Journal of Industrial and Systems Engineering
Vol. 11, No. 4, pp. 190-203
Autumn (November) 2018

A comprehensive model of demand prediction based on hybrid artificial intelligence and metaheuristic algorithms: A case study in dairy industry

Alireza Goli[1], Hasan Khademi Zare[1*], RezaTavakkoli-Moghaddam[2], Ahmad Sadeghieh[1]

[1] *Department of Industrial Engineering, Yazd University, Saffayieh, Yazd, Iran*
[2] *School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran*

a.goli@stu.yazd.ac.ir, hkhademiz@yazd.ac.ir, tavakoli@ut.ac.ir, sadegheih@yazd.ac.ir

Abstract

This paper presents a multi-stage model for accurate prediction of demand for dairy products (DDP) by the use of artificial intelligence tools including Multi-Layer Perceptron (MLP), Adaptive-Neural-based Fuzzy Inference System (ANFIS), and Support Vector Regression (SVR). The innovation of this work is the improvement of artificial intelligence tools with various meta-heuristic algorithms including Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Invasive Weed Optimization (IWO), and Cultural Algorithm (CA). First, the best combination of factors with can affect the DDP is determined by solving a feature selection optimization problem. Then, the artificial intelligent tools are improved with the goal of making a prediction with minimal error. The results indicate that demographic behavior and inflation rate have the greatest impact on dairy consumption in Iran. Moreover, PSO still exhibits a better performance in feature selection in compare of newcomer meta-heuristic algorithms such as IWO and CA. However, IWO shows the best performance in improving the prediction tools by achieving an error of 0.008 and a coefficient of determination of 95%. The final analysis demonstrates the validity and reliability of the results of the proposed model, as it supports the simultaneous analysis and comparison of the outputs of different tools and methods.

Keywords: Multi-layer perceptron, adaptive-neural-based fuzzy inference system, support vector regression, invasive weed optimization algorithm, cultural algorithm, feature selection

1- Introduction

Strong decision-making ability is one of the most important requirements of investment management. In essence, every decision must be made based on a prediction of future, which sometimes is mired in uncertainty because of unpredictable changes in the business or organizational environment. By predicting the demand in the entire industry and in the company, the company's market share can be calculated and the business performance can get improved. On the other hand, one of the earliest stages of budget planning is sales prediction. With sales prediction, the revenue of a firm can be determined and the cost of the firm can be adjusted to evaluated revenue. So if the prediction is not sufficiently accurate, it affects other dependent financial and business variables. Also, by predicting sales and determining the distance from reality, management can fill the gap with optimal profitability by analyzing and applying strategic and tactical planning (Hashim et al., 2017).

*Corresponding author
ISSN: 1735-8272,

Hence, the ability to predict the future and forthcoming changes is of great importance for proper decision-making. This ability is also an essential necessity for effective planning and plays a key role in the decisions of intra- and extra-organizational beneficiaries and actors. One of the areas where accurate prediction plays a particularly important role is the financial and economic affairs and especially the capital markets, where it is imperative to foresee the future trends of prices in order to adopt an appropriate buying or selling strategy (Esfahani et al., 2016).

The notable traditional models of prediction include linear or polynomial regression techniques, moving average, Box-Jenkins models, structural equation models, and time series. However, these models suffer from several weaknesses that undermine their utility for any prediction involving complex and non-linear determinants. Recent decades have seen significant developments in a new prediction approach known as artificial intelligence, which is able to discover the complex and nonlinear relationships of variables through an intelligent mechanism inspired by the process of learning in the human brain (Chung, 1999). Because of this extraordinary ability, this approach has found growing use in many fields of science.

In this research, using improved artificial intelligence tools is the main idea to have best accurate predict for DPD. The benefits of using artificial intelligence tools instead of other prediction methods can be included as follow (Anand & Suganthi, 2018).

1. Adaptive learning: artificial intelligence tools have the ability to learn how to perform its tasks on the basis of the given information.
2. Self-organization: artificial intelligence tools automatically organizes and deliver the given information.
3. Real-Time Operators: the calculations of these methods can be designed and implemented in parallel. This will result in optimal results from the artificial intelligence tools.
4. Classification: these methods are able to categorize inputs to get the outputs appropriately.
5. Generalization: These methods enable the network to only obtain a general rule by dealing with a limited number of instances, and generalize the results of these learning to previously data.
6. Flexibility: An artificial intelligence method is also to maintain its captured information and can accept new information without losing any previous information.

Overall, this study has added the following improvements: (a) an integrated model is presented for demand prediction. (b) The significant variables to the DPD are identified and filtered by the feature selection problem. (c) Improving the learning phase of MLP, ANFIS and SVR with some novel meta-heuristic algorithms (e.g., IWO, CA).

The rest of this study is organized as follows: Section 2 represents the literature review. Sections 3 and 4 introduce the artificial intelligence tools and feature selection problem respectively. Section 5 presents the optimization methods for improving artificial intelligence methods. The configuration of the DPD prediction model is described in Section 6. Section 7 discusses the results of the real case study and finally, section 8 concludes the paper.

2- Literature review

Since the prediction is one of the key tools in manufacturing system planning, various researchers have tried to develop this area well. In many ways, demand prediction is focused on behavioral patterns of demand in the past. In these methods, the use of historical data is of particular importance. New researches have focused on the use of artificial intelligence instead of traditional methods.

In a research by Caniato et al. (2005), they presented a method for the prediction of customer demands that do not follow any certain trend. In this research, the causes of demand variations were categorized into three classes of structural (e.g. seasonal changes), managerial (e.g. advertising), and random, and customers were clustered according to their correlation with structural and managerial factors. Then, the demand of each cluster of customers was predicted separately. The results showed that with this approach, random demand variations can be properly predicted and controlled.

In a study conducted by Kochak and Sharma (2015), the 2011-2013 sales data pertaining to oil filters in Madhya India was processed by a back propagation neural network to predict the future demand for this product. The results of this study demonstrated the less than 20% error of the aforementioned tool in the prediction of future demand.

Slimani et al. (2015) used a multilayer perceptron (MLP) neural network to predict the demand of supermarkets in Morocco. This neural network was designed to receive the demand of the past one, two, and three days and predict the demand of the following day accordingly. This study also examined the effect of the number of neurons on the prediction power and found that increasing the number of neurons to 10 can reduce the MSE of prediction to less than 1%.

In a study carried out by Lei et al. (2016), the future demand for the Chinese science and technology services was predicted with 14 factors identified as demand determinants used as model inputs. The predictions of this study were made by six different models belonging to the Group Method of Data Handling (GMDH) family, and ultimately the results were compared.

Wood and Hartzel (2017) proposed several indicators for improving the accuracy of demand prediction based on the point-of-sale. They argued that order size, order diversity, and order frequency are good indicators of sales demand and that relating the total demand to these indicators will improve the demand prediction accuracy. They used a hierarchical linear modeling tool to investigate the validity of this argument and achieved 11% improvement in the prediction power.

In a study by Yong et al. (2017), they collected the data pertaining to three months of daily electricity consumption in South Wales, Australia, and then used a combination of back propagation (BP) neural network and differential evolution algorithm (DE) to predict the future electricity demand. In this study, the BP neural network was used as the base model and the DE algorithm was tasked to find the optimal network weights during the training. The prediction results were compared with the classic BP neural network and adaptive-network-based fuzzy inference system (ANFIS). Ultimately, it was found that the hybrid network has 63% better RMSE than the classic network.

Jiang et al. (2017) proposed a new method of hospital demand prediction based on feature selection and demand foresting. In this method, feature selection was conducted by a genetic algorithm and demand foresting was made by a feed-forward neural network. The method was used for demand prediction in China and the outcomes were compared with the results of the feature selection with the PCA method and demand forecasting with the ARIMA method.

In a study by Rangel et al. (2017), several neural networks were employed to predict the Barcelona's demand for drinking water in the next 24 hours. To reach an improved performance, the training phase of the neural networks was enhanced with the help of a genetic algorithm. Ultimately, evaluations showed that the use of the genetic algorithm improved the prediction performance by up to 1.5%.

Perea et al. (2018) presented a short-term forecasting model for daily irrigation water demand by using Artificial Neural Networks architecture, the Bayesian framework, and Genetic Algorithms. The model was applied in Southern Spain. The presented model improved the prediction accuracy by between 3% and 11%.

Araromi et al. (2018) used sing ANFIS and generalized linear model (GLM) regression to model an activated sludge process for effluent prediction. The results of the research indicate that the presented model has an enormous potential in the estimation of the time lag before the desired effluent quality can be realized. Wu and Shen (2018) proposed the least squares support vector machine model based on grey related analysis to predict natural gas consumption demand. They also used weighted adaptive second-order particle swarm optimization algorithm to optimize the model's parameters. The results showed that the presented method has better generalization ability and training effect in compare of traditional methods.

After reviewing the main contributed papers in the field of demand prediction with artificial intelligence tools, it has been cleared that the main contribution of this paper is to present a new hybrid artificial intelligence tool (e.g., MLP and ANFIS and SVR) with a novel meta-heuristic algorithm (e.g., CA and IWO). At the first step, the significant variables are identified then the hybrid feature selection and artificial intelligence specify the most correlated variable to DPD. This method causes to have less error in the prediction. Then several hybrid artificial intelligence and novel meta-heuristic algorithms are proposed to have the best accurate demand prediction for DPD. These hybrid methods are implemented with a real case study and the compatibility and the performance of these hybrid methods is described.

3- Artificial intelligence tools for demand prediction

The artificial intelligence assisted prediction based on the deep analysis of data and hidden inter-variable relationships is a relatively recent development in the area of prediction and forecasting. One of the most prominent tools of artificial intelligence is the artificial neural network. An artificial neural network, or simply a neural network, is a computational system or method for machine learning, knowledge representation, and applying knowledge to predict the outputs of complex systems. The core idea of these networks is inspired by the way the biological nervous systems process information to learn and create knowledge.

ANFIS is another artificial intelligence tool, which utilizes fuzzy theory for prediction. In ANFIS, the training phase is typically implemented either by back propagation or by a hybrid method (combination of back propagation and the least squares method). The three most widely used types of fuzzy inference systems are the Mamdani fuzzy models, the Sugeno fuzzy models, and the Tsukamoto fuzzy models. The difference between these systems is in their fuzzy rules and therefore their method of aggregation and defuzzification (Sobhani et al., 2010).

The Support Vector Machine (SVM) is a supervised learning method that can be used for both classification and regression of information. In essence, SVM is a two-class classifier that separates the classes with a linear boundary. Originally developed in 1963 by Vapnik, SVM was further extended in 1995 by Vapnik and Cortes to account for the nonlinear state (Adankon & Cheriet, 2009). In recent decades, this method has shown excellent classification and regression performance in several fields (Yazdi et al., 2012). SVM models are very similar to neural networks. In fact, an SVM model with a sigmoid core can be considered an equivalent to a dual-layer neural network (Sherrod, 2008).

4- Feature selection

Feature selection methods attempt to find the best subset of features from among 2^N candidate subsets. All of these methods search for the subset that can optimize the value of an evaluation function, which varies depending on the application. Considering the breadth of possible solutions and the fact that the number of solutions subsets increases exponentially with *N*, it is difficult and costly to find the optimal solution with traditional methods, especially for larger N.

To overcome this issue, meta-heuristic algorithms can be used to accelerate the search for the best combination of features. Each solution of these algorithms is a subset of chosen features. The goal is to find the best combination of features in a way that prediction error is minimized and the least number of features is selected. Figure 1 demonstrates the flowchart of optimizing the feature selection problem.

The main decision to design a framework to solve the feature selection problem with meta-heuristic algorithms is the choice of solution representation and fitness function. In this research, the feature selection problem is solved with the help of IWO, CA, PSO and GA The following subsection describes the solution representation and the fitness function used in each of these algorithms to solve the feature selection problem.

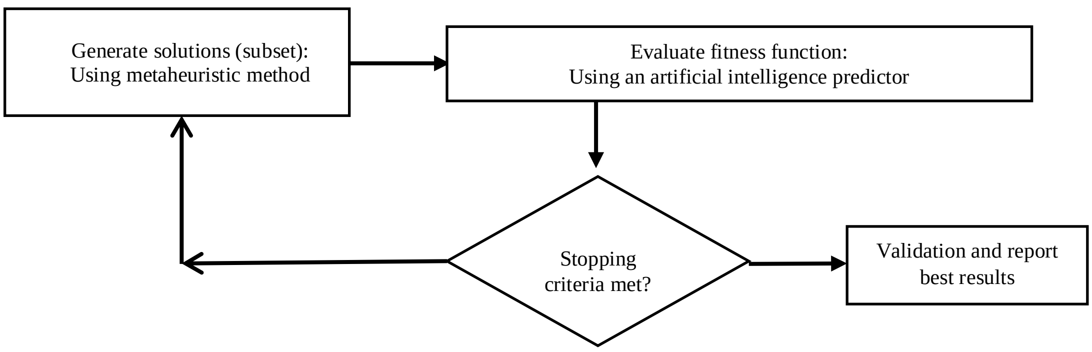

Fig 1. The flowchart of solving feature selection problem

4-1- Solution representation

Each valid and feasible solution to the feature selection problem is represented by a solution string of selected features. Each solution string can be described as a vector with N cells, where N is the total number of possible features. Each cell can take a numeric value between 0 and 1, which when is greater than 0.5, means that the corresponding feature is chosen for the subset. Figure 2 shows an example solution string for a problem where selection must be made from a group of seven features.

0.493	0.017	0.539	0.891	0.366	0.248	0.754

Fig 2. An example solution string for the feature selection problem

In the string shown in figure 2, features 3, 4 and 7 are selected to be included in the solution. The selected features will be used as inputs of the prediction tools (MLP/SVM/ANFIS), and then the resulting MSE will be calculated.

4-2- Fitness function

Each metaheuristic algorithm seeks to optimize the value of a particular function known as the fitness function. The fitness function used in this study is in the form of equation (1). This equation is, in fact, a linear combination of the error and the number of features.

$$Fitness = \alpha \ MSE \ + \beta \ |SF| \tag{1}$$

In equation (1), $|SF|$ is the number of selected features, and α and β are the (importance) coefficients of the error and the number of features, respectively. In this study, α is set to 0.9 and β is set to 0.1.

5- Prediction network optimization

Parameters of the neural networks (weights and biases) used in this part of the study are predetermined. These models are composed of three interconnected simple data processing layers, where the relationship between the output (y_t) and the inputs ($x1_t$, $x2_t$, …, xn_t) is in the form of equation (2).

$$y_t = b_0 + \sum_{j=1}^{q} w_j . g(b_0 + \sum_{i=1}^{n} w_{i,j} x_{i,t}) \tag{2}$$

Where w_j and w_{ij} are the weights between the inputs and the hidden layers, N is the number of inputs, and q is the number of hidden nodes. As implied in Eq. (2), there is only one (output) neuron in the third layer for single-step prediction. Since the random set of network weights is likely to lead to underperformance and production of low-quality outputs, a recommended approach is to optimize the weights of input and hidden layers with the help of meta-heuristic algorithms. In this study, Cultural Algorithm (CA), Invasive Weed Optimization (IWO), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA) are used to optimize the prediction network. In the following subsections, first, the structure of the solution strings of these algorithms is described and then the steps of each algorithm are explained.

5-1- Solution string for network optimization algorithms

For the network optimization algorithms, each solution string represents a set of weights corresponding to the neural network architecture. Figure 3 shows an example solution string for a network architecture with four neurons in the input layer, two neurons in the middle layer and one neuron in the output layer. In this representation, there are 8 (4×2) weights linking the first and second layers (W^1) and 2 weights linking the second and third layers (W^2). The values of all these weights range from zero to one.

W_1^1	W_2^1	W_3^1	W_4^1	W_5^1	W_6^1	W_7^1	W_8^1	W_1^2	W_2^2

Fig 3. Structure of the solution string for network optimization algorithms

5-2-Fitness function

The next step after determining the network weights is to obtain a set of predictions ($\hat{Y}$), compare them with their counterparts among real observations (Y), and measure the difference in terms of mean squared error (Montajabiha et al.). The fitness function used in this step attempts to minimize the sum of MSE, which is obtained from equation (3).

$$MSE = \frac{\sum_t (\hat{Y} - Y)^2}{N} \tag{3}$$

5-3- Cultural algorithm

The cultural algorithm has been introduced by Reynolds in 1994 (Reynolds, 1994). The cultural algorithm is inspired by the evolution of the human culture and its influence on individuals as well as future generations. These algorithms consider two search spaces in their optimization step. In other words, besides the population space of the classic genetic algorithm, cultural algorithms utilize an additional space called the belief space to perform a search in different domains of knowledge. The belief space will be updated by the group of elite people (those who have a higher weight or are more interesting). The effect of belief space is transferred to the population space via an influence function. The addition of knowledge domains improves the efficiency of evolutionary algorithms and leads to a more intelligent search. From this perspective, cultural algorithms can be viewed as an extension of genetic algorithms. Cultural algorithms are able to use various knowledge domains in their belief space to improve the search. Overall, the main components of a typical cultural algorithm include: a) population space, b) belief space, c) accept function and d) influence function.

5-4- Invasive weed optimization algorithm

The Invasive weed optimization is an algorithm introduced by Mehrabian and Lucas that takes inspiration from the growth and dispersal of weeds in nature (Mehrabian & Lucas, 2006).

By definition, weed refers to any unwanted plant with aggressive growth behavior that may threaten the growth of the plants under human cultivation. Although being very simple, IWO has shown excellent performance in the rapid tracking of optimums. As mentioned, this algorithm is inspired by the weeds' natural way of life including seed production, growth, and struggle for survival. The steps of the IWO algorithm can be summarized as follows:

A) Generating a random initial population (seeding) in the target space and evaluating their objective function values.

B) Commencing fitness-based reproduction and updating the standard deviation (seed dispersal).

6- Proposed hybrid method

By developing a hybrid method, we aim to investigate the performance of various artificial intelligence tools and their combinations with several meta-heuristic algorithms. The authors believe that the newer meta-heuristic algorithms may be more effective in the improvement of artificial intelligence tools and result in higher prediction performance.

This paper provides a comprehensive hybrid approach for the selection of features associated with the demand for dairy products and the prediction of this demand. The general structure of the proposed method is presented in figure 4.

As shown in figure 4, first, the feature selection phase of the method is implemented with the help of PSO, IWO, GA, and CA algorithms. The inputs required for this phase are the historical data related to each feature associated with the demand and the historical data of the demand itself. Depending on the algorithm mechanism, a number of effective features are filtered. Then, ANFIS, MLP, and SVM models are used to predict the future demand. The MSE of the predicted values is then calculated and used as the fitness function of the meta-heuristic algorithm. For each algorithm, the goal is to change the selected features so that minimum MSE is obtained. Given the use of four algorithms for optimization and three tools for prediction, the results are reported for 12 states (algorithm-tool combinations). After reporting the results of all states, the one resulting in the best outcome is identified. The features selected in this state are considered as the inputs of the prediction network while the demand is considered as the output.

To check the validity of the predictions made with the artificial intelligence tools, the data must be divided into the test and train groups. Based on the authors' prior experience, 70% of data is assigned to the training group and the remaining 30% is used for the test.

The demand for dairy products is predicted separately by SVM, MLP, and ANFIS with the PSO, IWO, GA, and CA algorithms used to improve the prediction network. The task of these meta-heuristic algorithms is to determine the weights of biases so that prediction MSE is minimized.

The main difference between the combinations of artificial intelligence tools with meta-heuristic algorithms in the first phase and the third phase is that, in the first phase, meta-heuristic algorithms must determine the features that will act as the network inputs and the bias weight are set randomly, but in the third phase, network inputs are constant and meta-heuristic algorithms must determine the bias weights.

In the third phase, 12 different states (algorithm-artificial intelligence combinations) are implemented, and ultimately the method providing the best demand prediction is reported.

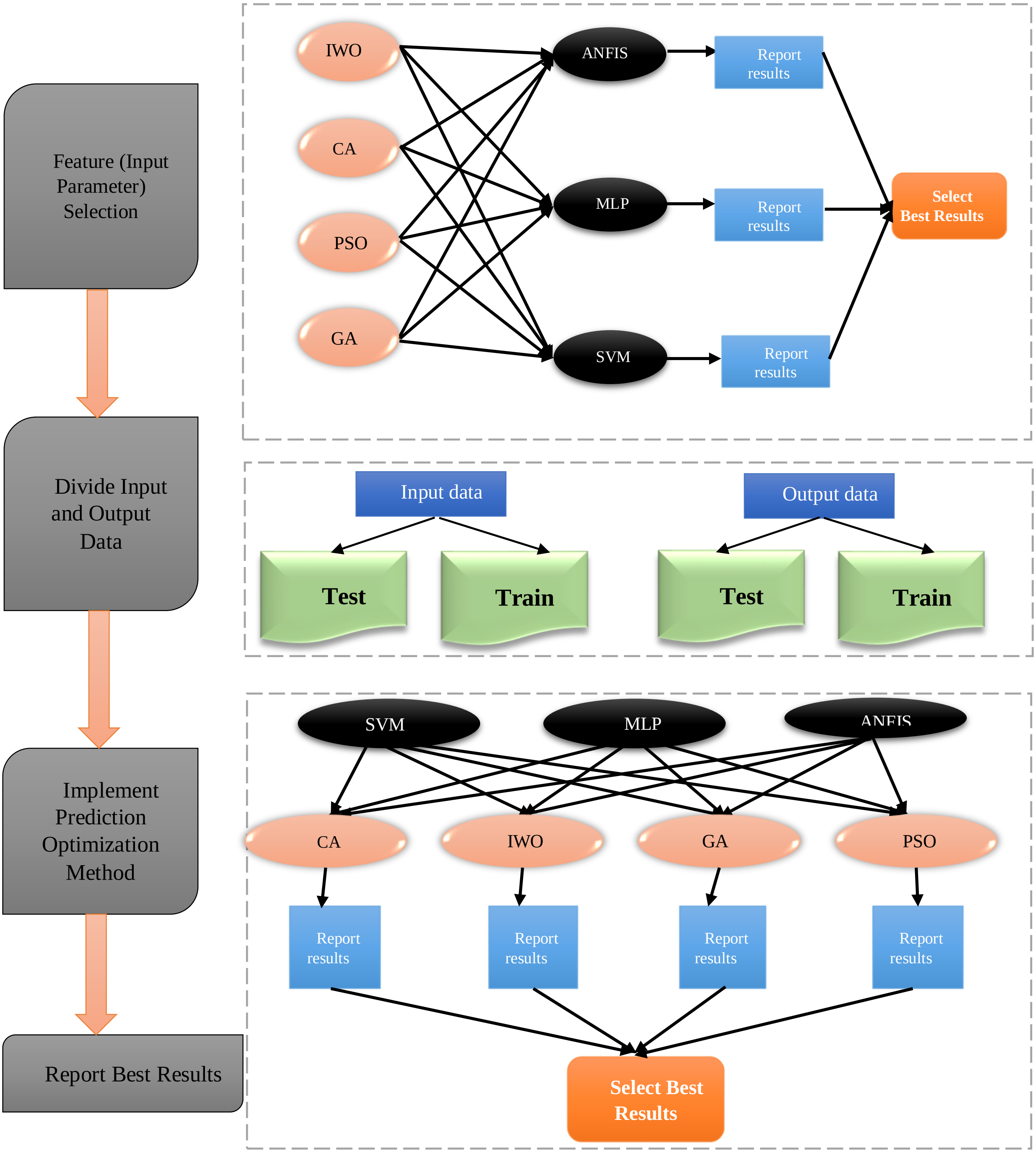

Fig 4. Proposed dairy demand prediction framework

7-Numerical results

In this section of the research, the numerical results of executing hybrid methods are discussed. At the first, the case study and the research data is described. After presenting significant variables, the existence of each variable as the input variables of the prediction network is tested and optimized by implementing hybrid feature selection and artificial intelligence tools. Finally, 12 hybrid artificial intelligence and novel meta-heuristic algorithms are implemented to predict DPD and the error indexes such as MSE, RMSE, MAE and R^2 are calculated and reported.

7-1- Research data

To study the trend of demand for dairy products in Iran, the database containing the monthly sales of dairy products of Pegah Golpayegan Dairy Company during a period of 60 months from 2013 to 2018 was collected. The trend of demand during these 60 months is displayed in figure 5. As shown in this figure, the demand for dairy products does not follow any specific pattern. Thus, to make a prediction about the demand in the coming months, it is necessary to identify a set of determinant factors and then apply an artificial intelligence tool on the collected data. After the examination of the literature, the authors identified 12 factors that affect the demand for dairy products. These factors are listed in table 1.

Table 1. List of the significant feature on dairy products demand

Feature	Symbol	Period	Explain
Inflation	IN	Monthly	General price increase rates
Consumer price index	CPI	Monthly	Average price levels of products portfolio purchased by consumers
Producer Price Index	PPI	Monthly	The average price of goods and services that producers pay for the production
The dairy products price index in urban areas	DPI	Monthly	Measuring the changes in the price of dairy products that are consumed by households in urban areas
Gini coefficient	GC	Monthly	Measuring the distribution of income within the country over a period of time
Average milk price	AMP	Yearly	Milk is the primary ingredient in the production of all dairy products and its price influences the price of other dairy products
Population	POP	annually	The total population of the Iran which is reported annually
Average number of households	ANH	annually	Measuring the distribution of the individuals among the households
Average annual gross income	AAGI	annually	The average income which the fixed costs are not reduced it.
Annual cost percentage for dairy products	ACPD	annually	The percentage of households’ average income that pays for dairy products
Legatum prosperity index	IPI	annually	This index determines the welfare level based on the economic situation (such as GDP), health, education, etc.
Industrial production index	IPI	Annually	Indicates the change in the number of goods and services produced by the enterprises.

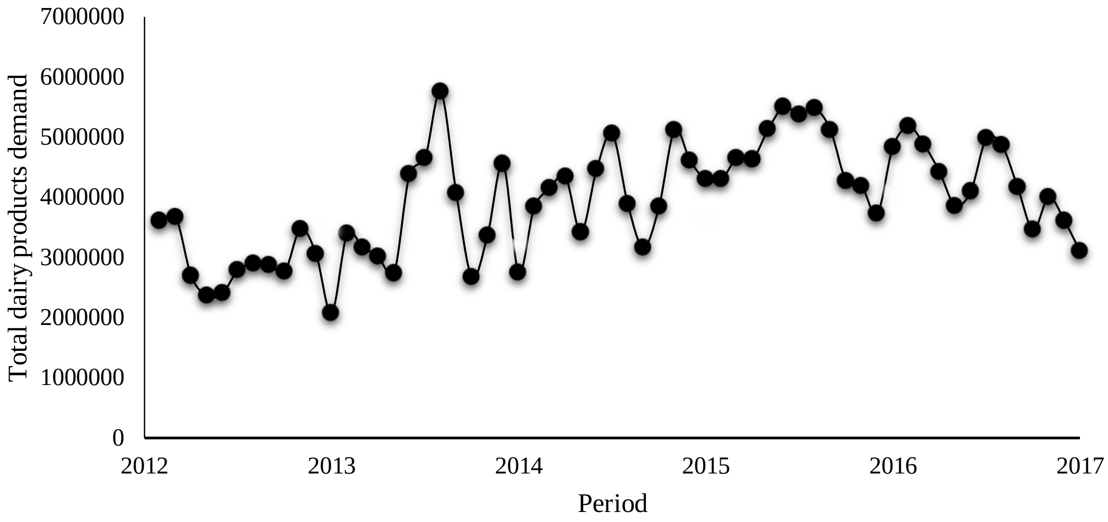

Fig 5. Total demand for dairy products from 2013 to 2018

7-2- Features selection

In this part of the study, the 12 different combinations of algorithms and AI tools to be used for prediction were coded in Matlab R2015. The output of each combination and the corresponding MSE are reported in table 2.

Table 2. The output of feature selection with different algorithms

Optimization	evaluation	IN	CPI	PPI	DPI	GC	AMP	POP	ANH	AAGI	ACPD	IPI	MSE
IWO	MLP	●	●	●		●		●	●	●		●	0.1147
	ANFIS	●		●		●		●			●		0.0320
	SVR	●			●		●	●	●	●			0.1404
CA	MLP					●		●			●		0.0933
	ANFIS	●		●		●					●		0.0352
	SVR						●	●			●	●	0.1575
PSO	MLP	●					●		●	●			0.0324
	ANFIS*	●				●	●	●			●		0.0085
	SVR	●		●				●	●				0.0662
GA	MLP	●				●		●	●				0.1822
	ANFIS			●	●		●	●					0.0091
	SVR	●						●	●		●		0.0913
Percentage		75%	8%	42%	17%	50%	42%	83%	50%	25%	50%	17%	-

As shown in table 2, each meta-heuristic algorithm reported a number of indicators as effective factors. Given the difference in the fit mechanism of MLP, ANFIS, and SVR, there are some differences in the results of each algorithm in each tool. Thus, to select the factors, the frequency of selection of each of the 11 indicators by each of the 12 combinations was examined. The highest frequency was related to the indicators POP and IN, which indicated that population and inflation have the greatest impact on dairy consumption. Meanwhile, the lowest frequency was obtained for CPI, DPI, and IPI. The lowest MSE was achieved while using the PSO for optimization and ANFIS for fitting. The MSE obtained with this combination is 0.008, which is significantly lower than those achieved with other combinations. This suggests that despite significant developments in the field of meta-heuristic algorithms, PSO remains one of the best choices for optimization.

7-3- Dairy product demand prediction

In the next step, the outputs of the PSO-ANFIS combination, i.e. the indicators IN, GC, IMP, POP, and ICPD, were used as the inputs of the prediction network. The output of this network was the total demand for dairy products in Iran. Of the collected data, 70% was assigned to the training set and the rest was used as test data. After using the MLP, ANFIS, and SVR tools to predict the total demand for dairy products, the quality of their output was evaluated with the help of quality indicators. A summary of the obtained results is provided in table 3.

Table 3. Summarized results of dairy product demand prediction

prediction	Optimization	MSE	RMSE	MAE	R^2
MLP	IWO	0.0943	0.30711	0.2794	0.93439
	CA	0.1195	0.34563	0.3210	0.86782
	GA	0.1246	0.35301	0.3477	0.85487
	PSO	0.0991	0.31487	0.2946	0.93628
ANFIS	IWO	0.0108	0.1041	0.2674	0.95511
	CA	0.0350	0.1870	0.3015	0.93585
	GA	0.1279	0.3577	0.4132	0.87089
	PSO	0.0295	0.1716	0.2967	0.93143
SVM	IWO	0.0997	0.3158	0.3786	0.89136
	CA	0.1365	0.3694	0.4016	0.84155
	GA	0.0985	0.3138	0.3004	0.90849
	PSO	0.1364	0.3694	0.3997	0.87726

In table 3, the indicator R^2 shows the level of compliance between the predicted values and the available observations. For all methods, this indicator fluctuates between 0.84 and 0.95. The best result is related to the combination of ANFIS with IWO, which has achieved 95% compliance. The weakest method is the combination of SVR with CA. A closer examination of the results will show that, compared to other algorithms, CA has failed to make a significant improvement in the performance of artificial intelligence tools. On the contrary, IWO and PSO have had very good impacts on the performance of prediction tools.

At the end, a comparison is made between the results of the best combination (ANFIS+IWO) and the actual observations. Figure 6 illustrates the error in different samples of ANFIS-IWO output, and figure 7 shows the level of compliance with these outputs with the target data.

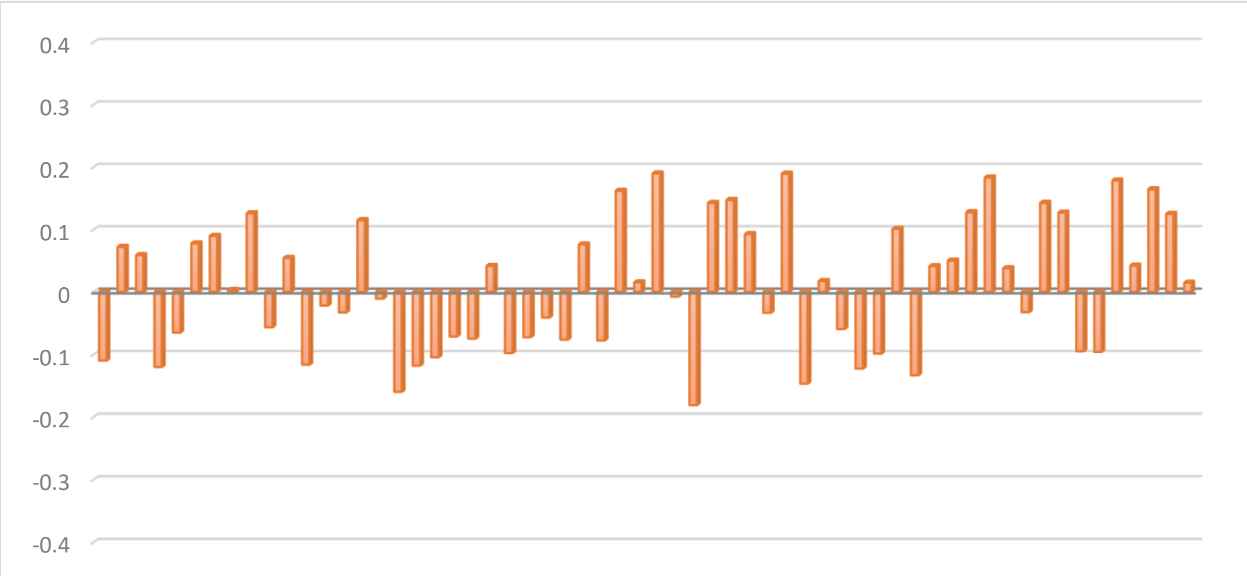

Fig 6. Error in different samples of ANFIS-PSO

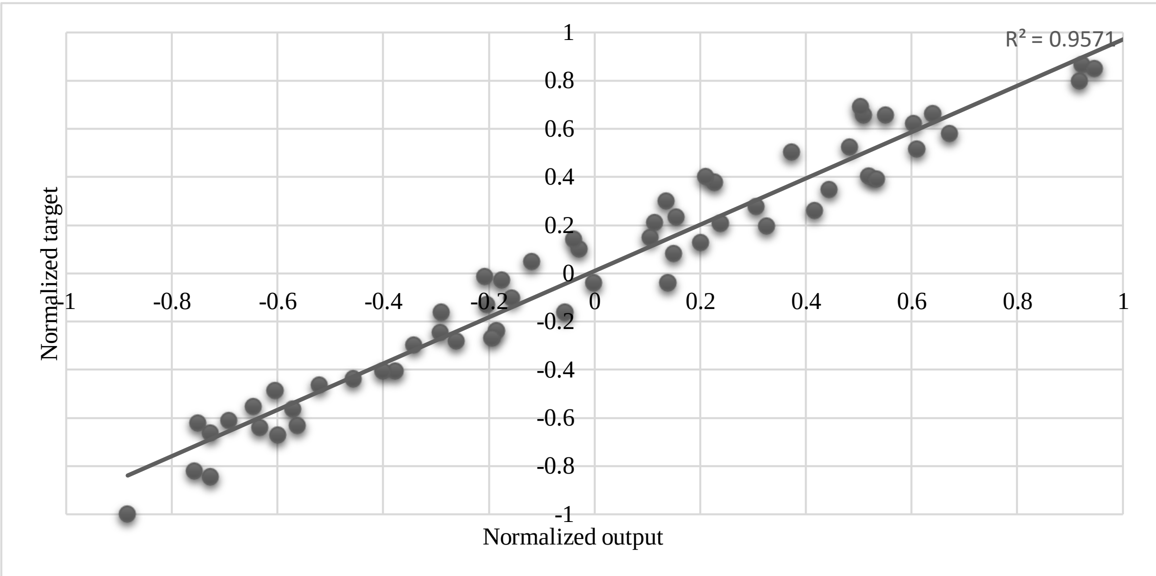

Fig 7. Compliance of ANFIS-PSO outputs with actual observations

Figure 6 shows that the error in different samples fluctuates between -0.2 and 0.2. These positive and negative deviations in the ANFIS-PSO prediction exhibit a symmetry, which can be considered as equivalence to a prediction error of 0. In figure 7, the horizontal axis shows the predictions of ANFIS-PSO and the vertical axis represents the normalized target values. Thus, the closer the points are to the Y=X line, the better (less erroneous) is the prediction. In this figure, the distance from this line represents the magnitude of the error. An examination of figure 7 shows that all points are at a minimum distance from their ideal line, which signifies the effectiveness of the ANFIS-PSO prediction tool.

Before improving theoretical methods on feature selection and artificial intelligence prediction, the results of this study have crucial implication on the strategic and tactical planning for DPD. Improved prediction accuracy ensures the quality and the robustness of obtained future demand as well as the obtained feature subsets. In order to provide some managerial insights for dairy products industry managers, some similarities and difference of feature combinations are presented in Table 2. Both IN and POP are selected as the most important variables that can affect the DPD demand in the future. Changes in these variables imply that the DPD would changes rapidly in a short time horizon.

8- Conclusion

This paper presented a comprehensive framework for improved prediction of DPD in Iran. The proposed framework consists of two parts, feature selection and prediction with the help of hybrid artificial intelligence and meta-heuristic algorithms. The advantage of the proposed framework is the use of different tools and algorithms to find the best possible prediction. While being slightly more time-consuming than the alternatives, this approach provides extremely reliable results, as it enables the user to analyze and compare the outputs of multiple tools and methods.

The main contribution of this study is the use of various meta-heuristic algorithms to improve the performance of artificial intelligence tools. This study made use of two popular meta-heuristic algorithms, i.e. GA and PSO, as well as two newer algorithms called IWO and CA for feature selection and demand prediction for dairy products.

In this proposed model, the feature selection problem was first implemented by using various predictive tools and various meta-heuristic algorithms. The results of this section indicate that the indexes such as IN and POPP are very important and have a significant effect on DPD in the future.

In the second part of the numerical results, 12 different hybrid artificial intelligence and novel meta-heuristic algorithms were used to predict DPD. Applying multiple methods to predict DPD implies that the proposed model has found the best forecasting method and best prediction accuracy. The results showed that IWO can very well compete with PSO, but CA yielded poorer results than other algorithms.

In the future, more effort will be made to improve the planning and implementing of the proposed methodology. One promising direction for improving this research is to use time series neural network to predict the feature value of significant variables and then using hybrid methods to explain the future value of DPD better. However, In order to improve this research, the proposed framework can be developed with other meta-heuristic algorithms, future studies are suggested to employ alternative methods for the framework.

References

Adankon, M. M., & Cheriet, M. (2009). Support vector machine. In Encyclopedia of biometrics (pp. 1303-1308): Springer.

Anand, A., & Suganthi, L. (2018). Hybrid GA-PSO Optimization of Artificial Neural Network for Forecasting Electricity Demand. Energies, 11, 728.

Araromi, D. O., Majekodunmi, O. T., Adeniran, J. A., & Salawudeen, T. O. (2018). Modeling of an activated sludge process for effluent prediction—a comparative study using ANFIS and GLM regression. Environmental Monitoring and Assessment, 190, 495.

Caniato, F., Kalchschmidt *, M., Ronchi, S., Verganti, R., & Zotteri, G. (2005). Clustering customers to forecast demand. Production Planning & Control, 16, 32-43.

Chung, K.-J. (1999). Bounds on the optimum order quantity in a DCF analysis of the EOQ model under trade credit. Journal of Information and Optimization Sciences, 20, 137-148.

Esfahani, H. N., Sobhiyah, M. h., & Yousefi, V. R. (2016). Project Portfolio Selection via Harmony Search Algorithm and Modern Portfolio Theory. Procedia - Social and Behavioral Sciences, 226, 51-58.

Hartzel, K. S., & Wood, C. A. (2017). Factors that affect the improvement of demand forecast accuracy through point-of-sale reporting. European Journal of Operational Research, 260, 171-182.

Hashim, F., Daud, N. N., Ahmad, K., Adnan, J., & Rizman, Z. (2017). Prediction of rainfall based on weather parameter using artificial neural network. Journal of Fundamental and Applied Sciences, 9, 493-502.

Jiang, S., Chin, K.-S., Wang, L., Qu, G., & Tsui, K. L. (2017). Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department. Expert Systems with Applications, 82, 216-230.

Kochak, A., & Sharma, S. (2015). Demand forecasting using neural network for supply chain management. International Journal of Mechanical Engineering and Robotics Research, 4, 96.

Lei, W., Wang, J., Huang, B., & Chen, F. (2016). Research on forecasting the market demand of Sci-tech Service Industry based on improved GMDH algorithm. In Guidance, Navigation and Control Conference (CGNCC), 2016 IEEE Chinese (pp. 2192-2197): IEEE.

Mehrabian, A. R., & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. Ecological informatics, 1, 355-366.

Montajabiha, M., Arshadi Khamseh, A., & Afshar-Nadjafi, B. (2017). A robust algorithm for project portfolio selection problem using real options valuation. International Journal of Managing Projects in Business, 10, 386-403.

Perea, R. G., Poyato, E. C., Montesinos, P., & Díaz, J. A. R. (2018). Optimisation of water demand forecasting by artificial intelligence with short data sets. Biosystems Engineering.

Rangel, H. R., Puig, V., Farias, R. L., & Flores, J. J. (2017). Short-term demand forecast using a bank of neural network models trained using genetic algorithms for the optimal management of drinking water networks. Journal of Hydroinformatics, 19, 1-16.

Reynolds, R. G. (1994). An introduction to cultural algorithms. In Proceedings of the third annual conference on evolutionary programming (pp. 131-139): World Scientific.

Sherrod, P. H. (2008). DTREG predictive modeling software. Users Manual. Disponível online no url:(http://www. dtreg. com/DTREG. pdf)(Acedido 28 Março 2014).

Slimani, I., El Farissi, I., & Achchab, S. (2015). Artificial neural networks for demand forecasting: Application using Moroccan supermarket data. In Intelligent Systems Design and Applications (ISDA), 2015 15th International Conference on (pp. 266-271): IEEE.

Sobhani, J., Najimi, M., Pourkhorshidi, A. R., & Parhizkar, T. (2010). Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models. Construction and Building Materials, 24, 709-718.

Wu, Y.-H., & Shen, H. (2018). Grey-related least squares support vector machine optimization model and its application in predicting natural gas consumption demand. Journal of Computational and Applied Mathematics, 338, 212-220.

Yang, Y., Chen, Y., Wang, Y., Li, C., & Li, L. (2016). Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting. Applied Soft Computing, 49, 663-675.

Yazdi, J. S., Kalantary, F., & Yazdi, H. S. (2012). Prediction of elastic modulus of concrete using support vector committee method. Journal of Materials in Civil Engineering, 25, 9-20.