PhUSE 2011

Paper PP09

Learning from critical values - adverse event identification and classification

Milena Balcerzak, Quanticate, Warsaw, Poland
Edyta Winciorek, Quanticate, Warsaw, Poland

ABSTRACT

Patient safety plays a crucial role in the health care industry. Information about serious adverse events comes from multiple assessments and randomized clinical trials are not optimal for detecting such rare and unexpected events. Meta-analysis of heterogeneous trial data is quite complex and the medical errors made at this stage can result in disability, decreased quality of life, or even death. Such potential negative outcomes emphasize the need for a medical decision support system, which could detect, classify and even predict adverse events. The semantic and neural-networks methods are becoming standard in most other professional industries. Why not have these artificial intelligence learning methods in place in the clinical laboratory? In one particular study, such methods were used to pilot an early detection system of unexpected patterns of occurrences of laboratory values. In this paper, the artificial intelligent algorithms based on fuzzy set theory and semantic neural networks (to fuse bio-signals and to identify adverse event) was applied.

INTRODUCTION

Several different symptoms could precede the occurrence of an adverse event. Early identification of such signals plays a crucial role for patient safety. Laboratory test results, vital signs information, ECG or EEG results are widely used to investigate potential disorders. However, doctors and nurses cannot monitor all the patients around the clock. In this paper, fuzzy inference system (FIS) is presented and applied to adverse events occurrence identification based on two bio-signals: systolic blood pressure and heart rate.

RELATED WORKS

To solve even the simplest real world problem, system modeling, based on conventional and traditional mathematical tools, is very often not a well-suited instrument. This limitation was a significant inducement for the intelligent algorithms development.

The milestone extension for classical set theory constituted fuzzy sets idea, introduced to literature by L. A. Zadeh [1] in a paper from 1965. One of the most popular thoughts based on this framework is the fuzzy inference systems concept.

A numerical analysis approach of such systems was, for the first time, presented by Takagi and Sugeno [2] in 1985. Two main methods in fuzzy modeling problem could be distinguished for model construction and parameters selection. The first one focuses on the ability of fuzzy logic to model natural language [3] and build rules on expert knowledge. This is called fuzzy expert systems or fuzzy controllers. The second approach is based on automatic learning from data and is known as neutral networks. In a paper from 1994, Hayashi [4] showed that any fuzzy rule based system could be approximated by feed forward neural network and vice versa [5].

Fuzzy modeling and fuzzy identification has found numerous practical applications in process simulation or control [6], prediction and inference [7]. Moreover, this concept was also used in traffic and transportation management [8] and in quantitative analysis of people's thought [9].

The semantic and neural-networks methods is becoming standard in most professional industries, however it is not so often used in clinical data. In the health care industry, examples of this approach are a perception-based medical decision support system for the diagnosis of heart diseases [10] or system for classification of electroencephalogram (EEG) signals [11].

FUZZY INFERENCE SYSTEM

In medicine, a fuzzy inference system (FIS) was introduced to formulate the mapping from a number of bio-signal inputs to a decision [12]. Such a system can automatically interpret the incoming signal and prompt about the emergent situation. The definition of categorical labels for bio-signals such as "high blood pressure" varies from patient to patient and can be affected by a doctor's opinion or diagnosis, therefore, imprecision and approximation is an intrinsic part of reasoning about the occurrence of an adverse event. By applying fuzzy logic into inference systems such imprecision can be taken into account [13].

In classical set theory an element either belongs or does not belong to the given set. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements to the given set; this is described with the aid of a ***membership function*** valued in the real unit interval [0, 1]. There is a various number of membership functions where the most popular ones are L-type, Pi-type, Triangular–type and Gamma-type presented in Figure 1 from left to right respectively.

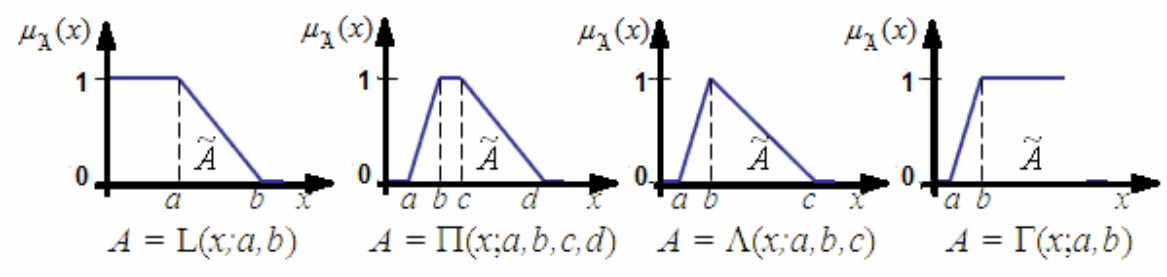

Figure 1. Types of Membership Functions.

Generally, the membership function shapes are described based on the knowledge of domain expert. However, taking into account number of bio-signals it can be difficult process. Therefore in practice very often the shape of membership functions alongside other parameters used in FIS are learned from data by means of neural networks [14].

Assuming that all bio-signals from patients are collected in a database, the fusion of this data can be used in the learning process in order to determine categorical labels for all bio-signals and assign them suitable membership functions. In the scope of FIS development this process is named ***fuzzification*** and it requires so called ***linguistic variables***.

For example, Figure 2 illustrates systolic blood pressure bio-signal defined as linguistic variable with 3 categorical labels (terms) {LOW, NORMAL, HIGH}. It is worth mentioning, that applying such representation, a fuzzy inference engine is able to process meanings - not only the numbers. This scenario can be applied to model human reasoning.

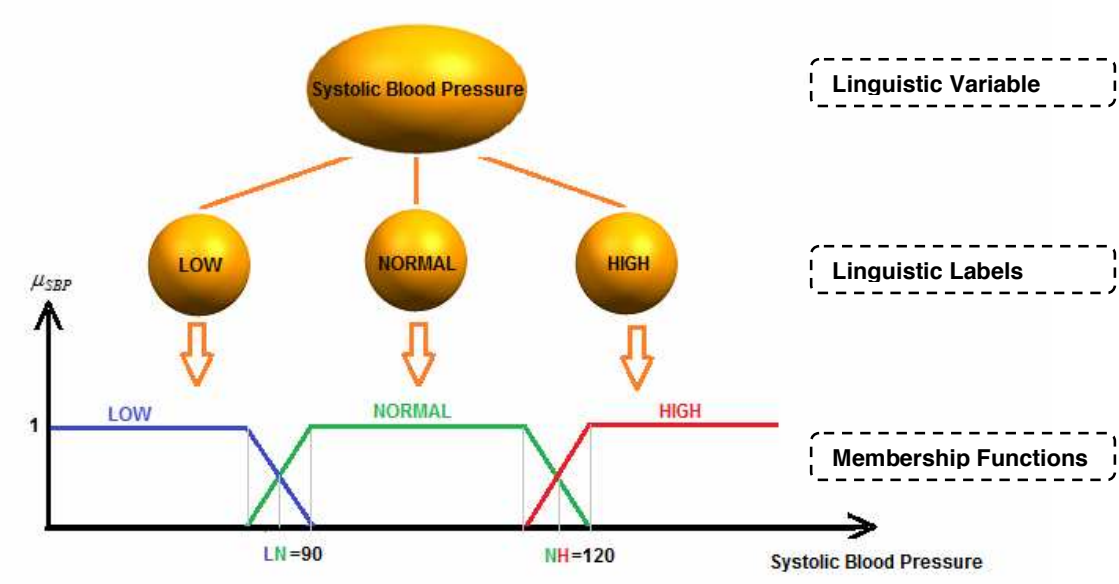

Figure 2. Linguistic Variable for Systolic Blood Pressure Bio-Signal.

The ***fuzzification*** process is the first step to the fuzzy inference based on the **rule base**. The rule base is a repository of all the knowledge used by FIS. Assuming that rules are predefined, the output strengths can be evaluated by combining degrees of input memberships. Therefore, before the rules can be inferred, it is necessary to fuzzify the inputs. It means that all bio-signals must be transformed into the fuzzy domain in the form of a ***linguistic variable***. The IF-THEN rules are then triggered and stored in the rule base. Fuzzy rules may be expressed in terms such as: *IF systolic blood pressure is normal and heart rate is normal THEN no risk of adverse event*, where the antecedents *blood pressure* and *heart rate* as well as the consequent *risk of adverse event* are imprecisely (fuzzily) defined quantities with fuzzy terms {“high”, “normal” , “low”} and {“high risk”, “medium risk”, “low risk”, “no risk”} respectively.

Important questions arise how to determine the optimal set of rules in terms of linguistic labels. Extraction of logical rules may be done manually in assistance of domain expert, but in a real situation, it is quite a difficult process and very often is supported by statistical, pattern recognition and machine learning methods, as well as neural network methods [15]. Neural networks are used for initial determination of linguistic variables and rule extraction, followed by minimization procedures for optimization of the sets of rules. The sample rules for the FIS system dedicated for adverse event identification are shown in Figure 3.

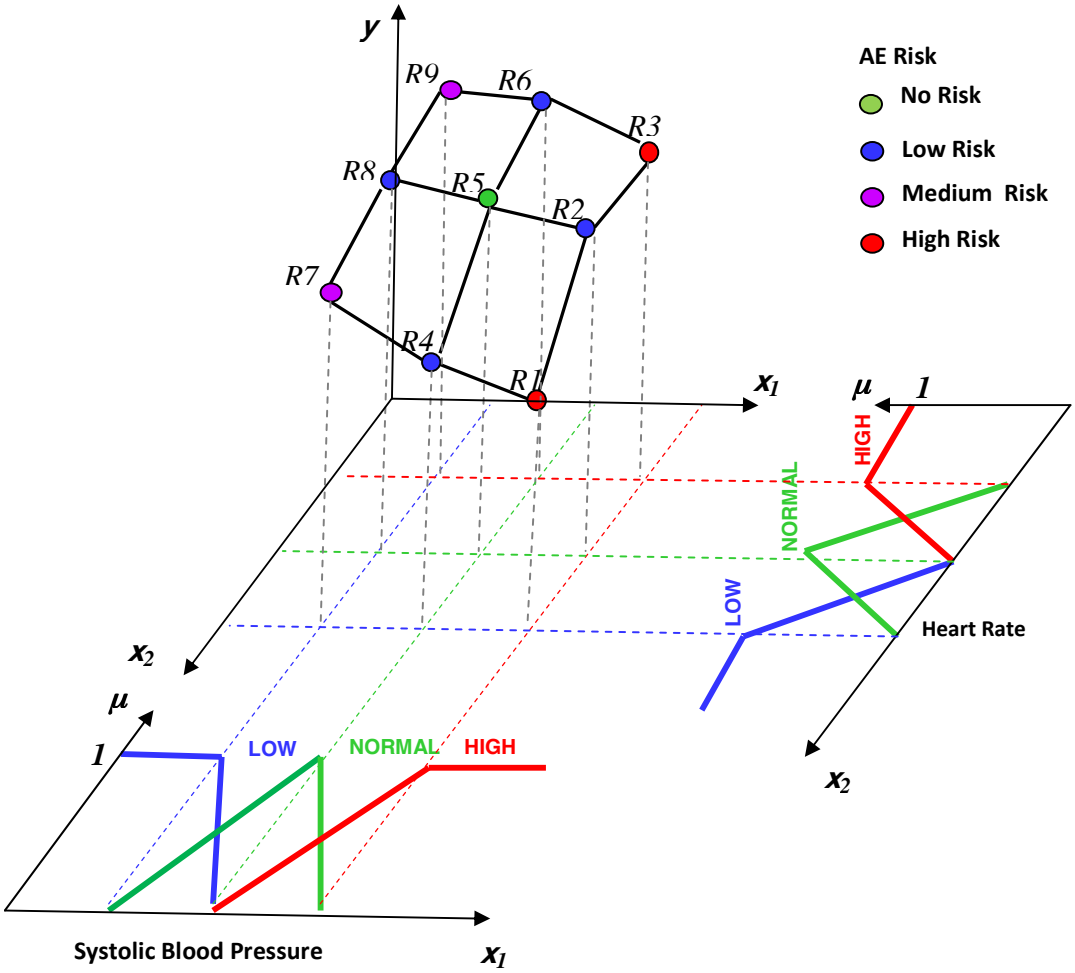

Figure 3. Sample Rules for the FIS System Dedicated for Adverse Event Identification.

After the inputs are fuzzyfied the inference algorithm calculates the degree to which each part of the antecedent is satisfied for each rule. It means that every rule has calculated a weight (a real number between 0 and 1), which is applied to the number given by the antecedent. If the antecedent of a given IF-THEN rule has more than one part, the operator AND/OR is applied to obtain one number that represents the result/strength for that rule.
In fuzzy logic the AND and OR operators are usually defined by means of the minimum and maximum function, respectively, so for the fuzzy variables x and y it can be denoted as:

x AND y = minimum(μ(x), μ (y))
x OR y = maximum(μ (x), μ (y))

where μ is the membership function of fuzzy variables.
There are also other operators, more linguistic in nature, called hedges that can be applied. These are generally adverbs such as "very", or "somewhat", which modify the meaning of a set using a mathematical formula.
The weighting assigned to each rule indicates the ***implication method***. The input for the implication process is a single number given by aggregated antecedents while a consequent is a fuzzy set represented by a membership function, which weights appropriately the linguistic characteristics that are attributed to it. Figure 4 illustrates in details the process of fuzzy implication for the above introduced rule *IF systolic blood pressure is normal and heart rate is normal THEN no risk of adverse event.*

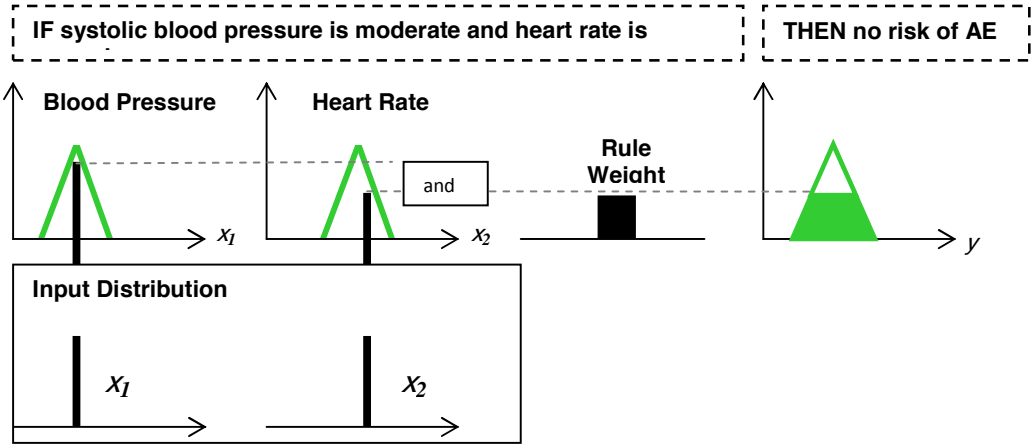

Figure 4. Process of Fuzzy Implication.

The final decision about the adverse event risk is based on the testing of all rules in the FIS. It means that all rules must be combined in some manner in order to make a final decision. **Aggregation** is the process by which the outputs of each rule are combined into a single fuzzy set and occurs only once for output variable.
As much as fuzziness helps the rule evaluation during the intermediate steps, the final desired output is generally a single number. The main objective of this process, so called ***defuzzification*** is to determine an adverse event risk based on the number of bio-signals (Figure 4). There is a fuzzy input but the end product is expected to be a crisp (discrete) value indicating the final decision, e.g. whether the given set of bio-signals indicates an adverse event or not.
There is no unique operation for defuzzification and different approaches are possible. The most popular one in literature is the centre of singleton method (COS) and centre of gravity method (COG).
In the COS method, the defuzzification can be strongly simplified if the membership functions of the conclusions are reduced to singletons. The centre of singletons is then calculated by using the average of singletons weighed by rules' strength.
For more complicated functions of the conclusion, the COG method is preferred. This approach is based on the idea of selecting a value that, on average, would lead to the smallest error in the sense of a criterion, e.g. the least squares means method can be used as COG.

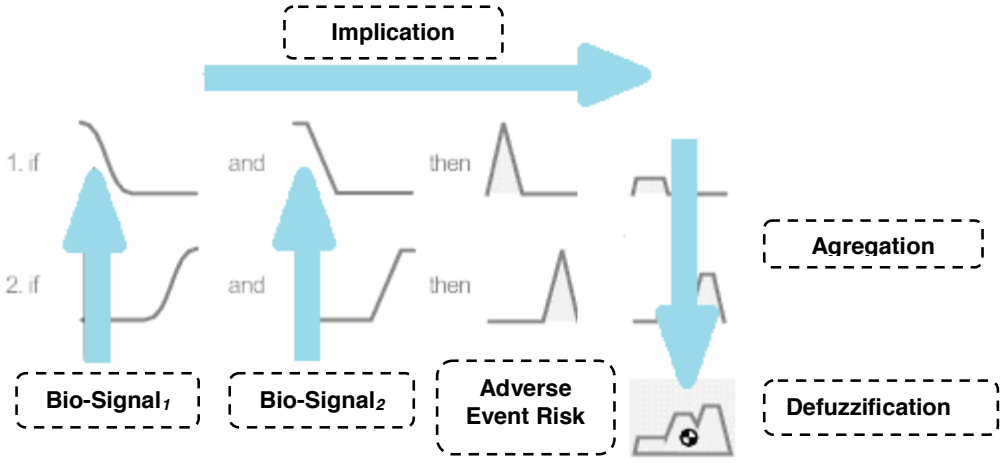

Figure 5. Fuzzy Inference Diagram.

The above presented structure of FIS (Figure 5) can be applied in order to analyze the bio-signals in context of adverse event risk classification. It is worth emphasizing that all of the above mentioned steps in FIS i.e. fuzzification, inference algorithm and defuzzification can be automatically implemented using the neural network algorithms.
In the next section, the algorithm of the FIS representation in terms of neural networks architecture is presented in more details.

FUZZY INFERENCE ALGORITHM

Let $\mathbf{F} = \{\mathbf{U}, \mathbf{A}\}$ be a FIS system where $\mathbf{A} = \{v^1, \ldots, v^m\}$ represents various of bio-signals such as blood pressure, pulse, age and so on, while $\mathbf{U} = \{P_1, P_2, \ldots, P_n\}$ denotes the set of patient's bio-signals. In terms of above-mentioned definition, P_i can be identified as a set of features for various bio-signals specific for the given patient, while set $\mathbf{A}$ represents the generic names for linguistic variables that can be identified in FIS. Initially, the granulation of the linguistic variable is set up to correspond with fuzzy semantic presentation such as {LOW, NORMAL, HIGH}. It means that initially we assume $k = 3$ categorical labels for each linguistic variable; however this can be changed during learning process of the neural network. It should be noticed that the categorical labels of linguistic variables are designed into network layers and all patients' bio-signals are analyzed in each layer through this granulation.

Let f_{jk} where $j = 1, \ldots, m$ represents the membership function for given categorical label of k linguistic variables of bio-signal $j = 1, \ldots, m$. In neural network architecture, the FIS is divided into five layers, namely: *input linguistic layer, input categorical term layer, rule layer, sub-rule layer* and *output layer* (Figure 6). The definitions of each layer are as follows:

Layer 1 ***(Input linguistic layer):*** The input of this layer is the set of patient's bio-signals $\{P_1, P_2, \ldots, P_n\}$. The output is the vector of patients' bio-signals organized as follows: $((P_{11}, P_{12}, \ldots, P_{1m}), \ldots, (P_{n1}, P_{n2}, \ldots, P_{nm}))$. The output of layer 1 comes into the next layer as an input.

Layer 2 (*Input categorical term layer):* In this layer for each patient's bio-signals P_{ij} the values of membership functions f_{jk} are calculated. Namely, the output of this layer is calculated as $\mu^2_{ijk}(P_{ij}) = f_{jk}(P_{ij})$ where $\mathrm{i} = 1, \ldots, \mathrm{n}$ and $\mathrm{j} = 1, \ldots \mathrm{m}$ and $k = \{1,2,3\}$ denotes the number of categorical labels of the linguistic variable assigned to bio-signal *j*-th.

Layer 3 (*Rule layer)*: The input of this layer is the vector of membership functions calculated for each patient's bio-signals $((f_{11}(P_{11}), \ldots, f_{1k}(P_{11}))$,…, $(f_{m1}(P_{nm}), \ldots, f_{mk}(P_{nm}))$). This is the step where the inference algorithm is initiated and aggregation using the AND operator is performed for each rule in FIS. The output for this layer is calculated as follows: $\mu^3_{ik} = min\{f_{1k}(P_{i1}), f_{2k}(P_{i2}), \ldots, f_{mk}(P_{im})\}$. The process of the aggregation is being continued in the next layer, however, the OR operator is used for relevant rules.

Layer 4 ***(Sub-rule layer):*** In this layer the aggregation process is being continued. As the input, the weights μ^3_{ik} calculated in the previous layer are used. The output is obtained as follows: $\mu^4_{ik} = max\{R_k(\mu^3_{11}), \ldots, R_k(\mu^3_{ik})\}$, where $R_k(\mu^3_{ik})$ is the value of k-th semantic rule of the μ^3_{ik}.

Layer 5 (Output layer): The output membership function is obtained by employing all the values from the layer 4.

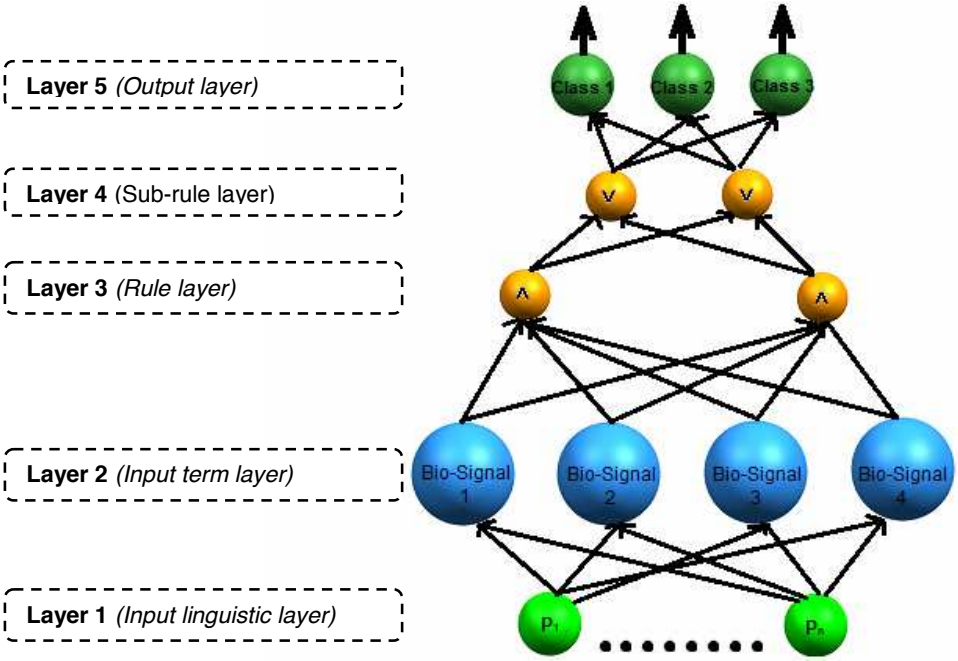

Figure 6. Granular Computing in Fuzzy Inference Algorithm.

EXPERIMENTAL EVALUATION

The main objective of this experiment is to show how using FIS to verify whether the given patient is at risk of occurrence of adverse event or not.

RULES MATRIX AND CLASSES DEFINITION

Potential heart disorders are: an irregular heartbeat (called arrhythmia) and abnormal heart rates. According to expert knowledge, **a normal heart rate ranges from 50 to 100 beats per minute**. Arrhythmias can occur with a normal heart rate, with slow heart rates (less than 50 beats per minute) or with rapid heart rates (greater than 100 beats per minute). It was established that there is a higher probability that arrhythmia takes place together with abnormal heart rate [16].

According to expert knowledge, a **normal blood pressure is below 120 over 80 (120/80)**. However, from fuzzy sets theory it is worth mentioning that most adults in the UK have blood pressure readings in the range from 140 over 90 (140/90). High blood pressure is a serious condition that can lead to different serious diseases. Unlike this state, low blood pressure can occur naturally in healthy, athletics persons. Many domain experts use the arbitrary cutoff levels: **a systolic (top) pressure of 90 or lower or a diastolic (bottom) pressure of 60 or less to define low blood pressure** [18].

Previously, diastolic blood pressure was thought to be a more important risk factor than systolic elevations, however according to specialists: *"Systolic blood pressure should become the principal clinical endpoint for detection, evaluation and treatment of hypertension, especially in middle-aged and older Americans"* [18].

In the example below, only two bio-signals were taken into account as potential signals of adverse event: systolic blood pressure and heart rate. The diastolic blood pressure results are not taken into account as input information. This is in accordance with the domain expert knowledge and it will simplify further considerations.

Based on the presented experts' knowledge and according to fuzzy semantic presentation, systolic blood pressure and heart rate could be divided into the following linguistic levels: **LOW**, **NORMAL** and **HIGH**, however, bounds between those categories are flexible.

With the use of expert knowledge, the rule matrix is created by dividing surface axes, determined by heart rate and systolic blood pressure, into nine sub-areas described in Figure 7 as Rule1 – Rule 9.

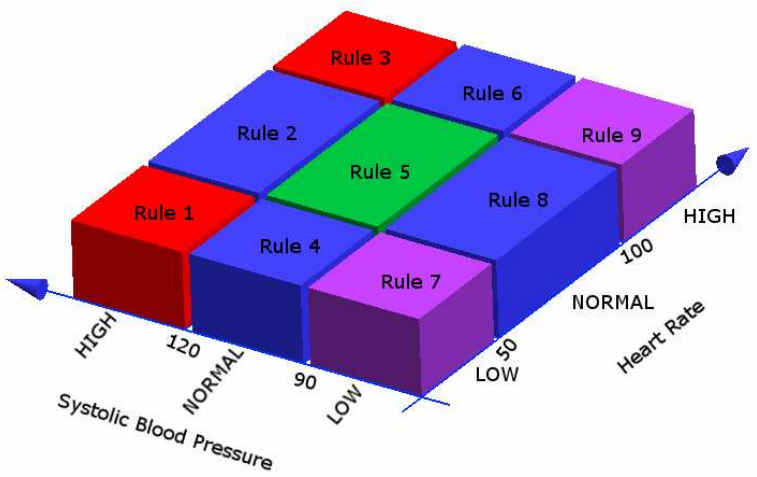

Figure 7. Defined Rules and Classes in Heart Rate versus Systolic Blood Pressure Diagram.

Furthermore, according to prior knowledge, the following compositions of linguistic (IF-THEN) rules can be described:

- **NO RISK** ← Rule 5
  - If a patient has normal systolic blood pressure and normal heart rate then there is no risk of adverse event.
- **LOW RISK** ← Rule 2 or Rule 4 or Rule 6 or Rule 8
  - If a patient has: (high systolic blood pressure and normal heart rate) or (normal systolic blood pressure and low heart rate) or (normal systolic blood pressure and high heart rate) or (low systolic blood pressure and normal heart rate) then there is a low risk of adverse event.
- **MEDIUM RISK** ← Rule 7 or Rule 9
  - If a patient has: (low systolic blood pressure and low heart rate) or (low systolic blood pressure and high heart rate) then there is medium risk of adverse event.
- **HIGH RISK** ← Rule 1 or Rule 3
  - If a patient has: (high systolic blood pressure and low heart rate) or (high systolic blood pressure and high heart rate) then there is high risk of adverse event.

MEMBERSHIP FUNCTION

The above defined linguistic rules are built in into the membership function, being a graphical representation of the magnitude of participation of each input and is illustrated in Figure 8:

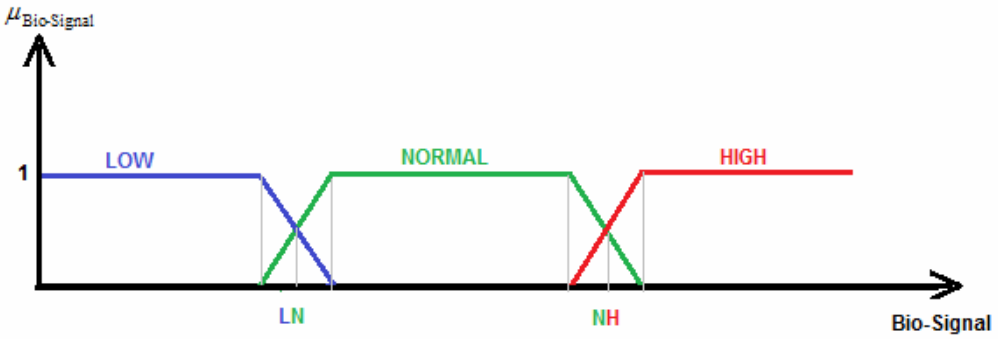

Figure 8. The Membership Function for All the Levels of Given Bio-Signal.

In this example it is assumed that membership functions are identical for all considered linguistic labels of the given bio-signal. Furthermore, it is assumed that maximum of cross-cut of adjacent levels (LN, NH) is determined according to expert knowledge. Thus for systolic blood pressure: LN is equal to 90, NH is equal to 120 and for heart rate: LN is equal to 50, NH is equal to 100.

INFERENCE ALGORITHM APPLICATION

One point (presented in Figure 9) is taken into account in further considerations: (heart rate; systolic blood pressure) = (116;78).

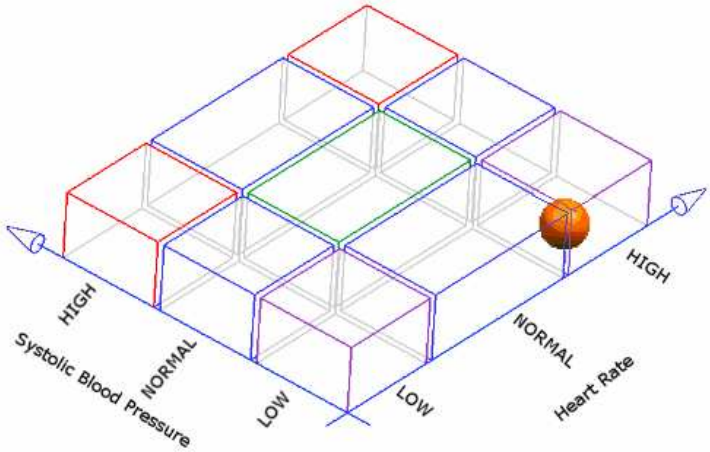

Figure 9. Heart Rate versus Systolic Blood Pressure Diagram.

Layer 1 & Layer 2:

Membership function defined for a selected point (input) determines its strength [0 to 1] for linguistic levels (labels) for both bio-signals as presented in Table 1.

	Heart Rate P_{11}	**Systolic Blood Pressure** P_{12}
Input Values:	116	78
Level	**Membership Function Values** f_{jk}	
LOW	0	0,8
NORMAL	0,1	0,2
HIGH	0,9	0

Table 1. The Membership Function Values.

Determination of membership function values is illustrated in Figure 10 and Figure 11.

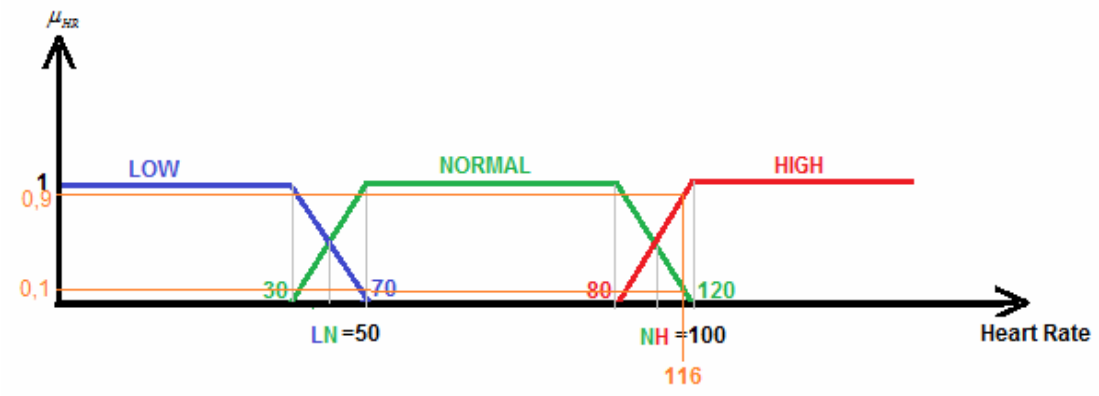

Figure 10. The Membership Function for Pointed Value of Heart Rate.

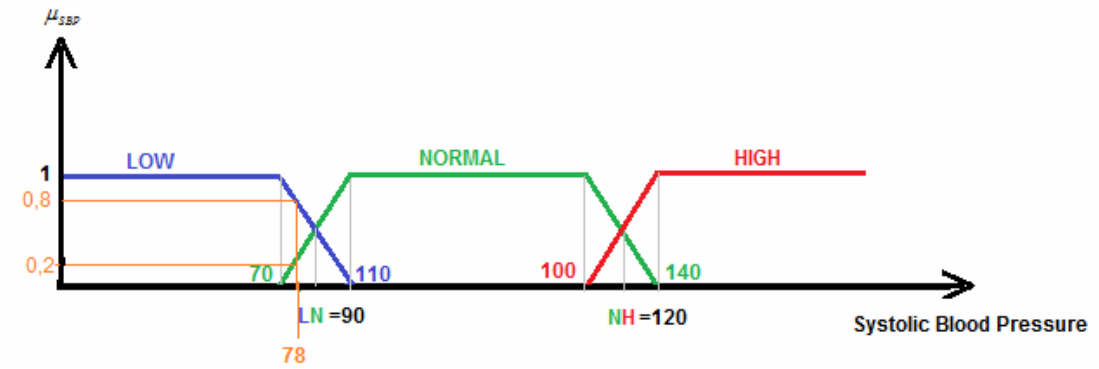

Figure 11. The Membership Function for Pointed Value of Systolic Blood Pressure.

Referring back to the linguistic rules and plugging in the membership function values from Table 1, the input term for min-max inference process is obtained as presented in Table 2.

Rule	Heart Rate	Systolic Blood Pressure	Heart Rate	Systolic Blood Pressure
	Level		Membership Function Values $\mu_{ijk}^{2}(P_{ij}) = f_{jk}(P_{ij})$	
Rule 5	*NORMAL*	*NORMAL*	0,1	0,2
Rule 2	*NORMAL*	*HIGH*	0,1	0
Rule 4	*LOW*	*NORMAL*	0	0,2
Rule 6	*HIGH*	*NORMAL*	0,9	0,2
Rule 8	*NORMAL*	*LOW*	0,1	0,8
Rule 7	*LOW*	*LOW*	0	0,8
Rule 9	*HIGH*	*LOW*	0,9	0,8
Rule 1	*LOW*	*HIGH*	0	0
Rule 3	*HIGH*	*HIGH*	0,9	0

Table 2. Cartesian Product of the Membership Function Values.

Layer 3 & Layer 4:
The logical product (Table 2) for each rule must be combined or inferred. In this example, min-max inference process was used. This method tests the magnitudes of each rule and selects the highest one.
Of the nine rules determined, only four have non-zero results (Table 3). This leave fuzzy output response magnitudes only for NO RISK, LOW RISK, and MEDIUM RISK classes.

Rule	Minimization $\mu_{ik}^{3} = \min\{f_{1k}(P_{i1}),....,f_{mk}(P_{im})\}$	Class	Maximization $\mu_{1k}^{4} = \max\{R_k(\mu_{11}^{3}),....,R_k(\mu_{ik}^{3})\}$
Rule 5	0,1	NO RISK	0,1
Rule 2	0	LOW RISK	0,2
Rule 4	0		
Rule 6	0,2		
Rule 8	0,1		
Rule 7	0	MEDIUM RISK	0,8
Rule 9	0,8		
Rule 1	0	HIGH RISK	0
Rule 3	0		

Table 3. Min-max Inference Process Results.

Layer 5:
As it is presented in Table 3 the following weights were obtained: 0,1; 0,2; 0,8 and 0 for NO RISK, LOW RISK, MEDIUM RISK and HIGH RISK classes respectively.
The last step of the presented algorithm is the defuzzification of the data into a crisp output. For this purpose, centre of singleton method (COS) is applied.
Boundary values of degree were defined as follows: 0 for NO RISK and 1 for HIGH RISK. Assuming that the degree is equally divided into three parts, then the range for a healthy person is between 0 and 0,33. The range for a patient at risk of an adverse event occurrence is from 0,33 to 1, with critical values between 0,66 to 1.

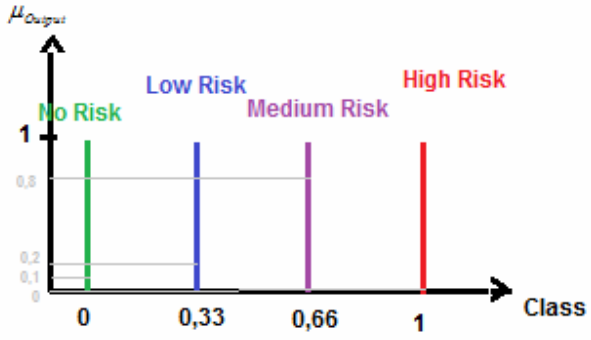

Figure 12. Defuzzification of the Data Into a Crisp Output.

$$output = \frac{0,1 \cdot 0 + 0,2 \cdot 0,33 + 0,8 \cdot 0,66 + 0 \cdot 1}{0,1 + 0,2 + 0,8 + 0} = \frac{0,066 + 0,528}{1,1} = \frac{0,594}{1,1} = 0,54$$

Referring to Figure 12 the crisp result equal to 0,54 was obtained thus it can be inferred that the considered patient is at risk of occurrence of an adverse event.

CONCLUSION

In the real world, people may not always clearly interpret changes in their health. In this paper, application of fuzzy inference system (FIS) in the adverse event occurrence identification process was presented to more objective identification of 'weaker' patients.

The main advantage of such a system is that it can automatically interpret the incoming bio-signals and prompt the doctor about the emergent situation through the system. This can be important especially when a doctor or nurse cannot assist the patient all the time - patient's bio-signals are immediately interpreted by the system.

REFERENCES

[1] Zedeh L. A., *Fuzzy Sets*, Information and Control 8, 1965, p. 338-353.
[2] Takagi T., Sugeno M., Structure *identification of systems and its application to modeling and control*, IEEE Trans. Systems Man Cybern. 15, 1985, p. 116–132.
[3] Mamdani E. H., Assilian S., *An experiment in linguistic synthesis with a fuzzy logic controller*, Int. J.Man-Mach. Stud. vol.7, 1975, p. 1–13.
[4] Hayashi Y., Buckley J.J., *Approximations Between Fuzzy Expert Systems and Neural Networks*, International Journal of Approximate Reasoning vol. 10, 1994, p. 63–73.
[5] Li X.H., Chen C.L.P., *The Equivalence Between Fuzzy Logic Systems and Feed forward Neural Networks*, IEEE Transactions on Neural Networks Vol 11 No. 2, 2000, p. 356–365.
[6] Pedrycz W., *Fuzzy Control and Fuzzy Systems*, New York: Willey, 1989; Sugeno M., Ed., *Industrial Application of Fuzzy Control*, New York: Elsevier, 1985.
[7] Kandel A., *Fuzzy Expert Systems*. Reading, MA: Addison-Wesley, 1988; Kandel A., *Fuzzy Expert Systems*, Baca Raton, FL: CRC Press, 1992.
[8] Wang F. Y., *Agent-based Control for Networked Traffic Management System*, IEEE Intelligent System Vol. 20 No. 5, 2005, p 92-96.
[9] Gerla G., *Effectiveness and Multivalued Logics,* Journal of Symbolic Logic vol 71, 2006, p.137-162.
[10] Yan H., Jiang Y., Zheng J., Peng C., Li Q., *A multiplayer Perceptron-based Medical Decision Support System for Heart Disease Diagnosis*, Expert Systems and Applications Vol. 30 No. 3, 2006, p.272-281.
[11] Gulera I., Ubeyli E. D., *Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients*, Journal of Neuroscience Methods, 2005.
[12] Chi-lu Yang, Chin-Huan Hsu, A *Fuzzy Inference Algorithm for Personal Bio-Informatics in Home-care Systems*, in Proceedings of 2008 International Computer Symposium (ICS 2008), Taipei, Taiwan, Nov. 2008, pp. 63-68.
[13] W. Duch, R. Adamczak, K. Grąbczewski, G. Zal, Y. Hayashi, *Fuzzy and crisp logical rule extraction methods in application to medical data*, Com- putational Intelligence and Applications. Springer Studies in Fuzziness and Soft Computing, Vol. 23, 1999.
[14] W. Duch, R. Adamczak, K. Grąbczewski, A new methodology of extraction, optimization and application of crisp and fuzzy logical rules, Neural Networks, IEEE, 2001, pp. 277 – 306.
[15] W. Duch, R. Adamczak, K. Grąbczewski, G. Żal, *Hybrid neural-global minimization method of logical rule extraction*, Int. Journal of Advanced Computational Intelligence (in print)
[16] http://www.medicinenet.com/arrhythmia_irregular_heartbeat/article.htm, available online 26 August 2011.
[17] http://www.bpassoc.org.uk/BloodPressureandyou/Thebasics/Whatisnormal, available online 26 August 2011.; http://www.medicinenet.com/low_blood_pressure/article.htm, available online 26 August 2011.
[18] http://www.sciencedaily.com/releases/2000/05/000508082111.htm, available online 26 August 2011.

CONTACT INFORMATION

Milena Balcerzak
Quanticate Polska Sp. z o.o
Hankiewicza 2
02 - 103 Warsaw, Poland
Tel. +48225762119
Fax: +48225762159
Email: milena.balcerzak@quanticate.com
Web: www.quanticate.com

Edyta Winciorek
Quanticate Polska Sp. z o.o
Hankiewicza 2
02 - 103 Warsaw, Poland
Tel. +48225762150
Fax: +48225762159
Email: edyta.winciorek@quanticate.com
Web: www.quanticate.com