Bayesian Methods Application for the Differential Diagnosis of the Chronic Obstructive Pulmonary Disease

Volodymyr I. Lytvynenko[1], Mariia O. Voronenko[1], Olena Kovalchuk[2], Ulzhalgas Zhunissova[3] and Luidmyla N. Lytvynenko[4]

[1] *Kherson National Technical University, Berislavske Shosse, 24, Kherson, 73008, Ukraine*
[2] *National Pirogov Memorial Medical University, st. Pirogova, 56, Vinnytsya, 21018, Ukraine*
[3] *Astana Medical University, st. Beibitshilik 49A, Astana, 010000, Kazakhstan*
[4] *Kherson city neuropsychiatric center, st. Morska, 1, Kherson, 73003, Ukraine*

Abstract

The paper proposes a methodology for attachment Bayesian models in the differential diagnosis of a disease such as chronic obstructive pulmonary disease in different age groups patients with the obligatory presence of 1, 2, or three concomitant diseases in anamnesis. The ways for building the Bayesian models structure are considered. The medical experts, pharmacists, specialists were screening of input data for creation the probabilistic predicting system.

Keywords 1

Diagnostic Methods, Accompanying Illnesses, Bayesian Networks, Feature Selection methods, Clustering, The algorithm MeanShift

1. Introduction

The medical diagnostics main task is to define possible diagnoses based on examination data. The main difficulties faced by doctors at the medical diagnosis stage are:

- the amount of present illnesses is greater than the amount of ways for their diagnosis, which rises the misdiagnosis risk;
- the diagnosis ambiguity, i.e. the presence of comorbidities that strongly distort the symptoms picture;
- a large indicators number, which also leads to data distortion;
- lack of individual approach to each patient [1].

The medical diagnosis peculiarity is that a qualified doctor for a certain group of patients can clearly diagnose any disease, but each of the patients may have not one disease, but several at the same time. That is, the patient may have several diagnoses (the doctor may prescribe for each diagnosis a subjective level of belonging).

There may also be situations where the physician is unable to diagnose a patient's illness due to a lack of experience or due to atypical or distorted patient symptoms. From the computational intelligence standpoint, this situation is the subject of active learning [2], when the diagnostic system in the process of its training processes both various observations, which know exactly the value of the training signal (reference signal) and a priori unclassified observations (self-learning mode). The chronic obstructive pulmonary illness is the reason of death among children and the elderly today.

MoMLeT+DS 2021: 3rd International Workshop on Modern Machine Learning Technologies and Data Science, June 5, 2021, Lviv-Shatsk, Ukraine
EMAIL: immun56@gmail.com (V. Lytvynenko); mary_voronenko@i.ua (M. Voronenko); elena.kovalchuk972@gmail.com (O. Kovalchuk); Ulzhalgaszhunisova@gmail.com (U. Zhunissova); llytvynenko58@gmail.com (L.Lytvynenko)
ORCID: 0000-0002-1536-5542 (V. Lytvynenko); 0000−0002−5392−5125 (M. Voronenko); 0000-0002-3102-7973 (O. Kovalchuk); 0000-0001-5255-9314 (U. Zhunissova); 0000-0001-8445-5704 (L.Lytvynenko)

The incidence of the chronic obstructive pulmonary disease in elderly patients increases dramatically while they are being treated in hospital facilities for another concomitant disease.

Modeling and differential diagnosis, taking into account the presence of several comorbidities, are quite difficult, since they are characterized by a complex structure of dependences. The computational problems are also obvious since the analysis of a large database requires a long processing time and can be performed only with adequate computational infrastructure. When resources are limited and research is mainly based on open data, there is a need for models that can be modified depending on the variability of functions, and thus can further use the experience to improve the predictability of the model. Bayesian networks (BNs) are the most suitable computational tools for this.

The study is dedicated to the creation of a system that makes it possible to achieve the effectiveness of drug therapy in the presence of the patient's main diagnosis and several concomitant diagnoses that aggravate the course of the disease. The task of this work is to make a methodology for constructing a BN in the differential diagnosis of a disease such as the chronic obstructive pulmonary disease.

The main contribution is as follows:

(i) realization of a comprehensive causal probabilistic model for the differential diagnosis of three types of pulmonary diseases,(ii) development of a specific methodology for planning the development of Bayesian networks for differential diagnosis problems, (iii) application of the MeanShift cluster analysis algorithm to identify specific signs, symptoms, laboratory and instrumental findings characteristic of the disease forms under study, and (iv) the use of algorithms to estimate the information entropy, which made it possible to assess the informativeness of the features about the corresponding class of the disease.

Thus, the main problem is related to improving the quality of differential diagnosis in medical decision support systems. The key problem focuses on improving the quality of differential diagnosis through the comprehensive use of methods for selecting the most informative features.

The article is drawn up like this. Section 2 defines the tasks that we will solve in our study. Section 3 shows the literature list on existing ways for the differential diagnosis. In Section 4 we discuss the general formulation of the solution to the problem. Section 5 describes the input data and methods for obtaining structural model indicators. After that, in the Section 6 we describe the sequence of building and validating a BN. We present the research process and its results. Section 7 presents the study results analysis. Section 8 summarizes and completes.

2. Problem Statement

Using Bayesian Approach the developed model will facilitate a faster diagnosis and thus help to pre-select the correct treatment method. This mathematical tool let us to describe the notions, which the patient uses when describing his health [3].

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

Based on a given training sample, you need to build an acyclic graph connecting the event sets $X_i, i=I,...,N$. In addition, each BN structure $g \in G$ is presented by a set N of predecessors $\left(P^{(1)},\text{K},P^{(N)}\right)$, that is, for each vertex $j = 1,\text{K},N, P^{(j)}$ it is a variety of parent vertices, such that $P^{(j)} \subseteq \left\{X^{(1)},\text{K},X^{(N)}\right\} \setminus \left\{X^{(j)}\right\}$. Research will be carried out in accordance with the following stages of Bayesian network development (Figure 1).

The aim of this engineering is to create the Bayesian-based model for the early diagnosis of the chronic obstructive pulmonary disease if the patient has the likelihood of concomitant diagnoses.

3. Review of the Literature

Many different fuzzy approaches to data clustering are currently used [4-6]. They are able to effectively cluster data in situations where clusters overlap, assuming that the cluster size is small, i.e. do not contain abnormal emissions. Real medical and biological data sets contain up to 20% of emissions. The data of these indicators allow forming clusters that allow to estimate dynamics of indicators, to simplify and accelerate the process of diagnostics.

Existing methods for estimating biological conditions are based on the values arrays analysis of the measured parameters set given in the vectors form. These can be immunological, biochemical, physiological, cytogenetic, cytomorphological, and other medical and biological data.

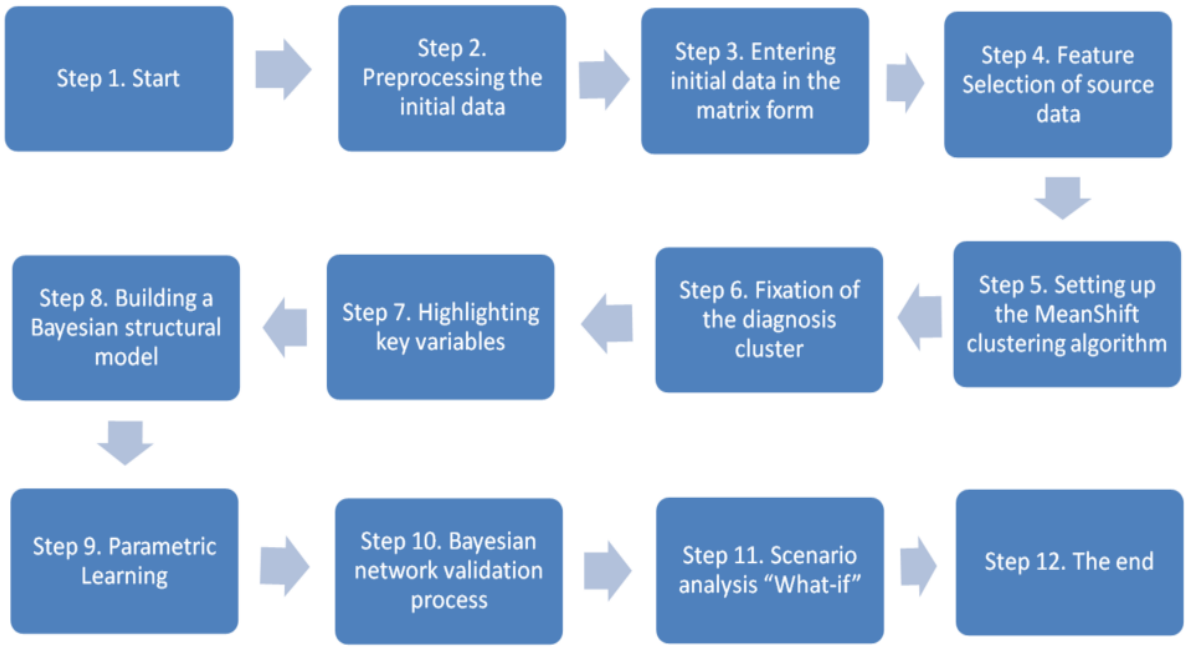


Figure 1: The stages according to which the research goal will be achieved.

The measurement data analysis of an individual condition of a separate organism consists in the definition of belonging of this condition to any of in advance known conditions (diagnoses). Each specialist in certain types of diseases makes "his" diagnosis. This is due to objective reasons related to the overlap and coincidence of different numbers of indicators of the state of the body for different diagnoses, and the widespread prevalence of so-called polysyndromic conditions.

Making correct and effective decisions in such situations requires a significant amount of time and money to organize consultations of highly qualified persons. The most stringent in terms of fuzzy clustering procedures are the so-called fuzzy clustering algorithms, based on objective functions [5, 7, 8] and made automatic classification (without teacher) by optimization predefined quality criteria. As a result of this procedure, the formed clusters have the shape of hyperspheres, which significantly limits the possibility of using the above methods to process data of more complex forms. In the artificial intelligence field, the computer system that imitates the possibility to make human decisions is called an expert system [9]. The first expert systems were created in the 1970s and became widespread in the 1980s. They were one of the first truly successful artificial intelligence forms. However, some experts note that expert systems were not part of true artificial intelligence, as they do not have the ability to learn independently of external data [10].

The MYCIN system, for example, has been designed to recommend the required amount of antibiotics depending on the patient's weight [11]. Studies at Stanford Medical School found that MYCIN offered acceptable therapies in approximately 69% of cases, better than infectious disease experts who were evaluated according to the same criteria.

However, when looking at the expert systems in real operating conditions, there are other problems, such as integration, access to large databases, and performance [12]. The CADUCEUS medical expert system was created to help diagnose blood infections as well as diagnose internal diseases. CADUCEUS was able to recognize up to 1,000 different diseases, as well as to recognize the several concomitant diseases presence. Fuzzy medical diagnosis in terms of increasing numbers signs and diagnoses are represented by online diagnosis systems. To date, there are medical online diagnostic systems DNFS and AWDNFN, designed on the basis of W-neuron. During their operation, there will be a worse criterion for the effectiveness of AWDNFN than in DNFS due to the longer processing time with better diagnostic accuracy [13].

In the past, many authors have tried to create computational intelligence systems for diagnosis using data sets from the medical repository as input data [14-18]. The inability to process data in real time, a fixed number of medical signs, and a low convergence rate are the disadvantages of these systems. In [19] was proposed the use of probabilistic methods of medical diagnosis, which raised expert systems to a new level of development.

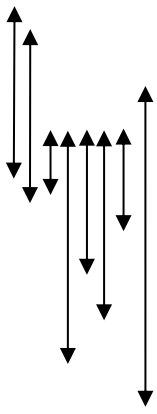
4. Materials and Methods

4.1. Data

The study involved 137 patients of different ages, different anamnesis, and varying disease severity. The data set describes the results of clinical tests of patients, symptoms, indicators of preliminary examination and complaints with which the patient consulted a doctor. All patients were given a preliminary diagnosis of the chronic obstructive pulmonary disease, and each patient has two more concomitant diagnoses that can complicate the course of the infection and the patient's treatment. Our primary task is to identify the presence of concomitant diagnoses and predict scenarios in which these diagnoses may complicate the course of the disease or aggravate the ongoing medical therapy. In total, the data set contains 104 parameters that describe the condition of each of 137 patients.

In situations where medical symptoms may include abnormal emissions, obstructions, and other artifacts, robust online procedures should be used that allow for consistent self-learning diagnostics using arbitrary cluster diagnoses. We carried out preprocessing of the data using the Feature Selection methods of optimization, and after clustering according to the algorithm MeanShift, the key indicators that have the greatest impact were identified (see Table 1). Among them, there are known to us input indicators and unknown indicators, the probability of occurrence of which we need to determine.

Table 1
The Input data

Designation of the node	Description	Interaction scheme
χ 9	Suffocation	
χ 18	Increased body temperature	
χ 40	Harmful working conditions	
χ 47	Chest shape	
χ 50	Breathing is weakened	
χ 51	Wheezing rales	
χ 52	Wet wheezing	
χ 53	Dry wheezing	
χ 57	The volume of wheezing	
χ 89	Focus of inflammation	
Y 1	Diagnosis 1 yes / no?	
Y 2	Diagnosis 2 yes / no?	
Y 3	Diagnosis 3 yes / no?	

The network contains ten nodes:

1. $\chi 9$ corresponds to the patient's suffocation. This indicator has two states s0 and s1:
   - s0 means no suffocation from the patient,
   - s1 means the patient has suffocation.
2. $\chi 18$ corresponds to the increased patient's body temperature. This indicator has two states s0 and s1:
   - s0 means the absence of information about increased patient's body temperature,
   - s1 means the presence of information about increased patient's body temperature.
3. $\chi 40$ corresponds to the harmful working conditions in which the patient works. This indicator has two states s0 and s1:
   - s0 means the absence of the harmful working,
   - s1 means the presence of the harmful working.
4. $\chi 47$ is the patient's chest shape. This indicator has two states s0 and s1:
   - s0 refers to the barrel-like shape of the patient's chest;
   - s1 refers to the patient's round chest.
5. $\chi 50$ is the patient's breathing is weakened/not weakened. This indicator has two states s0 and s1:
   - s0 means that the patient's breathing isn`t weakened;
   - s1 means that the patient's breathing is weakened.
6. $\chi 51$ corresponds to the wheezing/no wheezing in the patient's lungs. This indicator has two states s0 and s1:
   - s0 means no wheezing in the patient's lungs,
   - s1 means wheezing in the patient's lungs.
7. $\chi 52$ means presence /absence wet wheezing in the patient's lungs. This indicator has two states s0 and s1:
   - s0 means the absence of wet wheezing in the patient's lungs,
   - s1 means the presence of wet wheezing in the patient's lungs.
8. $\chi 53$ means the presence/absence of dry wheezing in the patient's lungs. This indicator has two states s0 and s1:
   - s0 means the absence of dry wheezing in the patient's lungs,
   - s1 means the presence of dry wheezing in the patient's lungs.
9. $\chi 57$ corresponds to the examination results of the patient's wheezing loudness. This indicator has two states s0 and s1:
   - s0 means low volume of wheezing;
   - s1 means high volume wheezing.
10. $\chi 89$ corresponds to the information whether there are foci of inflammation. This indicator has two states s0 and s1:
    - s0 means that there are no foci of inflammation;
    - s1 means that there are foci of inflammation.

4.2. Future Selection

Modern data arrays, to which certain Data Mining methods can be applied, can be described by a large amount of data that form a large-dimensional feature space. Therefore, the proportion of such a space reducing to a dimension that allows data processing and/or visualization without unnecessary difficulties is very urgent. The solution to such a problem is called the optimization of the feature space or the search for significant features (Feature Selection, or Feature Engineering).

Data preprocessing is the most important stage, the quality of which determines the possibility of obtaining high-quality results of the entire data analysis process. Feature Selection for the task being implemented consists of choosing the most informative, useful features and excluding uninformative

features from consideration without transforming the original data space [20, 21]. We used the ID3 (Iterative Dichotomizer) algorithm proposed by D. Quinlan, which determines the order of a variable and its attributes through their informational significance (informational entropy) [22]. To do this, find the entropy of all unused features and their attributes relative to test specimens and choose the one for which the entropy is minimal (and the information content is maximal). The entropy under the condition of not equiprobable events p_i is found by the well-known Shannon formula:

$$I = -\sum_i p_i \log_2 p_i \ , \tag{1}$$

where I is the amount of information in bits that can be transmitted using m elements in the message with n letters in the alphabet, and $p_i=m/n$.

4.3. Clustering algorithm MeanShift

MeanShift is a nonparametric modus for determining the location of the probability density maximum [23]. The mean shift algorithm basically assigns data points to clusters iteratively, shifting the points towards the highest data points density, that is, the centroid of the cluster. The Rosenblatt – Parsen estimate is one of the most widely used for nonparametric data density estimation [24].

The density is estimated as the total influence of the sample elements, while the contribution of each element is described by the bell-shaped function $K(x)$, which depends on the distance to this element. The formula for calculating the density estimate $f(x)$ with the smoothing parameter (bandwidth) P at an arbitrary point x has the form:

$$\hat{f}(x) = \frac{1}{NP^d} \sum_{i=1}^{N} K\left(\frac{x - x_i}{P} \right). \tag{2}$$

As $K(x)$, we can use the classical Gaussian kernel:

$$K_G\left(\frac{x - x_i}{P} \right) = \exp\left(-\frac{\|x - x_i\|}{2P^2} \right). \tag{3}$$

However, in practice, in order to reduce computational costs, limited kernels are used, such as, for example, the Epanechnikov kernel:

$$K_{Ep}\left(\frac{x - x_i}{P} \right) = \left(1 - \frac{\|x - x_i\|}{P^2} \right) \cdot I\left(\|x - x_i\| \le P^2 \right), \tag{4}$$

where $I(x)$ is the indicator function.

In this approach, clusters correspond to local maxima of the density estimation function (modes). And the data elements refer to clusters using the MeanShift procedure) [25], converging along the gradient to the corresponding local maximum. An iterative procedure, starting its work from a point, sequentially moves to a shift point $x_{k+1}=m(x_k)$ until convergence, where:

$$m(x) = \frac{\sum_{i=1}^{N} x_i \cdot K(x - x_i)}{\sum_{i=1}^{N} K(x - x_i)}. \tag{5}$$

The vector is called the "mean shift" vector and its direction coincides with the direction of the maximum density growth at the point x. Clustering algorithms based on the use of the mean shift procedure allow obtaining high-quality partitions, however, the main problem for using this approach is the high computational complexity [26].

4.4. Bayesian network methods

Let $G = (V, Bi)$ be a graph in which the ending V is a set of variables; B_i is non-reflexive binary relation on [27]. Each variable v has a kit of parent variables $c(v) \subseteq V$ and a kit of all descendants

$d(v) \subseteq V$. A set $s(v)$ is a set of child variables for a variable *v* and $s(v)$ is a subset of $d(v)$. Let's also mark, that:

$$(a(v) \subseteq V) = V - (d(v) \cup \{v\}). \quad (6)$$

That is, $a(v)$ is a kit of propositional signs from the set *V*, excluding the variable *v* and its descendants. The set of variables B, is the contexture of parameters defining the model. It constitution $Q_{x^i|pa(X^i)} = P(x^i \mid pa(X^i))$ for each x^i amount from X^i and $pa(X^i)$ from $Pa(X^i)$, where $Pa(X^i)$ means the variable X^i parents set in G . Each sign X^i in graph G is proposed as an apex. If we have more than one graph, then we use the notation to recognize the parents $Pa^G(X^i)$ in graph G [28]. The total probability B of Bayesian model is specified by the formula:

$$P_B(X^1,...,X^N) = \prod_{i=1}^{N} P_B(X^i \mid Pa(X^i)). \quad (7)$$

The parametric learning procedure purpose is to discover the most likely θ variables that interpret the data [29].

For the validation procedure, we chose the algorithm presented in [30]. The method of maximum expectation EM is a procedure of iterations, which was created for solve optimization tasks of some functionality, using an analytical search for the objective function extremes. This way is divided into two steps. At the first step of "expectation" (E - expectation) on the basis of available observations (patients) the expected values for each incomplete observation are calculated. After receiving the filled data set, the basic statistical parameters are estimated. In the second stage, "maximization" (M - maximization) maximizes the degree of compliance of the expected and actually substituted data [31].

5. Experiments and Results

The Bayesian model was established via the GeNIe 2.4 Academic software environment. The structural static Bayesian model is shown in Figure 2. The Bayesian formula is used in Bayesian networks as an inference tool to find a solution. If the Bayesian network is used to recognize (identify) objects, then many factors are replaced by factors or characteristics of a particular object.

Selecting a set of instantiated variables separately has its advantages and disadvantages. The advantage of this representation is that it prevents looping when forming the output. If the output is not selected separately, there is a risk that the messages will affect each other and the network will become unstable.

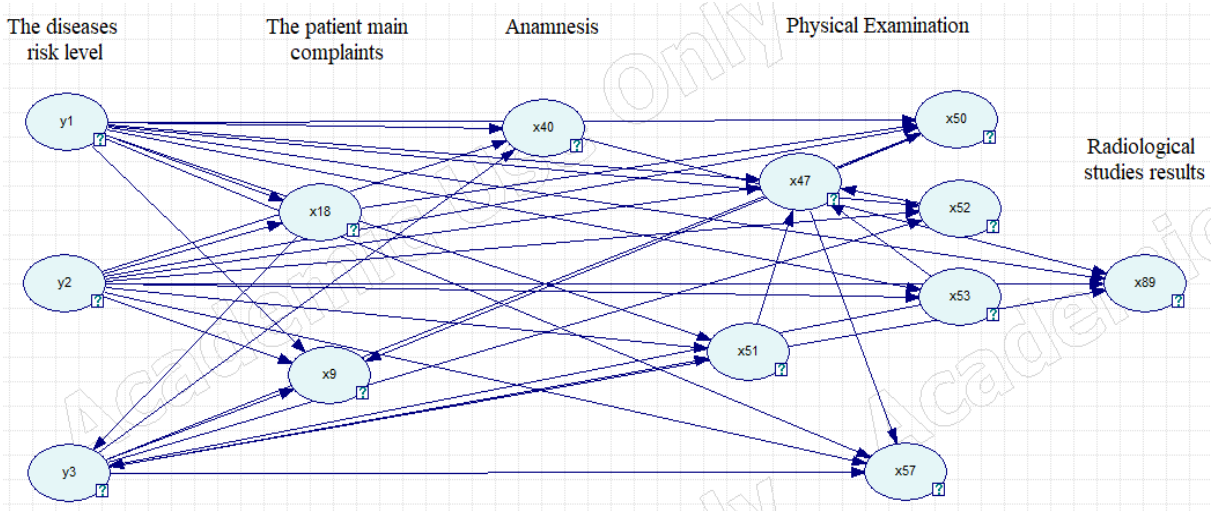


Figure 2: The static Bayesian network structural model

The disadvantage of this representation is that computing costs increase. However, the advantages are so great that the allocation of instantiated variables is completely justified. The final decision to confirm the effect between pollution data and test results, as well as the appointment of treatment, is made by the doctor. In Table 2, we present the results of modeling each specific case from a sample of data. The table shows the predicting result of confirming or refuting an early diagnosis.

Table 2
The scenario analysis results

Condition	Result	Description of cases, of interest
Clinical case 1: χ 9 max	χ50 will decrease by 11% (54 to 43) χ52 will increase by 3% (47 to 50) χ57 will increase by 3% (46 to 49) Y 1 will decrease by 3% (44 to 41) Y 3 will decrease by 10% (61 to 51)	When choking is at its maximum, the likelihood of breathing faster will decrease by 11%. The risk of being diagnosed with Y3 will decrease by 10%
Clinical case 2: χ 18 max	Y 2 will decrease by 6% (43 to 37) Y 3 will increase by 4% (61 to 65)	When the body temperature has reached its maximum, the risk of a diagnosis of Y2 will decrease by 6%, and a diagnosis of Y3 will increase by 4%.
Clinical case 3: χ 40 max	Y 2 will increase by 7% (43 to 50) Y 1 will decrease by 5% (44 to 39) Y 3 will increase by 2% (61 to 63)	In the presence of harmful working conditions, the risk of diagnosis Y2 will increase by 7%, and Y1 will decrease by 5%.
Clinical case 4: χ 50 min	Y 1 will decrease by 5% (44 to 43) Y 3 will increase by 4% (61 to 64) χ9 will increase by 10% (45 to 64)	In the presence of impaired breathing, the risk of choking increases by 10%
Clinical case 5: χ 50 max	Y 1 will increase by 4% (44 to 48) Y 3 will decrease by 3% (61 to 58) χ9 will decrease by 9% (45 to 36)	If the patient's breathing is weakened as much as possible, the risk of a diagnosis Y1 increases by 4%.
Clinical case 6: χ 52 min	Y 2 will increase by 4% (43 to 47) Y 3 will decrease by 6% (61 to 55) χ9 will increase by 3% (45 to 48) χ57 will increase by 3% (46 to 49)	If the level of wet wheezing decreases, the risk of diagnosis Y3 will decrease by 6%. In this case, choking and wheezing volume may increase by 3%
Clinical case 7: χ 52 max	Y 2 will decrease by 3% (43 to 40) Y 3 will increase by 6% (61 to 67)	In the presence of pronounced wet wheezing, the risk of the diagnosis of Y3 increases by 6%
Clinical case 8: χ 51 min	Y 1 will increase by 4% (44 to 48) Y 2 will decrease by 4% (39 to 43) Y 3 will increase by 4% (61 to 65)	If the level of wheezing is minimal, the risk of diagnosis Y2 will be reduced by 4%
Clinical case 9: χ 53 min	Y 1 will increase by 6% (44 to 50) Y 2 will decrease by 4% (43 to 39) Y 3 will increase by 2% (61 to 63)	If the level of dry wheezing is minimal, the risk of diagnosis Y1 increases by 6%
Clinical case 10: χ 53 max	Y 1 will decrease by 6% (44 to 38) Y 2 will increase by 4% (43 to 47) Y 3 will decrease by 2% (61 to 59)	In the presence of pronounced dry wheezing, the risk of the diagnosis of Y2 increases by 4%, while the risk of making a diagnosis of Y1 will decrease by 6%
Clinical case 11: χ 57 max	Y 2 will increase by 5% (43 to 48) Y 3 will increase by 3% (61 to 64)	At the maximum volume of wheezing, the risk of making a diagnosis of Y2 increases by 5%

6. Discussion

Now let's analyze clinical cases. The analysis results are shown in Table 2 and in Figures 3-6. To illustrate case 1 from Table 2: When choking is at its maximum, the likelihood that breathing will be faster will decrease by 11%. At the same time, the risk of making a diagnosis of $Y3$ will decrease by 10%, as shown in the Figure 3.

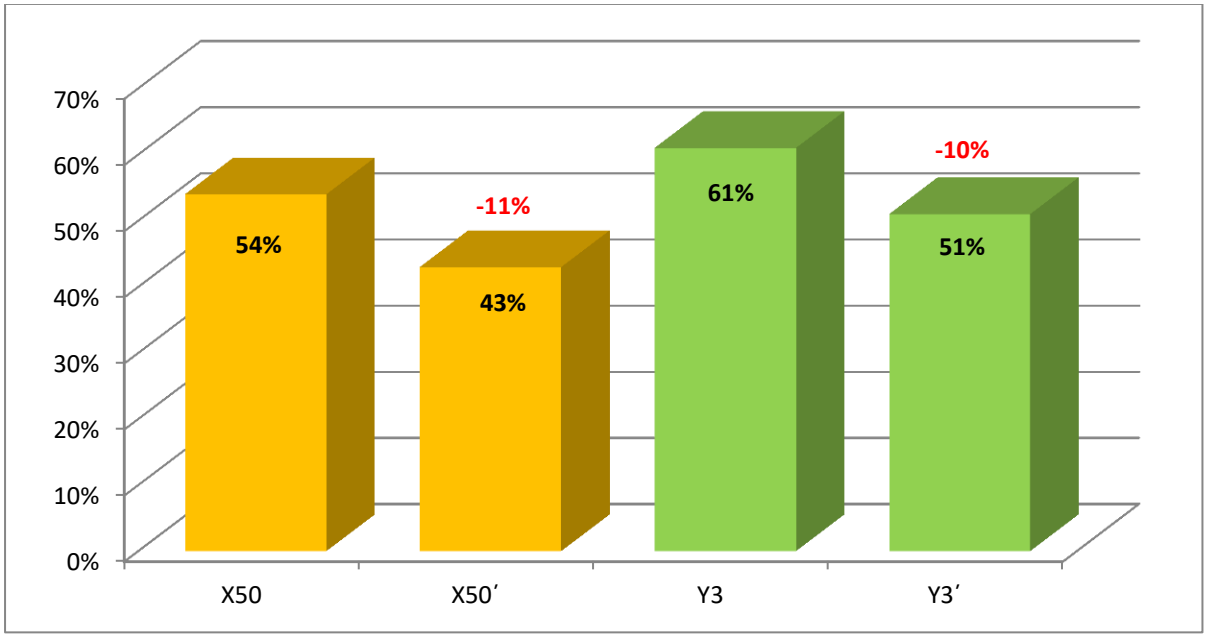


Figure 3: The internal profitability index tends to the maximum

Case 2: Upon reaching the maximum level of body temperature, the risk of a diagnosis of $Y2$ will decrease by 6%, and a diagnosis of $Y3$ will increase by 4%, as shown in the Figure 4.

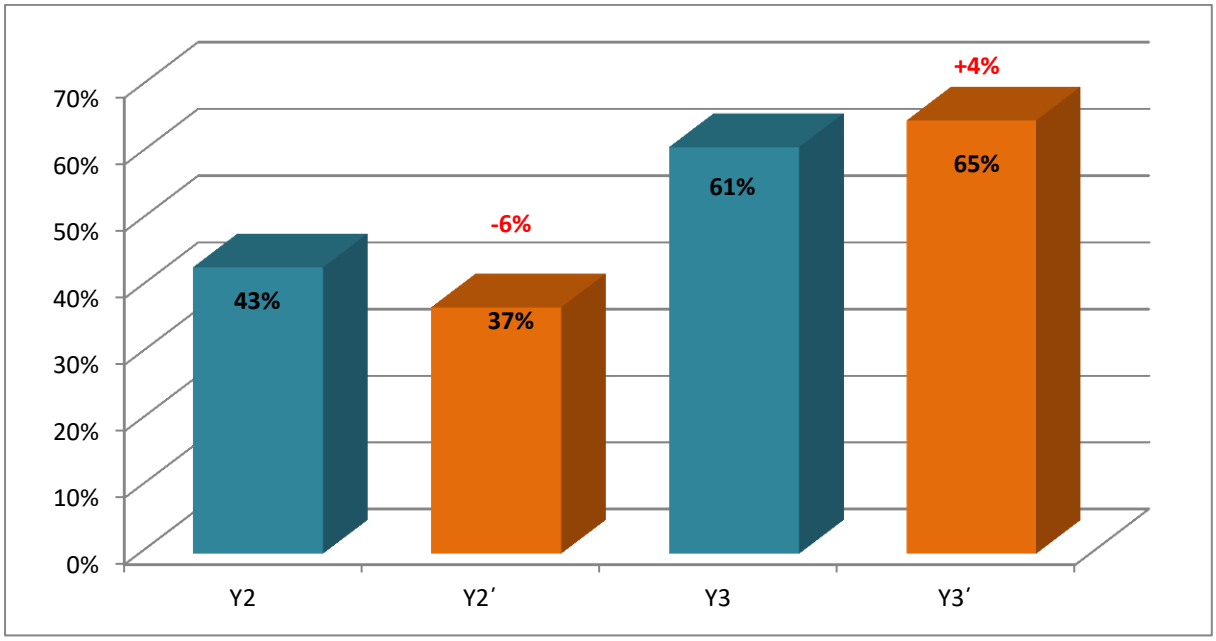


Figure 4: Clinical case 2

Cases 4,11 and 12: If the patient's professional activity takes place in the presence of the maximum harmful working conditions, the risk of diagnosis $Y2$ will increase by 7%, and $Y2$ will decrease by 5%, as shown in the Figure 5.

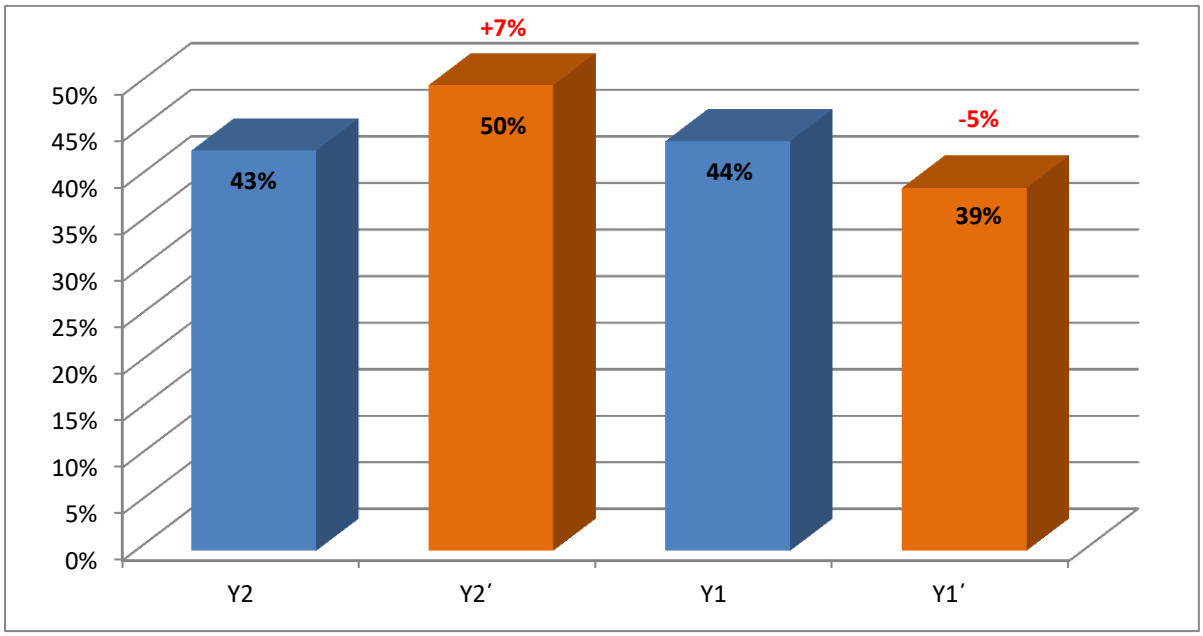


Figure 5: Clinical cases 4,11, and 12

Results of the analysis of case 10 are shown in the Figure 6.

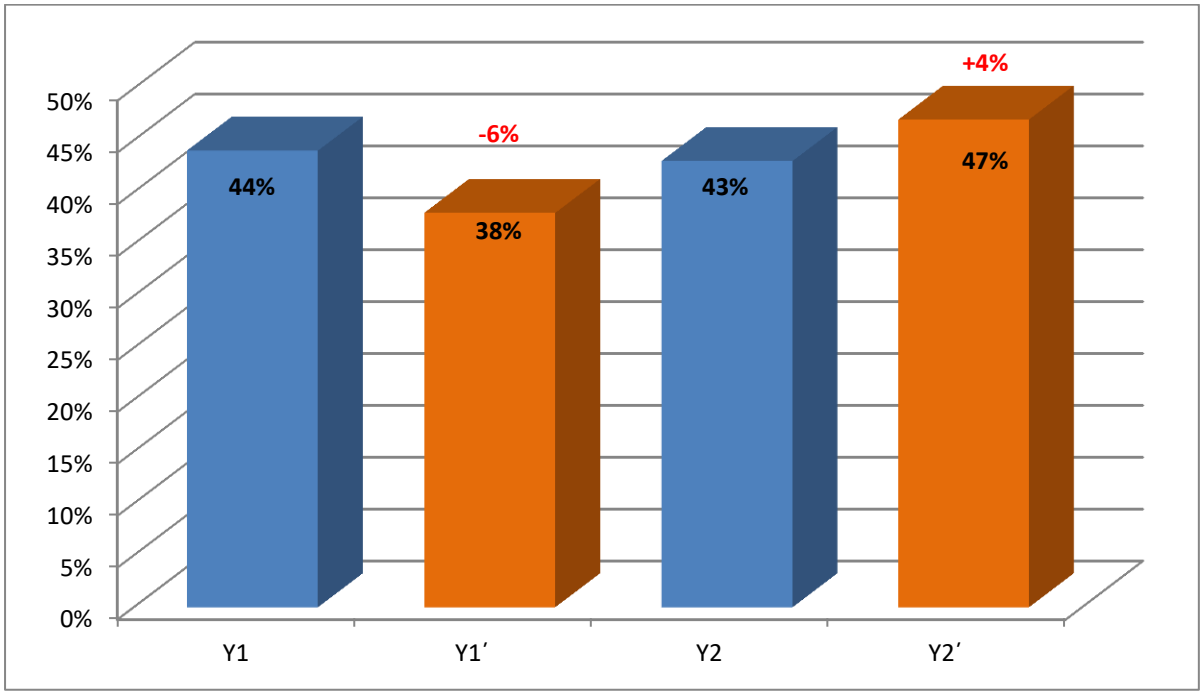


Figure 6: Clinical case 10

7. Conclusion

The chronic obstructive pulmonary disease gives dangerous complications to various organs of a sick person, that is, patients with hypertension, diabetes, and coronary heart disease are at risk of increased adverse outcomes.

The incidence of the chronic obstructive pulmonary disease depends on many factors: the standard of living, social and marital status, working conditions, contact with animals, travel, the presence of

bad habits, contact with sick people, individual characteristics of a person, the geographic prevalence of a particular pathogen.

Our study received comments from physicians and community physicians regarding the applicability of our probabilistic estimates of diagnosis as a working tool of daily practice. Our Bayesian system has also been evaluated by non-medical independent experts for comments from a patient perspective.

8. References

[1] O. Turuta, I. Perova, A. Deineko, Evolving flexible neuro-fuzzy system for medical diagnostic tasks, International Journal of Computer Science and Mobile Computing IJCSMC (2015) vol.4, Issue 8, 475-480.

[2] H. Kahramanli, N. Allahverdi, Design of a hybrid system for the diabetes and heart diseases, Expert Systems with Applications (2008) vols. 1-2, no. 35, 82–89.

[3] E. Lughofer, Single pass active learning with conflict and ignorance, Evolving systems (2012) vol. 4, no. 3, 251-271.

[4] V. Lytvyn, A. Hryhorovych, V. Hryhorovych, L. Chyrun, V. Vysotska, M. Bublyk, Medical Content Processing in Intelligent System of District Therapist, in: Proceedings of the 3rd International Conference on Informatics & Data-Driven Medicine (IDDM 2020), Växjö, Sweden, November 19 - 21, 2020, 415-429. http://ceur-ws.org/Vol-2753/paper29.pdf.

[5] B. Badiani, A. Messori, Targeted Treatments for Pulmonary Arterial Hypertension: Interpreting Outcomes by Network Meta-analysis, PharmD Heart, lung & circulation, (2015), vol.25 (1), 46-52.

[6] G. Battineni, G. Sagaro, N. Chinatalapudi, F. Amenta, Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis, Journal of personalized medicine, 2020-06-01, vol.10 (2), p.21.

[7] D. Luciani, M. Marchesi, G. Bertolini, The role of Bayesian Networks in the diagnosis of pulmonary embolism Journal of thrombosis and haemostasis, 2003-04, vol.1 (4), 698-707.

[8] D. Luciani, S. Cavuto, L. Antiga, M. Miniati, S. Monti, M. Pistolesi, G. Bertolini, Bayes pulmonary embolism assisted diagnosis: a new expert system for clinical use, Emergency medicine journal : EMJ, 2007-03, vol.24 (3), 157-164.

[9] Y. Bodyanskiy, Computational intelligence techniques for data analysis, in: Lecture Notes in Informatics, 2005, vol.5(72), pp. 15-36.

[10] K. Kaplan and M. Haenlein, Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence, Business Horizons, 2019, pp. 15-25.

[11] L. Chyrun, E. Leshchynskyy, V. Lytvyn, A. Rzheuskyi, V.Vysotska, Y. Borzov, Intellectual Analysis of Making Decisions Tree in Information Systems of Screening Observation for Immunological Patients. In: CEUR Workshop Proceedings, Vol-2362, (2019), 281-296.

[12] S. Kendal, M. Creen, An introduction to knowledge engineering, London: Springer, 2007, p. 287.

[13] S. Sahan, K. Polat, H. Kodaz, S. Gunes, The medical applications of attribute weighted artificial immune system (AWAIS): diagnosis of hearts and diabetes diseases, in: Proceedings of the 4th international conference on Artificial Immune Systems, Banff, AB, Canada, (2005).

[14] S. Lekkas, L. Mikhailov, Evolving fuzzy medical diagnosis of Pima Indians diabetes and of dermatological diseases, Artificial Intelligence in Medicine, (2010), no. 50, 117–126.

[15] K. Polat, S. Gunes, A. Arshlan, A cascade learning system for classification of diabetes disease: generalized discriminant analysis and least square support vector machine, Expert Systems with Applications, 2008, no. 34, pp. 482–487.

[16] P. Lucas, Bayesian networks in medicine : a model-based approach to medical decision making (2001). https://doi.org/10.1.1.22.4103

[17] H. Temurtas, N. Yumusak, F. Temurtas, A comparative study on diabetes disease diagnosis using neural networks, Expert Systems with Applications, 2009, no. 36, 8610–8615.

[18] C. Bojarczuk, H. Lopes, A. Freitas, E. Michalkiewicz, A constrained-syntax genetic programming system for discovering classification rules: application to medical data sets, Artificial Intelligence in Medicine, 2004, vol. 1, no. 10, 27-48.
[19] D.-C. Li, C.-W. Liu, S. C. Hu, A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets, Artificial Intelligence in Medicine, 2011, vol. 1, no. 45, pp. 45-52.
[20] M.L. Raymer, W.F. Punch, E.D. Goodman, L.A. Kuhn, L.C. Jain, Dimensionality reduc-tion using genetic algorithms, IEEE Trans. on Evolutionary Computation, 2000, no. 4 (2), 164-171.
[21] V. Lytvynenko, M. Voronenko, D. Nikytenko, N. Savina, O. Naumov, Assessing the possibility of a country's economic growth using dynamic bayesian network models. In: IEEE-2019 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). vol. CFP19D36-PRT, (2020), pp. 60-63.
[22] A. Magrini, D. Luciani, F. Stefanini, A probabilistic network for the diagnosis of acute cardiopulmonary diseases, Biometrical journal, 2018-01, vol.60 (1), 174-195.
[23] J. Amaral, A. Lopes, J. Jansen, A. Faria, P. Melo, Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease, Computer methods and programs in biomedicine, 2011, vol.105 (3), 183-193.
[24] Y. Cheng, Mean Shift, Mode Seeking, and Clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, August, vol. 17, issue 8, doi: 10.1109 / 34.400568.
[25] I.A. Pestunov, V.B. Berikov, Yu.N. Sinyavsky, Segmentation of multispectral images based on an ensemble of nonparametric clustering algorithms, Vestn. Sib. State aerospace un-that them. Academician M.F. Reshetnev. 2010, no. 5 (31), 56-64.
[26] I.A. Pestunov, Yu.N. Sinyavsky, Analysis and synthesis of signals and images, a nonparametric algorithm for clustering remote sensing data based on the grid approach, Avtometriya, 2006, vol. 42, no. 2, 90-99.
[27] N. Friedman, D. Koller, Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian Networks, Machine Learning, 2003, vol. 50, 95–125.
[28] A. A. Darwiche, Differential approach to inference in Bayesian networks, In Uncertainty in Artificial Intelligence: Proceedings of the Sixteenth Conference (UAI 2000), pp. 123–132. San Francisco, CA: Morgan Kaufmann Publishers, 2000.
[29] A.N. Terentyev, P.I. Bidyuk, The method of probabilistic inference in Bayesian networks according to training data, Cybernetics and system analysis, 2007, no. 3, 93—99 (in Rusian).
[30] M. Voronenko et al., Dynamic Bayesian Networks Application for Economy Competitiveness Situational Modelling. In: Advances in Intelligent Systems and Computing (2020), CSIT 2020, vol. 1293, 210-224.
[31] A.S. Cofino, R. Cano, C.Sordo, J.M. Gutierrez, Bayesian Networks For Probabilistic Weather Prediction, In Proceedings Of The 15th European Conference On Artificial Intelligence, Ios Press, 2002, 695-700.