Comparative Study of Noisy-MAX Nodes and General Nodes in Bayesian Network Models

Volodymyr Lytvynenko[1][0000-0002-1536-5542], Dmytro Nikytenko [2][0000-0003-4989-0879], Mariia Voronenko[1][0000-0002-5392-5125], Jan Krejci[3] [0000-0003-4365-5413], Nataliia Savina[2][0000-0001-8339-1219]

[1]Kherson National Technical University, Kherson, Ukraine,
[2] National University of Water and Environmental Engineering, Rivne, Ukraine,
[3] Jan Evangelista Purkyne University in Usti nad Labem, Czech Republic

immun56@gmail.com, mary_voronenko@i.ua, n.b.savina@nuwm.edu.ua, d.v.nikytenko@nuwm.edu.ua, jan.krejci@ujep.cz

Abstract. This article is devoted to the use of Bayesian networks for analyzing the growth of gross domestic product (GDP) of Ukraine and offers a comparative description of the use of various structural learning algorithms. A comparative study of the behavior of the Noisy-MAX nodes and the General nodes in the design of the Bayesian network was carried out. It has been shown that Noisy-max nodes in comparison with General nodes provide a relatively high initial accuracy. General nodes require retesting. However, Noisy-MAX nodes entail an increase in time and computational cost.

Keywords: Gross Domestic Product (GDP); General nodes; Noisy-MAX nodes; Bayesian networks; Structural learning; Sensitivity analysis; Validation.

1 Introduction

Theories of economic growth evolve over time, dependent on the stage of economics, and the improvement of mathematical and statistical tools have had a significant impact on the formulation of New Concepts.

Joseph E. Stiglitz and Andrew Weiss [1] argue that the existence of financial resources, not its value, is decisive in determining private investment, and therefore the economic growth of the country. Greenwald A.G. at al. [2] claimed that the exchange rate can also be impacted by economic development through the activation of investments. As suggested by Kenneth A. Froot and Jeremy C. Stein [3], devaluation encourages foreign investment, facilitating the acquisition of local assets by foreign companies at a much lower price.

In determining the impact of external investment, indicators such as foreign direct investment in Ukraine and the average annual dollar rate are used. In determining the impact of domestic investment potential, we take into account the average propensity to save and interest rates on attracted term deposits. The level of financial development depends on loans in national and foreign currencies. The level of manufacturability (or innovation) affects the profitability of the operating activities of industrial enterprises, the proportion of enterprises engaged in innovation, as well as the share of revenue of enterprises.

The aim of the work is to compare the use of General and Noisy-MAX nodes in designing a model of a static Bayesian network for assessing Ukraine's economic growth of economic indicators.

2 Problem Statement

For a set of events $X^{(i)}, i = 1, ..., N$ that are related, and a set of learning data $D = (d_1, ..., d_n), d_i = \{x_i^{(1)} x_i^{(2)} ... x_I^{(N)}\}$, is given. Here the subscript is the observation number, and the upper one is the variable number, n –is the number of observations, each observation consists of $N (N \geq 2)$ variables, and each j-th variable $(j = 1, ..., N)$ has $A^{(j)} = \{0, 1, ..., a^{(j)} - 1\}$ $(a^{(j)} \geq 2)$ conditions. Based on a given training sample, we need to build an acyclic graph connecting the event sets $X_i, i = 1, ..., N$. Having a set of input indicators that interact with each other as shown in Fig. 1, it is necessary to carry out a study of the construction of Bayesian networks using the nodes General and Noisy-MAX in order to assess the possibility of economic growth of the country.

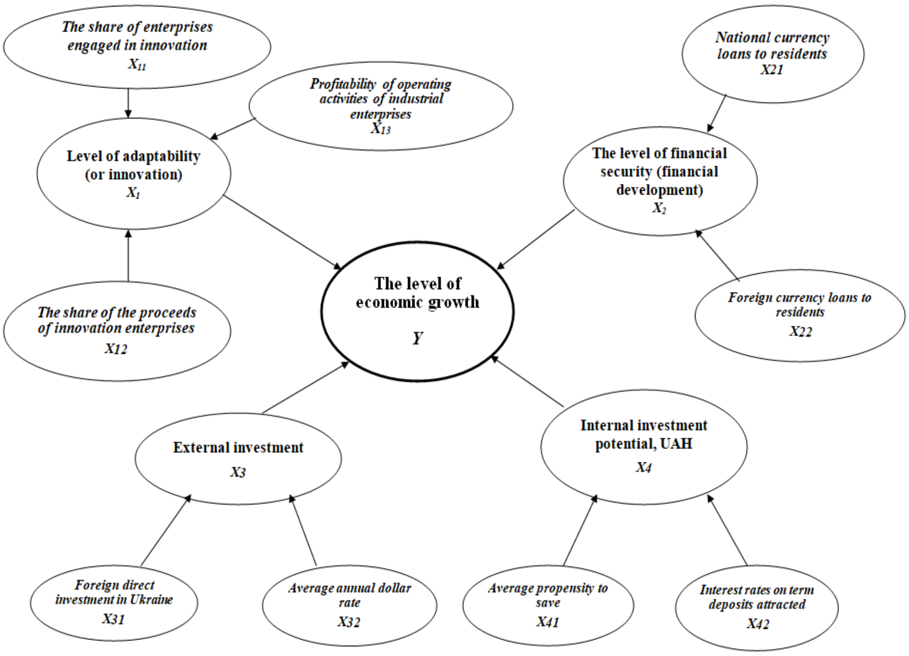

Fig. 1. Conceptual model of a static BN for assessing a country's economic growth.

We'll get BN structure $g \in G$, which is represented by a set N of predecessors $P^{(1)},...,P^{(N)}$, that is, for each vertex $j = 1,...,N, P^{(j)}$ it is a variety of parent vertices, such that $P^{(j)} \subseteq \{X^{(1)},...,X^{(N)}\} | \{X^{(j)}\}$. We have events $X^{(i)}, i = 1,...,N$ that are affected by the uncertainties of a different nature. And also we have data describing these events.

3 Review of the Literature

When we have an increase in the amount of parents, we also will have the exponential growth of parameters. This is one of the main difficulties of Bayesian network models. By using Noisy-MAX nodes [4, 5], since they use multi-valued variables, we can solve the problem of increasing the dimension and, as a consequence, the problem of increasing computational complexity.

This approach has proven itself in many real applications [6-8]. The main advantage of it is the small amount of parameters.

This reduces the spatial and temporal complexity of the algorithms [9] for Bayesian network models and improves the quality of distributions extracted from the data [10, 11]. The main and most important advantage of using Bayesian networks is their resistance to incomplete, inaccurate and noisy information. In these cases, the result will reflect the most likely outcome of events [12].

The use of Bayesian networks for the synthesis, prediction, and analysis of uncertainty is considered in [13], where they are one of the mathematical tools for eutrophication models, risk counting and cost-based performance analysis.

In [14] BNs are presented as a decision support tool for politicians that can be used to formulate strategic recommendations to improve the innovative level of a country's economy.

4 Materials and Methods

4.1 Data

As experimental data for assessing the economic growth of Ukraine, macroeconomic indicators were taken into account that the statistical capabilities and the modification of existing methods for studying the financial activity of an enterprise (Table 1).

In our study, we use a set of statistical data that are interconnected (Fig. 1) consisting of 14 indicators for the period 2005-2018.

The matrix of indicators is divided into four blocks that most fully characterize the financial economic and business activity of enterprises, as well as the course of economic processes in the country (Table 1). The resulting indicator Y is an integral indicator of the level of economic growth of Ukraine.

Table 1. Matrix of economic indicators.

Indicators	Appointment
Y	**The level of economic growth (**nominal GDP at actual prices)
X1	**Level of adaptability (or innovation)**
X11	The share of enterprises engaged in innovation
X12	The share of the proceeds of innovation enterprises
X13	Profitability of operating activities of industrial enterprises, %
X2	**The level of financial security (financial development), UAH**
X21	National currency loans for a term of 5 years to residents (excluding deposit-taking corporations), average value, UAH million
X22	Foreign currency loans to residents (excluding deposit-taking corporations) for a term of 5 years, average value, UAH million
X3	**External investment, UAH**
X31	Foreign direct investment in Ukraine
X32	Interest rates on term deposits attracted, %
X4	**Internal investment potential, UAH**
X41	Average propensity to save
X42	Average annual dollar rate, UAH

4.2 Materials and Methods

A Bayesian network (BN) is a pair <G, B>, in which the first component G is a directed acyclic graph corresponding to random variables [14,15]. Each variable is independent of its parents in G. So, the graph is written as a set of independence conditions. The set of parameters defining the network is the second component B. It contains parameters $Q_{x^i|pa(X^i)} = P(x^i \mid pa(X^i))$ for each possible x_i value from X_i and $pa(X^i)$ from $Pa(X^i)$, where $Pa(X^i)$ denotes the set of parents of the variable X_i in G . Each variable X_i is represented as a vertex. We use the notation to identify the parents $Pa^G(X^i)$ if we consider more than one graph.The total joint probability of BN is calculated by the formula $P_B(X^1,...,X^N) = \prod_{i=1}^{N} P_B(X^i \mid Pa(X^i))$.

BN is a probabilistic model for representing probabilistic dependencies, as well as the absence of these dependencies. At the same time, the A→B relationship is causal, when event A causes B to occur, that is, when there is a mechanism whereby the value accepted by A affects the value adopted by B.

Validation was proposed for the first time in 1977 in [16]. Validation of the network that we design was carried out according to the algorithm for maximizing expectations. The algorithm finds local optimal estimates of the maximum likelihood of arguments. The concept of the algorithm is that if we knew the values of all nodes, then training would be simple at some step M. Therefore, at stage E, estimations of the expected likelihood value are made, including latent variables, as if we were able to observe them. In step M, the maximum likelihood values of the parameters are estimated using the maximization of the expected likelihood values obtained in step

E. Then, the algorithm performs step E using the parameters obtained in step M again and so on.

A whole series of such algorithms was developed, based on the algorithm of maximizing the expectation [17,18].

GeNIe 2.4 Academic Bayesian Network Design Software implements three dicretization methods [19,20]:

Uniform Widths - a method with uniform width (discretization on the same width of classes), which makes the width of the sampling intervals the same,

Uniform Counts - the method of unit graphs (discretization on the same number of points inside the classes), which determine the number of values in each of the sampling registers,

Hierarchical - hierarchical method (hierarchical discretization), which is an uncontrolled method of discretization associated with clustering.

We will successively apply each discretization method to the experimental data set and carry out their comparative study on two types of General and Noisy-MAX nodes.

4.3 Noisy-MAX Nodes

The Noisy-MAX node consists of a child node, Y, taking on n_Y possible values that can be labeled from 0 to n_Y-1, and N parents, $Pa(Y) = \{X_1,\dots, X_N\}$, which represent the causes of Y. Each X_i has a certain zero value, so that $X_i = 0$ represents the absence of X_i. Two basic axioms define the Noisy-MAX [11]:

1. When all the causes are absent, the effect is absent:

$$P\left(Y = 0 \mid X_i = 0_{[\forall_i]}\right) = 1 \tag{1}$$

2. The degree reached by Y, is the maximum of the degrees produced by the X, if they were acting independently:

$$P\left(Y \le y \mid x\right) = \prod_i P\left(Y \le y \mid \boldsymbol{X}_i = x_i, X_j = 0_{\left[\forall_{j,j\ne i}\right]}\right) \tag{2}$$

where x represents a certain configuration of the parents of Y, x = $(x_1,\dots, x_N)$.

5 Experiments and Results

When developing the BN, the GeNIe 2.4 Academic software environment was used. As can be seen from Fig. 1, the network contains 4 key nodes:

- X1 - the level of manufacturability (innovation),
- X2 - the level of financial security (financial development), UAH
- X3 - external investment, UAH,
- X4 - domestic investment potential, UAH.

It should be noted that due to the specifics of the Bayesian networks, all the conclusions of this model regarding the information sought are probabilistic in nature and are presented in the form of a ranked list (according to the values of the probability of fidelity of a particular conclusion).

Data taken from 2005 to 2018. The dynamics of changes in the initial indicators for the observed period are presented in table 2. All nodes have five states: s1, s2, s3, s4, s5. For example, for the node X1 (as shown in Fig. 2), the intervals of state discretization will be as follows:

- s1_below_85879;
- s2_85879_114478;
- s3_114478_150998;
- s4_150998_218982;
- s5_218982_up

The resulting static BN of the country's economic growth is presented in Fig. We perform parameterization and validation on the General nodes. The initial accuracy of the result was 42%, while the overall accuracy of the network was 48.8%. After conducting a sensitivity analysis, the overall accuracy of the network remained almost unchanged at 47.56%. However, the accuracy of the result increased from 42% to 64%. At the next stage of the study, we changed the type of all nodes to Noisy – MAX with five states s1-s5, and the resulting node Y. The network remains the same, the data file also does not change. We carry out parametric learning, primary validation.

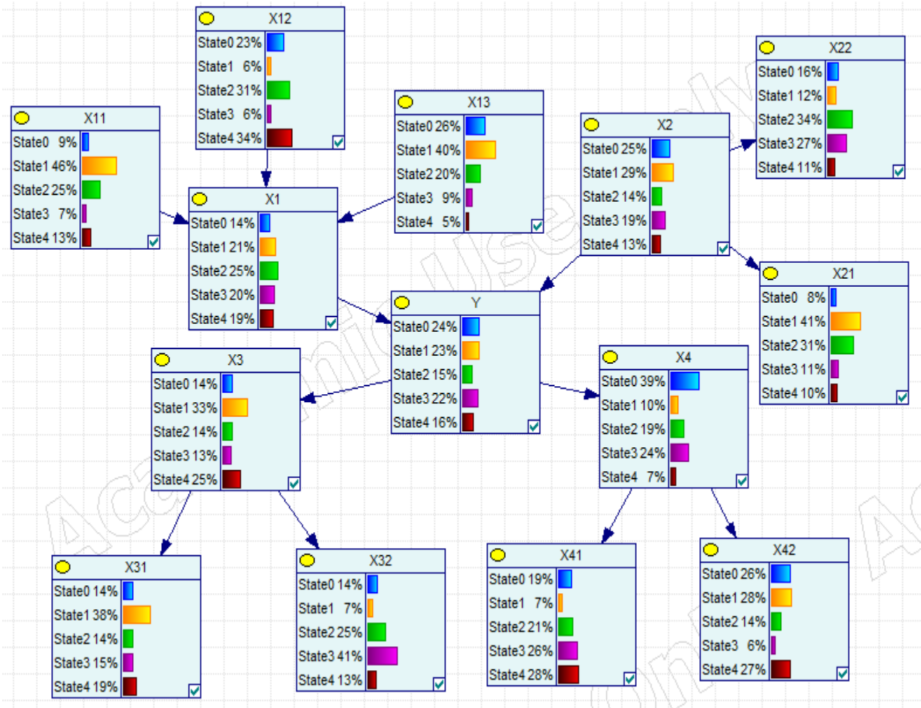

Fig. 2. Static BN of the country's economic growth.

The initial overall accuracy of the network when using Noisy – MAX nodes was immediately quite high and amounted to 55.4%, and the accuracy of the opposite result was very low and amounted to 46%.

After a sensitivity analysis, the overall accuracy of the network remained almost unchanged (increased by 2% - from 47% to 57.11%), but the accuracy of the result increased by 14% to 60% compared to the initial 46%. A comparison of the results is shown in table 2:

Table 2. Comparison of results after primary and re-validation.

	Initial accuracy		**Accuracy after a sensitivity analysis**	
	Overall network accuracy,%	Accuracy of the result ,%	Overall network accuracy,%	Accuracy of the result ,%
General nodes	48,8	42,0	47,6	**64,0**
Noisy-MAX nodes	55,4	46,0	57,1	60,0

Further, at the second stage of the study, we will apply each discretization method in the experimental data set and compare the results. In the beginning, we apply discretization using the Uniform Weights method. We need to discretize the existing data set and also generate a 100-line file for GeNie. We repeat all the steps first for the General nodes and then for the Noisy nodes: structural learning, parametric learning, validation, sensitivity analysis, and re-validation.

Next, we apply discretization using the Uniform Counts method. We need to rediscretize the existing data set and also generate a 100-line file for GeNie. We repeat all the steps first for the General nodes and then for the Noisy nodes: structural learning, parametric learning, validation, sensitivity analysis, and re-validation.

Finally, we discretize the available data using the Hirerical sampling method. A comparison of the accuracy of the three methods is given in table 3.

Table 2. Comparison of accuracy after changing the discretization method.

	Method *Uniform Counts*		**Method *Uniform Weigths***		**Method *Hirerical***	
	Overall network accuracy,%	Accuracy of the result , %	Overall network accuracy,%	Accuracy of the result , %	Overall network accuracy,%	Accuracy of the result , %
General nodes	40,0	**64,3**	44,3	42,9	47,6	**64,0**
Noisy-MAX nodes	31,4	28,6	45,7	50,0	57,1	**60,0**

6 Discussion

Based on the obtained experimental results, it is clear that the use of General nodes requires the use of sensitivity analysis procedures, repeated parameterization, and repeated validation since it significantly increases the resulting accuracy (in our case, by 10% - from 47% to 57%).

With Noisy-MAX nodes, the required resulting accuracy is achieved immediately after the initial validation, with a very small difference of 4%. This suggests that for a network with this type of nodes there is no need for sensitivity analysis and re-validation (Fig. 3).

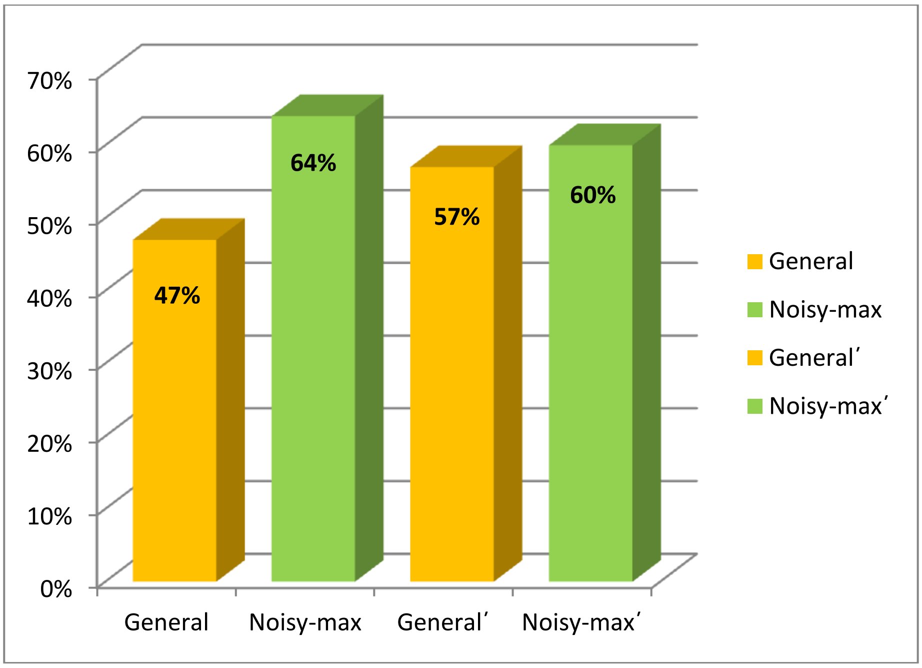

Fig. 3. The obtained experimental results.

After the successive application of the three discretization methods, the following conclusions can be drawn:

1. If you compare by the accuracy of the result, then the General nodes are better than Noisy-MAX in Uniform Counts by 35.7% (General = 64.3% and Noisy-MAX = 28.6%) and Hirerical (General = 64% and Noisy-MAX = 60%), but worse than Uniform Weights by 7% (General = 42.9% and Noisy-MAX = 50%). It is shown on fig. 4:

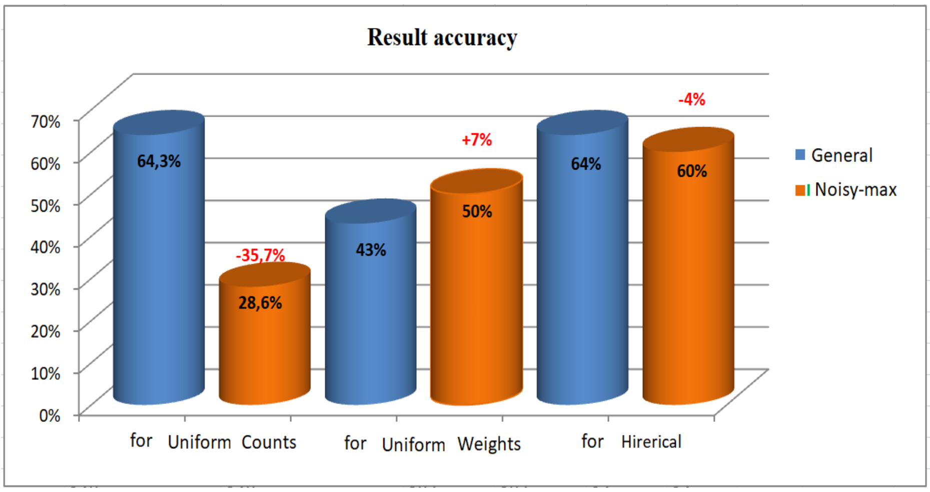

Fig. 4. The result accuracy of the three discretization methods.

2. If we compare in terms of the overall accuracy of the network, then Noisy-MAX nodes are vice versa better than General nodes in Uniform Weights by 2% (General = 44.3% and Noisy-MAX = 45.7%) and Hirerical by 9% (General = 47.6% and Noisy-MAX = 57.1%), but worse in Uniform Counts by 9% (General = 40% and Noisy-MAX = 31.4%). This is shown in Figure 5.

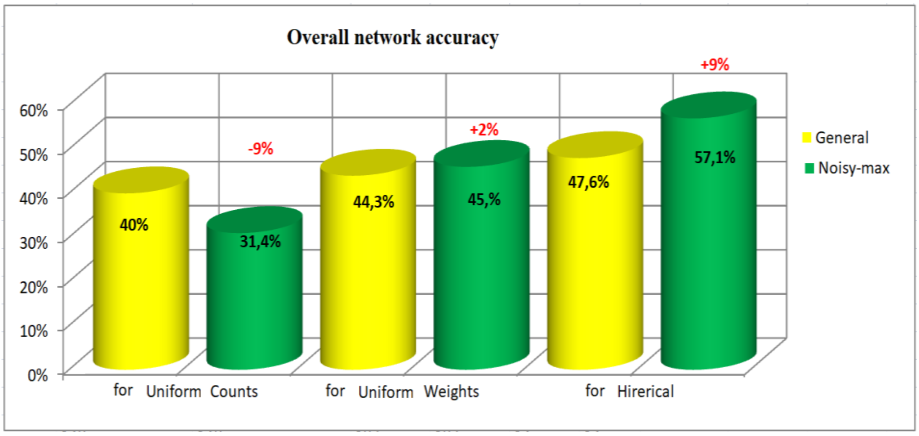

Fig. 5. The network accuracy of the three discretization methods

7 Conclusion

When comparing the use of General and Noisy-MAX nodes in designing a model of a static Bayesian network for assessing Ukraine's economic growth, we can conclude the following. When using the Hirerical discretization method, high rates of overall network accuracy and result accuracy are observed both with General nodes (64%) and Noisy-MAX nodes (60%), therefore both types of nodes work equally well with this method (table 3). As can be seen from the table 3, the Uniform Weigths dicretization method is poorly applicable for General nodes (accuracy below 50%), the Uniform Count sampling method is not applicable for Noisy-MAX nodes (accuracy below 40%).

Noisy-MAX nodes, compared to General nodes, provide relatively high initial accuracy. General nodes require sensitivity analysis, re-parameterization, and re-validation procedures.

In our future research, we will apply the proposed model to the design of a dynamic BN to assess general trends in increasing levels of economic growth at different time periods.

References

1. Stiglitz, J. E., Weiss, A.: The American Economic Review, Vol. 71/3, 393-410. (1981)
2. Greenwald, A.G., Schute, P.C., Shiveley, J.L.: Brachial plexus birth palsy: A 10 -year report on the incidence and prognosis, Pediatr Orthop, Vol.4, 689-692.(1984)
3. Froot, K. A., Stein, J. C.: Exchange Rates and Foreign Direct Investment: An Imperfect Capital Markets Approach, The Quarterly Journal of Economics, Vol. 106/4, 1191-1217. (1991)
4. Henrion, M., Kanal, L., Levitt, T., Lemmer, J.: Some practical issues in constructing belief networks. In: Uncertainty in Artificial Intelligence 3, New York: Elsevier, 161-173. (1989)
5. Díez, F.: Parameter adjustment in Bayes networks. The generalized noisy or-gate. In: Proc. 9th Conf. Uncertainty Artif. Intell., Washington, DC, 99-105. (1993)
6. Díez, F., Mira, J., Iturralde, E., Zubillaga, S.: DIAVAL, a Bayesian expert system for echocardiography. In: Artif. Intell. Med., Vol. 10/1, 59-73. (1997)
7. Pradhan, M., Provan, G., Middleton, B., Henrion, M.: Knowledge engineering for large belief networks. In: Proc. 10th Annu. Conf. UAI, San Francisco, CA, 484-490. (1994)
8. Shwe, M., Middleton, B., Heckerman, D., Henrion, M., Horvitz, E., Lehmann, H., Cooper, G.: Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base: I. The probabilistic model and inference algorithms. In: Methods Inf. Med.,Vol. 30/ 4, 241-255. (1991)
9. Onisko, A., Druzdzel, M., Wasyluk, H.: Learning Bayesian network parameters from small data sets: Application of noisy-OR gates. In: Int. J. Approx. Reason., Vol. 27/2, 165–182. (2001)
10. Zhang, N., Poole, N.:Exploiting causal independence in Bayesian network inference. In: J. Artif. Intell. Res., Vol. 5/1, 301-328. (1996)
11. Díez, F., Galán, S.: Efficient computation for the Noisy-MAX, Int. J. Int. Syst., Vol. 18/2, 165–177. (2004)
12. Kjærulff, U., van der Gaag, J.: Making sensitivity analysis computationally efficient. In: Proceedings of the Sixteenth Conference Uncertainty in Artificial Intelligence (UAI). San Francisco, CA: Morgan Kaufmann Publishers, 317–325. (2000)
13. Borsuk, M.E., Stow, C.A., Reckhow, K.H. : A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecological Modelling, Vol.2/173, 219–239. (2004)

14. Cheeseman, P.S., Kelly, M., Taylor, W., Freeman, D., Stutz, J.: Bayesian classification, In: Proceedings of AAAI, St. Paul, MN, 607-611. (1988)
15. McCann, R.K., Marcot, B.G, Ellis, R.: Bayesian belief networks: applications in ecology andnatural resource management. Canadian Journal of Forest Research, Vol. 36/12, 3053-3062. (2006)
16. Friedman, N.: The Bayesian structural EM algorithm / Fourteenth conference on Uncertainty in Artificial Intelligence (UAI'98), Madison, Wisconsin, USA, 24–26 July, 1998. – SF.: Morgan Kaufmann, 129-138. (1998)
17. van der Gaag, L. C., Coup´ e, V. M.: Sensitivity analysis for threshold decision making with Bayesian belief networks. In Lamma, E., and Mello, P., eds., AI*IA 99: Advances in Artificial Intelligence, Lecture Notes in Artificial Intelligence. Berlin: Springer-Verlag. 37 – 48. (2000)
18. Zhang, Z., Kwok, J., Yeung D.: Surrogate maximization (minimization) algorithms for AdaBoost and the logistic regression model. In: Proceedings of the twenty-first international conference on machine learning (ICML), 117 p. (2004)
19. Beinlich, N. , Suermondt, N. , Chavez, N. , Cooper, G: The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks, In: Proceedings of the Second European Conference on Artificial Intelligence in Medicine, London, England, 247-256. (1989)
20. Darwiche, N.: A differential approach to inference in Bayesian networks. In Uncertainty in Artificial Intelligence: Proceedings of the Sixteenth Conference (UAI 2000), San Francisco, CA: Morgan Kaufmann Publishers, 123–132. (2000)