Electronics and Information Technologies, 2025, 32, 55–66
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.9:519.26(075.8)

CONCEPT OF ADAPTIVE SELECTION OF STRUCTURED LEARNING ALGORITHMS OF BAYESIAN NETWORKS BASED ON THEIR CHARACTERISTICS

Mariia Voronenko,
Odesa Polytechnic National University,
Shevchenko Avenue, 1, Odesa, 65044, Ukraine

Voronenko, M.O. (2025). Concept of Adaptive Selection of Structured Learning Algorithms of Bayesian Networks Based on Their Characteristics. *Electronics and Information Technologies*, 32, 55–66. https://doi.org/10.30970/eli.32.4

ABSTRACT

Background. Modern intelligent systems require efficient mechanisms for analysis, prediction, and decision-making. Bayesian networks allow for the efficient representation of causal relationships between variables. The quality and reliability of the constructed network directly correlate with the effectiveness of the final intelligent system.

Materials and Methods. Structural training of a Bayesian network involves determining the structure of a directed acyclic graph in which variables are related to each other. The quality of the structure has a decisive impact on the ability of the model to accurately represent conditional probabilities and on the efficiency of the training algorithms and the reliability of the model. The main problem limiting the structural learning of Bayesian networks is the computational complexity of the model. This fundamental complexity means that for multidimensional problems, it is impossible to perform a complete search of all possible structures and find a global optimum. This forces reliance on heuristic search methods and approximation algorithms and creates a constant need to balance the quality of structure determination and computational resources.

Results and Discussion. The formalized concept of adaptive selection of algorithms for structural learning is based on a systematic analysis of algorithm characteristics and data properties, which allows you to choose the most suitable algorithm for a particular case, optimize the trade-off between model quality and computational resources, and increase the generalizability of the model in practical scenarios.

Conclusion. The proposed concept of adaptive selection of algorithms for structural learning is a timely contribution to the field of stochastic dependence modeling. It successfully translates the process of selecting the optimal algorithm from a routine, heuristic, brute force method to a systematic, multivariate analysis. Its full implementation has the potential to significantly increase the reliability, accuracy, and computational efficiency of building Bayesian models in complex analytical domains.

Keywords: adaptive selection concept, Bayesian network, structured learning algorithm, model accuracy

INTRODUCTION

Modern intelligent systems are increasingly used for analysis, forecasting, and decision-making in complex multifactorial environments. In conditions of incompleteness, uncertainty, and the presence of noisy data, Bayesian networks (BNs) play a special role, which provides the ability to model cause-and-effect relationships between variables and

make logical inferences. Bayesian networks are successfully used in such areas as medical diagnostics, data analysis in bioinformatics, financial forecasts, technical diagnostics, decision support systems, etc.

A key stage in building a BN is structural learning, which includes determining the graph of dependencies between variables, i.e., which variables in the network are related and in which direction the connections go. The quality of this structure has a decisive impact on the model's ability to accurately represent conditional probabilities, on the efficiency of inference algorithms, and on the ability to generalize. There are different approaches to structural learning: constraint-based, score-based, and hybrid methods, each of which has its own advantages and limitations depending on the data size, the noise level, the presence of missing values, or expert knowledge. For example, constraint-based methods analyze conditional independence between variables, while score-based approaches search for a structure that maximizes a certain value function [1]. Another important problem is computational complexity, which is the problem of structural learning that is NP-hard even with restrictions on the parent's number for each variable [1].

In this regard, there is a need to formalize the concept of algorithms adaptive selection for BN's structural learning, which would be based on a systematic analysis of the algorithm's characteristics (for example, time complexity, search space, sensitivity to noise, type of variables, missing data, expert estimates limitations) and on the data properties (number of variables, sample size, priori knowledge availability, etc.). Such an approach will allow choosing the most appropriate algorithm for a specific case, optimizing the trade-off between model quality and computational resources, and increasing the model's generalizability in practical scenarios.

Therefore, the purpose of this work is to develop a concept of adaptive algorithm selection for structural learning of Bayesian networks based on their characteristics, formalizing the criteria and parameters that determine the algorithm correspondence to the specific conditions of the problem, as well as empirical verification of the proposed approach.

The scientific novelty lies in the fact that:

1. the structured approach to the classification and structural learning algorithms comparison by a set of characteristics is proposed;
2. the set of adaptive selection criteria is defined, which considers not only the model quality, but also resources (time, memory), noise level, type of variables, and missing data;
3. the solution to the adaptive selection problem of a structural learning algorithm is proposed using the example case.

The implementation of the adaptive selection concept will contribute to increasing the efficiency of building Bayesian networks in various domains with complex data conditions, reducing time and computational costs, as well as increasing the reliability and accuracy of models.

REVIEW OF LITERATURE

Today, Bayesian models have become widespread and are comprehensively used in various industries. In [2], the authors carried out monitoring and diagnostics of a multi-stage production process using Bayesian methods. To perform medical diagnostics and differential diagnosis based on fuzzy and partially correct statistics, the authors [3] designed a static Bayesian network model. The authors [4-6] used probabilistic methods to model the supplier selection procedure based on the application of stability parameters. Dünder E., Cengiz M.A., and Koç H. investigated the impact of constraint-based algorithms on the Bayesian network structure quality in hybrid algorithms for medical research [7], and Ziegler V. investigated approximation algorithms for constrained Bayesian network structures [8]. There is also a number of studies devoted to learning the structure of Bayesian networks [9-13].

The main advantage of Bayesian networks is their robustness to incomplete, inaccurate, and noisy data [14,15]. In such complex cases, they are able to determine the most probable

outcome of events [16,17]. The study of the robustness of algorithms to noise (which is critically important), the ability to work with data with missing values, and the ability to take into account expert knowledge or a priori conditions, which are often present in practical applications, is devoted to the work [18]. Despite the large arsenal of existing methods, the question of adaptive selection of a structure learning algorithm for a specific task remains open. Many works compare the performance characteristics of methods: accuracy of structure recovery, time costs, data dimensionality, and others. For example, Scanagatta et al. (2018) analyze approximation algorithms for structure learning for large BNs, comparing the training time and quality of structure reproduction [19]. In the study [20], different score functions and heuristics are compared, and the noise influence is also studied.

PROBLEM STATEMENT

For a set of events $X^{(i)}, i = 1, \dots, N$ that are related, and a set of learning data $D = (d_1, \dots, d_n), d_i = \left\{x_i^{(1)} x_i^{(2)} \dots x_i^{(N)}\right\}$, is given. Here, the subscript is the observation amount, and the upper one is the variable amount, n–is the number of observations, each observation consists of $N(N \geq 2)$ variables, and each j -th variable $(j = 1, \dots, N)$ has $A^{(j)} = \{0,1, \dots, \alpha^{(j)} - 1\} \quad (\alpha^{(j)} \geq 2)$ conditions.

Based on a given training sample, you need to build an acyclic graph connecting the event sets $X_i,\ i = 1, \dots, N$. In addition, each BN structure $g \in G$ is represented by a set N of predecessors $\left(P^{(1)}, \dots, P^{(N)}\right)$, that is, for each vertex, $j = 1, \dots, N,\ P^{(j)}$ it is a variety of parent vertices, such that $P^{(j)} \subseteq \{X^{(1)}, \dots, X^{(N)}\} \backslash \{X^{(j)}\}$. We have events $X^{(i)}, i = 1, \dots, N$ that are affected by the uncertainties of a different nature. And also, we have data describing these events.

MATERIALS AND METHODS

The BN structure is learned using search algorithms at the 4th stage of the Bayesian network design sequence (**Fig. 1a**). Most of the existing methods for building the BN structure can be conditionally divided into two categories:

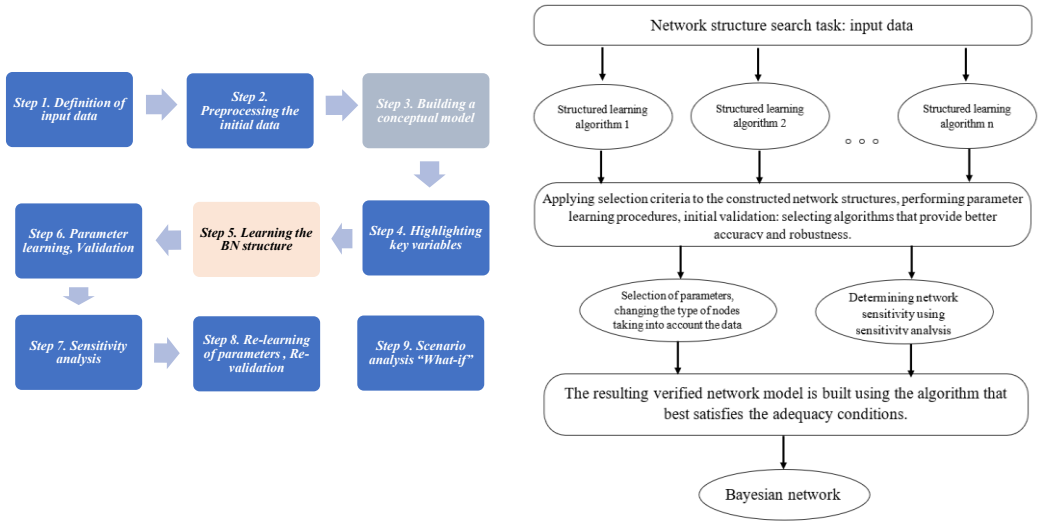
- methods based on evaluation functions (search & scoring);
- methods based on the application of the conditional independence test (dependency analysis) [21,22].

The adaptive selection concept is based on the ensemble method of selecting a structural learning algorithm that best fits the available data set (**Fig. 1b**). The ensemble method in structural learning of BN is an approach that, based on the results of several structural learning algorithms or several networks learned on the same data, allows obtaining a more accurate final structure of the network model than that built using a single algorithm.

The main idea is that different algorithms can fall into different local optima or find structures that are close in estimate, especially on limited, incomplete, or noisy data. The choice is made in favour of the model that, after repeated validation and sensitivity analysis, has the highest accuracy.

The advantages of the ensemble method are to reduce the influence of noise in the data and randomness in the selection of the initial structure. The final model demonstrates higher predictive accuracy. Due to the generation of structures by different algorithms, the ensemble is more likely to cover a wider space of structure search.

The main idea of the Path Condition (PC) algorithm is to obtain a set of conditionally independent and dependent nodes according to some statistical tests. PC checks the statistical tests for conditional independence for all variable pairs, except for the specified restrictions. An undirected link is added between each pair of variables in which conditional independence was not found.



a) b)

Fig. 1. a) Scheme of the BN's phased design, b) detailing the 5th stage "Learning the BN structure".

The undirected graph obtained in this way is called the skeleton of the learning structure [23]. Then, pairs of directed links are determined in such a way that they meet in a node, and that ensure the undirected cycles absence.

The next step is to find the directions of the links, which can be obtained from the conditional probability table. The remaining links will be arbitrarily directed, ensuring that no directed cycles will occur. In their work, Dempster and Druzdzel experimentally confirmed that the PC algorithm is quite robust to multivariate analysis. For comparison, they performed a graph structure search first using the PC algorithm and then using the Bayesian search approach [23]. Suppose we have a set of variables $X = (X_1, \ldots, X_n)$ with a global probability distribution over them P. We will denote by the letter A the subset of the variables X. By $I(A,B/C)$ we will denote that the sets A and B are conditionally independent of C. The PC algorithm assumes confidence probabilities.

This means that there exists a directed acyclic graph G such that the independence relations between the variables in X are exactly those represented in G by the d-partition criterion [23]. The PC algorithm assumes a procedure that can recognize when $I(A,B|C)$ is verified (tested) on a graph G. At first, it tries to find the skeleton which underlies the undirected graph, and in the last step, it does the edge orientation. The pseudocode of the algorithm PC can be represented as follows:

```
1. Start with a complete undirected graph G'
2. i = 0
3. Repeat
4.     For each X ∈ X
5.       For each Y ∈ ADJX
6.          Test whether ∃S ⊆ ADJX – {Y} with |S| = i and I (X, Y |S)
7.          If this set exists
8.             Make SXY = S
9.             Remove X – Y link from G'
10.         End If
11.      End For
12.   End For
13.   i = i + 1
14. Until |ADJX| ≤ i, ∀X.
```

The Greedy Thick Thinning (GTT or Greedy) structure learning algorithm is described by Cheng in [22]. GTT starts with an empty graph and repeatedly adds an edge without creating a cycle, maximizing the marginal likelihood P(D|S) until the addition of an edge leads to a positive increase (this is the “thickening” phase). It then repeatedly removes edges until the removal of an edge leads to a positive increase in P(D|S) (this is the “thinning” phase). The algorithm is quite efficient due to its susceptibility to the local maxima trap. This is a universal method for learning the structure of a graph. The probability tables are populated using expectation maximization [22]. The Greedy algorithm implements a “greedy” search for the existence of possible arcs from and the attribute variable A. If all attributes depend on the class variable, then the Bayesian network will have the structure of an extended simple Bayesian classifier. However, the purpose of Greedy is to determine which of these dependencies are really necessary. In addition, at each step, Greedy extends the revealed structure of dependencies between the class variable and the attribute variables by applying the Augmenter operator. It starts its work with an empty set of descendants and at each step adds a new descendant that is optimal in accordance with the value q of the network structure measure. The descendants of a node C are the elements of the set γ, and the set λ contains those attribute variables that do not directly depend on C. For each configuration of descendants γ, the dependencies between the attribute variables A are determined by an appropriately chosen Augmenter operator. The input information for the Augmenter operator is the dependencies between the class variable C and the attribute variables A, which are represented by the set $\gamma_{i-1} \cup \{X\}$. It learns additional dependencies between attribute variables using a set of training data D. Greedy selects the network that corresponds to the configuration of descendants λ with the largest value of the quality function q. The pseudocode of the GTT algorithm can be represented as follows:

Greedy $(D, q,$ Augmenter$)$

1. $\gamma \leftarrow \emptyset$ { a set of variables that depend on a class variable }
2. $\lambda \leftarrow \mathbf{A} = \{A_1, \ldots, A_{n-1}\}$ { a set of variables that do not depend on a class variable }
3. $\hat{q} \leftarrow -\infty$ { the highest value of the quality measure }
4. **for** $i = 1, \ldots, n-1$ **do**
5. Choose the attribute $A \in \lambda$, that maximizes the quality measure $q(B, D)$, for the network $B \leftarrow$Augmenter$(\gamma \cup \{A\}, \lambda \backslash \{A\}, D, q)$
6. **if** $q(B, D) > \hat{q}$ **then**
7. $\hat{B} \leftarrow B$
8. $\hat{q} \leftarrow q(B, D)$
9. $\gamma \leftarrow \gamma \cup \{A\}$
10. $\lambda \leftarrow \lambda \backslash \{A\}$
11. **return**

The Naive Bayes network structure learning algorithm is a structure learning method that is included in the category of structured learning algorithms only because it creates the structure and parameters of a Bayesian network directly from the data. The structure of a Naive Bayes network is not learned, but rather fixed by the assumption that the class variable is the sole parent of all remaining functional variables, and there are no other connections between the nodes of the network. The Naive Bayes structure assumes that the functions are independent of the class variable, which leads to inaccuracies. The pseudocode of the Naive Bayes algorithm can be represented as follows:

Class-Depend (S, γ, D)

1. $\hat{S} \leftarrow S$
2. **for** each variable $\gamma_i \in \gamma$ **do**

3. Add a directed arc $(C \to \gamma_i)$ to the network structure $\hat{S}$
4. **return** $\hat{S}$

The operation of the Naive Bayes Tree-Augmenter (TAN) algorithm is described by Friedman in [24,25]. The Tree Augmented structure learning algorithm starts with a Naive Bayes structure (i.e., one in which the class variable is the only parent of all remaining object variables) and adds a relationship between the feature variables to account for the possible dependency between them due to the class variable. The algorithm restricts only one additional parent of each feature variable (except for the class variable, which is the parent of each feature variable). Note that the Naive Bayes structure assumes independence of the feature from the class variable, which leads to inaccuracies when they are not independent. The TAN algorithm is simple and has been found to perform reliably better than Naive Bayes. The pseudocode of the Tree Augmented structure learning algorithm can be represented as follows:

Tree-Augmenter $(D, q,$ Augmenter$)$

1. $G \leftarrow \emptyset$
2. **for** each pair of variables $\{A_i, A_j\} \subset \gamma \cup \lambda$ such that $A_i \neq A_j$ **do** $w_{ij} \leftarrow I_\gamma(A_i A_j)$
3. add an undirected arc $(A_i - A_j)$ to the graph G
4. $T \leftarrow Maximum - Spanning - Tree(G, \boldsymbol{w})$
5. Order the arcs of an undirected tree T, by choosing one node as the root and setting the directions of all arcs from it. Then transform it into a Bayesian structure S.
6. $\hat{B} \leftarrow Class - Depend(S, \gamma, D)$
7. return $\hat{B}$

The Augmented Naive Bayes (ANB) structure learning algorithm is a semi-scientific structure learning method based on the Bayesian search approach. The ANB algorithm starts with a Naive Bayes framework (i.e., in which the class variable is the sole parent of all remaining feature variables) and adds a link between the feature variables to account for possible dependency between them due to the class variable. There is no limit to the number of additional connections included in each of the feature variables unless imposed by one of the algorithm's parameters. Note that the Naive Bayes framework assumes that the features are independent of the class variable, which leads to inaccuracies when they are not independent. The pseudocode of the ANB algorithm can be represented as follows:

ANB $(D, q,$ Augmenter$)$

1. $\gamma \leftarrow \emptyset$ { a set of variables that depend on a class variable }
2. $\lambda \leftarrow \boldsymbol{A} = \{A_1, \ldots, A_{n-1}\}$. {a set of variables that do not depend on a class variable }
3. $\hat{q} \leftarrow -\infty$ {the highest value of the quality measure }
4. **for** $i = 1, \ldots, n-1$ **do**
5. choose the attribute $A \in \lambda$, that maximizes the quality measure $q(B, D)$, for the network $B \leftarrow$Augmenter$(\gamma \cup \{A\}, \{A\}, \emptyset, D)$
6. **if** $q(B, D) > \hat{q}$ **then**
7. $\hat{B} \leftarrow B$
8. $\hat{q} \leftarrow q(B, D)$
9. $\gamma \leftarrow \gamma \cup \{A\}$
10. $\lambda \leftarrow \lambda \backslash \{A\}$
11. **return**

Fig. 2 presents the results of designing Bayesian network models using five structured learning algorithms.

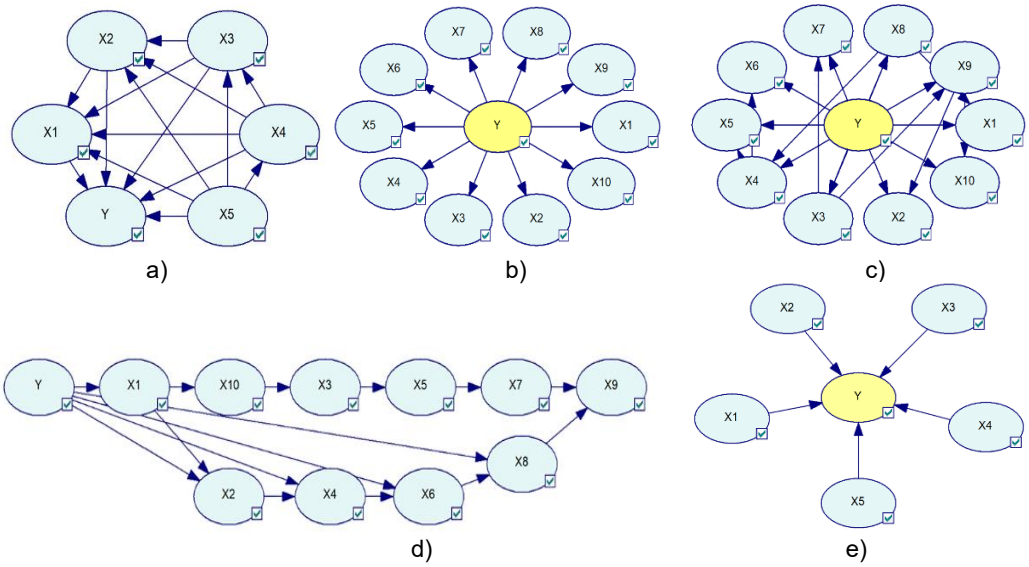


Fig. 2. Bayesian network models built using the algorithm: a) PC, b) ANB, c) TAN, d) GTT, e) NB.

RESULTS AND DISCUSSION

We will consider the implementation of the concept of choosing a structural learning algorithm that best fits the available data using the research example in [26]. This research concerns the modelling of investments in the transport sector and is best suited for experiments on selecting the best structure learning algorithm. Conceptual model of a Bayesian network for calculating the dependence of GDP growth of Ukraine on the volume of investments in the transport sector is shown in **Fig. 3**.

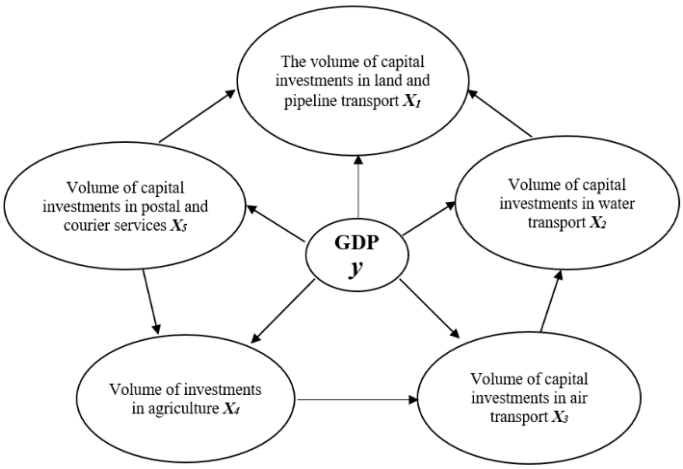


Fig. 3. Conceptual model of a Bayesian network for calculating the dependence of GDP growth of Ukraine on the volume of investments in the transport sector [26].

At the first stage of BN design in the GeNie2.3, we take the initial type of nodes to be General, each node has 5 states from s_1 to s_5. The following macroeconomic indicators for a period of 5 years were taken as experimental data to calculate the dependence of Ukraine's GDP growth on the volume of investments in the transport sector:

- x_1 – the volume of investments in land and pipeline transport in actual prices;
- x_2 – the volume of investments in water transport;
- x_3 – the volume of investments in air transport;
- x_4 – the volume of investments in warehousing and auxiliary activities in the transport sector;
- x_5 – the volume of investments in postal and courier activities.

The set of available data can be divided into two sets: 16 measurements is learning sample A, 8 measurements is test sample B [26].

At the first stage, we designed a structural model using the algorithms presented in the GeNie2.3. **Fig. 4** shows the first 3 algorithms that were applied to the available data: these are PC, TAN i ANB.

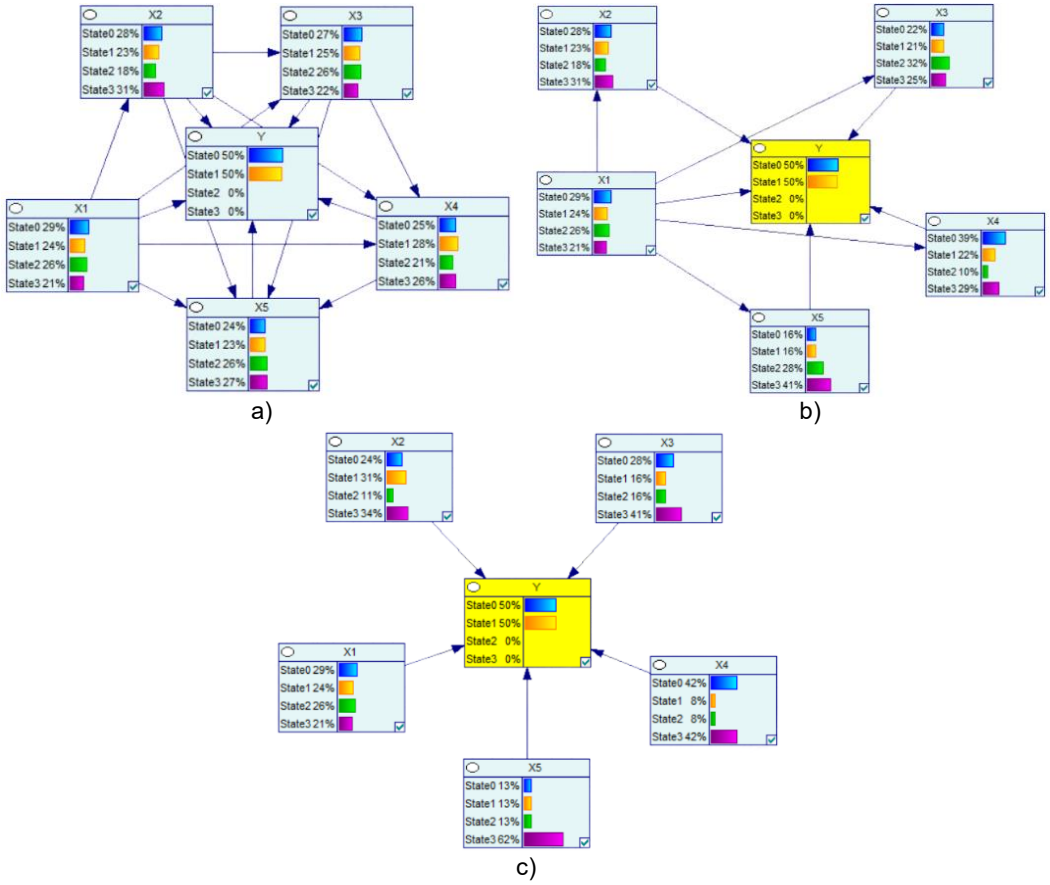


Fig. 4. Bayesian network models built using the algorithm: a) PC, b) TAN, c) ANB.

Similarly, structural models were built using the Bayesian search and Greedy algorithms. Then, the primary parameterization and validation of each network were sequentially carried out. As a result of the experiment on the structural learning method selection, five Bayesian networks were obtained, each of which consists of 6 nodes. After parametric learning, the primary validation of each network was carried out. Among the designed networks, using different structural learning algorithms, we select the network that meets the requirements for accuracy and adequacy of the model. The Greedy algorithm turned out to be an adequate method when working with the existing data set. The selection results are presented in the **Fig. 5**.

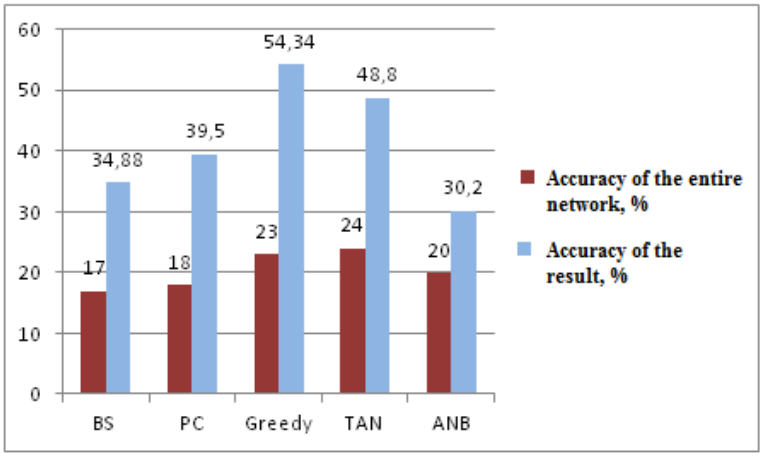


Fig. 5. Selection of the structural learning method [26].

Fig. 5 shows the results of the validation of five Bayesian network models built using five structured learning algorithms. We see that the three algorithms TAN, Greedy, and PC have the highest accuracy, so we will continue to work with them, and we discard the others. At the next stage, the type of all nodes was changed to Noisy-MAX with four states $s_1 \dots s_4$. This may slightly increase the computational complexity, but such a replacement helps to improve accuracy. The fact is that when the model is small and has a small number of nodes and arcs, the increase in computational complexity will be minimal, so in the case of small models, replacing the GENERAL node type with the Noisy-MAX node type is quite justified.

We have three network models, the most adequate, built using the TAN, Greedy, and PC structural learning algorithms; the data file remains unchanged. We conduct re-learning of parameters, re-validation, and sensitivity analysis. The accuracy comparison is shown in the **Fig. 6**.

The accuracy of the result in this case corresponds to the accuracy of forecasting the category of growth of the country's gross domestic product, depending on the volume of investments in the transport sector.

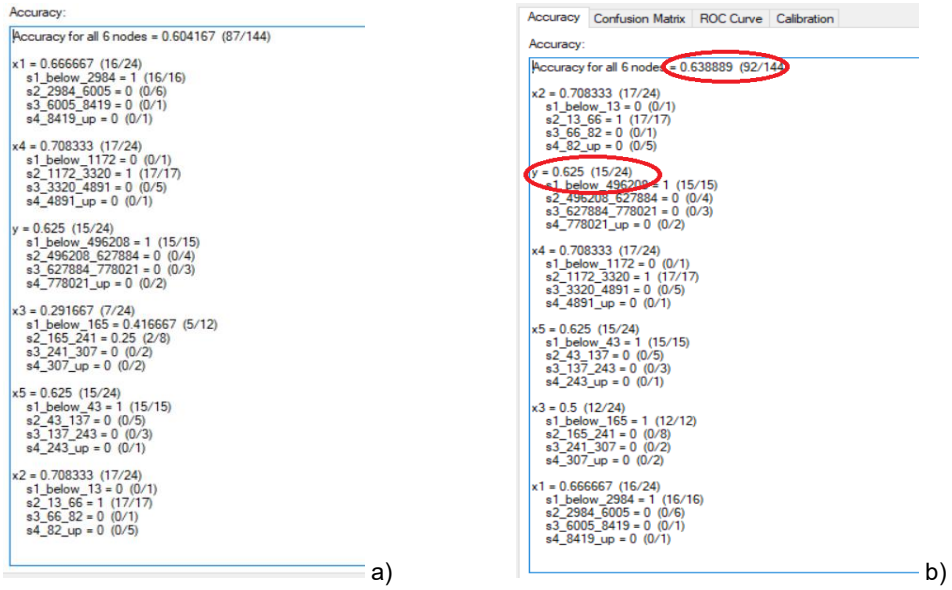


Fig. 6. Overall accuracy of the network and accuracy of the result after: a) primary and b) revalidation.

CONCLUSION

For the investment modelling problem with a critically small sample size (N = 16), the Greedy algorithm, as a score-based heuristic, turned out to be the most adequate, which indicates its advantage in combating overtraining compared to algorithms that rely on statistical power, such as PC. Further introduction of strong prior assumptions at the parametric level using Noisy-MAX nodes led to an increase in the overall accuracy of the network from 60.42% to 63.89%, confirming that the parametric knowledge inclusion is an important criterion for an adaptive approach in the data absence.

The proposed concept of adaptive selection of Bayesian network structural learning algorithms is a timely contribution to the field of stochastic dependence modeling. It successfully translates the process of selecting the optimal algorithm from routine, heuristic, brute force to systematic, multivariate analysis. Its full implementation has the potential to significantly improve the reliability, accuracy, and computational efficiency of BN construction in complex analytical domains.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The author received no financial support for the research, authorship, and/or publication of this article.

COMPLIANCE WITH ETHICAL STANDARDS

The author declares that she has no competing interests.

AUTHOR CONTRIBUTIONS

The author has read and agreed to the published version of the manuscript.

REFERENCES

[1] Kitson, N.K., Constantinou, A.C., Guo, Z. et al. A survey of Bayesian Network structure learning. Artif Intell Rev 56, 8721–8814 (2023). https://doi.org/10.1007/s10462-022-10351-w.

[2] Wolbrecht E., D'Ambrosio B., Paasch R., Kirby D. Monitoring and diagnosis of a multistage manufacturing process using Bayesian networks. Ai Edam. 1998;14:53–67. https://doi.org/10.1017/S0890060400141058.

[3] Nikovski D. Constructing Bayesian networks for medical diagnosis from vague and partially correct statistics. IEEE Trans. Knowl. Data Eng. 2000;12:509–516. https://doi.org/10.1109/69.868904.

[4] Hosseini S., Barker K. A Bayesian network model for resilience-based supplier selection. Int. J. Prod. Econ. 2016; 180: 68–87. https://doi.org/10.1016/j.ijpe.2016.07.007.

[5] Hosseini S., Barker K. Modeling infrastructure resilience using Bayesian networks: A case study of inland waterway ports. Comput. Ind. Eng. 2016; 93: 252–266. https://doi.org/10.1016/j.cie.2016.01.007.

[6] Hosseini S., Barker K. A Bayesian network model for resilience-based supplier selection. Int. J. Prod. Econ. 2016; 180: 68–87. https://doi.org/10.1016/j.ijpe.2016.07.007.

[7] Dünder E., Cengiz M.A., Koç H. Investigation of the impacts of constraint-based algorithms to the quality of Bayesian network structure in hybrid algorithms for medical studies. J. Adv. Sci. Res. 2014; 5: 8–12. https://sciensage.info/index.php/JASR/article/view/175.

[8] Ziegler V. Approximation algorithms for restricted Bayesian network structures. Inf. Process. Lett. 2008;108:60–63. https://doi.org/10.1016/j.ipl.2008.03.015.

[9] Li S., Zhang J., Sun B., Lei J. An incremental structure learning approach for Bayesian network; Proceedings of the 26th Chinese Control and Decision Conference; Changsha, China. 31 May–2 June 2014; pp. 4817–4822. https://doi.org/10.1109/CCDC.2014.6853036
[10] Chickering D.M., Dan G., Heckerman D. Learning Bayesian networks is NP-hard. Networks. 1994; 112: 121–130.
[11] Xu J.-G., Zhao Y., Chen J., Han C. A structure learning algorithm for Bayesian network using prior knowledge. J. Comput. Sci. Technol. 2015;30:713–724. https://doi.org/10.1007/s11390-015-1556-8.
[12] Amirkhani H., Rahmati M., Lucas P., Hommersom A. Exploiting experts' knowledge for structure learning of Bayesian networks. IEEE Trans. Pattern Anal. Mach. Intell. 2016; 39: 2154–2170. https://doi.org/10.1109/TPAMI.2016.2636828.
[13] Zhang Y., Zhang W., Xie Y. Improved heuristic equivalent search algorithm based on maximal information coefficient for bayesian network structure learning. Neurocomputing. 2013; 117: 186–195. https://doi.org/10.1016/j.neucom.2013.02.015.
[14] Voronenko M., Mikhailova E., Savina N., Lurie I., and Lytvynenko V., "A Bayesian Network Approach to Measure the Impact of Factors Characterizing the State of Educational Institutions on the Gross Domestic Product", in Proc. 13th International Conference on Advanced Computer Information Technologies, ACIT 2023, pp. 316–320, 2023. https://doi.org/10.1109/ACIT58437.2023.10275641.
[15] Voronenko M., Zhunissova U., Smailova S., Lytvynenko L., Savina N., Mulesa P., and Lytvynenko V., "Using Bayesian Methods in the Task of Modeling the Patients' Pharmacoresistance Development", IAPGOŚ, 2, pp.77-82, 2022, http://doi.org/10.35784/iapgos.2968.
[16] Romanko O., Voronenko M., Savina N., Zhorova I., Wójcik W. and Lytvynenko V., "The Use of Static Bayesian Networks for Situational Modeling of National Economy Competitiveness", in Proc. IEEE International Conference on Advanced Trends in Information Theory (ATIT), pp. 501-505, Kyiv, Ukraine, 2019, https://doi.org/10.1109/ATIT49449.2019.9030515.
[17] Voronenko M., Lurie I., Boskin O., Zhunissova U., Baranenko R. and Lytvynenko V., "Using Bayesian Methods for Predicting the Development of Children Autism", in Proc. IEEE International Conference on Advanced Trends in Information Theory (ATIT), pp. 525-529, Kyiv, Ukraine, 2019, https://doi.org/10.1109/ATIT49449.2019.9030523.
[18] Li H, Guo H. A Hybrid Structure Learning Algorithm for Bayesian Network Using Experts' Knowledge. Entropy (Basel). 2018 Aug 20; 20 (8): 620. https://doi.org/10.3390/e20080620. PMID: 33265709; PMCID: PMC7513154
[19] Scanagatta, M., Corani, G., de Campos, C. P., & Zaffalon, M. (2018). Approximate structure learning for large Bayesian networks. Machine Learning, 107(8-10), 1209-1227. https://doi.org/10.1007/s10994-018-5701-9.
[20] Stefano Beretta, Mauro Castelli, Ivo Goncalves, Roberto Henriques, Daniele Ramazzotti, "Learning the structure of Bayesian Networks: A quantitative assessment of the effect of different algorithmic schemes", 2018, https://doi.org/10.48550/arXiv.1704.08676.
[21] Cheng J., Bell D. A. Liu W. Learning belief networks from data: an information theory based approach // Proceedings of the sixth international conference on information and knowledge management (CIKM 1997), Las Vegas (Nevada), November 10-14. 1997. P. 325-331. https://doi.org/10.1145/266714.266920.
[22] Cheng J., Greiner R., Kelly J., Bell D. A., Liu W. Learning Bayesian networks from data: an information-theory based approach. The artificial intelligence journal (AIJ). 2002. Vol. 137. P. 43-90.
[23] Dempster A. P., Laird N. M., Rubin D. B. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society: series B (methodological). 1977. Vol. 39, № 1. P. 1-22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

[24] Freedman D., Kisilev P. Fast mean shift by compact density representation // Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2009. P. 1818-1825. https://doi.org/10.1109/CVPR.2009.5206716.
[25] Friedman N., Murphy K., Russell S. Learning The Structure Of Dynamic Probabilistic Networks // Proc. Fourteenth Conference On Uncertainty In Artificial Intelligence. 1998. https://doi.org/10.48550/arXiv.1301.7374.
[26] Lytvynenko V., Savina N., Krejci J., Voronenko M., Yakobchuk M., and Kryvoruchko O. Bayesian Networks' Development Based on Noisy-MAX Nodes for Modeling Investment Processes in Transport. In Proc. 2-nd International Workshop «Modern Machine Learning Technologies and Data Science» (MoMLeT&DS-2019), pp. 1-10, Shatsk, Ukraine, 2019, http://ceur-ws.org/Vol-2386/.

КОНЦЕПЦІЯ АДАПТИВНОГО ВИБОРУ АЛГОРИТМІВ СТРУКТУРНОГО НАВЧАННЯ БАЙЄСІВСЬКИХ МЕРЕЖ НА ОСНОВІ ЇХНІХ ХАРАКТЕРИСТИК

Марія Вороненко
Національний університет «Одеська політехніка»,
проспект Шевченка, 1, м. Одеса, 65044, Україна

АНОТАЦІЯ

Вступ. Сучасні інтелектуальні системи потребують ефективних механізмів аналізу, прогнозування та прийняття рішень. Байєсівські мережі дозволяють ефективно представляти причинно-наслідкові зв'язки між змінними.

Матеріали та методи. Структурне навчання байєсівської мережі включає визначення структури орієнтованого ациклічного графа, в якому змінні пов'язані між собою. Якість структури має вирішальний вплив на здатність моделі точно представляти умовні ймовірності і на ефективність алгоритмів навчання та надійність моделі. Основною проблемою є обчислювальна складність моделі. Це означає, що для багатовимірних задач неможливо виконати повний пошук усіх можливих структур, і знайти глобальний оптимум. Це змушує покладатися на евристичні методи пошуку та алгоритми апроксимації і створює постійну потребу в балансуванні якості визначення структури та обчислювальних ресурсів.

Результати. Концепція адаптивного вибору алгоритмів для структурного навчання базується на систематичному аналізі характеристик алгоритму та на властивостях даних, що дозволяє вибрати найбільш підходящий алгоритм для конкретного випадку, оптимізувати компроміс між якістю моделі та обчислювальними ресурсами, підвищити узагальнюваність моделі в практичних сценаріях.

Висновки. Запропонована концепція адаптивного вибору алгоритмів структурного навчання переводить процес вибору оптимального алгоритму з рутинного, евристичного методу грубої сили на систематичний, багатовимірний аналіз. Її повна реалізація має потенціал для значного підвищення надійності, точності та обчислювальної ефективності побудови Байєсівських моделей у складних аналітичних областях.

Ключові слова: концепція адаптивного вибору, Байєсівська мережа, алгоритм структурного навчання, точність моделі.

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
06 November, 2025	05 December, 2025	08 December, 2025	25 December, 2025