Volume XV, Issue (1) (2019) 422-431

International May Conference on Strategic Management

MACHINE LEARNING METHODS IN STRATEGIC MANAGEMENT

Krichevsky Mikhail[1], Serova Elena[2]

[1]Saint-Petersburg State University of Aerospace Instrumentation;
[2]National Research University Higher School of Economics, Russia

Abstract: One of the main objectives of strategic management is the development and selection of strategies to achieve the desired results. The main goal of this paper is the analysis of the main domains or areas of machine learning application to support the process of strategic planning and decision making. The scientific methodology of the research studies is methods and procedures of modeling and intelligent analysis. This is theoretical and empirical paper in equal measure. This paper deals with the issues of machine learning implementation and how intellectual models and systems can be used to support the process of strategic planning in the context of theory of economic growth and development. At the preprocessing stage on the basis of a modeled base of examples of strategy options, the use of clustering methods for forming groups of similar parameters that influence the choice of strategies and groups of similar enterprise objects, each of which has a certain type of strategy, are demonstrated. On the next step the selection of ranked characteristics that affect the choice of strategy is made. At the stage of solving the problem of choosing strategies, neural network and neuro-fuzzy approaches are used. The advantage of this hybrid method is based on the fact that the hybrid technology can combine the advantages of neural networks as well as the advantages of fuzzy logic.

Keywords: Strategic Management, Machine Learning, Intelligent Technologies and Systems, Neural Networks, Hybrid Modeling

1. INTRODUCTION

For any enterprise the choice of strategy is an important task, on the decision of which it further activity depends. In essence, a strategy is a long-term specific direction in an organizational development in terms of relationships system within an organization and position of an enterprise among competitors.

In other words, a strategy is a set of rules for making decisions that guide an organization in its activities [1-5]. Each of the strategies, chosen by an enterprise, leads to different outcomes; therefore, the selection of strategy should be made taking into account all available information at the decision-making stage. As stated in [1], strategy can be considered as a subtle and somewhat abstract concept; however, strategy is a tool that can seriously help a firm in conditions of instability. As a result, strategy, as a management tool, serves as an important guideline for enterprises of various fields of activity.

When choosing a type of strategy, it is necessary to define and substantiate the list of indicators by which the strategy type is assessed, to identify the most suitable method for solving this task, to test the performance of the proposed methodology.

The paper is structured as follows: first, the authors consider machine learning methods which can be used to solve the problem of strategy choosing. Then they describe the data preprocessing employing cluster analysis, consider the principle of neural networks, specify a way to reduce the dimension of the feature space, analyze the tools when executing enterprise development strategies taken from the Ansoff matrix, provide an assessment of the strategy using the neuro-fuzzy system.

2. METHODOLOGY AND LITERATURE REVIEW

Analysis of any data and working with them implies building a model from observations and its further use, for example, in classification, forecasting, etc. Methods that are used in Artificial Intelligence (AI) to work with data are part of Machine Learning (ML), which is a subset of AI.

According to Kim, Machine Learning is a modeling technique that involves data. This definition may be too short for first-timers to capture what it means. Machine Learning is a technique that figures out the “model” out of “data.” Here, the data literally means information such as documents, audio, images, etc. The “model” is the final product of Machine Learning [6, p.2]. ML is suitable for task solutions related to intelligence, in particular, in situations where the laws of physics or mathematical equations do not make it possible to build a model. The process of building a model from training data is shown in Figure 1.

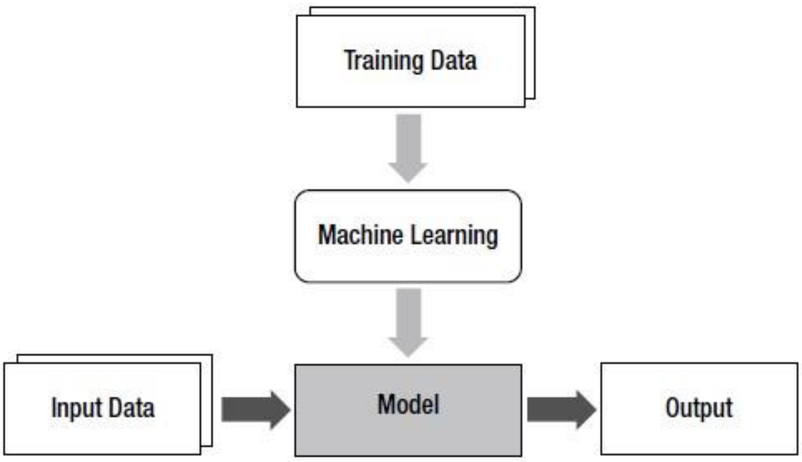


Figure 1. Applying a model based on field data [6]

The vertical flow of the figure 1 indicates the learning process, and the trained model is described as the horizontal flow, which is called inference.

It must be emphasized that data to build model and its application are different. Some of the Machine Learning methods used in this work are briefly described below.

2.1. CLUSTER ANALYSIS

Cluster Analysis (CA) is a convenient way to identify homogeneous groups of objects. United similar objects are called clusters. Objects (or cases, observations) in a particular cluster (group) have similar characteristics, but differ from objects that do not belong to this group. The purpose of CA is to identify a number of objects that are similar among themselves in a certain set of parameters (cluster variables). After selecting such variables, it is necessary to select a clustering procedure to form groups of objects.

There are many different clustering procedures, but the practical difference is the differentiation between hierarchical methods and centroid-based clustering (primarily, the k-means clustering) [7]:

- hierarchical clustering;
- k-means clustering.

Each of these procedures has its own approach to grouping the most similar objects into clusters. Agglomerative procedures are probably the most widely used of the hierarchical methods. They produce a series of partitions of the data: the first consists of n single member ‘clusters’; the last consists of a single group containing all n individuals. The basic operation of all such methods is similar, and will be illustrated for two specific examples, single linkage and centroid linkage. At each stage the methods fuse individuals or groups of individuals which are closest (or most similar). Differences between the methods arise because of the different ways of defining distance (or similarity) between an individual and a group containing several individuals, or between two groups of individuals [7, p. 72].

The k-means clustering method forms exactly k different clusters located at the greatest possible distances from each other. Its essence lies in the fact that the classification process begins with the assignment of some initial conditions (for example, the number of clusters being formed, threshold for classification process completion, etc.). As in hierarchical cluster analysis, in iterative methods there is the problem of determining the number of clusters. In general, their number may not be known. Not all iterative methods require the initial specification the number of clusters. But for the final decision about structure of set under study, several algorithms can be tried by changing either the number of clusters formed or the established proximity threshold for combining objects into clusters. Then there is the opportunity to choose the best partition.

2.2. FEATURE SELECTION

One of the widespread tasks of the ML is predictors selecting from a large list of candidates. Through a large number of input parameters, neural networks (NN) learning is hampered, time spent increases. Therefore, task of selecting from a large list of input variables those that are most useful for predicting the outputs of dependent variables is very important.

The *Feature Selection and Variable Screening* module of software *Statistica 13* can serve as an ideal preprocessor in ML, allowing to select abbreviated predictor sets for further analysis. The methods implemented in this module are designed to processing large sets of continuous and/or categorical predictors in tasks like regression or classification.

2.3. NEURAL NETWORKS

The neural network can be considered as "a black box" that reflects the situation with a completely unknown process, but there are observations (examples). Here the inputs and the outputs are known, but a base of examples is required for the network training. In general, an NN is a machine that simulates method of processing of specific task by the brain [8-10, 12, 13]. Such a network is implemented by using electronic components or modeled by a computer program. Due to its training and generalization capabilities, neural networks can be expressed as a mathematical mapping of the human brain architecture.

Let us explain the principle of neural network technology on a single-layer network with R inputs and S neurons in the layer. In this network, each element of the input vector **p** is connected to each neuron input through the weight matrix **W**. The *i*th neuron has a summer that gathers its weighted inputs and bias to form its own scalar output $n(i)$. The various $n(i)$ taken together form an S-element net input vector **n**. Finally, the neuron layer outputs form a column vector **a** [14].

The principle of neural network operation remains unchanged except the case when outputs of the hidden layer are the inputs of the output layer. After data collection and forming a database of examples, the choice of the type of network is solved quite simply.

2.4. HYBRID NETWORKS

The hybrid network, integrated the principles of neural networks and fuzzy logic, is a multilayered neural network with a special structure without feedback, which uses conventional (not fuzzy) signals, weights and activation functions. The main idea used in the hybrid network model is applying an existing data sample to determine the parameters of membership functions that best fit the particular fuzzy inference system. To find the parameters of the membership functions, we use well-known neural network learning procedures. The essence of such model is to determine the parameters of fuzzy systems through training methods adopted in neural networks [9-13].

The fuzzy logic mechanism can be implemented using an algorithm like Mamdani or Sugeno. The main difference between Mamdani and Sugeno output systems is the way to get a crisp output decision. In Mamdani algorithm, a crisp output value is obtained by defuzzification of the output fuzzy set. In Sugeno algorithm, a weighted average value is formed at the output.

The Adaptive Network-Based Fuzzy Inference System (ANFIS) implements Sugeno's fuzzy inference system in the form of a five-layer neural network. ANFIS architecture consists of five layers as shown in Figure 2. Each layer contains several nodes described by the node function [9, 12].

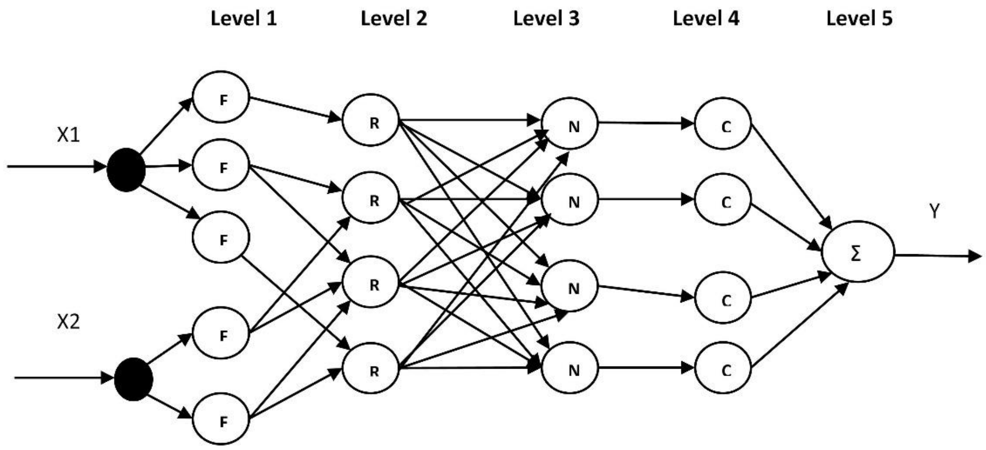


Figure 2. Architecture of ANFIS hybrid model [12]

According to [12, pp. 452-458]:

A. Layer 1

Each node in the first layer of ANFIS architecture processes the inputs coming by using node functions.

B. Layer 2

This is a layer of rules. The outputs of the first layer constitute the inputs of this one.

C. Layer 3

This is a layer of normalization. All the outputs of the nodes in the layer of rules are used as input. The proportion of the ignition power of the node *i* in the layer of rules to the sum of the ignition power of all the nodes gives the normalized ignition rate of node I

D. Layer 4

This is the clarification layer. Node *i* in this layer compute the contribution of i-th rule toward the overall output.

E. Layer 5

There is one node in this layer. This node sums the output values of each node in the layer 4. This summation is the output value of the ANFIS system.

3. EXPERIMENTS AND CONCLUSION

Various enterprise development strategies were presented as early as 1957 by I. Ansoff. However, until now this strategic management tool is actively used to determine the main directions of business development [1]. The most important strategies of activity and development of the company are as follows:

1. Market penetration.
2. Market development.
3. Product development.
4. Diversification.

Each of the strategies is characterized by the following indicators:

- x1 - expansion of market share;
- x2 - increase in the number of purchases of goods;

- x3 - increase the frequency of purchases of goods;
- x4 - discovery of new possibilities for using the product;
- x5 - the use of new distribution channels;
- x6 - search and penetration of new market segments;
- x7 - product sales in new regions;
- x8 - modernization of existing products;
- x9 - expansion of the product range;
- x10 - the creation of a new generation (models) of the product;
- x11 - development and production of a fundamentally new product;
- x12 - availability of competences for the development of a new business;
- x13 - an opportunity to grow in current markets with current staff

These indicators determine the tools that influence on strategy choice. The set of these tools may differ slightly in different sources, but it does not significantly effect on the purposes of this work.

To create a database of examples on which the neural network is trained, two approaches are possible:

- application of real data;
- use of "toy" datasets.

In machine learning, it is important to learn how properly use toy datasets, because learning the algorithm on real data is difficult and can end in failure [15]. Toy sets of data are crucial for algorithms understanding. If there is a simple synthetic data sample, it is sufficient easy to evaluate whether the algorithm has learned the correct rule or not. On real data to obtain such assessment is difficult. Here the authors have used Monte Carlo method to form a toy data sample.

Based on the analysis of the proposed strategies and characteristics that influence on these strategies, the following assumptions were accepted:

- Implementation of the strategy, 1 (market penetration) requires high values of predictors x1 - x4.
- Implementation of the strategy, 2 (market development) requires high values of predictors x5 – x8.
- Implementation of the strategy, 3 (product development) requires high values of predictors x1, x9 – x11.
- Implementation of the strategy, 4 (diversification) requires high values of predictors x4, x12, x13.

Each predictor was generated according to normal distribution, based on a 10-point scale, taking into account the “high” and “low” values of parameters. As a result, 40 variants were considered; therefore, the observation matrix consists of 40 rows and 13 columns. Figure 3 shows the dendrogramma of all 40 options that determine the chosen strategy.

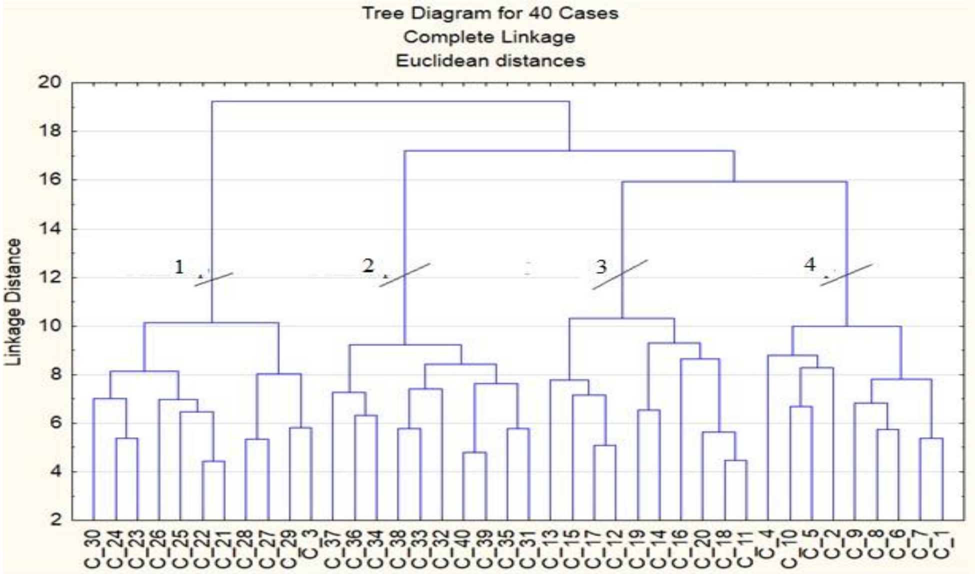


Figure 3. Dendrogramma of strategies

As can be seen from Figure 3, all options form four explicit clusters.

Assessment of the predictors significance was due to the help of the *Statistica Feature Selection and Variable Screening* module. The results of this ranking are shown in Figure 4.

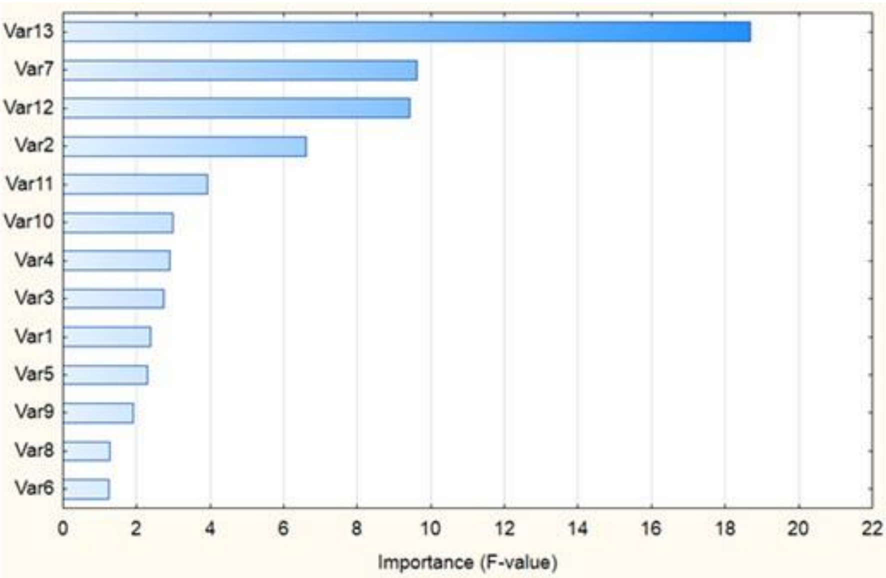
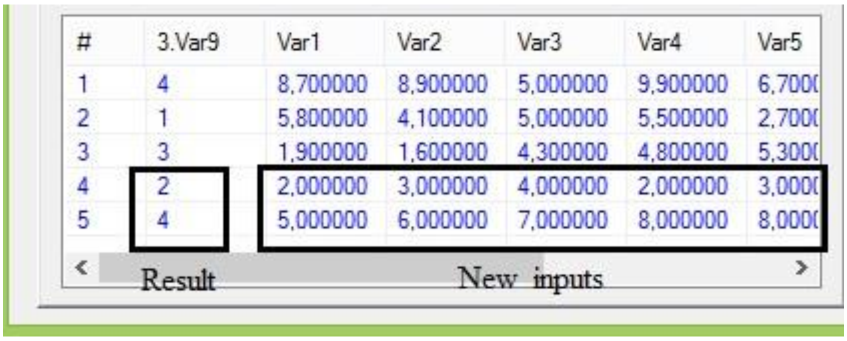


Figure 4. Histogram of the most important predictors

Based on the graph of the predictors importance, the authors reduced their number from 13 to 8. In this case the set includes 13, 7, 12, 2, 11, 10, 4, 3 predictors.

In the next step of the study, the possibility of selecting strategies was evaluated with the help of neural networks. Using *Statistica* software, the authors constructed networks of two types: a multilayer perceptron and a radial-basis function neural network. Network training was conducted on the same toy dataset, supplemented by a column indicating the type of strategies. After training, it was found that the best of the five saved neural networks is a radial-basic network, which is consisting of 8 neurons in the input layer, 10 neurons in the hidden layer and 2 neurons in the output layer. This process is illustrated in Figure 5.



#	3.Var9	Var1	Var2	Var3	Var4	Var5
1	4	8,700000	8,900000	5,000000	9,900000	6,700(
2	1	5,800000	4,100000	5,000000	5,500000	2,700(
3	3	1,900000	1,600000	4,300000	4,800000	5,300(
4	2	2,000000	3,000000	4,000000	2,000000	3,000(
5	4	5,000000	6,000000	7,000000	8,000000	8,000(

Figure 5. The result of NN work

The first three lines of Figure 5 show the results of network classification of the data that network has already seen during the training. The column with the index 3.Var9 demonstrates the strategy number. The last two lines define the new inputs, and the network refers them to one of the strategies.

Finally, the task of the strategies effectiveness evaluating was solved by using ANFIS system in software MatLab 2018b. Dividing the sample into two parts: the training (28 options) and test (12 options), the neuro-fuzzy system was created. This system is shown in Figure 6.

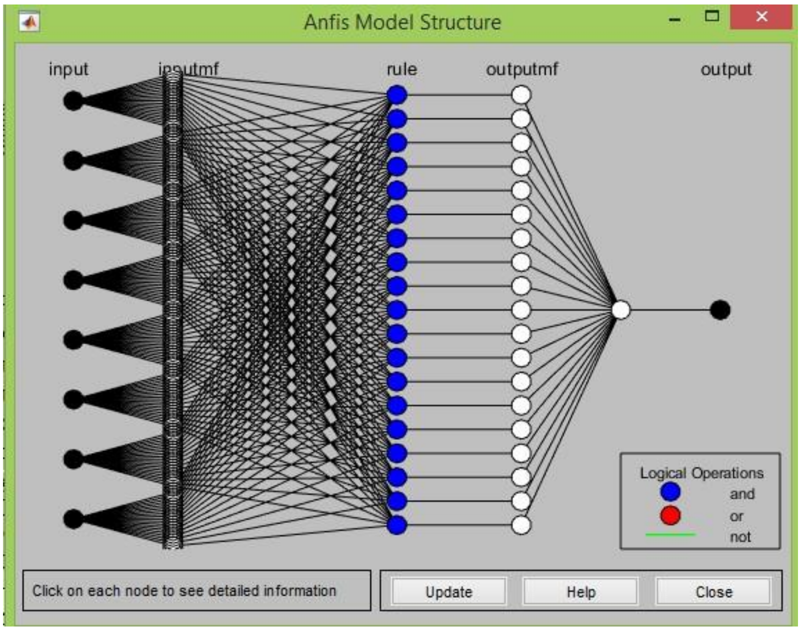


Figure 6. Neuro-fuzzy system

ANFIS system automatically generates a base of rules, which in this case consists of 19 rules (circles in the layer indicated by the "rule").

The test results of this system are shown in Figure 7. As can be seen from Figure 8, the test data is close to the results that the ANFIS system produces.

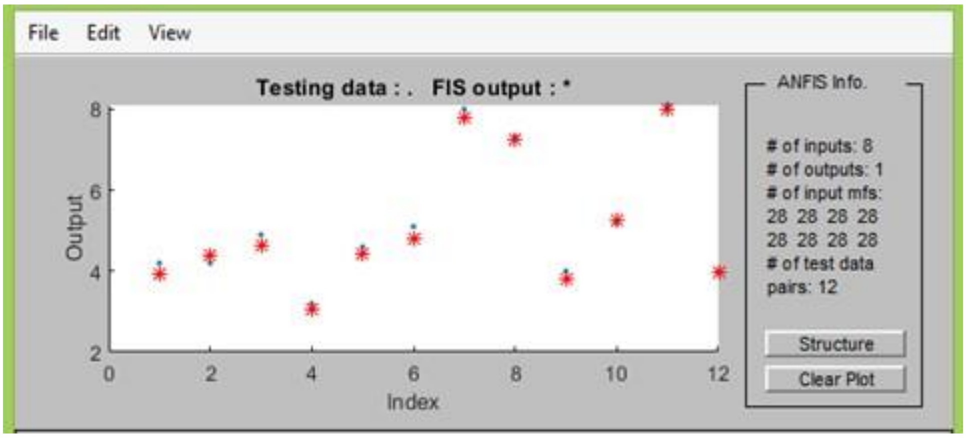


Figure 7. Test results

Finally, using the fuzzy part of the ANFIS system, it is possible to evaluate the effectiveness of the strategy through the option to view the rules (a 10-point scale is chosen here). Figure 8 shows a fragment of the figure that reflects this result: when applied to the input of the system the values of eight predictors, which are listed above the columns of inputs, the system evaluates this strategy type about 4 points.

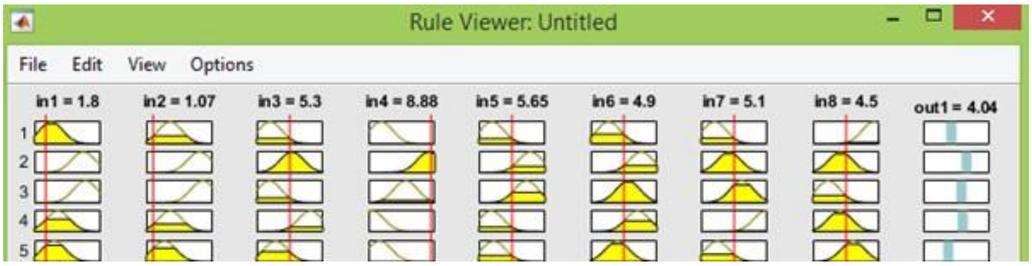


Figure 8. Strategy assessment

In conclusion, it should be state that machine learning methods can be successful implement to support the processes of strategic planning and decision making. With the help of using the advantages of clustering, neural network and hybrid neuro-fuzzy approaches it can possible to choose the competitive development strategy of an enterprise.

4. DISCUSSION

In further research it is planning to pay more attention to the development of strategies assessment systems.

ACKNOWLEDGMENT

The work was supported by the grant of the Russian Fundamental Research Fund No. 18-010-00338.

REFERENCES

1. Ansoff I., Strategic management: Classic edition, Palgrave Macmillan, New York, 2007.
2. Barney J., Hesterly W. Strategic Management and Competitive Advantage. Pearson, 2010.
3. Kotler. F., Berger, R., and Bickhoff N. The Quintessence of Strategic Management: What You Really Need to Know to Survive in Business, Springer, 2010.
4. Leibwitz, J., Strategic Intelligence: Business Intelligence, Competitive Intelligence, and Knowledge Management, Auerbach Publications, 2006.
5. Verbeke, A., International business strategy, Cambridge University Press, 2013.
6. Kim P., MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence., Soul-t'ukpyolsi, Seoul, 2017.
7. Everitt B.S., Landau S., Leese M. et al. Cluster Analysis, Wiley&Sons. New York, 2011.
8. Haykin S., Neural Networks and Learning Machines, Pearson Education, New York, 2009.
9. Jang, J-S. R., Sun, C-T., and Mizutani, E., Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, NJ: Prentice-Hall, Englewood Cliffs, 1997.
10. Kecman, V., Learning and Soft Computing - Support Vector Machines, Neural Networks, and Fuzzy Logic Models, The MIT Press, London, UK, 2001.
11. Ross, T.J., Fuzzy Logic with Engineering Applications, John Wiley & Sons Ltd, 2010.
12. Serova, E., System Models for Strategic Spatial Planning and Regional Development, Proceedings of the 13th European Conference on Management, Leadership and Governance – ECMLG 2017, Academic Conferences and Publishing International Limited, London, 2017, 452-458.
13. Serova, E. and Krichevsky, M., Intelligent Models and Systems in Spatial Marketing Research, The Electronic Journal Information Systems Evaluation 18 (2), 2014, 159-171.
14. Beale M., Hagan M., Demuth H. Neural Network Toolbox User's Guide. R2015b (online) https://mafiadoc.com/neuralnetworktoolboxuser39sguide_59c39e691723dd285cbc5dae.html.
15. Ramsundar B., Zadeh R., Tensor Flow for Deep Learning, O'Reilly, Beijing-Boston, 2018.