Journal of Computer and Knowledge Engineering, Vol.4, No.2. 2021.(17-22)
DOI: 10.22067/cke.2021.70486.1012

DAMP: Decision-Making with the Combination of Analytical Hierarchy Process and Deep Learning (Case study: Car Sales Forecasting) *

Short Paper

Mostafa Sabzekar[1] Farhad Afarideh[2] Arash Deldari[3] Atieh Rezaei[4]

Abstract: Nowadays, prediction and decision making are two inseparable principles in the management and two distinct roles of managers. The organizations spend a large part of their budgets on predictions from past data. They will lose their money if they are neglected. On the other hand, decision-making is the most critical step in problem-solving. Moreover, it is considered the main task of a manager as a problem solver. Making decision becomes more complicated when we are faced with multi-criteria decision-making issues. Combining prediction and decision-making approaches helps researchers to make a better choice utilizing prior knowledge. One of the most essential and comprehensive systems designed for multi-criteria decision-making is Analytical Hierarchy Process (AHP) process. Deep learning as a valuable extension of artificial neural networks has been the focus of many researchers. In this paper, AHP is used to classify, compare, and determine the weights of a deep learning approach. In order to evaluate the efficiency of the proposed method, the prediction of vehicle price application is chosen, and the results are compared with neural networks. The data set is related to the sale of Hyundai and Kia Motors cars in the United States and Canada. It is emphasized that the data are used only to evaluate the proposed method and can be generalized to solve all similar issues. The sales forecasting data of two car companies showed that the proposed method is superior to other regression methods. To extend the proposed methos as our future work, the aim will be to develop a comprehensive decision-making and forecasting system by combining these two approaches.

Keywords: Analytical hierarchy process, deep learning, decision-making, regression

1. Introduction

The problem-solving procedure can undoubtedly be called the most complex and, at the same time, the most sophisticated part of any thought process. All human beings are solving the problem every moment. Our minds and bodies are constantly, and even unconsciously, hosts of various problems, and we are all born with the ability to solve problems. As a general definition, it can be called a high-level cognitive process that requires the integration and control of a set of fundamental skills [1]. One of the most critical stages in problem-solving is decision-making. Prioritization and decision-making are defined as problem-solving activities that lead to an optimal or at least satisfying response [2]. The importance of decision-making is to the extent that in management, it is referred to as the principal task of a manager in the role of a problem solver. [3] considers decision-making equivalent to management and [4] consider management quality a function of decision-making quality. In any prioritization and decision-making process, there are factors known as a criterion that measures the desirability of that decision. These criteria may be expressed in terms of attributes or objectives. They can consider as performance parameters that are used to select decision options. Attributes can be quantitative or qualitative. Objectives consist of decision makers' desires and tendencies that can be expressed in maximizing profits or minimizing costs. Decision-making models can be either single-criteria or multi-criteria. In the single-criterion decision model, a quantitative objective is the basis of decision-making that can be solved using various mathematical methods such as linear programming.

However, in many decision-making problems, the problem solver seeks to optimize multiple criteria simultaneously. In this case, the decision problem is called multi-criteria; one of the most critical issues in mathematics, management, economics, engineering sciences, etc. In many cases, these criteria are not comparable and sometimes even contradictory. Consequently, to solve the problem, we must seek a state with the most significant advantage in terms of all criteria for the decision-maker. Whenever multi-criteria decision-making is based on multiple attributes, it is called multi-attribute decision-making, and multi-objective decision-making if it is based on multiple objectives [5]. One of the most essential and comprehensive systems designed for multi-criteria decision-making is the hierarchical analysis process introduced by [6]. This method considers decision-making issues that are used to solve ranking, selection, evaluation, and the prediction problems.

A critical issue in this research is whether, by proper training and determination of logical weights, by hierarchical analysis, one can design a predictive system so that the most minor error is achieved for accurate estimations close to actual statistics. For this purpose, in this study, a combination of the hierarchical analytical process with deep learning is proposed. In order to measure the efficiency of the proposed method, the automobile sales data set for Hyundai and Kia Motors in the USA and Canada from 2010 to 2014 was used. Therefore, it will be a matter of predicting car sales.

The automotive industry encompasses all parts of the design, development, production, market, and sales of motor vehicles. Companies and factories involved in designing,

* Manuscript received, April, 23, 2021; accepted, December, 29, 2021.
[1] Corresponding Author: Assistant Professor, Department of Computer Engineering, Birjand University of Technology, Birjand, Iran
Email: sabzekar@birjandut.ac.ir
[2] M.Sc. Department of Statistics, Allameh Tabataba'i University, Tehran, Iran
[3] Assistant Professor, Department of Computer Engineering, University of Torbat Heydarieh, Torbat Heydarieh, Iran.
[4] M.Sc. Department of Information and Communication Technology (ICT), Technical and Vocational University (TVU), Tehran, Iran.

manufacturing, marketing, and selling motor vehicles are part of the industry. In 2008, more than 70 million motor vehicles, including ordinary cars and commercial vehicles, were produced worldwide. In 2007, 71.9 million cars were sold worldwide, of which 22.9 million in Europe, 21.4 million in the Asia-Pacific, 19.4 million in the US and Canada, 4.4 million in Latin America, 4.4 million, 2 million in the Middle East, and 1.4 million in Africa were sold. When markets were stagnant in the US and Japan, Asia and South America grew and became powerful. The large markets of Russia, Brazil, India, and China also appear to have overgrown. The automotive industry, as one of the largest in the world, with vast amounts of financial and time capital invested in it, will require careful and accurate prediction of its futures and competitors to make significant and sensitive decisions. The automotive industry is affected not only by macro variables but also by hundreds of other factors. Many of these factors complicate decisions about the future of production and sales in the automotive industry. Producing any automotive product like any other industrial product requires preliminary investment and study. In recent years, the automotive and its related industries have taken on economic and political aspects, to the extent that the import or export of a country is sometimes subject to the trade of the automobile industry, and the trade balance is measured by this criterion.

Therefore, to predict car sales, which is also the subject of this study, a set of experts in North America, specified the priorities effective in car sales using questionnaires. Then, the processed weights obtained from the comments were presented as input to the neural network. The proposed conceptual model first finds the weights of the factors affecting sales, then attempts to discover the intrinsic relationship between the data, which finally achieves a more accurate prediction. Therefore, our main issue is defined as predicting the sale of automotive products to formulate and implement strategic decisions for the manufacturing and the distribution of the products by combining hierarchical analytical process and deep learning approaches.

Analytical Hierarchy Process (AHP) has shown great attention in the past decades for solving multi-criteria decision-making problems. For example, it has been applied to routing [7] and found the backup channels [8] in wireless networks, scheduling in the cloud [9], customer relationship management (CRM) [10], project quality management [11], medicine [12], e-learning [13], robotics [14], etc. On the other hand, the deep neural network has been attracted by different research communities during these recent years, from radiotherapy [15] and agriculture [16] to image processing [17] and network security in the Internet of Things (IoT) [18].

Due to the discussed applications of AHP and neural networks, there are several works in the literature that have tried to combine these two categories of methods. For example, in [19], the authors combined neural networks and AHP to choose the best place. This idea also was utilized by [20]. In another research, which we will compare with our method, the authors used this idea in car sales prediction [21]. Moreover, it has been applied to solve other similar problems [22-24].

In this paper, we utilized the benefits of deep neural networks for solving the problem. First, the opinions of experts are extracted, and AHP is used for weighting the criteria. Then, a deep learning approach is initialized with these weights. To the best of our knowledge, no study has utilized the AHP for initializing the weights of deep learning networks. Studies have usually used the optimization methods such as evolutionary algorithms to find the optimal weights for artificial neural networks (ANNs) and deep neural networks. Determining the initial weights for deep neural networks utilizing the experts' opinion can lead to better and more interpretable results.

2. Theoretical Foundations

A. Analytical Hierarchy Process (AHP)

As mentioned bifore, one of the most critical steps in problem-solving is decision-making. Prioritization and decision-making are defined as the problem-solving activities leading to an optimal or at least satisfying response. Decision-making models can be either single-criteria or multi-criteria. In the single-criterion decision model, only a quantitative goal is the basis of decision-making that can be computed using various mathematical methods such as linear programming. However, in many decision-making issues, the problem solver seeks to optimize multiple criteria simultaneously. In this case, the decision problem is called multi-criteria, one of the most critical issues in mathematics, management, economics, engineering sciences, etc. One of the most important and comprehensive systems designed for multi-criteria decision-making is the Analytical Hierarchy process first introduced by [6]. This method is used to solve problems such as ranking, selection, evaluation, preparation, and the prediction that are all considered as decision-making problems. The advantages of this approach include the possibility of formulating the problem in a hierarchical way, the possibility of considering different quantitative and qualitative criteria in the problem, as well as the possibility of incorporating different options in decision-making and sensitivity analysis on the criteria. In addition, since this method is based on pairwise comparisons, it facilitates judgment and computation. The AHP can also express the degree of consistency and inconsistency of the decision, which is a prominent feature of this approach in solving multi-criteria decision-making problems. Finally, it should be noted that this method has benefited from a solid theoretical foundation [3].

In the next step, the optimal weights of the edges are calculated, and the decision compatibility is examined. Some of the essential features and advantages of the AHP method can be summarized as follows [6]:

- Uniqueness and the simplicity of the model;
- Complexity: This approach uses both systematic and detailed analysis simultaneously to solve complex problems;
- Hierarchical structure (like human thinking);
- Consistency: it calculates and presents the logical consistency of judgments.

B. Deep learning

Deep structures, unlike shallow ones that usually have a hidden layer, have more hidden layers in their architecture. In supervised learning, after the last hidden layer in both types of structures, a layer with linear activation is placed to produce desirable outputs [25]. Many shallow structures such as Gaussian combinations and neural networks with a hidden layer are general approximators. In other words, they can represent any function, but there is a fundamental limitation considering these structures. These structures can represent any function that has enough variables in the hidden layer. In practice, this constraint is not always possible to meet. Specifically, for functions with high fluctuations, the number of parameters required increases exponentially as the input data dimension increases [26].

In contrast, deep structures that employ a greater number of hidden layers than shallow structures while being general approximators can provide more efficient representations, simultaneously [27]. In practice, deep structures lead to representations with significant features that can resist transformations such as displacement and rotation [28]. Representations obtained from deep structures are mainly distributed (which lead to non-local generalizations) and sparse. Moreover, deep structures can learn hierarchical representations that are very similar to the visual structure in humans [29].

In this study, deep feedforward networks are used. The traditional belief about feedforward networks was that training in these networks with backpropagation is difficult [30]. However, [31] claim that good classification functions can be achieved in these networks. In this research, the best classification result of handwritten numbers of the MNIST database was achieved using many hidden layers, many neurons per layer, and numerous deformed training images to avoid overfitting. The number of parameters in the proposed structure was between 1.34 and 11.12 million, leading to low generalization capability and very high computational overhead. In order to overcome the above challenges, different studies such as [13] were performed.

Therefore, deep learning is neural networks that model high-level abstract concepts at different levels and layers. The main benefits of this learning method can be stated as follows:

- Learning representation: the primary requirement of any learning algorithm is to extract features from the inputs. These features may be manual (supervised methods) or automated (unsupervised methods). Manually extracting features is usually time-consuming, inaccurate, incomplete, or overly expensive. Deep learning is a way to extract features automatically.
- Multilayer learning representation: deep learning enables us to build high-level abstract concepts using bottom-up multilevel learning that leads to high accuracy.

In the following section, we will present the proposed structure that is based on a deep learning approach.

3. The Proposed Method

The neural networks and, by nature, deep networks are considered as black-box models. The reason is that there is no direct and simple link between their trained weights and the function being approximated by them. The black-box models are created and designated directly from data. It means that no one (even those who design them) can understand how their variables and weights are being combined to make predictions. The performance of prediction is directly associated with how these weights is determined. In most of the training algorithms, they start with random weights and refine them based on training data. Therefore, they are sensitive to initial random weights and a good initialization have positive effects on the performance of the algorithm.

Many studies in the literature have tried to find appropriate weights. For example, in some researches, the optimal weights were determined by combining the learning algorithm with evolutionary algorithms such as genetic algorithms and so on. It should be emphasized that obtaining the optimal weights in training stage does not necessarily guarantee high accuracy in testing stage. Thus, we cannot really talk about the best weights and therefore the best solution.

As mentioned above, the main contribution of the paper is to investigate whether the obtained weights from the experts' opinion can improve the prediction results in car sales forecasting application. For this purpose, we first determine how much each input factor affects the car sales using the Analytical Hierarchy Process (AHP). We believe that the obtained weights are more reliable in comparison to calculate them using the optimization algorithms, such as Bayesian optimization, evolutionary optimization algorithms, and etc. The reason is that the weights are obtained from the experts' opinion. In the next phase, we use these weights fed to a deep learning predictive system as initial weights. Figure 1 shows the schematic of the proposed structure:

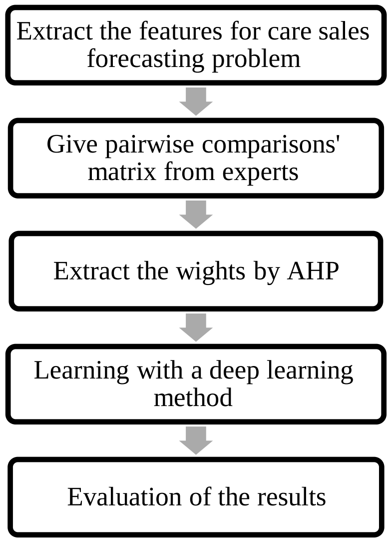

Figure 1. The structure of the proposed method

As Figure 1 shows, the proposed architecture consists of five phases. In the first phase, essential and influential criteria for car sales must be extracted to calculate the AHP weights. The variables of this study are divided into five categories of external factors, including economic dimension, performance, safety, driver and passenger comfort, body and interior size, and an internal factor called season and month influence on car sales. Next, in order to prepare the matrix consisted of score pairs in the AHP method, a questionnaire must be prepared and completed by experts in the automotive field. Therefore, the authors of [28] have benefited from the help of UCLA university professors. Then, using the Expert Choice software, the required weights are calculated by the AHP method, and deep learning structure is used for training of the extracted data. Finally, the trained network will be evaluated. Hence, the steps can be described in more detail as follows:

1. First, we acquire the influential variables in selecting the best model through questionnaires and interviews with the experts. According to the questionnaire, the necessity and importance of each criterion are asked by the Likert scale from the expert. Thus, at this stage, the criteria that are effective in the evaluation process and their importance are determined.
2. The hierarchical structure of the problem is constructed.
3. Determining the statistical population and samples: A set of questionnaires are distributed among the samples to collect information, including the extent of various factors on car sales. The importance of each of these factors is judged by the extremely important, very important, important, slightly important, and unimportant options.
4. Data is normalized and converted into appropriate network inputs.
5. Use of AHP for determining the initial weights: To calculate the initial weights in the neural network using weights derived from the Likert questionnaires, the paired comparisons table must be constructed. The pairwise comparisons table is compiled by dividing the weights obtained for each of the factors in the questionnaire and comparing the individual elements of each level relative to the levels. Then, using AHP rules and using the *Expert Choice* software, the final weight of each criterion, sub-criteria, and options are calculated.
6. Architecture selection for the network: The deep learning network has been used for the proper architecture selection.

In the next section, we compare the results of the proposed architecture with the neural network-based method [21].

4. Experimental Results

A. Analytical Hierarchy Process (AHP)

Similar to the research conducted in [28], the statistical population of this study is the market of Kia and Hyundai products in the US and Canada between 2010 and 2014. The data was extracted from official US industry oversight databases, as well as from Kia and Hyundai. Table 1 shows the target data assumed in this research.

Table 1. Monthly sales

Month	2010	2011	2012	2013	2014
January	52626	65003	78211	80015	81016
February	58056	76339	96189	93816	90221
March	77524	106052	127233	117431	121782
April	74059	108828	109814	110871	119783
May	80476	107426	118790	120685	130994
June	83111	104253	115139	115543	118051
July	89525	105065	110095	115009	119320
August	86068	99693	111127	118126	124670
September	76627	87660	108130	93105	96638
October	73855	90092	92723	93309	94775
November	67324	86617	94542	101416	98608
December	75246	94155	98613	96636	11009

B. Data preparation

The first five exogenous variables were used as neural network inputs, and the network was prepared to enter the sixth variable, which was the effect of season and month on sales. The monthly impact was then normalized and used as the main input of the network. In the previous researches, the effective weights on car sales have been extracted, but this study needs to reproduce these effective weights due to the specific geographical area. In order to study the impact of seasonal and monthly inputs, the best-selling and low-selling months and the ranking of these months based on the sales statistics from 2010 to 2014 were extracted and analyzed and then normalized using the Min-Max method as follows.

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \quad (1)$$

As for the Expert Panel Analysis, it should be noted that this study benefited from the opinion of the experts who are directly involved with the car industry.

C. Analytical Hierarchy Process

The importance of the factors affecting vehicle sales and their weights were extracted by AHP.

D. Training using deep learning methods

The first five variables of the first category are used as inputs for network training. They are normalized between 0 and 1 using Equation (1) (x_{min}=0 and x_{max}=1). Due to the consistency of criteria selection for human beings, the data that influence one's choice over time are unchanged. For example, a person who cares about safety, according to the theory of personality and choice stability, is unlikely to change his mind about his choice in the following five years. However, after considering fixed weights, this research requires a dynamic index to improve the network training accuracy. For this purpose, seasonal and monthly data were used to input the sixth variable, dynamically. The method of extracting the monthly data for each country was independent and unique because the coefficients of the months differ in the two countries. To extract valid monthly data and seasonal impact on purchases by classifying the data, the ranking of the best-selling months was done, and then the obtained rankings were normalized. The results obtained from the analysis of Table 2 are then normalized and presented to the network for training.

Table 2. Ranking based on the yearly sales

Month	2010	2011	2012	2013	2014
January	12	12	12	12	12
February	11	11	9	9	11
March	5	3	1	3	3
April	8	1	6	6	4
May	4	2	2	1	1
June	3	5	3	4	6
July	1	4	5	5	5
August	2	6	4	2	2
September	6	9	7	11	9
October	9	8	11	10	10
November	10	10	10	7	8
December	7	7	8	8	7

E. Results Comparison

In order to evaluate the proposed method, the MSE criterion is considered, which is defined as Equation (2):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}i) \quad (2)$$

where Y is the actual sales value, and $\hat{Y}$ is the predicted sales value. The lower the error, the better the performance of the method. The results of this parameter are then compared with the exponential regression, linear regression, support vector regression (SVR), and AHP + ANN methods [21]. The results are summarized in Table 3.

There are many supervised regression models. We chose SVR for comparing. The SVR model is the promising extension of SVM to solve regression problems [32]. In ε-SVR, the goal is to find a function *f(x)* that has at most ε deviation from the true y_i for all the training data, and is as smooth as possible. In the other words, we do not care about errors as long as they are less than ε. By introducing the slack variables ξ_i, ξ_i^*, some errors are allowed in the constraints. Hence, the SVR can be formulated as the following optimization problem:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &\text{subject to} \quad y_i - (w^T\varphi(x_i) + b) \le \varepsilon + \xi_i, \\ &\qquad w^T\varphi(x_i) + b - y_i \le \varepsilon + \xi_i^*, \\ &\qquad \xi_i, \xi_i^* \ge 0 \end{aligned} \quad (3)$$

The constant C determines the trade-off between the flatness of f and the amount up to which deviations larger than ε are tolerated.

Table 3. Comparison of different methods

Method	**MSE**	**R²**
Exponential Regression	37.2×10^8	0.73
Linear Regression	1.64×10^8	0.77
SVR	1.0×10^8	0.80
AHP+ANN [21]	0.44×10^8	0.84
ANN-GA	0.84×10^7	0.87
The proposed method	0.6×10^7	0.91

As Table 3 indicates, the proposed method performs ten times better than the AHP + ANN method, which demonstrates the superiority of deep learning in conditions similar to artificial neural networks. It is also observed that in a method such as SVR, which is one of the most important forecasting methods, not considering the weights obtained from AHP has reduced the prediction efficiency.

As final discussion, we should emphasize that the AHP does not find the optimal weights for ANN. However, it should be noted that obtaining the optimal weights in training stage does not necessarily guarantee high accuracy in testing stage. Thus, we cannot really talk about the best weights and therefore the best solution. Thus, we compared our proposed method with ANN-GA to prove our claim. In ANN-GA, we obtained the optimal weights of a neural network predictor by genetic algorithm. We can conclude from Table 3 that the proposed method outperforms NN-GA. The obtained results proved our claim.

For statistical analysis of the obtained results by different approaches, we utilized Wilcoxon's signed-rank test [33] with significance level of 0.05 for ten independent runs of each method. This test is used for pairwise performance evaluation between the proposed method and the others. Table 4 shows the test results in terms both *MSE* and R^2.

Table 4. P-value for different methods in terms of MSE and R^2

Method	**MSE**	**R²**
Exponential Regression	0.0001	0.0004
Linear Regression	0.0025	0.0012
SVR	0.0056	0.0041
AHP+ANN [21]	0.0067	0.0069
ANN-GA	0.0057	0.0054

As Table 4 shows, the proposed AHP + Deep method showed a significant difference in comparison with the other approaches.

5. Conclusion and Future Works

This paper presented a hybrid method combining the Analytical Hierarchy Process (AHP) and deep learning for car sales forecasting. One of the main challenges in neural networks and by nature deep networks is determining their weights. No one can understand how their variables and weights are being combined to make predictions. We utilized AHP to feed the obtained weights to the neural network as input weights. These weights reflect the experts' opinions about the factors affecting car sales and provide better choices for initialization of network training. Thus, we first acquire the influential variables in selecting the best model through interviews and questionnaires from the experts. Using the questionnaire, the necessity and importance of each criterion were asked and ranked by the Likert scale from the expert. Thus, the criteria that are effective in the evaluation process and their importance were determined. Then, a set of questionnaires were distributed among the samples to collect information, including the extent of various factors on car sales. Next, the AHP was used to calculate the initial weights of networks. The sales forecasting results for two car companies showed that the proposed method was superior to other regression methods. To extend and improve our proposed method as a future work, the aim will be to develop a comprehensive decision-making and forecasting system by combining these two approaches. Thus, it can be left as a future study for the researchers.

References

[1] F.C. Goldstein, and H. S. Levin, "Disorders of reasoning and problem-solving ability," Guilford Press, 1987.

[2] R. R. Yager, "Bidirectional possibilistic dominance in uncertain decision making," Knowledge-Based Systems, (133): 269–277, 2017.

[3] H. Simon, "Making Management Decisions: The Role of Intuition and Emotion," The Academy of Management Executive, 1(1), 1987.

[4] D. Lee, P. Newman and R. Price, "Decision making in organisations," Financial Times/Pitman Pub, 1999.

[5] C. Zopounidis and M. Doumpos, "Multiple Criteria Decision Making," Springer International Publishing, 2017.

[6] T. L. Saaty, "The analytic hierarchy process : planning, priority setting, resource allocation," McGraw-Hill International Book Co., 1980.

[7] N. Dharani Kumari, B. Shylaja, "AMGRP: AHP-based Multimetric Geographical Routing Protocol for Urban

environment of VANETs," *Journal of King Saud University - Computer and Information Sciences*, 31(1):72–81, 2019.

[8] C. Salgado, C. Hernandez, V. Molina, F. Beltran-Molina, "Intelligent Algorithm for Spectrum Mobility in Cognitive Wireless Networks," *Procedia Computer Science*, 83, 278–283, 2016.

[9] S. Nayak, C. Tripathy, "Deadline sensitive lease scheduling in cloud computing environment using AHP," *Journal of King Saud University - Computer and Information Sciences*, 30(2): 152–163, 2018.

[10] M. Moshref Javadi, Z. Azmoon, Z. (2011). "Ranking branches of System Group company in Terms of acceptance preparation of electronic Customer Relationship Management using AHP method,"*Procedia Computer Science*, 3: 1243–1248, 2011.

[11] Y. Chang, H. Ishii, "Fuzzy Multiple Criteria Decision-Making Approach to Assess the Project Quality Management in Project," *Procedia Computer Science*, 22: 928–936, 2013.

[12] D. Byun, R. Chang, M. Park, H. Son, C.Kim, "Prioritizing Community-Based Intervention Programs for Improving Treatment Compliance of Patients with Chronic Diseases: Applying an Analytic Hierarchy Process," *Int. J. Environ. Res. Public Health*, 18: 455, 2021.

[13] Sudaryono, U. Rahardja, Masaeni, "Decision Support System for Ranking of Students in Learning Management System (LMS) Activities using Analytical Hierarchy Process (AHP) Method," *J. Phys.: Conf. Ser.* 1477: 022022, 2020.

[14] C. Kim, Y. Kim, H. Yi, "Fuzzy Analytic Hierarchy Process-Based Mobile Robot Path Planning," *Electronics 9*:290, 2020.

[15] P. Meyer, V. Noblet, C. Mazzara, C., A. Lallement, "Survey on deep learning for radiotherapy. Computers in Biology and Medicine," 98:126–146, 2018.

[16] X. Jin, X. Yu, X. Wang,Y. Bai, T. Su, J. Kong, "Deep Learning Predictor for Sustainable Precision Agriculture Based on Internet of Things System," Sustainability,12:1433, 2020.

[17] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence,2021.

[18] A. Diro, N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Generation Computer Systems*, 82:761–768, 2018.

[19] R. Kuo, S. Chi, S. Kao, "A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network," *Computers in Industry*, 47(2): 199–214,2002.

[20] S. Tang, N. Hakim, W. Khaksar, M. Ariffin, S. Sulaiman, P. Pah, "A Hybrid Method using Analytic Hierarchical Process and Artificial Neural Network for Supplier Selection," *International Journal of Innovation, Management and Technology*, 4(1): 109–111, 2013.

[21] D. Farahani, M. Momeni, N. Sayyed Amiri, "Car Sales Forecasting Using Artificial Neural Networks and Analytical Hierarchy Process," *The Fifth International Conference on Data Analytics*, 57–62, 2016.

[22] G. Kabir, M. Hasin, "Multi-criteria inventory classification through integration of fuzzy analytic hierarchy process and artificial neural network," *International Journal of Industrial and Systems Engineering*, 14(1): 74, 2013.

[23] A. Kar, "A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network," *Journal of Computational Science*, 6:23–33, 2015.

[24] M. Ilunga, "Analytic Hierarchy Process (AHP) in selecting rainfall forecasting models," Proceedings of the 20th World Multi-Conference on Systemics, Cybernetics and Informatics, 225–229, 2016.

[25] Y. Bengio, O. Delalleau, and N. Roux, "The Curse of Highly Variable Functions for Local Kernel Machines," Proceedings of the 18th International Conference on Neural Information Processing Systems, 107–114, 2007.

[26] B. Zou, X. Mi, P. Tighe, G. Koch, F. Zou, "On kernel machine learning for propensity score estimation under complex confounding structures,". *Pharmaceutical Statistics*. 1– 13, 2021.

[27] Y. Liu, P. Sun, N.Wergeles, Y. Shang, "A survey and performance evaluation of deep learning methods for small object detection," *Expert Systems with Applications*, 172, 114602, 2021.

[28] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling and R. Yang, "Salient Object Detection in the Deep Learning Era: An In-depth Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

[29] A. Dhillon, G. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Prog Artif Intell*, 9: 85–112, 2020.

[30] X. Glorot, Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–256, 2010.

[31] D. Cireşan, U. Meier, L. Gambardella, J. Schmidhuber, "Deep, Big, Simple Neural Nets for Handwritten Digit Recognition," *Neural Computation*, 22 (12): 3207–3220, 2010.

[32] M. Sabzekar and S.M.H. Hasheminejad, "Robust regression using support vector regressions," Chaos, *Solitons and Fractals*, 14, 110738, 2021.

[33] F. Wilcoxon, "Individual Comparisons by Ranking *Methods*," pp. 196–202, 1992, doi: 10.1007/978-1-4612-4380-9_16.