World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:13, No:5, 2019

Use of Hierarchical Temporal Memory Algorithm in Heart Attack Detection

Tesnim Charrad, Kaouther Nouira, Ahmed Ferchichi

***Abstract*—In order to reduce the number of deaths due to heart problems, we propose the use of Hierarchical Temporal Memory Algorithm (HTM) which is a real time anomaly detection algorithm. HTM is a cortical learning algorithm based on neocortex used for anomaly detection. In other words, it is based on a conceptual theory of how the human brain can work. It is powerful in predicting unusual patterns, anomaly detection and classification. In this paper, HTM have been implemented and tested on ECG datasets in order to detect cardiac anomalies. Experiments showed good performance in terms of specificity, sensitivity and execution time.

***Keywords*—HTM, Real time anomaly detection, ECG, Cardiac Anomalies.

I. Introduction

ANOMALY detection is the identification of an unexpected behavior or patterns that are not conform to the normal or seems to be unusual. It is defined by Hawkins [2] as ”an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism”. Later, Edgeworth [7] defined the anomaly as ”discordant observations may be defined as those which present the appearance of differing in respect of their law of frequency from other observations with which they are combined”. We can define three types of anomalies [1]:

1) **Point Anomaly:** is one object different from all other observations
2) **Contextual Anomalies:** are anomalous objects in some specific context.
3) **Collective Anomaly:** are linked objects seemed to be regarded as anomalies compared to others.

Anomaly detection is a very important task applied to resolve problems especially in intrusion detection, fraud detection, financial domain and even in health monitoring systems.

In this purpose, many anomaly detection algorithms have been developed. In 2013, Chandore P. R. et al. [3] developed a combined cluster based approach and distance based approach to detect outliers in real time over wireless sensor networks. In 2015, Malhotra P. et al. [4] designed a Long Short Term Memory Networks for Anomaly detection in time series. LSTMAD uses stacked LSTM networks in order to detect anomalies in time series by training the model on normal data behavior and then detect deviations from the learned normal behavior.

In 2017, Jankov D. et al. [5] introduced an efficient high performance anomaly detection system for sensor data streams generated by manufacturing equipment. It is based essentially on Kmeans and Markov chain. In 2018, Loganathan G. et al. [6] developed a Sequence to Sequence Pattern Learning Algorithm for Real-Time Anomaly Detection in Network Traffic. It is a multi-attribute model for predicting a network packet sequence using a sequence to sequence encoder-decoder model as an application to intrusion detection problem.

In 2018, Chen X. et al. [22] created a hybrid supervised-unsupervised learning schemes. it is based on a clustering module (DCM) using DBSCAN and a deep neural network (DNN) based classifier and regressor.

However, human brain still very efficient in solving cognitive tasks compared to computers. So, the idea is to focus on the brain's characteristics, its operating mode and study the neocortex in order to develop new algorithms inspired by this. In 2011, Hawkin et al. are the first to develop this idea [18]. HTM networks are based on realistic neuron models with nonlinear active dendrites and thousands of synapses [14] to learn continuous online time-based sequences [19], [20]. HTM is very powerful in streaming data anomaly detection tasks [9], [10] as well as in sequence prediction tasks [21].

The primary goal in this paper is to give a presentation of HTM anomaly detection algorithm and apply it on cardiac datasets in order to detect heart attack.

This paper is organized as follows: Section II provides an overview of HTM anomaly detection algorithm and its principle. Section III deals with the performance and the results of HTM while applied in real-world cardiac datasets. Discussion is given in Section IV. Finally, we will conclude in Section V.

II. Hierarchical Temporal Memory

HTM is an Artificial Neural Network (ANN) derived from neuroscience that models spatial and temporal streaming data [11]. HTM learns from historical sequences, the association between these sequences, their context and the representations of patterns [12]. It is represented as a tree-shaped hierarchy of nodes [13]. Each node applies a common learning and memory function. HTM continuously learns and predicts what will happen and then makes inference. For each point x_t, HTM makes predictions for x_{t+1}. Besides, it compares the predicted value to the current value and deducts if this can be an anomaly or not. The algorithm adapts the inference model to track the change in continuous stream of incoming inputs.

First of all, HTM uses Sparse Distributed Representation (SDR) encoder in order to represent data stream. This representation has many benefits such as having strong resistance to noise, generalizability across data stream type

Tesnim Charrad is with the Higher Institute of Management, Tunisia (e-mail: charradtesnim@gmail.com).

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:13, No:5, 2019

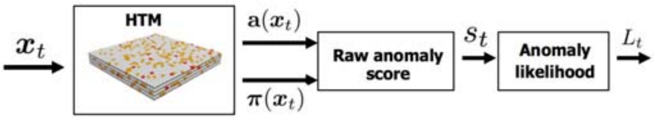


Fig. 1 HTM's functional steps [10]

and the attachment of semantic meaning to data points [14], [15]. Moreover, in contrary to dense representation, SDR few cells are active at the same time. This means that an SDR consists of thousands of bits where at any point in time a small percentage of the bits are 1s and the rest are 0s. A bit equal to 1 is an active bit. Each bit has a meaning. If two SDRs have active bits in the same locations, they share the semantic attributes represented by those bits. SDRs are binary and enable extremely fast computation.

Every neuron is only connected to a subset of other neurons. HTM activates this column if these bits occur together. In fact, SDRs offer an extremely high capacity, are robust to random deletions, can recognize patterns in the presence of noise and represent a dynamic set of patterns in a single fixed structure [16].

HTM networks constantly learn and model the spatio-temporal characteristics of their inputs. HTMs give good results for prediction, but HTM networks do not directly produce an anomaly score. In order to perform an anomaly detection, HTM uses two different internal representations: Raw anomaly score and anomaly likelihood (see Fig. 1).

The raw anomaly score (given by (1)) is calculated from two vectors $a(x_t)$ and $\pi(x_t)$. $a(x_t)$ represents SDR input vector at time t. $\pi(x_t)$ is the predicted vector of the next input (at time t+1). The raw anomaly score measures the difference between the actual input ($a(x_t)$) and its prediction ($\pi(x_t)$) [10].

$$s_t = 1 - \frac{\pi(x_{t-1}).a(x_t)}{|a(x_t)|} \tag{1}$$

The raw anomaly score is equal to 0 if the input is perfectly planned and1 if the input is completely unpredictable or is far away from the expectation.

Defining a threshold to limit the raw score leads to false positives. In this sense, HTM computes the likelihood score. HTM models the distribution of anomaly scores and uses this distribution to determine the probability that the current state is abnormal. The anomaly probability measures the anomaly as a function of the prediction history of the HTM model. Likelihood score measures how anomalous the current state is. HTM holds a window of the last raw anomaly score in order to calculate the likelihood score. HTM assumes that the raw anomaly score follows a rolling normal distribution and that it continuously updates the mean and the variance of the distribution [10]. The anomaly likelihood is the complement of the tail probability:

$$L_t = 1 - Q(\frac{\tilde{\mu_t} - \mu_t}{\sigma_t}) \tag{2}$$

with

$$\tilde{\mu_t} = \frac{\sum_{i=0}^{i=W'-1} s_{t-i}}{j} \tag{3}$$

TABLE I
COMPARISON OF LATENCY AND NAB SCORE FOR ALGORITHMS IMPLEMENTED AND TESTED ON NAB. LATENCY IS THE AVERAGE TIME OVER THREE RUNS ON A SINGLE DATA FILE. NAB SCORE REFLECTS THE STANDARD PROFILE SCORES [8]

Detector	Latency	NAB Score
HTM	11.3	70.1%
Relative Entropy	0.05	54.6%
Twitter ADVec	3.0	47.1%
Etsy Skyline	414.2	35.7%
Sliding Threshold	0.4	30.7%
Bayesian Changepoint	3.5	17.7%
EXPoSE	2.6	16.4%

W' is a window for a short term moving average, where $W' \ll W$. L_t is thresholded; HTM detects an anomaly if L_t is close to 1. $L_t \geq 1 - \varepsilon$.

This measures how much the model is able to make a prediction based on historical patterns. In order to evaluate HTM anomaly detection algorithm, we take two measures: The sensitivity and the specificity. The ratio Sensitivity represents the number of detected anomalies compared to the number of missed ones [17]. It can be defined as:

$$Sensitivity = \frac{AnomaliesDetected}{Anomalies} \tag{4}$$

However, the Specificity represents the number of significant detected anomalies compared to irrelevant ones [17]:

$$Specificity = \frac{TruePositiveAlarms}{TruePositiveAlarms + FalsePositiveAlarms} \tag{5}$$

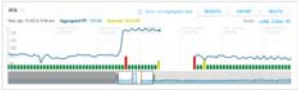
Indeed, a recent research [9] applying a set of real time anomaly detection algorithms (Twitter ADVec, Etsy Skyline, Random and HTM) shows that HTM achieves the best overall scores. Another research [10] shows that HTM is the leader in term of reducing false positives and true negatives. HTM anomaly detection algorithm is proved to be the best real time algorithm compared to the other real time anomaly detection algorithms [8] (results are shown in Table I).

The complexity of the HTM anomaly detection algorithm is equal to O(n) where n is the number of records. The execution time depends on the number of columns, cells per column and segments per cell for Temporal Memory, and potential pool size for Spatial Pooling. But regardless of the parameters, the time-per-record will converge to some constant value when the algorithm has hit the upper bounds.

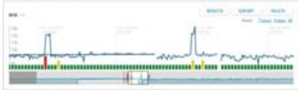
III. RESULTS

HTM have been implemented and tested on eleven datasets from the UCI [23] machine learning repository. Results are detailed in Table II where we calculated the performance representing the sensitivity and the specificity of the algorithm. Experiments show the high performance of the HTM anomaly detection algorithm. The sensitivity is between 50% and 100% (100% in most of cases). In the two last datasets, HTM only missed one anomaly. While, the specificity is between 99.28% and 100%. In fact, in most cases, the algorithm gives a high sensitivity and specificity. Accordingly, in most cases, HTM has successfully detected the spike in the heart rate measure

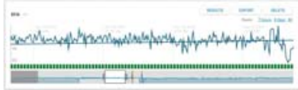
(a) Dataset 1



(b) Dataset 2



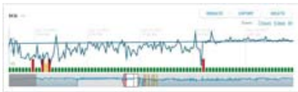
(c) Dataset 3



(d) Dataset 4

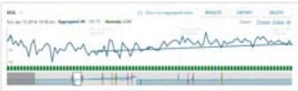
(e) Dataset 5

(f) Dataset 6



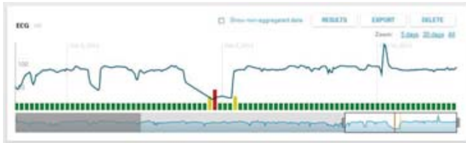
(g) Dataset 7

(h) Dataset 8



(i) Dataset 9

(j) Dataset 10



(k) Dataset 11

Fig. 2 Several datasets of 11 patients showing a variety of results since applying HTM anomaly detection algorithm. All of these datasets represents the maximum heart rate

TABLE II
THE PERFORMANCE OF HTM APPLIED INTO 11 UCI MACHINE LEARNING REPOSITORY DATASETS

Detector	Sensitivity	Specificity
Dataset 1	100%	99.95%
Dataset 2	100%	99.60%
Dataset 3	100%	99.80%
Dataset 4	100%	99.77%
Dataset 5	100%	99.85%
Dataset 6	100%	100%
Dataset 7	80%	100%
Dataset 8	100%	99.40%
Dataset 9	100%	99.80%
Dataset 10	80%	99.28%
Dataset 11	50%	99.87%

(see Fig. 2). Fig. 2 represents the plots of the HTM anomaly detection algorithm while applied to 11 datasets representing ECG signals.

IV. DISCUSSION

This study investigated the effectiveness of a cortical based anomaly detection algorithm "HTM". As hypothesized, HTM could provide promising solution for early heart attack prevention. It has a sensitivity that reach 100% in most cases and a specificity between 99.28% and 100%. This means that HTM has the ability to detect the maximum of anomalies while avoiding false alarms. Among the 11 experiments, HTM had two low sensitivity values (50% and 80%). In the first case, it detected one anomaly and missed one (50%=1/2) and in the second, it detected four anomalies and missed one (80%=4/5). Knowing that datasets' duration, in which HTM is applied, is 14 days.

V. CONCLUSION

Implementing cardiac anomaly detection algorithm able to prevent heart attack is very promising. In this paper, we have shown that HTM can succeed in this challenge. HTM paves the road towards a highly accelerated execution because of the use of Sparse Distributed Representation and its hierarchical model. In addition, HTM gives promising results in term of sensitivity and specificity which is a necessary and unavoidable criteria in healthcare systems.

REFERENCES

[1] Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly detection: A survey, ACM Computing Surveys (CSUR). 41(3), 1–72.

[2] Acuna, E. and Rodriguez, C. (2004). A meta analysis study of outlier detection methods in classification, Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez.

[3] Chandore, P. R. and Chatur, D. P. N. (2013). Hybrid approach for outlier detection over wireless sensor network real time data. International Journal Of Computer Science And Applications, 6(2).

[4] Malhotra, P., Vig, L., Shroff, G. and Agarwal, P. (2015, April). Long short term memory networks for anomaly detection in time series. In Proceedings (p. 89). Presses universitaires de Louvain.

[5] Jankov, D., Sikdar, S., Mukherjee, R., Teymourian, K. and Jermaine, C. (2017, June). Real-time High Performance Anomaly Detection over Data Streams: Grand Challenge. In Proceedings of the 11th ACM International Conference on Distributed and Event-based Systems (pp. 292-297). ACM.

[6] Loganathan, G., Samarabandu, J. and Wang, X. (2018, May). Sequence to Sequence Pattern Learning Algorithm for Real-Time Anomaly Detection in Network Traffic. In 2018 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1-4). IEEE.

[7] Edgeworth, F. Y. (1887). On discordant observations, Philosophical Magazine. To Appear in ACM Computing Surveys, 09 2009. 23(5), 364–375.

[8] Ahmad, S., Lavin, A., Purdy, S. and Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. Neurocomputing.

[9] Lavin, A. and Ahmad, S. (2015). Evaluating Real-Time Anomaly Detection Algorithms–The Numenta Anomaly Benchmark. In Machine Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on (pp. 38-44). IEEE.

[10] Ahmad, S. and Purdy, S. (2016). Real-Time Anomaly Detection for Streaming Analytics. arXiv preprint arXiv:1607.02480.

[11] Kerner, M. and Tammeme K. (2017). Hierarchical temporal memory implementation on FPGA using LFSR based spatial pooler address space generator. In Design and Diagnostics of Electronic Circuits and Systems (DDECS), 2017 IEEE 20th International Symposium on pp. 92–95. IEEE.

[12] M. Putic,A. J. Varshneya and M. R. Stan (2017). Hierarchical Temporal Memory on the Automata Processor. IEEE Micro, 37(1), 52-59.

[13] Fan, D., Sharad, M., Sengupta, A., and Roy, K. (2016). Hierarchical temporal memory based on spin-neurons and resistive memory for energy-efficient brain-inspired computing. IEEE transactions on neural networks and learning systems, 27(9), 1907-1919.

[14] Hawkins, J., and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. Frontiers in neural circuits, 10.

[15] Ahmad, S., and Hawkins, J. (2016). How do neurons operate on sparse distributed representations? A mathematical theory of sparsity, neurons and active dendrites. arXiv preprint arXiv:1601.00720.

[16] Wu, J., Zeng, W., Chen, Z., and Tang, X. F. (2016, December). Hierarchical Temporal Memory Method for Time-Series-Based Anomaly Detection. In Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on (pp. 1167-1172). IEEE.
[17] Nouira, K. and Trabelsi, A. (2012). Intelligent monitoring system for intensive care units. Journal of medical systems, 36(4), 2309-2318.
[18] Hawkins, J., Ahmad, S. and Dubinsky, D. (2011). Cortical learning algorithm and hierarchical temporal memory, Numenta Whitepaper, 1-68.
[19] Antic, S. D., Zhou, W. L., Moore, A. R., Short, S. M. and Ikonomu, K. D. (2010). The decade of the dendritic NMDA spike, Journal of neuroscience research, 88(14), 2991-3001.
[20] Major, G., Larkum, M. E. and Schiller, J. (2013). Active properties of neocortical pyramidal neuron dendrites. Annual review of neuroscience, 36, 1-24.
[21] Cui, Y., Ahmad, S. and Hawkins, J. (2016). Continuous online sequence learning with an unsupervised neural network model. Neural computation, 28(11), 2474-2504.
[22] Chen, X., Li, B., Shamsabardeh, M., Proietti, R., Zhu, Z. and Yoo, S. J. B. (2018, March). On real-time and self-taught anomaly detection in optical networks using hybrid unsupervised/supervised learning. In European Conf. Optical Communications.
[23] A. Frank and A. Asuncion (2013). UCI machine learning repositor, http://archive.ics.uci.edu/ml.