NOVA IMS Information Management School

MGI

Master Degree Program in

Information Management

Predictive Factors of Major Adverse Cardiovascular Events, Myocardial Infraction and Stroke after Carotid Endarterectomy under logo-regional anesthesia: A Machine Learning Model

Filipa Mendes Afonso

Master Thesis

presented as partial requirement for obtaining the Master Degree in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Predictive Factors of Major Adverse Cardiovascular Events, Myocardial Infraction and Stroke after Carotid Endarterectomy under logo-regional anesthesia: A Machine Learning Model

by

Filipa Mendes Afonso

Master Thesis presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Business Intelligence and Knowledge Management

Supervised by

Nuno António, PhD, Nova Information Management School, Ana Maria Duarte Inácio Marreiros , PhD, Universidade do Algarve

February, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 28 of February of 2025

ACKNOWLEDGEMENTS

Writing this thesis has been a challenging yet incredibly rewarding journey, and I would like to express my deepest gratitude to those who have supported me throughout this process. Their guidance, encouragement, and unwavering support have been instrumental in completing this work.

First and foremost, I extend my sincere gratitude to my advisors, Professor Nuno António and Professor Ana Maria Duarte Inácio Marreiros. Their invaluable expertise and insightful feedback have guided me throughout this research. Their dedication and patience have played a crucial role in shaping my academic growth, and for that, I am truly grateful.

I would also like to thank Dr. João Neves, whose knowledge and guidance have greatly contributed to my understanding and approach to this study. I am especially grateful for the opportunity to be part of this project and contribute to its development, which has been an invaluable learning experience.

To my family, especially my father, João, my mother, Vera, and my siblings, Tiago and Bernardo, thank you for your unconditional love and support. Your encouragement and belief in my abilities have been a constant source of motivation.

Lastly, to my boyfriend, Bernardo, thank you for supporting me throughout this journey. Your patience, encouragement, and endless belief in me have meant more than words can express.

To all of you, I am profoundly grateful. This achievement would not have been possible without your support.

ABSTRACT

Carotid endarterectomy (CEA) is a surgical procedure aimed at reducing the risk of Stroke in patients with Carotid Artery disease. However, postoperative complications, including Major Adverse Cardiovascular Events (MACE), Myocardial Infarction (MI), and Stroke, remain concerns. This study focuses on developing a Machine Learning (ML) predictive model to identify pre- and post-surgery factors influencing the occurrence of these adverse events after CEA under loco-regional anesthesia. Data from patients who underwent CEA under loco-regional anesthesia at the Centro Hospitalar de São João between January 2012 and January 2018 was used. Two predictive models were developed: one using only pre-surgery factors and another exclusively considering post-surgery factors. The classification models evaluated included K-Nearest Neighbors (kNN), Naive Bayes, Logistic Regression, Random Forest, Extreme Gradient Boosting (XGBoost), Decision Trees, Support Vector Machine (SVM), and Neural Networks and feature importance evaluation was performed using Shapley Additive Explanations (SHAP). For the preoperative model, Random Forest achieved the highest predictive performance with an accuracy of 65%, followed by XGBoost with 60%. In the post-surgery model, XGBoost, Random Forest, and Logistic Regression demonstrated the best results, each reaching an accuracy of 62.5%. The strongest predictors factors identified were the number of days since surgery and the patient age. The results demonstrated that the more days that pass since surgery and the older the patient the higher is the risk of adverse events. In pre-surgery model body mass index and the presence of chronic kidney disease were found to be highly relevant, while in the post-surgery side, the occurrence of a stroke 30 days after the surgery emerged as a crucial factor.

KEYWORDS

Carotid Endarterectomy; Carotid clamping; Machine Learning; Outcomes; Stroke; Myocardial Infarction; Major Adverse Cardiovascular Events; Risk factors; Loco-regional anesthesia

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ABIM American Board of Internal Medicine. Organization that certifies internists and subspecialists who demonstrate the knowledge and skills essential for patient care in internal medicine.

ACSS Administração Central do Sistema de Saúde. Public Portuguese entity responsible for managing, coordinating, and developing the National Health Service, ensuring the efficiency, quality, and sustainability of health services in Portugal.

AdaBoost Adaptive Boosting. Ensemble Machine Learning algorithm that combines the outputs of weak learners to create a strong classifier by focusing on the errors of previous models and adjusting their weight accordingly.

ANOVA Analysis of Variance. Statistical method used to compare the means of three or more groups to determine if there are any statistically significant differences among them by analyzing variance within and between groups.

AUC-ROC Area Under the Receiver Operating Characteristic Curve. Performance metric for classification models that quantify the ability to distinguish between positive and negative classes.

BMI Body Mass Index. Numerical value derived from an individual's weight and height that is used to categorize them as underweight, normal weight, overweight, or obese.

CAS Carotid Artery Stenting. Minimally invasive procedure in which a stent is placed in the Carotid Artery to widen it and improve blood flow.

CEA Carotid Endarterectomy. Surgical procedure to remove atherosclerotic plaques from Carotid Arteries.

COPD Chronic Obstructive Pulmonary Disease. Progressive respiratory condition characterized by persistent airflow limitation.

CRISP-DM Cross-Industry Standard Process for Data Mining. Structured methodology that outlines the steps necessary for successful data mining projects, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment.

DBP Diastolic Blood Pressure. Pressure exerted by the blood on the walls of the arteries when the heart is at rest between beats.

DM Diabetes Mellitus. Chronic medical condition characterized by high levels of glucose in the blood due to the body's inability to produce enough insulin or effectively use the insulin it produces.

DT Decision Trees. Machine Learning algorithm that uses a tree-like model of decisions and their possible consequences to make predictions or classifications based on input data.

ECST European Carotid Surgery Trial. Clinical study that demonstrated that Carotid endarterectomy significantly reduces the risk of ischemic stroke and stroke-related death in patients with symptomatic carotid artery stenosis.

IND Intraoperative Neurological Deficits. Adverse neurological events occur during Surgery.

IQR Interquartile range. Statistical measure of variability that represents the range within which the central 50% of the data points lie, calculated as the difference between the first quartile (Q1) and the third quartile (Q3).

KNN K-Nearest Neighbors. Non-parametric Machine Learning algorithm used for classification and regression, where a data point is classified based on the majority class among its k-nearest neighbors in the feature space.

MACE Major Adverse Cardiovascular Events. Combined outcome that includes Mortality, Myocardial Infarction and Stroke.

MBP Mean Arterial Pressure. Average pressure in arteries during one cardiac cycle, providing an indicator of overall blood flow and tissue perfusion.

MI Myocardial infarction. Cardiac adverse event characterized by blockage of blood supply.

ML Machine Learning. Subset of Artificial Intelligence where computers are trained to recognize patterns and make decisions based on data, improving their performance over time**.**

NASCET North American Symptomatic Carotid Endarterectomy Trial. Clinical study that demonstrated the benefits of Carotid Endarterectomy in reducing the risk of Stroke and Stroke-related death in patients with symptomatic Carotid Artery stenosis.

NSQIP National Surgical Quality Improvement Program. Program initiated by the American College of Surgeons to measure and improve the quality of surgical care by collecting and analyzing clinical data from participating hospitals.

RFE Recursive Feature Elimination. Feature selection method that iteratively fits a model and removes the least important features to improve model performance by selecting the most relevant features for prediction.

RS Random Search. Hyperparameter optimization technique that selects hyperparameter values randomly from a predefined distribution to efficiently explore the search space.

SBP Systolic Blood Pressure. Pressure exerted by the blood on the walls of the arteries when the heart contracts and pumps blood into the circulatory system.

SHAP Shapley Additive Explanations. A method that explains the results of Machine Learning models by assigning a value to the contribution of each feature based on its input to the model's predictions.

SMOTE Synthetic Minority Oversampling Technique. Resampling method that generates synthetic samples for the minority class by interpolating between existing data points, helping to address class imbalance in Machine Learning.

SVM Support Vector Machine. Supervised Machine Learning algorithm that finds the optimal hyperplane to separate data points of different classes in a high-dimensional space, maximizing the margin between the closest points of each class.

U-CEA Urgent Carotid Endarterectomy. Surgical procedure to remove atherosclerotic plaques from Carotid Arteries, performed promptly.

VIF Variance Inflation Factor. Measure used to detect multicollinearity in regression analysis by quantifying how much the variance of an estimated regression coefficient increases when your predictors are correlated.

XGBoost Extreme Gradient Boosting. Scalable Machine Learning algorithm that uses an ensemble of weak decision trees and gradient boosting techniques to improve prediction accuracy and model performance.

1. INTRODUCTION

Carotid Endarterectomy (CEA) is a surgical procedure designed to remove atherosclerotic plaques from the Carotid Arteries. Introduced in the 1950s (Saha et al., 2014), its purpose was to prevent Strokes and has remained the gold standard in the treatment of Carotid Artery disease since then (Pereira-Neves et al., 2021). Nowadays, this disease persists, affecting 3% of the world's population (Dossabhoy & Arya, 2021), and is responsible for 15-20% of recorded Strokes, one of the leading global causes of death (Joh &Cho, 2020).

In the Carotid Arteries, atherosclerotic plaques most commonly stand in the carotid bifurcation region, where the Common Carotid Artery branches into the Internal and External Carotid Arteries (Allain et al., 2005). Composed of fat and calcium, these plaques narrow the arteries, reducing blood flow and increasing the risk of Stroke (Bir & Kelley, 2022). Greater artery narrowing (degree of stenosis) indicates a higher risk of Stroke development.

The disease in the Carotid Arteries often remains silent/asymptomatic for extended periods. However, the worsening increased plaque/stenosis severity can rapidly make it symptomatic, manifesting as cerebral infarction or cerebral ischemia (Song et al., 2020). Although CEA has shown better results in symptomatic patients, it can also be performed in asymptomatic patients who meet specific criteria. According to Pinheiro (2010), eligible patients have stenosis over 50% in symptomatic cases without severe neurological deficits or stenosis over 60% in asymptomatic cases. However, a pre-surgery evaluation is performed to assess the urgency of the procedure and the patient's risk factors, essentially evaluating the cost-benefit of the surgery.

CEA is performed using general or loco-regional anesthesia, involving an incision in the affected Carotid Artery and removal of atherosclerotic plaque (Howell, 2007). Although both types of anesthesia can be used, loco-regional anesthesia offers the advantage of allowing neurological monitoring of the patient during the procedure while they remain awake. Additionally, Rocha-Neves (2021) indicates that this approach is associated with less shunt dependency, faster surgeries, and a lower chance of coronary incidents.

Before initiating the process, it is necessary to halt blood flow in the artery by clamping its three branches, with the brain being exclusively irrigated by the contralateral Carotid Artery during this period. This has been reported as a significant factor in the development of Intraoperative Neurological Deficits (IND), due to potential cerebral hypoperfusion (Rocha-Neves et al., 2021). Furthermore, CEA also presents a risk of 13% of developing major adverse cardiovascular events (MACE), including Stroke, Myocardial Infarction (MI), and cardiovascular death (Mekke et al., 2023). However, the major adverse events that occur after CEA are Stroke and MI, being reported 3-4% and 0-5% of the time, respectively (Bai et al., 2020). Despite factors such as gender (women more prone) and age (higher age, higher risk)

being identified in the literature as ones of the principal contributors to adverse events, there is little consensus on the pre-and post-operative risk factors related to the role of artery clamping in adverse events development (Pereira-Neves et al., 2021; Rocha-Neves et al., 2021; Saha et al., 2010). Establishing the prediction of outcomes after clamping during CEA in different patients with various characteristics or factors becomes imperative.

In recent years, Artificial Intelligence (AI) has been increasingly used in healthcare, particularly in diagnosing accuracy and new treatment methods (Matsuo et al., 2022). One of its domains, Machine Learning (ML), serves as a highly useful tool in outcome prediction, allowing exploration/analysis and defining patterns in extensive datasets (Buchlak et al., 2020).

In today's context, there is an increasing number of studies on CEA and the occurrence of adverse events, such as MACE, MI, and Stroke. Despite this, the literature exploring the contributing factors to these events remains limited, with little consensus on the matter (Lobo et al., 2015). To address this gap, this retrospective study included patients undergoing CEA under loco-regional anesthesia at São João Hospital Center, Porto, Portugal, between January 2012 and January 2018, who experienced or did not MACE, MI, or Stroke after clamping. Based on the recorded characteristics of these patients, a predictive ML model will be created to predict the risk of each one of the outcomes mentioned before occurring after clamping during CEA. Thus, the present dissertation aims to answer the research question: "What are the predictors of pre- and post-adverse events, such as MACE, MI, and Stroke after clamping the Carotid Artery during Carotid Endarterectomy?".

2. LITERATURE REVIEW

Although many studies have been exploring CEA and its outcomes, the search on the factors that trigger adverse events after this surgery, namely MACE, MI and Stroke, is still on going. The lack of agreement represents one of the main challenges in predicting these postoperative outcomes, making it difficult to identify patients at greater risk and to implement more effective preventive strategies.

Most studies have focused primarily on the outcomes of CEA without delving into the underlying factors. For example, the renowned studies conducted by North American Symptomatic Carotid Endarterectomy Trial (NASCET) and European Carotid Surgery Trial (ECST) in the 1990s evaluated the efficacy and results of CEA in patients with Carotid stenosis. NASCET, conducted in North America, focused on symptomatic patients, while ECST, conducted in Europe, included both symptomatic and asymptomatic patients. Both studies compared the efficacy of Endarterectomy with conservative clinical treatment, finding a significant reduction in the risk of cerebrovascular events in patients undergoing surgery (Ferguson et al., 1999; "Randomised Trial of Endarterectomy for Recently Symptomatic Carotid Stenosis: Final Results of the MRC European Carotid Surgery Trial (ECST)," 1998).

Building on these foundational studies, Rockman (2000) aimed to analyze neurological events post-CEA to identify which patients could benefit from further surgery. They identified causes such as clamping ischemia (13.2%), intracerebral hemorrhage (13.2%), unrelated factors (10.5%), and thromboembolic events (63.2%), with thromboembolic events being predominant within the first 24 hours post-surgery.

Years later, Bekelis (2013) developed a predictive model based on risk factors to predict 30-day postoperative outcomes such as Stroke, MI, and Death that included 35698 patients. Their results highlighted significant differences in the prevalence of risk factors like gender, chronic obstructive pulmonary disease (COPD), history of peripheral vascular disease, congestive heart failure, previous coronary angioplasty, coronary artery bypass graft surgery, history of bleeding disorders, and increased body mass index (BMI). The incidences 30 days after CEA were 1.64% for Stroke, 0.69% for MI, 0.75% for death, and 2.79% for MACE (Figure 2.1).

Table 2.1 – Outcomes registered 30 days after CEA. Adapted from Bekelis et al (2023)

	All Patients	Symptomatic	Asymptomatic	*P* Value
Stroke	585 (1.64%)	365 (2.33%)	220 (1.10%)	<0.001
MI	247 (0.69%)	122 (0.78%)	125 (0.63%)	0.09
Death	267 (0.75%)	163 (1.04%)	104 (0.52%)	<0.001
Composite	994 (2.79%)	581 (3.70%)	413 (2.06%)	<0.001

Barbetta (2014) assessed the outcomes of Urgent Carotid Endarterectomy (U-CEA) in symptomatic patients, noting a higher rate of Stroke (70%) and mortality (4.4%) in patients with acute symptoms compared to those undergoing elective surgery.
Dua (2016) studied factors impacting outcomes in CEA and Carotid Artery Stenting (CAS) procedures, highlighting specific patient factors like gender and procedural volume that significantly influence risks and outcomes.
In the same year, Chaudhry (2016) developed a risk index for CEA patients to predict in-hospital mortality and other outcomes. They identified factors such as age over 70, atrial fibrillation, congestive heart failure, smoking, symptomatic status, and chronic renal failure, which were linked to an increased rate of stroke and/or death in CEA patients.
Volkers (2018) conducted a systematic literature review of studies developing prediction models or risk scores after Carotid revascularization to further enrich the understanding of risk factors. They identified common predictors like age, diabetes, heart failure, and contralateral carotid stenosis.

More recently, Dasenbrock (2019) developed a predictive scale using data from the National Surgical Quality Improvement Program (NSQIP) registry from 2011 to 2015. This scale was based on 14 variables, such as insulin-dependent diabetes and high-risk cardiac physiological characteristics. It predicted adverse events with a concordance of 0.67 for general events, 0.69 for stroke or death, 0.76 for mortality, 0.73 for prolonged hospitalization, and 0.83 for non-routine discharge. This scale is effective in predicting adverse events after CEA and can stratify the risk of patients with symptomatic and asymptomatic Carotid stenosis. It serves as a valuable tool in risk stratification and patient counseling, aiding in better clinical decision-making.

Building on traditional analyses, Feng (2020) aimed to assess clinical outcomes and identify risk factors for MACEs, such as Stroke, MI, and death. Their study showed a MACE rate of 4.6%, with 4.02% of Strokes, 0.28% of MI and death (Figure 2.2). Univariate analysis revealed that factors such as symptomatic injury, diabetes mellitus, and elevated diastolic blood pressure (DBP) were significantly associated with the occurrence of MACEs.

Table 2.2 – Outcomes. Adapted from Feng et al. (2020)

Complications	n (%)
MACEs	16 (4.60)
Death	1 (0.28)
Stroke	14 (4.02)
Ischemic stroke	8 (2.30)
Hemorrhagic stroke	6 (1.72)
Composite	1 (0.28)

In addition, Bai (2020) explored the application of ML methods to identify factors related to early Cerebral and Myocardial Infarction after CEA. They collected extensive perioperative data for 443 CEA patients, registering 6 cases of Cerebral Infarction (1.4%) and 10 cases of MI (2.3%). The algorithm with the highest accuracy was XGBoost (Figure 2.3).

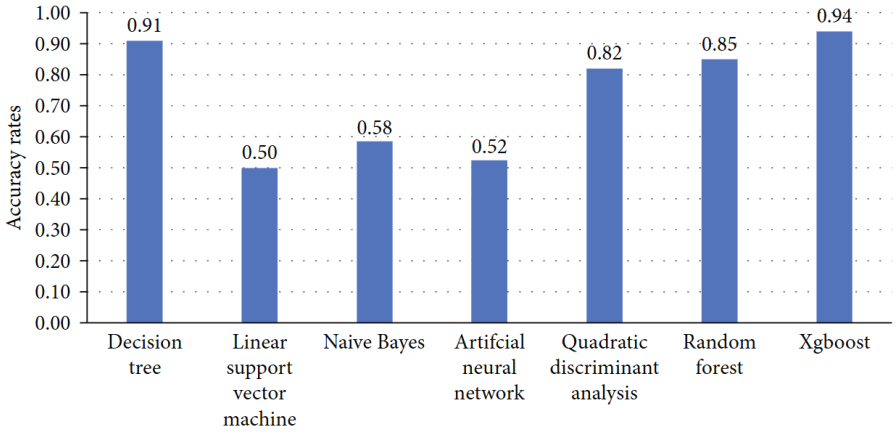

Figure 2.1 – Algorithms prediction accuracy from Bai et al (2020)

Based on the importance score, the study identified eight factors as the most influent on the outcomes, mean arterial pressure during occlusion (T4 MBP), BMI, mean arterial pressure after the operation (T6 MBP), the standard deviation of systolic pressure during occlusion (T4 SBP standard deviation), diastolic pressure during occlusion (T4 DBP), mean arterial pressure in the operating room (T2 MBP), systolic pressure during occlusion (T4 SBP), and age (Figure 2.4).

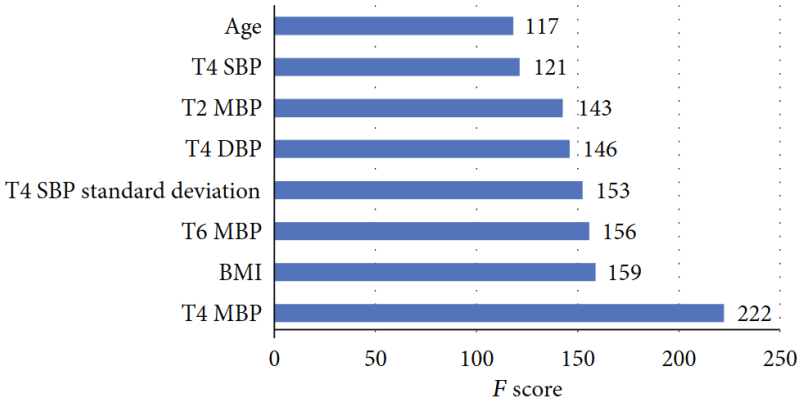

Figure 2.2 – Principal factors of the outcome prediction from Bai et al. (2020)

Finally, Li (2023) performed a systematic review focusing on the impact of diabetes mellitus (DM) on patients undergoing CEA. Including 19 studies with a total of 122,003 patients, they found that DM is associated with an increased risk of both short-term and long-term MACE,

death/Stroke, death, and MI. The study highlighted that patients with DM, especially those with insulin-dependent DM, have higher risks of adverse outcomes after CEA.

There is a significant body of literature exploring the CEA and its outcomes throughout these studies. However, there is still a notable gap in the understanding of adverse events and their predictive factors, mainly due to the lack of consensus around them and the time until their occurrence. Table 2.3 summarizes all the articles discussed here, highlighting their key findings and contributions.

A search was conducted using the PubMed and Google Scholar databases to find articles on predictive models for detecting factors associated with the occurrence of stroke, MI, and MACE after CEA. The following keywords were used: "Carotid Endarterectomy", "Carotid clamping", "Machine Learning", "Outcomes", "Stroke", "Myocardial Infarction", "MACE", "Risk factors", and "Loco-regional anesthesia".

Table 2.3 – Literature Review - Overview

Author	Objective	Methods	Results	Conclusions
European Carotid Surgery Trialists' Collaborative Group (1998)	Evaluate the advantages and disadvantages of CEA, particularly focusing on its effectiveness in preventing Strokes, in individuals with recently symptomatic Carotid stenosis.	3024 patients were involved in the trial, including men and women of any age, with some degree of Carotid stenosis, who had experienced at least one transient or mild symptomatic ischemic vascular event in the previous 6 months, affecting one or both Carotid Arteries. The control group comprised 40% of patients.	No significant difference in the overall occurrence of major Stroke or death between the surgery and control groups. The risk of major Stroke on the same side as the untreated symptomatic Carotid Artery increased with stenosis severity, but only for the first 2-3 years post-randomization. For stenosis exceeding 80%, immediate surgery reduces the long-term Stroke risk. At 3 years, major Stroke or death occurred in 26.5% of the control group compared to 14.9% of the surgery group. Age and sex variations influenced the patient selection for surgery.	CEA is indicated for patients with symptomatic stenosis greater than 80%. Age and gender are important factors when deciding whether to operate.
Ferguson et al. (1999)	Report surgical outcomes in patients who underwent CEA in NASCET.	The study involved 1415 patients and calculated the rates of perioperative Stroke and death at 30 days and the final assessment of Stroke severity at 90 days, identified the variables that increased or decreased perioperative risk using a regression model, documented non-resultant surgical complications and analyzed the durability of CEA.	The study recorded 92 perioperative events. At 30 days, the rates were: death 1.1%, disabling Stroke 1.8%, and non-disabling Stroke 3.7%. At 90 days, death remained at 1.1%, disabling Stroke decreased to 0.9%, and non-disabling Stroke increased to 4.5%. Thromboembolism was the primary cause of most Strokes. A history of Coronary Artery disease with previous cardiac procedures reduced risk. There were 30 intraoperative and 62 postoperative events. Eight years post-surgery, the risk of disabling ipsilateral Stroke was 5.7%, and the overall risk of any ipsilateral Stroke was 17.1%.	While the combined rate of perioperative Stroke and death stood at 6.5%, the incidence of permanently disabling Stroke and death was notably lower at 2.0%. Notably, CEA was demonstrated to be an effective procedure.

Rockman et al. (2000)	Investigate perioperative neurological events and assess whether a detailed examination of their timing and underlying causes could identify cases that may benefit from reoperation.	A review of 38 patients (1.9%) who manifested IND during the perioperative period in 2024 patients who underwent CEAs between 1985 and 1997. Statistical analysis was carried out using the SPSS program, and results with a $P < 0.05$ were considered significant, using tests such as the $\chi 2$ or Fisher's exact test for comparisons.	IND had various causes: clamping ischemia (13.2%), thromboembolic events (63.2%), intracerebral hemorrhage (13.2%), and unrelated factors (10.5%). Those occurring within 24 hours were mostly due to thromboembolic events (88.0%), while later deficits were linked to other causes. Immediate exploration helped resolve or notably improve deficits in 66.7% of cases where intraluminal thrombus was noted.	Deficits within 24 hours are likely due to intraluminal thrombus, warranting immediate reexploration to remove embolic sources and correct technical issues, potentially improving neurologic outcomes by preventing ongoing embolization.
Bekelis et al. (2013)	Develop a risk factor-based predictive model for outcomes in CEA by analyzing patient characteristics and identifying independent risk factors associated with postoperative Stroke, MI, or death.	The study conducted a retrospective cohort analysis of patients who underwent CEA between 2005 and 2010, using data from the American College of Surgeons National Quality Improvement Project database. Multivariate analysis was performed to identify independent risk factors associated with postoperative Stroke, Myocardial Infarction, or death.	The study analyzed 35,698 patients, of whom 56.1% were asymptomatic and 43.9% were symptomatic. The observed risks within 30 days after CEA were 1.64% for Stroke, 0.69% for MI, and 0.75% for death. Multivariate analysis identified factors such as advanced age, male sex, and a history of cardiovascular and pulmonary conditions as independent predictors of adverse outcomes. A predictive model was developed, highlighting the significant impact of age on the risk of Myocardial Infarction and death.	The risks associated with CEA vary based on patient-specific characteristics. The model is not suitable for individual risk assessment but serves as a baseline for future improvements and the development of more accurate predictive models.

Barbetta et al. (2014)	Evaluate the results of U-CEA procedures performed on acutely symptomatic patients using a straightforward selection protocol.	From January 2002 to January 2012, in a total of 193 patients, 90 had acute symptoms and underwent U-CAE, where 27 registered transient ischemic attacks, 52 had Mild to Moderate strokes and 11 had a Stroke in evolution. The statistical analysis involved descriptive, inferential, and significant testing methods to evaluate data.	Acute patients had higher Stroke rates and more Coronary Artery disease history compared to elective patients. U-CEAs performed within 48 hours from symptom onset had a lower postoperative Stroke rate (4.4% vs 8.8%; $P < .3$). Among those with Strokes, NIH Stroke Scale assessment at discharge showed improvement in 79% (25% with significant reduction). Patients with moderate Strokes improved slightly better than those with mild Strokes ($P < .001$).	U-CEA shows a higher risk compared to elective surgery. U-CEA is safe for patients with transient ischemic attacks. Acceptable outcomes were seen in minor to moderate Strokes.
Dua A. et al. (2016)	Identify factors related to patients and hospitals that negatively influence outcomes.	Between 1998 and 2012, 1,756,445 patients who underwent CEA or CAS were identified from healthcare databases. A regression analysis was conducted to determine odds ratios. A subgroup analysis matched asymptomatic and symptomatic patients based on Chalson scores in a 1:1 risk-matched comparison.	Symptomatic disease led to worse outcomes for both CEA and CAS. In risk-matched asymptomatic and symptomatic patients, female gender ($P < .001$) and performing 1-2 cases annually ($P < .05$) were associated with higher Stroke risk. Predictors of MI included congestive heart failure and peripheral artery disease ($P < .05$), while bleeding predictors included peripheral artery disease and COPD for symptomatic patients.	Patients opting for CEA consistently demonstrate better outcomes compared to matched groups undergoing CAS.
Chaudhry et al. (2016)	Develop and validate a new index to provide risk adjustment and to predict in-hospital patient outcomes undergoing CEA.	Using Nationwide Inpatient Sample data, the study investigated Stroke, cardiac complications, and mortality rates before CEA. Multivariate logistic regression was applied. A risk index score was developed.	Several predictors, including age, atrial fibrillation, congestive heart failure, smoking, symptomatic status, and chronic renal failure, were identified. Validation in a separate cohort showed that patients with higher risk index scores had higher rates of the composite endpoint. The risk index had a receiver operating characteristic curve of 68.5%.	A risk index method in comparative studies for CAS placement versus CEA would facilitate more informed decision and improve the quality of the services.

Volkers et al. (2018)	Identify and summarize existing prediction models used to forecast outcomes following CAS or CEA in patients with Carotid stenosis.	Systematic literature review of studies that developed a prediction model or risk score published until 22 December 2016. Eligible prediction models had to predict the risk of vascular events with at least one patient characteristic.	37 studies developing 46 prediction models were found, mainly for CEA patients predicting short-term risks like Stroke or death. Age, diabetes, heart failure, and Carotid stenosis were common predictors. Around 61% provided full risk formulas, 46% were internally validated and 26% were externally validated. Model performance ranged from 0.66 to 0.94 for CAS and from 0.58 to 0.74 for CEA. External validations showed c-statistic ranges from 0.55 to 0.	Incomplete reporting, methodological constraints, and the absence of external validation obscure the clinical value of many prediction models for outcomes following CAS or CEA.
Dasenbrock et al. (2019)	Develop an integer-weighted predictive scale for adverse events after CEA to enhance risk stratification and improve patient counseling.	Using data from the NSQIP registry, the study applied multivariable logistic regression to identify predictors of 30-day adverse events. A weighted predictive scale was constructed based on effect estimates and validated in a separate population.	The derivation and validation cohorts present adverse event rates of 4.0% and 3.7%, respectively. The final CEA risk scale assigned the highest risk points to factors such as insulin-dependent diabetes and high-risk cardiac conditions. The model demonstrated a concordance of 0.69 for Stroke or death, and 0.76 for mortality.	The NSQIP CEA risk scale effectively predicts post-CEA adverse outcomes and allows for risk stratification of patients, supporting personalized clinical decision-making.
Feng et al. (2020)	Evaluate the 30-day outcomes and risk factors for MACEs in elderly patients undergoing CEA through a retrospective single-center analysis.	Patients who underwent CEA at Xuanwu Hospital, Capital Medical University, between 2001 and 2017 were analyzed. The primary outcome was the incidence of MACEs within 30 days post-surgery. Univariable and multivariable analyses were conducted.	Among the 348 elderly patients studied, the 30-day incidence of MACEs was 4.6%. Univariate analysis identified symptomatic lesions, DM, and higher DBP as significant risk factors. Multivariable analysis confirmed DM (OR = 2.882, p = 0.049) and higher DBP (OR = 1.079, p = 0.004) as independent predictors of MACEs, while symptomatic lesions were not found to be an independent risk factor.	CEA can be safely performed in elderly patients, but those with diabetes and higher diastolic blood pressure are at increased risk of adverse events.

Bai P. et al. (2020)	Aims to use ML techniques to analyze factors associated with early Cerebral Infarction and MI following CEA.	443 patients undergoing CEA under general anesthesia were included. It collected demographic data, medical history, vascular stenosis degree, perioperative blood pressure, occlusion duration, shunt placement, hospital stay, and occurrence of Cerebral and Myocardial Infarction. An ML was built, selecting variables through single-factor analysis.	The incidence of Cerebral Infarction was 1.4% (6/443), and MI was 2.3% (10/443). Boost identified stable factors: average arterial pressure during occlusion, body mass index, postoperative arterial pressure, systolic pressure variability during occlusion, diastolic pressure during occlusion, mean arterial pressure after entry, systolic pressure during occlusion, and age, in descending order of importance.	Blood pressure, BMI, and age, among other factors, could potentially link to postoperative Cerebral and MI patients undergoing CEA. Further investigation is needed into these factors using ML methods.
Li et al. (2023)	Evaluate the impact of diabetes mellitus on adverse outcomes in patients undergoing CEA for carotid stenosis.	A systematic review and meta-analysis were conducted using 19 studies (n = 122,003) published between 2000 and 2023 from multiple databases. The study analyzed short-term and long-term MACEs, including death, Stroke, MI, and composite outcomes. Subgroup analyses were performed based on symptomatic/asymptomatic carotid stenosis and insulin/noninsulin-dependent DM.	DM was associated with an increased risk of short-term and long-term adverse outcomes following CEA. In the short term, DM significantly raised the risk of MACEs, death or Stroke, Stroke alone, death, and MI. Regarding long-term outcomes, DM was linked to a higher risk of MACEs. Subgroup analysis revealed that asymptomatic patients with DM had an elevated risk of short-term MACEs, death/Stroke, Stroke, and MI. In contrast, symptomatic patients with DM exhibited an increased risk only for short-term MACEs. Additionally, insulin-dependent DM was associated with a greater impact on adverse outcomes, including a higher risk of death, Stroke, and MI, compared to noninsulin-dependent DM.	DM is a significant risk factor for both short-term and long-term adverse events after CEA, with a greater impact on asymptomatic patients. Insulin-dependent DM presents a higher risk than noninsulin-dependent DM.

3. METHODOLOGY

This chapter describes the materials and methods used to conduct the study and was divided into Study Population and Machine Learning Model Construction, where Python programming language and Pandas software were used to build the models, manipulate, and analyze the data.

3.1 Study Population

The initial dataset comprises 228 patients who underwent CEA under loco-regional anesthesia at the Centro Hospitalar de São João between January 2012 and January 2018, who presented or not, adverse events after carotid clamping, and 191 variables related to patient data, surgery data, and hospital information (described in Table 3.1). The study protocol and the guarantee of data confidentiality/protection were presented to and approved by the Ethics Committee of Centro Hospitalar de São João/Faculdade de Medicina da Universidade do Porto.

3.2 Machine Learning Model Construction

The construction of the ML models based on the dataset available followed the Cross-Industry Standard Process for Data Mining (CRISP-DM). This type of methodology consists of six iterative phases (Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment), that help to plan and execute a well-organized data mining process (Schröer et al., 2021).

Since the study's purpose is to identify the pre- and post-predictive factors of Stroke, MI, and MACE after CEA, two models were created: one combining only variables considered before surgery (Model 1) and another with post-surgery variables (Model 2). The creation of the two models aimed to reduce leakage and better understand the results.

Table 3.1 – Variables description and type

Variable	Description	Type
Número de processo (ND)	ID	Pre and Post
Dadoshemodinamica	Hemodynamic monitoring data	-
Data de Nascimento	Birth Date	Pre and Post
Data da Cirurgia	Surgery Date	Pre and Post
AVCposop30D	Stroke 30 days after Surgery	Post
Tissue Stroke - 0; Watershed stroke 1 - Pos	Type of Stroke	Post
TC-data	CT scan Date	Pre

Complicações30D	Outcome 30 days after Surgery	Post
ComplicaçõesMajorCod30D	Major complications 30 days after Surgery	Post
ComplicaçõesMinor	Minor complications 30 days after Surgery	Post
Complicações pós-operatórias	Complications after Surgery	Post
Défice inta-opCod	Intraoperative Neurological Deficit Code	Pre
Timing Afundamento (0-sem afundamento-; 1-primeiros minutos/precoce 3; 2 afundamento tardio 3 a 10; 3 afundamento tardio depois de 10)	Time it took to manifest symptoms after Carotid Clamping (Minutes)	Pre
Défice intra-op	Manifested signs of Cerebral Ischemia	Pre
Anestesia	Type of Anesthesia	Pre
AnestesiaCod	Anesthesia type Code	Pre
MétodoAvaliaçãoHipoIntra	Type of Intraoperative Cerebral Hypoperfusion assessment method used	Pre
Shunt	Shunt utilization	Pre
Lateralidade	Laterality	Pre
LateralidadeCod	Laterality Code	Pre
Doença circulação posterior	Presence of Posterior Circulation Disease	Pre
Reestenose pós-operatória	Recurrence of Stenosis after Surgery	Post
Reestenosecod30	Restenosis of 30%	Post
RestenoseCod50	Restenosis of 50%	Post
RestenoseCod70	Restenosis of 70%	Post
PCR	C-Reactive Protein values	Pre
trop I	Troponin I Test value	Pre
trop I alta senssibilidade	High sensitivity Troponin I Test value	Pre
MINS	Myocardial Injury after non-Cardiac Surgery	Pre
UPA/UCI – posCEA	Unit days after Surgery	Post
Internamento – PosCEA	Hospitalization days after Surgery	Post
FC pre op	Heart rate before Surgery	Pre
FC pos op	Heart rate after Surgery	Post
ECG	Presence of ECG alterations	Pre
Arritmia	Presence of Arrhythmia on ECG	Pre
Bloqueios ramo	Presence of Branch Block	Pre
Hemiplegia	Paresis Presence	Pre
alcoolismo	Alcoholic/non-Alcoholic patient	Pre

LeucócitosPosOp	Number of Leukocytes after Surgery	Post
hb pos op	Hemoglobin value after Surgery	Post
RDW Cvpos op	Index of the variability of Erythrocyte size as variation coefficient	Post
RDW Sdposop	Index of the variability of Erythrocyte size as standard deviation	Post
Linf%posop	Percentage of Lymphocytes after Surgery	Post
Linfabsposop	Number of Lymphocytes after Surgery	Post
Neut%posop	Percentage of Neutrophils after Surgery	Post
Neutposop	Number of Neutrophils after Surgery	Post
eosinofilos%posop	Percentage of Eosinophils after Surgery	Post
eosinofilosposop	Number of Eosinophils after Surgery	Post
monocitos%posop	Percentage of Monocytes after Surgery	Post
monocitosposop	Number of Monocytes after Surgery	Post
plaquetasposop	Number of Platelets after Surgery	Post
MPVposop	Average Platelet Volume after Surgery	Post
pdw_paletetPosOp	Platelet Distribution Width after Surgery	Post
CreatPos	Creatine after Surgery	Post
UreiaPos	Urea after Surgery	Post
RRT30D	Dialysis realization after Surgery	Post
Colesterol total	Cholesterol before Surgery	Pre
HDL	High-Density Lipoprotein Cholesterol before Surgery	Pre
LDL	Low-Density Lipoprotein Cholesterol before Surgery	Pre
Triglicerideos	Triglyceride before Surgery	Pre
intolerânciaglicose	Glucose intolerance	Pre
Insulina1; ADO 2	Type of Drug Treatment for Diabetes	Pre
HBA1c	Blood Test values for diagnosing Diabetes Type 2	Pre
BNP pre op	Level of B-type Natriuretic Peptide in Blood	Pre
AAS	Aspirin medication	Pre
Clopidogrel-preop	Clopidogrel medication	Pre
peso	Patient Weight (Kg)	Pre

altura	Patient Height (Cm)	Pre
Bd_mass_ind	Body Mass Index (Kg/m2)	Pre
calc_adb_perim	Abdominal Perimeter Calculator (Cm)	Pre
MetSyn	Metabolic Syndrome	-
INVOSprelpsilat	Ipsilateral oximetry measured by INVOS system	-
INVOSpreContra	Contralateral oximetry measured by the INVOS system	-
INVOSposclamplpsi	Ipsilateral oximetry after clamping measured by the INVOS system	-
INVOSposClampContra	Contralateral oximetry after clamping measured by the INVOS system	-
Subida com shunt	Oximetry values with shunt utilization	-
Pos desclampagem Ipsi	Ipsilateral oximetry values after de-clamping	-
Hemoglobina	Hemoglobin before Surgery	Pre
Eritrocitos pre op	Number of Erythrocytes before Surgery	Pre
Hct pre op	Hematocrit before Surgery	Pre
MCHC	Mean Corpuscular Hemoglobin Concentration before Surgery	Pre
RDW-CV	Index of the variability of Erythrocyte size as variation coefficient before Surgery	Pre
RDW-SD	Index of the variability of Erythrocyte size as standard deviation	Pre
Leucócitos	Number of Leukocytes before Surgery	Pre
Linfócitos %	Percentage of Lymphocytes before Surgery	Pre
Linf abs	Absolute Number of Lymphocytes before Surgery	Pre
Neutrofilos %	Percentage of Neutrophils before Surgery	Pre
Neutrofilos abs	Number of Neutrophils before Surgery before Surgery	Pre
MonocitepercentAbsol	Absolute Percentage of Monocytes before Surgery	Pre
MonciteabsolPre	Absolute Number of Monocytes before Surgery	Pre
Plaquetas	Number of Platelets before Surgery	Pre

Plaquetas(MPV)	Average Platelet Volume before Surgery	Pre
Plquetas(PDW)	Platelet Distribution Width before Surgery	Pre
ABO - o-0; A-1; B2; AB3	Blood type	Pre
RH+	Presence of the Rhesus factor	Pre
Creatinina	Creatinine before Surgery	Pre
ureia	Urea before Surgery	Pre
sodio	Sodium before Surgery	Pre
potassio	Potassium before Surgery	Pre
pele a pele	Time from skin incision to skin closure (Hours and Minutes)	Pre
pele a pele Min	Time from skin incision to skin closure (Minutes)	Pre
tempo de clampagemContado	Clamping time (Hours and Minutes)	Pre
tempo de clampagem (Min)	Clamping time (Minutes)	Pre
PAD – agudização	Episode of Acute Peripheral Arterial Disease	Post
PADagudização data	Occurring Date of Acute Peripheral Arterial Disease	Post
EAM	Episode of Acute Myocardial Infarction	Post
EAM data	Occurring Date of Acute Myocardial Infarction	Post
ICA	Episode of Acute Heart Failure	Post
ICA data	Occurring Date of Acute Heart Failure	Post
AVC	Episode of Stroke	Post
AVC data	Occurring Date of Stroke	Post
Mortalidade	Death	Post
DeathDate	Date of Death	Post
Causa	Cause of Death	Post
MACE	Episode of Major Adverse Cardiovascular Events	Post
MACEDate	Occurring Date of Major Adverse Cardiovascular Events	Post
MACE+CerebroVascular	Major Adverse Cardiovascular Events and Cerebrovascular Events Episode	Post
MACE+CerebrovascularDate	Occurring Date of Major Adverse Cardiovascular Events and Cerebrovascular Events	Post
Idade	Patient's age (Years)	Post
Idade Superior a 75	Patients aged 75 or over	Pre

Sexo F=0 M=1	Patient's sex	Pre
HTA	Hypertensive patient	Pre
Fumador	Smoker	Pre
DM	Diabetic	Pre
Dislipidemia/estatina	Statin medication	Pre
DRC - creat sup. 1.5	Chronic Kdney Disease	Pre
Obesidade	Obesity	Pre
PADpreop	Peripheral Arterial Disease	Pre
DAC	Coronary Artery Disease	Pre
Atherosclerosis	Presence of Atherosclerosis	Pre
FA pre op	Atrial Fibrillation	Pre
DPOC/SAHOS/DPRC	Chronic Obstructive Pulmonary Disease, Obstructive Sleep Apnea Syndrome	Pre
FE<40	Ejection Fraction less than 40%	Pre
FEVE	Left Ventricular Ejection Fraction	Pre
CABG	Coronary Artery Bypass Graft Surgery	Pre
Stentcoronario	Presence of a Stent in the Coronary Artery	Pre
CardiacStressTest2Y (0-not done; 1-normal; 2-altered)	Cardiac Stress Test	Pre
RevascNonCoronaryPreop	Absence of Coronary Revascularization	Pre
Cx Cardíaca mesmo internamento?	Cardiac Surgery in the same Hospitalization	Pre
mRankim	Stroke Assessment Scale	Pre
Status funcional	Patient's Functional Capacity	Pre
2-Status Funcional	Patient's Functional Capacity (StatusFunctional inversion)	Pre
Nihss	Stroke severity quantification by Nihss Scale	Pre
Estenose	Type of stenosis	Pre
SintomaticCod	Stenosis type Code	Pre
Se sintomática - Quanto tempo evento-cirurgia	Time passed since surgery until the stenosis manifests as symptomatic (days)	Pre
SintomaticSimp	Stenosis type Code	Pre
Técnica cirúrgica	Technique used	Pre
TecnicaCod	Technique code used	Pre
PatchCod	Patch utilization (Code)	Pre
Grau de estenose	Stenosis Degree (%)	Pre

Grau de estenose contra-lateral	Degree of stenosis on the opposite side of the Index Surgery	Pre
EstenoseIpsiCod	Range of Ipsilateral Stenosis Degree	Pre
EstenoseContraCod	Range of Contra-lateral stenosis Degree	Pre
Fibrinólise	Fibrinolysis	Pre
ASA	Patient's state according to ASA classification system	Pre
Numero anti-HTA	Number of Antihypertensives the patient is taking until Surgery	Pre
IECAs/ARAS	Hypertensive Vitamins	Pre
BB	Beta Blockers medication	Pre
BCC	Calcium Chanel Blockers	Pre
DiureticoTiazidicos	Thiazide diuretic medication	Pre
Nitratos	Nitrates medication	Pre
AngioTC_data	Date of CT Angio	Pre
CT pre-op	CT realization	Pre
CT_data	CT scan Date	Pre
ARWC – córtex	Fazekas Score	Pre
ARWC - basal ganglia	Fazekas Score	Pre
Tissue stroke - 0; Watershed stroke 1	Stroke subtype	Pre
ACM	Stroke affects ACM territory	Pre
ACA	Stroke affects ACA territory	Pre
ACP	Stroke affects ACP territory	Pre
Fetal_ACP	Presence of Fetal Posterior Cerebral Artery	Pre
Fetal_ACP_lado	Side of the Fetal Posterior Cerebral Artery	Pre
PWAntSuf	Anterior part of the Willis Polygon	Pre
PWpostSuf	Back of the Willis Polygon	Pre
Tissue stroke - 0; Watershed stroke 1 - Pre	Stroke subtype	Pre
PAD internamentp	Diastolic Blood Pressure during Hospitalization	Pre
PAS internamentp	Systolic Blood Pressure during Hospitalization	Pre
intern_pas	Systolic Blood Pressure before Hospitalization	-
Incisao	Incision time	-
clamp_missing	Lack of clamping	-
clamp	Clamping time	-
desclamp_missing	Lack of de-clamplification	-

Desclamp	Time to de-clamping	-
tempo_clamp	Clamping Time	-
fimcir	Surgery end Time	-
tempo_cir	Surgery Duration	-
upa_missing	Missing data for Hospital entry Time	-
upa_in	Time of entry in the Hospital	-

3.2.1 Data Exploration

Data Exploration allows us to understand the dataset by examining and visualizing it, as well as identifying necessary adjustments to improve data quality and model performance (Schröer et al., 2021). As mentioned above, the dataset is made up of 191 variables (described in Table 3.1), of which 55 are categorical (object), 133 are numerical (float and int) and 3 are dates (datetime).

Table 3.2 presents the statistical information from the three outcomes of interest to the study and that are going to be combined to form the target variable. As is possible to see, all the variables have a total of 228 observations, indicating no missing values. Each variable has three distinct categories, which doesn't make sense, since all of them are binary variables. The most frequent value in all cases is 0, suggesting that the majority of instances do not exhibit the corresponding event. Specifically, for EAM, the value 0 appears in 211 cases (92.5%), for AVC, in 197 cases (86.4%), and for MACE, in 159 cases (69.7%), highlighting a strong class imbalance. This imbalance can impact modeling, resulting in biased results, thus it will need to be addressed. Besides this, the variables are represented as categorical, however, as mentioned before, they are all binary, which means that datatypes are also incorrect.

Table 3.2 – Statistical Information for the outcomes in both models

Variable	Count	Unique	Top	Frequency
EAM	228	3	0	211
AVC	228	3	0	197
MACE	228	3	0	159

3.2.2 Data Pre-Processing

During Data Pre-Processing the data is processed to eliminate any inconsistencies detected on Data Exploration, as well as processing missing values and outliers (Schröer et al., 2021). Before any Data Preparation, duplicate data was checked and the variable Número de Processo (ND) was set to index in both models.

3.2.2.1 Data Cleaning

Besides the elimination of black columns and the variables considered post-surgery in Model 1 and pre-surgery in Model 2, both models had columns with some kind of problem, namely redundancy and irrelevance. At the same time, some patients had missing values in more than 50% of the variables, and one patient had not undergone CEA. The Data Cleaning phase was therefore carried out in both models to eliminate all these columns/ID numbers (Table 3.3).

Table 3.3 – Data Exclusions During Cleaning Phase

Model	Criteria for eliminating	Column's Name/ID Number
1	Redundant Columns	Lateralidade, Linfócitos %, Neutrofilos %, 2-Status Funcional, Técnica cirúrgica, Idade Superior a 75, Data de Nascimento, Grau de estenose, MCHC
	Columns with no meaning	Anestesia, AnestesiaCod, MétodoAvaliaçãoHipoIntra, Défice intra-op, TC-data, CT_data, CT pre-op, AngioTC_data
	Columns with no relevance	Dadoshemodinamica, MetSync, INVOSpreIpsilat, INVOSpreContra, INVOSposclampIpsi, INVOSposClampContra, Subida com shunt, Pos desclampagem Ipsi, intern_pas, incisao, clamp_missing, clamp, desclamp_missing, desclamp, tempo_clamp, fimcir, tempo_cir, upa_missing, upa_in, PAD – agudização, PADagudização data, ICA, ICA data, Mortalidade, DeathDate, Causa, MACE+CerebroVascular, MACE+CerebrovascularDate, EAM data, AVC data, MACEDate
	Columns with no sense of values	FE<40
	IDs with missing values in more than 50% of the columns	6029821, 23014428, 1005994, 17012730, 4019783, 93078862, 1005994, 98028667, 5035373, 94048239
	Didn't realize CEA	93071715
2	Redundant Columns	Linf%posop, Neut%posop, eosinofilosposop, monocitosposop
	Columns with no relevance	Dadoshemodinamica, MetSync, INVOSpreIpsilat, INVOSpreContra, INVOSposclampIpsi, INVOSposClampContra, Subida com shunt, Pos desclampagem Ipsi, intern_pas, incisao, clamp_missing, clamp, desclamp_missing, desclamp, tempo_clamp, fimcir, tempo_cir, upa_missing, upa_in, PAD – agudização, PADagudização data, ICA, ICA data, Mortalidade, DeathDate, Causa, MACE+CerebroVascular, MACE+CerebrovascularDate, EAM data, AVC data, MACEDate
	IDs with missing values in more than 50% of the columns	6029821, 23014428, 1005994, 17012730, 4019783, 93078862, 1005994, 98028667, 5035373, 94048239
	Didn't realize CEA	93071715

3.2.2.2 Missing Data

Most of the ML algorithms cannot deal with missing data, besides that the presence of missing values complicates data analysis as well as the validation of results and conclusions. There are many ways to manage missing values (Acuna & Rodriguez, 2004).

After analyzing the missing values in both models, variables with over 15% missingness were removed and for the rest, the strategy adopted was replacing the missing values with mode on categorical variables and mean in numerical variables to maintain the original distribution of the data.

3.2.2.3 Outliers

Outliers are data points that deviate significantly from most observations in a dataset. In predictive modeling, outliers can heavily skew the results, leading to biased parameter estimates, reduced model accuracy, and poor performance of ML algorithms (Aggarwal, 2017).

Given the context of this dataset, which involves clinical data, it was acknowledged that outliers might not always be the result of errors. In healthcare datasets, outliers may represent patients with specific medical conditions, making these data points clinically relevant. Therefore, in both models, to identify possible outliers, histograms were constructed for each of the numerical variables and after that, it was considered:

- Reference values for each variable that appeared to have outliers by consulting the review published by Zahorec (2021) and the documents Intervalos de referência (convencionais) and Laboratory Tests Reference Ranges published by the Administração Central do Sistema de Saúde (ACSS) and the American Board of Internal Medicine (ABIM), respectively (the reference values for the interest variables are detailed in Table 3.4).
- Number of apparently extreme values registered.
- Values of the correlated variables within the same patient.

Table 3.4 – Reference values by Zahorec (2021), ACSS and ABIM Clinical Reference Ranges

Marker	Reference values		
	Lower	Normal	Upper
Cardiac Frequency (bpm)	<60	60-100	>100
Cholesterol (mg/dL)	<40	40-200	>200
Creatinine (mg/dL)	<0.6	0.6-1.3	>1.3
Eosinophils (%)	<1	1-3	>3
Erythrocytes (million/µL)	<4.1	4.1-5.9	>5.9
FEVE (%)	<52	52-74	>74
HCT	<41	41-50	>50
HDL (mg/dL)	<40	≥40	-
Hemoglobin (g/dL)	<13	13-18	>18
LDL (mg/dL)	-	100-129	>129
MCHC (g/dL)	<33	33-36	>36
Monocytes (%)	<3	3-7	>7
MPV (fl)	<7	7-9	>9
NLR	<1	1-2	>2
PDW (fl)	<9	9-17	>17
Potassium (mEq/L)	<3.5	3.5-5.0	>5.0
Platelet (x109/L)	<150	150-400	>400
RDW Cv (%)	<11.5	11.5-14-5	>14.5
RDW Sd (fl)	<39	39-46	>46
Sodium (mEq/L)	<135	135-145	>145
Triglycerides (mg/dL)	-	<150	>150
Troponin I (ng/mL)	<0	0-0.04	>0.04
Urea (mL/min)	<40	40-60	>60

The identification of the possible outliers in the different variables was done using the interquartile range (IQR) method, a common technique in exploratory data analysis, and was also counted the frequency of outliers in each variable to identify patterns in the different rows. The IQR method defines outliers as observations that fall outside the range of $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$, where Q1 and Q3 represent the first and third quartiles of the data, respectively. The IQR, or the difference between these quartiles, captures the middle 50% of the data, while points falling far outside this range are flagged as potential outliers (Tukey, 1977).

This approach strikes a balance between the need to maintain data integrity and avoid bias, while recognizing the potential importance of extreme values in clinical settings (Aggarwal, 2017; Osborne & Overbay, 2004).

None of the variables were eliminated at this stage; however, in each of the models, the extreme values that were considered outliers after evaluation were eliminated, which can be consulted in Table 3.5.

Table 3.5 – Outliers and elimination criteria

Model	Variable	Elimination Criteria
1	trop I	== 5.69
	HDL	== 14
	Eritrocitos pre op	== 11.9
	Hct pre op	== 10.9
	Hct pre op	== 88
	ureia	< 1
	FEVE	== 0
2	hb pos op	< 5
	RDW Cvpos op	> 18.9
	RDW Sdposop	< 10 and > 70
	Linfabsposop	> 30
	Neutposop	>16
	eosinofilos%posop	> 10
	monocitos%posop	> 25
	Plaquetasposop	< 50
	MPVposop	< 2

3.2.2.4 Features Engineering

In the Feature Engineering phase, the target variable was created, consisting of a multi-class variable constructed from the outcomes to be predicted: MI, Stroke, and MACE. These outcomes initially represented as individual columns in the dataset, were combined to form a single categorical variable – Outcome - which classifies patients based on the occurrence of one or more of these events.

The construction of the variable followed a logic of mutually exclusive conditions. The reference group, corresponding to patients with none of the events, was assigned the value 0. The different scenarios were then defined: patients with only MI (1), only Stroke (2) or only MACE (3). For those with multiple events, additional categories were created: MI and Stroke (4), MI and MACE (5), Stroke and MACE (6), and finally, patients with all three events

simultaneously (7). Each patient was assigned to the class using the np.select() function, ensuring efficient classification without overlapping.

Besides constructing the target variable, the variable originally represented as a date, Data da Cirurgia, was converted into numerical value, creating DaysSinceSurgery. In addition, the Ratio of Neutrophils and Lymphocytes (NLR pre and NLR pos) was calculated, based on the variables Neutposop and Linfabsposop, to assess the patient's inflammatory status.

Finally, the variables corresponding to laboratory parameters were discretized into bins based on the reference values, as already shown in Table 3.3. This approach made it possible to transform continuous variables into interpretable categories, facilitating analysis and potentially improving the performance of predictive models.

3.2.3 Train Test Split

For the development of a predictive model, it is essential to divide the dataset into training and test sets. This separation allows the model to learn patterns from the training data while keeping an independent test set for evaluation (Hastie, Tibshirani, & Friedman, 2009).

During this process, the class imbalance in the dataset became even more evident, also revealing that classes 1 and 4 had no records at all (Figure 3.1 and Figure 3.2). To address this issue, firstly, class 7 was eliminated, since it presents only 2 registers, which is not enough for the model to learn any kind of pattern. Secondly, the Synthetic Minority Oversampling Technique (SMOTE) was applied. SMOTE generates synthetic examples for the minority class by interpolating between existing data points, rather than merely duplicating them (Chawla et al., 2002).

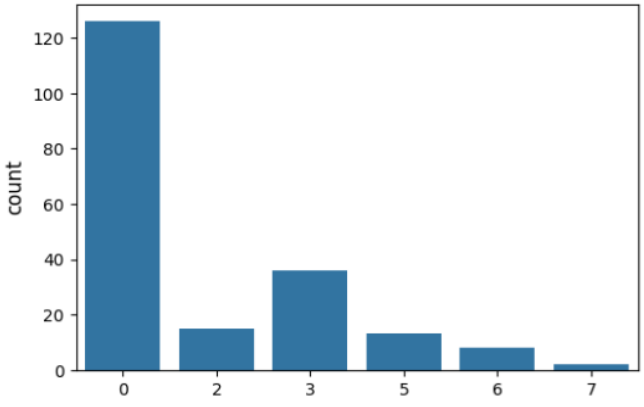

Figure 3.1 – Target variable class imbalance Model 1

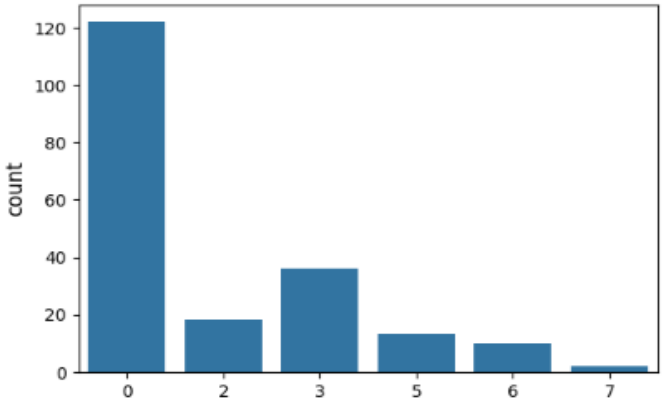

Figure 3.2 – Target variable class imbalance Model 2

3.2.4 Feature Selection

Feature selection is a crucial step in developing a predictive model, as it involves reducing the input variables based on their significance in predicting the outcome (Li et al., 2018). To ensure a comprehensive approach, several methods were used for the two models: Spearman Correlation, Analysis of Variance (ANOVA), Decision Trees (DT), Random Forest, Lasso Regression, Adaptive Boosting (AdaBoost), and Recursive Feature Elimination (RFE). This multiplicity of methods avoids reliance on a single technique and ensures robust selection.

Spearman correlation (Ali & Al-Hameed, 2022) and the Variance Inflation Factor (VIF) were used to assess multicollinearity. A threshold of 0.7 was set for Spearman correlation, and a threshold of 50 was established for VIF. Variables that exhibited a high correlation (above 0.7) and a high VIF (above 50) were removed. VIF (Salmerón et al., 2020) is a metric that detects multicollinearity among independent variables. A VIF close to 1 indicates low collinearity, while values above 5 or 10 suggest moderate to high collinearity. A VIF above 50, as used in this analysis, indicates severe collinearity, warranting variable removal to maintain model stability and interpretability.

ANOVA (Kim, 2014) is a statistical test that helps determine which features have a significant impact on the target variable, making it a valuable tool for selecting relevant predictors. Similarly, Decision Trees (DT) (Gupta et al., 2017) assist in identifying the most influential variables for classification while also providing a clear visualization of their relative importance. Expanding on this approach, Random Forest (Boulesteix et al., 2021) leverages an ensemble of decision trees to improve accuracy and mitigate overfitting, offering a robust method for estimating feature importance. Meanwhile, Lasso Regression (Muthukrishnan et al., 2016) enhances model simplicity by performing automatic feature selection, eliminating non-significant coefficients, and retaining only the most impactful variables. AdaBoost (Cao et al., 2013), refines model accuracy by giving more weight to the most critical features, progressively improving classification performance. Finally, Recursive Feature Elimination

(RFE) (Dadebayev et al., 2021) systematically removes less important variables to identify the subset that contributes most to the model's predictive power.

3.2.4.1 Final Feature Selection

To enhance the robustness of variable selection and improve the reliability and accuracy of predictions, only the variables identified as relevant by more than two methods were retained to construct the final predictive model. Additionally, any selected variables not originally named in English were renamed based on their respective characteristics (Table 3.6).

Table 3.6 – Final features selected

Model	Features Selected	Final Feature Name
1	PADpreop	PreOp PAD
	Doença circulação posterior	Posterior Circulation Disease
	FA pre op	PreOp Atrial Fibrillation
	RDW-CV	PreOp RDW CV
	RDW-SD	PreOp RDW SD
	DRC - creat sup. 1.5	Chronic Kidney Disease
	Hemiplegia	PreOp Paresis
	altura	Height
	Status funcional	Functional State
	DAC	CAD
	DaysSinceSurgery	DaysSinceSurgery
	EstenoseContraCod	Contralateral stenosis
	Bd_mass_ind	Body Mass Index
	Idade	Age
	AAS	Acetylsalicylic Acid
	Défice inta-opCod	IND
	ureia	PreOp Urea
	EstenoseIpsiCod	Ipsilateral stenosis
	Numero anti-HTA	Anti-Hypertensors Number
	Hemoglobina	PreOp Hemoglobin
	CardiacStressTest2Y (0-not done; 1-normal; 2-altered)	Cardiac Stress Test
2	FC pos op	PostOp Heart Rate
	DaysSinceSurgery	DaysSinceSurgery
	Age	Age
	RDW Cvpos op	PostOp RDW CV
	RDW Sdposop	PostOp RDW SD
	RestenoseCod70	Restenose 70%
	monocitos%posop	PosOp Monocytes
	AVCposop30D	30D PostOp Stroke

3.2.5 Model Development and Selection

During the model creation phase, various algorithms were employed to compare their performances and select the most suitable one. The algorithms used included K-Nearest Neighbors (kNN), Naive Bayes, Logistic Regression, Random Forest, Extreme Gradient Boosting (XGBoost), Decision Trees, Support Vector Machine (SVM), and Neural Networks. These algorithms were chosen for their different approaches and modeling capabilities, allowing a comprehensive analysis of variables and ensuring the final model is robust and accurate. The modeling phase was equal to the different models, to guarantee a more precise comparison.

K-Nearest Neighbors (KNN) (Taunk et al., 2019) was chosen for its ability to handle multiple classes effectively and for its strong performance on small datasets, which aligns well with the study's dataset characteristics. Similarly, Naïve Bayes (Wickramasinghe & Kalutarage, 2021) was included due to its simplicity, speed of implementation, and effectiveness. Despite this assumption rarely holding in practice, Naïve Bayes often delivers robust results. Logistic Regression (Smita M., 2021) was selected for its interpretability, as it allows for the direct examination of how each explanatory variable influences the probability of the outcome. This characteristic is particularly valuable in medical studies, where understanding the relationship between risk factors and clinical events is crucial. To capture more complex interactions between features, Random Forest (Shaik & Srinivasan, 2019) was chosen for its ability to introduce variation into the learning process, reducing overfitting and providing a reliable measure of feature importance. Likewise, Decision Trees (DTs) (Gupta et al., 2017) were used due to their interpretability and capacity to model non-linear relationships, making them especially useful for understanding which variables are most critical in predicting adverse outcomes. XGBoost (Budholiya et al., 2022) was included because of its high efficiency, particularly in scenarios requiring strong predictive performance. Its ability to handle complex data structures makes it a powerful choice for medical outcome predictions. Support Vector Machines (SVMs) (Cervantes et al., 2020) were also considered, as they perform well in high-dimensional spaces and allow for the use of different kernels (linear, polynomial, RBF), enabling the capture of non-linear relationships within the data. Finally, Neural Networks (Dastres & Soori, 2021) were integrated into the analysis due to their flexibility in learning complex patterns. This characteristic is particularly relevant in a medical context, where interactions between clinical variables can be intricate and difficult to model using traditional statistical methods.

3.2.5.9 Evaluation Metrics

Various metrics are essential for evaluating the performance of ML algorithms and gaining a comprehensive understanding of their effectiveness. Common evaluation metrics include Accuracy, Precision, Recall, F1-score, and ROC curve (AUC-ROC) (M & M.N, 2015).

The study presented used Accuracy as the primary metric offering a straightforward assessment of the model's performance (Ferri et al., 2009).

A classification report was generated to better analyze the model's performance. This tool provides a detailed breakdown of performance metrics for each class, including precision - the ratio of true positive predictions to the total predicted positives, indicating the accuracy of positive predictions -, recall - the ratio of true positive predictions to all actual positives, measuring the model's ability to identify all relevant cases within a dataset -, F1-score - mean of precision and recall, balances the two metrics -, and support - the number of actual occurrences of each class in the dataset rates (Sokolova & Lapalme, 2009).

3.2.5.10 Hyperparameter Optimization

Hyperparameter optimization is a crucial step in developing ML models, as these parameters significantly impact the model's performance (Yang & Shami, 2020). Before optimizing the parameters, all the algorithms were tested with the default ones, and all those with an accuracy above 55% were selected for optimization, using Randomized Search (RS), which explores the search space and randomly selects a predefined number of hyperparameter combinations (Yang & Shami, 2020). This selection criterion focused optimization efforts on high-potential models, ensuring that time and resources are used effectively. Table 3.7 details the hyperparameters used for the models that perform better and went through hyperparameter optimization phase.

Table 3.7 – Hyperparameters used for performance optimization

Algorithm	Hyperparameters
DT	criterion, max_depth, min_samples_split, min_samples_leaf, max_features, class_weight, min_impurity_decrease, max_leaf_nodes
Logistic Regression	C, solver, penalty, max_iter, multi_class, fit_intercept, intercept_scaling, class_weight, l1_ratio, tol, verbose, warm_start
Random Forest	n_estimators, max_depht, min_samples_split, min_samples_leaf, max_features, class_weight, criterion, max_leaf_nodes, verbose
XGBoost	Objective, num_class, booster, alpha, lambda, max_depht, learning_rate, n_estimators, subsample, colsample_bytree, min_child_weight, grow_policy

3.2.6 Feature Importance

In this study, the feature importance stage was conducted using the Shapley Additive Explanations (SHAP) method and was done in each class individually, since it was a multi class model. SHAP is an interpretative approach that aims to explain the predictions of ML models by quantifying the individual contribution of each feature to the model's output. SHAP's main objective is to provide consistent and local explanations of model predictions, allowing for better interpretation (Scavuzzo et al. 2022).

4. RESULTS AND DISCUSSION

After addressing all preprocessing challenges, the development of the predictive models was carried out. As previously mentioned, this study aims to identify pre- and post-surgery factors that influence the occurrence of adverse events (MACE, MI, and Stroke) following CEA under loco-regional anesthesia. To achieve this, two distinct models were developed: one considering only pre-factors and the other based exclusively on post-factors.

4.1 MODEL 1

4.1.1. Pre-Hyperparameter Tunning Results

This initial approach tested multiple algorithms using default hyperparameters to assess their baseline performance. The results highlighted varying levels of effectiveness across the models, with XGBoost and Random Forest emerging as the top performers, being the only models to surpass 55% accuracy. Random Forest achieved the highest accuracy at 62.5% and a ROC AUC of 0.59, demonstrating a moderate ability to distinguish between classes. However, while it almost perfectly classified class 0, it struggled significantly with the remaining classes, particularly classes 2, 3, 5, and 6, failing to correctly identify them. XGBoost followed closely, achieving an accuracy of 57.5% and a ROC AUC of 0.63, indicating a slightly better overall ability to differentiate between classes. Despite this, its classification remained heavily skewed towards class 0, while it exhibited difficulties in correctly predicting the other classes, particularly classes 2, 5, and 6.

The remaining models exhibited moderate to low performance, failing to exceed the 55% accuracy threshold. Given these results, only XGBoost and Random Forest were selected for further fine-tuning and optimization, as they demonstrated the highest potential for predictive accuracy despite their misclassification of minority classes.

4.1.2. Post-Hyperparameter Tunning Results

After performing hyperparameter optimization by testing various parameter combinations, the results show a moderate improvement in model performance compared to the baseline evaluation. The optimized hyperparameters and their respective values are described in Table 4.1.

Table 4.1 – Hyperparameters optimized and respectively value in Model 1

Algorithm	Hyperparameters	Values
Random Forest	n_estimators	140
	max_depth	96
	min_samples_split	10
	min_samples_leaf	7
	max_features	0.986450070336335
	class_weight	None
	criterion	gini
	max_leaf_nodes	40
	verbose	9
XGBoost	objective	multi:softprob
	num_class	5
	booster	gbtree
	alpha	0
	lambda	0
	max_depth	5
	learning_rate	0.05
	n_estimators	100
	subsample	0.6
	colsample_bytree	1.0
	min_child_weight	1
	grow_policy	deptwise

Random Forest achieved the highest accuracy (65.0%) and the best ROC AUC score (0.70), indicating a stronger ability to distinguish between classes compared to XGBoost. This model perfectly classified class 0, but struggled with minority classes, particularly class 2, 5, and 6, which were not classified correctly. XGBoost followed with an accuracy of 60% and a ROC AUC score of 0.65. It also performed well for class 0, but failed to effectively classify classes 2, 3, 5, and 6, leading to lower overall recall and precision. The XGBoost model benefited from tuning parameters such as increasing max_depth to 10, setting the booster to 'gbtree', adjusting the learning rate to 0.05, and optimizing regularization parameters (alpha and lambda set to 0). These adjustments aimed at improving model generalization, though challenges remained in predicting minority classes. For Random Forest, the best results were achieved with 140 estimators, a maximum depth of 96, a max_features value of 0.99, and fine-tuned min_samples_split and min_samples_leaf parameters. These optimizations contributed to an overall improvement in predictive performance, particularly for the majority class.

Given these results, Random Forest was selected as the primary model due to its superior classification performance for feature importance evaluation.

Table 4.2 – Results after hyperparameters optimization in Model 1

Algorithm	Class	Evaluation Metric					
		Precision	Recall	F1-Score	Support	Accuracy	ROC AUC
Random Forest	0	0.71	1.00	0.83	25	65%	0.70
	2	0.00	0.00	0.00	3		
	3	0.33	0.14	0.20	7		
	5	0.00	0.00	0.00	3		
	6	0.00	0.00	0.00	2		
XGBoost	0	0.68	0.92	0.78	25	60%	0.65
	2	0.00	0.00	0.00	3		
	3	0.25	0.14	0.18	7		
	5	0.00	0.00	0.00	3		
	6	0.00	0.00	0.00	2		

4.1.3. Feature Importance Evaluation

In terms of feature importance, Random Forest revealed distinct patterns across the different classes, presenting DaysSinceSurgery, Age and Body Mass Index variables as the most important for the majority of the classes. As described above, class 0 corresponds to no event, class 2 to Stroke, class 3 to MACE and classes 5 and 6 to combine events, MI and MACE, and Stroke and MACE, respectively.

In class 0, the model identified DaysSinceSurgery, Age and Body Mass Index variables as the most important, showing that lower values in the DaysSinceSurgery and Age variables increase the probability of no adverse event occurring, while in the Body Mass Index variable, higher values increase the model's prediction. For class 2, Random Forest presented Height as the most important variable, being followed by Age and DaysSinceSurgery. In this class Age showed a opposite behavior compared to class 0, in the sense that higher values increase the risk of the event occurring. Class 3 returns the three same variables as class 0 as being the most important, however all of them presented a different behavior when compared to this last class. From the results, higher values of DaysSinceSurgery and Age increase the probability of the event occur, while higher Body Mass Index values make it less likely. In class 5, the model spotted Chronic Kidney Disease as the second variable with bigger impact, in addition to DaysSinceSurgery in first place and Body Mass Index in third place. The model showed that patients with Chronic Kidney Disease have a higher probability of developing MACE. Lastly, class 6 is the only one in which the two most important features were not yet identified in first

places in the other classes, which were PreOp RDW CV and IND, followed by DaysSinceSurgery. For this class, the PreOp RDW CV variable proves to be the most important feature, where higher values increase the likelihood of patients developing the combined adverse event, Stroke and MACE.

Overall, the results showed that the number of days passed since surgery, patient age and body mass index play a central role in predicting adverse events, although their impact varies depending on the class analyzed. Additionally, the presence of chronic kidney disease, the values of RDW CV and the occurrence of IND proved to be particularly relevant for specific classes, emphasizing the need for a personalized approach in risk assessment. Figures 4.1 present the feature importance plots obtain with Random Forest model.

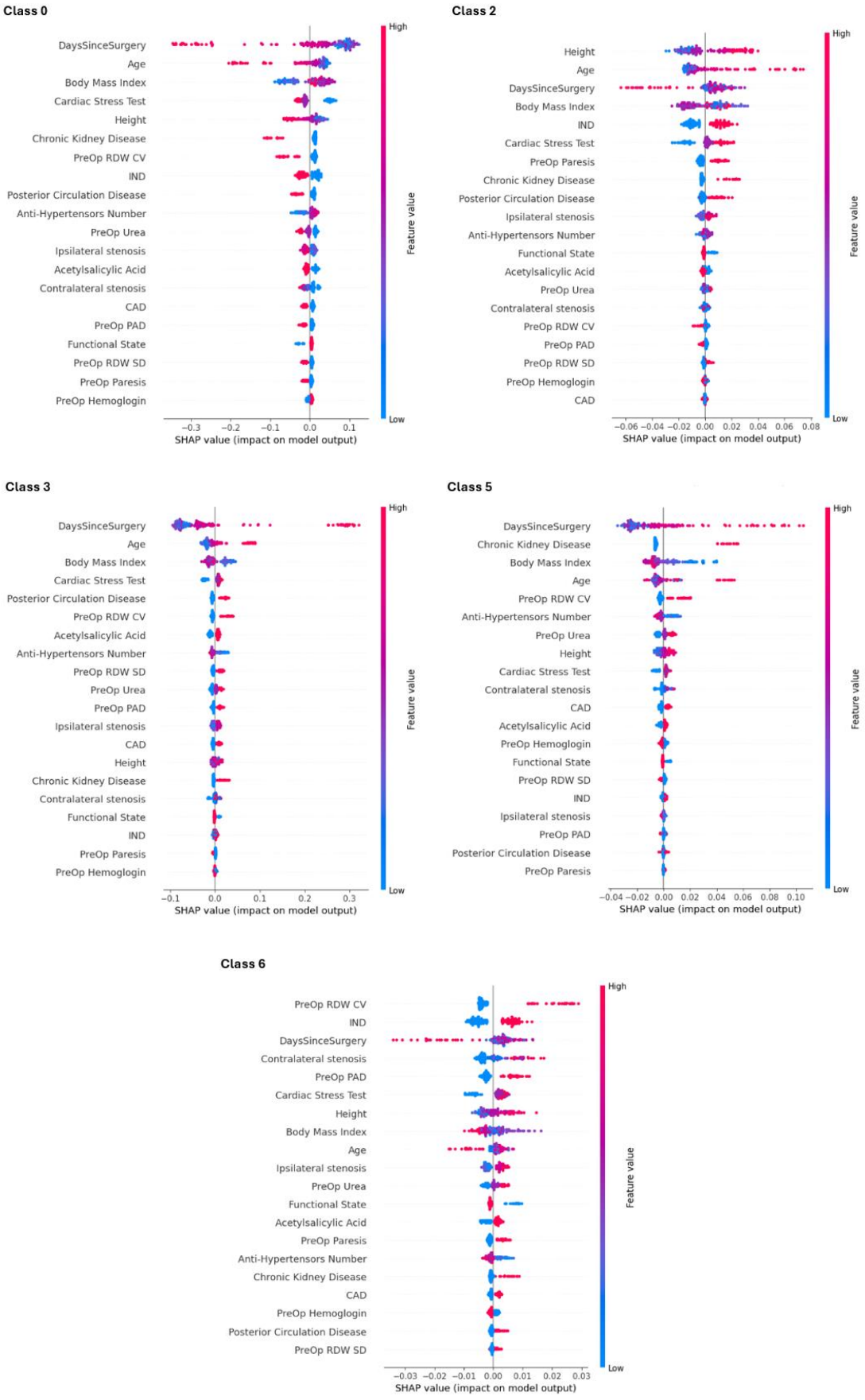

Figure 4.1 – Feature Importance with Random Forest in each class for Model 1

4.2 MODEL 2

4.2.1 Pre-Hyperparameter Tunning Results

With the default hyperparameters, XGBoost, Random Forest, and Logistic Regression achieved the best results in Model 2, all reaching 62.5% accuracy and being selected for further hyperparameter optimization.

XGBoost demonstrated the highest overall performance in terms of ROC AUC, reaching 0.65, suggesting a slightly stronger ability to distinguish between classes compared to the other models. It accurately classified most instances of class 0, but struggled with minority classes, particularly class 2, 5, and 6, which were almost entirely misclassified. Random Forest also achieved 62.5% accuracy, with a ROC AUC of 0.62, showing similar performance trends. It correctly classified most instances of class 0, had some success with class 3, but performed poorly for classes 2, 5, and 6, failing to classify them effectively. Logistic Regression, despite having the same accuracy (62.5%), achieved the highest ROC AUC at 0.72, indicating a better ability to distinguish between classes overall. It perfectly classified class 0, but struggled heavily with classes 2, 3, and 5, leading to lower recall and precision for these categories.

4.2.2 Post-Hyperparameter Tunning Results

After performing hyperparameter tuning by testing multiple parameter combinations, the models demonstrated low variance compared to the results obtained before, with XGBoost and Random Forest maintaining the same performance as before and Logistic Regression slightly decreasing the performance. Table 4.3 presents the optimized hyperparameters and their respective values that yielded the best results for each model.

Table 4.3 – Hyperparameters optimized and respectively value in Model 2

Algorithm	Hyperparameters	Values
Logistic Regression	C	10
	solver	lbfgs
	penalty	l2
	max_iter	100
	multi_class	auto
	fit_intercept	False
	intercept_scaling	1
	class_weight	None
	l1_ratio	None
	tol	0.001
	verbose	0
	warm_start	False
Random Forest	n_estimators	160
	max_depht	88
	min_samples_split	9
	min_samples_leaf	4
	max_features	0.6482933281662953
	class_weight	None
	criterion	Entropy
	max_leaf_nodes	50
	verbose	6
XGBoost	objective	multi:softprob
	num_class	5
	booster	gbtree
	alpha	0
	lambda	0
	max_depht	5
	learning_rate	0.1
	n_estimators	100
	subsample	0.6
	colsample_bytree	0.9
	min_child_weight	3
	grow_policy	deptwise

XGBoost and Random Forest maintained an accuracy of 62.50% but showed a slight increase in the ROC AUC Score, reaching 0.64. In both models, class 0 remained the most accurately classified, while classes 5 and 6 achieved a reasonable level of identification. However, the predictive ability for classes 2 and 6 remained weak, indicating challenges in distinguishing these groups. For XGBoost, the adjustments included setting the max_depth to 5, a learning rate of 0.1, using the 'gbtree' booster, and optimizing the regularization parameters. These configurations contributed to improved model stability and slight enhancements in generalization. Meanwhile, the Random Forest model benefited from tuning the number of estimators to 160, setting a maximum depth of 88, and optimizing the selection of features considered at each split, being possible to achieve a better ROC AUC. Regarding Logistic Regression, accuracy slightly decreased to 62% and ROC AUC to 0.59. The model still struggles with classifying classes 3 and 6, especially.

Overall, hyperparameter tuning yielded mixed results across models. While Logistic Regression saw a decrease in the overall performance, Random Forest demonstrated enhanced model discrimination, while XGBoost maintained stable performance, reinforcing its robustness. The most persistent challenge across all models remains the correct classification of minority classes, particularly class 5. The detailed performance of each model for each class is summarized in Table 4.

Table 4.4 – Results after hyperparameters optimization in Model 2

Algorithm	Class	Evaluation Metric					
		Precision	Recall	F1-Score	Support	Accuracy	ROC AUC
Logistic Regression	0	0.66	0.88	0.75	24	62%	0.59
	2	0.50	0.25	0.33	4		
	3	0.40	0.29	0.33	7		
	5	0.00	0.00	0.00	3		
	6	1.00	0.50	0.67	2		
Random Forest	0	0.66	0.88	0.75	24	62.5%	0.64
	2	0.50	0.25	0.33	4		
	3	0.40	0.29	0.33	7		
	5	0.00	0.00	0.00	3		
	6	1.00	0.50	0.67	2		
XGBoost	0	0.67	0.92	0.77	24	62.5%	0.64
	2	0.00	0.00	0.00	4		
	3	0.60	0.43	0.50	7		
	5	0.00	0.00	0.00	3		
	6	0.00	0.00	0.00	2		

4.2.3 Feature Importance Evaluation

To evaluate feature importance, both XGBoost and Random Forest were selected, since they were the ones performing better.

In Class 0, XGBoost identified Age, DaysSinceSurgery and PostOp Monocytes as the most important variables, while Random Forest had identified PostOp Monocytes in first place followed by PostOp RDW SD and PostOp RDW CV. To XGBoost, which the most important variable is Age, younger patients tend to increase the probability of no event occurring, while to Random Forest, which the most important variable is PostOp Monocytes, bigger values showed having more impact on the probability of no event occur. For class 2, the two models continued to show different results, with XGBoost identifying DaysSinceSurgery as the most important variable, and Random forest identifying 30D PostOp Stroke. In class 3, DaysSinceSurgery and Age are the key preditors to XGBoost, in which the Shap values of these features behave in the opposite way when compared with Class 0. To Random Forest, similar in class 2, the two most important features are 30D PostOp Stroke and DaysSinceSurgery. Class 5, in XGBoost is primarily driven DaysSinceSurgery followed by Age, as in the last class, while in Random Forest, PostOp RDW SD and DaysSinceSurgery are the principal predictors. In both models, DaysSinceSurgery behave the same way, with biggest values showing more impact on the event occurrence. Finally, class 6, is the only one where both models agree in the most important variable, PostOp RDW CV, showing, also, the same behavior in both – bigger values showed more impact in predicting this combine event.

In conclusion, while both XGBoost and Random Forest were the best-performing models, their feature importance rankings varied across classes. XGBoost consistently highlighted DaysSinceSurgery and Age as key predictors, whereas Random Forest placed more emphasis on hematological variables like PostOp Monocytes and PostOp RDW SD. Despite these differences, DaysSinceSurgery emerged as a crucial predictor in multiple classes, showing a consistent impact on event occurrence in both models.

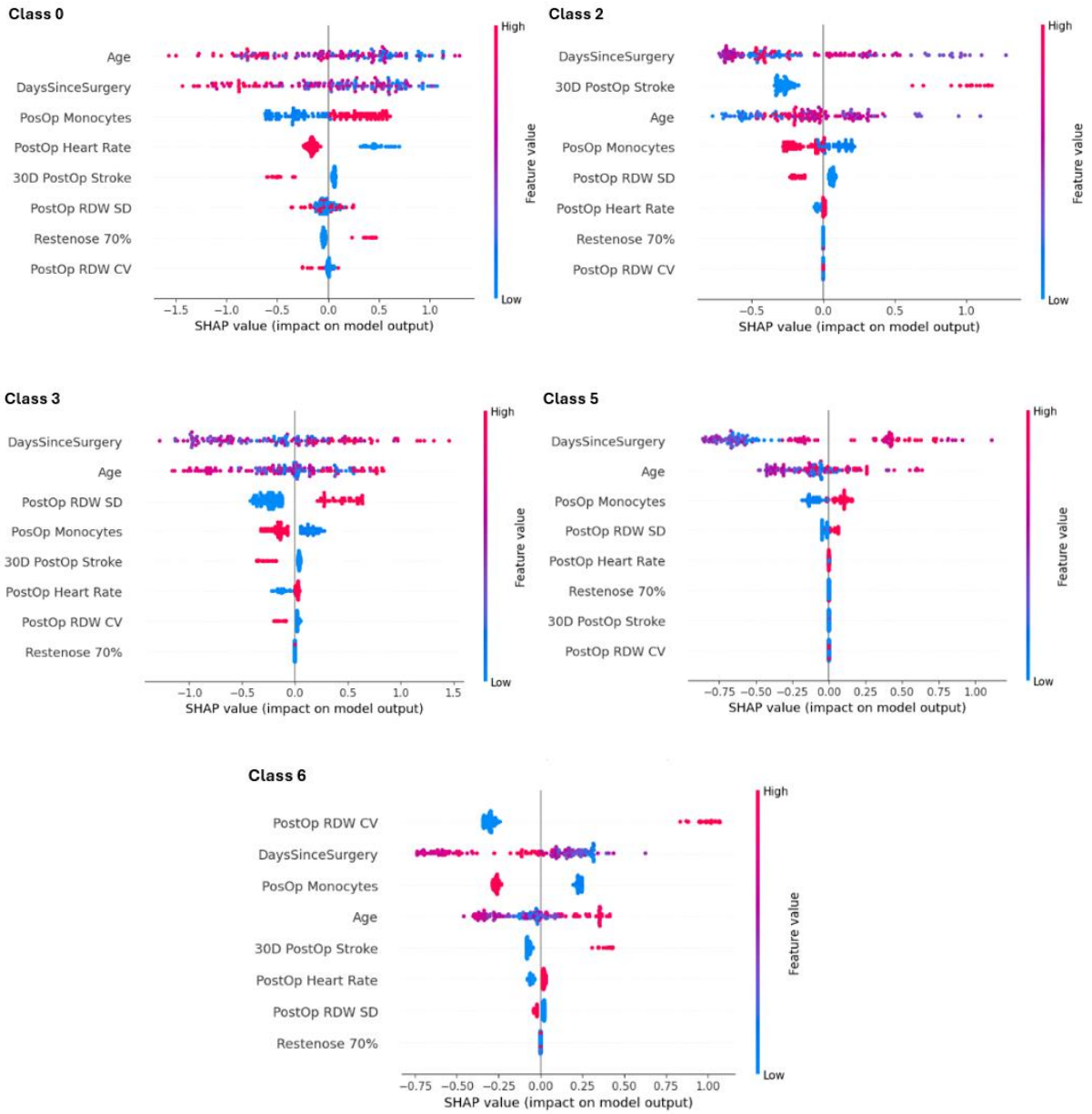

Figure 4.2 – Feature Importance with XGBoost in each class for Model 2

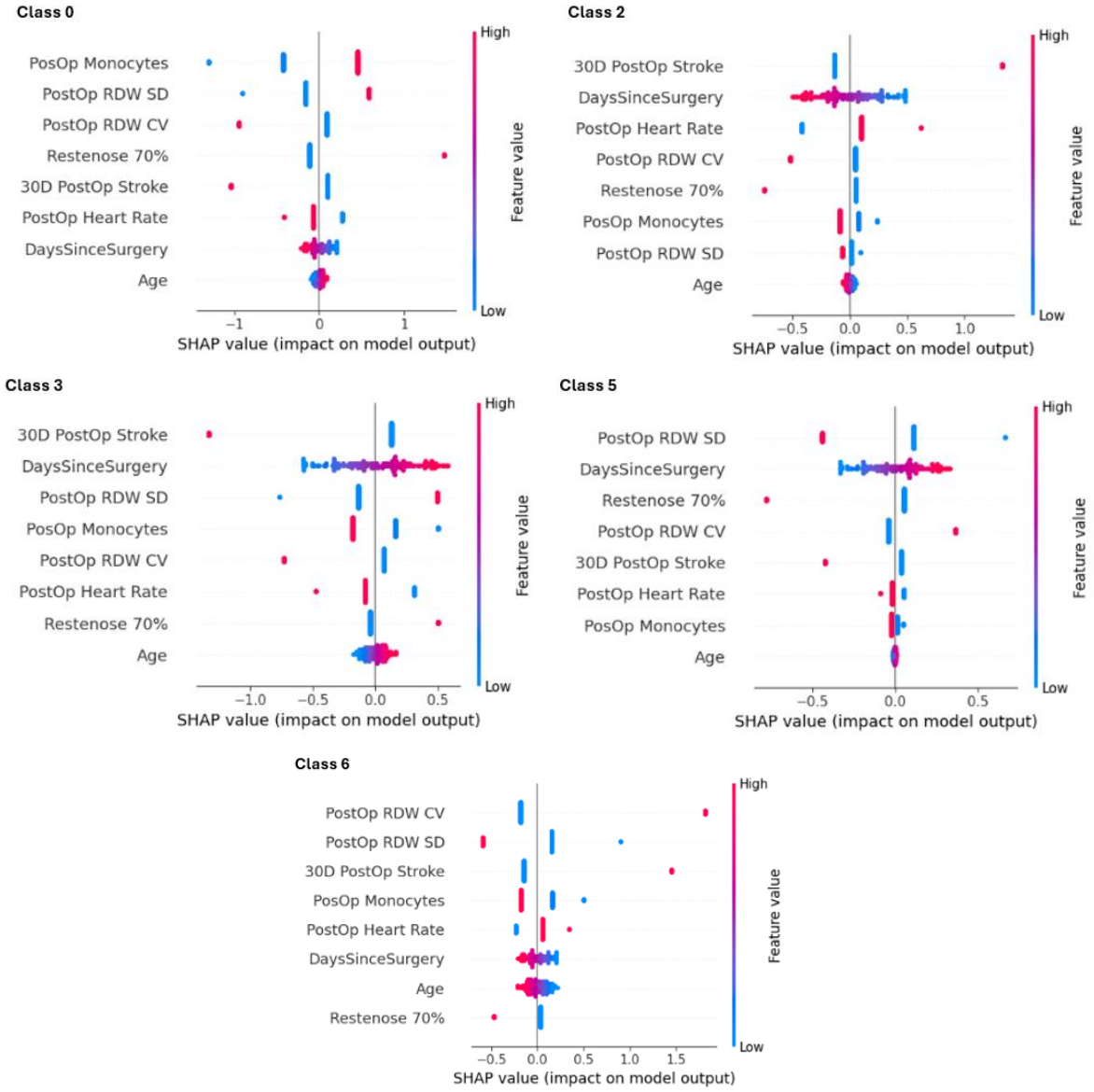

Figure 4.3 – Feature Importance with Random Forest in each class for Model 2

5. CONCLUSIONS AND FUTURE WORKS

The occurrence of adverse events following CEA remains a significant clinical challenge, highlighting the need for robust predictive models that can identify key risk factors. This study aimed to determine the pre- and post-surgical factors that influence the occurrence of adverse events, specifically MACE, MI, and Stroke, in patients undergoing CEA under loco-regional anesthesia. To achieve this, two distinct predictive models were developed: one considering only pre-surgery factors and the other focusing exclusively on post-surgery factors.

The results demonstrated that XGBoost and Random Forest were the most effective models, consistently achieving the highest accuracy and ROC AUC scores. Notably, both models identified DaysSinceSurgery as the most influential variable, reinforcing its importance in post-surgical risk assessment. In general, higher DaysSinceSurgery values were associated with an increased risk of adverse events, underlining the time-sensitive nature of post-operative monitoring. In addition to DaysSinceSurgery, the models highlighted the importance of patient age. However, the effect of age varied depending on the event: age was protective in the absence of events (Class 0), meaning younger patients were more likely to remain event-free, while it was a risk factor for Stroke (Class 2) and MACE (Class 3), where older patients had a higher likelihood of developing these complications. Beyond these temporal and demographic factors, pre-operative predictors such as BMI and Chronic Kidney Disease (CKD) were found to be highly relevant, particularly in the context of combined cardiovascular events. On the post-operative side, 30D PostOp Stroke emerged as a crucial factor, suggesting that the occurrence of a stroke within 30 days after surgery significantly increases the risk of subsequent adverse outcomes.

This study contributes to the growing body of literature on risk stratification following CEA by leveraging ML models to identify key prognostic factors based on both pre- and post-surgical characteristics. The developed models dynamically integrate temporal and physiological variables, enhancing predictive accuracy and supporting a more data-driven approach to patient management. By recognizing the influence of the days since surgery and the occurrence of a Stroke 30 days after surgery, this study reinforces the importance of ongoing post-operative surveillance, as these factors can provide early warning signs for complications. Additionally, the identification of chronic kidney disease as a critical risk factor suggests the need for specialized management protocols for patients with renal comorbidities. By analyzing separate pre- and post-operative predictors, this study helps bridge the gap between pre-surgical risk assessment and post-operative monitoring, promoting a more comprehensive and holistic approach to patient care.

The study had several limitations that must be acknowledged and besides that the results obtained have potential for improvement. Firstly, the dataset used was relatively small, which may limit the generalizability of the findings. A larger and more diverse patient cohort is necessary to enhance the reliability of the results and ensure broader applicability in clinical practice. Another limitation was the high-class imbalance in the target variables, which

affected the models' ability to correctly classify the minority classes. Future research should explore other advanced resampling techniques or cost-sensitive learning to improve model performance in predicting less frequent but clinically significant outcomes. Additionally, the integration of longitudinal patient data across multiple healthcare institutions could provide a richer dataset, allowing for a more refined analysis of risk factors over time.

BIBLIOGRAPHICAL REFERENCES

Acuña, E., & Rodriguez, C. (2004). The Treatment of Missing Values and its Effect on Classifier Accuracy. Classification, Clustering, and Data Mining Applications, 639–647. doi: 10.1007/978-3-642-17103-1_60

Administração Central do Sistema de Saúde (2018). Intervalos de referência (convencional). https://www.acss.min-saude.pt/wp-content/uploads/2018/09/Tabela_Final.pdf

Aggarwal, C. C. (2017). Outlier Analysis. Springer International Publishing. doi: 10.1007/978-3-319-47578-3

Ali, K., & Al-Hameed, A. (2022). Spearman's correlation coefficient in statistical analysis. Int. J. Nonlinear Anal. Appl, 13, 2008–6822. doi: 10.22075/ijnaa.2022.6079

Allain, R., Marone, L. K., Meltzer, J., & Jeyabalan, G. (2005). Carotid Endarterectomy. International Anesthesiology Clinics, 15–38, 43(1). doi: 10.1097/01.aia.0000150279.00355.6c

American Board of Internal Medicine. (2024). Laboratory Tests Reference Ranges. https://www.abim.org/Media/bfijryql/laboratory-reference-ranges.pdf

Bai, P., Zhou, Y., Liu, Y., Li, G., Li, Z., Wang, T., & Guo, X. (2020). Risk Factors of Cerebral Infarction and Myocardial Infarction after Carotid Endarterectomy Analyzed by Machine Learning. *Computational and Mathematical Methods in Medicine*, 1–8. doi: 10.1155/2020/6217392

Barbetta, I., Carmo, M., Giulio Mercandalli, P. Lattuada, Mazzaccaro, D., Settembrini, A. M., Raffaello Dallatana, & Piergiorgio Settembrini. (2014). Outcomes of urgent carotid endarterectomy for stable and unstable acute neurologic deficits. *Journal of Vascular Surgery*, 440–446, *59*(2). doi: 10.1016/j.jvs.2013.08.035

Boulesteix, A.-L., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(6), 493–507. doi: 10.1002/widm.1072

Buchlak, Q. D., Esmaili, N., Leveque, J.-C., Farrokhi, F., Bennett, C., Piccardi, M., & Sethi, R. K. (2020). Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. Neurosurgical Review,1235–1253, 43(5). doi: 10.1007/s10143-019-01163-8

Budholiya, K., Shrivastava, S. K., & Sharma, V. (2020). An optimized XGBoost based diagnostic system for effective prediction of heart disease. Journal of King Saud University - Computer and Information Sciences. doi: 10.1016/j.jksuci.2020.10.013

Cao, Y., Miao, Q.-G., Liu, J.-C., & Gao, L. (2014). Advance and Prospects of AdaBoost Algorithm. Acta Automatica Sinica, 39(6), 745–758. doi: 10.3724/sp.j.1004.2013.00745

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 408(1), 189–215. doi: 10.1016/j.neucom.2019.10.118

Chaudhry, S., Mohammad Rauf Afzal, Kassab, A., Hussain, S., & Qureshi, A. I. (2016). A New Risk Index for Predicting Outcomes among Patients Undergoing Carotid Endarterectomy in Large Administrative Data Sets. *Journal of Stroke and Cerebrovascular Diseases*, 1978–1983, *25*(8). doi: 10.1016/j.jstrokecerebrovasdis.2016.01.023

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16(16), 321–357. doi: 10.1613/jair.953

Dadebayev, D., Wei Wei, G., & Ee Xion, T. (2021). A Machine Learning Approach to EEG-based Prediction of Human Affective States Using Recursive Feature Elimination Method. MATEC Web of Conferences, 335, 04001. doi: 10.1051/matecconf/202133504001

Dasenbrock, H. H., Smith, T. R., Gormley, W. B., Castlen, J. P., Patel, N. J., Frerichs, K. U., M. Ali Aziz-Sultan, & Du, R. (2019). Predictive Score of Adverse Events After Carotid Endarterectomy: The NSQIP Registry Carotid Endarterectomy Scale. Journal of the American Heart Association, 8(21). doi: 10.1161/JAHA.119.013412

Dastres, R. (2021). Artificial Neural Network Systems [Review of Artificial Neural Network Systems]. International Journal of Imaging and Robotics, 21(2).

Dossabhoy, S., & Arya, S. (2021). Epidemiology of atherosclerotic carotid artery disease. Seminars in Vascular Surgery, 3–9, 34(1). doi: 10.1053/j.semvascsurg.2021.02.013

Dua, A., Romanelli, M., Upchurch, G. R., Pan, J., Hood, D., Hodgson, K. J., & Desai, S. S. (2016). Predictors of poor outcome after carotid intervention. Journal of Vascular Surgery, 663–670, 64(3). doi: 10.1016/j.jvs.2016.03.428

Feng, Y., Bai, X., Wang, T., Wang, Y., Chen, Y., Li, M., Hua, Y., Ling, F., Ma, Y., & Jiao, L. (2020). Thirty-day outcomes of carotid endarterectomy in the elderly: A 17-year single-center study. Journal of Clinical Neuroscience, 78, 86–90. doi: 10.1016/j.jocn.2020.06.007

Ferguson, G. G., Eliasziw, M., Barr, H. W. K., Clagett, G. P., Barnes, R. W., Wallace, M. C., Taylor, D. W., Haynes, R. B., Finan, J. W., Hachinski, V. C., & Barnett, H. J. M. (1999). The North American Symptomatic Carotid Endarterectomy Trial. *Stroke*, 1751–1758, *30*(9). doi: 10.1161/01.STR.30.9.1751

Ferri, C., Hernández-Orallo, J., & Modroiu, R. (2009). An experimental comparison of performance measures for classification. Pattern Recognition Letters, 30(1), 27–38. doi: 10.1016/j.patrec.2008.08.010

Gupta, B., Rawat, A., Jain, A., Arora, A., & Dhami, N. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. International Journal of Computer Applications, 163(8), 15–19. doi: 10.5120/ijca2017913660

Hastie, T., Friedman, J., & Tibshirani, R. (2001). The Elements of Statistical Learning : Data Mining, Inference, and Prediction. Springer. doi: 10.1007/978-0-387-84858-7

Howell, S. J. (2007). Carotid endarterectomy. British Journal of Anaesthesia, 119–131, 99(1). doi: 10.1093/bja/aem137

Joh, J. H., & Cho, S. (2020). Cardiovascular risk of carotid atherosclerosis: global consensus beyond societal guidelines. The Lancet Global Health, e625–e626, 8(5). doi: 10.1016/S2214-109X(20)30132-7

Kelley, R., & Bir, S. (2022). Carotid atherosclerotic disease: A systematic review of pathogenesis and management. Brain Circulation, 127, 8(3). doi: 10.4103/bc.bc_36_22

Kim, H.-Y. (2014). Analysis of Variance (ANOVA) Comparing Means of More than Two Groups. Restorative Dentistry & Endodontics, 39(1), 74. doi: 10.5395/rde.2014.39.1.74

Kimon Bekelis, Bakhoum, S. F., Desai, A., Mackenzie, T. A., Goodney, P., & Nicos Labropoulos. (2013). A Risk Factor-based Predictive Model of Outcomes in Carotid Endarterectomy. Stroke, 44(4), 1085–1090. doi: 10.1161/STROKEAHA.111.674358

Li, F., Zhang, R., Di, X., Niu, S., Rong, Z., Liu, C., & Ni, L. (2023). Diabetes mellitus and adverse outcomes after carotid endarterectomy: A systematic review and meta-analysis. Chinese Medical Journal, 136(12), 1401–1409. doi: 10.1097/CM9.0000000000002730

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2018). Feature Selection: A Data Perspective. ACM Computing Surveys, 50(6), 1–45. doi: 10.1145/3136625

Lobo, M., Mourão, J., & Afonso, G. (2015). Endarterectomia carotídea: revisão de 10 anos de prática de anestesia geral e locorregional num hospital terciário em Portugal. Brazilian Journal of Anesthesiology, 65(4), 249–254. doi: 10.1016/j.bjan.2014.03.010

Matsuo, K., Fujita, A., Hosoda, K., Tanaka, J., Taichiro Imahori, Ishii, T., Masaaki Kohta, Tanaka, K., Yoichi Uozumi, Kimura, H., Takashi Sasayama, & Eiji Kohmura. (2022). Potential of machine learning to predict early ischemic events after carotid endarterectomy or stenting: a comparison with surgeon predictions. *Neurosurgical Review*, 607–616, *45*(1). doi: 10.1007/s10143-021-01573-7

Mekke, J. M. et al. Plasma Lipoprotein Lipase Is Associated with Risk of Future Major Adverse Cardiovascular Events in Patients Following Carotid Endarterectomy. European Journal of Vascular and Endovascular Surgery, v. 65, n. 5, p. 700–709, 26 jan. 2023. doi: 10.1016/j.ejvs.2023.01.035.

M, H., & M.N, S. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process, 5(2), 01-11. doi: 10.5121/ijdkp.2015.5201

Muthukrishnan, R. & Rohini, R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning. 18-20. doi: 10.1109/ICACA.2016.7887916

Osborne, J., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). Practical Assessment, Research, and Evaluation, 9(6), 6. doi: 10.7275/qf69-7k43

Pereira-Neves, A., Rocha-Neves, J., Fragão-Marques, M., Duarte-Gamas, L., Jácome, F., Coelho, A., Cerqueira, A., Andrade, J. P., & Mansilha, A. (2021). Red blood cell distribution width is associated with hypoperfusion in carotid endarterectomy under regional anesthesia. Surgery (United States), 1536–1543, 169(6). doi: 10.1016/j.surg.2021.01.004

Pinheiro, I. M., Marques, G., Duarte, N., Gonçalves, A., Barroso, P., Gonzalez, A., & Ferreira, M. J. (2010). Patologia da carótida. Revista Portuguesa de Clínica Geral, 496–501, 26(5). doi: 10.32385/rpmgf.v26i5.10784

Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). (1998). The Lancet, 1379–1387, 351(9113). doi: 10.1016/s0140-6736(97)09292-1

Rocha-Neves, J., Pereira-Macedo, J., Ferreira, A., Dias-Neto, M., Andrade, J. P., & Mansilha, A. (2021). Impact of intraoperative neurologic deficits in carotid endarterectomy under regional anesthesia. Scandinavian Cardiovascular Journal, 180–186, 55(3). doi: 10.1080/14017431.2021.1874509

Rockman, C. B., Jacobowitz, G. R., Lamparello, P. J., Adelman, M. A., Woo, D., Schanzer, A., Gagne, P. J., Landis, R., & Riles, T. S. (2000). Immediate reexploration for the perioperative neurologic event after carotid endarterectomy: Is it worthwhile? Journal of Vascular Surgery, 1062–1070, 32(6). doi: 10.1067/mva.2000.111284

Román Salmerón Gómez, Beatriz, C., & García, J. (2020). Overcoming the inconsistences of the variance inflation factor: a redefined VIF and a test to detect statistical troubling multicollinearity. doi: 10.48550/arxiv.2005.02245

Saha, S. P., Saha, S., & Vyas, K. S. (2015). Carotid Endarterectomy: Current Concepts and Practice Patterns. The International Journal of Angiology, 223–235, 24(3). doi: 10.1055/s-0035-1558645

Scavuzzo, C. M., Scavuzzo, J. M., Micaela Natalia Campero, Melaku Anegagrie, Amor, A., Benito, A., & Periago, V. (2022). Feature importance: Opening a soil-transmitted helminth machine learning model via SHAP. Infectious Disease Modelling, 7(1), 262–276. doi: 10.1016/j.idm.2022.01.004

Schröer, Christoph, et al. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. Procedia Computer Science, vol. 181, no. 1, pp. 526–534. doi: 10.1016/j.procs.2021.01.199

Shaik, A. B., & Srinivasan, S. (2019). A Brief Survey on Random Forest Ensembles in Classification Model. International Conference on Innovative Computing and Communications, 56, 253–260. doi: 10.1007/978-981-13-2354-6_27

Sharma, S., Sharma, S., & Athaiya, A. (2020). ACTIVATION FUNCTIONS IN NEURAL NETWORKS. International Journal of Engineering Applied Sciences and Technology, 4(12), 310–316. doi: 10.33564/IJEAST.2020.v04i12.054

Smita, M. (2021). LOGISTIC REGRESSION MODEL –A REVIEW. International Journal of Innovative Science and Research Technology, 6(5)

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427–437. doi: 10.1016/j.ipm.2009.03.002

Song, P., Fang, Z., Wang, H., Cai, Y., Rahimi, K., Zhu, Y., Fowkes, F. G. R., Fowkes, F. J. I., & Rudan, I. (2020). Global and regional prevalence, burden, and risk factors for carotid atherosclerosis: a systematic review, meta-analysis, and modelling study. The Lancet Global Health, e721–e729, 8(5). doi: 10.1016/S2214-109X(20)30117-0

Taunk, K., De, S., Verma, S., & Swetapadma, A. (2019, May 1). A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. IEEE Xplore. doi: 10.1109/ICCS45141.2019.9065747

Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley Publishing Company. doi: 10.1002/bimj.4710230408

Volkers, E. J., Algra, A., Kappelle, L. J., Jansen, O., Howard, G., Hendrikse, J., Halliday, A., Gregson, J., Fraedrich, G., Eckstein, H.-H., Calvet, D., Bulbulia, R., Brown, M. M., Becquemin, J.-P., Ringleb, P. A., Mas, J.-L., Bonati, L. H., Brott, T. G., & Greving, J. P. (2018). Prediction Models for Clinical Outcome After a Carotid Revascularization Procedure. *Stroke*, 1880–1885, *49*(8). doi: 10.1161/STROKEAHA.117.020486

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295–316. doi: 10.1016/j.neucom.2020.07.061

Wickramasinghe, I., & Kalutarage, H. (2020). Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation. Soft Computing, 25(3), 2277–2293. doi: 10.1007/s00500-020-05297-6

Zahorec, R. (2021). Neutrophil-to-lymphocyte ratio, past, present and future perspectives. Bratislavske Lekarske Listy, 122(7), 474–488. doi: 10.4149/BLL_2021_078

APPENDIX A

Project No.: DDMKT2024-4-218792

Project Title: Neurological deficits after carotid cross-clamping during carotid endarterectomy under loco-regional anesthesia, study of pre-, intra-, and post-operative prognostic predictors

Principal Researcher: Filipa Mendes Afonso

According to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered APPROVED on 10/05/2024.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change.

- Any unforeseen events or unexpected developments that merit notification.

- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 10/05/2024

NOVA IMS Ethics Committee

ethicscommittee@novaims.unl.pt

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa