Scientia Africana, Vol. 15 (No. 1), June 2016. Pp 1-10
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ***ISSN 1118 – 1931***

ONLINE CREDIT CARD FRAUD PREDICTION BASED ON HIERARCHICAL TEMPORAL MEMORY MODEL

[1]L. U. Oghenekaro and [2]P. Enyindah

[1]*Department of Computer Sciences,*
University of Port Harcourt,
Choba, Nigeria.
E-mail: linda.igwebike@uniport.edu.ng
Telephone: +2348059039323

[2]*Department of Computer Sciences,*
University of Port Harcourt,
Choba, Nigeria.
E-mail: promise.enyindah@uniport.edu.ng
Telephone: +2348036710489

Received: 12-01-16
Accepted: 24-05-16

ABSTRACT

Recent studies of the human brain have brought about a new understanding of the structural and algorithmic property of the neocortex. This understanding gave birth to the Hierarchical Temporal Memory (HTM) which holds a lot of promises in the area of time-series prediction and anomaly detection problems. This paper demonstrates the behaviour of an HTM model with respect to its learning and prediction of online credit card fraud. The model was designed using the object oriented analysis and design methodology. Java programming language was used for implementation and Matlab was used to carry out simulations. The resulting model demonstrated learning like that of the human brain using sparse data, hence, the model required less resources to evaluate big data. The model recorded higher accuracy of 92.3% compared to Artificial Neural Network model that recorded 89.6%, hence reducing cases of misclassification.

Key words: Duty Cycle, Boost Value, Synapses, Learned Index

INTRODUCTION

The rapid and steady increase of online activities and electronic commerce in general has led to an increase in the number of credit card users. The increasing number of credit card users brings about an increase in the crimes committed with the credit card. The nature of the online credit card transactions unlike other means of transaction is very convenient and has a lot of advantages which includes; a quick way to borrow funds, borrowing for free, incentives, useful in times of emergencies and flexibility. However, its major disadvantage is that it is vulnerable to fraud. This paper is delimited in scope to predicting credit card fraud in the financial sector under the online credit card fraud category. This is because this category is of greater impact and perpetuators tend to carry out fraud in this manner as it is more convenient for them, in the sense that they do not need to buy weapons to rob a shop, neither do they need to make known their location or identity. All they need to carry out their criminal act is an internet connection and a computer. Credit card fraud poses a big threat to business establishments of today and is a serious problem throughout the world, with the increasing card not present scenario, activities such as shopping online and e-commerce transactions are no longer protected with the advantages of physical mandate verification. According to a survey, Bhatla et. al. (2003), the rate at which internet fraud occurs is 12 to 15 times higher than physical world fraud and detecting CCF have two typical characteristics: first, the huge amount of credit card transactions to be screened within a given time knowing that only a few or no fraudulent transaction will be detected and secondly, the very limited time span in which the rejection or acceptance decision has to be made, Patidar and Sharma (2011). Companies and institutions loose very huge amount of money annually due to fraud, and fraudsters are continuously devising new means in carrying out fraudulent activities using the credit card, Patidar and Sharma (2011).

Hierarchical Temporal Memory (HTM) Theory

The HTM theory came to light from recent studies done on the human brain; these studies bring a new understanding of the brain which in turn will lead to the creation of machines that are truly intelligent, Hawkins and Blakeslee (2007). The brain is a very complex organ and also a power house, composed of various structures. The HTM theory draws its study from a structure in the brain that is in charge of intelligence, a structure referred to as the seat of intelligence, called the neocortex. It is located at the outer part of the brain, where it surrounds and envelops almost all other parts of the brain. Figure 1 depicts the position of the neocortex in the human brain. The HTM is particular about the neocortex because of its function in (intelligence) and also, because all the essential aspect of intelligence also occur at the neocortex, although, two other regions, the thalamus and the hippocampus play vital roles too, Ganong (2003). The neocortex of the human brain is larger than that of other mammals. It is a thin sheet of neural tissue, made up of six layers that are formed by variation in the densities of cell bodies, cell types and their connection, and each layer is loaded with nerve cells that are tightly packed. Figure 2 reveals the layers of the neocortex. These tightly packed cells contain knowledge, memories and accumulated life experiences, Ganong (2003). Physically, they are arranged in an irregular patchwork quilt and it varies a little for different individuals. Functionally, they are arranged in a branching hierarchy. The lowest of the functional regions is where the sensory information first arrives in the cortex and is processed at its rawest form and passed on to higher regions. However, information can also flow from higher regions to lower regions via other connections. Figure 3 shows the flow of information in a four-level hierarchy of an HTM.

Scientia Africana, Vol. 15 (No. 1), June 2016. Pp 1-10
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

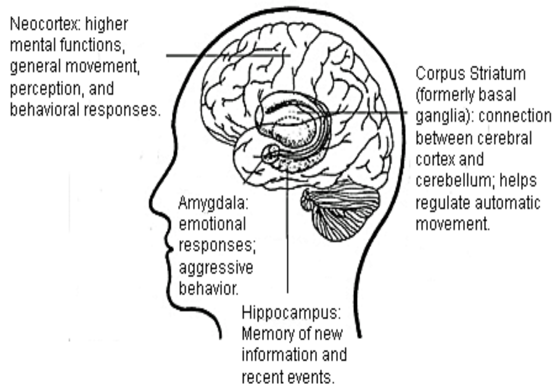

Figure 1: The position of the neocortex in the human brain, Ganong (2003)

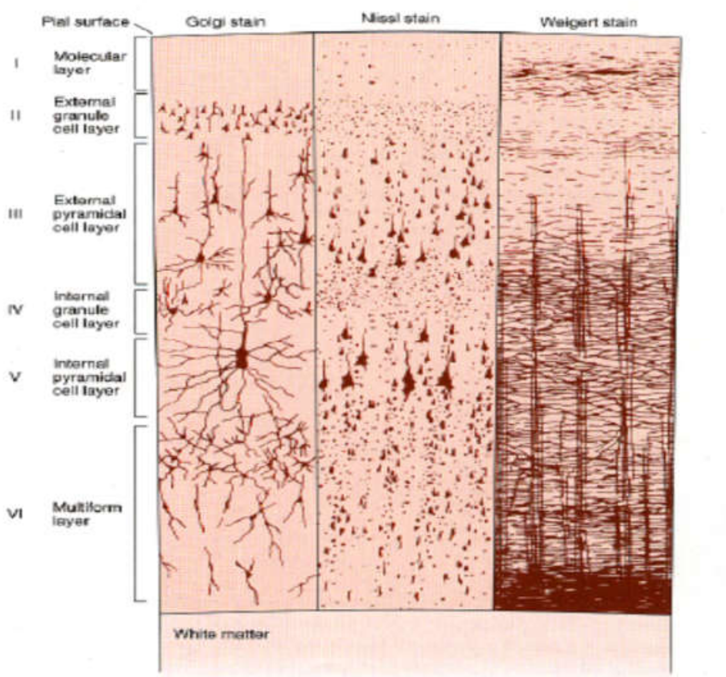

Figure 2: Different staining methods revealing the various layers in the neocortex. Ganong (2003)

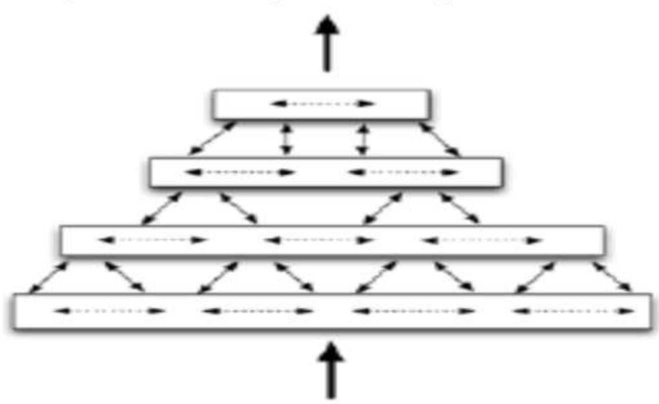

Figure 3: A four-level hierarchy in an HTM, Numenta (2011).

The HTM theory is the concept that gave rise to the cortical learning algorithm (CLA). J. Hawkins and D. George of Numenta Inc developed the HTM which is an online machine learning model; the model copies some of the algorithmic and structural properties of the neocortex. In the book On Intelligence, Hawkins and Blakeslee (2007), Hawkins developed the initiative to create a simple model of the neocortex by simulating its core function rather every function, Galetzka (2014), however the idea of simulating a part of the brain have been tried before, for instance, the neural networks did this by using artificial neurons and synapses as its building blocks, Hopfield (1988). The complexity of the brain in terms of physiological and anatomical details are the previous challenges faced in earlier algorithms trying to model a part of the brain, hence it is difficult to determine exactly what is required in allowing proper functioning in a living being, George (2008). Due to this complexity, the derived models often lack some basic biological features such as feedback connection, and even those that have this feature have trouble with values varying over time, hence, they are used for only static problems such as classification. Chappelier et. al. (2001).

The hierarchical temporal memory gives a better representation of the biological neurons in the human brain than the neural network.

The following are core elements of the HTM architecture:

1. CELL: This is the smallest element of the hierarchy and it is responsible for most of the computation. The cells can be likened to the neurons of ANNs. A cell uses the input from the dendrites in computing its output with a threshold function, Galetzka (2014).
2. COLUMN: A column is made up of several cells and the combined output of these cells either activates the column or not. Once a column is active, it hinders columns around it from becoming active, hence creating a sparse distributed representation (SDR). The columns are likened to columnar organized structures of the neurons in the neocortex, Numenta (2011).
3. REGION: A region is made up of fixed number of columns. The output of a region is a sparse distributed representation which is as a result of how columns in one region influence each other.
4. LAYER: The layer is the biggest organizational unit in the HTM architecture and they can be likened to the layers present in ANNs, however, multiple layers are not needed for a HTM model to work effectively, with large data, as it is with the ANN, since each layer in a HTM model comprises of one or more regions.
5. DENDRITE: The dendrite which is similar to the synaptic connections in ANNs is used to connect different part of the HTM, however, the dendrite in the HTM architecture are more complex as it more closely modeled to the biological structure of the brain. The dendrites in the HTM exists in two types: first is the proximal dendrites which carry the unified output of columns to individual cells in the next higher level and the second type is the distal dendrites that carry output of cells to other cells within same region. Galetzka (2014).

Scientia Africana, Vol. 15 (No. 1), June 2016. Pp 1-10
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

MATERIALS AND METHOD

Experiment/Implementation

The methodology employed in this study was the object oriented analysis and design methodology because of its recursive nature and object-oriented development method. The Cortical Learning Algorithm that the HTM region uses in learning patterns as it keeps account of spatial and temporal variability was used in developing the HTM model. The algorithm is as follows, Galetzka (2014):

1. Sensor input is to be encoded into a sparse representation and received into the HTM region.
2. The Spatial Pooler (SP) function finds a single representation for patterns that are spatially similar, Numenta (2011), the result is a sparse vector of active columns in the region, the function achieves this by following the sub steps;

   i. Every column is connected with a proximal dendrite to a unique subset of the input bits using synapses. Each synapse has a permanence value which is used to determine if the synapse is active. The overlap function then computes the overlap score for each column by counting the synapses with high permanence value.
   ii. Column with a high overlap score are activated and hinders the activation of columns near them. Learning then takes place by adjusting synapse permanence value of all columns, the permanence of synapses is increased if they are connected to active input bits, otherwise it is decreased.
3. The Temporal Pooler (TP) function represents the input's context and form predictions, the function achieves this by following the sub steps;

   i. Each active column activates cells in their predictive state from the previous input, if no cell is in its predictive state the all cells of the active column are activated to represent a novel input.
   ii. By adjusting the permanence of the synapses connected to cells in the predictive state, learning takes place. If connections are to active cells, permanence is increased, otherwise it is decreased.
4. The output of the region is a set of all active columns.

These routines have some dependencies and have been implemented as a core suite of matlab functional classes categorized as Spatial Pooler Functions and Temporal Pooler Functions.

The program used in developing the model was Java programming language because of its open-source support and object-oriented feature and Matlab was used for analytic simulation experiments because of its technicality, rapid software development prototype feature, functional programming feature and flexibility as an analytic and graphing tool.

Data used was obtained from UCI repository, the dataset consists of a thousand instances with twenty attributes for each instance. The CLA algorithm applied in developing the model makes use of unsupervised learning and hence does not require an output data set for training i.e. the system learns only on the input data. This has the advantage of reducing the overall system cost by half. The feature of sparse distribution exhibited by HTM models also reduces the data handling workload significantly.

In carrying out the experiment, the administrator launches the application and activates the system. The system is expected to be linked to an already existing credit card database. Once the system is activated, streaming credit card data comes into the system with which predictions are made based on the threshold set and permanencies used in the Cortical Learning Algorithm.

RESULTS

Table 1: Simulation tests results for the HTM/CLA system

Run Attributes		**HTM-CLA Prediction Accuracy**
Cell size (increasing columns)	**Number of Runs**	**% Accuracy**
5	1	55
10	1	65
15	1	60
20	1	70
21	1	70
22	1	60
23	1	60
24	1	70
25	1	70
50	1	75
70	1	80
71	1	75
72	1	85
73	1	80
74	1	80
75	1	80
76	1	80
77	1	80
78	1	80
79	1	85
80	1	85
81	1	70
82	1	80
83	1	80
84	1	75
85	1	75
86	1	75
87	1	80
88	1	80
89	1	75
90	1	75
500	1	88

Scientia Africana, Vol. 15 (No. 1), June 2016. Pp 1-10
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

1000	1	92
1500	1	93

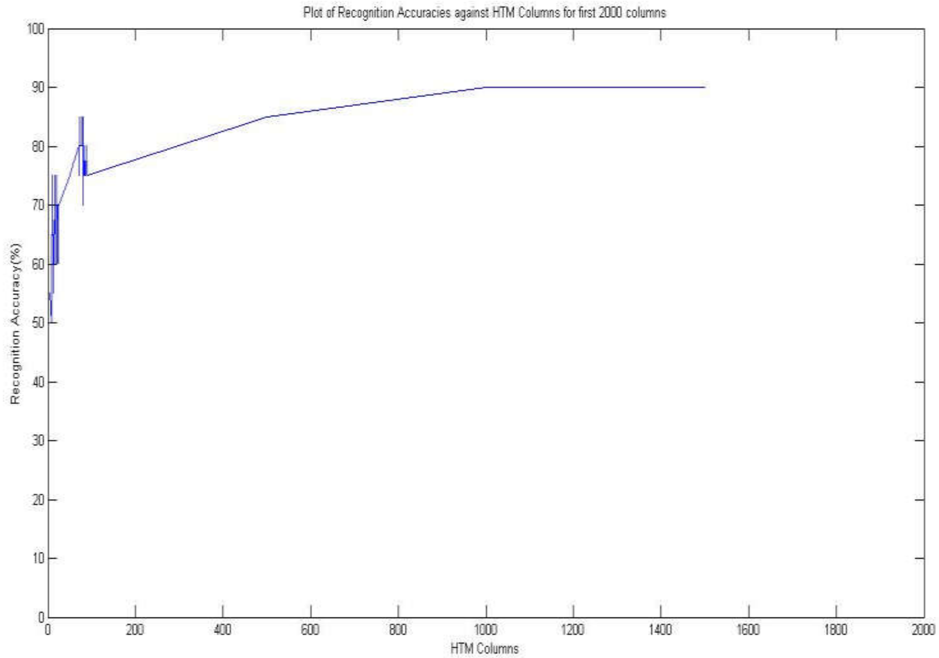

Figure 4: Graphical representation of test results for the HTM/CLA system

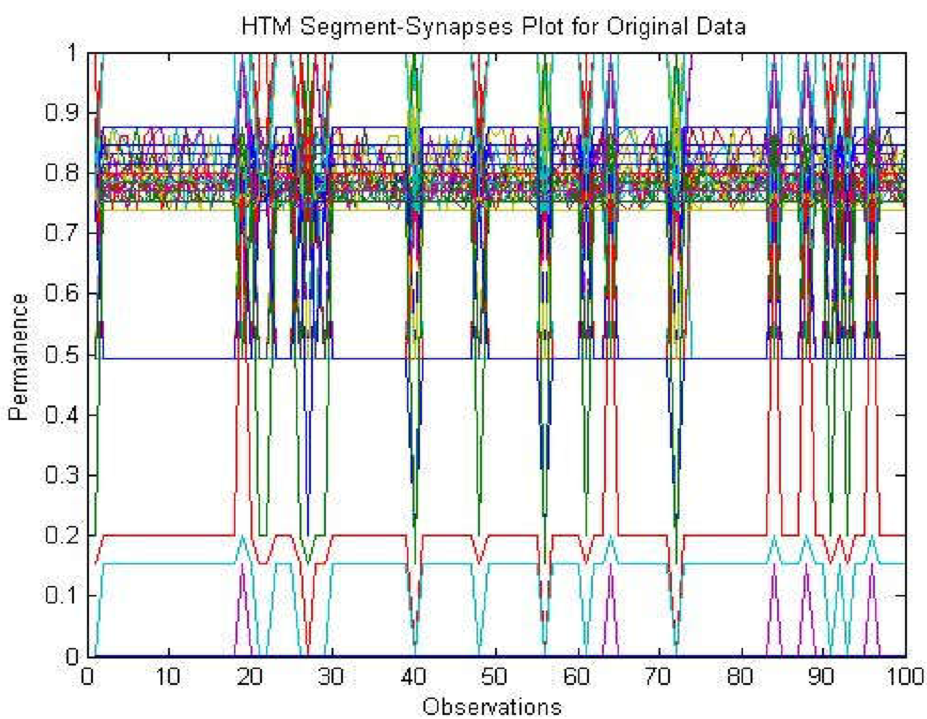

Figure 5: Visualization of HTM Segment Synapses of CCF data for Training

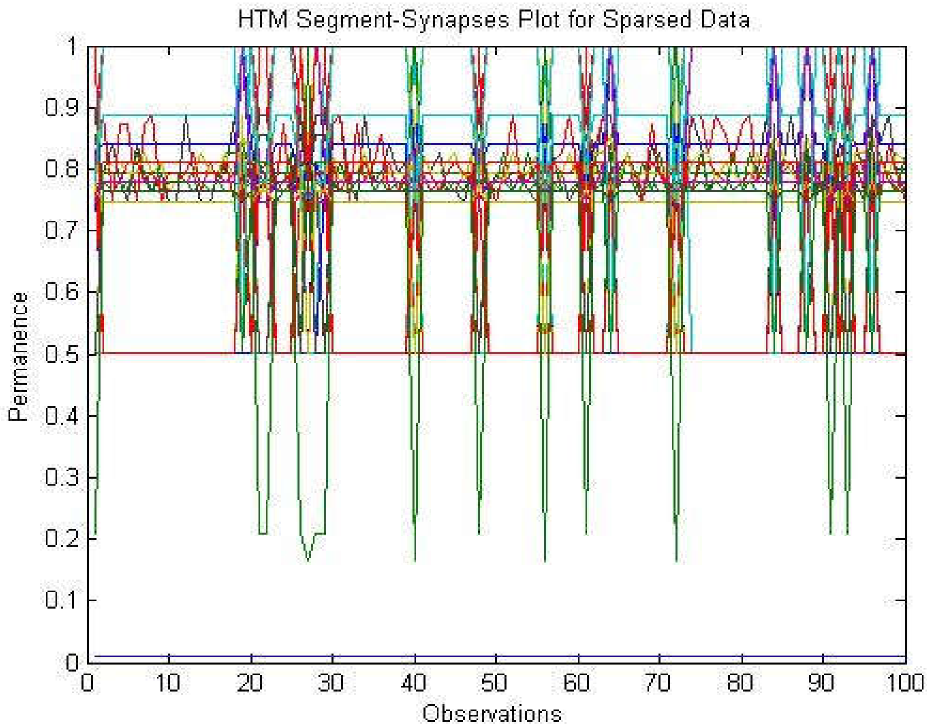

Figure 6: Visualization of Sparse HTM Segment Synapses of CCF data after Training

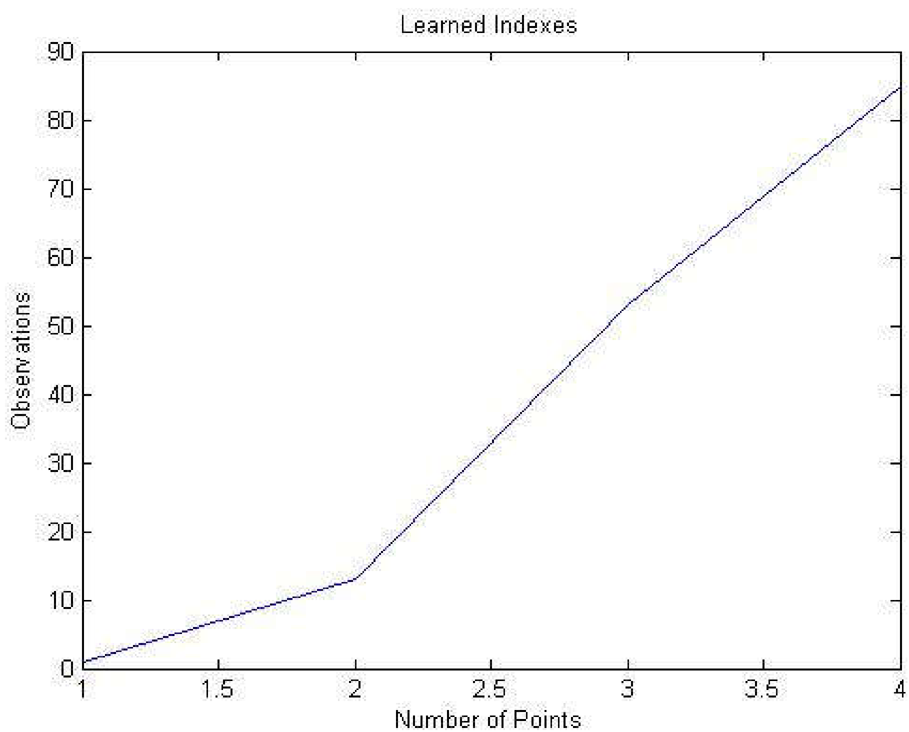

Figure 7: Learned indexes after Training

Scientia Africana, Vol. 15 (No. 1), June 2016. Pp 1-10
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

DISCUSSION

The simulations carried out on the HTM model revealed various behaviours of hierarchical temporal memory. These behaviours have been demonstrated on the following graphical visualization. The increase in the number of learning columns shows a significant increase in prediction accuracy as demonstrated in table 1 and figure 4. Figure 5 shows visualization of HTM segment synapses of CCF data for Training, the figure illustrates a high level of activity within permanence of 0.7 and 0.9 for the original data. This same learning activity is visualized in its sparse distributed representation and the illustration is given in figure 6, from figure 6, it is seen that training can also be recorded within same permanence range as in the case of the original data, the visualization of the training with sparse distributed representation also account for the handling of irregular and noisy data by HTM models, which is a typical characteristics of the credit card data. The HTM model learns more with increase in its observation during training just as in the human brain as shown in figure 7.

The paper has succeeded in demonstrating how the hierarchical temporal memory model learns like the neocortex of the human brain with regards to sparse distributed representation of data. This way of learning has not in any way reduced the prediction accuracy results generated; on the contrary the HTM model produced higher prediction accuracy of 92.3% as compared to the ANN model that recorded 89.6%. The behavioural analysis of the HTM model developed to predict credit card fraud in terms of its training, learning and prediction accuracy were reviewed. The hierarchical structure of the HTM model also allows for more intelligent prediction and anomaly detection. Using functional object-oriented approach, the ideas of neuroscience and biology, simple logical reasoning blocks and complex data structures can lead to better neural models of the brain.

REFERENCES

Bhatla, T., Vikram, P. and Amit, Dua. (2003). Understanding Credit Card Frauds. Tata Consultancy Services, 1-14.

Chappelier, J., Gori, M. and Grumbach, A. (2001). Time in Connectionist models. In Sequence Learning, Springer. 105-134.

Chaudhary, K., Yadav, J. and Mallick, B. (2012). A review of Fraud Detection Techniques: Credit Card. International Journal of Computer Applications. Vol. 45. 975-988.

Galetzka, M. (2014). Intelligent Prediction: An empirical study of the Cortical Learning Algorithm. Master Thesis, University of Applied Sciences Mannheim.

Ganong, F. (2003). Review of Medical Physiology. 21st Edition, McGraw-Hill. 196, 271-298.

George, L. (2008). Systemic Casualty in Complex Human Systems. In Proceedings of the 18th Conference on Behaviour Representation in Modelling and Simulation. 9-14.

Ghosh, S. and Reilly, D. (1994). Credit Card Fraud Detection with a Neural-Network. In Proceedings 27th Hawaii International Conference on System

Sciences: Information Systems: Decision Support and Knowledge-Based Systems, Vol. 3. 621-630.

Hawkins, J. and Blakeslee, S. (2004). On Intelligence. Henry Holt and Company.

Hopfield, J. (1988). Artificial Neural Networks. In Circuits and Devices Magazine, Institute of Electrical and Electronics Engineers. 3-10.

Numenta (2011). Hierarchical Temporal Memory. URL: http://numenta.org/cla-whitepaper.html

Patidar, R. and Sharma, L. (2011). Credit Card Fraud Detection Using Neural Network. International Journal of Soft Computing and Engineering (IJSCE), Vol. 1, 32-38.