AN OPTIMAL DESIGN FOR FACT DETECTION THROUGH SELF-ORGANIZING MAP BASED CLUSTERING FROM INDUSTRIAL PERSPECTIVE

Valarmathi.V[1], Swathi.M[2], Ukesh.S[3]

[1]Assistant Professor, [2,3]M.Sc IT Student
Department Of Information Technology
Sri Krishna Art and Science College
Coimbatore -8, Tamil Nadu, India

***ABSTRACT*: In this paper, predicted data is clustered and the determined facts are useful in deriving optimal operational set points needed for performance enhancement. The work illustrates how Self-Organizing Map, an unsupervised learning method suitable for complex system analysis to visualize high dimensional data and discover new facts based on the trained data. The knowledge obtained from proposed analytical model generates facts where exist fewer characteristics reside in each resultant clusters that states the general behavior of attributes dominated during training process.**

***Keywords*: Fact detection, Self-Organizing Map, Optimal design, High dimensional data.**

1.INTRODUCTION

Nowadays, the performance of the manufacturing system depends widely on raw data and variety of design with evaluation criteria. Solving the complex problems which arises dynamically that affects overall process really needs to formulate optimal objective.Several computing methodologies provide suggestible solution but challenging task is to select the best among available options. Based on the working engineering process the user needs to have knowledge about theimpact of operational parameters. First, theuser can select simple and flexible algorithmwhich satisfies prescribed procedural constraints. Second, choose the decision attributes among the raw data to execute chosen algorithm leads to result in set of good solutions. Theevaluator then decides on optimal solution among the set using hidden or obtained additional knowledge.

Data mining clustering algorithm were applied to generate control signatures from the sample instances to improve combustion efficiency. Neural network model were used to validate the boiler efficiency by using control signatures.Virtual testing procedure overcomes the complexities such as cost and time exist in real time testing [1]. Data mining approach is highly suitable for predicting best patterns applied to execute optimization of boiler process. Analysis reported that the importance of a particular operational variable is not stable. The value of each variable depends on number of factors in different period of times [2].

Optimization model needs proper input to be feed to execute. Data mining technique plays vital role in generating best patterns that act as a base to make optimization output to be an efficient one [3].Hybrid combination of evolutionary computation, data mining and model predictive control forms a new framework to provide more exact dynamic information about operational process. Research suggested focusing on multi objective optimization where more performance related variables to be considered [4]. The applications of data mining techniques for identification and prediction of status patterns in wind turbines are confidently presented. A prediction model was built using operational and status data collected at working environment and different analyzing technique were applied to derive required patterns [5]. The practical application of soft computing techniques to various selected industrial applications is reviewed and stated its generic characteristics [6]. Thermal analysis leading to excellent prediction of equipment essential parameters based on their correlations. Thermal performance was evaluated by creating thermal modeling. The impact of net present value was depicted with obtained results [7].

Quantitative measures are derived based on past happenings and its severity with the intention to identify the location of the plant where more monitoring is needed to support safety programs and to avoid future complexities [8]. Power plant system is comprised into various sub systems and had equal contribution to the overall plant efficiency and different methods for effective energy savings are stated and results shown the impact of analysis [9].The unreliable and uncertain failure data associated with 500 MW loads for five years were collected and reliability estimation were carried out and come out with the measures to improve plant performance [10]. The performance study of various plant units with the data associated to different period of time were carried out and resulted in representing the causes for unavailability of energy and the ways to overcome it [11].

A major unit of power plant called boiler drum were considered for investigation and required controllable parameters are chosen to derive optimal strategies to design secured automatic control system. Deviations are automatically determined and clearly visualized in comparison to standard operations [12].

Theoretical investigation proves that the contribution of each individual equipment's of the plant need to be assessed periodically to monitor, control and fault diagnosis [13]. An efficiency of neural network was demonstrated over power plant applications by selecting single parameter belongs to specific superheater and defines how modeled network is suitable to monitor and control performance based on the defined standard [14].

2. SELF-ORGANIZING MAP

2.1 Need for SOM

The Neural Network model is classified as supervised and unsupervised learning methods.In supervised, the overall data flow must be predefined and a desired outcome for each input samples represented in the vectors is required to train the constructed network.Here,human intervention is more and essential to have in depth domain knowledge to provide sufficient information and have control over the selected parameters.This supervised learning method is suitable for the application where the data is in small volume. It is not suited to handle the applications with enormous amount of data, automatically the performance deviates.

The unsupervised learning method is suggestible to deal with large amount of data with expected efficiency, purely based on data and predefined information’s are not required.Several methodologies can emerge for this purpose; the user can select suitable approach for their applications. In this paper, we selected Self-Organizing Map (SOM), a clustering based neural network model an unsupervised learning method for our application.

SOM algorithm need not expect to specify the starting seeds for iterations. Based on the features incorporated within data, it randomly chooses the initial seeds and starts iteration without knowing the class membership.SOM provides a structure by preserving the semantics when obtaining low dimensionality of data reduced from high dimensional topology.

2.2 SOM Architecture

Basically network has two layers,

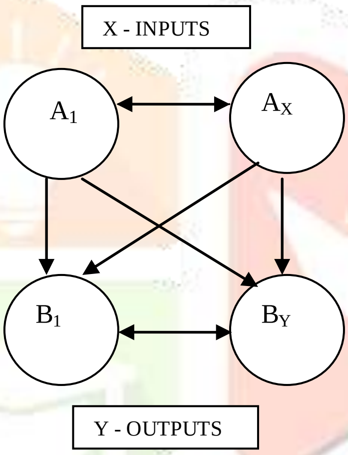

Fig - 1 SOM Schematic Representation

2.3Steps in SOM Algorithm

STEP: 1 – Select required input parameters and initialize current neighborhood distance with positive integer as DIST (0).

STEP: 2 – Initialize weight vector WV_{ij}using small random values (n=1).

$WV_{1,1}$ $WV_{1,2}$…… $WV_{1,x}$

$WV_{2,1}$ $WV_{2,2}$…… $WV_{2,x}$

STEP: 3 – Select an input sample A_i from AV_{ij}input vector.

$A_{1,1}$ $A_{1,2}$ …….. $A_{1,x}$

$A_{2,1}$ $A_{2,2}$ …….. $A_{2,x}$

$A_{P,1}A_{p,2}$ …….. $A_{p,x}$

STEP: 4 – Calculate square of Euclidian distance of $A_i$from weight vector WV_{ij}corresponding to each output unit B.

$$\sum_{m=0}^{x} (A_{i,j} - WV_{j,m}(n))^2$$

STEP: 5 – Select the winning node among output units based on minimal value.

B_1, B_2, B_y

STEP:6 – Perform weight updates to all units from winning node as DIST(n) based on the following rule,

$WV_{i,j}(n+1) = WV_{i,j}(n) + (ή (n) A_i - WV_{i,j}(n))$

STEP: 7 – Increment n by 1.

STEP: 8 – Terminate when criteria are met.

2.4 RESULTS AND DISCUSSION

The following table shows the characteristics of resultant clusters.

X, Y ....▶ Target Units

CID ....▶ Cluster ID

TABLE: 1 Resultant Clustered Characteristics

CID		Min	Max	Mean	Median	Mode	Stddev
C1	**X**	0.184	0.234	0.210	0.210	0.184	0.0146
	Y	0.186	0.239	0.214	0.214	0.186	0.0154
C2	**X**	0.235	0.262	0.248	0.248	0.234	0.0079
	Y	0.239	0.265	0.253	0.253	0.239	0.0078
C3	**X**	0.262	0.279	0.271	0.270	0.262	0.0052
	Y	0.266	0.283	0.275	0.274	0.266	0.0050
C4	**X**	0.280	0.298	0.289	0.289	0.280	0.0054
	Y	0.283	0.301	0.293	0.292	0.283	0.0053
C5	**X**	0.324	0.377	0.350	0.349	0.324	0.0158
	Y	0.328	0.389	0.357	0.355	0.328	0.0179
C6	**X**	0.298	0.323	0.311	0.310	0.298	0.0073
	Y	0.302	0.327	0.315	0.314	0.302	0.0075

Obtained Values:

Min ....▶ Minimum value of input

Max ....▶ Maximum value of input

Mean ....▶ Centroid value

Median ....▶ Central value

Mode ....▶ Form value

Stddev ....▶ Standard deviation

The following figures show the optimal design value for each SOM clusters and linearity of the selected parameters are visualized.

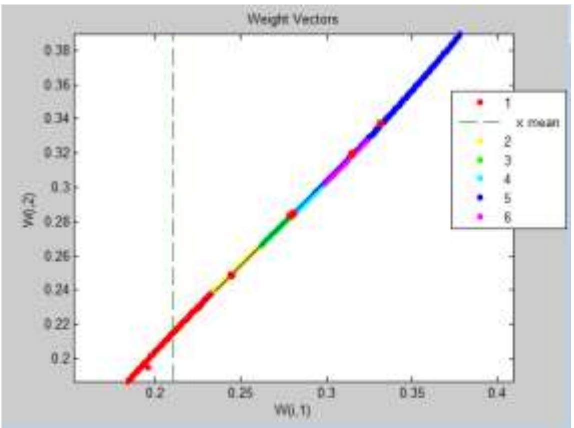

Fig -2 SOM_Cluster_1

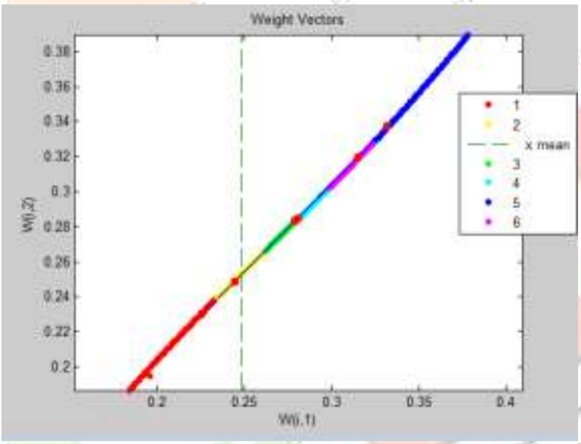

Fig -3 SOM_Cluster_2

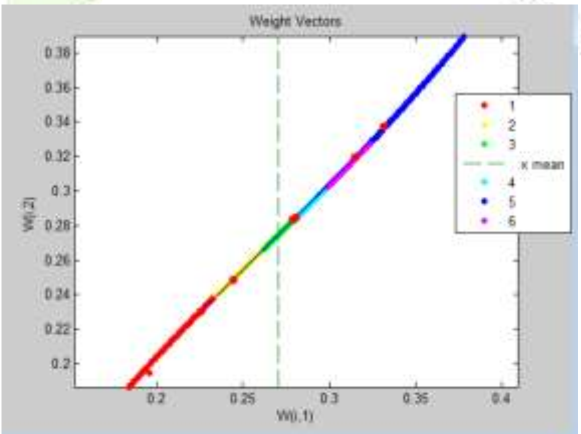

Fig -4 SOM_Cluster_3

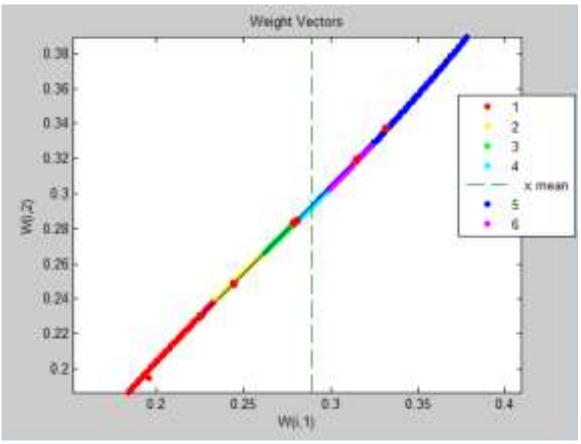

Fig - 5 SOM_Cluster_4

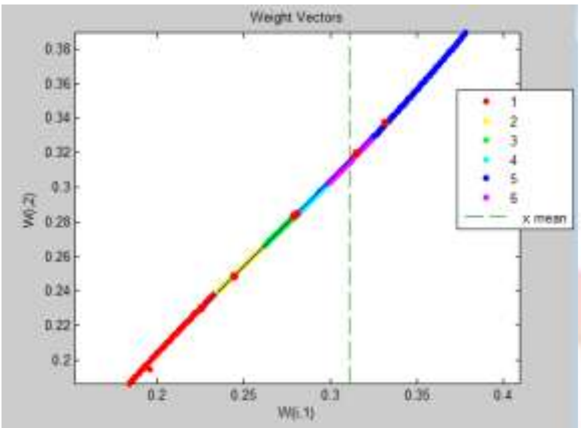

Fig - 6 SOM_Cluster_5

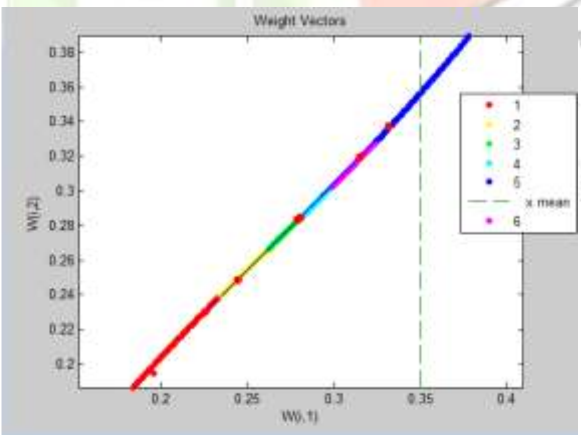

Fig - 7 SOM_Cluster_6

CONCLUSION

The practical application of SOM in engineering process analysis visually represents the facts essential to derive process control strategies that makes evaluation set points more accurate and acceptable. The SOM algorithm therefore leads to an optimal design in which effects are taken to improve industrial outcomes. The proposed method envisagesanalysis through operational values is more effective rather than raw data.

ACKNOWLEDGMENT

The authors would like to express thanks to the management of Sri Krishna Institutions for giving incessant support and encouragement for the success of our work.

REFERENCES

[1] Andrew Kusaik and Zhe Song, 2006, “Combustion Efficiency optimization and virtual testing – a data mining approach”.
[2] Andrew Kusaik and Zhe Song, 2007, “Constraint based control of boiler Efficiency – a data mining approach”.
[3] Andrew Kusaik and Zhe Song , 2008, “Clustering based performance optimization of boiler – turbine system”
[4] Zhesong and Andrew Kusiak “ Optimization of temporal process : A model predictive control approach”
[5] Andrew Kusaik ,AnoopVarma 2011 ,” Prediction of Status Patterns of Wind Turbines: A Data mining Approach”,Journal of Solar Energy Engineering
[6] Yasuhiko dote and SteppoJ.ovaska,” Industrial application of soft computing – a Review “.
[7] RavinderKumar,AvdeshKr.Sharma and P.C Tewari, a Research Article (2014) “Thermal Performance and Economic Analysis of 210 MWe Coal Fired Boiler”,Journal of Thermodynamics.
[8] Abhaynath Kumar, Jain N.K & Praveen Patel, 2015, “Analysis of Safety Performance Rating in Thermal Power Plant” International Journal of Emerging Technology and Advanced Engineering, vol. 5, no. 1, pp. 120-128.
[9] Aditya Sindhu, Aashish Bhaskar&Arbind Singh, 2014, “An Analysis of Different Methods for Major Energy Saving in Thermal Power Plant” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 3, no. 9, pp. 12153-12161.
[10] Dewangan.D.N, Manoj Kumar Jha&Banjare.Y.P,2014, “Reliability Investigation of Steam Turbine used in Thermal Power Plant” International Journal of Innovative Research in Science ,Engineering and Technology, vol. 3, no. 7, pp . 14915-14923.
[11] Ogieva.F.E,Ike.S.A&Anyaeji.C.A, 2015, “Sample Thermal Power Station Generator Availability and Units Performance Studies” International Refereed Journal of Engineering and Sciences, vol. 4, no. 6, pp. 01-17.
[12] Padma Priya.K, Naveen Kumar.P&Murthy.S.E 2015, “Real Time Monitoring and Controlling of Boiler Drum Parameters in Thermal Power Station” International Journal of Engineering Development and Research, vol. 3, no. 2, pp. 897-904.
[13] Panda.S, Barpanda.N.K&Pradhan.S, 2014, “A New Approach for Optimal Operation of Thermal Power Plant” International Journal of Innovative in Science, Engineering and Technology, vol. 3, no. 3, pp. 43-46.
[14] Sai.T.K& Reddy K.A,2015, “Neural Network Applications in a Power Station” International Journal of Soft Computing and Engineering, vol. 4, no. 6, pp. 112-120.