# Colombian Energy Market: An approach of Anfis and Clustering Techniques to an Optimal Portfolio

Olga Lucía Quintero
Department of Basic Sciences
EAFIT University
Medellín, Colombia
e-mail: oquinte1@eafit.edu.co

Marcela Giraldo Uribe
Department of Basic Sciences
EAFIT University
Medellín, Colombia
e-mail: mgiral41@eafit.edu.co

Alejandro Palacio Vásquez
Department of Basic Sciences
EAFIT University
Medellín, Colombia
e-mail: apalac19@eafit.edu.co

***Abstract*—This paper focuses on the study of a first approach to an optimal portfolio in the Colombian Energy Market using Artificial Intelligence. Specifically, ANFIS and Clustering techniques are applied. The methodology is implemented using the Matlab Toolboxes for clustering and FIS generation. Te results are presented, as well as the analysis of them. A first approximation to an optimal portfolio obtained with this methodology is shown. Consequently, some conclusions of the different techniques available for the same purpose are discussed. Finally the future work is proposed.**

***Index Terms*—Energy Markets, Artificial Intelligence, Fuzzy Modeling, Neural Networks, ANFIS.**

## I. INTRODUCTION

In the world of economics and finance, the term market means the aggregate or set of possible buyers and sellers of a certain good or service and the transactions between them. The term market is sometimes used for what are more strictly exchanges, organizations that facilitate the trade in financial securities, for example, a stock exchange or commodity exchange. This may be a physical location or an electronic system. Most trading of stocks takes place on an exchange; nevertheless, corporate actions are outside an exchange, while any two companies or people, for whatever reason, may agree to sell stock from the one to the other without using an exchange [1].

Based on the concept of market, comes a new concept: the financial market. A financial market is a market in which people and entities can trade financial securities, commodities, and other fungible items of value at low transaction costs and at prices that reflect supply and demand. Securities include stocks and bonds, and commodities include precious metals, agricultural goods and energy. There are general markets (where many commodities are traded), and specialized markets (where only one commodity is traded). Markets work by placing many interested buyers and sellers; including households, firms, and government agencies, in one same place, thus making it easier for them to find each other [2]. An economy which relies primarily on interactions between buyers and sellers to allocate resources is known as a market economy, in contrast either to a command economy or to a non-market economy. In finance, depending on the financial markets, it facilitates:

- The raising of capital.
- The transfer of risk.
- Price discovery.
- Global transactions with integration of financial markets.
- The transfer of liquidity.
- International trade.

As mentioned before, there are several types of financial markets depending on the financial asset that is traded on it. Also, the type of the financial market is given depending on the level of the three most important variables that an investor should quantify: yield, risk and liquidity. The most important financial markets are, among others:

- Capital markets which consist of stock markets, and bond markets. Stock markets provide financing through the issuance of shares or common stock, and enable the subsequent trading thereof. Bond markets, which provide financing through the issuance of bonds, and enable the subsequent trading thereof.
- Commodity markets, which facilitate the trading of commodities.
- Money markets, which provide short term debt financing and investment.
- Derivatives markets, which provide instruments for the management of financial risk.
- Futures markets, which provide standardized forward contracts for trading products at some future date (see also forward market).
- Insurance markets, which facilitate the redistribution of various risks.

- Foreign exchange markets, which facilitate the trading of foreign exchange.

Promptly, energy markets are commodity markets that deal specifically with the trade and supply of energy and that trade in the Energy Sector. The energy sector is a category of stocks that are related to producing or supplying energy. This sector includes companies involved in the exploration and development of oil or gas reserves, oil and gas drilling, or integrated power firms [3].

Energy market refers to an electricity market where electricity (both power and energy) is a commodity capable of being bought, sold and traded. An electricity market is a system for effecting purchases, through bids to buy and offers to sell. Bids and offers use supply and demand principles to set the price [3].

In Colombia, in order to get the final rate of the energy, the asset must go through four processes, which are: generation, transmission, distribution, and commercialization. The process of generating electric energy in Colombia, has very specific technical as well as economical characteristics which make the market behave as an oligopoly. Some of these characteristics are: high costs associated to the installation of new plants, long construction periods, restrictions when transporting the energy, and impossibility to store the energy in efficient quantities, among others[4]. As it can be seen in table I, the 86% of the generation of electric energy in the country, was focused on just 6 agents, among the 44 agents that trade on stock.

| Agent | Share |
|---|---|
| EPM | 25.8% |
| EMGESA | 22% |
| ISAGEN | 16% |
| GECELCA | 9% |
| EPSA - CELSIA | 6% |
| AES Chivor | 7.7% |

Table I
SHARE OF THE TOTAL GENERATED ENERGY OF COLOMBIA FOR EACH OF THE BIGGEST AGENTS IN 2012[4]

The technology used to generate electricity in Colombia is as well crucial when determining the prices, because it focuses mainly in hydraulic technology (64% against a 30% from thermoelectric plants). The principal types of plants that exist in Colombia for generating energy are:

1) Hydraulic plants.
2) Thermal power plants.
3) Smaller plants (less than 20 MW).

All the energy produced in the country is traded in the *Mercado Mayorista de Electricidad en Colombia* (MEM), all the generator companies (agents) are linked to the *Sistema Interconectado Nacional* (SIN), who is in charge of satisfying the demand of all the final users who are connected as well to this system. These transactions are made daily through an auction made by the *Administrador del Sistema de Intercambios Comerciales* (ASIC), who acts on behalf of the final consumers[5].

Concerning the plants that are bigger than 20 MW, each agent must present daily its available capacity for each hour of the following day and a selling price for the same day, to the *Centro Nacional de Despacho* (CND). This has to be done for each one of its resources. The CND is in charge of making the “economical dispatch”, which consists of sorting from lowest to highest the generation plants according to the price, until they reach the demand[5].

The smaller plants provide smaller selling price, but they do not present it to the CND, so they are all included in the “real dispatch” which is a variation of the “economical dispatch” after adding the smaller plants to the list. The operation and dispatch is done by XM S.A E.S.P.

The electrical sector in Colombia is controlled by the *Comisión de Regulación de Energía y Gas* (CREG), *Unidad de Planeación Minero Energética* (UPME), and the *Superintendencia de Servicios Públicos Domiciliarios* (SSPD)[4].

Now, reviewing the background of general techniques of artificial intelligence (such as fuzzy logic, neural networks or a hybrid between the two named ANFIS) applications to energy markets, especially optimal investment portfolios, there is no specific application in this field. However, there are different applications of artificial intelligence techniques focused on solving problems of financial markets.

Reviewing the available papers, publications and past studies where an ANFIS structure is applied to the problem of finding an optimal portfolio, specifically in the energy market, no such job has been done. Nevertheless, there are some interesting works related with it and that prove the excellent development of ANFIS techniques in this field and generally in the process of forecasting prices and valuing assets.

In [6], the authors suggest stock portfolio optimization using the combination of Adaptive Neuro-Fuzzy Inference System (ANFIS) and Capital Asset Pricing Model (CAPM). Stock portfolio optimization aims to determine which of the stocks should be added to a portfolio based on the investor’s needs, changing economic and market conditions. The ANFIS is used to take decisions for forecasting the stock’s price using historical data and some technical indicators. CAPM has been incorporated for portfolio optimization, particularly to find the combination of stocks to offer an investor trade-off between expected return and risk of a portfolio. ANFIS-CAPM plays a decisive role in discovering portfolio strategies for investors and creates the optimal portfolio from a combination of stocks. Finally, with some experimental results the authors show that the proposed hybrid intelligent system ANFIS-CAPM yields better performance than existing portfolio models. This work is very interesting since it probes that ANFIS technique can be used in combination with one of the most recognized pricing models, and it provides great results in the field of finding an optimal portfolio.

Another really important field in the Asset Management or generally in the markets is the concept of forecast or forecasting. The authors in [7] propose a hybrid ANFIS based on n-period moving average model to forecast the price of a particular stock. The proposed model is verified by root mean squared error (RMSE) and they use a ten-year period of the stock selected as experiment datasets. The results show that the proposed model is superior to the other forecasting models, proving once again the efficiency and effectiveness of the ANFIS technique in this field. Other really important subject in corporate finance or finance generally is the valuation of assets and pricing of an asset. In [8] the authors do a house selling price assessment using two different adaptive Neuro-fuzzy techniques: ANFIS with grid partition (ANFIS-GP) and ANFIS with sub clustering (ANFIS-SC). The results were satisfactory. A comparison made indicated that the ANFIS-GP models performed better than the ANFIS-SC models and the conclusion is that the ANFIS-GP technique can be successfully used in the estimation of prices, or in the process of pricing assets. The last topic that should be covered when making investments and evaluating an optimal portfolio is the Risk involved. The authors in [9] present a technique capable of enhancing risk-adjusted performance of stock market intraday trading, the technique off course, uses adaptive Neuro-Fuzzy systems. The conclusion is that combining multiple risk-adjusted objective functions using an ANFIS ensemble yields promising results, and that in general, the ANFIS technique is a good tool for managing risks in all levels of investments.

Finally, since this work is fully developed on the energy market which is a commodity market, there are some authors that have indeed used ANFIS technique to determine the price of other commodities. For example, authors in [10] made a study on crude oil prices modeled by Neuro-fuzzy networks. They prove that the Neuro-fuzzy approach based on ANFIS networks compare favorably with respect to other standard and neural models, and it is able to achieve useful performances in terms of accurate prediction of prices and their probability distribution. This shows that ANFIS is an excellent tool to be used in the general commodity market.

Therefore, since ANFIS is an effective method for these types of problems, and it has not been used yet for them, the objective of this paper is to apply the theory and techniques of Artificial Intelligence, to build an optimal portfolio for the Colombian Energy Market, representing a game of strategies. In particular, special attention is payed to the Adaptive Neuro Fuzzy Inference System (ANFIS), which is the hybrid of two other techniques of Artificial Intelligence, namely Fuzzy Modeling and Neural Networks. The study will focus on the first part of the process defined for Colombia, it means, generation; and taking into consideration the biggest market, which can in a first approximation, represent as well the behavior of the other markets. This market is EPM (*Empresas Públicas de Medellín*).

The paper is organized as follows: in section 2 the definition of the problem that will be studied is defined, section 3 explains the two main techniques that will be used: clustering and ANFIS, the methodology to follow is presented in section 3, giving a brief introduction to the elements necessary for the implementation of these techniques in Matlab. All the results obtained by implementing and developing the methodology, as well as the pertinent discussions and analysis, are exposed in Section 4. Finally, in section 5 the conclusions are given and some future work is proposed. At the end the references consulted are presented.

## II. Problem Definition

The methodology is applied on the basis of an optimal portfolio for the Colombian Energy Market. An *optimal portfolio* is the one that minimizes the risk of the agent, while striving for the highest return possible, or said in other words, maximizing its utility. The theory states that investors will act rationally, always making decisions aimed at maximizing their return for their acceptable level of risk [11].

*1) Inputs and Outputs:* In order to evaluate de network, we defined the following inputs:

$u_1 =$ Marginal price of the system in $t$.
$u_2 =$ Price of the resource 1 (hydraulic) in $t$.
$u_3 =$ Price of the resource 2 (thermal power) in $t$.
$u_4 =$ Availability of resource 1 in $t$.
$u_5 =$ Availability of resource 2 in $t$.

Whereas the output is an objective function, which should be maximized in a future work:

$$J = (u_1 - u_2) \times u_4 + (u_1 - u_3) \times u_5$$

## III. CLUSTERING AND ANFIS TECHNIQUES

### A. Clustering

Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group, or cluster, are more similar to each other than to those in other groups or clusters. It is a main task of exploratory data mining, and a common technique for statistical data analysis [12]. Clustering is used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics. Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them [13]. Some notions of clusters include groups with small distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem.

The appropriate clustering algorithm and parameter settings depend on the individual data set and intended use of the results. Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. It will often be necessary to modify data preprocessing and model parameters until the result achieves the desired properties. Besides the term clustering, there are a number of terms with similar meanings, including automatic classification, numerical taxonomy, botryology and typological analysis. The subtle differences are often in the usage of the results: while in data mining, the resulting groups are the matter of interest, in automatic classification the resulting discriminative power is of interest. This often leads to misunderstandings between researchers coming from the fields of data mining and machine learning, since they use the same terms and often the same algorithms, but have different goals.

Cluster analysis was originated in anthropology by Driver and Kroeber in 1932 and introduced to psychology by Zubin in 1938 and Robert Tryon in 1939, and famously used by Cattell in 1943 for trait theory classification in personality psychology.

As it has been mentioned, there are several algorithms of clustering with several objectives and with different properties that make them work different depending on the problem [14]. The most representative off-line clustering techniques presented in [13] are:

- K-means (or Hard C-means) Clustering.
- Fuzzy C-means Clustering.
- Mountain Clustering.
- Subtractive Clustering.
- Partition Simplification Fuzzy C-means Clustering.

These approaches solve the problem of categorizing data by partitioning a data set into a number of clusters based on some similarity measure so that the similarity in each cluster is larger than among clusters. For more information about the presented clustering techniques or algorithms please check [14], [12], [13].

### B. Adaptive Neuro Fuzzy Inference System (ANFIS)

The Adaptive Neuro Fuzzy Inference System ($ANFIS$) is a kind of neural network that is based on Takagi–Sugeno fuzzy inference system. Since it integrates both neural networks and fuzzy logic principles, it has potential to capture the benefits of both in a single framework. Its inference system corresponds to a set of fuzzy "$If-Then$" rules that have learning capability to approximate nonlinear functions. Hence, ANFIS is considered to be a universal estimator [15].

First of all, an explanation of the Sugeno Fuzzy model is presented. It was proposed by Takagi Sugeno in an effort to formalize a systemic approach to generating fuzzy rules from an input-output data set. A fuzzy rule in a Sugeno fuzzy model has the format:

$If\ x\ is\ A\ and\ y\ is\ B\ then\ z = f(x,y)$

Where $A$ and $B$ are fuzzy sets in the antecedent; $z = f(x,y)$ is a crisp function in the consequent. Usually it is a polynomial in the input variables $x$ and $y$, but it can be any other functions that can that can appropriately describe the output of the system within the fuzzy region specified by the antecedent rule. If $f(x,y)$ is a first order polynomial, it is a **first order** Sugeno fuzzy model. When $f$ is a constant, it is the **zero order** Sugeno fuzzy model, which can be viewed either as a special case of the Mamdami fuzzy inference system or as a special case of Tukamoto fuzzy model [16].

To formalize an ANFIS structure, consider a first-order Sugeno fuzzy inference system which contains two rules:

1) If $X$ is $A_1$ and $Y$ is $B_1$, then $f_1 = p_1x + q_1y + r_1$
2) If $X$ is $A_2$ and $Y$ is $B_2$, then $f_2 = p_2x + q_2y + r_2$

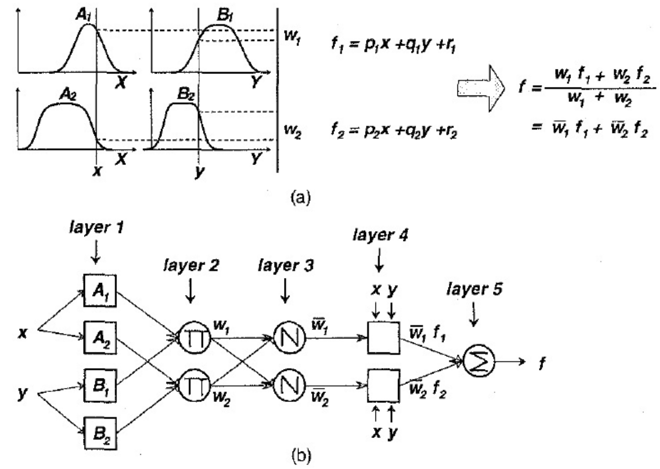


Figure 1. (a) Sugeno Fuzzy Model; (b) Corresponding ANFIS Architecture

To facilitate the learning or adaptation of the Sugeno fuzzy model, it is inserted into the framework of adaptive networks that can compute gradient vectors systematically. The resultant network architecture is called ANFIS, where nodes within the same layer perform functions of the same type described as follows:

- Layer 1: Generates a membership grade of a linguistic label.
- Layer 2: Calculates the firing strength of a rule w via multiplication.
- Layer 3: Calculates the ratio of the i-th rule's firing strength.
- Layer 4: Computes the contribution of i-th rule toward the overall output.
- Layer 5: computes the overall output as the summation of contribution from each rule.

In the figure 1, (a) illustrates graphically the fuzzy reasoning mechanism to derive an output $f$ from a given input vector $[x, y]$, while (b) shows the Adaptive Neuro Fuzzy Inference System (ANFIS) structure presented above.

Recently, the training of all the parameters in the ANFIS structure has become the main problem to solve. That is why different heuristic methods have been proposed and studied, concluding that the results are as good as the expected. For example, in [17], the authors present a training ANFIS structure with the modified PSO (Particle Swarm Optimization) heuristic algorithm. The results showed are quite satisfactory and prove the functionality of this specific algorithm in this task.

### C. METHODOLOGY

Data from 2250 days is available, from the different Colombian agents and resources. The implementation is done for the case of EPM (*Empresas Públicas de Medellín*), for the sake of simplicity.

The methodology is implemented and applied in the software Matlab, taking advantage of the complete toolboxes and plotting tools that it possesses. The methodology applied is quite simple and understandable. It is divided into two parts:

the Clustering techniques and the ANFIS techniques. The results of both parts are discussed in the next section.

*1) Clustering:* First the subtractive clustering method is used to determine the cluster centers, or in other words, the total number of clusters in the data set. The subtractive clustering method assumes each data point is a potential cluster center and calculates a measure of the likelihood that each data point would define the cluster center, based on the density of surrounding data points. The algorithm does the following:

1) Selects the data point with the highest potential to be the first cluster center.
2) Removes all data points in the vicinity of the first cluster center in order to determine the next data cluster and its center location.
3) Iterates on this process until all of the data is within the ratio of a cluster center.

It is valid to say that the subtractive clustering method is an extension of the mountain clustering method.

After the number of clusters is established, the Fuzzy C-Means clustering method is used to obtain the matrix of final cluster, specifically, the center coordinates. The Fuzzy c-means (FCM) is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. This technique was originally introduced by Jim Bezdek in 1981 as an improvement on earlier clustering methods. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters. The algorithm starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster, the initial guess for these cluster centers is most likely incorrect. Then it assigns every data point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point, it moves the cluster centers to the right location within a data set. This iteration is based on minimizing an objective function that represents the distance from any given data point to a cluster center weighted by that data point's membership grade.

*2) ANFIS:* For the ANFIS part, first, three algorithms can be used to generate a Sugeno-type FIS structure used as initial conditions (initialization of the membership function parameters) for ANFIS training.

The first method (function *genfis1* in Matlab) generates a single-output Sugeno-type fuzzy inference system using a grid partition on the data. However, this method is not shown in this paper, due to the poor results it provides.

The second one, generates a Sugeno-type FIS structure using Subtractive Clustering (in Matlab, *genfis2*) and requires separate sets of input and output data as input arguments. When there is only one output, this method may be used to generate an initial FIS for ANFIS training. This is accomplished by extracting a set of rules that models the data behavior. The rule extraction method first uses the Subtractive Clustering method to determine the number of rules and antecedent membership functions and then uses linear Least-Squares estimation to determine each rule's consequent equations. This function returns a FIS structure that contains a set of fuzzy rules to cover the feature space.

The third method generates a Sugeno FIS using Fuzzy C-Means clustering (FCM) by extracting a set of rules that models the data behavior (it is called *genfis3* in Matlab). It requires separate sets of input and output data as input arguments. When there is only one output, this method can be used to generate an initial FIS for ANFIS training. The rule extraction method first uses the FCM function to determine the number of rules and membership functions for the antecedents and consequents. This function allows to make as well a Mamdani FIS using FCM, however, it cannot be used as an initial FIS for the ANFIS.

Afterwards, by performing fuzzy inference calculations to the three methods described it is possible to make tangible the results and compare them with the real output of the system.

Finally after obtaining the Sugeno-type FIS structure with any of the three methods described it is possible to implement and use the ANFIS structure. It is the major training routine for Sugeno-type fuzzy inference systems. ANFIS uses a hybrid learning algorithm to identify parameters of Sugeno-type fuzzy inference systems and it applies a combination of the Least-Squares method and the back-propagation gradient descent method for training FIS membership function parameters to emulate a given training data set. The Sugeno-type systems used in the ANFIS structure must have the following properties:

- Be first or zeroth order Sugeno-type systems.
- Have a single output, obtained using weighted average defuzzification. All output membership functions must be the same type and either be linear or constant.
- Have no rule sharing. Different rules cannot share the same output membership function, namely the number of output membership functions must be equal to the number of rules.
- Have unity weight for each rule.

The ANFIS technique is used to estimate the output of the system under study using the inputs and any of the FIS structures obtained with the methods described. The results are presented and analyzed in the next section.

## IV. RESULTS

### A. Clustering

Following the methodology proposed in previous sections, for the clustering section, first the Subtractive Clustering method is used to find the number of clusters that better suit the problem. Then, the Fuzzy C-Means (FCM) technique is used to obtain the coordinates of these centers and an improve the results.

*1) Subtractive Clustering:* The subtractive clustering technique works as defined on previous sections. The subtractive clustering method (in Matlab *subclust*) assumes each data point is a potential cluster center and calculates a measure of the likelihood that each data point would define the cluster center, based on the density of surrounding data points. The function parameters are as follow:

- Data to be clustered.
- Vector of entries between 0 and 1 that specifies a cluster center's range of influence in each of the data dimensions.

| Center 1 | Center 2 | Center 3 | Center 4 | Center 5 | Center 6 |
|---|---|---|---|---|---|
| 0.1918 | 0.1469 | 0.1995 | 0.1837 | 0.1634 | 0.3025 |
| 0.0959 | 0.1686 | 0.1230 | 0.2179 | 0.0541 | 0.1934 |
| 0.0925 | 0.1114 | 0.0924 | 0.1068 | 0.0784 | 0.0045 |
| 0.9299 | 0.9233 | 0.9061 | 0.8010 | 0.4781 | 0.8101 |
| 0.6791 | 0.9412 | 0.3529 | 0.9412 | 0.6791 | 0.6516 |

Table II
CENTERS PROVIDED BY *subclust*

| Center 1 | Center 2 | Center 3 | Center 4 | Center 5 | Center 6 |
|---|---|---|---|---|---|
| 0.2660 | 0.5164 | 0.3339 | 0.1837 | 0.2119 | 0.1658 |
| 0.1774 | 0.4962 | 0.2002 | 0.1555 | 0.1277 | 0.1225 |
| 0.1028 | 0.0463 | 0.0900 | 0.1274 | 0.0978 | 0.1252 |
| 0.8615 | 0.8086 | 0.8466 | 0.8849 | 0.8940 | 0.9147 |
| 0.6626 | 0.3348 | 0.6456 | 0.9429 | 0.3298 | 0.6727 |

Table III
CENTERS PROVIDED BY *fcm*

- Vector specifying clustering algorithm parameters to override the default values.

The function returns the cluster center (the number of clusters) in a Matrix C. Each row of C contains the position of a cluster center. Also the function returns a vector “S” which contains the sigma values that specify the range of influence of a cluster center in each of the data dimensions. This function gives us the number of clusters so that we can use it as a parameter on the next function, Fuzzy c-means, which will be explained below.

The mentioned function provided 6 centers of the data, which are shown in table II.

*2) Fuzzy c-means (FCM):* The Fuzzy C-Means clustering technique (in Matlab, *fcm*) works as described on previous sections. The function starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster. The initial guess for these cluster centers is most likely incorrect. The function also assigns every data point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point it moves the cluster centers to the right location within a data set. This iteration is based on minimizing an objective function that represents the distance from any given data point to a cluster center weighted by that data point’s membership grade. The function parameters are as follow:

- Data set to be clustered.
- Number of clusters (in this approach, the one given by the function *subclust*).

The function returns the matrix of final cluster centers where each row provides the center coordinates, the final fuzzy partition matrix (or membership function matrix) and the values of the objective function during iterations. As mentioned above, using the number of cluster given by the subclust function as a parameter, the FCM returns much better coordinates of this centers so that it is much easier to build up the clusters. In other words, once the exact number of clusters is known, given by subtractive clustering, FCM gives much better approximations of the centers. The centers provided by *fcm*, are shown in table III.

*3) Analysis of the clustering techniques:* The result of using the Subtractive Clustering method established that on the data under study, there are 6 clusters and when applying this result to the Fuzzy C-Means (FCM) better coordinates of the centers of those clusters are obtained. Not losing the point of this work let us remember that what is studied, is the price of some resources that together will establish the change of the price the next day. The data then is a matrix full of prices that are observed on the same time point. Analyzing clusters with this idea, is quite helpful since according to the theory, elements that belong to a same cluster are most likely to have the same properties between them than with elements that belong to other clusters. Having six clusters may be a sign that there are six moments in the time (remember the observations are daily) in which the prices of the resources were highly similar between them and have a high level of correlation between them meaning that if the price of one resource goes up then most likely the other resources price would also rise. The existence of clusters prove that there are periods were the prices were highly similar between them helping the company under study to establish price dynamics and be able to estimate the price of one resource based on the price of other resources. Also this information can be used to avoid eventual problems since if there is a period were one resource price decrease for instance, the company must be able to make other resources to increase their prices in order to break a possible future “wrong” cluster.

### *B. ANFIS*

*1) Sugeno FIS using Subtractive Clustering:* The second experiment consisted of testing the ANFIS structure for the data using the Sugeno type FIS generated by the function *genfis2*. This function, as explained on section3, generates a Sugeno-type FIS structure using subtractive clustering and requires separate sets of input and output data as input arguments. When there is only one output, *genfis2* may be used to generate an initial FIS for ANFIS training. The arguments for *genfis2* are as follows:

- The data inputs.
- The data outputs.
- The cluster center’s range of influence in each of the data dimensions.

Figure 2 shows the obtained FIS system, figure 3 presents the membership functions, obtained for the inputs, and figure 4 plots the real output against the output obtained with the mentioned function. It can be seen that the output provided by Sugeno FIS using Subtractive Clustering is very close to the real output, so this is a good option for a FIS.

*2) Mamdani - Sugeno FIS using FCM:* *genfis3*, as explained on section 3, generates a Mamdani or Sugeno-type FIS structure using Fuzzy C-Means (FCM) clustering by extracting a set of rules that models the data behavior. When there is only one output, and Sugeno type is chosen, *genfis3* may be used to generate an initial FIS for ANFIS training. The arguments for *genfis3* are as follows:

- The data inputs.
- The data outputs.
- The input and output membership function type.
- The number of clusters.

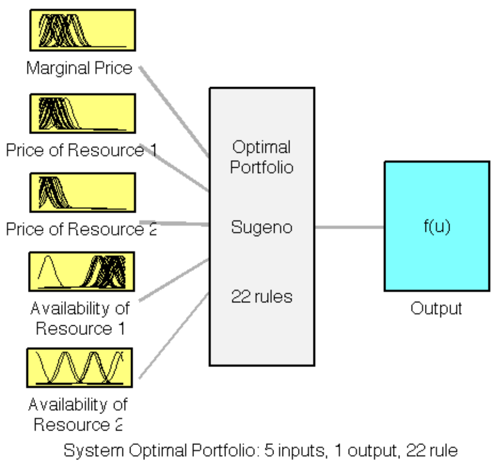


Figure 2. Sugeno FIS using Subtractive Clustering

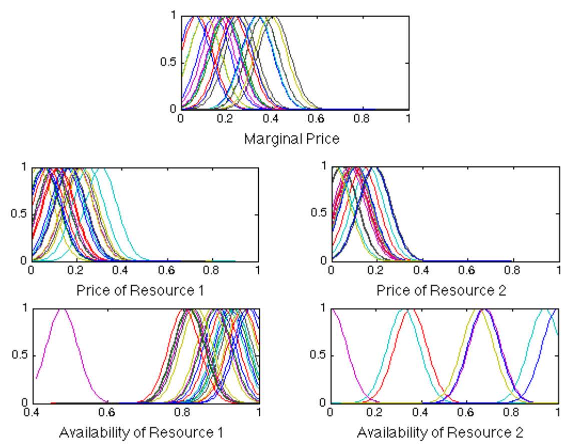


Figure 3. Membership functions of the inputs - Sugeno FIS using Subtractive Clustering

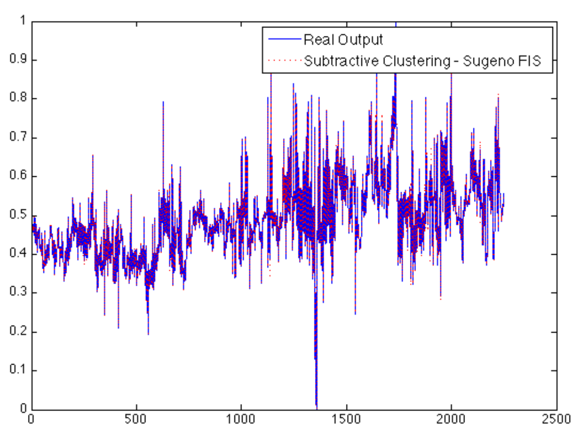


Figure 4. Real Output Vs. Subtractive Clustering - Sugeno FIS Output

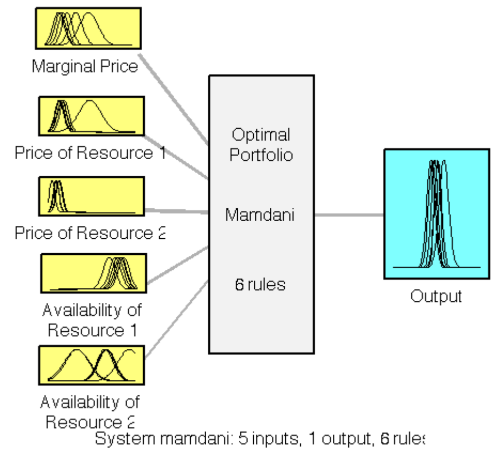


Figure 5. Mamdani FIS using FCM

This function gives the user two options: either get a Mamdani FIS or a Sugeno FIS, both using FCM. Results are shown for both cases.

The results of the first experiment, a Mamdani FIS using FCM are obtained after performing fuzzy inference on the generated FIS (figures 5 and 6) and are presented in figure 7, where it can be seen the real output against the one estimated with this generated Mamdani FIS. It can be seen that the Mamdani FIS using FCM tries to follow the movements of the real output, however it does it in a bad way since although at some points it follows correctly the movements, at some others it is completely out of phase and the error is quite high. Concluding that this type of Mamdani FIS doesn't works correctly on our data.

Nevertheless when changing the generated Mamdani FIS to a Sugeno-type FIS, the results are much better. The same results are presented on figures 8, 9 and 10, and it can be seen that they are quite satisfactory. This method follows correctly the real output movements and the error within is really low. The results in this experiment are quite satisfactory and allow to conclude that it is possible to work in this problem with an ANFIS structure and a Sugeno FIS using FCM as an initial FIS, as will be shown below.

*3) Anfis:* After performing fuzzy inference on the Sugeno FIS using Subtractive Clustering, the results were good enough to be tested on the ANFIS structure, as well as the Sugeno FIS using FCM. This is the reason why, in order to evaluate the ANFIS, two experiments were defined:

- Experiment 1: ANFIS with Sugeno FIS using Subtractive Clustering as initial FIS.
- Experiment 2: ANFIS with Sugeno FIS using FCM as initial FIS.

The function *anfis* is the training routine for Sugeno-type Fuzzy Inference System and works as explained on section 3. The arguments for anfis are as follows:

- The training data set (inputs, outputs).
- The fuzzy inference system (FIS) used to provide ANFIS with an initial set of membership functions for training. In this case is the one generated by *genfis2*.

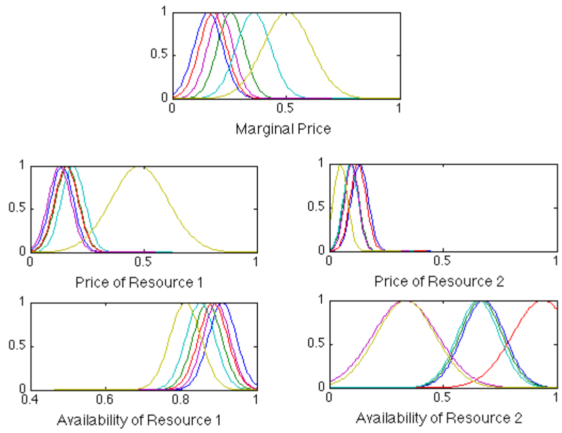


Figure 6. Membership functions of the inputs - Mamdani FIS using FCM

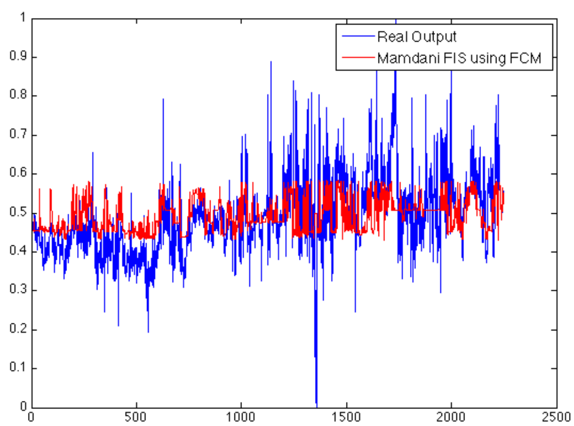


Figure 7. Real Output Vs. Mamdani FIS using FCM Output

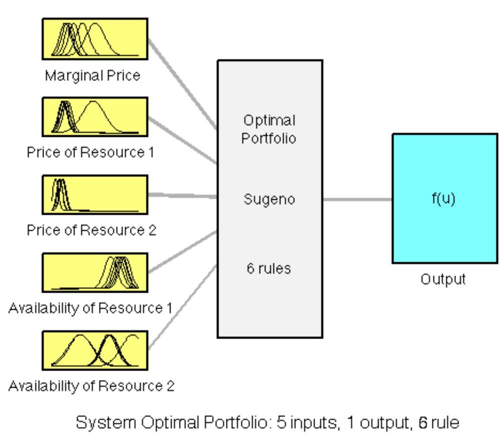


Figure 8. Sugeno FIS using FCM

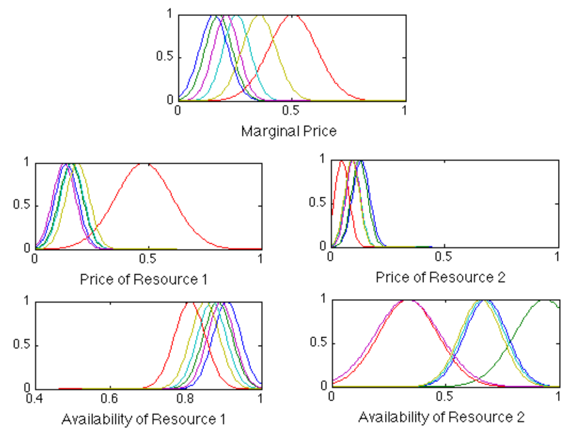


Figure 9. Membership functions of the inputs - Sugeno FIS using FCM

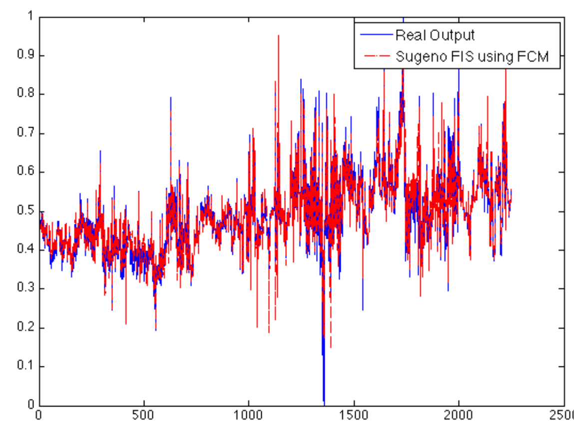


Figure 10. Real Output Vs. Sugeno FIS using FCM Output

- The number of membership functions.

The results of these experiments are presented on figures 13 and 16. Figure 13 shows the real output plotted against the output provided by the ANFIS of experiment 1, and the output of the Sugeno FIS using Subtractive Clustering. Analogously, figure 16 shows the same plot but for experiment 2. Clearly the results of both algorithms follow correctly the real output movements and the error within is really low. The results in these experiments are quite satisfactory and may allow to conclude that it is possible to work this kind of problems with the ANFIS with either *genfis2* or *genfis3* (but with a Sugeno FIS) as initial solutions. In order to show the systems obtained for each experiment, and the form of the membership functions, figures 11, 14, 12 and 15 are presented as well.

*4) General Analysis:* Table 2 IV exposes a comparison of the errors obtained when using any of those techniques. The error was computed with the following formula:

$$E = \frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}_i| \tag{1}$$

Where $N$ is the number of patterns, $y_i$ is the real output,

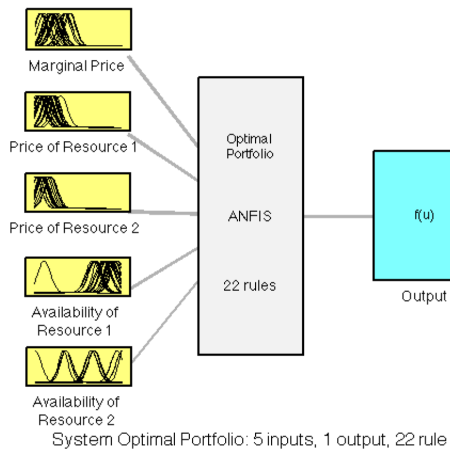


Figure 11. ANFIS - Experiment 1

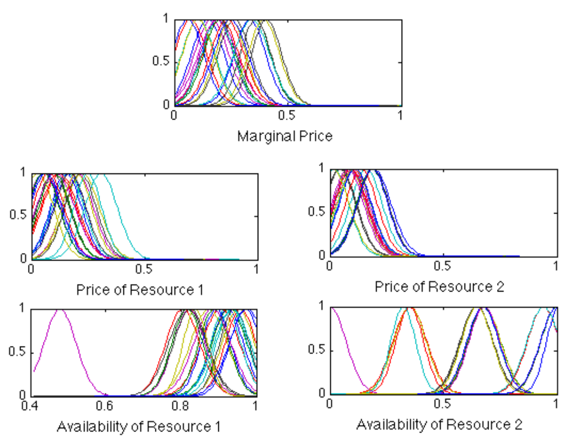


Figure 12. Membership functions of the inputs - ANFIS - Experiment 1

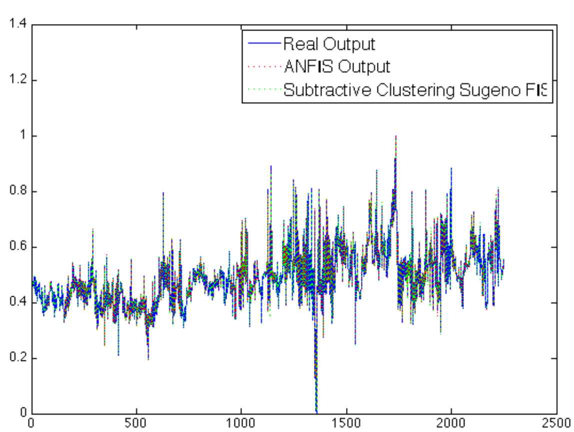


Figure 13. Real Output Vs. ANFIS Output - Experiment 1

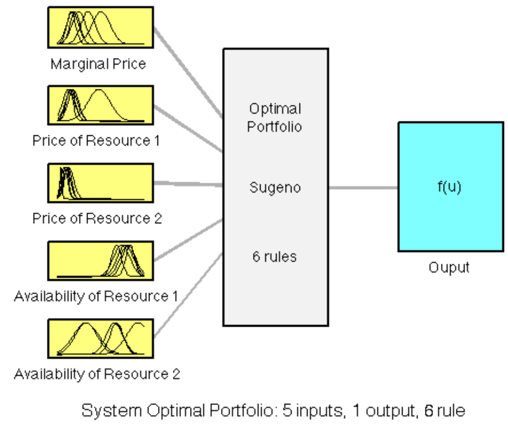


Figure 14. ANFIS - Experiment 2

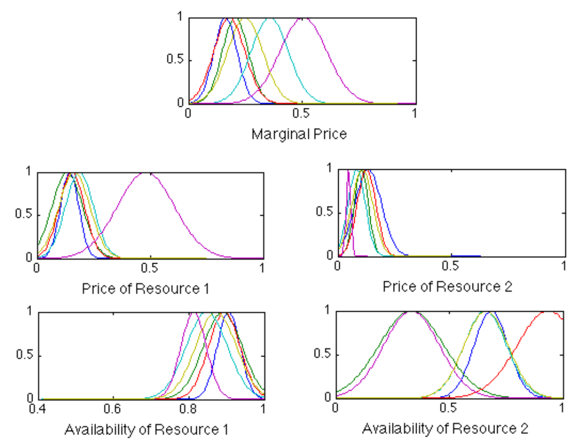


Figure 15. Membership functions of the inputs - ANFIS - Experiment 2

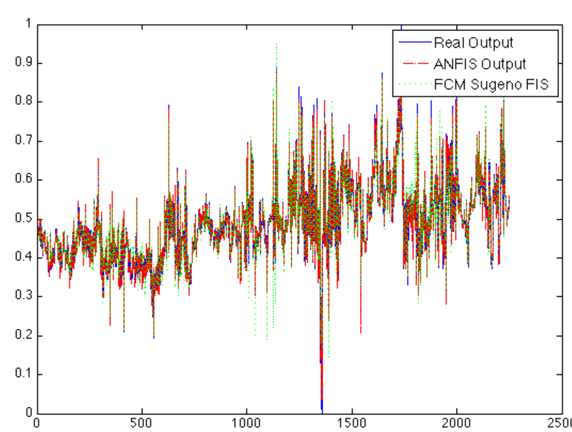


Figure 16. Real Output Vs. ANFIS Output - Experiment 2

| FIS | Error |
|---|---|
| Sugeno FIS using Subtractive Clustering | 0.0020 |
| Mamdani FIS using FCM | 0.0656 |
| Sugeno FIS using FCM | 0.0163 |
| ANFIS with Sugeno FIS using Subtractive Clustering | 0.0016 |
| ANFIS with Sugeno FIS using FCM | 0.0047 |

Table IV
COMPARISON OF ERRORS

and $\bar{y}_i$ is the estimated output.

All the initial FIS that were considered and that were built with the different functions presented in past sections (Sugeno FIS using Subtractive Clustering, Sugeno FIS using FCM and Mamdani FIS using FCM) work correctly on the task of estimating the real output from our problem after performing fuzzy inference on them.

The worst FIS for this approach is Sugeno FIS (in Matlab *genfis1*), however it is not included in this methodology and analysis due to the bad performance it shows. A reason could be that the other two functions, use Subtractive Clustering or FCM to find the clusters of the data, while the Sugeno FIS uses only the Sugeno structure. In consequence, *genfis1* is not recommended, for approaches similar to this one. The second worst result was obtained with Mamdani FIS using FCM, as it can be seen in the mentioned table. From figure 7, it can be observed that the estimated output tries to follow the real one, and in some cases it successes, but in most of the cases, it does not make it. In general, Sugeno structures had a better result (figure 10), for example Sugeno FIS using FCM, which is the exact same function as for Mamdani FIS using FCM, but the option *'sugeno'* is chosen in the function's parameters, instead of *'mamdani'*. The best error is provided by Sugeno FIS using Subtractive Clustering (figure 4). This is the reason why it was the first option to provide as initial FIS for the ANFIS, and as it was expected, the resulting ANFIS provided the best performance (an error of 0.0016), and it can be seen graphically in figure 13. Finally, an ANFIS with Sugeno FIS using FCM is also used, just to see the result and how the error changed. It showed not bad results (figure 16), with an error of 0.0047, which is better than the error of its respective initial FIS (0.0163).

When selecting an appropriate initial FIS, the most recommended one is the Sugeno FIS using Subtractive Clustering, since it throws the best results among all the experiments.

## V. CONCLUSIONS

Generally, determining an optimal portfolio is a really interesting subject in many fields and since there are many researchers trying to do it, the artificial intelligence techniques worked on this paper are a really solid and helpful tools to do so. The main objective of this paper was to help into the procedure of applying correctly these techniques to obtain a first approach to an optimal portfolio in Colombian energy market and to find interesting results during the process.

In the literature, there are no other authors who tried to apply artificial intelligence techniques (specifically clustering and ANFIS) to solve the problem of finding an optimal portfolio in an energy market, giving this work a huge added value in this topic and making it totally original.

Using Subtractive Clustering for the available data, six centers were obtained. And after knowing this, the *fcm* function provided more appropriate centers to work with.

The implementation of the ANFIS and FIS structures was made satisfactorily as well. The results of all the experiments presented previously demonstrated the functionality of this tool in this problem. On those experiments where the results were not good enough it can be justified by the Sugeno-type FIS generator, when it was by a grid partition the results were really poor while in the other two cases subtractive clustering and Fuzzy C-Means (using Sugeno not Mamdani) followed correctly the real output and worked as excellent estimators. In this cases the ANFIS structure as well worked good on the estimation of the output. The effects were captured correctly and the error was almost null.

We propose thus, as a first step to select an Optimal Portfolio for the Colombian Energy Market, an ANFIS structure, with Sugeno FIS using Subtractive Clustering as initial FIS, since it provided, in general, the best results.

It can be concluded that artificial intelligence tools and techniques are quite innovator and useful in the task of solving problems of the real world. In this case, FIS, Neural Networks and ANFIS structures work correctly on the task of estimating an optimization function, as the one defined in this work. These techniques are as well excellent tools to be explored and used to solve new problems. As future work, first it can be proposed to give a little plus to the ANFIS structure by using for example the PSO heuristic algorithm in the input selection problem or using any other hybridized technique of the artificial intelligence in order to get better results. As a second future work, the optimization function proposed in this paper should be optimized, and its output used as a basis to define an optimal portfolio for the Colombian Energy Market.

## REFERENCES


[1] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann, *Modern Portfolio Theory and Investment Analysis*. Wiley, 2009.

[2] T. E. Copeland, J. F. Weston, and K. Shastri, *Financial Theory and Corporate Policy*. Prentice Hall, 2004.

[3] *Energy Security: Managing Risk in a Dynamic Legal and Regulatory Environment*. Oxford University Press, 2004.

[4] R. A. H. Anfossi, "Determinación de precios vía equilibrio de cournot en el mercado eléctrico spot colombiano," *Trabajo de Grado Presentado como Requisito para Optar por el Título de Maestría en Economía*, 2013.

[5] S. D. Greiff and O. L. Quintero, "Redes neuronales para la estimación del precio ofertado de la energia en hidroeléctricas a filo de agua," 2013.

[6] M. Gunasekaran and K. Ramaswami, "A hybrid intelligent system of anfis and capm for stock portfolio optimization," *Journal of Intelligent and Fuzzy Systems*, vol. 26, pp. 277–286, 2014.

[7] L.-Y. Wei, C.-H. Cheng, and H.-H. Wu, "A hybrid anfis based on n-period moving average model to forecast taiex stock," *Applied Soft Computing Journal*, vol. 19, pp. 86–92, 2014.

[8] I. Gerek, "House selling price assessment using two different adaptive neuro-fuzzy techniques," *Automation in Construction*, vol. 41, pp. 33–39, 2014.

[9] V. Vella and W. Ng, "Enhancing risk-adjusted performance of stock market intraday trading with neuro-fuzzy systems," *Neurocomputing*, 2014.

[10] M. Panella, L. Liparulo, F. Barcellona, and R. D Ecclesia, "A study on crude oil prices modeled by neurofuzzy networks," in *IEEE International Conference on Fuzzy Systems*, 2013.

[11] March 2014.

[12] M. Steinbach, G. Karypis, and V. Kumar, “A comparison of document clustering techniques,” *Department of Computer Science and Engineering University of Minnesota*.

[13] V. S. Moertini, “Introduction to five data clustering algorithms,” 2002.

[14] K. Hammouda, “A comparative study of data clustering techniques,” *University of Waterloo*.

[15] J.-S. R. Jang, “Anfis: adaptive-network-based fuzzy inference system,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, pp. 665–685, 1993.

[16] Jyh, “Input selection for anfis learning,” *Department of Computer Science*.

[17] V. Seydi Ghomsheh, I. A. Univ., Kermanshah, M. Shoorehdeli, and M. Teshnehlab, “Training anfis structure with modified pso algorithm,” *Control & Automation*, pp. 1–6, 2007.