Appendix - A Supplementary Information For Evaluation Results

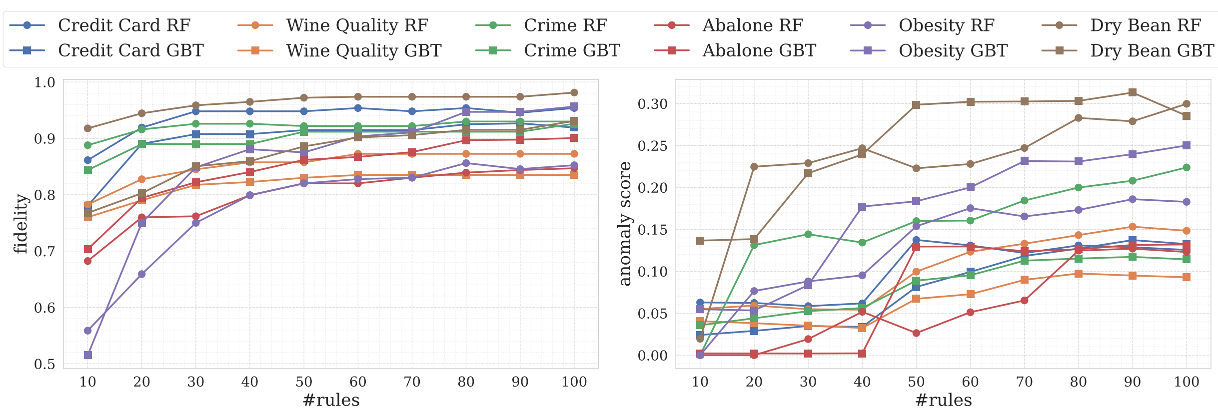
1 Sensitivity Analysis of Parameter M

We conducted a sensitivity analysis on various datasets and models to examine how varying the number of rules (M) affects the fidelity and anomaly scores of the extracted rules. Specifically, we increased M from 10 to 100 with a step of 10, performed the model reduction, and evaluated the fidelity and anomaly scores of the extracted rules.

As shown in Fig. S1(a), the fidelity curves initially improve as M increases, and then stabilizes. This trend is expected because a larger M allows the extracted rules to better represent the entire rule set, thus improving fidelity. However, the rate of stabilization varies across different datasets and models. It is influenced by multiple factors, such as the dataset characteristics, the total number of rules, and the similarity between rules. Nonetheless, by $M = 80$, the majority have already stabilized.

Fig. S1(b) illustrates the effect of M on the anomaly score. Similar to fidelity, the anomaly scores first increase and then stabilize as M increases. This trend is also expected because a larger M allows the selection of more anomaly rules without compromising much fidelity.

These results support our choice of $M = 80$ as a unified setting. This value achieves a balance between the quality of extracted rules and screen constraints, as both fidelity and anomaly scores tend to stabilize beyond this point across most datasets.



(a) (b)

Fig. S1. The relationship between the number of rules and (a) fidelity scores, (b) anomaly scores.

2 Examples of Different Values of Parameter τ in Section V-C

We present the matrix reordering results using different values of τ (0.05, 0.10, and 0.15), which controls whether two numerical conditions are similar enough to be placed adjacently. As shown in Fig. S2, while varying τ has a minor overall impact on the reordering results, it still influences the results in certain cases. For example, with $\tau = 0.05$, some rules with similar conditions are not placed adjacently (Fig. S2A) due to the strict similarity threshold. With $\tau = 0.1$, these rules are placed adjacently (Fig. S2B). With $\tau = 0.15$, the higher threshold causes rules with more distinct conditions to be considered similar and placed adjacently (Fig. S2C). After examining the reordering results across different datasets using different values of τ, we choose $\tau = 0.1$, which generates satisfactory results in most cases.

3 Explanation of Financial Attributes in Section VI-B-2

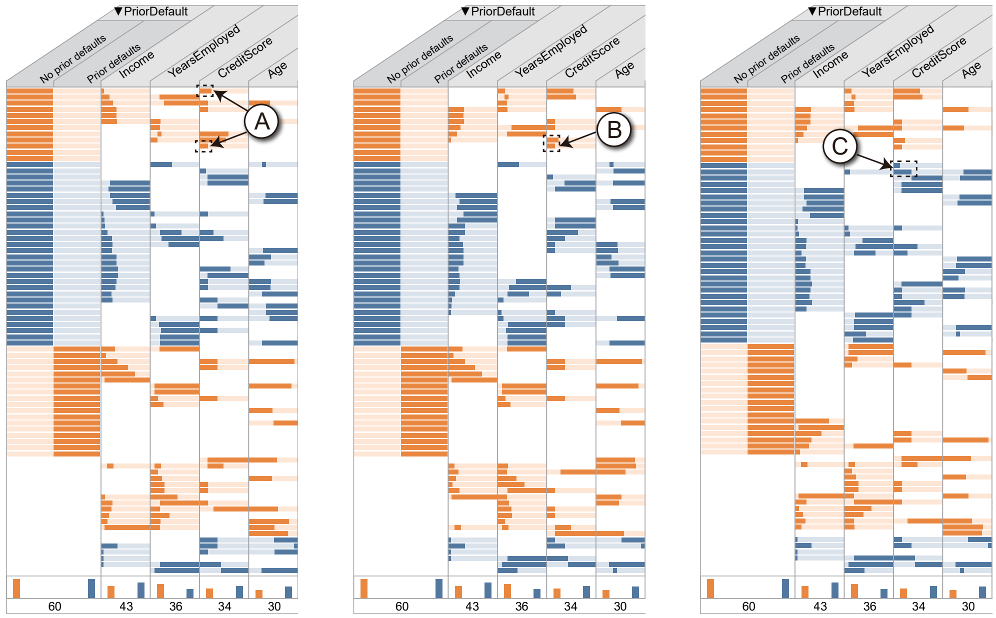
In Section VI-B-2 of the main paper, we present a case study involving various attributes related to financial instruments and market dynamics. Due to space constraints, we were unable to provide a detailed explanation of these attributes within the main body of the paper. To ensure clarity and comprehension for readers who may not be familiar with these specific financial concepts, we have included this supplementary material to offer concise explanations of the attributes mentioned in Section VI-B-2. The following list provides a brief description of each attribute:

STD10: The standard deviation of the stock's price over the past 10 trading days, divided by the latest close price. It measures the stock's short-term price volatility. A higher *STD10* indicates more significant short-term price fluctuations.

STD60: The standard deviation of the stock's price over the past 60 trading days, divided by the latest close price. It measures the stock's medium-term price volatility. Compared to *STD10*, *STD60* captures price fluctuations over a more extended period.

RESI20: The difference between the actual stock price and the price estimated by a linear regression model over the past 20 trading days, divided by the latest close price. It measures how much the stock price has deviated from its expected short-term linear trend.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx



(a). τ = 0.05 (b). τ = 0.1 (c). τ = 0.15

Fig. S2. Results for three different values of parameter τ.

RESI60: The difference between the actual stock price and the price estimated by a linear regression model over the past 60 trading days, divided by the latest close price. Similar to *RESI20*, *RESI60* measures how much the stock price has deviated from its expected medium-term linear trend.

KLEN: A technical analysis indicator that measures volatility by creating a channel around the stock's price. Higher *KLEN* values indicate increased volatility, while lower *KLEN* values suggest reduced volatility.

BETA5: The slope of the linear regression line of the stock's close price over the past 5 trading days, divided by the latest close price. It measures the stock's sensitivity to market movements and the average daily rate of price change in the short term. A higher *BETA5* value suggests stronger short-term price momentum, either in an upward or downward direction.

NATR: The normalized average true range, a volatility measure that takes into account gaps and limit moves in the stock's price. A lower *NATR* indicates lower volatility and potentially more stable price action, while a higher *NATR* suggests higher volatility and potentially more dramatic price movements.

VMA20: The average trading volume of the stock over the past 20 trading days, divided by the latest trading volume.

revenueGrowth: The percentage increase in a company's revenue compared to the previous year. It indicates the annual rate at which the company's sales are growing.

4 Data Augmentation Process in Section VI-B-1

To address potential spurious associations identified during our analysis, we employed a targeted data augmentation technique using the Synthetic Data Vault (SDV) toolkit [3]. This targeted augmentation approach helped us expand our training dataset with additional diverse samples. Our process involved the following three steps:

Model Training. We trained the CTGANSynthesizer model from the SDV toolkit on our original dataset to learn its statistical properties and relationships between variables.

Rule-based Synthetic Data Generation. For each of our 16 anomaly rules, we used the trained model to generate synthetic samples. We applied rejection sampling [1] to ensure each generated sample complied with its corresponding anomaly rule. We generated 15 new samples per rule, creating a total of 240 new training data samples. This augmentation size is comparable to our original training set size (690*0.75).

Expert Validation and Refinement. E_1 (a bank account manager) assessed whether the generated data, especially customer information, aligned with real-world scenarios. Most samples were found to align well. Samples that did not align with real-world situations were discarded and regenerated until they met both rule conditions and real-world plausibility. E_1 then reviewed and corrected the labels of all generated data to ensure they conformed to the bank's actual approval criteria.

5 Results for Drybean and Obesity Datasets in Section VI-A

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, 2006. doi: 10.5555/1162264

[2] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. LightGBM: A highly efficient gradient boosting decision tree. In *Proceedings of the Advances in Neural Information Processing Systems*, pp. 3146–3154, 2017. doi: 10.5555/3294996.3295074

[3] N. Patki, R. Wedge, and K. Veeramachaneni. The synthetic data vault. In *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics*, pp. 399–410, 2016. doi: 10.1109/DSAA.2016.49

[4] U. M. L. Repository. Estimation of obesity levels based on eating habits and physical condition . UCI Machine Learning Repository, 2019. doi: 10.24432/C5H31Z

[5] U. M. L. Repository. Dry Bean Dataset. UCI Machine Learning Repository, 2020. doi: 10.24432/C50S4B

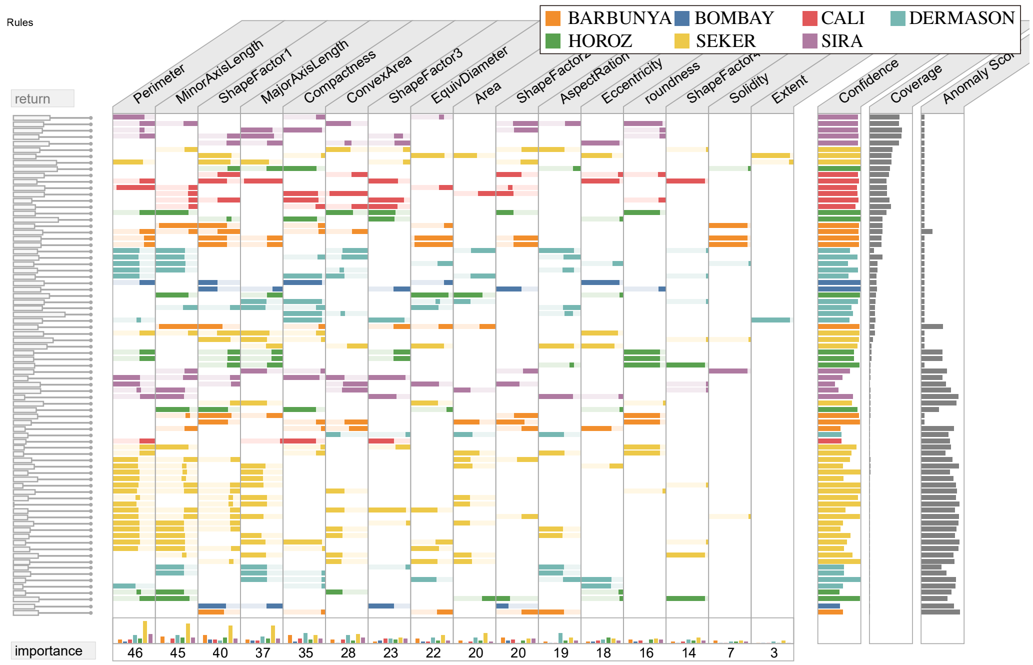


Fig. S3. Matrix view in RuleExplorer showing rules from the random forest model trained on the Dry Bean [5] dataset, as mentioned in TABLE I from the main paper.

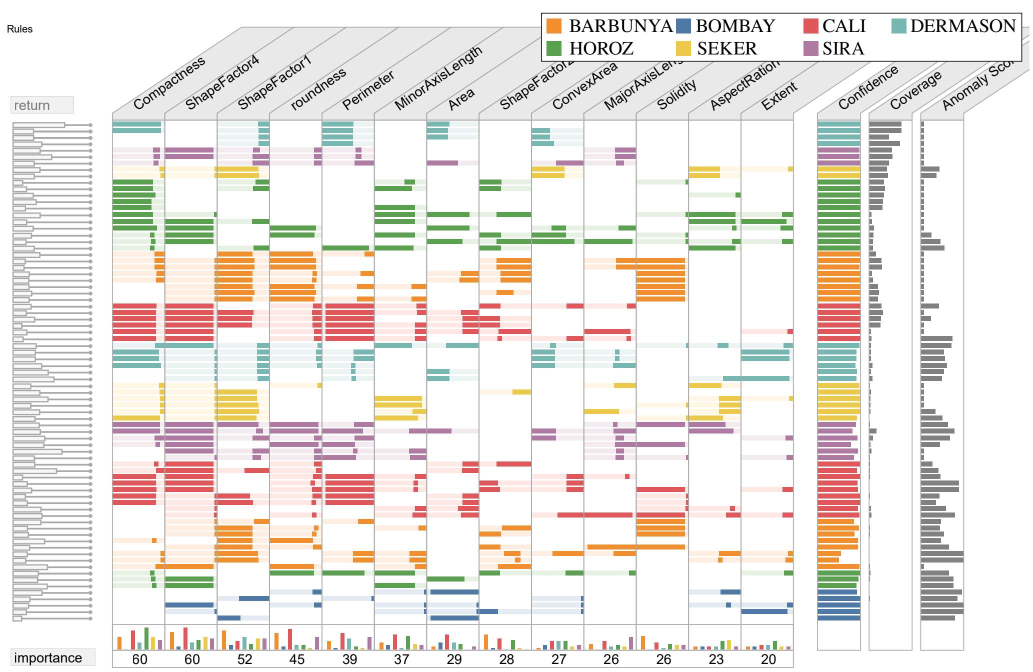


Fig. S4. Matrix view in RuleExplorer showing rules from the LightGBM model [2] trained on the Dry Bean [5] dataset, as mentioned in TABLE I from the main paper.

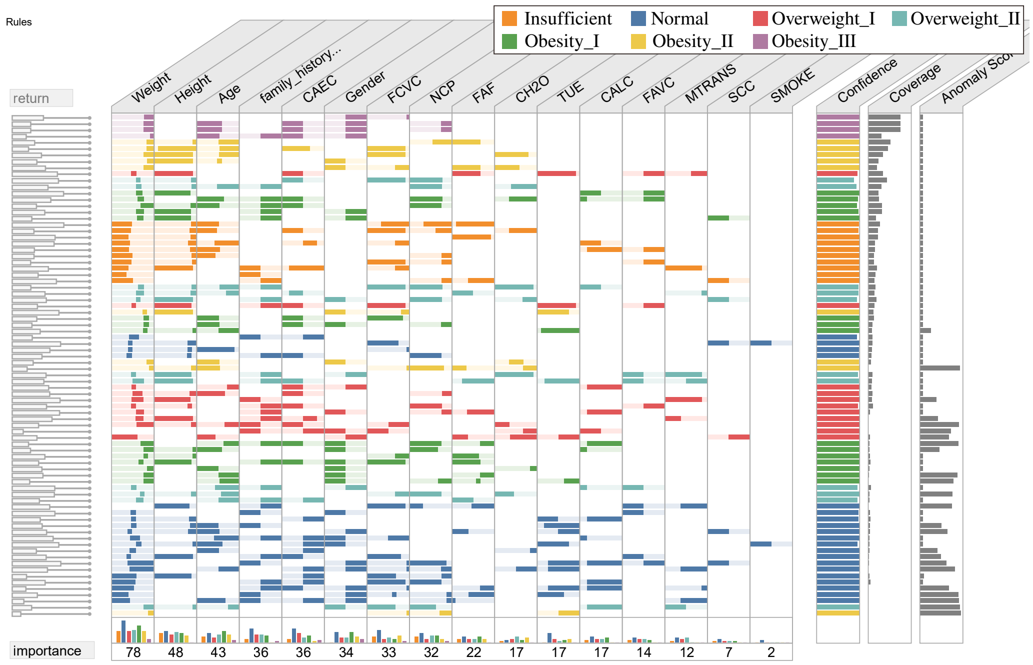


Fig. S5. Matrix view in RuleExplorer showing rules from the random forest model trained on the Obesity [4] dataset, as mentioned in TABLE I from the main paper.

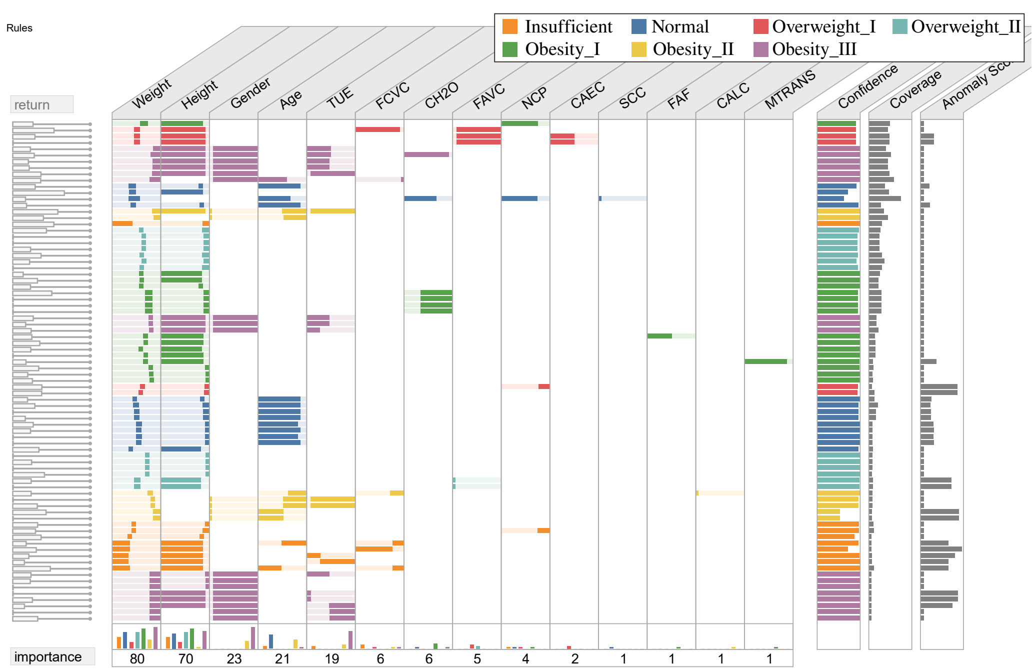


Fig. S6. Matrix view in RuleExplorer showing rules from the LightGBM model [2] trained on the Obesity [4] dataset, as mentioned in TABLE I from the main paper.