Apriori Meets LLMs: Interpretable Rule Mining from Continuous Healthcare Data

Asma Abboura
University of Hassiba Benbouali
Chlef, Algeria
a.abboura@univ-chlef.dz

Imane Hocine
Snt, University of Luxembourg
Esch-sur-Alzette, Luxembourg
imane.hocine@uni.lu

Yacine Hakimi
École supérieure d'informatique
Algiers, Algeria
y_hakimi@esi.dz

Soror Sahri
Université de Paris
Paris, France
soror.sahri@u-paris.fr

Grégoire Danoy
Snt and FSTM-DCS, University of Luxembourg
Esch-sur-Alzette, Luxembourg
gregoire.danoy@uni.lu

***Abstract*—Mining meaningful patterns from numerical healthcare data is challenging, as continuous lab values are difficult to analyze directly and traditional association rule mining often generates arbitrary thresholds. We introduce Threshold-Aware Association Rules (TAAR), a framework that converts continuous lab values into semantic intervals and extracts interpretable rules using an enhanced Apriori algorithm. Large Language Models (LLMs) are employed to refine support and confidence thresholds, filter implausible rules, and produce natural-language explanations. Applied to blood test data, TAAR improves clinical usability, guides actionable follow-up recommendations, and supports informed decision-making.**

***Index Terms*—Association Rule Mining, Large Language Models (LLMs), Blood Test Analysis, Threshold-Aware Association Rules (TAAR).**

I. Introduction

Timely diagnosis and follow-up remain persistent challenges in healthcare. Patients often face delays in appointments and test results, while clinicians spend considerable time interpreting numerical lab values. Automating reasoning over continuous data can accelerate clinical workflows and improve patient outcomes [9], [13], [14].

For instance, low hemoglobin levels in anemia may trigger follow-up tests (ferritin, B12, folate). Traditional association rule mining (ARM) often relies on arbitrary discretization and fails to capture meaningful thresholds [10], [17]. Although fuzzy or quantitative ARM variants exist [2], [3], [10], [12], [17], their usability and clinical relevance remain limited, often due to arbitrary thresholds or rules that are difficult to directly translate into actionable follow-up tests.

We propose *TAAR*, a framework that converts continuous lab values into semantic intervals, mines actionable rules with an enhanced Apriori algorithm, and uses LLMs to refine thresholds, filter implausible rules, and generate explanations. In this context, "interpretable" means that the rules are easy for clinicians to use for follow-up test recommendations and are presented in natural-language explanations, making them both actionable and understandable in clinical practice.

II. Related Work

Association rule mining (ARM), particularly the Apriori algorithm, has been widely used in healthcare for uncovering disease patterns, including ophthalmologic disorders [5] and blood test interpretation [13], [14]. Most existing approaches rely on categorical or discretized data, which limits their ability to capture nuances in continuous laboratory measurements.

To address this, prior research explored discretization, fuzzy association rules, and quantitative rules with inequality constraints [2], [3], [10], [12], [17], as well as threshold optimization via evolutionary algorithms [1]. While these approaches improve handling of continuous values, thresholds are often arbitrary and rules may not directly guide actionable follow-up.

Hybrid approaches have incorporated semantic constraints and domain expertise, for example, interpretable rule-based models for chronic disease risk prediction [9]. More recently, ARM has been combined with Large Language Models (LLMs). Li et al. [6] applied LLMs for interpretable factor and rule mining in finance, while Naduvilakandy et al. [8] used retrieval-augmented prompting for causal rule extraction.

Building on these advances, TAAR uses clinically established thresholds and LLMs to mine actionable, interpretable rules from continuous data.

III. Threshold-Aware Association Rules (TAAR)

Blood test data often contain continuous measurements whose interpretation depends on clinical thresholds. Classical association rule mining cannot represent such threshold-based reasoning. TAAR extends this by introducing inequality constraints and recommendation-type rules that link abnormal findings to follow-up tests (see Fig. 1).

Let $D = \{T_1, \ldots, T_n\}$ be a dataset of *transactions*, where each transaction T_i represents a patient record containing laboratory measurements $\mathcal{A} = \{A_1, \ldots, A_m\}$. A TAAR rule r is defined as:

$$r : \text{LHS} \Rightarrow \text{RHS},$$

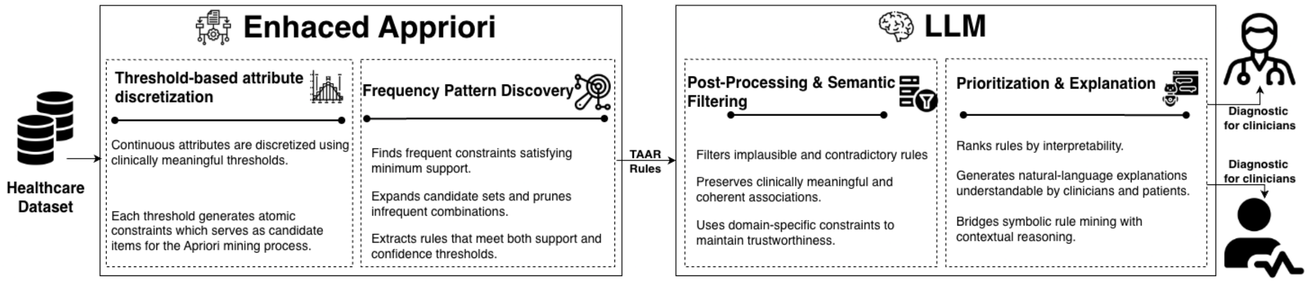

Fig. 1. TAAR mining workflow from raw data to interpretable rules, including LLM-based threshold refinement.

where the LHS is a set of attribute–value constraints $A_i\ \theta\ \tau$, with $\theta \in \{\leq, <, >, \geq\}$ and τ a clinically meaningful threshold. The RHS either specifies another constraint or a recommendation for follow-up tests. When a patient's results satisfy the LHS, TAAR recommends the RHS tests.

Example: $\{\text{TSH} > 4.5\} \Rightarrow \{\text{Recommend T3, T4 tests}\}$.

Support and confidence are computed as in classical ARM, ensuring comparability with standard methods:

$$\text{support}(r) = \frac{|\text{LHS} \cup \text{RHS}|}{|D|}, \quad \text{conf}(r) = \frac{\text{support}(\text{LHS} \cup \text{RHS})}{\text{support}(\text{LHS})}.$$

In healthcare, thresholds are derived from standard clinical reference intervals defined by medical guidelines and laboratory norms, ensuring semantic and diagnostic consistency in mined rules.

IV. LLM-Enhanced TAAR Mining

The system extracts TAARs from clinical datasets using the TAARs-Apriori algorithm. Continuous test values are first discretized into semantic intervals: below normal (-1), normal (0), and above normal $(+1)$, replacing the traditional binary encoding. The Apriori algorithm then mines frequent patterns under inequality constraints, generating candidate rules that correspond to clinically meaningful thresholds.

LLMs further refine these rules by filtering implausible or redundant patterns, adjusting support and confidence thresholds, and generating natural-language explanations. By linking thresholds to recommended follow-up tests, TAAR translates complex lab data into actionable clinical guidance.

V. Real-Time Blood Test Recommendation

For a new patient, initial blood test values are first mapped to semantic intervals $(-1,\ 0,\ +1)$. The system then scans the mined TAARs to identify which rules are activated, that is, those whose left-hand side (LHS) conditions match the patient's current results. The corresponding right-hand sides (RHS) of these activated rules specify recommended follow-up tests or clinical actions. LLMs refine this process by prioritizing recommendations according to clinical relevance and generating natural-language explanations. This combination of symbolic rule activation and LLM-based reasoning ensures that the resulting recommendations are both actionable and interpretable for clinical use.

VI. Experimental Results Summary

TAAR was evaluated on 800 patient records from the *NHANES 2011–2012* dataset[1] . The integrated LLM analyzed the mined rules and suggested optimal support and confidence thresholds based on rule coherence and clinical plausibility. The performance of the hybrid *TAAR + LLM* framework was compared to an LLM-only baseline (*GPT-4o-mini*, OpenAI, 2025), which directly generated recommendations without rule mining. True positives (TP), false positives (FP), and false negatives (FN) were computed by comparing system recommendations to actual follow-up tests present in the dataset.

Strategy	TP	FP	FN
LLM-only	4.20	7.50	5.80
TAAR + LLM (support 0.02, confidence 0.70)	4.35	7.50	5.50
TAAR + LLM (support 0.08, confidence 0.70)	4.45	7.50	5.50
TAAR + LLM (support 0.20, confidence 0.70)	4.54	7.45	5.60

TABLE I
Follow-up test recommendation performance of LLM-only vs. TAAR + LLM.

The TAAR + LLM system achieved higher true positives and fewer false negatives compared to the LLM-only baseline, improving patient safety by minimizing missed follow-up recommendations. False positives remained stable for most thresholds and slightly decreased at higher support (0.20), showing that stricter rule filtering improves precision. By limiting LLM use to rule refinement and explanation, TAAR also reduces computational cost for clinical decision support.

VII. Conclusion and Future Directions

We introduced TAAR, an Apriori-based framework augmented with LLMs for generating clinically meaningful rules from numerical healthcare data. By combining semantic intervals, inequality-based rule mining, and LLM-driven threshold optimization and explanations, TAAR enhances rule clarity and usability. Future work will investigate context-aware thresholding that adapts to patient-specific factors (e.g., pregnancy), integrate LLMs more dynamically to improve robustness and scalability, and extend TAAR to other domains such as finance and manufacturing with tailored threshold criteria and validation. Beyond healthcare, we also plan to apply TAAR to satellite data analysis, where adaptive thresholding can enhance environmental and geospatial monitoring.

[1]https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/

Acknowledgment

This work is partially funded by the Luxembourg National Research Fund (FNR), CORE International programme under the SERENITY Project (ref. C22/IS/17395419/SERENITY) and the Polish National Centre for Research and Development (NCBR) (ref. POLLUX-XI/15/Serenity/2023).

References

[1] J. Azevedo, M. Silva, and C. Soares, "Hybrid evolutionary algorithms for quantitative association rule mining," *Applied Soft Computing*, 2024.

[2] D. Dubois and H. Prade, "Fuzzy association rules in data mining," *International Journal of Approximate Reasoning*, vol. 50, no. 7, pp. 979–997, 2010.

[3] I. Fišter, M. Laznik, and D. Fister, "NARMviz: Quantitative association rule mining and visualization," *Knowledge-Based Systems*, 2024.

[4] A. Gerussi, F. Barbieri, and J. Camacho-Collados, "Large language models for symbolic reasoning: A survey," *arXiv preprint arXiv:2210.12345*, 2022. [Online]. Available: https://arxiv.org/abs/2210.12345

[5] S. Gulzar, M. A. Khan, and M. Hussain, "An efficient Apriori-based framework for mining disease patterns in ophthalmology," *Journal of Healthcare Engineering*, 2023.

[6] Y. Li, H. Chen, and Z. Zhou, "LLM-based interpretable factor and rule mining for financial data," in *Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 2024, pp. 3891–3902.

[7] X. Liang, C. Xue, and M. Huang, "Improved Apriori algorithm for mining association rules of many diseases," in *International Symposium on Intelligence Computation and Applications*, Springer, 2010, pp. 272–279.

[8] T. M. Naduvilakandy, H. Jang, and M. Al Hasan, "Retrieval augmented generation based large language models for causality mining," *arXiv preprint arXiv:2505.23944*, 2025. [Online]. Available: https://arxiv.org/abs/2505.23944

[9] M. Piran, S. Rahimi, and H. Karami, "Interpretable rule-based prediction of diabetic foot ulcer risk using laboratory data," *Computers in Biology and Medicine*, 2024.

[10] S. Rajasekar and K. Venkatesan, "Application of discretization methods in mining medical data," *International Journal of Computer Science*, 2009.

[11] M. Rashed and J. Streit, "Exploring discretization techniques for medical data mining," in *Proceedings of IEEE EMBS*, 2003.

[12] S. Rai and R. Kumar, "Association rules for COVID-19 and chronic disease risk prediction," *Computers in Biology and Medicine*, 2023.

[13] A. Santos, J. Carvalho, and P. Ferreira, "Artificial intelligence for blood test interpretation: A survey of association rule-based methods," *Artificial Intelligence in Medicine*, 2024.

[14] S. Shaik and R. Patel, "Remote monitoring of chronic patients using association rule mining on laboratory data," *BMC Medical Informatics*, 2023.

[15] A. Vashishtha, K. Patel, and S. Singh, "Causal inference with LLMs: Opportunities and challenges," in *Proceedings of NeurIPS Workshop on Causality*, 2023.

[16] J. Wang, Q. Liu, and S. Zhang, "Retrieval-augmented generation for causal and rule-based knowledge discovery," *arXiv preprint arXiv:2503.01234*, 2025. [Online]. Available: https://arxiv.org/abs/2503.01234

[17] Y. Yang and G. Webb, "Discretization for numerical data in association rule mining," *Data Mining and Knowledge Discovery*, vol. 20, no. 3, pp. 400–416, 2010.