ESC European Society of Cardiology

European Heart Journal: Acute Cardiovascular Care (2021) **10**, 866–868
doi:10.1093/ehjacc/zuab074

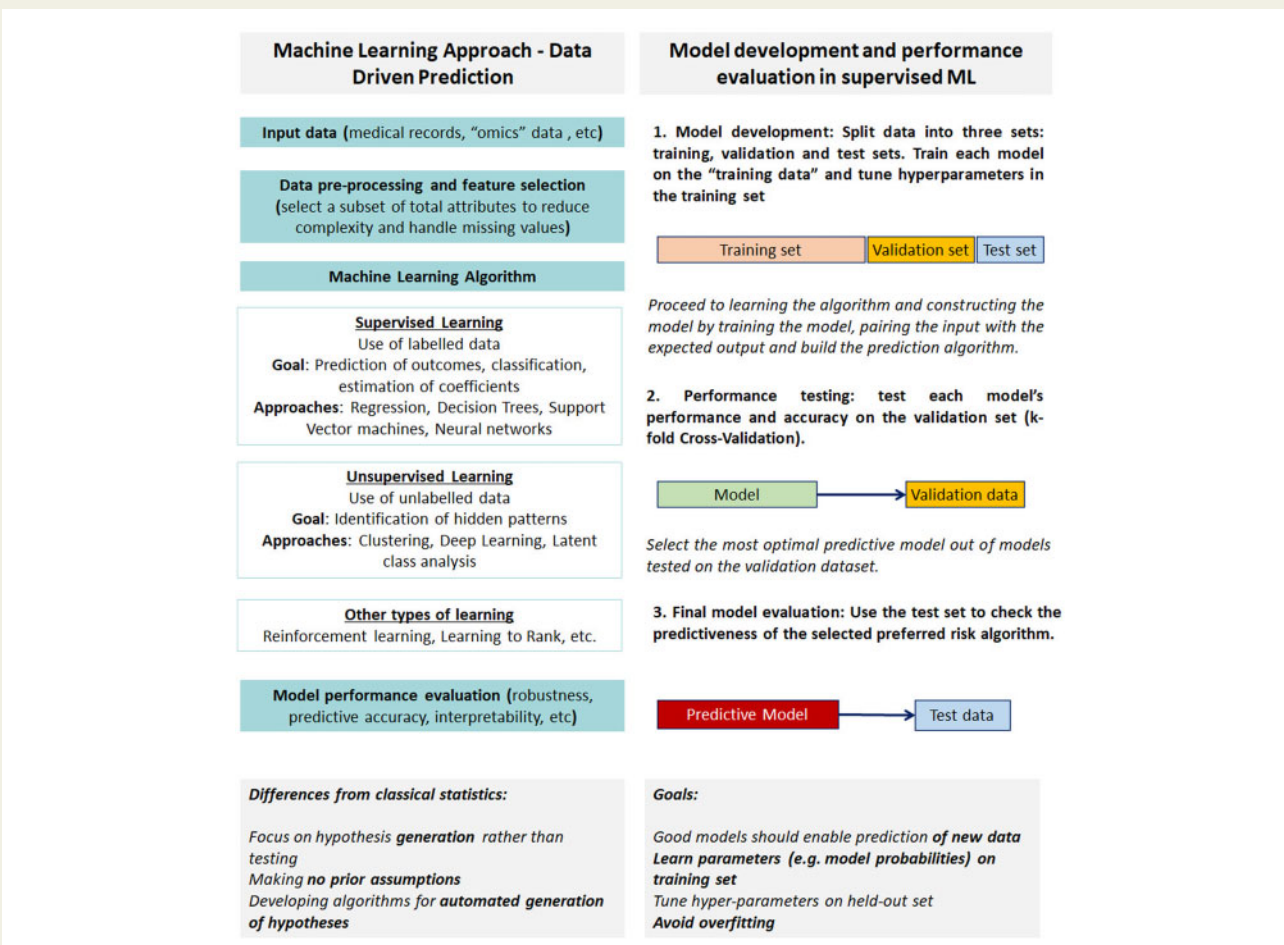
EDITORIAL

Challenges and promises of machine learning-based risk prediction modelling in cardiovascular disease

Maribel González-Del-Hoyo[1] and Xavier Rossello [1,2,3,4]*

[1]Cardiology Department, Institut d'Investigació Sanitària Illes Balears (IdISBa), Hospital Universitari Son Espases, Palma, Spain; [2]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain; [3]Facultad de Medicina, Universitat de les Illes Balears (UIB), Palma, Spain; and [4]Medical Statistics Department, London School of Hygiene & Tropical Medicine (LSHTM), London, UK

Online publish-ahead-of-print 28 August 2021



Graphical abstract Machine learning (ML) approach and model development and performance in supervised ML. ML workflow begins with data acquisition and preparation, feature selection, application of the ML algorithm, followed by model training, validation, and testing.

The opinions expressed in this article are not necessarily those of the Editors of the *European Heart Journal: Acute Cardiovascular Care* or of the European Society of Cardiology.

* Corresponding author. Tel: +34 871 20 52 34; Email: fjrossello@cnic.es

This editorial refers to 'Improving 1-year mortality prediction in ACS patients using machine learning', by S. Weichwald *et al.*, doi: 10.1093/ehjacc/zuab030.

The era of big data

'Big data' refers to data with high volume, variety, value, and rapid accumulation.[1] The exponential availability of electronic health records, large clinical studies, and biobanks, as a source of 'big data', has been increasing in the past decade.[2] To make the most of this vast amount of information, novel computational techniques have replaced conventional statistics, improving diagnostic accuracy (e.g. heart failure phenotypes[3]), and risk prediction modelling (e.g. death in patients with coronary heart disease[4]). Unlike standard multivariable regression methods, big data analysis can potentially use millions of variables, with billions of permutations, to develop dynamic predictive models that might need to be constantly updated by newly added information. Artificial intelligence (AI) in general, and machine learning (ML) in particular, have enabled us to tackle this challenge.[5]

Machine learning models

A risk prediction model refers to the mathematical function relating an outcome to a set of predictors (covariates). This approach estimates the probability (or risk) for an outcome to happen within a specific time period in a subject with a particular predictor profile.[6] Traditional regression models assume that each predictor is related in a linear fashion to the outcome, oversimplifying complex relationships with non-linear interactions. In contrast, the techniques from the field of AI, such as ML, offer an alternative approach to more generalizable risk prediction models. Machine learning reduces the error between predicted and observed events through the study of pattern recognition and computational learning (so-called 'artificial intelligence')—'learning' from an input training dataset to extract all complex and non-linear interactions between predictors.[7] Machine learning is related to automatic analysis ('data-mining'), which can be performed with minimal human input and incorporates new data to update and optimize its algorithms.[8] This approach can consider a greater number of variables and obtain better model performance compared to traditional techniques.[9]

Machine learning models usually split data into a training dataset (aimed to test an algorithm to recognize a pattern), a validation set (aimed to select and tune the ML model), and finally, the testing set (aimed to evaluate the performance of the algorithm). In ML, there are a set of hyperparameters, which are completely external to the model (remain unchanged) and must be tuned before the learning process begins; and there are also parameters internal to the model, which are the weights or coefficients estimated purely from the training data and are updated during the training process. Then, the model is iteratively trained and validated. To reduce the effects of randomness k-fold cross-validation is used. Once a model has been derived, its performance and external validation should be evaluated in other datasets. Guidelines for optimal reporting of prediction models when using ML techniques are expected to be published soon.[10] For those unfamiliar with ML, recent tutorials and reviews about ML are available elsewhere.[11,12]

Machine learning approaches can be driven by supervised and unsupervised learning. In supervised learning, an algorithm is trained or 'supervised' on a labelled dataset to recognize patterns or predict outcomes (Graphical Abstract).[13] Labelling should be performed by humans. The model then uses what it has learned from the training dataset to assign test data into specific categories or for regression modelling. The most commonly used supervised ML models are logistic regression, random forests/decision trees, artificial neuronal networks, gradient boosting machines, and support vector machines.[12] In unsupervised learning, however, there is no distinction between training and test data, and there is no data labelling from humans. We observe only the features, and make no outcome measure, as the goal is to describe the associations and patterns among a set of input data. A commonly used unsupervised ML model is clustering.

Artificial intelligence in cardiology

Traditional risk prediction tools have modest predictive power to accurately identify high-risk patients and an increasing body of evidence is supporting the use of big data techniques to predict outcomes in cardiovascular research.[4,14–16] In this issue of the journal, Weichwald *et al.*[17] described the application of ML to create a new risk score—the SPUM-ACS Score—outperforming the GRACE 2.0 score in terms of risk stratification in acute coronary syndrome (ACS) patients. The complex robust 8-variable model was chosen through iterative model development (1.4 billion linear models) by selecting out of 56 candidate predictors, 17 GRACE-outperforming quintuplets of practical relevant variables which were highly predictive and robust in a multitude of models. Compared with the performance of the existing GRACE 2.0 score, ML exhibited higher discrimination (c-statistic) for predicting 1-year all-cause mortality after five-fold cross-validation. Importantly, the five-fold cross-validated area under the curve score provided an estimate of the out-of-sample performance less prone to overfitting and upwards bias. Generalized Mallows rank model was used for distributions on permutations of the 56 variables and for model selection and a validation set was used to estimate overall model performance. By calculating the fraction of GRACE-outperforming models containing the respective variables, a ranking was elaborated which allowed to identify heart failure and inflammation markers, as the major prognostic determinants of 1-year combined events. The authors describe that, in contrast to other ML predictive models,[18] their approach uses multivariate interactions to optimally identify predictors of death, which otherwise would stay unobservable. Machine learning-based correlation analysis through cluster analysis identified which high-ranked variables were not selected in the eight-variable model due to overlap in prognostic value with other variables and which predictor variables were stable across all models and complemented another specific feature. The article of Weichwald *et al.*[17] is an excellent example of novel state-of-the-art ML prediction approaches outperforming already an existing traditional regression model-based score, by not being restricted to fixed linear associations between covariates.

Limits of artificial intelligence

However, the mere presence of large databases and innovative computational techniques has not (yet) been enough to translate their use in concrete meaningful clinical changes, such as a reduction in clinical outcomes through a better discrimination and accuracy in risk prediction. ML limitations are as follows (i) the bias–variance trade-off: to achieve a model with low variance and low bias to be highly accurate, too complex models are prone to overfitting. Independent training and cross-validation data sets for model testing and verification tackle this issue.[14] (ii) Machine learning algorithms require large datasets to reach acceptable performance levels, but inadequate data acquisition and bias will lead to incorrect results. (3) The 'black box' nature of unsupervised ML based decision making (not being able to know how the algorithm is achieving what it is achieving) is far from being understood, however, designing models that are inherently interpretable could potentially solve this issue.[19] (4) As big data analysis is not generally focused on causal inference, but rather on correlation or on identifying unseen or invisible patterns, the human factor is yet critical to detect spurious associations due to modifiable variables. (5) The lack of validated AI models makes it necessary to develop prospective randomized trials to ensure the accuracy and safety of ML before extending their application in clinical practice.

Artificial intelligence promises

Artificial intelligence is a rapidly evolving discipline with the potential to revolutionize the cardiovascular research, as ML approaches enable to analyse massive, large datasets, uncover new associations, and improve previous prognostic models, and diagnostic accuracy. However, most researchers have not adequately trained for the 'Big Data' era, as basic principles of epidemiology and biostatistics do not apply in this new setting. As we move on to the next era, future cardiology fellows-in-training should learn skills and gain knowledge in big data concepts. As illustrated by Weichwald *et al.*[17] with the development of a highly accurate SPUM-ACS Score through ML algorithms and feature ranking analysis, the development of accurate prediction tools is yet an unmet clinical need.[6] Cardiovascular medicine has already been challenged to grow with AI—the 'invisible' insights generated from these new techniques can lead to highly accurate predictive models .

Conflict of interest: none declared.

References

1. Wang L, Alexander CA. Big data in medical applications and health care. *Curr Res Med* 2015;**6**:1–8.
2. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, van Thiel GJM, Cronin M, Brobert G, Vardas P, Anker SD, Grobbee DE, Denaxas S; Innovative Medicines Initiative 2nd programme, Big Data for Better Outcomes, BigData@Heart Consortium of 20 academic and industry partners including ESC. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. *Eur Heart J* 2018;**39**:1481–1495.
3. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiade M, Bonow RO, Huang C-C, Deo RC. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015;**131**:269–279.
4. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, Andreini D, Budoff MJ, Cademartiri F, Callister TQ, Chang HJ. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;**38**:500–507.
5. Mayer-Schönberger V. Big data for cardiology: novel discovery? *Eur Heart J* 2016;**37**:996–1001.
6. Rossello X, Dorresteijn JA, Janssen A, Lambrinou E, Scherrenberg M, Bonnefoy-Cudraz E, Cobain M, Piepoli MF, Visseren FL, Dendale P. Risk prediction tools in cardiovascular disease prevention: a report from the ESC Prevention of CVD Programme led by the European Association of Preventive Cardiology (EAPC) in collaboration with the Acute Cardiovascular Care Association (ACCA) and the Association of Cardiovascular Nursing and Allied Professions (ACNAP). *Eur Heart J Acute Cardiovasc Care* 2020;**9**:522–532.
7. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform* 2002;**35**:352–359.
8. Mjolsness E, DeCoste D. Machine learning for science: state of the art and future prospects. *Science* 2001;**293**:2051–2055.
9. Dimopoulos AC, Nikolaidou M, Caballero FF, Engchuan W, Sanchez-Niubo A, Arndt H, Ayuso-Mateos JL, Haro JM, Chatterji S, Georgousopoulou EN, Pitsavos C, Panagiotakos DB. Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. *BMC Med Res Methodol* 2018;**18**:179.
10. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, Logullo P, Beam AL, Peng L, Van Calster B, van Smeden M, Riley RD, Moons KG. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* 2021;**11**:e048008.
11. Liu Y, Chen P-HC, Krause J, Peng L. How to read articles that use machine learning: users' guides to the medical literature. *JAMA* 2019;**322**:1806–1816.
12. Al'Aref SJ, Anchouche K, Singh G, Slomka PJ, Kolli KK, Kumar A, Pandey M, Maliakal G, van Rosendael AR, Beecy AN, Berman DS, Leipsic J, Nieman K, Andreini D, Pontone G, Schoepf UJ, Shaw LJ, Chang H-J, Narula J, Bax JJ, Guan Y, Min JK. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 2019;**40**:1975–1986.
13. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, Carter R, Chareonthaitawee P, Jouni H, Kapa S, Lerman A, Luong C, Medina-Inojosa JR, Noseworthy PA, Pellikka PA, Redfield MM, Roger VL, Sandhu GS, Senecal C, Friedman PA. Artificial intelligence in cardiology: present and future. *Mayo Clin Proc* 2020;**95**:1015–1039.
14. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;**12**:e0174944.
15. Sánchez-Cabo F, Rossello X, Fuster V, Benito F, Manzano JP, Silla JC, Fernández-Alvira JM, Oliva B, Fernández-Friera L, López-Melgar B, Mendiguren JM, Sanz J, Ordovás JM, Andrés V, Fernández-Ortiz A, Bueno H, Ibáñez B, García-Ruiz JM, Lara-Pezzi E. Machine learning improves cardiovascular risk definition for young, asymptomatic individuals. *J Am Coll Cardiol* 2020;**76**:1674–1685.
16. Hernesniemi JA, Mahdiani S, Tynkkynen JA, Lyytikäinen L-P, Mishra PP, Lehtimäki T, Eskola M, Nikus K, Antila K, Oksala N. Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome - the MADDEC study. *Ann Med* 2019;**51**:156–163.
17. Weichwald S, Candreva A, Burkholz R, Klingenberg R, Räber L, Heg D, Manka R, Gencer B, Mach F, Nanchen D, Rodondi N. Improving 1-year mortality prediction in ACS patients using machine learning. *Eur Heart J Acute Cardiovasc Care* 2021;**10**:855–865.
18. D'Ascenzo F, De Filippo O, Gallone G, Mittone G, Deriu MA, Iannaccone M, Ariza-Solé A, Liebetrau C, Manzano-Fernández S, Quadri G, Kinnaird T, Campo G, Simao Henriques JP, Hughes JM, Dominguez-Rodriguez A, Aldinucci M, Morbiducci U, Patti G, Raposeiras-Roubin S, Abu-Assi E, De Ferrari GM; PRAISE Study Group. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. *Lancet* 2021;**397**:199–207.
19. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019;**1**:206–215.