BMC Bioinformatics

SOFTWARE

Open Access

R.ROSETTA: an interpretable machine learning framework

Mateusz Garbulowski[1], Klev Diamanti[1,2†], Karolina Smolińska[1†], Nicholas Baltzer[1,3], Patricia Stoll[1,4], Susanne Bornelöv[1,5], Aleksander Øhrn[6], Lars Feuk[2] and Jan Komorowski[1,7,8,9*]

*Correspondence: jan.komorowski@icm.uu.se
†Klev Diamanti and Karolina Smolińska contributed equally to this work
[1] Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden
Full list of author information is available at the end of the article

Abstract

Background: Machine learning involves strategies and algorithms that may assist bioinformatics analyses in terms of data mining and knowledge discovery. In several applications, viz. in Life Sciences, it is often more important to understand how a prediction was obtained rather than knowing what prediction was made. To this end so-called interpretable machine learning has been recently advocated. In this study, we implemented an interpretable machine learning package based on the rough set theory. An important aim of our work was provision of statistical properties of the models and their components.

Results: We present the R.ROSETTA package, which is an R wrapper of ROSETTA framework. The original ROSETTA functions have been improved and adapted to the R programming environment. The package allows for building and analyzing non-linear interpretable machine learning models. R.ROSETTA gathers combinatorial statistics via rule-based modelling for accessible and transparent results, well-suited for adoption within the greater scientific community. The package also provides statistics and visualization tools that facilitate minimization of analysis bias and noise. The R.ROSETTA package is freely available at https://github.com/komorowskilab/R.ROSETTA. To illustrate the usage of the package, we applied it to a transcriptome dataset from an autism case–control study. Our tool provided hypotheses for potential co-predictive mechanisms among features that discerned phenotype classes. These co-predictors represented neurodevelopmental and autism-related genes.

Conclusions: R.ROSETTA provides new insights for interpretable machine learning analyses and knowledge-based systems. We demonstrated that our package facilitated detection of dependencies for autism-related genes. Although the sample application of R.ROSETTA illustrates transcriptome data analysis, the package can be used to analyze any data organized in decision tables.

Keywords: Transcriptomics, Interpretable machine learning, Big data, Rough sets, Rule-based classification, R package

Background

Machine learning approaches aim at recognizing patterns and extracting knowledge from complex data. In this work, we aim at supporting the knowledge-based data mining with an interpretable machine learning framework [1, 2]. Recently, understanding the complex machine learning classifiers that explain their output is a highly important topic [3]. Here, we implemented an R package for a non-linear interpretable machine learning analysis that is based on rough set theory [4]. Moreover, we enriched our tool with basic statistical measurements that is a unique development with comparison to current state-of-the-art tools. At the end of this section, we also briefly introduce the mathematical theory behind rough sets. For a complete presentation of rough sets the reader is recommended to consult the tutorial [5] or other literature [1, 4, 6–8].

Classification models are trained on labeled objects that are a priori assigned to them. The universal input structure for machine learning analyses is a decision table or decision system [9–11]. This concept is similar to feature matrix that is well-known in image analysis [12, 13]. These structures organize data in a table that contains a finite set of features, alternatively called attributes or variables. However, decision tables are adapted for supervised learning. Specifically, a set of objects, also called examples or samples, is labelled with a decision or outcome variable. Decision system is defined as $\mathcal{D} = \left(U, A \cup \left\{d\right\}\right)$ where U is a non-empty finite set of objects called universe, A is a non-empty finite set of features and d is the decision such that $d \notin A$ [5]. Importantly, most of the omics datasets can be represented as decision tables, and machine learning analysis can be applied to a variety of problems such as, for instance, case–control discrimination. When analyzing ill-defined decision tables, i.e. tables where $|A| \gg |U|$, an appropriate feature selection step is necessary prior to the machine learning analysis [14–16]. The main goal of feature selection is to reduce the dimensionality to the features that are relevant to outcome. Thus, it is recommended to consider feature selection as a standard step prior to the machine learning analysis, especially for big omics datasets such as transcriptome data.

The ROSETTA software is an implementation of a framework for rough set classification [17]. It was implemented in C++ as a graphical user interface (GUI) and command line version. ROSETTA has been successfully applied in various studies to model biomedical problems [8, 18, 19]. Here, we present a more accessible and flexible implementation of ROSETTA that was used as the core program of the R package. R.ROSETTA substantially extends the functionality of the existing software towards analyzing complex and ill-defined bioinformatics datasets. Among others, we have implemented functions such as undersampling, estimation of rule-statistical significance, prediction of classes, merging of models, retrieval of support sets and various approaches to model visualization (Fig. 1). To the best of our knowledge, there is no framework that allows for such broad analysis of interpretable classification models. Overall, rough set-based algorithms proved successful in knowledge and pattern discovery [20–22]. Here, we illustrated the functionality of R.ROSETTA by exploring rule-based models for synthetically generated datasets and transcriptomic dataset for patients with and without autism, hereafter called the autism-control dataset.

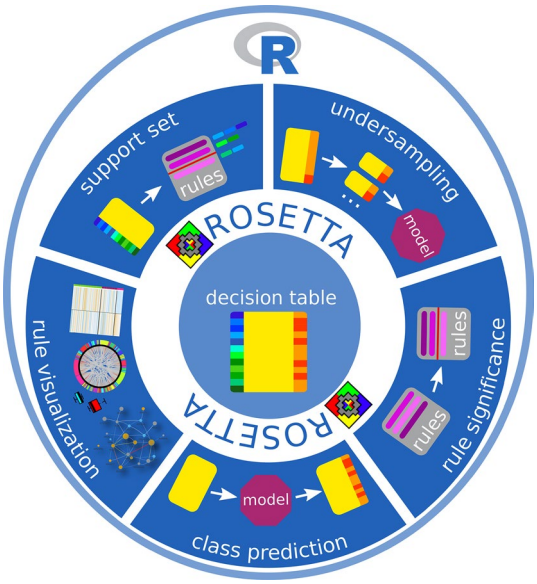

Fig. 1 Overview of R.ROSETTA and the major components that were implemented to enhance the ROSETTA functionality. (Icon with R logo included in this figure was released by the R Foundation under the CC BY-SA 4.0 license)

Rough sets

Rough set theory has become an inherent part of interpretable machine learning. In recent years, the rough sets methodology has been widely applied to various scientific areas, e.g. [23–25]. It supports artificial intelligence research in classification, knowledge discovery, data mining and pattern recognition [4, 7]. One of the most important properties of rough sets is the discovery of patterns from complex and imperfect data [26, 27]. The principal assumption of rough sets is that each object x such that$x \in X$, where$X \subseteq U$, is represented by an information vector. In particular, objects identified with the same information vectors are indiscernible. Let $\mathcal{D} = \left(U, A \cup \left\{d\right\}\right)$ be a decision system. For any subset of features $B \subseteq A$ there is an equivalence relation$IND_{\mathcal{D}}(B)$, called B-indiscernibility relation [5]:

$$IND_{\mathcal{D}}(B) = \left\{\left(x, x'\right) \in U^2 : \forall a \in Ba(x) = a\left(x'\right)\right\} \tag{1}$$

where $\left(x, x'\right)$ are objects that are indiscernible from each other by features from B if $\left(x, x'\right) \in IND_{\mathcal{D}}(B)$.

Consider three subsets of features: $B_1 = \left\{gene_1\right\}$, $B_2 = \left\{gene_2\right\}$ and $B_3 = \left\{gene_1, gene_2\right\}$ for the decision system in Table 1. Each of the indiscernibility relations defines a partition of U (1) $IND_{\mathcal{D}}(B_1) = \{\{x_1, x_2, x_3\}, \{x_4, x_5\}\}$ (2) $IND_{\mathcal{D}}(B_2) = \{\{x_1, x_3, x_5\}, \{x_2, x_4\}\}$ and (3) $IND_{\mathcal{D}}(B_3) = \{\{x_1, x_3\}, \{x_2\}, \{x_4\}, \{x_5\}\}$. For example, using B_1, objects $\{x_1, x_2, x_3\}$ are indiscernible and thus belong to the same equivalence class [5].

Let us consider decision system $\mathcal{D}$ with a subset of features $B \subseteq A$ and subset of objects$X \subseteq U$. We can then approximate X using features from B by constructing the so-called B-lower and B-upper approximations of B expressed as BX and $\overline{B}X$, respec-

Table 1 An example decision table $\mathcal{D} = (U, A \cup \{d\})$ where A is a set of two genes, U is a set of five objects $x_1, \ldots, x_5$ and d is case or control diagnosis. The values in U are discrete gene expression levels "up" or "down"

Object	$gene_1$	$gene_2$	*diagnosis*
x_1	Up	Up	Case
x_2	Up	Down	Case
x_3	Up	Up	Control
x_4	Down	Down	Control
x_5	Down	Up	Case

Table 2 An example generalized decision table $\mathcal{D} = (U, A \cup \{d\})$ for case–control study of autism, where A is a set of three genes $\{g_1, g_2, g_3\}$ and a risk factor $\{rf\}$, and U is a set of objects that belong to equivalence classes $[q_1], \ldots, [q_8]$. The values in U are discrete gene expression levels "low", "medium" or "high" and a presence of undefined risk factor "yes" or "no". For simplicity we omit the brackets in the notation in the table. For the equivalence classes q_4 and q_5 both diagnoses are written since some of the indiscernible objects belong to the boundary region

Equivalence class	g_1	g_2	g_3	*rf*	*diagnosis*
q_1	Low	Low	Medium	Yes	Autism
q_2	Medium	Medium	Medium	Yes	Autism
q_3	Medium	Low	Medium	Yes	Autism
q_4	Medium	Low	High	No	Autism or control
q_5	Low	Low	High	No	Autism or control
q_6	Low	High	Medium	No	Control
q_7	Medium	High	Medium	Yes	Control
q_8	Medium	High	High	No	Control

tively, where $\underline{B}X = \{x : [x]_B \subseteq X\}$ and $\overline{B}X = \{x : [x]_B \cap X \neq \varnothing\}$. The B-lower approximation contains objects that certainly belong to X and the B-upper approximation contains objects that may belong to$X$. The set $Bnd_B(X) = \overline{B}X - \underline{B}X$ is called B-boundary region of X. The set X is called rough if $Bnd_B(X) \neq \varnothing$ and crisp otherwise. The objects that certainly do not belong to X are in the B-outside region and their set is defined as$U - \overline{B}X$. For example, if$X = \{x : diagnosis(x) = case\}$, as in Table 1, then the approximation regions are$\underline{A}X = \{x_2, x_5\}$,$\overline{A}X = \{x_1, x_2, x_3, x_5\}$, $Bnd_A(X) = \{x_1, x_3\}$ and$U - \overline{A}X = \{x_4\}$.

For such example (Table 2) we can define another table called decision-relative discernibility matrix M as shown in the literature [5, 28]. From M we can construct a discernibility function $f_{\mathcal{D}}M(A)$ that is a Boolean [29] function in a conjunctive normal form of disjunctive literals where the literals are the names of features that discern in a pairwise fashion equivalence classes with different decisions. For example, $(g_1 \vee g_3 \vee rf)$ discerns between q_1 and q_4. Next, the Boolean formula is minimized and called a reduct. The discernibility function for our decision system is $f_{\mathcal{D}}M(A) = (g_1 \vee g_3 \vee rf)(g_2 \vee g_3 \vee rf)(g_3 \vee rf)(g_3 \vee rf)(g_1 \vee g_2 \vee g_3 \vee rf)(g_1 \vee g_3 \vee rf)(g_2 \vee rf)(g_1 \vee g_2 \vee rf)(g_1 \vee g_2 \vee rf)(g_1 \vee g_2 \vee g_3)(g_2 \vee g_3)(g_1 \vee g_2)(g_2)(g_2 \vee g_3 \vee rf)(g_1 \vee g_2 \vee g_3 \vee rf)(g_1 \vee g_2 \vee g_3 \vee rf)(g_2 \vee g_3 \vee rf)(g_2 \vee g_3 \vee rf)(g_2)(g_1 \vee g_2)$ that after a simplification results in two reducts

$f_{\mathcal{D}}M(A) = \left(g_2 \wedge rf\right) \vee \left(g_2 \wedge g_3\right)$. From the construction it follows that the reducts have the same discernibility as the full set of features. This study investigates two algorithms of computing reducts, called reducers, the Johnson reducer [30] which is a deterministic greedy algorithm, and the Genetic reducer [31] which is a stochastic method based on the theory of genetic algorithms. The reader may notice that this process is a form of feature selection. Finally, each reduct gives rise to rules by overlaying it over all objects in $\mathcal{D}$ (Table 2). For example, the first reduct $\left(g_2 \wedge rf\right)$ and the equivalence class q_1 give a rule IF $g_2 =$ low AND rf = yes THEN diagnosis = autism. The IF-part of the rule consists of conjuncts and is called the condition (or predecessor, or left-hand side) of the rule and the THEN-part is a conclusion (or successor, or right-hand side). Importantly, rules can have an arbitrary, but finite number of conjuncts.

Numerical characterization of rules

Rules are frequently described with measurements of support, coverage and accuracy. The rule support represents the number of objects that fulfill the rule conditions. Left-hand side support (LHS support) is the number of objects that satisfy the rule conjuncts i.e. IF-part of the rule. Right-hand side support (RHS support) is the number of the LHS objects of the respective classes i.e. of the THEN-part of the rule. The rule coverage can be explicitly determined from the LHS or RHS support as a percentage of objects contributing to the rule. We discern between RHS and LHS coverage:

$$coverage_{RHS}(rule) = \frac{support_{RHS}(rule)}{n_d} \tag{2}$$

$$coverage_{LHS}(rule) = \frac{support_{LHS}(rule)}{n_d} \tag{3}$$

where n_d is the total number of objects from U for a decision class d defined by the rule. Accuracy of the rule represents its predictive strength that is computed based on the support values. Specifically, accuracy for a rule is calculated as:

$$accuracy(rule) = \frac{support_{RHS}(rule)}{support_{LHS}(rule)} \tag{4}$$

Johnson reducer

The Johnson reducer belongs to the family of greedy algorithms. For the decision table $\mathcal{D} = \left(U, A \cup \{d\}\right)$, the main aim of the Johnson algorithm is to find a feature $a \in A$ that discerns the highest number of object pairs [32]. Computing reducts with Johnson approach has time complexity $O(k \bullet m^2 \bullet |R|)$, where k is the number of features, m is the number of objects and R is the computed reduct [32]. The Johnson algorithm for computing a single reduct is expressed as follows [33]: (1) Let $R = \varnothing$.(2) Let $a_{max} \in A$ be the feature that maximizes $\sum w(S)$ where $w(S)$ denotes a weight for subsets $S \subseteq \mathcal{S}$ for set $\mathcal{S}$ obtained from discernibility matrix. The sum is taken over all S from $\mathcal{S}$ that contain

a_{max}. (3) Add a_{max} to R. (4) Remove all S from $\mathcal{S}$ that contain a_{max}. (5) If $\mathcal{S} = \varnothing$ return R. Otherwise, go to step 2.

Genetic reducer

The genetic algorithm is based on Darwin's theory of natural selection [31]. This is a heuristic algorithm for function optimization that follows the "survival of the fittest" idea [34]. It simulates the selection mechanism with a fitness function f [33, 34] that rewards hitting sets$B$:

$$f(B) = (1-\alpha) \times \frac{cost(A) - cost(B)}{cost(A)} + \alpha \times min\left\{\varepsilon, \frac{|[S \subseteq \mathcal{S} : S \cap B \neq \varnothing]|}{|\mathcal{S}|}\right\} \quad (5)$$

where B are hitting sets such that $B \subseteq A$ found through the search by the fitness function, $\mathcal{S}$ is a set obtained from discernibility matrix, α is a control parameter for weighting between subset cost and hitting fraction, and ε is the degree of approximation, i.e. hitting sets B that have a hitting fraction at least ε are kept in the list. For the Genetic reducer, the most time-consuming part is the fitness computation. The time complexity for the fitness function is $O(k \bullet m^2)$ [31]. A more detailed description of applying genetic algorithm for estimating reducts can be found in [31].

Implementation

R.ROSETTA was implemented under R [35] version 3.6.0 and the open-source package is available on GitHub (https://github.com/komorowskilab/R.ROSETTA). The R.ROSETTA package is a wrapper (Additional file 1: Package architecture) of the command line version of the ROSETTA system [17, 36]. In contrast to ROSETTA, R.ROSETTA is an R package with multiple additional functionalities (Fig. 1). The following sections cover a detailed description of the new functions.

Undersampling

Class imbalance issue may lead to biased performance of the machine learning models [37, 38]. Ideally, each decision class shall contain approximately the same number of objects. To tackle this, we suggested to randomly sample a sufficient number of times the majority class without replacement in order to achieve an equal representation of classes. This approach of balancing the data is generally known as undersampling [37].

To build a balanced rule-based model, we have implemented an option that divides the dataset into subsets of equal sizes by undersampling the larger sets. By default, we require each object to be selected at least once, although the user can specify a custom number of sampled sets, as well as a custom size for each set. Classification models for each undersampled set are merged into a single model that consists of unique rules from each classifier. The overall accuracy of the model is estimated as the average value of the sub-models. Finally, the statistics of the merged rule-set shall be recalculated on the original training set using the function *recalculateRules*. Herein, the recalculation procedure compares each rule from trained model to the features from original data and calculates adjusted statistics.

Rule significance estimation

The P value is a standard measure of statistical significance in biomedical studies. Here, we introduced P value estimation as a quality measure for the rules. Classification models generated by R.ROSETTA consist of sets of varying number of rules estimated by different algorithms. In the case of the Johnson algorithm, this set contains a manageable number of rules (Table 3, Additional file 1: Table S1), while in the case of the Genetic algorithm this set can be considerably larger (Additional file 1: Tables S1, S2). In both cases, supervised pruning of rules from the models would not heavily affect the overall performance of the classifier. To better assess the quality of each rule we assume a hypergeometric distribution to compute P values [39] followed by multiple testing correction. The hypergeometric distribution estimates the representation of the rule support against the total number of objects. When estimating the P value for a rule, the hypergeometric distribution is adapted to the rule concepts:

$$P(X = x) = \frac{\binom{n_d}{x}\binom{n_o}{y-x}}{\binom{N}{y}} \tag{6}$$

where x is the RHS support of the rule, y is the LHS support of the rule, n_d is the total number of objects matching the decision class d defined by the rule,n_o is the number of objects for the decision class(es) opposite to the given rule and N is the total number of objects. Models enriched with rule P values can be pruned based on significance levels to illustrate the essential co-predictive mechanisms among the features. Additionally, the user may apply multiple testing correction. Herein, we used rigorous Bonferroni correction in order to protect from type I error and to account for the large number of rules generated by the Genetic reducer. To compare both reducers upon the same assumptions, Bonferroni correction was used also for rules generated with the Johnson reducer. However, this parameter can be tuned in R.ROSETTA for a less stringent correction that can be more adequate for models generated with the Johnson reducer only. We also implemented additional model-tuning statistical metrics for rules including risk ratio, risk ratio P value and risk ratio confidence intervals that are estimated with the R package fmsb [40]. Full set of statistical measurements is included in the output of the R.ROSETTA model.

Table 3 Performance evaluation of rules for the Johnson reduction method with undersampling. The average statistic values of rule support and accuracy are presented in the table. For the rule statistics, the most significant co-predictors (Bonferroni-adjusted $P \leq 0.05$) were selected

Class	Control		Autism	
Total number of rules	207		194	
Rule statistics	Basic	Recalculated	Basic	Recalculated
Number of rules ($P \leq 0.05$)	150	89	128	94
LHS support	13	18	13	16
RHS support	13	17	13	14
Accuracy	0.97	0.94	0.98	0.85
Top co-predictors	PPOX, NCS1	MAP7, NCKAP5L	RHPN1, ZFP36L2	NCS1, CSTB

Vote normalization in the class prediction

Rule-based models allow straightforward class prediction of unseen data using voting. Every object from the provided dataset is fed into the pre-trained machine-learning model and the number of rules for which their LHS is satisfied are counted in. In the final step, the votes from all rules are collected for each individual object. Typically, an object is assigned to the class with the majority of votes. However, for some models an imbalanced number of rules for each decision class over another may have been generated. For example, Johnson model generated more rules for control than the autism class (Table 3). This imbalance may impact the voting procedure. For such cases, we proposed adjusting for the rule-imbalance by normalizing the result of voting. Herein, vote counts represent the number of rules from trained model that match features and their discrete values from an external test set. We implemented various vote normalization methods in R.ROSETTA. Vote normalization can be performed by dividing the number of counted votes by its mean, median, maximum, total number of rules or square root of the sum of squares. We compared the performance of these methods in (Additional file 1: Table S3).

Rule-based model visualization

The model transparency is an essential feature that allows visualization of co-predictive mechanisms in a local (single rule) and global (whole model) scale. The package provides several ways for visualizing single rules, including boxplots and heatmaps (Additional file 1: Figs. S1d, S2) that illustrate the continuous levels of each feature of the selected rule for each object. Such rule-oriented visualizations gather the objects into those that belong to the support set for the given class, those that do not belong to the support set for the given class and the remaining objects for the other classes. Such graphic representations can assist towards the interpretation of individual rules of interest and visualization of interactions with respect to their continuous values.

A more holistic approach displays the entire model as an interaction network [41]. The R.ROSETTA package allows exporting the rules in a specific format which is suitable with rule visualization software such as Ciruvis [42] or VisuNet (Additional file 1: Fig. S1b) [43, 44]. Such model can be pruned to display only the most relevant co-predictive features and their levels. These approaches provide a different point of view on the interpretation of machine learning models that allow discovering known proof-of-concept and novel co-predictive mechanisms among features [45, 46].

Recapture of support sets

R.ROSETTA is able to retrieve support sets that represent the contribution of objects to rules (Additional file 1: Figs. S1d, S2). As a result, each rule is characterized by a set of objects that fulfill the LHS or RHS support. For example, in case of the gene expression data, gene co-predictors will be represented with the list of corresponding samples (patients). There are several advantages into knowing this information for the corresponding objects. Support sets contribute to uncovering objects whose levels of features might have shared patterns. Such sets may be further investigated to uncover specific subsets within decision classes. Moreover, non-significant support sets allow

detecting objects that may potentially introduce a bias to the model and might be excluded from the analysis.

Synthetic data

To evaluate rule-based modelling with R.ROSETTA, we implemented a function to create synthetic data. The synthetic dataset can be generated with a predefined number of features, number of objects and proportion of classes. Additionally, the user may choose between continuous and discrete data. The synthetic data structure is formulated as a decision table that follows the description in the introduction. A synthetic dataset is constructed from the transformation of randomly generated features computed from a normal distribution. In this approach, the randomly generated features are multiplied by the Cholesky decomposition of positive-definite covariance matrix [47]. The Cholesky decomposition D of the matrix L is calculated as:

$$D = LL^T \tag{7}$$

where L is a lower triangular covariance matrix, and L^T is conjugate transpose of L.

Results and discussion

Benchmarking

We benchmarked R.ROSETTA against three other R packages that perform rule-based machine learning including C50 [48], RoughSets [49] and RWeka [50] (Additional file 1: Benchmarking, Table S4). Using the autism-control dataset, we compared the efficacy of the classification algorithms by measuring the accuracy, the area under the ROC curve (AUC), the running time and the total number of rules (Fig. 2). To perform compatible benchmarking across algorithms, we standardized the classification procedure for equal frequency discretization and tenfold cross validation (CV). Additionally, to account for the stochasticity introduced by sampling in CV, each algorithm was executed 20 times with different seed values.

Even though R.ROSETTA produced one of the highest quality models (Fig. 2a, b), its runtime, especially for the genetic algorithm, was higher than of the other algorithms (Fig. 2d). We highlight that computing single reduct is linear in time while finding all minimal reducts is an NP-hard problem [5, 34]. In contrast to other systems, R.ROSETTA computes all minimal reducts. We believe this is an important feature since biological systems are robust and usually have alternative ways of achieving the outcome. Clearly, as a consequence of estimating multiple reducts, R.ROSETTA algorithms tend to produce more rules in comparison to other methods (Fig. 2c). It is then natural to remove the weakest rules and obtain simpler and more interpretable models. To this end, we suggest to prune the set of rules using the quality measurements such as, for instance, support, coverage or P value. Furthermore, we observed that the surveyed packages do not provide straightforward quality-statistic metrics for the model and the rules. The R.ROSETTA package includes a variety of quality and statistical indicators for models (accuracy, AUC etc.) and rules (support, P value, risk ratio etc.) in an effortlessly and R-friendly inspectable output. Notably, the evaluated packages do not include newly implemented R.ROSETTA features such as undersampling, support sets retrieval and rule-based model visualizations.

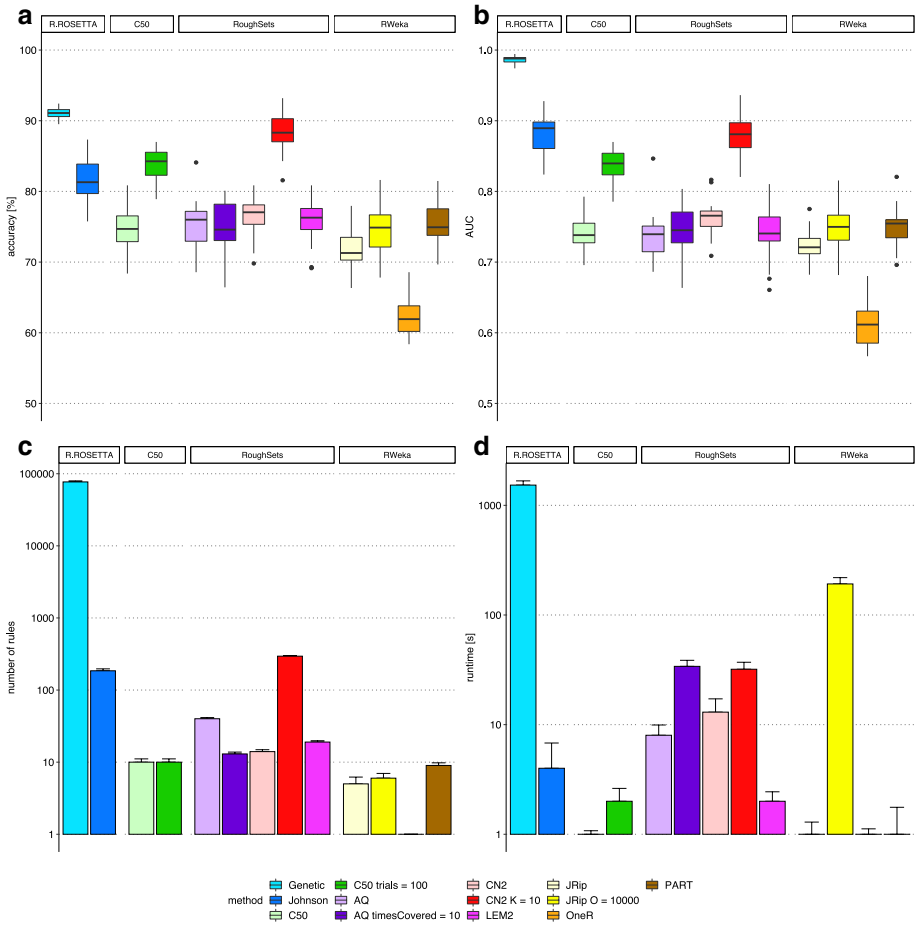

Fig. 2 Benchmarking the R packages for rule-based machine learning applied onto the autism-control dataset. The packages were evaluated with various methods for default and tuned parameters. For the R.ROSETTA package, Johnson reducer (Johnson) and Genetic reducer (Genetic) were used. For the package C50, the C5.0 method (C50) was used. For the package RoughSets, Learning from Examples Module (LEM2), CN2 rule induction (CN2) and Quasi-optimal covering Algorithm (AQ) were used. For the package RWeka Repeated Incremental Pruning to Produce Error Reduction (JRip), 1-rule classifier (OneR) and partial decision trees-based (PART). Several methods were tuned for the number of boosting iterations (*trials*), times covering objects by rules (*tc*), algorithm complexity (*K*) and number of optimizations (*O*). Other methods were evaluated with default parameters. The results of benchmarking are presented for **a** accuracy distribution of classifiers, **b** ROC AUC distribution of classifiers, **c** number of estimated rules (logarithmic scale) and **d** average runtime of the algorithms (logarithmic scale). Two standard deviations were marked above each bar. The time was measured from inputting a decision table to receiving a model

In addition, we benchmarked R.ROSETTA against methods based on decision trees, which is a concept closely related to the rule-based systems [1] (Additional file 1: Fig. S3). Both are considered as highly interpretable approaches that are able to capture non-linear dependencies among features. However, the main advantage of rough sets over the decision trees is an improved stability of models [51, 52]. To compare the performance of R.ROSETTA with tree-based methods, we investigated regression trees from package rpart [53], bagging from package ipred [54], random forest from package randomForest [55] and generalized boosted regression models from package GBM [56]. The evaluation has been performed with the autism-control dataset using the same discretization

and CV approach as for the rule-based packages. The results showed that R.ROSETTA performance is, in terms of accuracy, similar to tree-based methods (Additional file 1: Fig. S3a). However, both R.ROSETTA reduction methods had the highest median AUCs among all tested approaches (Additional file 1: Fig. S3b). Moreover, increasing the number of trees or replications for the tree-based methods resulted in a time complexity similar to the Genetic reducer, although without outperforming the rule-based classifiers (Additional file 1: Fig. S3c). Majority of benchmarked methods showed that the dataset is well-predictable. However, we emphasize that diverse datasets can perform in various ways. Furthermore, the choice of feature selection method can also play a major role in the final performance.

Sample application of R.ROSETTA on transcriptome data

To illustrate R.ROSETTA in a bioinformatics context, we applied the tool to a sample transcriptome analysis task. We examined gene expression levels of 82 male children with autism and 64 healthy male children (control) (Additional file 1: Table S5) downloaded from the GEO repository (GSE25507) [57]. The expression of 54,678 genes was measured from the peripheral blood material with the Affymetrix Human Genome U133 Plus 2.0 array. Previously, it has been reported that blood can be used as effectively as brain for transcriptomic studies of autism [58, 59]. Importantly, while obtaining the samples, blood is less invasive tissue than brain [58]. Other studies suggested that blood–brain barrier and immune system are altered in subjects with autism [60, 61]. The dataset was preprocessed (Additional file 1: Data preprocessing) and corrected for the effect of age (Additional file 1: Fig. S4). The decision table was ill-defined with the number of genes being much larger than the number of samples. To handle such high-dimensional data, we employed the Fast Correlation-Based Filter (FCBF) [62] that is a classifier-independent method of feature filtration. FCBF belongs to the group of filter-based methods, thus can be used prior to the learning processes [63, 64]. Furthermore, we favored the FCBF method as an algorithm with low time complexity and operating on pre-discretized data (Additional file 1: Feature selection, Table S6). The final decision table was reduced to 35 genes (Additional file 1: Table S7), which allowed us to generate classifiers with a reasonably low time complexity.

We constructed two models (Additional file 1: Classification) with R.ROSETTA for Johnson and Genetic reducers with 80% and 90% (0.88 and 0.99 area under the ROC curve) accuracy, respectively. Model significance was determined using a permutation test. The labels for the decision class were randomly shuffled 100 times and a new model was constructed on each modified dataset. We compared these shuffled models to the original (non-shuffled) models, and found that none of the random models resulted in a better accuracy or AUC ($P \leq 0.01$) (Additional file 1: Fig. S5). To test the influence of undersampling, we generated balanced models with 82% and 90% accuracy (0.85 and 0.98 area under the ROC curve), respectively (Fig. 3a, Additional file 1: Fig. S1a, Table S1). As the difference between unbalanced and balanced data performance was small, we analyzed these models interchangeably. For simplicity and regarding the computation time complexity, undersampling was turned off in models used for permutation tests (Additional file 1: Fig. S5) and benchmarking (Fig. 2, Additional file 1: Fig. S3).

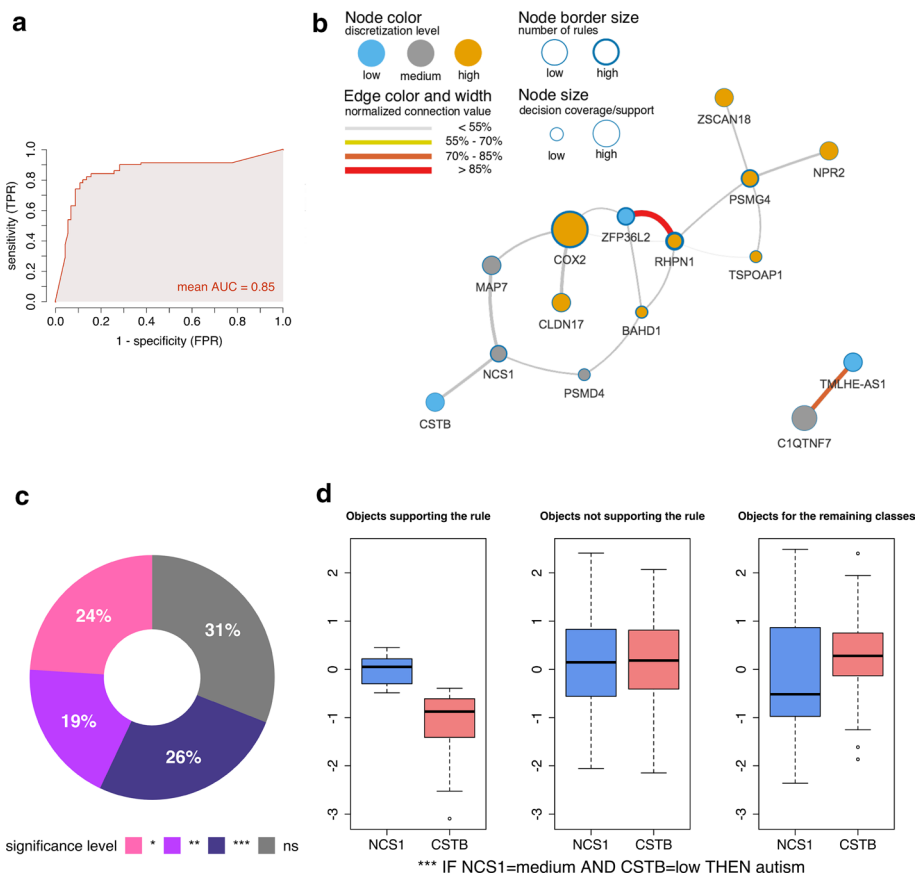

Fig. 3 Rule-based model evaluation for the autism-control dataset performed with the Johnson reduction method. Discretization levels were obtained from the equal frequency method by categorizing the features into three bins. **a** ROC AUC for the model. Sensitivity that is as a true positive rate (TPR) and 1-specificity that is a false positive rate (FPR). **b** VisuNet network of co-predictive features for the autism class. Connection values represent the strength of node or edge. These values were estimated based on the rule support and accuracy. Rules were selected based on their statistical significance (Bonferroni-adjusted $P \leq 0.05$). **c** Distribution of the significance of rules in the model. Bonferroni-adjusted P values were marked as ns($P > 0.05$), *($P \leq 0.05$), **($P \leq 0.01$) and ***($P \leq 0.001$). **d** Distribution of support sets for the top-ranked rule from the recalculated model. Support sets represent sets of objects that fulfil the RHS of the rule (THEN-part). Boxplots display scaled gene expression values for objects supporting and non-supporting the given rule

The overall performance of the Genetic algorithm was better than Johnson's. However, its tendency to generate numerous rules reduced the significance of individual rules after correcting for multiple testing (Fig. 3c, Additional file 1: Fig. S6, Table S1). To identify the most relevant co-predictors among genes, we selected several significant (Bonferroni-adjusted $P \leq 0.05$) top rules (Fig. 3d, Table 3) from the Johnson model. In addition, for the same model, we presented a sample set of strongly significant (Bonferroni-adjusted $P < 0.001$) rules in Additional file 1: Table S8. The highest ranked co-predictors include the medium expression levels of neuronal calcium sensor 1 (*NCS1*) and low expression levels of cystatin B (*CSTB*). The *NCS1* gene is related to the calcium homeostasis control [65] and is predominantly expressed in neurons [66]. In previous studies, dysregulated expression and mutations in *NCS1* have been linked to neuropsychiatric disorders [65, 66]. Moreover, another study has

demonstrated that calcium homeostasis is altered in autism disorders [67]. *CSTB* is a second component of the rule and its elevated expression have been linked to immune response [68]. Furthermore, the reduced expression of *CSTB* has been linked to the mechanism of pathogenesis in epilepsy [69].

We also utilized the VisuNet framework that supports visualization and exploration of rule-based models. Moreover, we displayed a pruned rule-based network for the significant rules for autism (Bonferroni-adjusted $P \leq 0.05$) (Fig. 3b). The largest node in the network is the cyclooxygenase 2 (*COX2*) gene and suggests a meaningful contribution to the prediction of young males with autism. Elevated expression of *COX2* has been earlier associated with autism [70]. The study reported that *COX2* carried the Single Nucleotide Polymorphism (SNP) rs2745557 and the GAAA haplotype that were significantly associated with autism [70]. Moreover, *COX2* is constitutively expressed in neuronal tissues of patients with psychiatric disorders [71]. Other studies have shown that *COX2*-deficient mice show abnormal expression of autism-related genes [72] and presented its possible therapeutic character for neuropsychiatric disorders [73, 74]. Based on the network, we can also observe a very strong co-prediction between the high expression levels of rhophilin rho GTPase binding protein 1 (*RHPN1*) and the low expression levels of *ZFP36* ring finger protein like 2 (*ZFP36L2*). The association of abnormalities in the GTPase signaling pathway and neurodevelopmental disorders has been previously reported [75]. Rho GTPases participate in the spectrum of signaling pathways related to neurodevelopment such as neurite extension or axon growth and regeneration [75]. The second component is a zinc-finger protein coding gene [76]. The enrichment of lowly expressed zinc fingers in the case–control studies of autism was also discovered by the authors of this dataset [57]. We investigated other autism-related genes that have been reported and described in Additional file 1: Feature validation. The described co-prediction mechanisms illustrate dependencies among the genes that may suggest biological interactions. Although we found relationships to neurodevelopmental and autistic pathways, given hypotheses shall be further verified experimentally.

Synthetic data evaluation

To explore the influence of the basic properties of the decision table onto the rule-based modelling, we implemented a function that generates synthetic decision tables. We used such synthetic data to describe the rule-based model performance with respect to the number of features, the number of objects and the decision-class imbalance (Additional file 1: Figs. S7, S8). Multiplying the number of features did not affect the quality of the model that remained stable across tests (Additional file 1: Fig. S7b, c). However, increasing the number of objects moderately improved the overall quality of the model (Additional file 1: Fig. S7e, f). To show that undersampling corrects biased performance that arose from the class imbalance, we generated random synthetic datasets with various imbalance proportions. We showed that the class imbalance issue biases the accuracy and the bias is corrected after applying the undersampling (Additional file 1: Fig. S8). We also confirmed that it is better to assess the performance with the AUC values, which are immune to uneven distribution of samples (Additional file 1: Fig. S8).

Conclusions

The R.ROSETTA is a response to the needs of developers of interpretable machine learning models for Life Sciences. It facilitates the access to the functionality of the R environment that is one of the major environments used in bioinformatics. To our knowledge, it is the first and only learning system that makes available a comprehensive toolbox of statistical measures essential in analyzing, validating and, potentially, certifying classifiers at the level of the models and their components. Furthermore, several original ROSETTA procedures were improved and/or adapted to the R environment and target bioinformatics applications. These improvements include undersampling methods to account for imbalance, estimation of the statistical significance of classification rules, retrieving objects from support sets, normalized prediction of external datasets and integration with rule-based visualization tools.

Rule-based models generated under the paradigm of rough sets have several attractive properties but also limitations. Models are built using a well-defined procedure of Boolean reasoning to obtain reducts i.e. minimal subsets of the original set of features that maintain discernibility of the decision classes; usually, approximate reducts are generated, both for the sake of computational efficiency but also for their often better generalization properties. Rough sets are especially useful in the applications to modelling biological systems since living organisms are robust and often have multiple ways of achieving their goals. These may be captured by multiple reducts. The price for finding all possible reducts is the complexity of the problem, which is NP-hard**,** while finding only one minimal subset of features (corresponding to finding one reduct in R.ROSETTA) is linear in time. Another disadvantage of R.ROSETTA may be the very large number of rules generated with the Genetic heuristics which makes such models more difficult to interpret. However, we showed that with the use of the toolbox it is possible to prune such model**s** and keep the statistically significant rules. Another feature of the rough set-based models is the need to discretize the values of the features. Depending on the application this at times may hamper the quality of classification, but equally well, this may improve interpretation of the models and their generalization power.

Herein, we investigated the package to describe the influence of the properties of decision tables on the rule-based learning performance. Next, a real sample application of analyzing autism and controls using gene expression data was introduced and the classifier interpreted. The autism-control dataset was also exploited to benchmark the package with a broad selection of the state-of-the-art methods within the rule- and decision tree-based domains. R.ROSETTA compared favorably with the other methods with the Genetic heuristics usually outperforming the Johnson heuristics; both heuristics compared favorably with the other systems. We showed that a rule set, be it generated with any of the two heuristics, can be visualized in the form of co-predictive rule-networks, which further enhance the interpretability of rule-based models. Finally, we investigated the performance of R.ROSETTA depending on the properties of the decision tables that are input to the system.

In contrast to methods that allow explaining black box approaches, so-called post hoc explanation methods, rough sets theory is a technique that directly produces interpretable models. On the other hand, commonly used algorithms, such as ELI5 [77], LIME [78] or SHAP [79], are able to explain most of machine learning models. However, recent

studies have shown that explanations of black box models may be affected by biases and their application has been questioned [80, 81]. Nevertheless, interpretable models and explainable methods have a common goal of elucidating classifications and are likely to complement each other.

Abbreviations
AUC: Area under the ROC curve; AQ: AQ method; C50: C.50 method; CN2: Clark and Niblett method; CV: Cross validation; FPR: False positive rate; GBM: Generalized boosted models; JRip: Repeated Incremental Pruning to Produce Error Reduction—RIPPER; Lem2: Learning from Examples Module—version 2; LHS: Left-hand site; ns: Not significant; OneR: 1R classifier method; PART: Partial decision trees-based method; *P*: *P* Value; RHS: Right-hand site; ROC: Receiver Operating Characteristic; TPR: True positive rate.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-021-04049-z.

Additional file 1. Supplementary Notes, Supplementary Figures S1–S8, Supplementary Tables S1–S8, Supplementary References

Acknowledgements
We would like to thank Fredrik Barrenäs, Stella Belin, Marco Cavalli, Zeeshan Khaliq, Behrooz Torabi Moghadam, Gang Pan, Claes Wadelius and Sara Younes for their insightful discussions and testing the package. Our thanks go also to the anonymous reviewers who helped us improve this manuscript.

Authors' contributions
MG and KS implemented the package and performed the analyses. MG and PS tested the first release of the package. MG, KD and JK wrote the manuscript. MG, KD, KS, NB, SB, AØ and JK designed the package components. MG, LF and JK contributed to results interpretation. All authors have read and approved the final manuscript.

Funding
Open access funding provided by Uppsala University. This research was supported in part by Foundation for the National Institutes of Health (Grant No. 0925-0001), Uppsala University, Sweden and the eSSENCE grant to MG, KS and JK. This work was also supported by the Swedish Research Council (Grant No. 2017-01861) to LF and by The Polish Academy of Sciences to JK. The funding was used to develop, implement and evaluate the framework. The funding bodies had no role in the design of the study, analysis, data collection and preparing the manuscript.

Availability of data and materials
The data that support the findings of this study are publicly available from the Gene Expression Omnibus repository, [GSE25507, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE25507]. The synthetic data that supports the findings of this study has been generated with the R.ROSETTA package. The package can be found at https://github.com/komorowskilab/R.ROSETTA.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden. [2] Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [3] Department of Research, Cancer Registry of Norway, Oslo, Norway. [4] Department of Biosystems Science and Engineering, ETH Zurich, Zurich, Switzerland. [5] Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK. [6] Department of Informatics, University of Oslo, Oslo, Norway. [7] Swedish Collegium for Advanced Study, Uppsala, Sweden. [8] Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland. [9] Washington National Primate Research Center, Seattle, WA, USA.

Received: 19 November 2020 Accepted: 24 February 2021
Published online: 06 March 2021

References

1. Molnar C. Interpretable Machine Learning: Lulu. com; 2020.

2. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv: 170208608 2017.
3. Azodi CB, Tang J, Shiu S-H. Opening the Black Box: Interpretable machine learning for geneticists. *Trends in Genetics* 2020.
4. Pawlak Z. Rough sets. Int J Comput Inform Sci. 1982;11(5):341–56.
5. Komorowski J, Pawlak Z, Polkowski L, Skowron A. Rough sets: a tutorial. In: Rough fuzzy hybridization: a new trend in decision-making 1999; pp. 3–98.
6. Pawlak Z, Skowron A. Rough sets and Boolean reasoning. Inf Sci. 2007a;177(1):41–73.
7. Pawlak Z, Skowron A. Rudiments of rough sets. Inf Sci. 2007b;177(1):3–27.
8. Komorowski J. Learning rule-based models — the rough set approach. Amsterdam: Comprehensive Biomedical Physics; 2014.
9. Kohavi R. The power of decision tables. In: European conference on machine learning. Springer, 1995; pp 174–189.
10. Huysmans J, Dejaeger K, Mues C, Vanthienen J, Baesens B. An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. Decis Support Syst. 2011;51(1):141–54.
11. Pawlak Z. Rough sets and intelligent data analysis. Inf Sci. 2002;147(1–4):1–12.
12. Zhang Y, Liu C, Wei S, Wei C, Liu F. ECG quality assessment based on a kernel support vector machine and genetic algorithm with a feature matrix. J Zhejiang Univ Sci C. 2014;15(7):564–73.
13. Wu C-M, Chen Y-C. Statistical feature matrix for texture analysis. CVGIP Graph Models Image Process. 1992;54(5):407–19.
14. Dash M, Liu H. Feature selection for classification. Intelligent data analysis. 1997;1(3):131–56.
15. Liu H, Motoda H. Feature selection for knowledge discovery and data mining, vol. 454. Berlin: Springer; 2012.
16. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: a new perspective. Neurocomputing. 2018;300:70–9.
17. Øhrn A, Komorowski J. Rosetta — a rough set toolkit for analysis of data. In: Proceedings of third international joint conference on information sciences 1997. Citeseer.
18. Setiawan NA, Venkatachalam PA, Hani AFM. Diagnosis of coronary artery disease using artificial intelligence based decision support system. arXiv preprint arXiv: 200702854 2020.
19. Gil-Herrera E, Yalcin A, Tsalatsanis A, Barnes LE, Djulbegovic B. Rough set theory based prognostication of life expectancy for terminally ill patients. In: 2011 annual international conference of the IEEE Engineering in Medicine and Biology Society: 2011. IEEE, pp 6438–6441.
20. Cao Y, Liu S, Zhang L, Qin J, Wang J, Tang K. Prediction of protein structural class with Rough Sets. BMC Bioinform. 2006;7(1):20.
21. Chen Y, Zhang Z, Zheng J, Ma Y, Xue Y. Gene selection for tumor classification using neighborhood rough sets and entropy measures. J Biomed Inform. 2017;67:59–68.
22. Maji P, Pal SK. Fuzzy–rough sets for information measures and selection of relevant genes from microarray data. IEEE Trans Syst Man Cybern Part B Cybern. 2009;40(3):741–52.
23. Kumar SS, Inbarani HH. Cardiac arrhythmia classification using multi-granulation rough set approaches. Int J Mach Learn Cybern. 2018;9(4):651–66.
24. Zhang J, Wong J-S, Li T, Pan Y. A comparison of parallel large-scale knowledge acquisition using rough set theory on different MapReduce runtime systems. Int J Approximate Reasoning. 2014;55(3):896–907.
25. Jothi G, Inbarani HH, Azar AT, Devi KR. Rough set theory with Jaya optimization for acute lymphoblastic leukemia classification. Neural Comput Appl. 2019;31(9):5175–94.
26. Bal M. Rough sets theory as symbolic data mining method: an application on complete decision table. Inform Sci Lett. 2013;2(1):35–47.
27. Bello R, Falcon R. Rough sets in machine learning: a review. In: Thriving rough sets. Springer; 2017; pp 87–118.
28. Skowron A, Rauszer C. The discernibility matrices and functions in information systems. In: Intelligent decision support. Springer; 1992, pp. 331–362.
29. Brown FM. Boolean reasoning: the logic of Boolean equations. Berlin: Springer; 2012.
30. Johnson DS. Approximation algorithms for combinatorial problems. J Comput Syst Sci. 1974;9(3):256–78.
31. Wroblewski J. Finding minimal reducts using genetic algorithms. In: Proceedings of the second annual join conference on information science, 1995, pp. 186–189.
32. Hoa NS, Son NH. Some efficient algorithms for rough set methods. In: Proceedings IPMU: 1996, pp 1541–1457.
33. Øhrn A. Rosetta technical reference manual. Trondheim: Norwegian University of Science and Technology, Department of Computer and Information Science; 2001.
34. Vinterbo S, Øhrn A. Minimal approximate hitting sets and rule templates. Int J Approx Reason. 2000;25(2):123–43.
35. Team RC. R: a language and environment for statistical computing. R Foundation for Statistical Computing. R version 3.6. 0. 2019.
36. Øhrn A, Komorowski J, Skowron A, Synak P. The design and implementation of a knowledge discovery toolkit based on rough sets-The ROSETTA system. 1998.
37. Liu X-Y, Wu J, Zhou Z-H 2008 Exploratory undersampling for class-imbalance learning. *IEEE Trans Syst Man Cybern Part B Cybern* 39(2):539–550.
38. Japkowicz N. The class imbalance problem: significance and strategies. In: Proceedings of the Int'l Conference on Artificial Intelligence. Citeseer; 2000.
39. Hvidsten TR, Wilczyński B, Kryshtafovych A, Tiuryn J, Komorowski J, Fidelis K. Discovering regulatory binding-site modules using rule-based learning. Genome Res. 2005;15(6):856–66.
40. Nakazawa M, Nakazawa MM: Package 'fmsb'. Retrieved from https://cran.r-project.org/web/packages/fmsb/ 2019.
41. Shmulevich I, Dougherty ER, Kim S, Zhang W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. Bioinformatics. 2002;18(2):261–74.

42. Bornelöv S, Marillet S, Komorowski J. Ciruvis: a web-based tool for rule networks and interaction detection using rule-based classifiers. BMC Bioinform. 2014;15(1):139.
43. Onyango SO. VisuNet: Visualizing Networks of feature interactions in rule-based classifiers. Uppsala: Uppsala University; 2016.
44. Smolińska K, Mateusz G, Klev D, Xavier D, Stephen O. O. A, Fredrik B, Susanne B, Jan K: VisuNet: an interactive tool for rule network visualization of rule-based learning models. https://github.com/komorowskilab/VisuNet. GitHub repository; 2021.
45. Dramiński M, Dabrowski MJ, Diamanti K, Koronacki J, Komorowski J. Discovering networks of interdependent features in high-dimensional problems. In: Big data analysis: new algorithms for a new society. Springer; 2016, pp. 285–304.
46. Enroth S, Bornelov S, Wadelius C, Komorowski J. Combinations of histone modifications mark exon inclusion levels. PLoS ONE. 2012;7(1):e29911.
47. Pourahmadi M. Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. Biometrika. 1999;86(3):677–90.
48. Kuhn M, Weston S, Culp M, Coulter N, Quinlan R: Package 'C50'. 2020.
49. Riza LS, Janusz A, Bergmeir C, Cornelis C, Herrera F, Śle D, Benítez JM. Implementing algorithms of rough set theory and fuzzy rough set theory in the R package "RoughSets." Inf Sci. 2014;287:68–89.
50. Hornik K, Buchta C, Zeileis A. Open-source machine learning: R meets Weka. Comput Statistics. 2009;24(2):225–32.
51. Li R-H, Belford GG. Instability of decision tree classification algorithms. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, 2002, pp. 570–5.
52. Dwyer K, Holte R. Decision tree instability and active learning. In: European conference on machine learning: 2007. Springer, pp. 128–39.
53. Therneau T, Atkinson B, Ripley B, Ripley MB. Package 'rpart'. 2015. https://cran.r-project.org/web/packages/rpart. Accessed 20 April 2016
54. Peters A, Hothorn T, Lausen B. ipred: improved predictors. R news. 2002;2(2):33–6.
55. Liaw A, Wiener M. Classification and regression by randomForest. R News. 2002;2(3):18–22.
56. Ridgeway G, Southworth MH. RUnit S: Package 'gbm.' Viitattu. 2013;2013(10):40.
57. Alter MD, Kharkar R, Ramsey KE, Craig DW, Melmed RD, Grebe TA, Bay RC, Ober-Reynolds S, Kirwan J, Jones JJ. Autism and increased paternal age related changes in global levels of gene expression regulation. PLoS ONE. 2011;6(2):e16715.
58. Ansel A, Rosenzweig JP, Zisman PD, Melamed M, Gesundheit B. Variation in gene expression in autism spectrum disorders: an extensive review of transcriptomic studies. Front Neurosci. 2017;10:601.
59. Enstrom AM, Lit L, Onore CE, Gregg JP, Hansen RL, Pessah IN, Hertz-Picciotto I, Van de Water JA, Sharp FR, Ashwood P. Altered gene expression and function of peripheral blood natural killer cells in children with autism. Brain Behav Immunity. 2009; 23(1):124–33.
60. Mead J, Ashwood P. Evidence supporting an altered immune response in ASD. Immunol Lett. 2015;163(1):49–55.
61. Kealy J, Greene C, Campbell M. Blood-brain barrier regulation in psychiatric disorders. Neurosci Lett. 2020;726:133664.
62. Novoselova N, Wang J, Pessler F, Klawonn F. Biocomb: feature selection and classification with the embedded validation procedures for biomedical data analysis. R Package Version 04. https://cran.r-project.org/web/packages/Biocomb. Accessed 1 Oct 2018.
63. Das S. Filters, wrappers and a boosting-based hybrid for feature selection. In: Icml: 2001; pp. 74–81.
64. Yu L, Liu H. Feature selection for high-dimensional data: a fast correlation-based filter solution. In: Proceedings of the 20th international conference on machine learning (ICML-03): 2003, pp 856–863.
65. Boeckel GR, Ehrlich BE. NCS-1 is a regulator of calcium signaling in health and disease. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 2018, 1865(11):1660–1667.
66. Handley MT, Lian L-Y, Haynes LP, Burgoyne RD. Structural and functional deficits in a neuronal calcium sensor-1 mutant identified in a case of autistic spectrum disorder. PLoS ONE. 2010;5(5):e10534.
67. Palmieri L, Papaleo V, Porcelli V, Scarcia P, Gaita L, Sacco R, Hager J, Rousseau F, Curatolo P, Manzi B. Altered calcium homeostasis in autism-spectrum disorders: evidence from biochemical and genetic studies of the mitochondrial aspartate/glutamate carrier AGC1. Mol Psychiatry. 2010;15(1):38–52.
68. Okuneva O, Li Z, Körber I, Tegelberg S, Joensuu T, Tian L, Lehesjoki A-E. Brain inflammation is accompanied by peripheral inflammation in Cstb−/− mice, a model for progressive myoclonus epilepsy. J Neuroinflammation. 2016;13(1):1–10.
69. Lalioti MD, Scott HS, Buresi C, Rossier C, Bottani A, Morris MA, Malafosse A, Antonarakis SE. Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. Nature. 1997;386(6627):847–51.
70. Yoo HJ, Cho IH, Park M, Cho E, Cho SC, Kim BN, Kim JW, Kim SA. Association between PTGS2 polymorphism and autism spectrum disorders in Korean trios. Neurosci Res. 2008;62(1):66–9.
71. Ibuki T, Matsumura K, Yamazaki Y, Nozaki T, Tanaka Y, Kobayashi S. Cyclooxygenase-2 is induced in the endothelial cells throughout the central nervous system during carrageenan-induced hind paw inflammation; its possible role in hyperalgesia. J Neurochem. 2003;86(2):318–28.
72. Wong CT, Bestard-Lorigados I, Crawford DA. Autism-related behaviors in the cyclooxygenase-2-deficient mouse model. Genes Brain Behav. 2019;18(1):e12506.
73. Sethi R, Gómez-Coronado N, Robertson ODA, Agustini B, Berk M, Dodd S. Neurobiology and therapeutic potential of cyclooxygenase-2 (COX-2) inhibitors for inflammation in neuropsychiatric disorders. Front Psychiatry. 2019;10:605.
74. Müller N, Schwarz M, Dehning S, Douhe A, Cerovecki A, Goldstein-Müller B, Spellmann I, Hetzel G, Maino K, Kleindienst N. The cyclooxygenase-2 inhibitor celecoxib has therapeutic effects in major depression: results of a double-blind, randomized, placebo controlled, add-on pilot study to reboxetine. Mol Psychiatry. 2006;11(7):680–4.

75. Reichova A, Zatkova M, Bacova Z, Bakos J. Abnormalities in interactions of Rho GTPases with scaffolding proteins contribute to neurodevelopmental disorders. J Neurosci Res. 2018;96(5):781–8.
76. Babaknejad N, Sayehmiri F, Sayehmiri K, Mohamadkhani A, Bahrami S. The relationship between zinc levels and autism: a systematic review and meta-analysis. Iranian J Child Neurol. 2016;10(4):1.
77. TeamHG-Memex: Explain like i'm five (ELI5), https://github.com/TeamHG-Memex/eli5. In. GitHub repository; 2019.
78. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining: 2016, pp. 1135–44.
79. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Advances in neural information processing systems*:* 2017, pp. 4765–74.
80. Slack D, Hilgard S, Jia E, Singh S, Lakkaraju H: Fooling lime and shap: Adversarial attacks on post hoc explanation methods. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society: 2020. 180–186.
81. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1(5):206–15.

Publisher's Note