based classifiers such as Decision Tree, Random Forest, K-Nearest Neighbor, Logistic Regression, AdaBoost algorithm, and XGBoost algorithm. All models have been evaluated using the publicly available dataset "Human Activity Recognition with Smartphones," which captures six daily activities: Lying, Standing, Sitting, Walking, Walking Upstairs, and Walking Downstairs. Among all the applied models, XGBoost accomplished the highest accuracy of 99.52% with 100% precision, 100% recall (100%), and 1.0 F1 score. Hyperparameter tuning on the ML models is implemented to attain the best accuracy.**

***Keywords*—Adaptive Boosting, Decision Tree, Explainable AI, Human Activity Recognition, Random Forest, XGBoost.**

I. INTRODUCTION

Human activity recognition (HAR) aims to automatically identify and classify human behaviors from wearable devices. The method of analyzing human motion with computer vision and other technologies is known as HAR. Sensor-recorded actions, gestures, or behaviors can be understood as human motion. HAR can be carried out via sensors, smartphones, or pictures [1]. The present models for activity recognition, such as recurrent, convolutional, or hybrid models, have difficulty extracting spatiotemporal context from the feature space of sensor reading sequences [2]. Since cooperation between robots and people needs artificial agents to recognize complicated environmental cues, HAR is a crucial problem in human-robot interaction. In this work, we have used several machine learning models to evaluate the activity of a human by wearable devices. Human activity recognition provides advantages such as personalized healthcare monitoring and fitness training, adaptive smart environments, enhanced security through intruder detection, and improved industrial safety by identifying hazardous activities.

Some recent articles on machine learning-based human activity recognition are briefly discussed in the subsequent paragraphs. For instance, Mahajan et al. [3] studied human activity recognition from videos using K-means and ANN models. The authors used a private dataset, and they made it using different devices. Most of those videos comprised noises, so they preprocessed activity videos. They also used mean squared error (MSE) to evaluate the performances of the mentioned models. The applied artificial neural networks attained an accuracy of 98.24%. Gumaei and colleagues [4] studied human activity from body sensors using deep learning models like RNN (GPU, SRU). They used the MHEALTH dataset, which is a public dataset. Body motion and different postures of 10 subjects were used in this dataset. They used precision and F1-score to evaluate the performances. They had an accuracy of 99.80% for deep SRUs-GRUs and 94.145% for deep SRUs.

Liu et al. [5] implemented new methodologies for 3D human action analysis. Large-scale human activity recognition datasets, comprising 56,000 video tests and 4 million edges, were gathered from 40 participants. There are sixty activity classes available, from daily to wellness focused. For the Berkeley MHAD dataset, the authors achieved 100% accuracy using HBRNN techniques. Nicolas et al. [6] studied human activity recognition using UCI-HAR dataset containing six activities and a light residual network for wearable sensor data. Their goal of using light residual network (ResNet) is to perform tasks such as modifying ResNet18 and reducing the complexity of the model and parameters.

Ghazal et al. [7] studied human activity classification utilizing 2D skeleton data. Two publicly available datasets were used for the performance evaluation. The authors evaluated the performance of various classifiers, including feed-forward, KNNs, SVM, discriminant analysis, and Naive Bayes. Naive Bayes achieved 90% accuracy in the first set of experiments, followed by SVM at 76%. Linear discriminant analysis (LDA) excelled in classifying the falls category, while Naive Bayes and KNN achieved 100% recall for standing activity. After data normalization, all classifiers showed improved efficiency, with KNN reaching 98% accuracy. Pathan and his colleagues [8] used an 8-second time series window with 75% overlap for frame-by-frame physiological and inertial data analysis, using wavelet packet transform for feature extraction. The Random Forest classifier achieved 97.8% accuracy in recognizing human activities, outperforming a previous SVM-CNN-based method, demonstrating efficiency and improved performance.

Leema and colleagues [9] studied human activity recognition using machine learning techniques and a private dataset with various sensor devices. To analyze human activity, they used a variety of machine classification algorithms. Manoj et al. [10] employed PCA and LDA techniques to study human activity recognition using the UCI ML dataset. The findings indicate that LDA outperforms PCA. The SVM technique attained an accuracy of 96.23% with optimized kernel. While dimensionality has improved model performance, simple datasets have shown low accuracy.

It can be inferred from the paragraphs above that there has been a lot of study on human activity recognition. Still, the number of comparisons between the machine learning models could be higher. These researchers made accurate predictions by utilizing a range of machine-learning approaches. Most of these studies examined the predictions made by machine learning models using various publicly and privately available datasets without using any comprehensible artificial intelligence techniques. These studies have thus recommended a variety of open-source and original datasets together with explainable AI techniques.

This work uses several machine learning models to predict possible human activity by wearable devices and compare the accuracy of the machine learning models. We used an open-source dataset from an online reliable repository with 563 features to identify whether a person is walking, lying, standing, sitting, walking upstairs, or downstairs. We have used multiple machine learning models in this research, such as decision tree, random forest, KNN, logistic regression, etc. Hyperparameter tuning for various ML models has been initiated by using GridSearchCV and RandomizedSearchCV frameworks. Finally, the explainable AI library, LIME, has been implemented to evaluate the outcome and identify the main contributing components that impact the prediction process.

The description of the dataset, dataset preprocessing, and machine learning models and architecture applied in the study are provided in Section II, which also discusses the proposed system. Section III presents the outcomes and a discussion of the results of the applied machine learning models. The comparative analysis of this research is shown in Section IV. And in Section V, the paper is conclude and presented some future works.

II. PROPOSED SYSTEM

In this part, the employed dataset is described, preprocessing system, and different models that have been used used in our work.

A. Dataset

In this work, an Android smartphone-based human activity recognition open-source dataset with six classes has been used [17]. Various accelerometers and gyroscope sensors have been employed to record the data.

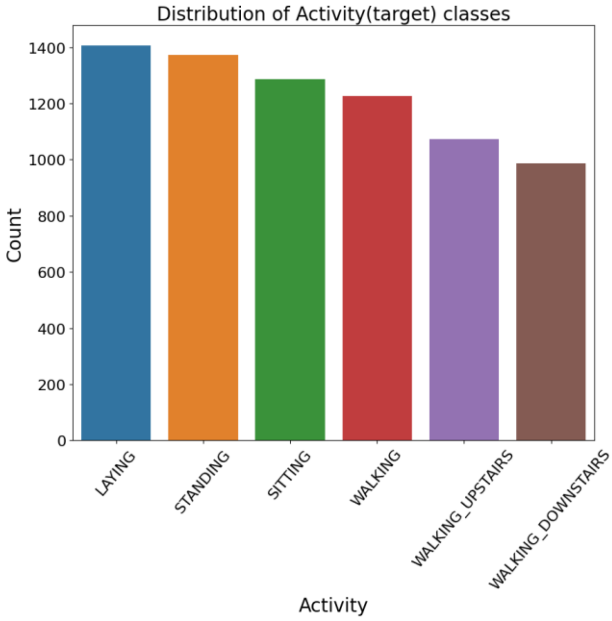

Fig. 1. Number of instances of various human activities in the employed dataset.

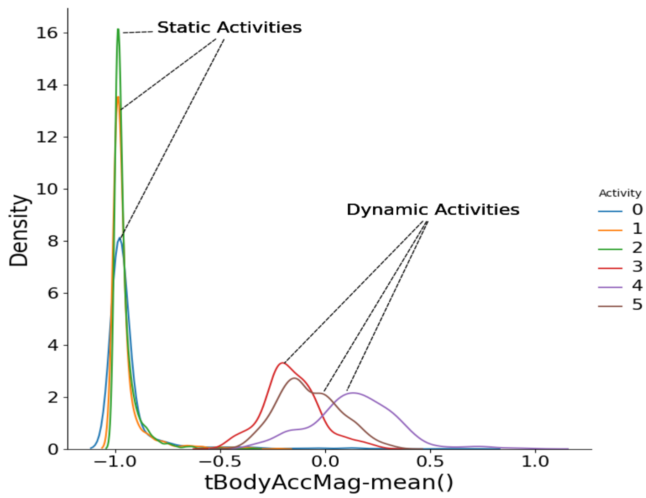

Fig. 2. Density functions of static and dynamic activities.

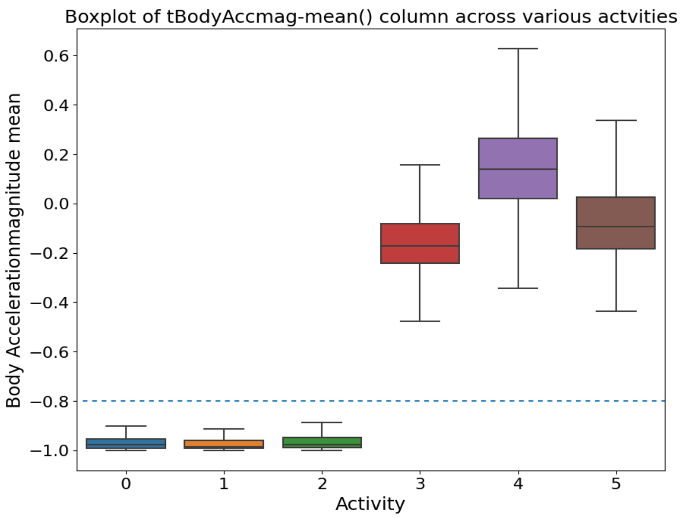

Fig. 3. Box plots of various human activities of the employed dataset.

Fig. 1 illustrates the distribution of instances across various human activities in the utilized dataset, i.e., laying, standing, walking, sitting, etc. As depicted in the figure, the dataset exhibits a balanced composition with an approximately equal number of instances across all activity classes. Density functions for distinct categories of both static and dynamic activities are presented in Fig. 2. Additionally, Fig. 3 displays box plots representing various human activities within the employed dataset.

B. Dataset Preprocessing

The employed dataset is balanced with no null or missing and duplicate values. Next, we use label encoding to transfer the categorical value to the numerical value. In the employed dataset, our target class, which means the "Activity," has a categorical value, which we changed into numerical values 0 to 5 using label encoding because we have six classes in the dataset. We also used 'MinMaxScaler' as feature scaling the categorical variables.

C. Preparing Machine Learning Algorithms

1) Hyperparameter tuning: This process provides the optimal model parameters by fine-tuning the hyperparameter. Unlike GridSearchCV, RandomizedSearchCV is a machine-learning method for tuning hyperparameter combinations randomly from predetermined ranges. This method preserves the search for the best model configurations while reducing computing expenses. RandomizedSearchCV effectively searches hyperparameter spaces by arbitrarily choosing combinations, offering a decent compromise between exploitation and exploration. This approach offers a more useful substitute for thorough grid searches when the search space is large. Through the use of a resource-efficient, randomized search procedure, RandomizedSearchCV improves model performance and aids in the optimization of machine learning models.

2) Explainable AI: Explainable artificial intelligence [18] is a technique that is helpful for people who are not experts in machine learning or need help understanding the output of algorithms. It is mainly used to explain any AI model. In our work, we applied LIME, which is very user-friendly.

3) LIME: "Local Interpretable Model-agnostic Explanations" is known as "LIME". The LIME approach uses a locally approximal interpretable model to faithfully explain the predictions of any classifier or regressor. In our work, we used it to understand the applied models.

D. Machine Learning Models

1) Decision tree: A decision tree which is a widely used machine learning technique because it can classify dependent variables into different classes [11]. It scans the data and identifies the most important independent factors. That is a simple structure where evaluation results are displayed in terminal nodes and checks for one or more functions are displayed in non-terminal nodes. Previous research on human action detection using decision tree classifiers has produced encouraging results.

2) Random forest: The random forest ensemble understanding technique creates many decision trees during training [12]. A subset of the bagging technique called the RF performs better when noise and poor discriminating data are not affected by parameter initialization overfitting [14]. With its excellent accuracy, ability to handle big datasets with plenty of features, and ability to provide feature-importance insights, random forest is very useful for classification and regression applications.

3) Logistic regression: This fundamental machine-learning technique is used in binary classification [13]. Despite its name, its purpose is to calculate the probability that an instance will fall into a given class. The program simulates this chance using the logistic function, which adjusts the input data linearly. Logistic regression is a versatile tool that may be applied in many different settings because of its effectiveness, interpretability, and simplicity. Predictions are based on a threshold, and coefficients are optimized by maximum likelihood estimation. In many machine learning contexts, the fundamental logistic regression method is utilized in finance and medicine to predict binary outcomes.

4) AdaBoost: AdaBoost technique is especially effective with weak learners or models that are better than random chance. It constructs a series of weak learners to correct errors and iteratively focuses on misclassified instances, assigning higher weights to them. A weighted average of these learners is used to make the final prediction, resulting in a robust and accurate model.

5) KNN: The KNN model generates predictions by comparing newly collected data points to labelled data points in the training set [15]. The algorithm determines the K nearest neighbors using weighted averaging or majority voting, computes distances, and forecasts the class label for classification or regression value. Depending on the parameter K selection, KNN is an instance-based, non-parametric model.

6) XGBoost: XGBoost model is trained by minimizing a specific objective function, incorporating regularization terms for enhanced predictive performance. XGBoost employs a novel regularization term, the "eXtreme Gradient," that helps control model complexity. It is widely used for classification, regression, and ranking tasks due to its efficiency, speed, and ability to handle large datasets. The XGBoost model has been developed to optimize multimode fiber architecture and to diagnose cause-aware network failures with improved estimate and precision. XGBoost has become a go-to algorithm in machine learning competitions and is known for achieving state-of-the-art results [16].

Fig. 4. Working sequences of the proposed HAR system.

Fig. 4 shows the working sequences of the proposed human activity detection system. In this work, we have used a public dataset to apply the machine learning models. Before applying the ML models, we pre-processed the dataset. Next, exploratory data analysis (EDA) is applied to analyze the data. Next, the data is split into the train (80%), test (20%) data, and implied feature scaling. After doing all these processes, we applied machine learning models and demonstrated the classification results. For better performance of the models, the hyperparameter is applied for tuning and evaluated the model. Finally, LIME-based explainable AI technique is applied to interpret the results.

III. RESULTS AND DISCUSSION

This section discusses the results of different machine learning algorithms for human activity recognition. six machine learning models are applied in the employed human activity recognition with smartphones dataset containing 563 features, 7,352 instances and six classes.

TABLE I. HYPERPARAMETER VALUES' RANGES FOR VARIOUS ML MODELS

Model	Hyperparameter Value Range	Optimized value
XGBoost	estimators #: [100, 200, 300], maximum depth: [3, 4, 5], learning rate: [0.01, 0.1, 0.2], sub sample: [0.7, 0.8, 0.9], % of features per tree: [0.7, 0.8, 0.9], objective: ['binary: logistic', 'multi: softmax']	subsample: 0.9, objective: 'multi: softmax', estimators #: 300, maximum depth: 5, learning rate: 0.1, % of features per tree: 0.9
Random Forest	estimators #: [100, 200, 300], Maximum depth: [None, 10, 20, 30], Min samples per split: [2, 5, 10], Min samples per leaf: [1, 2, 4], criterion: ['Gini', 'entropy']	estimator #: 200, min samples per split: 5, min samples per leaf: 1, max depth: None, criterion: 'entropy'
Logistic Regression	'C': [0.001, 0.01, 0.1, 1, 10], 'penalty': ['l1', 'l2'], 'solver': ['liblinear', 'newton-cg', 'lbfgs', 'saga']	'C': 10, 'penalty': 'l1' 'solver': 'liblinear'
Decision Tree	max features: [auto, sqrt, log2], max depth: np.linspace (20, 1000, 50), min samples per split: [2, 5, 10, 14], min samples per leaf: [1, 2, 4, 6, 8], criterion: [entropy, Gini]	max features: [sqrt], max depth: 520, min samples per split: 10, min samples per leaf: 4, criterion: entropy
KNN	p: [1,2], neighbors #: [1-30], leaf size: [1-50]	p: 1, neighbors #: 1, leaf size: 1
AdaBoost	estimators #: [50, 200, 500], learning rate: [0.01, 0.1, 1.0],	estimators #: 200, learning rate: 0.01

Table I illustrates the hyperparameter values' ranges and the corresponding optimized values for all the applied machine learning models.

TABLE II. PERFORMANCE METRICS OF VARIOUS CLASSIFIERS FOR DEFAULT PARAMETERS

Classifier	Precision	Recall	F1 Score	Accuracy
XGBoost	99%	99%	99%	99.39%
Random Forest	98%	98%	98%	98.23%
Logistic Regression	99%	99%	99%	99%
Decision Tree	94%	94%	94%	94%
KNN	96%	96%	96%	95.93%
AdaBoost	88%	80%	75%	80.16%

Table II illustrates the performance metrics of various classifiers with default parameters. As per this table, XGBoost outperforms all the machine learning models with the highest accuracy of 99.39% and AdaBoost has the lowest accuracy of 80.16%.

TABLE III. PERFORMANCE METRICS OF VARIOUS CLASSIFIERS WITH OPTIMIZED HYPERPARAMETERS

Classifier	Precision	Recall	F1 Score	Accuracy
XGBoost	100%	100%	100%	99.52%
Random Forest	99%	99%	99%	98.64%
Logistic Regression	99%	99%	99%	99%
Decision Tree	94%	94%	94%	93.55%
KNN	98%	98%	98%	98.31%
AdaBoost	97%	97%	97%	96.61%

Table III summarizes the performance metrics of various classifiers with optimized hyperparameters. As expected, the performances of all the applied models improve with optimized hyperparameters. According to this table, XGBoost outperforms all the machine learning models with the highest accuracy of 99.52% and decision tree has the lowest accuracy of 93.55 percent.

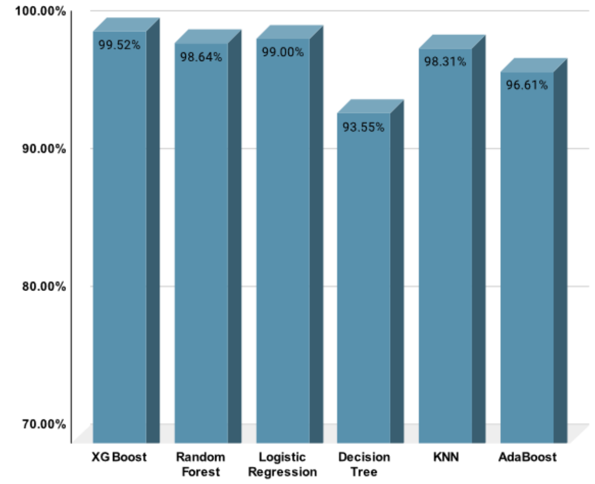

Fig. 5. Accuracy of all ML models with optimized hyperparameters.

Fig. 5 shows the accuracy of the ML models in the form of a bar graph with optimized hyperparameters using GridSearchCV or RandomizedSearchCV. It shows that XGBoost has the best accuracy of 99.52% for the dataset and Decision Tree has the lowest accuracy of 93.55%.

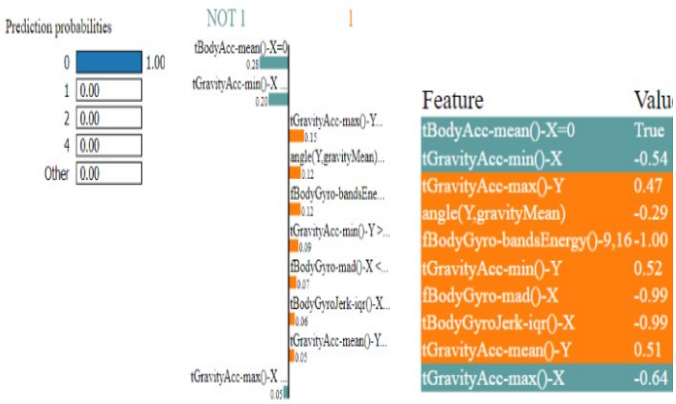

Fig. 6. Machine learning model prediction interpretation by LIME explainable AI library.

Fig. 6 depicts the interpretation of the XGBoost model's predictions using the LIME explainable AI framework. In the case of a particular sample, the prediction for the walking class is attributed to triaxial body acceleration and gravity acceleration, as highlighted in the first two rows of the feature table.

TABLE IV. COMPARISON OF THIS WORK WITH SIMILAR HUMAN ACTIVITY RECOGNITION SYSTEMS

Ref.	Dataset	Model	Accuracy	Recall
[3]	Private	ANN	98.24%	N/A
[4]	MHEALTH	RNN	99.80%	N/A
[5]	MHAD	HBNN	100%	N/A
[6]	UNI-HAR	ResNet	97.60%	N/A
[7]	Public	SVM	98%	100%
[8]	Public	Random Forest	97.80%	N/A
[9]	Private	Random Forest	99.55%	N/A
[10]	Public (UCI ML)	SVM	96.78%	N/A
This work	Smartphone -based	XGBoost	99.52%	100%

Table IV presents a comparative analysis between this study and other human activity recognition systems. Notably, the performance of this work is outstanding, leveraging the XGBoost ensemble technique with optimized hyperparameters for the smartphone-based human activity detection dataset.

IV. CONCLUSIONS

This research analyzes human activity recognition by extensively exploring machine learning models applied to data from wearable devices. The study employs diverse classifiers, assessing their performance under default and hyperparameter-tuned settings. The results showcase XGBoost's superior performance, emphasizing its adaptability and resilience. However, the study also highlights the nuanced strengths of other models, such as the robustness of logistic regression and random forest and the competitive accuracy of decision tree and KNN. The significance of hyperparameter tuning is underscored, with notable improvements observed across models, particularly in the case of AdaBoost. Moreover, the research integrates Explainable AI, utilizing LIME to interpret the predictions of the logistic regression model. This emphasis on interpretability contributes to the broader learning of how machine learning models arrive at their predictions. The study concludes by emphasizing the need for tailored approaches in configuring models, acknowledging the specific nuances of datasets, and recognizing the ongoing refinement required for optimal model performance. Overall, this work advances the understanding of HAR and provides valuable insights into the intricate interplay between machine learning models, hyperparameter tuning, and interpretability in the context of wearable device-centric human activity recognition.

In future, the feature engineering techniques can be refined further for extracting relevant information from sensor data. Exploring advanced model architectures, including convolutional and recurrent neural networks, will enhance the recognition accuracy.

REFERENCES

[1] K. Suwannarat and W. Kurdthongmee, "Optimization of deep neural network-based human activity recognition for a wearable device," Heliyon, vol. 7, 2021.

[2] K. V. Kumar and J. Harikiran, "Privacy preserving human activity recognition framework using an optimized prediction algorithm," International Journal of Artificial Intelligence, vol. 11, pp. 254-264, 2022.

[3] S. Mahajan and H. Amhia, "Improvement of Video Human Activity Recognition by using ANN Approach," International Conference on Communication and Signal Processing, pp. 829-832, 2019.

[4] A. Gumaei, M. M. Hassan, A. Alelaiwi and H. Alsalman, "A Hybrid Deep Learning Model for Human Activity Recognition Using Multimodal Body Sensing Data," IEEE Access, vol. 7, pp. 99152-99160, 2019.

[5] J. Liu, A. Shahroudy, D. Xu, and G. Wang, "Spatio-temporal LSTM with trust gates for 3D human action recognition," European Conference on Computer Vision, pp 816–833, 2016.

[6] F. M. C. Nicolás and O. M. Mozos, "Light Residual Network for Human Activity Recognition using Wearable Sensor Data," IEEE Sensors Letters, vol. 7, pp. 1-4, 2023.

[7] S. Ghazal, U. S. Khan, M. M. Saleem, N. Rashid, and J. Iqbal, "Human activity recognition using 2D skeleton data and supervised machine learning," IET Image Processing, vol. 13, pp. 2572–2578, 2019.

[8] N. S. Pathan, M. T. F. Talukdar, M. Quamruzzaman and S. A. Fattah, "A Machine Learning based Human Activity Recognition during

Physical Exercise using Wavelet Packet Transform of PPG and Inertial Sensors data," International Conference on Electrical Information and Communication Technology, 2019, pp. 1-5.

[9] Z., Gulzar, A. A. Leema and I. Malaserene, "Human Activity Analysis Using Machine Learning Classification Techniques." International Journal of Innovative Technology and Exploring Engineering, vol. 9, 2019, pp. 3252–3258,

[10] M. Kumar, *et al.*, "Dimensionality Reduction-Based Discriminatory Classification of Human Activity Recognition Using Machine Learning." Lecture Notes in Networks and Systems, 2022, pp. 581–593.

[11] M. T. Huyut and Z. Huyut, "Effect of ferritin, INR, and D-dimer immunological parameters levels as predictors of COVID-19 mortality: A strong prediction with the decision trees," Heliyon, vol. 9, 2023.

[12] N. E. J. Asha, E. U. Islam and R. Khan, "Low-Cost Heart Rate Sensor and Mental Stress Detection Using Machine Learning," International Conference on Trends in Electronics and Informatics, pp. 1369-1374, 2021.

[13] F. Hidayat and T. M. S. Astsauri, "Applied random forest for parameter sensitivity of low salinity water Injection (LSWI) implementation on carbonate reservoir," Alexandria Engineering Journal, vol. 61, pp. 2408-2417, 2022.

[14] I. Tasin, T. U. Nabil, S. Islam and R Khan, "Diabetes prediction using machine learning and explainable AI techniques", Healthcare Technology Letters, vol. 10, 2023.

[15] M. M. Islam, M. A. Kashem and J. Uddin, "Fish survival prediction in an aquatic environment using random forest model", International Journal of Artificial Intelligence, vol. 10, pp. 614-622, 2021.

[16] S. Siddique, S. Islam, E. E. Neon, T. Sabbir, I. T. Naheen and R. Khan, "Deep learning-based Bangla sign language detection with an edge device," Intelligent Systems with Applications, vol. 18, 2023.

[17] R. O. Jorge, A. Davide, G. Alessandro, O. Luca, and P. Xavier, "Human Activity Recognition Using Smartphones," UCI Machine Learning Repository, 2012.

[18] S. Solayman, S. A. Aumi, C. S. Mery, M. Mubassir and R. Khan, "Automatic COVID-19 prediction using explainable machine learning techniques," International Journal of Cognitive Computing in Engineering, vol. 4, pp. 36-46, 2023.