Bayesian Networks : A State-Of-The-Art Survey

Nelda Kote
Department of Fundamentals of Computer Science, Faculty of Information Technology, Polytechnic University of Tirana, Albania
nkote@fti.edu.al

Marenglen Biba
Department of Computer Science, Faculty of Information Technology, New York University of Tirana, Albania
marenglenbiba@unyt.edu.al

Elena Canaj
Department of Computer Engineering, Faculty of Information Technology, Polytechnic University of Tirana, Albania
elenacanaj@gmail.com

Abstract

Over the last decade, Bayesian Networks (BNs) have become an increasingly popular Artificial Intelligence approach. BNs are a widely used method in the modelling of uncertain knowledge. There have been many important new developments in this field. This paper presents a review and classification scheme for recent researches on Bayesian Networks. This is achieved by reviewing relevant articles published in the recent years. The articles are classified based on a scheme that consists of three main Bayesian Networks topics: Bayesian Networks Structure Learning, Advanced Application of Bayesian Networks and Bayesian Network Classifiers. This review provides a reference source and classification scheme for researchers interested in BNs, and indicates under-researched areas as well as future directions.

1 Introduction

This paper presents a review of recent researches in the area of Bayesian Networks. BNs are a popular class of probabilistic graphical models for researches and applications in the field of Artificial Intelligence. BNs are built on Bayes' theorem and allow to represent a joint probability distribution over a set of variables in the network. In Bayesian probabilistic inference, the joint distribution over the set of variables in a Bayesian Network can be used to calculate the probabilities of any configuration of these variables given fixed values of another set of variables, called observations or evidence [Rus09].

Bayesian Networks can be built from human knowledge, i.e. from theory, or, they can be machine learned from data. Thus, they cover the entire spectrum in terms of their model source. Also, due to their graphical structure, machine-learned Bayesian Networks are intuitively interpretable, thus facilitating human learning and theory building. Bayesian Networks allow human learning and machine learning to interact efficiently. This way, Bayesian Networks can be developed from a combination of human and artificial intelligence.

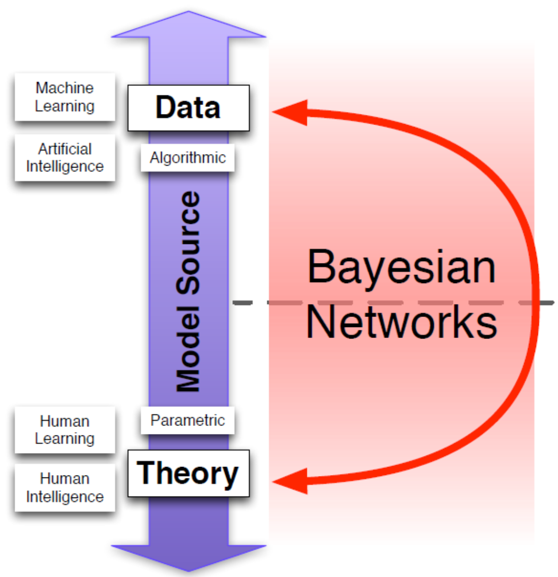

Figure 1: Bayesian Networks spanning theory and data [Con13]

Figure 1 illustrates the role and position of Bayesian Networks between theory and data in Artificial Intelligence. This paper addresses most of the recent research works of three main **Bayesian Networks fields: Bayesian Networks Structure Learning, Advanced Application of Bayesian Networks and Bayesian Networks Classifiers**.

The structure of the paper is as follows: Section 2 presents the research methodology; Section 3 presents results and analysis of the searches in a quantitative perspective; Section 4 gives the detailed description and evaluation of the reviewed papers and finally we conclude our work in Section 5.

2 Methodology

The scope of this review is to identify and evaluate the recent research fields on Bayesian Networks. Over fifty papers were first extracted from searches made on three major research databases for computer science: IEEE Xplore, CiteSeerX and Google Scholar, for the following keywords: Bayesian Networks, data classification, learning structure, data mining, Bayes Theorem. The date range for this search was limited from 2011 until 2018. We kept our scope wider to consider all topics of Bayesian Networks. The challenges related to the structure learning methods and algorithms, implementation of different applications and classification methods and algorithms were all within the scope of this review paper. The citation-references of the selected papers were checked, and additional papers were found to be necessary to add to this review based on the criteria mentioned above. From the numerous research publications, around thirty papers were selected for this review.

The papers are categorized based on their main focus in three groups: Bayesian Networks Structure Learning, Advanced Application of Bayesian Networks and Bayesian Networks Classifiers.

3 Literature Review: Quantitative Results and Analysis

In this section we present the results of our study, based on the methodology explained in section 2. All thirty selected publications are analyzed and evaluated based on their research contributions. The articles are noted by their type as Review, Survey, Improvements in existing Technology, New Proposal and special attention is given to real experiments, simulation/emulation and system implementation made by authors. Table 3 shows all the selected papers for this review.

Based on the classification scheme, we give the results on the total number of publications per domain and their percentage on the total numbers of reviewed papers, shown in table 1. The results that we found is that researches are equally focused on these three main BNs fields.

Table 1: Distribution of covered network aspects

Domain	**Total no. of publications**	**Percentage Papers / Domain (%)**
BNs Structure Learning	9	30
Application of BNs	10	33
BNs Classifiers	11	37
Total	**30**	**100**

Table 2 shows the total number of publications per year and their percentage on the total numbers of reviewed papers.

Table 2: Representation of the total number of publications per year

Year	**Total nr. of publications**	**Percentage Papers /Year (%)**
2011	3	10
2012	2	7
2013	4	13
2014	4	13
2015	6	20
2016	3	10
2017	5	17
2018	3	10
Total	**30**	**100**

Table 3: List of articles

Article	BNs Learning Structure	Application of BNs	BNs Classifiers	Article Type	Measures / Experiments
[Kos12]	Main Topic	X		Review	
[Dal11]	Main Topic	X		Review/Improvements	
[Mal15]	Main Topic	X		Review/Improvements	X
[Zha14]	Main Topic	X		Improvements/New Design	X
[Mil15]	Main Topic	X		New Design	X
[Tsc15]	Main Topic	X	X	New Design	X
[Li17]	Main Topic	X		New Design	X
[Kar16]	Main Topic	X		New Design	X
[Zha18]	Main Topic	X		New Design	X
[Per14]		Main Topic		Improvements/New Design	X
[Yua11]	X	Main Topic		Review/Improvements	X
[Vle15]	X	Main Topic		New Design	X
[Oku12]		Main Topic		Review/Improvements	
[Kle15]	X	Main Topic		Review/Improvements	X
[Cay11]		Main Topic		New Design	X
[Lan13]		Main Topic		Review/Improvements	X
[Ren13]		Main Topic		New Design	X
[Urs17]	X	Main Topic		New Design	X
[Wee18]		Main Topic		New Design	X
[Bie14]		X	Main Topic	Survey	
[Ang16]	X	X	Main Topic	New Design	X
[Vij13]			Main Topic	Comparative Analysis	X
[Suc14]		X	Main Topic	New Design	X
[Cho16]		X	Main Topic	Review/Improvements	X
[Liu13]		X	Main Topic	New Design	X
[Tsc15]		X	Main Topic	New Design	X
[Xu17]		X	Main Topic	New Design	X
[Ans17]		X	Main Topic	Improvements	X
[Kan17]		X	Main Topic	Improvements	X
[Wu18]		X	Main Topic	Improvements/Comparative Analysis	X

4 Literature Review: Topics-Related Analysis

This section is an overview of each of the domains. The 30 publications are mapped based on main topic and their contributions on Bayesian Networks, as well as references and possible analysis. First, we give a general description of Bayesian Networks.

4.1 The Bayesian Network

A Bayesian Network is a form of probabilistic graphical model. Structurally, a Bayesian Network is a directed acyclic graph where nodes represent variables and arcs represent dependency relations between the variables (nodes). An arc from node A to another node B is called: A is a parent of B. A node can represent any kind of random variable.

A Bayesian network with parameters is a graphical representation of the joint distribution over all the variables represented by nodes in the graph. If the variables are $X_1,..., X_n$ we let "parents(A)" be the parents of the node A. Then the joint distribution for X_1 through X_n is represented as the product of the probability distributions:

$$P(X_1, ... , X_n) = P(X_i \ \text{parents}\ (X_i)) \text{ for } i = 1 \text{ to } n.$$

To fully specify the Bayesian Network and to carry out numerical calculations, it is necessary to further specify for each node X the probability distribution for X conditional on its parents. In this way a Bayesian Network could be used to perform any probabilistic inference over the domain variables [Rus09].

Important usage of Bayesian Networks is made in modeling, where the structure of the Bayesian network is generated by software. Learning the structure of a Bayesian Network is a very important task in machine learning. To find the structure of the network, a scoring function should be maximized through a search algorithm. We review this topic in section 4.2.

Bayesian Networks are used for modeling knowledge in many domains with uncertain knowledge, like medicine, engineering, text analysis, image processing, data fusion, decision support systems, and data classification. The recent researches on these topics are reviewed in sections 4.3 and 4.4.

4.2 Bayesian Networks Structure Learning

In this section we review the recent research works in Bayesian Networks structure learning and analyze their characteristics. We have reviewed nine papers in terms of Bayesians Network Structure Learning.

Bayesian Networks Structure Learning problem takes the data as input and produces a directed acyclic graph as the output. There are roughly three main approaches to the learning problem: score-based learning, constraint-based learning, and hybrid methods. These approaches are reviewed in detail in three papers [Kos12], [Dal11], [Mal15]. Score-based learning methods evaluate the quality of Bayesian Network structures using a scoring function and select the one that has the best score. These methods basically formulate the learning problem as a combinatorial optimization problem. They work well for datasets with not too many variables but may fail to find optimal solutions for large datasets. Constraint-based learning methods typically use statistical tests to identify conditional independence relations from the data and build a Bayesian Network structure that best fits those independence relations. Constraint-based methods mostly rely on results of local statistical tests, so they can often scale to large datasets. However, they are sensitive to the accuracy of the statistical tests and may not work well when there are insufficient or noisy data. In comparison, score-based methods work well even for datasets with relatively few data points. Hybrid methods aim to integrate the advantages of the previous two approaches and use combinations of constraint-based and/or score-based methods for solving the learning problem. One popular strategy is to use constraint-based learning to create a skeleton graph and then use score-based learning to find a high-scoring network structure that is a subgraph of the skeleton.

Authors in [Kos12] and [Dal11] take a broad look at the literature on learning Bayesian Networks in particular their structure from data.

Authors in [Mal15] present results from an empirical evaluation of the impact of Bayesian Network structure learning strategies on the learned structures. They investigate how learning algorithms with different optimality guarantees compare in terms of structural aspects and generalizability of the produced network structures.

Articles [Zha14], [Mil15], [Tsc15], [Li17], [Kar16], [Zha18] give further details on learning structures and evaluate algorithms used for data learning.

Authors in [Zha14] aim to provide a timely review on this area with emphasis on state-of-the-art multi-label learning algorithms. Firstly, fundamentals on multi-label learning including formal definition and evaluation metrics are given. Secondly and primarily, eight representative multi-label learning algorithms are scrutinized under common notations with relevant analyses and discussions. Thirdly, several related learning settings are briefly summarized.

In the work presented in [Mil15] a set of experiments are performed to compare the performance of two Bayesian Student Models, whose parameters have been specified by experts and learnt from data respectively. Results show that both models are able to provide reasonable estimations for knowledge variables in the student model, in spite of the small size of the dataset available for learning the parameters.

Article [Tsc15] presents generative and discriminative learning algorithms for Bayesian network classifiers relying only on reduced-precision arithmetic. For several standard benchmark datasets, these algorithms achieve classification-rate performance close to that of Bayesian Network classifiers with parameters learned by conventional algorithms using double precision floating-point arithmetic.

Authors in [Li17] by combining the advantages of constraint-based and score-based algorithms, proposed a hybrid distributed Bayesian Network structure learning algorithm from large-scale dataset using MapReduce. The algorithm reuses the statistical results of MapReduce that makes it possible for learning structures accurately. The experimental results show that the proposed solution has good results in both efficiency and accuracy.

In [Kar16], the authors proposed a new approach to accelerate the exact structure learning of Bayesian Networks. This approach leverages relationship between a partial network structure and the remaining variables to constrain the number of ways in which the partial network can be optimally extended. Experimental results show that the proposed method performs extremely well in practice, even though it does not improve the worst-case complexity.

Authors in [Zha18] present a new algorithm for learning BNs based on the hybrid ACO and differential evolution (DE). In this hybrid algorithm, the entire ant colony is divided into different groups, among which DE operators are adopted to lead the evolutionary process. Experimental results show that this algorithm outperforms the basic ACO in learning BN structure in terms of convergence and accuracy.

At the end we observed that score-based exact structure learning has become an active research topic in recent years. In this context, a scoring function is used to measure the goodness of the data fitting a structure. The goal is to find the structure which optimizes the scoring function, and it has been shown a NP-hard problem.

4.3 Application of Bayesian Networks

Bayesian Networks are used for modeling knowledge in many domains with uncertain knowledge, like medicine, engineering, text analysis, image processing, data fusion, decision support systems, and data classification. Ten papers that address different application of BN are reviewed in this section.

The first article of this domain [Per14], presents an approach to directly infer individual differences related to subjective mental representations within the framework of Bayesian models of cognition. In this approach, Bayesian data analysis methods are used to estimate cognitive parameters and motivate the inference process within a Bayesian cognitive model. Authors illustrate this integrative Bayesian approach on a model of memory. They apply the model to behavioral data from a memory experiment involving the recall of heights of people. A cross-validation analysis shows that the Bayesian memory model with inferred subjective priors predicts withheld data better than a Bayesian model where the priors are based on environmental statistics. In addition, the model with inferred priors at the individual subject level led to the best overall generalization performance, suggesting that individual differences are important to consider in Bayesian models of cognition.

Authors in [Yua11] introduce a method called Most Relevant Explanation (MRE) which finds a partial instantiation of the target variables that maximizes the generalized Bayes factor (GBF) as the best explanation for the given evidence. This study shows that GBF has several theoretical properties that enable MRE to automatically identify the most relevant target variables in forming its explanation. In particular, conditional Bayes factor (CBF), defined as the GBF of a new explanation conditioned on an existing explanation, provides a soft measure on the degree of relevance of the variables in the new explanation in explaining the evidence given the existing explanation. As a result, MRE is able to automatically prune less relevant variables from its explanation. Authors show that CBF is able to capture well the explaining-away phenomenon that is often represented in Bayesian networks. Moreover, they define two dominance relations between the candidate solutions and use the relations to generalize MRE to find a set of top explanations that is both diverse and representative. Case studies on several benchmark diagnostic Bayesian networks show that MRE is often able to find explanatory hypotheses that are not only precise but also concise.

The article [Vle15] proposes to combine Bayesian Networks with a narrative approach to reasoning with legal evidence, the result of which allows a juror to reason with alternative scenarios while also incorporating probabilistic information. The proposed method aids both the construction and the understanding of Bayesian networks, using scenario schemes.

Authors in [Oku12] use Bayesian Networks to determine the probabilistic influential relationships among software metrics and defect proneness.

In [Kle15] authors have made a systematic review that investigates the psychometric analysis of performance data of simulation-based assessment (SBA) and game-based assessment (GBA).

In [Cay11], Bayesian networks are used to extract the effects of data mining algorithm parameters on the final model obtained, both in terms of efficiency and efficacy in a given situation. Based on this knowledge, authors propose to infer future algorithm configurations appropriate for situations. Instantiation of the approach for association rules is also shown in the paper and the feasibility of the approach is validated by the experimentation.

Authors in [Lan13] review several BN-based ecosystem service (ESS) models developed in the last decade. A SWOT analysis highlights the advantages and disadvantages of BNs in ESS modelling and pinpoints remaining challenges for future research. The existing BN models are suited to describe, analyze, predict and value ESS. Nevertheless, some weaknesses must be considered, including poor flexibility of frequently applied software packages, difficulties in eliciting expert knowledge and the inability to model feedback loops.

In [Ren13] the authors used a hierarchical Bayesian network to build a model for the analysis of the human beings' emotions. It finds complex emotions in the document by establishing a relationship between the topic modeling and analyzing the emotions. The experimental results show that the proposed method has good performance and can be used in complex domains.

Authors in [Urs17] proposed to use a Bayesian networks mathematical model to evaluate the software quality, from the reliability point of view. This model evaluates the reliability of a software system for EMS (Energy Management Systems) and DMS (Distribution Management System) that are the core of national energy system as they are used the National Dispatch Control Center. To evaluate the performance of the proposed approach the authors perform a simulation to obtain some practical results and draw important conclusions if this model can improve the EMS and DMS software systems.

In [Wee18] is used a combination of Bayesian Network and fuzzy cognitive maps (FCM) for modeling and analyzing network intrusions. First, the BN is learnt from network intrusion data; following this, an FCM is generated from the BN, using a migration method. The proposed method of network intrusion analysis using both BN and FCM consists of several stages, in order to leverage the capabilities of each approach in building the causal model and performing causal analysis.

The application of Bayesian networks for modeling knowledge domain is most researched on recent years. 50% of our reviewed papers, implement and propose new designs for modeling knowledge in many domains with uncertain knowledge.

4.4 Bayesian Network Classifiers

This section reviews the theory and implementation of Bayesian Networks in the context of classification. Bayesian networks provide a very general and yet effective graphical language for factoring joint probability distributions which in turn make them very popular for classification.

Figure 2 depicts the possible structure of a Bayesian network used for classification. The dotted lines denote potential links, and the blue box is used to indicate that additional nodes and links can be added to the model, usually between the input and output nodes. In order to perform classification with a Bayesian Network such as the one depicted in Figure 2, first evidence must be set on the input nodes, and then the output nodes can be queried using standard Bayesian network inference. The result will be a distribution for each output node, so that you can not only determine the most probable state for each output, but also see the probability assigned to each output state. [Xu13]

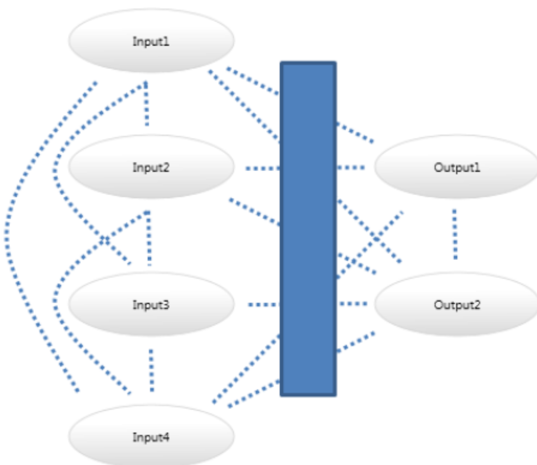

Figure 2: Generic structure of a Bayesian Network classifier [Xu13]

Authors in [Bie14] survey the whole set of discrete Bayesian Network classifiers devised to date, organized in increasing order of structure complexity: Naive Bayes, selective Naive Bayes, Seminaive Bayes, One Dependence Bayesian classifiers, k-dependence Bayesian classifiers, Bayesian network-augmented naive Bayes, Markov blanket-based Bayesian classifier, unrestricted Bayesian classifiers, and Bayesian multinets. Issues of feature subset selection and generative and discriminative structure and parameter learning are also covered.

In [Ang16], the authors show the accuracy of a General Bayesian Network (GBN) used with the Hill-Climbing learning method, which does not impose any restrictions on the structure and better represents the dataset. The results show that it gives equivalent performances or even outperforms Naive Bayes and Tree Augmented Naive Bayes in most of the data classification.

In the research work of [Vij13], authors have analyzed the performance of Bayesian and Lazy classifiers for classifying the files which are stored in the computer hard disk. There are two algorithms in Bayesian classifier namely BayesNet, and Naïve Bayes. In lazy classifier has three algorithms namely IBL, IBK and Kstar. The performances of Bayesian and lazy classifiers are analyzed by applying various performance factors. From the experimental results, it is observed that the lazy classifier is more efficient than Bayesian classifier.

In [Suc14] authors introduce a method for chaining Bayesian classifiers that combines the strengths of classifier chains and Bayesian networks for multi-label classification. A Bayesian Network is induced from data to represent the probabilistic dependency relationships between classes, constrain the number of class variables used in the chain classifier by considering conditional independence conditions, and reduce the number of possible chain orders. The effects in the Bayesian chain classifier performance of considering different chain orders, training strategies, number of class variables added in the base classifiers, and different base classifiers, are experimentally assessed. In particular, it is shown that a random chain order considering the constraints imposed by a Bayesian Network with a simple tree-based structure can have very competitive results in terms of predictive performance and time complexity against related state-of the art approaches.

Authors in [Cho16] propose the structured Naive Bayes (SNB) classifier, which augments the ubiquitous Naive Bayes classifier with structured features. SNB classifiers facilitate the use of complex features, such as combinatorial objects (e.g., graphs, paths and orders) in a general but systematic way. Underlying the SNB classifier is the recently proposed Probabilistic Sentential Decision Diagram (PSDD), which is a tractable representation of probability distributions over structured spaces. They illustrate the utility and generality of the SNB classifier via case studies. First, they show how to distinguish players of simple games in terms of play style and skill level based purely on observing the games they play. Second, they show how to detect anomalous paths taken on graphs based purely on observing the paths themselves.

In paper [Liu13], the scalability of Naıve Bayes classifier (NBC) is evaluated in large datasets. Instead of using a standard library (e.g., Mahout), authors implemented NBC to achieve fine-grain control of the analysis procedure. A Big Data analyzing system is also design for this study. The result is encouraging in that the accuracy of NBC is improved and approaches 82% when the dataset size increases. The authors have demonstrated that NBC is able to scale up to analyze the sentiment of millions movie reviews with increasing throughput.

In [Tsc15], authors investigate the effect of precision reduction of the parameters on the classification performance of Bayesian Network classifiers (BNCs). The probabilities are either determined generatively or discriminatively. Discriminative probabilities are typically more extreme. However, the results indicate that BNCs with discriminatively optimized parameters are almost as robust to precision reduction as BNCs with generatively optimized parameters. Furthermore, even large precision reduction does not decrease classification performance significantly. These results allow the implementation of BNCs with less computational complexity. This supports application in embedded systems using floating-point numbers with small bit-width. Reduced bit-widths further enable to represent BNCs in the integer domain while maintaining the classification performance.

Traditional Bayes Network classifiers have a fixed structure that are very difficult to reflect the relationships among nodes (attributes). The authors in [Xu17] proposed a self-adaptive Bayesian Network classifier based on genetic optimization. Genetic optimization is used to realize the self-adaptiveness, which means the network structure can be gradually optimized when constructing Bayesian Network classifier. Experimental results show that the proposed method leads to a high classification accuracy than traditional classifier on some benchmarks.

Authors in [Ans17] proposed a framework to detect the hypervisor attacks in virtual machines using Bayesian classifier on the publicly available dataset. They have characterized vulnerabilities of two Hypervisors XEN and VMware, based on real-time attacks. Three attributes namely authentication, integrity impact and confidentiality impact were considered for the input feature vector. Experimental results show the parameters of the used attributes that have more density for being classified as a hypervisor attack.

In [Kan17] it is proposed a model using a Bayesian classifier for airborne point cloud classification fusing multiple data types. The authors based on the analysis of the characteristics of LiDAR dataset point clouds and aerial images, they extract the geometric features from the point clouds and the spectral features from the optical images. Then the BN structure is trained using an improved mutual-information-based K2 algorithm to obtain the optimal BN classifier for point cloud classification. Experiment results show that the BN classifier can effectively distinguish four types of basic ground objects, including ground, vegetation, trees, and buildings, with a high accuracy. Moreover, compared with other classifiers, the proposed BN classifier can achieve the highest overall accuracies, and in particular, the classifier demonstrates its advantage in the classification of ground and low vegetation points.

Authors in [Wu18] to improve the safety of bus driving, classify the specific types of latent abnormal driving behavior, which include sudden braking, lane changing casually, quick turn, fast U-turn and longtime parking, and propose a method to identify the abnormal driving behavior of the bus. After collecting the data, they extract features in thirteen dimensions and then train the Naive Bayesian classifier, which is employed to detect and identify abnormal driving behaviors. They evaluate through experiments the performance of NB and support vector machine. NB has better performance than support vector machine on detecting and identifying various types of the abnormal bus driving behavior with the accuracy at 98.40%.

5 Conclusion

In this paper we have reviewed recent research work on Bayesian Networks. Over the last decade, Bayesian Networks have become an increasingly popular Artificial Intelligence approach. We have reviewed a pool of most recent works done classifying these based on a scheme that consists of three main Bayesian Networks topics: Bayesian Networks Structure Learning, Advanced Application of Bayesian Networks and Bayesian Network Classifiers. We have found that these fields are being deeply investigated and interesting approaches are being proposed in the field leading also to open directions for further potential research.

References

[Kos12] T. J.T. Koski, J. Noble. A Review of Bayesian Networks and Structure Learning. *Journals of the Polish Mathematical Society*, 40(1):51-103, 2012. doi: 10.14708/ma.v40i1.278

[Dal11] R. Daly, Q. Shen, S. Aitken. Learning Bayesian networks: approaches and issues. *The Knowledge Engineering Review*, 26 (2):99–157, 2011. doi:10.1017/S0269888910000251

[Mal15] B. Malone, M. Järvisalo, P. Myllymäki. Impact of learning strategies on the quality of Bayesian networks: an empirical evaluation. *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*: 562-571, 2015.

[Zha14] M.-L. Zhang, Zh.-H. Zhou. A Review on Multi-Label Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 26(8):1819 – 1837, 2014. doi: 10.1109/TKDE.2013.39

[Mil15] E. Millán, G. Jiménez, M.-V. Belmonte, J.-L. Pérez-de-la-Cruz. Learning Bayesian Networks for Student Modeling. *AIED: International Conference on Artificial Intelligence in Education*, *LNAI*, 9112:718–721, 2015. doi: 10.1007/978-3-319-19773-9_100

[Tsc15] S. Tschiatschek, F. Pernkopf. Parameter Learning of Bayesian Network Classifiers Under Computational Constraints. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, *LNCS*, 9284:86-101, 2015. doi:10.1007/978-3-319-23528-8 6

[Li17] Sh. Li, B. Wang. A Method for Hybrid Bayesian Network Structure Learning from Massive Data Using MapReduce. *IEEE 3rd International Conference on Big Data Security on Cloud*, 2017. doi: 10.1109/BigDataSecurity.2017.42

[Kar16] S. Karan, J Zola, Exact Structure Learning of Bayesian Networks by Optimal Path Extension. *IEEE International Conference on Big Data*, 2016. doi: 10.1109/BigData.2016.7840588

[Zhan18] X Zhang, S. Jia, X. Li, C. Guo. Learning the Bayesian networks structure based on ant colony optimization and differential evolution. *4th International Conference on Control, Automation and Robotics (ICCAR)*, 2018. doi: 10.1109/ICCAR.2018.8384700

[Per14] P. Hemmer, S. Tauber, M. Steyvers. Moving beyond qualitative evaluations of Bayesian models of cognition. *Psychon Bull Rev*, 22(3):614-28, 2014. doi: 10.3758/s13423-014-0725-z

[Yua11] C. Yuan, H. Lim, T. C. Lu. Most Relevant Explanation in Bayesian Networks. *Journal of Artificial Intelligence Research*, 42: 309-352, 2011.

[Vle15] C. Vlek, H. Prakken, S. Renooij, B. Verheij. Constructing and Understanding Bayesian Networks for Legal Evidence with Scenario Schemes. *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, 128-137, 2015, doi: 10.1145/2746090.2746097

[Oku12] A. Okutan, O. T. Yıldız. Software defect prediction using Bayesian networks. *Empirical Software Engineering*, 19(1):154–181, 2014(2012). doi: https://doi.org/10.1007/s10664-012-9218-8

[Kle15] S. de Klerk, B. P. Veldkamp, Th. J.H.M. Eggen. Psychometric analysis of the performance data of simulation-based assessment: A systematic review and a Bayesian network example. Computers & Education, 85: 23-34, 2015. Doi: https://doi.org/10.1016/j.compedu.2014.12.020

[Cay11] A. Cayci, S. Eibe, E. Menasalvas, Y Saygin. Bayesian Networks to Predict Data Mining Algorithm Behavior in Ubiquitous Computing Environments. *MUSE/MSM*, *LNAI*, 6904:119–141, 2011, Springer

[Lan13] D. Landuyt, S. Broekx, R. D'hondt, G. Engelen, J. Aertsens, P. L.M. Goethals. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling & Software*, 46:1-11, 2013,

[Ren13] F. Ren, X. Kang. Employing hierarchical Bayesian networks in simple and complex emotion topic analysis. *Computer Speech &*

Language, 27(4): 943-968, 2013. doi: https://doi.org/10.1016/j.csl.2012.07.012

[Urs17] V. Ursianu, F. Moldoveanu, R. Ursianu, E. Ursianu. Bayesian Networks Applications Extended to Evaluate the Reliability of EMS and DMS Software Systems. *21st International Conference on Control Systems and Computer Science (CSCS)*, 2017. doi: 10.1109/CSCS.2017.65

[Wee18] Y. Y. Wee, W. P. Cheah, Sh. Y. Ooi, Sh. Ch. Tan, K. Wee. Application of Bayesian belief networks and fuzzy cognitive maps in intrusion analysis. *Journal of Intelligent & Fuzzy Systems*, 35(1):111-122, 2018. doi:10.3233/JIFS-169572

[Bie14] C. Bielza, P. Larrañaga. Discrete Bayesian Network Classifiers: A Survey. *Journal ACM Computing Surveys (CSUR) Surveys*, 47(1), 2014. doi:10.1145/2576868

[Ang16] S.L. Ang, H. C. Ong, H. C. Low. Classification Using the General Bayesian Network. *Pertanika J. Science & Technology*, 24(1): 205 – 211, 2016.

[Vij13] S. Vijayarani, M. Muthulakshmi. Comparative Analysis of Bayes and Lazy Classification Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(8): 2319-5940, 2013.

[Suc14] L. E. Sucar, C. Bielza, E. F. Morales, P. Hernandez-Leal, J. H. Zaragoza, P. Larrañaga. Multi-label classification with Bayesian network-based chain classifiers. *Pattern Recognition Letters*, 41: 14-22, 2014. doi: https://doi.org/10.1016/j.patrec.2013.11.007

[Cho16] A. Choi, N. Tavabi, A. Darwiche. Structured Features in Naive Bayes Classification. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 2016

[Liu13] B. Liu, E. Blasch, Y. Chen, D. Shen. Scalable sentiment classification for big data analysis using Naive Bayes Classifier. *2013 IEEE International Conference on Big Data*, 2013. doi: 10.1109/BigData.2013.6691740

[Tsc15] S. Tschiatschek, F. Pernkopf. On Bayesian Network Classifiers with Reduced Precision Parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(4): 774 - 785, 2015. doi: 10.1109/TPAMI.2014.2353620

[Xu17] H. Xu, W. Huang, J. Wang, D. Wang. A self-adaptive Bayesian network classifier by means of genetic optimization. *8th IEEE International Conference on Software*, 2017. doi: 10.1109/ICSESS.2017.8343007

[Ans17] S. Ansari, K. Hans, S. K. Khatri. A Naive Bayes classifier approach for detecting hypervisor attacks in virtual machines. *2nd International Conference on Telecommunication and Networks (TEL-NET)*, 2017. doi: 10.1109/TEL-NET.2017.8343551

[Kan17] Zh. Kang, J. Yang, R. Zhong. A Bayesian-Network-Based Classification Method Integrating Airborne LiDAR Data with Optical Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(4): 1651 – 1661, 2017. doi: 10.1109/JSTARS.2016.2628775

[Wu18] X. Wu, J. Zhou, J. An, Y. Yang. Abnormal driving behavior detection for bus based on the Bayesian classifier. *2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)*, 2018. doi: 10.1109/ICACI.2018.8377618

[Con13] S. Conrady, L. Jouffe. *Introduction to Bayesian Networks & BayesiaLab A Practical Introduction for Researchers*. Bayesia USA, 2015.

[Rus09] S. J. Russell, P. Norvig. Artificial *Intelligence A Modern Approach Book - 3d edition*, Pearson 2009.

[Xu13] G. Xu, Y. Zong, Zhenglu Yang. *Applied Data Mining*. CRC Press Taylor & Francis Group, 2013.