Soft computing's development and full load forcasting of electrical output power of a combined cycle power plant

Hanieh Maleki / Master Student of Artifical Intelligence
Department of Computer Engineering
Islamic Azad University, Central Branch
Tehran, Iran
hanieh.maleki87@gmail.com

Seyyed Mohammad Hossein Dadgar / Master Student of Artifical Intelligence
Department of Computer Engineering
Islamic Azad University, Central Branch
Tehran, Iran
dadgar_1370@yahoo.com

Mostafa Kaardann / Master Student of Software Engineering
Department of Computer Engineering
Islamic Azad University, Central Branch
Tehran, Iran
Mostafa.kaardaan@yahoo.com

Zeinab Sadate Hosseini / Master Student of Artifical Intelligence
Department of Computer Engineering
Islamic Azad University, Central Branch
Tehran, Iran
zshhadis@gmail.com

***Abstract*— soft computing is a result of new scientific efforts that makes modeling, analysis and finally controlling of complex systems possible with more ease and success. Full load forcasting of electrical output power of a power plant is important to maximize benefit. Load operation of a power plant is affected by four main parameters as input including ambient temperature, ambient humidity, ambient pressure and exhaust vapor pressure. Soft computing methods are assessed in this paper in summary then using a body of data's which have been gathered more than 6 years a definite learning machine was designed using descend gradient method to predict load of output power then a probability machine was designed using bayes algorithm in order to increase this prediction. Conducted comparison between definite learning machine and probability learning machine shows that probability learning machine can predict load of output power of this power plant more accurate with 96.89 % efficiency.**

Keywords*— *soft computing, load prediction, definite learning machine, probability learning machine

I. INTRODUCTION

Soft computing, is a word in computer science and is to determine none accurate and estimated solutions to solve issues that their solving is hard calculation wise and there is no known algorithm for their accurate solve in multiple sentence time. Soft computations is a group of computational methods such as Fuzzy logic, artificial neural networks and genetic algorithm. This computational technic in contrast of hard computational methods which use all of their efforts and power to be accurate and are for the sake of complete modeling of fact, is based on acceptance of none accuracy, partial and incomplete facts and absence of reliability. It is as clear as possible comprehending of why, how and also concept of these kind of new computations that put light on new perspectives in future complicated sciences [1].

Applied scientific methods in previous centuries just had been able to model and analyzing of relatively simple systems in mechanic, physic and some of engineering and practical fields. More complicated issues such as systems which are dependent to biological, medical, social science, management science and others like them have been remained away from main and successful scope of mathematical and analytical methods. It is notable that attributes such as simplicity and complexity are relative concepts and definitely most of the past successful mathematical and scientific modelings were also very important and complicated in absolute meaning of the word. Soft computations is a group of new computational methods in computer science, artificial intelligence, machine learning and many other practical fields. All of these fields require study, modeling and analyzing of very complicated phenomenon's which past accurate scientific methods had not been successful for their easy, analytical and complete solvation.

This paper assess soft computational methods in summary and learning machines are used for prediction of a thermodynamic system. This system is a combined cycle power plant with two gas turbine, one vapor turbine and two thermal recycling systems that aims to predict full load electrical output power from it and maximizing of benefit due to megawatt [2].

In continue, this paper assesses methods of soft computations in 2nd section and in 3rd section algorithms of definite and indefinite machines are described and in 4th section the results of implementation of definite and indefinite machines are shown and in 5th section final conclusion is stated.

II. SOFT COMPUTATION METHODS

In general, soft computations includes Fuzzy logic methods, artificial neural network, genetic algorithm and algorithms of machine learning which in following will be assessed in summary.

A. *Fuzzy logic*

At the first time, Professor Lotfizadeh introduced theory of Fuzzy sets and Fuzzy logic within a lecture called "Fuzzy sets, information and control" in 1965[3]. His primary aim at that time was to develop a more efficient model for description of processing procedure of natural languages. He entered concepts and phrases such as Fuzzy sets, fuzzy incidents, Fuzzy numbers and Fuzzification into the mathematical and engineering sciences. Fuzzy logic is one of strongest control methods which is known based on laws and consideration of multiple variants and provides results faster compared to neural network and genetic algorithm [4, 5].

Fuzzy logic is based on fuzzy sets theory. This theory is an extension of classic sets theory in mathematic science. In classic theory of sets an element is either member of set or not. In fact, element's membering follows a zero, one and bayesian pattern but theory of fuzzy sets expands this concept and introduces classified membering in a way that a member can be not a full member of a set in some level. Most interesting application of fuzzy logic is an interpretation which this science provides from decision making's structure of intelligent livings and at top of them of human intelligence.

General structure of a fuzzy logic controller (FLC) or a fuzzy controller (FC) composes of three basic part in summary: 1- fuzzification (a part which is related to inputs) 2- inference engine which is based on fuzzy rules 3- defuzzification (a part which is related to outputs). These three stage are shown in figure 1.

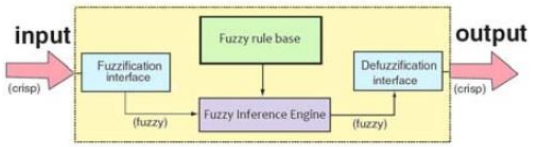

Fig. 1. General structure of a fuzzy logic controller

B. *Artificial neural network*

Artificial Neural Network (ANN) or in a simpler word, neural networks are new computational systems and methods for machine learning, knowledge display and finally applying of obtained knowledge for prediction of out coming responses from complex systems. The main idea of these kinds of networks is (almost) inspirited of working method of bio neural system in data and information processing for learning and creation of knowledge [6]. In these networks data structure is designed using programming knowledge that is called neuron. Then by creating of a network between these neurons and applying of an educational algorithm into it, network is being educated. In this memory or neural network, neurons have two mode of active (on or 1) and inactive (off or 0) and each synapse (connection between nods) has one weight. Positive weighted synapse causes stimulation or activation of next inactive nod and negative weighted synapse inactivates or controls next connected nod (if it was active).

An artificial neuron is a system with lots of inputs and only one output. Neuron has two modes, learning mode and operation mode. In learning mode, neuron learns to be stimulated by special input patterns or open fire idiomatically. In operation mode, when an identified input pattern enters, its corresponding output is provided. If input is not among already identified ones, firing rules decides to be started or not.

A neural network composes layer and weight maker members. Network's behavior also depends to relation between members. Generally, there are three neuron layer types in neural networks: input layer, hidden layer and output layer.

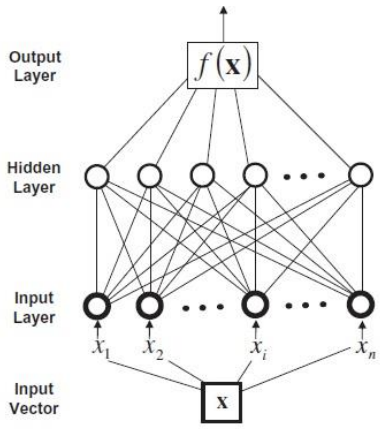

Fig. 2. Structure of artificial neural network

These networks have shown a very high efficiency in estimation and approximation. Application scope of these mathematical models is very vast, such as performance of human brain, which as some small instances we can indicate to usage of this mathematical tool in processing of biological signals, telecommunication and electronical up to help in astronomy and space exploration.

C. *Genetic algorithm*

Genetic algorithm (GA) is a search technic in computer science to find optimum solution and search issues. Genetic algorithms are one of types of evolving algorithms which is inspired by biology science such as inheritance, jump, sudden selection (biology), natural selection and combination. Genetic algorithm which is known as one of random optimization methods, is innovated by John Holland in 1967. Its main goal is

to obtain best results via removal of bad results during creation of population from one generation to the next one [7].

Genetic algorithms often are implemented as a computer simulator in which population of an abstract sample (chromosomes) of nominees of an optimization issue's solutions lead to a better solution. A solution for concerned issue is shown with a list of parameters that are called chromosome or genome. Chromosomes often are shown as a simple row of data's. However structures types of other data's can also be used. Firstly multiple attributes are generated randomly to create first generation. During each generation, each attribute is assessed and fitness value is measured by fitness function. Next step is to create second generation of society that is based on selection procedures, production from selected attributes with genetic operator: connection of chromosomes to head of each other and change. A parent couple is selected per person. Selections are in a way that most appropriate elements be selected to even poorest elements also have selection chance in order to prevent from reaching to a local result.

D. Machine learning

To analyze the system using thermo dynamical method, so many assumptions are required, without these assumptions, thermo dynamical analysis from a real plan required so many none linear equations which their solutions are either impossible or requires lots of time and effort in terms of computation. To eliminate these obstacles, machine learning methods are often used as a substitute of thermo dynamical method [8]. Prediction of real value, which is known as regression, is most common research problem in machine learning. For this reason, machine learning algorithms are used for prediction to control system's response [9]

Machine learning algorithms would be fully described in next section.

III. DESIGN OF LEARNING MACHINES

In this section, two definite and probability learning machines are designed using descend gradient and Bayesian algorithms.

A. Definite learning machine

This machine was designed using descend gradient method. Descend gradient is an important general sample of learning. This algorithm is a strategy to explore big and infinite hypothetical spaces. This algorithm can be used if continuous hypothetical space (for instance, weights space in leaner unit) and errors be explicit based on parameters of this hypothesis. Problems of using of descend gradient is that convergence to a local minimum value takes long time sometimes. For example, hundreds of steps are required to be converged to a specific value and if multiple local minimum exist there is no guaranty for algorithm to be tended to absolute minimum. Descend gradient method is one of oldest and simplest parameter's correction methods in neural networks. In this method, network's parameters are corrected in opposite direction of error gradient.

How largest error reduction can be found? This direction is obtained by implicit differentiation from error value E per all variants of vector$\vec{W}$. This vector is called E's gradient and is shown as $\Delta(\vec{W})$.

$$\Delta E(\vec{W}) \equiv [\frac{\partial E}{\partial W_0}, \frac{\partial E}{\partial W_1}, \dots, \frac{\partial E}{\partial W_n}] \tag{1}$$

As gradient determines side of largest reduction of E, learning law for descend gradient would be as below:

$$w \leftarrow w + \Delta w \tag{2}$$

Where

$$\Delta W = -\eta(\frac{\partial E}{\partial W}) \tag{3}$$

Here, η is a positive value that is called learning rate. This value determines steps' value in descend gradient's algorithm. Negative sign is because we want to move weight's vector toward reduction of E value. This law can be written simpler on variants of weight vector too:

$$w_i \leftarrow w_i + \Delta w_i \tag{4}$$

Where

$$\Delta W_i = -\eta(\frac{\partial E}{\partial W_i}) \tag{5}$$

This shows that to reach to largest reduction, each variant should be changed in proportion with $\frac{\partial E}{\partial W_i}$ value.

To change this process to algorithm and repeating of steps by above correlation it is required to have effective way for calculation of gradient in each step:

$$\frac{\partial E}{\partial W_i} = \frac{\partial}{\partial W_i}\frac{1}{2}\sum_{d\in D}(t_d - o_d)^2$$

$$= \frac{1}{2}\sum_{d\in D} 2\,(t_d - o_d)\frac{\partial}{\partial W_i}(t_d - o_d)$$
$$= \frac{1}{2}\sum_{d\in D}\frac{\partial}{\partial W_i}(t_d - o_d)^2 \qquad (6)$$

$$= \frac{1}{2}\sum_{d\in D} 2(t_d - o_d)\frac{\partial}{\partial W_i}(t_d - \vec{W}.\vec{X}_d)$$

$$\frac{\partial E}{\partial W_i} = \sum_{d\in D}(t_d - o_d)\,(-X_{id})$$

Correlation of change of weight's values for descend gradient is obtained:

$$\Delta W_i = \eta \sum_{d\in D}(t_d - o_d)(X_{id}) \qquad (7)$$

In summary, descend gradient algorithm for educating of leaner units is as below:

First, select a desired vector for weights. Then calculate leaner unit value for each educational sample and calculate ΔW_i for each weight. Change each weight by adding of ΔW_i. And repeat this process till finishing of educational samples.

1) Method of threshold determination

Result of multiplication of each sample by final obtained weight is a number which initially consider minimum value as threshold and consider all of data's samples which result of their multiplication by final obtained weight is lower than this value if real class of that sample also is zero, as proper estimated samples of class zero. Consider all of data's samples which result of their multiplication by final obtained weight is higher than this threshold value if real class of that sample also is one, as proper estimated samples of class one. In next step, add a small value e.g. 0.1 to threshold value that is lowest value at first and obtain the accuracy based on mentioned method and finally report largest accuracy and threshold.

B. Probability learning machine

Classification methods are divided into two general category of definite and indefinite. In indefinite approach, it is tried to directly estimate dividing function. This approach is looking for a function in form of:

$$gi\,(x) = f(p(Ci \mid x)) \qquad (8)$$

i.e. divide is a function of next probability. Indefinite approach is looking for finding of decision making level and decision making level is where in which probability of classes be equal. Means that:

$$gi\,(x) - gj\,(x) = 0 \qquad (9)$$

Or in a more complete way:

$$gij(x) = gi\,(x) - gj\,(x) \qquad (10)$$

If $gij(x) > 0$, x is in I category and if $gij(x) < 0$ x is in j category .

We want to see if distribution of x is normal, how bayes classifier works using distinguishing approach? Now as it is a gaussian distribution, we consider discriminant function as natural logarithm. i.e.:

$$gi(x) = \ln(p(Ci \mid x)) = \ln(p(x \mid Ci)p(Ci) = \ln(p(x \mid Ci)) + \ln(p(Ci)) \qquad (11)$$

Now, with inserting into multidimensional distribution function we have:

$$b_i = logp_i - \frac{1}{2}\log|\sum i| \qquad (12)$$

Where

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1d} \\ \dots & & \sigma_{ij} & \dots \\ \sigma_{d1} & & \dots & \sigma_{dd} \end{bmatrix} \qquad (13)$$

$$\sigma_{ij} = E[(x_i - m_i)(x_j - m_j)]$$

IV. Results and Assessment

This section is describing results that is obtained from definite learning machine and indefinite learning machine. Description related to data group is stated in following and in continue obtained results are shown.

A. Data group

Used data group is combined cycle power plant that is taken from UCI site [16]. This data group is related to combined cycle power plant and have 4 characteristics that by using them power output hours from this power plant is calculated. Also this data group has 9568 samples.

TABLE I. INFORMATION RELATED TO DATA GROUP

Attribute	Value
Temperature	1.81°C and 37.11°C
Ambient Pressure	992.89-1033.30 milibar
Relative Humidity	25.56% to 100.16%
Exhaust Vacuum	25.36-81.56 cm Hg
Net hourly electrical energy	420.26-495.76 MW (output)

B. Results related to definite learning machine

Definite learning machine using descend gradient method was educated with 30 percent of data's and was tested with 70 percent of data's. Using provided method for determination of threshold, threshold value was determined as 1.63 and using this value accuracy value as of 95.71% for prediction was obtained.

Taking into account of made assessments in training section, machine will lead to over learning by beyond limit error reduction that desired response will not obtained in test section. Via sometimes repeating by stop of different errors the best error value using this method was 0.015. error reduction curve considering obtained results are shown in figure 3, 4.

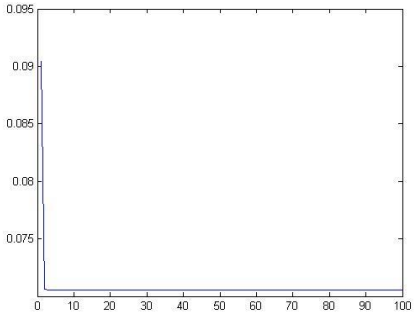

Fig. 3. error reduction curve

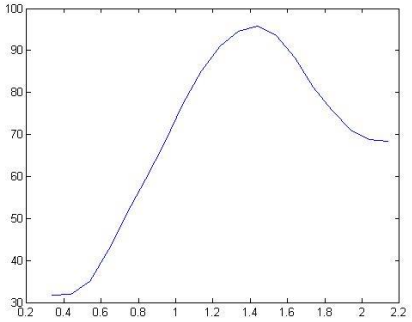

Fig. 4. curve of threshold values per related accuracies

C. Results related to indefinite learning machine

Indefinite learning machine was educated with 30 percent of data's and was tested with 70 percent of data's. Identification value in class 1 is 95% and in class two is 97.66%. Obtained accuracy value in this machine is 96.89%. Obtained results are seen in table 2.

TABLE II. OBTAINED RESULTS OF INDEFINITE LEARNING MACHINE

	Sensitivity	Specify	Accuracy
Class 1	%95	% 97.66	% 96.89
Class 2	%97.66	%95	% 96.89

D. Assessment

Table 3 shows comparison between two definite and probability method for prediction of full load of electrical output power. Generated results by this paper shows that probability machine is doing better in prediction.

TABLE III. COMPARISON BETWEEN TWO DEFINITE AND PROBABILITY METHOD

Method	Train (%)	Test (%)	Accuracy (%)
Definite	70	30	95.71
Probability	70	30	96.98

V. CONCLUSION

Soft computation is an evolving group of inaccurate computational methods of artificial intelligence which is able to analyze and model very complicated problems. Prediction of full load electrical output power of a power plant for benefit maximizing is important. Machine learning methods often are used as substitutes of thermo dynamical method. In this paper, soft computation methods was assessed in summary and a definite learning machine was designed using deset gradient

method to predict load value of output power then a probability machine was designed using bayes algorithm in order to increase this prediction and then was assessed using a body of data's which have been gathered more than 6 years which accuracy value as of 95.71% for definite learning machine and as of 96.89% for probability learning machine were obtained that made comparison shows that probability machine is more successful in prediction.

Otherwise, you may use this as just an instruction set. It is remarked that you do not have to follow this style file when your works is submitted for the initial review stage.

REFERENCES

[1] Yun Huang, Yi Lan, Joseph Thomson, Ami Fang, Wicky Hoffmann and Roland Lacey, "Development of soft computing and applications in agricultural and biological engineering", *Computers and Electronics in Agriculture*, vol. 71, pp. 107–127, 2010.

[2] Puck Tufekci, "Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods", *Electrical Power and Energy Systems,* vol. 60, pp. 126–140, 2014.

[3] Ebrahim Zadehlotfi, "Fuzzy sets", *Information and Control*, vol. 8: pp. 338–353, 1965.

[4] Caber Salah and Mohm Ouali, "Comparison of fuzzy logic and neural network in maximum power point tracker for PV systems", *Electric Power Systems Research*, vol. 81, pp. 43–50, 2011.

[5] Caber Salah, Mohm Chaabene and Mirda Ammar, "Multi-criteria fuzzy algorithm for energy management of a domestic photovoltaic panel", *Renewable Energy* vol. 33, pp. 993–1001, 2008.

[6] Hen Yoon, Sijitcu Jun, Yun Hyun, Geo Bae and Kim Lee, "A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer", *Journal of Hydrology,* vol. 396, pp.128–138, 2011.

[7] Michel Albayrak and Ned Allahverdi, "Development a new mutation operator to solve the Traveling Salesman Problem by aid of Genetic Algorithms", *Expert Systems with Applications,* vol. 38, pp. 1313–1320, 2011.

[8] Uikh Kesgin and Hamer Heperkan, "Simulation of thermodynamic systems using soft computing techniques", *Int J Energy Res,* vol. 29, pp. 581–611, 2005.

[9] Harf Guvenir and Isak Uysal, "Regression on feature projections", *Knowl-Based Syst,* vol.13, pp. 207–14, 2000.

[10] Lee Malago, Merd Matteo and Gerard Pistone, "Stochastic Natural Gradient Descent by estimation of empirical covariances", *IEEE International Conference on Evolutionary Computation (CEC*), koln, Germany, pp. 949 – 956, 2011.

[11] Rab Arablouei, Sid Werner and Kim Dogancay, "Diffusion-based distributed adaptive estimation utilizing gradient-descent total least-squares", *IEEE Congress on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, pp. 5308 – 5312, 2013.

[12] Fang Liu, Wee Keong Ng and Wei Zhang, "Encrypted Gradient Descent Protocol for Outsourced Data Mining", *IEEE 29th International Conference on Advanced Information Networking and Applications (AINA),* Jhingen, China, pp. 339 – 346, 2015.

[13] Jin Wu, Zhang Cai and Xheng Zhu, "Self-adaptive probability estimation for Naive Bayes classification", *International Joint Conference on Neural Networks,* Netherland, pp. 1 – 8, 2013.

[14] Mart Martinez-Arroyo and Liong Sucar, "Learning an Optimal Naive Bayes Classifier", *18th International Conference on Pattern Recognition*, (ICPR) 2006, pp. 958.

[15] Fred Klawonn and Pilton Angelov, "Evolving Extended Naive Bayes Classifiers", *Sixth IEEE International Conference on Data Mining Workshops*, (ICDM Workshops) China, pp. 643 – 647, 2006.

[16] Jack Nilsson, "Learning Machines", pp. 43-63, 1965.

[17] Data set combined cycle power plant over 6 years (2006-2011), https://archive.ics.uci.edu/ml/datasets/ Combined + Cycle+Power+Plant, Accessed dec 8, 2015.