Task-Based Visual Interactive Modeling: Decision Trees and Rule-Based Classifiers

Dirk Streeb, Yannick Metz, Udo Schlegel, Bruno Schneider, Mennatallah El-Assady, Hansjörg Neth, Min Chen, *Member, IEEE*, and Daniel A. Keim, *Member, IEEE*

Abstract—Visual analytics enables the coupling of machine learning models and humans in a tightly integrated workflow, addressing various analysis tasks. Each task poses distinct demands to analysts and decision-makers. In this survey, we focus on one canonical technique for rule-based classification, namely decision tree classifiers. We provide an overview of available visualizations for decision trees with a focus on how visualizations differ with respect to 16 tasks. Further, we investigate the types of visual designs employed, and the quality measures presented. We find that (i) interactive visual analytics systems for classifier development offer a variety of visual designs, (ii) utilization tasks are sparsely covered, (iii) beyond classifier development, node-link diagrams are omnipresent, (iv) even systems designed for machine learning experts rarely feature visual representations of quality measures other than accuracy. In conclusion, we see a potential for integrating algorithmic techniques, mathematical quality measures, and tailored interactive visualizations to enable human experts to utilize their knowledge more effectively.

Index Terms—Decision trees, rule-based classification, visual analytics, interactive machine learning, interactive model analysis, survey, visualization

1 Introduction

Interactive machine learning has gained large interest in the visualization and visual analytics community [1], [2]. However, visualizations need to match the demands of distinct analysis tasks [3]. Visual analytics promises to provide exceptional matches by offering specialized, bi-directional interfaces between analysts and machine learning models [4], [5]. From the assertions above we can expect that distinct visualizations, suitable for different analysis tasks, have been developed.

Visualization can facilitate steps along the analysis workflow in at least two ways: First, visualizing the data aids in spotting outliers in training data and predictions. Second, representing abstract models visually can support understanding [6]. For example, the data flow through a decision tree can be explicated by augmenting nodes with class distributions [7]. Additionally, visual analytics introduces the direct manipulation of the underlying model to enable the integration of domain knowledge in model construction, for instance by adjusting split values of a decision tree [8]. Further, targeted what-if analyses facilitate the diagnosis of malfunctions. Once sources of errors are identified, problems can be fixed and resulting changes can be observed immediately. Within classification, several distinct analysis tasks/steps can be identified, including model building and refinement. All steps can benefit from the close involvement of human analysts via visual interfaces.

As a result, (interactive) visualization, and visual analytics in particular, are central angles of attack for improving machine learning models [4], [5]. At the same time, visualization can provide the foundation for the utilization of constructed models, as well as the extraction and communication of new insights gathered from modeling a classification problem. Therefore, it is in the interest of the visualization and visual analytics community to obtain an overview of what kinds of visualizations are available for solving tasks in interactive modeling and machine learning [9].

In this survey, we investigate whether visual designs actually diversify and are tailored more closely to individual tasks, or whether general-purpose visualizations flourish. We restrict the survey to decision trees and rule-based classifiers, which are one of the canonical types of classifier models. (More details on this choice follow in Section 2). This restriction lowers the barrier for readers who use this survey as an entry to visual interactive modeling/machine learning, and highlights a topic that has attracted repeated interest by scholars and practitioners. Furthermore, the specialized focus avoids potential confusion introduced by mixing miscellaneous types of models. In consequence, it is straightforward to compare differences between visualizations across tasks, while an abstract and model-agnostic set of tasks enables the generalization of results. Future evaluations may build on our survey in order to substantiate design defaults and guidelines, which potentially can be transferred to the visualization of other types of machine learning models.

- *Dirk Streeb, Yannick Metz, Udo Schlegel, Bruno Schneider, Mennatallah El-Assady, Hansjörg Neth, and Daniel A. Keim are with the Universität Konstanz, 78464 Konstanz, Germany. E-mail: {dirk.streeb, yannick.metz, udo.3.schlegel, bruno.schneider, Mennatallah.El-Assady, h.neth, keim}@uni-konstanz.de.*
- *Min Chen is with the University of Oxford, Oxford OX1 2JD, U.K. E-mail: min.chen@oerc.ox.ac.uk.*

Manuscript received 18 June 2020; revised 12 Dec. 2020; accepted 14 Dec. 2020. Date of publication 13 Jan. 2021; date of current version 1 Aug. 2022. (Corresponding author: Dirk Streeb.)
Recommended for acceptance by R. Maciejewski.
Digital Object Identifier no. 10.1109/TVCG.2020.3045560

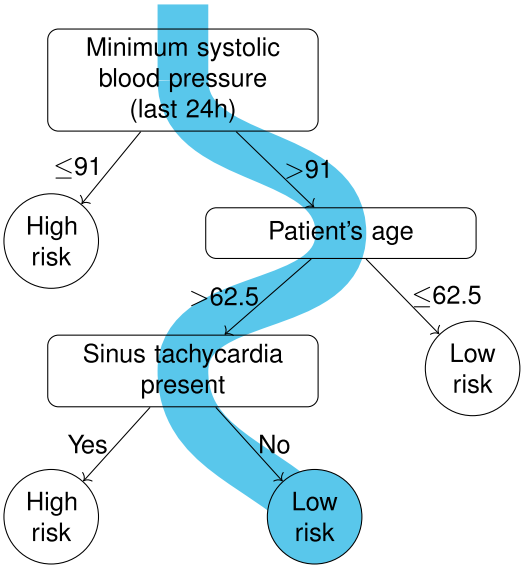


Fig. 1. Decision tree addressing a risk assessment task in an emergency room. Based on observed symptoms, patients are classified into a risk class: *high risk* or *low risk*. How a single patient is classified is transparent as depicted by the blue trace representing one exemplary patient, who is classified as having a low risk. Data by Breiman *et al.* [10].

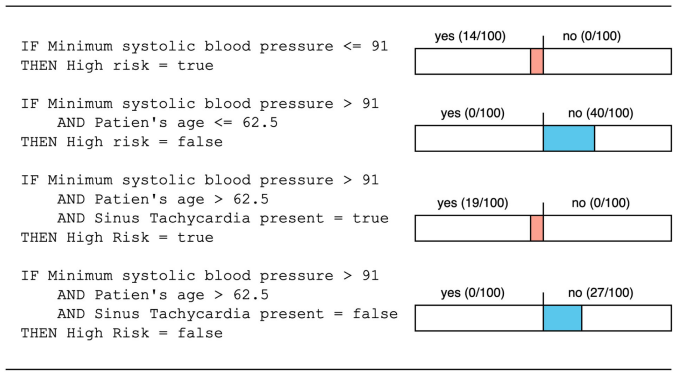


Fig. 2. Rule set equivalent to the decision tree in Fig. 1. The bars at the right show how many of 100 patients each rule covers. Figure inspired by Fürnkranz *et al.* [25].

For clarification, we cover decision tree classifiers, which we call *decision trees* throughout this paper (see also Section 2). In the supplementary material, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TVCG.2020.3045560, we provide a more detailed primer for readers who are not familiar with using decision tree models for classification. Other notions of the term "decision tree" as, for example, used in decision theory [11], [12], expert systems [13], [14], operations research [15], or decision support systems using forecasting [16] lie beyond this scope. Similarly, diagnostic trees that depict class prevalence, and the Recall/Sensitivity and Specificity of binary classifiers [17] share some aspects, but cannot be reasonably covered. To complement our survey, we present a brief comparison of these other meanings and how visualizations have been part of their history in the supplementary material, available online.

In contrast to previous work on tree visualizations [18], [19], [20], [21], [22], our focus is not only on the visual designs. Instead, we focus on how visualizations match the analysis tasks in classification. Endert *et al.* [23] present a broad view on integrating machine learning into visual analytics for dimension reduction, clustering, classification, and regression from the perspectives of models and frameworks, techniques, and application areas. More generally, Jiang *et al.* [9] summarize recent advances in interactive machine learning. Finally, Sacha *et al.* [2] propose an ontology integrating visualization and machine learning. By contrast, our survey details on available visual representations. The survey of Liu and Salvendy [24] is most closely related to our work. In 2007, they surveyed the aspects visualization of tree models, visualization of tree evaluation, and visual interactive tree construction. We present an updated overview of the topic and consider a broader scope of analysis tasks. Thereby, we advance the research on task-based visualization and point out open questions about how to tailor visualizations and visual analytics systems closely to task demands. In particular, we contribute:

- a survey of visualizations for decision tree classifiers,
- a categorization of visualizations from 152 publications by 16 tasks and a comparison across these tasks, and
- an outlook on open questions and opportunities in visual interactive modeling and machine learning.

Additionally, we include categorizations by the types of visual designs employed, as well as the numeric quality measures displayed, and briely discuss the lack of evaluation studies. In the next section, we introduce the concept of decision tree classifiers and provide details on our choice for focusing on this type of models. In Section 3, we describe our methodology and present an overview of results. Afterwards, we present detailed results grouped by three perspectives, namely Classifier Development (Section 4), Classifier Utilization (Section 5), and the Descriptive Modeling of Classification Processes (Section 6). Based on these results, we discuss the role of visualization and visual analytics across analysis tasks, visual designs and quality measures, in Section 7. Resulting findings lead us to Open Questions and Opportunities, which we present in Section 8.

2 Decision Trees and Decision Rules

Compared to other types of classifier models, decision trees closely resemble human reasoning. Hence, they are more transparent and easier to understand [26]. Further, classification trees lend themselves to visual representation, for example, as a node-link diagram in Fig. 1, which supports comprehensibility. In the training process, decision trees require comparatively few observations and can be refined interactively [27]. Decision trees can also be robust and fast in application [28]. For instance, features are applied sequentially and only need to be measured on demand.

Other types of machine learning classifiers, for instance, based on deep learning [29], achieve promising results in several scenarios [30]. However, their inherent lack of interpretability can be problematic [31]. Thus, decision trees are especially useful when human comprehension or interaction with the model is required. This includes applications in contexts that

demand high levels of trust requiring a thorough understanding and validation of classification processes [32], the manual execution of a classification procedure to enable decision-making [33], as well as gaining new insights from data [34].

For our survey, decision trees in combination with rule-based classifiers are a perfect choice for several reasons. Beyond the more general positive aspects summarized above, decision trees are a canonical part of most introductions to classification with machine learning. Second, the combination of a long history of investigation and up-to-date research is beneficial. Finally, decision trees are widely used in visualization and practical application. In this paper, we consider rule-based classifiers as a subset of decision trees, namely the non-branching trees with nodes created from the list of rules. Alternatively, every tree can be represented as a set of rules by formulating each path through the tree as a rule. Classification rules are not to be confused with *association rules*, which do not target classification [37]. Fig. 2 shows a set of classification rules that is equivalent to the decision tree in Fig. 1. For example, the decision rule "If *Minimum systolic blood pressure* > 91 and *Patient's Age* > 62.5 and *Sinus tachycardia* is not present, then *High risk* = false" is equivalent to the path highlighted in Fig. 1. The bars on the right-hand side of Fig. 2 depict how many instances of the underlying dataset are classified by each rule (i.e., their coverage).

Apart from extracting classification rules out of a classification tree, they can also be induced directly from datasets. This process is called *rule set induction*, for which a variety of algorithms have been proposed [25]. Compared to the exclusive paths in decision trees, in principle, decision rules may overlap. As a result, multiple rules cover the same instances, which demands for a mechanism to break ties of overlapping rules predicting contrary classes. For example, rules can be ordered as a list and the first rule that applies determines the prediction [38], [39], which leads to a structure that, again, can be represented by decision trees.

3 Methodology and Overview of Results

In this survey, we cover visualization journals such as TVCG, CG&A, CGF, IV as well as the VIS, EuroVis, and Diagrams conferences. We complement the publications from these established venues with publications from outside the visualization and visual analytics community in order to provide a broad overview. In particular, we apply a sampling strategy as depicted in Fig. 3, which also provides an overview of the categorization process.

Main Sources. We primarily examined eight sources for publications. For a rough comparison, the numbers in parenthesis below indicate the sizes of initial result sets based on the search terms *decision tree* or *rule-based classification*, and *visual**. The wildcard term *visual** covers relevant terms such as *visualization*, *visualisation*, and *visual analytics*. The number of publications that actually present visualizations of decision trees is smaller. Further, some publications are listed in multiple of the following sources:

- IEEE Digital Library (87, including VIS, TVCG, CG&A)
- ACM Digital Library (75, including CHI, IUI)
- Eurographics Digital Library (50, including EuroVis, CGF)
- Pubmed (211)
- PsycInfo (43)
- ArXiv (37)
- Information Visualization Journal (11)
- Diagrams Conference (5)

Then, we manually filtered publications based on contents, keeping all that demonstrate a decision tree as described in Section 2.

Additional Sources. Based on this initial set, we considered references that are related to the topic. We added relevant referenced publications, which are not already in our sample. Finally, we included additional publications from other sources, for example, recommendations of colleagues. Obviously, we could not incorporate all published visualizations of decision trees. Nonetheless, extending our sample beyond visualization and visual analytics venues, which we cover extensively, adds a valuable outlook and records from (scientific) practice.

Publication Sample. Following this strategy, we identified 152 publications featuring visualizations of decision trees. They were published between 1989 and January 2020, with most published after the year 2005 (112 of 150, 75 percent, 2 NA). Fig. 4 shows the distribution of publications in our sample over time.

Categorization. Based on an additional screening, the first author distributed publications such that each was coded by one of the authors. We categorized each decision tree visualization identified in the publications along three

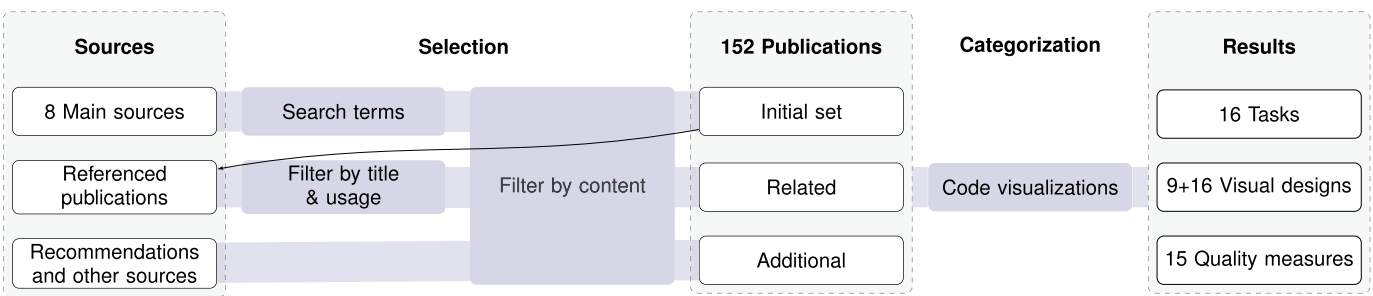


Fig. 3. Selection of eligible publications and categorization of visualizations. Eight main sources covering major visualization venues build the foundation of our sample. Based on a keyword search and manual filtering, we derive an initial set of publications. Following references, we add related publications. Additional publications from other sources such as colleagues' recommendations broaden the sample. For the categorization of visualizations, we distributed publications among the authors. We categorize each visualization based on targeted tasks (cf. [35], [36]), applied visual designs (cf. [24]), and represented quality measures (cf. [3]).

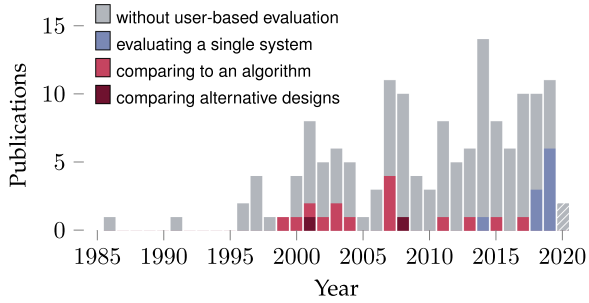


Fig. 4. Number of publications over time. The development of novel visualization techniques in the 1980s and early 1990s [18], [40], [41], [42] lays ground for the visualization of decision trees, starting in the late 1990s. There are some user-based evaluations of single systems and comparisons to automatic algorithms, but comparative experiments between visual designs are rare. References to individual publications can be found in Section 8 and the supplementary material, available online.

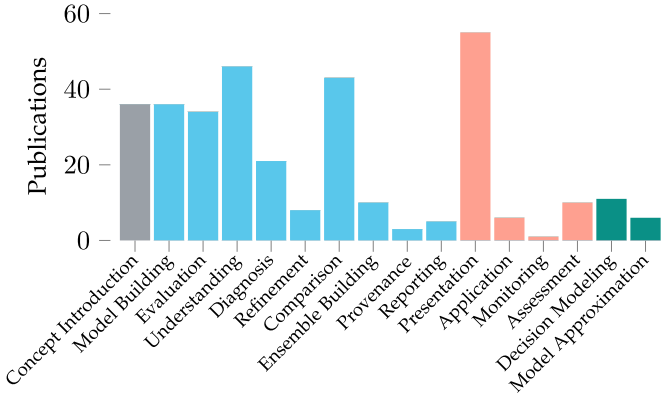


Fig. 6. Number of publications per analysis task. Most visualizations are designed for the steps in Classifier Development (blue). Except for Presentation, tasks in Classifier Utilization (peach) and the Descriptive Modeling of Classification Processes (green) are addressed rarely. Furthermore, visualizations are commonly used to introduce the concept of decision trees (gray). Task descriptions and individual references can be found in Sections 3.1–6 and the supplementary material, available online, respectively.

dimensions: i) Analysis task, ii) Type of visual designs, and iii) Quality measures displayed. Once all visualizations were coded, the first author double checked codings. We present results aggregated by publication, as observations on the level of visualizations are heavily influenced by publications that present many, identically designed, visualizations for presenting and comparing multiple decision trees.

Analysis Tasks. We distinguish between 16 analysis tasks in total. Most of these tasks directly relate to the steps in Classifier Development and Classifier Utilization (cf. [35], [36]). Further, we identified two tasks in the context of the Descriptive Modeling of Classification Processes and the Concept Introduction task. Fig. 5 provides a structured overview. Detailed descriptions of individual analysis tasks can be found below, in Sections 3.1–6. We assigned visualizations designed for, or practically used for, dealing with a task to its category. Fig. 6 shows the distribution of publications across tasks.

Visual Designs. We distinguish between two types of visual designs. First, the tree structure can be represented by node-link diagrams, treemaps, or the like (cf. [24]). Second, more complex systems integrate additional visual components, for example, by encoding class distributions as pipe diagrams (see Fig. 10 below). Some visualizations and most visual analytics systems combine different designs.

Quality Measures. As a third dimension, we tracked which numeric measures for the quality of classifiers, such as Accuracy, Recall/Sensitivity, Gini-index, or Frugality visualizations depict (cf. [3]). Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TVCG.2020.3045560, briefly introduces the measures displayed in our sample.

We present the results along the analysis tasks. The tasks are grouped by the steps in Classifier Development (Section 4), the tasks in Classifier Utilization (Section 5), and approaches to the Descriptive Modeling of Classification Processes (Section 6). Sub-sections detail on individual analysis tasks. Before diving into Classifier Development, we have a look at the visualizations used for introducing the concept of decision trees in the following Section 3.1. In Section 7, we present findings across tasks, which are complemented by a tabular overview in the supplementary material, available online.

3.1 Concept Introduction

The goal of the *Concept Introduction* task is to help novices to understand how decision tree models work in general, not focused on a particular tree. For instance, a visualization outlines the hierarchical structure of decision trees, or explains the sequential application as in Section 2 and Fig. 1.

There are mainly two scenarios for using visualization when introducing the concept of decision trees: On the one hand, introducing decision trees to people and domains that have not been using them before [45], [46], [47]. On the other hand, researchers explain new concepts going beyond the state-of-the-art [48], [49], [50]. Visualizations used for Concept

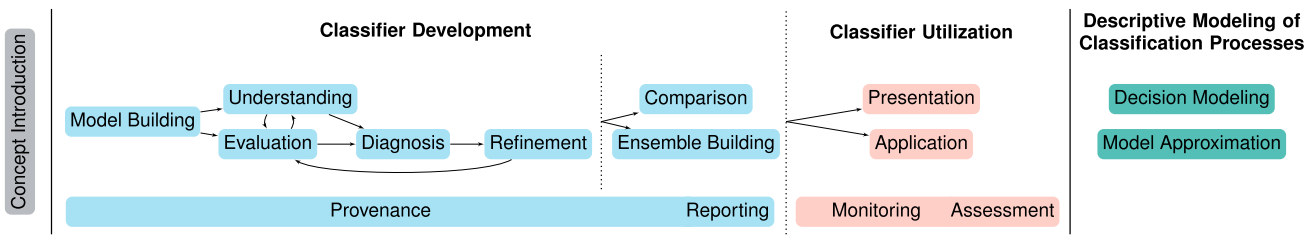


Fig. 5. Structured overview of analysis tasks. Concept Introduction (gray, see Section 3.1) deals with educating novices about decision trees. In predictive scenarios, there are two main stages. First, a classifier is developed (blue, Section 4), then, it is utilized in the target environment (peach, Section 5). In descriptive scenarios (green, Section 6), decision trees are used to describe observed decision processes or black-box classifiers as if they were decision trees. While some tasks like Model Building and Application can be (partially) automated, others including Understanding heavily involve human analysts. The iterative Refinement is central to the interactive modeling process. Provenance and Monitoring run in parallel to the main workflow, whereas Reporting and Assessment summarize the process/performance within a pre-defined period of time.

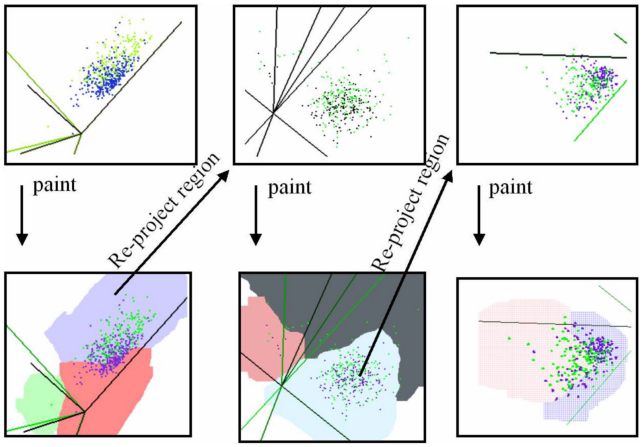


Fig. 7. Workflow and visualization of the *StarClass/PaintingClass* system [43], [44]. During model building, analysts use re-projection and painting of regions at the different levels of the hierarchy to effectively partition classes in the instance space. Image by Teoh and Ma [43, Fig. 6]. Copyright © 2003 Society for Industrial and Applied Mathematics. Reprinted with permission. All rights reserved.

Introduction are clean and focus on the tree structure consisting of decision nodes and leaf nodes. Predominantly, a small decision tree is visualized as node-link diagram. No additional information beyond the attributes considered at decision nodes and applied cutoff values are presented. Sometimes group sizes are added, which explicates the splitting into sub-groups. Restricting visualizations to the bare minimum is in line with the goal of explaining fundamental mechanisms instead of peculiarities of a particular example.

4 Classifier Development

Classifier Development, in machine learning, is driven by a machine training a model based on a dataset. While established algorithms automatically generate a decision tree, the whole process is iterative, including the manual tasks of Evaluation, Diagnosis, and Refinement to name a few. Thus, practical approaches are often semi-automatic. Visual analytics can provide powerful interfaces for interactive machine learning.

4.1 Model Building

Model Building is the process of generating classification trees, either automatically [51], or interactively by an analyst [52]. The only human inputs required by fully-automated, algorithmic approaches are a training dataset and global parameters, such as splitting criteria [53]. Visualization and visual analytics facilitate interactive Model Building by providing rich interfaces [7]. In this case, analysts use the training data and quality measures, but steer the building process through manual intervention. This way, the analyst introduces domain knowledge to the model, which can improve its effectiveness, and can enforce domain-specific requirements [8], [54].

Prior to Model Building, visualization and visual analytics can play an important role in data preparation and exploratory analysis [55]. Here, we assume that the analyst has already prepared a dataset. There are several approaches for manual and interactive Model Building. Successful visualizations enable analysts to keep track of the growing tree by providing an overview, zooming and highlighting functions. Liu and Salvendi [24], as well as van den Elzen and van Wijk [7] present powerful visual analytics systems. At the same time, the main effort in Model Building is to choose nodes to expand and to determine appropriate split attributes and values. Visualizations facilitate the selection of features by showing the distributions of values and by displaying effects of potential partitionings [56], [57], [58].

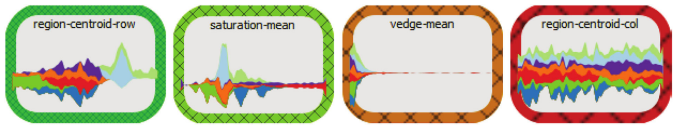


Fig. 8. The *BaobabView* system supports analysts with algorithmic support for selecting split attributes and presents suggestions visually. Border color indicates the goodness of the split as measured by the Gain-ratio. Image by van den Elzen and van Wijk [7, Fig. 6].

A notable example for an interactive visualization targeted at Model Building is the *StarClass/PaintingClass* system by Teoh and Ma [43], [44]. Analysts draw decision boundaries in two-dimensional projections of the instance space to separate classes. An algorithm then builds a decision tree that splits the data according to these decision boundaries. Fig. 7 shows the workflow used to construct a tree based on continuous attributes.

Van den Elzen and van Wijk [7] present *BaobabView*, an extensive visual analytics system offering numerous integrated views and interaction mechanics. Additionally, algorithmic support suggests good options, for example, for split attributes and split values. Fig. 8 shows suggested split attributes including the distributions of values.

4.2 Evaluation

Evaluating the quality of constructed trees is crucial. While the general predictive qualities of classifiers can be evaluated automatically using quality measures and a separate test set of previously unseen data (i.e., cross-validation), more sophisticated *Evaluation* requires the involvement of human analysts [3]. The broad Evaluation of decision trees covers multiple objectives, such as global performance, performance regarding a class of special interest, tree size or structure, and application cost. Human analysts aim at figuring out how well decision trees match these demands. Based on the Evaluation, they decide on further steps, for instance, whether or not refinements are necessary.

The confusion matrix is a simple model-agnostic tool for evaluating the global performance [28], [59]. It can be enriched visually, for example, by mapping the number of instances in each cell to colored areas [7]. Another common visual tool is the ROC plot, which depicts the prediction quality as measured by Recall/Sensitivity and Specificity [28], [60]. These model-agnostic techniques are invaluable complements to visualizations particularly designed for decision trees. Alsallakh *et al.* [61] present an overview of visual approaches for the Evaluation of classifiers. Visualizations targeted at decision trees enable analysts to inspect the tree structure after (automatic) Model Building [24], [62], [63], [64]. Interaction capabilities, like pan-and-zoom, are key to handling large trees. In practice, Evaluation is interwoven tightly with other tasks. For example, in interactive Model Building, the continuous Evaluation of the model is commonplace. Although quality measures play an important role in Evaluation, only a few visualizations present

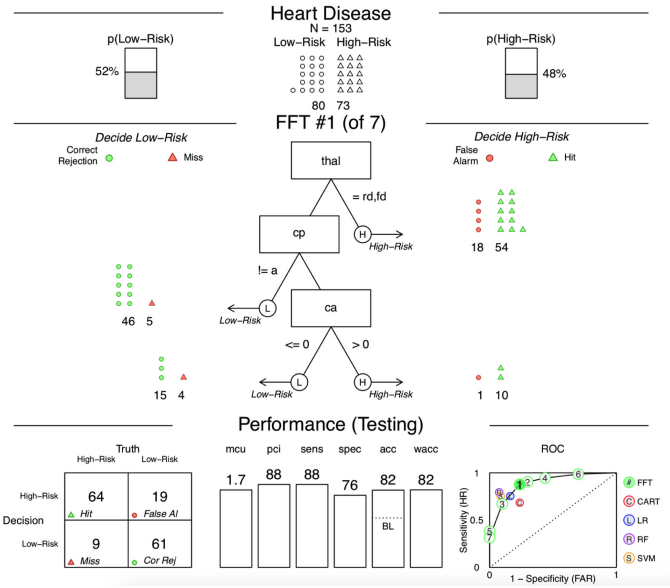


Fig. 9. Compact, yet comprehensive, visualization of numerous performance metrics to evaluate the predictive qualities of a classification tree. Image by Philipps *et al.* [28, Fig. 6] .

quality measures. While Accuracy is the most prominent measure, some visualizations also include some other quality criteria (e.g., [7], [65], [66]). Noteably, the visualization of Philipps *et al.* [28], in Fig. 9, presents a multitude of quality measures, next to a confusion matrix and an ROC plot.

4.3 Understanding

Understanding describes the task of generating an overview of a decision tree and its underlying data, as well as browsing its complete structure. This includes the overall comprehension of the model and its fitness to the classification task at hand [67]. Generally, Understanding is tightly interwoven with other stages like Model Building, Diagnosis, and Refinement. However, in contrast to these stages, Understanding considers a more abstract level without investigating particular data instances or classes.

When investigating classification processes, getting an overview of all decision nodes is essential. Node-link diagrams are the most commonly used technique for depicting the tree structure [24], [68], [69]. They enable analysts to follow the classification process of decision trees [64]. *BaobabView* [7] also implements another approach by representing decision nodes as labels in front of class distributions. Links show the data flowing from parent to child nodes. This type of visualizations is known as pipe diagram [42]. Pipe diagrams provide a complete overview of the classification process, as depicted in Fig. 10. Icicle plots [40] are a more compact option [56].

Understanding is a central step in scientific research. With the adoption of decision trees for data analysis, more and more visualizations of decision trees are published. While the primary visualization task in the publication is Presentation, readers need to fully understand the decision tree in order to comprehend its scientific value. Taking the constraints of publication media into account, most visualizations are minimalistic node-link diagrams [70], [71]. Sometimes nodes are augmented with additional visualizations, such as pie charts or bar charts [72], [73]. In print media, visualizations with multiple views are rare [74].

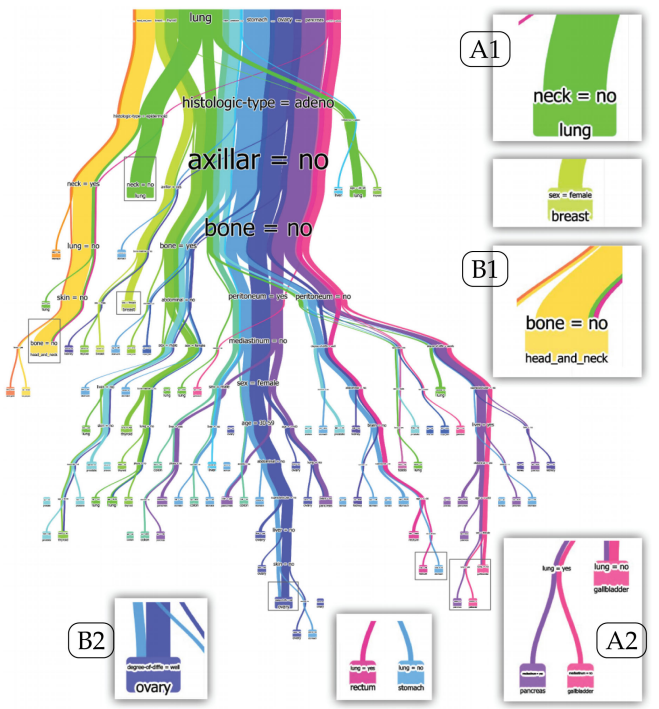


Fig. 10. *BaobabView* system showing the partitioning of instances. Correct predictions are visible (A1, A2) and mis-classifications stand out (B1, B2). Image by van den Elzen and van Wijk [7, Fig. 13].

While these visualization approaches are intended to enable the Understanding of the classification process, multiple or complex decision trees demand more than a single visualization technique to highlight all important aspects. Commonly, to be more expressive, visual analytics systems combine different views and link them through interactions, such as cross-filtering or linking-and-brushing. *RuleMatrix* [75] is an example of such a system for visualizing classifier rules. More recently, Jia *et al.* [76] visualize surrogate decision trees of convolutional neural networks. Similarly, *GBRTVis* [77] integrates views for analyzing gradient boosting regression trees.

4.4 Diagnosis

Diagnosis describes the process of unveiling failures of classifiers and errors in datasets when solving a problem with a classifier model [78]. Problems in trained models or the underlying dataset need to be identified. With the gained knowledge, analysts can correct wrong data labels and anticipate means to improve model performance [35]. Already in interactive Model Building, it is possible to spot errors in the dataset by continuously predicting instances and inspecting mis-classifications [7]. The focus on identifying sources of errors and potential remedies distinguishes Diagnosis from the more general tasks of Evaluation, and Understanding.

In many cases, visualizations are employed to diagnose the decision paths followed to classify instances. Often, an inspection of an individual path unveils problems of the training dataset towards the test dataset and vice versa. Similarly, over-fitting issues can be accounted for by pruning [62]. Mainly, node-link diagrams and pipe diagrams are used to identify wrong paths and find pruning targets

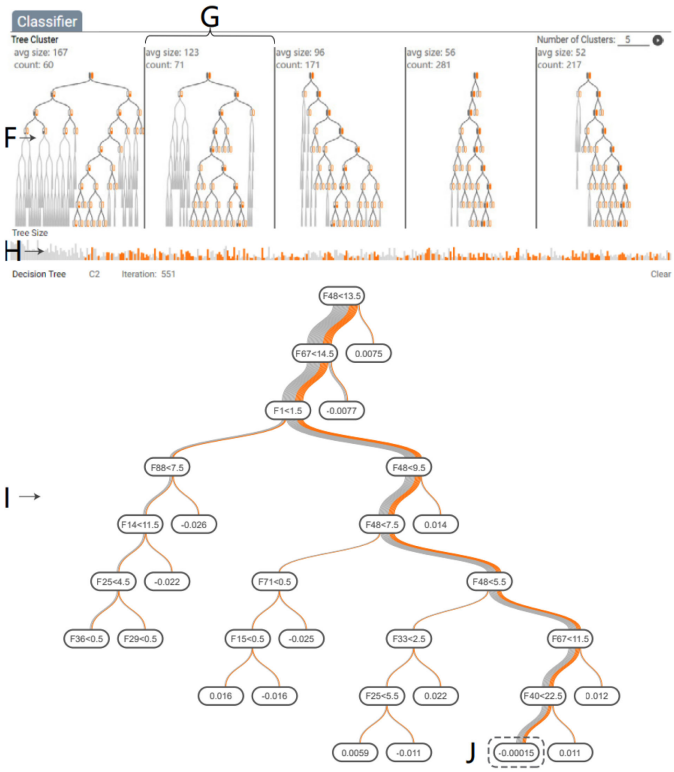


Fig. 11. The *BOOSTVis* system enables analysts to inspect boosted ensembles of decision trees and to diagnose why certain combinations perform better. Image by Liu *et al.* [78, Fig. 1].

[7], [78], [79]. Interaction helps to investigate splits leading to poor partitioning [80]. Enhancing these visualizations with data about the attributes and, for example, their distribution [7], [81], enables the generalization from particular samples to the whole dataset. On the level of single splits, diagnosing errors in attributes is possible [82]. Further, failures in techniques, such as ensemble learning, can be identified [78].

For instance, the *BaobabView* system [7] features various interaction techniques and visualizations. The interactions analysts can perform and the visual mapping to pipe diagrams enable a deep-dive into the inner-workings of decision trees. Further, complementary visualizations (similar to Fig. 8) at the nodes help to learn more about the distribution of instances throughout the decision tree. The *BOOSTVis* system [78], shown in Fig. 11, generalizes this approach to the Diagnosis of boosting tree ensembles combining several decision trees. It utilizes pipe diagrams and node-link diagrams to highlight weak decision trees, and attribute splits. With these visualizations, it is possible to identify attributes that are more heavily used for a split after subsampling.

4.5 Refinement

As available training data and automatic algorithms (with all their assumptions) rarely match perfectly, trained models often need further improvements to solve targeted classification problems. Such Refinement regularly leads to optimized and improved models [35]. Most refinements originate from findings of a previous Diagnosis and tackle specific problems. For instance, pruning decision trees may increase generalization [62]. In most cases, Refinement incorporates human domain knowledge to automatically built models *post hoc*. These refinements steer models in directions that analysts can relate to their mental models [35].

Interactive pruning is a common capability of visual analytics systems [7], [62]. With a more in-depth Diagnosis, it is possible to identify intricate problems, such as a leaf node that should be split, but is not split correctly due to global parameter settings. In such a case, an analyst is able to resolve the problem by selecting a split by custom attributes and split values. Overall, we find few visulizations designed for Refinement. Most likely, this is the case as Refinement is especially closely coupled with Diagnosis and the interaction capabilities required for Model Building. Meanwhile, visual analytics systems excel, as they are build for directly interacting with data and models. The *BaobabView* system [7] mentioned above incorporates such interaction tools. In Fig. 10, highlighted leaf nodes can be pruned as they hardly separate the instances, but over-fit the training dataset.

4.6 Comparison

Comparison supports model selection for a particular classification problem at hand. Analysts may compare the overall performance, or even the performance within subsets of the data, between two or more models. For example, the ROC plot at the bottom right of Fig. 9 can be utilized to compare different classifiers. The decision trees' structures can also be of interest. In this case, the Comparison goes beyond reporting on quality measures for selecting an appropriate model. For instance, comparing how meaningful the splits in one data dimension are enables selecting a model that segments the data in a way that makes most sense for domain experts.

Comparing decision trees visually enables domain experts to manage trade-offs, such as balancing performance and cost. By putting the human in the loop, domain experts can perform comparisons along multiple dimensions, which are hard to automate. With an increasing number of trees to compare, there remains less space for each individual tree. As a result, some visualizations build on compact representations such as treemaps [66] and icicle plots [65]. Comparing trees (and other classifiers) solely based on quality measures is also common [28], [83], [84]. However, often the differences between two trees need to be investigated in detail. Node-link diagrams show the trees' structures more explicitly and are common when only two decision trees are on display [8], [85], [86].

The *TreePOD* system [66] enables analysts to explore a broad set of automatically generated candidate trees. Besides the typical node-link diagram to show tree structures, it offers compact pixel-based treemaps (see Fig. 12). Presented in a small-multiples layout, the treemaps facilitate the comparison of a variety of trees. They convey qualitative aspects of the accuracy (measured by F1-score) and complexity (Size) of Pareto optimal trees and thus provide analysts with the information needed to find a suitable tree from the generation algorithm's parameter space.

4.7 Ensemble Building

Combining models in an ensemble often leads to an ensemble model with strong predictive qualities. In *Ensemble Building*, there are several strategies for combining the outputs of different classification models, as well as a multitude

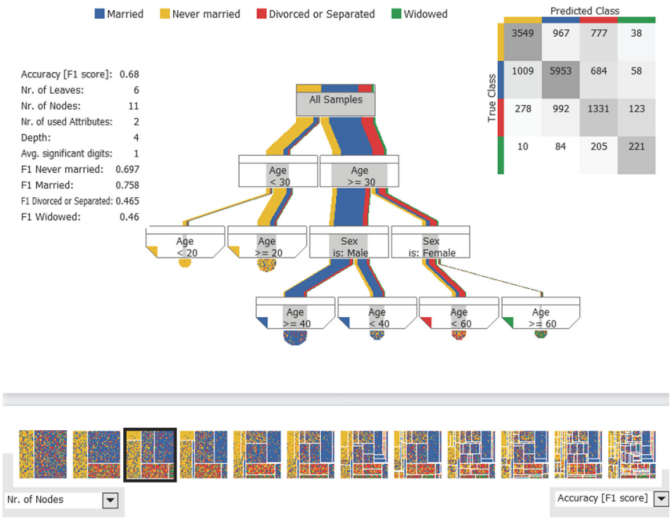


Fig. 12. *TreePOD* system: Analysts can compare several candidate trees in compact pixel-based treemaps, which encode qualitative aspects of Pareto optimal trees. Image by Mühlbacher *et al.* [66, Fig. 1].

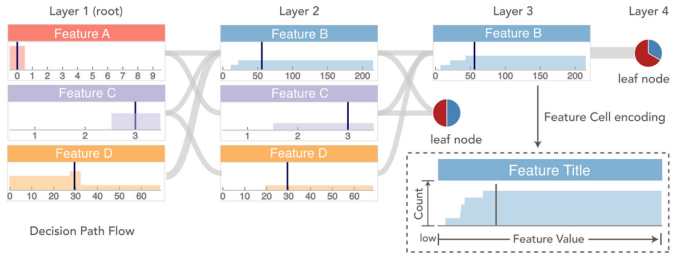


Fig. 13. The *iForest* system represents the *decision path flow* merged into a digraph to let analysts understand and compare all decision paths in random forest models. Image by Zhao *et al.* [87, Fig. 3].

of approaches to generate individual models. For instance, random forests, in their most common form, are ensembles of decision trees trained using different random subsets of data features. Model developers build ensembles using ready-made algorithms or by customizing their outputs. Visualizations can support the ensemble building by showing details of the set of models constituting the ensemble.

Visualization and improved interpretation can support the Diagnosis and Refinement of ensembles [78] (see also Fig. 11). For instance, analysts visualize and compare feature importance on different trees to help feature engineering [88]. As with the Comparison task, the large number of trees within the ensemble is challenging. Thus, abstracting from individual trees and focusing on the ensemble's prediction can be a reasonable strategy for visualization [89]. The *iForest* system [87] offers analysts visualizations to understand and compare decision paths in random forests (see Fig. 13). Analysts use the system to calibrate their trust in an ensemble's predictions by inspecting how the ensemble works, and analyzing training data that is most similar to new inputs. Solving a problem using similar instances is known as case-based reasoning [90].

4.8 Provenance and Reporting

Provenance captures the Classifier Development process over time. Resulting timelines can be especially useful for analysts to track progress, resume, and return to earlier model states. However, we find few visualizations of provenance, except for a table of quality measures tracking recent changes [91], and some visualizations contrasting training performance of novel techniques against established methods [92]. For very large datasets and more complex ensemble models, supervision of the training process can be useful, for example, via a line chart showing the classifiers prediction quality over training time [92, Fig. 1]). Looking beyond decision trees, there are more general approaches for visualizing changes in hierarchical structures [93]. A notable feature of decision trees is that the tree itself is a representation of the training process, as deeper nodes are expanded later in the building process. Utilizing a hierarchical visualization like a node-link diagram, therefore, enables analysts to track the progress of an automated algorithm.

By contrast to Provenance, *Reporting* generates an aggregated summary of the process up to a specific point in time. Reports can be used to update managers on recent changes or to aid developers in resuming. For instance, a report may include the model structure, performance characteristics, and a list of important issues that were diagnosed and fixed.

Similar to Provenance, Reporting is not in the focus of research. In our sample of publications, we could only find one visualization designed for Reporting on the Classifier Development process. Fig. 9 on page 6 shows the visualization developed by Phillips *et al.* [28] for Reporting on a developed classifier. The visualization combines a description of the problem at the top, details on the tree in the center, and quality measures (also in comparison to alternative classifiers) at the bottom.

5 Classifier Utilization

Having constructed a classifier, automatically or interactively, there are a number of tasks regarding its utilization [7]. By contrast to Classifier Development, now the decision trees are fixed and ready for application. Adapting to the target environment and not the training setup is a particular challenge (see also [3], [94]). Except for the Presentation of decision trees, this area attracted much less interest from researchers than Classifier Development in the past.

5.1 Presentation

Presentation is often stated as one of the main tasks for visualization [95]. In the case of decision trees, presentation goals range from the description of a tree's structure [64] via lessons learned in the construction to presenting potential improvements of applying a new classifier compared to the *status quo*. One particular use case for Presentation is the visualization of decision trees in scientific publications that make the classifier explicit for readers [96], [97], [98].

Most of the visualizations for Presentation that we find target wide audiences. As a result, they apply the node-link diagrams, which are well known. However, there are more compact summaries as, for instance, by Kreiser *et al.* [100]. They target a small audience and develop a special encoding for their particular tree. A more typical example is given by Tam *et al.* [99], who present the decision tree they created for distinguishing between facial expressions. As shown in Fig. 14, they highlight class distributions at nodes visually by using color. Additionally, they explicate the rather complex rules of the decision nodes.

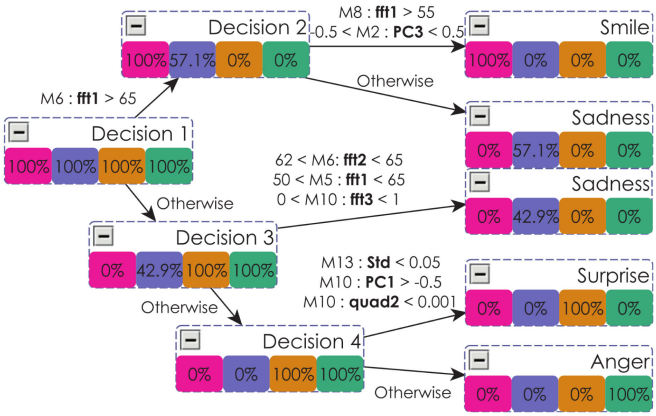


Fig. 14. Presentation of a manually created decision tree for discriminating between facial expressions. Class distributions and decision rules are displayed in nodes and next to links, respectively. Image by Tam *et al.* [99, Fig. 9].

5.2 Application

While classifiers are applied automatically from a machine learning perspective, there are situations in which it is beneficial to involve humans in the *Application* of a decision tree. The (manual) Application can be relevant when access to computing devices cannot be guaranteed or available time is not sufficient to input measurements, for example, in the case of emergencies. But even in case the execution of the algorithm is performed automatically, involving humans may be necessary to establish trust in predictions in general, and to provide a reasoning for individual predictions. Domain experts often know how a particular tree takes specific constraints posed by their tasks and domain into account, and can estimate how well the training dataset reflects the population in a particular application.

There are only few visualizations in our sample that are tailored to Application. One such visualization is shown in Fig. 15. It depicts the decision tree of the *START* triage procedure [33] as an indented list. The visualization is intended to be the size of a credit card and usable in the field. Starting from the top, emergency responders can follow the procedure to quickly identify those people who need immediate treatment.

5.3 Monitoring and Assessment

Checking whether a classifier works in practical application can be done in two ways, either by the *Assessment* of performance up to a specific point in time, or by *Monitoring* the classifier continuously. Both tasks aim at rating how well a classifier extrapolates beyond the training environment to the real application environment [94]. They consume the training and application data, next to the classifier as inputs. Outputs include reports for managers in the case of Assessment and permanent feedback to operators, which can be used for spotting problems, in the case of Monitoring.

As with Provenance and Reporting, we rarely find visualizations for Monitoring and Assessment. The examples we find are not from productive utilization of deployed decision trees, but on validation datasets [102]. For example, a second node-link diagram of the same tree showing class proportions based on the validation data can be presented next to the diagram based on the training data [96]. A more tailored visualization shows class proportions in leaf nodes in, both,

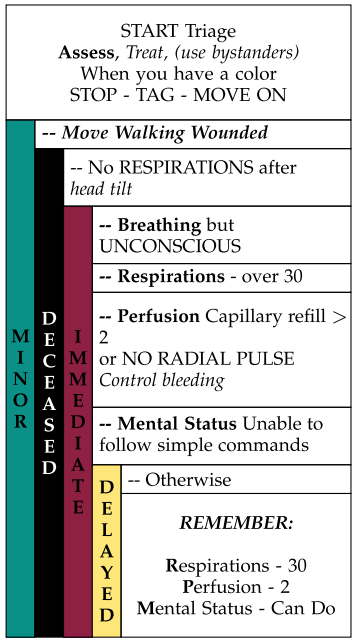


Fig. 15. *START* triage decision tree [33]. Reproduced from Critical Illness and Trauma Foundation, Inc. [101].

training and validation data [79]. In our sample, we only find one visual analytics tool for building rules covering the Monitoring task shown in Fig. 16 [103], despite the fact that dataset shift [104], [105] and other limitations to generalization demand for the Assessment and Monitoring of classifier models in practice.

6 Descriptive Modeling of Classification Processes

In contrast to the previous sections, which deal with Classifier Development and Classifier Utilization for prediction, this section is about investigating classification processes by means of modeling them as decision trees. In this context, we identify two main tasks. Decision Modeling deals with gaining insight into postulated or observed decision-making. By contrast, Model Approximation focuses on explaining previously constructed opaque classifier models.

6.1 Decision Modeling

Describing observed decision-making processes by decision trees is the goal of *Decision Modeling*. It aims at matching the

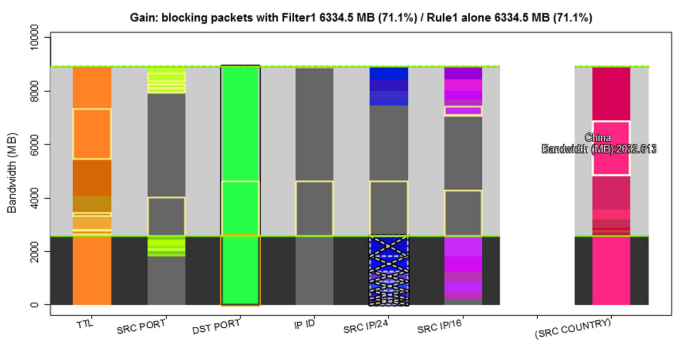


Fig. 16. View inspired by parrallel coordinates for monitoring the applicability of filter rules to counter attacks on computer networks. Image by Aupetit *et al.* [103, Fig. 5].

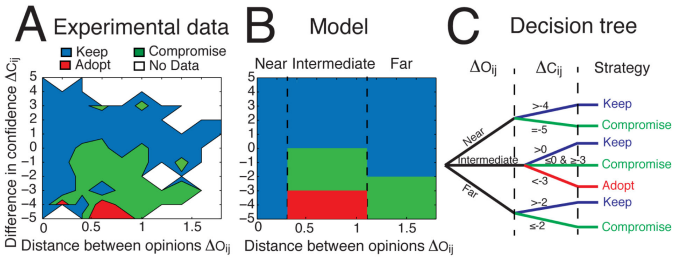


Fig. 17. A descriptive decision tree modeling the conditions under which a person is willing to change an opinion when confronted with another. The decision tree model is presented as a treemap (B) and a node-link diagram (C). Image by Moussaïd *et al.*- [74, Fig. 3].

outcomes of decisions with a suitable decision tree, and investigating the observed process descriptively *as if* the underlying decision process was the execution of a decision tree. By contrast to prescriptive usages of decision trees, primarily descriptive decision trees aim to enable insights and are not intended to be applied.

Examples of descriptive trees explain how British courts decide whether to make a punitive bail decision [106], and how people decide whether or not to forgive another person for an offense committed during social interactions [107]. Visualizations of proposed descriptive trees are typically simple node-link diagrams that serve to illustrate the steps required by the decision algorithm and to explicate the threshold values for selecting branches [106]. In that respect, the presentations do not differ from those designed for Presentation (discussed in Section 5.1). However, some visualizations for descriptive modeling compare multiple tree variants, for example, created by systematically varying the exit structure at decision nodes [107], [108], [109].

Moussaïd *et al.* [74] employ a more advanced visualization that provides two alternative views on their descriptive model (see Fig. 17). They investigate when people are willing to change their opinion by using a decision tree that captures two dimensions: i) How different is one's currently held opinion from the another? ii) Is the other person more confident? The relation between the data and the model is illustrated by a treemap an a node-link diagram.

6.2 Model Approximation

Explaining black-box machine learning models has been a very prominent task in recent years [67]. Explainable Artificial Intelligence (XAI) discusses approaches to make complex and opaque models more interpretable, while interactive machine learning tries to provide solutions for analysts to apply domain knowledge to models and refine existing classifiers. Both can be seen as methods to address the problem of not easily accessible classifiers. One common method to explain black-box models and to make a complex model interpretable is *Model Approximation* by simplifying their internal processes to surrogate decision trees. These surrogates are utilized as a proxy for the Understanding, Diagnosis, Comparison, and Evaluation of opaque models. In addition, such surrogates can also be used as an interaction interface for the target models, enabling the Refinement of opaque models through visual analytics.

Surrogate decision tree models are some of the most prominent approaches for increasing the interpretability of neural networks [87]. However, finding the appropriate degree of simplification remains a challenge [117], as the decision trees should approximate the process and performance of the opaque target models, while remaining interpretable. Approaches for visualizing classification processes in neural networks using surrogate decision trees range from using hashing neural networks [118], to the analysis of convolutional neural networks [76], and gradient boosting regression trees [77].

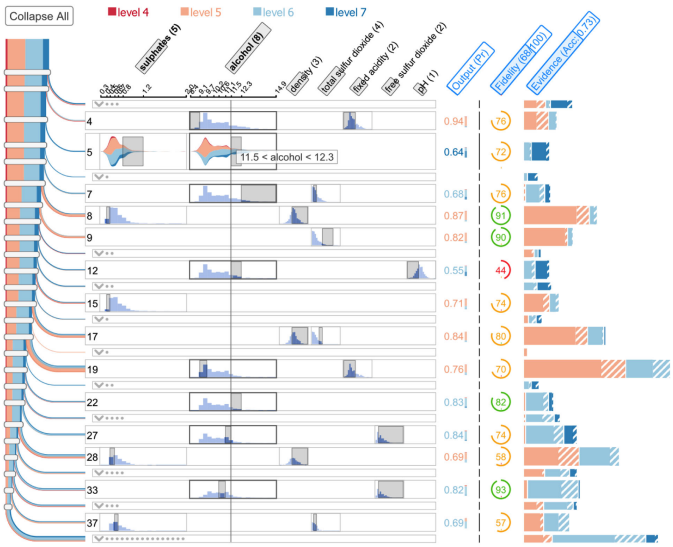


Fig. 18. Surrogate approximation of a neural network model using rule-based explanations in the *RuleMatrix* system. Rows represent individual rules and columns depict involved attributes. The pipe diagram on the left explicates how ordered lists of rules coincide with the branching structure of decision trees. Image by Ming *et al.* [75, Fig. 1b].

In contrast to visualizing surrogate models as simplifications, the *RuleMatrix* system [75] approximates complex models (here neural networks) by a list of classification rules. It enables analysts to interact with the visual interface to explore, as well as refine the opaque target model. Fig. 18 shows the interactive visual interface. In each row, it displays one classification rule, which is composed of different attributes, depicted as columns. This interactive approach is based on user-defined rule filters to adjust the application of the underlying neural network to boost its performance. Such interactive feedback is essential to enable model steering. However, integrating the feedback into the opaque model is still an open challenge, and an opportunity for future research.

7 The Role of Visualization and Visual Analytics

Since visualization gained interest in the early 1990s the number of visualizations of decision trees increased over time (see also Fig. 4 on page 3). Especially in machine learning, visualizations facilitate tasks such as Model Building, Evaluation and Comparison. While the introduction of new techniques (e.g., [41], [114]) spurred novel visualizations of decision trees [119], [120], node-link diagrams remain the most common visual design by far. Only interactive visual analytics systems regularly offer multiple views on tree structures [65], [66], [79]. Going beyond tree structures and showing more detailed visualizations as well is less common than we expected [7], [74], [120].

TABLE 1
Cross-Tabulation of Tasks and Visual Designs Employed in the 152 Surveyed Publications

	Visual design of tree										Visual design of further components																
Task	Circle packed layout	Flow chart [110]	Icicle plot [40]	(Indented) list	Node-link diagram	Parallel coordinates plot [41]	Pipe diagram [42]	Sunburst diagram [111]	Treemap [18]	Other	Area chart	Bar chart/histogram	Bar chart (stacked)	Box plot [112]	Dot plot [113]	Icon array	Image-like	Kernel density plot	Line chart	Matrix view	Parallel coordinates plot [41]	Pie chart	Pixel-based view [114]	Scatter plot	Star coordinates plot [115]	Stream graph [116]	Other	*Total*
Concept Introduction		1	1	3	35			1	5	3	1	1							1	2	1	2	2	1				*36*
Model Building		1	4	6	20		3	1	2	3	1	4	6		1		1		4	6	3	1	7	10	1	2	3	*36*
Evaluation			5	5	17		4			3		8	6		1	1		1	6	3	1	1	4	12		2	2	*34*
Understanding		2	5	8	28	2	4	1	1	3	1	13	5				1		5	4	3	7	6	15	1	1	4	*46*
Diagnosis		1	2	2	15		5			2	1	6	2		1				2	3	2	2	2	8		2	1	*21*
Refinement		1		1	5		1		1	1	1	2	2	1					2	3		1		1		1	1	*8*
Comparison			4	2	29		3		3	3	2	11	3	1	1	1		1	9	4	3		5	14		1	2	*43*
Ensemble Building			1	2	4		1		1	2	1	4	2						5	1	1	2	1	4		1	2	*10*
Provenance					2								1						1	1	1			1				*3*
Reporting					4		1					1	1			1		1	1	2	1			5				*5*
Presentation		2	1	1	52	1	3		3	1	2	9	3	1		1		1	4	5		2	5	11		1		*55*
Application			1	1	5	1	1			1		1	2			1		1		1		1		2				*6*
Assessment			1	2	3		3				1	4	1						3	1		2	1	2				*10*
Monitoring										1			1															*1*
Decision Modeling					10	1			1			1											1	1				*11*
Model Approximation				1	2		2		1		1	1	1						2	1	1	2		2				*6*
Total	*0*	*5*	*10*	*14*	*110*	*2*	*6*	*2*	*9*	*10*	*4*	*22*	*10*	*1*	*2*	*1*	*1*	*1*	*14*	*13*	*7*	*9*	*16*	*33*	*1*	*2*	*6*	

Totals count unique publications in each row/column. The node-link diagram is the most prominent visual representation of the tree structure across all tasks. Standard visualizations like bar charts, line charts and scatter plottes are most commonly used to augment the tree structure with additional information.

Interactive visualizations play an increasing role across many tasks. In Model Building, they support the effective involvement of domain experts, for instance, by defining splits [58], [99], [120]. Furthermore, highlighting single decision paths can ease Understanding and Diagnosis [121], [122]. In Classifier Development, tasks are particularly closely interwoven. For example, direct interaction with visual analytics systems can aid in diagnosing a decision tree and immediately applying a refinement. More generally, the very nature of machine learning is an iterative process. Effectively going back and forth between tasks calls for well integrated visualizations [7], [66].

The visual designs used for the different tasks is depicted in Table 1. As is also visible from Fig. 6 on page 4, some tasks are more common than others. As noted above, the node-link diagram is by far the most prominently used design for representing the structure of the decision tree. By contrast, we did not observe any visualization that uses a circle packed layout. While there is a diverse mixture of designs for further components, the well-known bar charts, line charts, and scatter plots are commonplace.

To our surprise, there is very little variation in quality measures displayed with decision trees. Basically, we find a small number of quality measures quantifying four different aspects: i) Prediction quality, including Accuracy, AUC, Balanced accuracy, F1-score, and Lift, ii) Aspects of prediction quality, like Precision, Recall/Sensitivity, and Specificity, iii) Group (im)purity, Gini-index and iv) Tree structure, including Size, Mean cues used, and Frugality. Overall, the integration of visualizations and numeric quality measures is limited. Basic Accuracy dominates all other quality measures, but most visualizations do not show any measures. Table 2 cross-tabulates tasks and displayed quality measures, but except for the general lack of displaying quality measures we do not identify any pattern.

Similarly, rule-based classification is rather a niche topic in our sample [75], [103], [117], [119], [123]. One possible explanation for this finding is that, as discussed above, sets of classification rules can be transformed to decision trees. Decision trees may provide more structure and thus be easier to visualize and comprehend in many cases [124], [125]. In Fig. 18, for example, the left-hand side explicitly shows a branching tree structure, despite the fact that the *RuleMatrix* system [75] is designed for a list of rules.

In Evaluation, many visual designs are agnostic to the type of classification model. Employing versatile visualizations enables comparisons across model types, which are relevant as the performance of decision trees often needs to be judged in comparison to other candidate models such as neural networks. As a result, few visualizations specialize on the evaluation of decision trees, for example, by highlighting split values. This observation resonates with the absence of quality measures in most visualizations. More examples in the direction of cross-type comparison come from Concept Introduction, as, for instance, shown in Fig. 19. Such comparisons across model types are not unique to classification. For instance, Rudin and Carlson [30] contrast regression trees with other regression techniques.

8 Open Questions and Opportunities

As mentioned in the previous section, we are surprised by how rarely quality measures are part of visualizations of decision trees. But even among those visualizations that

TABLE 2
Cross-Tabulation of Tasks and Quality Measures Displayed in the 152 Publications We Surveyed

Perspective	Column				Row				Marginal/Mixture					n.a.			
Task \ Quality Measure	AUC	Recall/Sensitivity	Specificity	Weighted accuracy	Gain-ratio	Gini-index	Information gain	Precision	Accuracy	F1-score	G-means	Lift	χ^2	Frugality	Mean cues used	Other	*Total*
Concept Introduction	1	1	1			1	1					1		1		2	*4*
Model Building	1	2			1	1	4	2	4	1				1		3	*11*
Evaluation	3	4	4	1		1	2	1	11	1	1		1	1	1	4	*18*
Understanding	2	3	1		2	3	2	3	14	1		1	3			6	*24*
Diagnosis		3			1		2	3	4	1			1			2	*7*
Refinement		2			1		2	2	3	1							*4*
Comparison	3	4	3	1			1	1	11		1	1	2	2	1	4	*18*
Ensemble Building	1	1				1	1	1	5			1	1			2	*8*
Provenance	1															1	*1*
Reporting		3	2	1				1	1					1	1		*3*
Presentation		5	3	1		1	1	2	12	1	1		1	1	1	5	*18*
Application		1	1	1		1			1					1	1		*2*
Assessment		2			1	1		2	2				1			2	*5*
Monitoring																	*0*
Decision Modeling									3							1	*4*
Model Approximation		2						2	1								*4*
Total	*6*	*6*	*4*	*1*	*2*	*4*	*5*	*3*	*26*	*1*	*1*	*2*	*3*	*2*	*1*	*12*	

Totals count unique publications in each row/column. Clearly, Accuracy is the most prominently displayed measure of quality. However, compared to the size of our sample quality measures are rarely displayed. There is no relationship apparent between tasks and the quality measures displayed. Quality measures are sorted according to inherent perspectives [3].

show quality measures only a tiny fraction shows multiple quality measures capturing different aspects of quality [7], [28], [66]. One exception to this is the ROC plot, which is a common and model-agnostic visualization showing Recall/Sensitivity over Specificity. Still, the question remains, *how can quality measures be integrated in visualizations of decision trees?* Especially for visual analytics and interactive machine learning, we expect that integrating quality measures offers analysts additional views on the trees. Just like linked views advanced data visualization a closer integration of quality measures and visualizations will advance visual analytics systems for interactive machine learning, and decision trees. They likely also will increase the acceptance of visual approaches within the machine learning community.

Likewise, we see a potential for utilizing algorithms developed in similar domains, for instance, integrating general tree comparison algorithms [127], [128] in visualizations for decision tree comparison. These algorithms work with hierarchically organized data in general and are likely to facilitate the comparison between large decision trees as well. Especially the Comparison of multiple trees and

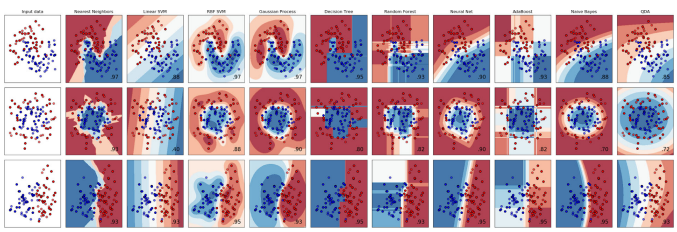


Fig. 19. Decision boundaries of ten types of classifiers across three datasets. The sixth column (center) depicts a decision tree with typical sharp boundaries in parallel to considered attributes. Source: https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html (accessed Jan. 2020, cf. [126]).

Ensemble Building are complex and difficult problems for which no standard visualization techniques have emerged. More generally, one may ask: *How can visualization algorithms and visual analytics systems provide better default layouts and assistance?* The tighter integration of mathematical and algorithmic approaches and visualization will not only help people, who are not visualization experts, to come up with better visualizations, but also analysts to use visual analytics systems more efficiently.

These potentials notwithstanding, there are also opportunities for integrating humans more closely in Classifier Development. Recent research shows that small decision trees perform competitively in noisy environments [28], [38], [39], [50], [108], [129]. Small trees, in particular, constitute a special opportunity for visualization and visual analytics as problems with visualizing large trees can be avoided and integrating analysts' domain knowledge becomes increasingly important [8]. This raises the question: *How can visualizations and visual analytics systems facilitate the externalization of domain knowledge?* The whole workflow will not only benefit from the externalization of domain knowledge and improved communication, but ultimately yield better decision trees and more accurate classifications. However, to date empirical studies on interactive visual Classifier Development are rare and usually do not involve domain experts [6], [8], [54], [56], [130]. Still, interactively constructed decision trees provide an alternative to deep learning classifiers, especially in scenarios that demand for the positive properties of decision trees summarized in Section 2.

Particularly to laypeople, who only get in touch with decision trees in basic Concept Introduction or by attending a presentation, the node-link diagram is omnipresent. But also analysts working with standard software default to simplistic node-link diagrams that only visualize the tree

structure without additional information, such as distributions of values or split qualities. *How can rich visualizations and visual analytics systems for dealing with decision trees become more accessible?* Spill-over of design knowledge from the visualization community will lead to more informative and aesthetic visualizations. Meanwhile, which visual designs are most accessible needs to be answered alongside the technical questions. Although there are some studies comparing visualizations of hierarchical data (e.g., [131], [132]), a number of comparisons between automated algorithms and interactive systems [6], [8], [43], [44], [52], [54], [56], [99], [120], [130], [133], [134], [135], [136], [137], as well as evaluations of individual interactive systems [66], [75], [76], [77], [87], [117], [123], [138], [139], [140], only few empirical experiments target alternative visual designs of decision trees [125], [141] (see also Fig. 4 on page 3). Hence there is an obvious need for comparative evaluations between different designs. Without such empirical investigations, it is difficult to formulate and substantiate design guidelines.

In resemblance to the prominence of the node-link diagram, we do not find major differences between the visualizations aimed at different tasks. Clearly, visual analytics systems covering large parts of the iterative Classifier Development offer a diverse set of interaction capabilities and more advanced visual displays [7], [66] than static visualizations for Presentation. Still, often multiple tasks are tackled from one general-purpose visualization, not a number of specialized views. Hence: *How can visual analytics systems integrate visualizations tailored more closely to the steps in Classifier Development?* For example, Evaluation and Diagnosis demand for distinct levels of detail, which may be integrated via semantic zooming. Visualizations for Provenance and Reporting will aid in transitioning between tasks and facilitate the supervision of Classifier Development efforts. One particular challenge will be to coordinate the tailored views in such a way that their potential benefits outweigh the friction induced by switching between views.

Beyond the Classifier Development workflow, there appears to be an exceptionally vast and empty space for creative utilization and future innovation. Both, Classifier Utilization and Descriptive Modeling of Classification Processes, attracted little attention, except for Presentation. It is much more difficult to deploy and evaluate visualizations that are aimed at actual utilization in the field. Those domains that spend the effort to evaluate decision aids and tools at a relevant scale, like medicine, tend to be conservative and, understandably, not too open for being a testbed for late-breaking visual designs and visual analytics systems. In the area of Descriptive Modeling of Classification Processes we can envision that the descriptive use of Decision Modeling will benefit from visualizations and visual analytics systems that spill over from Model Approximation, which started to attract researchers only recently in the domain of Explainable Artificial Intelligence (XAI). Hence, we ask: *How can visualizations and visual analytics systems for utilization tasks be developed and evaluated?* Adopting visualizations and visual analytics systems in real-world contexts will be one potential response to societal demands for an accountable and transparent decision-making when delegated to (partially) automated classification algorithms.

Taken together, these open questions highlight that even in a domain that has been researched for decades, there remain quite fundamental gaps, which align astonishingly well with more general visualization research. To begin with, the transparent depiction of diverse quality measures as well as the provision of good defaults and assistance are challenging. Our review clearly highlights that there is a lack of empirical research that prevents the proposition of substantiated guidelines. At the same time, a study of observations made by practitioners in their daily work would complement our work by going beyond our analysis of how scientists from other domains present their results (see Section 5.1). Offering workflows that are accessible to different stakeholders goes well beyond the development of typical research prototypes. Despite the prominent discussions on the task-dependency of visualizations, dedicated tailoring to task demands is not apparent in our sample. Especially when it comes to the utilization of classifiers in practice, there is little research. In consequence, practitioners are left alone in choosing the right visual tool that supports their requirements.

An obvious next step is to extend our survey to the more general field of visualization to support classification based on machine learning. While visual designs can be expected to be different, we expect our set of tasks as well as our focus on performance measures to closely match the workflows and demands across modeling approaches. Surveys centered around visualization, like ours, will complement surveys that are structured along high-level tasks and questions [9], [23], [36]. Regarding potential results, on the one hand, we would expect to find lacks of comparative evaluations of visual designs and visual analytics systems. On the other hand, more complicated techniques may attract expert audiences that are capable of working with richer visualizations, such that basic visualizations comparable to node-link diagrams are less prominent. Beyond classification, for instance, the use of visualization in regression modeling is an active field of research and a promising candidate for a similar survey.

9 Conclusion

In this survey, we compiled a broad overview on available visualizations of decision trees and rule-based classifiers from a task-based perspective. The long history of decision trees and their close relationship to visualization renders them a perfect class of prediction models for investigating the differences between visualizations designed for distinct tasks. We surveyed eight main sources of publications covering major visualization venues and extend our sample based on references, recommendations and additional keyword searches. In total, our sample consists of 152 publications dating back to 1986.

To our surprise, visualization designs are rather general and homogeneous across tasks, instead of being highly specialized and tailored to particular tasks. By contrast, there is a big difference between visualizations designed for different audiences. In visualizations designed for audiences of laypeople, the node-link diagram is omnipresent. Machine learning model developers, on the other hand, often are confronted with a number of complementary designs organized

in linked views. But even in visualizations designed for model developers, quality measures, except for Accuracy, are rarely presented in a visual fashion, and alternative indicators of model quality are mostly lacking.

In consequence, we see substantial opportunities for integrating visualizations more closely with algorithms and mathematical measures of model quality. At the same time, increasing interaction capabilities will lead to an improved accessibility and the utilization of domain experts' knowledge in model construction. The lack of visualizations for Classifier Utilization and Decision Modeling uncovers that there still is a considerable gap between research and practical application in areas that are more distant to visualization researchers daily business. Finally, the question remains, why do alternative (tailored) visual designs not match the ubiquitous use of node-link diagrams?

Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the projects A03 of TRR 161 Project-ID 251654672 and Knowledge Generation in Visual Analytics Project-ID 350399414.

References

[1] D. Sacha *et al.*, "Human-centered machine learning through interactive visualization: Review and open challenges," in *Proc. Eur. Symp. Artif. Neural Netw. Comput. Intell. Mach. Learn.*, 2016, pp. 641–646.

[2] D. Sacha, M. Kraus, D. A. Keim, and M. Chen, "VIS4ML: An ontology for visual analytics assisted machine learning," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 385–395, Jan. 2019.

[3] H. Neth, N. Gradwohl, D. Streeb, D. A. Keim, and W. Gaissmaier, "Perspectives on the 2×2 matrix: Solving semantically distinct problems based on a shared structure of binary contingencies," *Front. Psychol.: Cogn.*, to be published, doi: 10.3389/fpsyg.2020.567817.

[4] D. A. Keim, J. Kohlhammer, G. P. Ellis, and F. Mansmann, *Mastering the Information Age Solving Problems With Visual Analytics*. Goslar, Germany: Eurographics Association, 2010.

[5] N. V. Andrienko *et al.*, "Viewing visual analytics as model building," *Comput. Graph. Forum*, vol. 37, no. 6, pp. 275–299, 2018.

[6] Y. Liu and G. Salvendy, "Design and evaluation of visualization support to facilitate decision trees classification," *Int. J. Man-Mach. Stud.*, vol. 65, no. 2, pp. 95–110, 2007.

[7] S. van den Elzen and J. J. van Wijk, "BaobabView: Interactive construction and analysis of decision trees," in *Proc. IEEE Conf. Vis. Analytics Sci. Technol.*, 2011, pp. 151–160.

[8] G. K. L. Tam, V. Kothari, and M. Chen, "An analysis of machine- and human-analytics in classification," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 71–80, Jan. 2017.

[9] L. Jiang, S. Liu, and C. Chen, "Recent research advances on interactive machine learning," *J. Vis.*, vol. 22, no. 2, pp. 401–417, 2019.

[10] L. Breiman, "Classification and regression trees," Wadsworth Internat. Group, Belmont, Calif., 1984.

[11] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 1944.

[12] K. Steele and H. O. Stefánsson, "Decision theory," in *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford Univ., 2016. [Online]. Available: https://plato.stanford.edu/archives/win2016/entries/decision-theory/

[13] B. Cong, S. Y. Shin, and A. R. Salehnia, "Expert systems and expert system languages," *J. Comput. Sci. Colleges*, vol. 7, no. 5, pp. 1–9, 1991.

[14] J. C. Giarratano and G. D. Riley, *Expert Systems: Principles and Programming*. Boston, MA, USA: Course Technol. Inc., 2004.

[15] D. L. Keefer, C. W. Kirkwood, and J. L. Corner, "Summary of decision analysis applications in the operations research literature, 1990– 2001," *Decis. Anal.*, vol. 1, no. 1, pp. 1–67, 2004.

[16] S. Afzal, R. Maciejewski, and D. S. Ebert, "Visual analytics decision support environment for epidemic modeling and response evaluation," in *Proc. IEEE Conf. Vis. Analytics Sci. Technol.*, 2011, pp. 191–200.

[17] G. Gigerenzer and U. Hoffrage, "How to improve Bayesian reasoning without instruction: Frequency formats," *Psychol. Rev.*, vol. 102, no. 4, pp. 684–704, 1995.

[18] B. Shneiderman, "Tree visualization with tree-maps: 2-D space-filling approach," *Trans. Graph.*, vol. 11, no. 1, pp. 92–99, 1992.

[19] M. Graham and J. Kennedy, "A survey of multiple tree visualisation," *Inf. Vis.*, vol. 9, no. 4, pp. 235–252, 2010.

[20] H.-J. Schulz, S. Hadlak, and H. Schumann, "The design space of implicit hierarchy visualization: A survey," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 4, pp. 393–411, Apr. 2011.

[21] H.-J. Schulz, "Treevis.net: A tree visualization reference," *IEEE Comput. Graphics Appl.*, vol. 31, no. 6, pp. 11–15, Nov./Dec. 2011.

[22] G. Li, M. Tian, Q. Xu, W. Zhang, M. J. McGuffin, and X. Yuan, "GoTree: A grammar of tree visualizations," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–13.

[23] A. Endert *et al.*, "The state of the art in integrating machine learning into visual analytics," *Comput. Graph. Forum*, vol. 36, no. 8, pp. 458–486, 2017.

[24] Y. Liu and G. Salvendy, "Visualization support to better comprehend and improve decision tree classification modelling process: A survey and appraisal," *Theor. Issues Ergonom. Sci.*, vol. 8, no. 1, pp. 63–92, 2007.

[25] J. Fürnkranz, D. Gamberger, and N. Lavrac, *Foundations of Rule Learning*. Berlin, Germany: Springer, 2012.

[26] S. B. Kotsiantis, "Decision trees: A recent overview," *Artif. Intell. Rev.*, vol. 39, no. 4, pp. 261–283, 2013.

[27] L. Rokach and O. Maimon, *Data Mining With Decision Trees: Theory and Applications*. Singapore: World Scientific, 2015.

[28] N. D. Phillips, H. Neth, J. K. Woike, and W. Gaissmaier, "FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees," *Judgment Decis. Making*, vol. 12, no. 4, pp. 344–368, 2017.

[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[30] C. Rudin and D. Carlson, "The secrets of machine learning: Ten things you wish you had known earlier to be more effective at data analysis," *Tuts. Operat. Res.*, 2019, doi: 10.1287/educ.2019.0200.

[31] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nat. Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019.

[32] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim, "The role of uncertainty, awareness, and trust in visual analytics," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 240–249, Jan. 2016.

[33] M. Benson, K. L. Koenig, and C. H. Schultz, "Disaster triage: START, then SAVE—A new method of dynamic triage for victims of a catastrophic earthquake," *Prehospital Disaster Med.*, vol. 11, no. 2, pp. 117–124, 1996.

[34] M. K. Dhami and P. Ayton, "Bailing and jailing the fast and frugal way," *J. Behav. Decis. Making*, vol. 14, no. 2, pp. 141–168, 2001.

[35] S. Liu, X. Wang, M. Liu, and J. Zhu, "Towards better analysis of machine learning models: A visual analytics perspective," *Vis. Informat.*, vol. 1, no. 1, pp. 48–56, 2017.

[36] F. Hohman, M. Kahng, R. Pienta, and D. H. Chau, "Visual analytics in deep learning: An interrogative survey for the next frontiers," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 8, pp. 2674–2693, Aug. 2019.

[37] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *SIGMOD Rec.*, vol. 22, no. 2, pp. 207–216, 1993.

[38] D. Bertsimas, A. Chang, and C. Rudin, "ORC: Ordered rules for classification a discrete optimization approach to associative classification," 2011. [Online]. Available: http://hdl.handle.net/1721.1/73166

[39] C. Rudin and Ş. Ertekin, "Learning customized and optimized lists of rules with mathematical programming," *Math. Program. Comput.*, vol. 10, no. 4, pp. 659–702, 2018.

[40] J. B. Kruskal and J. M. Landwehr, "Icicle plots: Better displays for hierarchical clustering," *Amer. Statistician*, vol. 37, no. 2, pp. 162–168, 1983.

[41] A. Inselberg and B. Dimsdale, "Parallel coordinates: A tool for visualizing multi-dimensional geometry," in *Proc. 1st IEEE Conf. Vis.*, 1990, pp. 361–378.

[42] C. Konold, "Representing probabilities with pipe diagrams," *Math. Teacher*, vol. 89, no. 5, pp. 378–382, 1996. [Online]. Available: http://www.jstor.org/stable/27969794
[43] S. T. Teoh and K.-L. Ma, "StarClass: Interactive visual classification using star coordinates," in *Proc. Int. Conf. Data Mining*, 2003, pp. 178–185.
[44] S. T. Teoh and K.-L. Ma, "PaintingClass: Interactive construction, visualization and exploration of decision trees," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 2003, pp. 667–672.
[45] V. Podgorelec, P. Kokol, B. Stiglic, and I. Rozman, "Decision trees: An overview and their use in medicine," *J. Med. Syst.*, vol. 26, no. 5, pp. 445–463, 2002.
[46] C. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[47] C. Kingsford and S. L. Salzberg, "What are decision trees?," *Nat. Biotechnol.*, vol. 26, no. 9, pp. 1011–1013, 2008.
[48] L. A. Breslow and D. W. Aha, "Simplifying decision trees: A survey," *Knowl. Eng. Rev.*, vol. 12, no. 1, pp. 1–40, 1997.
[49] B. Liu, M. Hu, and W. Hsu, "Intuitive representation of decision trees using general rules and exceptions," in *Proc. 17th Nat. Conf. Arti. Intell. 12th Conf. Innovative Appl. Artif. Intell.*, 2000, pp. 615–620.
[50] L. Martignon, O. Vitouch, M. Takezawa, and M. R. Forster, "Naive and yet enlightened: From natural frequencies to fast and frugal decision trees," in *Thinking: Psychological Perspective on Reasoning, Judgment, and Decision Making*. Hoboken, NJ, USA: Wiley, 2003.
[51] W.-Y. Loh, "Classification and regression trees," *Wiley Interdisciplinary Rev., Data Mining Knowl. Discov.*, vol. 1, no. 1, pp. 14–23, 2011.
[52] M. Ware, E. Frank, G. Holmes, M. A. Hall, and I. H. Witten, "Interactive machine learning: Letting users build classifiers," *Int. J. Hum.-Comput. Stud.*, vol. 55, no. 3, pp. 281–292, 2001.
[53] L. Rokach and O. Maimon, "Top-down induction of decision trees classifiers – A survey," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 4, pp. 476–487, Nov. 2005.
[54] Y. Liu and G. Salvendy, "Interactive visual decision tree classification," in *Proc. Int. Conf. Hum.-Comput. Interact.*, 2007, pp. 92–105.
[55] K. Li, "On integrating information visualization techniques into data mining: A review," 2015, *arXiv:1503.00202*, [Online]. Available: https://arxiv.org/abs/1503.00202
[56] M. Ankerst, M. Ester, and H.-P. Kriegel, "Towards an effective cooperation of the user and the computer for classification," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 2000, pp. 179–188.
[57] F. Poulet, "Cooperation between automatic algorithms, interactive algorithms and visualization tools for visual data mining," in *Proc. Int. Workshop Vis. Data Mining*, 2002, pp. 67–79.
[58] S. Urbanek, "Interactive construction and analysis of trees," in *Proc. Joint Statist. Meetings*, 2003.
[59] S. Amershi, M. Chickering, S. M. Drucker, B. Lee, P. Simard, and J. Suh, "ModelTracker: Redesigning performance analysis tools for machine learning," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2015, pp. 337–346.
[60] F. Provost and T. Fawcett, "Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 1997, pp. 43–48.
[61] B. Alsallakh, A. Hanbury, H. Hauser, S. Miksch, and A. Rauber, "Visual methods for analyzing probabilistic classification data," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1703–1712, Dec. 2014.
[62] K. N. Pham and N. T. Do, "Interactive exploration of decision tree results," in *Proc. Int. Symp. Appl. Stochastic Models Data Anal.*, 2008, pp. 152–160.
[63] V. Bhatnagar, E. Zaman, Y. Rajpal, and M. Bhardwaj, "VisTree: Generic decision tree inducer and visualizer," in *Proc. Int. Workshop Databases Netw. Inf. Syst.*, 2010, pp. 225–243.
[64] T. Parr and P. Grover, "How to visualize decision trees," Visited: Sep. 23, 2019. [Online]. Available: https://explained.ai/decision-tree-viz/
[65] L. Padua, H. Schulze, K. Matković, and C. Delrieux, "Interactive exploration of parameter space in data mining: Comprehending the predictive quality of large decision tree collections," *Comput. Graph.*, vol. 41, pp. 99–113, 2014.
[66] T. Mühlbacher, L. Linhardt, T. Möller, and H. Piringer, "TreePOD: Sensitivity-aware selection of pareto-optimal decision trees," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 174–183, Jan. 2018.
[67] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52 138–52 160, 2018.
[68] T. D. Nguyen, T. B. Ho, and H. Shimodaira, "A visualization tool for interactive learning of large decision trees," in *Proc. 12th IEEE Int. Conf. Tools Artif. Intell.*, 2000, pp. 28–35.
[69] C. Chen, F. Ibekwe-SanJuan, E. SanJuan, and C. Weaver, "Visual analysis of conflicting opinions," in *Proc. IEEE Symp. Vis. Analytics Sci. Technol.*, 2006, pp. 59–66.
[70] L. Tanner *et al.*, "Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness," *PLoS Negl Trop Dis*, vol. 2, no. 3, p. e196, 2008. [Online]. Available: https://doi.org/10.1371/journal.pntd.0000196
[71] U. Niemann, H. Völzke, J.-P. Kühn, and M. Spiliopoulou, "Learning and inspecting classification rules from longitudinal epidemiological data to identify predictive features on hepatic steatosis," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5405–5415, 2014.
[72] J. F. Sydow, F. Lipsmeier, V. Larraillet, M. Hilger, and B. Mautz, "Structure-based prediction of asparagine and aspartate degradation sites in antibody variable regions," *PLoS One*, vol. 9, 2014, Art. no. e100736.
[73] H. Handa *et al.*, "Exhaled breath analysis for lung cancer detection using ion mobility spectrometry," *PLoS One*, vol. 9, 2014, Art. no. e114555.
[74] M. Moussa∂d, J. E. Kämmer, P. P. Analytis, and H. Neth, "Social influence and the collective dynamics of opinion formation," *PLoS One*, vol. 8, no. 11, 2013, Art. no. e78433.
[75] Y. Ming, H. Qu, and E. Bertini, "RuleMatrix: Visualizing and understanding classifiers with rules," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 342–352, Jan. 2019.
[76] S. Jia, P. Lin, Z. Li, J. Zhang, and S. Liu, "Visualizing surrogate decision trees of convolutional neural networks," *J. Vis.*, vol. 23, pp. 141–156, 2020.
[77] Y. Huang, Y. Liu, C. Li, and C. Wang, "GBRTVis: Online analysis of gradient boosting regression tree," *J. Vis.*, vol. 22, no. 1, pp. 125–140, 2019.
[78] S. Liu, J. Xiao, J. Liu, X. Wang, J. Wu, and J. Zhu, "Visual diagnosis of tree boosting methods," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 163–173, Jan. 2018.
[79] T. Barlow and P. Neville, "Case study: Visualization for decision tree analysis in data mining," in *Proc. IEEE Symp. Inf. Vis.*, 2001, pp. 149–152.
[80] K. Hornbaek and A. Oulasvirta, "What is interaction?," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2017, pp. 5040–5052.
[81] T. May and J. Kohlhammer, "Towards closing the analysis gap: Visual generation of decision supporting schemes from raw data," *Comput. Graph. Forum*, vol. 27, no. 3, pp. 911–918, 2008.
[82] D. Mott and R. Tomsett, "Illuminated decision trees with lucid," 2019, *arXiv:1909.05644*, [Online]. Available: https://arxiv.org/pdf/1909.05644
[83] T. Wang, C. Rudin, F. Doshi-Velez, E. Klampfl, and P. MacNeille, "A Bayesian framework for learning rule sets for interpretable classification," *J. Mach. Learn. Res.*, vol. 18, no. 70, pp. 1–37, 2017.
[84] X. Hu, C. Rudin, and M. Seltzer, "Optimal sparse decision trees," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 7267–7275.
[85] G. Stiglic, S. Kocbek, I. Pernek, and P. Kokol, "Comprehensive decision tree models in bioinformatics," *PLoS One*, vol. 7, no. 3, 2012, Art. no. e33812.
[86] I. C. Hostettler *et al.*, "Decision tree analysis in subarachnoid hemorrhage: Prediction of outcome parameters during the course of aneurysmal subarachnoid hemorrhage using decision tree analysis," *J. Neurosurgery*, vol. 129, pp. 1499–1510, 2018.
[87] X. Zhao, Y. Wu, D. L. Lee, and W. Cui, "iForest: Interpreting random forests via visual analytics," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 407–416, Jan. 2019.
[88] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[89] S. H. Welling, H. H. F. Refsgaard, P. B. Brockhoff, and L. H. Clemmensen, "Forest floor visualizations of random forests," 2016, *arXiv:1605.09196*, [Online]. Available: https://arxiv.org/abs/1605.09196
[90] D. Leake, "Problem solving and reasoning: Case-based," in *Proc. Int. Encyclopedia Soc. Behav. Sci.*, 2001, pp. 12117–12120.
[91] T. Lee, J. Johnson, and S. Cheng, "An interactive machine learning framework," 2009, *arXiv:1610.05463*, [Online]. Available: https://arxiv.org/pdf/1610.05463

[92] G. Ke *et al.*, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 3149–3157.

[93] J. Guerra-Gómez, M. L. Pack, C. Plaisant, and B. Shneiderman, "Visualizing change over time using dynamic hierarchies: TreeVersity2 and the stemview," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2566–2575, Dec. 2013.

[94] S. Linn, "A new conceptual approach to teaching the interpretation of clinical tests," *J. Statist. Edu.*, vol. 12, no. 3, pp. 1–11, 2004.

[95] M. O. Ward, G. G. Grinstein, and D. A. Keim, *Interactive Data Visualization: Foundations, Techniques, and Applications*. Boca Raton, FL, USA: CRC Press, 2015.

[96] G. C. Fonarow, K. F. Adams, W. T. Abraham, C. W. Yancy, and W. J. Boscardin, "Risk stratification for in-hospital mortality in acutely decompensated heart failure classification and regression tree analysis," *J. Amer. Med. Assoc.*, vol. 293, no. 5, pp. 572–580, 2005.

[97] E. T. Brown *et al.*, "Finding Waldo: Learning about users from their interactions," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1663–1672, Dec. 2014.

[98] D. Sakurai *et al.*, "Interactive visualization for singular fibers of functions f: R3 → r2," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 1, pp. 945–954, Jan. 2016.

[99] G. K. L. Tam *et al.*, "Visualization of time-series data in parameter space for understanding facial dynamics," *Comput. Graph. Forum*, vol. 30, no. 3, pp. 901–910, 2011.

[100] J. Kreiser, A. Hann, E. Zizer, and T. Ropinski, "Decision graph embedding for high-resolution manometry diagnosis," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 873–882, Jan. 2018.

[101] Start flowchart, 2001. [Online]. Available: http://citmt.org/Start/flowchart.htm

[102] R. Kagan, L. Parlee, B. Beckett, J. B. Hayden, K. R. Gundle, and Y.-C. Doung, "Radiographic parameter-driven decision tree reliably predicts aseptic mechanical failure of compressive osseointegration fixation," *Acta Orthopaedica*, vol. 91, pp. 171–176, 2020.

[103] M. Aupetit, Y. Zhauniarovich, G. Vasiliadis, M. Dacier, and Y. Boshmaf, "Visualization of actionable knowledge to mitigate DRDoS attacks," in *Proc. IEEE Symp. Vis. Cyber Secur.*, 2016, pp. 1–8.

[104] J. Quionero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset Shift in Machine Learning*. Cambridge, MA, USA: MIT Press, 2009.

[105] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla, and F. Herrera, "A unifying view on dataset shift in classification," *Pattern Recognit.*, vol. 45, no. 1, pp. 521–530, 2012.

[106] M. K. Dhami, "Psychological models of professional decision making," *Psychol. Sci.*, vol. 14, no. 2, pp. 175–180, 2003.

[107] J. H. Tan, S. Luan, and K. V. Katsikopoulos, "A signal-detection approach to modeling forgiveness decisions," *Evol. Hum. Behav.*, vol. 38, no. 1, pp. 27–38, 2017.

[108] S. Luan, L. J. Schooler, and G. Gigerenzer, "A signal-detection analysis of fast-and-frugal trees," *Psychol. Rev.*, vol. 118, no. 2, pp. 316–338, 2011.

[109] J. H. Tan, S. Luan, T. Gonzalez, and E. Jablonskis, "Testing error-management predictions in forgiveness decisions with cognitive modeling and process-tracing tools," *Evol. Behav. Sci.*, vol. 12, no. 3, pp. 206–217, 2018.

[110] F. B. Gilbreth and L. M. Gilbreth, *Process Charts*. New York, NY, USA: ASME, 1921.

[111] J. T. Stasko, R. Catrambone, M. Guzdial, and K. McDonald, "An evaluation of space-filling information visualizations for depicting hierarchical structures," *Int. J. Hum.-Comput. Stud.*, vol. 53, no. 5, pp. 663–694, 2000.

[112] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA, USA: Addison-Wesley, 1977.

[113] L. Wilkinson, "Dot plots," *Amer. Statistician*, vol. 53, no. 3, pp. 276–281, 1999.

[114] D. A. Keim and H.-P. Kriegel, "VisDB: Database exploration using multidimensional visualization," *Comput. Graph. Appl.*, vol. 14, no. 5, pp. 40–49, 1994.

[115] E. Kandogan, "Star coordinates: A multi-dimensional visualization technique with uniform treatment of dimensions," in *Proc. IEEE Inf. Vis. Symp., Late Breaking Hot Topics*, pp. 9–12, 2000.

[116] L. Byron and M. Wattenberg, "Stacked graphs – geometry aesthetics," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1245–1252, Nov./Dec. 2008.

[117] F. Di Castro and E. Bertini, "Surrogate decision tree visualization: Interpreting and visualizing black-box classification models with surrogate decision tree," in *Proc. Conf. Intell. User Interfaces Workshops*, 2019.

[118] R. Balestriero, "Neural decision trees," 2017. [Online]. Available: https://arxiv.org/pdf/1702.07360

[119] J. Han and N. Cercone, "RuleViz: A model for visualizing knowledge discovery process," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 2000, pp. 244–253.

[120] M. Ankerst, C. Elsen, M. Ester, and H.-P. Kriegel, "Visual classification: An interactive approach to decision tree construction," in *Proc. Int. Conf. Knowl. Discov. Data Mining*, 1999, pp. 392–396.

[121] T. Zuk and S. Carpendale, "Visualization of uncertainty and reasoning," in *Proc. Int. Symp. Smart Graph.*, 2007, pp. 164–177.

[122] J. Zhang, L. Gruenwald, and M. Gertz, "VDM-RS: A visual data mining system for exploring and classifying remotely sensed images," *Comput. Geosci.*, vol. 35, no. 9, pp. 1827–1836, 2009.

[123] J. Forberg, A. Mitschick, M. Voigt, and R. Dachselt, "Interactive exploration of large decision tree ensembles," in *Proc. Hum.-Centered Mach. Learn. Perspectives Workshop*, 2019.

[124] G. Ilczuk and A. Wakulicz-Deja, "Visualization of rough set decision rules for medical diagnosis systems," in *Proc. Int. Conf. Rough Sets Fuzzy Sets Data Mining Granular Comput.*, 2007, pp. 371–378.

[125] A. Wlodyka, R. Mlynarski, G. Ilczuk, E. Pilat, and W. Kargul, "Visualization of decision rules – from the cardiologist's point of view," *Comput. Cardiol.*, vol. 35, pp. 645–648, 2008.

[126] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[127] T. Munzner, F. Guimbretière, S. Tasiran, L. Zhang, and Y. Zhou, "TreeJuxtaposer: Scalable tree comparison using focus+context with guaranteed visibility," *Trans. Graph.*, vol. 22, no. 3, pp. 453–462, 2003.

[128] D. Holten and J. J. van Wijk, "Visual comparison of hierarchically organized data," in *Proc. Joint Conf. Vis.*, 2008, pp. 759–766.

[129] Y. Park, J. Ho, and J. Ghosh, "ACDC: α-carving decision chain for risk stratification," in *Proc. Int. Conf. Mach. Learn.*, 2016

[130] T.-N. Do, "Towards simple, easy to understand, an interactive decision tree algorithm," College Information Technology Can tho University, Can Tho, Vietnam, 2007.

[131] T. Dang, P. Murray, R. Etemadpour, and A. G. Forbes, "A user study of techniques for visualizing structure and connectivity in hierarchical datasets," in *Proc. Int. Workshop Vis. Interact. Ontologies Linked Data*, 2017.

[132] N. H. Müller, B. Liebold, D. Pietschmann, P. Ohler, and P. Rosenthal, "Visualizations for hierarchical data: Analyzing user behavior and performance with eye tracking," *Int. J. Advances Softw.*, vol. 4, no. 3/4, pp. 385–396, 2017.

[133] J. Wang, B. Yu, and L. Gasser, "Concept tree based clustering visualization with shaded similarity matrices," in *Proc. IEEE Int. Conf. Data Mining*, 2002, pp. 697–700.

[134] D. Liu, A. P. Sprague, and J. G. Gray, "PolyCluster: An interactive visualization approach to construct classification rules," in *Proc. Int. Conf. Mach. Learn. Appl.*, 2004, pp. 280–287.

[135] Y. Xu, W. Hong, N. Chen, X. Li, W. Liu, and T. Zhang, "Parallel filter: A visual classifier based on parallel coordinates and multivariate data analysis," in *Proc. Int. Conf. Intell. Comput.*, 2007, pp. 1172–1183.

[136] W. Castillo-Rojas, F. Medina-Quispe, and C. Meneses-Villegas, "Augmenting decision tree models using self-organizing maps," in *Proc. Latin Amer. Conf. Hum. Comput. Interact.*, 2013, pp. 148–155.

[137] W. Castillo-Rojas, F. Medina Quispe, and C. Meneses Villegas, "Augmented visualization for data-mining models," *Procedia Comput. Sci.*, vol. 55, pp. 650–659, 2015.

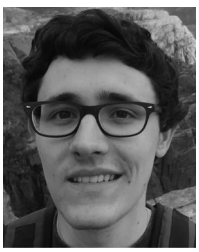
[138] C. Xie, W. Chen, X. Huang, Y. Hu, S. Barlowe, and J. Yang, "VAET: A visual analytics approach for E-transactions time-series," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1743–1752, Dec. 2014.

[139] D. Szücs and F. Schmidt, "Decision tree visualization for high-dimensional numerical data," in *5th Proc. Int. Conf. Soc. Netw. Anal. Manage. Secur.*, 2018, pp. 190–195, doi: 10.1109/SNAMS.2018.8554961.

[140] M. Cavallo and Ç. Demiralp, "Clustrophile 2: Guided visual clustering analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 267–276, Jan. 2019.

[141] T. Barlow and P. Neville, "A comparison of 2-D visualizations of hierarchies," in *Proc. IEEE Symp. Inf. Vis.*, 2001, pp. 131–138.
[142] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.
[143] K. Pearson, "X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *London Edinburgh Dublin Philosoph. Mag. J. Sci.*, vol. 50, no. 302, pp. 157–175, 1900.

Dirk Streeb received the MSc degree in social and economic data analysis from the University of Konstanz, Konstanz, Germany, in 2016. He is currently working toward the PhD degree with the Graduate School of Decision Sciences, University of Konstanz, Konstanz, Germany. He is a member of the Data Analysis and Visualization Research Group, University of Konstanz, Konstanz, Germany. His research interests include theoretical aspects of information visualization and visual analytics as aids in human decision-making processes.



Yannick Metz received the BSc degree in computer science from the University of Konstanz, Konstanz, Germany, in 2019. He is currently working toward the master's degree with the University of Konstanz, Konstanz, Germany. His research interests include explainability in machine learning, deep (reinforcement) learning, and interactive data visualization.

Udo Schlegel received the MSc degree in computer and information science from the University of Konstanz, Konstanz, Germany, in 2018 with his thesis Towards Crime Forecasting Using Deep Learning. He is currently working toward the PhD degree with the Data Analysis and Visualization Group, University of Konstanz, Konstanz, Germany. His research interests include visual analytics for machine learning, and explainable AI.

Bruno Schneider received the MSc degree from the School of Applied Math, FGV, Rio de Janeiro, Brazil, in 2014. He is currently working toward the PhD degree with the Data Analysis and Visualization Group, University of Konstanz, Konstanz, Germany. His main research interests include information visualization and interactive machine learning.

Mennatallah El-Assady received the MSc degree in information engineering from the University of Konstanz, Konstanz, Germany, in 2015. She is currently a research associate with the Data Analysis and Visualization Research Group, University of Konstanz, Germany, and with the Visualization for Information Analysis Lab, University of Ontario Institute of Technology, Canada. Her research interests include visual text analytics, user-steerable topic modeling, and discourse/conversational text analysis.

Hansjörg Neth received the PhD degree in psychology from Cardiff University, Cardiff, U.K. He is currently researcher in social psychology and decision sciences with the University of Konstanz, Germany. His research focuses on the theoretical analysis of adaptive behavior and ecological rationality, as well as applied aspects of choice and heuristic decision-making under risk and uncertainty. He has served as acting chair of cognition, emotion, and communication with the University of Freiburg, was a member of the ABC research group at the Max-Planck-Institute for Human Development, Berlin, and taught cognitive and decision sciences with the University of Göttingen, and was research assistant professor in cognitive science with the Rensselaer Polytechnic Institute.

Min Chen (Member, IEEE) received the PhD degree from the University of Wales, Cardiff, U.K., in 1991. He is currently a professor of scientific visualization with Oxford University and a fellow of Pembroke College. Before joining Oxford, he held research and faculty positions with Swansea University, U.K. His research interests include visualization, computer graphics, and human-computer interaction. His services to the research community include papers co-chair of the IEEE Visualization 2007 and 2008, IEEE VAST 2014 and 2015, and Eurographics 2011; co-chair of Volume Graphics 1999 and 2006, and EuroVis 2014; associate editor-in-chief of the *IEEE Transactions on Visualization and Computer Graphics*; and co-director of Wales Research Institute of Visual Computing. He is a fellow of BCS, EG, and LSW.

Daniel A. Keim (Member, IEEE) received the PhD degree in computer science from the University of Munich, Munich, Germany. He is currently a professor and head of the Data Analysis and Visualization Research Group, Computer Science Department, University of Konstanz, Germany. He has been actively involved in data analysis and information visualization research for more than 25 years and developed novel visual analysis techniques for large data sets. His services to the research community include papers cochair of IEEE InfoVis 1999 and 2000, ACM SIDKDD 2002, IEEE VAST 2006 and 2019; general chair of InfoVis 2003; and an associate editor of the *IEEE Transactions on Visualization and Computer Graphics*, *IEEE Transactions on Knowledge and Data Engineering*, and the *Sage Information Visualization Journal*. Before joining the University of Konstanz, he was an associate professor with the University of Halle, Germany, and technology consultant with AT&T Shannon Research Labs, New Jersey.