

Prediction of In-Hospital Atrial Fibrillation After Acute Myocardial Infarction

Matteo Bulloni¹, Guadalupe García-Isla¹, Pedro Moreno-Sánchez², Erica Rurali³, Alice Bonesi³,
Mattia Chiesa³, Pablo J Werba³, Giancarlo Marenzi³, Valentina Corino¹, Claudio Tondo³,
Mark van Gils², Linda Pattini¹, Luca Mainardi¹

¹ Politecnico di Milano, Milan, Italy

² Tampere University, Tampere, Finland

³ IRCCS Centro Cardiologico Monzino, Milan, Italy

Abstract

Atrial fibrillation (AF) is a relatively frequent complication of acute myocardial infarction (AMI). While AF prediction has been extensively studied, the identification of risk factors for early, new-onset AF (NOAF) after AMI in the intensive cardiac care unit (ICCU) remains less explored. Specifically, to our knowledge, there are no reported attempts at predicting in-hospital NOAF after AMI using machine learning. In this study, we developed a machine learning model to predict in-hospital NOAF following AMI. The dataset used for model development included 2445 consecutive AMI patients admitted to the ICCU of Centro Cardiologico Monzino, out of which 241 (9.9%) developed NOAF prior to ICCU discharge. Fifty-six features encompassing demographic and clinical variables were retrospectively collected and analysed. Several data balancing, feature selection and classification techniques were evaluated and compared by means of area under the ROC curve (AUROC) through nested cross-validation. The best-performing model combined an undersampling step, based on the Edited Nearest Neighbors algorithm, a mutual-information-based feature selection and a logistic regression model. The model achieved an AUROC of 0.765 (95% CI: 0.732 – 0.795), exploiting both known and previously unreported markers.

1. Introduction

Acute myocardial infarction (AMI) represents one of the leading causes of death and the main cause for cardiovascular death worldwide [1]. Atrial fibrillation (AF) is a relatively common complication of AMI, affecting between 2% and 22% of AMI cases [2], and is associated with increased mortality [3]. New-onset AF (NOAF), without history of AF prior to the AMI, has been shown to remain independently associated with increased risk of stroke, recurrence and death [3–6]. Although AF prediction has

been widely researched [7, 8], limited attention has been given to identifying risk factors for early, new-onset AF (NOAF) following AMI in the intensive cardiac care unit (ICCU) [2, 9–13]. In particular, to our knowledge, no studies have employed machine learning techniques to predict in-hospital NOAF after AMI. We aimed, therefore, at filling this gap with the present study.

2. Materials and Methods

2.1. Study population

A dataset of 2445 consecutive patients admitted between 2010 and 2018 to the ICCU of Centro Cardiologico Monzino following AMI (ICCU admission within 24 hours from the AMI event) was retrospectively gathered. The dataset included fifty-six features encompassing demographic, clinical history, therapy at admission, in-hospital complications and other clinical data. Out of 2445 patients, 241 (9.9%) developed NOAF prior to ICCU discharge, with NOAF defined as occurrence of AF in the period after hospital admission and before hospital discharge in a patient without known chronic permanent AF or AF at admission. Table 1 provides an overview of the main clinical variables' value distributions of the dataset, grouped by prediction class - did or did not present NOAF.

2.2. Learning pipeline

To develop the prediction model, an extensive set of pre-processing, data balancing, feature selection and classification techniques were evaluated. Figure 1 displays the general sequential structure of the learning pipeline; depending on the classifier, the feature and data balancing processes could be implicit. The different learning configurations were evaluated via nested cross-validation (nested CV), to overcome test set selection bias, while performing hyperparameter selection with balanced accuracy as optimization metric. Randomized search, limited to 100

Table 1. Value distributions of the main clinical variables by outcome class. Continuous variables are presented as either mean $\pm$ standard deviation or with median and interquartile ranges, in case of particularly skewed distributions. Categorical data are presented as the number of subjects fulfilling the criterion, with the corresponding percentage value. P-values were obtained comparing the SR (sinus rhythm) and AF distributions via the Mann-Whitney U test for continuous variables, and the Chi-squared test for binary ones.

Variable	SR (n = 2204)	AF (n = 241)	p-value
Age (years)	66 $\pm$ 12	75 $\pm$ 10	<0.0001
Male sex, <i>n</i> (%)	1647 (75%)	159 (66%)	0.003
Body weight (Kg)	76 $\pm$ 14	74 $\pm$ 16	0.01
Hypertension, <i>n</i> (%)	1398 (63%)	194 (80%)	<0.0001
Smoking history, <i>n</i> (%)	1191 (54%)	98 (31%)	<0.0001
Prior myocardial infarction, <i>n</i> (%)	571 (26%)	77 (32%)	0.004
Prior coronary artery bypass graft, <i>n</i> (%)	255 (12%)	39 (16%)	0.03
Prior percutaneous coronary intervention, <i>n</i> (%)	578 (26%)	68 (28%)	0.48
Left ventricular ejection fraction (%)	51 $\pm$ 11	43 $\pm$ 13	<0.0001
Atrioventricular Block/Pacemaker, <i>n</i> (%)	69 (3%)	18 (7%)	0.001
Laboratory values at hospital admission			
Serum creatinine (mg/dL)	1.02 $\pm$ 0.44	1.23 $\pm$ 0.82	<0.0001
Estimated glomerular filtration rate (mL/min/1.73 m ²)	79 $\pm$ 26	68 $\pm$ 29	<0.0001
Hemoglobin (g/dL)	13.8 $\pm$ 1.8	13.2 $\pm$ 1.9	<0.0001
Blood glucose (mg/dL)	147 $\pm$ 60	166 $\pm$ 72	<0.0001
High-sensitivity troponin I (ng/L)	400 (70–2340)	1064 (130–6690)	0.0001*
High-sensitivity C-reactive protein (mg/L)	3.39 (1.37–10.71)	7.06 (2.12–36.41)	<0.0001*
Medication at hospital admission			
Aspirin, <i>n</i> (%)	822 (37%)	100 (41%)	0.2
Statins, <i>n</i> (%)	747 (34%)	99 (42%)	0.02
Beta-blockers, <i>n</i> (%)	772 (35%)	104 (43%)	0.001
Angiotensin converting enzyme/Angiotensin receptor blockers, <i>n</i> (%)	859 (39%)	121 (50%)	0.001
In-hospital complications			
Cardiogenic shock, <i>n</i> (%)	86 (4%)	35 (15%)	<0.0001
Acute pulmonary edema, <i>n</i> (%)	150 (7%)	76 (32%)	<0.0001
Ventricular tachycardia/ventricular fibrillation, <i>n</i> (%)	147 (7%)	32 (13%)	0.0001
Acute Kidney Injury, <i>n</i> (%)	191 (9%)	63 (26%)	<0.0001
Major bleeding, <i>n</i> (%)	50 (2%)	26 (11%)	<0.0001

combinations, was employed for hyperparameter tuning. Reported performances throughout the paper are those obtained on unseen data during nested CV, with confidence intervals obtained through bootstrapping (1000 iterations). The final model was obtained by re-training the best performing pipeline in nested CV on the entire dataset, with hyperparameter values tuned with a regular 10-fold CV on the whole dataset. The list of features identified as predictive and the hyperparameter values mentioned throughout the paper are those of this final model.

All code was implemented in Python 3.11.5, mostly relying on the Scikit-Learn 1.3.0 and Imbalanced-Learn 0.12.0 libraries.

3. Results

The best-performing model resulted to be the one employing a mutual-information-based feature selection (nonparametric implementation of Scikit-Learn) and a logistic regression classifier, and achieved an AUROC of 0.765 (95% CI: 0.732 - 0.795). Among oversampling (SMOTE [14] algorithm) and undersampling (Edited Nearest Neighbors [15] algorithm (ENN)), undersampling resulted the most effective data balancing approach for all combinations of feature selection techniques and classifiers evaluated, with the number of neighbors for ENN tuned to 5 for the final model. This rather heavy undersampling is understandable due to the high unbalance in the dataset, with a ratio between AF and no AF patients equal to 10.9%. No correlation filtering ended up being ap-

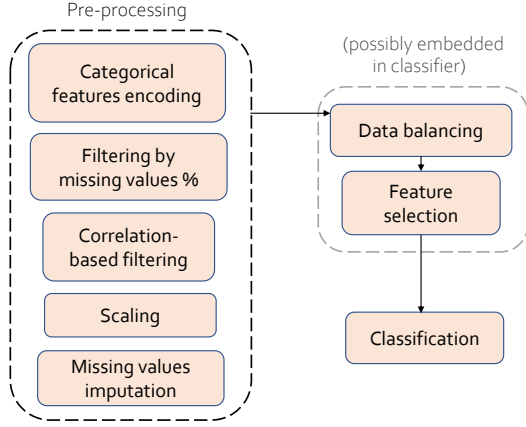


Figure 1. General structure of the learning pipeline.

plied to the features in the final model. Figure 2 displays the ROC curve of all pipelines that achieved an AUROC greater than 0.73. Additional metrics with respective confidence intervals for the best-performing model are reported in Table 2. Figure 3 reports instead the list of features employed by the final model, with the respective importance values.

4. Discussion and Conclusion

The best performing model, consisting of a pipeline of an ENN undersampling, a mutual-information-based feature selection and a Logistic Regression classifier, achieved an AUROC of 0.765. This result appears aligned

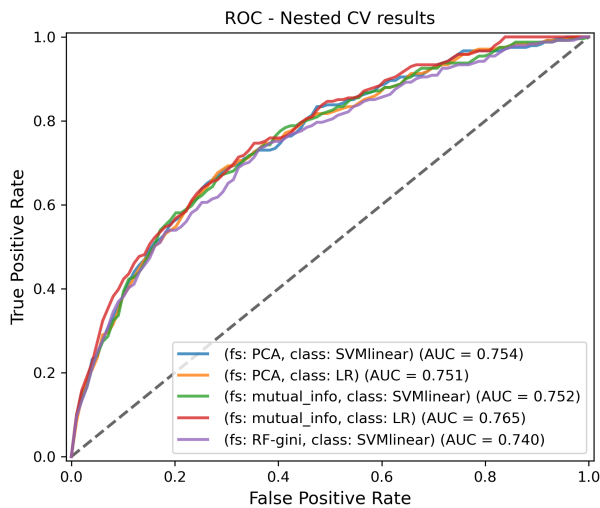


Figure 2. ROC curves of the best performing pipelines in nested CV.

Table 2. Performance of the best pipeline, with 95% confidence intervals.

Metric	Value [95% CI]
AUROC	0.765 [0.732 - 0.795]
Balanced accuracy	0.692 [0.661 - 0.722]
Sensitivity	0.652 [0.592 - 0.713]
Specificity	0.731 [0.712 - 0.750]
Precision	0.209 [0.180 - 0.240]
F1-score	0.317 [0.278 - 0.355]

with what obtained by two other studies predicting NOAF after AMI [2, 11]. Fu et al. [11] reached an AUROC of 0.791 in their work, although the smaller dataset implied much larger confidence intervals (0.692–0.891). Tu et al. [2] reported an AUROC of 0.818, although working on a dataset of only 333 patients and possibly adopting biased evaluation procedure. Our result, obtained on a dataset of 2445 patients, provides arguably stronger evidence of the actual predictive power of routine clinical variables for NOAF after AMI. Predictably, age and ejection fraction resulted among the most significant predictors, as already demonstrated in other works’ multivariate analyses [2, 10, 12]. Interestingly, sex and high-sensitivity C-reactive protein (HsCRP) were not selected as predictors, although being so in other studies [9, 11]. Among the novel markers selected by our model, it is worth to highlight acute pulmonary edema, ranked first among all predictive variables, and serum creatinine, present both with its value at ICCU admission and with its delta value between admission and peak recorded one, where either variable represents a proxy of a high inflammatory status. In spite of these findings and considerations, it is worth to underline how the proposed model’s performance remained modest, far from an acceptable level for clinical applications, similarly to the other proposed predictors. NOAF prediction after AMI, especially in absence of ECG data, remains extremely challenging, and further research is needed to address the problem.

Acknowledgments

This work is funded by the project PerCard (Personalised Prognostics and Diagnostics for Improved Decision Support in Cardiovascular Diseases) in ERA PerMed supported by the Research Council of Finland (decision number 351846) and by the Fondazione Regionale della Ricerca Biomedica (FRRB).

References

- [1] Salari N, Morddarvanjoghi F, Abdolmaleki A, Rasoulpoor S, Khaleghi AA, Hezarkhani LA, Shohaimi S, Mohammadi

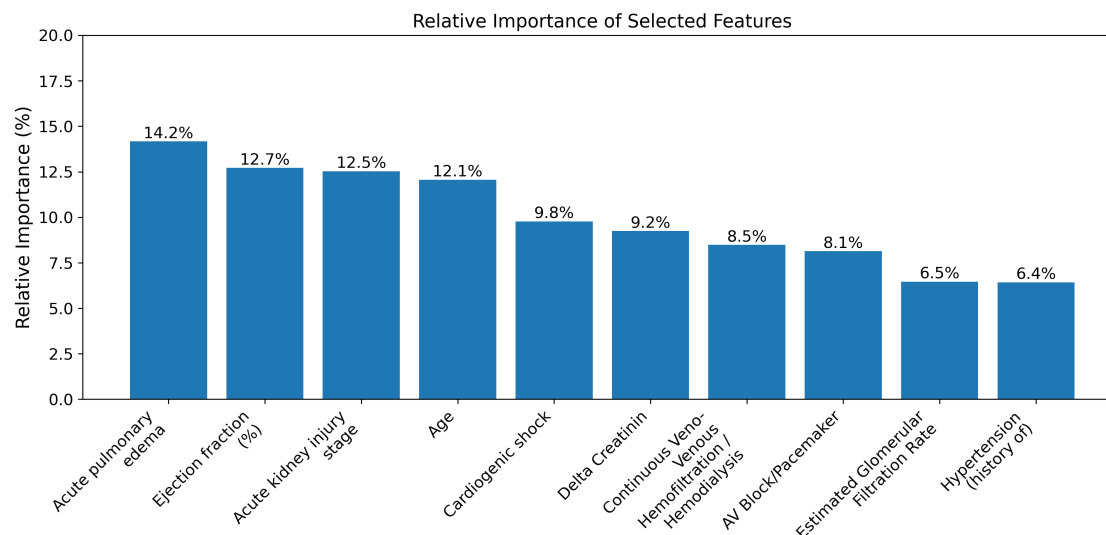


Figure 3. Relative importance of the features retained by the final model (values add up to 100%).

- M. The Global Prevalence of Myocardial Infarction: a Systematic Review and Meta-Analysis. *BMC Cardiovascular Disorders* 12 2023;23:1–12.
- [2] Tu J, Ye Z, Cao Y, Xu M, Wang S. Establishment and Evaluation of a Nomogram for In-Hospital New-Onset Atrial Fibrillation After Percutaneous Coronary Intervention for Acute Myocardial Infarction. *Frontiers in Cardiovascular Medicine* 2024;11:1370290.
 - [3] Jabre P, Roger VL, Murad MH, Chamberlain AM, Prokop L, Adnet F, Jouven X. Mortality Associated with Atrial Fibrillation in Patients with Myocardial Infarction: a Systematic Review and Meta-Analysis. *Circulation* 4 2011; 123:1587.
 - [4] Bishara R, Telman G, Bahouth F, Lessick J, Aronson D. Transient Atrial Fibrillation and Risk of Stroke After Acute Myocardial Infarction. *Thrombosis and haemostasis* 11 2011;106:877–884.
 - [5] Siu CW, Jim MH, Ho HH, Miu R, Lee SW, Lau CP, Tse HF. Transient Atrial Fibrillation Complicating Acute Inferior Myocardial Infarction: Implications for Future Risk of Ischemic Stroke. *Chest* 2007;132:44–49.
 - [6] Schmitt J, Duray G, Gersh BJ, Hohnloser SH. Atrial Fibrillation in Acute Myocardial Infarction: a Systematic Review of the Incidence, Clinical Features and Prognostic Implications. *European heart journal* 5 2009;30:1038–1045.
 - [7] Tseng AS, Noseworthy PA. Prediction of Atrial Fibrillation Using Machine Learning: a Review. *Frontiers in Physiology* 10 2021;12:752317.
 - [8] Matias I, Garcia N, Pirbhulal S, Felizardo V, Pombo N, Zacarias H, Sousa M, Zdravevski E. Prediction of Atrial Fibrillation Using Artificial Intelligence on Electrocardiograms: a Systematic Review. *Computer Science Review* 2 2021;39:100334.
 - [9] Shiyovich A, Axelrod M, Gilutz H, Plakht Y. Early Versus Late New-Onset Atrial Fibrillation in Acute Myocardial Infarction: Differences in Clinical Characteristics and Predictors. *Angiology* 11 2019;70:921–928.
 - [10] Raczkowska-Golanko M, Raczak G, Gruchała M, Daniłowicz-Szymanowicz L. Comprehensive Use of Routine Clinical Parameters to Identify Patients at Risk of New-Onset Atrial Fibrillation in Acute Myocardial Infarction. *Journal of Clinical Medicine* 8 2021;10:3622. Early, in-hospital.
 - [11] Fu Y, Pan Y, Gao Y, Yang X, Chen M. Predictive Value of CHA2DS2-VASc Combined with hs-CRP for New-Onset Atrial Fibrillation in Elderly Patients with Acute Myocardial Infarction. *BMC cardiovascular disorders* 12 2021;21.
 - [12] Sagnard A, Guenancia C, Mouhat B, Maza M, Fichot M, Moreau D, Garnier F, Lorgis L, Cottin Y, Zeller M. Involvement of Autonomic Nervous System in New-Onset Atrial Fibrillation During Acute Myocardial Infarction. *Journal of clinical medicine* 5 2020;9.
 - [13] Luo J, Li Z, Qin X, Zhang X, Liu X, Zhang W, Xu W, Liu B, Wei Y. Association of Stress Hyperglycemia Ratio with In-Hospital New-Onset Atrial Fibrillation and Long-Term Outcomes in Patients with Acute Myocardial Infarction. *DiabetesMetabolism Research and Reviews* 2 2024;40:e3726.
 - [14] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. Synthetic Minority Over-Sampling Technique. *Journal Of Artificial Intelligence Research* 6 2011;16:321–357.
 - [15] Edited Nearest Neighbors, Accessed August 29th, 2024. URL imbalancedlearn.org/stable/references/generated/imblearn.under_sampling.EditedNearestNeighbours.html.

Address for correspondence:

Matteo Bulloni
Building 21, via Golgi 39, 20133, Milan, Italy
matteo.bulloni@polimi.it