Journal of
*Sensor and
Actuator Networks*

MDPI

Article

Sensitivity of Machine Learning Approaches to Fake and Untrusted Data in Healthcare Domain

Fiammetta Marulli *, Stefano Marrone and Laura Verde

Dipartimento di Matematica e Fisica, Universitdella Campania "Luigi Vanvitelli", 81100 Caserta, Italy; stefano.marrone@unicampania.it (S.M.); laura.verde@unicampania.it (L.V.)
* Correspondence: fiammetta.marulli@unicampania.it

Abstract: Machine Learning models are susceptible to attacks, such as noise, privacy invasion, replay, false data injection, and evasion attacks, which affect their reliability and trustworthiness. Evasion attacks, performed to probe and identify potential ML-trained models' vulnerabilities, and poisoning attacks, performed to obtain skewed models whose behavior could be driven when specific inputs are submitted, represent a severe and open issue to face in order to assure security and reliability to critical domains and systems that rely on ML-based or other AI solutions, such as healthcare and justice, for example. In this study, we aimed to perform a comprehensive analysis of the sensitivity of Artificial Intelligence approaches to corrupted data in order to evaluate their reliability and resilience. These systems need to be able to understand what is wrong, figure out how to overcome the resulting problems, and then leverage what they have learned to overcome those challenges and improve their robustness. The main research goal pursued was the evaluation of the sensitivity and responsiveness of Artificial Intelligence algorithms to poisoned signals by comparing several models solicited with both trusted and corrupted data. A case study from the healthcare domain was provided to support the pursued analyses. The results achieved with the experimental campaign were evaluated in terms of accuracy, specificity, sensitivity, F1-score, and ROC area.

Keywords: data poisoning attacks; fake and untrusted data; healthcare monitoring; resilient artificial intelligence

check for updates

Citation: Marulli, F.; Marrone, S.; Verde, L. Sensitivity of Machine Learning Approaches to Fake and Untrusted Data in Healthcare Domain. *J. Sens. Actuator Netw.* **2022**, *11*, 21. https://doi.org/10.3390/jsan11020021

Academic Editors: Antonio Coronato and Giovanni Paragliola

Received: 31 January 2022
Accepted: 28 March 2022
Published: 30 March 2022

1. Introduction

In recent years, the emergence of intelligent systems enhanced by Artificial Intelligence (AI) techniques has led to several benefits in different aspects of human life. These systems can play a vital role in finance, marketing, data analysis, healthcare, and much more. In fact, AI systems predict what a user is typing into a search engine by speeding up their search. They serve personalized ads based on previous purchases and browsing history. In addition, advanced AI systems can especially help medical experts diagnose and monitor patients' illnesses [1–3].

Unfortunately, AI algorithms are vulnerable to several attacks: (1) poisoning of training data can decrease model accuracy or lead to specific errors; (2) a carefully designed disturbance in the test input (adversarial examples) can cause the model to fail in correctly predicting the obtained result; (3) model inversion attacks and membership inference attacks can recover sensitive training data or steal model parameters; (4) a well-designed backdoor in training data can set dangerous consequences for a system [4].

These security threats can lead to serious consequences for intelligent systems, especially in data security-critical applications such as those belonging to the healthcare sector. Any breach of patient data security could be very critical. Security attacks can be performed in several scenarios. Intelligent network systems present, certainly, high security vulnerability. A common practice for these systems is the data poisoning, in which the attacker manipulates the observations (e.g., sensory data) in the wireless sensor networking system.

Fake data were used in training [5] or during the testing [6] phase of the model, corrupting all systems. In addition to compromising the performance of the algorithm through misdiagnosis, this can lead users to distrust the AI algorithm, increasing the distrust of experts to use such AI techniques as a valid support to diagnoses. The resulting erroneous conclusions can lead to a serious negative impact on people, institutions, and healthcare services. The propagation of fake news can influence public opinion and become dangerous in some cases [7].

Resilience can be defined as the ability of the system to continue the required function under adverse operating conditions. A resilient system first detects an adverse condition (e.g., data poisoning) that causes it to perform poorly; then, it recovers the required function of the service by mitigating the impact of the adverse condition. In this study, we evaluated the resilience and reliability of several AI approaches to classify correctly appropriate samples in the presence of poisoned data.

The remaining sections of the paper are organized as follows. Section 2 identifies the motivation of this study and some aspects of data poisoning and its effects on AI, while Section 3 discusses related defense mechanisms against security attacks. The use-case is described in Section 4, while the conclusions are presented in Section 5.

2. Motivation and Background

Recently, Machine Learning (ML) has become of paramount importance in several domains, including healthcare, where the increasing availability of solutions able to support the early detection of diseases and monitoring of patient vital signs has contributed to increase interesting from health research and the industry. Unfortunately, ML can be subject to numerous attacks, and, in security-sensitive applications, the success of Machine Learning depends on a thorough checking of their resistance and resilience to adversarial data.

Organizations should ensure the resilience of such systems, just as they would for any other critical asset. However, the "black box" approach typical for ML and, more generally, for AI, may make assessing and ensuring resilience different when compared to traditional IT systems. This section provides an overview of the emerging field of resilient ML research, along with a brief literature review on the main security and reliability issues currently affecting ML-based systems, and, more precisely, adversarial attacks to ML solutions and strategies.

2.1. Adversarial Attacks to ML

The robustness and resilience to attacks of AI-based systems [8] have received increasing attention, following the evidence of vulnerabilities exhibited by Deep Neural Networks (DNNs) and, more generally, by ML-based systems, to small perturbations.

Adversarial Machine Learning (AML), which achieved great popularity thanks to the work of [9], has become more and more subtle and specific in addressing its attacks to precise targets and by adopting specific strategies.

ML system security threats can be classified into three dimensions: attack influence, security violations, and attack specificity. Influence attacks can be of two types: causative, in which one seeks to gain control over the data; or exploratory, in which ML model misclassification is used for the attack without intervening in model training. Instead, security violations relate to the availability and integrity of services. They can be categorized into three types: integrity attack, in which an attempt is made to alter the performance of the classifier by increasing the number of false negatives; availability attack, in which the number of false positives is increased in order to alter the performance of the classifier; and privacy violation attack, in which sensitive and confidential information such as training data is violated. Finally, attack specificity can be defined as a targeted attack, where the attack is targeted at a specific input sample or a group of samples, or an indiscriminate attack that causes the failure of the ML model indiscriminately [10].

To identify vulnerabilities in ML models, adversarial attacks have been devised. The generation of contradictory examples through the addition of small, carefully crafted perturbations in ML model samples to attack their integrity is the main objective of an adversarial attack. Two types of adversarial attacks are mainly used: poisoning and evasion attacks. The first includes attacks that affect the model training, manipulating the training data, for example, to alter the performance of the ML trained model. Instead, evasion attacks act on the inference phase of the training process. In these cases, test data are manipulated to compromise the reliability of ML models. In healthcare applications, Poisoning Attacks (PA) are very critical because the manipulation of training data and its detection can be very difficult [11].

To drive both theoretical and methodological research on adversarial attacks and potential countermeasures to design more robust and reliable ML-based systems, the work of [12] provides a significant systematic literature review, in which the concepts, types, harms, and ongoing trends of adversarial attacks along with several common defense mechanisms for ML, are discussed.

2.2. Evasion Attacks to ML

The widespread usage of ML in several application domains has boosted the focus of research toward adversarial threats to these models, and, consequently, toward dependable and secure ML systems. ML models have shown vulnerabilities not only to training-time poisoning and evasion attacks but also to model inference attacks [13]. During an evasion attack, an adversary may attempt to evade a deployed system at test time by carefully crafting a legitimate input to obtain an adversarial sample that cheats the target model to lead an incorrect prediction and affect the model performances.

In the work of [14], a gradient-based approach is presented, aiming to systematically assess the security of the most widely used classification algorithms against evasion attacks. This work was based on a framework for security evaluation, where attack scenarios, with different risk levels, are simulated in order to break a classifier by increasing the attacker's knowledge of the system and its ability to craft attack samples. This experiment provides useful insights to design and tune more robust and reliable classification systems under evasion attack conditions.

2.3. Poisoning Attacks to ML

PA to machine learning [15] are an old security problem that is currently making a comeback. There are two main types of PA, which are distinguished according to its target: (1) those targeting ML availability and (2) those targeting ML integrity (better known as "backdoor attacks). PA have been studied in several application fields, e.g., malware detection, sentiment analysis, intrusion detection, and social media chatbots. There are some key dimensions to consider in PA scenarios that identify the capability of attackers to penetrate and manipulate a victim system. Likely information access, adversarial access can be grouped in levels, as follows: logic corruption, data manipulation, data injection, and transfer learning.

PA [16,17] have found fertile ground when targeted against ML and Deep Learning (DL) systems. A data poisoningattack [18,19] occurs when an attacker injects bad data into a model's training data set, thus leading to a decline of the overall ML model, which produces erroneous results [20,21].

Fake data and, more generally, false and misleading information are artificially crafted, pursuing the goal of deceiving users typically by creating streams of fake data and opinions to influence an idea upon a specific subject, thus impairing the platform's integrity. These Fake Data Checker and Detection Support Systems are suitable candidates targeted by adversarial attacks, since several defense systems rules are based on automatic learning of behaviors and classifications provided by ML or DL systems. In this way, when Fake News from a source are detected as reliable ones, according to a wrong prediction of a classifier, this situation would lead to learn different rules by permissions or defense systems.

The uncontrolled spread of Fake News has generally negative effects on the general well-being of the society, but these effects can develop in catastrophic and severe consequences when this misleading information deals with health, medicine, and critical fields. Tangible evidence of the Fake News effects has been provided mostly during the starting phase of the recent COVID-19 pandemic, when several *infodemics* surrounding the coronavirus made it difficult for people to find reliable guidance and take the appropriate measures without resorting to panic or falling into complacency. The uncertainty and fears caused by the lack of knowledge of the disease, as well as by the measures of social distancing used to contain it, mean that this threat to public health is potentially even higher.

Data poisoning is particularly dangerous in healthcare and medical applications. In fact, such attacks can not only lead to incorrect analyses and consequent misdiagnoses with possible fatal repercussions on patients' lives, just as a false positive classification can cause unnecessary concern, but also reduce the confidence of medical experts and patients in the proposed AI system. Therefore, it is opportune to adopt reliable countermeasures to reduce the effects of these attacks [22–24]. For example, several studies report adversarial attacks to the ML model in medical image processing to alter results by adding noise [25,26].

3. Related Works

Intelligent smart systems are integrating into healthcare, reshaping the healthcare industry with multiple health monitoring devices and applications. Solutions such as smart sensors, wearable devices, and health monitoring systems are playing significant roles in the development of healthcare systems (smart hospitals, mobile healthcare) and the healthcare industry. Resilience is the ability of such systems to continue their normal operations and respond effectively in the event of any unexpected or unforeseen situation. The importance of resilience is multiplied in sensitive systems such as healthcare where systems continuing their normal operations and responding effectively in the event of any unexpected or unforeseen situation is desirable from diagnosis to treatment and the care process.

A lightweight security scheme for ensuring both information confidentiality and transmission resiliency in the Internet-of-Things (IoT) communication was proposed in [27]. A single antenna transmitter communicates with a half-duplex single-antenna receiver in the presence of a sophisticated multiple antenna-aided passive eavesdropper and a multiple-antenna-assisted hostile jammer (HJ). A Blockchain-based Authentication and Key Management Scheme for Internet of Medical Things (BAKMP-IoMT) was described in [28]. This provides a secure key management between cloud servers and personal servers as well as between personal servers and implantable medical devices. The authors demonstrate the resilience of the proposed system against several known active/passive attacks. Instead, Strielkina et al. [29] proposed a Markov Queuing approach for taking into account the security and safety issues of healthcare IoT infrastructure. Meanwhile, in [30], a preliminary study aiming to investigate the improvement of reliability in the Machine Learning-based classification by extending Random Forests with Bayesian Network models was performed.

Furthermore, from the literature, there are several studies describing the possibility of defending against adversarial attacks by exploiting generative models. For example, a framework named Defense Generative Adversarial Networks (GAN) was presented in [31]. This approach was trained on the distribution of unperturbed samples, while during the testing phase, it finds similar output without adversarial perturbations that are given as input to the original DL model. In addition, in [32], GANs models, trained on the same dataset, were used to clean adversarial examples.

GANs, properly trained, are one of the most efficient methods of defending AI approaches from various attacks. Other methods are also described in the literature, such as the augmentation of the training dataset with adversarial examples [9,33]. This approach was particularly effective to defend against adversarial noise. The adversarial examples are generated using an attack model and added to the training dataset. When the attack model used to generate the augmented set is the same used by the attacker, the model reacts with promising robustness to the attack. However, this performance does not occur when the attack pattern is not the same.

4. Use Cases

An evaluation of effects of corrupted data on several ML approaches to corrupted data, in order to assess their resilience and reliability, was performed in this study. In detail, an extension of study conducted in [20] was presented. The resilience and reliability of the most widely used ML approaches in the literature were evaluated in terms of classification accuracy to distinguish healthy and pathological subjects in the presence of data poisoning. Several health time series were analyzed to assess the effects of poisoned signals on these ML approaches and whether these effects depend on the data and/or the noise used to corrupt the data. In particular, two of the main biomedical signals analyzed in the literature were evaluated: voice and Electrocardiographic (ECG) signals. These were processed to estimate the presence, respectively, of psychogenic dysphonia and REM Behavior Disorder (RBD).

Appropriate features were extracted by these signals, as indicated in the following Sections 4.1 and 4.2. Instead, ML approaches evaluated in this study are indicated in Section 4.3, while performance metrics and the obtained results were discussed in Sections 4.4 and 4.5.

4.1. Voice Signals

Voice signals were evaluated to estimate the presence of psychogenic dysphonia. The effects of this disorder usually manifest as voice alterations; a very breathy or whispery, aphonic, or hyperacute voice is typical of psychogenic dysphonia. An universal etiology of psychogenic dysphonia is not accepted. Psychogenic voice disorders are often multifactorial, resulting from psychosocial stress and triggered by conflicting events in the family or at work. Often, people suffering from psychogenic voice disorders are struggling with negative emotions and are reluctant to express them [34].

Voice quality was evaluated by estimating appropriate acoustic features extracted by voice signals. In detail, the sound of vowel /a/ was evaluated. The recordings of 182 subjects were analyzed. Ninety-one voices of patients suffering from psychogenic dysphonia (mean age, 49.6 ± 10.6 years) and 91 ones of healthy subjects (mean age, 33.3 ± 17.3 years) were selected from an available database, the Saarbruecken Voice Database (SVD) [35]. This is a collection of over 2000 recordings of sustained /a/, /i/, and /u/ vowels and a speech sequence, which was recorded at the Caritas Clinic St. Theresia in Saarbruecken by the Institute of Phonetics of the University of Saarland together with the Department of Phoniatrics and Ear, Nose, and Throat (ENT). Only adult voices were considered in this study in order to limit influences and alterations due to the instabilities of younger voices [36,37].

Several acoustic features were considered as inputs of ML approaches. In addition to the features considered in [20], such as Fundamental Frequency (F_0), jitter, shimmer, and Harmonic to Noise Ratio (HNR), Mel-Frequency Cepstral Coefficients (MFCC) and their derivatives were added. These were chosen because they constitute the most relevant features useful to identify specific changes in voice quality, each of them representing a characteristic of the speech production process and therefore able to evaluate the presence of possible alterations [38–40]. F_0, in fact, represents the rate of oscillating of vocal folds useful to evaluate the laryngeal function. Instead, the variations from cycle to cycle of frequency and amplitude of voice signal are estimated through, respectively, jitter and shimmer. Meanwhile, HNR is the ratio of signal information over noise due to the turbulent airflow caused by an incomplete vocal fold closure, which is calculated according to de Krom's algorithm. Finally, MFCC are useful to analyze the vocal tract independently of the vocal folds. F_0, jitter, shimmer, and HNR are estimated by using Praat, which is a software widely used in clinical and research practice to estimate acoustic parameters [41]. Meanwhile, MFCCs and their derivatives are calculated adopting the function *audioFeatureExtractor* by using MATLAB version R2021b [42].

All voice samples were recorded with a sampling frequency of 50 KHz at 16-bit resolution. In order to reproduce the ”poisoned” data, the noise was added to selected recordings. In detail, noise due to the speech bubble of a crowd of people with a Signal to Noise Ratio (SNR) of 5 dB, selected from AURORA database [43], was added to clean signals by using the Audacity tool [44].

4.2. Electrocardiographic Signals

Analysis of ECG signals constitutes the most frequently used technique to diagnose and reveal several pathologies. Cardiovascular diseases are the most commonly diagnosed diseases from the evaluation of these signals, but recently, several studies discussed the possibility of detecting sleep disorders by analyzing ECG signals [45,46]. Since each sleep phase has different cardiac dynamics, it is possible to assess Heart Rate Variability (HRV) to evaluate the presence of sleep disorders [47,48]. Although the standard method for diagnosing sleep disorders is Polysomnography (PSG), analysis of ECG signals is preferable, due to it being less time-consuming and costly, and the acquisition of these signals is less invasive and troublesome for patients [49].

Several sleep disorders can be diagnosed, such as insomnia, Restless Legs Syndrome (RLS), RBD, or Sleep-Disordered Breathing (SDB). In this study, ECG signals were analyzed for RBD detection. Rapid eye movement (REM) sleep behavior disorder is a condition in which abnormal behavior occurs during sleep. Sudden body movements and vocalizations during REM sleep are typical symptoms of this disorder. Samples selected by the Cyclic Alternating Pattern (CAP) database [50], acquired at the Sleep Disorders Center of the Ospedale Maggiore of Parma, Italy, available on Physionet website [51], were processed to evaluate the resilience and reliability of the considered ML algorithms. In detail, ECG samples of 14 healthy subjects (mean age, 32.2 ± 5.5 years) and 22 subjects suffering from RBD disease (mean age, 70.7 ± 6.4 years) were analyzed.

HRV features were extracted from ECG signals, distinguishing between non-linear features and those calculated in the frequency and time domain. In this study, the considered HRV features estimated in time domain are: Mean of RR intervals (Mean RR), Standard deviation of the RR intervals (SDRR), Standard deviation of NN intervals (SDNN), Standard deviation of the successive difference RR intervals (SDSD), and Square root of the mean of the squares of the successive differences between adjacent NNs (RMSSD), where NN intervals refer to the intervals between normal R-peaks. Meanwhile, three spectral components, such as Very Low Frequency (VLF), Low Frequency (LF), and High Frequency (HF) were considered as features estimated in the frequency domain, where the HF band is from 0.15 Hz to 0.40 Hz, LF band is from 0.04 to 0.15 Hz, and the VLF band is from 0 to 0.04 Hz. Finally, it is possible to estimate other HRV parameters using non-linear methods, through, for example, the Ponicaré plot. Poincaré plot analysis is a quantitative analysis that can provide visual information about the behavior of the cardiovascular system; it is a representation of a time series in a phase space, where each RR interval is plotted against the next RR interval. In this study, the standard deviations of the instantaneous beat-to-beat R-R variability and the long term R-R interval variability, respectively, indicated as Standard Deviation1 (SD1) and Standard Deviation2 (SD2), were considered. These HRV parameters were estimated using Kubios HRV [52], which is a tool widely used in research and clinical practice.

All samples were recorded with a sampling frequency of 512 Hz. The “poisoned” data were obtained by adding the baseline wander noise to the selected samples. Baseline wander is a low-frequency noise having non-linear and non-stationary nature, mainly due to movement during breathing, patient movements, and poor contact between the electrode cables due to inadequate preparation of the skin where the electrode is placed and dirty electrodes. This noise was selected from the MIT-BIH Noise Stress Test Database [53] available on the Physionet website [51]. MATLAB, version R2021b [42], was used to "poison" clean ECG signals.

4.3. Machine Learning Techniques

The capability to distinguish healthy and pathological subjects of several ML techniques were evaluated. In addition to the performance of the Random Forest algorithm, evaluated in [20], the reliability and resilience of other used ML approaches were analyzed. The performance of the following ML models was analyzed:

- Support Vector Machine (SVM) [54]: this is one of the most used ML approaches in the literature due to its reliability in classification task and resistence to overfitting. SVM represents a discriminative approach defined by a hyperplane that divides data of different classes.
- BayesNet [55]: this method belongs to the class of Bayesian classifiers, which is a type of probabilistic graphical model where probability computations were performed by using Bayesian inference.
- Random Forest (RF) [56]: this is an ensemble model among methods used in literature classification tasks as an SVM algorithm. A multitude of decision trees was constructed in training time. The class selected by most trees constitutes the output of the model.
- k-Nearest Neighbor (Ibk) [57]: this represents an instance-based learning algorithm where the k-nearest neighbors are evaluated to indicate the class of relevance.
- Adaptive Boosting (Adaboost) [58]: this is an ensemble learning algorithm that uses the boosting method. The algorithm formulates H hypotheses via the ensemble boosting algorithm from a training set of N examples. It assigns a weight z to each hypothesis to measure its effectiveness. Finally, the algorithm formulates the final weighted majority hypothesis.

4.4. Performance Metrics

The resilience of selected ML approaches was evaluated in terms of classification accuracy to distinguish pathological and healthy subjects also in the presence of poisoned data. The accuracy is defined as the number of correct predictions over all the datasets, which was evaluated according to the following equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where the True Positive (TP) and True Negative (TN) are the samples correctly classified, respectively, as pathological and healthy. Meanwhile, the False Positive (FP) and False Negative (FN) represent the samples incorrectly classified, respectively, as pathological and healthy.

Additionally, the performance of the considered ML algorithms was assessed in terms of sensitivity, specificity, F1-score, precision, and Receiver Operating Characteristic (ROC) area. The sensitivity and specificity represent the percentage of, respectively, pathological or healthy cases correctly predicted over all the pathological or healthy cases in the dataset. These were estimated as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

Defining the precision as the number of pathological predictions correctly classified, this was calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

The F1-score is the harmonic mean of the sensitivity and precision, and it was estimated according to the following equation:

$$F1 - score = 2 * \frac{precision * sensitivity}{precision + sensitivity} \quad (5)$$

Finally, the area under the ROC curve was estimated. When this is minimum, the algorithm classified all samples incorrectly, while when this area is maximum, all samples were classified correctly.

For both analyses, on voice and ECG signals, three scenarios were observed: tests were performed only on clean data (scenario 1), on poisoned data (scenario 2), and, finally, each ML model was trained with poisoned data and tested on clean ones (scenario 3).

4.5. Results and Discussion

Tables 1–6 show the results obtained evaluating the performance of several ML classifiers in the absence or presence of poisoned data for two different biomedical signals, voice and ECG, in three scenarios described previously. The results obtained by analyzing voice signals were reported in Tables 1–3, while those achieved evaluating ECG samples were shown in Tables 4–6. In both cases, for scenarios 2 and 3, in addition to the percentages obtained for performance metrics, the difference between the values obtained for each metric between the observed scenario (2 or 3) and scenario 1 is reported, which are shown respectively as Δ_{2-1} and Δ_{3-1}.

Table 1. Results obtained considering clean **voice** signals in the training and testing sets (*scenario 1*).

Classifier	Set	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)	F1-Score (%)	AUC
SVM [54]	*Training*	95.89	100.00	97.95	100.00	97.90	0.979
	Testing	72.22	77.78	75.00	76.47	74.29	0.750
BayesNet [55]	*Training*	84.93	78.08	81.51	79.49	82.12	0.911
	Testing	83.33	72.22	77.78	75.00	78.95	0.858
RF [56]	*Training*	100.00	100.00	100.00	100.00	100.00	1.000
	Testing	83.33	66.67	75.00	71.43	76.92	0.855
Ibk [57]	*Training*	100.00	100.00	100.00	100.00	100.00	1.000
	Testing	50.00	83.33	66.67	75.00	60.00	0.667
Adaboost [58]	*Training*	95.89	80.82	88.36	83.33	89.17	0.966
	Testing	88.89	55.56	72.22	66.67	76.19	0.809

Table 2. Results obtained considering poisoned **voice** signals in the training and testing sets (*scenario 2*).

Classifier	Set	Sensitivity (%)		Specificity (%)		Accuracy (%)		Precision (%)		F1-Score (%)		AUC	
		2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}		Δ_{2-1}
SVM [54]	*Training*	100.00	+4.11	100.00	0	100.00	+2.05	100.00	0	100.00	+2.1	1.000	−0.02
	Testing	66.67	−5.55	55.56	−22.23	61.11	−13.89	60.00	−16.47	63.16	−11.13	0.611	−0.139
BayesNet [55]	*Training*	87.67	+2.74	73.97	−4.11	80.82	−0.68	77.11	−2.38	82.05	−0.07	0.904	−0.007
	Testing	72.22	−11.12	66.67	−5.56	69.44	−8.33	68.42	−6.58	70.27	−8.68	0.824	−0.034
RF [56]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	72.22	−11.12	77.78	+11.12	75.00	0	76.47	+5.04	74.29	−2.64	0.795	−0.059
Ibk [57]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	44.44	−5.56	72.22	−11.12	58.33	−8.33	61.54	−13.46	51.61	−8.39	0.583	−0.084
Adaboost [58]	*Training*	90.41	−5.48	97.26	+16.44	93.84	+5.48	97.06	+13.72	93.62	+4.44	0.993	−0.03
	Testing	66.67	−22.23	72.22	+16.67	69.44	−2.78	70.59	+3.92	68.57	−7.62	0.762	−0.047

Observing the results achieved for voice and ECG signals, it is possible to note a decrease in the performance of the different algorithms in terms of correct classification between healthy and pathological subjects between scenarios 1 and 2, i.e., when clean and poisoned data are evaluated. The presence of "noise" inevitably has an impact on the reliability of the classifiers. This is particularly evident observing the values of accuracy achieved into two scenarios (1 and 2). A decrease of accuracy to distinguish healthy and pathological samples when poisoned data are processed is observed. This decrease is also achieved in the case where malicious data corrupted the training model, while the testing is performed on clean data (scenario 3). The poisoned training models of the considered ML algorithms achieve lower accuracy values than other scenarios.

Table 3. Results obtained considering poisoned **voice** signals in the training set and clean signals in the testing set (*scenario 3*).

Classifier	Set	Sensitivity (%)		Specificity (%)		Accuracy (%)		Precision (%)		F1-Score (%)		AUC	
		3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}		Δ_{3-1}
SVM [54]	*Training*	100.00	+4.11	100.00	0	100.00	+2.05	100.00	0	100.00	+2.10	1.000	+0.021
	Testing	72.22	0	61.11	−16.67	66.67	−8.34	65.00	−11.47	68.42	−5.86	0.667	−0.083
BayesNet [55]	*Training*	87.67	+2.74	73.97	−4.11	80.82	−0.68	77.11	−2.38	82.05	−0.07	0.904	−0.007
	Testing	22.22	−61.12	100.00	+27.78	61.11	−16.67	100.00	+25	36.36	−42.58	0.759	−0.099
RF [56]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	38.89	−44.45	94.44	+27.78	66.67	−8.33	87.50	+16.07	53.85	−23.08	0.792	−0.063
Ibk [57]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	22.22	−27.78	83.33	0	52.78	−13.89	57.14	−17.86	32.00	−28	0.528	−0.139
Adaboost [58]	*Training*	90.41	−5.48	97.26	+16.44	93.84	5.48	97.06	+13.72	93.62	4.44	0.993	+0.027
	Testing	44.44	−44.45	83.33	+27.78	63.89	−8.34	72.73	+6.06	55.17	−21.02	0.765	−0.044

Table 4. Results obtained considering clean **ECG** signals in the training and testing sets (*scenario 1*).

Classifier	Set	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)	F1-Score (%)	AUC
SVM [54]	*Training*	100.00	45.45	78.57	73.91	85.00	0.798
	Testing	100.00	33.33	75.00	71.43	83.33	0.667
BayesNet [55]	*Training*	88.24	90.91	89.29	93.75	90.91	0.960
	Testing	80.00	100.00	87.50	100.00	88.89	0.967
RF [56]	*Training*	100.00	100.00	100.00	100.00	100.00	1.000
	Testing	100.00	66.67	87.50	83.33	90.91	1.000
Ibk [57]	*Training*	100.00	100.00	100.00	100.00	100.00	1.000
	Testing	80.00	66.67	75.00	80.00	80.00	0.733
Adaboost [58]	*Training*	100.00	90.91	96.43	94.44	97.14	1.000
	Testing	80.00	66.67	75.00	80.00	80.00	0.933

Table 5. Results obtained considering poisoned **ECG** signals in the training and testing sets (*scenario 2*).

Classifier	Set	Sensitivity (%)		Specificity (%)		Accuracy (%)		Precision (%)		F1-Score (%)		AUC	
		2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}	2	Δ_{2-1}		Δ_{2-1}
SVM [54]	*Training*	94.12	−5.88	45.45	0	75.00	−3.57	72.73	−1.19	82.05	−2.95	0.698	−0.100
	Testing	80.00	−20.00	33.33	0	62.50	−12.50	66.67	−4.76	72.73	−10.61	0.667	0
BayesNet [55]	*Training*	82.35	−5.88	81.82	−9.09	82.14	−7.14	87.50	−6.25	84.85	−6.06	0.941	−0.019
	Testing	80.00	0	66.67	−33.33	75.00	−12.50	80.00	−20.00	80.00	−8.89	0.833	−0.134
RF [56]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	100.00	0	66.67	0	87.50	0	83.33	0	90.91	0	0.933	−0.067
Ibk [57]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	80.00	0	66.67	0	75.00	0	80.00	0	80.00	0	0.733	0
Adaboost [58]	*Training*	100.00	0	100.00	+9.09	100.00	+3.57	100.00	+5.55	100.00	+2.85	1.000	0
	Testing	60.00	−20.00	33.33	−33.33	50.00	−25.00	60.00	−20.00	60.00	−20.00	0.558	−0.375

Concerning scenario 2, RF achieved the best performance compared to other ML techniques. In this scenario, both for voice and ECG signals, the best classification accuracy was, in fact, achieved by RF, recording not very high differences between the scenario in which data are clean and when they are poisoned. Additionally, observing other performance metrics, it is possible to note that not only the sensitivity is promising but also the obtained specificity is higher than other algorithms. This means that RF is able to distinguish accurately healthy subjects despite the poisoned data, both by evaluating voice and ECG signals. In particular, considering ECG signals, also, the Ibk algorithm achieved good performance in terms of resilience, obtaining minimal differences between scenarios 1 and 2.

Table 6. Results obtained considering poisoned **ECG** signals in the training set and clean signals in the testing set *(scenario 3)*.

Classifier	Set	Sensitivity (%)		Specificity (%)		Accuracy (%)		Precision (%)		F1-score (%)		AUC	
		3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}	3	Δ_{3-1}		Δ_{3-1}
SVM [54]	*Training*	94.12	−5.88	45.45	0	75.00	−3.57	72.73	−1.19	82.05	−2.95	0.698	−0.100
	Testing	80.00	−20.00	33.33	0	62.50	−12.50	66.67	−4.76	83.33	0	0.627	−0.040
BayesNet [55]	*Training*	82.35	−5.88	81.82	−9.09	82.14	−7.14	87.50	−6.25	84.85	−6.06	0.941	−0.019
	Testing	60.00	−20.00	66.67	−33.33	62.50	−25.00	75.00	−25.00	66.67	−22.22	0.800	−0.167
RF [56]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	80.00	−20.00	66.67	0	75.00	−12.50	80.00	−3.33	80.00	−10.91	0.933	−0.067
Ibk [57]	*Training*	100.00	0	100.00	0	100.00	0	100.00	0	100.00	0	1.000	0
	Testing	40.00	−40.00	66.67	0	50.00	−25.00	66.67	−13.33	50.00	−30.00	0.533	−0.200
Adaboost [58]	*Training*	100.00	0	100.00	9.09	100.00	3.57	100.00	5.56	100.00	2.86	1.000	0
	Testing	40.00	−40.00	100.00	+33.33	62.50	−12.50	100.00	+20.00	57.14	−22.86	0.576	−0.357

However, considering the results obtained in the presence of poisoned data, RF obtains the most resilient and reliable algorithms compared with the other algorithms.

Concerning scenario 3, SVM, instead, obtained the best performance. In the voice signals evaluation, although SVM and RF algorithms achieved the same accuracy value to discriminate healthy and pathological samples (about 67%), the sensitivity and specificity values obtained by SVM are better than those achieved by RF. Therefore, SVM distinguishes both healthy and pathological samples, unlike RF, which cannot discriminate well between pathological voices as demonstrated by the low sensitivity value. Observing the results obtained for theECG signals, the classification accuracy achieved by the algorithms SVM and RF, as well as the sensitivity and specificity, there were no differences between the case when the algorithms were trained using clean data (scenario 1) or using poisoned data (scenario 3).

The results obtained in the different scenarios observed, considering different types of health data, demonstrate that RF is the most resilient and reliable technique able to classify with higher accuracy the presence of affected subjects despite corrupted data among the analyzed ML techniques.

5. Conclusions

Nowadays, the continuous use of intelligent and smart sensors and devices contributes to generate a large amount of data in various fields, such as healthcare. AI provides valuable tools for processing and analyzing this wide variety of data, which is very often useful for supporting the correct diagnosis and treatment of people's health status. Unfortunately, while the wide availability of data allows for more robust and reliable analysis, the continuous transfer of sensitive data could be subject to malicious attacks that could affect the performance of AI algorithms.

In this study, we investigated the resilience and reliability of different ML approaches when the analyzed data are poisoned. In detail, the reliability of these techniques in correctly distinguishing healthy and pathological subjects was evaluated. The presence of this disorder has been estimated by evaluating vocal signals, from which appropriate acoustic features have been extracted, used as input of the ML algorithms. The results obtained show that the best performance, in the presence of data poisoning, was achieved by the RF model compared to other algorithms.

In future plans, we will explore efficient defense methods to safeguard the security of AI models, which can reduce the sensitivity of these approaches to malicious attacks and improve their ability to correctly classify samples. Additionally, we will extend the original dataset by exploiting a data augmentation strategy, based on harnessing a generative neural network to obtain coherent artificial samples.

Author Contributions: F.M., S.M. and L.V. have contributed to the conception and the design of this study as well as carrying out the analysis. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results was funded by the Research Program "Vanvitelli V:ALERE 2020 WAIILD TROLS", financed by the University of Campania "Luigi Vanvitelli", Italy.

Data Availability Statement: All data considered in this study were selected from the Saarbruecken Voice Database (SVD) available online on http://www.stimmdatenbank.coli.uni-saarland.de/help_en.php4 (accessed on 15 December 2021) and CAP Sleep Database available online on https://physionet.org/content/capslpdb/1.0.0/. (accessed on 22 March 2021).

Acknowledgments: The research described in this work is funded and realized within the activities of the the Research Program "Vanvitelli V:ALERE 2020—WAIILD TROLS", financed by the University of Campania "L. Vanvitelli", Italy.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ML	Machine Learning
DNNs	Deep Neural Networks
AML	Adversarial Machine Learning
PA	Poisoning Attacks
DL	Deep Learning
IoT	Internet-of-Things
BAKMP-IoMT	Blockchain-based Authentication and Key Management Scheme for Internet of Medical Things
GAN	Generative Adversarial Networks
SVD	Saarbruecken Voice Database
F_0	Fundamental Frequency
HNR	Harmonic to Noise Ratio
MFCC	Mel-Frequency Cepstral Coefficients
SNR	Signal to Noise Ratio
SVM	Support Vector Machine
RF	Random Forest
ibk	k-Nearest Neighbor
Adaboost	Adaptive Boosting
ROC	Receiver Operating Characteristic
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
PSG	Polysomnography
HRV	Heart Rate Variability
SDB	Sleep-Disordered Breathing
RBD	REM Behavior Disorder
RLS	Restless Legs Syndrome
REM	Rapid eye movement
CAP	Cyclic Alternating Pattern
Mean RR	Mean of RR intervals
SDRR	Standard deviation of the RR intervals
SDNN	Standard deviation of NN intervals

SDSD	Standard deviation of the successive difference RR intervals
RMSSD	Square root of the mean of the squares of the successive differences between adjacent NNs
VLF	Very Low Frequency
LF	Low Frequency
HF	High Frequency
SD1	Standard Deviation1
SD2	Standard Deviation2

References

1. Verde, L.; De Pietro, G. A neural network approach to classify carotid disorders from heart rate variability analysis. *Comput. Biol. Med.* **2019**, *109*, 226–234. [CrossRef] [PubMed]
2. Agliari, E.; Barra, A.; Barra, O.A.; Fachechi, A.; Vento, L.F.; Moretti, L. Detecting cardiac pathologies via machine learning on heart-rate variability time series and related markers. *Sci. Rep.* **2020**, *10*, 8845. [CrossRef] [PubMed]
3. Paragliola, G.; Coronato, A. An hybrid ECG-based deep network for the early identification of high-risk to major cardiovascular events for hypertension patients. *J. Biomed. Inform.* **2021**, *113*, 103648. [CrossRef]
4. Xue, M.; Yuan, C.; Wu, H.; Zhang, Y.; Liu, W. Machine learning security: Threats, countermeasures, and evaluations. *IEEE Access* **2020**, *8*, 74720–74742. [CrossRef]
5. Wang, C.; Chen, J.; Yang, Y.; Ma, X.; Liu, J. Poisoning attacks and countermeasures in intelligent networks: Status quo and prospects. *Digit. Commun. Netw.* **2021**. [CrossRef]
6. Newaz, A.I.; Haque, N.I.; Sikder, A.K.; Rahman, M.A.; Uluagac, A.S. Adversarial attacks to machine learning-based smart healthcare systems. In Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6.
7. Lara-Navarra, P.; Falciani, H.; Sánchez-Pérez, E.A.; Ferrer-Sapena, A. Information management in healthcare and environment: Towards an automatic system for fake news detection. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1066. [CrossRef]
8. Eigner, O.; Eresheim, S.; Kieseberg, P.; Klausner, L.D.; Pirker, M.; Priebe, T.; Tjoa, S.; Marulli, F.; Mercaldo, F. Towards Resilient Artificial Intelligence: Survey and Research Issues. In Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Rhodes, Greece, 26–28 July 2021; pp. 536–542.
9. Goodfellow, I.J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. *arXiv* **2014**, arXiv:1412.6572.
10. Fredrikson, M.; Jha, S.; Ristenpart, T. Model inversion attacks that exploit confidence information and basic countermeasures. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015; pp. 1322–1333.
11. Qayyum, A.; Qadir, J.; Bilal, M.; Al-Fuqaha, A. Secure and robust machine learning for healthcare: A survey. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 156–180. [CrossRef]
12. Kong, Z.; Xue, J.; Wang, Y.; Huang, L.; Niu, Z.; Li, F. A Survey on Adversarial Attack in the Age of Artificial Intelligence. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 4907754. [CrossRef]
13. Amich, A.; Eshete, B. Explanation-Guided Diagnosis of Machine Learning Evasion Attacks. In *International Conference on Security and Privacy in Communication Systems*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 207–228.
14. Biggio, B.; Corona, I.; Maiorca, D.; Nelson, B.; Šrndić, N.; Laskov, P.; Giacinto, G.; Roli, F. Evasion Attacks against Machine Learning at Test Time. In *Machine Learning and Knowledge Discovery in Databases*; Blockeel, H., Kersting, K., Nijssen, S., Železný, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 387–402.
15. Jagielski, M.; Oprea, A.; Biggio, B.; Liu, C.; Nita-Rotaru, C.; Li, B. Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning. In Proceedings of the 2018 IEEE Symposium on Security and Privacy, SP 2018, San Francisco, CA, USA, 20–24 May 2018; IEEE: Washington, DC, USA, 2018; pp. 19–35. [CrossRef]
16. Muñoz-González, L.; Biggio, B.; Demontis, A.; Paudice, A.; Wongrassamee, V.; Lupu, E.C.; Roli, F. Towards poisoning of deep learning algorithms with back-gradient optimization. In Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security, Dallas, TX, USA, 3 November 2017; pp. 27–38.
17. Marulli, F.; Visaggio, C.A. Adversarial Deep Learning for Energy Management in Buildings. In Proceedings of the SummerSim '19: Proceedings of the 2019 Summer Simulation Conference, Berlin, Germany, 22–24 July 2019; pp. 50–51.
18. Ahmed, I.M.; Kashmoola, M.Y. Threats on Machine Learning Technique by Data Poisoning Attack: A Survey. In *International Conference on Advances in Cyber Security*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 586–600.
19. Steinhardt, J.; Koh, P.W.W.; Liang, P.S. Certified defenses for data poisoning attacks. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3520–3532 .
20. Verde, L.; Marulli, F.; Marrone, S. Exploring the Impact of Data Poisoning Attacks on Machine Learning Model Reliability. *Procedia Comput. Sci.* **2021**, *192*, 2624–2632. [CrossRef]
21. Marulli, F.; Verde, L.; Campanile, L. Exploring Data and Model Poisoning Attacks to Deep Learning-Based NLP Systems. *Procedia Comput. Sci.* **2021**, *192*, 3570–3579. [CrossRef]
22. Rahman, A.; Hossain, M.S.; Alrajeh, N.A.; Alsolami, F. Adversarial examples—Security threats to COVID-19 deep learning systems in medical IoT devices. *IEEE Internet Things J.* **2020**, *8*, 9603–9610. [CrossRef]

23. Ahmed, A.; Latif, R.; Latif, S.; Abbas, H.; Khan, F.A. Malicious insiders attack in IoT based multi-cloud e-healthcare environment: A systematic literature review. *Multimed. Tools Appl.* **2018**, *77*, 21947–21965. [CrossRef]
24. Mozaffari-Kermani, M.; Sur-Kolay, S.; Raghunathan, A.; Jha, N.K. Systematic poisoning attacks on and defenses for machine learning in healthcare. *IEEE J. Biomed. Health Inform.* **2014**, *19*, 1893–1905. [CrossRef] [PubMed]
25. Finlayson, S.G.; Bowers, J.D.; Ito, J.; Zittrain, J.L.; Beam, A.L.; Kohane, I.S. Adversarial attacks on medical machine learning. *Science* **2019**, *363*, 1287–1289. [CrossRef]
26. Newaz, A.I.; Sikder, A.K.; Rahman, M.A.; Uluagac, A.S. A survey on security and privacy issues in modern healthcare systems: Attacks and defenses. *ACM Trans. Comput. Healthc.* **2021**, *2*, 1–44. [CrossRef]
27. Letafati, M.; Kuhestani, A.; Wong, K.K.; Piran, M.J. A lightweight secure and resilient transmission scheme for the Internet of Things in the presence of a hostile jammer. *IEEE Internet Things J.* **2020**, *8*, 4373–4388. [CrossRef]
28. Garg, N.; Wazid, M.; Das, A.K.; Singh, D.P.; Rodrigues, J.J.; Park, Y. BAKMP-IoMT: Design of blockchain enabled authenticated key management protocol for internet of medical things deployment. *IEEE Access* **2020**, *8*, 95956–95977. [CrossRef]
29. Strielkina, A.; Kharchenko, V.; Uzun, D. Availability models for healthcare IoT systems: Classification and research considering attacks on vulnerabilities. In Proceedings of the 2018 IEEE 9th international conference on dependable systems, services and technologies (DES;SERT), Kyiv, Ukraine, 24–27 May 2018; pp. 58–62.
30. de Biase, M.S.; Marulli, F.; Verde, L.; Marrone, S. Improving Classification Trustworthiness in Random Forests. In Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Rhodes, Greece, 26–28 July 2021; pp. 563–568.
31. Samangouei, P.; Kabkab, M.; Chellappa, R. Defense-gan: Protecting classifiers against adversarial attacks using generative models. *arXiv* **2018**, arXiv:1805.06605.
32. Santhanam, G.K.; Grnarova, P. Defending against adversarial attacks by leveraging an entire GAN. *arXiv* **2018**, arXiv:1805.10652.
33. Moosavi-Dezfooli, S.M.; Fawzi, A.; Frossard, P. Deepfool: A simple and accurate method to fool deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2574–2582.
34. Rosen, D.C.; Shmidheiser, M.H.; Sataloff, J.B.; Hoffmeister, J.; Sataloff, R.T. Psychogenic Dysphonia. *Psychol. Voice Disord.* **2020**, 187.
35. Pützer, M.; Koreman, J. A German database of patterns of pathological vocal fold vibration. *Phonus* **1997**, *3*, 143–153.
36. Sataloff, R.T.; Linville, S. *The Effect of Age on the Voice*; Plural Publishing: San Diego, CA, USA, 2005 .
37. Latoszek, B.B.v.; Ulozaitė-Stanienė, N.; Maryn, Y.; Petrauskas, T.; Uloza, V. The influence of gender and age on the acoustic voice quality index and dysphonia severity index: A normative study. *J. Voice* **2019**, *33*, 340–345. [CrossRef] [PubMed]
38. Teixeira, J.P.; Fernandes, P.O. Acoustic analysis of vocal dysphonia. *Procedia Comput. Sci.* **2015**, *64*, 466–473. [CrossRef]
39. Kosztyła-Hojna, B.; Moskal, D.; Łobaczuk-Sitnik, A.; Kraszewska, A.; Zdrojkowski, M.; Biszewska, J.; Skorupa, M. Psychogenic voice disorders. *Otolaryngol. Pol.* **2018**, *72*, 26–34. [CrossRef]
40. Verde, L.; Raimo, G.; Vitale, F.; Carbonaro, B.; Cordasco, G.; Marrone, S.; Esposito, A. A Lightweight Machine Learning Approach to Detect Depression from Speech Analysis. In Proceedings of the 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), Washington, DC, USA, 1–3 November 2021; pp. 330–335.
41. Boersma, P.; Weenink, D. Praat: Doing Phonetics by Computer (Version 5.1. 05). Computer Program. Retrieved 1 May 2009. Available online: https://www.praat.org/ (accessed on 25 January 2021).
42. Matlab. audioFeatureExtractor Function. 2020. Available online: https://it.mathworks.com/help/audio/ref/audiofeatureextractor.html/ (accessed on 25 January 2021).
43. Hirsch, H.G.; Pearce, D. The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In Proceedings of the ASR2000-Automatic Speech Recognition: Challenges for the New Millenium ISCA Tutorial and Research Workshop (ITRW), Beijing, China, 16–20 October 2000.
44. Audacity. 2021. Available online: https://https://www.audacityteam.org// (accessed on 22 November 2021).
45. Schumann, A.Y.; Bartsch, R.P.; Penzel, T.; Ivanov, P.C.; Kantelhardt, J.W. Aging effects on cardiac and respiratory dynamics in healthy subjects across sleep stages. *Sleep* **2010**, *33*, 943–955. [CrossRef]
46. Widasari, E.R.; Tanno, K.; Tamura, H. Automatic sleep disorders classification using ensemble of bagged tree based on sleep quality features. *Electronics* **2020**, *9*, 512. [CrossRef]
47. Kantelhardt, J.; Havlin, S.; Ivanov, P.C. Modeling transient correlations in heartbeat dynamics during sleep. *EPL Europhys. Lett.* **2003**, *62*, 147. [CrossRef]
48. Penzel, T.; Kantelhardt, J.W.; Bartsch, R.P.; Riedl, M.; Kraemer, J.F.; Wessel, N.; Garcia, C.; Glos, M.; Fietze, I.; Schöbel, C. Modulations of heart rate, ECG, and cardio-respiratory coupling observed in polysomnography. *Front. Physiol.* **2016**, *7*, 460. [CrossRef]
49. Kushida, C.A.; Littner, M.R.; Morgenthaler, T.; Alessi, C.A.; Bailey, D.; Coleman Jr, J.; Friedman, L.; Hirshkowitz, M.; Kapen, S.; Kramer, M.; et al. Practice parameters for the indications for polysomnography and related procedures: An update for 2005. *Sleep* **2005**, *28*, 499–523. [CrossRef] [PubMed]
50. Terzano, M.G.; Parrino, L.; Sherieri, A.; Chervin, R.; Chokroverty, S.; Guilleminault, C.; Hirshkowitz, M.; Mahowald, M.; Moldofsky, H.; Rosa, A.; et al. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. *Sleep Med.* **2001**, *2*, 537–553. [CrossRef]

51. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef] [PubMed]
52. Tarvainen, M.P.; Niskanen, J.P.; Lipponen, J.A.; Ranta-Aho, P.O.; Karjalainen, P.A. Kubios HRV–heart rate variability analysis software. *Comput. Methods Programs Biomed.* **2014**, *113*, 210–220. [CrossRef]
53. Moody, G.B.; Muldrow, W.; Mark, R.G. A noise stress test for arrhythmia detectors. *Comput. Cardiol.* **1984**, *11*, 381–384.
54. Schölkopf, B.; Burges, C.J.; Smola, A.J. Introduction to support vector learning. In *Advances in Kernel Methods: Support Vector Learning*; MIT Press: Cambridge, MA, USA, 1999; pp. 1–15.
55. John, G.H.; Langley, P. Estimating Continuous Distributions in Bayesian Classifiers. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, 18–20 August 1995; Morgan Kaufmann: San Mateo, CA, USA, 1995; pp. 338–345.
56. Venkatesan, N.; Priya, G. A study of random forest algorithm with implementation using WEKA. *Int. J. Innov. Res. Comput. Sci. Eng.* **2015**, *1*, 156–162.
57. Aha, D.W.; Kibler, D.; Albert, M.K. Instance-based learning algorithms. *Mach. Learn.* **1991**, *6*, 37–66. [CrossRef]
58. Dietterich, T.G. Ensemble methods in machine learning. In *International Workshop on Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 1–15.