ISEG
School of Economics and Management
TECHNICAL UNIVERSITY OF LISBON

DYNAMIC APERIODIC NEURAL NETWORK FOR TIME SERIES PREDICTION

Chiu-Che Tseng
Department of Computer Science and Information Systems
Texas A&M University-Commerce

Abstract

There are many things that humans find easy to do that computers are currently unable to do. Tasks such as visual pattern manipulating objects by touch, and navigating in a complex world are easy for humans. Yet, despite decades of research, we have no viable algorithms for performing these and other cognitive functions on a computer. In this study, we used a bio-inspired neural network called a KA-set neural network to perform a time series predictive task. The results from our experiments showed that the predictive accuracy with this method was better in most markets than results obtained using a random walk method.

Keywords: Kset neural network, Time series, Prediction

INTRODUCTION

Neural Network technology was developed in an attempt to artificially reproduce the acquisition of knowledge and the organization skills of the human brain. It offers significant support in terms of organizing, classifying, and summarizing data, but with few assumptions and a high degree of predictive accuracy. Neural networks are less sensitive to error term assumptions; they can tolerate noise, chaotic components, and heavy tails better than most other methods.

In this paper we apply a novel type of neural network, called KA sets. We selected the KA set neural network from traditional neural networks because results

Correspondence Address: Chiu-Che Tseng, Department of Computer Science and Information Systems – Texas A&M University-Commerce – Commerce, Texas 75429 USA. E-mail: chiu-che_tseng@tamu-commerce.edu

from experiments conducted by other researchers have shown that the KAIII model is a promising computational methodology for financial time series prediction [3], and it compares very well with alternative prediction methods.

A KA set is an abstract version of an already existing K set. Named after Aharon Katchalsky, an early pioneer of neurodynamics who tragically died in the early 70s, KA sets are strongly biologically motivated. KA models represent a family of models of increasing complexity that describe various aspects of. vertebrate brain functions. Similar to other neural network models, KA models consist of neurons and connections, thus their structure and behavior are more biologically plausible. The KA model family includes KA0, KAI, KAII, KAIII and KAIV.

Time series prediction takes an existing series of data. X(0),.......X(t-1),X(t), and forecasts the future values X(t+1),X(t+2)....etc. The goal is to model the history data in order to forecast future unknown data values accurately. Due to high noise level and the non-stationary nature of the data, financial forecasting is a challenging application in the domain of time series prediction. In this work we used the KAIII model to predict the one step direction of daily currency exchange rates.

RELATED WORKS

There are several studies using KA-set neural networks. Beliaev and Kozma [1] stored binary data in a KIII model and a Hopfield model. They then tried to retrieve the stored data by giving noisy input. The results of their study suggest that the K-model had greater potential for memory capacity than the Hopfield model.

Kozma and Beliaev [2] developed a methodology to use KIII for multi-step time series prediction and applied the method to the IJCNN CATS benchmark data.

Li and Kozma [3] applied a KIII dynamic neural network to the prediction of complex temporal sequences. In their paper, their KIII model gives a step-by-step prediction of the change in direction of a currency exchange rate.

Our approach differs from that of Li and Kozma in the following ways: we have not fixed the lateral nodes to 40; we use a myopic strategy to try all possible results, varying lateral nodes from 2 to 60. The best value for the lateral nodes is therefore determined by the results. We also increase the training set by using the sliding window input approach. Our results from eight markets are comparable to theirs.

FOUNDATIONS OF THE KAIII NEURAL NETWORK

The architecture of artificial neural networks is inspired by the biological nervous system. It captures information from data by learning, and stores the information among its weights. Compared to other symbolic computational models, this computation model is especially good with regard to generalization and error tolerance. In a departure from traditional neural networks, the structure and behavior of the KA models are closer to the nervous system. Each KA set models some biological part of the nervous system, such as sensory pathways, cortical areas etc. Another prominent difference from other generally known Neural Networks is that the activity of the K model is oscillatory, which is typically found within the chaotic regime.

The basic KA-unit, called KAO set, models a neuron population of about 10^4 neurons. It is described by a second order ordinary differential equation:

$$(a*b)\frac{d^2P(t)}{d^2t}+(a+b)\frac{dP(t)}{dt}+P(t)=F(t). \quad (1)$$

Here a and b are biologically determined time constants; a = 0.22, b = 0.72. P(t) denotes the activation of the nodes as function of time; F(t) is the summed activation from neighbor nodes.

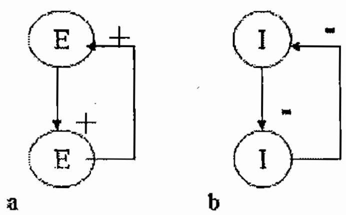
The KAO set has a weighted input and an asymptotic sigmoid function for the output. The sigmoid function Q(x) is given by the equation:

$$Q(x) = Q_m *\{1 - e^{[-1/Qm * (ex-1)]}\} \qquad (2)$$

where Q_m = 5, is the parameter specifying the slope and maximal asymptote of the curve. This sigmoid function is modeled from experiments on biological neural activation.

FIGURE 1

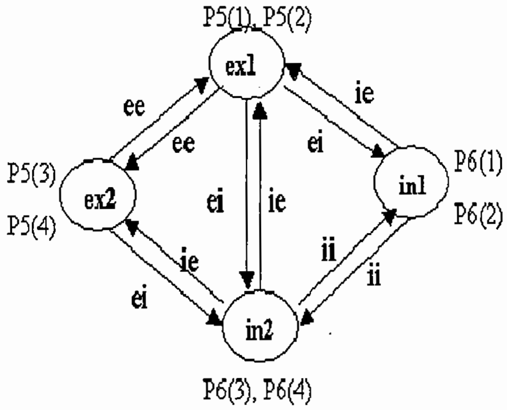
KI sets (a) excitatory KI, (b) inhibitory KI



Coupling two or more KO sets with excitatory connections, we get a KAI set.

FIGURE 2

A KAII network



The next step in the hierarchy is the KAII model. KAII is a double layer of excitatory or inhibitory units. In the simplest architecture there are 4 nodes: two excitatory and two inhibitory nodes;

Figure 2 shows the schema of the KAII set with two excitatory nodes (ex1, ex2), and two inhibitory nodes (in1, in2). P5(i) and P6(i) are the activation levels and the derivatives of the activations of the excitatory and inhibitory nodes. They are called P5 and P6 for historical reasons.

In order to achieve a certain stability level on KAII, we conducted experiments to search for weight parameters, namely, ee, ii, ei, and ie, so that KII can sustain oscillatory activation in impulse-response tests. Biological experiments have shown that when there is a single pulse of small random input perturbation, there are three types of KII based on the activation trajectory of the excitatory node P5(1): positive attractor KAII, zero attractor KAII, and negative attractor KAII. The selection of weight parameters is based on satisfying these criteria. Three sets of KAII weight parameters are selected as in the following table:

TABLE I

KAII weights parameters

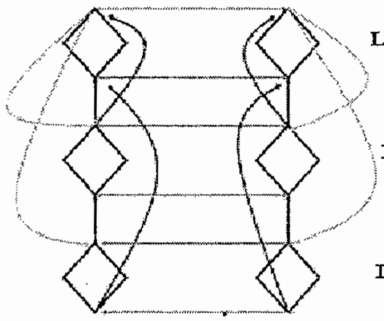
	ei	*ie*	*ee*	*ii*
Zero attractor KAII	1.0	2.0	1.8	0.8
Positive attractor KAII	1.6	1.5	1.6	2.0
Negative attractor KAII	1.9	0.2	1.6	1.0

ei, ie, ee, ii are the connections between excitatory node and inhibitory node

The KAIII model is designed to be a dynamic computational model that simulates the sensory cortex. It can perform pattern recognition and classification. Based on the structure of the cortex, KAIII consists of three layers connected by feed forward/feedback connections. Each layer has multiple KAII sets, connected by lateral weights between corresponding P5(1) and P6(3) nodes. Based on the structure and dynamics of the sensory cortex, KAII sets in each layer should be zero attractor KAII, positive attractor KAII, and negative attractor KAII, respectively.

FIGURE 3

A sample KAIII network



The schema of a KAIII set with three layers of KAII units with feed forward, feedback, and lateral connections is shown in Figure 3. It only shows two KAII sets in each layer. In order to get a homeostatic balanced state in the KAIII level, we list the activation equations for all 12 nodes in a single column in KAIII. The equation of P5(1) node in the first layer is:

$$P_{15(1)} = ee1 * f(P_{15(3)}) - ie1 * f(P_{16(1)}) - ie1*f(P_{16(3)}) + W_{ee} * f(P_{15(1)}) + W_{P25(1)P15(1)} * f(P_{25(1)}) \quad (3)$$

In the above equation $P_{15(1)}, P_{15(3)}, P_{16(1)}, P_{16(3)}$ are the activations of the node P5(1) , P5(3),P6(1) and P6(3) in layer 1; $P_{25(1)}$ is the activation of P5(1) node in layer 2; ee1 and ie1 are the connections between KAII sets in layer 1. Wee is the lateral connection between KAII sets in layer1; $W_{P25(1)P15(1)}$ is the feedback connection from $P_{25(1)}$ node in layer 2 to $P_{15(1)}$ node in layer 1; f(x) is the asymmetric sigmoid function.

After plugging in the parameters in table I, we balance 12 nonlinear equations for each node. As a result, we get the parameters for all the lateral connections between layers .

TABLE II

Lateral connections of three layers in KIII

	W_{ee}	W_{ii}
Layer 1	0.15	0.1
Layer 2	0.2	0.2
Layer 3	0.15	0.1

Wee is the lateral connections between excitatory nodes; Wii is the connections between inhibitory nodes

TABLE III

Lateral connections of three layers in KIII

$W_{P25(1)P15(1)}$	0.05	Feedback Connection
$W_{P25(1)P16(1)}$	0.25	Feedback Connection
$W_{P36(1)P16(1)}$	0.05	Feedback Connection
$W_{P15(1)P25(1)}$	0.15	Feedforward Connection
$W_{P35(1)P26(1)}$	0.2	Feedback Connection
$W_{P15(1)P35(1)}$	0.6	Feedforward Connection

NEURAL NETWORK STRUCTURE AND LEARNING RULE

After the set up of KAIII to a homeostatic balanced state, its periodic dynamic can be sustained if the inputs are within a certain range. We constructed KAIII with different lateral nodes, varying from 2 to 60. This parameter decides the dimension of the input time series. For example, in a KAIII with 40 lateral nodes, the inputs for each excitatory node are X(t-40), X(t-39).....X(t-1). If the previous 40-day currency exchange rates are entered into the system, our model will predict the change in direction of the next day's rate.

The learning rule in KAIII is associative Hebbian learning. Since the system is always in a dynamic state, there is no single converged value. The activation standard deviation (σi) of each node in a certain duration T is used.

$$\Delta Wij = L^*(\sigma i - \sigma)^*(\sigma j - \sigma) \quad (4)$$

where L is the learning rate; σi and σj are the activation standard deviation of two nodes i, j; Δ Wij is the weight change between these two nodes.

$$\sigma_i = \frac{1}{T}\sqrt{\int_0^T\left(P35_i(t) - \frac{1}{T}\sum_{t=1}^{T} P35_i(t)\right)^2} \qquad (5)$$

In the discrete version we use

$$\sigma_i = [\, 1 / T \, \Sigma_{i=1}{}^{T} \, (P35_i(t) - 1/T \, \Sigma_{t=1}{}^{T} P35_i(t))^2]^{1/2}$$

where T = 200 ms, $P35_i(t)$ is the activation of excitatory node at instant t in layer 3; σ is defined as the mean of the σi.

$$\sigma = 1 / L \, \Sigma_{i=1}{}^{L} \, \sigma i \qquad (6)$$

where L is the number of lateral nodes.

MYOPIC STRATEGY

In the domain of time series prediction, in order to predict the next output data, we need to determine what is the right input size to use. For example, should we use the prior 30 days' index data as input and predict the data on the 31st day or should we use a different input size? In this study, we applied a myopic strategy to find the right input size to use for predicting the following day's output. The strategy starts with 2 inputs and constructs the KAII network accordingly. It will then increase the input size and so on. The strategy will stop when the accuracy rate produced by the network has reached a preset threshold. The result network will be selected for the particular data set.

This approach saves a considerable amount of computational time over the trial and error approach for selecting the right input size to use for the data set.

EXPERIMENTAL SETUP

The input consisted of eight different markets' stock index data from the 01/02/1986 to 12/30/2004. We used the sliding window input approach in the learning phase. In sliding window input, if the number of nodes is 40, we first input the values from 1 to 40. In the next step we input the values from 2 to 41 and so on. This increases the training set, i.e. the number of values to train the network. Two thirds of data are used in the learning phase and the remaining third is used in both the validation and testing phases.

We did not fix the L value, i.e. the learning rate, at 40 nodes. Instead, we used a myopic strategy to try all possible input sizes, varying L value from 2 to 60 nodes. The best L value was then determined on the basis of the results.

DATA NORMALIZATION

The data used were the stock index data of eight different countries from the period of 1986 to 2004. We normalized the data so that they were in the range of [-1,1]. This was done by taking the difference between two corresponding data values and dividing them with the maximum of all difference values present. We had more than four thousand data points of the daily closing bids of each country's stock index data. Due to the nature of the KAIII, the same input needs to be repeated for a certain duration, then followed by a period of relaxation (zero input). This is a biologically inspired process to mimic the activity of a sniff and other sensory activities. We chose both the learning and relaxation durations to be 25minutes in this work.

LEARNING PHASE

Prior to putting learning data into KAIII, each input was labeled as UP/SAME/DOWN according to the next data. The Hebbian learning process happens in the lateral connections among the excitatory nodes in the third layer. As explained in the previous section, the learning phase lasted 25minutes. In our experiment, the first 5 minutes' response activations were skipped because initial activations are considered unstable when the signal is first perceived by the system. Thus, only the following 20minutes' activations of excitatory nodes P35(1) in layer 3 were collected for learning purposes. We plugged these data into the Hebbian learning equation. The lateral connections were adjusted accordingly. We called this process a cycle. The learning cycle was repeated for two thirds of the data. Observing the weight change indicates that some weights increased, while some decreased during learning.

VALIDATION PHASE

The validation phase was basically the same as the training phase, except no learning was applied. In each cycle, the filtered standard deviations of all the excitatory nodes in the third layer were stored separately, according to the input data label. This enabled us to have reference for each category. We used these references to classify patterns in the testing phase. From the point of view of aperiodic dynamics, the values of the spatially distributed oscillation intensities encode the input data. When a given input is applied, the high dimensional dynamics collapse to a lower dimensional subspace. In other words, these oscillatory patterns generate the clusters for different categories.

TESTING PHASE

During the testing phase, the filtered standard deviations were recorded for each group of testing data. These were defined as the activation points for the testing data. Then, the distances from all references to the activation point were calculated. The classification algorithm selected the N references with the smallest distances. Our classification algorithm required that the difference between the numbers of references belonging to each pattern be greater than others in order to claim a "winner". The "winner" pattern decides which pattern the testing data belongs to. For example, for each case there are 9 nearest neighbors. In Case 1, 6 references belong to pattern UP and 3 references belong to pattern DOWN. The system would classify the UP category as the "winner" for this testing point.

In Case 2, if 5 references are pattern UP, and 4 references are pattern DOWN, the system classifies the testing data "winner" as the UP category. The size of the neighborhood N which decides the "winner" is acquired from the experiment in the testing phase. We used the first one third of testing data to try some N values. We checked the classification performance for every N configuration. The N configurations with the best performance were selected for the following two-thirds of the testing data. Once N was decided, they would remain constant for all of the following testing data. Therefore, only two-thirds of the testing data were counted as performance evaluation in each test run.

EXPERIMENTAL RESULTS

We ran through all eight markets' stock index data using out myopic KAII neural network and we got different results from different markets. Most of the predictive accuracy rates were more than 50%, only Japan's market did not attain more than 50%.

Let us take the USA data as an example. For our given set of data, the best results occurred when the number of nodes was 4. We got 52% correct predictions and 47% incorrect predictions. That means we were using the previous four days of data to predict the fifth day's trend.

TABLE IV

USA index predictive accuracy

No_of_nodes	*Correct%*	*Incorrect%*
2	51.3233%	48.6767%
3	47.5%	52.5%
4	52.2096%	47.7904%
5	50.4372%	49.5628%
6	49.6164%	50.3836%
7	48.503%	51.497%
8	49.3789%	50.6211%
9	46.4995%	53.5005%
10	48.5577%	51.4423%
11	46.2398%	53.7602%
12	46.2556%	53.7444%
13	46.4359%	53.5641%
14	48.0227%	51.9773%
15	30.023%	69.977%
16	39.9957%	60.0043%

TABLE V

UK index predictive accuracy

No_of_Nodes	*Correct %*	*Incorrect %*
2	50%	50%
3	50.3788%	49.6212%
4	49.6843%	50.3157%
5	50.2732%	49.7268%
6	50.6262%	49.3738%
7	49.3584%	50.6416%
8	47.4767%	52.5233%
9	46.0467%	53.9533%
10	40.2244%	59.7756%
20	23.609%	76.391%
30	15.5642%	84.4358%
40	11.2004%	88.7996%

TABLE VI

Canada index predictive accuracy

No_of_Nodes	*Correct %*	*Incorrect %*
2	48.6616%	51.3384%
3	50.7645%	49.2355%
4	51.4066%	48.5934%
5	48.956%	51.044%
6	49.7101%	50.2899%
7	49.9785%	50.0215%
8	48.9875%	51.0125%
9	47.5986%	52.4014%
10	42.4834%	57.5166%
20	22.9135%	77.0865%
30	16.3424%	83.6576%
40	11.7659%	88.2341%

TABLE VII

Germany index predictive accuracy

No_of_Nodes	*Correct %*	*Incorrect %*
2	47.7055%	52.2945%
3	47.9358%	52.0642%
4	46.0358%	53.9642%
5	49.2857%	50.7143%
6	47.6329%	52.3671%
7	50.6221%	49.3779%
8	51.7523%	48.2477%
9	49.7133%	50.2867%
10	40.7616%	59.2384%
20	21.9737%	78.0263%
30	15.0454%	84.9546%
40	11.2004%	88.7996%

TABLE VIII

Japan index predictive accuracy

No_Of_Nodes	*Correct %*	*Incorrect%*
2	42.8112%	57.1888%
3	47.0284%	52.9716%
4	45.4741%	54.5259%
5	46.8111%	53.1889%
6	45.207%	54.793%
7	46.8975%	53.1025%
8	46.4096%	53.5904%
9	47.6534%	52.3466%
10	44.1667%	55.8333%
20	20.3292%	79.6708%
30	13.9111%	86.0889%
40	9.91706%	90.08294%

TABLE IX

Spain index predictive accuracy

No_Of_Nodes	*Correct%*	*Incorrect%*
2	46.3585%	53.6415%
3	49.2117%	50.7883%
4	47.1698%	53.8302%
5	52.8455%	47.1545%
6	50.5007%	49.4993%
7	50.0952%	49.9048%
8	47.7841%	52.2159%
9	48.5714%	51.4286%
10	47.9024%	52.0976%
20	18.1921%	81.8079%
30	14.1408%	85.8592%
40	9.50893%	90.49107%

TABLE X

Taiwan index predictive accuracy

No_of_Nodes	*Correct %*	*Incorrect %*
2	50.2342%	49.7658%
3	48.7828%	51.2172%
4	48.0408%	51.9592%
5	48.6486%	51.3514%
6	49.4069%	50.5931%
7	47.3379%	52.6621%
8	47.2222%	52.7778%
9	47.8128%	52.1872%
10	37.8175%	62.1825%
20	19.819%	80.181%
30	13.6426%	86.3574%
40	10.6429%	89.3571%

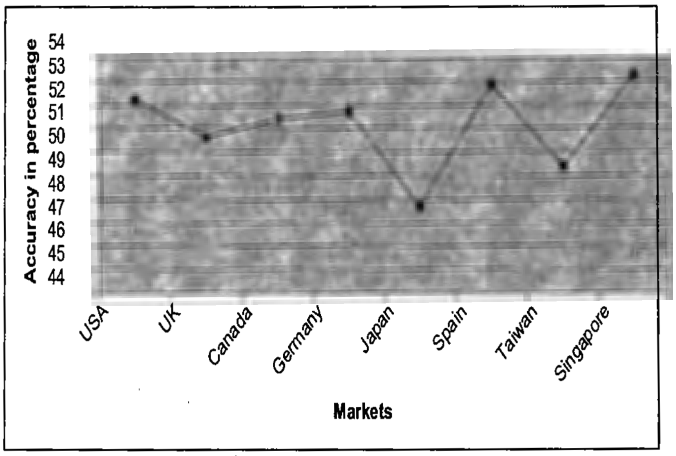
TABLE XI

Singapore index predictive accuracy

No_of_Nodes	*Correct %*	*Incorrect %*
2	49.4715%	50.5285%
3	48.1293%	51.8707%
4	51.8362%	48.1638%
5	48.693%	51.307%
6	50.0534%	49.9466%
7	53.3078%	46.6922%
8	52.921%	47.079%
9	48.6014%	51.3986%
10	41.2635%	58.7365%
20	20.8025%	79.1975%
30	14.7407%	85.2593%
40	11.7654%	88.2346%

FIGURE 4

The best predictive accuracy from KAIII neural network in different markets



CONCLUSION

Our initial research on using a neurodynamic KIII network on different stock markets' index data prediction produced good results. We plan to further our research by applying it to other domains, including health and human activity modeling.

References

Atiya, A., and Parlos, A, (1995), "Identification of nonlinear dynamics using a general spatio-temporal network, ", Math. Compute Modeling J. Vol. 21, no. 1, pp. 55-71, Jan.

Bassi, D., (1995) *Stock price predictions by current by recurrent multiplayer neural network architectures in* the Capital markets Conf., A. Reference Ed. London, U.K., October 1995, London Business School, pp. 331 – 340.

Beliaev, Igor, Kozma, Robert (2007) "Time Series Prediction Using chaotic Neural Networks: Case Study of Ijcnn Cats Benchmark Test", Department of Mathematical Sciences, University of Memphis, Memphis, TN.

Beliaev, Igor, Kozma, Robert (2006), "Studies on the Memory Capacity and Robustness of Chaotic Dynamic Neural Networks", International Joint Conference on Neural Networks, Vancouver, BC, Canada, July 16 – 21.

Carney, J.G., Cunningham (1996), "Neural networks and currency exchange rate prediction", Foresight Business Journal, http://www.maths.tcd.ie/pub/fbj/index.html.

Chang, H.Z., Freeman, W.J. and Burke, B.C. (1998), "Optimization of Olfactory Model in Software to Give 1/f Spectra Reveals Numerical Instabilities in Solutions Governed by Aperiodic (chaotic) Attractors", Neural networks, 11, 449 -466.

Freeman, W.J. (1975) *Mass Action in the Nervous System,* Academic Press, N.Y.

Gardner, E., Derrida B., (1988), "Optimal storage properties of neural network models", Journal of Physics A: Mathematical and general, Volume 21, Issue 1, pp. 271 – 284.

Giles, C.L., Lawrence, S. and Tsoi A.C. (2001), *Noisy time series prediction using a recurrent neural network and grammatical inference,* Machine earning, 44 (1-2), pp. 161-183.

Hopfield, J.J.,(1982), " Neural Networks and Physical Systems with Emergent Collective Computational Abilities", Proc. Natl. Acad. Sci., USA, Volume 79, pp. 2554-2558.

Kozma, R., Freeman W.J., and Erdi, P., (2003), *The Kiv Model – Nonlinear spatio – temporal dynamics of the primordial vertebrate forebrain;* Neurocomputing, 2003 in press.

Li Haizhon, Kozma, Robert (1982), "A dynamic Neural Method for time Series Prediction Using the KIII Model" Division of Computer Science, University of Memphis, Memphis, TN.

McEliece, R.J., Posner, E.C. Posner, Rodemich, E.R. and Venkatesh, S.S., (1987) "The capacity of the Hopfield associative memory", IEEE Trans. Inform. Theory, Vol. 33, pp. 461, July 1987.

Moody, J. and Wu L. (1996), *Optimization of trading systems and portfolios, in* Proc. Neural Networks Capital Markets Conf. Pasadena, CA.

Mueller, B., Reinhardt (1990), "Neural Networks: An Introduction", Springer.

Muthu, S., Kozma, R., Freeman W.J. (2004), "Applying KIV dynamic neural network model for real time navigation by mobile robot Aibo", IEE/INNS 2004 Int. Joint Conference on Neural networks IJCNN'04, July 25-29, 2004 Budapest, Hungary.

Toomarian, N. and Barthen, J. (1991), "Adjoint -functions and temporal learning algorithm in neural networks", Advances in neural information Processing Systems, 3, pp. 113 – 120.

Resumo

Há muitas coisas que o Homem tem facilidade em fazer e que os computadores actuais não têm ainda essa capacidade. Por exemplo, manipular objectos através do tacto e navegação num mundo complexo são tarefas fáceis para os seres humanos. Apesar da investigação realizada ao longo de várias décadas, não há algoritmos suficientemente válidos para realizar estas e outras funções cognitivas no computador.

Neste estudo utiliza-se um sistema "bio – inspirado" numa rede neuronal designada por "KA-set neural network" para efectuar previsões de séries temporais. Os resultados do estudo demonstram que a precisão de previsão deste método foi melhor na maioria dos mercados do que os resultados obtidos por um método aleatório.

Palavras Chave: Ket rede neural; séries temporais; previsão.