ASSOCIATIVE SYSTEM TO PREDICT STRUCTURES IN THE IONOSPHERE

Elena Acevedo
Instituto Politécnico Nacional, ESIME Zacatenco, Av. IPN s/n Lindavista, 07738, Mexico City, Mexico
Fabiola Martínez
Instituto Politécnico Nacional, ESIME Zacatenco, Av. IPN s/n Lindavista, 07738, Mexico City, Mexico
Erick Villanueva
Instituto Politécnico Nacional, ESIME Zacatenco, Av. IPN s/n Lindavista, 07738, Mexico City, Mexico
Rodrigo Velázquez
Instituto Politécnico Nacional, CIC, Av. Juan de Dios Bátiz, Esq. Miguel Othón de Mendizábal S/N, Nueva Industrial Vallejo, 07738, Mexico City, Mexico

ABSTRACT

Communications are the most important part of our daily life. The ionosphere play an important role in communications due to the conditions of the ionosphere can affect severely the transmitting and receiving information. Therefore, we propose an intelligent system that can predict accurately structures in the ionosphere. We use a morphological associative model. The obtained results of effectiveness from the Leave One out, Hold Out and Ten-Fold Cross validation test were: 89.45%, 97.77% and 95.83%, respectively, when we use only the *max* memory because *min* memory showed a bad performance.

Keywords: Artificial Intelligence, Associative Memories, Pattern Recognition, Prediction, Ionosphere

INTRODUCTION

Actual globalization and the importance of terrestrial and satellite communications make important any element that affects these types of communications. One of these elements is the condition of the ionosphere, therefore it is constantly monitored. However, current methods require many infrastructure or a high budget due to the constant maintenance.

In this work, we propose a useful tool to identify whether there are structures in the ionosphere in a specific moment in order to predict if a signal can be affected. For example, there is a current error in positioning lecture from a GPS that is about 10 or 20 meters depending on the features of the ionosphere at that time. An accurate prediction could allow us to have a better idea of the situation in a specific time.

We propose to use an associative model that is the Morphological Associative Memory which is based on the basic operations of the mathematical morphology: erosion and dilatation.

Prediction of structures in the Ionosphere

Basic concepts of Ionosphere

The ionosphere (Hunsucker and Hargreave, 2003) is the ionized component of the atmosphere, comprising free electrons and positive ions, generally in equal numbers, in a medium that is electrically neutral. Though the charged particles are only a minority amongst the neutral ones, they nevertheless exert a great influence in electrical properties of the medium, and it is their presence that brings about the possibility of radio communication over large distances by making use of one or more ionospheric reflections.

Looked at most simply, the ionosphere acts as a mirror situated between 100 to 400 Km above the Earth's surface, as in Figure 1, which allows reflected signals to reach points around the bulge of the Earth. The details of how reflection occurs depends on the radio frequency of the signal, but most usual mechanism, which applies in the high frequency (HF) band (3-30 MHz), is actually a gradual bending of the ray towards the horizontal as the refracting index of the ionospheric medium decreases with altitude. Under good conditions, signals can be propagated in this way for several thousand kilometres by mean of repeated reflections between ionosphere and ground.

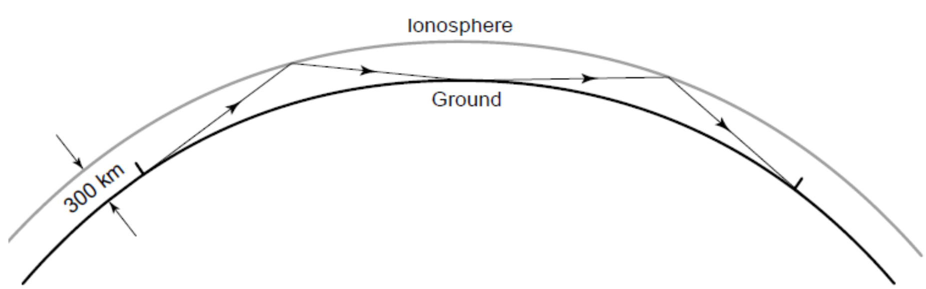

Figure 1. Long distance propagation by multiple hops between the ionosphere and the ground.

Typical vertical profiles of the ionosphere are showed in Figure 2. The identification of the regions was much influenced by their signatures on ionograms, which tend to emphasize inflections in the profile, and it is not necessarily the case that distinct minima separate the various layers.

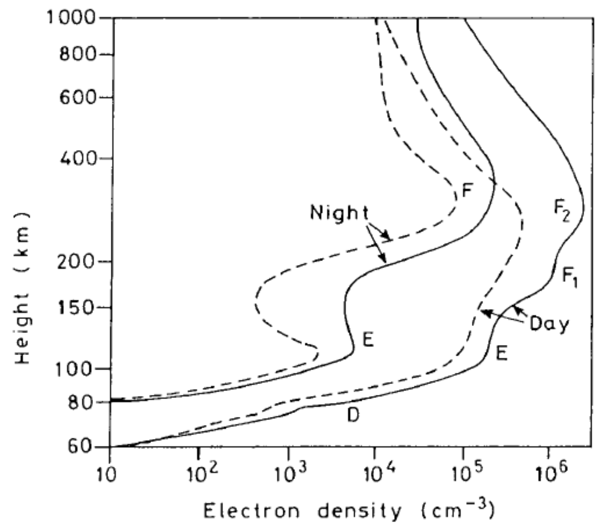

Figure 2. Typical vertical profiles of electron density in the mid-altitude ionosphere: sunspot maximum (continuous line) and minimum (dashed line).

The main regions are designated D, E, F1, and F2, with the following daytime characteristics:

- D region, 60-90 Km: electron density 10^8-10^{10} m^{-3} (10^2-10^4 cm^{-3});
- E region, 105-160 Km: electron density of several times 10^{11} m^{-3} (10^5 cm^{-3});
- F1 region, 160-180 Km: electron density of several times 10^{11} to about 10^{12} m^{-3} (10^5 – 10^6 cm^{-3});
- F2 region, height of maximum variable around 300 Km: electron density up to several times 10^{12} m^{-3} (10^6 cm^{-3});

All these ionospheric regions are highly variable, and in particular there is generally a large change between day and night. The D and F1 regions vanish at night, and E region becomes much weaker. The F2 region, however, tends to persist though at reduced intensity.

During the day, the intensity of ionizing radiation varies with the elevations of the Sun. At night, the source of radiation is remove and so the electron density decays.

Related work
A hybrid classifier (Raymer *et al*, 2003) combine with an evolutionary algorithm was applied to classify some datasets, among them it is the Ionosphere dataset from the UCI machine learning data set repository (https://archive.ics.uci.edu/ml/datasets/Ionosphere) . This data was also used in this work. The algorithm performs feature selection and extraction to isolate the salient features from large data sets. They report the 87.5% of effectiveness when they use 8 features.

Kim and Park (Kim and Park, 2004) proposed a kernelized ionic interaction (IoI) model for data reduction in support vector machines. The instance-based algorithm (IB2) was already applied to select points in the input space. Here, it is further developed it into a kernelized-IB2 (KIB2), where IB2 is applied in the high dimensional feature space based on the kernel. The third is a hybrid KIB2-IoI method where IoI is used as a complementary method to select data points in addition to the data points selected by KIB2. The results show that with the full data set they obtained 95.20%, with IoI the effectiveness was o 94.31%, the 92.26% was the result of applying KIB2 and with KIB2-IoI the effectiveness was of 94.59%. The authors used de 10-Fold Cross validation for testing the performance of their proposal.

The need for finding the number of clusters in conjunction with feature selection, and the need for normalizing the bias of feature selection criteria with respect to dimension was accomplish by the means of FSSEM (Feature Subset Selection using Expectation-Maximization (EM) clustering) and through two different performance criteria for evaluating candidate feature subsets: scatter separability and maximum likelihood (Dy and Brodley, 2004). They obtained the 83.8% of effectiveness with the 10-Fold cross validation test algorithm.

Genetic Programming (Eggermont, Kok and Kosters, 2004) is used to evolve decision trees for data classification, search spaces tend to become extremely large. In this work, authors use different variables to classify the Ionosphere data set. The best result they obtained was the 93.5% of effectiveness with the 10-Fold cross validation test algorithm.

Chung-Jui *et al.* (Chung-Jui *et al.*, 2007) proposed the Particle Swarm Optimization (PSO) for feature selection and SVM as a fitness function to classify some data sets from UCI database, among them is the Ionosphere set. They used the Hold Out method to measure the performance of their algorithm. They reported the 97.33% of effectiveness with only 15 features from the 34 original features.

Clustering Algorithm from the Radial Basis Function (RBF) network architecture was incorporated it into the conventional Hybrid Multilayer Perceptron HMLP network architecture to improve the performance (Mat and Fahmi, 2011). One of the dataset used in this work was the Ionosphere. They applied the Hold-Out method with the 80/20 training/test and they obtained the 96.14% of effectiveness.

The authors proposed a sequential learning algorithm for a neural network classifier based on human meta-cognitive learning principles (Sateesh and Suresh, 2012). The network, referred to as Meta-cognitive Neural Network (McNN). McNN has two components, namely the cognitive component and the meta-cognitive component. Friedman test followed by the Benferroni–Dunn test is used to establish the statistical significance of McNN classifier. The authors used 100 samples for training and 251 for testing, and the best result they obtained was 95.62%.

Recently, a data-based prediction method, Random Bits Regression (RBR) was proposed (Wang *et al.*, 2016). The method first generates a large number of random binary intermediate/derived features based on the original input matrix, and then performs regularized linear/logistic regression on those intermediate/derived features to predict the outcome. The authors used the 10-Fold cross validation test and they obtained the 94.58% of effectiveness.

METHODS AND MATERIALS

In this section, we present the basic concepts of an Associative Memory and the training and recalling phases of the Morphological Associative Memories. Finally, we describe the algorithm of the proposed system.

Associative Memories

An Associative Memory (AM) is a system that works as human memory: it associates patterns to recall them later. For example, we associate faces with names and when we meet a friend we can call him by his name because we recognize the face. We can say that we are eating a guava due to the smell and the taste because we associate that smell and taste from the first time we had a guava. Also, a doctor can diagnose a knee fracture with just to observe an X-ray because he has associated the pattern of a fracture with that event. We can predict the rain when we see a cloudy sky. Therefore, when we have associated stimuli with responses we can recall the response if the stimulus is presented. That is what an associative memory does. In the learning phase, it associates input patterns (stimuli) with output patterns (responses) and then, in the recalling phase, it recalls the corresponding response when a specific stimulus is presented.

The input and output patterns for an AM can be images, strings, or numbers (real, integer or binary), they can be any kind of patterns that can be represented with a number. These patterns are stored in vectors. The task of association of these vectors is called Training Phase and the Recognizing Phase allows recovering patterns. The stimuli are the input patterns represented by the set $\mathbf{x} = \{x^1, x^2, x^3, \ldots, x^p\}$ where p is the number of associated patterns. The responses are the output patterns and are represented by $\mathbf{y} = \{y^1, y^2, y^3, \ldots, y^p\}$. Representation of vectors x^μ is $=$ where n is the cardinality of x^μ. The cardinality of vectors y^μ is m, then $=$. The set of associations of input and output patterns is called the fundamental set or training set and is represented as follows: $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, ..., p\}$.

Morphological Associative Memories

The basic computations occurring in the proposed morphological network (Ritter, Sussner and Diaz de León, 1998) are based on the algebraic lattice structure ($\mathbf{R}$, $\vee$, $\wedge$, +), where the symbols $\vee$ and $\wedge$ denote the binary operations of maximum and minimum, respectively. Using the lattice structure ($\mathbf{R}$, $\vee$, $\wedge$, +), for an m x n matrix A and a p x n matrix B with entries from $\mathbf{R}$, the matrix product $C = A \nabla B$, also called the *max* product of A and B, is defined by equation (1).

$$= \quad + \quad = (\quad + \quad) \quad (\quad + \quad) \tag{1}$$

The *min product* of A and B induced by the lattice structure is defined in a similar fashion. Specifically, the *i,j*th entry of $C = A \Delta B$ is given by equation (2).

$$= \quad + \quad = (\quad + \quad) \qquad (\quad + \quad) \tag{2}$$

Henceforth, let $(\mathbf{x}^1, \mathbf{y}^1), (\mathbf{x}^2, \mathbf{y}^2), \ldots, (\mathbf{x}^p, \mathbf{y}^p)$ be p vector pairs with $x^k = (x_1^k, x_2^k, \ldots, x_n^k)^t \in \mathbf{R}^n$ and $y^k = (y_1^k, y_2^k, \ldots, y_m^k)^t \in \mathbf{R}^m$ for $k = 1, 2, \ldots, p$. For a given set of pattern associations $\{(\mathbf{x}^k, \mathbf{y}^k) \mid k = 1, 2, \ldots, p\}$ we define a pair of associated pattern matrices $(\mathbf{X}, \mathbf{Y})$, where $X=(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^p)$ and $Y=(\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^p)$. Thus, $\mathbf{X}$ is of dimension n x p with *i,j*th entry x_i^j and Y is of dimension m x p with *i,j*th entry y_i^j. Since $\mathbf{y}^k \nabla (-\mathbf{x}^k)^t = \mathbf{y}^k \Delta (-\mathbf{x}^k)^t$, the notational burden is reduced by denoting these identical morphological outer vector products by $\mathbf{y}^k \times (-\mathbf{x}^k)^t$. With these definitions, we present the algorithms for the training and recalling phase.

Training Phase

1. For each p association $(\mathbf{x}^\mu, \mathbf{y}^\mu)$, the minimum product is used to build the matrix $\mathbf{y}^\mu \Delta (-\mathbf{x}^\mu)^t$ of dimensions m x n, where the input transposed negative pattern $\mathbf{x}^\mu$ is defined as $= (\qquad)$.
2. The maximum and minimum operators ($\vee$ and $\wedge$) are applied to the p matrices to obtain M and W memories as equations (3) and (4) show.

$$= \quad (\quad (\quad)) \tag{3}$$

$$= \quad (\quad (\quad)) \tag{4}$$

Recalling phase

In this phase, the minimum and maximum product, Δ and ∇, are applied between memories $\mathbf{M}$ or $\mathbf{W}$ and input pattern $\mathbf{x}^\omega$, where $\omega \in \{1, 2, \ldots, p\}$, to obtain the column vector $\mathbf{y}$ of dimension m as equations (5) and (6) shows:

$$\mathbf{y} = M \oplus \mathbf{x}^\omega \tag{5}$$

$$\mathbf{y} = W \oplus \mathbf{x}^\omega \tag{6}$$

Ritter, G. X., Sussner, P. and Diaz de León, J. L. (1998). Morphological Associative Memories. IEEE Transactions on Neural Networks 9(2): 281-293.

Description of the System

Dataset

The radar data were collected by the Space Physics Group of The Johns Hopkins University Applied Physics Laboratory. The radar system, located in Goose Bay,

Labrador, consists of a phased array of 16 high-frequency antennas, with a total transmitted power about 6.4 kW and an antenna gain of about 30 dBm at frequency ranges of 8 to 20 MHz. The radar returns are used to study the physics of the ionosphere at the E- and F-layers (100- to 500-km altitude). The targets, free electrons in the ionosphere, have small cross sections about 10^{-30} m^2. A typical number density of electrons would be about $10^8/m^3$, and the total volume could be as large as 10^6 m^3.
In general, good returns are indicated by well defined signals, which are evidence of the presence of some type of structure in the ionosphere. Bad returns can be caused by the absence of identifiable structure (the signal passes through the ionosphere); by incoherent scattering (signals are reflected from too many structures, resulting in phase cancellation); by the absorption of radar pulses; and by interference from other transmitters. Bad returns are more diverse than good ones.
Received signals were processed using an autocorrelation function whose arguments are the time of a pulse and the pulse number. There were 17 pulse numbers for the Goose Bay system. In this dataset, instances are described by two attributes per pulse number, corresponding to the complex values returned by the function resulting from the complex electromagnetic signal. Therefore, there are 34 continuous attributes and additional attribute is used for the good or bad returns classification.

There are 351 instances from which, 225 are classified as good returns and 126 instances are bad returns. This dataset is unbalanced.

Prediction system
The input patterns for the Morphological Associative Memory (MAM) will be the instances from dataset. Then, each vector will be of dimension of 34. The dimension of the output patterns will depend on the number of instances that will be used to train the memory, for example, if we train the MAM with 100 instances, then the dimension of the output patterns will be of 100. The values of the elements of the output vectors will depend on the type of the memory, i.e., if we train a *max* memory then just one of the elements will be 500 and the remain will be 0. If we are training a *min* MAM, one of the elements will have the value of -500 and the remain will have the value of 0. We illustrate the training and recalling phases with the following example. We have the input patterns,

= , = , =

For *max* MAM, the output patterns are,

= , = , =

And for the *min* Morphological Associative Memory,

$$- = \quad , \quad - = \quad , \quad - =$$

Now we apply equation (3) to build *max* memory.

$$= \quad (\quad) = \quad =$$

$$=$$

$$= \quad (\quad) = \quad =$$

$$=$$

$$= \quad (\quad) = \quad =$$

$$=$$

$$=$$

$$=$$

Then we apply equation (4) to build *min* memory

$$=$$

$$=$$

Finally, we recall pattern y^1 by presenting input pattern x^1 and applying equations (5) and (6).

$$= \qquad\qquad =$$

$$= \quad + \quad + \quad + \quad =$$
$$+ \quad + \quad +$$
$$+ \quad + \quad +$$

$$=$$

After this, we set the values different of 500 to 0, then we obtain the corresponding output pattern y^1.

$$= \quad \rightarrow \quad =$$

Now we present the pattern x^1 to the *min* MAM.

$$= \qquad\qquad =$$

$$= \quad + \quad + \quad =$$
$$+ \quad +$$
$$+ \quad +$$

$$=$$

In a similar way, we set to 0 all the values in the result vector that are different to -500, then we obtain the corresponding pattern y^1.

$$= \quad \rightarrow \quad = -$$

In both cases, *min* and *max* memories recalled the corresponding output patterns.

In the following section we present the results from the tests we accomplished.

EXPERIMENTS AND RESULTS

We implement the software in a laptop with an Intel CORE i5 processor and with language programming Visual Studio 2013 C#.

We used three algorithms to test our proposal: Leave-One-Out, Hold-Out and K-Fold Cross Validation.

In the first case, we took the first record out for testing and we trained the MAM with the 350 records remain. Then, we took the second record out for testing and we use the other 350 for training. We did 351 tests, the average is presented in Table I. We can observe that the *good* class had a better recall than *bad* class with both memories. The total recall was of 89.45% with *max* memory and 78.63% with *min* memory.

Table I. Results from the Leave-One-Out test.

Class	Rec. Max	Max	Rec. Min	Min
Good	221/225	98.22%	218/225	96.88%
Bad	93/126	73.80%	58/126	46.03%
Total	314/351	89.45%	276/351	78.63%

For the Hold-Out test, we used the 10% of the records for training and the 90% for testing. Then, we took the 20% of the records for training and the 80% for testing, and so on. The selection of the records is done randomly. The results are presented in Table II.

In general, we can observe from Table II that we had better recall in *good* class and that the *max* memory presented the best recall in both classes. The result from the 80/20 (training/test) is 97.77% of recalling.

In the final test, we chose $K = 10$, this means that we built ten blocks of 35 records (we did not include one record from the *good* class due to is the class with more records), the blocks were built in a random way. After, we used the first block for testing and the remaining nine blocks were used to train the MAM. Then, we used the second block for testing and the remaining records for training and so on. The results can be observed in Table III. In the first column, we can see the number of the block that was used to test the memory.

Table II. Results from the Hold-Out test.

Training (%)	**Good Class**				**Bad Class**			
	Max recall		*Min* recall		*Max* recall		*Min* recall	
10	173/202	85.64%	183/202	90.59%	108/114	94.73%	66/114	57.89%
20	151/180	83.88%	157/180	87.22%	83/101	82.17%	72/101	71.28%
30	148/157	94.42%	149/157	94.90%	72/88	81.81%	48/88	54.54%
40	129/135	95.55%	131/135	97.03%	62/76	81.57%	38/76	50%

50	107/112	95.53%	109/112	97.32%	52/63	82.53%	32/63	50.79%
60	87/90	96.66%	87/90	96.66%	41/50	82%	26/50	52%
70	64/67	95.52%	64/67	95.52%	34/38	89.47%	20/38	52.63%
80	44/45	97.77%	44/45	97.77%	23/25	92%	13/25	52%
90	22/22	100%	22/22	100%	11/13	84.61%	5/13	38.46%

Table III. Results from the 10-Fold Cross Validation.

Block	Good Class				Bad Class			
	Max recall		*Min* recall		*Max* recall		*Min* recall	
1	23/23	100%	23/23	100%	8/12	66.66%	6/12	50%
2	23/23	100%	23/23	100%	11/12	91.66%	8/12	66.66%
3	22/23	95.65%	22/23	95.65%	10/12	83.33%	9/12	75%
4	23/23	100%	22/23	95.65%	5/12	41.66%	5/12	41.66%
5	22/23	95.65%	22/23	95.65%	7/12	58.33%	3/12	25%
6	22/23	95.65%	23/23	100%	7/12	58.33%	4/12	33.33%
7	23/23	100%	23/23	100%	4/12	33.33%	4/12	33.33%
8	22/23	95.65%	22/23	95.65%	10/12	83.33%	5/12	41.66%
9	23/23	100%	22/23	95.65%	11/12	91.66%	8/12	66.66%
10	18/18	100%	18/18	100%	12/17	70.58%	9/17	52.94%
Average		98.26%		97.82%		67.88%		48.62%

From Table III, we can observe that in this case, the best results are found in good class and they are obtained with both memories. The average shows a 98.26% of effectiveness with only one mistake in the *good* class, the best result was of 100% of recalling. In the case of the *bad* class, the average with *max* memory was of 67.88% of recalling and the best result was of 91.66%. The Table III showed that the worst performance for *bad* class was obtained with the *min* memory.

In Table IV we show the comparisons with the other approaches mentioned in the sub section of Related work.

From Table IV, we can observe that when we used the Hold Out method test (80/20) we obtained the 97.77% of effectivity while the PSO-SVM (15 features) and the Clustered-Hybrid Multilayer Perceptron obtained 97.33% and 96.14%, respectively. In the case of the 10-Fold cross validation test, the average of our best results is 95.83% of effectiveness, almost the same result from McNN.

Table IV. Comparisons with the results from different approaches.

Year	Algorithm	Method test	Efectiveness (%)
2003	Hybrid classifier	----	87.5
2004	Full set IoI KIB2 KIB2-IoI	Ten-Fold cross validation	95.20 94.31 92.26 94.59
2004	FSSEM	Ten-Fold cross validation	83.8

2004	Genetic programming	Ten-Fold cross validation	93.5
2007	PSO-SVM (15 features)	Hold-Out (80/20)	97.33
2011	Clustered-Hybrid Multilayer Perceptron	Hold-Out (80/20)	96.14
2012	McNN	Friedman and Benferroni–Dunn tests	95.62
2016	Random Bits Regression	Ten-Fold cross validation	94.58
2016	Our proposal	Leave One Out Hold Out (80/20) Ten-Fold cross validation	89.45 97.77 95.83

CONCLUSIONS

Morphological associative memories (MAM) are suitable tools to predict structures in the ionosphere. Our results are competitive with other approaches and our algorithm is less complex than others as we can see with the illustrative example. We showed that we can predict structures in the ionosphere with an accuracy of 97.77%.

We observed that, in this case, *min* MAM had a poor performance.

In the three cases of the test methods, the worst results are for the *bad* class. The reason is because the dataset is unbalanced, i.e., the *good* class has more records than the *bad* class, almost the double of the records.

REFERENCES

Chung-Jui, T., Li-Yeh, C., Jun-Yang, C., and Cheng-Hong, Y. (2007). Feature Selection using PSO-SVM. *IAENG International Journal of Computer Science*. 33:1.

Dy, J. G. and Brodley, C. E. (2004). Feature Selection for Unsupervised Learning. *Journal of Machine Learning Research*. 5: 845–889.

Eggermont, J., Kok, J.N. and Kosters, W.A. (2004). Genetic Programming for Data Classification: Partitioning the Search Space. *Proceedings of the 2004 ACM symposium on Applied computing*. 1001-1005.

Hunsucker, R. D. and Hargreaves, J. K. (2003). *The high-latitude ionosphere and its effects on radio propagation.* Cambridge University Press 2003.

Kim, H. and Park, H. (2004). Data Reduction in Support Vector Machines by a Kernelized Ionic Interaction Model. *Conference: Proceedings of the Fourth SIAM International Conference on Data Mining*, Lake Buena Vista, Florida, USA, April 22-24, 2004.

Mat, N. A. & Fahmi, W.M. (2011). Clustered-Hybrid Multilayer Perceptron network for pattern recognition application. *Applied Soft Computing*. 11: 1457–1466.

Raymer, M., Doom, T., Kuhn, L. & Punch, F. (2003). Knowledge Discovery in Medical and Biological Datasets using a Hybrid Bayes Classifier/Evolutionary Algorithm, *IEEE Transactions on systems, man and cybernetics Part* B.

Ritter, G. X., Sussner, P. and Diaz de León, J. L. (1998). Morphological Associative Memories. *IEEE Transactions on Neural Networks*. 9(2): 281-293.

Sateesh, G. & Suresh, S. (2012). Meta-cognitive Neural Network for classification problems in a sequential learning framework. *Neurocomputing*. 81: 86–96.

Wang, Y., Li, Y., Xiong, M., Shugart ,Y.Y., and Jin., L. (2016). Random bits regression: a strong general predictor for big data. *Big Data Analytics*. 1:12.