Machine learning approaches as an alternative to traditional statistical methods in cardiovascular risk prediction

Machine learning approaches as an alternative to traditional statistical methods in cardiovascular risk prediction

Michela Sperti,[1] Fabrizio D'Ascenzo,[2] Luca Navarini,[3] Giacomo Di Benedetto,[4] Antonella Afeltra,[3] Roberto Giacomelli,[5] Marco A. Deriu[1]

[1]PolitoBIOMedLab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin; [2]Division of Cardiology, Cardiovascular and Thoracic Department, Città della Salute e della Scienza, Turin; [3]Unit of Allergology, Immunology, Rheumatology, Department of Medicine, Università Campus Bio-Medico di Roma, Rome; [4]7HC srl, Rome; [5]Rheumatology Unit, Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, Aquila, Italy

Abstract

Machine Learning (ML) algorithms have proven promising methodologies in improving Cardiovascular (CV) risk predictors based on traditional statistics. In the present work, two case studies are reported: CV risk prediction in patients affected by Inflammatory Arthritis (IA), with attention to Psoriatic Arthritis (PsA), and patients who experienced Acute Coronary Syndrome (ACS).

Introduction

CV diseases are the leading cause of death worldwide. Early detection is of crucial importance to correctly treat patients and prioritized decisions. Addressing behavioral risk factors (such as tobacco use, unhealthy diet, and obesity, physical inactivity, etc.) can prevent most CV diseases. CV comorbidities deeply affect the quality of life of patients affected by IA, while patients which experienced ACS are at high risk for ischemic and bleeding events. In both scenarios, traditional CV risk scores underestimated the CV risk and provided an overall insufficient accuracy.[1,2] ML methods represent a promising approach to overcome some of the limitations of traditional techniques. They are applied to large datasets and can better capture the high-dimensional, non-linear relationships among clinical features. This study aims at comparing the performances of already existing models with results obtained by ML techniques in two medical scenarios and discuss their potentialities.[3-5]

Materials and Methods

For the IA case study, Random Forest (RF), Support Vector Machine (SVM) And K-Nearest Neighbors (KNN) ML classifiers were compared to Framingham Risk Score (FRS) performances. The algorithms were trained on the freely available Framingham dataset and evaluated on two cohorts of patients affected by PsA (155 subjects). For the ACS case study, KNN and Naïve Bayes (NB) classifiers were used to develop the final model. The algorithms were trained on a cohort of 19826 adult patients with ACS and evaluated on an external cohort of 3444 patients. Algorithms' performances were evaluated both in terms of their discrimination abilities (receiver operating characteristic curve and area under the curve, sensitivity, specificity, accuracy) and calibration abilities (calibration plot). Relative variables importances were calculated in both scenarios.

Correspondence:Michela Sperti, PolitoBIOMedLab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy.
michela.sperti@polito.it

Key words: Machine learning; cardiovascular risk prediction; cardiology; rheumatology; decision support systems.

Disclosures: The authors have declared no conflicts of interest.

Conference presentation: This paper was presented at the Third Centro 3R Annual Meeting - L'era delle 3R: modelli *in silico, in vitro* e *in vivo* per promuovere la ricerca traslazionale - 30 September - 1 October 2021, Evento online organizzato dal Politecnico di Torino.

Received for publication: 9 July 2021.
Accepted for publication: 7 September 2021.

Biomedical Science and Engineering 2021; 4(s1):195
doi:10.4081/bse.2021.195

Results

Results of IA case study are shown in Figure 1 and Table 1. ML models revealed a higher sensitivity in patients affected by PsA. The features analysis results shed light on the most important variables in prediction, which are also considered by the traditional score FRS, such as Systolic Blood Pressure (SBP), hypertension treatment, and age. Moreover, several psoriatic arthritis features such as C-Reactive Protein (CRP) and Psoriatic Area Severity Index (PASI) emerged as key predictors.

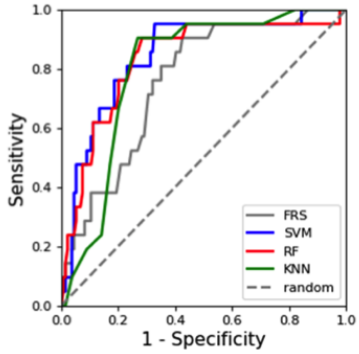

Figure 1. Receiver operating characteristic curves for algorithms evaluation and comparison in the case of patients affected by PsA.

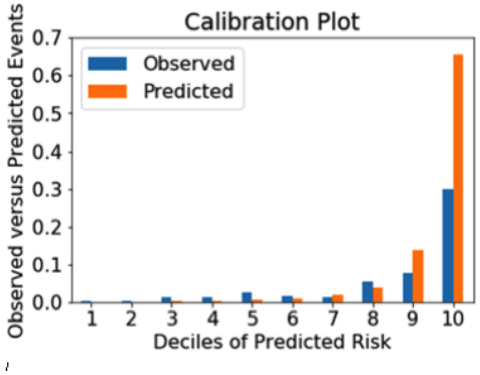

Figure 2. Calibration plot for death prediction. Probabilities have been obtained by applying the NB classifier.

Table 1. Performance metrics for algorithms evaluation and comparison in the case of patients affected by PsA.

	SVM	RF	KNN	FRS
Sensitivity (%)	86	86	90	38
Specificity (%)	68	73	61	86
Accuracy (%)	70	75	65	79

A limitation of this study is the PsA dataset's size (155 patients), too small to allow the training and validation of an ML algorithm specifically on the PsA dataset. Further studies should consider a larger PsA dataset.

In the ACS case study, ML models presented excellent performances in terms of discrimination abilities for the prediction of patients' death (Figure 2), ischemic risk (reami), and bleeding risk (barc-MB).

Concerning calibration ability, the stratification of risk deciles for all three outcomes was considered as follows: low risk - first to sixth deciles; intermediate risk - seventh to ninth deciles; and high risk - tenth decile. A gradual and progressive increase in absolute event rates was observed across risk classes for the three outcomes.

A limitation of this study is the overestimation of risk in the last decile by the NB classifier. Further algorithm improvements may consider mixing this strategy with others that can privilege underestimation, like the Adaptive Boosting classifier.

Discussion and Conclusions

ML techniques are proposed as a promising approach to improve traditional methods in terms of cardiovascular events classification, probabilities calibration, and elucidation of the importance of clinical variables for diagnosis. Two different case studies have been reported as successful examples of ML application in the clinical field.

References

1. Navarini L, Margiotta DPE, Caso F, Currado D, Tasso M, Angeletti S, Ciccozzi M, Scarpa R, Afeltra A, Costa L. Performances of five risk algorithms in predicting cardiovascular events in patients with Psoriatic Arthritis: An Italian bicentric study. PLoS One 2018;13:e0205506.
2. Raposeiras-Roubína S, Abu-Assia E, D'Ascenzo F, et al. Annual incidence of confirmed stent thrombosis and clinical predictors in patients with acs treated with ticagrelor or prasugrel. Rev Esp Cardiol (Engl Ed) 2019;72:215-23.
3. Navarini L, Sperti M, Currado D, et al. A machine-learning approach to cardiovascular risk prediction in psoriatic arthritis. Rheumatol (Oxford) 2020;59:1767-9.
4. Navarini L, Caso F, Costa L, et al. Cardiovascular risk prediction in ankylosing spondylitis: from traditional scores to machine learning assessment. Rheumatol Ther 2020;7:867-82.
5. D'Ascenzo F, De Filippo O, Gallone G, et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. Lancet 2021;397:199-207.