



Article Evaluation of Blood Biomarkers and Parameters for the Prediction of Stroke Survivors' Functional Outcome upon Discharge Utilizing Explainable Machine Learning

Aimilios Gkantzios ^{1,2,*}, Christos Kokkotis ³, Dimitrios Tsiptsios ¹, Serafeim Moustakidis ^{3,4}, Elena Gkartzonika ⁵, Theodoros Avramidis ², Nikolaos Aggelousis ³ and Konstantinos Vadikolias ¹

- ¹ Department of Neurology, School of Medicine, University Hospital of Alexandroupolis, Democritus University of Thrace, 68100 Alexandroupolis, Greece
- ² Department of Neurology, Korgialeneio—Benakeio "Hellenic Red Cross" General Hospital of Athens, 11526 Athens, Greece
- ³ Department of Physical Education and Sport Science, Democritus University of Thrace, 69100 Komotini, Greece
- ⁴ AIDEAS OÜ, Narva mnt 5, 10117 Tallinn, Estonia
- ⁵ School of Philosophy, University of Ioannina, 45110 Ioannina, Greece
- * Correspondence: aimilios.gkantzios@yahoo.com

Abstract: Despite therapeutic advancements, stroke remains a leading cause of death and long-term disability. The quality of current stroke prognostic models varies considerably, whereas prediction models of post-stroke disability and mortality are restricted by the sample size, the range of clinical and risk factors and the clinical applicability in general. Accurate prognostication can ease poststroke discharge planning and help healthcare practitioners individualize aggressive treatment or palliative care, based on projected life expectancy and clinical course. In this study, we aimed to develop an explainable machine learning methodology to predict functional outcomes of stroke patients at discharge, using the Modified Rankin Scale (mRS) as a binary classification problem. We identified 35 parameters from the admission, the first 72 h, as well as the medical history of stroke patients, and used them to train the model. We divided the patients into two classes in two approaches: "Independent" vs. "Non-Independent" and "Non-Disability" vs. "Disability". Using various classifiers, we found that the best models in both approaches had an upward trend, with respect to the selected biomarkers, and achieved a maximum accuracy of 88.57% and 89.29%, respectively. The common features in both approaches included: age, hemispheric stroke localization, stroke localization based on blood supply, development of respiratory infection, National Institutes of Health Stroke Scale (NIHSS) upon admission and systolic blood pressure levels upon admission. Intubation and C-reactive protein (CRP) levels upon admission are additional features for the first approach and Erythrocyte Sedimentation Rate (ESR) levels upon admission for the second. Our results suggest that the said factors may be important predictors of functional outcomes in stroke patients.

Keywords: stroke; prognosis; functional outcome; poststroke disability; blood biomarkers; clinical data; artificial intelligence; explainability

1. Introduction

Despite significant breakthroughs in early detection, due to high-quality imaging and therapeutic interventions such as thrombolysis and thrombectomy, stroke continues to be a major cause of acquired disability in adults and the second leading cause of death worldwide, for individuals of all ages [1,2]. Approximately twenty-five percent of people over the age of twenty-five will experience a stroke in their lives and more than fifty percent of stroke survivors continue to have adverse outcomes, including chronic functional reliance and even death, despite the standard care provided in specialized stroke units. Therefore, it



Citation: Gkantzios, A.; Kokkotis, C.; Tsiptsios, D.; Moustakidis, S.; Gkartzonika, E.; Avramidis, T.; Aggelousis, N.; Vadikolias, K. Evaluation of Blood Biomarkers and Parameters for the Prediction of Stroke Survivors' Functional Outcome upon Discharge Utilizing Explainable Machine Learning. *Diagnostics* 2023, 13, 532. https://doi.org/10.3390/ diagnostics13030532

Academic Editors: Semen A. Kurkin and Alexander E. Hramov

Received: 9 January 2023 Revised: 25 January 2023 Accepted: 29 January 2023 Published: 1 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). is evident that the development of new rehabilitation-improving therapies is necessary [3,4]. Currently, there are several data points about the factors that affect the prognosis and functional results of stroke. These variables pertain to the patients' concurrent conditions, genetic and demographic traits, as well as the data collected during the acute phase of the stroke and connected to the stroke as an entity. More specifically, it is well documented and described in numerous studies that the vascular profile of such patients is determined by the presence of diseases such as hypertension, diabetes, hypercholesterolemia, a previous transient ischemic attack, atrial fibrillation, a history of ischemic heart disease (angina or myocardial infarction), internal carotid artery hemodynamically significant stenosis and peripheral arterial disease [5–7]. In terms of demography, age and gender have been identified as key prognostic risk factors. In addition, elevated CRP and creatinine values, the number of cardiovascular disease sites (cerebrovascular, coronary and peripheral arterial disease) and the presence of an infection at the onset of stroke appear to be independent risk factors for a high risk of early stroke recurrence during hospitalization [8–11]. Using MRI and, more specifically, the DWI sequence, which can be used to measure the cerebral infarct volume, it is further possible to anticipate the patient's disability, thus serving as an independent predictor of the 90-day modified Rankin scale (mRS) outcome for neurologic disability in the acute ischemic stroke (AIS) [12].

To date, there have been no significant breakthroughs comparable to those noticed in acute stroke treatment. We understand what causes a stroke considerably better than what promotes recovery of function. Hence, it becomes vital to comprehend the underlying pathophysiological processes, in order to support rehabilitation. Rehabilitation therapies and services are essential for restoring patients' abilities, independence and quality of life. However, a lack of consensus among measurements hinders the ability to fully maximize clinical findings and define a sufficient level of evidence for treatments, given the expanding availability of rehabilitation therapies. Thus, in order to study in depth the factors that influence patients' recovery and to support the clinical choice regarding tailored treatment, a precise and comprehensive assessment is required [13–15].

In this direction, blood biomarkers have emerged as a useful diagnostic tool and a new method for accurately predicting functional outcomes following stroke [16], while several risk-prediction models have been developed thus far for detecting functional outcomes following an acute stroke. The majority of the latter base their prognosis on comparable input factors such as age, the severity of the first stroke, and comorbidities. However, none of these predictive models has been generally utilized in clinical practice, most probably due to implementation challenges [17–20]. As a result, the most difficult aspect of stroke rehabilitation research is the optimizing of rehabilitation regimens, based on an early prognosis [21,22]. A precise prognosis may aid in post-stroke discharge planning and the development of a more individualized rehabilitation program. However, it may be difficult to accurately predict the patients' results, most particularly in the early post-admission period [23–25].

The far more recent rise in the respective studies' interest might be ascribed to the growing complexity and quantity of data, as well as to the existence of multivariate data with heterogeneous origins, for which conventional approaches have proven incapable of delivering accurate results [22,26,27]. In recent years, machine learning (ML) methodologies have been widely used to address complex problems in a range of scientific fields, including medicine [28–32]. Machine learning is a subset of artificial intelligence in which a computer acquires knowledge from data, without explicit programming of rules. Machine learning algorithms identify patterns within data and use them to make predictions on new data. These models are beneficial for utilizing enormous datasets, due to the computers' capacity to rapidly process vast volumes of electronic data. Where people are expected to become overwhelmed by ever-growing volumes and complexities of data inputs, machine learning models frequently thrive. To develop machine-learning solutions for clinical usage, one must have a solid grasp of clinical care, data science and implementation science. The idea that machine learning might quickly and drastically improve health care by automating

routine processes and increasing clinical decision-making is enticing today. Finding effective machine-learned solutions necessitates careful consideration of poor classification and collection of health care data, as well as the complexity of clinical choices and workflows. In fact, machine learning has the potential to significantly improve health care. To genuinely assist the patients we serve, however, a disciplined, inclusive, collaborative and iterative approach to the development and use of new technologies is essential. Nonetheless, the practice of medicine will remain relatively unchanged at its foundation. Medicine entails communicating with patients and assisting them at, frequently, the most difficult phases of their life. These higher cognitive activities, such as understanding a patient's goals and assessing competing priorities, are unlikely to be delegated to a computer. The future will likely be characterized by augmented intelligence, in which computers may become indispensable instruments for patient care, allowing physicians to dedicate more time to the patient care [33,34].

Hence, in its attempt to prevent strokes, the scientific community focuses on the creation of models that can anticipate their occurrence. In this context and since its implementation, artificial intelligence (AI) has played a key role in disease prevention and is now ubiquitous [35–37]; a growing number of studies indicate that the ML technique is more accurate in predicting stroke outcomes than statistical methods or scoring systems [38–40]. According to a recent review, numerous studies have been undertaken to generate models for stroke diagnosis [41–43] or predict treatment responses and clinical outcomes, with the aim of developing tailored rehabilitation regimens [44–47]. The emergence of machine learning (ML) and deep learning has facilitated the incorporation of a wide variety of structured data for the data-driven prediction of clinically meaningful outcomes in stroke patients [20,48,49].

To this end, the aim of the study at hand is to collect and analyze data from the hyperacute and acute phases of stroke, with the goal of developing an explainable machine learning model for predicting functionality at discharge, as measured by the mRS. Even though there is no commonly accepted stroke outcome measurement scale, we decided to use the mRS to assess the poststroke functional outcome of the patients, having considered the advantages of this scale. More specifically, the mRS is extensively used to assess residual disability and functional outcome in stroke patients. Both its reliability and its effectiveness in evaluating residual handicap in stroke patients have been thoroughly demonstrated. The mRS provides a rapid and straightforward evaluation of a patient's stroke results, activities and social contribution, encompassing aspects of cognitive function, language, visual function and pain, that are not explicitly evaluated. Particularly, the 7-level mRS has several important strengths: its categories are simple and easily comprehended by both doctors and patients, its concurrent validity is established by a strong connection with measures of stroke pathology and agreement with other stroke scales, while its usage has delineated effective and ineffective acute stroke treatment. Having fewer levels than other stroke scales, the mRS may be less sensitive to change; yet, a single-point change on the mRS is clinically significant [50–54].

Explainable artificial intelligence (AI) is a type of AI, able to provide interpretable and transparent reasoning for its predictions or decisions. In this case, the aim is not only to accurately predict functionality at discharge but also to identify the variables from patients' medical histories and admission biomarkers and parameters that contribute to the mRS result at discharge. This information can be used to support rehabilitation decision-making and improve patient outcomes by tailoring rehabilitation plans to the specific needs and characteristics of each individual patient. The use of explainable AI in this context is particularly important, as it can help healthcare professionals understand and trust the predictions and decisions made by the model, ultimately leading to better patient care.

2. Materials and Methods

2.1. Participants

This study included a total of 470 patients hospitalized with acute stroke at Korgialeneio—Benakeio "Hellenic Red Cross" General Hospital of Athens from July 2017 to June 2018. The study was approved by the Scientific Council of the hospital (protocol nr. 6673/08-03-2018). All patients were observed retrospectively until discharge. Stroke severity at admission was assessed using the National Institutes of Health Stroke Scale (NIHSS) and functional outcomes at discharge were assessed using the modified Rankin Scale (mRS). We determined a study population consisted of patients aged over 18, with ischemic or hemorrhagic stroke, without previous functional deficits (mRS before stroke, 0).

2.2. Data Description

In the present study, we collected patient data on a total of 35 different parameters or variables, as follows: demographic characteristics (age and gender), the type of stroke (ischemic or hemorrhagic), NIHSS scores and admission levels of systolic blood pressure, glucose, CRP and ESR. We also recorded data from patients' medical history, including hypertension, smoking, diabetes, dyslipidemia, atrial fibrillation, previous stroke, previous myocardial infarction or coronary heart disease, history of heart failure, history of mechanical or bioprosthetic heart valve, history of alcoholism, history of taking antiplatelet drugs and history of taking anticoagulant drugs. We also recorded data from the first 72 h after a stroke, including intubation, initial diagnosis of hypertension, diabetes, dyslipidemia and atrial fibrillation, as well as total cholesterol, LDL, HDL, low T3 thyroid hormone, B12 vitamin, folic acid levels and the presence or absence of respiratory infection. For the same period of time, we also recorded stroke localization based on blood supply and hemispheric stroke localization according to CT or MRI brain imaging. At discharge, we assessed functional outcomes for each patient using the mRS.

2.3. Problem Definition

In this study, we consider the task of mRS grade prediction as a binary classification problem. Under this prism, we worked with two different approaches. In the first approach we divided the post-stroke patients into two categories: (i) Independent: Post-stroke patients with 0, 1 and 2 scores in mRS at discharge from hospital and (ii) Non-Independent: Post-stroke patients who scored 3, 4, 5 and 6 in mRS at discharge from hospital. In a second approach, we divided the employed post-stroke patients into two classes: (i) Non-disability: Post-stroke patients with 0 and 1 scores in mRS at discharge from hospital, who have no disabilities and (ii) Disability: Post-stroke patients who scored 2, 3, 4, 5 and 6 in mRS at discharge from hospital, who have no disabilities and (ii) Disability: Post-stroke patients who scored 2, 3, 4, 5 and 6 in mRS at discharge from hospital, who have no disabilities and (ii) Disability: Post-stroke patients who scored 2, 3, 4, 5 and 6 in mRS at discharge from hospital, who have no disabilities and (ii) Disability: Post-stroke patients who scored 2, 3, 4, 5 and 6 in mRS at discharge from hospital, who have no disabilities.

1st Approach: Independent for daily activities post-stroke patients versus nonindependent for daily activities post-stroke patients, employing data from the baseline.

- Class A: involves 219 patients with 0, 1 and 2 mRS grades at discharge and
- Class B: comprises 247 patients with 3, 4, 5 and 6 mRS grades at discharge.

2nd Approach: Post-stroke patients without disabilities versus post-stroke patients with mild to severe disabilities, using data from the baseline.

- Class A: involves 153 patients with 0 and 1 mRS grades at discharge and
- Class B: comprises 313 patients with 2, 3, 4, 5 and 6 mRS grades at discharge.

For both approaches, the classification outputs 0 and 1 correspond to assignments to classes A and B, respectively.

2.4. Machine Learning Workflow Methodology

To identify robust biomarkers and parameters associated with the prediction of the functional outcome as assessed by mRS at discharge in the aforementioned approaches, a hierarchal ML pipeline was implemented (Figure 1). In the first stage, the data were normalized. The standardization of the features was performed by removing the mean

and scaling to unit variance, in order to create a common basis for the feature selection stage and the learning process. For feature selection, we used the Boruta library, which is a wrapper technique based on a Random Forest classifier. This process ranked our clinical data by feature importance [55,56]. Five well-established classifiers were employed, including Random Forest (RF), XGBoost, Multi-Layer Perceptron (MLP), Support Vector Machines (SVMs) and Logistic Regression (LR). The purpose of employing five wellestablished classifiers was to determine their applicability to the specific problem. To minimize overfitting and maximize the performance of the ML models, hyperparameter tuning was applied. RF and XGboost classifiers belong to ensemble learning algorithms and were utilized as a result of their increased performance and speed [57,58]. MLP is a type of simple neural network that was included in this study because it has the ability to handle complex data and has a structure similar to the human brain [59]. In addition to MLP, we also employed SVMs and LR models in our proposed methodology [60,61]. Both of these models have been widely used in various applications and have shown good performance in classification tasks, which is why they were included in this study. SVMs are particularly useful when dealing with high-dimensional and complex datasets, while LR is a simple and efficient model that is particularly useful for binary classification problems.

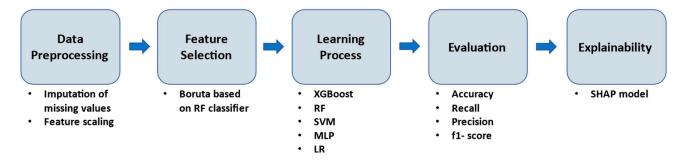


Figure 1. Workflow of the proposed methodology.

We employed a 70% training and 30% testing data split to validate the proposed ML models. During the training phase, we used an internal 10-fold cross-validation to identify the consistent hyperparameters for each model. To balance the class distribution, we applied a random under-sampling reduction technique, in which the majority class was reduced to the same size as the minority class. To evaluate the performance of the proposed classifiers, we used several metrics including accuracy, f1-score, sensitivity (also known as recall) and precision. For the best-performing models, we also presented the confusion matrix. Finally, in order to understand the decision-making mechanism and the contribution of each selected feature in our medical task, we used Shapley Additive Explanations (SHAP) to interpret the output of the ML model. SHAP is a mini-explainer that provides a robust way to interpret ML models, which are often treated as black boxes [62–64]. By using SHAP, we become able to gain insights into the relative importance of each feature and the way it contributes to the overall prediction made by the model. This can be particularly useful for understanding the underlying mechanisms of the model and for identifying potential areas for improvement.

3. Results

In this section, we present the prediction performance results, the selected features and the explainability results for the best ML models, separately in each approach. We will provide detailed information on the performance of the models, the features that were most important in making predictions and the ways in which these features contributed to the overall prediction made by the model.

3.1. Prediction Performance

This sub-section focused on the testing performance results of the employed classifiers. These results are based on the classifiers trained on the first (1, 2, 3, etc.) most informative biomarkers. We will show the testing performance in relation to the number of selected biomarkers and compare it to the performance of the best model for each approach. The results will be presented separately for each approach.

First Approach: Independent and Non-Independent Categorization

In the first approach, all the competitive classifiers presented a rising trend regarding selected biomarkers in the first 2–6 features, with a maximum of 88.57% (RF classifier) at eight features (Figure 2). Furthermore, the inclusion of the additional biomarkers kept the performance of the XGBoost classifier constant. MLP and LR presented a decrease in accuracy with the addition of extra features.

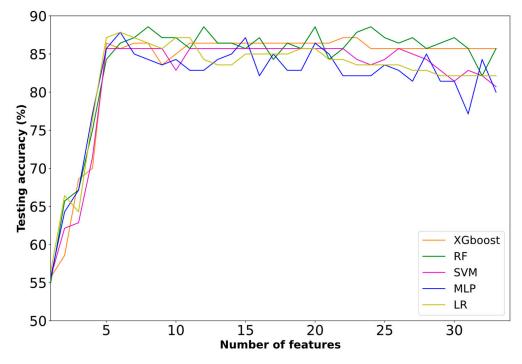


Figure 2. Learning curves based on the testing performance of the first approach.

Table 1 summarizes the best performance results of the employed ML models. The highest accuracy (88.57%) was achieved from RF at eight features. Furthermore, RF achieved 94.74% recall, 85.71% precision and a 90.00% f1-score. On the other hand, the lowest accuracy (85.71%) was achieved by the SVM classifier at five features.

Table 1. Best scores of the competitive classifiers.

Classifiers	Accuracy (%)	Recall	Precision	f1-Score	Num of Features
XGboost	87.14	96.05	82.96	89.02	22
RF	88.57	94.74	85.71	90.00	8
SVM	85.71	93.42	82.56	87.65	5
MLP	87.86	96.05	83.91	89.57	6
LR	87.86	93.42	85.54	89.03	6

In bold the model with the highest accuracy.

In addition, the confusion matrix for the best ML model (RF classifier) is presented below in Figure 3.

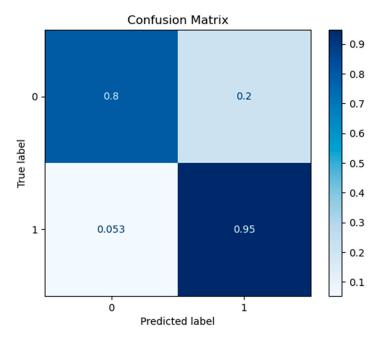


Figure 3. Confusion matrix for RF model at eight features.

Second Approach: Disability and Non-Disability Categorization

A corresponding trend to the aforedescribed first approach is observed in the prediction task for the binary problem of disability and non-disability. All the competitive classifiers presented a rising trend regarding selected biomarkers in the first 3–7 features, with a maximum of 89.29% at seven features (Figure 4). Furthermore, the inclusion of the additional biomarkers kept the performance in the majority of the classifiers constant. As observed, only the SVM presented non-constant performance with the additional biomarkers.

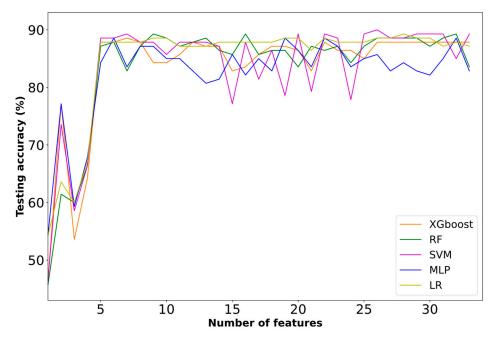


Figure 4. Learning curves based on the testing performance of second approach.

Table 2 demonstrates the best performance results of the employed ML models. The highest accuracy (89.29%) was achieved by SVM at seven features. Furthermore, SVM also achieved 89.47% recall, 94.44% precision and a 91.89% f1-score. On the other hand, the

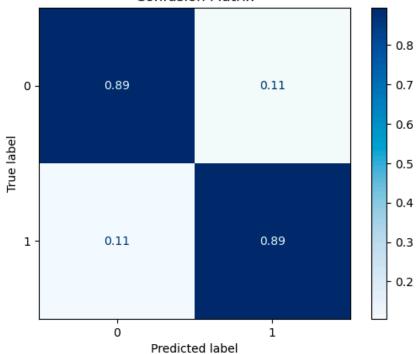
lowest accuracy (88.6%) was achieved by the XGBoost, MLP and LR classifiers at seven, six and nine features, respectively.

Classifiers	Accuracy	Recall	Precision	F1-Score	Num of Features
XGBoost	88.59	87.35	95.41	91.21	7
RF	89.27	89.45	94.43	91.88	9
SVM	89.29	89.47	94.44	91.89	7
MLP	88.57	87.37	95.42	91.22	6
LR	88.56	87.36	95.40	91.23	9

Table 2. Best performance results of the employed ML models.

In bold the model with the highest accuracy.

The confusion matrix for the best-performing machine learning model (SVM classifier) is shown below in Figure 5.



Confusion Matrix

Figure 5. Confusion matrix for SVM model at seven features.

3.2. Selected Features

In this section, the importance of the selected features for both of the binary problems will be presented separately. As described in Section 2.4, we utilized the Boruta feature selection method to identify all relevant features in our dataset. Boruta is an all-relevant feature selection algorithm that compares the importance of a feature to the importance of a random shadow feature, using the RF algorithm to measure feature importance. The features were ranked based on their importance, and those with greater importance than the random shadow feature were considered relevant. The algorithm iteratively continues this process, until all the features have been either selected or rejected. As a wrapper method, Boruta uses the base estimator to compute feature importance and make a decision about feature selection. This technique is powerful for datasets with a large number of features and works well with both continuous and categorical features. In Tables 3 and 4, the feature name, the description and the type of variable (numerical or categorical) are presented.

Ranking	Features	Туре
1	Age	Categorical
2	Hemispheric stroke localization	Categorical
3	Stroke localization based on blood supply	Categorical
4	Development of respiratory infection	Categorical
5	NIHSS upon admission	Categorical
6	CRP levels upon admission	Categorical
7	Systolic blood pressure levels upon admission	Categorical
8	Intubation	Categorical

Table 3. Ranking according to the feature significance of the selected characteristics from the best ML model (RF) at the binary problem independent vs. non-independent.

Table 4. Ranking according to the feature significance of the selected characteristics from the best ML model (RF) at the binary problem disability vs. non-disability.

Ranking	Features	Туре
1	Age	Categorical
2	Hemispheric stroke localization	Categorical
3	Stroke localization based on blood supply	Categorical
4	Development of respiratory infection	Categorical
5	NIHSS upon admission	Categorical
6	Systolic blood pressure levels upon admission	Categorical
7	ESR levels upon admission	Categorical

The feature importance results for both binary classification problems are presented in Tables 3 and 4. These tables include the name of each selected feature and the type of the variable (whether it is numerical or not). These results provide insight into the factors that are most influential in determining the functional outcome of stroke patients at discharge, as assessed by the modified Rankin Scale. Understanding these factors can help inform the development of targeted rehabilitation strategies and interventions.

3.3. Explainability Analysis

In this section, the impact of different biomarkers and parameters on the output of the best-performing model (RF classifier) is evaluated. Figure 6 shows the relationships of these biomarkers and parameters with the binary outcome of being independent or non-independent in daily activities at discharge. The SHAP summary values of these biomarkers and parameters are displayed in descending order, with the most impactful features at the top. The color coding in Figure 6 indicates the extent of the impact on the model's output, with red indicating the high values and blue indicating the low ones for each given observation.

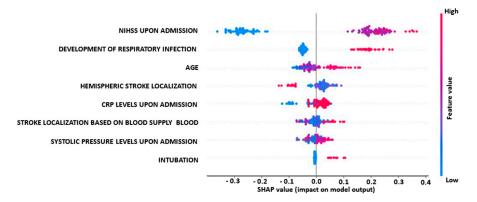


Figure 6. Biomarkers' impact on RF model output for the prediction of mRS at independent vs. non-independent task.

According to the figure, NIHSS upon admission, the development of respiratory infection, age, CRP levels upon admission, stroke localization based on blood supply, systolic blood pressure levels upon admission and intubation all have a strong positive impact on the functional outcome for post-stroke patients, indicating that an increase in these values is closely associated with non-independence at discharge. On the other hand, hemispheric stroke localization has a negative correlation with non-independence at discharge, with left and bilateral localizations having better functional outcomes than most right hemispheric localizations.

Figure 7 depicts the impact of biomarkers and parameters on the output of the bestperforming model (SVM classifier) across test instances for the prediction task of disability versus non-disability. Factors such as NIHSS upon admission, the development of respiratory infection, systolic blood pressure levels upon admission, stroke localization based on blood supply and age have a high and positive impact on the pure functional outcome for post-stroke patients (e.g., the development of respiratory infection pushes the model output towards the disability class), thus indicating that the presence of respiratory infection is closely associated with mild to severe disabilities in post-stroke patients. On the other hand, hemispheric stroke localization is negatively correlated with the disability class of post-stroke patients, with left and bilateral hemispheric stroke localizations generally resulting in better functional outcomes at discharge compared to the majority of right hemispheric stroke localizations.

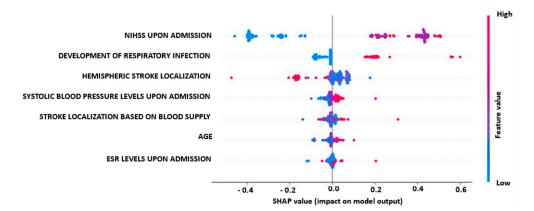


Figure 7. Biomarkers' impact on MLP model output for the prediction of mRS at disability vs. non-disability task.

4. Discussion

In this study, having identified and utilized the aforementioned 35 parameters upon admission, in the first 72 h and from the medical history of patients with stroke, we attempted to develop an explainable machine learning model for the prediction of functional outcome at discharge. Under this prism, we worked with two different approaches, which resulted in age, hemispheric stroke localization, stroke localization based on blood supply, development of respiratory infection, NIHSS upon admission, CRP levels upon admission, systolic blood pressure levels upon admission, intubation and ESR levels upon admission being the more crucial and influential features that appear in determining the functional outcome of stroke patients at discharge. What follows is a literature-based discussion of the association between those relevant criteria, the functional outcome prognosis of stroke patients and the findings of our analysis, taking into consideration that, hitherto, researchers have separately evaluated each of the aforementioned factors but have not juxtaposed them with each other. In particular: age is considered the primary non-modifiable risk factor for stroke [65,66]. The mortality and morbidity rates as well as the functional recovery rates for elderly stroke patients are higher than those for younger patients [67]. In several previous epidemiological studies, age, along with baseline stroke severity as depicted in NIHSS, is another significant predictor of post-stroke outcome [68]. In the respective literature it is

also claimed that as age increases, evidence-based stroke care guidelines are less likely to be followed [66,69,70].

Stroke localization, in which hemispheric detection or blood supply are involved, appears to have a complex relationship with the spatial distribution of the damage and its relevance to functional recovery. Several studies [71-74] have demonstrated that hemisphere lateralization has variable effects on functional outcomes. It is generally recognized that ischemic strokes affecting the dominant hemisphere, which in the majority of the population is the left- as opposed to the non-dominant or right hemisphere, can result in unique clinical syndromes and deficits depending on the regions affected. Multiple studies [74–82] have also shown that the right hemisphere's involvement is predictive of inferior functional outcomes in patients with AIS. Given that the connection between lesion location, clinical impairments, recovery and functional outcome, as measured by mRS, is expected to be altered by the site of the lesion, differentiation is crucial [82–87]. On the other hand, there have also been reports indicating that hemisphere lateralization has no effect on functional outcomes and that there is no significant difference between left- and right-hemispheric strokes in terms of favorable result rates [72,73,83,84]. From the above, it is obvious that infarct location can have a significant impact on the prognosis of individuals, who have suffered an acute ischemic stroke; therefore, prediction models must account for it [85].

As highlighted in several studies, following a stroke, respiratory infection and, more specifically, pneumonia are closely related to an increased risk of poor outcome or mortality. Within the first three months following a stroke, two of every three cases of pneumonia occur in the first week, with the highest frequency occurring on the third day. Evaluating preventative measures for pneumonia is most effective during the first four days after a stroke. Pneumonia that develops later has also been associated with poor functional outcomes or mortality [86–91].

The NIHSS is a quantitative tool used to evaluate the severity of a stroke. It has been found to be a robust predictor of functional outcome 90 days after stroke. Low (<5) and high (>22) NIHSS cut-off points are highly predictive of positive (mRS 0-1) and negative (mRS 4–5 or death) outcomes, respectively. Results are less clear for intermediate initial NIHSS and symptom duration criteria. However, it favors anterior circulation stroke over posterior circulation stroke, resulting in larger results for the former. The mean NIHSS in the group with a favorable functional outcome is considerably lower for posterior circulation stroke than for anterior circulation stroke, highlighting the need to base acute stroke emergency management decisions on a comprehensive neurological examination, as opposed to a single scale, especially in the cases where posterior circulation stroke is in question [92–95].

Studies have shown that elevated CRP levels in the early stages of ischemic or hemorrhagic strokes are related to stroke severity, poor clinical outcomes and long-term prognosis. In particular, after an ischemic stroke, CRP has been found to be an independent predictor of long-term death [96,97].

Hypertension is a well-documented modifiable stroke risk factor. J- and U-shaped curves have been seen between systolic blood pressure (SBP) and clinical outcomes in acute ischemic stroke patients (a satisfactory clinical result at 3 months was construed as an mRS score between 0 and 2); both low and high SBP have been associated with adverse outcomes (mRS 3–6). According to a number of studies, elevated blood pressure after a stroke may increase infarct volume and is associated with a poorer clinical outcome (mRS 3–6). There are other reports indicating that a higher diastolic blood pressure (DBP) at admission is predictive of in-hospital mortality among patients with acute ischemic stroke. Elevation in mean arterial pressure (MAP) is indicative of increased mortality over the long term. Higher entry blood pressure (BP) in patients suffering from acute ischemic stroke, who were treated with endovascular therapy (EVT), is associated with reduced odds of effective reperfusion and poor clinical outcomes. Other studies affirm that, in patients with AIS, a low SBP on admission is associated with an increased risk of in-hospital mortality and comorbidities such as heart failure, gastrointestinal bleeding and sepsis [98–106].

Regarding intubation in the hyperacute phase of stroke or in the early days after stroke, recent research, despite being retrospective, suggests that intubation is associated with an increase in mortality. Moreover, various studies argue that poor outcomes may be associated with an increased risk of aspiration pneumonia and sepsis in intubated patients as opposed to non-intubated patients [107–111].

Lastly, the levels of ESR at admission should not be used as a direct predictors of inflammation in ischemic stroke. While other markers of inflammation, such as CRP, fibrinogen and triglycerides, may be elevated in ischemic stroke patients, ESR levels alone may not be sufficient to identify patients with persistent ischemic stroke episodes. It is worth noting that ischemic stroke patients often have a wide range of risk factors and it is possible that they may have a pre-existing proinflammatory procoagulant state, which could contribute to the shift in ESR values shortly after a cerebral infarction, with the exception of hematological parameters [112].

In our study, for the first approach, the biomarkers' and parameters' impact on the best-performing model's (RF classifier) output across test instances for independent vs. non-independent prediction classes, with an accuracy of 88.57%, revealed that NIHSS upon admission, the development of respiratory infection, age, CRP levels upon admission, stroke localization based on blood supply, systolic blood pressure levels upon admission and intubation show a high and positive impact on the pure functional outcome for post-stroke patients. On the other hand, hemispheric stroke localization is negatively correlated with non-independent post-stroke patients. More specifically, hemispheric stroke localizations such as left and bilaterally have better functional outcomes at discharge than the majority of the right hemispheric stroke localizations. For the second approach, the biomarkers' and parameters' impact on the best-performing model's (SVM classifier) output across test instances for the disability vs. non-disability prediction task, with an accuracy of 89.29%, revealed that NIHSS upon admission, the development of respiratory infection, systolic blood pressure levels upon admission, stroke localization based on blood supply and age show a high and positive impact on the pure functional outcome for poststroke patients. On the other hand, hemispheric stroke localization is negatively correlated with the disability class post-stroke patients. In particular, hemispheric stroke localizations such as left and bilaterally have better functional outcomes at discharge than the majority of the right hemispheric stroke localizations.

4.1. Limitations

Our research was concluded under the assumption of the following limitations: Although the mRS is the most commonly used outcome scale in stroke research, it is primarily dictated by motor impairment and is rather insensitive to cognitive impairment. As a measure of global functional status, the mRS is highly dependent on ambulatory status and lacks sensitivity for deficits in other domains, such as language or neglect, which influence functional outcomes. Due to the inherent limitations of the mRS, disparate methods are required to evaluate functional results. In addition, due to the fact the study at hand is retrospective, namely the database was created before the widespread application of intravenous thrombolysis to patients with ischemic stroke and the mechanical thrombectomy era, our results require further replication in a prospective population. Despite our confident belief that the same parameters will remain able to predict the functional outcome, we expect that the proportion of patients with severe or moderate stroke is likely to be reduced, resulting in a superior functional recovery, given the administration of intravenous thrombolysis to a number of stroke patients. Finally, it shall be noted that the data pertain to one hospital, even though the Neurological Clinic handles a sufficient number of stroke patients of all ages—a feature that distinguishes it from the majority of neurological clinics in Greece. Hence, it should finally be taken into consideration that the sample represents particular geographic and socioeconomic features.

4.2. Future Work

In future work, we will employ state-of-the-art ratios of biomarkers and parameters for the long-term prediction (\geq 3 months) of the functional outcome at the acute phase of stroke patients utilizing robust feature selection techniques and explainable machine learning (XAI) techniques. An external validation dataset will be employed in order to assess the best model's reproducibility and generalizability. In addition, we intend to create a database in which Brain Reserves, such as brain volume and leukoaraiosis, will be examined and evaluated.

5. Conclusions

In conclusion, this study identified 35 parameters upon admission, in the first 72 h and from the medical history of patients with stroke and attempted to develop an explainable machine learning model for the prediction of functional outcome at discharge using the mRS as a binary classification problem. Two approaches were employed and competitive classifiers showed a rising trend in relation to selected biomarkers in the first 2-6 features of the first approach and in the first 3–7 features of the second approach. The maximum accuracy was achieved at eight and seven features, respectively, with six of the features being common to both approaches. Up to today, the predictive models for post-stroke outcome that have been developed using machine learning have focused on large-vessel ischemic strokes that underwent mechanical thrombectomy, and the functional result of patients was assessed three months after the stroke [29,38,113–115]. To the best of our knowledge, this study is the first attempt to collect data during the acute phase of stroke in conjunction with the detection of parameters from the medical history of these stroke patients and to analyze the data using machine learning, in an effort to develop a prognostic model of the functional outcome of the patients upon hospital discharge, thereby potentially contributing to timely, but also more accurate and individualized, diagnosis and treatment. This information can also be used to support rehabilitation decision-making and improve patient outcomes by tailoring rehabilitation plans to the specific needs and characteristics of each individual patient. While the results of this study appear to be promising in predicting functional outcome in stroke patients, it is important to note the limitations of the mRS as an outcome measure and the retrospective nature of the study. Further replication in a prospective population with the inclusion of intravenous thrombolysis and mechanical thrombectomy is necessary to confirm the robustness of these findings. Withal, this study's sample size was restricted to a single institution and may not be representative of stroke patients in different geographic and socioeconomic contexts.

Author Contributions: Conceptualization, A.G.; Data curation, A.G.; Funding acquisition, N.A. and K.V.; Methodology, A.G.; Software, C.K. and S.M.; Supervision, T.A., N.A. and K.V.; Writing—original draft, A.G., C.K. and S.M.; Writing—review & editing, D.T., E.G., T.A., N.A. and K.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the MIS 5047286 grant from Greek and European funds (EYD-EPANEK).

Institutional Review Board Statement: The collection of the data included in the present study was conceptualized and materialized by Dr. Aimilios Gkantzios, in the context of his master thesis for the Master Program "Stroke" of Democritus University of Thrace, respectively approved by the Scientific Council of the Korgialeneio—Benakeio "Hellenic Red Cross" General Hospital of Athens under the permission for scientific use protocol nr. 6673/08-03-2018. The study was conducted according to the guidelines of the Declaration of Helsinki.

Informed Consent Statement: Patient consent requirement was waived due to the fact that our research presents minimal, if any, risk to its subjects and it does not adversely affect their rights and welfare: The study is entirely retrospective; De-identified, archival data from medical records created for administrative non-research purposes have been used, thus (a) no study results could affect the subjects of our study or/and any clinical decision about the individual's care, (b) it was practically impossible to obtain consent from anonymous subjects and (c) it was inappropriate to attempt to

identify and contact them. For the reasons mentioned, this study could not practically be carried out without the waiver.

Data Availability Statement: The dataset generated and/or analyzed during the current study is not publicly available.

Acknowledgments: We acknowledge the support of the work at hand by the project "Study of the interrelationships between neuroimaging, neurophysiological and biomechanical biomarkers in stroke rehabilitation (NEURO-BIO-MECH in stroke rehab)" (MIS 5047286), which is implemented under the Action "Support for Regional Excellence"- Operational Program "Competitiveness, Entrepreneurship and Innovation" (NSRFm2014-2020), under the auspices of Greece and the European Union (European Regional Development Fund).

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Donkor, E.S. Stroke in the 21st century: A snapshot of the burden, epidemiology, and quality of life. *Stroke Res. Treat.* 2018, 2018, 3238165. [CrossRef]
- Roth, G.A.; Mensah, G.A.; Johnson, C.O.; Addolorato, G.; Ammirati, E.; Baddour, L.M.; Barengo, N.C.; Beaton, A.Z.; Benjamin, E.J.; Benziger, C.P.; et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. J. Am. Coll. Cardiol. 2020, 76, 2982–3021. [CrossRef]
- Benjamin, E.J.; Virani, S.S.; Callaway, C.W.; Chamberlain, A.M.; Chang, A.R.; Cheng, S.; Chiuve, S.E.; Cushman, M.; Delling, F.N.; Deo, R.; et al. Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation* 2018, 137, e67–e492.
- Christidi, F.; Tsiptsios, D.; Sousanidou, A.; Karamanidis, S.; Kitmeridou, S.; Karatzetzou, S.; Aitsidou, S.; Tsamakis, K.; Psatha, E.A.; Karavasilis, E.; et al. The Clinical Utility of Leukoaraiosis as a Prognostic Indicator in Ischemic Stroke Patients. *Neurol. Int.* 2022, 14, 952–980. [CrossRef]
- Xu, M.; Amarilla Vallejo, A.; Cantalapiedra Calvete, C.; Rudd, A.; Wolfe, C.; O'Connell, M.D.L.; Douiri, A. Stroke Outcomes in Women: A Population-Based Cohort Study. *Stroke* 2022, 53, 3072–3081. [CrossRef]
- 6. Wańkowicz, P.; Nowacki, P.; Gołąb-Janowska, M. Risk factors for ischemic stroke in patients with non-valvular atrial fibrillation and therapeutic international normalized ratio range. *Arch. Med. Sci.* **2019**, *15*, 1217–1222. [CrossRef]
- Wańkowicz, P.; Gołąb-Janowska, M.; Nowacki, P. Risk factors for death by acute ischaemic stroke in patients from West-Pomerania, Poland. Neurol. Neurochir. Pol. 2020, 54, 150–155. [CrossRef]
- Ekker, M.S.; Verhoeven, J.I.; Schellekens, M.M.I.; Boot, E.M.; van Alebeek, M.E.; Brouwers, P.J.A.M.; Arntz, R.M.; van Dijk, G.W.; Gons, R.A.R.; van Uden, I.W.M.; et al. Risk Factors and Causes of Ischemic Stroke in 1322 Young Adults. *Stroke* 2022, 54, 439–447. [CrossRef]
- Xu, J.; Yalkun, G.; Wang, M.; Wang, A.; Wangqin, R.; Zhang, X.; Chen, Z.; Mo, J.; Meng, X.; Li, H.; et al. Impact of Infection on the Risk of Recurrent Stroke Among Patients With Acute Ischemic Stroke. *Stroke* 2020, *51*, 2395–2403. [CrossRef]
- Li, J.; Pan, Y.; Xu, J.; Li, S.; Wang, M.; Quan, K.; Meng, X.; Li, H.; Lin, J.; Wang, Y.; et al. Residual Inflammatory Risk Predicts Poor Prognosis in Acute Ischemic Stroke or Transient Ischemic Attack Patients. *Stroke* 2021, 52, 2827–2836. [CrossRef]
- Gynnild, M.N.; Hageman, S.H.J.; Dorresteijn, J.A.N.; Spigset, O.; Lydersen, S.; Wethal, T.; Saltvedt, I.; Visseren, F.L.J.; Ellekjær, H. Risk Stratification in Patients with Ischemic Stroke and Residual Cardiovascular Risk with Current Secondary Prevention. *Clin. Epidemiol.* 2021, *13*, 813–823. [CrossRef] [PubMed]
- 12. Wong, K.K.; Cummock, J.S.; Li, G.; Ghosh, R.; Xu, P.; Volpi, J.J.; Wong, S.T.C. Automatic Segmentation in Acute Ischemic Stroke: Prognostic Significance of Topological Stroke Volumes on Stroke Outcome. *Stroke* 2022, *53*, 2896–2905. [CrossRef] [PubMed]
- 13. Grefkes, C.; Fink, G.R. Recovery from stroke: Current concepts and future perspectives. *Neurol. Res. Pract.* 2020, 2, 17. [CrossRef] [PubMed]
- Winstein, C.J.; Stein, J.; Arena, R.; Bates, B.; Cherney, L.R.; Cramer, S.C.; Deruyter, F.; Eng, J.J.; Fisher, B.; Harvey, R.L.; et al. Guidelines for adult stroke rehabilitation and recovery: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2016, 47, e98–e169. [CrossRef] [PubMed]
- 15. Campagnini, S.; Liuzzi, P.; Mannini, A.; Basagni, B.; Macchi, C.; Carrozza, M.C.; Cecchi, F. Cross-validation of predictive models for functional recovery after post-stroke rehabilitation. *J. Neuroeng. Rehabil.* **2022**, *19*, 96. [CrossRef]
- 16. Gkantzios, A.; Tsiptsios, D.; Karatzetzou, S.; Kitmeridou, S.; Karapepera, V.; Giannakou, E.; Vlotinou, P.; Aggelousis, N.; Vadikolias, K. Stroke and Emerging Blood Biomarkers: A Clinical Prospective. *Neurol. Int.* **2022**, *14*, 784–803. [CrossRef]
- 17. GBD 2016 Lifetime Risk of Stroke Collaborators. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *N. Engl. J. Med.* **2018**, 379, 2429–2437. [CrossRef] [PubMed]
- 18. Campbell, B.C.V.; Khatri, P. Stroke. Lancet 2020, 396, 129–142. [CrossRef] [PubMed]
- 19. Drozdowska, B.A.; Singh, S.; Quinn, T.J. Thinking about the future: A review of prognostic scales used in acute stroke. *Front. Neurol.* **2019**, *10*, 274. [CrossRef]

- 20. Sung, S.F.; Chen, C.H.; Pan, R.C.; Hu, Y.H.; Jeng, J.S. Natural Language Processing Enhances Prediction of Functional Outcome After Acute Ischemic Stroke. *J. Am. Heart Assoc.* **2021**, *10*, e023486. [CrossRef]
- 21. Winters, C.; Kwakkel, G.; van Wegen, E.E.H.; Nijland, R.H.M.; Veerbeek, J.M.; Meskers, C.G.M. Moving stroke rehabilitation forward: The need to change research. *NeuroRehabilitation* **2018**, *43*, 19–30. [CrossRef] [PubMed]
- Campagnini, S.; Arienti, C.; Patrini, M.; Liuzzi, P.; Mannini, A.; Carrozza, M.C. Machine learning methods for functional recovery prediction and prognosis in post-stroke rehabilitation: A systematic review. J. Neuroeng. Rehabil. 2022, 19, 54. [CrossRef] [PubMed]
- Cho, J.S.; Hu, Z.; Fell, N.; Heath, G.W.; Qayyum, R.; Sartipi, M. Hospital Discharge Disposition of Stroke Patients in Tennessee. South Med. J. 2017, 110, 594–600. [CrossRef]
- 24. Luker, J.A.; Bernhardt, J.; Grimmer, K.A.; Edwards, I. A qualitative exploration of discharge destination as an outcome or a driver of acute stroke care. *BMC Health Serv. Res.* **2014**, *14*, 193. [CrossRef]
- Bacchi, S.; Oakden-Rayner, L.; Menon, D.K.; Jannes, J.; Kleinig, T.; Koblar, S. Stroke prognostication for discharge planning with machine learning: A derivation study. J. Clin. Neurosci. 2020, 79, 100–103. [CrossRef]
- Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* 2019, 110, 12–22. [CrossRef]
- Senders, J.T.; Staples, P.C.; Karhade, A.V.; Zaki, M.M.; Gormley, W.B.; Broekman, M.L.D.; Smith, T.R.; Arnaout, O. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurg.* 2018, 109, 476–486.e1. [CrossRef]
- Lin, W.-Y.; Chen, C.-H.; Tseng, Y.-J.; Tsai, Y.-T.; Chang, C.-Y.; Wang, H.-Y.; Chen, C.-K. Predicting post-stroke Activities of Daily Living through a Machine Learning-Based Approach on Initiating Rehabilitation. Int. J. Med. Inform. 2018, 111, 159–164. [CrossRef]
- Van Os, H.J.A.; Ramos, L.A.; Hilbert, A.; van Leeuwen, M.; van Walderveen, M.A.A.; Kruyt, N.D.; Dippel, D.W.J.; Steyerberg, E.W.; van der Schaaf, I.C.; Lingsma, H.F.; et al. Predicting Outcome of Endovascular Treatment for Acute Ischemic Stroke: Potential Value of Machine Learning Algorithms. *Front. Neurol.* 2018, *9*, 784. [CrossRef]
- Debs, N.; Rasti, P.; Victor, L.; Cho, T.-H.; Frindel, C.; Rousseau, D. Simulated Perfusion MRI Data to Boost Training of Convolutional Neural Networks for Lesion Fate Prediction in Acute Stroke. *Comput. Biol. Med.* 2020, 116, 103579. [CrossRef]
- Fang, G.; Liu, W.; Wang, L. A Machine Learning Approach to Select Features Important to Stroke Prognosis. Comput. Biol. Chem. 2020, 88, 107316. [CrossRef] [PubMed]
- 32. Fang, G.; Huang, Z.; Wang, Z. Predicting Ischemic Stroke Outcome Using Deep Learning Approaches. *Front. Genet.* 2022, 12, 827522. [CrossRef]
- Hofer, I.S.; Burns, M.; Kendale, S.; Wanderer, J.P. Realistically Integrating Machine Learning Into Clinical Practice: A Road Map of Opportunities, Challenges, and a Potential Future. *Anesth. Analg.* 2020, 130, 1115–1118. [CrossRef] [PubMed]
- 34. Verma, A.A.; Murray, J.; Greiner, R.; Cohen, J.P.; Shojania, K.G.; Ghassemi, M.; Straus, S.E.; Pou-Prom, C.; Mamdani, M. Implementing machine learning in medicine. *CMAJ* **2021**, *193*, E1351–E1357. [CrossRef] [PubMed]
- 35. Kokkotis, C.; Moustakidis, S.; Papageorgiou, E.; Giakas, G.; Tsaopoulos, D.E. Machine learning in knee osteoarthritis: A review. *Osteoarthr. Cartil. Open* **2020**, *2*, 100069. [CrossRef]
- 36. Rajpurkar, P.; Chen, E.; Banerjee, O.; Topol, E.J. AI in health and medicine. Nat. Med. 2022, 28, 31–38. [CrossRef]
- Hügle, M.; Omoumi, P.; van Laar, J.M.; Boedecker, J.; Hügle, T. Applied machine learning and artificial intelligence in rheumatology. *Rheumatol. Adv. Pract.* 2020, 4, rkaa005. [CrossRef]
- Heo, J.; Yoon, J.G.; Park, H.; Kim, Y.D.; Nam, H.S.; Heo, J.H. Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. Stroke 2019, 50, 1263–1265. [CrossRef]
- Jang, S.K.; Chang, J.Y.; Lee, J.S.; Lee, E.J.; Kim, Y.H.; Han, J.H.; Chang, D.I.; Cho, H.J.; Cha, J.K.; Yu, K.H.; et al. Reliability and Clinical Utility of Machine Learning to Predict Stroke Prognosis: Comparison with Logistic Regression. *J. Stroke* 2020, 22, 403–406. [CrossRef]
- Lin, C.H.; Hsu, K.C.; Johnson, K.R.; Fann, Y.C.; Tsai, C.H.; Sun, Y.; Lien, L.M.; Chang, W.L.; Chen, P.L.; Lin, C.L.; et al. Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry. *Comput. Methods Programs Biomed.* 2020, 190, 105381. [CrossRef]
- 41. Murray, N.M.; Unberath, M.; Hager, G.D.; Hui, F.K. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review. *J. Neurointerv. Surg.* **2020**, *12*, 156–164. [CrossRef]
- Zhao, Y.; Fu, S.; Bielinski, S.J.; Decker, P.A.; Chamberlain, A.M.; Roger, V.L.; Liu, H.; Larson, N.B. Natural Language Processing and Machine Learning for Identifying Incident Stroke From Electronic Health Records: Algorithm Development and Validation. *J. Med. Internet Res.* 2021, 23, e22951. [CrossRef]
- 43. McDermott, B.; Elahi, A.; Santorelli, A.; O'Halloran, M.; Avery, J.; Porter, E. Multi-frequency symmetry difference electrical impedance tomography with machine learning for human stroke diagnosis. *Physiol. Meas.* **2020**, *41*, 075010. [CrossRef]
- 44. Bivard, A.; Churilov, L.; Parsons, M. Artificial intelligence for decision support in acute stroke—Current roles and potential. *Nat. Rev. Neurol.* **2020**, *16*, 575–585. [CrossRef]
- 45. Wang, W.; Kiik, M.; Peek, N.; Curcin, V.; Marshall, I.J.; Rudd, A.G.; Wang, Y.; Douiri, A.; Wolfe, C.D.; Bray, B. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS ONE* **2020**, *15*, e0234722. [CrossRef]
- 46. Sirsat, M.S.; Fermé, E.; Câmara, J. Machine Learning for Brain Stroke: A Review. J. Stroke Cerebrovasc. Dis. 2020, 29, 105162. [CrossRef]
- 47. Kokkotis, C.; Giarmatzis, G.; Giannakou, E.; Moustakidis, S.; Tsatalas, T.; Tsiptsios, D.; Vadikolias, K.; Aggelousis, N. An Explainable Machine Learning Pipeline for Stroke Prediction on Imbalanced Data. *Diagnostics* **2022**, *12*, 2392. [CrossRef]

- 48. Ding, L.; Liu, C.; Li, Z.; Wang, Y. Incorporating artificial intelligence into stroke care and research. Stroke 2020, 51, e351-e354. [CrossRef]
- 49. Matsumoto, K.; Nohara, Y.; Soejima, H.; Yonehara, T.; Nakashima, N.; Kamouchi, M. Stroke prognostic scores and data-driven prediction of clinical outcomes after acute ischemic stroke. *Stroke* **2020**, *51*, 1477–1483. [CrossRef]
- 50. Harrison, J.K.; McArthur, K.S.; Quinn, T.J. Assessment scales in stroke: Clinimetric and clinical considerations. *Clin. Interv. Aging* 2013, *8*, 201–211. [CrossRef]
- 51. Broderick, J.P.; Adeoye, O.; Elm, J. Evolution of the Modified Rankin Scale and Its Use in Future Stroke Trials. *Stroke* 2017, 48, 2007–2012. [CrossRef]
- 52. Siniscalchi, A. Use of stroke scales in clinical practice: Current concepts. Turk. J. Emerg. Med. 2022, 22, 119–124. [CrossRef]
- 53. Alijanpour, S.; Mostafazdeh-Bora, M.; Ahmadi Ahangar, A. Different Stroke Scales; Which Scale or Scales Should Be Used? *Casp. J. Intern. Med.* **2021**, *12*, 1–21. [CrossRef]
- 54. De Haan, R.; Horn, J.; Limburg, M.; Van Der Meulen, J.; Bossuyt, P. A comparison of five stroke scales with measures of disability, handicap, and quality of life. *Stroke* **1993**, *24*, 1178–1181. [CrossRef] [PubMed]
- Iman, A.N.; Ahmad, T. Improving Intrusion Detection System by Estimating Parameters of Random Forest in Boruta. In Proceedings of the 2020 International Conference on Smart Technology and Applications (ICoSTA), Surabaya, Indonesia, 20 February 2020; pp. 1–6. [CrossRef]
- 56. Poona, N.K.; van Niekerk, A.; Nadel, R.L.; Ismail, R. Random Forest (RF) Wrappers for Waveband Selection and Classification of Hyperspectral Data. *Appl. Spectrosc.* 2016, *70*, 322–333. [CrossRef]
- 57. Fernandez-Lozano, C.; Hervella, P.; Mato-Abad, V.; Rodríguez-Yáñez, M.; Suárez-Garaboa, S.; López-Dequidt, I.; Estany-Gestal, A.; Sobrino, T.; Campos, F.; Castillo, J.; et al. Random forest-based prediction of stroke outcome. *Sci. Rep.* **2021**, *11*, 10071. [CrossRef]
- 58. Yu, D.; Liu, Z.; Su, C.; Han, Y.; Duan, X.; Zhang, R.; Liu, X.; Yang, Y.; Xu, S. Copy number variation in plasma as a tool for lung cancer prediction using Extreme Gradient Boosting (XGBoost) classifier. *Thorac. Cancer* **2020**, *11*, 95–102. [CrossRef]
- 59. Chan, L.C.; Li, H.H.T.; Chan, P.K.; Wen, C. A machine learning-based approach to decipher multi-etiology of knee osteoarthritis onset and deterioration. *Osteoarthr. Cartil. Open* **2021**, *3*, 100135. [CrossRef]
- 60. Mohr, M.; von Tscharner, V.; Emery, C.A.; Nigg, B.M. Classification of gait muscle activation patterns according to knee injury history using a support vector machine approach. *Hum. Mov. Sci.* **2019**, *66*, 335–346. [CrossRef]
- 61. Ali, A.A. Stroke Prediction using Distributed Machine Learning Based on Apache Spark. Stroke 2019, 28, 89–97.
- 62. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing* Systems 30, Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017.
- 63. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]
- 64. Janzing, D.; Minorics, L.; Blöbaum, P. Feature relevance quantification in explainable AI: A causal problem. In Proceedings of the 23th International Conference on Artificial Intelligence and Statistics, PMLR Means Proceedings of Machine Learning Research, Online, 26–28 August 2020.
- 65. Boehme, A.K.; Esenwa, C.; Elkind, M.S. Stroke Risk Factors, Genetics, and Prevention. Circ. Res. 2017, 120, 472–495. [CrossRef]
- 66. Sharrief, A.; Grotta, J.C. Stroke in the elderly. Handb. Clin. Neurol. 2019, 167, 393–418. [CrossRef]
- 67. Porcello Marrone, L.C.; Diogo, L.P.; de Oliveira, F.M.; Trentin, S.; Scalco, R.S.; de Almeida, A.G.; Gutierres Ldel, C.; Marrone, A.C.; da Costa, J.C. Risk factors among stroke subtypes in Brazil. *J. Stroke Cerebrovasc. Dis.* **2013**, *22*, 32–35. [CrossRef]
- 68. Bhaskar, S.; Stanwell, P.; Bivard, A.; Spratt, N.; Walker, R.; Kitsos, G.H.; Parsons, M.W.; Evans, M.; Jordan, L.; Nilsson, M.; et al. The influence of initial stroke severity on mortality, overall functional outcome and in-hospital placement at 90 days following acute ischemic stroke: A tertiary hospital stroke register study. *Neurol. India* 2017, 65, 1252–1259. [CrossRef] [PubMed]
- Fonarow, G.C.; Reeves, M.J.; Zhao, X.; Olson, D.M.; Smith, E.E.; Saver, J.L.; Schwamm, L.H.; Get With the Guidelines-Stroke Steering Committee and Investigators. Age-related differences in characteristics, performance measures, treatment trends, and outcomes in patients with ischemic stroke. *Circulation* 2010, *121*, 879–891. [CrossRef]
- Furlan, N.E.; Luvizutto, G.J.; Hamamoto Filho, P.T.; Zanati Bazan, S.G.; Modolo, G.P.; Ferreira, N.C.; Miranda, L.A.; de Souza, J.T.; Winckler, F.C.; Vidal, E.I.O.; et al. The Impact of Age on Mortality and Disability in Patients With Ischemic Stroke Who Underwent Cerebral Reperfusion Therapy: A Brazilian Cohort Study. *Front. Aging Neurosci.* 2021, 13, 649902. [CrossRef]
- 71. Denes, G.; Semenza, C.; Stoppa, E.; Lis, A. Unilateral spatial neglect and recovery from hemiplegia: A follow-up study. *Brain* **1982**, *105 Pt 3*, 543–552. [CrossRef]
- 72. Yoo, A.J.; Romero, J.; Hakimelahi, R.; Nogueira, R.G.; Rabinov, J.D.; Pryor, J.C.; González, R.G.; Hirsch, J.A.; Schaefer, P.W. Predictors of functional outcome vary by the hemisphere of involvement in major ischemic stroke treated with intra-arterial therapy: A retrospective cohort study. *BMC Neurol.* **2010**, *10*, 25. [CrossRef]
- 73. Etherton, M.R.; Rost, N.S.; Wu, O. Infarct topography and functional outcomes. J. Cereb. Blood Flow Metab. 2018, 38, 1517–1532. [CrossRef]
- 74. Cheng, B.; Forkert, N.D.; Zavaglia, M.; Hilgetag, C.C.; Golsari, A.; Siemonsen, S.; Fiehler, J.; Pedraza, S.; Puig, J.; Cho, T.H.; et al. Influence of stroke infarct location on functional outcome measured by the modified rankin scale. *Stroke* 2014, 45, 1695–1702. [CrossRef] [PubMed]
- 75. Johansson, B.B.; Jadbäck, G.; Norrving, B.; Widner, H.; Wiklund, I. Evaluation of long-term functional status in first-ever stroke patients in a defined population. *Scand J. Rehabil. Med. Suppl.* **1992**, *26*, 105–114. [PubMed]

- 76. Laufer, Y.; Sivan, D.; Schwarzmann, R.; Sprecher, E. Standing balance and functional recovery of patients with right and left hemiparesis in the early stages of rehabilitation. *Neurorehabil. Neural Repair* **2003**, *17*, 207–213. [CrossRef]
- 77. Kalra, L.; Smith, D.H.; Crome, P. Stroke in patients aged over 75 years: Outcome and predictors. *Postgrad. Med. J.* **1993**, 69, 33–36. [CrossRef]
- Ween, J.E.; Alexander, M.P.; D'Esposito, M.; Roberts, M. Factors predictive of stroke outcome in a rehabilitation setting. *Neurology* 1996, 47, 388–392. [CrossRef]
- Wu, O.; Cloonan, L.; Mocking, S.J.; Bouts, M.J.; Copen, W.A.; Cougo-Pinto, P.T.; Fitzpatrick, K.; Kanakis, A.; Schaefer, P.W.; Rosand, J.; et al. Role of Acute Lesion Topography in Initial Ischemic Stroke Severity and Long-Term Functional Outcomes. *Stroke* 2015, 46, 2438–2444. [CrossRef]
- 80. Rangaraju, S.; Streib, C.; Aghaebrahim, A.; Jadhav, A.; Frankel, M.; Jovin, T.G. Relationship Between Lesion Topology and Clinical Outcome in Anterior Circulation Large Vessel Occlusions. *Stroke* 2015, *46*, 1787–1792. [CrossRef] [PubMed]
- Menezes, N.M.; Ay, H.; Zhu, M.W.; Lopez, C.J.; Singhal, A.B.; Karonen, J.O.; Aronen, H.J.; Liu, Y.; Nuutinen, J.; Koroshetz, W.J.; et al. The real estate factor: Quantifying the impact of infarct location on stroke severity. *Stroke* 2007, *38*, 194–197. [CrossRef]
- Königsberg, A.; DeMarco, A.T.; Mayer, C.; Wouters, A.; Schlemm, E.; Ebinger, M.; Cho, T.H.; Endres, M.; Fiebach, J.B.; Fiehler, J.; et al. Influence of stroke infarct location on quality of life assessed in a multivariate lesion-symptom mapping study. *Sci. Rep.* 2021, *11*, 13490. [CrossRef]
- 83. Paolucci, S.; Antonucci, G.; Pratesi, L.; Traballesi, M.; Lubich, S.; Grasso, M.G. Functional outcome in stroke inpatient rehabilitation: Predicting no, low and high response patients. *Cerebrovasc. Dis.* **1998**, *8*, 228–234. [CrossRef]
- 84. Ring, H.; Feder, M.; Schwartz, J.; Samuels, G. Functional measures of first-stroke rehabilitation inpatients: Usefulness of the Functional Independence Measure total score with a clinical rationale. *Arch. Phys. Med. Rehabil.* **1997**, *78*, 630–635. [CrossRef] [PubMed]
- 85. Ernst, M.; Boers, A.M.M.; Forkert, N.D.; Berkhemer, O.A.; Roos, Y.B.; Dippel, D.W.J.; van der Lugt, A.; van Oostenbrugge, R.J.; van Zwam, W.H.; Vettorazzi, E.; et al. Impact of Ischemic Lesion Location on the mRS Score in Patients with Ischemic Stroke: A Voxel-Based Approach. *AJNR Am. J. Neuroradiol.* **2018**, *39*, 1989–1994. [CrossRef] [PubMed]
- 86. de Jonge, J.C.; van de Beek, D.; Lyden, P.; Brady, M.C.; Bath, P.M.; van der Worp, H.B. Temporal Profile of Pneumonia After Stroke. *Stroke* 2022, *53*, 53–60, Erratum in *Stroke* 2022, *53*, e129. [CrossRef]
- Heuschmann, P.U.; Kolominsky-Rabas, P.L.; Misselwitz, B.; Hermanek, P.; Leffmann, C.; Janzen, R.W.; Rother, J.; Buecker-Nott, H.J.; Berger, K. German Stroke Registers Study Group. Predictors of in-hospital mortality and attributable risks of death after ischemic stroke: The German Stroke Registers Study Group. Arch. Intern. Med. 2004, 164, 1761–1768. [CrossRef]
- Sellars, C.; Bowie, L.; Bagg, J.; Sweeney, M.P.; Miller, H.; Tilston, J.; Langhorne, P.; Stott, D.J. Risk factors for chest infection in acute stroke: A prospective cohort study. *Stroke* 2007, *38*, 2284–2291. [CrossRef] [PubMed]
- 89. Gujjar, A.R.; Deibert, E.; Manno, E.M.; Duff, S.; Diringer, M.N. Mechanical ventilation for ischemic stroke and intracerebral hemorrhage: Indications, timing, and outcome. *Neurology* **1998**, *51*, 447–451. [CrossRef]
- 90. Armstrong, J.R.; Mosher, B.D. Aspiration pneumonia after stroke: Intervention and prevention. *Neurohospitalist* 2011, 1, 85–93. [CrossRef]
- Bustamante, A.; Giralt, D.; García-Berrocoso, T.; Rubiera, M.; Álvarez-Sabín, J.; Molina, C.; Serena, J.; Montaner, J. The impact of post-stroke complications on in-hospital mortality depends on stroke severity. *Eur. Stroke J.* 2017, 2, 54–63. [CrossRef]
- 92. Kazi, S.; Siddiqui, M.; Majid, S. Stroke Outcome Prediction Using Admission Nihss In Anterior And Posterior Circulation Stroke. J. Ayub Med. Coll. Abbottabad 2021, 33, 274–278.
- 93. Wouters, A.; Nysten, C.; Thijs, V.; Lemmens, R. Prediction of Outcome in Patients With Acute Ischemic Stroke Based on Initial Severity and Improvement in the First 24 h. *Front. Neurol.* **2018**, *9*, 308. [CrossRef]
- 94. Sablot, D.; Belahsen, F.; Vuillier, F.; Cassarini, J.F.; Decavel, P.; Tatu, L.; Moulin, T.; Medeiros de Bustos, E. Predicting acute ischaemic stroke outcome using clinical and temporal thresholds. *ISRN Neurol.* **2011**, 2011, 354642. [CrossRef] [PubMed]
- Mahdy, M.E.; Ghonimi, N.A.; Elserafy, T.S.; Mahmoud, W. The NIHSS score can predict the outcome of patients with primary intracerebral hemorrhage. *Egypt. J. Neurol. Psychiatry Neurosurg.* 2019, 55, 21. [CrossRef]
- Mazaheri, S.; Reisi, E.; Poorolajal, J.; Ghiasian, M. C-Reactive Protein Levels and Clinical Outcomes in Stroke Patients: A Prospective Cohort Study. Arch. Iran. Med. 2018, 21, 8–12.
- Idicula, T.T.; Brogger, J.; Naess, H.; Waje-Andreassen, U.; Thomassen, L. Admission C—Reactive protein after acute ischemic stroke is associated with stroke severity and mortality: The 'Bergen stroke study'. BMC Neurol. 2009, 9, 18. [CrossRef] [PubMed]
- 98. Bager, J.E.; Hjalmarsson, C.; Manhem, K.; Andersson, B. Acute blood pressure levels and long-term outcome in ischemic stroke. *Brain Behav.* **2018**, *8*, e00992. [CrossRef] [PubMed]
- 99. van den Berg, S.A.; Uniken Venema, S.M.; Mulder, M.J.H.L.; Treurniet, K.M.; Samuels, N.; Lingsma, H.F.; Goldhoorn, R.B.; Jansen, I.G.H.; Coutinho, J.M.; Roozenbeek, B.; et al. Admission Blood Pressure in Relation to Clinical Outcomes and Successful Reperfusion After Endovascular Stroke Treatment. *Stroke* 2020, *51*, 3205–3214. [CrossRef]
- Verschoof, M.A.; Groot, A.E.; Vermeij, J.D.; Westendorp, W.F.; van den Berg, S.A.; Nederkoorn, P.J.; van de Beek, D.; Coutinho, J.M. Association Between Low Blood Pressure and Clinical Outcomes in Patients With Acute Ischemic Stroke. *Stroke* 2020, 51, 338–341. [CrossRef] [PubMed]

- 101. Maïer, B.; Gory, B.; Taylor, G.; Labreuche, J.; Blanc, R.; Obadia, M.; Abrivard, M.; Smajda, S.; Desilles, J.P.; Redjem, H.; et al. Mortality and Disability According to Baseline Blood Pressure in Acute Ischemic Stroke Patients Treated by Thrombectomy: A Collaborative Pooled Analysis. J. Am. Heart Assoc. 2017, 6, e006484, Erratum in J. Am. Heart Assoc. 2017, 6, e004193. [CrossRef]
- 102. Mulder, M.J.H.L.; Ergezen, S.; Lingsma, H.F.; Berkhemer, O.A.; Fransen, P.S.S.; Beumer, D.; van den Berg, L.A.; Lycklama, À.; Nijeholt, G.; Emmer, B.J.; et al. Baseline Blood Pressure Effect on the Benefit and Safety of Intra-Arterial Treatment in MR CLEAN (Multicenter Randomized Clinical Trial of Endovascular Treatment of Acute Ischemic Stroke in the Netherlands). *Stroke* 2017, 48, 1869–1876, Erratum in *Stroke* 2017, 48, e187. [CrossRef]
- 103. Goyal, N.; Tsivgoulis, G.; Iftikhar, S.; Khorchid, Y.; Fawad Ishfaq, M.; Doss, V.T.; Zand, R.; Chang, J.; Alsherbini, K.; Choudhri, A.; et al. Admission systolic blood pressure and outcomes in large vessel occlusion strokes treated with endovascular treatment. *J. Neurointerv. Surg.* 2017, *9*, 451–454. [CrossRef]
- 104. Tziomalos, K.; Giampatzis, V.; Bouziana, S.D.; Spanou, M.; Papadopoulou, M.; Kostaki, S.; Dourliou, V.; Papagianni, M.; Savopoulos, C.; Hatzitolios, A.I. Elevated diastolic but not systolic blood pressure increases mortality risk in hypertensive but not normotensive patients with acute ischemic stroke. Am. J. Hypertens. 2015, 28, 765–771. [CrossRef] [PubMed]
- 105. Lin, M.P.; Ovbiagele, B.; Markovic, D.; Towfighi, A. Systolic blood pressure and mortality after stroke: Too low, no go? *Stroke* 2015, 46, 1307–1313. [CrossRef] [PubMed]
- 106. Cao, Y.; Li, R.; Jiang, S.; Guo, J.; Luo, X.; Miao, J.; Liu, J.; Zheng, B.; Du, J.; Zhang, Y.; et al. The Relationship Between Admission Blood Pressure and Clinical Outcomes for Acute Basilar Artery Occlusion. *Front. Neurosci.* 2022, 16, 900868. [CrossRef] [PubMed]
- 107. Abou-Chebl, A.; Lin, R.; Hussain, M.S.; Jovin, T.G.; Levy, E.I.; Liebeskind, D.S.; Yoo, A.J.; Hsu, D.P.; Rymer, M.M.; Tayal, A.H.; et al. Conscious sedation versus general anesthesia during endovascular therapy for acute anterior circulation stroke: Preliminary results from a retrospective, multicenter study. *Stroke* 2010, *41*, 1175–1179. [CrossRef]
- 108. Jumaa, M.A.; Zhang, F.; Ruiz-Ares, G.; Gelzinis, T.; Malik, A.M.; Aleu, A.; Oakley, J.I.; Jankowitz, B.; Lin, R.; Reddy, V.; et al. Comparison of safety and clinical and radiographic outcomes in endovascular acute stroke therapy for proximal middle cerebral artery occlusion with intubation and general anesthesia versus the nonintubated state. *Stroke* 2010, *41*, 1180–1184. [CrossRef]
- 109. Nichols, C.; Carrozzella, J.; Yeatts, S.; Tomsick, T.; Broderick, J.; Khatri, P. Is periprocedural sedation during acute stroke therapy associated with poorer functional outcomes? *J. Neurointerv. Surg.* **2010**, *2*, 67–70. [CrossRef] [PubMed]
- 110. Takahashi, C.; Liang, C.W.; Liebeskind, D.S.; Hinman, J.D. To Tube or Not to Tube? The Role of Intubation during Stroke Thrombectomy. *Front. Neurol.* **2014**, *5*, 170. [CrossRef]
- 111. Hassan, A.E.; Chaudhry, S.A.; Zacharatos, H.; Khatri, R.; Akbar, U.; Suri, M.F.; Qureshi, A.I. Increased rate of aspiration pneumonia and poor discharge outcome among acute ischemic stroke patients following intubation for endovascular treatment. *Neurocrit. Care* **2012**, *16*, 246–250. [CrossRef]
- 112. Kaur, K.; Kaur, A.; Kaur, A. Erythrocyte Sedimentation Rate: Its Determinants and Relationship with Risk Factors Involved in Ischemic Stroke. *Korean J. Clin. Lab. Sci.* 2022, *54*, 1–8. [CrossRef]
- Jabal, M.S.; Joly, O.; Kallmes, D.; Harston, G.; Rabinstein, A.; Huynh, T.; Brinjikji, W. Interpretable Machine Learning Modeling for Ischemic Stroke Outcome Prediction. *Front. Neurol.* 2022, 13, 884693. [CrossRef]
- 114. Brugnara, G.; Neuberger, U.; Mahmutoglu, M.A.; Foltyn, M.; Herweh, C.; Nagel, S.; Schönenberger, S.; Heiland, S.; Ulfert, C.; Ringleb, P.A.; et al. Multimodal Predictive Modeling of Endovascular Treatment Outcome for Acute Ischemic Stroke Using Machine-Learning. *Stroke* 2020, *51*, 3541–3551. [CrossRef] [PubMed]
- 115. Jiang, B.; Zhu, G.; Xie, Y.; Heit, J.J.; Chen, H.; Li, Y.; Ding, V.; Eskandari, A.; Michel, P.; Zaharchuk, G.; et al. Prediction of Clinical Outcome in Patients with Large-Vessel Acute Ischemic Stroke: Performance of Machine Learning versus SPAN-100. *AJNR Am. J. Neuroradiol.* 2021, 42, 240–246. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.