RMI-RRG: A Soft Protocol to Postulate Monotonicity Constraints for Tabular Datasets

Iko Vloothuis and Wouter Duivesteijn(✉)

Technische Universiteit Eindhoven, Eindhoven, The Netherlands
i.vloothuis@student.tue.nl, w.duivesteijn@tue.nl

Abstract. Ensuring that a predictive model respects monotonicity constraints can enhance societal acceptance of such models. Literature on monotone classification shows that it can even improve classifier performance. However, a set of applicable monotonicity constraints is often assumed as input for the model. We propose RMI-RRG: a soft protocol that can be employed to postulate monotonicity constraints for any tabular dataset. The protocol encompasses consensus from scientific literature, aggregating the strength of (anti-)monotonicity relations in an RMI Table, aggregating the effect of imposing more constraints on the number of relabelings required to fully monotonize the dataset in a Required Relabelings Graph (RRG), and inspecting the effect on the comparability rate. We illustrate the deployment of the protocol on six datasets, arriving at some conclusions that deviate from conclusions from (mutually disagreeing) existing literature, and showing how individual steps in the protocol each have their role to play in arriving at a final postulate.

Keywords: Monotone Classification · Monotonicity Constraints · RMI-RRG Protocol · Rank Mutual Information · Required Relabelings Graph

1 Introduction

If Person A and Person B both apply for a mortgage at the same mortgage lender, if the data we have on both persons is identical except that we know that Person A has a higher income, then it would be strange if Person B is approved for the mortgage while Person A is not. In data mining we call such an event a *violation* of a *monotonicity constraint*: all else equal, if the value of certain input variables increase, we cannot have a decrease in the value of the target. Literature (e.g. [6]) on monotone classification and regression has shown that making it mandatory for a prediction model to respect given monotonicity constraints can keep predictive performance at the same level while guaranteeing no monotonicity violations, or even increase predictive performance.

Humans react quite viscerally to monotonicity constraints and their violations: we automatically connect this to a feeling of (un-)fairness, sometimes in

I. Miliou et al. (Eds.): IDA 2024, LNCS 14641, pp. 16–27, 2024.
https://doi.org/10.1007/978-3-031-58547-0_2

an almost axiomatic way (such as in the mortgage example above). However, given any tabular dataset, it is not necessarily obvious which monotonicity constraints are reasonable to postulate. Two paths exist in literature. The *subjective* approach inspects the domain of the dataset, reasons about the intrinsic meaning of the target variable, connects this to reasoning about the intrinsic meaning of input variables, and draws a conclusion about which monotonicity constraints to impose. In the mortgage example this makes sense, but it comes with two drawbacks. On the one hand, many datasets exist where monotonicity constraints might be helpful but it is not that easy to reason about the domain of the dataset, so this is hard to put in practice. On the other hand, the subjective nature of this process may lead different researchers to draw different conclusions. The *objective* approach measures something objective about the relation between input variables and the output variable, and puts a threshold on that measurement: any input variables whose relation to the output variable surpasses the threshold, are deemed to have a monotonicity relation with the output variable. Such a procedure is clear, objective, and ensures that anybody who runs this procedure will always find the same answer; said answer can also be completely arbitrary (due to the need to fix the threshold level somewhere) and lead to suboptimal results. No existing solution is satisfactory.

In this paper, we introduce the *RMI-RRG protocol*: a soft protocol to answer the question: "Given any tabular dataset, which monotonicity constraints should we postulate?" It incorporates aspects of both approaches: the subjective and the objective. We do not claim that all researchers that follow this protocol will derive the exact same conclusions; this is why we call it a soft protocol. But we show how a scientific consensus can give us a good first insight. We show that a summary of the direction and strength of monotonicity relations as provided in an *RMI Table* (cf. Sect. 5) can sharpen the image. We show how we can correct misleading conclusions from RMI Tables, by inspecting the *Required Relabelings Graph* (RRG, cf. Sect. 6): a visualization of the effect that imposing additional constraints has on the number of relabelings required to fully monotonize the dataset. We show how inspection of comparability rates and mRMR values can finetune the conclusions. In the end, we deploy the RMI-RRG protocol on six datasets, resulting in postulating monotonicity constraints on each dataset.

2 Related Work

Monotonicity is a fundamental concept spanning mathematics and various disciplines, denoting the preservation of order. In simple terms, a function or relation is deemed to satisfy (anti-)monotonicity if it maintains the (reversed) order of its inputs in the outputs. Specifically, for any pair of values, if one value is greater than or equal to the other, the corresponding output must retain the same (or, in case of antimonotonicity, reversed) order.

Monotonicity finds applications across diverse fields such as economics, statistics, optimization, mathematics, and computer science. In economics, it primarily models preferences and utility functions [5], while statistics employs it for non-parametric regression and analyzing ordered categorical data [21]. In optimization, monotonicity often serves as a foundation for convergence and efficiency [26]. In machine learning and pattern recognition, monotonicity assumptions can enhance model interpretability and generalization. By introducing monotonicity constraints, meaningful trends and relationships in data can be captured. This proves invaluable in domains like credit scoring, risk assessment, or fraud detection, where monotonic patterns are expected [17].

Classification is one domain that notably benefits from monotonicity. Leveraging monotonic relationships empowers classifiers to align with domain-specific expectations, refine model performance, and furnish more dependable and interpretable predictions [15,19].

For a survey on classification with monotonicity constraints, see [4]. The paper encompasses extensively employed datasets, monotonic preprocessing, relabeling techniques, and a diverse set of classifiers respecting monotonicity constraints. One of the earlier applications of monotone constraints to kNN classification [6] showed that adapting kNN to respect monotonicity constraints can be achieved without loss of predictive accuracy; in fact, predictive accuracy increases on some datasets. Monotonicity constraints have also been imposed when employing kNN as a kernel method [16], yielding markedly superior results compared to the baseline kernel estimators. Other notable monotone classifiers are MonoBoost [2] and MonGel [8].

3 Preliminaries and Notation

Let X denote the $n \times p$ matrix of input variables, where each row represents an observation with values x. Let $\pi_S(X)$ denote the projection of X on any subset $S \subseteq \{1, \ldots, p\}$ of the input variables. Let Y denote an $n \times 1$ vector representing the output variable, taking on values from a one-dimensional space Y, where $Y \subseteq \mathbb{R}$. Let $D = \{(x_i, y_i)\}_{i=1}^{n}$ denote the bag of observed data points.

A function f is said to be monotone if it preserves order, meaning that the ordering of the input values is preserved in the output values. Two types of monotone functions exist: increasing or decreasing (also known as non-decreasing or non-increasing, respectively). Specifically, a function is monotone increasing if, for any two values x and x' in the domain of f, it holds that $x \leq x' \Rightarrow f(x) \leq f(x')$, or, for a decreasing monotone function: $x \leq x' \Rightarrow f(x) \geq f(x')$.

A partial order is assumed for X, denoted by $\leq_X$. The partial order establishes criteria for determining the order between input values in X. A pair of points (x_i, x_j) in a dataset are said to be *comparable* if $x_i \geq x_j \vee x_j \geq x_i$, and *incomparable* if not. A pair of comparable data points (x_i, y_i) and (x_j, y_j) is considered monotone if the ordering of the input variables is preserved in the output labels. In other words, if $x_i \geq_X x_j$, then it must hold that $y_i \geq y_j$. Dataset D is considered monotone if all combinations of pairs of data points are either monotone or incomparable. We overload the partial order notation with any subset S of input variables, such that $x_i \leq_S x_j$ evaluates the comparability of x_i and x_j only on $\pi_S(X)$.

Definitions stated above assume a direct monotone restriction on the data, but an inverse relation is also possible. In that case the definition for an antimonotone pair of data points would be $x_i \geq x_j \Rightarrow y_i \leq y_j$.

For any data point x_i and subset S of input variables, let the *upset* $\uparrow_{(i,S)}$ and *downset* $\downarrow_{(i,S)}$ be:

$$\uparrow_{(i,S)} = \{x_j \in D \mid x_i \leq_S x_j\} \qquad \downarrow_{(i,S)} = \{x_j \in D \mid x_j \leq_S x_i\}$$

The upset of x_i consists of all points whose input values are considered 'higher', when limited to variables in S, than x_i. The downset consists of all data points 'lower' than x_i. Again, for variables with an antimonotone relation these definitions would also be inverted.

3.1 Rank Mutual Information

Rank Mutual Information (RMI) [12] evaluates the monotonic consistency between variables. The RMI between variables A and B is:

$$\mathrm{RMI}(\{A\},\{B\}) = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{\mid \uparrow_{(i,\{A\})} \mid \cdot \mid \uparrow_{(i,\{B\})} \mid}{n \cdot \mid \uparrow_{(i,\{A\})} \cap \uparrow_{(i,\{B\})} \mid}$$

Applying this formula to an input and an output variable will deliver a value that can range from $[-\infty, \infty]$. In practice this value rarely ranges beyond $(-1.5, 1.5)$. The RMI gives the properties of strength (sign of the RMI) and direction (magnitude of the RMI) between one or more variables and a target. RMI measures the type of information needed while also being fairly robust against noise [11].

3.2 Relabeling

Some data mining algorithms (e.g., [6]) that guarantee respecting monotonicity constraints in their predictions, require as input a dataset that fully respects those same constraints. If the training data respects monotonicity constraints, any prediction such a method makes on test/validation data will necessarily also respect the constraints. A problem is that many datasets contain some monotonicity violations. A typical approach then is to relabel values in the affected target columns, so as to not violate monotonicity. Several optimal relabeling methods have been proposed, including Feelders relabeling [7,10], Single-pass Optimal Ordinal relabeling [20], and Optimal Flow Network relabeling [6,7]. These methods minimize the number of instances to relabel, and the relabeled dataset can be viewed as a monotone classifier that minimizes the error rate on the training data [6].

4 Main Direct Competitors

We will briefly go over our main competitors in this section. Keep in mind that the final goal of each method in this section is to come up with a set of *postulates*:

these are the input variables for which we have decided that a monotonicity constraint can be imposed between the input and the output variable. Along the way, we will often measure some *monotonicity relations* between input variables and the output variable that can come with a given *strength*. The final outcome, however, will have to turn these fuzzy monotonicity relations into a set of hard postulated monotonicity constraints.

4.1 Subjective Approaches

The traditional approach is the subjective one, where scientists reason about the domain of the datasets under consideration. One typically goes by the common sense approach as exemplified by the first paragraph of Sect. 1 of this paper: if two persons are identical except that Person A has a higher income than Person B, it would be ridiculous if Person B gets a mortgage while Person A does not. This "it stands to reason" test has been the gold standard in classification under monotonicity constraints. An example is provided by [6]. The most extensive example of this procedure, involving a careful evaluation of each dataset and the identification of variables based on domain knowledge, is given in [24].

4.2 Objective Approaches

Most of the objective strategies assume the existence of monotonicity relations between input variables and the output variable, and will then use a combination of the strength and direction of this relation to decide whether to include it as a postulate. The most common and simple strategy here is to assume that all variables have a relation with the target and assign a direction to them. This can be seen in [22] and [27], where tables containing information on the datasets clearly show the inclusion and direction of all chosen datasets; how the direction of the relation was established remains opaque.

Rank Mutual Information (RMI) can represent the strength of all monotonicity relations in the dataset; subsequently, a hard constraint on its values can be imposed to arrive at a set of monotonicity constraints. The authors of [4] have used $|\text{RMI}| > 0.1$ as the cutoff value: a monotonicity constraint will be postulated on an input variable and the output variable if and only if their absolute RMI value surpasses 0.1. Personal communication with the authors revealed that there seems to be no substantial reason for choosing this particular cutoff value.

The authors of [13] propose to use the RMI in combination with the *min-Redundancy Max-Relevance* (mRMR) algorithm to calculate optimal subsets of variables. The goal here is to iteratively add the variable with the highest ranked RMI value to the chosen set, while taking into account whether doing so would merely introduce redundant information. When to stop adding variables is not specified: the authors recommend to choose a fixed number of variables and apply that to each dataset, which strikes us as a blunt, arbitrary instrument.

Table 1 lists methods used in selecting which monotonicity constraints to postulate during experimentation, with an example of a paper deploying that strategy. Clearly, there is a wide range of strategies without any clear consensus.

5 RMI Tables and Required Relabelings Graphs

Our protocol employs two forms of data aggregation. On the one hand, for each dataset, we will order the input variables by decreasing absolute RMI values (cf. Table 3 for examples of such *RMI Tables* on real datasets); denote this ordering of input variables by $x_{(1)}, x_{(2)}, \ldots, x_{(p)}$. It stands to reason that in terms of monotonicity relations with the target variable, $x_{(1)}$ will be the input variable with the strongest relation, and $x_{(p)}$ will be the input variable with the weakest relation. On the other hand, subsequently, for each dataset, we will generate the *Required Relabelings Graphs* (RRGs). This is a line graph consisting of p observations. For each observation $1 \leq i \leq p$, we compute the number of relabelings required to make the dataset fully monotone, postulating a monotonicity relation between the output variable and *all* of the top-i strongest-related variables $\{x_{(1)}, \ldots, x_{(i)}\}$. Informally, as our index i increases, we keep expanding the set of input variables for which we postulate a monotonicity constraint (starting with the strongest such relations). As this set expands, one could reasonably expect the constraints to become more pressing, but the number of comparable pairs of observations will reduce. See Fig. 1 for examples of RRGs on real datasets. From RRGs one would hope to observe behavior such as in Fig. 1a: the number of required relabelings decreases, but the curve flattens beyond some point, indicating that postulating more constraints likely has limited benefits.

The point of these data aggregations is: we think RMI has something interesting to say about monotonicity relations, but boiling that message down to a threshold on the RMI value is likely too blunt an instrument. Inspecting the RMI Tables and the RRGs should provide more information.

6 The RMI-RRG Protocol

The RMI-RRG Protocol is a soft protocol for postulating monotonicity constraints on any tabular dataset, encompassing the following four steps:

1. Literature consistency check: identify input variables that are consistently recognized in the existing literature.
2. RMI alignment: cross-reference the input variable with their respective RMI values (cf. Table 3).

Table 1. Monotonicity constraint postulation strategies.

Strategy	Example reference
Selecting constraints based on a priori information	[2]
Selecting constraints based on other literature	[24]
Postulating all constraints	[9]
Selecting constraints based on RMI	[4, 18]
No explanation given	[1]

Table 2. Dataset metadata.

D	n	p	Y	Y type	Relabels	comp. %	Source
Breast Cancer	683	9	Class	bin	12	0.821	[14]
Car	1728	6	Decision	cat (4)	466	0.528	[14]
CMC	1473	9	contraceptive	cat (3)	865	0.657	[14]
Pasture	36	22	pasture-prod-class	cat (3)	2	0.694	[23]
PIMA	768	8	Outcome	bin	198	1.0	[14]
Windsor	546	11	price	num	241	0.414	[14]

3. RRG analysis: address any discrepancies between literature consensus and RMI values by inspecting the Required Relabelings Graph (cf. Fig. 1).
4. Comparability assessment: evaluate whether the inclusion of a variable impacts the comparability of the dataset or necessitates a substantial number of relabels; exclude those variables with an outsized negative impact.

Whether input variables are included as a monotone or an antimonotone relation is decided by the sign of the RMI value.

7 Experimental Results

We collected a set of datasets, including the ten most commonly used ordinal datasets in monotone classification literature [4], the datasets used in the original kNN classification under monotonicity constraints paper [6], and four additional ordinal datasets [3]. From this set of 16 datasets, we report results on a selection of six datasets in this paper: these are the six datasets from which the most interesting conclusions could be drawn. Results on the other ten datasets are available in the MSc thesis [25, Chapter 5] from which this paper was derived.

Table 2 provides metadata on the datasets in this paper. The column 'relabels' contains the number of relabels necessary to make the dataset fully monotone (cf. Sect. 3.2), and 'comp. %' shows the fraction of all possible pairs in the dataset that are comparable. The last column indicates the source from which the dataset was obtained.

7.1 Breast Cancer

The RMI values of the first few attributes (cf. Table 3a) are fairly close together and seem promising. Figure 1a reveals that imposing constraints on variables after *Bland_ Chromatin* flattens the curve, which could be an indication of these subsequent variables being redundant. Checking the redundancy between the variables with the highest RMI scores and these latter ones confirms this. We postulate monotonicity constraints on *Uniformity_ Size*, *Uniformity_ Shape*, *Bare_ Nuclei*, and *Bland_ Chromatin*.

Table 3. RMI Tables for every dataset in Table 2; only a limited number of attributes with highest absolute RMI are shown, due to space limitations.

Variable	Uniform_Size	Uniform_Shape	Bare_Nuclei	Bland_Chromatin	Single_Cell_Size
RMI	0.501	0.499	0.473	0.459	0.442

(a) Breast Cancer.

Variable	safety	no. persons	buying price	maintenance cost	lug_boot	no. doors
RMI	0.246	0.175	-0.107	-0.095	0.078	0.032

(b) Car.

Variable	wife_age	wife_edu	husband_occup	num_child	SOL_index	wife_working	husband_edu
RMI	-0.194	0.052	0.035	0.028	0.025	0.023	0.021

(c) CMC.

Variable	HFRG-pct-mean	OlsenP	Leaf-P	Eworms-main-3	MinN	NFIX-mean
RMI	0.659	0.598	0.592	0.561	0.54	0.526

(d) Pasture.

Variable	Glucose	BMI	Pregnancies	Age	DPF	Insulin	BloodPressure	SkinThickness
RMI	0.316	0.187	0.139	0.138	0.121	0.096	0.093	0.087

(e) PIMA.

Variable	lotsize	stories	bathrooms	garage	bedrooms	aircon	prefer	recreation
RMI	0.631	0.349	0.347	0.316	0.299	0.291	0.192	0.134
mRMR	0.63	0.31	0.30	0.29	0.29	0.28	0.14	0.13

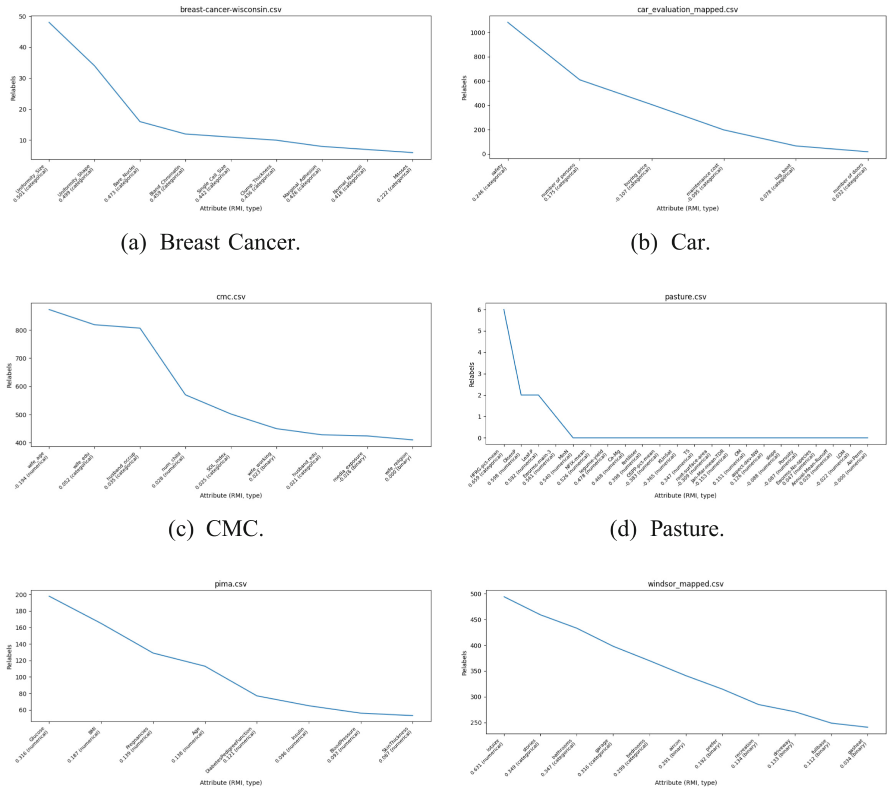
(f) Windsor.

7.2 Car

This dataset proved to be rather controversial. Most literature [4, 22] claims that all input variables have a direct relation with the target variable. The RMI values (cf. Table 3b, Fig. 1b) disagree: only two or three variables seem even remotely fit to consider. Assuming constraints on all input variables, the fraction of observation pairs that is comparable is 14%. Limiting the constraints to only *safety*, *number of persons*, and *buying price*, increases the comparability rate to 52%. Hence, we postulate only those monotonicity constrains.

7.3 CMC

The CMC dataset illustrates the added value of the RRG on top of the RMI Table. From Table 3c, the *wife_ age* variable seems to be the only sensible candidate for a monotonicity constraint: existing objective approaches [4] would draw that conclusion. But Fig. 1c shows that something more interesting is happening. While assuming constraints on *wife_ edu* and *husband_ occup* does not influence the number of required relabels much, subsequent variables do have an outsized



(a) Breast Cancer. (b) Car.

(c) CMC. (d) Pasture.

(e) PIMA. (f) Windsor.

Fig. 1. Required Relabelings Graphs (RRGs) for every dataset in Table 2.

influence. Hence, we postulate monotonicity constraints on *wife_ age*, *wife_ edu*, *husband_ occup*, and *num_ child*. As part of a reason why this RRG moves in this strange shape, we hypothesize this is caused by *wife_ edu* and *husband_ occup* being categorical variables while *wife_ age* and *num_ child* are numerical. Postulating monotonicity constraints on numerical variables will most likely lead to smoother behavior than postulating such constraints on categorical variables, since the changes in the values for numerical variables can be much more subtle.

7.4 Pasture

This dataset only contains 37 observations, which means that the comparability and the number of relabels becomes more important than the raw values in Table 3d. Figure 1d shows that imposing constraints on variables after *FRG-pct-mean*, *OlsenP*, and *Leaf-P*, would lower the number of relabels to 0. This suggests that the number of comparable pairs of observations decreases drastically, which is indeed confirmed by measurement: comparability rate drops from 69% to

43% upon inclusion of a fourth constraint. Hence, we postulate monotonicity constraints on *FRG-pct-mean*, *OlsenP*, and *Leaf-P*.

7.5 PIMA

Literature disagrees on treatment of this dataset: [24] assumes monotonicity constraints on four variables, where [4] uses six. Only the *Glucose* variable has a substantially outlying RMI value (cf. Table 3e). Figure 1e shows that constraints on additional variables do not impact the number of required relabels in any unusual way. By lack of reason to decide otherwise, we exclude from consideration all variables whose absolute RMI value is lower than 0.1, hence postulating monotonicity constraints on *Glucose*, *BMI*, *Pregnancies*, *Age*, and *DPF*.

7.6 Windsor

There is no literature that specifically states which variables and directions are to be constrained in this dataset. The RMI values (cf. Table 3f) indicate that any variable lower than 0.2 is probably best excluded. Figure 1f shows that there are no unusual hitches for any specific variable. The mRMR values do show that redundant information is introduced for any variable after *aircon*, which can be seen in the drop in value. Hence, we postulate monotonicity constraints on all variables except *prefer* and *recreation*.

8 Conclusions

Monotonicity constraints make sense to human intuition. Ensuring that predictive models respect monotonicity constraints can increase societal acceptability of such models, and even increase their predictive performance. However, there is no consensus in scientific literature on how to determine which monotonicity constraints to assume for a given dataset. This paper provides RMI-RRG, a soft protocol to postulate monotonicity constraints for any tabular dataset.

Section 5 introduces two forms of data aggregation employed in the protocol. On the one hand, the RMI Table orders the input variables of a dataset by their perceived strength of (anti-)monotone relation to the target variable; it also provides the perceived direction of this relation. On the other hand, the Required Relabelings Graph (RRG) shows more detailed effects, of imposing constraints on additional variables, on the number of relabelings required to fully monotonize the dataset. The RMI-RRG protocol (cf. Sect. 6) incorporates these two aggregation methods, along with consensus from literature and assessment of comparability rates, into a final judgment.

In Sect. 7 we illustrate application of the RMI-RRG protocol on six datasets. Results encompass fairly typical behavior (*Breast cancer* dataset), a clear conclusion that disagrees with literature (*Car*), the added value of the RRG over existing methods (*CMC*), the necessity to look at comparability rates (*Pasture*) and mRMR values (*Windsor*), and a decisive conclusion where existing literature draws multiple distinct conclusions (*PIMA*).

RMI-RRG is a soft protocol: while we think it will often lead in a fairly clear direction, it is still possible that multiple scientists can arrive at multiple distinct conclusions. We think that this is unavoidable: monotonicity is, to a degree, in the eye of the beholder. We claim that the ambiguity of RMI-RRG is better than the objective arbitrariness of setting a hard threshold of 0.1 on the absolute RMI values [4], while RMI-RRG provides more information leading to more informed decisions than existing subjective approaches [6,24]. Future improvements might include separate treatment for categorical and numerical input variables (cf. Sect. 7.3), and a tradeoff between the number of required relabelings and the comparability rate.

References

1. Alcala-Fdez, J., Alcala, R., Gonzalez, S., Nojima, Y., Garcia, S.: Evolutionary fuzzy rule-based methods for monotonic classification. IEEE Trans. Fuzzy Syst. **25**(6), 1376–1390 (2017)
2. Bartley, C., Liu, W., Reynolds, M.: A novel framework for constructing partially monotone rule ensembles. In: Proceedings ICDE, pp. 1320–1323 (2018)
3. Bellmann, P., Lausser, L., Kestler, H.A., Schwenker, F.: A theoretical approach to ordinal classification: feature space-based definition and classifier-independent detection of ordinal class structures. Appl. Sci. **12**(4), 1815 (2022)
4. Cano, J.R., Gutiérrez, P.A., Krawczyk, B., Woźniak, M., García, S.: Monotonic classification: an overview on algorithms, performance measures and data sets. Neurocomputing **341**, 168–182 (2019)
5. Carlier, G., Dana, R.A.: Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints. Stat. Risk Model. **24**(1), 127–152 (2006)
6. Duivesteijn, W., Feelders, A.: Nearest neighbour classification with monotonicity constraints. In: Daelemans, W., Goethals, B., Morik, K. (eds.) Machine Learning and Knowledge Discovery in Databases. Lecture Notes in Computer Science(), vol. 5211, pp. 301–316. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-87479-9_38
7. Feelders, A.: Monotone relabeling in ordinal classification. In: Proceedings ICDM, pp. 803–808 (2010)
8. García, J., Fardoun, H.M., Alghazzawi, D.M., Cano, J.R., García, S.: MoNGEL: monotonic nested generalized exemplar learning. Pattern Anal. Appl. **20**(2), 441–452 (2015)
9. González, S., García, S., Li, S.T., John, R., Herrera, F.: Fuzzy k-nearest neighbors with monotonicity constraints: moving towards the robustness of monotonic noise. Neurocomputing **439**, 106–121 (2021)
10. González, S., Herrera, F., García, S.: Monotonic random forest with an ensemble pruning mechanism based on the degree of monotonicity. New Gener. Comput. **33**(4), 367–388 (2015)
11. Hao, H., Wang, M., Tang, Y.: Feature selection based on improved maximal relevance and minimal redundancy. In: Proceedings IMCEC, pp. 1426–1429 (2016)

12. Hu, Q., Che, X., Zhang, L., Zhang, D., Guo, M., Yu, D.: Rank entropy-based decision trees for monotonic classification. IEEE Trans. Knowl. Data Eng. **24**(11), 2052–2064 (2012)
13. Hu, Q., Pan, W., Song, Y., Yu, D.: Large-margin feature selection for monotonic classification. Knowl.-Based Syst. **31**, 8–18 (2012)
14. Kelly, M., Longjohn, R., Nottingham, K.: The UCI machine learning repository. http://archive.ics.uci.edu/ml
15. Kotlowski, W., Slowinski, R.: Rule learning with monotonicity constraints. In: Proceedings ICML, pp. 537–544 (2009)
16. Li, Z., Liu, G., Li, Q.: Nonparametric kNN estimation with monotone constraints. Economet. Rev. **36**(6–9), 988–1006 (2017)
17. Pan, W.: Fraudulent firm classification using monotonic classification techniques. In: Proceedings ITAIC, pp. 1773–1776 (2020)
18. Pei, S., Hu, Q.: Partially monotonic decision trees. Inf. Sci. **424**, 104–117 (2018)
19. Potharst, R., Feelders, A.J.: Classification trees for problems with monotonicity constraints. ACM SIGKDD Explor. Newslett. **4**(1), 1–10 (2002)
20. Rademaker, M., De Baets, B., De Meyer, H.: Optimal monotone relabelling of partially non-monotone ordinal data. Optim. Methods Softw. **27**(1), 17–31 (2012)
21. Saarela, O., Rohrbeck, C., Arjas, E.: Bayesian non-parametric ordinal regression under a monotonicity constraint. Bayesian Anal. **18**(1), 193–221 (2023)
22. Suárez, J.L., González-Almagro, G., García, S., Herrera, F.: A preliminary approach for using metric learning in monotonic classification. In: Fujita, H., Fournier-Viger, P., Ali, M., Wang, Y. (eds.) Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence. Lecture Notes in Computer Science(), vol. 13343, pp. 773—784. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08530-7_65
23. Vanschoren, J., van Rijn, J.N., Bischl, B., Torgo, L.: OpenML: networked science in machine learning. SIGKDD Explor. **15**(2), 49–60 (2013)
24. Verbeke, W., Martens, D., Baesens, B.: RULEM: a novel heuristic rule learning approach for ordinal classification with monotonicity constraints. Appl. Soft Comput. **60**, 858–873 (2017)
25. Vloothuis, I.: Imputing missing data using k-Nearest Neighbors under monotonicity constraints. Master's thesis, Technische Universiteit Eindhoven (2023)
26. Zhou, J., Mayne, R.W.: Monotonicity analysis and the reduced gradient method in constrained optimization. J. Mech., Transmiss., Autom. Des. **106**(1), 90–94 (1984)
27. Zhu, H., Liu, H., Fu, A.: Class-weighted neural network for monotonic imbalanced classification. Int. J. Mach. Learn. Cybern. **12**(4), 1191–1201 (2021)