The Artificial Intelligence and Business Decisions

Abstract

The aim of the use of artificial intelligence in business is in the field of optimization - minimization the costs and maximization the profit.

The mostly used methods of artificial intelligence are fuzzy logic, artificial neural networks and evolutionary algorithms. The artificial intelligence have had successful applications also in business. Nowadays the new theories of artificial intelligence are used for these purposes. The applications in business, economics, and finance have specific features in comparison with others. The processes are focused on private corporate attempts at money making or decreasing expenses; therefore the details of applications, successful or not, are not published very often. The optimization methods help in decentralization of decision-making processes to be standardized, reproduced, and documented. The optimization plays very important roles especially in business because it helps to reduce costs that can lead to higher profits and to success in the competitive fight.

1 Introduction

There are various optimization methods appropriate to use in business and economics: classical ones and methods using soft computing such as fuzzy logic, neural networks, genetic algorithms, and the theory of chaos.

Artificial intelligence differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty, partial truth, and approximation. In effect, the role model for soft computing is the human mind. The guiding principle of artificial intelligence is: Exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost. The basic ideas underlying artificial intelligence in its current incarnation have links to many earlier influences, among them Zadeh's 1965 paper on fuzzy sets. The inclusion of neural computing and genetic computing in artificial intelligence came at a later point.

At this juncture, the principal constituents of Artificial intelligence (AI) are Fuzzy Logic (FL), Neural Computing (NC), Evolutionary Computation (EC) Machine Learning (ML) and Probabilistic Reasoning (PR), with the latter subsuming belief networks, chaos theory and parts of learning theory. What is important to note is that artificial intelligence is not a melange. Rather, it is a partnership in which each of the partners contributes a distinct methodology for addressing problems in its domain. In this perspective, the principal constituent methodologies in SC are complementary rather than competitive. Furthermore, artificial intelligence may be viewed as a foundation component for the emerging field of conceptual intelligence.

The mentioned applications in this chapter are as follows:

- datamining
- prediction
- stock market
- risk management
- decision making

2 Fuzzy logic

In classical logic, a theory defines a set as a collection having certain definite properties. Any element belongs to the set or not according to clear-cut rules; membership in the set has only the two values 0 or 1. Later, the theory of fuzzy logic was created by Zadeh in 1965. Fuzzy logic defines a variable degree to which an element x belongs to the set. The degree of membership in the set is denoted $\mu(x)$; it can take on any value in the range from 1 to 0, where 0 means absolute non-membership and 1 full membership. The use of degrees of membership corresponds better to what happens in the world of our experience. Fuzzy logic measures the certainty or uncertainty of how much the element belongs to the set. People make analogous decisions in the fields of mental and physical behaviour. By means of fuzzy logic, it is possible to find the solution of a given task better than by classical methods.

The fuzzy logic system consists of three fundamental steps: fuzzification, fuzzy inference, and defuzzification. See Fig. 1.

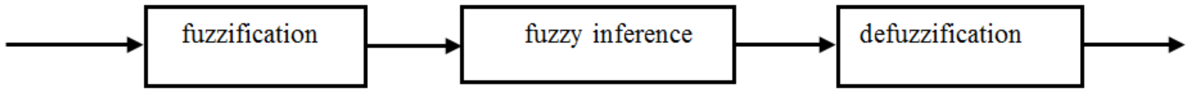

Fig.1 Decision making solved by means of fuzzy logic

The first step (fuzzification) means the transformation of ordinary language into numerical values. For variable risk, for example, the linguistic values can be no, very low, low, medium, high, and very high risk. The variable usually has from three to seven attributes (terms). The degree of membership of attributes is expressed by mathematical functions. There are many shapes of membership functions. For example, for mf_1, P = [0 0 3]; mf_2, P = [2 4 6]; mf_3, P = [4 6 7 9]; mf_4, P = [8 10 10]; and so forth. See Fig. 2.

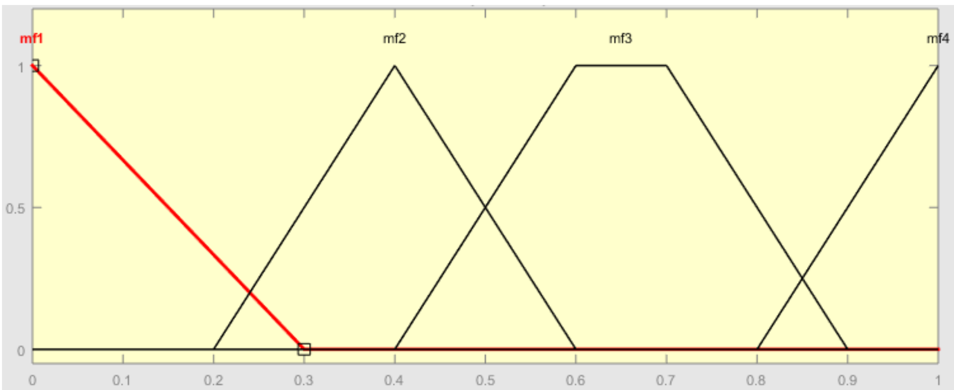

Fig. 2 The types of membership functions Λ and Π

The types of membership functions that are used in practice are for example Λ and Π. There are many other types of standard membership functions on the list including spline ones. The attribute and membership functions concern input and output variables.

The second step (fuzzy inference) defines the system behaviour by means of the rules such as <IF>, <THEN>, <WITH>. The conditional clauses create this rule, which evaluates the input variables. These conditional clauses have the form

<IF> I_1 is mf_a <OR> I_2 is mf_b . . . <OR > I_{N-1} is mf_y <OR >I_N is mf_z <THEN> O_1 is mf_{O1} <WITH> s.

The written conditional clause could be described by words: If the input I_1 is mf_a or I_2 is mf_b or . . . or I_{N-1} is mf_y or I_N is mf_z then O_1 is mf_{O1} with the weight *s*, where the value *s* is in the range <0–1>. These rules must be set up and then they may be used for further processing.

The fuzzy rules represent the expert systems. Each combination of attribute values that inputs into the system and occurs in the condition <IF>, <THEN>, <WITH> represents one rule. Next it is necessary to determine the degree of supports for each rule; it is the weight of the rule in the system. It is possible to change the weight rules during the process of optimization of the system. For the part of rules behind <IF>, it is necessary to find the corresponding attribute behind the part <THEN>. These rules are created by experts. The <AND > could be instead <OR >.

The third step (defuzzification) means the transformation of numerical values to linguistic ones. The linguistic values can be, for example, for variable *Risk* very low, low, medium, high, and very high. The purpose of defuzzification is the transformation of fuzzy values of an output variable so as to present verbally the results of a fuzzy calculation. During the consecutive entry of data the model with fuzzy logic works as an automat. There can be a lot of variables on the input.

3 Neural networks

The history of the development of neural networks started in the first half of the twentieth century. The first publications were by McCulloch. Later Pitts worked on the simplest model of a neuron, and after that Rosenblatt created a functional perception that solves only problems involving areas that are linearly separable. When the multilayer network was discovered by Rumelhart, then Hinton and Williams created back-propagation methods for multi-layer networks. A great boom of neural network applications has been ongoing since the mid-1970s.

The neural network model represents the thinking of the human brains. The model is described as a “black box.” It is simulated by a “black box” that enables us to describe the behaviour of the system by the function that performs transformation of input and output. It is suitable to use neural networks in cases where the influences on searched phenomena are random and deterministic relations are very complicated. In these cases we are not able to separate and analytically identify them. They are suitable for simulation of complicated and often irreversible strategic decision making. The biological neuron can be presented in a simple way that consists of many inputs (dendrites), body (soma), and one output (axon) as shown in Fig. 3. The inputs are processed by neurons. The output information is spread by the axon to terminals that are called “synapses.” The synapsis communicates with the dendrites of other neurons.

The history of the development of neural networks started in the first half of the twentieth century. The first publications were by McCulloch. Later Pitts worked on the simplest model of a neuron, and after that Rosenblatt created a functional perception that solves only problems involving areas that are linearly separable. When the multilayer network was discovered by Rumelhart, then Hinton and Williams created back-propagation methods for multi-layer networks. A great boom of neural network applications has been ongoing since the mid-1970s.

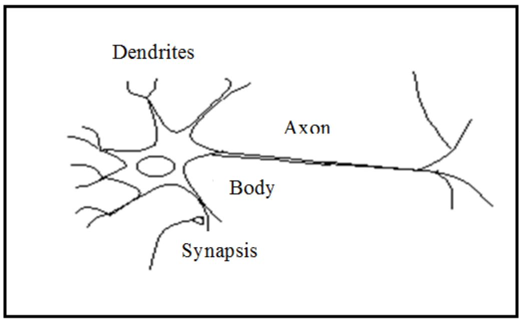

Fig. 3 The biological neuron

The neural network works in two phases. In the first phase the network presents a model of a complicated system as a "curious pupil"; it tries to set up parameters so as to best correspond to the topology of neural networks. In the second phase, the neural network becomes an "expert" to produce the outputs based on the knowledge obtained in the first phase. During the building up of a neural network, the layers of the network must be defined (input, hidden, output); single input and output neurons specified, and the method of connecting the neurons among them identified (the setup of transfer functions among neurons).

The simplest neural network is called a "perceptron." It may have an input of R variables p_1, p_2, p_3, . . . , p_R. These variables are multiplied by weight coefficients w_1, w_2, w_3, . . . , w_R. The threshold value b influences the output; it increases the value of sum just about this value. The formula is

$$a = w_1 * p_1 + w_2 * p_2 + w_3 * p_3 + \cdots + w_R * p_R + b = \sum_{i=1}^{R} w_i p_i + b .$$

Figure 4 shows the single-layer neural network: perceptron.

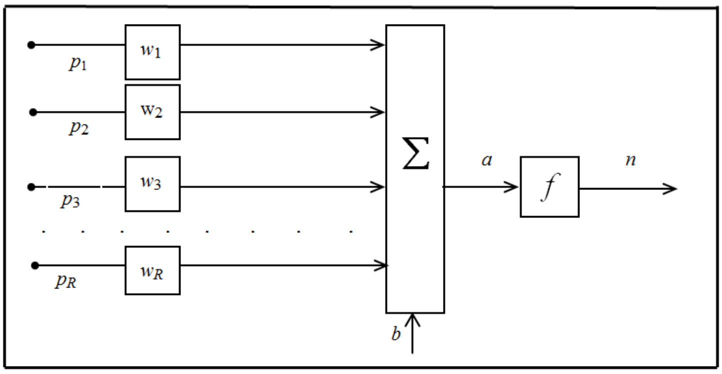

Fig. 4 Perceptron

The most important MATLAB functions are hardlim n= 0 for a<0 and n=1 for a≥0, purelin n=a, logsig n=1/(1+e^{-a}) and tansig n=(e^{a}-e^{-a}) /(e^{a}+e^{-a}). See Fig. 5.

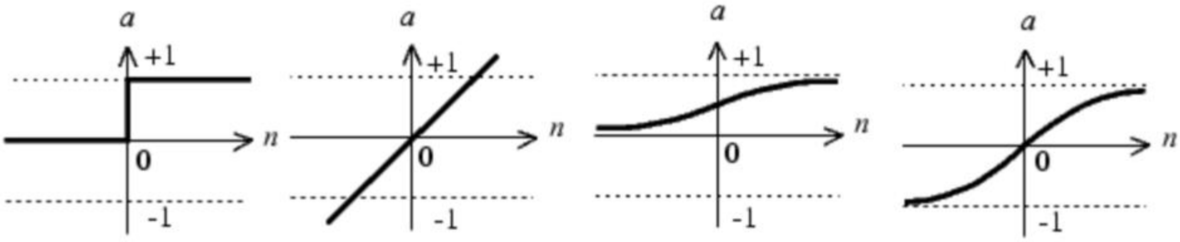

Fig. 5 Transfer functions hardlim, purelin, logsig, and tansig

The simplest transfer function is hardlim, when output value is equal 1 or 0, according to whether the value *a* is less than 0 or equal and bigger than 0. With this function the transformation gives us standardized values. The function logsig has values in the interval from 0 to 1. The function tansig has values in the interval from −1 to 1. The output values could achieve high values in case of not using such or similar transfer functions, which is the problem especially of multi-layer neural networks. The value *b* is the so called threshold value that increases the value of sum (sum of inputs *x* weight coefficient) just about this value.

Fig. 6 presents a multi-layer network with input layer, hidden layers, and output layer. For perception and multi-layer networks it is possible to write the equation in the matrix form $\boldsymbol{n} = \boldsymbol{f}\,(\boldsymbol{w} * \boldsymbol{p} + \boldsymbol{b})$.

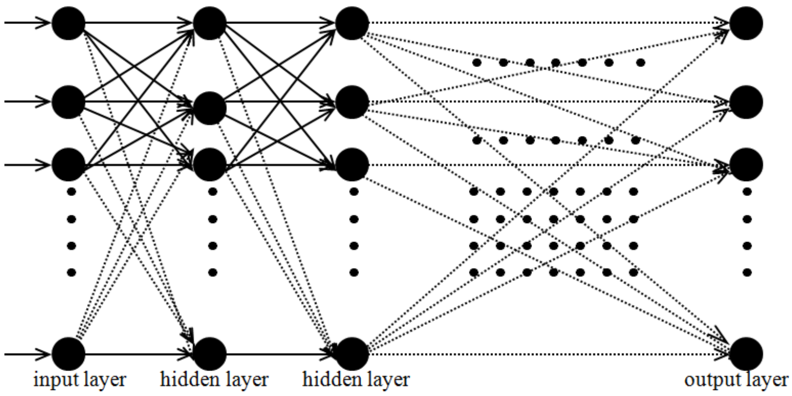

Fig. 6 The diagram of multi-layer network

The back-propagation method is used for calculation of weights of neural networks. It consists of two steps. First, it is necessary to make a calculation of outputs on the basis of inputs and weights (forward step). Next it is necessary to calculate the error *E* as a square difference of calculated output *n* and expected output *o* over all outputs, using an equation in the form $E = \Sigma\ (n_i - o_i\ )^2$. The value *E* is used for the backward calculation of weight (backward step). The process is repeated till the values of *E* converge to an acceptable value (the problem of learning is an optimization task, in which the error function *E* is a fitness function that must be minimized).

4 Evolutionary algorithms

Evolutionary algorithms includes the genetic algorithms. The genetic algorithms are used in studies where exact solution by systematic searching would be extremely slow, which is well suited for solving complicated problems. Genetic processes in nature were discovered in the nineteenth century by Mendel and developed by Darwin. The computer realization of genetic algorithms discovered in the 1970s, is connected with the names of J. Holland and Goldberg. Recently there has been considerable expansion of genetic algorithms in the spheres of economic applications and the decision making of firms and companies.

Let us mention a few terms that are used in the branch of genetics: chromosomes, selection, crossover, mutation, population, parents, and offspring. The chromosomes consist of genes (bits). Every gene inherits one or several bits and its position in chromosomes. We say that the chromosomes have locus. The information coded in chromosomes consists of phenotypes. Most of the implementations of genetic algorithm work with the original representation of chromosomes is binary representation: 0 and 1. A chromosome is represented by a binary string, e.g. 0101. These binary strings mostly represent coded decimal numbers. The operators of selection, crossover, and mutation are most often used in genetic algorithms. The diagram is then chained, where the permitted symbol occurs in at least one position (in case of binary representation it is 0 or 1). For the handling of chromosomes, several genetic operators have been proposed. The most used operators are selection, crossover, and mutation.

The process of selection involves the choice of chromosomes that become parents. The fitness of the parents plays an important role in the process of selection. The process of selection is presented in Table 1 when the number 7 (binary 0111) is bigger than 2 (binary 0010). In this case the chromosome with number 7 (binary 0111) progresses to the next generation as the strongest specimen (it leads in computations to better solutions and thus to higher profit).

Table 1 Selection

Population I.		Population II.
0111	>	0010
7	>	2

The process of crossover involves the exchange of part of two or more parent's chromosomes, which causes the modification of the chromosomes of the offspring. The crossover is presented in Table 2.

Table 2 Crossover

Parents	Offspring
1|001	1|100
1|100	1|001

The crossover can be improveed by selection of more parents. The crossover can have more crossing points. This process is called a multi-point cross when more division points are generated. The generalized crossover is done according to the pattern of zeros and ones generated by alternative divisions.

The process of mutation involves the modification of chromosomes. The modification is done by random change of some bit. The mutation is not frequent. The mutation is presented in Table 3.

Table 3 Mutation

Before	After
1001	0101

The genetic algorithms work in such a way that the population of chromosomes is created at first. Then the population is changed by means of genetic operators until it is found that the parents are the same (the value of fitness function does not change after some number of iterations). The process of reproduction repeats. Each epoch of population (one generation) represents three steps: selection, crossover, and mutation. See Fig. 7.

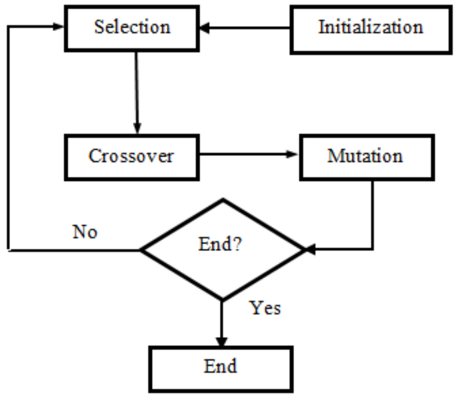

Fig. 7 The reproduction process

Note: A fitness function is a function that prescribes the optimality of a solution and correlates closely with the algorithm's goal (for example maximum profit or minimum costs in business.)

When the genetic algorithm is applied to a problem in the decision making of firms (whose decisions are not reversible), each chromosome codes some solution of the problem (thus a chromosome is a genotype and the corresponding solution is a phenotype). Chromosomes with higher fitness functions in genetic algorithms are preferred. The higher the value of the fitness function during the process of iteration, the better the solution of the economic problem. The commercially sold programs do not demand knowledge of binary algebra.

For processes where optimization is required, it is possible to use not only genetic algorithms but also ant colony, particle swarms, bee hive colony algorithms, and others.

5 Application of artificial intelligence

Data mining. Data mining is a branch that covers a wide range of techniques that are used for obtaining information from great amounts of economic and financial data. It is used in various branches such as insurance companies, direct mailing, retailing, wholesaling, etc. For example, one of the aims of companies is to decrease the risk of losses. For this purpose data mining is used to obtain the data and evaluate them. It is possible to search these bases not only for new customers but also existing customers that could be lost or to reveal risky customers. Data mining is used to solve these problems by means of advanced mathematical methods. Figure 8 presents the realization of cluster analysis with the use of data obtained from the process of data mining. Cluster S_1 presents a group of people with low and medium income whose savings are low and medium, cluster S_2 presents a group of people with low and medium income and high savings, and cluster S_3 presents a group of people with high income and low, medium, and high savings.

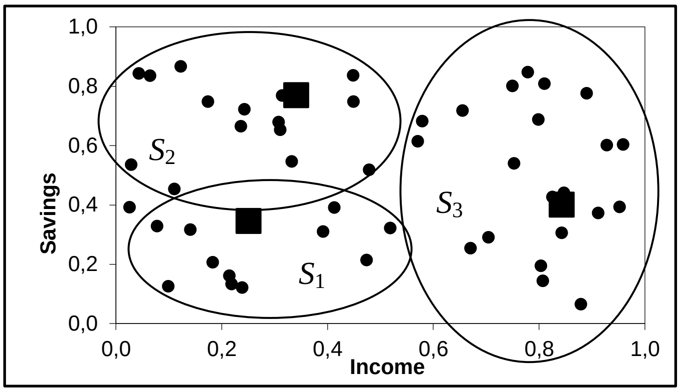

Fig. 8 Data mining

Prediction. It is very important to know the future behaviour of various variables in many branches of human activities. Various methodologies have been developed. Some of them are based on the principle of algorithms, other methods are heuristic, and still others use abilities of learning. During previous years many different and very powerful prediction methods have been developed and used in practice.

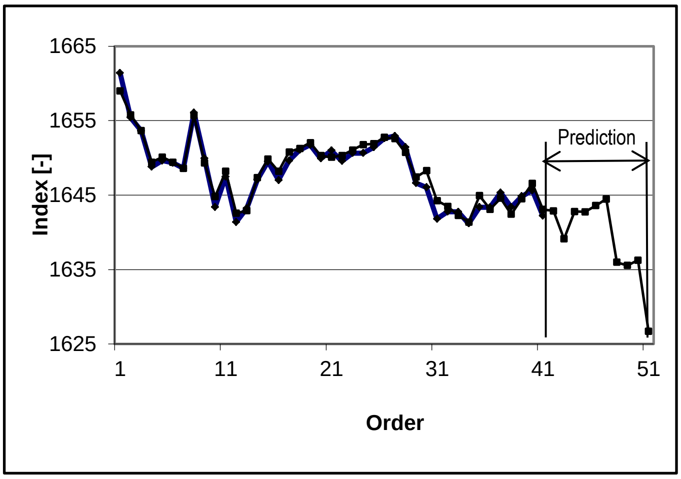

Fig. 9 Prediction

Business and management processes can be very complicated because society creates these phenomena with a considerable degree of chaotic behaviour. In these cases it is suitable to use advanced mathematical methods. Figure 9 presents a time series with past and predicted values. The future development of financial and economic indicators, prices of shares, commodities, values of indexes, currency ratios, etc., will be emphasized in business and management.

Stock market. The prices of shares and commodities, currency ratios, and values of indexes in the stock market are in the form of time series. The time series consist of the close, open, maximum, and minimum prices or indexes, as well as the volume of trade. As the courses of time series are influenced by complex economic and psychological phenomena that contain high degrees of chaos, advanced mathematical methods are required for processing and evaluation of information and data from this area. In this area many analyses and technical, fundamental, and psychological methods have been developed whose aim is to determine sell and buy signals to get a maximum profit. It is also important to set up a portfolio with its percentage composition of single items in the portfolio to be optimum as presented on Fig. 10.

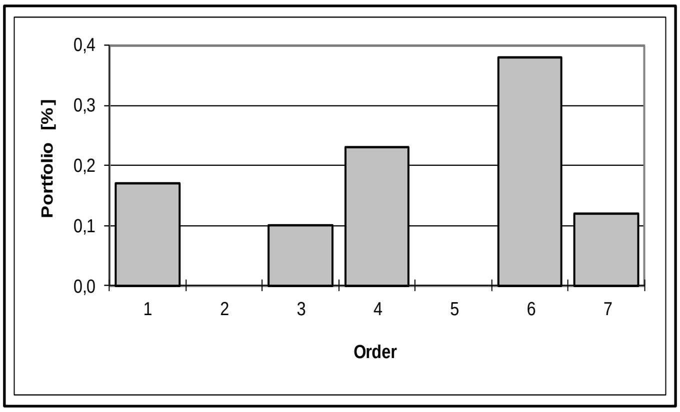

Fig. 10 Stock market

Risk management. Risk management is a branch that covers a wide range of methods used in many endeavours. Estimation of risk must be balanced against the hope of achievement of the best economic results. Control of risk involves minimizing the danger of failures or losses that can affect the stability of the firm or lead to bankruptcy. This branch has become inevitable to maintain competitiveness. This branch extends to many fields such as banking, insurance, direct mailing, etc. The aim of the firm is to achieve the best economic results and to decrease failures and/or losses. Risk arises because of lack of information and insufficient understanding of phenomena, the use of unsuitable and unreliable data, the use of unsuitable methods, or the influence of random processes.

Figure 11 presents a realization of cluster analysis in risk management. The cluster S_1 presents a group of people with low income and low, medium, and high risk of nonpayment of debts; the cluster S_2 presents a group with medium income and low risk of nonpayment; the cluster S_3 presents a group with high income and low risk of nonpayment.

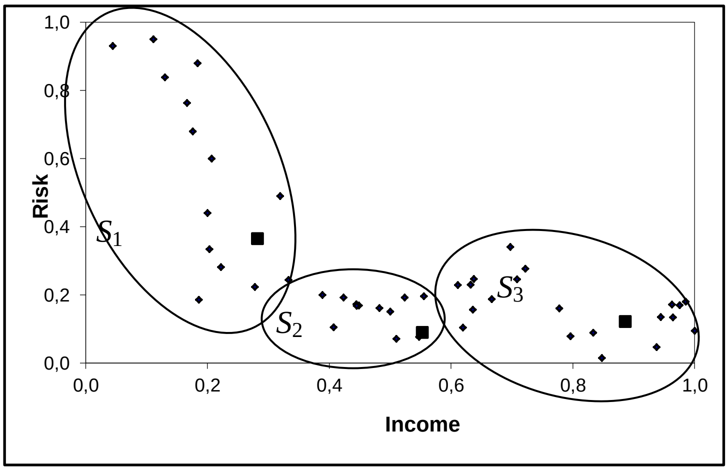

Fig. 11 Risk management

Decision making. To up-to-date access in the process of decision making in firms belongs the procedural access based on the comprehension activities of searched organizations, psychological-social access that comes from interpersonal relations, quantitative access using mathematical methods and means of computer techniques, and a systems approach with complex comprehension of partial processes and empirical access based on practice.

Decision-making methods do not provide homogenous opinions; only through experience can one choose for this purpose the relevant access and thus obtain in a creative way the knowledge from diverse directions of guidance. The elements of advanced methods can be placed among quantitative, complex, and also empirical accesses. Complicated decision-making processes have to be solved by complex methods in which multiple techniques are used. So we can use as a partial model, e.g. the theory of fuzzy logic, neural networks, evolutionary algorithms, and chaos. We can use a combination of these methods. This structure is presented in Fig. 12.

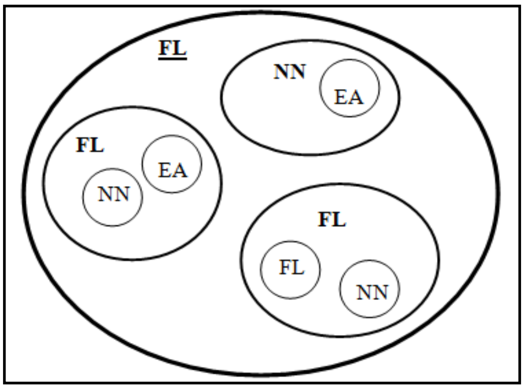

Fig. 12 Decision making

It is necessary to use various tools for decision making (such as programs, consultants, etc.) during the process of decision. The tools are used by the manager who makes decisions. This means that the tools can help the manager to make his/her own decisions, but he/she (the manager) bears the responsibility for his/her decision, not the other people—the designer of the program tools for support of decision making, the consultant, etc. As an illustration of a possible means of decision making with the aid of combination of methods, it is possible to mention speculative operation in world stock markets where the aim is the highest profit. These complicated applications need the use of combined models where, in the process of decision making, the fuzzy logic model is a supervisor of partial models that are created by one or more partial models of fuzzy logic, neural networks, evolutionary algorithms, or other models.

Advanced methods such as fuzzy logic, neural networks, evolutionary algorithms, and the theory of chaos are very useful because business and economic phenomena are very complex and complicated.

CONCLUSION

The advanced decision-making methods play very important roles in companies because they help to reduce costs that can lead to higher profit and they can help to compete successfully, or decrease expenses in institutions. The decision-making processes in business are very complicated because they include political, social, psychological, economic, financial, and other factors. Many variables are difficult to measure; they may be characterized by imprecision, uncertainty, vagueness, semi-truth, approximations, and so forth.

The use of the theories mentioned above is in the sphere of analyses and simulation. Except the applications such as risk investment, risk management, optimization, prediction of time series, journey optimization, description of phenomena it could be many other applications it could be mentioned mortgage loan risk evaluation, direct mailing decision making, stock market decision, stock trading decision, decision making evaluation, client risk evaluation, supplier risk evaluation, etc.

The use of these computing methods can lead to higher quality of analyses and simulations that can be used for decision-making and control processes in business, economy, financial areas, or the public sector.

REFERENCES

Aliev, A., Aliev, R. (2002). *Soft computing and its applications*. World Scientific Publishing.

Altroc, C. (1996). *Fuzzy logic & neurofuzzy—applications in business & finance*. Prentice Hall, USA.

Azoff, E.M. (1994). *Neural network time series forecasting of financial markets*. John Wiley, USA.

Bose, K., Liang, P. (1996). *Neural networks, fundamentals with graphs, algorithms and applications*. McGraw-Hill, USA.

Davis, L. (1991). *Handbook of genetic algorithms*. Int. Thomson Com. Press, USA.

Dostál, P. (2011). *Advanced decision making in business and public services*. CERM, Brno.

Dostál, P. (2008). *Advanced economic analyses*. VUT – FP – Brno.

Dostál, P., Rukovanský I. Economy and Application of Chaos Theory. In *International Journal of Business, Humanities and Technology*, Vol.5, No.5, 2015, ISSN: 2162-1357(Print), ISSN: 2162-1381(Online).

Ehlers, F.J. (2004). *Cybernetic analysis for stock and futures*. John Wiley, USA.

Franses, P.H. (2001). *Time series models for business and economic forecasting*. Cambridge University Press, UK.

Gately, E. (1996). *Neural networks for financial forecasting*. John Wiley, USA.

Hagan, T., Demuth, B. (1996). *Neural network design*. PWS Publishing, USA.

Chen, G. (2000) *Controlling chaos and bifurcations in engineering systems*. CRC Press, China.

Chen, P., Jain, L., Tai, C. (2005). *Computational economics: a perspective from computational intelligence*. Idea Group Publishing.

Chen, S., Wang, P., Wen, T. (2004). *Computational intelligence in economics and finance*. Springer, New York.

Chen, S., Wang, P., Wen, T. (2007). *Computational intelligence in economics and finance*. Volume II, Springer, New York.

Kazabov, K., Kozma, R. (1998). *Neuro-fuzzy—techniques for intelligent information systems*. Physica-Verlag, Germany.

Klir, G.J., Yuan, B. (1995). *Fuzzy sets and fuzzy logic, theory and applications*. Prentice Hall, New Jersey, USA.

Li, Z., Halong, W.A., Chen, G. (2006). *Integration of fuzzy logic and chaos theory*. Springer, New York.