Visualizing Rule-based Classifiers for Clinical Risk Prognosis

Dario Antweiler*
Fraunhofer IAIS, Fraunhofer
Center for Machine Learning

Georg Fuchs†
Fraunhofer IAIS

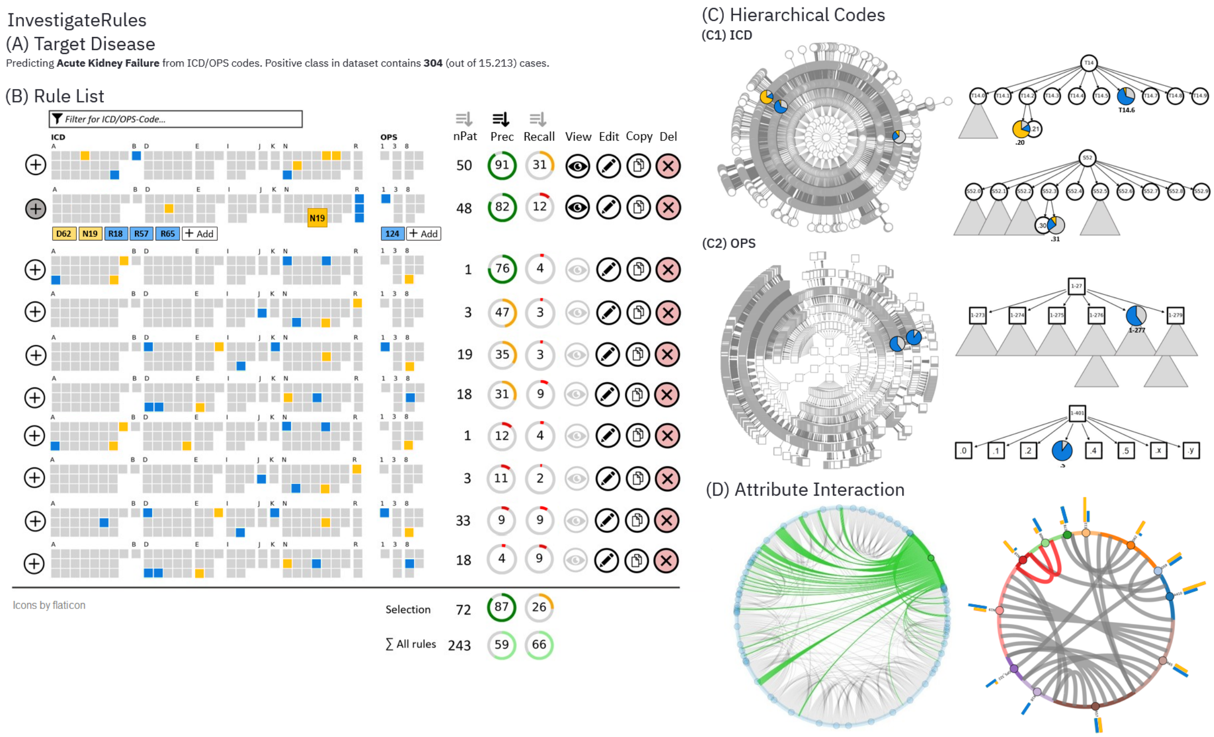

Figure 1: Overview of the our proposed system as a first prototype. It consists of (A) control panels for the selection of the dataset, trained model and logging (B) the main rule list view containing rule attributes as well as quality metrics (C) hierarchical tree view containing ICD and OPS codes and (D) feature interaction view showcasing how code combination are distributed across a user-selected subset of rules.

Abstract

Deteriorating conditions in hospital patients are a major factor in clinical patient mortality. Currently, timely detection is based on clinical experience, expertise, and attention. However, healthcare trends towards larger patient cohorts, more data, and the desire for better and more personalized care are pushing the existing, simple scoring systems to their limits. Data-driven approaches can extract decision rules from available medical coding data, which offer good interpretability and thus are key for successful adoption in practice. Before deployment, models need to be scrutinized by domain experts to identify errors and check them against existing medical knowledge. We propose a visual analytics system to support healthcare professionals in inspecting and enhancing rule-based classifier through identification of similarities and contradictions, as well as modification of rules. This work was developed iteratively in close collaboration with medical professionals. We discuss how our tool supports the inspection and assessment of rule-based classifiers in the clinical coding domain and propose possible extensions.

*e-mail: dario.antweiler@iais.fraunhofer.de
†e-mail: georg.fuchs@iais.fraunhofer.de

Index Terms: Information systems applications [Decision support systems]: Data analytics—; Human computer interaction (HCI) [HCI design and evaluation methods]: User studies— [Applied computing]: Life and medical sciences—Health care information systems

1 Introduction

Automatic risk prognosis systems that predict relevant complications based on patient data and subsequently alert clinicians are highly desirable and have been the goal of many research endeavors [14,32]. Approaches can be divided into prediction of in-hospital mortality on one hand, and prognosis of specific medical conditions including acute kidney failure (AKF), respiratory failure, sepsis, and need for transfusion on the other hand. Prediction is based on heterogenous data such as admission details, anamnesis, comorbidities, and laboratory results. Other major factors are pre-existing or developing medical conditions together with carried out operations and procedures within the hospital. Due to international efforts and driven by medical controlling departments within healthcare systems, medical conditions, operations and procedures are coded in a standardized, machine-readable way in many countries worldwide. Specifically, hospitals in Germany are required by law to use the *International Statistical Classification of Diseases and Related Health Problems* (ICD) [35] and the *Operationen- und Prozedurenschlüssel* (OPS, based on WHO's ICPM [34]) classification systems for all patients.

Rule learning is a widely used machine learning technique that extracts sets of rules from structured datasets [9] in the form IF X = TRUE and Y = FALSE $\Rightarrow$ Z = TRUE. Rule-based models are notably used in decision-critical fields like healthcare and security as domain experts can understand and review rules as opposed to black-box models.

This paper proposes a visual analytics system that supports healthcare professionals (HCPs) in inspecting classification rule sets for clinical risk prediction based on ICD and OPS coding. The work has been developed in close collaboration with HCPs in a research project aiming for a data-driven early risk identification system in a clinical setting.

2 Related Work

The challenge of explaining model behaviour in a variety of fields is well recognized, and especially so in health-related applications [20]. A number of frameworks supporting reasoning about black-box models have been proposed, often focussing on specific model classes like DNN [2]. Within this body of work, our contribution is closely related to the research areas of visual analytics for risk prognosis in healthcare, as well as rule set visualization.

2.1 Visualizing Clinical Risk

Visualization for Clinical Decision Support Systems (CDSS) plays a key role in the interaction between the system and the clinical users. Early works visualized medical patient records [6, 22]. Some approaches focus on specific risk types, such as visualizing surgery complications [4] or adverse drug effects [12, 29]. Later works adressed specific disease trajectories by outcomes [17], often in combination with the identification and visualization of similarities between patients [24]. The work most similar to ours explores clinical code visualization for intensive care unit (ICU) patients with node-link diagrams and dimensionality reduction techniques [1]. In comparison, we focus on inspecting and checking rule-based classifiers for risk prognosis regarding patients in the general ward of a hospital. This poses an additional challenge, as the breadth of possible clinical codes is far larger than in the narrow ICU setting.

2.2 Visualizing Rule Sets

While rule sets can be visualized by a simple list or table, a large body of literature emerged to enable a diverse set of visual analytics tasks. Some focus on the instances covered by rules, rule exceptions [25] or differences between rules [19]. Other works emphasize the visualization on rule metrics such as *confidence* or *support* [19, 31] while some highlight the relations between attributes through chord diagrams [5], parallel coordinates [36] or directed graphs [16]. Visualization idioms deployed include matrices, grids and lattices [3, 11, 27, 30] as well as mosaic plots [10]. Innovative approaches use 3D visualizations [21, 33, 37], self-organizing maps [7] or dimensionality reduction techniques [19]. Specialized visualizations were created for rules regarding certain data types, such as temporal data [8, 38] or spatial data [23]. Due to the fact that decision trees can be represented as sets of rules, one can make use of tree visualization approaches as well [11, 39]. A comprehensive overview for pattern visualization can be found in [13].

None of the identified approaches address the need for a visualization system targeted at medical professionals to inspect and interact with a rule-based classifier. Mainly, they lack the ability to study individual rules as well as combinations and interactions between them. To the best of our knowledge, our approach of a multi-view system with both rule-specific and hierarchical visualization is novel.

3 Concept

Our proposed visualization system consists of multiple views and is based upon the well-known principle *Overview first, zoom and filter, details on demand* [28].

3.1 Data

The data used to develop and test the proposed system consists of 42m anonymized records of hospital visits in Germany between the years 2011 and 2020 across multiple hospitals. Each datapoint contains information on demographics (age, gender), medical conditions as well as procedures, which are coded according to the ICD [35] and German OPS classification systems, respectively. Additionally, the dataset contains year of admission and length of stay.

ICD and OPS codes are particularly suitable starting points for risk prognosis in hospitals because they are readily available from documentation. Their usage is required by law in Germany and many other countries worldwide enforce identical or similar classifications. Each of the 16,674 alphanumeric ICD codes describes a specific medical condition (e.g. *N17 - Acute Kidney Injury*) and is part of the ICD code hierarchy, organized into layers divided into chapters and groups. During a hospital visit, either doctors themselves or dedicated *medical coders* translate written clinical documentation (e.g. admission records, surgery protocols) into ICD codes for downstream controlling and statistical analysis tasks. It is important to note only leaf nodes in the ICD tree are used for coding, while inner nodes are used for aggregation purposes only. Both code systems are updated regularly by the governing body with codes being added, removed, or split, such that even HCPs find it challenging to remember the leaf-level in full detail. Nevertheless, we acknowledge that medical coding data has its limitations, as the data quality depends on specific hospital circumstances and is sometimes affected by financial rather than purely medical drivers.

In the specific project at hand, the goal was the development of an interpretable set of decision rules for early identification of clinical risks including multiple types of organ failure, ventilation, and sepsis based on the provided historical dataset.

3.2 Users & Tasks

We developed our system in close collaboration with HCPs, specifically, two hospital doctors. Based on the growth of available structured patient data and the track record of interpretable machine learning applications in other domains, an increasing number of systems try to combine traditional knowledge-driven expert systems with data-driven approaches in clinical medicine. Projects with this goal are typically established bringing together a group of data scientists with HCP and organized through proven industry processes such as CRISP-DM [26]. A key element in the process is the evaluation of trained models not only statistically, but to also validate them *semantically* wrt. the domain goal. After partaking in multiple similar projects, the necessity for tools supporting this process step became apparent.

Currently, the modus operandi is as follows: (1) The HCP selects a dataset of patient visits, a condition of interest (e.g. "kidney failure") as well as a subset of medical codes (from ICD + OPS), that might hold predictive power for that specific condition from a clinical perspective. (2) A data scientist preprocesses the data, selects one or multiple suitable model architectures (e.g. rule learning) that satisfy interpretability requirements and evaluates the model(s) statistically. (3) Lastly, the HCP examines the learned model, which may contain a large number of rules, by probing specific aspects of the results, such as individual rules or risk predictions and gives feedback to the data scientist. This iterative process is repeated multiple times until the HCP is satisfied with the results of the learned model both in terms of predictive power and comprehensibility. Based on the described workflow, in which the rule learner is external to our system, we identified the following key tasks that we aim to support with our visualization system:

$\mathbf{T_1}$: Inspect rules separately as well as in combination to understand their structure

$\mathbf{T_2}$: Identify and observe similarities and differences between rules

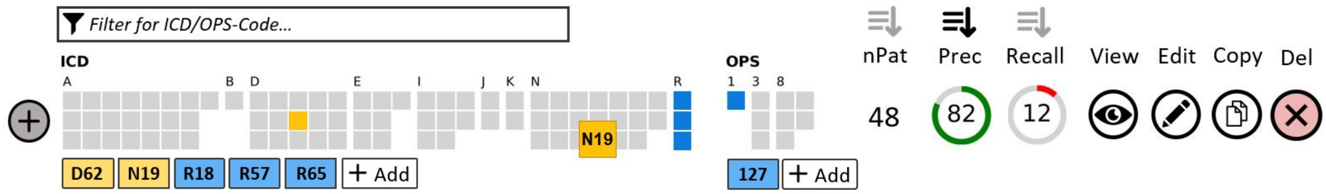

Figure 2: System detail: The main view contains a list of all rules, describing their contents and metrics in detail including (1) a *rule fingerprint* displaying which ICD and OPS codes are inclusive (blue) or exclusive (yellow) conditions within the rule, (2) relevant *rule metrics* consisting of coverage, precision and recall regarding the training data and (3) *rule editing* elements to add or remove individual codes as well as duplicate or remove entire rules from the set. The rule set can be filtered and sorted via respective description fields and buttons to select for rules of interest.

T_3: Support HCPs in checking extracted rules against pre-existing knowledge from a medical and coding domain

T_4: Modify and remove rules according to identified patterns

As reported above, these tasks are not addressed by the tools currently deployed by HCPs. In addition, any proposed solution needs to fulfill a set of requirements to be a viable option:

R_1: Widely known visualization idioms are used to appeal to a diverse group of HCPs without any visual analytics background

R_2: Linked multiple view interface integrating all task-related aspects according to the information seeking mantra [28]

R_3: Computation and visualization is done locally to reduce risks when working with privacy sensitive data

We will refer back to the tasks and requirements throughout the paper to showcase our efforts to address them.

The implementation was based on the requirements gathered from the discussions with HCPs and focused the development of visual elements on solving the use case at hand. The system consists of three main interactive displays (Fig. 1) that we describe in detail in the following sections. Rule list and attribute interaction view generate a high-level overview while the hierarchical code view offers a drill-down on a selected subset of rules.

3.3 Visualization & Refinement of Clinical Risk Rule Sets

The main view of our system displays an overview of the entire rule set, consisting of individual *rule fingerprints* together with relevant *rule metrics* as well as *rule interaction* elements. The *rule fingerprint* (cf. [15]) on the left showcases the contained attributes for each rule. Recall that rules in our scenario consist of a combination of included or excluded medical codes from the ICD and OPS hierarchies. Within the fingerprint, code chapters are divided up into separate waffle chart segments with color indicating the given code's inclusion within this rule, i.e., whether a code must (positive – blue), must not (negative – yellow), or may optionally (grey) be present in a patient record (**T_1**). Based on the fingerprints, a user can easily identify similarities and differences between a selection of rules (**T_2**) as confirmed by feedback from HCPs. For instance, two rules with a large overlap but some distinct codes, encourage the HCP to reflect on the most sensible pathomechanisms and lead to the subsequent editing of a rule. Individual codes are displayed on mouse-over, while the entire attribute set can be observed by activating a detailed mode on the "⊕" button. This facilitates the examination of a specific rule, especially when the distribution of included attributes reveals patterns that are unexpected or run against premeditated assumptions by the HCP (**T_3**). We show the most important *rule metrics* next to the fingerprint, namely *coverage* (in our case number of patients), *precision* and *recall*. The latter two ratios are visualized by colored circle arcs, inspired by [18]. All three metrics are important for sorting or comparing rules and often the starting point of an analysis (**T_2**). A subset of rules can be selected via the "view" button, which itself controls the visualization in the connected view of hierarchical code sets (**R_2**). Motivated by user feedback, buttons on the right enable modifying the entire rule set by adding or removing, as well as editing individual rules (**T_4**).

Adding or removing individual attributes (codes) is achieved simply clicking on the corresponding waffle chart segment in edit mode to cycle between its positive, negative, or non-inclusion in the given rule. Including a new code can also be done via the corresponding "+ Add" button (cf. Fig. 2) and directly entering the code as text prefixed with '+' or '−', respectively. By editing a copy of an existing rule, HCPs can further augment the rule set with code combinations matching their professional expertise or assumptions. The modified rule set can then be re-applied to the data set under investigation (cf. Fig. 1A). By observing the performance metrics for the modified selection of rules and the rule set as a whole, the HCP can assess their individual validity as well as the impact of these modifications on the overall model, and whether they represent an improvement over the previous model and its comprising rule set.

Our system focuses on non-leaf codes (e.g. "N17") in the hierarchy due to the use case at hand. During discussion with HCPs, we recognized that codes involved in risk prognosis usually include higher level nodes as well as all of their subsidiary nodes. Attention to individual leaf codes may require different visualization approaches not yet covered by our system.

3.4 Visualization of Hierarchical Code Sets

We target the task of understanding the hierarchical relationships between rule attributes (**T_1**, **T_3**) by connected views on the right of the system (**R_2**). Both ICD and OPS can be described as trees with nodes representing codes and edges representing parent-child relations, where a child is typically a more disease-specific code (e.g. acute kidney injury with tubular necrosis) than its parent (acute kidney injury). The overall structure of this hierarchy is widely known and used among HCPs who deal with these codes daily. Therefore, we designed our view with the requirement in mind that codes and their relative position to each other are an important aspect in reviewing a rule-based machine learning model. Specifically, we combine a high-level radial tree view of the entire code hierarchy with a selection-driven local view of a set of subtrees.

Within the radial tree view, all codes contained in the rule subset selected in the main view, are highlighted for an immediate overview. One of the most important aspects in rule inspection in our use case is the observation whether codes are either more often part of a rule as an inclusion or as an exclusion criterion (**T_2**). To address this need, we display these ratios as a small pie chart in place of the concerned nodes. The same color coding as for the rule fingerprints is used for the pie segments. Alongside, the local view is purely focused on the subparts of the hierarchy tree that contain codes from rules selected in the main view. Affected subtrees are visualized in a simple hierarchical top-down manner (**R_1**), while parts which do not contain any relevant nodes are compressed into triangular glyphs. They can be expanded on-click to reveal their structure if needed.

3.5 Feature Dependencies

The third and last component puts an emphasis on the interaction of codes within a subset of rules by showing how often codes are present in combination (**T$_2$**). The *feature dependencies view* is comprised of a chord diagram, connecting nodes (codes) via color coded arcs, if they share an appearance within selected rules. This allows the user to quickly identify codes that are highly congruent, even if they do not share a close common ancestor in the respective code hierarchy. The chord diagram includes both ICD and OPS codes. HCPs emphasized that precisely the interaction between both code sets is important to observe. Typically, each procedure code is triggered by an ICD code, i.e. a condition this procedure addresses.

3.6 Technical details

We deploy various filtering capabilities to enable efficient analysis of rules (**T$_1$**). The rule list in the main view can be filtered by code keywords, while the chord view provides filtering via selection of a node on the outer circle. Analysis is done locally on device in the browser (**R$_3$**). Backend computation is done in *python*, while the user interface uses the *matplotlib*, *bokeh* and *streamlit* libraries.

4 Use Case

4.1 Workflow

This section showcases an exemplary workflow that utilizes the proposed system system. In our scenario, a HCP wants to explore a rule set that has been generated by a machine learning algorithm to predict acute kidney failure from clinical code data. First, the dataset together with a learned rule set is selected as an input and constitutes the base for the analysis. Initially the main view displays each rule individually as part of a rule list (cf. Fig. 2) to give the user an overview. The *rule fingerprints* can be used to identify similar rules and the *rule metrics* depict a general notion of rule quality. From exchange with HCPs, we discovered that the order of importance is *precision, coverage, recall*. If needed, rule fingerprints can be further expanded to show individual codes, following the idea of "details on demand". The user selects a subset of rules from the list due to further exploration.

With this objective, the focus shifts to the hierarchical code view on the right (Fig. 3). The main driving questions by users regard the *relatedness* of codes and how codes are *distributed* across the rule set. The radial tree view gives the user an idea of affected hierarchy levels and distribution among code chapters. In contrast, the detailed tree view is used to explore a small group of related codes. During interviews, users mentioned the need for analysis of "code families" consisting of parent nodes (more general) and child nodes (more specific) to decide how the rule should be configured and achieve a balance in precision and recall. The tree can be expanded by clicking on collapsed subtrees, visualized by gray triangles.

Lastly, the chord view enables the identification of code combinations that attract or oppose each other. Codes connected through a large number of chords in the default view are instantly recognized as "partner codes", while more subtle relationships can be detected by clicking on a node of interest which highlights only connected nodes and fades out all other. In one instance, the test user identified a divergence of a rule from pre-existing medical coding knowledge through the hierachical tree view: The rule-learning algorithm identified *J91 - Pleural effusion* as an inclusion criterion for acute kidney failure, without acknowledging the fact that this code cannot be coded stand-alone and needs an accompanying code such as *I50.- Heart failure* or *J01.- Acute Sinusitis*. Hence, the user duplicated the rule and added the respective codes to enhance the rule set. After editing, the final rule set is saved and exported for subsequent usage.

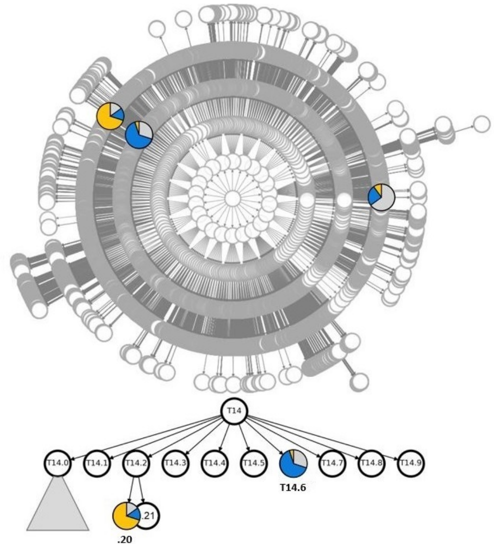

Figure 3: The hierarchical code view juxtaposes global and local structure. The radial tree view displays codes and their relationships. Codes from selected rules in main view are replaced by pie charts depicting the distribution of positive (blue), negative (yellow) or no incidence (gray). Below, the top-down tree highlights dependencies.

4.2 Discussion

We discussed our system with two HCPs in multiple structured interviews each, and received strong positive feedback. The combined view of the rule set together with the hierarchical code view supports the inspection of individual rules as well as entire rule sets. The attribute interaction view promotes faster assessment of rules against pre-existing medical or coding knowledge. Specifically, both HCP stated that unexpected code combinations are explored more easily than with basic rule tables. During consultations we also discussed several variations to the visual encoding within all three views. We intend to do a more in-depth evaluation of design alternatives and their impact in future work. While experimenting with the prototype, both HCPs found it very helpful to create different variations of specific rules and observe their impact on model prediction, before settling on the final rule set. We intend to support this by further augmenting the system with an "editing timeline" to showcase the impact of rule changes on relevant metrics, allow rollback of rule edits, thereby facilating exploration of branching rule set variations.

5 Conclusion and Future Work

In this paper, we contributed a novel visual system to support HCPs in understanding and reviewing rule-based classifiers for clinical risk prediction. Our system visualizes similarities and differences between rules and contextualizes their attribute distribution within hierarchical code structures. Additionally, the interactions between rule attributes can be explored. We evaluated our initial design with HCPs and received positive feedback. As future work, we aim to expand our framework to further optimize the visual encoding of our co-occurrence diagram, since initial user feedback specifically indicated the importance of evaluating rule structure across code hierarchies. Currently, our system is limited to datasets consisting of hierarchical code sets. In future, CDSS will additionally include demographic and laboratory records in their models, and we plan to investgate their inclusion into our system as well.

Acknowledgements

The work has been supported by the Fraunhofer Center for Machine Learning @ Fraunhofer Cluster for Cognitive Internet Technologies.

REFERENCES

[1] D. Alcaide and J. Aerts. A visual analytic approach for the identification of ICU patient subpopulations using ICD diagnostic codes. *PeerJ Computer Science*, 7:e430, Apr. 2021. doi: 10.7717/peerj-cs.430

[2] G. Alicioglu and B. Sun. A survey of visual analytics for explainable artificial intelligence methods. *Computers & Graphics*, 102:502–520, 2022. doi: 10.1016/j.cag.2021.09.002

[3] F. Z. Benhacine, B. Atmani, and F. Z. Abdelouhab. Contribution to the association rules visualization for decision support: A combined use between boolean modeling and the colored 2d matrix. *IJIMAI*, 5(5):38, 2019. doi: 10.9781/ijimai.2018.09.002

[4] A. Bihorac, T. Ozrazgat-Baslanti, A. Ebadi, A. Motaei, M. Madkour, P. M. Pardalos, G. Lipori, W. R. Hogan, P. A. Efron, F. Moore, L. L. Moldawer, D. Z. Wang, C. E. Hobson, P. Rashidi, X. Li, and P. Momcilovic. MySurgeryRisk: Development and validation of a machine-learning risk algorithm for major complications and death after surgery. *Annals of Surgery*, 269(4):652–662, Apr. 2019. doi: 10.1097/sla.0000000000002706

[5] S. Bornelöv, S. Marillet, and J. Komorowski. Ciruvis: a web-based tool for rule networks and interaction detection using rule-based classifiers. *BMC Bioinformatics*, 15(1), May 2014. doi: 10.1186/1471-2105-15-139

[6] A. A. T. Bui, D. R. Aberle, and H. Kangarloo. TimeLine: Visualizing integrated patient records. *IEEE Transactions on Information Technology in Biomedicine*, 11(4):462–473, July 2007. doi: 10.1109/titb.2006.884365

[7] W. Castillo-Rojas, C. Vargas, and C. M. Villegas. Interactive visualization of association rules model using SOM. In *Proceedings of the XV International Conference on Human Computer Interaction - Interacción '14*. ACM Press, 2014. doi: 10.1145/2662253.2691319

[8] C. Combi and A. Sabaini. Extraction, analysis, and visualization of temporal association rules from interval-based clinical data. In *Artificial Intelligence in Medicine*, pp. 238–247. Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-38326-7_35

[9] J. Fürnkranz, D. Gamberger, and N. Lavrač. *Foundations of Rule Learning*. Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-540-75197-7

[10] H. Hofmann, A. P. J. M. Siebes, and A. F. X. Wilhelm. Visualizing association rules with interactive mosaic plots. In *Proc. ACM SIGKDD 2000*. ACM Press, 2000. doi: 10.1145/347090.347133

[11] K. huat Ong, K. leong Ong, W.-K. Ng, and E.-P. Lim. CrystalClear: Active visualization of association rules. In *In ICDM'02 International Workshop on Active Mining AM2002*. Press, 2002.

[12] L. Jean-Baptiste, H. Berthelot, and M. Favre. Rainbow boxes: A technique for visualizing overlapping sets and an application to the comparison of drugs properties. In *2016 20th International Conference Information Visualisation*. IEEE, July 2016. doi: 10.1109/iv.2016.26

[13] W. Jentner and D. A. Keim. Visualization and visual analytic techniques for patterns. In *Studies in Big Data*, pp. 303–337. Springer International Publishing, 2019. doi: 10.1007/978-3-030-04921-8_12

[14] M. Jones. NEWSDIG: The national early warning score development and implementation group. *Clinical Medicine*, 12(6):501–503, Dec. 2012. doi: 10.7861/clinmedicine.12-6-501

[15] D. A. Keim and D. Oelke. Literature fingerprinting: A new method for visual literary analysis. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pp. 115–122, 2007. doi: 10.1109/VAST.2007.4389004

[16] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A. I. Verkamo. Finding interesting rules from large sets of discovered association rules. In *Proceedings of the third international conference on Information and knowledge management*. ACM Press, 1994. doi: 10.1145/191246.191314

[17] G. E. Marai, C. Ma, A. T. Burks, F. Pellolio, G. Canahuate, D. M. Vock, A. S. R. Mohamed, and C. D. Fuller. Precision risk analysis of cancer therapy with interactive nomograms and survival plots. *IEEE TVCG*, 25(4):1732–1745, Apr. 2019. doi: 10.1109/tvcg.2018.2817557

[18] Y. Ming, H. Qu, and E. Bertini. Rulematrix: Visualizing and understanding classifiers with rules. *IEEE TVCG*, 25(1):342–352, 2019. doi: 10.1109/TVCG.2018.2864812

[19] A. Mukherji, X. Lin, J. Whitehouse, C. R. Botaish, E. A. Rundensteiner, and M. O. Ward. FIRE: a two-level interactive visualization for deep exploration of association rules. In *Proceedings of the 22nd ACM CIKM '13*. ACM Press, 2013. doi: 10.1145/2505515.2505631

[20] J. Ooge, G. Stiglic, and K. Verbert. Explaining artificial intelligence with visual analytics in healthcare. *WIREs Data Mining and Knowledge Discovery*, 12(1):e1427, 2022. doi: 10.1002/widm.1427

[21] M. S. Ounifi, H. Amdouni, R. B. Elhoussine, and H. Slimane. New 3d visualization and validation tool for displaying association rules and their associated classifiers. In *2016 20th International Conference Information Visualisation*. IEEE, July 2016. doi: 10.1109/iv.2016.78

[22] S. Powsner. Graphical summary of patient status. *The Lancet*, 344(8919):386–389, Aug. 1994. doi: 10.1016/s0140-6736(94)91406-0

[23] A. Savinov. Mining spatial rules by finding empty intervals in data. In *Lecture Notes in Computer Science*, pp. 1058–1063. Springer Berlin Heidelberg, 2003. doi: 10.1007/978-3-540-45224-9_141

[24] J. Schaaf, M. Sedlmayr, H.-U. Prokosch, N. Tegtbauer, D. Kadioglu, J. Schaefer, M. Boeker, and H. Storf. Visualization of similar patients in a clinical decision support system for rare diseases – a focus group study. In *German Medical Data Sciences: Bringing Data to Life*. IOS Press, May 2021. doi: 10.3233/shti210050

[25] G. Sharma, G. Shroff, A. Pandey, P. Agarwal, and A. Srinivasan. Interactively Visualizing Summaries of Rules and Exceptions. In M. Pohl and J. Roberts, eds., *EuroVis Workshop on Visual Analytics*. The Eurographics Association, 2014. doi: 10.2312/eurova.20141146

[26] C. Shearer. The crisp-dm model: The new blueprint for data mining. *Journal of Data Warehousing*, 5(4), 2000.

[27] X. Shen, L. Bao, and L. Zhang. Research on visualization algorithm of association rules based on concept lattice. In *CIAT 2020*, p. 22–27. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3444370.3444542

[28] B. Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. In *Proc. of the IEEE Symposium on Vis. Languages*, pp. 336–343, Sep 1996. doi: 10.1109/VL.1996.545307

[29] A. F. Simpao, L. M. Ahumada, B. R. Desai, C. P. Bonafide, J. A. Gálvez, M. A. Rehman, A. F. Jawad, K. L. Palma, and E. D. Shelov. Optimization of drug–drug interaction alert rules in a pediatric hospital's electronic health record system using a visual analytics dashboard. *JAMIA*, 22(2):361–369, Oct. 2014. doi: 10.1136/amiajnl-2013-002538

[30] K. Techapichetvanich and A. Datta. VisAR : A new technique for visualizing mined association rules. In *Advanced Data Mining and Applications*, pp. 88–95. Springer Berlin Heidelberg, 2005. doi: 10.1007/11527503_12

[31] A. Unwin, H. Hofmann, and K. Bernt. The TwoKey Plot for multiple association rules control. In L. De Raedt and A. Siebes, eds., *Principles of Data Mining and Knowledge Discovery*, pp. 472–483. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[32] J. L. Vincent, R. Moreno, J. Takala, S. Willatts, A. D. Mendonça, H. Bruining, C. K. Reinhart, P. M. Suter, and L. G. Thijs. The SOFA score to describe organ dysfunction/failure. *Intensive Care Medicine*, 22(7):707–710, July 1996. doi: 10.1007/bf01709751

[33] B. Wang, T. Zhang, Z. Chang, T. Ristaniemi, and G. Liu. 3d matrix-based visualization system of association rules. In *2017 IEEE International Conference on Computer and Information Technology (CIT)*. IEEE, Aug. 2017. doi: 10.1109/cit.2017.39

[34] WHO. International classification of procedures in medicine, 1978.

[35] WHO. ICD-10 : international statistical classification of diseases and related health problems : tenth revision, 2004.

[36] L. Yang. Visualizing frequent itemsets, association rules, and sequential patterns in parallel coordinates. In *Computational Science and Its Applications — ICCSA 2003*, pp. 21–30. Springer Berlin Heidelberg, 2003. doi: 10.1007/3-540-44839-x_3

[37] Z. yun Zheng, J. zhuang Guo, Y. Fu, L. Li, and D. Li. Visualization of association rules based on a gray correlation analysis model. In *ICNC-FSKD 2017*. IEEE, July 2017. doi: 10.1109/fskd.2017.8393345

[38] K. Zhao and B. Liu. Visual analysis of the behavior of discovered rules. 2001.

[39] X. Zhao, Y. Wu, D. L. Lee, and W. Cui. iForest: Interpreting random forests via visual analytics. *IEEE TVCG*, 25(1):407–416, Jan. 2019. doi: 10.1109/tvcg.2018.2864475