Apparel Recommender System Based on Feature Selection and Self Organized Map

Priya Gupta, Snigdha Agrawal, Narina Thakur

Abstract— With the explosive growth of online shopping the buyer has day by day become more vigilant examining and comparing the products available with various users' views about the various brands/ vendor, so that they can get the best deal. Self organizing Maps is a well-known, unsupervised learning approach of neural network used for clustering and classifying high dimensional large data. This evaluation requires certain machine learning tools to handle large and complex data set which helps in enabling the customers to take certain decisions among the available choices. This paper proposes a framework for recommending apparel based on apparel buying behavior of females using rank based feature selection approach, self organized map and UCI repository Dresses Attribute Data Set comprise of 500 training data sets with 13 different attributes. These attributes are further reduced to key attributes using feature selection algorithms and further a set of dresses are recommended to the buyer

Index Terms— Clustering; Unsupervised learning; Competitive learning; Euclidean Distance.

I. INTRODUCTION

Self-organizing map (SOM) is a type of artificial neural network (ANN) which uses unsupervised learning typically to produce a two-dimensional, discretised representation of the input space of the training samples. SOM is one of the clustering method that were proposed by Teuvo Kohonen in 1982 [1]. It belongs to the category of competitive learning networks. No target results for the input data vectors are provided and the training of the network is data-driven in unsupervised learning. The SOM model has two variants: the first variant performs unsupervised learning through a controlled growth process and the second variant of the model is a supervised learning method with the radial basis function (RBF) approach.

In this paper, we propose rank based feature selection method on Dresses Attribute Data Set comprising 500 training data sets with 13 different attributes. Here we will be clustering our data set according to the attributes available and then produce low feature output space domain so that a set of dresses are recommended to the buyer.

Manuscript received March 26, 2015

Priya Gupta, Department of Information Technology, Bharati Vidyapeeth College Of Engineering New Delhi, India

Snigdha Agrawal, Department of Information Technology, Bharati Vidyapeeth College Of Engineering New Delhi, India

Narina Thakur, Associate professor in Department of Computer Science Engineering , Bharati Vidyapeeth College Of Engineering New Delhi, India

The principal of SOM can be extended to higher levels of processing by relationships between items semantically rather than imprecise essential features using *symbolic expressions. [2]*

A model was proposed by James Smith [3] which is based on the SOM that allows various structure i.e. One-to-one, many-to one or one-to-many of the desired state-action mapping which is to be captured

An algorithm that describes Grow When Required (GWR) network was presented by Marsland, Shapiro, and Nehmzow [4]. This network has properties that it maintains a set of neighborhood connections between nodes that match similar perceptions and adds neurons whenever the current input is not matched sufficiently well by any of the current nodes.

A post-training method for the interpolation of a trained SOM was proposed by Yin and Allinson [5] so that the map can be enlarged without retraining from scratch or the need to use the original data. SOM is trained on the data, which are then projected onto the map using U-matrix colors to reflect the inter-neuron or inter-sample distances.

The network self-organizing map (NetSOM) model which is an extension to the SOM model, is capable of effectively decomposing complex, large-scale pattern classification problems into a number of units, each of which is manageable with a local classification device. [6]

In section 2, we briefly describe about the SOM, its process and its algorithm. In section 3, we describe the data set. Section 4 discusses our proposed scheme. In section 5, we discuss the theoretical analysis and empirical estimation of training and testing time of both the proposed schemes. Section 6 includes the conclusion. In section 7, we cite the references.

II. SELF-ORGANIZING MAP

SOM consists of input and output layers which are completely connected with the help of weights. A topology is defined in output layer for weight updating. The main objective of the SOM is to obtain a bijective application, i.e. Neighborhood in the input space and neighborhood in the output space (as in figure 1). SOM is characterized by the formation of a topographic map [7] of the input patterns in which the spatial locations of the neurons in the lattice are indicative of intrinsic statistical features contained in the input patterns. This network represents a feed forward structure with a single computational layer consisting of neurons arranged in a grid.

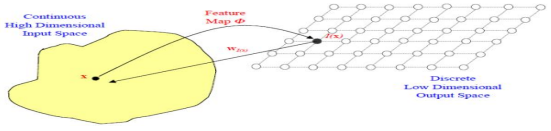

Figure 1: Mapping of continuous input space into discrete output space.

SOM also works with the principal of Competitive Learning but it is different from it. There is a spatial organization in the distribution of new neurons. In SOM, we talk about the lattice of output neurons. The lattice that can be 1D, 2D or higher dimensions. 1D or 2D is preferred over higher dimensions. We feed the input patterns which act as stimuli to neurons present at the output layer. So that when the stimulus is presently one of the neurons will be the winner and synaptic weights will be updated in such a way that the Euclidean distance between the input and the winning neuron will be minimized due to which synaptic weight updating will disturb the lattice structure and would move towards winning neuron. The neurons present at the output layer act in a competitive manner in the sense that they inhibit the responses of each other. Around the winning neuron, an exciting response is created, whereas inhibitory response is created for the neurons that are farther from the winning neurons. This mechanism will encourage long large inhibition and short range excitation.

The growing hierarchical self-organizing map (GHSOM) was presented by Dittenbach, Merkl, and Rauber [8] that evolves dynamically into a hierarchical structure according to the requirements of the input data during an unsupervised training process.

A global orientation of the independently growing maps, i.e. GHSOM in the individual layers of the hierarchy, navigation across branches was provided by Rauber, Merkl, and Dittenbach [9] which offers problem-dependent architecture, and the hierarchical relations representation of the data.

A compromise SOM can be designed that enables us to treat several variables that are numerical and categorical in a single model [10]. This provides an integrated model that can be used to detect suspicious observations and to impute values.

A model was given by Hagenbuchner, Sperduti, and Tsoi [11] which is an extension of traditional SOMs, is capable of exploiting both information conveyed on the labels attached to each node of the input DAGs and information encoded in the DAG topology.

Due to the SOM model's first variant, there is no need to choose the network size in advance instead the growth process can be continued until a performance criterion is met and due to the second variant, the current classification error can be used to determine the locations of new RBF units as the positioning of RBF units and supervised training of connection weights [12] is performed in parallel.

The efficacy of SOMs was investigated by Bezerianos ,Vladutu and Papadimitriou [13] which is an unsupervised learning approach for the automatic clustering of XML documents.

A powerful relevance feedback mechanism can be implemented by using SOM's inherent property of topology-preserving mapping from a high-dimensional feature space to a two-dimensional grid of artificial neurons. [14]

A. SOM Process

The self-organization process includes four major phase [15]:

Initialization: In the initialization phase, all the connection weights are initialized with small random values

Competition: After completion of initialization phase, the neurons compute their respective values of a discriminant function which provides the basis for competition for each input pattern. The particular neuron with the smallest value of the discriminant function is declared the winner. Discriminant function is defined as in equation(1)-

$$d_j(x) = \sum_{i=1}^{p}(x_i - w_{ji})^2 \qquad (1)$$

Cooperation: The coordinates of a topological neighborhood of excited neurons are determined by winning neuron, thereby providing the basis for cooperation among neighboring neurons during cooperation phase. The topological neighborhood for the neurons is defined as in equation(2)-

$$T_{j,I(x)} = \exp(-S^2_{j,I(x)}/2\sigma^2) \qquad (2)$$

Adaptation: In adaptation phase, the excited neurons decrease their individual values of the discriminant function in relation to the input pattern by making suitable adjustment in the associated connection weights so that the response of the winning neuron to the subsequent application of a similar input pattern is improved. The appropriate equation for weight update is defined as in equation(3)-

$$\Delta w_{ji} = \eta(t).T_{j,I(x)}(t).(x_i - w_{ji}) \qquad (3)$$

B. SOM Algorithm

The SOM algorithm [16] has two phases-training and testing.

Training: In the training phase, first select output layer topology and then trains, weights connecting inputs to outputs; Topology is used to define which weights will be updated in conjunction with current mapping of inputs to outputs. Topology defines which output layer units are neighbors with which others. The advantages are distance measure using the topology is reduced over time; reduces the number of weights that get updated per iteration and learning rate is also reduced over time.

Testing: In testing phase, use weights from training to test the data.

SOM algorithm offers various advantages as easy interpretation of data mapping and the projection of the high-dimensional data onto a two-dimensional map where as the associated issues are lack of qualitative training data, No standard measure for identifying SOM, No fixed criteria to determine which input weights can be used, though the mapping process can result in clusters segregations which computationally expensive.

C. SOM'S MODELS

Von der Malsburg model*:* It works on the same principle as retina, optic mapping which is from retina to the visual cortex where the retina is the pre synaptic layer and the visual cortex is the post synaptic layer. In this, input dimensions are same as the output dimensions as in figure 2.

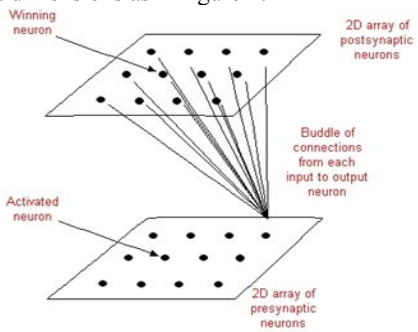

Figure 2: Von der Malsburg model [12]

Kohonen model: Kohonen model is considered as one of the most general models as it permits the data deduction. In this model, only the output layer is organized as in seen in fig 3. It belongs to the vector coding algorithm which optimally places the fixed number of vectors in a higher dimensional input space. It is also known as topology preserving map because there is a topological structure imposed on the nodes in the network.

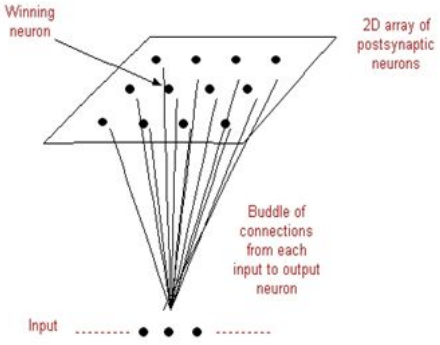

Figure 3: Kohonen model [12]

III. DATASET DESCRIPTION

The data set contains the information about the dresses sale based on the various attributes taken from UCI Repository. The different attributes are: Style, Price, Rating, Size, Season, Neck Line, Sleeve Length, waistline, Material, Fabric Type, Decoration, Pattern Type and Recommendation. Out of all, we are using only four main attributes which are Sleeves Length, Neck Line, Price, and Style. It includes 500 instances and 14 attributes as shown in Table I Attribute Information below.

Table I: Attribute Information-

S.No.	Parameters	Types
1	Style	Bohemia, flare, novelty, OL, party, sexy, brief, casual, cute, fashion, vintage, work
2	Price	Low, High, Average, Medium, Very-High
3	Rating	1-5
4	Size	S, XL, M, L, Free
5	Season	Autumn, Summer, winter, Spring.
6	Neck Line	O-neck, backless, board-neck, Bow neck, halter, mandarin collar, open, ruffled, scoop, slash-neck, square collar, sweetheart, turn down collar, V-neck, Peterpan-color,
7	Sleeve Length	Full, half sleeves, butterfly, sleeveless, short, half, three quarters
8	Waistline	Dropped, princess, empire, null, natural.
9	Material Fabric Type	Wool, cotton, mix, chiffon, Dobby, knitted, jersey, flannel, poplin, satin, corduroy
10	Decoration	Appliqué, beading, bow, button, draped, embroridary, feathers, cascading, crystal, flowers.
11	Pattern type	Solid, leopard, animal, dot.
12	Recommendation	0,1

IV. IMPLEMENTATION

Data Processing Stages

The data processing stages of SOM are-

1. *Data acquisition* involves making a database query, measuring variables; etc. If the data are coded in a non metric scale the coding must be transformed. The coding must be in harmony with the similarity measure used.
2. *Data preprocessing* stage removes or corrects erroneous data (NANs). The preprocessing operation is filtered using fixed or adaptive conditions. The filters are typically built using a priori knowledge of the problem domain.
3. *Segmentation* means dividing the input data into two separate subsets according to recommendation criteria, which are often determined using a priori knowledge.
4. *Feature Selection* is used for better classification accuracy. Feature is a distinctive attribute or aspect of something based on which analysis is done. There are two main approaches for feature selection : Wrapper methods, in which the features are selected using the classifier, and Filter method, in which the selection of features is independent of the classifier used. We have used univariant filter model for analysis. Filter model is a better approach than the Wrapper model as it is less complex. It has two variants-univariant and multivariant. Univariant is ranking method technology, which gives a better score and a better rank given by equation 4

$$\text{Rank} = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \qquad (4)$$

5. *Normalization* gives input data components a desired weight in the SOM training. This assures that each component has equal influence in training.

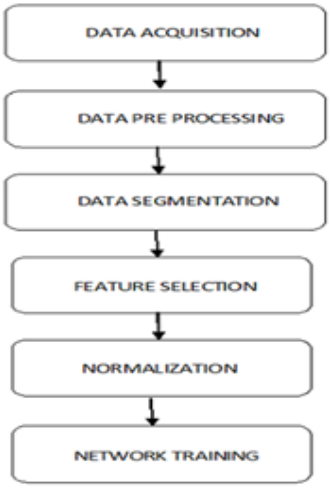

Figure 4: Data Processing Stages

V. RESULTS AND ANALYSIS

We can specify different topologies for the original neuron locations with the functions such as hextop, gridtop, and randtop. Our proposed scheme is implemented in MATLAB.

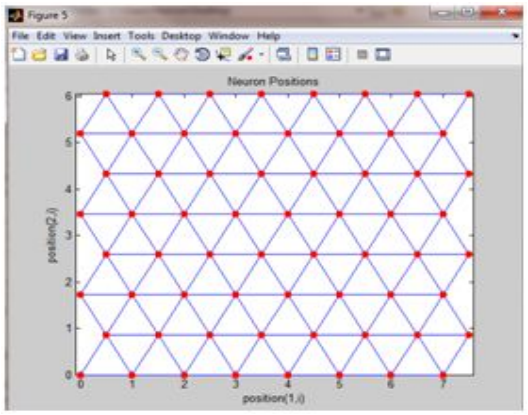

Figure 5: Neuron Positions

In this figure, the red dots represent the positions of the neurons in a hexagonal arrangement and each of the hexagons represents a neuron. The grid is 4-by-4, so there is a total of 16 neurons in this network. There are four basics in each input vector, so the input space becomes four-dimensional.

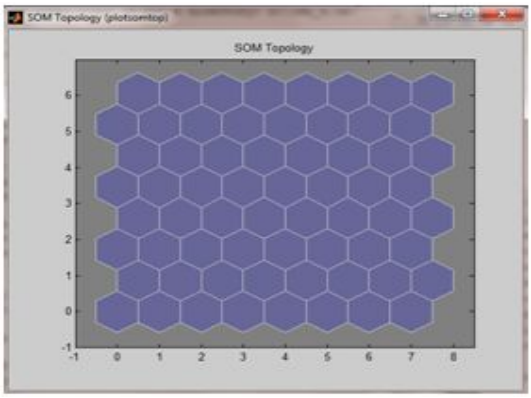

Figure 6: SOM Topology

In this figure 7, the blue hexagons represent the neurons. It shows the topological mapping of SOM.

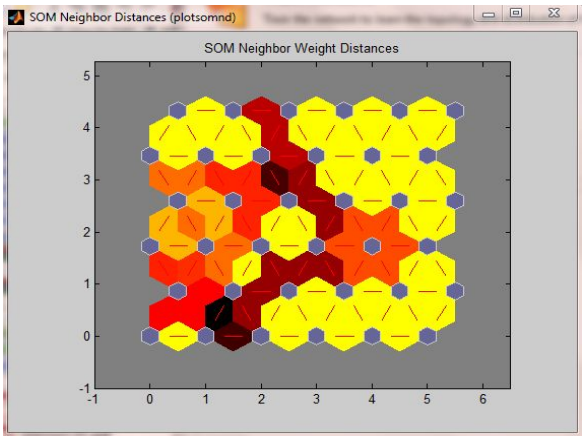

Figure 7: SOM Neighbor Distances

A group of light segments appear in the upper-left part of figure 7 which is bounded by some dark color segments. This represents that the network has clustered the data into two groups and this color difference indicates that data points in this region are farther apart. Here, the colors in the regions containing the red lines indicate the distances between neurons. The darker colors indicate larger distances and the lighter colors represent smaller distances. This shows the locations of the data points and the weight vectors.

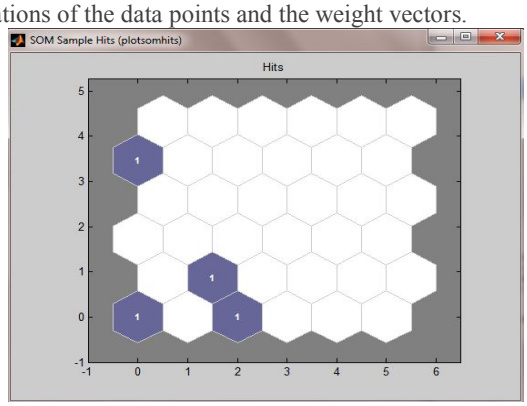

Figure 8: SOM Sample Hits

This figure 8 tells us the number of data points which are associated with each neuron. The data is concentrated a little more in the lower-left neurons, but the overall distribution is fairly even.

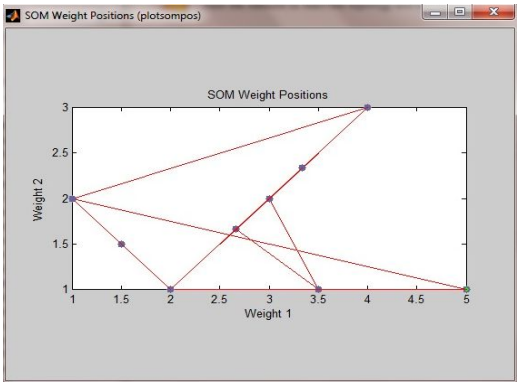

Figure 9: SOM Weight Positions

When the input space is high dimensional, we are not able to visualize all the weights at the same time. In this case, click SOM Neighbor Distances, which indicate the distances between neighboring neurons.

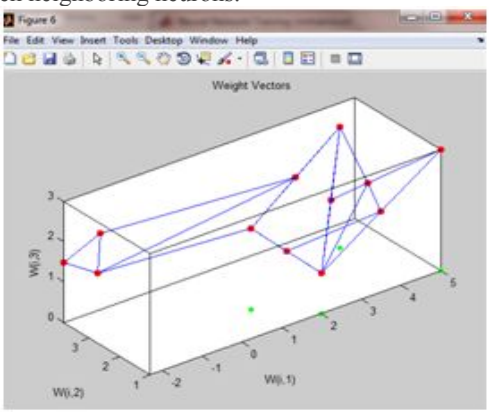

Figure 10: Weight Vectors

CONCLUSIONS

The proposed framework utilizes a Self-organizing Neural Network model. Prior to the apparel recommendation the rank based feature selection method is chosen to extract four key features from thirteen inputs of Dresses Attribute Data Set from UCI repository. The used SOM algorithm in the domain of apparel buying behavior of females enables; easy interpretation of mapped data and the projection of the high-dimensional data onto a two-dimensional map. SOM helps in cluster the apparel data for recommendations to the buyer, so that the recommender system can suggest apparel and it further discover the other similar apparels in the domain using the Euclidean measure. Future work will examine the topology preservation capabilities of the networks and the associated qualitative training data issues. The proposed method decreases the effort buyers' in searching apparel and enable them to in reducing the purchase decision time and increases the quality of decisions.

REFERENCES

[1] Teuvo Kohonen, Self-Organized Formation of Topologically Correct Feature Maps, Biol. Cybern. 43, 59-69 (1982).
[2] H. Yin and N.M. Allinson, Interpolating self-organizing map (ISOM), ELECTRONICS LETTERS 16th September 1999 Vol. 35 No. 79, pno 1649.
[3] Andrew James Smith, Applications of the Self-Organizing Map to Reinforcement Learning, Institute for Adaptive and Neural Computation, 2003
[4] Stephen Marsland, Jonathan Shapiro, Ulrich Nehmzow, A self-organizing network that grows when required, Neural Networks 15 (2002) 1041–1058.
[5] H.Yin and N.M.Allinson, Bayesian self-organising map for Gaussian mixtures, IEE Proc.-Vis. Image Signal Process., Vol. 148, No. 4, August 2001.
[6] F. Trentini, and M. Hagenbuchner, A Self-Organizing Map Approach for Clustering of XML Documents, 2006 International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.
[7] Marc M. Van Hulle, Self-Organizing Maps, 2009
[8] Michael Dittenbach, Dieter Merkl, Andreas Rauber, The Growing Hierarchical Self-Organizing Map, Proceedings of the Int'l Joint Conference on Neural Networks (IJCNN'2000), Como, Italy, July 24-27, 2000, pp VI-15 - VI-19. IEEE Computer Society Press, Los Alamitos, CA, ISBN 0-7695-0619-4.
[9] Andreas Rauber, Dieter Merkl, and Michael Dittenbach, The Growing Hierarchical Self-Organizing Map: Exploratory Analysis of High-Dimensional Data, IEEE TRANSACTIONS ON NEURAL NETWORKS, PRE-PRINT, ACCEPTED FOR PUBLICATION.
[10] Jorma Laaksonen, Markus Koskela, Sami Laakso and Erkki Oja, Self-Organizing Maps as a Relevance Feedback Technique in Content-Based Image Retrieval, Pattern Analysis & Applications (2001)4:140–152 2001 Springer-Verlag London Limited.
[11] Markus Hagenbuchner, A. Sperduti, Ah Chung Tsoi, A self-organizing map for adaptive processing of structured data, IEEE Transactions on Neural Networks, May 2003, 14(3), 491-505. Copyright IEEE 2003.
[12] Dr. Stathis Kasderidis, WK6 – Self-Organizing Networks: CS 476: Networks of Neural Computation, Spring Semester, 2009
[13] Bezerianos I L. Vladutu I S. Papadimitriou, Hierarchical state space partitioning with a network self-organising map for the recognition of ST-T segment Changes, Med. Biol. Eng. Comput., 2000, 38, 406-415.
[14] Francoise Fessant and Sophie Midenet, Neural Comput & Applic (2002)10:300–310 Ownership and Copyright 2002 Springer-Verlag London Limited.
[15] H. Ritter and T. Kohonen, "Self-Organizing Semantic Maps", Biological Cybernetics, Springer-Verlag 1989.
[16] BERND FRITZKE, Growing Cell Structures A Self-Organizing Network for Unsupervised and Supervised Learning, Neural Networks, Vol. 7, No. 9, PP. 1441-1460, 1994