Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 7 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i07.61221

PAPER

A Data-Driven Framework for Stress Detection from Wearable Sensors in Smart Healthcare

Hayat Jebbar[1] (✉),
Mohamed El Mohadab[2],
Omar Boutkhoum[1]

[1]Chouaib Doukkali University, El Jadida, Morocco

[2]Sultan Moulay Slimane University, Beni Mellal, Morocco

jebbar.hayat@ucd.ac.ma

ABSTRACT

With the development of technology and increasing demands, the integration of Internet of Things (IoT) and wearable sensors into urban infrastructure including smart healthcare becomes a pillar of smart cities. This work presents a data-driven framework for stress level predicting by using physiological and behavioral data collected from wearable devices. The methodology proposed follows four key stages: (1) data preprocessing, (2) feature engineering, (3) model development and (4) metrics evaluation. Also, four machine learning (ML) algorithms are trained and tested on a multi-modal dataset including: decision tree (DT), random forest (RF), light gradient boosting machine (LightGBM), and Categorical Boosting algorithm (CatBoost). The results show that boosting algorithms, particularly LightGBM and CatBoost, outperformed the traditional DT and achieved the highest ROC-AUC values (0.6354 and 0.6326) respectively. This study underlines the importance of integrating wearable technologies and ML in order to enhance preventive healthcare within smart cities.

KEYWORDS

Internet of Things (IoT), stress-level prediction, wearable sensor data, smart healthcare, boosting algorithms, smart cities

1 INTRODUCTION

With the rapid development of technology, and the continuous growth of urban populations, the concept of smart city has emerged as a dynamic area to enhance the quality of life and optimize urban services [1]. Smart cities rely on the integration of technologies such as Internet of Things (IoT), machine learning (ML) and data analytics in order to enable decision-making and efficient resource management [2], [3]. Among the different dimensions of smart cities, smart healthcare plays an important role in enhancing public health outcomes while reducing operational costs [4].

Smart healthcare systems integrate IoT-enabled devices, particularly wearable sensors, to monitor individuals' physiological and behavioral signals [5], [6]. These sensors generate real-time data including heart rate, blood oxygen levels, sleep patterns

Jebbar, H., El Mohadab, M., Boutkhoum, O. (2026). A Data-Driven Framework for Stress Detection from Wearable Sensors in Smart Healthcare. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(7), pp. 169–184. https://doi.org/10.3991/ijoe.v22i07.61221

Article submitted 2026-02-22. Revision uploaded 2026-04-21. Final acceptance 2026-04-23.

© 2026 by the authors of this article. Published under CC-BY.

and physical activity, providing valuable information about an individual's health status [7]. In this regard, stress detection has gained significant attention due to its impact in both mental and physical health. Chronic stress is associated with critical health conditions such as cardiovascular diseases, hypertension and metabolic disorders [8].

Previous approaches to stress detection are based on self-reported questionnaires or clinical assessments which are not suitable for continuous monitoring. Nowadays, wearable sensors combined with machine learning techniques offer an efficient solution for stress detection [9], [10]. However, predicting stress from wearable devices is still a challenging task due to data heterogeneity and overlapping physiological patterns across stress levels [11].

Various researches have discussed the application of machine learning algorithms for stress prediction using wearable sensors [10], [12]. These approaches demonstrate results, however, most of them are based on binary classification without providing comprehensive analysis of multi-class stress prediction. Particularly, researches have been given limited attention to the difficulty of identifying intermediate stress levels, specifically the "moderate" class which overlaps with both low and high stress patterns.

To overcome these challenges, this paper proposes a data-driven framework based on multimodal wearable sensors data for multi-class stress level prediction. The proposed study integrates various machine leaning algorithm including Decision Tree (DT), random forest (RF), LightGBM and CatBoost.

The main contributions of this paper are summarized as follows:

- A data-driven framework for stress level prediction using multimodal wearable sensors data in smart healthcare.
- A comparative analysis of boosting-based and tree-based machine learning algorithms for multi-class stress classification.
- An in-depth analysis of multi-class stress prediction with a focus on the challenges associated with moderate stress level which remains underexplored in existing studies.

The remainder of the paper is organized as follows: Section two discusses related work and literature review of smart healthcare and the integration of machine learning. Section three is reserved to present wearable sensors as a part of IoT. The methodology and the approach followed in this work are presented in the fourth section. Section fifth discusses the results and sixth section concludes the paper.

2 RELATED WORK

Recently, healthcare prediction has become an important factor to save lives. There is a significant development in the healthcare domain in terms of intelligent systems to analyze complicated data and transform it for the prediction process [11]. The use of machine learning in healthcare data management has led to significant contributions made by many researchers and authors.

Researchers have highlighted a comprehensive review of both machine learning and deep learning techniques that are used in healthcare prediction [13]. They also give a background on machine learning as well as deep learning by explaining approaches employed in healthcare prediction [11]. Moreover, the integration of big data analysis and machine learning toward clinical assessment is discussed in [14]; it also highlights the drawbacks of data heterogeneity, privacy, and computational complexity.

The authors in [15] discuss how the COVID-19 pandemic exposed critical shortcomings of traditional medicine and highlighted the necessity for connected devices that generate health data by posing challenges in management and decision-making. In [11], the authors highlighted the role of ML and DL in healthcare prediction by discussing their utility in disease diagnosis from clinical data and images. In [4], the authors present the relation between IoT especially wireless sensors, and healthcare systems by explaining the role of machine learning in analyzing data generated by IoT devices. In paper [9], recent trends in wearable sensors are presented. The authors are focused on mental health by using machine learning techniques to detect and monitor stress. Also, in [10] authors study as well different machine learning classifiers to predict stress. Moreover, the authors in [12] explore the importance of wearable devices to collect data for various diseases. They also explain the importance of machine learning in healthcare in order to determine whether disease is present or not and this by using classification process. In [16], the authors aim to analyze the factors impacting healthcare demand by using machine learning algorithms based on mathematical models which establish the relationship between variables and clinical diagnosis. Anomaly detection has become an essential component in digital healthcare analytics because of difficulties associated with the integration of smart technologies, as discussed by authors in [17]. Also, the study introduced a methodology aimed at securing heterogeneous healthcare systems. Table 1 presents literature work in recent years.

Table 1. Literature review

Study	Year	Description
[18]	2018	• Authors discussed the need of IoT in healthcare applications in order to ensure an end-to-end communication process. • Different diseases, such as cardio-vascular and infectious diseases, are used to present many applications scenarios.
[19]	2019	• Data comprehensive overview of IoT applications within smart healthcare. • Various services and applications are described based on their main ideas in smart healthcare systems.
[20]	2020	• Authors proposed a comprehensive architecture of smart healthcare and highlighted the key requirements. • Different advanced technologies used to improve the performance of smart healthcare systems including machine learning, block chain, and big data.
[21]	2021	• The concept of smart healthcare infrastructure and the advanced technologies in smart healthcare systems are discussed. • The challenges of smart healthcare are introduced in order to give an understanding of the requirements of smart healthcare systems. • Some technologies are reviewed such as ML and block chain that researchers discussed for enhancing smart healthcare systems.
[22]	2022	• The authors presented an overall overview of the main functions of wearable technology for health monitoring, disease detection and prediction. They also discussed the quality and the efficiency of wearable devices.
[23]	2022	• This study provides an overall understanding to healthcare professionals, government, researchers and students of the different digital technologies and their applications during COVID-19 pandemic. • The use of different industry tool and digital technologies such as cloud computing, telemedicine and Internet of Things provide opportunities for effective healthcare services, work from home and online education.
[24]	2024	• The paper presented a comprehensive analysis of ML technologies for digital health industry and wearable devices. • Authors proposed a taxonomy for digital healthcare, transforming it into a developed model in order to highlight key research challenges such as data collection, health monitoring and data privacy issues.

3 SMART DEVICES AND INTERNET OF THINGS IN SMART HEALTHCARE

Internet of Things is defined as the network that includes different physical objects equipped with embedded technologies for sensing and communication. IoT consist of providing improvements to the population in order to make their lives easier by IoT devices including sensors, smartphones, etc. With the development in communication technologies, the growth of mobile users, and the increasing capacities of cloud technologies, smart city, smart healthcare, and smart home concepts are increasingly integrated into people's daily routines. Moreover, IoT enables an ecosystem where devices can transmit their collected data with cloud servers in order to achieve their goals with no human intervention [21]. A variety of protocols and communication types can be utilized in IoT network such as Bluetooth, ZigBee, Wi-Fi, 5G in order to address user satisfaction and wireless communication ranges.

The main role of IoT is to enable objects that generate data and transfer it, in order to be analyzed and used for classification decision-making [25]. IoT plays a significant role in healthcare by improving access to care, increasing quality, and reducing the cost of care. Based on an individual's unique biological, behavioral, social, and cultural characteristics, IoT has the potential to enable a wide range of medical applications such as remote health monitoring, chronic disease, elderly care, stress detection, and well-being [26]. As an example of patient's heart rate can be collected by sensors from time to time and then transmitted to the doctor's office. However, different medical sensors, diagnostics, and imaging devices can be seen as smart devices constituting a core part of Internet of Things.

3.1 Wearable devices

Wearable health devices and sensors constitute a revolutionary technology that enhances remote monitoring of patients health, data collection and transmission to medical providers [27]. Smart sensors become important tools because of their ability to give higher quality, reliable, and real-time data according to an individual's medical condition and surrounding conditions [28]. Nowadays, sensor technology has made remarkable progress, reflected in downsizing, sensitivity, precision and affordability [29]. These sensors give a deeper understanding of an individual's well-being and lifestyle by monitoring various physiological, environmental and psychological indicators. Figure 1 depicted the different connected wearable devices.

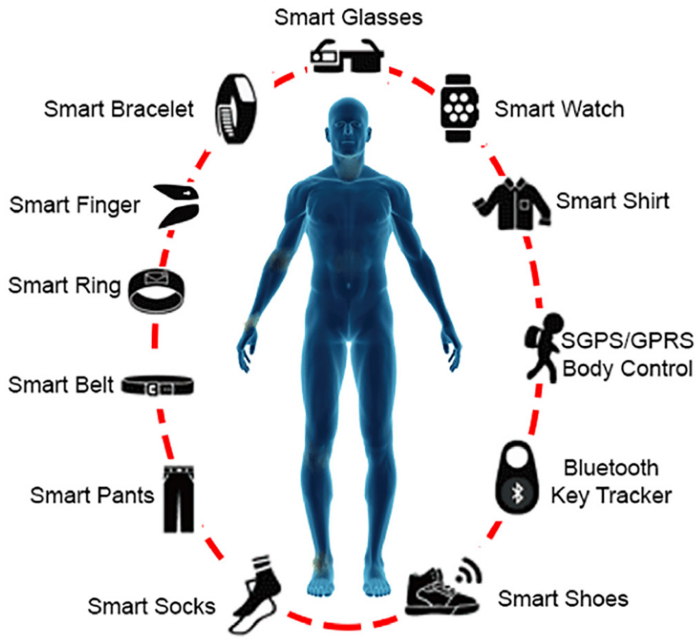

Fig. 1. Different types of wearable sensors in smart healthcare [30]

In smart healthcare, wearable devices can be categorized into three functions:

- Physiological monitoring
- Activity and sleep tracking
- Specialized medical sensing

In fact, physiological devices are designed to measure signal and provide a real-time status for a patient's health. Smart watched and connected bracelets equipped with photo plethysmography (PPG) and electrocardiography (ECG) were both used for heart rate monitoring, while integrated optical sensors are used for blood oxygen saturation (SpO2) measurement [31]. Moreover, blood pressure monitors and skin-contact are also considered as a wearable platform in order to detect body temperature and blood pressure.

Another function of wearable devices is to measure physical activity and rest patterns. Accelerometers and gyroscopes are embedded in wristbands and smart rings in order to detect daily step count, movement intensity and posture [32]. These devices also play an important role for sleep monitoring, where heart rate and motion variability data are combined in order to predict sleep duration, quality and stages. Therefore, monitoring sleep patterns and activity not only support wellness and fitness applications but also generates data for detecting anomalies and issues such as insomnia, stress-related disorders, anxiety, and obstructive sleep.

The third category of wearable sensors are adopted for clinical applications and chronic disease managements and it defined as specialized wearables. Actually, continuous glucose monitoring sensors (CGM) are used to detect real-time measurements of blood glucose levels in diabetic patients as show in Figure 3, significantly reducing the need of invasive finger-prick [33]. In the same way, portable ECG and EEG devices offer long-term cardiac and neurological monitoring outside clinical settings and this is highly beneficial for detecting epilepsy, seizure disorders or arrhythmias. The integration of wearable sensors into IoT infrastructure guarantee that critical health issue can be detected by a data transmitted securely to healthcare providers for timely interventions.

One of the most benefits of wearable devices is the real-time monitoring, which continuous monitoring of vital parameters such as blood oxygen saturation, heart rate and glucose levels. This capability prevents medical crises by transmitting early alerts for abnormal behavior and thus to facilitate timely interventions. Another advantages is the early detection of chronic diseases. Valuable data are collected by wearables which can reveal various changes in physiological or behavioral patterns which can serve as early indicators to detect cardiovascular disease, diabetes or sleep issues. Also, wearable devices are made to be worn comfortably which make it easy of use for individuals. Furthermore, by using wearable devices the cost in healthcare systems is reduced. Taking into consideration that remote patient monitoring reduces the frequency of hospital visits and then lowering healthcare expenditures for both providers and patients. Different studies have shown that telemedicine supported by wearable devices lead to enhance patient's outcomes and decreased hospitalization especially in chronic diseases.

Despite all the benefits of wearable devices in smart healthcare, they faced challenges that limited their adoption. A primary drawback is energy consumption, most of wearable rely on small batteries with limited autonomy which make their life period short and this disrupts continuous monitoring. Another critical issue is interoperability. Nowadays, wearable devices are produced by different manufacturers using heterogeneous standards, data format and communication protocols. This lack of standardization complicated seamless integration into healthcare information systems. And finally, the major critical challenge is data privacy and security. As discussed in this section, wearable devices generate sensitive data which contain medical information, making them attractive targets for cyberattacks. To ensure confidentiality, integrity and availability of health data requires robust encryption and secure authentication. Addressing these challenges is important in order to build trust between patients and clinicians and that for realizing the full potential of wearable healthcare technologies.

While wearable sensors monitor individual's health by generating large volume of data streams, which are often heterogeneous, high-dimensional and complex. To transform raw sensor data into clinically meaningful knowledge, advanced techniques are required where machine learning (ML) is introduced. ML plays a significant role by offering predictive and diagnostics capabilities that enhance the efficiency of healthcare systems. The following section highlight the integration of ML in smart healthcare and presented the most relevant algorithms applications.

4 METHODOLOGY

This section is reserved to present the proposed methodology for stress prediction using multimodal data collected from wearable sensors. The objective of this work is to develop a framework based on machine learning capable of analyzing physiological, behavioral and lifestyle signal to predict stress levels (low, moderate and high). As depicted in Figure 2, the methodology is divided into four main stages: data collection, data pre-processing, and feature engineering and performance evaluation.

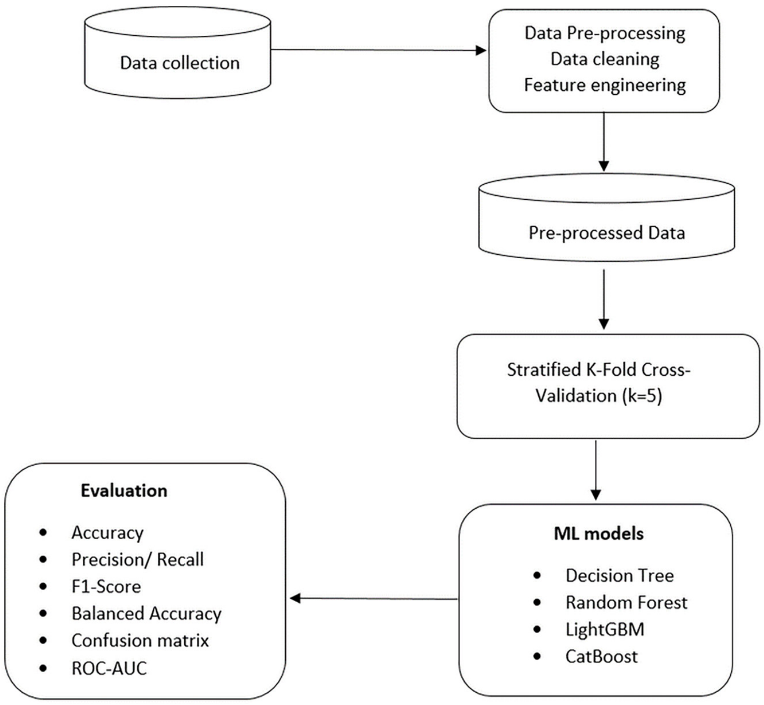

Fig. 2. Proposed approach for stress level prediction

4.1 Data collection

The dataset used in this study is collected through wearable sensors and user self-reports, publicly available on Kaggle, and widely used for research purposes in stress detection and health monitoring.

The dataset consists of approximately 10000 instances with 28 attributes. It includes multimodal health and lifestyle data from multiple users across many times of day.

As shown in Table 2, wearable sensors used on this study capture physiological signals (heart rate, blood oxygen saturation, and skin temperature), sleep-related metrics, lifestyle behaviors and demographic information. In this study, our variable target is “stress_level” categorized into three discrete levels: low, moderate, and high.

Table 2. Dataset attributes

Demographic and health background	User_ID, Age, Gender, Medical_Conditions, Medication, Smoker, Alcohol_Consumption.
Temporal Information	Timestamp, Day_of_Week.
Sleep Metrics	Sleep_Duration (hours), Deep_Sleep_Duration (hours), REM_Sleep_Duration (hours), Wakeups (count).
Physiological Measurements	Heart_Rate (bpm), Blood_Oxygen_Level (%), Skin_Temperature (°C), ECG (Normal/Abnormal).
Lifestyle & Behavior	Calories_Intake (kcal), Water_Intake (liters), Snoring (Yes/No), Mood (Happy, Sad, Anxious, and Neutral).
Body Composition	Weight (kg), Height (cm), Body_Fat_Percentage (%), Muscle_Mass (%).
Target Variable	Stress_Level (categorical: Low, Moderate, and High).

4.2 Data pre-processing and feature engineering

In order to enhance robustness of the predictive models, various preprocessing steps are applied. Categorical variables are encoded using the appropriate encoding in order to ensure compatibility with machine learning models. Missing values are handled using suitable techniques and numerical variables are normalized.

In addition, correlation analysis is performed to detect redundant variables and reduce the risk of multicollinearity, which can negatively influence the model. To improve model generalization and reduce overfitting high correlation features should be removed.

Moreover, multi-class stress data suffer sometimes from class imbalance, suitable strategies are applied to overcome this issue. In our study, class weighting techniques are included during model training in order to ensure balanced learning across all stress levels. This approach ensures model’s robustness to correctly classify moderate stress category.

4.3 Modeling and analytical approach

In order to improve the robustness of the experimental results, hyperparameter tuning was conducted using GridSearchCV in combination with stratified 5-fold cross-validation. A set of key hyperparameters was explored for each model, including the number of estimators and tree depth for tree-based methods, as well as learning rate and leaf parameters for boosting models. The optimal configuration was selected based on the highest F1-score. This technique ensures reliable estimation of model performance compared to a single train-test split. The four supervised classification algorithms were employed: RF, DT, Light Gradient Boosting Machine (LightGBM) and Categorical Boosting (CatBoost). Each model is designed for its advantages in terms of robustness against noisy data, efficiency, and interpretability.

- Decision tree: DT is a supervised learning algorithm which integrate the values of attributed based on their order (ascending or descending) [34]. DT is worked as flowchart in order to make prediction, where internal nodes represent attribute test, branch represent attribute value and the leaf node represent the final prediction. The impurity of a node is measured using the Gini index as defined in the equation:

$$Gini(D) = 1 - \sum_{k=1}^{K} p_k^2$$

Where:
D represents the set of samples at a given node.
K is the number of classes.
p_k denotes the probability (or proportion) of samples in D belonging to class k.

The Gini index measures the impurity of the node, where lower values indicate better class separation.

However, a lower Gini index shows a pure distribution while a higher value indicates an impure distribution. In DT, the Gini index evaluates the quality of a split by calculating the difference between the impurity of the parent node and the impurity of the child ones.

- Random forest: Random forest is a supervised machine learning algorithm. It consists to combine the predictions of multiple decision trees and blending them. The main idea of RF is that each DT is trained separately using random selected data subsets, and the general performance is determined by the average of trees predictions [35] in order to achieve high accuracy and robustness. RF improves predictive power by developing tree which considered the best feature from a randomly selected subset of feature that than evaluating all possible feature at each split. This method produces models that attaint high accuracy. Therefore, the risk of overfitting is decreased and the overall generalization is improved. Due to its scalability and working well with high dimensional data, RF is used in different healthcare studies. The output of RF is defined as:

$$\hat{y} = mode\left\{h_1(x), h_2(x), \ldots, h_T(x)\right\}$$

Where h_t represent the prediction the t-th tree and T is total number of tree
x is the input sample
T denotes the number of trees in the ensemble.
$h_T(x)$ represents the prediction of the t-th decision tree.

- Light gradient boosting machine: Light gradient boosting machine is a variant of gradient-boosting algorithm widely used for supervised tasks such as classification and ranking [36]. LightGBM is a framework that developed by Microsoft and Peking university as a fast and efficient algorithm for ML tasks. The lightGBM technique is based on gradient boosting framework which consists of training multiple weak learners sequentially, with each learner trying to correct the mistake of their predecessors [37]. Then the final model is the sum of these weak learners that combine on order of form a strong learner as described in the equation below:

$$F_{m+1}(x) = F_m(x) + \alpha \sum_{t=1}^{T} \gamma_t h_t(x)$$

Where:
$F_m(x)$ is the current model at iteration m.
$F_{m+1}(x)$ is the updated model, and x represents the input feature vector.

h_t is the weak learner.
γ_t denotes the weight associated with the t-th tree.
α represents the learning rate.

- Categorical Boosting algorithm (CatBoost): CatBoost is a field of gradient boosting DT algorithm [38]. The main idea behind this algorithm is to manage categorical feature more efficiently. CatBoost used improved and revised target which call it Greedy TBS. in this process, categorical feature is replaced by their corresponding average label values which are served as the criteria for node splitting during the application of decision trees [37]. It is defined by equation:

$$L = \sum_{i=1}^{N} l(y_i, F(x_i))$$

Where:
N is the number of training instances.
y_i is the ground truth label of the i-th sample.
$F(x_i)$ represents the model prediction.
$l(.)$ denotes the loss function.

4.4 Evaluation metrics

The predictive performance of ML is evaluated using multiple metrics to identify factors affecting healthcare services [17]. Accuracy measures the proportion of correctly classified instances over the total number of instances, while precision evaluates the ability of the model to correctly identify positive instances among all predicted positives, Recall (Sensitivity) which measures the proportion of actual positive instances that are correctly identified and F1-score which represents the harmonic mean of precision and recall, providing a balanced metric when there is a trade-off between them. The Area Under the ROC Curve (AUC) aims to distinguish between classes, the higher score is the superior performance [39]. The Table 3 summarizes the mathematical expression of the metrics used by ML healthcare predicting service, where:

TP (True Positive): denotes correctly predicted positive instances.
TN (True Negative): denotes correctly predicted negative instances.
FP (False Positive): represents incorrectly predicted positive instances.
FN (False Negative): represents incorrectly predicted negative instances.

Table 3. Mathematics metrics formula

Metrics	Formula
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Recall	$Recall = \frac{TP}{TP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
F1-Score	$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}$

5 RESULTS AND DISCUSSION

In order to analyze the potential of stress prediction, three classification models were trained using physiological and lifestyle feature as input, the target variable is the stress level (Low, Moderate, and high). The dataset was evaluated using stratified 5-fold cross-validation to ensure robust and reliable performance estimation. The ML algorithms were evaluated by using metrics: Accuracy, Recall, precision, F1-scroe and ROC-AUC.

5.1 Algorithms performances

The experimental results show that stress prediction from wearable devices becomes a challenging task particularly in a real-world multiclass setting. Moreover, all ML models demonstrate comparable outcomes.

- Decision tree performance: Decision tree model achieved an accuracy of 38.65%, with F1-score of 0.3851. The limitations of single-tree models in handling complex and non-linear relationships in wearable sensors are the first parameter of low performance. Despite their interpretability, decision trees tend to overfit the training dataset, resulting in reduced generalization performance, especially when dealing with multi-class classification problems.
- Random forest performance: Random forest model achieved an accuracy of 36.8% with F1-score of 0.3648. The results show that RF does not significantly outperform the DT model. This can be explained by the overlap between stress classes. Although, RF enhances generalization by aggregation multiple trees, it is still limited in capturing complex non-linear relationships in multi-class stress prediction.
- LightGBM performance: LightGBM achieved an accuracy of 44.87%, with F1-score around 0.4473. This improvement over traditional models shows the effectiveness of gradient boosting models in which that it provided the best overall performance among the previous models. The high performance of LightGBM can be attributed to its leaf-wise tree growth strategy, which allows efficient learning of non-linear patterns in high-dimensional data. Additionally, its ability to handle heterogeneous features makes it well-suited for wearable sensor data.
- CatBoost performance: CatBoost model achieved and accuracy of 45.20% with F1-score of 0.4467. The model shows robustness and strong generalization capability across all stress classes.

  The effectiveness of CatBoost is primarily due to its ordered boosting mechanism and its ability to handle categorical features efficiently, reducing overfitting and improving prediction stability.

  A global comparison of model performance is summarized in Table 4:

Table 4. Metrics result for each model

Model	Accuracy	Precision	Recall	F1-Score	Balanced Accuracy
Decision Tree	0.3865 ± 0.0098	0.3861 ± 0.0111	0.3865 ± 0.0098	0.3851 ± 0.0101	0.3860 ± 0.0098
Random Forest	0.3686 ± 0.0113	0.3655 ± 0.0123	0.3686 ± 0.0113	0.3684 ± 0.0124	0.3679 ± 0.0114
LightGBM	0.4487 ± 0.0088	0.4463 ± 0.0094	0.4487 ± 0.0088	0.4473 ± 0.0091	0.4478 ± 0.0089
CatBoost	0.4520 ± 0.0092	0.4445 ± 0.0101	0.4520 ± 0.0092	0.4467 ± 0.0096	0.4509 ± 0.0092

5.2 Comparative discussion

Based on the results shown in Table 4, boosting-based models such as light GBM and CatBoost, outperform traditional tree-based models. This confirms that gradient boosting algorithms show effectiveness in modeling complex and non-linear relationships in wearable sensors. Additionally, the low performance of RF compared to DT can be explained by the overlap particularly in multi-class stress classification.

The overall performance remains moderate across all models, demonstrating the inherent difficulty of stress prediction in a multi-class setting. The relatively low performance is due to various factors including the overlapping physiological patterns across stress levels, the subjective nature of stress labeling, and the limitations of the dataset, which may not reflect real clinical conditions.

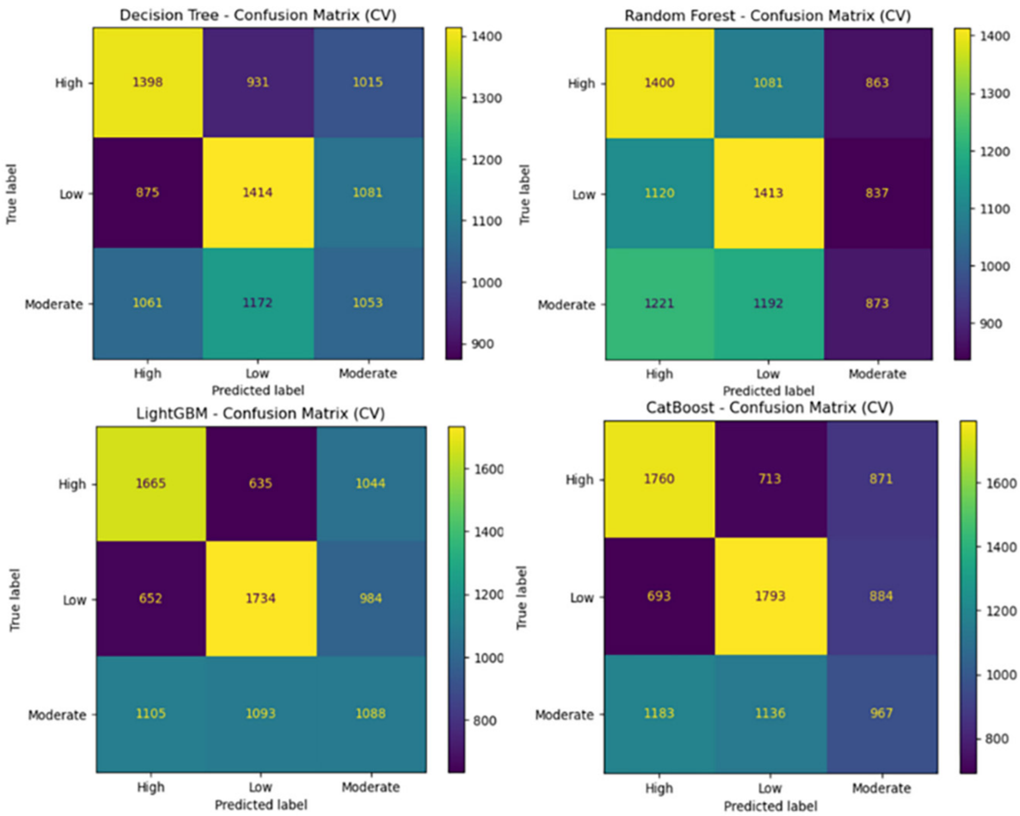

Fig. 3. Confusion Matrix for stress classification models

Figure 3 describes the confusions matrices of each classification models applied for a three class stress. The confusion matrices provide deeper insights into the classification behavior of the evaluated models. Across all models, a consistent pattern of misclassification is observed, especially involving the "moderate" stress level.

For both DT and RF models, significant confusion exists between all classes, demonstrating their limited ability to separate stress levels. In particular, the "moderate" class is widely distributed across predictions, which confirm difficulty of modeling intermediate stress states. However, LightGBM and CatBoost shows better classification performance, with higher correct predictions along the diagonal of the confusion matrix. These models demonstrate better discrimination of "high" and "low" stress levels. Also, even for these boosting models, the "moderate" class remains the most challenging to classify accurately due to the overlapping of physiological signals such as heart rate, sleep patterns, and temperature across stress levels. The "moderate" class represents a transitional state between low and high stress, making it inherently ambiguous and difficult for machine learning models to distinguish.

Overall, the confusion matrix analysis confirms that while boosting-based models improve performance, stress prediction in a multi-class setting remains a complex problem due to class overlap and data variability.

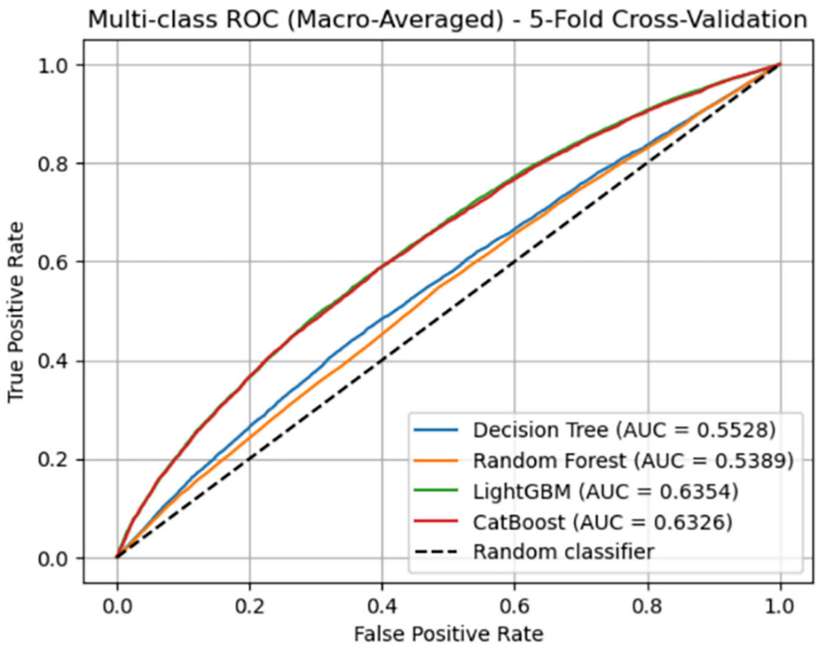

Fig. 4. ROC-AUC for stress classification models

Figure 4 illustrates ROC-AUC curves. The ROC curves demonstrates a comprehensive evaluation of the discriminative performance of the models in a multi-class setting using a one-vs-rest strategy. As shown in Figure 4, the boosting-based models: LightGBM and CatBoost, show the highest ROC-AUC values of 0.6354 and 0.6326, respectively. This indicates their ability to distinguish between stress classes compared to traditional tree-based methods.

On the other hand, the DT and RF models achieve lower ROC-AUC values of 0.5528 and 0.5389, respectively, reflecting limited discriminative capability. Their performance remains only slightly above the random classifier baseline (AUC = 0.5), suggesting difficulty in capturing complex patterns within the data.

Overall, all models perform above random classification, the ROC curves confirm that stress prediction in a multi-class context is still challenging. The moderate AUC values highlight the presence of overlapping physiological features across stress levels, which reduces class separability. Nevertheless, the results demonstrate that boosting techniques significantly improve the model's ability to learn discriminative patterns in wearable sensor data.

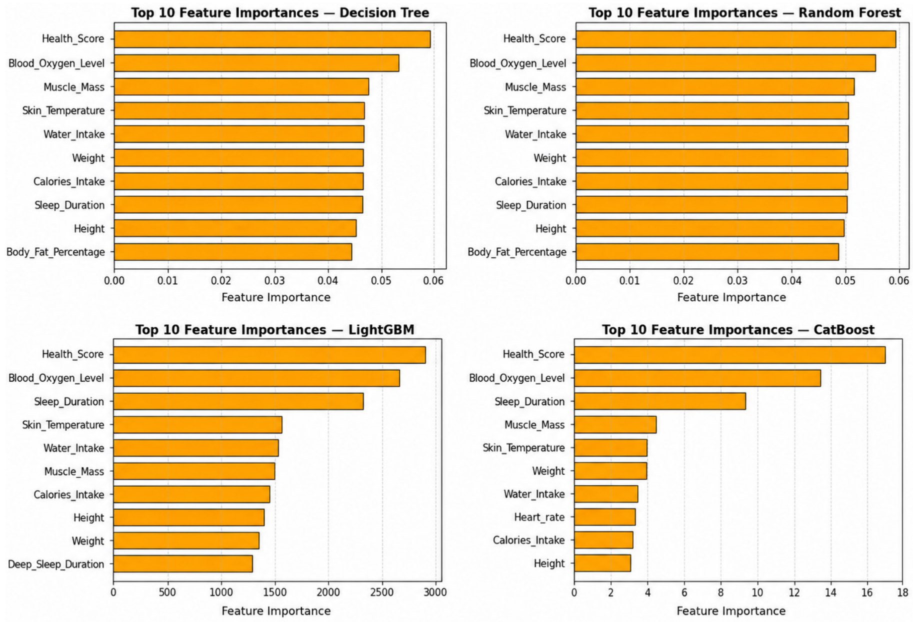

Fig. 5. Features engineering for each models

As depicted in Figure 5, the consistency of the top-ranked features across all algorithms indicates that variables such as Health Score, Blood Oxygen Level, Sleep Duration, Skin Temperature, and Muscle Mass play a significant role in stress classification. These results are consistent with previous studies demonstrating that sleep patterns and thermoregulation are closely associated with psychological stress responses.

6 CONCLUSION

This work presented a data-driven framework for multi-class stress level prediction using multimodal wearable sensor data in smart healthcare systems. By using physiological, behavioral and lifestyle features, the study discussed the effectiveness of several machine learning models such as: DT, RF, LightGBM and CatBoost.

The experimental results showed that boosting based models, especially LightGBM and CatBoost, perform better than traditional tree-based approaches in capturing complex and non-linear relationships within heterogeneous data. The overall result remains moderate, highlighting the inherent complexity of stress prediction in multi-class scenarios.

A key contribution of this study lies in the detailed analysis of multi-class stress prediction with a particular focus on the moderate stress level which remains underexplored in existing studies.

Nevertheless, the results highlight the potential of integrating wearable device data with machine learning techniques for continuous and non-invasive stress monitoring. This work further provides an initial validation of the proposed framework rather than a full-scale clinical evaluation. The use of public dataset may limit the clinical validity of the result as a real-world medical data often exhibit higher variability and noise.

Future work will focus on enhancing prediction performance by integrating advanced approaches including deep learning models, multimodal data fusion and explainable artificial intelligence (XAI). Also, the integration of clinical datasets will be important to improve the reliability of stress detection systems in smart healthcare. In addition, feature importance analysis was conducted, future work will further explore advanced interpretability techniques such as SHAP to provide deeper insights into model predictions.

7 REFERENCES

[1] Z. Mohammadzadeh, H. R. Saeidnia, A. Lotfata, M. Hassanzadeh, and N. Ghiasi, "Smart city healthcare delivery innovations: A systematic review of essential technologies and indicators for developing nations," *BMC Health Serv. Res.*, vol. 23, no. 1, p. 1180, 2023. https://doi.org/10.1186/s12913-023-10200-8

[2] Y. Kaluarachchi, "Implementing data-driven smart city applications for future cities," *Smart Cities*, vol. 5, no. 2, pp. 455–474, 2022. https://doi.org/10.3390/smartcities5020025

[3] M. K. Al-Azzam and M. Bader, "Smart city and smart-health framework, challenges and opportunities," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 2, 2019. https://doi.org/10.14569/IJACSA.2019.0100223

[4] T. M. Ghazal *et al.*, "IoT for smart cities: Machine learning approaches in smart healthcare—a review," *Future Internet*, vol. 13, no. 8, p. 218, 2021. https://doi.org/10.3390/fi13080218

[5] H. Ziwei *et al.*, "The applications of internet of things in smart healthcare sectors: A bibliometric and deep study," *Heliyon*, vol. 10, no. 3, p. e25392, 2024. https://doi.org/10.1016/j.heliyon.2024.e25392

[6] M. Hamzei, S. Khandagh, and N. Jafari Navimipour, "A quality-of-service-aware service composition method in the internet of things using a multi-objective fuzzy-based hybrid algorithm," *Sensors*, vol. 23, no. 16, p. 7233, 2023. https://doi.org/10.3390/s23167233

[7] X. Liu, J. Zhao, J. Li, B. Cao, and Z. Lv, "Federated neural architecture search for medical data security," *IEEE Trans. Ind. Inform.*, vol. 18, no. 8, pp. 5628–5636, 2022. https://doi.org/10.1109/TII.2022.3144016

[8] Health and Safety Executive (HSE), "Work-related ill health and occupational disease in Great Britain," 2021. https://www.hse.gov.uk/statistics/causdis/

[9] A. Pinge, V. Gad, D. Jaisighani, S. Ghosh, and S. Sen, "Detection and monitoring of stress using wearables: A systematic review," *Front. Comput. Sci.*, vol. 6, p. 1478851, 2024. https://doi.org/10.3389/fcomp.2024.1478851

[10] P. Garg, J. Santhosh, A. Dengel, and S. Ishimaru, "Stress detection by machine learning and wearable sensors," in *26th International Conference on Intelligent User Interfaces*, College Station TX USA: ACM, 2021, pp. 43–45. https://doi.org/10.1145/3397482.3450732

[11] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: A survey," *J. Electr. Syst. Inf. Technol.*, vol. 10, no. 1, p. 40, 2023. https://doi.org/10.1186/s43067-023-00108-y

[12] H. S. Saad, J. F. W. Zaki, and M. M. Abdelsalam, "Employing of machine learning and wearable devices in healthcare system: Tasks and challenges," *Neural Comput. Appl.*, vol. 36, no. 29, pp. 17829–17849, 2024. https://doi.org/10.1007/s00521-024-10197-z

[13] N. F. A. Ahmad Sukri, W. M. A. F. Wan Hamzah, M. K. Yusof, I. Ismail, H. M. Yusoff, A. Yacob, "A systematic literature review on machine learning in healthcare prediction," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 21, no. 6, pp. 155–177, 2025. https://doi.org/10.3991/ijoe.v21i06.54211

[14] S. Rani *et al.*, "Machine learning-powered smart healthcare systems in the era of big data: Applications, diagnostic insights, challenges, and ethical implications," *Diagnostics*, vol. 15, no. 15, p. 1914, 2025. https://doi.org/10.3390/diagnostics15151914

[15] W. Li *et al.*, "A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system," *Mob. Netw. Appl.*, vol. 26, no. 1, pp. 234–252, 2021. https://doi.org/10.1007/s11036-020-01700-6

[16] F. Orhan and M. N. Kurutkan, "Predicting total healthcare demand using machine learning: Separate and combined analysis of predisposing, enabling, and need factors," *BMC Health Serv. Res.*, vol. 25, no. 1, p. 366, 2025. https://doi.org/10.1186/s12913-025-12502-5

[17] M. Tabassum, S. Mahmood, A. Bukhari, B. Alshemaimri, A. Daud, and F. Khalique, "Anomaly-based threat detection in smart health using machine learning," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, p. 347, 2024. https://doi.org/10.1186/s12911-024-02760-4

[18] M. M. Alam, H. Malik, M. I. Khan, T. Pardy, A. Kuusik, and Y. Le Moullec, "A survey on the roles of communication technologies in IoT-based personalized healthcare applications," *IEEE Access*, vol. 6, pp. 36611–36631, 2018. https://doi.org/10.1109/ACCESS.2018.2853148

[19] N. Hema Rajini, "A comprehensive survey on internet of things based healthcare services and its applications," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, 2019, pp. 483–488. https://doi.org/10.1109/ICCMC.2019.8819642

[20] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare internet of things: A survey of emerging technologies," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 2, pp. 112–1167, 2020. https://doi.org/10.1109/COMST.2020.2973314

[21] M. A. Tunc, E. Gures, and I. Shayea, "A survey on IoT smart healthcare: Emerging technologies, applications, challenges, and future trends," *arXiv preprint arXiv:2109.02042*, 2021. https://doi.org/10.48550/ARXIV.2109.02042

[22] S. Canali, V. Schiaffonati, and A. Aliverti, "Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness," *PLOS Digit. Health*, vol. 1, no. 10, p. e0000104, 2022. https://doi.org/10.1371/journal.pdig.0000104

[23] M. Chandra, K. Kumar, P. Thakur, S. Chattopadhyaya, F. Alam, and S. Kumar, "Digital technologies, healthcare and Covid-19: Insights from developing and emerging nations," *Health Technol.*, vol. 12, no. 2, pp. 547–568, 2022. https://doi.org/10.1007/s12553-022-00650-1

[24] V. Tandel, A. Kumari, S. Tanwar, A. Singh, R. Sharma, and N. Yamsani, "Intelligent wearable-assisted digital healthcare industry 5.0," *Artif. Intell. Med.*, vol. 157, p. 103000, 2024. https://doi.org/10.1016/j.artmed.2024.103000

[25] A. Rejeb *et al.*, "The Internet of Things (IoT) in healthcare: Taking stock and moving forward," *Internet Things*, vol. 22, p. 100721, 2023. https://doi.org/10.1016/j.iot.2023.100721

[26] B. G. Mohamme and D. S. Hasan, "Smart healthcare monitoring system using IoT," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 17, no. 1, pp. 141–152, 2023. https://doi.org/10.3991/ijim.v17i01.34675

[27] W. Shafik, "Smart biomedical devices for smart healthcare," in *Machine Learning Models and Architectures for Biomedical Signal Processing*, Elsevier, 2025, pp. 421–448. https://doi.org/10.1016/B978-0-443-22158-3.00017-X

[28] J. Ali, A. Jusoh, N. Idris, A. G. Airij, and R. Chandio, "Wearable devices in healthcare services. Bibliometrix analysis by using R package," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 18, no. 8, pp. 61–86, 2022. https://doi.org/10.3991/ijoe.v18i08.31785

[29] D. Syed, N. Islam, M. H. Shabbir, and S. B. Manzar, "Applications of big data in smart health systems," in *Advances in Healthcare Information Systems and Administration*, D. Samanta and D. Singh, Eds., IGI Global, 2022, pp. 52–85. https://doi.org/10.4018/978-1-6684-4580-8.ch004

[30] A. Shaikh, D. Syed, M. S. Al Reshan, and S. M. Alqhtani, "Intelligent sensors for smart healthcare," in *Advanced Sensors for Smart Healthcare*, 2025, pp. 347–361. https://doi.org/10.1016/B978-0-443-24790-3.00022-3

[31] Z. Lv and Y. Li, "Wearable sensors for vital signs measurement: A survey," *J. Sens. Actuator Netw.*, vol. 11, no. 1, p. 19, 2022. https://doi.org/10.3390/jsan11010019

[32] M. De Zambotti, N. Cellini, A. Goldstone, I. M. Colrain, and F. C. Baker, "Wearable sleep technology in clinical and research settings," *Med. Sci. Sports Exerc.*, vol. 51, no. 7, pp. 1538–1557, 2019. https://doi.org/10.1249/MSS.0000000000001947

[33] J. Y. Stone and T. S. Bailey, "Benefits and limitations of continuous glucose monitoring in type 1 diabetes," *Expert Rev. Endocrinol. Metab.*, vol. 15, no. 1, pp. 41–49, 2020. https://doi.org/10.1080/17446651.2020.1706482

[34] D. Dhall, R. Kaur, and M. Juneja, "Machine learning: A review of the algorithms and its applications," in *Proceedings of ICRIC 2019*, in Lecture Notes in Electrical Engineering, vol. 597, P. K. Singh, A. K. Kar, Y. Singh, M. H. Kolekar, and S. Tanwar, Eds., Springer, Charm, 2020, pp. 47–63. https://doi.org/10.1007/978-3-030-29407-6_5

[35] M.-F. Avram *et al.*, "Random forests algorithm using basic medical data for predicting the presence of colonic polyps," *Front. Surg.*, vol. 12, p. 1523684, 2025. https://doi.org/10.3389/fsurg.2025.1523684

[36] T. O. Omotehinwa, D. O. Oyewola, and E. G. Moung, "Optimizing the light gradient-boosting machine algorithm for an efficient early detection of coronary heart disease," *Inform. Health*, vol. 1, no. 2, pp. 70–81, 2024. https://doi.org/10.1016/j.infoh.2024.06.001

[37] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," *arXiv preprint arXiv:1706.09516*, 2017. https://doi.org/10.48550/ARXIV.1706.09516

[38] K. Ileri, "Comparative analysis of CatBoost, LightGBM, XGBoost, RF, and DT methods optimised with PSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks," *Int. J. Mach. Learn. Cybern.*, vol. 16, no. 9, pp. 6937–6956, 2025. https://doi.org/10.1007/s13042-025-02654-5

[39] A. J. Bowers and X. Zhou, "Receiver operating characteristic (ROC) area under the curve (AUC): A diagnostic measure for evaluating the accuracy of predictors of education outcomes," *J. Educ. Stud. Placed Risk JESPAR*, vol. 24, no. 1, pp. 20–46, 2019. https://doi.org/10.1080/10824669.2018.1523734

8 AUTHORS

Hayat Jebbar is currently pursuing a Ph.D. degree in Computer Science at the Department of Computer Science, Chouaib Doukkali University, El Jadida, Morocco. Her research focuses on smart cities and smart healthcare, with particular emphasis on the application of machine learning to wearable sensor data. Her interests also include semantic interoperability and intelligent data-driven solutions for healthcare systems (E-mail: jebbar.hayat@ucd.ac.ma).

Mohamed El Mohadab is an Associate Professor at the Computer Science department in the High School of technologies of Sultan Moulay Slimane University, Beni Mellal, Morocco. He received his Ph.D. degree in Computer Science from the Faculty of Sciences and Techniques of Sultan Moulay Slimane University, Beni Mellal. His research interests are Data Mining, Machine Learning and Information System (E-mail: m.elmohadab2@usms.ma).

Omar Boutkhoum is an Associate Professor at the Computer Science department in the Faculty of Sciences of Chouaib Doukkali University, EL Jadida, Morocco. He received his Ph.D. degree in Computer Science from the Faculty of Sciences and Techniques of Caddi Ayyad University, Marrakesh. His research interests are in the application of decision support systems, knowledge graphs, and artificial intelligence (E-mail: boutkhoum.o@ucd.ac.ma).