Enhancing Cost-Sensitive Tree-Based XAI Surrogate Method: Exploring Alternative Cost Matrix Formulation*

Marija Kopanja[1,*], Miloš Savić[1] and Luca Longo[2]

[1]*Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovića 3, 21000 Novi Sad, Serbia*

[2]*Artificial Intelligence and Cognitive Load Research Lab, The Centre of Explainable Artificial Intelligence, Technological University Dublin, Central Quad, CQ-214 Grangegorman Campus, D07 ADY7 Dublin, Ireland*

Abstract

This study investigates how different cost matrix formulations influence cost-sensitive tree extraction method performance within the post-hoc model-agnostic XAI framework. As an input parameter, the cost matrix is essential in building cost-sensitive tree models. The initial, default version of the cost matrix is defined to reflect the class imbalance ratio among each pair of classes. Here, two different formulations of the alternative cost matrix are proposed: centroid distance-based and medoid distance-based cost matrix. The cost-sensitive tree method with different formulations of cost-matrix is compared against other tree-based and rule-based XAI methods as a surrogate model for the underlying black-box model. Evaluation metrics are employed to assess the generated explanations, and results demonstrate that rule sets extracted from cost-sensitive trees are smaller with shorter rules on average across different datasets with varying number of classes.

Keywords

Explainable artificial intelligence, Cost-sensitive decision tree, Surrogate modeling, Rule extraction, Tree-based methods, Model-agnostic explanations, Rule-based systems, Interpretability, Machine Learning

1. Introduction

Explainable artificial intelligence (XAI) is one of the fastest emerging sub-fields of AI dedicated to developing methods for making machine learning (ML) models more understandable and transparent [1, 2]. To extract information from already trained models, several different methods are developed for explaining their inferential process post-hoc (after the model has been trained), without modifying the internal structure or training process of the model. Creating a surrogate model is a post-hoc approach [3] used to approximate the decision-making process of the original model by using simple models such as decision trees, linear or rule-based models, which are typically interpretable and offer a more understandable and transparent view of the decision-making process.

Decision trees and rule sets are graphical and textual representations types of explanations that are easily understandable and interpretable [4]. The cost-sensitive rule and tree extraction method CORTEX [5] investigated in this study provides two easily understandable and interpretable explanation forms: a cost-sensitive tree model and a set of rules. Cost-sensitive trees are an important category of tree methods created by using a cost-sensitive supervised approach that considers various costs during the learning process, such as misclassification costs (incorrectly classifying a sample), feature costs (the cost of obtaining the feature values) or other related costs [6]. By incorporating misclassification costs for each class into the learning process, cost-sensitive algorithms can effectively address the class imbalance problem, a well-known issue in the ML community that occurs when the number of samples is uneven across classes. In a class-dependent cost matrix, samples from the same class have the same costs, as opposed to a sample-dependent cost matrix, in which each sample may have a different cost. The CORTEX is grounded in a cost-sensitive decision tree algorithm introduced for the binary

Late-breaking work, Demos and Doctoral Consortium, colocated with The 3rd World Conference on eXplainable Artificial Intelligence: July 09–11, 2025, Istanbul, Turkey

*Corresponding author.

marija.kopanja17@gmail.com (M. Kopanja); svc@dmi.uns.ac.rs (M. Savić); luca.longo@tudublin.ie (L. Longo)

0000-0003-1176-5677 (M. Kopanja); 0000-0003-1267-5411 (M. Savić); 0000-0002-2718-5426 (L. Longo)

classification framework [7] with a sample-dependent cost matrix, where each sample has its cost matrix defined. Specifically, the two-class sample-dependent cost-sensitive framework has been adapted into a multi-class class-dependent framework by introducing an n-dimensional class-dependent cost matrix. In our previous research, the default (ratio-based) cost matrix was initially developed based on class imbalance ratios, providing a foundational approach to address the skewness of the class distribution. However, the CORTEX method can effectively operate with a balanced distribution of target variables since cost matrix definitions allow for generating a symmetric matrix in such cases.

The CORTEX has limitations related to the ratio-based cost matrix in multi-class classification as it treats the costs of minority samples equally across equally sized majority classes. Therefore, CORTEX fails to consider that minority samples might be more similar to one majority class and fails to make errors more acceptable for that class. To address this issue, as the main contribution of this paper, we propose an alternative cost matrix formulation that utilizes the distance between class centroids or medoids to reflect the similarities/differences among classes and have a more accurate representation of each class cluster.

Consistent with the evaluation framework of our previous study [5], the CORTEX method with different formulations of the cost matrix is compared to other tree-based and rule-based XAI methods, serving as a surrogate model for the underlying black-box model (neural network). Developing a tree-based/rule-based model as an explanation for a neural network model is accomplished on a relabeled target variable without using internal elements of the network. The experimental results obtained that CORTEX offers competitive performance while addressing key limitations of existing tree-based and rule-based methods, such as reduced interpretability due to deep trees, many rules, and long rule lengths on average.

The remainder of the paper is structured as follows: Section 2 reviews related work. Section 3 introduces the concept of an n-dimensional class-dependent cost matrix. In the first part of Section 4 are reported cost-sensitive tree models extracted by the CORTEX method with three different cost matrix formulations followed by a comprehensive comparative evaluation of the CORTEX method with other tree-based and rule-based XAI methods. Finally, Section 5 summarizes our key findings.

2. Related work

In numerous applications, complex neural network models are often the preferred choice due to the high-performance capacity of these models. Nevertheless, higher accuracy comes at the cost of these models' incomprehensible and non-understandable decision-making process. Tree-based models are considered self-interpretable, transparent, and comprehensible [3, 8]. Several approaches have been proposed to explain deep learning classification models, including using decision tree methods as surrogate models and extracting rule sets from the resulting tree [9]. Surrogate models can be created globally or locally [3], where global surrogate models aim to explain the model as a whole, and the local surrogate model explains a single instance.

Local Interpretable Model-Agnostic Explanations (LIME) [10] is a widely used local surrogate method. Another popular post-hoc method that can provide local and global explanations is the Shapley Additive Explanations (SHAP) method proposed in [11]. Both model-specific and model-agnostic versions of the SHAP have been proposed for tree-based models [12], including also cost-sensitive models [13, 14].

Tree-based algorithm C4.5-PANE [15] is an extension of a C4.5 decision tree algorithm [16], capable of extracting if-then rules from ensembles of neural networks, and its performance is compared to other rule-extractors in study [17]. Rule Extraction From Neural Network Ensemble (REFNE) was developed to extract symbolic rules from neural networks [18]. Another rule-based method that relies on a reverse engineering technique to extract rules from neural networks is Rule Extraction by Reverse Engineering (RxREN) [19]. Finally, the TREPAN [20] method generates a decision tree by querying the underlying network using a query and sampling approach.

3. Design and Methods

The cost-sensitive rule and tree extraction method CORTEX [5] is a cost-sensitive multi-class tree-building algorithm where misclassification costs are incorporated using a pre-defined class-dependent cost matrix. The learning phase consists of stratifying feature space into regions in a recursive manner (top-down greedy search). The CORTEX method classifies the sample into the least costly class, equivalent to classifying a sample with the highest cost-sensitive probability. In study [14], cost-sensitive probabilities are introduced into the cost-sensitive decision tree method for a two-class classification framework and later generalized in [5] for an arbitrary number of classes. By introducing cost-sensitive probabilities, it is possible to access information about confidence in the prediction by persevering the cost-dependence of labels. A detailed description of the CORTEX method is given in [5].

The misclassification costs are typically represented as elements of a cost matrix. The cost matrix can be class-dependent or sample-dependent, where the costs are associated with the classes or samples, respectively. The former assumption of constant costs across classes is more substantial and widespread through the application of most cost-sensitive learning algorithms [21, 22] since, in many real-life problems, the values in the matrix are unknown and are not given by experts. Throughout this paper, the term 'cost matrix' will refer to the class-dependent type.

The cost matrix is a function **C** of the actual and predicted classes, defined as $\mathbf{C} = [C_{ij}] \quad i, j = 1, \ldots, K$ where K represents number of classes, while i and j represent actual and predicted class, respectively. Accordingly, $C_{ij} = C(i, j)$ is the cost of predicting class i when the actual (true) class is j.

3.1. Ratio-based cost matrix

In the CORTEX method [5], the default (ratio-based) version of the cost matrix is defined by using class imbalance ratios among classes. If N_i is the number of samples in class i, the values of a ratio-based cost matrix are defined as $C_{ij} = \frac{N_i + N_j}{N_i}$ which reflects class imbalance ratio among the classes i and j. The cost matrix in CORTEX is intentionally defined to reflect the proportions of the samples in classes, since otherwise, with equal costs, CORTEX would not have the advantage over some other algorithm (assuming other differences between them are negligible) since minimizing cost would be equivalent to minimizing the error rate, leading to inappropriate, biased classifier towards the majority class.

One drawback of the ratio-based cost matrix can be noticed in the multi-class classification framework where one class is under-represented. Namely, suppose there is the same number of samples in majority classes (or nearly the same). In that case, the costs for misclassifying minority samples in either of the majority classes will be the same. However, the minority samples might be more similar to those in one majority class, and making such an error might be more acceptable than wrongly classifying minority samples in other, more dissimilar majority class(es). Consequently, to reflect the similarity/dissimilarity among classes, an alternative approach is proposed to use distance among their centroids or medoids.

3.2. Distance-based cost matrix

Two different formulations of the alternative cost matrix are proposed: centroid distance-based and medoid distance-based cost matrix. The centroid of a class is the point corresponding to the geometric mean of all samples in the class, while the medoid is the existent sample from the class that minimizes the average dissimilarity (in our study, Euclidean distance) to other samples in the class. Accordingly, the centroid c_i of a class i is obtained as the mean vector of all samples belonging to the class i. In contrast, the medoid m_i of a class i is the sample within the class i with the minimum average distance to all other samples in the class i. By calculating Euclidean distance among centroids/medoids, the symmetric cost matrix is obtained, where $C_{ij} = C_{ji} = d(c_i, c_j)$ or $C_{ij} = C_{ji} = d(m_i, m_j)$. Afterwards, the obtained matrix must be multiplied with weights to reflect that the minority class(es) has fewer samples (and, therefore, a higher cost). This is accomplished by scaling distances between centroids/medoids by the size of the corresponding class. Accordingly, the centroid and medoid distance-based cost matrix are defined as $C_{ij} = d(c_i, c_j) * \sqrt{N_j}$ and $C_{ij} = d(m_i, m_j) * \sqrt{N_j}$. In our implementation, the weights are proportional to the square root of the class size since we want to prevent the classifier from being too biased towards the minority class.

The intuition behind the distance-based cost matrix is that the further the two classes are, the higher the cost of making the wrong classification for samples in the class with fewer samples should be. The distance is measured between centroids or medoids of classes where the central tendency of a cluster with outliers or skewed distribution can be more accurately reflected with the medoids as a more robust measure [23], especially in the presence of class sub-concepts commonly observed in class imbalance frameworks [24]. Depending on the target distribution, a centroid distance-based or medoid distance-based cost matrix might be more suitable than a ratio-based cost matrix.

4. Results

In the experimental part of the study, we compared the performance of the CORTEX method with different formulations of cost matrix with other tree-based and rule-based XAI methods. The CORTEX and other methods are used as a post-hoc XAI method by creating a surrogate tree model for a simple neural network model and automatically extracting a set of rules from the obtained tree. Eight datasets with varying class sizes ranging from 2 to 29 are considered, as in other studies [25, 17].

The first step of the experimental setup is training a simple neural network model (feed-forward with two fully connected hidden layers) on 70% of data with early-stopping to prevent overfitting. For all network hyperparameters, optimal values are obtained from Table 2 reported by [17]. Afterwards, the post-hoc surrogate models are created using 30% of test data, and predictions given by the neural network. The cost-sensitive tree model is trained using the CORTEX algorithm. Due to space limitations, tree topologies for CORTEX with three cost matrix formulations are given in Table 4 for several datasets. The CORTEX method with centroid distance-based cost-matrix (CORTEX-c) gives a smaller tree for all three datasets. For other datasets, neither matrix formulation provides a constantly smaller tree. The CORTEX method with a medoid distance-based cost matrix (CORTEX-m) performs as well as or worse than CORTEX-c or CORTEX with a ratio-based cost matrix. Notably, the CORTEX method generates tree models with different topologies depending on the formulation of the cost matrix.

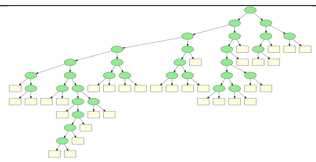
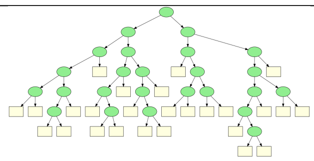
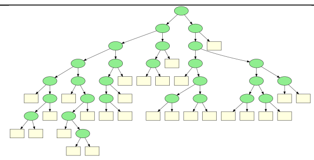
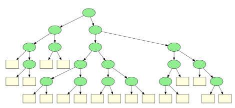
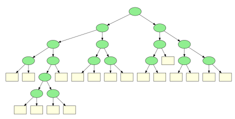
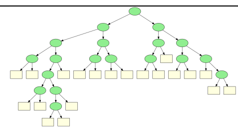
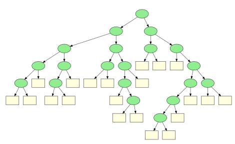
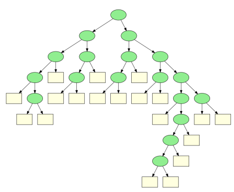
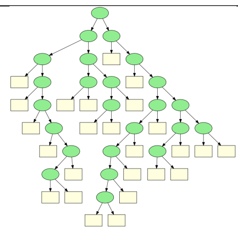
Dataset	CORTEX	CORTEX-c	CORTEX-m
abalone			
contraceptive			
page_blocks			

Table 1
Tree topology of CORTEX method with different cost matrix formulation across datasets.

The transformation from a tree model into a set of rules is essential to facilitate the comparison of CORTEX with other tree-based and rule-based XAI methods. For comprehensive comparative analysis are considered five rule exaction methods, where four rule-extractors, C4.5-PANE, REFNE, RxREN, and TREPAN, are extensively studied in the literature [17] in similar framework and therefore considered as a strong baseline model in our study. Furthermore, considering the CORTEX method is a tree-based algorithm, the selected subset of benchmarking methods is extended with a traditional decision tree classifier (DT) to provide a more comprehensive evaluation. In our work, we have used the scikit-learn implementation of DT with weights that are automatically adjusted to be inversely proportional to class frequencies in the weighted impurity gain measure in order to effectively take into account the class imbalance ratios of the datasets[1].

Six metrics were selected to assess the degree of explainability of the rule sets, including completeness, correctness, fidelity, robustness, number of rules, and average rule length. The formal definitions and detailed description of these measures can be found in [17, 5].

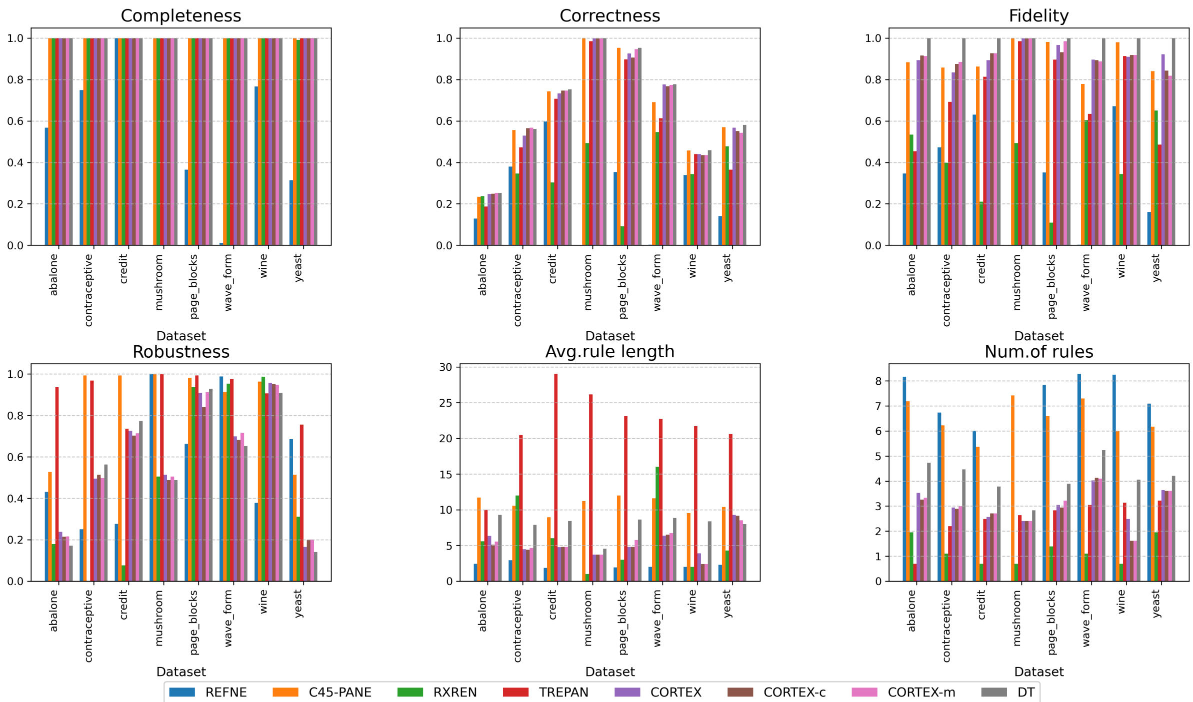


Figure 1: Comparative performance evaluation of tree-based and rule-based XAI extraction methods against CORTEX method with different cost matrix formulation across eight datasets.

In Figure 1 are reported evaluation results, where only the number of rules is converted into the logarithmic scale to enhance the visibility of the results. Notably, all methods except REFNE produce a set of rules covering all samples across all datasets, reaching 100% completeness. Regarding correctness, the CORTEX method, with different formulations of cost matrices, performs equally well or better than other methods. The reported results for the fidelity measure show that the DT model outperforms other methods across all datasets. However, for most datasets, the CORTEX, CORTEX-c, and CORTEX-m can be ranked second-best, right after the DT method. Results also reveal that CORTEX models are less robust than other tree-based extractors, such as C4.5-PANE and TREPAN. At the same time, CORTEX is competitive with other rule extractors or better than them in terms of robustness, depending on the dataset. The surpassed robustness of C4.5-PANE over other methods could be due to the augmentation of training data with synthetic data in its training process. On the other hand, the good robustness of the TREPAN can be explained by a user-specified minimum number of samples available at a node before choosing a splitting feature for that node. Nonetheless, the robustness of TREPAN and C4.5-PANE comes with a trade-off regarding average rule length. As noted, both TREPAN and C4.5-PANE produce the highest average rule length. The CORTEX, CORTEX-c, and CORTEX-m produce rule sets significantly shorter than rules generated by TREPAN, C4.5-PANE, and DT, but still not shorter than those extracted from REFNE. Despite generating the shortest rules, REFNE generates sets with the highest number

[1]Other rule extractors are obtained from https://github.com/giuliavilone/rule_extractor

of rules, followed by C4.5-PANE. While CORTEX may not have the lowest average rule length nor the smallest set of rules, it clearly shows the ability to balance different metrics, establishing effective performance.

A non-parametric Friedman test [26] is used to assess whether a specific tree-based or rule-based XAI method performs significantly differently than others according to the six analyzed metrics across eight datasets. The results of the Friedman test are evaluated using a significance level of 0.05. Results show that six p-values are lower than the significance level of 0.05, meaning there is evidence supporting the null hypothesis for 6 out of 8 datasets, that some method performs consistently better (or worse).

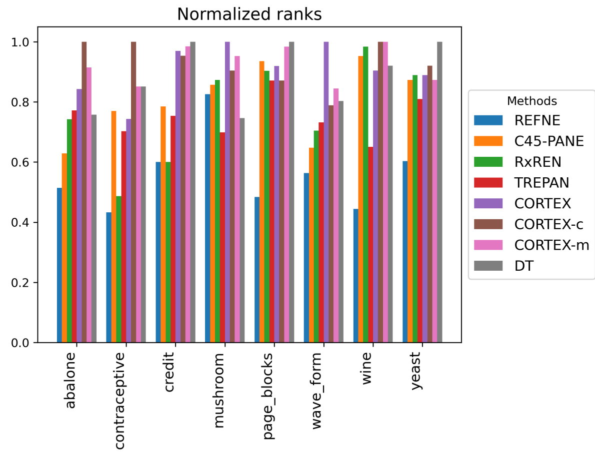


Figure 2: Normalized ranks of eight tree-based and rule-based XAI extraction methods across eight datasets.

The subsequent phase of the experimental procedure involves ranking the selected XAI methods according to the six metrics. Initially, rankings were determined for each metric. These rankings were then aggregated across all datasets, and the sum of ranks for each metric was normalized, yielding the final normalized rankings. The results reported in Figure 2 indicate that CORTEX-c is ranked as the best method for *abalone* and *contraceptive* datasets. Baseline CORTEX method with ratio-based cost matrix achieves the highest rank for *mushroom* and $wave\_form$ datasets. The CORTEX-m method is top-ranked only for *wine* dataset altogether with CORTEX-c. Therefore, for 5 out of 8 datasets, the CORTEX method with different cost matrix formulations is ranked as the best method considering all six measures used for performance assessment. For the other 3 datasets, the CORTEX method is ranked as the second-best model. However, choosing the best cost matrix definition isn't straightforward; it largely depends on the specific dataset.

The CORTEX method with different cost matrix formulations demonstrates competitive performance compared to other tree-based models, showcasing its effectiveness in handling black-box models on diverse datasets. Furthermore, it surpasses the capabilities of some inherent rule-extraction techniques, delivering superior results in terms of analyzed quantitative measures of the degree of explainability. Specifically, extracting shorter rule sets with shorter rule length, on average, suggests the advantages of using the CORTEX method over alternative methods. However, this advantage comes with the trade-off of having a less accurate and robust model, although it effectively balances this trade-off. Overall, the results underscore the potential of CORTEX as a powerful XAI tool for scenarios requiring clear, human-understandable rules while maintaining good predictive performance.

5. Concluding remarks

In this paper, we have explored alternative cost matrix formulations in cost-sensitive rule and tree extraction method (CORTEX) using centroid and medoid distance-based cost matrix. By using distance among centroids or medoids of classes, the distance-based costs will differ for wrongly classifying minority samples into majority classes. Instead of centroids, medoids could be used as more representative objects of each class cluster, especially in the presence of outliers and class sub-concepts commonly observed in class imbalance frameworks. The CORTEX method is compared against other tree-based and rule-based XAI methods as a surrogate model for the underlying black-box model (neural network). Our study demonstrates that CORTEX offers competitive performance while addressing key limitations of existing tree-based and rule-based methods, such as reduced interpretability due to deep trees, many rules, and long rule lengths on average. Depending on the cost matrix used in CORTEX, smaller rule sets with shorter rules can be produced at the cost of slightly reduced accuracy and robustness. To enhance the robustness of the CORTEX method in future research, the training set could be augmented with synthetic data. Overall, CORTEX effectively balances the interpretability-accuracy trade-off since it can generate understandable tree models without significantly compromising other performance measures. Therefore, CORTEX is a valuable XAI tool for generating understandable rules while retaining good predictive performance as a surrogate model for complex models in class imbalance frameworks.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] L. Longo, M. Brcic, F. Cabitza, J. Choi, R. Confalonieri, J. D. Ser, R. Guidotti, Y. Hayashi, F. Herrera, A. Holzinger, R. Jiang, H. Khosravi, F. Lecue, G. Malgieri, A. Páez, W. Samek, J. Schneider, T. Speith, S. Stumpf, Explainable artificial intelligence (xai) 2.0: A manifesto of open challenges and interdisciplinary research directions, Information Fusion 106 (2024) 102301. doi:https://doi.org/10.1016/j.inffus.2024.102301.

[2] G. Vilone, L. Longo, Development of a human-centred psychometric test for the evaluation of explanations produced by xai methods, in: L. Longo (Ed.), Explainable Artificial Intelligence, Springer Nature Switzerland, Cham, 2023, pp. 205–232.

[3] S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, F. Herrera, Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence, Information Fusion 99 (2023) 101805. doi:https://doi.org/10.1016/j.inffus.2023.101805.

[4] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models 51 (2018). URL: https://doi.org/10.1145/3236009. doi:10.1145/3236009.

[5] M. Kopanja, M. Savić, L. Longo, Cortex: A cost-sensitive rule and tree extraction method, 2025. URL: https://arxiv.org/abs/2502.03200. arXiv:2502.03200.

[6] P. D. Turney, Types of cost in inductive concept learning, 2002. URL: https://arxiv.org/abs/cs/0212034. arXiv:cs/0212034.

[7] B. A. Correa, Example-dependent cost-sensitive decision trees, Expert Systems with Applications 42 (2015) 6609–6619. doi:10.1016/j.eswa.2015.04.042.

[8] G. Vilone, L. Rizzo, L. Longo, A comparative analysis of rule-based, model-agnostic methods for explainable artificial intelligence, in: L. Longo, L. Rizzo, E. Hunter, A. Pakrashi (Eds.), Proceedings of The 28th Irish Conference on Artificial Intelligence and Cognitive Science, Dublin, Republic of Ireland, December 7-8, 2020, volume 2771 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020, pp. 85–96. URL: http://ceur-ws.org/Vol-2771/AICS2020_paper_33.pdf.

[9] E. Mekonnen, P. Dondio, L. Longo, Explaining deep learning time series classification models using a decision tree-based post-hoc xai method, volume 3554, CEUR-WS, 2023. doi:https://doi.org/10.21427/9YKT-WZ47, publisher Copyright: © 2023 CEUR-WS. All rights reserved.; Joint 1st World Conference on eXplainable Artificial Intelligence: Late-Breaking Work, Demos and Doctoral Consortium, xAI-2023: LB-D-DC ; Conference date: 26-07-2023 Through 28-07-2023.

[10] M. T. Ribeiro, S. Singh, C. Guestrin, "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, KDD '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 1135–1144. doi:10.1145/2939672.2939778.

[11] S. Lundberg, S. Lee, A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems 30, Curran Associates, Inc., 2017, pp. 4765–4774.

[12] S. Lundberg, G. Erion, H. Chen, A. DeGrave, J. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S. Lee, From local explanations to global understanding with explainable ai for trees, Nature Machine Intelligence 2 (2020) 2522–5839.

[13] M. Kopanja, S. Brdar, S. Hacko, Uncovering decision-making process of cost-sensitive tree-based classifiers using the adaptation of treeshap., in: "Joint Proceedings of the xAI-2023 Late-breaking Work, Demos and Doctoral Consortium co-located with the 1st World Conference on eXplainable Artificial Intelligence (xAI-2023)", 2023, pp. 95–100.

[14] M. Kopanja, S. Hačko, S. Brdar, M. Savić, Cost-sensitive tree shap for explaining cost-sensitive tree-based models, Computational Intelligence 40 (2024) e12651. doi:https://doi.org/10.1111/coin.12651.

[15] Z.-H. Zhou, Y. Jiang, Medical diagnosis with c4.5 rule preceded by artificial neural network ensemble, IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society 7 (2003) 37–42. doi:10.1109/TITB.2003.808498.

[16] J. R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.

[17] G. Vilone, L. Longo, A quantitative evaluation of global, rule-based explanations of post-hoc, model agnostic methods, Frontiers in Artificial Intelligence 4 (2021). doi:10.3389/frai.2021.717899.

[18] Z.-H. Zhou, Y. Jiang, S.-F. Chen, Extracting symbolic rules from trained neural network ensembles, AI Commun. 16 (2003) 3–15.

[19] M. Augasta, T. Kathirvalavakumar, Reverse engineering the neural networks for rule extraction in classification problems, Neural Processing Letters (2011). doi:10.1007/s11063-011-9207-8.

[20] M. W. Craven, J. W. Shavlik, Using sampling and queries to extract rules from trained neural networks, in: W. W. Cohen, H. Hirsh (Eds.), Machine Learning Proceedings 1994, Morgan Kaufmann, San Francisco (CA), 1994, pp. 37–45. URL: https://www.sciencedirect.com/science/article/pii/B9781558603356500131. doi:https://doi.org/10.1016/B978-1-55860-335-6.50013-1.

[21] G. K. Haomin Wang, Y. Peng, Multi-class misclassification cost matrix for credit ratings in peer-to-peer lending, Journal of the Operational Research Society 72 (2021) 923–934. doi:10.1080/01605682.2019.1705193.

[22] B. Krawczyk, M. Woźniak, G. Schaefer, Cost-sensitive decision tree ensembles for effective imbalanced classification, Applied Soft Computing 14 (2014) 554–562. URL: https://www.sciencedirect.com/science/article/pii/S1568494613002895. doi:https://doi.org/10.1016/j.asoc.2013.08.014.

[23] H.-S. Park, C.-H. Jun, A simple and fast algorithm for k-medoids clustering, Expert Systems with Applications 36 (2009) 3336–3341. URL: https://www.sciencedirect.com/science/article/pii/S095741740800081X. doi:https://doi.org/10.1016/j.eswa.2008.01.039.

[24] Y. Sun, A. Wong, M. S. Kamel, Classification of imbalanced data: a review, International Journal of Pattern Recognition and Artificial Intelligence 23 (2011). doi:10.1142/S0218001409007326.

[25] G. Vilone, L. Rizzo, L. Longo, A comparative analysis of rule-based, model-agnostic methods for explainable artificial intelligence, Proceedings for the 28th AIAI Irish Conference on Artificial Intelligence and Cognitive Science, Dublin, Ireland 2771 (2020). doi:10.21427/z4x3-3f86.

[26] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, Journal of the American Statistical Association 32 (1937) 675–701. doi:10.1080/01621459.1937.10503522.