Health Prediction: A Comprehensive IoT-Driven Health Monitoring System with Machine Learning Analysis and XAI Insights

Marcelo Colares da Silva
Department of Computer Science
Institute Federal of Ceará
Fortaleza, CE 60040-215
marcelo.colares06@aluno.ifce.edu.br

Caio Marques Silva
Institute Federal of Ceará
Fortaleza, CE 60040-215
caio.marques.silva05@aluno.ifce.edu.br

Suane Pires P. da Silva
Institute Federal of Ceará
Fortaleza, CE 60040-215
suanepires@lapisco.ifce.edu.br

Elizângela de Souza Rebouças
Institute Federal of Ceará
Fortaleza, CE 60040-215
elizangela.reboucas@ifce.edu.br

Caio C. H. Cunha
AtilsLab
Fortaleza, CE 60160-230
cchc32@gmail.com

Cilis A. Benevides
AtilsLab
Fortaleza, CE 60160-230
cilis.benevides@gmail.com

Roger M. Sarmento
Institute Federal of Ceará
Fortaleza, CE 60040-215
rogerms@ifce.edu.br

Houbing Hebert Song
University of Maryland,
Baltimore County (UMBC),
songh@umbc.edu

Pedro Pedrosa Rebouças Filho
Institute Federal of Ceará
Fortaleza, CE 60040-215
pedrosarf@ifce.edu.br

Abstract

Stroke causes significant damage due to the interruption or reduction of blood supply to an area of the brain. This condition can result in severe sequelae, including cognitive impairment, paralysis, speech and coordination difficulties, directly affecting the patient's quality of life. The brain damage resulting from a stroke can have irreversible impacts on physical and mental capacity, highlighting the importance of preventive measures, rapid interventions, and rehabilitation to minimize adverse consequences. In this study, we propose a method for monitoring and predicting strokes that integrates Internet of Things (IoT) devices for remote patient monitoring. Recognizing the severity of stroke-associated sequelae, our approach aims to mitigate adverse impacts through preventive measures and timely interventions. Using a combination of machine learning algorithms, including Naive Bayes, Multilayer Perceptron, Support Vector Machine, k-Neighbors, Decision Tree, XGBoost, and Random Forest, we aim to assess the risk of stroke occurrence, with XGBoost standing out with an Accuracy of 98.52% and a testing time of 0.076ms.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

1 Introduction

According to the World Health Organization (WHO) (1), 15 million people worldwide suffer a stroke annually, with 5 million fatalities and another 5 million left permanently disabled. A stroke occurs when the brain's blood supply is interrupted, leading to brain damage due to lack of oxygen and nutrients. Strokes are classified as ischemic, caused by blood vessel blockage, or hemorrhagic, due to brain bleeding (2). Symptoms include weakness, speech difficulties, severe headache, and loss of consciousness.

Risk factors include prior stroke, transient ischemic attack (TIA), heart conditions such as heart failure (3) and atrial fibrillation (4), and being over 55 years old (5). Other risk factors are smoking, high cholesterol, diabetes (6), obesity (7), sedentary lifestyle, excessive alcohol (8), blood clotting disorders, estrogen therapy, and psychoactive substances. Work-related stress, extreme temperatures, long working hours (9), and socioeconomic factors (10) also increase stroke risk.

Given stroke's severe consequences, predicting its occurrence is crucial. Wearable devices integrated with IoT allow for remote monitoring, and IoT health systems improve patient-doctor communication and data visualization (11). This article proposes a novel approach to classify stroke risk using machine learning algorithms in conjunction with wearable devices. To enhance the transparency of the decision-making process, Explainable AI (XAI) techniques are employed, allowing both healthcare professionals and patients to understand the risk predictions. The system comprises a web platform for healthcare providers and a mobile application for patients, facilitating real-time monitoring and alerts.

2 Explained Artificial Intelligence

In this section, we will explore the two explainable artificial intelligence methods used in the proposed approach: SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). These methods will explain which features are most important for the decision-making of the models, as well as identify which attributes contribute positively or negatively to classifying the patient as belonging to the stroke-affected class.

2.1 SHapley Additive exPlanations

The SHapley Additive exPlanations (SHAP) method (12) is a technique used to interpret machine learning models, based on Shapley value theory, which ensures a fair distribution of the importance of each input feature in the model's predictions. In classifications, SHAP provides explanations for the overall importance of features by calculating the marginal contribution of each feature. Compatible with various models, such as decision trees and neural networks, SHAP enhances transparency, helps detect issues, and reduces biases in the model, offering a clear and detailed understanding of predictions. The Shapley values (ϕ_i) for a given attribute i are calculated using the following mathematical expression:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \left(v(S \cup \{i\}) - v(S)\right) \tag{1}$$

where:

- N is the set of all features.
- S is a subset of N that does not include feature i.
- $|S|$ is the number of elements in S.
- $v(S)$ is the value of the value function (in this context, the model prediction) with the subset of features S.
- $v(S \cup \{i\})$ is the value of the value function with the subset S plus feature i.

In the proposed method, SHAP reveals which features were most important for the model's decision-making, providing a detailed and transparent understanding of each feature's contributions to the model predictions. SHAP helps reduce biases in machine learning by explaining the contribution

of each feature to the model's predictions. It reveals biases, such as disproportionate influences of features related to demographic groups, and identifies when the model overly relies on irrelevant information. This enables adjustments to data and models for fairer and more transparent decisions.

2.2 Local Interpretable Model-Agnostic Explanations

Local Interpretable Model-agnostic Explanations (LIME) method explains individual AI model predictions by creating local interpretations around specific points, such as a single prediction. It fits simple models like linear regressions to the data around that point to identify influential features in that context (13). The following steps illustrate its operation:

1. **Perturbed Points:** The first step is to generate perturbed points around the example we want to explain. Let x' be the original example, and x_i a perturbed example. The goal is to generate x_i so that they are similar to x' with small changes.
2. **Importance Weights:** Importance weights are assigned to each perturbed example x_i, indicating its relevance in the explanation. These weights are determined based on the proximity between x_i and the original example x', using a similarity function such as Euclidean distance or Manhattan distance.
3. **Interpretable Model:** An interpretable model, such as a linear regression, is trained using the perturbed examples and the importance weights assigned to them. This model is trained to predict the outputs of the original machine learning model.
4. **Final Explanation:** Finally, the final explanation for the original model prediction is derived from the coefficients of the trained interpretable model. These coefficients indicate the relative contribution of each feature to the original model prediction.

Although LIME offers a flexible approach applicable to various machine learning models, it has limitations that should be considered, such as the potential instability of explanations with small changes in input data and the reliance on the choice of perturbed points. These limitations can affect the consistency of the generated explanations and their representativeness in relation to the model's global behavior (13).

3 Connection with Wearable Devices

In this section, we will explore the operation of the connection with wearable devices within the context of the developed system. A mobile application has been developed, compatible with various smartphones, which is made available to the patients. Data collection is performed through the mobile application, exclusively developed for the patients, which connects to compatible devices such as smartbands, blood pressure monitors, or glucose meters, via Bluetooth. Obtaining information from these devices is carried out through their respective Bluetooth profile implementations (14). Table 1 presents the UUIDs of the services and characteristics used for data acquisition.

Table 1: Summary of Service and Characteristic UUIDs.

Function	Service UUID	Characteristic UUID
Heart Rate	0x180D	0x2A37
Glucose	0x1808	G0x2A18
Blood Pressure	0x1810	0x2A35

Periodically, the application checks the new data collected, being responsible for sending it to a cloud-based platform where classification is carried out by the pre-trained model as detailed in Section 3.0.1. The application and its testing were conducted using a Fitbit Inspire 2 smart band, an Omron HEM-6232T blood pressure monitor, and a G-Tech Lite glucose meter. The Fitbit Inspire 2 smart band utilizes an optical heart rate sensor that employs photoplethysmography technology to measure heartbeats by monitoring changes in blood flow. The Omron HEM-6232T blood pressure monitor has an oscillometric sensor that detects pressure fluctuations in the cuff as it inflates and deflates, capturing pulse waves generated by blood flow. The G-Tech Lite glucose meter uses an amperometric sensor that measures the electric current, which the device processes to calculate and display the blood glucose concentration.

3.0.1 Proposed Methodology

In this section, we present the methodology developed for classifying patients as either healthy or afflicted with stroke. Initially, we provide a detailed description of the dataset employed in this study. Subsequently, we highlight the preprocessing applied. Next, we introduce the classification techniques implemented, describing the stages of the training process. Finally, we define the metrics used in evaluating the results. Figure 1 graphically illustrates the proposal, with each stage explained in detail in the following subsections.

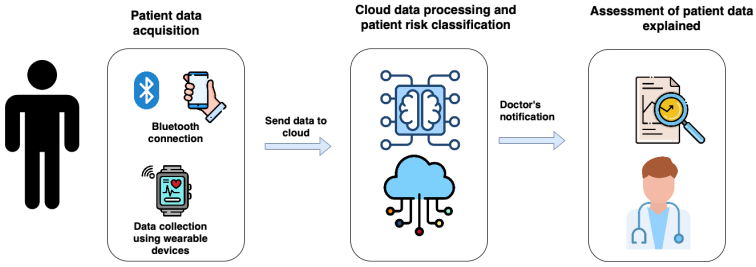

Figure 1: Proposed Methodology.

3.1 Dataset

For the development of the solution, we employed a clinical dataset (15) that includes information from 5110 patients, covering 10 attributes described below: **Gender:** Refers to the patient's gender. The number of males is 1260, while the number of females is 1994. **Age (years):** Refers to the age of the participants. **Heart Disease:** Refers to the presence of heart disease in the patient, with 6.33% of participants having heart diseases. **Hypertension:** Refers to the presence of hypertension in the patient, with 12.54% of participants having hypertension. **Ever Married:** Indicates marital status of participants, with 79.84% of patients being married. **Work Type:** Represents the type of employment of the patient, with 4 categories: self-employed 19.21%, private 65.02%, public 15.67%, and never worked 0.1%. **Residence Type:** Represents the type of residence of the patient, with 2 categories: rural 48.86% and urban 51.14%. **Avg Glucose Level (mg/dL):** Indicates the average glucose level of the patient. **BMI (Body Mass Index):** Indicates the body mass index of the participants. **Smoking Status:** Indicates whether the patient is a smoker 22.37%, never smoked 52.64%, or an ex-smoker 24.99%. **Stroke:** Indicates whether the patient has had a stroke.

3.2 Data Preprocessing

The first step in preprocessing involved converting nominal features into numerical values using the One-Hot Encoding technique (16). The next step was to balance the dataset, as only 249 participants had suffered a stroke. Balancing was achieved by randomly selecting 249 healthy participants to form a balanced subset. Additionally, to improve the generalization capacity of the trained models, the SMOTENC technique (17) was used to generate new samples of stroke patients and equalize them with the healthy ones. Several tests were conducted to determine the optimal number of new samples generated with the mentioned classifiers, and it was found that the best accuracy was achieved when 140 new samples per class were generated. Thus, the new dataset consists of balanced classes, with 389 samples per class. The use of the imbalanced dataset, prior to balancing, caused the model to tend to classify all patients as healthy, impairing its ability to correctly identify patients with the condition.

3.3 Classification Steps

To classify the patients, we chose machine learning methods frequently cited in the literature: Bayes (18), k-Nearest Neighbors (kNN), Multilayer Perceptron (MLP) (19), (20), Optimal Path Forest (OPF) Classifier (21), and Support Vector Machine (SVM) (22), eXtreme Gradient Boosting (XGB) (23), and Random Forest (RF) (24).

The classification is performed in three steps: i) model training, ii) model testing, and iii) repetition of steps i) and ii). Each subset of data consists of samples from the original dataset.

1. *Model Training*: In this stage, we use 80% of the subset to perform model training. We consider the hyperparameter configurations presented in Table 2 to find the classifier settings on the training set. Classifiers configured for random search perform a 10-iteration search. Hyperparameters for all classifiers, except the Bayes classifier, are determined after 10-fold cross-validation.
2. *Model Testing*: In this stage, we perform testing on the remaining 20% of the data subset using the saved classifiers. The system determines a class for each sample in the data subset. Additionally, metrics are calculated in this stage.
3. *Repeat Steps 1) and 2)*: The data subsets are randomly split into other training and testing sets. These sets are differentiated from each other by the seed used. We then perform ten repetitions of steps 1) and 2).

Table 2: Hyperparameter Tuning for Classifiers.

Classifier	Search Type	Parameter	Setup
SVM(Linear)	Random	C	2^{-5} to 2^{15}
SVM(RBF)	Random	C, γ	2^{-5} to 2^{15}
KNN	Grid	Number of neighbors	3, 5, 7, 9, 11, 13, 15
Bayes(Normal)	–	–	Gaussian probability density function
OPF	–	–	Euclidian distance
XGB	Random	learning rate and Nº of estimators	1^{-4} to 1^{-2}, 40 to 100
MLP	Random	Number in hidden layer	2 to 1000 Levenberg-Marquardt method
RF	Random	Number of estimators	2 to 100

3.4 Remote Patient Monitoring

The patient classification model was implemented on a cloud server, developed using a Microservices-based Architecture. In this model, each functionality, such as risk classification, user management, and integration with wearable devices, was implemented as an independent microservice. The server was developed using Django (Python) to ensure robustness and scalability, while communication between services is handled through REST APIs, enabling efficient integration between the system's various modules.

New data from the application is sent to a cloud platform for processing and checking discrepancies. For example, heart rate is considered healthy if it follows the rule $HR_{max} = 208 - 0.7age$ (25); otherwise, it indicates irregular heartbeats. According to the European Society of Hypertension (26) and the European Society of Cardiology (ESC), blood pressure for individuals aged 16 and over is classified as follows, as illustrated in Table 3.

Table 3: Blood Pressure Classification Ranges.

Category	Systolic (mmHg)	Diastolic (mmHg)
Optimal	<120	<80
Normal	$120 - 129$	$80 - 84$
Elevated	$130 - 139$	$85 - 89$
Hypertension - Stage 1	$140 - 159$	$90 - 99$
Hypertension - Stage 2	$160 - 179$	$100 - 109$
Hypertension - Stage 3	≥ 180	≥ 110

This classification is used within the website as a metric for evaluating blood pressure measurements obtained from patients. The monitoring system proposed in this approach is aimed at adults or older people, which is why these measures become valid for use.

In terms of glucose measurements, the guidelines from the American Diabetes Association (ADA) (27) and the World Health Organization (WHO) are important references. The ranges for classifying fasting and postprandial glucose align with the recommendations of these organizations and are widely adopted in clinical practice for the diagnosis and monitoring of diabetes and prediabetes. Table 4 presents the classification of glucose based on observed values.

The comparison of glucose levels may vary depending on the protocol established by the physician, being customized for each patient, including the frequency of measurements. If data outside the established standards is received, a notification will be sent to the attending physician. It is worth

Table 4: Classification of Blood Glucose Levels.

Classification	Anonymous Authors Address	Anonymous Authors Address
Healthy	Less than 100	Less than 140
Pre-diabetes	100 - 125	140 - 199
Diabetes	126 or more	200 or more

noting that patient data is encrypted to anonymize it and is in compliance with the LGPD (General Data Protection Law) (28) in force in Brazil.

4 Results and Discussions

In this section, the results achieved by the stroke patient monitoring system are investigated. We analyzed the results of this article using the following metrics: Accuracy, Recall and Precision. We present the data obtained from the training of the machine learning model, highlighting its effectiveness in predicting stroke risk. Additionally, we discuss the insights provided by the SHAP method, which allowed us to interpret and understand the model's decisions in more detail. The system infrastructure used was an Intel i7, 16 GB of RAM, running Ubuntu 20.04 Linux without a graphical processing unit (GPU). Table 5 presents the metrics and their standard deviations from the results after 10 iterations of the steps described in Section 3.3, with the best results highlighted in green.

Observing the obtained results, it is possible to verify that the XGB classifier stood out with an Accuracy of 98.51%, followed by the OPF and SVM(RBF) classifiers with 97.11% and 96.05%, respectively. A comparison between the XGB, SVM with RBF kernel, and OPF classifiers highlights XGB as the most effective, with an accuracy of 98.51%, due to its ability to handle complex and imbalanced data through boosting. The SVM with RBF kernel (96.05%) performs well but relies on fine-tuning and can be sensitive to outliers. The OPF (97.11%) uses a graph-based approach, efficient for structured data but sensitive to data distribution. The performance differences reflect the characteristics of each method, and factors like dataset biases and generalizability should also be considered.

Still observing the results from Table 5, the XGB classifier showed the highest score in terms of Recall, with 98.06%, meaning that the model correctly identified a good percentage of positive cases of strokes compared to the total number of actual stroke cases in the dataset. In other words, the model captured the vast majority of stroke cases, missing only a small proportion of them. This indicates a high level of sensitivity of the model in detecting stroke cases. Following, the OPF and SVM(RBF) classifiers stood out with 97.03% and 96.25%, respectively.

Regarding Precision, the XGB classifier stood out with a score of 98.01%, followed by the OPF classifier with 96.39% and SVM(RBF) with 95.18%, respectively. These results indicate that these classifiers were able to identify most of the actual stroke cases, minimizing false positives. This demonstrates the ability of the models to make accurate and reliable predictions about stroke cases.

Table 5: Metrics Obtained by Each Classifier.

Classifier	Accuracy(%)	Recall(%)	Precision(%)
SVM(Linear)	92.88 ± 0.31	92.19±0.27	92.15±0.45
SVM(RBF)	**96.05±0.48**	**96.24±0.36**	**95.18±0.64**
kNN	94.03±0.52	94.12±0.42	94.09 ± 0.58
Bayes(Normal)	94.08±0.67	92.34 ± 0.58	92.28±0.73
OPF	**97.11±0.79**	**97.03±0.67**	**96.39±0.85**
XGB	**98.51±0.55**	**98.06±0.33**	**98.01±0.39**
MLP	95.18±0.47	95.23±0.38	95.17±0.57
RF	93.11±0.55	91.06±0.33	92.01±0.39

Table 6: Training and Testing Times Obtained for Each Classifier.

Classifier	Training Time(s)	Testing Time(ms)
SVM(Linear)	**17.43±3.189**	1.412±0.021
SVM(RBF)	19.457±0.147	1.509±0.018
kNN	**0.371±0.049**	2.248±0.042
Bayes(Normal)	**0.094±0.016**	**0.006±0.001**
OPF	765.178±4.421	0.934±0.008
XGB	98.585±19.343	**0.076±0.002**
MLP	1420.715±29.453	**0.096±0.008**
RF	442.715±13.343	0.098±0.012

Evaluating the training and testing times of classifiers is crucial to ensure the efficiency and practical viability of the models. The training time affects the scalability of the model, while the testing time influences its ability to predict in real-time. Table 6 presents the training and testing times obtained for the evaluated classifiers.

Observing the results, the Bayesian (Normal) classifier had the shortest training time, at 0.094s, followed by the kNN and Linear SVM classifiers, which completed the task in 0.371s and 17.43s, respectively. Regarding test time, the Bayesian classifier also stood out with 0.006ms, followed by the XGB and MLP classifiers, with times of 0.076ms and 0.096ms, respectively. Although the Bayesian classifier provides faster responses for prediction requests, the evaluation of accuracy, precision, and recall values indicates that the XGB classifier is more suitable for the task, even though it requires significant more time.

4.1 XAI Performace

To assess the impact of each predictor on the model outcome, we calculated the average values using the SHAP method for the XGB classifier. We observed that the seven factors with the greatest influence are Age, Average glucose level, Type of job, Type of residence, Body mass index, Gender, and Marital status.

(a) Average Feature Impact on Model Output Magnitude

(b) SHAP Value Impact on Model Output (Each Class).

Figure 2: Comparative Analysis of Feature Impacts on Model Outputs.

Through Figure 2a, it is possible to observe the average importance of each feature. Additionally, the Shapley values (on the x-axis) indicate the individual impact of each variable on the prediction. Positive values indicate that the feature increased the probability of class 1 (Stroke), while negative values indicate an increase in the probability of class 0 (Healthy). A Shapley value of 0 means that the variable had no influence on the prediction. The age had the greatest weight for the prediction, while the type of residence was the least relevant variable.

In Figure 2b, when observing the considered variables, it is evident that the attributed impact makes sense. For example, it is natural for strokes to be more common in older individuals and those with higher glucose levels, as observed in Figure 2. As higher values of these quantities are observed, higher Shapley values are also observed, indicating a positive contribution to the classification of the stroke class (illustrated in red). The same pattern can be observed for body mass index, as found in the study conducted by (29). . Female patients (labeled as 1) contributed more to the stroke class (30). Married patients (labeled as 1) also showed a higher contribution to the positive class, as noted in (31). Regarding residence type, patients living in rural areas (labeled as 1) had a slight contribution to the stroke class, according to (32).

Table 7: Explanations of Two Instances from The Dataset.

Feature	Healthy	Stroke
Gender	0.00	0.00
Age	25.00	65.00
Work type	0.00	0.00
Avg glucose level	96.00	120.00
Ever married	0.00	0.00
Body mass index	19.00	26.00
Residence type	1.00	1.00

Table 7 illustrates the explanations of two instances from the dataset. For the Healthy prediction, the result indicates the analysis of feature values, we can identify which characteristics are most relevant for this prediction. Similarly, in the prediction associated with the diagnosis of a stroke, XAI provides

a detailed explanation, highlighting the features and essential information for doctors and patients to understand and trust computer-assisted diagnoses.

We have the instance of a healthy patient: a 25-year-old woman residing in an urban environment, with an average glucose level within normal range and a body mass index considered healthy. Based on this information, the model classified the patient as healthy. In contrast, in the Stroke column, we have the instance of a patient who suffered a stroke. The classification was Stroke due to the advanced age, above-normal glucose level, and unhealthy body mass index highlighted in orange. It is observed that both patients reside in urban areas, which may facilitate access to controlling risk factors for stroke. This is considered a positive factor in managing this type of occurrence, since mortality due to stroke is higher in rural areas than in urban areas, due to difficulty in accessing medical treatment, especially related to poverty.

4.2 Comparisons With Related Work

As shown in Table 8, we compared the proposed approach with the studies of (33), (34), (35) and (36). In (33), the authors used stacking classification to predict strokes, achieving an accuracy of 98%. Despite the good results obtained, the method adds an extra layer of complexity to the model, as it involves training multiple base models and then combining their predictions into a meta-model. This can make model implementation and maintenance more challenging. Furthermore, the computational cost for using this model also increases.

Table 8: Comparison of Other Works with the Proposed Methodology.

Author	**Method**	**Accuracy**	**XAI Techniques**
Our Approach	XGB	98.51%	SHAP and LIME
(33)	Stacking	98.00%	No
(34)	Random Forest	94.70%	No
(35)	Random Forest	98.94%	No
(36)	Random Forest	98.94%	SHAPE and LIME

Still observing the Table 8, in (34) and (35), similar approaches were used for the classification task, in which the Random Forest classifier stood out as the best, presenting accuracies of 94.7% and 98.94%, respectively. In the study presented in (36), an accuracy of 90.36% was achieved in classifying stroke patients. To provide insights into machine learning models considered black boxes, two explainable techniques were also studied: SHAP and LIME. The results were in line with what was presented in the work, with the same features being selected as the most relevant for decision-making in the algorithms.

5 Conclusions and Future Works

In this paper, we present an approach for classifying patients regarding the risk of stroke based on their clinical data, utilizing the XGBoost classifier. We achieved an accuracy of 98.51% and performed the classification task in just 0.076ms. Although the system exhibited high accuracy and sensitivity, it is fundamental to evaluate the clinical implications of these results. Accurately predicting stroke risk, for instance, could enable earlier interventions, enhanced patient monitoring, and the development of personalized treatment plans, contributing to improved clinical outcomes.

Additionally, we developed an application that presents the features considered in decision-making through the LIME algorithm. For patients with potential stroke risk, remote monitoring, including alerts, is possible. It is important to note that our method does not replace medical expertise. Instead, it can serve as an auxiliary diagnostic resource, automating and expediting the process of evaluating patients with potential stroke risk.

For future work, we plan to increase the size of the dataset by incorporating information collected in clinics or hospitals. We also intend to explore other machine learning methods, such as the Minimal Learning Machine and Minimal Learning Machine Nearest Neighbours (37; 38). Additionally, we aim to improve the interpretability of model predictions, potentially integrating Symbolic AI techniques to explain decisions through rules based on medical knowledge. This hybrid approach aims to increase healthcare professionals' confidence by providing clearer and more justified predictions and enabling continuous adjustments based on new findings and clinical guidelines.

References

[1] W. H. Organization. (2024) Stroke (cerebrovascular accident). Accessed in: 20-12-2023. [Online]. Available: https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html

[2] A. Bustamante, A. Penalba, C. Orset, L. Azurmendi, V. Llombart, A. Simats, E. Pecharroman, O. Ventura, M. Ribó, D. Vivien *et al.*, "Blood biomarkers to differentiate ischemic and hemorrhagic strokes," *Neurology*, vol. 96, no. 15, pp. e1928–e1939, 2021.

[3] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt, A. K. Boehme, A. E. Buxton, A. P. Carson, Y. Commodore-Mensah *et al.*, "Heart disease and stroke statistics—2022 update: a report from the american heart association," *Circulation*, vol. 145, no. 8, pp. e153–e639, 2022.

[4] X. Xia, W. Yue, B. Chao, M. Li, L. Cao, L. Wang, Y. Shen, and X. Li, "Prevalence and risk factors of stroke in the elderly in northern china: data from the national stroke screening survey," *Journal of neurology*, vol. 266, pp. 1449–1458, 2019.

[5] K. M. Rexrode, T. E. Madsen, A. Y. Yu, C. Carcel, J. H. Lichtman, and E. C. Miller, "The impact of sex and gender on stroke," *Circulation research*, vol. 130, no. 4, pp. 512–528, 2022.

[6] A. Alloubani, A. Saleh, and I. Abdelhafiz, "Hypertension and diabetes mellitus as a predictive risk factors for stroke," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 12, no. 4, pp. 577–584, 2018.

[7] S. Elsayed and M. Othman, "The effect of body mass index (bmi) on the mortality among patients with stroke," *Eur. J. Mol. Clin. Med*, vol. 8, pp. 181–187, 2021.

[8] J. Patra, B. Taylor, H. Irving, M. Roerecke, D. Baliunas, S. Mohapatra, and J. Rehm, "Alcohol consumption and the risk of morbidity and mortality for different stroke types-a systematic review and meta-analysis," *BMC public health*, vol. 10, no. 1, pp. 1–12, 2010.

[9] M. Yang, H. Yoo, S.-Y. Kim, O. Kwon, M.-W. Nam, K. H. Pan, and M.-Y. Kang, "Occupational risk factors for stroke: A comprehensive review," *Journal of Stroke*, vol. 25, no. 3, p. 327, 2023.

[10] A. M. Cox, C. McKevitt, A. G. Rudd, and C. D. Wolfe, "Socioeconomic status and stroke," *The Lancet Neurology*, vol. 5, no. 2, pp. 181–188, 2006.

[11] A. Onasanya and M. Elshakankiri, "Smart integrated iot healthcare system for cancer care," *Wireless Networks*, vol. 27, pp. 4297–4312, 2021.

[12] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, 2017, pp. 4765–4774.

[13] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you? explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1135–1144.

[14] Bluetooth specifications. Accessed in 2024-02-12. [Online]. Available: https://www.bluetooth.com/specifications/specs/

[15] Fedesoriano, "Stroke prediction dataset," https://www.kaggle.com/datasets/fedesoriano/stroke-dataset, 2023.

[16] J. Buckman, A. Roy, C. Raffel, and I. Goodfellow, "Thermometer encoding: One hot way to resist adversarial examples," in *International conference on learning representations*, 2018.

[17] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. Ieee, 2008, pp. 1322–1328.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[19] S. Haykin, *Redes Neurais - Princípios e Prática*, 2nd ed. Porto Alegre: Boolman, 10 2000, 900 páginas.

[20] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *Information Theory, IEEE Transactions on*, vol. 13, no. 1, pp. 21–27, 1967.

[21] T. M. Nunes, A. L. Coelho, C. A. Lima, J. P. Papa, and V. H. C. De Albuquerque, "Eeg signal classification for epilepsy diagnosis via optimum path forest–a systematic assessment," *Neurocomputing*, vol. 136, pp. 103–123, 2014.

[22] S. Theodoridis and K. Koutroumbas, "Pattern recognition–fourth edition, 2009."

[23] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 785–794.

[24] L. Breiman and A. Cutler, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[25] R. A. Robergs and R. Landwehr, "The surprising history of the "hrmax= 220-age" equation," *J Exerc Physiol*, vol. 5, no. 2, pp. 1–10, 2002.

[26] A. V. Chobanian, G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo Jr, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright Jr *et al.*, "Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure," *hypertension*, vol. 42, no. 6, pp. 1206–1252, 2003.

[27] American Diabetes Association. (2022) Standards of medical care in diabetes. Accessed: May 26, 2024. [Online]. Available: https://professional.diabetes.org/sites/professional.diabetes.or.pdf

[28] "General data protection law (lgpd)," law No. 13,709, of August 14, 2018.

[29] Y. G. Wang, G. Li, Y. S. Han, X. H. Jin, and Y. S. Xu, "Body mass index and risk of stroke: A dose-response meta-analysis of prospective studies," *Medicine (Baltimore)*, vol. 96, no. 12, p. e5711, 2017.

[30] K. M. Rexrode, T. E. Madsen, A. Y. Yu, C. Carcel, J. H. Lichtman, and E. C. Miller, "The impact of sex and gender on stroke," *Circulation research*, vol. 130, no. 4, pp. 512–528, 2022.

[31] K. Andersen and T. Olsen, "Stroke case-fatality and marital status," *Acta Neurologica Scandinavica*, vol. 138, no. 4, pp. 377–383, 2018.

[32] J. Joubert, L. F. Prentice, T. Moulin, S.-T. Liaw, L. B. Joubert, P.-M. Preux, D. Ware, E. Medeiros de Bustos, and A. McLean, "Stroke in rural areas and small communities," *Stroke*, vol. 39, no. 6, pp. 1920–1928, 2008.

[33] E. Dritsas and M. Trigka, "Stroke risk prediction with machine learning techniques," *Sensors*, vol. 22, no. 13, p. 4670, 2022.

[34] H. Al-Zubaidi, M. Dweik, and A. Al-Mousa, "Stroke prediction using machine learning classification methods," in *2022 International Arab Conference on Information Technology (ACIT)*. IEEE, 2022, pp. 1–8.

[35] N. S. Adi, R. Farhany, R. Ghina, and H. Napitupulu, "Stroke risk prediction model using machine learning," in *2021 International Conference on Artificial Intelligence and Big Data Analytics*. IEEE, 2021, pp. 56–60.

[36] K. Mridha, S. Ghimire, J. Shin, A. Aran, M. M. Uddin, and M. F. Mridha, "Automated stroke prediction using machine learning: An explainable and exploratory study with a web application for early intervention," *IEEE Access*, 2023.

[37] A. H. de Souza Junior, F. Corona, G. A. Barreto, Y. Miche, and A. Lendasse, "Minimal learning machine: A novel supervised distance-based approach for regression and classification," *Neurocomputing*, vol. 164, pp. 34–44, 2015.

[38] D. P. P. Mesquita, J. P. P. Gomes, and A. H. S. Junior, "Ensemble of minimal learning machines for pattern classification," in *Advances in Computational Intelligence: 13th International Work-Conference on Artificial Neural Networks, IWANN 2015, Palma de Mallorca, Spain, June 10-12, 2015. Proceedings, Part II 13*. Springer, 2015, pp. 142–152.