International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

The Use of Bayesian Networks in Public Administration of the Economy

Sergii Lekar, Dmytro Shumeiko, Volodymyr Lagodiienko, Valerii Nemchenko, Olena Nikoliuk, Olena Martynyuk

Abstract: ***The state, exercising its managerial function in the economic sphere, pursues specific goals, the main of which is to ensure for society and the decent state welfare, as well as to increase the level of material production. Bayesian trust networks can be used in government project management as diagrams reflecting the causal relationships between events or influence diagrams. Besides, Bayesian trust networks allow integrating into the model the data obtained at each stage of project development, thereby providing feedback, which is an essential component of the public administration.***

Index Terms: ***Bayesian networks, public administration, economy, management.***

I. INTRODUCTION

Public administration in the economic sphere is the management activity of the country's executive authorities in the areas of industry, agriculture, transport, construction, finance and some other relations related to the economy in nature and content [1].

Large volumes of information accumulate around the world, which requires proper processing and decision-making based on the results of this processing. Data mining methods (DMM) provide the ability to automatically search for patterns characteristic of multidimensional data [2-7]. Most of the data mining tools are based on two technologies: machine learning [8] and visualization (visual presentation of information). Bayesian networks combine these two technologies.

II. THE RATIO OF PUBLIC ADMINISTRATION AND ECONOMY

The state, exercising its managerial function in the economic sphere, pursues specific goals, the main of which is to ensure for society and the decent state welfare, as well as to increase the level of material production. In this case, the result of the implementation of economic management should be the provision of human needs. Public administration in the field of economy pursues the implementation of many important tasks, which include:

- ensuring economic growth and development of the country;
- the provision of adequate and full employment;
- the desire to achieve efficiency in the field of economics;
- stability in the price segment; freedom of economic activity;
- equity in the distribution of income;
- ensuring the balance of foreign trade.

Activities related to the management of economic processes entail a solution to a wide range of issues, which is achieved by using specific methods, including financial and administrative (power) and mathematical methods.

III. BAYESIAN NETWORK - A TOOL OF INTELLECTUAL DATA ANALYSIS

The Bayesian network (belief network) is a probabilistic graph model, which is a set of variables and their probabilistic Bayesian dependencies.

When using Bayesian networks (BN) as a DMM tool, it is necessary to solve two mathematical problems:

1. constructing the BN structure
2. forming a probabilistic inference.

The task of constructing a BN for given training data is NP-hard, that is, a problem of nonlinear polynomial complexity. The number of all possible non-cyclical models that need to be analyzed is calculated using the Robinson recurrent formula [9]:

$$f\,(n) = \sum_{i=1}^{n} (-1)^{i+1} \cdot C_n^{i+1} \cdot 2^{i\cdot(n-1)} \cdot f(n-1) \qquad (1)$$

where n is the number of vertices, and f (0) = 1.

However, in practice, it is possible to perform a complete search of models only for networks with no more than seven vertices (nodes), because otherwise, there are not enough computational resources.

The task of forming a probabilistic inference in BN is essential and complex and belongs to the class of decision-making tasks.

Manuscript published on 30 June 2019.
***** Correspondence Author (s)
Sergii Lekar*, Department of Tax Policy, University of the State Fiscal Service of Ukraine, Irpin, Ukraine
Dmytro Shumeiko, Department of Criminal Procedure, National Academy of Internal Affairs, Kyiv, Ukraine
Volodymyr Lagodiienko, Department of Marketing, Business and Trade, Odessa National Academy of Food Technologies, Odessa, Ukraine
Valerii Nemchenko, Department of Accounting and Audit, Odessa National Academy of Food Technologies, Odessa, Ukraine
Olena Nikoliuk, Department of Management and Logistics, Odessa National Academy of Food Technologies, Odessa, Ukraine
Olena Martynyuk, Department of Management, International Humanitarian University, Odessa, Ukraine

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

It is evident that the existing methods of BN construction and formation of a conclusion require time-consuming computations. Therefore, the development of ways to reduce computational complexity is relevant and in demand when modelling processes of different nature by Bayesian networks.

A Bayesian network is a pair$< G, B >$ in which the first component of G is a directed non-cyclic graph that corresponds to random variables and is written as a set of independence conditions: each variable is independent of its parents in G. The second component of B is the set of parameters defining the network. The component contains parameters $\theta_{X^{(i)}|pa(X^{(i)})} = P(X^{(i)}|pa(X^{(i)}))$ for each possible value $x^{(i)} \in X^{(i)}$ and $pa\ (X^{(i)}) \in Pa(X^{(i)})$, where $Pa(X^{(i)})$ denotes a set of parent variables $(X^{(i)}) \in G$. Each variable $(X^{(i)}) \in G$ is represented as a vertex. If consider more than one graph, then to identify the parents of the variable $X^{(i)}$ in the graph G column use the designation $Pa^G(X^{(i)})$. The total BN joint probability is calculated by the formula $P_B(X^{(1)}, \dots, X^{(N)}) = \prod_{i=1}^{N} P_B(X^{(i)})|Pa(X^{(i)})$).

Currently, DMM technologies are developing rapidly, and their areas of application are expanding.

IV. Bayesian Networks in Public Administration

There are several ways to build Bayesian networks. The choice of method depends on the input parameters, laboriousness, relevance and demand. For the tasks of government regulation, we propose to use the probabilistic inference algorithm in BN based on the training data as a less difficult method compared to the Heuristic method of building Bayesian networks.

A. Input data required for the output generation algorithm.

1. The set of training data$D = \{d_1, \dots, d_n\}$, where $d_1 = \{x_i^{(1)} x_i^{(2)} \dots x_i^{(N)}\}$ (subscript is the observation number, and the upper one is the variable number), n – the number of observations. Each observation consists of N $(N \geq 2)$variables$X^{(1)}, X^{(2)}, \dots X^{(N)}$, each j-th variable $(j = 1, \dots, N)$ has $A^{(j)} = \{0,1, \dots, a^{(j)} - 1\}(\alpha^{(j)} \geq 2)$ conditions.

2. The structure of the BN g is represented by a set of N ancestors $(\Pi^{(1)}, \dots, \Pi^{(N)})$, i.e. for each vertex $j = 1, \dots, N$; $\Pi^{(j)}$ is the set of parent vertices, wherein $\Pi^{(j)} \subseteq \{X^{(1)}, \dots, X^{(N)}\} \backslash \{X^{(j)}\}$ (the vertex cannot be an ancestor for itself, that is, the loop in box must be absent).

3. The set of instantiated vertices $\{X^{(P_1)} = x^{(P_1)}, \dots, X^{(P_v)} = x^{(P_v)}\}$, i.e. vertices that are in a certain state with a unit probability. If the set of instantiated vertices is empty, then classical probabilistic inference should be used. Algorithm for the formation of output.

B. Algorithm for the formation of output

Step 1. A set of empirical values of the joint probability distribution of the entire network $P(X^{(1)}, \dots, X^{(N)})$ is calculated from the set of training data. According to the formula $P_{matrix}(X^{(1)} = x^{(1)}, \dots, X^{(N)} = x^{(N)}) = \frac{n[X^{(1)}=x^{(1)}, \dots, X^{(N)}=x^{(N)}]}{n}$, where n is the number of teaching observations, $x^{(j)} \in A^{(j)}$, and $n[X^{(1)} = x^{(1)}, \dots, X^{(N)} = x^{(N)}] = \sum_{j=1}^{n} I(X^{(1)} = x_j^{(1)}, \dots, X^{(N)} = x_j^{(N)})$, where the function $I(E) = 1$ when the predicate $E = true$, otherwise $I(E) = 0$.

Step 2. We iterate through all the vertices of the BN. If the vertex is not instantiated, then it is necessary to calculate the probability values of all possible states of this vertex. For this, a sequential enumeration of all rows of the matrix of empirical values of the joint probability distribution of the entire network is done. If the benefits of the vertices of the string coincide with the interests of the instantiated vertices and state 6 of the analyzed vertex, then the corresponding value $P_{\text{matrix}}(X^{(1)}, \dots, X^{(N)})$ is added to the probability value of the relevant state of the analyzed vertex. After that, the benefits of the probabilities of the states of the analyzed vertex are normalized.

Imagine an algorithm for calculating the probability values of all possible states of uninstantiated vertices:

$$for\ j = 1\ to\ N\ if\ X^j \notin \{X^{(P_1)}, \dots, X^{(P_v)}\}\ then$$

$$begin$$

$$sum\ = 0;$$

$$\forall x^{(j)} \in A^{(j)}\ do$$

$$begin$$

$$for\ k = 1\ to\ last\_string\_matrix\ do$$

$$begin$$

$$if\ (X_{matrix}^{(P_1)} = x^{(P_1)})\ and \dots and\ (X_{matrix}^{(P_v)} = x^{(P_v)})\ and\ (X_{matrix}^{(j)} = x^{(j)})\ then$$

$$begin$$

$$P(X^{(j)} = x^{(j)}) = P(X^{(j)} = x^{(j)}) + P_{matrix}(X_{matrix}^{(1)}, \dots, X_{matrix}^{(N)})$$

$$end;$$

$$end;$$

$$sum = sun + P(X^{(j)} = x^{(j)});$$

$$end;$$

$$\forall x^{(j)} \in A^{(j)}\ do$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

$$begin$$

$$P\left(X^{(j)} = x^{(j)}\right) = \frac{P\left(X^{(j)} = x^{(j)}\right)}{sum}$$

$$end;$$

$$end;$$

Output.

The output data are the probabilities of all possible states of all non-instantiated vertices.

V. EXAMPLES OF PRACTICAL USE THE BAYESIAN NETWORKS IN PUBLIC ADMINISTRATION

Example 1. For the development of an effective implementation plan, for example, a social project, we can use not only such factors as cost, effect and advertising. We can and should take into account personal information about the target residents, such as their habits, age, income, and so on. Solving such a task allows us to customize the implementation strategy of any government project for each consumer group separately. This allows for a more loyal attitude towards innovation and government agencies in general.

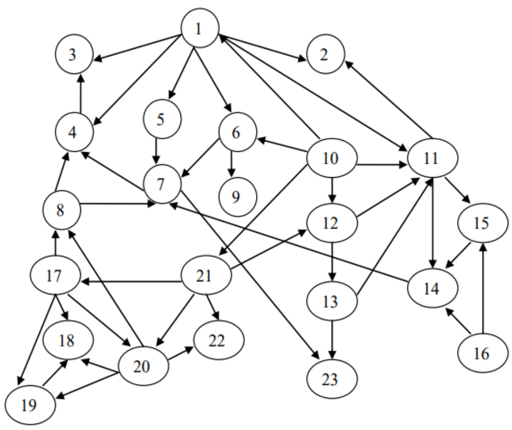

Fig. 1 The Bayesian network

1 – age of the client; 2 – retired; 3 – the number of children; 4 – marital status; 5 – how many years live at the address specified in the questionnaire; 6 – work experience in the current workplace; 7 – uses informational news; 8 – landline telephone; 9 – mobile phone; 10 – the total length of service; 11 – income; 12 – project costs; 13 – project category; 14 – the TV; 15 – CD player; 16 – Blue Ray; 17 – wireless communication; 18 – fax; 19 – organizer; 20 – computer; 21 – level of education; 22 – uses the Internet; 23 - consumer category.

Example 2. The city authorities plan to implement the project "Healthy Heart". The medical database has accumulated information on 1,605 patients with cardiovascular diseases. Fig. 2 shows a medical BN showing the relationship between the patient's diseases, surgery and health status.

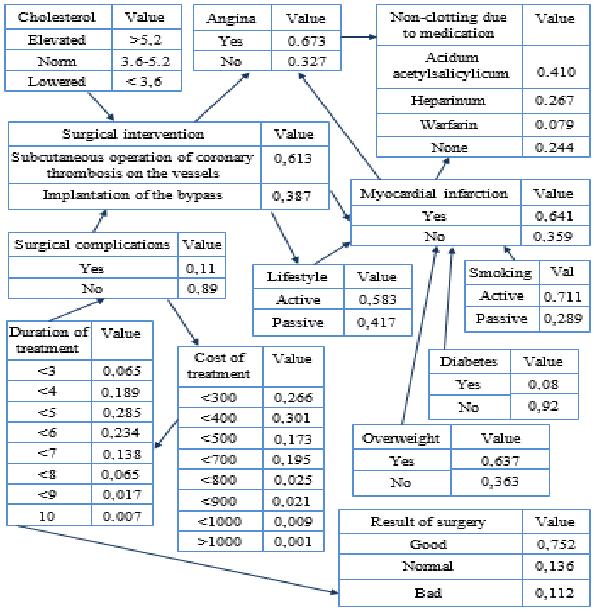

Fig. 2 The Bayesian network to predict the implementation of a social project

It is evident that the listed types of tasks are relevant for almost all branches and areas of human activity.

VI. CONCLUSION

Bayesian trust networks can be used in government project management as diagrams reflecting the causal relationships between events or influence diagrams. Besides, Bayesian trust networks allow integrating into the model the data obtained at each stage of project development, thereby providing feedback, which is a crucial component of the public administration.

REFERENCES

1. Ponomarenko T., Khudolei V., Prokopenko O., Klisinski J. "Competitiveness of the information economy industry in Ukraine", Problems and Perspectives in Management, Vol. 16, Issue № 1, 85-95, 2018
2. Karshiev Zaynidin, Ivan Kholod, "Block structure of data mining algorithms", IEEEE Nirma University International Conference on Engineering (NUiCONE), 2013
3. Lagodiienko Volodymyr, Malanchuk Marina, Gayvoronska Inna and Sedikov Denys. "Selection of criteria for key performance indicators by the matrix method", International Journal of Mechanical Engineering and Technology, 10(1), 1303-1311, 2019
4. Chintan Shah, Anjali Jivani, "Comparison of Data Mining Clustering Algorithms", IEEEE, 2016.
5. Iryna Bashynska, Marina Malanchuk, Olena Zhuravel, Kateryna Olinichenko, "Smart Solutions: Risk Management of Crypto-Assets and Blockchain Technology", International Journal of Civil Engineering and Technology (IJCIET) 10(2), 1121-1131, 2019
6. David E Booth. "Data Mining Methods and Models", Technometrics 49(4), 2012 https://doi.org/10.1198/tech.2007.s697
7. P.I. Bidyuk, V.H. Huskova. "Analysis of solvency using data mining methods", Èlektron. model. 41(2), 111-120, 2018 https://doi.org/10.15407/emodel.41.02.111
8. Svitlana Bondarenko, Volodymyr Lagodienko, Iryna Sedikova and Olga Kalaman, "Application of Project Analysis Software in Project Management in the Pre-Investment Phase", Journal of Mechanical Engineering and Technology, 9(13), 676-684, 2018
9. Joshua W. Robinson, Alexander J. Hartemink. "Learning non-stationary dynamic bayesian networks". J. of Mach. Learn. Res., 11, 3647-3680, 2010

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)