IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 15, 2021, accepted May 22, 2021, date of publication May 26, 2021, date of current version June 4, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3084063

Arterial Disease Computational Prediction and Health Record Feature Ranking Among Patients Diagnosed With Inflammatory Bowel Disease

DAVIDE CHICCO[1] **AND GIUSEPPE JURMAN**[2]
[1]Institute of Health Policy Management and Evaluation, University of Toronto, Toronto, Canada
[2]Data Science for Health Unit, Fondazione Bruno Kessler, Trento, Italy

Corresponding author: Davide Chicco (davidechicco@davidechicco.it)

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was originally granted by the Commission Nationale de l'Informatique et des Libertés (CNIL—1 104 603) in France to the original dataset curators, as explained in article [23].

ABSTRACT Inflammatory bowel diseases (IBDs) are a group of disorders causing chronic inflammation of small intestine and colon, and include Chron's disease and ulcerative colitis as most common occurrences. Patients suffering from IBD have more chances to experience an arterial event, such as a stroke or an acute coronary syndrome. In this setting, computational data mining methods applied to electronic medical records of patients diagnosed with IBD can provide useful information regarding the possibility for them to develop arterial diseases, in few minutes and at small cost. In this manuscript, we analyzed a dataset of 90 patients diagnosed with IBD, 30 of which having an arterial disease. After detecting the capability of predicting the arterial event and the most important features of the whole dataset, we repeated the analysis only on the subset of 30 patients suffering from both IBD and arterial disease. Our results show that machine learning can precisely predict both the occurrence of an arterial event and its type (stroke or acute coronary syndrome) from medical records, and can provide useful rankings about the most important clinical variables available in the dataset. Our otherwise unobservable findings can have a strong impact in the clinical settings, allowing physicians and medical doctors to make better decisions regarding prognoses and therapies of patients suffering from this disease.

INDEX TERMS Arterial disease, inflammatory bowel disease, binary classification, computational intelligence, supervised machine learning, machine learning, feature ranking, electronic health records.

I. INTRODUCTION

Dealing with comorbidities is a major task in healthcare nowadays, with multi-morbidity being more the norm than the exception [1]. Additionally, as the recent Covid-19 pandemics has brought to public attention worldwide [2], comorbidities have relevant structural and economic impact on healthcare systems [3], [4].

The importance of designing correct data science strategies in modeling comorbidities conditions has been known for the last 25 years [5], but a number of factors such as the lack of a shared consensus on definitions [1] has prevented reaching consistently good results in prediction. Promising trends are arising, however, and they rely on integration of clinical records with omics data and the recent advances in artificial intelligence algorithms [6]–[8].

Here we will focus on a well known comorbidity, linking the raised risk for cardiovascular (CV) events to a quite heterogeneous condition known as inflammatory bowel disease (IBD). The umbrella definition of IBD, in fact, encompasses a number of inflammatory pathologies of the colon and small intestine, Crohn's disease and ulcerative colitis being the principal phenotypes. The relation between a pre-existing IBD condition and the occurrence of a CV event has been extensively studied throughout the last decades, with a steady flow of published studies still ongoing nowadays. While there is a widespread consensus about the increased risk for cardiovascular events such as stroke and acute coronary syndrome in patients with IBD, with higher prevalence in women and in younger patients [9], it is less clear whether also the risk

The associate editor coordinating the review of this manuscript and approving it for publication was Syed Muhammad Anwar.

of worse outcomes from CV events such as sudden cardiac death and overall mortality is also increased by IBD [10]. Furthermore, the underlying mechanism inducing the raised risk is still far from being fully unveiled [11]–[14]. Together with genetic predisposition, the effects of chronic or acute inflammation has been indicated by several authors as the most likely culprits [15], [16]. In fact, inflammation induces a procoagulative state coupled with systemic endothelial dysfunction leading to the onset of thromboembolic adverse events [17]. Further, chronic inflammation-based risk factors due to IBD activity has been shown to be strongly associated to an increased risk of CV events even if conventional risk factors are not over-represented in patients with IBD [18]. Consequences on treatment are relevant, too: standard therapies targeting IBD with anti-inflammatory effect such as anti-TNF or thiopurines were shown also to reduce incidence and outcome of CV events, while other IBD therapies addressing different processes may be detrimental for the CV-related aspects, as in the case of tumor necrosis factor alpha antagonist [14]. Interestingly, a line of clinical research concerns the dual problem of tracking the IBD development among patients affected by CV diseases, through the definition of ad-hoc predictive scores such as the CHA 2DS 2-VASc [13].

In the current manuscript we are interested in the prediction of two CV events, namely stroke and acute coronary syndrome (ACS), in patients with IBD. Definition of both conditions varies along the years due to their heterogeneity and complexity [19]–[22]. Following the original study [23], we indicate as stroke a CV event inducing neurologic clinical modification and confirmed ischemia in magnetic resonance imaging [24], while acute coronary syndrome is identified by electrocardiogram modification, documented elevation in cardiac troponin and presence of a culprit coronary lesion [25].

Prediction of stroke occurrence and outcome has its roots in the final decades of the last century, when a number of risk scores started being introduced to estimate CV event hazard from demographic data [26], and newer scores are appearing recently, too [27]. The first models making their way to the clinical practice were heavily relying on statistical methods like the Cox proportional hazards, and, in particular, they adopted linearity as the basic mechanism relating risk to prevalence. However, this interaction was soon shown to be far from linear, and more complex statistical models appeared on stage [28] and they are still effective nowadays [29]–[31], keeping the pace with newer strategies.

Since 2010s, two game changers in this context have been the fast evolution of machine learning algorithms [32] and the growing availability of health electronic records (EHRs) [33]. As stated the recent review [34], the use of EHR features as predictor has boosted the performance of machine learning algorithms, whose improvement over statistical models has been quite limited when based solely on administrative claims, also allowing the detection through feature selection of critical biomarkers in the prediction, such as the demographic age, frailty score and the clinical serum urea nitrogen, serum creatinine and serum potassium [34], [35]. Among the most recent developments, the novel paradigm of deep learning has shown good diagnostic results for large cohorts [36], especially in integrated models [37] complementing EHR with high-throughput omics data [38] and imaging [39]. And, as in all fields where artificial intelligence is involved, model interpretability play a crucial role [28], [40].

Moving to ACS prediction, the publication pattern resembles closely the one for stroke, only on a smaller scale. Although statistical and score-based methods are still considered [41], [42], the machine learning methods are taking over, showing better performances [43]–[45]. Also for ACS prediction, EHRs play a crucial role [46], as well as the application of novel deep artificial neural networks [47], [48] and the quest for key biomarkers [49]. When dealing with the scenario of an underlying IBD condition, predicting CV events is still an uncharted territory, especially if relying only on EHR, let alone detecting significant biomarkers. The available references are in fact essentially those aforementioned and those cited in [23], while integration of bioimaging data [15], [50] and, even more, of metagenomics data [51], [52] are providing interesting and encouraging results.

Starting from the EHR dataset of 90 individuals originally introduced in [23], here we aim at simultaneously identify key predictive features and forecast CV events in a cohort of patients diagnosed with IBD. Given the limited sample size, we rely on classical well established machine learning models, namely Support Vector Machine, Random Forest, XGBoost, while we adopt Recursive Feature Elimination as the ranking algorithm. Model robustness and reproducibility are warranted by repeated stratified cross-validation cycles. As a major outcome, we name clinical activity, gender and the corticosteroids treatment dosage as most relevant features to identify stroke and ACS, providing novel information useful to the clinical practitioners for diagnostic purposes.

We organize the rest of the article as follows. After this Introduction, we describe the datasets analyzed (section II) and the methods employed (section III). We then report the complete results we obtained (section IV) and discuss their relevance and impact (section V). We finally outline the main messages, the limitations, and the future developments of this study (section VI).

II. DATASET

In this study, we analyze a dataset of electronic health records of 90 patients (37 women and 53 men) collected at Hôpital Saint-Antoine hospital (Paris, France) between 1996 and 2015 [23].

This dataset contains twelve clinical variables; eleven of them are binary and one (corticosteroids) is categorical (Table 1). This dataset contains a binary feature called ''arterial event'' that indicates of the patient had a stroke or an arterial coronary syndrome (value 1) or not (value 0). We use this feature as classification target for this dataset, and therefore we call this global dataset ''arterial event dataset''. With 66.67% negative data instances (patients with arterial event)

TABLE 1. Meaning, measurement unit, and possible values of each feature of the dataset. ACS: acute coronary syndrome. IBD: inflammatory bowel disease.

feature	explanation	measurement unit	values
arterial hypertension	if the patient had arterial hypertension or not	boolean	[0, 1]
clinical activity	quiescent (0) or IBD clinically active (1)	boolean	[0, 1]
corticosteroids (CTC)	dose level of CTC taken by the patient	category	[0, 1, 2]
diabetes	if the patient had diabetes or not	boolean	[0, 1]
dyslipidemia	if the patient had dyslipidemia or not	boolean	[0, 1]
immunosuppressants	if the patient has taken them or not	boolean	[0, 1]
mean CRP over 5 mg/L in 1 year	level of C-reactive protein in the last year	boolean	[0, 1]
mean CRP over 5 mg/L in 3 years	level of C-reactive protein in the last 3 years	boolean	[0, 1]
sex	if the patient was a woman (0) or a man (1)	binary	[0, 1]
smoker	if the patient was a smoker or not	boolean	[0, 1]
tumour necrosis factor inhibitors (TNFi)	if the patient has taken them or not	boolean	[0, 1]
[target] arterial event	if the patient had an acute arterial event or not	boolean	[0, 1]
[target] event type	if the patient had an ACS (0) or a stroke (1)	binary	[0, 1]

TABLE 2. Binary features quantitative characteristics for the arterial event dataset. All the binary features have meaning *true* for the value 1 and *false* for the value 0, except sex (0 = female and 1 = male) and arterial event (0 = patient who did not have an arterial event and 1 = patient who had an arterial event). The arterial event dataset contains 90 subjects, all with IBD.

feature	value	#	%
arterial hypertension	0	59	65.556
arterial hypertension	1	31	34.444
clinical activity	0	66	73.333
clinical activity	1	24	26.667
diabetes	0	81	90
diabetes	1	9	10
dyslipidemia	0	72	80
dyslipidemia	1	18	20
immunosuppressants	0	64	71.111
immunosuppressants	1	26	28.889
mean CRP over 5 mg/L in 1 year	0	51	56.667
mean CRP over 5 mg/L in 1 year	1	39	43.333
mean CRP over 5 mg/L in 3 years	0	50	55.556
mean CRP over 5 mg/L in 3 years	1	40	44.444
sex	0	37	41.111
sex	1	53	58.889
smoker	0	52	57.778
smoker	1	38	42.222
tumour necrosis factor inhibitors (TNFi)	0	84	93.333
tumour necrosis factor inhibitors (TNFi)	1	6	6.667
[target] arterial event	0	60	66.667
[target] arterial event	1	30	33.333
total		90	100

TABLE 3. Numeric feature quantitative characteristics. s.d.: standard deviation. corticosteroids (CTC): value 0 for 0 mg/j, value 1 for values between 0 and 20 mg/j, and value 2 for any value greater than 20 mg/j.

feature	median	mean	range	s.d.
corticosteroids (CTC)	0.000	0.300	[0, 2]	0.507

and 33.33% positive data instances (individuals without arterial event), we can say that the dataset is negatively unbalanced (Table 2 and Table 3).

Afterwards, we decided to consider only the subset of 60 patients who had an arterial event, and to consider the event type feature as target variable for our binary classification. The event type factor has value 0 if the patient had an acute coronary syndrome, or value 1 if the patient had a stroke (Table 4 and Table 5). We call this global dataset "event type dataset". This dataset contains data of 30 patients, and is negatively imbalanced, too: 73.33% negative elements and 26.67% positive elements.

TABLE 4. Binary features quantitative characteristics for the event type dataset. All the binary features have meaning *true* for the value 1 and *false* for the value 0, except sex (0 = female and 1 = male) and event type (0 = acute coronary syndrome and 1 = stroke). The event type dataset contains 30 individuals, all with IBD.

feature	value	#	%
arterial hypertension	0	14	46.667
arterial hypertension	1	16	53.333
clinical activity	0	14	46.667
clinical activity	1	16	53.333
diabetes	0	23	76.667
diabetes	1	7	23.333
dyslipidemia	0	20	66.667
dyslipidemia	1	10	33.333
immunosuppressants	0	23	76.667
immunosuppressants	1	7	23.333
mean CRP over 5 mg/L in 1 year	0	12	40
mean CRP over 5 mg/L in 1 year	1	18	60
mean CRP over 5 mg/L in 3 years	0	12	40
mean CRP over 5 mg/L in 3 years	1	18	60
sex	0	12	40
sex	1	18	60
smoker	0	16	53.333
smoker	1	14	46.667
tumour necrosis factor inhibitors (TNFi)	0	28	93.333
tumour necrosis factor inhibitors (TNFi)	1	2	6.667
[target] event type	0	22	73.333
[target] event type	1	8	26.667
total		30	100

TABLE 5. Numeric feature quantitative characteristics. s.d.: standard deviation. corticosteroids (CTC): value 0 for 0 mg/j, value 1 for values between 0 and 20 mg/j, and value 2 for any value greater han 20 mg/j.

feature	median	mean	range	s.d.
corticosteroids (CTC)	0.000	0.333	[0.000, 2.000]	0.547

The original article by Le Gall and colleagues [23] highlights the importance of the IBD clinical activity. Regarding

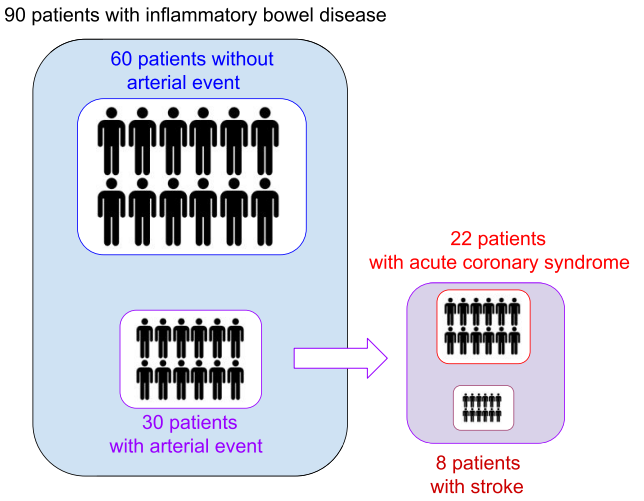

FIGURE 1. Dataset representation. Representation of the arterial event whole dataset and of the event type subset.

the meaning of this clinical variable, we contacted the corresponding author Le Gall *et al.* [23] who kindly provided this information privately via email:

0) Quiescence;
1) Minimal chronic or intermittent symptoms compatible with life (including work), doubtful imputability to IBD;
2) Clinical signs (chronic or intermittent) probably attributable to IBD but compatible with normal activity (including work);
3) Flair or active chronic disease incompatible with an activity (including work);
4) Hospitalization for more than one day for relapse or complication;
5) Hospitalization for abdominal surgery (resection, bypass).

The value 0 for clinical activity in this dataset refers then to the "Quiescence" condition listed above, while the value 1 refers to all the other conditions.

Several variables of the electronic health records of this dataset state if the patient has particular diseases or not: arterial hypertension, diabetes, and dyslipidemia (Table 1), where dyslipidemia means excessive presence of lipids in the blood. Other factors say if the patient has taken inhibitors (tumour necrosis factor inhibitors), steroids (corticosteroids), or immunosuppressive drugs (immunosuppressants): all drugs usually taken by patients diagnosed with inflammatory bowel disease [53]–[55]. Another variable indicates if the patient was a smoker, and another one indicates its biological sex.

Two variables provide information about the level of C-reactive found it the blood tests: one states if the mean CRP was over 5 mg/L in the last year, and another one refers to the same quantity in in the last three years. CRP, in fact, is used in the medical community to evaluate infection, inflammation or tissue injury. A value below 5 mg/L means lower risk of cardiovascular disease, while a value above that threshold means higher risk of cardiovascular disease.

More information about the dataset can be found in the original article [23].

III. METHODS

In this section, we first described the machine learning techniques we used for binary classification (subsection III-A), and the machine learning and biostatistics techniques we employed to compute the medical feature rankings (subsection III-B).

A. BINARY CLASSIFICATION

For both the whole arterial event dataset and the event type subset of patients, we trained a suite of binary classifiers and assessed their performances on a set of replicated experiments. In detail, the set of experiments consists of 1,000 runs of Monte Carlo stratified Cross Validation [56], with 80%/20% train/test splits: in each run, the training portion includes 80% of the data, randomly sampled preserving the original proportion between classes; then, classifier performances are evaluated on the remaining set of samples constituting the test portion. As classifiers, we employed the following algorithms:

- Support Vector Machines (SVM) [57], linear and gaussian kernel, with $C = 10^3$ for both;
- XGBoost [58], gbtree and gblinear booster, $\eta = 0.001$ and 1, early stopping rounds 10, subsample 0.75 and 0.85, $\gamma = 3$, max depth 5, binary: logistic objective, error evaluation metric and nrounds 10,000;
- Random Forests [59];
- Logit, binomial object model family.

Common best practices in machine learning require the additional split of the dataset into training set, validation set, and test set, where the validation set is used to test several hyper-parameter configurations through a grid search [60]. In this case, however, we decided to avoid the hyper-parameter optimization because the of the small size of the dataset, and decided to stick with the default parameters of the corresponding `R` function, unless otherwise specified.

We reported more details about the `R` functions and packages used in the Supplementary information.

B. FEATURE RANKING

To compute the ranking of the most predictive clinical features of the two datasets, we employed a typical recursive feature elimination (RFE) [61]. Applying this approach, we performed a loop cycle where, at each iteration, we removed a different feature and applied the XGBoost classifier as described before (subsection III-A). Then, for each iteration, we saved the prediction scores expressed as confusion matrix indicators. We finally ranked each feature based on the MCC obtained when that specific feature was removed [62]: the lower the MCC, the more important the feature is.

TABLE 6. Arterial event prediction results. s.d.: standard deviation. Positives: patients who had an arterial event. Negatives: patients who did not experience an arterial event. SVM: Support Vector Machine. MCC: Matthews correlation coefficient. MCC worst value = −1 and best value = +1. Sensitivity: true positive rate, recall. Specificity: true negative rate. PR: precision-recall curve. Precision: positive predictive value (PPV). NPV: negative predictive value. ROC: receiver operating characteristic curve. AUC: area under the curve. Confusion matrix threshold cut-off: 0.5. F_1 score, accuracy, TP rate, TN rate, PPV, NPV, PR AUC, ROC AUC: worst value = 0 and best value = +1. We highlighted in blue and with an asterisk * the top result for each statistical indicator. We report the formulas of these rates in the supplementary information.

	target: arterial event								
method	MCC	F_1 score	accuracy	sensitivity	specificity	precision	NPV	PR AUC	ROC AUC
XGBoost mean	**+0.406***	0.816*	0.744*	0.864*	0.503*	0.782*	0.682*	0.558*	0.683*
XGBoost s.d.	0.217	0.070	0.088	0.109	0.199	0.071	0.208	0.137	0.104
Logit mean	**+0.342**	0.798	0.718	0.847	0.459	0.762	0.629	0.521	0.653
Logit s.d.	0.232	0.074	0.095	0.112	0.192	0.071	0.223	0.140	0.108
Gaussian SVM mean	**+0.264**	0.760	0.677	0.781	0.468	0.749	0.535	0.470	0.624
Gaussian SVM s.d.	0.233	0.085	0.101	0.127	0.197	0.078	0.197	0.127	0.111
Random Forests mean	**+0.255**	0.788	0.692	0.864*	0.349	0.729	0.591	0.475	0.607
Random Forests s.d.	0.233	0.065	0.086	0.105	0.187	0.061	0.261	0.131	0.100

For both the datasets, we applied only the XGBoost classifier because it was the one which obtained the highest prediction scores.

Feature ranking through machine learning methods can capture the importance of the dataset variables while they interact between each other. Although it generates interesting results, it does not express the relevance of each feature alone in relationship with the target.

For a better understanding of the importance of each clinical variable, we therefore applied traditional univariate statistical tests such as *chi-squared test* [63] and *Kruskal-Wallis test* [64].

These two tests check how likely an observed distribution is due to chance between two binary variables (chi-squared test) or between two multi-class variables (Kruskal-Wallis test). If the test generate a low *p*-value (close to 0), it means that the two features have a strong relation, while if it generates a high p-value (close to 1), it means that the null hypothesis of independence cannot be discarded. These univariate statistical tests capture the importance of the relationship between each feature alone and the target feature, without considering the interactions between the features.

IV. RESULTS

In this section, we first report the results we obtained for the binary classification of the arterial event dataset and of the event type subset (subsection IV-A), and the feature ranking results achieved through biostatistics and machine learning techniques for the arterial event dataset (subsection IV-B) and for the event type dataset (subsection IV-C).

A. BINARY CLASSIFICATION RESULTS

We employed the binary classification methods previously described (section III) to the general arterial event dataset and the event type subset, and report the results in Table 6 and Table 7, respectively. We also represented the MCC results as barplots in Figure 2. For each dataset, we employed both the SVM with linear kernel and the SVM with Gaussian kernel, and reported only the kernel which generated the best results.

Among the confusion matrix statistical indicators employed, we gave more importance to the Matthews correlation coefficient (MCC) because it is the only metric that generates a high score only if the classifier achieved high results on all sensitivity, specificity, precision, and negative predictive value at the same time [60], [65]–[68].

Regarding arterial event, all the four machine learning methods employed were able to obtained good prediction results, with MCC ranging from +0.255 (Random Forests) to +0.406 (XGBoost). The gradient boosting method XGBoost earned the top score for all the statistical indicators, tied for first with Random Forests regarding the recall. XGBoost was also the only method able to correctly predict most of the negatives (sensitivity = 0.503), while the other three techniques correctly predict only less than 50% of them (Table 6).

Regarding event type, the classifiers obtained lower prediction scores, with MCC ranking from +0.050 (logit) to +0.222 (XGBoost).

As one can notice, the gradient boosting XGBoost method achieved the top results for this task as well, with top MCC, specificity, precision, NPV, precision-recall AUC, and receiver operating characteristic AUC Table 7). No method here was capable of correctly predicting most of the negative data instances (all specificity scores are lower than 0.5): this result is probably due to the small size of the negative class of this subset, with only 8 elements. Also, the other methods did not generate overall reliable predictive results: the MCC attained by logit (+0.050) is close to random guessing, while the ones gained by Random Forests and Linear SVM are just above that value (+0.111 and +0.139, respectively).

Because of the imbalance of both the datasets, all the four methods obtained higher scores when correctly classifying positive data instances (sensitivity) and making correct positive predictions (precision), than negative data instances (specificity and negative predictive value). During training, in fact, all the methods were able to observe and learn more positive data instances than negative elements, and therefore they performed better on the former class than on the latter class.

Overall, we can state that the machine learning method XGBoost was able to correctly predict patients with arterial

TABLE 7. Event type prediction results. s.d.: standard deviation. Positives: patients who had a stroke. Negatives: patients who had an acute coronary syndrome. SVM: Support Vector Machine. MCC: Matthews correlation coefficient. MCC worst value = −1 and best value = +1. Sensitivity: true positive rate, recall. Specificity: true negative rate. PR: precision-recall curve. Precision: positive predictive value (PPV). NPV: negative predictive value. ROC: receiver operating characteristic curve. AUC: area under the curve. Confusion matrix threshold cut-off: 0.5. F_1 score, accuracy, TP rate, TN rate, PPV, NPV, PR AUC, ROC AUC: worst value = 0 and best value = +1. We highlighted in blue and with an asterisk * the top result for each statistical indicator. We report the formulas of these rates in the supplementary information.

	target: event type								
method	MCC	F_1 score	accuracy	sensitivity	specificity	precision	NPV	PR AUC	ROC AUC
XGBoost mean	**+0.222***	0.793	0.708	0.765	0.482*	0.864*	0.369*	0.385*	0.623*
XGBoost s.d.	0.496	0.164	0.198	0.217	0.500	0.147	0.376	0.303	0.270
Random Forests mean	**+0.139**	0.827*	0.734*	0.839*	0.315	0.839	0.336	0.338	0.577
Random Forests s.d.	0.455	0.123	0.166	0.179	0.465	0.116	0.398	0.289	0.244
Linear SVM mean	**+0.111**	0.744	0.648	0.703	0.428	0.842	0.263	0.311	0.565
Linear SVM s.d.	0.459	0.172	0.195	0.224	0.495	0.150	0.310	0.239	0.260
Logit mean	**+0.050**	0.713	0.613	0.669	0.391	0.825	0.220	0.282	0.530
Logit s.d.	0.448	0.188	0.200	0.237	0.488	0.158	0.290	0.222	0.255

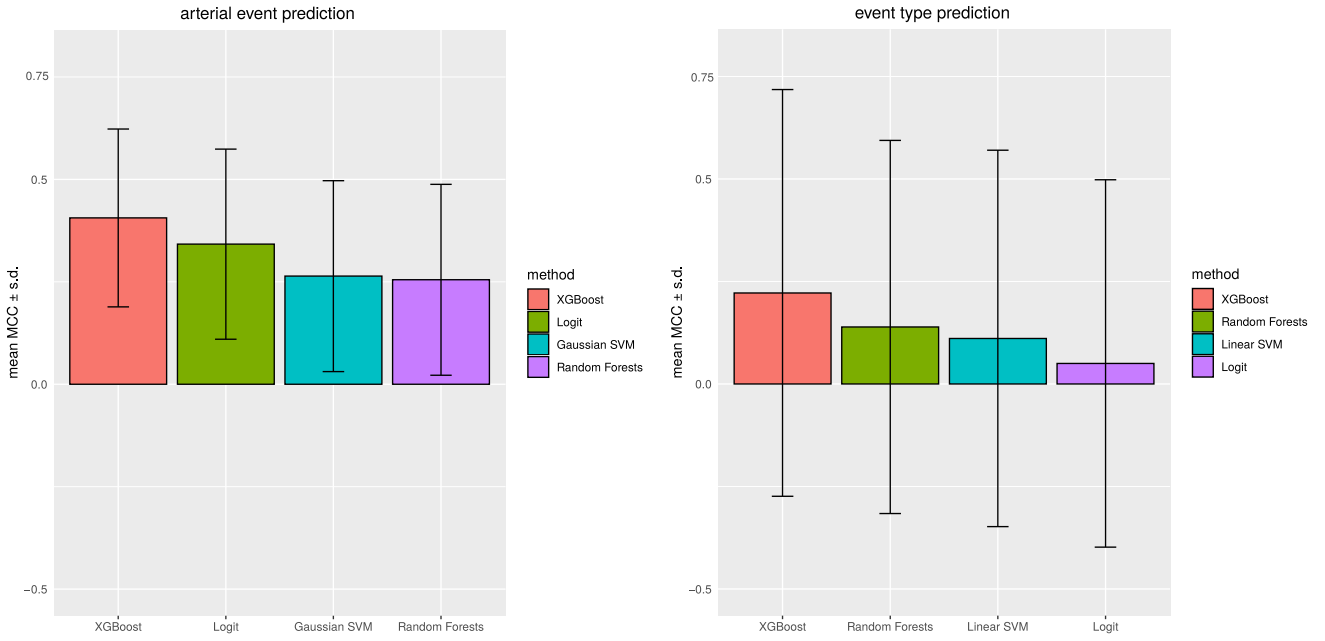

FIGURE 2. Classification results representation. Representation of the classification results reported as mean Matthews correlation coefficients ± the corresponding standard deviations for each method. We reported the precise values in Table 6 and in Table 7.

event, subjects who had no arterial event, and patients with stroke in the event type dataset.

The barplots show that each method employed to predict the arterial event type had a large standard deviation (Figure 2). We believe this result is due to the small size of the dataset, which made the results being unstable.

B. FEATURE RANKING RESULTS FOR ARTERIAL EVENT

After assessing the capability of computational intelligence to predict if a patient had an arterial event or not, we decided to investigate which clinical features are the most relevant in this prediction. As described earlier (subsection III-B), we employed both a machine learning recursive feature elimination technique and a traditional biostatistics approach.

Both the machine learning (Table S1) and the biostatistics approach (Table S2) identified clinical activity as the most relevant clinical feature to predict if a patient will have an arterial event. Clinical activity resulted being the medical feature whose removal caused the biggest MCC drop in the XGBoost ranking, and the feature obtaining the lowest *p*-value among the statistical tests. This result confirms the finding of the original curators of this dataset regarding clinical activity [23].

Regarding the other features, the two approaches generated significantly different results for biological sex and mean CRP over 5 mg/L in 1 year. The biological sex of a patient results being the second most important feature in the machine learning ranking (Table S1), but the least important variable in the biostatistics ranking (Table S2). Mean CRP over 5 mg/L in 1 year, instead, results being unimportant according to the computational intelligence ranking (Table S1), but forth most important feature in the biostatistics ranking (Table S2).

C. FEATURE RANKING RESULTS FOR EVENT TYPE

Regarding event type, the computational intelligence approach identified biological sex, corticosteroids, diabetes and tumour necrosis factor inhibitor as the top most important medical features to discriminate patients with stroke from patients with acute coronary syndrome (Table S3). The univariate biostatistics tests, instead, detected biological sex, mean CRP over 5 mg/L in 1 year, and clinical activity as the most predictive clinical factors.

Even if both the rankings agreed to recognize the biological sex of the patient as the most important medical factor, they showed different behaviors for other features. The biostatistics ranking, for example, shows mean CRP over 5 mg/L in 1 year as second most important variable, while the machine learning ranking inserted it at the penultimate position. Clinical activity, also, results being in the third position of the biostatistics ranking, but only on eighth rank in the XGBoost standing.

V. DISCUSSION

A. ARTERIAL EVENT PREDICTION AND EVENT TYPE CLASSIFICATION

Our results show that machine learning can effectively predict patients with arterial event and patients with strokes in few minutes, with small computational resources, from electronic clinical records of subjects with IBD. The gradient boosting XGBoost method, in particular, was able to correctly identify most of patients with arterial event and without it, and to make correct positive predictions and negative predictions, achieving an average MCC of +0.406 (subsection IV-A). Also for the event type, the best results were achieved again by XGBoost, which obtained an MCC of +0.222 (subsection IV-A). Both the arterial event prediction and the event type classification tasks were not included in the original study of Le Gall and colleagues [23].

B. ARTERIAL EVENT FEATURE RANKING

Our feature ranking analysis confirmed clinical activity as the most predictive factor for arterial events among patients with IBD, as indicated by Le Gall *et al.* [23]. This result was highlighted both by the machine learning standing and by the univariate biostatistics tests standing (subsection IV-B). The two rankings, however, gave discordant messages regarding biological sex: our XGBoost ranking positioned it second by importance, while the biostatistics ranking listed it in the last rank. According to Lawton [69], the biological sex of an individual is a relevant piece of information regarding her/his risk of arterial event, supporting the finding of our machine learning feature ranking approach. On other side of the ranking, our gradient boosting technique listed tumour necrosis factor inhibitors as less important factor in predicting arterial events in patients.

C. EVENT TYPE FEATURE RANKING

The XGBoost feature ranking approach applied to the event type dataset identified biological sex, corticosteroids, and diabetes as the top three most relevant clinical features to predict if a patient will have a stroke or an acute coronary syndrome (subsection IV-C). The importance of biological sex and corticosteroids was detected by the biostatistics ranking, too, while diabetes was not in the top positions there. In their study, Sealy and colleagues [70] highlighted the role of biological sex in stroke risk. The intake of corticosteroids can be due to an ongoing treatment had by the patients at risk of stroke [71]. Our machine learning approach, however, successfully identified diabetes as a top third most relevant factor, since this disease is known to increase the stroke risk [72]. In the last positions of our machine learning ranking, having hypertension resulted being the as least important factor in discriminating between stroke and acute coronary syndrome.

VI. CONCLUSION

Inflammatory bowel diseases affect millions of people worldwide, and can worse the conditions of patients when present with arterial diseases. In this study, we showed that computational intelligence applied to medical records can be an effective tool both to diagnoses prediction (presence of arterial event versus absence of arterial event, and stroke versus acute coronary syndrome) and to rank the clinical features resulting being more predictive for them. Our machine learning methods were able to unveil new findings that differ from the ones obtained by traditional univariate statistical tests and the ones found in the medical literature, suggesting computational intelligence as an additional support means for clinical decisions.

The main limitation of our study is the employment of a single dataset: our analysis would have been enhanced by the usage of an external validation cohort, where we could have found confirmation of our findings. We looked for an alternative public dataset of patients with IBD and cardiovascular diseases but unfortunately we could not find any with the same features.

Other studies applied machine learning techniques to detect the most important features of clinical records of patients with cardiovascular events in the past [73], [74]. Unfortunately the medical records of the datasets analyzed in these articles did not include the features considered in our study.

In the future, we plan to investigate further the inflammatory bowel disease by applying our computational intelligence and statistical methods to genomics and transcriptomics datasets of patients diagnosed with this condition [75]–[77].

LIST OF ABBREVIATIONS

AUC: area under the curve. CRP: C-reactive protein. CTC: corticosteroids. IBD: inflammatory bowel disease. PR: precision-recall. ROC: receiver operating characteristic. TN rate: true negative rate. TNFi: tumour necrosis factor inhibitors. TP rate: true positive rate. s.d.: standard deviation.

COMPETING INTERESTS

The authors declare they have no competing interest.

FUNDING

Not applicable.

ACKNOWLEDGMENT

The authors thank Guillaume Le Gall (Hôpital Saint Antoine) for his help with the dataset feature understanding.

DATA AND SOFTWARE AVAILABILITY

The dataset file used in this project is publicly available on FigShare under the Creative Commons Attribution 4.0 International (CC BY 4.0) license at: https://figshare.com/articles/Clinical_activity_is_an_independent_risk_factor_of_ischemic_heart_and_cerebrovascular_arterial_disease_in_patients_with_inflammatory_bowel_disease/7036235

Our software code is publicly available under GNU General Public License v3.0 at: https://github.com/davidechicco/arterial_events_and_IBD

AUTHORS' DETAILS

Davide Chicco (ORCID: 0000-0001-9655-7142) is with University of Toronto, Toronto, Ontario, Canada.
Giuseppe Jurman (ORCID: 0000-0002-2705-5728) is with Fondazione Bruno Kessler, Trento, Italy.

Correspondence should be addressed to Davide Chicco: davidechicco@davidechicco.it

REFERENCES

[1] J. M. Valderas, B. Starfield, B. Sibbald, C. Salisbury, and M. Roland, "Defining comorbidity: Implications for understanding health and health services," *Ann. Family Med.*, vol. 7, no. 4, pp. 357–363, Jul. 2009.

[2] A. Sanyaolu, C. Okorie, A. Marinkovic, R. Patidar, K. Younis, P. Desai, Z. Hosein, I. Padda, J. Mangat, and M. Altaf, "Comorbidity and its impact on patients with COVID-19," *Social Netw. Comprehensive Clin. Med.*, vol. 2, no. 8, pp. 1069–1076, Aug. 2020.

[3] K. Barnett, S. W. Mercer, M. Norbury, G. Watt, S. Wyke, and B. Guthrie, "Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study," *Lancet*, vol. 380, no. 9836, pp. 37–43, Jul. 2012.

[4] S. Cortaredona and B. Ventelou, "The extra cost of comorbidity: Multiple illnesses and the economic burden of non-communicable diseases," *BMC Med.*, vol. 15, no. 1, Dec. 2017, Art. no. 216.

[5] L. L. Roos, R. C. James, and J. Li, "Complications, comorbidities, and mortality: Improving classification and prediction," *Health Services Res.*, vol. 32, no. 2, pp. 229–238, 1997.

[6] E. Capobianco and P. Liò, "Comorbidity: A multidimensional approach," *Trends Mol. Med.*, vol. 19, no. 9, pp. 515–521, Sep. 2013.

[7] P. Akram and L. Liao, "Prediction of comorbid diseases using weighted geometric embedding of human interactome," *BMC Med. Genomics*, vol. 12, no. S7, pp. 1–15, Dec. 2019.

[8] Z. Qian, A. Alaa, A. Bellot, M. van der Schaar, and J. Rashbass, "Learning dynamic and personalized comorbidity networks from event data using deep diffusion processes," in *Proc. 23rd Int. Conf. Artif. Intell. Statist. (AISTATS)*, vol. 108, S. Chiappa and R. Calandra, Eds., New York, NY, USA, 2020, pp. 3295–3305.

[9] A. Bigeh, A. Sanchez, C. Maestas, and M. Gulati, "Inflammatory bowel disease and the risk for cardiovascular disease: Does all inflammation lead to heart disease?" *Trends Cardiovascular Med.*, vol. 30, no. 8, pp. 463–469, Nov. 2020.

[10] T. Abe, T. Ajose, N. N. Atem, and V. S. Effoe, "Cardiovascular outcomes in patients with coexisting coronary artery disease and inflammatory bowel disease," *Circulation*, vol. 140, no. 1, 2019, Art. no. A11770.

[11] R. Schicho, G. Marsche, and M. Storr, "Cardiovascular complications in inflammatory bowel disease," *Current Drug Targets*, vol. 16, no. 3, pp. 181–188, Mar. 2015.

[12] G. Tsigkas, P. Davlouros, S. Despotopoulos, S. F. Assimakopoulos, G. Theocharis, and G. Hahalis, "Inflammatory bowel disease: A potential risk factor for coronary artery disease," *Angiology*, vol. 68, no. 10, pp. 845–849, Nov. 2017.

[13] W.-S. Hu and C.-L. Lin, "CHA 2 DS 2 -VASc score in the prediction of ischemic bowel disease among patients with atrial fibrillation: Insights from a nationwide cohort," *Int. J. Cardiol.*, vol. 235, pp. 56–60, May 2017.

[14] S. D. Thapa, H. Hadid, W. Imam, J. Schairer, and S.-M. Jafri, "Effect of inflammatory bowel disease–related characteristics and treatment interventions on cardiovascular disease incidence," *Amer. J. Med. Sci.*, vol. 350, no. 3, pp. 175–180, Sep. 2015.

[15] D.-M. Bunu, C.-E. Timofte, M. Ciocoiu, M. Floria, C.-C. Tarniceriu, O.-B. Barboi, and D.-M. Tanase, "Cardiovascular manifestations of inflammatory bowel disease: Pathogenesis, diagnosis, and preventive strategies," *Gastroenterol. Res. Pract.*, vol. 2019, pp. 1–14, Jan. 2019.

[16] R. B. Biondi, P. S. Salmazo, S. G. ZanatiBazan, J. C. Hueb, S. A. R. D. Paiva, and L. YukieSassaki, "Cardiovascular risk in individuals with inflammatory bowel disease," *Clin. Exp. Gastroenterol.*, vol. 13, pp. 107–113, Apr. 2020.

[17] M. Pepe, E. Carulli, C. Forleo, M. Moscarelli, O. D. Cillo, A. S. Bortone, P. L. Nestola, G. Biondi-Zoccai, A. Giordano, and S. Favale, "Inflammatory bowel disease and acute coronary syndromes: From pathogenesis to the fine line between bleeding and ischemic risk," *Inflammatory Bowel Diseases*, vol. 27, no. 5, pp. 725–731, Apr. 2021.

[18] S. Singh, I. J. Kullo, D. S. Pardi, and E. V. Loftus, "Epidemiology, risk factors and management of cardiovascular diseases in IBD," *Nature Rev. Gastroenterol. Hepatol.*, vol. 12, no. 1, pp. 26–35, Jan. 2015.

[19] A. P. Coupland, A. Thapar, M. I. Qureshi, H. Jenkins, and A. H. Davies, "The definition of stroke," *J. Roy. Soc. Med.*, vol. 110, no. 1, pp. 9–12, 2017.

[20] B. C. V. Campbell and P. Khatri, "Stroke," *Lancet*, vol. 396, no. 10244, pp. 129–142, 2020.

[21] P. Klootwijk and C. Hamm, "Acute coronary syndromes: Diagnosis," *Lancet*, vol. 353, no. 2, pp. 10–15, 1999.

[22] I. N. Findlay and A. D. Cunnigham, "Definition of acute coronary syndrome," *Heart*, vol. 91, no. 7, pp. 857–859, Jul. 2005.

[23] G. Le Gall, J. Kirchgesner, M. Bejaoui, C. Landman, I. Nion-Larmurier, A. Bourrier, H. Sokol, P. Seksik, and L. Beaugerie, "Clinical activity is an independent risk factor of ischemic heart and cerebrovascular arterial disease in patients with inflammatory bowel disease," *PLoS ONE*, vol. 13, no. 8, Aug. 2018, Art. no. e0201991.

[24] R. L. Sacco, S. E. Kasner, J. P. Broderick, L. R. Caplan, J. J. B. Connors, A. Culebras, M. S. V. Elkind, M. G. George, A. D. Hamdan, R. T. Higashida, B. L. Hoh, L. S. Janis, C. S. Kase, D. O. Kleindorfer, J.-M. Lee, M. E. Moseley, E. D. Peterson, T. N. Turan, A. L. Valderrama, and H. V. Vinters, "An updated definition of stroke for the 21st century," *Stroke*, vol. 44, no. 7, pp. 2064–2089, Jul. 2013.

[25] K. Thygesen *et al.*, "Third universal definition of myocardial infarction," *J. Amer. College Cardiol.*, vol. 60, no. 16, pp. 1581–1598, 2012.

[26] P. A. Wolf, R. B. D'Agostino, A. J. Belanger, and W. B. Kannel, "Probability of stroke: A risk profile from the framingham study," *Stroke*, vol. 22, no. 3, pp. 312–318, Mar. 1991.

[27] K. Matsumoto, Y. Nohara, H. Soejima, T. Yonehara, N. Nakashima, and M. Kamouchi, "Stroke prognostic scores and data-driven prediction of clinical outcomes after acute ischemic stroke," *Stroke*, vol. 51, no. 5, pp. 1477–1483, May 2020.

[28] R. Zhang, "Interpretable machine learning methods for stroke prediction," M.S. thesis, Massachusetts Inst. Technol., Cambridge, MA, USA, 2019.

[29] D. G. Manuel, M. Tuna, R. Perez, P. Tanuseputro, D. Hennessy, C. Bennett, L. Rosella, C. Sanmartin, C. van Walraven, and J. V. Tu, "Predicting stroke risk based on health behaviours: Development of the stroke population risk tool (SPoRT)," *PLoS ONE*, vol. 10, no. 12, Dec. 2015, Art. no. e0143342.

[30] T. J. Quinn and B. A. Drozdowska, "Stroke prediction and the future of prognosis research," *Nature Rev. Neurol.*, vol. 15, no. 6, pp. 311–312, Jun. 2019.

[31] G. M. De Marchis, T. Dankowski, I. R. König, J. Fladt, F. Fluri, H. Gensicke, C. Foerch, O. Findling, R. Kurmann, U. Fischer, A. Luft, D. Buhl, S. T. Engelter, P. A. Lyrer, M. Christ-Crain, M. Arnold, and M. Katan, "A novel biomarker-based prognostic score in acute ischemic stroke," *Neurology*, vol. 92, no. 13, pp. e1517–e1525, Mar. 2019.

[32] W. Wang, M. Kiik, N. Peek, V. Curcin, I. J. Marshall, A. G. Rudd, Y. Wang, A. Douiri, C. D. Wolfe, and B. Bray, "A systematic review of machine learning models for predicting outcomes of stroke with structured data," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0234722.

[33] B. A. Goldstein, A. M. Navar, M. J. Pencina, and J. P. A. Ioannidis, "Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review," *J. Amer. Med. Inform. Assoc.*, vol. 24, no. 1, pp. 198–208, Jan. 2017.
[34] R. J. Desai, S. V. Wang, M. Vaduganathan, T. Evers, and S. Schneeweiss, "Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes," *J. Amer. Med. Assoc. Netw. Open*, vol. 3, no. 1, Jan. 2020, Art. no. e1918962.
[35] E. Kogan, K. Twyman, J. Heap, D. Milentijevic, J. H. Lin, and M. Alberts, "Assessing stroke severity using electronic health record data: A machine learning approach," *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, Dec. 2020, Art. no. 8.
[36] S. Cheon, J. Kim, and J. Lim, "The use of deep learning to predict stroke patient mortality," *Int. J. Environ. Res. Public Health*, vol. 16, no. 11, p. 1876, May 2019.
[37] Y. Xie, B. Jiang, E. Gong, Y. Li, G. Zhu, P. Michel, M. Wintermark, and G. Zaharchuk, "Use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information," *Amer. J. Roentgenol.*, vol. 212, no. 1, pp. 44–51, Jan. 2019.
[38] G. Abraham, R. Malik, E. Yonova-Doing, A. Salim, T. Wang, J. Danesh, A. S. Butterworth, J. M. M. Howson, M. Inouye, and M. Dichgans, "Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke," *Nature Commun.*, vol. 10, no. 1, pp. 1–10, Dec. 2019.
[39] A. Jamthikar, D. Gupta, N. N. Khanna, L. Saba, T. Araki, K. Viskovic, H. S. Suri, A. Gupta, S. Mavrogeni, M. Turk, J. R. Laird, G. Pareek, M. Miner, P. P. Sfikakis, A. Protogerou, G. D. Kitas, V. Viswanathan, A. Nicolaides, D. L. Bhatt, and J. S. Suri, "A low-cost machine learning-based cardiovascular/stroke risk assessment system: Integration of conventional factors with image phenotypes," *Cardiovascular Diagnosis Therapy*, vol. 9, no. 5, pp. 420–430, Oct. 2019.
[40] E. Zihni, V. I. Madai, M. Livne, I. Galinovic, A. A. Khalil, J. B. Fiebach, and D. Frey, "Opening the black box of artificial intelligence for clinical decision support: A study predicting stroke outcome," *PLoS ONE*, vol. 15, no. 4, Apr. 2020, Art. no. e0231166.
[41] F. Sanchis-Gomar, C. Perez-Quilis, R. Leischik, and A. Lucia, "Epidemiology of coronary heart disease and acute coronary syndrome," *Ann. Transl. Med.*, vol. 4, no. 13, p. 256, Jul. 2016.
[42] J. Stepinska, M. Lettino, I. Ahrens, H. Bueno, L. Garcia-Castrillo, A. Khoury, P. Lancellotti, C. Mueller, T. Muenzel, A. Oleksiak, R. Petrino, M. R. Guimenez, D. Zahger, C. J. M. Vrints, S. Halvorsen, E. de Maria, G. Y. H. Lip, R. Rossini, M. Claeys, and K. Huber, "Diagnosis and risk stratification of chest pain patients in the emergency department: Focus on acute coronary syndromes," *Eur. Heart J., Acute Cardiovascular Care*, vol. 9, no. 1, pp. 76–89, 2020.
[43] J. A. Hernesniemi, S. Mahdiani, J. A. Tynkkynen, L.-P. Lyytikäinen, P. P. Mishra, T. Lehtimäki, M. Eskola, K. Nikus, K. Antila, and N. Oksala, "Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome—The MADDEC study," *Ann. Med.*, vol. 51, no. 2, pp. 156–163, Feb. 2019.
[44] A. J. Steele, S. C. Denaxas, A. D. Shah, H. Hemingway, and N. M. Luscombe, "Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease," *PLoS ONE*, vol. 13, no. 8, Aug. 2018, Art. no. e0202344.
[45] J. P. Costabel, L. Poler, C. Garmendia, F. Lambardi, P. Ariznavarreta, R. Echegoyen, and M. Trivi, "Risk prediction of suffering an acute coronary syndrome through the use of machine learning classifiers," *J. Amer. College Cardiol.*, vol. 75, no. 11, p. 247, Mar. 2020.
[46] D. Hu, W. Dong, X. Lu, H. Duan, K. He, and Z. Huang, "Evidential MACE prediction of acute coronary syndrome using electronic health records," *BMC Med. Informat. Decis. Making*, vol. 19, no. S2, pp. 9–17, Apr. 2019.
[47] H. Duan, Z. Sun, W. Dong, and Z. Huang, "Utilizing dynamic treatment information for MACE prediction of acute coronary syndrome," *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, Dec. 2019, Art. no. 5.
[48] Z. Huang, W. Dong, H. Duan, and J. Liu, "A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health records," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 956–968, May 2018.
[49] E. Mahran, N. E. Hassan, T. K. El-Ddin, and T. H. N. Kafafy, "Platelet indices and blood cell ratios in acute coronary syndrome and their predictive values," *J. Current Med. Res. Pract.*, vol. 5, no. 1, p. 57, 2020.
[50] S. Patil, K. Gonuguntla, M. Kumar, C. Rojulpote, A. Aujla, A. Achunair, and K. Chen, "Myocarditis complicating inflammatory bowel disease: A comparison between ulcerative colitis and Crohn's disease from the national inpatient sample," *J. Amer. College Cardiol.*, vol. 75, no. 11, p. 1077, Mar. 2020.
[51] T. Toya, M. T. Corban, E. Marrietta, I. E. Horwath, L. O. Lerman, J. A. Murray, and A. Lerman, "Coronary artery disease is associated with an altered gut microbiome composition," *PLoS ONE*, vol. 15, no. 1, Jan. 2020, Art. no. e0227147.
[52] N. Kazemian, M. Mahmoudi, F. Halperin, J. C. Wu, and S. Pakpour, "Gut microbiota and cardiovascular disease: Opportunities and challenges," *Microbiome*, vol. 8, no. 1, p. 36, Dec. 2020.
[53] W. A. Faubion, Jr., E. V. Loftus, W. S. Harmsen, A. R. Zinsmeister, and W. J. Sandborn, "The natural history of corticosteroid therapy for inflammatory bowel disease: A population-based study," *Gastroenterology*, vol. 121, no. 2, pp. 255–260, Aug. 2001.
[54] O. H. Nielsen and M. A. Ainsworth, "Tumor necrosis factor inhibitors for inflammatory bowel disease," *New England J. Med.*, vol. 369, no. 8, pp. 754–762, Aug. 2013.
[55] F. Goldstein, "Immunosuppressant therapy of inflammatory bowel disease: Pharmacologic and clinical aspects," *J. Clin. Gastroenterol.*, vol. 9, no. 6, pp. 654–658, Dec. 1987.
[56] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," pp. 1–49, 2018, *arXiv:1811.12808*. [Online]. Available: http://arxiv.org/abs/1811.12808
[57] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[58] T. Chen and C. Guestrin, "XGBoost," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794.
[59] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[60] D. Chicco, "Ten quick tips for machine learning in computational biology," *BioData Mining*, vol. 10, no. 1, pp. 1–17, Dec. 2017.
[61] X.-W. Chen and C. Jeong, "Enhanced recursive feature elimination," in *Proc. 6th Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2007, pp. 429–435.
[62] L. Oneto, *Model Selection and Error Estimation in a Nutshell*. Heidelberg, Germany: Springer, 2020.
[63] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*, vol. 280. Hoboken, NJ, USA: Wiley, 1996.
[64] N. Breslow, "A generalized Kruskal–Wallis test for comparing k samples subject to unequal patterns of censorship," *Biometrika*, vol. 57, no. 3, pp. 579–594, 1970.
[65] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochim. et Biophys. Acta, Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975.
[66] G. Jurman, S. Riccadonna, and C. Furlanello, "A comparison of MCC and CEN error measures in multi-class prediction," *PLoS ONE*, vol. 7, no. 8, Aug. 2012, Art. no. e41882.
[67] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, Dec. 2020.
[68] D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Mining*, vol. 14, no. 1, pp. 1–22, Dec. 2021.
[69] J. S. Lawton, "Sex and gender differences in coronary artery disease," *Seminars Thoracic Cardiovascular Surg.*, vol. 23, no. 2, pp. 126–130, Jun. 2011.
[70] S. Sealy-Jefferson, J. J. Wing, B. N. Sánchez, D. L. Brown, W. J. Meurer, M. A. Smith, L. B. Morgenstern, and L. D. Lisabeth, "Age- and ethnic-specific sex differences in stroke risk," *Gender Med.*, vol. 9, no. 2, pp. 121–128, 2012.
[71] N. Qizilbash, S. Lewington, and J. López-Arrieta, "Corticosteroids for acute ischaemic stroke," *Cochrane Database Syst. Rev.*, vol. 2002, no. 2, 2002, Art. no. CD000064.
[72] N. N. Tun, G. Arunagirinathan, S. K. Munshi, and J. M. Pappachan, "Diabetes mellitus and stroke: A clinical update," *World J. Diabetes*, vol. 8, no. 6, p. 235, 2017.
[73] P. Melillo, R. Izzo, A. Orrico, P. Scala, M. Attanasio, M. Mirra, N. D. Luca, and L. Pecchia, "Automatic prediction of cardiovascular and cerebrovascular events using heart rate variability analysis," *PLoS ONE*, vol. 10, no. 3, Mar. 2015, Art. no. e0118504.

[74] B. Ambale-Venkatesh, X. Yang, C. O. Wu, K. Liu, W. G. Hundley, R. McClelland, A. S. Gomes, A. R. Folsom, S. Shea, E. Guallar, D. A. Bluemke, and J. A. C. Lima, "Cardiovascular event prediction by machine learning: The multi-ethnic study of atherosclerosis," *Circulat. Res.*, vol. 121, no. 9, pp. 1092–1101, Oct. 2017.

[75] F. Imhann, K. J. Van Der Velde, R. Barbieri, R. Alberts, M. D. Voskuil, A. V. Vila, V. Collij, L. M. Spekhorst, K. W. J. Van der Sloot, V. Peters, H. M. Van Dullemen, M. C. Visschedijk, E. A. M. Festen, M. A. Swertz, G. Dijkstra, and R. K. Weersma, "The 1000IBD project: Multi-omics data of 1000 inflammatory bowel disease patients; data release 1," *BMC Gastroenterol.*, vol. 19, no. 1, pp. 1–10, Dec. 2019.

[76] M. Abbas, J. Matta, T. Le, H. Bensmail, T. Obafemi-Ajayi, V. Honavar, and Y. EL-Manzalawy, "Biomarker discovery in inflammatory bowel diseases using network-based feature selection," *PLoS ONE*, vol. 14, no. 11, Nov. 2019, Art. no. e0225382.

[77] C. J. Cardinale, Z. Wei, J. Li, J. Zhu, M. Gu, R. N. Baldassano, S. F. A. Grant, and H. Hakonarson, "Transcriptome profiling of human ulcerative colitis mucosa reveals altered expression of pathways enriched in genetic susceptibility loci," *PLoS ONE*, vol. 9, no. 5, May 2014, Art. no. e96153.

DAVIDE CHICCO received the Bachelor of Science and Master of Science degrees in computer science from the Università di Genova, Genoa, Italy, in 2007 and 2010, respectively, and the Ph.D. degree in computer engineering from the Politecnico di Milano University, Milan, Italy, in 2014. He also spent a semester with the University of California at Irvine, USA, as a Visiting Doctoral Scholar. From September 2014 to September 2018, he has been a Postdoctoral Researcher with the Princess Margaret Cancer Centre, and a Guest of the University of Toronto. From September 2018 to December 2019, he was a Scientific Associate Researcher with the Peter Munk Cardiac Centre, Toronto, Canada. From January 2020 to January 2021, he was a Scientific Associate Researcher with the Krembil Research Institute, Toronto. Since January 2021, he has been a Scientific Research Associate with the Institute of Health Policy Management and Evaluation, University of Toronto.

GIUSEPPE JURMAN received the Ph.D. degree in algebra from the Università di Trento, Italy, in 1998. After two years as a Postdoctoral Fellow with The Australian National University (ANU), Canberra, in 2002, he moved to Fondazione Bruno Kessler (FBK), Trento, where he is currently a Senior Researcher in data science, working mainly on computational biology. Since 2008, he has been the Co-Director of the WebValley, the FBK summer school for dissemination of interdisciplinary research for high school students. He is also an Expert in scientific programming with R/Python and other computing languages. He teaches data visualization for the Master of Science degree in data science with the Università di Trento. His main research interests include machine learning, mathematical modeling, and network analysis.