Elements About Exploratory, Knowledge-Based, Hybrid, and Explainable Knowledge Discovery

Miguel Couceiro, Amedeo Napoli

Elements about Exploratory, Knowledge-Based, Hybrid, and Explainable Knowledge Discovery

Miguel Couceiro and Amedeo Napoli

Université de Lorraine, CNRS, Inria, LORIA
F-54000 Nancy, France
{miguel.couceiro,amedeo.napoli}@inria.fr

Abstract. Knowledge Discovery in Databases (KDD) and especially pattern mining can be interpreted along several dimensions, namely data, knowledge, problem-solving and interactivity. These dimensions are not disconnected and have a direct impact on the quality, applicability, and efficiency of KDD. Accordingly, we discuss some objectives of KDD based on these dimensions, namely exploration, knowledge orientation, hybridization, and explanation. The data space and the pattern space can be explored in several ways, depending on specific evaluation functions and heuristics, possibly related to domain knowledge. Furthermore, numerical data are complex and supervised numerical machine learning methods are usually the best candidates for efficiently mining such data. However, the work and output of numerical methods are most of the time hard to understand, while symbolic methods are usually more intelligible. This calls for hybridization, combining numerical and symbolic mining methods to improve the applicability and interpretability of KDD. Moreover, suitable explanations about the operating models and possible subsequent decisions should complete KDD, and this is far from being the case at the moment. For illustrating these dimensions and objectives, we analyze a concrete case about the mining of biological data, where we characterize these dimensions and their connections. We also discuss dimensions and objectives in the framework of Formal Concept Analysis and we draw some perspectives for future research.

1 Introduction

Knowledge discovery in databases (KDD) consists in processing possibly large volumes of data in order to discover patterns that can be significant and reusable. It is usually based on three main steps: data preparation, data mining, and interpretation of the extracted patterns (Figure 1) [56,48]. KDD is interactive and iterative, controlled by an analyst who is a specialist of the domain and is in charge of selecting data and patterns, setting thresholds (frequency, confidence), replaying the process at each step whenever needed, depending on the interpretation of the selected patterns.

In the following, we consider four main dimensions within KDD which are based on data, knowledge, problem-solving, and interactivity.

- The data dimension: knowledge discovery is data-oriented by nature. This dimension is related to the input of knowledge discovery and involves data preparation, e.g. feature selection, dimensionality reduction, and data transformation. The exploration of the data space and of pattern space are main operations. Moreover, the diversity and quality of data have an influence on the whole KDD process.
- The knowledge dimension [34]: data are related to a particular domain. Hence knowledge discovery is knowledge-oriented and depends on domain knowledge that can be expressed, e.g., as constraints, relations, and preferences. The knowledge dimension is also attached to the control of KDD, possibly involving "meta-mining" [11]. Moreover, the output of KDD, i.e. the discovered patterns, may be represented as actionable knowledge units.
- The problem-solving dimension [9]: knowledge discovery is intended to solve various tasks for human or software agents and may be guided by the task at hand. The problem-solving dimension is dependent on iteration and search strategies. It can be data-directed, pattern-directed or goal-directed, and it can rely on declarative or procedural approaches.
- Interactivity [10]: knowledge discovery is interactive as the analyst may integrate constraints and preferences for guiding the data exploration, especially for minimizing the exploration of the data and pattern spaces. Interaction plays also a role in the evaluation of the quality of the patterns and in the activation of the different replay loops.

These four dimensions are interconnected and correspondences between them can be made explicit, especially in the framework of pattern mining [1]. Such correspondences support the following objectives of KDD:

- KDD is exploratory: the data dimension is related to an interactive exploration of the data and pattern spaces. We should be able to identify "seeds" or "prototypes" for guiding the pattern and data space exploration. Moreover, this exploration should be consistent w.r.t. domain knowledge and associated constraints. Threshold issues w.r.t. analyst queries can be addressed thanks to a skyline analysis within the pattern space or by integrating preferences.
- KDD is knowledge-based: domain knowledge is related to control, i.e. meta-mining, constraints and preferences, explanations, and production of knowledge, materializing the links between knowledge discovery

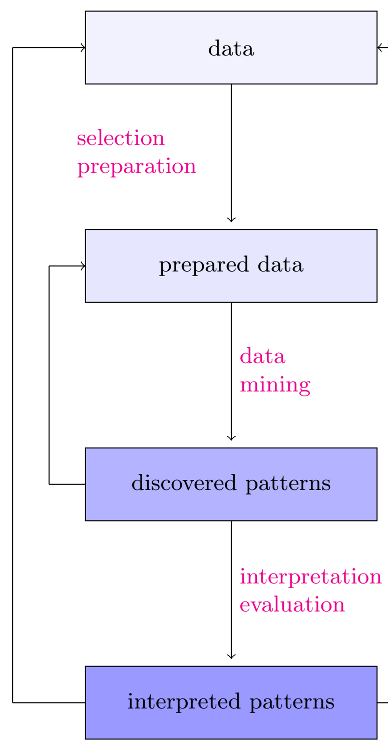

Fig. 1: The KDD loop.

and knowledge engineering. We should be able to define environments within which knowledge discovery and knowledge engineering can be combined in an efficient and operational way.

- KDD is hybrid: and may rely on the combination of numerical and symbolic data mining algorithms for solving problems. In addition, supervised and unsupervised learning methods can interact and be combined as well.
- KDD is expected to be explainable: the output of knowledge discovery may be of different types, e.g. rules and classes or concepts, which can be reused for solving problems and decision making. In this way, elements supporting a decision –and especially an algorithmic decision– should be available [44,55,27].

In the following sections, we survey in more detail these different objectives and discuss their importance and their materialization. In the last section of this paper, we propose a synthesis that illustrates how Formal

Concept Analysis (FCA [23]) can, more or less, fulfill these objectives. The paper terminates with a large bibliography which illustrates the content of this paper and some relations existing between pattern mining and FCA.

2 KDD should be Exploratory

In the KDD loop, exploration is related to data mining where the data space is searched and the patterns are mined. Exploration is also related to the notions of interaction and iteration, involving "replay", which is of main importance within knowledge discovery. This is discussed under different names in the literature, e.g. "exploratory data mining" and the loop "Mine, Interact, Learn, and Repeat" in [8,52], "interactive data mining" in [32], "declarative approaches in data mining" [9], and "exploratory knowledge discovery" in [3] (this list is certainly not exhaustive).

All these approaches are based on interaction and go back to the ideas underlying "exploratory data analysis" (EDA [50]). The goal of EDA is to improve data analysis and result interpretation, providing the analyst with suitable techniques based on computational power, data exploration and visualization methods.

In the context of KDD, exploration can be achieved in various ways, using either data-directed or pattern-directed methods [13], particular interestingness measures [12], and visualization procedures [3]. Nonetheless, the knowledge discovery process should be efficient and automated as much as possible, while keeping facilities for interaction and iteration.

In the same way, let us quote some recent variations about the current exploratory approaches in pattern mining, namely constraint-based pattern mining, subgroup discovery, and exceptional model mining. Constraint-based pattern mining is based on skylines in [51] in which preferences are expressed w.r.t. a dominance relation.

The goal of subgroup discovery [41,5] is to find particular descriptions of subsets of a population that are sufficiently large and statistically unusual, i.e. subsets of the population that deviate from the norm. Such deviations are measured in terms of a relatively high occurrence, e.g. frequent itemset mining, or an unusual distribution for one designated target attribute. The latter is the common use of subgroup discovery, which is, in turn, related to exceptional model mining. Exceptional Model Mining (EMM [19,39,6]) is aimed at capturing a general notion of interestingness in subsets of a dataset. EMM can be considered as a supervised local pattern mining framework, where several target attributes are selected, and a model over these targets is chosen to be the target concept.

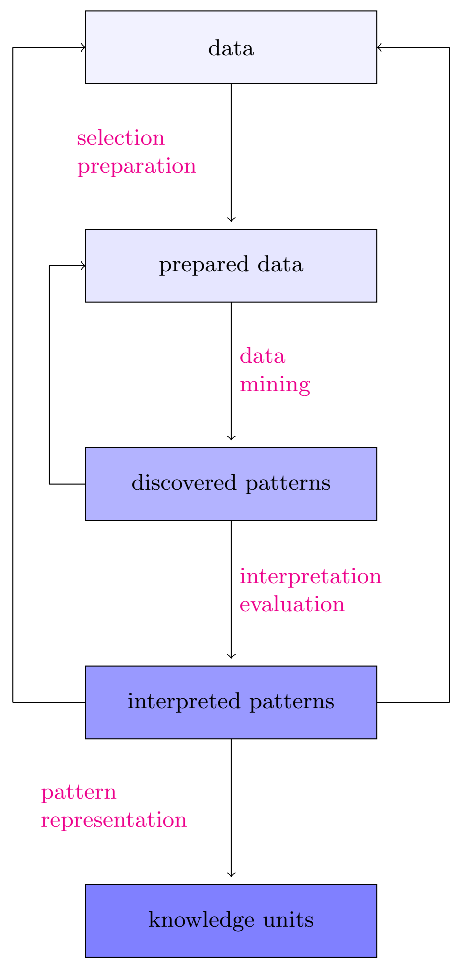

Fig. 2: The augmented KDD loop.

3 KDD should be Knowledge-Based

There are many links between knowledge discovery and knowledge engineering. At each step of KDD, domain knowledge can be used to guide and to complete the process given in Figure 1. Actually, a fourth step can be considered in the KDD process, where selected patterns are represented as "actionable knowledge units" (see Figure 2). This involves a knowledge representation (KR) formalism and, subsequently, actionable units can be reused –by human or software agents– in knowledge graphs or knowledge systems for problem-solving.

Recently, efforts have been made to automate several tasks in KDD. Indeed, this is one main objective of meta-learning to design principles that can make algorithms adaptive to the characteristics of the data [11]. In [31,43], authors consider meta-learning as the application of machine learning techniques to meta-data describing past learning experiences for adapting the learning process and improve the performance of the current model [31,43] (a kind of "analogy"). Meta-learning is often related to the problem of dynamically solving or adjusting learning constraints. Most of the time, if there are no restrictions on the space of hypotheses to be explored by a learning algorithm and no preference criteria for comparing candidate hypotheses, then no inductive method can do better on average than random guessing. Authors in [31,43] make a distinction between two main types of "biases", the representational bias restricts the hypothesis space whereas the preference-bias gives priority to certain hypotheses over others in this space. The most widely addressed meta-learning tasks are algorithm selection and model selection. Algorithm selection is the choice of the appropriate algorithm for a given task, while model selection is the choice of a specific parameter settings that produces a good performance for a given algorithm on a given task.

In the framework of KDD, meta-mining is not only meta-learning, since the KDD loop involves data preparation and pattern interpretation. Thus, the performance of the process and the quality of the discovered patterns and their usage are not only depending on the mining operation but also on data preparation –data selection, data cleaning, and feature selection–, pattern interpretation and possible pattern representation.

4 KDD should be Hybrid

Knowledge discovery should be able to work on various datasets with various characteristics, e.g. data can be either symbolic or numerical, or even more complex structured data such as sequences, trees, graphs, texts, linked data... The data management is based on different operations, including for example feature selection, dimensionality reduction, and noise reduction. Also, the mining approaches can be supervised or unsupervised, depending on the task at hand, and the availability of examples and counterexamples. As in every task, there is no universal approach which may be used alone to tackle and to solve all problems. Hence, a reasonable strategy is to design a hybrid process based on several tactics whose output is a combination of outputs of the involved procedures.

Accordingly, in [28], we aimed at discovering in metabolomic data a small set of relevant predictive features using a hybrid and exploratory knowledge discovery approach. This approach relies on adapted classification techniques which should deal with high-dimensional datasets, composed of small sets of individual and large sets of complex features. There are many possible classifiers that can be used and their application induces a bias on the results, calling for the simultaneous use of several classifiers. Hence, we adopted a kind of "ensemble approach" and we designed a set of classifiers instead of using a single one, to reach complementarity. The whole process is exploratory and hybrid as it combines numerical and symbolic classifiers. More details are given farther in Section 6.

To conclude, the combination of numerical and symbolic data mining algorithms remains rather rare, even if ensemble methods [18,46] are available without being always satisfactory.

5 KDD should be Explainable

Many recent progress in Machine Learning (ML) are mostly due to the success of Deep Learning methods in recognition tasks. However, Deep Learning and other numerical ML approaches are based on complex models, whose outputs and proposed decisions, as accurate as they are, cannot be easily explained to the layman [30]. Indeed, it is interesting to study hybrid ML approaches that combine complex numerical models with explainable symbolic models, in order to make ML methods more "interpretable". The objective is to attach what we could call "integrity constraints" –or kinds of "pre" and "post-conditions" to be fulfilled– to build and then deliver understandable explanations on the work of numerical ML models.

The objective of supervised ML is to learn how to perform a task (e.g. recognition) based on a sufficient number of training examples. The ML algorithm builds a model from the training examples which is then used on test examples. A model can be either symbolic, e.g. a set of rules or a hierarchy of concepts, or numerical, e.g. a set of weights associated with a structure as in neural networks. In practice, numerical models often prove to be more flexible and better suited to capture the complexity of some tasks such as recognition. However, symbolic ML models are more often used in pattern mining and in domains where the learning model should be understandable by human experts, or be related to domain ontologies for knowledge representation and reasoning purposes.

Moreover, ML models approaches based on numerical models are more and more used for complex tasks such as decision making with a strong impact for human users, e.g. make a decision for a student orientation at university. In the latter case it can be very difficult to provide the necessary explanations justifying the decision using these numerical ML models, and especially models based on Deep Learning. Thus there is an emerging research trend whose goal is to provide interpretation and explanations about the decision of numerical ML algorithms such as Deep Learning.

Here, we are interested in understanding the different ways of providing explanations and facilitating interpretation of the outputs of numerical ML models [44]. There are several attempts to build explainable and trustable ML models. As mentioned above, one is to combine symbolic and numerical approaches and to build interactions between both approaches, as in [29] which is based on a combination of numerical learning methods and Formal Concept Analysis [23], or in [47] which is based on a combination of first-order logic and neural network learning.

There are also other initiatives, as in [31,43], on the understandability of a mining process in terms of core components (modules), underlying assumptions, cost functions and optimization strategies being used. One subsequent idea is to understand how Deep Learning models can be decomposed w.r.t. such modules and then to integrate adapted explanation modules .

These are some other possible directions for analyzing a numerical ML model and provide plausible explanations about its output, as for example "neural-symbolic learning and reasoning" combinations [17,49] that should be carefully examined and adapted.

6 An Application in the Mining of Metabolomic Data

In [28], we presented a case study about the mining of metabolomic data using a combination of symbolic and numerical mining methods. Given a dataset composed of individuals described by features, the objective of the experiment is to discover subsets of features that can be *discriminant* and *predictive*. The discrimination power allows to build classes of individuals, where classes include similar individuals and separate dissimilar individuals at the best. The prediction power allows to determine the potential class membership of individuals, e.g. people who will develop the disease under study.

The classification process is split into two main procedures, one being supervised and the other unsupervised (see Figure 3). The supervised classification procedure is based on the design of NC numerical classifiers (including Random Forest and SVM in the present case) which are completed by preprocessing and postprocessing operations. This first procedure is applied to a bidimensional dataset composed of individuals and features. The output of this first procedure provides a set of ranked features RF for each of the NC classifiers.

Then the second procedure is based on a unsupervised mining operation applied to RF, the set of ranked features, for discovering the most frequent features, given a threshold set by the analyst. This second procedure is applied to a bidimensional dataset composed this time of $|RF|$ features and NC numerical classifiers. The output of this second procedure provides a set FF of frequent features. These frequent features are the best candidates for becoming best discriminant features. Actually, there is a change of the representation space between the two classification procedures.

Following the classification process and the selection of most discriminant features, the latter are tested for evaluating their predictive capabilities using a ROC analysis [22].

For summarizing, this knowledge discovery strategy relies on two main steps: (i) a concurrent use of multiple classifiers producing a stable set of discriminant features, (ii) a classification of features based on FCA through a change of the problem space representation, where a small set of most relevant features is retained. In this classification process, we can distinguish the following elements:

- KDD is exploratory: two search strategies are run, the first on a data space $individuals \times features$ and the second on a data space $features \times classifiers$. The combination of these two exploration operations produces a set of most discriminant features which are then tested for prediction.
- KDD is knowledge-based: the analyst is asked to control the process and the design of classifiers, to adjust the value of thresholds for dimensionality reduction and feature selection. The analyst plays also a similar role in the prediction analysis.
- KDD is hybrid and involves numerical classifiers as well as more symbolic pattern mining methods. The latter are used for analyzing top-k ranked features w.r.t. discrimination and prediction.
- KDD is explainable thanks to visualization. In particular, the pattern mining procedure involves FCA and the design of a concept lattice

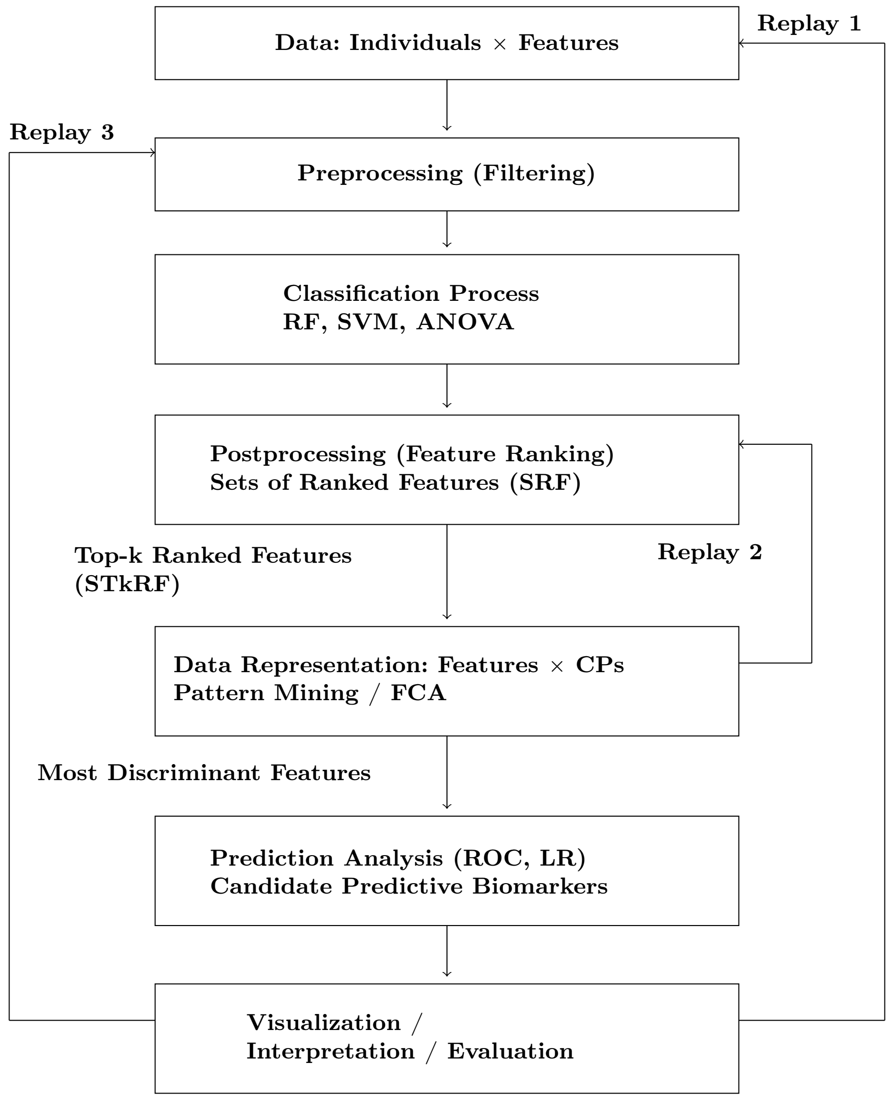

Fig. 3: A hybrid and exploratory approach to metabolomic data mining, with a classification process based on two procedures applied to two different bidimensional datasets. In addition, the "Replay" arrows show that the whole process is interactive and iterative.

where concepts represent classes of top-k ranked features. The distribution of the features within the lattice can be analyzed by the analyst for suggesting further test for prediction analysis.

This experiment shows how and why the different dimensions and objectives of KDD are contributing to deliver a more complete and qualitative analysis of metabolomic data.

7 Discussion: What about FCA?

We have introduced dimensions on which KDD is based, and objectives associated with these dimensions that may characterize KDD. Below we discuss how FCA could be considered w.r.t. these dimensions and objectives.

Starting from a binary table or context "object $\times$ attributes", FCA allows the discovery of concepts, where each concept materializes two interrelated views. Concepts have an extent which is composed of a set of objects and which stands for a class of individuals. Concepts have an intent which is composed of a set of attributes and which stands for a (class) description. Moreover, concepts can be organized within a poset, actually a concept lattice, based on subsumption relation. Thanks to the double view of concepts, the concept lattice can be related to a hierarchy of concepts in Description Logics [4]. In addition, from the set of concepts, implications and association rules can be discovered and reused for knowledge discovery, representation and explanation purposes (see practical examples, e.g., in [7,14,20,26,25]).

Now, when it is necessary to deal with complex structured data, pattern structures [24,38] extend the capabilities of FCA, while keeping the good properties of FCA. Many applications are based on pattern structures where object descriptions are intervals, sequences, trees and graphs [36]. Another extension of FCA, namely Relational Concept Analysis [45], allows to explicitly deal with relations between objects and to build object intents with relational attributes.

Hence FCA is naturally exploratory and exploration is based on the concept lattice structure. The exploration is often data-directed but there are some attempts to design pattern-directed processes [13]. The exploration is also aimed at selecting concepts of interest w.r.t. interestingness measures such as stability for example [12,40]. More recently, there is a trend of research on exploration directed by the MDL principle [53,42], exhibiting the links existing with subgroup discovery and exceptional model mining. In particular the selection of initial seeds and the construction of good object coverings are important questions.

Continuing on exploration and interaction, there is a whole line of work on the use of graphical tools for interacting with the lattice structure, for visualizing concepts, and selecting concepts of interest [21,3]. Visualization has always been a main concern for FCA practitioners as the lattice structure can be rather easily understood and interpreted. In the same way, the exploration of the concept lattice allows different types of information retrieval [16,15] and related operations such as recommendation and biclustering [37,35].

But FCA is also knowledge-based as it allows, especially in the case of pattern structures and RCA, to take into account domain knowledge under different forms. For example there is a number of studies on the use of attribute hierarchies within a context (this was initiated in [14], actualized in [24], and then in [3]). Furthermore, there are also other attempts for discovering definitions in the web of data that can be reused as concept definitions in ontologies and knowledge graphs [2].

Furthermore, FCA shows many links with knowledge engineering, while, until now, there does not exist any type of meta-level in FCA. Such a meta-level could take several forms, for example introducing a kind of meta-concepts for providing knowledge about the management of the concepts, or meta-rules for generating or controlling implications and association rules. Still on knowledge construction, we should mention attribute exploration [25] which enables the completion of a context and the exploration of rule sets and concept sets among others.

However, FCA is not hybrid and does not offer any explanation facility strictly speaking. There are some experiments already mentioned showing that FCA can be combined with numerical machine learning methods for performing knowledge discovery tasks. In [29] the concept lattice is also used in an interactive way for concept selection and in a certain sense for providing plausible explanations. Other attempts on hybridization can be found in [33,54].

To conclude, let us mention some research topics that have not received too much attention within FCA, and that we think should deserve more attention in the future, namely the construction of a meta-level, the hybridization with numerical methods, and the production of explicit explanations. This is a vast program, and a series of articles will hopefully emerge in a near future to tackle these open subjects.

References

1. Charu C. Aggarwal and Jiawei Han, editors. *Frequent Pattern Mining*. Springer, 2014.
2. Mehwish Alam, Aleksey Buzmakov, Víctor Codocedo, and Amedeo Napoli. Mining Definitions from RDF Annotations Using Formal Concept Analysis. In Qiang Yang and Michael Wooldridge, editors, *Proceedings of IJCAI*, pages 823–829. AAAI Press, 2015.
3. Mehwish Alam, Aleksey Buzmakov, and Amedeo Napoli. Exploratory Knowledge Discovery over Web of Data. *Discrete Applied Mathematics*, 249:2–17, 2018.
4. Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, Cambridge, UK, 2003.
5. Aimene Belfodil, Adnene Belfodil, and Mehdi Kaytoue. Anytime Subgroup Discovery in Numerical Domains with Guarantees. In Michele Berlingerio, Francesco Bonchi, Thomas Gärtner, Neil Hurley, and Georgiana Ifrim, editors, *Proceedings of ECML-PKDD*, LNCS 11052, pages 500–516. Springer, 2018.
6. Ahmed Anes Bendimerad, Marc Plantevit, and Céline Robardet. Mining exceptional closed patterns in attributed graphs. *Knowledge Information Systems*, 56(1):1–25, 2018.
7. Karell Bertet, Christophe Demko, Jean-François Viaud, and Clément Guérin. Lattices, closures systems and implication bases: A survey of structural aspects and algorithms. *Theoretical Compututer Science*, 743:93–109, 2018.
8. Tijl De Bie. Subjective Interestingness in Exploratory Data Mining. In Allan Tucker, Frank Höppner, Arno Siebes, and Stephen Swift, editors, *Proceedings of IDA*, LNCS 8207, pages 19–31. Springer, 2013.
9. Hendrik Blockeel. Data Mining: From Procedural to Declarative Approaches. *New Generation Computing*, 33(2):115–135, 2015.
10. Ronald J. Brachman and Tej Anand. The Process of Knowledge Discovery in Databases. In Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padraic Smyth, and Ramasamy Uthurusamy, editors, *Advances in Knowledge Discovery and Data Mining*, pages 37–57. AAAI Press/MIT Press, 1996.
11. Pavel Brazdil, Christophe G. Giraud-Carrier, Carlos Soares, and Ricardo Vilalta. *Metalearning – Applications to Data Mining*. Cognitive Technologies. Springer, 2009.
12. Aleksey Buzmakov, Sergei O. Kuznetsov, and Amedeo Napoli. Scalable Estimates of Stability. In Cynthia Vera Glodeanu, Mehdi Kaytoue, and Christian Sacarea, editors, *Proceedings of ICFCA*, LNAI 8478, pages 157–172. Springer, 2014.
13. Aleksey Buzmakov, Sergei O. Kuznetsov, and Amedeo Napoli. Fast Generation of Best Interval Patterns for Nonmonotonic Constraints. In Annalisa Appice, Pedro Pereira Rodrigues, Vítor Santos Costa, João Gama, Alípio Jorge, and Carlos Soares, editors, *Proceedings of ECML-PKDD*, LNCS 9285, pages 157–172. Springer, 2015.
14. Claudio Carpineto and Giovanni Romano. *Concept Data Analysis: Theory and Applications*. John Wiley & Sons, Chichester, UK, 2004.
15. Víctor Codocedo, Ioanna Lykourentzou, and Amedeo Napoli. A semantic approach to concept lattice-based information retrieval. *Annals of Mathematics and Artificial Intelligence*, 72:169–195, 2014.
16. Víctor Codocedo and Amedeo Napoli. Formal Concept Analysis and Information Retrieval – A Survey. In Jaume Baixeries, Christian Sacarea, and Manuel Ojeda-Aciego, editors, *Proceedings of ICFCA*, LNCS 9113, pages 61–77. Springer, 2015.

17. Artur S. d'Avila Garcez, Tarek R. Besold, Luc De Raedt, Peter Földiák, Pascal Hitzler, Thomas Icard, Kai-Uwe Kühnberger, Luís C. Lamb, Risto Miikkulainen, and Daniel L. Silver. Neural-Symbolic Learning and Reasoning: Contributions and Challenges. In *AAAI Spring Symposium*, 2015.
18. Thomas G. Dietterich. Ensemble Methods in Machine Learning. In Josef Kittler and Fabio Roli, editors, *First International Workshop on Multiple Classifier Systems (MCS)*, LNCS 1857, pages 1–15. Springer, 2000.
19. Wouter Duivesteijn, Ad Feelders, and Arno J. Knobbe. Exceptional Model Mining - Supervised descriptive local pattern mining with complex target concepts. *Data Mining and Knowledge Discovery*, 30(1):47–98, 2016.
20. Vincent Duquenne. Latticial Structures in Data Analysis. *Theoretical Computer Science*, 217:407–436, 1999.
21. Peter W. Eklund and Jean Villerd. A Survey of Hybrid Representations of Concept Lattices in Conceptual Knowledge Processing. In Léonard Kwuida and Baris Sertkaya, editors, *Proceedings of ICFCA*, LNCS 5986, pages 296–311. Springer, 2010.
22. Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
23. Benhard Ganter and Rudolf Wille. *Formal Concept Analysis - Mathematical Foundations*. Springer, 1999.
24. Bernhard Ganter and Sergei O. Kuznetsov. Pattern Structures and Their Projections. In Harry S. Delugach and Gerd Stumme, editors, *Proceedings of ICCS*, LNCS 2120, pages 129–142. Springer, 2001.
25. Bernhard Ganter and Sergei A. Obiedkov. *Conceptual Exploration*. Springer, 2016.
26. Bernhard Ganter, Gerd Stumme, and Rudolf Wille, editors. *Formal Concept Analysis, Foundations and Applications*, LNCS 3626. Springer, 2005.
27. Nina Grgic-Hlaca, Muhammad Bilal Zafar, Krishna P. Gummadi, and Adrian Weller. Beyond Distributive Fairness in Algorithmic Decision Making: Feature Selection for Procedurally Fair Learning. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of AAAI-18*, pages 51–60. AAAI Press, 2018.
28. Dhouha Grissa, Blandine Comte, Mélanie Pétéra, Estelle Pujos-Guillot, and Amedeo Napoli. A hybrid and exploratory approach to knowledge discovery in metabolomic data. *Discrete Applied Mathematics*, 2019. To be published.
29. Dhouha Grissa, Blandine Comte, Estelle Pujos-Guillot, and Amedeo Napoli. A Hybrid Knowledge Discovery Approach for Mining Predictive Biomarkers in Metabolomic Data. In *Proceedings of ECML-PKDD*, LNCS 9851, pages 572–587. Springer, 2016.
30. Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, 51(5), 2018.
31. Melanie Hilario, Phong Nguyen, Huyen Do, Adam Woznica, and Alexandros Kalousis. Ontology-Based Meta-Mining of Knowledge Discovery Workflows. In *Meta-Learning in Computational Intelligence*, pages 273–315. Springer, 2011.
32. Andreas Holzinger, Matthias Dehmer, and Igor Jurisica. Knowledge Discovery and interactive Data Mining in Bioinformatics – State-of-the-Art, future challenges and research directions. *BMC Bioinformatics*, 15(S-6):I1, 2014.
33. Anna Hristoskova, Veselka Boeva, and Elena Tsiporkova. An Integrative Clustering Approach Combining Particle Swarm Optimization and Formal Concept Analysis. In Christian Böhm, Sami Khuri, Lenka Lhotská, and M. Elena Renda, editors, *Third International Conference Information on Technology in Bio- and Medical Informatics (ITBAM)*, LNCS 7451, pages 84–98. Springer, 2012.

34. Krzysztof Janowicz, Frank van Harmelen, James A. Hendler, and Pascal Hitzler. Why the Data Train Needs Semantic Rails. *AI Magazine*, 36(1):5–14, 2015.
35. Mehdi Kaytoue, Víctor Codocedo, Jaume Baixeries, and Amedeo Napoli. Three interrelated FCA methods for mining biclusters of similar values on columns. In Karell Bertet and Sebastian Rudolph, editors, *Proceedings of CLA*, pages 243–254. CEUR Workshop Proceedings 1252, 2014.
36. Mehdi Kaytoue, Víctor Codocedo, Aleksey Buzmakov, Jaume Baixeries, Sergei O. Kuznetsov, and Amedeo Napoli. Pattern Structures and Concept Lattices for Data Mining and Knowledge Processing. In Albert Bifet, Michael May, Bianca Zadrozny, Ricard Gavaldà, Dino Pedreschi, Francesco Bonchi, Jaime S. Cardoso, and Myra Spiliopoulou, editors, *Proceedings of ECML-PKDD*, LNCS 9286, pages 227–231. Springer, 2015.
37. Mehdi Kaytoue, Sergei O. Kuznetsov, Juraj Macko, and Amedeo Napoli. Biclustering meets triadic concept analysis. *Annals of Mathematics and Artificial Intelligence*, 70(1–2):55–79, 2014.
38. Mehdi Kaytoue, Sergei O. Kuznetsov, Amedeo Napoli, and Sébastien Duplessis. Mining Gene Expression Data with Pattern Structures in Formal Concept Analysis. *Information Science*, 181(10):1989–2001, 2011.
39. Mehdi Kaytoue, Marc Plantevit, Albrecht Zimmermann, Ahmed Anes Bendimerad, and Céline Robardet. Exceptional contextual subgraph mining. *Machine Learning*, 106(8):1171–1211, 2017.
40. Sergei O. Kuznetsov and Tatiana P. Makhalova. On interestingness measures of formal concepts. *Information Sciences*, 442-443:202–219, 2018.
41. Nada Lavrac, Branko Kavsek, Peter A. Flach, and Ljupco Todorovski. Subgroup Discovery with CN2-SD. *Journal of Machine Learning Research*, 5:153–188, 2004.
42. Tatiana P. Makhalova, Sergei O. Kuznetsov, and Amedeo Napoli. A First Study on What MDL Can Do for FCA. In Dmitry I. Ignatov and Lhouari Nourine, editors, *Proceedings of CLA*, CEUR Workshop Proceedings 2123, pages 25–36, 2018.
43. Phong Nguyen, Melanie Hilario, and Alexandros Kalousis. Using Meta-mining to Support Data Mining Workflow Planning and Optimization. *Journal of Artificial Intelligence Research (JAIR)*, 51:605–644, 2014.
44. Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu C. Aggarwal, Dou Shen, and Rajeev Rastogi, editors, *Proceedings of SIGKDD*, pages 1135–1144. ACM, 2016.
45. Mohamed Rouane-Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. Relational concept analysis: mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67(1):81–108, 2013.
46. Omer Sagi and Lior Rokach. Ensemble Learning: A Survey. *Wiley Interdisciplinary Review on Data Mining and Knowledge Discovery*, 8(4), 2018.
47. Gustav Sourek, Vojtech Aschenbrenner, Filip Zelezný, Steven Schockaert, and Ondrej Kuzelka. Lifted Relational Neural Networks: Efficient Learning of Latent Relational Structures. *Journal of Artificial Intelligence Research*, 62:140–151, 2018.
48. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar. *Introduction to Data Mining (Second Edition)*. Pearson, 2018.
49. Son N. Tran and Artur S. d'Avila Garcez. Deep Logic Networks: Inserting and Extracting Knowledge From Deep Belief Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 29(2):246–258, 2018.
50. John W. Tukey. *Exploratory Data Analysis.* Addison-Wesley Publishing Company, 1977.

51. Willy Ugarte, Patrice Boizumault, Bruno Crémilleux, Alban Lepailleur, Samir Loudni, Marc Plantevit, Chedy Raïssi, and Arnaud Soulet. Skypattern mining: From pattern condensed representations to dynamic constraint satisfaction problems. *Artificial Intelligence*, 244:48–69, 2017.
52. Matthijs van Leeuwen. Interactive Data Exploration Using Pattern Mining. In Andreas Holzinger and Igor Jurisica, editors, *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics – State-of-the-Art and Future Challenges*, LNCS 8401, pages 169–182. Springer, 2014.
53. Jilles Vreeken and Nikolaj Tatti. Interesting Patterns. In Aggarwal and Han [1], pages 105–134.
54. Yuka Yoneda, Mahito Sugiyama, and Takashi Washio. Learning graph representation via formal concept analysis. *CoRR*, abs/1812.03395, 2018.
55. Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, Krishna P. Gummadi, and Adrian Weller. From Parity to Preference-based Notions of Fairness in Classification. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Proceedings of NIPS*, pages 228–238, 2017.
56. Mohamed J. Zaki and Wagner Meira Jr. *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, NY, USA, 2014.