Assessing the Possibility of a Country's Economic Growth Using Static Bayesian Network Models

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

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

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

Abstract: This article is devoted to the use of Bayesian networks to analyze the possibility of economic growth in Ukraine. It was found that at the maximum level of external investment, direct investment in Ukraine will increase and this creates the conditions for increasing the country's economic growth. It has also been shown that Noisy-max nodes, compared to 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: Economic growth; Innovative development; General nodes; Noisy-max nodes; Bayesian networks; Structural learning; Sensitivity analysis; Validation

1 Introduction

Economic growth can be considered a major factor in the well-being and prosperity of the country. Industrialization, technology development, and innovation activity are widening the gap between developed countries and developing countries. The innovative development of enterprises is one of the basic needs of the national economy.

The activities of the enterprise reveal innovations by transforming and reforming production through the use of inventions or various opportunities for the release of new goods, the opening of new sources of raw materials, markets, modernization of production, etc., ie the implementation of new combinations of factors of production. Innovative activity is a factor that gives a dynamic character to the economy and has a two-sided influence: on the one hand, it opens new opportunities for economic expansion, on the other hand, it requires a change of traditional directions for further development [1].

Today, there are 42,564 industrial enterprises in Ukraine, representing 12.4% of their total. In the countries of the European Union (EU), the share of enterprises engaged in innovation activity is about 53%. The largest number of innovative enterprises among EU countries is in Germany (79.3% of the total number of enterprises), the smallest in Bulgaria (27.1% of the total number of enterprises) [2]. Theories and models of economic growth highlight the ways in which current economic activity can influence future economic events. Therefore, it will be advisable to determine the informative economic indicators that have the greatest impact on the dynamics of the economic growth of Ukraine.

This will create the necessary prerequisites for the growth of production and expanded reproduction of GDP in order to increase the welfare of the country's population. This paper presents the results of studies on the development of probability-determined models based on Bayesian networks to assess the degree of economic development of Ukraine. Analysis of the country's economic growth is associated with the level of external investment resource, the level of internal investment potential, financial development and the level of manufacturability (innovation) of industrial enterprises.

The aim of the work is to develop static Bayesian network models based on noisy-MAX nodes to analyze the country's economic growth trends.

2 Problem Statement

A mathematical model that, when analyzing the financial, investment, and economic indicators of an enterprise would help assess the level of economic growth of a country, requires the availability of input data. Bayesian network methods, with a certain degree of probability, make it possible to achieve the goal.

Having input indicators, such as level of manufacturability (innovativeness), level of financial security (financial development), UAH, an indicator of external investment, UAH, internal investment potential, UAH, which interact with each other as shown in Figure 1, it is necessary to design a static Bayesian network for assessing the country's economic growth opportunities.

Considering that one of the problems in the development of Bayesian networks is the exponential increase in the number of parameters in conditional probability tables (CPT), this study proposes a technique for using noisy-MAX nodes to model economic processes.

The Noisy-MAX node, which in the case of noisy variable reduces to the noisy-OR, 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, $Pa(Y) = \{X_1, \ldots, X_N\}$, which usually 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 [3].

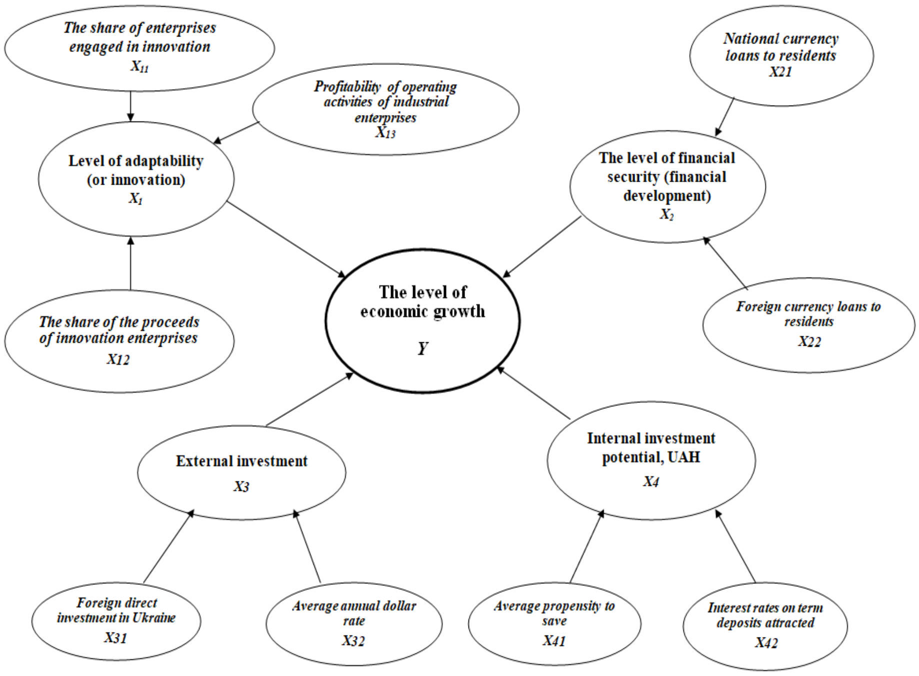


Fig. 1. Conceptual model of a static BN for assessing a country's economic growth

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

$$P = \left(Y = 0 \middle| X_i = 0_{[\forall_i]}\right) = 1, \quad (1)$$

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

$$P\left(Y \le y \middle| x\right) = \prod P\left(Y \le y \middle| X_i = x_i,\, X_j = 0_{[\forall_j,\, j \ne i]}\right), \quad (2)$$

where x represents a certain configuration of the parents of Y, x = (x_1,…, x_N). The parameters for link $X_i \rightarrow Y$ – are the probabilities that the effect assumes a certain value y, when X_i takes on the value x_i, and all the other causes of Y are absent:

$$c_y^{x_i} = P\left(Y = y \middle| X_i = x_i, \cdots X_j = 0_{[\forall_j, j \ne i]}\right), \quad (3)$$

If X_i has n_{X_i} values, the number of parameters required for the link $X_i \rightarrow Y$ is $(n_{X_i}-1)\times(n_Y-1)$ – because of Equation 1. Since all the variables involved in a noisy OR are binary, this model only requires one parameter per link. Alternatively, it is possible to define new parameters:

$$C_y^{x_i} = P\left(Y \le y \middle| X_i = x_i, X_j = 0_{[\forall_j, j \ne i]}\right) = \sum_{y'}^{y} c_y^{x_i} \, , \tag{4}$$

so that Equation 2 can be rewritten as:

$$P\left(Y \le y \middle| x_1, \ldots, x_n\right) = \prod_i C_y^{x_i} \, , \tag{5}$$

The CPT is obtained by taking into account that:

$$P\left(y \middle| x\right) = \begin{cases} P\left(Y \le 0 \middle| x\right) & if \;\; y = 0 \\ P\left(Y \le y \middle| x\right) - P\left(Y \le y - 1x \middle| \right) & if \;\; y > 0 \end{cases} \, , \tag{6}$$

3 Review of the Literature

An analysis of the current state, tendency of innovation activity of industrial enterprises of Ukraine and generalization of theoretical approaches, directions, and measures of increasing the innovative activity of the country is covered in [1].

Authors [4,5] concluded that the actions of large financial institutions have significant implications for the stability of the entire economic system and may be a threat to economic risk.

Network models have become attractive for modeling dependencies in real-world phenomena because of their ease of representation and the ability to provide an intuitive way visualization and interpretation of complex economic relations [6].

To understand vulnerabilities in the financial system, the idea of what-if network analysis has proven to be a promising tool that can help monitor the interconnectedness of financial institutions and markets. This has led to a significant increase in research on the statistical properties of network indicators to analyze systemic risk and the possibility of economic development of the country [7].

Directional acyclic graphical models and graphical Gaussian models were dealt with by the authors [8-10] and achieved results in their research.

However, when analyzing and modeling economic processes, risks, and to predict future results, Bayesian network models have proven themselves best [11, 12]. They are currently widely used for working with discrete data [13].

4 Materials and Methods

4.1 Data

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

Table 1. Matrix of economic indicators

Indicators	Appointment
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 potential, 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
Y	**The level of economic growth (**nominal GDP at actual prices)

In our study, we are dealing with 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.

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.

The goal of parametric learning is to find the most likely θ variables that explain the data. Let $D=\{D_1,D_2,\ldots,D_N\}$ be a composition of the learning data, where $D_l=\{x_1[l],x_2[l],\ldots,x_n[l]\}$ consists of instances of Bayesian network nodes. So the learning parameter is quantified by a log-likelihood function, denoted as $L_D(\theta)$ [17].

5 Experiments

When developing the BN, the GeNIe 2.4 Academic software environment was used. The original Bayesian network model is built on General nodes. The block diagram of the BN of the country's economic growth is presented in Fig. 2.

When developing the BN, the GeNIe 2.4 Academic software environment was used. As can be seen from Fig. 1, the network contains 5 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,
Y – the level of economic growth.

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 were taken from 2005 to 2018. The dynamics of changes in the initial indicators for the observed period are presented in Figure 3. All nodes have five states: s1, s2, s3, s4, s5.

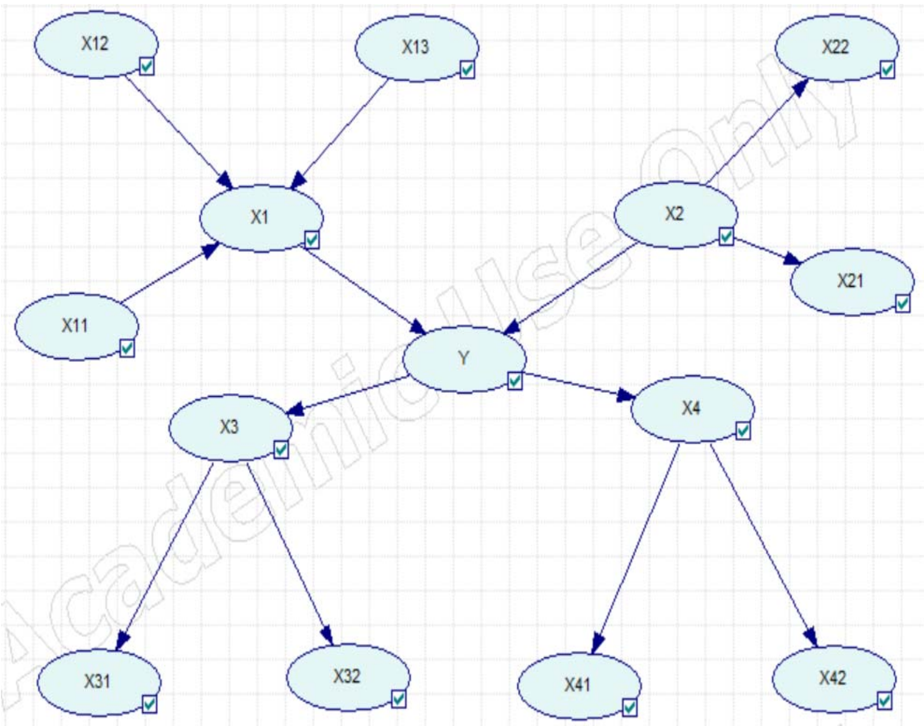


Fig. 2. Structural model of a static BN of the country's economic growth

For example, for the node X1, the intervals of state discretization will be as follows:

s1_below_85879;
s2_85879_114478;
s3_114478_150998;
s4_150998_218982;
s5_218982_up

We carry out, parameterization, and network validation on the nodes of General. The initial overall accuracy of the network was 48.8%, the accuracy of the result was 42%. 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 training in parameters, primary validation. The overall accuracy of the network increased by 6.6% (and amounted to 55.4%), the accuracy of the result increased by 4% (and amounted to 46%).

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Y	441452	544153	720731	948056	913345	1079346	1299991	1404669	1465198	1586915	1988544	2383182	2982920	2502130
X1	**0.253**	**0.355**	**0.44**	**0.39**	**0.26**	**0.308**	**0.352**	**0.307**	**0.295**	**0.21**	**0.154**	**0.188**	**0.218**	**0.399**
X11	0.082	0.1	0.115	0.108	0.107	0.115	0.128	0.136	0.136	0.121	0.152	0.166	0.143	0.166
X12	0.065	0.067	0.067	0.059	0.048	0.038	0.038	0.033	0.033	0.025	0.014	0.005	0.007	0.059
X13	0.016	0.035	0.058	0.049	0.018	0.035	0.047	0.034	0.03	0.016	0.009	0.042	0.066	0.034
X2	**0.335**	**0.365**	**0.18**	**0.28**	**0.414**	**0.338**	**0.273**	**0.222**	**0.197**	**0.215**	**0.196**	**0.163**	**0.143**	**0.168**
X21	41802	51979	13590	34691	41311	47002	53498	62168	68997	72530	72763	90300	116728	146944
X22	46798	67256	64037	124453	185337	171670	159721	125079	104503	132053	161182	142275	139388	252627
X3	**0.995**	**1**	**1**	**0.998**	**0.76**	**0.978**	**0.94**	**0.978**	**0.592**	**0.175**	**0.513**	**0.669**	**0.529**	**0.487**
X31	7808	5604	9891	10913	4816	6495	7207	8401	4499	410	2961	3284	2202	1526
X32	0.147	0.197	0.1	0.116	0.159	0.158	0.147	0.184	0.18	0.197	0.215	0.182	0.149	0.18
X4	**0.512**	**0.445**	**0.567**	**0.587**	**0.586**	**0.578**	**0.616**	**0.543**	**0.516**	**0.305**	**0.279**	**0.33**	**0.335**	**0.305**
X41	0.2028	0.2003	0.243	0.1951	0.1184	0.1479	0.1209	0.1177	0.1039	-0.0749	-0.0801	-0.0194	0.0362	0.0362
X42	5.12	5.05	5.05	5.27	7.79	7.94	7.97	7.99	7.99	11.89	21.84	25.55	26.6	27.2

Fig. 3. The dynamics of changes in the initial indicators for the observed period

Next, we analyze the sensitivity [18] of the network using influence charts. Repeated training in parameters and repeated validation led to an increase in overall accuracy by 2% (from 55.4% to 57.4%), and the accuracy of the result increased by 14% (from 46% to 60%). The results are shown in Table 2:

Table 2. Comparison of the initial accuracy of the model with accuracy after sensitivity analysis on nodes Noisy - MAX

	Initial accuracy		**Accuracy after a sensitivity analysis**	
	Overall network accuracy,%	Accuracy of the result ,%	Overall network accuracy,%	Accuracy of the result ,%
Noisy-MAX nodes	55,4	46,0	57,4	60,0

6 Results and Discussion

If the level of innovation is increased to the maximum, the share of enterprises' bargaining will increase by 7% (from 34% to 41%), and this, in turn, will lead to an increase in the country's economic growth by 12% (from 24% to 36%), as shown in Figure 4.

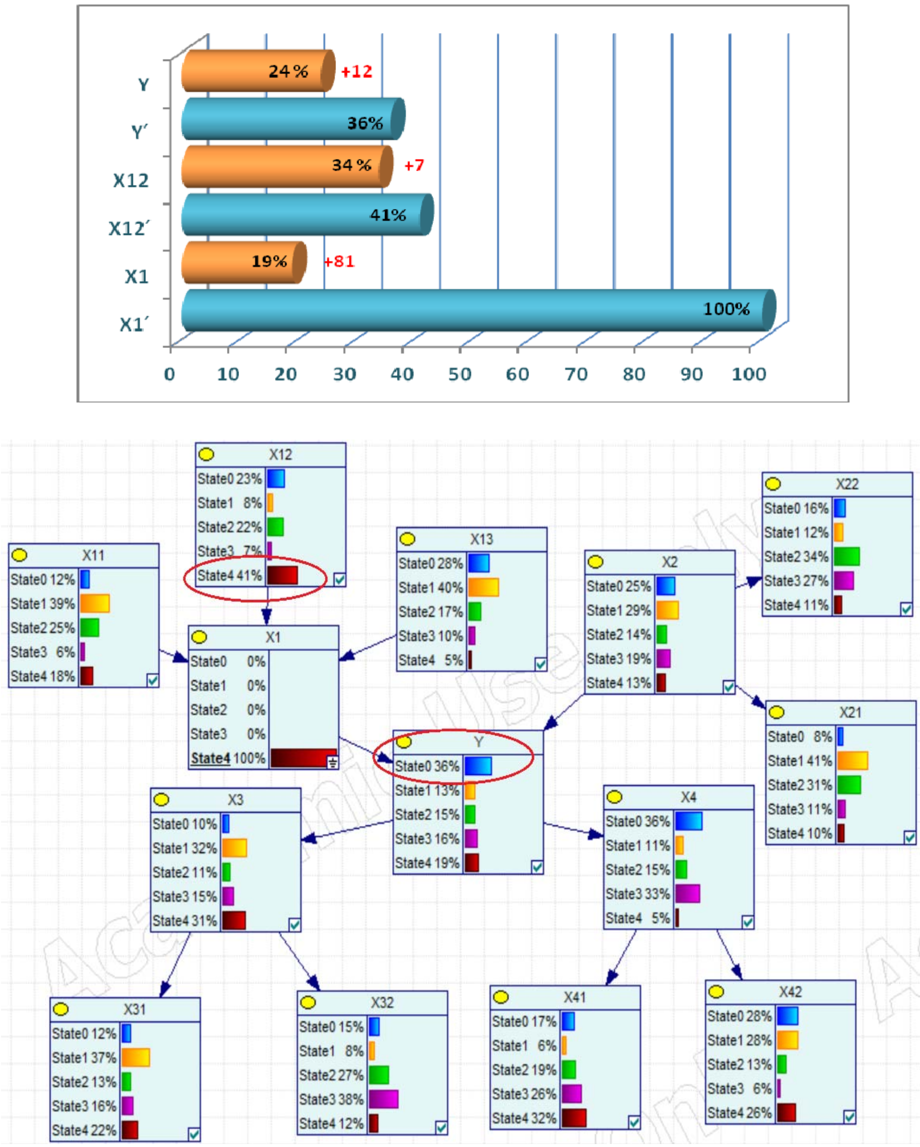
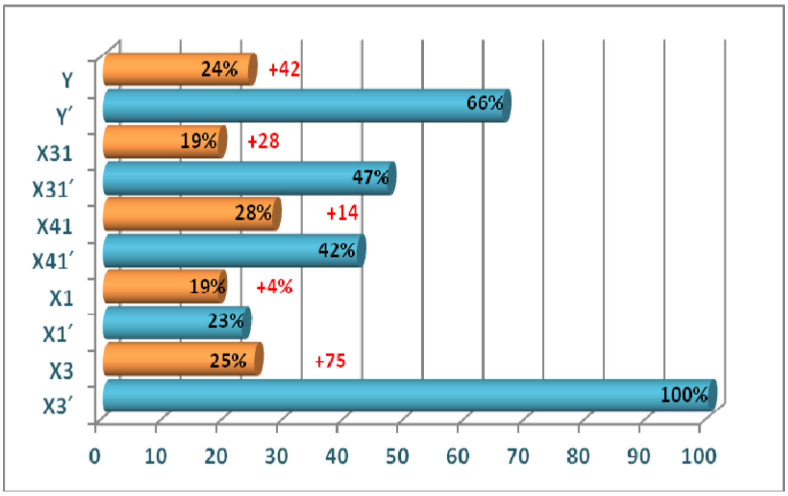


Fig. 4. Experiment Results

At the maximum level of external investment, the level of innovation will increase by 4% (from 14% to 23%), direct investment in Ukraine will increase by 28% (from 19% to 47%), the tendency of the population to save will increase by 14% (from 28% to 42%). All this together will create conditions for increasing the country's economic growth, which will increase by 42% (from 24% to 66%), as shown in Figure 5.



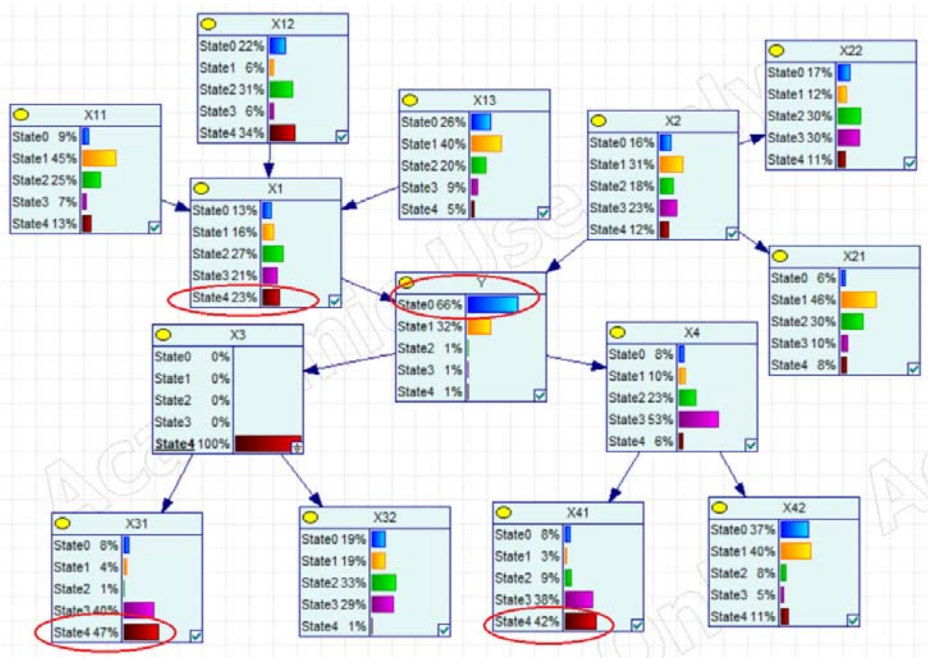


Fig. 5. Conditions for ensuring economic growth

7 Conclusion

This article has conducted a comparative study of the behavior of Noise Max nodes and common nodes when designing a Bayesian network. Noisy-max nodes have been shown to provide relatively high initial accuracy compared to conventional nodes.

Shared nodes require retesting. However, Noisy-max nodes entail an increase in time and computational cost.

Along with this, experiments were conducted to assess the general trends of the country's economic growth potential. It was found that at the maximum level of external investment, direct investment in Ukraine will increase by 28% (from 19% to 47%) and this creates the conditions for increasing the country's economic growth, which will increase by 42% (from 24% to 66%).

In our future research, we will try to trace the country's economic growth over time using the dynamic Bayesian network tool.

References

1. Shumpeter, Y.: Teoriia ekonomichnoho rozvytku. Prohress, 455p. (in Ukrainian) (1982).
2. Naukova ta innovatsiina diialnist v Ukraini u 2012–2017 rokakh. Statystychnyi zbirnyk. KDP, 289p. (in Ukrainian) (2017).
3. Díez, F., Galán, S.: Efficient computation for the Noisy-MAX, Int. J. Int. Syst., vol. 18/2, 165–177. (2004)
4. Billio, M., Getmansky, M., Lo, A., Pelizzon, L.: Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors. Journal of Financial Economics, vol. 104 (3), 535 – 559. (2012)
5. Battiston, S., Delli-Gatti, D., Gallegati, M., Greenwald, B., Stiglitz, J.: Liaisons Dangereuses: Increasing Connectivity, Risk Sharing, and Systemic Risk. Journal of Economic Dynamics and Control, vol. 36 (8), 1121–1141. (2012)
6. Brunnermeier, M., Pedersen, H.: Market Liquidity and Funding Liquidity/ Review of Financial Studie, vol. 22 (6), 2201–2238. (2009)
7. Granger, C. W.: Investigating Causal Relations by Econometric Models and Cross-spectral Methods. Econometric, vol. 37 (3), 424–438. (1969)
8. Geiger, D., Heckerman, D.: Parameter Priors for Directed Acyclic Graphical Models and the Characterization of Several Probability Distributions. Annals of Statistics, vol. 30 (5), 1412– 1440. (2002)
9. Giudici, P., Green, P.: Decomposable Graphical Gaussian Model Determination. Biometrika, vol. 86 (4), 785–801. (1999)
10. Miyamura, M., Kano, Y.: Robust Gaussian Graphical Modelin. Journal of Multi-variate Analysis, vol. 97 (7), 1525–1550. (2006)
11. Heckerman, D., Chickering, D.: Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. Machine Learning, 20–197. (1995)
12. Heckerman, D., Geiger, D.: Learning Gaussian Networks. Uncertainty in Artificial Intelligence, 235–243. (1994)
13. Madigan, D., York, J.: Bayesian Graphical Models for Discrete Data. International Statistical Review, vol. 63 (2), 215–232. (1995)
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, SF.: Morgan Kaufmann, 129-138. (1998)
17. Pollino, C. A., Woodberry, O., Nicholson, A., Korb, K., Hart, B.: Parameterisation and Evaluation of a Bayesian Network for Use in an Ecological Risk Assessment. Environmental Modelling and Software, February 5: In proceedings, vol. 22(8), 1-13. (2007)
18. 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)