Identification of the Head-and-Shoulders Technical Analysis Pattern with Neural Networks

Achilleas Zapranis and Prodromos Tsinaslanidis

Department of Accounting and Finance, University of Macedonia of Economic and Social Sciences, P.O. Box 1591, 54006 Thessaloniki, Greece
{Zapranis,tsinpro}@uom.gr

Abstract. In this paper we present a novel approach for identifying the head-and-shoulders technical analysis pattern based on neural networks. For training the network we use actual patterns that were identified in stochastically simulated price series by means of a rule-based algorithm. Then the patterns are being converted to binary images, in a manner similar to the one used in hand-written character and digit recognition. Our approach is tested on new simulated price series using a rolling window of variable size. The results are very promising with an overall correct classification rate of 97.1%.

Keywords: Pattern Recognition, Technical Analysis, Head-and-Shoulders, Neural Networks, Identification.

1 Introduction

In this paper we look at the pattern recognition problem from a financial perspective. To be more precise a multiple layer feed-forward neural network is incorporated for identifying within simulated price series the Head-and-Shoulders (HS) pattern. HS is one of the most known price pattern of technical analysis (TA). The preparation of the training dataset is being done with methods similar to those used in hand-written character and digit recognition. Patterns of TA include a high level of subjectivity. This makes in practice the pattern “unobservable”. Usually technical patterns are identified by rule-based mechanisms [1]. One way to train the network would be with cases where HS pattern was identified by known investment advisors, or other financial intermediaries. Such a procedure would be time-consuming and impractical. In this paper, the Geometric Brownian Motion (GBM) is used to simulate an adequate number of simulated stochastic price series. After that, actual patterns identified in stochastically simulated price series by means of a rule-based algorithm are used to train the network. Before training the network these patterns are converted into binary images, then into sub-total matrixes, and finally into column vectors. The network is tested on new simulated price series using a rolling window of variable size (‘multi-sized rolling window’). The network seems to perform near perfect with an overall classification rate of 97.1%.

The remainder of the paper is organized as follows. Section 2 describes the Technical analysis and the HS pattern. Section 3 describes how neural networks deal with

K. Diamantaras, W. Duch, L.S. Iliadis (Eds.): ICANN 2010, Part III, LNCS 6354, pp. 130–136, 2010.

pattern recognition challenges. Section 4 describes the input data preparation. Section 5 describes the creation of the multi-layer feed-forward network, and section 6 concludes.

2 Technical Analysis and the Head-and-Shoulders Pattern

Some see TA as a powerful tool in order to predict future prices and returns of stocks and other investment products. According to [2] "*technical analysis is the science of recording, usually in graphic form, the actual history of trading (price changes, volume of transactions, etc.) in a certain stock or in "the Averages" and then deducing from that pictured history the probable future trend"*. The main tools of TA are indicators and patterns. The later are distinctive formations created by the movements of security's prices on a chart. Two well known technical patterns are the HS and the Cup-with-Handle. More information on TA indicators and chart patterns can be found in [2, 3].

The HS pattern's name comes from its resemblance of the upper part of human body. It consists of three peaks, with two intervening troughs between the head and each shoulder. The neckline is drawn through the lowest points of these intervening troughs and may slope upward or downward. When the price breaks downwards the neckline, a fall of the price equal to the distance between the head and the neckline (head's height) is expected. There are four categories of the HS pattern. The distinction is based on the trend (bull or bear trend) before the pattern identified, and on the pattern's form (normal or inverse). When the HS pattern has its normal form we expect a down-trend when the "neckline" is crossed. Reversely, when it has the inverse form we expect an up-trend. When the normal form occurs and an up-trend preexists the pattern is characterized as a trend-reversal pattern. Down-trend preexistence characterizes the pattern as trend-continuation. Similarly, when the inverse form takes place the pattern is characterized as a trend-reversal or as a trend-continuation when a down-trend or an up-trend preexists respectively.

3 Neural Networks and Pattern Recognition

The concept of neural networks started as an effort to describe the human brain behavior. An artificial neural network is a system emulating a biological neural system. Artificial neural networks emerged after the introduction of simplified neurons [4]. Today's architectures are able to make systems that perform similar with the human brain. Pattern recognition is one of the main actions performed by the human brain. Personal identification from handwriting, voice, fingerprints, facial images etc. can be achieved by identifying patterns hidden in biological data. Hand-written character recognition [5-9] and digit recognition [10] are well-known pattern recognition challenges.

The concept of *classification* involves the learning of likenesses and differences of objects in a population of no identical objects [11]. Pattern *recognition* is the process defining that an object from a population P belongs to a known subpopulation S. The recognition of an object as a unique singleton class is called *identification*. On the other hand, classification is the process of grouping objects into classes according to their similarities. Recognition and classification are both included in pattern recognition. Additionally, [11] presents a taxonomy of pattern recognition methods. The three main classes of pattern recognition methods are:

1) *Decision-theoretic methods* which consist of statistical, graph-theoretic, and rule-based methods; [5], [1].

2) *Structural/Syntactic methods* which consist of automata, Hopfield recurrent neural networks and bidirectional associative maps.

3) *Associative mappings (neural and fuzzy mappings*) which consist of feed-forward neural networks, self-organizing networks and hybrid networks;[6, 7, 10, 12-14].

In our network, the training patterns were preprocessed with methods similar to those used in handwritten character and digit recognition. In [10], a combination of Self Organizing Maps (SOMs) with fuzzy rules was applied for handwritten numerical recognition. The training patterns were rescaled and centered onto a 64×64 binary grid. This method is also adopted in this paper at the data preprocessing stage where the HS patterns identified in stochastic price series were converted into 64×64 binary matrixes before being used as inputs for the network. Similar methods were used in many studies in the field of hand-written character recognition; [5] and [12].

4 Data Preparation

The HS is one of the most known patterns in TA. However, it is considered to occur very seldom. As a result very few verified time series exist and a satisfactory large training dataset cannot be created in order to use it as an input to the feed-forward neural network. Simulated data is used to rectify this. More precisely, the GBM[1] is used to create a large number of simulated price paths. The GBM is given by:

$$dS = S\mu dt + S\sigma\varepsilon\sqrt{dt} \tag{1}$$

where μ is the expected return, σ is the volatility of the price series, dt is a short interval of time and ε follows a standard normal distribution. GBM is considered as an accurate representation of the underlying stock price (S) generation mechanism.

First, we construct 250,000 price paths of 500 observations each with different combinations of drift rate and volatility rate. On each price series a rule-based mechanism was applied to identify the HS pattern by applying the criteria presented in [1]. The rule-based algorithm identifies the HS pattern in a price path and returns the characteristic points (local peaks and troughs) of the pattern. Cases where the pattern is identified or not, are categorized into two groups in which the target pattern is assigned to value 1 or 0 respectively.

However, the rule-based mechanism based on the criteria presented in [1] is limited and can identify only the pattern that has the normal form and uptrend preexistence. On the other hand, in our proposed methodology the network is trained to identify all four cases mentioned in section 2. To do so, after the rule based script returns the characteristic points (locals) of the pattern (Fig. 1), the points of the trend

[1] For further details about the Geometric Brownian Motion see [15].

preexistence (P0 and T0) are excluded and the intersection between the neckline and the uptrend that preexists (P**) is added. In other words, from the initial formation only the part above the neckline remains. For the identification of the HS pattern's inverted form, the network can reexamine the same price series after multiplying the series with minus one.

In order for all the inputs of the network to have the same formation, the method mentioned in [10] is applied. Each part of the price series is converted into a binary 64×64 matrix. This can be achieved by scaling the prices and the days of the price series between 1 and 64 and rounding them toward the closest integer. This results in the coordinates which prices are represented by ones on the 64×64 matrix. The value of zero is assigned to the remaining prices of the matrix. For the improvement of the network the initial binary matrix is bordered into 4×4 sub-blocks resulting in a 16×16 matrix where each point expresses the sum of each sub-block. So each point of the latest matrix has values between 0 and 16. Finally, the matrix is converted into a vector ready to be used as an input to the network.

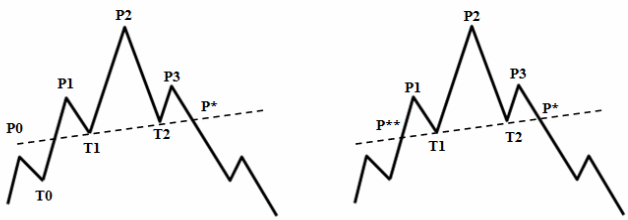

Fig. 1. The locals of the trend reversal HS pattern with the normal form. On the left pattern are illustrated the characteristic points of the pattern given by the rule based mechanism. On the right pattern only the locals that are used as inputs of the network, are presented.

The network, after the training stage will try to identify the pattern in new price series at different time intervals. This will have an impact in the 64×64 binary matrix. For larger period of times after the scaling of the price series there will be more points in the matrix with the value of one. In other words, this will affect the thickness of the price series when viewed in the binary matrix. To overcome this problem the network is trained with inputs of different thicknesses.

Before the scaling, a linear connection between the characteristic points of the pattern is applied. By this, the noise between those locals is extracted and the main structure of the pattern is isolated without taking into account the price's fluctuation between the locals. The thickness parameter mentioned before helps to deal with the noise realized at the price series when the network is used for identification.

Another key point to the preparation of the data is the slope of the neckline that can vary. We exclude 30% of the cases with the highest and another 30% of the cases with the lowest necklines' slopes, in order to include more symmetrical HS patterns in our training dataset. This is an addition to the criterion presented in [1].

For the network's training, two thousands different input vectors are used. Half targeted with ones, which represent HS pattern cases, and the other half targeted

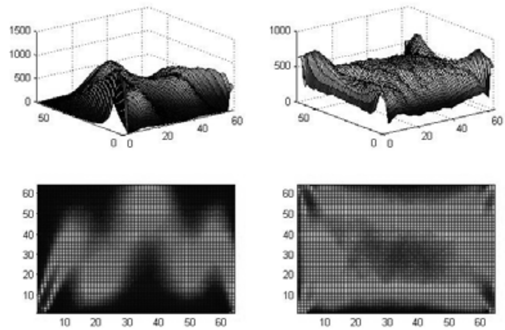

Fig. 2. Top sub-figures; cumulative, three dimensional distributions of the input binary matrixes. Down sub-figures; ground plans of the above 3D sub-figures. Left sub-figures; Inputs targeted with ones (HS patterns). Right sub-figures; Inputs targeted with zeros (no HS patterns).

with zeros that represent cases where the pattern was not identified. Fig. 2 illustrates cumulative, three dimensional (3D) distributions of the input binary matrixes before they are converted into column vectors. Left sub-figures present the inputs targeted with one where the pattern is clear, and right sub-figures the inputs with target zero where a very noisy picture can be seen. The top sub-figures show a three dimensional illustration while the bottom ones show a ground plan of the above 3D figures. These figures help illustrate the inputs used to feed the network.

5 The Neural Network

A two-layer feed-forward network, with tan-sigmoid transfer functions in both the hidden layer and the output layer can classify vectors sufficiently, given enough neurons in its hidden layer. The network's aim is to decide whether a specific price series is a HS pattern or not. For that reason, the network has two output neurons because there are two categories associated with each input vector. With the Neural Network Pattern Recognition Tool (nprtool) it is possible for the user to create, train a network, and evaluate its performance using mean square error and confusion matrices[2].

The Scaled Conjugate Gradient algorithm is used for training. 60% of the input patterns is used for training, 20% is used to validate that the network is generalizing and to stop training before overfitting, and 20% is used as an independent test of network generalization.

In tables 1-4 the confusion matrices for training, testing, validation and an overall view are illustrated. Each table consist of three columns (a, b and c) and three rows (1, 2 and 3). The network's outputs are near perfect, as shown by the high numbers of correct responses in the bordered cells (cells a1 and b2) and the low numbers of incorrect responses in the not bordered cells (cells a2 and b1). The lower right bold percentages (cell c3) illustrate the overall accuracies. Results of the training, validation, and test state indicate that the network is able to classify the input data with great

[2] For further details about nprtool see the product help of *Matlab*. 2009b. The Mathworks, I.

accuracy. This is confirmed by the Receiver Operating Characteristic (ROC) curve which indicates that the network performs near perfectly.

The next step is to apply a rolling window of variable length on the price series and then convert the parts of the series into forms similar to the ones mentioned in the training stage. Variable length is used to capture all possible sizes of the pattern. First, the rolling window has the same size as the price series. Then the window is halved, and rolls with a step of 5 days[3]. By incorporating this "divide and roll" procedure the network seeks in the price series the HS pattern with different views. It can be argued that a technical analyst does the same, by focusing on the series at different view levels. Overall, our network identifies most patterns found by the rule-based mechanism.

Table 1. Training Confusion Matrix

Output Class	1	597	9	98.5%
		49.8%	0.8%	1.5%
	2	6	588	99.0%
		0.5%	49.0%	1.0%
		99.0%	98.5%	**98.8%**
		1.0%	1.5%	**1.2%**
		1	2	
		Target Class		

Table 2. Validation Confusion Matrix

Output Class	1	192	18	91.4%
		48.0%	4.5%	8.6%
	2	6	184	96.8%
		1.5%	46.0%	3.2%
		97.0%	91.1%	**94.0%**
		3.0%	8.9%	**6.0%**
		1	2	
		Target Class		

Table 3. Test Confusion Matrix

Output Class	1	194	14	93.3%
		48.5%	3.5%	6.7%
	2	5	187	97.4%
		1.3%	46.8%	2.6%
		97.5%	93.0%	**95.3%**
		2.5%	7.0%	**4.7%**
		1	2	
		Target Class		

Table 4. All Confusion Matrix

Output Class	1	983	41	96.0%
		49.1%	2.1%	4.0%
	2	17	959	98.3%
		0.9%	47.9%	1.7%
		98.3%	95.9%	**97.1%**
		1.7%	4.1%	**2.9%**
		1	2	
		Target Class		

6 Conclusion

The aim of this study is to create a neural network that identifies the HS technical price pattern in a price series of a stock. To prepare our training data we use a method similar to the ones presented in other hand-written character or digit recognition studies. We overcome the small observation frequency of the pattern with a rule-based pattern identification mechanism and a large number of simulated stochastic price series that represent the price generation mechanism of real stock prices. The network classified the input data near perfectly (97.1%) while at the same time outperformed the rule-based mechanism. More precisely, the neural network in contrast to the rule-based algorithm identifies all four different forms of the HS pattern. Our future

[3] We prefer a 5-days step than a daily step for accelerating the network's identification process.

studies will focus on the measurement of the network's identification power and determine statistically if it outperforms the rule-based mechanism. The aforementioned system can be used in the future as a tool for weak-form market efficiency[4] test in financial markets.

References

1. Bernd, L.: Are Technical Trading Rules Profitable? Evidence for head-and-shoulder rules. Applied Economics 35, 33–40 (2003)
2. Edwards, R.D., Magee, J.: Technical Analysis of Stock Trends. 7th edn. (1997)
3. Achelis, S.B.: Technical Analysis From A to Z (1995)
4. McCulloch, W.S., Pitts, W.H.: A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133 (1943)
5. Li, T.F., Yu, S.S.: Handprinted Chinese character recognition using the probability distribution feature. Int. J. Pattern Recogn. Artif. Intell. 8(5), 1241–1258 (1994)
6. Tsay, M.-K., Shyo, K.-H., Chang, P.-C.: Feature Transformation with Generalized Learning Vector Quantization for Hand-Written chinese Character Recognition. IEICE Trans. Inf. & Syst., E82-D(3) (1999)
7. Camastra, F., Vinciarelli, A.: Cursive character recognition by learning vector quantization. Pattern Recognition Letters 22, 625–629 (2001)
8. Liu, C.-L., Nakagawa, M.: Evaluation of prototype learning algorithms for nearest-neighbor classifier in application to handwritten character recognition. Pattern Recognition 34, 601–615 (2001)
9. Liu, C.-L., Sako, H., Fujisawa, H.: Performance evaluation of pattern classifiers for handwritten character recognition. International Journal on Document Analysis and Recognition 4, 191–204 (2002)
10. Chi, Z., Wu, J., Yan, H.: Handwritten Numeral Recognition Using Self-Organizing Maps and Fuzzy Rules. Pattern Recognition 28(1), 59–66 (1995)
11. Looney, C.G.: Pattern Recognition Using Neural Networks (1997)
12. Tseng, D.C., Chiu, H.P., Cheng, J.H.: Invariant handwritten Chinese character recognition using fuzzy ring data. Image and Vision Computing 14, 647–657 (1996)
13. Kohonen, T.: Self-Organizing Maps. In: Huang, T.S., Kohonen, T., Schroeder, M.R. (eds.) 2nd edn. Springer series in information sciences, vol. 30 (1997)
14. Cho, S.-B.: Ensemble of structure-adaptive self-organizing maps for high performance classification. Information Sciences 123, 103–114 (1999)
15. Hull, J.C.: Options, Futures, and Other Derivatives. 6th edn. (2006)
16. Fama, E.: 'Efficient Capital Markets: A Review of Theory and Empirical Work'. Journal of Finance 25(2), 383–417 (1970)
17. Fama, E.: 'Efficient Capital Markets II'. Journal of Finance 46(5), 1575–1617 (1991)

[4] For market efficiency hypothesis see [16], [17].