

# Clustering of Sedimentary Basins Using Associative Neural memories (ART2)

LAKSHMIPRASAD BOPPANA

Student, (M.Tech)

Department of Computer Science&Engineering, Chaitanya Engineering College  
Visakhapatnam, Andhra Pradesh, India

E-mail: [laxmiprasadboppana@gmail.com](mailto:laxmiprasadboppana@gmail.com)

B.POORNA SATYANARAYANA

Associate Professor

Department of Computer Science&Engineering, Chaitanya Engineering College  
Visakhapatnam, Andhra Pradesh, India

E-mail: [poornasatyanarayana@gmail.com](mailto:poornasatyanarayana@gmail.com)

**Abstract--Associative Memory (AM) research covers technologies enabling implementation of associative memory which enables thought process and links previous experience to novel situations. Each neural network system requires a memory for storing and retrieval of associated concepts, based on a combination of the base concept and the context. Adaptive Resonance Theory is a kind of associative neural memory model as unsupervised neural network model. The aim of this article is to present identification and recognition of Magneto-telluric data for sedimentary basins using associative neural memory with Adaptive Resonance Theory (ART2).The ART2 is an unsupervised learning algorithm where the network is provided with inputs but not with desired outputs. The system itself to decide what features it will use to group the input data. Several sets of data consisting of 17 phases and 17 apparent resistivity values and their respective tag values are given. These sets of data are used for training the network, and other sets of data are used to test the network for clustering. The testing will result in the approximate identification of the data patterns with tag value of 1 where there is sediment of hydrocarbon and a tag value of 0 where there is no sediment of hydrocarbon in the given data set. The recognition rate in the proposed system lies between 90% and 100%.**

**Keywords: Sedimentary Basins, Artificial Neural Networks, ART2, Associative Memory**

## I. INTRODUCTION

This paper describes two major areas of research like “Geo Physical Sciences” and “Neural Network Applications”. These two areas are very broad in nature[1].The present system of testing the sedimentary basins at National Geophysical Research Institute (NGRI) uses the manual approach. The data collected when processed manually needs a lot of time since the data needs to be converted into normalized values and then the resultant data is studied to identify the presence of sediment. This process requires lot of flat files to be processed manually. The retrieval of past data at any certain point of time results in a manual search through the vast number of data records. Since the data is stored and retrieved manually there is always the scope for error increases, data validation also becomes very hard. In the first phase the data (sediment) is segregated into various clusters which are characterized into a number which specifies the pattern of one specific value. These values are gathered into the file which we called as “original file”. This file is processed through manual approach to make it into normalized approach. After the normalized data gets divided into various patterns which is again the task of the human being. Finally at the later stage this is verified for the presence of hydrocarbons using software application.

## II. ANALYSIS

The analysis of this paper has been done mainly on two contrasting subjects.

**1. Magneto** – telluric data

**2. Neural Networks**

Magneto telluric data deals with a branch of Geophysics, which is used by Geophysicists to identify the formation and classification of deposits. Neural Network is used to automate the identification of deposits with either a positive answer (the identification of deposits) or a negative answer (no deposits present). Neural networks are also used to speed up the process of deposit identification which actually takes lots of time if done

manually, there by the objective of this paper is to identify the presence of deposits and speed up the process of identification there by reducing the time delay [2]. In the first phase Magneto telluric data is being identified with presence of phase and apparent resistivity values which are divided into 17 values each respectively in the given data set. The data i.e. the phase and apparent resistivity values should be converted into logarithmic values and the resultant data should be normalized. The data also consists of a tag value, which represents the presence of sediment with a tag value of 1(one) and identification of no sediment with a tag value of 0 (zero). In the second phase, the Neural Network analysis to the analysis on the subject has been done to identify the function to be used for computation in the networks

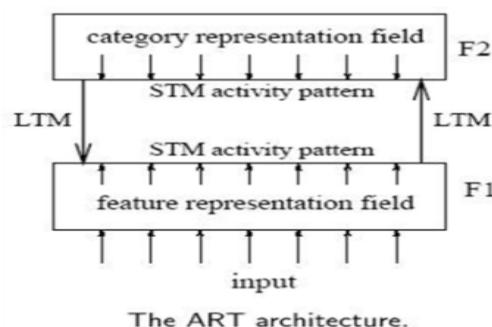
### III. ADAPTIVE RESONANCE THEORY

A feature of Cognitive systems is that they can be receptive to new patterns (as Plasticity) but remain unchanged to others (as Stability). This is referred to as the *stability/plasticity dilemma* (Grossberg 1976) developed a theory called *adaptive resonance theory* (ART). In terms of design, ART networks are the most complex Artificial Neural Networks (ANN). Caudill & Butler (1990) regard this process as a type of hypothesis test. A pattern presented at an input layer is passed to a second layer, which is interconnected to the first. The second layer makes a guess about the category to which the original pattern belongs, and this hypothetical identity is passed back to the first layer. The hypothesis is compared with the original pattern and if found to be a close match, the hypothesis and original pattern reinforce each other (resonance takes place). But if the hypothesis is incorrect, the second layer produces another guess. If the second layer cannot eventually provide a good match with the pattern, the original pattern is learned as the first example of the new category.

The model has three crucial properties:

- a *normalization* of the total network activity. Biological systems are usually very adaptive to large changes in their environment. For instance, the human eye can adapt itself to large variations in light intensities;
- *contrast enhancement* of input patterns. The awareness of subtle differences in input patterns can mean a lot in terms of survival. Distinguishing a hiding panther from a resting one makes all the difference in the world. The mechanism used here is contrast enhancement;
- *short-term memory* (STM) storage of the contrast-enhanced pattern. Before the input pattern can be decoded, it must be stored in the short-term memory. The *long-term memory* (LTM) implements an arousal mechanism (i.e., the classification), whereas the STM is used to cause gradual changes in the LTM.

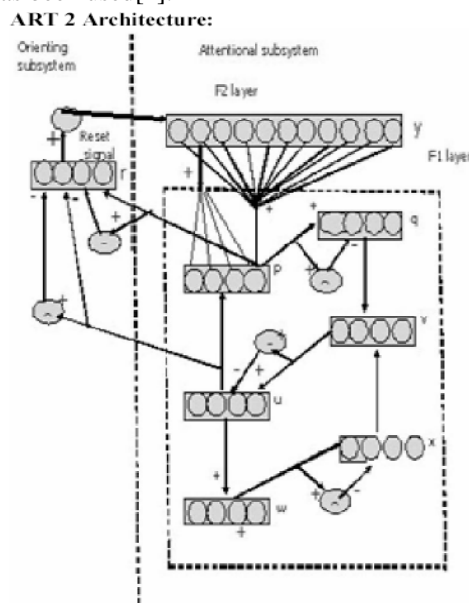
The system consists of two layers,  $F1$  and  $F2$ , which are connected to each other via the LTM. The input pattern is received at  $F1$ , whereas classification takes place in  $F2$ . As mentioned before, the input is not directly classified. First a characterization takes place by means of extracting features, giving rise to activation in the feature representation field. The expectations, residing in the LTM connections, translate the input pattern to a categorization in the category representation field. The classification is compared to the expectation of the network, which resides in the LTM weights from  $F2$  to  $F1$ . If there is a match, the expectations are strengthened, otherwise the classification is rejected. A simple model of the adaptive resonance theory is presented below.



A various forms of ART model now a-days to solve pattern-classification domains, are ART, ART1, ART2, ART3,... and so on that depends upon the task to implement. Among these models I considered ART2 algorithm for clustering the sedimentary basins is presented below.

**3.1 ART2 [3]** is a Neural Network and is operated in an unsupervised learning mode. It typically consists of a comparison field and a recognition field composed of neurons, a vigilance parameter, and a reset module. The vigilance parameter has considerable influence on the system: higher vigilance produces highly detailed memories (many, fine-grained categories), while lower vigilance results in more general memories (fewer, more general categories). The comparison field takes an input vector (a one-dimensional array of values) and transfers it to its best match in the recognition field. Its best match is the single neuron whose set of weights (weight vector) most closely matches the input vector. Each recognition field neuron outputs a negative signal proportional to that neuron's quality of match to the input vector) to each of the other recognition field neurons and inhibits their output accordingly. In this way the recognition field exhibits lateral inhibition, allowing each neuron in it to represent a category to which input vectors are classified. After the input vector is classified, the reset module compares the strength of the recognition match to the vigilance parameter. If the vigilance threshold is met, training commences. Otherwise, if the match level does not meet the vigilance parameter, the firing recognition neuron is inhibited until a new input vector is applied; training commences only upon the completion of a search procedure. In the search procedure, recognition neurons are disabled one by one by the reset function until the vigilance parameter is satisfied by a recognition match. If no committed recognition neuron's match meets the vigilance threshold, then an uncommitted neuron is committed and adjusted towards matching the input vector. We have two methods of training ART-based neural networks: slow and fast. In the slow learning method, the degree of training of the recognition neuron's weights towards the input vector is calculated to continuous values with differential equations and is thus dependent on the length of time the input vector is presented. While fast learning is effective and efficient for a variety of tasks, the slow learning method is more biologically plausible and can be used with continuous-time networks. There are 2 types of ART. ART 1 is the simplest variety of ART networks, accepting only binary inputs. ART 2 extends network capabilities to support continuous inputs.

In this paper ART 2 architecture has been used[4].



#### ART 2 Architecture:

##### Algorithm:

- Let  $M$  be the number of units in each F1 sublayers and  $N$  be the number of the units on F2 layer.
- Parameters are chosen according to following constraints:

$$a, b > 0, 0 \leq d \leq 1, \left( \frac{a+d}{1-d} \right) \leq 1, 0 \leq \theta \leq 1, 0 \leq p \leq 1, \theta \leq 1$$

- Top down weights all initialized to zero,  $Z_{ij}(0) = 0$
- Bottom up weights are initialized,  $Z_{ij}(0) < 1/(1-d)M$

**STEP 1:** Initialize all layers and sub-layers output to zero vectors, and establish a cycle counter initialized to a value of one.

**STEP 2:** Apply an input pattern,  $I$  to the  $w$  layer of F1. The output of this layer is

$$W_i = I_i + a * u_i$$

- STEP 3:** Propagate forward to the x sub layer,  $X_i = W_i / (p + |w|)$
- STEP 4:** Propagate forward to the v sub layer,  $V_i = f(X_i) + b * f(q_i)$
- STEP 5:** Propagate to the u sub layer,  $U_i = V_i / (s + |v|)$
- STEP 6:** Propagate forward to the p sub layer,  $P_i = U_i + d * Z_{ij}$
- STEP 7:** Propagate to the q sub-layer,  $q_i = P_i / (s + |v|)$
- STEP 8:** Repeat steps 2 through 7 as necessary to stabilize the values on F1
- STEP 9:** Calculate the output of the r layer,  $R_i = (U_i + cP_i) / (s + u + cp)$
- STEP 10:** Determine whether a reset condition is indicated. If  $p / (s + r) > 1$ , then send a reset signal to F2
- STEP 11:** Propagate the output of the p sub-layer to the F2 layer.  
Calculate the net inputs to F2
- $$T_j = \sum (p_i * Z_{ij}), \text{ where } i = 1, 2, \dots, m.$$
- STEP 12:** Only the winning F2 node has nonzero output.
- $$G(T_j) = d * T_j = \begin{cases} \max(T_j) \\ 0, & \text{otherwise} \end{cases}$$
- STEP 13:** Repeat steps through 6 to 10
- STEP 14:** Modify bottom up weights on the winning F2 Unit,  $Z_{ij} = u_i / (1 - d)$
- STEP 15:** Modify top down weights coming from the winning F2 Unit.
- STEP 16:** Remove the input vector. Restore all inactive F2 units.  
Return to step 1 with new input pattern

#### IV. DATA NORMALIZATION

From the analysis we arrived at an understanding that the given data set needs to be normalized for further processing. The normalization of the data is achieved with the help of java program[5]

- The given data set is stored in a file
- The file is given as input to the java program
- The program generates an output in the form of normalized data
- The output is stored in another file.

$$\text{Normalized data} = \frac{(H_i - H_{min})}{(H_{max} - H_{min})}$$

#### V. PATTERN CREATION

The following is the pattern file format

```
#Input Pattern i
Values -----
# Input Pattern i
Values -----
```

Since the above described is the format for a pattern file. The normalized file needs to be converted into the pattern file format. To achieve this java program is coded which takes normalized file as input and generates a Pattern file as output.

#### VI. NETWORK CREATION

After creating pattern files, we need to create ART2 network. ART2 network is divided into two subsystems

1. Attentional subsystem
2. Orienting subsystem

##### 6.1 Intentional Subsystem

The Attentional Subsystem consists of two layers

- F1 layer
- F2 layer

The input is given to F1 layer, where normalizing of the input vector is performed. In F2 layer the process of clustering takes place (pattern creation)

**F1 layer**

The F1 layer consists of six sub layers. They are

**Layer 1-W layer:** Input is given to this layer. The output of this layer is,  $W_i = I_i + a * u_i$

**Layer 2-X layer:** The output of the W layer is given to this layer. The output of this layer is

$$X_i = W_i / (s + |w|)$$

**Layer 3-V layer:** The output of the X layer is given to this layer. The output of this layer is

$$V_i = f(X_i) + b * f(q_i)$$

**Layer 4-U layer:** The output of V layer is given to this layer. The output of this layer is

$$U_i = V_i / (s + |v|)$$

**Layer 5-P layer:** The output of U layer is given to this layer. The output of this layer is

$$P_i = U_i + d * Z_{ij}$$

**Layer 6-Q layer:** The output P layer is given to this layer. The output of this layer is

$$q_i = P_i / (s + |v|)$$

**F2 layer**

In this layer, the process of clustering is done. Each neuron in this layer consists of a unique pattern. If the entered pattern matches with anyone of the existing patterns in neurons then it is placed in the same neuron. Otherwise a new neuron is created and the entered pattern is stored in that.

**6.2 Orienting Subsystem**

This is used to reset the layers of attentional subsystem. During the comparison, whenever the mismatch occurs, the orienting subsystem gets activated and resets the layers of attentional subsystems are, deactivating the neurons in the layers.

## VII. TRAINING

Here we try different values of parameters and train the network to get the desired output.

## VIII. TESTING

The network which is stabilized with the parameters is tested with new data files for clustering.

**8.1 System Testing**

All the modules were first tested individually using both test data and live data. After each module was ascertained that it was working correctly it was integrated with the system, again the system was tested as a whole.

## IX. RESULTS

The results of this work are carried out in several phases. The following is the sequence of the procedure involved in getting the results.

1. In the first phase sedimentary data which is in the form of text file is supplied to the application/program.
2. The application/program using ART2 algorithm which is implemented through Java program normalizes the text file.
3. After the data normalization it gets divided into various clusters
4. Testing Sedimentary Basins Using Adaptive Resonance Theory
5. Once the data gets into clusters we supply this data to the neural network for its creation
6. Once the network gets trained it can be tested for the results.

The above representation which exhibits the values from the neural network which lie in between 0 and 1 and can emphasize the presence of magneto telluric data in the sedimentary soils.

## X. CONCLUSION

The article "clustering the sedimentary basins using associative neural memories ART2" is carried out to support the field of oil explorations in a precise way to help in the identification of hydrocarbon sediment deposits in any area. This paper is optimized to help in overcoming the drawbacks of traditional ways of processing data manually, thereby resulting in advantages such as no time delaying calculating the presence of sediment and faster computerized report generation and also helpful in faster retrieval of past data in data records. Accuracy of the Pattern classification of Magneto telluric data at 95% success rate in identifying either

the presence of hydrocarbon deposits or in the identification that there is no presence of hydrocarbon deposits. The error rate of 3% is considered to be very negligible. This paper is made capable to process data and make its output available in a matter of seconds as against the traditional way of manually computing which may end days together.

#### XI. REFERENCES

- [1] "Introduction to Geomagnetism" by W.D.Parkinson-Scottish Academic press-1983.
- [2] "Neural Networks" by Simon Haykin, Pearson Education Asia Publication-1999.
- [3] "Adaptive Resonance Theory " by Carpenter, G.A.& Grossberg, S. (2003), In M.A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, Second Edition (pp. 87-90). Cambridge,MA: MIT Press.
- [4] Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps Carpenter, G.A., Grossberg, S.,Markuzon, N., Reynolds, J.H., & Rosen, D.B.(1992), , IEEE Transactions on Neural Networks, 3, 698-713.
- [5] "Neural Networks A Classroom Approach" by Satish Kumar, Tata-McGraw Hill.

#### AUTHORS PROFILE

**Lakshmiprasad Boppana** is M.Tech student at Chaitanya Engineering College, Visakhapatnam, India.

**B.Poorna Satyanarayana** is Associate Professor at Chaitanya Engineering College, Visakhapatnam, India