Journal of Ubiquitous Systems & Pervasive Networks
Volume 18, No. 2 (2023) pp. 49-59

Predicting Cardiovascular Events with Machine Learning Models and Heart Rate Variability

Mohammad Moshawrab [a]*, Mehdi Adda [a], Abdenour Bouzouane [b], Hussein Ibrahim [c], Ali Raad [d]

[a] *Département de Mathématiques, Informatique et Génie, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, G5L 3A1, Québec, Canada*
[b] *Département D'informatique et de Mathématique, Université du Québec à Chicoutimi, Chicoutimi, 555 boulevard de l'Université, Chicoutimi, QC G7H 2B1, Canada*
[c] *Institut Technologique de Maintenance Industrielle, 175 Rue de la Vérendrye, Sept-Îles, G4R 5B7, Québec, Canada*
[d] *Dean of the Faculty of Science and Arts, Islamic University of Lebanon, Wardaniyeh, Lebanon*

Abstract

Artificial Intelligence (AI) is increasingly becoming a potential answer to many of science's most challenging problems. In this context, healthcare is using this technology and its advancement to improve the quality of services provided, including cardiac healthcare services. According to studies, Cardiovascular Diseases (CVDs) are among the most common and deadly diseases in the world. However, Artificial Intelligence and its branches such as Machine Learning (ML) and Deep Learning (DL) offer tremendous potential to improve disease diagnosis and even predict its occurrence. In this study, eight Machine Learning and Deep Learning models are created and trained with "PhsyioNet Smart Health for Assessing the Risk of Events via ECG Database" to analyze the characteristics of Heart Rate Variability and predict the occurrence of heart disease and cerebrovascular events. The results support the use of Artificial Intelligence in cardiology, with five of the proposed models outperforming previous implementations. Specifically, Support Vector Machines, TabTransformers, Deep Neural Networks, AdaBoost, and XGBoost achieved accuracy rates of 91.80%, 90.38%, 90.19%, 89.50%, and 89.10%, respectively. Further performance metrics are presented throught the article such as precision, recall and others.

Keywords: *Artificial Intelligence, Machine Learning, Deep Learning, Cardiovascular Diseases, Heart Rate Variability*

1. Introduction

Cardiovascular Disease causes the most deaths and is therefore considered the most dangerous disease in the world. According to the latest data from the World Health Organization (WHO) in the field of heart disease, the number of deaths caused by these diseases has increased from 12.1 million in 1990 to 18.6 million in 2019, accounting for 32% of global mortality in 2019. In addition, CVDs is a significant source of health conflict and economic hardship. Based on the Medical Expenditure Panel Survey, the cost of CVDs in the United States between 2017 and 2018 was estimated at $378.0 billion, including $226.0 billion in expenditures and $151.8 billion in lost future productivity [1, 2].

1.1. AI in Healthcare: A New Cardiology Era

The potential for AI to automate processes, enhance decision-making, and enable new discoveries has broad implications, with possible applications in healthcare [3],transportation [4], industry [5], luxury [6] and more. Smart health, for instance, is the use of computational methods, data analysis, and artificial intelligence to the healthcare industry with the goal of enhancing patient care, administrative efficiency, and clinical results [3] and in enabling diseases prediction. However, the deadly nature of Cardiovascular Diseases necessitates the development of effective solutions that can help in the early detection of these diseases and, if possible, even predict their development. Electrocardiogram, Echocardiogram, Coronary Angiography, stress test, Magnetic Resonance Imaging or Intracoronary Ultrasound are traditional methods to

*Corresponding author. Tel.: +1234567890
Fax: +9876543210; E-mail: Author.Name@iasks.org
© 2011 International Association for Sharing Knowledge and Sustainability.
DOI: 10.5383/JUSPN.18.02.001

detect these diseases. However, technological advances, particularly Information and Communication Technologies (ICTs) and the growth of Artificial Intelligence and its derivatives, are improving the quality of healthcare services and facilitating the detection of CVDs. In addition, AI technologies are considered the next revolution in cardiology because they can accelerate and improve patient care outcomes. Moreover, AI will soon transform the field of cardiovascular health, as its tools could outperform specialists in detecting or even predicting CVDs [7, 8].

1.2. Heart Rate Variability as a CVD Indicator

Recently, there has been a rise in interest in using Heart Rate Variability (HRV) as a predictor of CVDs, particularly with the advent of AI and the data analysis capabilities afforded by its branches: Machine Learning and Deep Learning. Furthermore, HRV is defined as the beat-to-beat variation in heart rate or the length of the RR peak interval, where R is a QRS complex wave taken from a cardiac ECG signal. Because changes in the autonomic control of the heart may be interpreted from temporal fluctuations in heart rate, the parameters retrieved from HRV data are classified into three types: time domain, frequency domain, and non-linear parameters [9]. Table 1 below lists these categories:

Table 1. Heart Rate Variability Parameters.

Group	Parameter	Unit	Description
Time Domain Parameters	Mean NN	(ms)	Mean of NN interval
	SDNN	(ms)	Standard deviation of NN intervals
	RMSSD	(ms)	Square root of the mean squared differences of successive NN intervals
	pNN50	(ms)	Proportion of interval differences of successive NN intervals greater than 50 ms
Frequency Domain Parameters	VLF	(ms^2)	Power in very low frequency range (0–0.04 Hz)
	LF	(ms^2)	Power in low frequency range (0.04–0.15 Hz)
	HF	(ms^2)	HF ms2 Power in high frequency range (0.15–0.4 Hz)
	LF/HF	(ratio)	Ratio of LF over HF
Non-Linear Parameters	SD1	(ms)	Standard deviation of points perpendicular to the axis of line of identity or the successive intervals scaled by $\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}var(RR_{\mathrm{n}} - RR_{\mathrm{n+1}})}$
	SD2	(ms)	Standard deviation of points along the axis of line of identity, or $\sqrt{2SDNN^2 - \frac{1}{2}SD1^2}$
	SD1/SD2	(ratio)	Ratio of SD1 over SD2

1.3. Prediction of CVDs with HRV; State of the Art

There has been a surge in recent years in the number of researches looking at the ability to diagnose CVDs by measuring HRV characteristics. AI has demonstrated its efficacy and precision in this field, and researchers are increasingly turning to AI models to examine a wide range of HRV data.

In [10], for instance, the authors constructed a model to assess several HRV variables and predict the onset of ventricular tachycardia (VT) using the Fast Fourier Transform (FFT) and the Blackman Harris window technique. Additionally, the authors in [11] created an Artificial Neural Networks (ANN) classifier to predict the incidence of VT and trained it using the "PhysioNet Spontaneous Ventricular Tachyarrhythmia Database" [12]. They used a number of different criteria to assess performance, recording rates of 76.60% for accuracy, 82.9% for sensitivity, and 71.4% for specificity. In addition, the authors of [13] employed Multilayer Perceptron (MLP), Radial Basis Function (RBF), and Support Vector Machines (SVM) to make predictions about cardiovascular risk. Accuracy of their best model was 96.67%. Additionally, authors in [14] employed SVM to create a prediction model to predict cardiovascular risk following Myocardial Infarction, and the model accuracy was 89%.

In addition, the authors of [15] used the k-Nearest Neighbor and Multilayer Perceptron Neural Network algorithms to develop models that predict Sudden Cardiac Death (SCD). Their models were trained using the "PhysioNet Sudden Cardiac Death Holter database" [16] and the "PhysioNet Normal Sinus Rhythm database" [17], and their results showed an accuracy of 99.73% for the first minute, 96.52% for the second minute, 90.37% for the third minute, and 83.96% for the fourth minute. And in [18], authors performed the same study using SVM and Probabilistic Neural Network (PNN) to predict SCD two minutes beforehand. SVM and PNN achieved 96.36% and 93.64% accuracy in predicting sudden cardiac death using the "PhysioNet Sudden Cardiac Death Holter database" [16] and the "PhysioNet MIT Normal Sinus Rhythm database" [17].

Besides, in [19], the authors developed a novel SVM, Tree-Based Classifier, Artificial Neural Network, and Random Forest models to automate cardiovascular risk classification for hypertension patients. Using the "Smart Health for Assessing the Risk of Events through ECG database" [20], the authors were able to train their data with a sensitivity of 71.4% and a specificity of 87.8%. In addition, authors in [21] created an Artificial Neural Networks model that examines respiratory rate in addition to HRV data to identify ventricular tachycardia an hour before it manifests. Their model has a sensitivity of 88%, a specificity of 82%, and an area under the curve of 93%. The authors in [22] also employed a statistical model called MIL to predict CVDs using HRV characteristics. As they noted, their model was quite accurate. In addition, the authors of [23] developed and trained a variety of classification methods, including K Nearest Neighbor, Decision Tree, Naive Bayes, Logistic Regression, Support Vector Machine, Neural Network, and Vote. They used the "UCI Heart Diseases Repository" [24], to train their models. It was shown that the models had an accuracy of 87.4% in predicting CVDs.

1.4. Outline & Main Contributions

Several AI models were developed in this study to predict CVDs and related events where eight different models were implemented. The dataset and the preparation processes that were performed to get the data ready for the models are described in Section 2. below. A description of the models developed may be found in Section 3., while Section 4. contains a listing and discussion of the results.

Despite the fact that several Machine Learning implementations have been performed in CVD detection and prediction, this article aims to propose ML models that have either never been used in this field or to propose models already in use and improve their performance. Therefore, this article aims to propose ML models capable of predicting CVDs with improved performance that outperforms previous implementations. The result obtained by the models is a binary result, stating whether a CVD is detected or not. The article thus contributes to the ML field in predicting CVDs:

- Proposing use of new models in the prediction of CVDs
- Enhancing and boosting the performance of ML in CVDs

2. Materials & Methods

2.1. Dataset

The dataset used in this study is the "PhysioNet Smart Health for Assessing the Risk of Events via ECG Database" (SHAREEDB) [20] that is offered by the PhysioNet online data repository. This dataset was collected to investigate the efficiency of classifying hypertensive patients at higher risk for cardiac and cerebrovascular events using HRV characteristics. It consists of 139 records of 24-hour Electrocardiographic (ECG) Holter recordings, each containing three ECG signals sampled at a rate of 128 samples per second with a precision of 8 bits. The population from which the data were gathered included 49 women and 90 men aged 55 and up. They were followed up for 12 months to record the occurrence of cardiovascular and cerebrovascular events. During the follow-up period, 17 patients experienced such event, including 11 Myocardial Infarctions, 3 strokes, and 3 syncopal events. The dataset also includes some demographic and clinical information about the subjects, such as their age, sex, any vascular events, and others. Figure 1 below describes the specifications of the dataset in use: SHAREEDB.

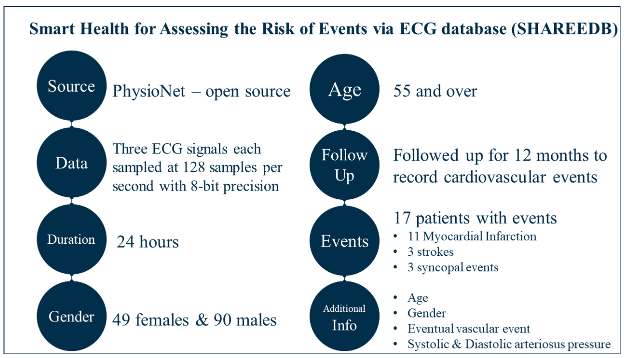

Fig. 1. SHAREEDB Description and Specs.

2.2. Data Filtering & Preprocessing

Since the SHAREEDB dataset contains Electrocardiogram (ECG) signals gathered in a laboratory, there may be substantial background noise that must be eliminated before the data is fed into the AI models. Before feeding the data into the models, it is crucial to clean the data and eliminate the noise in order to produce high-quality ECG signals. Briefly described below are the procedures used to clean and prepare the data for this study:

2.2.1. Filtering & Artifacts Removal

The 3-channel ECG Holter device was used to record the data included in the dataset files. A normal ECG signal has a frequency range of 0.05 Hz to 100 Hz. However, there are a number of signals interreferences that may affect ECG recordings, including baseline drift, channel interference, power line interference, muscle movement interference, and electrode contact interference. Raw ECG readings often include two forms of noise [25]:

- High frequency noise: current conduction noise, white Gaussian noise, Electromyogram or motion noise.
- Low-frequency noise: baseline drift and electrode contact loss

We can successfully identify the kind of noise and then pick the approaches to employ to decrease the noise or eradicate the artifact if we have a thorough grasp of each noise artifact. Various sounds are caused by various things, including:

- Power Line Interference: is caused by harmonics of electromagnetic interference through the power line and the electromagnetic field of nearby electrical equipment and is between 50 Hz and 60 Hz.
- White Gaussian Noise: is similar to channel noise in nature but is difficult to identify its sources because they occur at different levels and are random in nature.
- Electromyogram/Motion Noise: generated by the electrical activity of the muscles or the change in the electrode-skin impedance due to changes in skin temperature, humidity, etc.
- Baseline Drift: low-frequency noise, typically around 1 Hz and caused by respiration and rapid body movements
- Electrode Contact Loss: is caused by loss of contact between the electrode and the skin

Because of this, the following filters are effective in getting rid of both low-frequency and high-frequency artifacts and has been adopted, in this study, to clean the data before being used:

- IIR Notch Filters: are used to remove power line interference and/or motion artifacts in a specific frequency spectrum
- FIR Filters: are very stable filters and operate in the range of 1 Hz to 100 Hz making them suitable for ECG data cleaning

2.2.2. R Peaks Detection

The electrical activity of the heart muscle may be seen in an ECG signal throughout time. The ECG represents the amplified sum of the electrical depolarization of muscle cells that causes the heart muscle to contract during a certain time period. Three components make up the electrocardiogram signal: The P-Wave, the QRS complex, and the T-wave. The ventricular depolarization represented by the QRS complex is the electrical impulse as it travels through the ventricles. Immediately following each other in rapid succession are the Q wave, the R wave, and the S wave. Because HRV is defined as the difference between two successive RR periods, the R Peaks are the peaks to be discovered in this investigation. The R Peaks may be found using any of the available detection methods [26, 27]. These algorithms include:

- Hamilton
- Christov
- Engelse and Zeelenberg
- Pan and Tompkins
- Stationary Wavelet Transform
- Two Moving Average

According to [27], Engelse and Zeelenberg was selected as the most accurate peak detection algorithm. Although the tests were performed on a different data set, Engelse and Zeelenberg was selected for R peak detection in this study based on the recommendation of authors.

2.2.3. Calculation of RR Intervals

Heart Rate Variability is defined as the RR intervals or the difference between two consecutive R peaks, which are then calculated using the required equations.

2.2.4. Outliers Removal

After the RR intervals are detected, the outliers, defined as points that are extremely far from the mean, are removed and replaced with the mean value.

2.2.5. Extract HRV features

Finally, the HRV features were calculated using the appropriate mathematical formulas. In this study, 26 HRV features were calculated, and despite the high number of features calculated, the use of all features gave good results.

2.3. Artificial Intelligence Models

Cardiology is defined as the healthcare sector concerned with heart health, and the usage of AI in this discipline is rapidly expanding. AI has showed excellent accuracy and efficiency in identifying CVDs, and owing to its strong capacity to evaluate cardiac data, it may sometimes go beyond professional diagnosis and even be utilized in predicting CVDs rather than detecting them [28, 29]. Furthermore, AI is notable for its diverse branches that are applied in various aspects of life all over the globe. Figure 2 below shows the different branches of AI. In this research, AI branches such as Machine Learning, Ensemble Learning, and Deep Convolutional Neural Networks were applied:

- **Classical Machine Learning Algorithms [30]:** are algorithms that give computers learning potential by training them with experimental data and generating models based on these data, enabling them to make decisions in new situations such as: Support Vector Machines, Naïve Bayes, Logistic and Linear Regression and others.
- **Ensemble Learning [31]:** is a special branch of ML where its algorithms are based on merging predictions from different models. Some of these models are XGBoost, AdaBoost, Gradient Boosting, LightGBM and others.
- **Deep Convolutional Neural Networks (DCNNs) [30]:** are a type of Neural Networks that are used to analyze data with a grid-like structure. However, these networks are intended for analyzing multidimensional data such as images and videos. Using these networks to analyze tabular data may require transforming the data used. Nevertheless, there are several models that offer transformation of tabular data for use in DCNNs, such as TabNet, GrowNet, TreeEnsemble Layers, TabTransformers, Self-Normalizing Neural Networks, Neural Oblivious Decision Ensembles (NODE), AutoInt, and Deep & Cross Neural Networks (DCNs) [32].

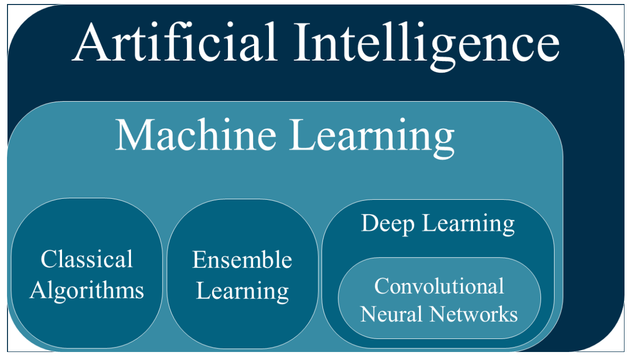

Fig. 2. SHAREEDB Description and Specs.

3. Construction of AI Models

In this study, different AI models were used to analyze HRV features to detect heart diseases and events. However, before passing the extracted features to the models, some data fitting steps were performed, as explained below.

3.1. Data Adjustment

Given that only 17 of the 139 patients in the research suffered a cardiovascular event throughout the 12-month follow-up period, the retrieved HRV features show an unbalanced identity, with the majority falling into the "no cardiovascular event" class. Because the percentage of non-defected subjects is 122 of 139, the performance of the prediction models may be harmed, implying the usage of data modifications such as balancing and scaling:

- **Synthetic Minority Over-sampling Technique (SMOTE):** a data expansion in which new samples are drawn from existing ones to oversample the minority class
- **Preprocessing Standard Scaling:** the standardization of characteristics is achieved by removing the mean of the data and scaling it to a unit variance

3.2. Building the Models; Hyperparameters to be Considered

After applying the necessary data fitting steps to the extracted HRV features, they are then passed to the models created for fitting with the thresholds listed below:

3.2.1. Support Vector Machines

SVM is a supervised Machine Learning algorithm that is fed labeled training data to learn how to assign labels to objects based on examples, and then gain the ability to predict the category of new example(s) [30]. The performance of the SVM model is affected by the following hyperparameters [33]:

- **Kernel:** the function that converts the input data into the required form such as linear, polynomial and radial basis function (RBF).
- **Regularization:** denotes the misclassification or error term and is expressed as hyperparameter "C".
- **Gamma:** interpret how far the effect of a single training sample extends
- **Class Weight:** used for imbalanced datasets and defines the weight of the classes to be predicted

3.2.2. TabTransformers

TabTransformers is a model based on transformers whose layers convert categorical feature embeddings into robust contextual embeddings to achieve higher prediction accuracy, and is affected by the following hyperparameters [34]:

- **Activation Function:** defines how the weighted sum of the input is converted into an output of a node in a network layer
- **Number of Heads:** specifies the number of heads of attention

- **Dropout:** regularization to reduce overfitting and improve generalization of deep neural networks
- **MLP Hidden Units Factors:** MLP hidden layer units, as factors of the number of inputs.
- **Learning Rate:** is the shrinkage step size used in updating to prevent overfitting

3.2.3. Deep Neural Networks

These networks are algorithms that mimic human brain cells called neurons. In general, these networks use brain simulations to improve their learning and increase the accuracy of the models. The structure of DNNs consists of more than two interconnected layers and is affected by the following hyperparameters [35]:

- **Number of Layers:** input, output, and the hidden layers that define the structure of the network.
- **Units:** denotes the output of each layer.
- **Activation Function:** also known as the "transfer function", which defines how the weighted sum of the input is converted into an output from one or more nodes in a layer of the network
- **Number of Epochs:** a complete pass through all rows of the training data
- **Batch Size:** samples that the model examines within each epoch before updating the weights
- **Learning Rates:** a variable that controls how the optimizer's learning rate changes over time
- **Momentum:** is the "delay" in learning the mean and variance

3.2.4. AdaBoost

AdaBoost is a meta-estimator that first fits a classifier to the original data and then fits additional copies of the classifier to the same data, changing the weights of misclassified instances so that subsequent classifiers examine them extensively, leading to an improved result [36]:

- **Number of Estimators:** the number of base estimators or weak learners to be used in the dataset
- **Learning Rate:** is the step size used in the update to prevent overfitting

3.2.5. XGBoost

XGBoost is an Ensemble Learning algorithm that also belongs also to the Machine Learning AI Branch. XGBoost, eXtreme Gradient Boosting package, is a scalable implementation of the gradient boosting framework built with an efficient linear model solver and a tree learning algorithm with hyperparameters [37]:

- **Booster:** the type of model to run at each iteration
- **Learning Rate:** is the step size shrinkage used during the update to prevent overfitting
- **Gamma:** specifies the minimum loss reduction required to perform splitting
- **Max Depth:** the parameter used to control overfitting, as a higher depth allows the model to learn relationships that are very specific to a given sample
- **Min Child Weight:** defines the minimum sum of weights of all observations required in a child
- **Max Delta Step:** makes updating more conservative
- **Sub Sample:** denotes the fraction of observations that are randomly selected for each tree
- **Lambdas:** is used to handle the regularization part
- **Alpha:** is used in case of very high dimensionality to make the algorithm run faster during implementation
- **Tree Method:** Algorithm for tree construction
- **Scale Weight:** control the weight of positive-negative classes
- **Objective:** defines the loss function to be minimized

3.2.6. Logistic Regression

Logistic Regression is a Machine Learning algorithm that analyzes data for classification and is a supervised algorithm that sorts data into two categories. The algorithm is named after the function that is at the core of the method, the logistic function. There are several forms for LR and in this article we will use binary logistic regression, where the target variable has only two possible outcomes. The performance of LR is affected by three important hyperparameters [38]:

- **Solver:** uses a Coordinate Descent (CD) algorithm that solves optimization problems by successively performing approximate minimization along coordinate directions or coordinate hyperplanes
- **Penalty (Regularization):** is any modification of a learning algorithm that aims to reduce its generalization error, but not its training error
- **C:** the inverse of the regularization strength in Logistic Regression
- **Class Weight:** weight of the classes to be predicted

3.2.7. TabNet

TabNet is a model that uses sequential attention to select which features to infer at each decision step, and is influenced by the following hyperparameters [39]:

- **Optimizer:** an algorithm that modifies the neural network attributes, such as weights and learning rate.
- **Learning Rate:** is the step size used in updating to prevent overfitting
- **Batch Size:** number of examples per batch.

3.2.8. Deep Convolutional Neural Networks: Neural Oblivious Decision Ensembles (NODE)

Neural Oblivious Decision Ensembles is a model with a layered structure built from differentiable oblivious trees, which are decision tables that decompose the data along dd-splitting features and compare each feature to a learned threshold. It was trained in an end-to-end manner using backpropagation and is affected by the following hyperparameters [40]:

- Number of Layers: Number of layers forming the Neural Network
- Number of Trees: Number of trees in each layer
- Depth: Depth of the tree
- Learning Rate: is the shrinkage step size used in the update to prevent overfitting

3.3. Technical Environment Specifications

To implement this study, the computer used carried the below mentioned specifications:

- **Hardware Specs:**
  - **CPU:** Intel(R) Core i7-7500U CPU @ 2.70GHz
  - **RAM:** 16.0 GB DDR4
- **Operating System:** Windows 10 Home
- **Programming Language Used:** python 3.9
- **Libraries Used:**
  - **wfdb:** used to read data from the PhysioNet binary files [41]
  - **Scipy Signal Library:** provides efficient functions for both IIR Notch and FIR filters [42]
  - **py-ecg-detectors:** provide R Peaks detection algorithms [43]
  - **Scipy Zscore:** used for outliers' removal [44]
  - **SMOTE:** to apply Synthetic Minority Over-sampling [45]
  - **SKLearn Preprocessing Standard Scaling:** to apply standard scaling [46]

3.4. Wrapping Up, Training, Prediction, and Optimization

Once the models were created, they were trained using the extracted HRV features. The models were then evaluated using several performance metrics, namely accuracy, precision, recall, F1 score, specificity, and negative predictive value. The results obtained are detailed and discussed in the next section. Figure 3 below describes the overall structure of the implemented system.

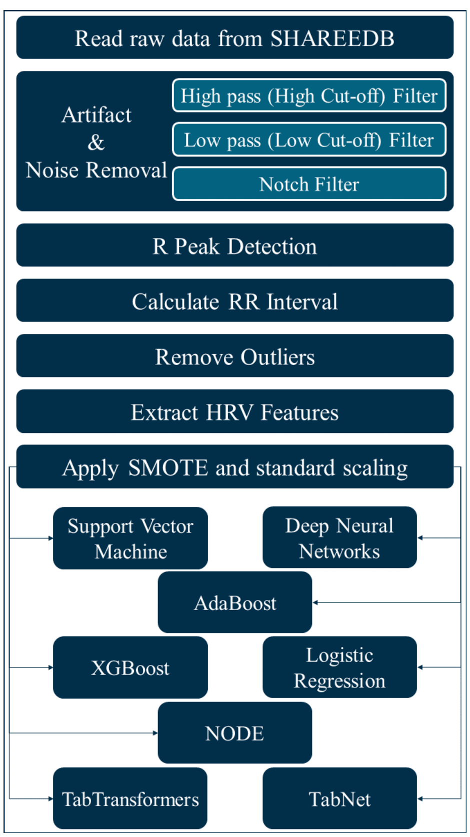

Fig. 3. Overall Architecture Followed in this Study.

4. Results & Discussion

The created models were trained with the HRV features. The eight models were evaluated with the metrics of Accuracy, Precision, Recall, Specificity, Negative Predictive Value NPV, and F1 Score. For better measurement, and to be aware of overfitting, Repeated K-fold Cross Validation [47] was implemented with 10 folds and repeated 5 times. Beside detection of overfitting, the use of K-fold cross validation ensure that the recorded results are not obtained from an optimistic execution. Consequently, the performance graphs are illustrated in Figure 4, 5 & 6 respectively, where the first shows the graphs for classical ML models, the second shows the graphs related to Ensemble ML models and the third shows the graphs of the DL models. In addition, the results are shown in Table 2 below, and the values of accuracy, precision, recall, specificity, negative predictive value, and F1 score are denoted as AC, PR, RE, SP, NPV, and F1, respectively. In addition, the values of the hyperparameters used are listed in the table.

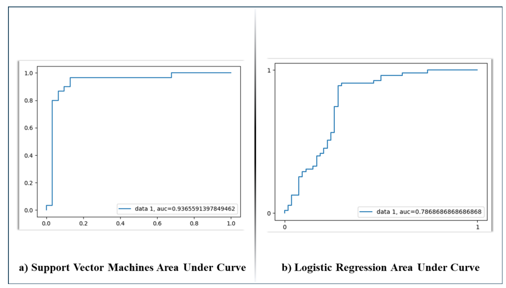

Fig. 4. Classical ML Models Performance Graphs.

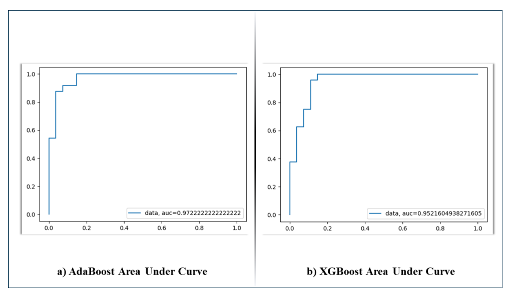

Fig. 5. Ensemble Learning ML Models Performance Graphs.

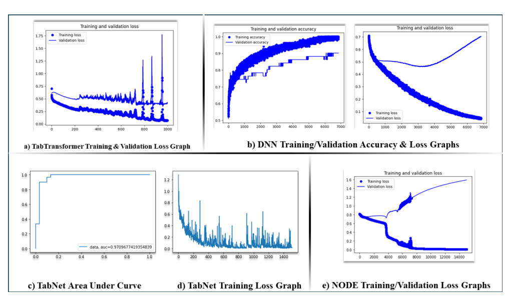

Fig. 6. Deep Learning Models Performance Graphs.

Table 2. AI Models Evaluation Metrics & Hyperparameters Used.

#	Model	Hyperparameters Used		AC	PR	RE	SP	NPV	F1
		Parameter	Value						
1	Support Vector Machines	Train/Test Split	0.75(0.25)	91.80%	87.87%	96.66%	87.09%	96.42%	92.06%
		Kernel	rbf						
		Regularization (C)	2.66						
		Gamma	0.141						
2	TabTransformers	Train/Test Split	0.72(0.28)	90.38%	86.66%	96.29%	91.22%	84%	95.45%
		Activation Function	Sigmoid						
		Number of Transformer Blocks	1024						
		Number of Attention Heads	1024						
		Dropout Rate	0.25						
		MLP Hidden Units	[1024,512]						
		Learning Rate	0.05						
		Epochs	1000						
3	Deep Neural Networks	Train/Test Split	0.79(0.21)	90.19%	85.18%	95.83%	85.18%	95.83%	90.19%
		Layers	Input/3 Hidden/Output						
		Units	512/256/128/64/1						
		Activation Function	tanh/tanh/tanh/sigmoid						
		Dropout	Before Output Layer 0.2						
		Optimizer	SGD						
		Epochs	6850						
		Batch Size	250						
		Learning Rate	0.005						
		Momentum	default						
4	AdaBoost	Train/Test Split	0.79(0.21)	89.50%	87.20%	94.60%	84.90%	93.60%	90.80%
		Number of Estimators	200						
		Learning Rate	1						
5	XGBoost	Train/Test Split	0.79(0.21)	89.10%	86.00%	93.80%	85.10%	92.50%	89.10%
		Booster	gbtree						
		Learning Rate	0.01						
		Gamma	0.1						
		Maximum Depth	10						
		Minimum Child Weight	0.01						
		Max Delta Step	0						
		Sub Sample	0.75						
		Lambda	1						
		Alpha	0.01						
		Tree Method	Auto						
6	Logistic Regression	Train/Test Split	0.71(0.29)	80.73%	76.56%	89.09%	72.22%	86.66%	82.35%
		Solver	newton-cg						
		Regularization (Penalty)	none						
		C	3.1						
7	TabNet	Train/Test Split	0.79(0.21)	76%	74.70%	82.60%	74.70%	80.50%	76.50%
		Learning Rate	0.9						
		Batch Size	1024						
		Virtual Batch Size	1024						
8	NODE	Train/Test Split	0.71(0.29)	76.92%	77.77%	73.68%	80%	76.19%	75.67%
		Number of Layers	5						
		Depth	10						
		Number of Trees (per layer)	1						
		Learning Rate	0.1						
		Batch Size	26						
		Epochs	9000						

4.1. Discussion

In this study, several models were created to analyze HRV characteristics to detect cardiovascular risks. The results obtained demonstrate the high efficiency of AI models in predicting cardiovascular disease. However, the results obtained in this study outperformed previous implementations.

First, the authors in [19] applied similar models to the same dataset. Nevertheless, the results obtained in this study exceeded their results. For example, their SVM model recorded accuracy, recall and specificity results were 89.00%, 86.30% and 91.80% respectively, whereas our results are 91.80%, 96.66% and 87.09% for the same performance metrics. In addition, the performance metrics of their Multi-Layer Perceptron model were Accuracy: 78.10%, Recall: 86.30%, Specificity: 69.90% and our model recorded 90.19%, 95.83% and 85.18% for the same metrics.

In addition, our SVM model achieved 91.80% accuracy, the highest performance among all previous implementations. For example, the SVM models in [13] recorded an accuracy of 88.64%, 82.95% and 82.58% for the Linear, Polynomial and RBF kernels, respectively. Moreover, the accuracy of SVM in [14, 19, 23] was 79.81%, 89.00% and 85.19%, respectively. Even though the accuracy is close, other metrics such as Precision and Recall clearly outperform the previous results by a large margin. Knowing that Recall measures how a model correctly classifies True Positives, the models presented in this study are more accurate in predicting whether a person will have a CVD in future. The high recall for SVM, DNN, and XGBoost, which are 96.66%, 95.83% and 93.80%, respectively, reflects the highest ability of all implementations to correctly predict that a person is in the cardiovascular risk zone.

Likewise, the DNN model presented in this article also outperforms all previous implementations. The accuracy of this model is 90.19%, whereas the multilayer perceptron in [13, 19] is 86.67% and 78.10%, and the accuracy of artificial neural networks in [11, 21] is 76.60% and 85.30%. Moreover, precision and recall are significantly higher than the previous implementations, which also reflects a higher capability in cardiovascular risk detection. Table 3 provides a detailed comparison between the results of the models presented in this article and the previous implementations. The symbols of the performance metrics used in this table are similar to those in Table 2, and an "NA" symbol indicates that the corresponding metric was not mentioned in the associated study.

On the other hand, none of the previous implementations used XGBoost, which also outperformed the previous implementations with an accuracy of 98.10% and a recall of 94.60%, reflecting high efficiency in predicting cardiovascular risk, in contrast to the implementation of NODE, which achieved an accuracy of 76.92%, which is not comparable with the previous implementations.

Finally, the SVM, DNN, and XGBoost models discussed in this study can be considered the most accurate models for predicting cardiac disease and events. Even the implementations in [15, 19] had higher accuracy and relatively higher recall, but their models were developed to detect Sudden Cardiac Death only minutes before its occurrence. For example, the model mentioned in [15] achieved 99.73% accuracy in predicting sudden cardiac death one minute before its onset, but the performance drops to 83.93% when the event is predicted four minutes before its occurrence. However, the models presented here are able to predict cardiovascular disease 12 months before its onset, demonstrating high efficiency in predicting cardiovascular disease and cardiac events long before their onset, thus increasing confidence in the use of AI in detecting and predicting cardiac disease and related events.

4.2. Challenges & Future Recommendations

Although Machine Learning are ready to play a significant role in predicting CVDs, there are a number of potential obstacles that might occur in the course of their deployment. What follows are some of the most typical problems that arise in such a setting:

- Data Readiness and Availability: Data determines machine learning model performance. The availability of more data will help in improving the performance of the smart models and therefore increase their accuracy in predicting CVDs. However, the availability of data is prone to different problems such as the legal or ethical restrictions. However, assumed available and accessible, the data to be used may be noisy since digital ECG recordings are more vulnerable to environmental noise. Artefacts—unwanted signals or signal distributions—interfere with the signal in noisy data. In this context, Intrinsic Artefacts come from the monitored body, whereas Extrinsic Artefacts come from their surroundings [48, 49]
- Data Privacy and Confidentiality: Although the technical structure of the models, data cleanliness and readiness, and other factors affect model accuracy, more data to train AI models usually improves their accuracy. For privacy and secrecy considerations, gathering data is the largest hurdle in constructing AI models in the real world. Society, governments, and organizations are enhancing data privacy and security. The European Union's General Data Protection Regulation (GDPR) [50], China's Cyber Security Law [51] and hundreds of other principles have been legislated worldwide. These restrictions safeguard private data, but also make it harder to gather data to train models, which makes it harder to increase model performance [49]
- Users Acceptability: User acceptability, adoption, and engagement are of the most significant obstacles to using AI and its branches to identify CVDs. Using those technologies to predict illnesses has met with mixed reception from users owing to concerns about privacy, discomfort, and other contextual factors
- Additional Computation-Cost: Due to the additional computing imposed by the added tasks such as data balancing and noise removal, an increase in computation time is obtained, and thus this imposes additional slowdowns that may impair the overall performance of the models

However, several approaches have been made to resolve those challenges in the attempt to enhance the feasibility of using AI and its descendants to predict cardiac illness. Those solutions are considered as hot topics that are being studied carefully nowadays:

- Automating Noise Removal: Before processing the signals, artifacts, both extrinsic and intrinsic, that obfuscate the signals should be eliminated or greatly reduced. This goal has already been accomplished by a number of existing solutions, some of which are discussed in Refs. [52]. Thus, research into automated noise reduction to clean and preprocess the data to enhance the precision of physical tiredness detection in the workplace is warranted
- Privacy Preserving: Data used in Machine Learning models training should be stored on a local server or distributed to decentralized storage and processing devices to construct and train the models. Thus, the model has complete access to the subject's data, whether anonymous or labelled by the subject. Federated learning (FL) may address this issue. Federated learning is a defined as collaborative distributed/decentralized machine learning privacy-preserving method that trains models without transferring data from edge devices to a central server. Instead, edge devices communicate learned models with the central server, which works as an aggregation station to create the global model without understanding the embedded data [53, 54]. The use of Federated Learning into CVDs prediction would help resolve privacy issues and therefore resolve the challenges in this regard
- Increase Accuracy, Explainability and Trust: Predicting the onset of cardiovascular disease is crucial in light of the growing health burden caused by this condition. The black box nature of the models used, however, must be reduced as much as possible, and the accuracy of AI tools and procedures in this area must be increased. Devices that are more accurate and easier to explain will be more likely to be employed as a CVDs prediction device. In this context, several technologies can be adopted such as the one mentioned in [55] that automates assessing the quality of a smart model

Table 3. Comparison with Previous Implementations.

Study	Model	AC	PR	RE	SP	NPV	F1
Our Study	Support Vector Machines	91.80%	87.87%	96.66%	87.09%	96.42%	92.06%
	TabTransformers	90.38%	86.66%	96.29%	91.22%	84.00%	95.45%
	Deep Neural Network	90.19%	85.18%	95.83%	85.18%	95.83%	90.19%
	AdaBoost	89.50%	87.20%	94.60%	84.90%	93.60%	90.80%
	XGBoost	89.10%	86.00%	93.80%	85.10%	92.50%	89.10%
	Logistic Regression	80.73%	76.56%	89.09%	72.22%	86.66%	82.35%
	TabNet	76.00%	74.70%	82.60%	74.70%	80.50%	76.50%
	NODE	76.92%	77.77%	73.68%	80.00%	76.00%	75.67%
[11]	Artificial Neural Network	76.60%	70.70%	82.90%	71.40%	NA	NA
[13]	Support Vector Machines (Linear Kernel)	88.64%	90.84%	86.36%	90.91%	86.96%	NA
	Support Vector Machines(Polynomial Kernel)	82.95%	80.85%	79.55%	86.36%	85.37%	NA
	Support Vector Machines (RBFKernel)	82.58%	79.45%	77.27%	87.88%	86.44%	NA
	Multi Layer Perceptron (Top15 Features)	86.67%	100%	73.33%	100%	78.95%	NA
[14]	Support Vector Machines	79.81%	21.15%	91.67%	79.08%	99.36%	NA
[15]	MLP (A Minute Before the SCD Event)	99.73%	NA	NA	NA	NA	NA
	K-NN (A Minute Before the SCDEvent)	98.32%	NA	NA	NA	NA	NA
[18]	SVM (2 minutes before VF Event)	96.36%	NA	NA	NA	NA	NA
	Penalized Neural Network	93.64%	NA	NA	NA	NA	NA
[19]	Support Vector Machines	89.00%	NA	86.30%	91.80%	NA	NA
	Multi Layer Perceptron	78.10%	NA	86.30%	69.90%	NA	NA
[21]	Artificial Neural Network	85.30%	83.30%	88.20%	82.40%	87.50%	NA
[22]	MIL Statistics Algorithm	85.47%	92.11%	86.42%	83.33%	NA	NA
[23]	Vote	87.41%	NA	NA	NA	NA	NA
	Naïve Bayes	84.81%	NA	NA	NA	NA	NA
	Support Vector Machines	85.19%	NA	NA	NA	NA	NA

5. Conclusion

Ultimately, AI will determine the fate of humans, they say. However, the widespread adoption and use of these technologies today proves that they are no longer science fiction. The field of cardiology, as well as methods for diagnosing and treating Cardiovascular Disease, will benefit from the development of AI, which could one day enable accurate prediction of disease. Research has produced a number of models that can accurately predict the occurrence of cardiac problems or events, boosting confidence in AI and its applications in medicine. If these models are operational in real time, this will undoubtedly contribute to the development of personalized and continuous monitoring that can be used to monitor the heart health of patients or even the health of workers who work in stressful environments or for extremely long periods of time.

Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) [funding reference number 06351], Fonds Québécois de la Recherche sur la Nature et les Technologies (FRQNT), and Centre d'Entrepreneuriat et de Valorisation des Innovations (CEVI).

References

[1] Roth, Gregory A., et al. "Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study." Journal of the American College of Cardiology 76.25 (2020): 2982-3021.

[2] Tsao, Connie W., et al. "Heart disease and stroke statistics—2022 update: a report from the American Heart Association." Circulation 145.8 (2022): e153-e639.

[3] Patel, Vishv, Devansh Shah, and Nishant Doshi. "Emerging Technologies and Applications for Smart Cities." J. Ubiquitous Syst. Pervasive Networks 15, no. 02 (2021): 19-24.

[4] Karakostas, B. (2023). Control of autonomous UAV using an onboard LSTM neural network. Journal of Ubiquitous Systems & Pervasive Networks, 18(1), 09-14.

[5] Haricha, Karim, Azeddine Khiat, Yassine Issaoui, Ayoub Bahnasse, and Hassan Ouajji. "Towards smart manufacturing: Implementation and benefits." J. Ubiquitous Syst. Pervasive Networks 15, no. 02 (2021): 25-31.

[6] Zaki, Marquos, Ali Alquraini, and Tarek R. Sheltami. "Home Automation using EMOTIV: Controlling TV by Brainwaves." J. Ubiquitous Syst. Pervasive Networks 10, no. 1 (2018): 27-32.

[7] National Heart Lung and Blood Institute. "In Brief: Your guide to living well with heart disease (NIH Publication 06-5716)." Washington, DC: National Institutes of Health (2005).

[8] Johnson, Kipp W., et al. "Artificial intelligence in cardiology." Journal of the American College of Cardiology 71.23 (2018): 2668-2679.

[9] Rajendra Acharya, U., et al. "Heart rate variability: a review." Medical and biological engineering and computing 44.12 (2006): 1031-1051.

[10] Baumert, Mathias, et al. "Forecasting of ventricular tachycardia using scaling characteristics and entropy of heart rate time series." World Congress on Medical Physics and Biomedical Engineering 2006. Springer, Berlin, Heidelberg, 2007.

[11] Joo, Segyeong, Kee-Joon Choi, and Soo-Jin Huh. "Prediction of ventricular tachycardia by a neural network using parameters of heart rate variability." 2010 Computing in Cardiology. IEEE, 2010.

[12] "Spontaneous Ventricular Tachyarrhythmia Database v1.0." Spontaneous Ventricular Tachyarrhythmia Database v1.0, 2 May 2022, https://physionet.org/content/mvtdb/1.0. Accessed on 01 March 2022

[13] Ramirez-Villegas, Juan F., et al. "Heart rate variability dynamics for the prognosis of cardiovascular risk." PloS one 6.2 (2011): e17060.

[14] Song, T., Qu, X. F., Zhang, Y. T., Cao, W., Han, B. H., Li, Y., ... & Da Cheng, H. (2014). Usefulness of the heart-rate variability complex for predicting cardiac mortality after acute myocardial infarction. BMC cardiovascular disorders, 14(1), 1-8.

[15] Ebrahimzadeh, E., Pooyan, M., & Bijar, A. (2014). A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. PloS one, 9(2), e81896.

[16] "Sudden Cardiac Death Holter Database v1.0.0." Sudden Cardiac Death Holter Database v1.0.0, 2 July 2004, https://physionet.org/content/sddb/1.0.0. Accessed 01 March 2022.

[17] "MIT-BIH Normal Sinus Rhythm Database v1.0.0." MIT-BIH Normal Sinus Rhythm Database v1.0.0, 3 Aug. 1999, https://physionet.org/content/nsrdb/1.0.0. Accessed 01 March 2022.

[18] Murukesan, L., et al. "Machine learning approach for sudden cardiac arrest prediction based on optimal heart rate variability features." Journal of Medical Imaging and Health Informatics 4.4 (2014): 521-532.

[19] Melillo, Paolo, et al. "Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis." PloS one 10.3 (2015): e0118504.

[20] "Smart Health for Assessing the Risk of Events via ECG Database v1.0.0." Smart Health for Assessing the Risk of Events via ECG Database v1.0.0, 19 May 2015, https://physionet.org/content/shareedb/1.0.0. Accessed 01 March 2022.

[21] Lee, Hyojeong, et al. "Prediction of ventricular tachycardia one hour before occurrence using artificial neural networks." Scientific reports 6.1 (2016): 1-7.

[22] Lan, Kun-chan, et al. "Toward hypertension prediction based on PPG-derived HRV signals: A feasibility study." Journal of medical systems 42.6 (2018): 1-7.

[23] Amin, Mohammad Shafenoor, Yin Kia Chiam, and Kasturi Dewi Varathan. "Identification of significant features and data mining techniques in predicting heart disease." Telematics and Informatics 36 (2019): 82-93.

[24] "UCI Machine Learning Repository: Data Set." UCI Machine Learning Repository: Data Set, https://archive.ics.uci.edu/ml/datasets/heart+disease. Accessed 01 March 2022.

[25] Luo, Shen, and Paul Johnston. "A review of electrocardiogram filtering." Journal of electrocardiology 43.6 (2010): 486-496.

[26] Ashley, Euan A., and Josef Niebauer. "Cardiology explained." (2004).

[27] Eilers, Justus, Jonas Chromik, and Bert Arnrich. "Choosing the Appropriate QRS Detector." BIOSIGNALS. 2021.

[28] Mathur, Pankaj, et al. "Artificial intelligence, machine learning, and cardiovascular disease." Clinical Medicine Insights: Cardiology 14 (2020): 1179546820927404.

[29] Mathur, Pankaj, et al. "Artificial intelligence, machine learning, and cardiovascular disease." Clinical Medicine Insights: Cardiology 14 (2020): 1179546820927404.

[30] Ertel, Wolfgang. Introduction to artificial intelligence. Springer, 2018.

[31] Dong, Xibin, et al. "A survey on ensemble learning." Frontiers of Computer Science 14.2 (2020): 241-258.

[32] Gorishniy, Yury, et al. "Revisiting deep learning models for tabular data." Advances in Neural Information Processing Systems 34 (2021): 18932-18943.

[33] Noble, William S. "What is a support vector machine?." Nature biotechnology 24.12 (2006): 1565-1567.

[34] Huang, Xin, et al. "Tabtransformer: Tabular data modeling using contextual embeddings." arXiv preprint arXiv:2012.06678 (2020).

[35] Anderson, James A. An introduction to neural networks. MIT press, 1995.

[36] Schapire, Robert E. "Explaining adaboost." Empirical inference. Springer, Berlin, Heidelberg, 2013. 37-52.

[37] Chen, Tianqi, et al. "Xgboost: extreme gradient boosting." R package version 0.4-2 1.4 (2015): 1-4.

[38] Kleinbaum, David G., et al. Logistic regression. New York: Springer-Verlag, 2002.

[39] Arik, Sercan Ö., and Tomas Pfister. "Tabnet: Attentive interpretable tabular learning." Proceedings of the AAAI

Conference on Artificial Intelligence. Vol. 35. No. 8. 2021.

[40] Popov, Sergei, Stanislav Morozov, and Artem Babenko. "Neural oblivious decision ensembles for deep learning on tabular data." arXiv preprint arXiv:1909.06312 (2019).

[41] "Wfdb." PyPI, 27 June 2022, https://pypi.org/project/wfdb. Accessed on 10 March 2022

[42] Signal Processing (scipy.signal) — SciPy v1.9.3 Manual, 1 June 2022, https://docs.scipy.org/doc/scipy/reference/signal.html. Accessed on 05 June 2022

[43] "Py-ecg-detectors." PyPI, 1 June 2022, pypi.org/project/py-ecg-detectors.

[44] Scipy.stats.zscore — SciPy v1.9.3 Manual, 1 June 2022, https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.zscore.html. Accessed on 15 March 2022.

[45] "SMOTE — Version 0.9.1." SMOTE — Version 0.9.1, 1 Sept.2000, https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html. Accessed on 15 March 2022.

[46] "Sklearn.Preprocessing.StandardScaler." Scikit-learn, 1 Jan.2000, https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html. Accessed on 15 March 2022.

[47] Fushiki, Tadayoshi. "Estimation of prediction error by using K-fold cross-validation." Statistics and Computing 21.2 (2011): 137-146.

[48] Islam, Md, Amir Rastegarnia, and Saeid Sanei. "Signal artifacts and techniques for artifacts and noise removal." Signal processing techniques for computational health informatics. Springer, Cham, 2021. 23-79.

[49] Talha, Muhammad, Nabil Elmarzouqi, and Anas Abou El Kalam. "Quality and Security in Big Data: Challenges as opportunities to build a powerful wrap-up solution." J. Ubiquitous Syst. Pervasive Networks 12, no. 1 (2020): 9-15.

[50] Albrecht, Jan Philipp. "How the GDPR will change the world." Eur. Data Prot. L. Rev. 2 (2016): 287.

[51] Parasol, Max. "The impact of China's 2016 Cyber Security Law on foreign technology firms, and on China's big data and Smart City dreams." Computer law & security review 34.1 (2018): 67-98.

[52] Daly, Ian, et al. "On the automated removal of artifacts related to head movement from the EEG." IEEE Transactions on neural systems and rehabilitation engineering 21.3 (2013): 427-434.

[53] Mammen, Priyanka Mary. "Federated learning: opportunities and challenges." arXiv preprint arXiv:2101.05428 (2021).

[54] Yang, Qiang, et al. "Federated machine learning: Concept and applications." ACM Transactions on Intelligent Systems and Technology (TIST) 10.2 (2019): 1-19

[55] Kara, Madjid, Olfa Lamouchi, and Amar Ramdane-Cherif. "Software Quality Assessment Algorithm Based on Fuzzy Logic." J. Ubiquitous Syst. Pervasive Networks 8, no. 1 (2017): 1-9.