IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sep 2018 E-ISSN: 2347-2693

Classifying Sequences of Market Profile using Deep Learning

Pranshu Rupesh Dave[1] and Priti Srinivas Sajja[2]

[1]Birla Vishvakarma Mahavidyalaya Engineering College, India
[2]Sardar Patel University, India

**Corresponding Author: pranshudave@gmail.com, Tel.: +91 8238330100*

Available online at: www.ijcseonline.org

Accepted: 16/Sept/2018, Published: 30/Sept/2018

Abstract— Since its inception, market profile has been used by traders as a way to assess the market value of a stock. By reading market profile charts, it is possible for traders to assess who is driving the market (buyers or sellers) and make trades accordingly. The spatiotemporal feature of market profile can be used to train a deep learning model for classifying sequences of market profile. This is a novel idea and one that needs to be examined and experimented upon. LSTM networks are structures capable of remembering long term dependencies in time series data. Convolutional Neural Networks, on the other hand help in figuring out patterns in multidimensional data. A python library is built to generate market profile from time series data. Leveraging the power of LSTMs and CNNs, two models are proposed for the classification: FC-LSTM and ConvLSTM. The results show that the proposed models are able to catch patterns amongst profiles and FC-LSTM performs better than ConvLSTM on this task.

Keywords—Market Profile, Machine Learning, ConvLSTM

NOMENCLATURE

LSTM: Long Short Term Memory Networks
FC-LSTM: Fully Connected LSTM
RNN: Recurrent Neural Networks
ConvLSTM: Convolutional Long Short Term Memory Networks
BTC: Bitcoin
USD: US Dollars
ROC: Receiver Operating Characteristic
OHLCV: Open High Low Close Volume
AUC: Area Under Curve
SGD: Sigmoid Gradient Descent

I. INTRODUCTION

With the recent advent of deep learning techniques, neural networks have been used to reach groundbreaking results in image recognition, self-driving cars, game playing, speech recognition, intrusion detection, natural language processing and financial trading [1, 2]. Recurrent Neural Networks and Long Short Term Memory Networks are advanced deep learning architectures for sequence learning tasks [3].

One task that has always piqued researchers and one that is notoriously difficult is that of predicting the movement of price in a financial time series. Lately, it has been established that various artificial intelligence techniques that employ machine learning help in getting better results at this task. Machine learning has been applied in various ways to analyze these data [4, 5, 6].

Market profile was first envisioned by J. Peter Steidlmayer, who was a trader at the Chicago Board of Trade (CBOT) in 1985. It came into being as a way to evaluate the market value of stocks at a particular time. Trading using market profile involves mastering the ability to read the profile to understand where the value lies and making trades accordingly [7].

Fig 1 shows an example of a market profile chart. A typical structure of a market profile has prices on the y-axis and time frames on the x-axis. Each time frame is represented by an alphabet. For example, the first time frame will be represented by A, the second by B, and so on. The Point of Control (POC) is the price which occurs in the highest number of time frames [8].

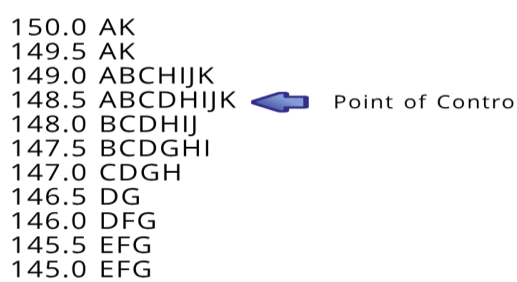

Fig 1. A typical market profile chart

This paper presents ideas on how the correlation between the structure of market profiles and price movement can be leveraged to train different machine learning models that predict price movements before they occur. Deep learning models based on LSTM networks that leverage their memory power are proposed. Furthermore, the applications of these models in predicting very high volatility time series data such as that of cryptocurrencies is assessed. The results of the models are displayed in the form of their confusion matrices and ROC curves.

The rest of the paper is organized as follows, Section I contains the introduction of market profile and use of machine learning in financial time series, Section II contains the related work of a number of researches that have been carried out in this area, Section III contains methodology, Section IV contains results and discussion, Section V contains conclusion and future scope of the research work.

II. RELATED WORK

Although the application of machine learning on a series of market profile has been unexplored, it is important to study other ways in which time series analysis in financial markets is carried out. Krauss used an ensemble a deep neural network, a gradient boosted tree and a random forest to achieve this task [9]. Also, Nikola proposed a number of machine learning models to evaluate if a company's value will go up by 10% in a year with 76.5% accuracy [5]. Kalyani proposed a model based on SVM to predict stock price movement by performing sentiment analysis of news [6]. Fischer leverages the memory power of LSTM networks to predict out-of sample directional movements of the stocks of S&P 500 [4]. Amudha proposes a framework that backtracks and modifies its previous predictions if earlier forecasts were incorrect [10]. Another time series data that are highly volatile is that of cryptocurrency time series. Recently, the price of Bitcoin has seen some of the biggest jumps both in upward as well as downward directions. Some of the techniques that have been used for predicting price movements in cryptocurrencies are: Velankar uses various features related to the Bitcoin price [11]. Kondor uses the bitcoin ledger network data to make predictions [12].

III. METHODOLOGY

III.I Data and Preprocessing

The dataset for Bitcoin was obtained from the Coinbase cryptocurrency exchange. The data were in the form: Open, High, Low, Close, and Volume (OHLCV) with a time interval of 1 minute. The prices were in USD and volume in BTC. Duration of the time series data is from 1/12/2014 to 27/3/2018.

For preprocessing of the data, a python library was developed (called market-profile). This library is made available on the PyPi package index. The library takes data in the OHLC format and produces market profile of different types namely, regular, compacted. For the purpose of this research, a regular market profile with prices as the first row and time periods as the columns was used.

Fig 2 shows an example of a sample generated profile. The 1's indicate the prices that occurred in a particular time frame. The prices are as shown in the first row. Since financial time series data is very volatile, there is no uniformity in the actual range of data for any particular time frame. This would result in each time frame having a different number of rows since the number of distinct prices would be different each time. Furthermore, the step size between two prices would be uniform for the same reason. One way to tackle this problem is by defining that all market profiles that the model is trained on being the same size (rows, columns). Also, the prices should be averaged out to get a uniform step size.

The model was trained on market profile with different number of rows and different length of time frames. Since the Bitcoin time series is highly volatile and there was a higher precision data available (1-minute interval), the market profiles were generated for 6 hours, 3 hours and 1-hour time slots. Furthermore, variations in the number of rows were tried as well, namely, 16 columns and 32 columns. To understand this more clearly, consider a 3-hour market profile. This 3-hour profile would be divided into 15 different time slots for a 16 column profile (1st column for the price).

	6	7	8	9	10	11	12	13	14	15
8	0	0	1	1	0	0	0	0	0	1
9	0	0	1	1	0	1	1	0	0	1
10	0	0	1	1	0	1	1	0	0	1
11	0	0	1	1	0	1	1	0	0	1
12	0	0	1	1	0	1	1	0	0	1
13	0	0	1	1	0	1	1	0	0	1
14	0	0	1	1	1	1	1	0	0	1
15	0	0	1	1	1	1	1	0	0	1
16	0	0	1	1	1	0	1	1	0	1
17	1	1	1	1	1	0	1	1	1	1
18	1	1	1	1	1	0	1	1	1	1
19	1	1	1	1	1	0	1	1	1	1
20	1	1	1	1	1	0	0	1	1	1
21	1	0	0	0	0	0	0	0	1	0
22	1	0	0	0	0	0	0	0	1	0
23	1	0	0	0	0	0	0	0	1	0
24	1	0	0	0	0	0	0	0	1	0

Fig 2. Sample generated Profile

Furthermore, to model the data as a supervised learning problem, the labels for sequences of the market profile need to be defined. These labels were defined in the following manner: If the highest price in the next profile exceeds the current highest price by an amount d, the label will be 1. In all other scenarios, the label will be 0. Thus, supervised learning problem is a binary classification problem.

III.II Models

The paper focuses mainly on two machine learning models. The first model is a Long Short Term Memory

Network. LSTM networks have found extensive usage for analyzing financial time series data. They serve as a memory cell and are able to maintain their state over time. Also they do not suffer from various difficulties in training that RNNs suffer from (vanishing gradients) [13]. This makes them ideal to be used as a model for this particular problem.

The problem with feeding a market profile to a LSTM is the fact that LSTM require a sequence of numbers, i.e. they are 1D structures but the profile is a 2D structure. One workaround is to unroll the entire profile into a 1-dimensional array and feed it to the LSTM. To account for multiple previous time frame, these layers of 1D arrays are stacked together.

The second model uses a variant of LSTM, known as a Convolutional LSTM. ConvLSTM differs from normal fully connected LSTM (FC-LSTM) in the sense that it has convolutional structures in both the input-to-state and state-to-state transitions. ConvLSTM helps capture the patterns in the shapes of input [14]. Since the shape of market profile is an important factor in making trading decisions, this should be an ideal structure for this classification.

Two main models were tested for the classification problem. The first model trains a stack of LSTM layers. Various time splits were tried out including daily, 6 hourly, 3 hourly and hourly. The model has two layers to LSTM stacked together and a fully connected layer to make a prediction. Adding more layers led to over fitting. The training test split was set to allocate 25% of the data as test data and the rest as training data. Adam optimizer was used for the training and the loss function used was binary cross entropy. The time steps used for feeding data into the model were 1, 2, 5, 10, and 20. The model was trained for 10 epochs.

The second model uses Convolutional LSTM with three ConvLSTM layers. Each layer had 20 filters with a kernel size of (4, 4). Experimental results show that smaller kernel sizes failed to catch the patterns in the profile. Batch normalization was used over each ConvLSTM layer. The final layer was a fully connected layer to which the flattened outputs of previous layers were given. The logarithm of the hyperbolic cosine was used as the loss function and sigmoid gradient descent as the optimizer. The test set size was set to 20% of the dataset. The summary of the two models is shown in the table 1.

Table 1. Model Characteristics

Model	Time Split (hours)	Time steps	Number of layers	Test Size	Optimizer
LSTM	3	5	2	25%	Adam
ConvLSTM	1	2	3	20%	SGD

IV. RESULTS AND DISCUSSION

Evaluation of any machine learning model is vital as it helps find inaccuracies and biases, if any, and assesses its performance in real world application. It is essential that the model does not under fit the data making it unable to catch patterns in the profile. Also, it is imperative that the model doesn't over fit training data as this might lead to poor performance in the application.

Confusion matrices and ROC charts find ubiquitous use in evaluating machine learning models. Precision and recall are metrics that aid in evaluation. Precision is the proportion of positive cases that were correctly identified and recall is the proportion of actual positive cases that were correctly identified. Area under ROC curve can be used as a single evaluation metric over accuracy as stated by Bradley [15].

IV.I FC-LSTM:

For the model based on FC-LSTM networks, the best results were achieved on the 3-hour time split. This might be because volatility drastically increases on lowering time split not allowing the model to grasp patterns in the profile. Also, increasing the time split has a detrimental effect on accuracy as it becomes difficult to capture the entire patterns over such a long period in just a few rows of the profile. Furthermore, lower time steps seemed to have greater accuracy associated with them. This was surprising as LSTM networks are used to capture patterns over a longer period time, but the best results were obtained over 5 time steps.

Fig 3. displays a model loss versus epoch chart for 12 epochs. Results of training the model show that the test set error begins to plateau around 9 epochs, but training set error continues to decrease. Training over more epochs results in over fitting. Thus, training is stopped at 12 epochs to prevent overfitting the data.

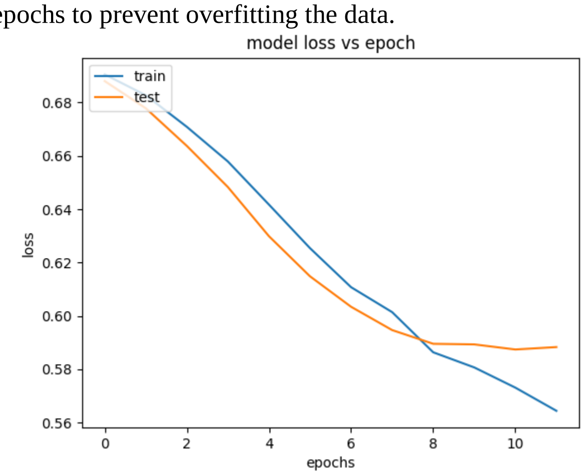

Fig 3. Model loss vs. epoch for FC-LSTM

Fig 4. on the other hand displays the Receiving Operating Characteristic of the LSTM based model. The area under the curve is 0.70. Finally, Fig 3. displays the confusion matrix of the model. Using the values from the matrix, precision and recall are evaluated. Precision is found to be 0.731 and recall 0.7054.

Table 2. Confusion matrix for FC-LSTM

	Predicted: 1	Predicted: 0	Total
Actual: 1	886	326	1212
Actual: 0	370	738	1108
Total	1256	1064	

$$Precision = proportion\ of\ positive\ cases\ that\ were\ correctly\ identified = \frac{886}{1212} = 0.731$$

$$Recall = proportion\ of\ actual\ positive\ cases\ correctly\ identified = \frac{886}{1256} = 0.7054$$

$$Area\ under\ Curve\ (AUC) = 0.70$$

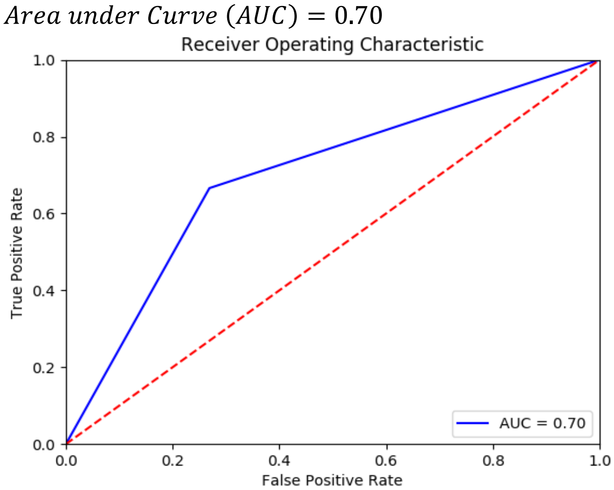

Fig 4. ROC curve for FC-LSTM

IV.II Convolutional LSTM:

An even shorter time split is considered for Convolutional LSTM (1 hour). Longer time splits lead to worse results in this model. Also, a larger filter size is seen to better capture patterns in the profile leading to a higher accuracy. Fig 5 displays model loss versus epoch for this model. Training is

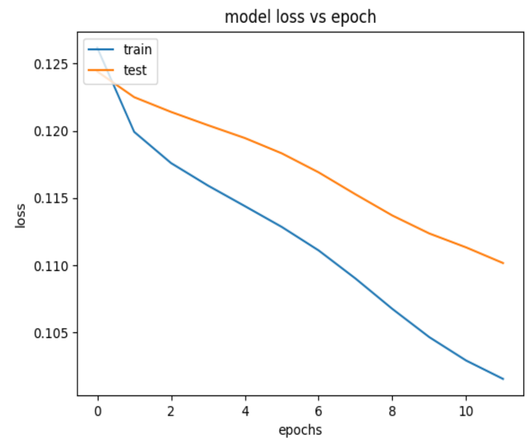

Fig 5. Model loss vs. epoch for ConvLSTM

stopped at 12 epochs once again to prevent over fitting. Fig 6 displays the ROC curve for this model. The area under the curve is 0.61 which is much worse than that for FC-LSTM. The confusion matrix is as shown in Table 3. The model displays a higher precision, but a lower recall than FC-LSTM.

$$Precision = proportion\ of\ positive\ cases\ correctly\ identified = \frac{2829}{3675} = 0.76$$

Table 3. Confusion matrix for ConvLSTM

	Predicted: 1	Predicted: 0	Total
Actual: 1	2829	846	3675
Actual: 0	1764	1438	3184
Total	4575	2284	

$$Recall = proportion\ of\ actual\ positive\ cases\ correctly\ identified = \frac{2829}{4575} = 0.6183$$

$$Area\ under\ Curve\ (AUC) = 0.61$$

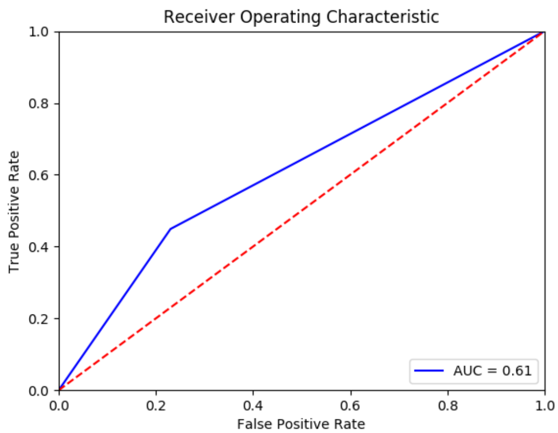

Fig 6. ROC curve for ConvLSTM

Since market profile is a spatiotemporal structure, it is only logical to assume that ConvLSTM should perform better. But the results show that FC-LSTM performs much better on this task. This might be due to the fact that the Bitcoin time series is highly volatile and behaves differently when compared to a normal stock market time series that would agree with the laws of finance.

CONCLUSION AND FUTURE SCOPE

This paper proposes a novel approach to make time series predictions of cryptocurrency data leveraging market profile. It successfully models the sequence classification of market profile as a deep learning problem. It proposes two models for this purpose: FC-LSTM and Convolutional LSTM. The FC- LSTM model is shown to give better results over Convolutional LSTM model, contrary to the initial reasoning that ConvLSTM should perform better as market profile is a type of spatiotemporal data. The fact that market profile is relevant for predicting time series price movement (for cryptocurrency data) even after 3 decades of its inception opens up immense research opportunities. Future work can involve building deep learning models using market profile for the prediction of stock market data and also building an entire trading algorithm using such models. Also, effect of incorporation of other trading parameters like price and volume in such models would need to be evaluated.

REFERENCES

[1] Ratnesh Kumar Shukla, Ajay Agarwal, Anil Kumar Malviya, "*An Introduction of Face Recognition and Face Detection for Blurred and Noisy Images*", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.3, pp.39-43, 2018.

[2] P. Rutravigneshwaran, "*A Study of Intrusion Detection System using Efficient Data Mining Techniques*", International Journal of Scientific Research in Network Security and Communication, Vol.5, Issue.6, pp.5-8, 2017

[3] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, No. 10, pp. 2222-2232, Oct. 2017.

[4] T. Fischer and C. Krauss, ‘‘Deep learning with long short-term memory networks for financial market predictions’’, FAU Discussion Papers in Economics, 2017.

[5] Milosevic, Nikola. “Equity forecast: Predicting long term stock price movement using machine learning.” Journal of Economics Library, Vol. 3, No. 2, pp. 288-294, Jun. 2016.

[6] K. Joshi, H.N. Bharathi, Jyothi Rao, “Stock Trend Prediction using News Sentiment Analysis”, International Journal of Computer Science & Information Technology (IJCSIT), Vol. 8, No. 3, pp67-76, July 2016.

[7] Steidlmayer, J. P. & Hawkins, S. B. “Steidlmayer on Markets: Trading with Market Profile”, John Wiley & Sons, Inc. NJ, 2003.

[8] James F. Dalton, Robert Bevan Dalton, and Eric T. Jones, “Markets in Profile”, John Wiley, New York, 2007.

[9] Krauss, C., Do, X. A., Huck, N., “Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500”, European Journal of Operational Research Vol 2, pp. 689-702, 2017.

[10] K. Amudha, S. Padmapriya, D. Saravanan, "*Ensemble Forecasting Through Learning Process*", International Journal of Computer Sciences and Engineering, Vol. 06, No. 02, pp. 332-335, 2018.

[11] S. Velankar, S. Valecha and S. Maji, "Bitcoin price prediction using machine learning," *20th International Conference on Advanced Communication Technology (ICACT)*, pp. 144-147, 2018

[12] Kondor, D., Posfai, M., Csabai, I. & Vattay, G. “Do the rich get richer? an empirical analysis of ´ the bitcoin transaction network.” Public Library of Science, Vol. 9, pp. 1–10, 2014.

[13] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, S. C. Kremer, J. F. Kolen, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies" A Field Guide to Dynamical Recurrent Networks, Piscataway, NJ, USA: IEEE Press, 2001.

[14] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, ‘‘Convolutional LSTM network: A machine learning approach for precipitation nowcasting,’’ Neural Information Processing Systems, pp. 802–810, 2015.

[15] Bradley, A. P., “The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms”, Pattern Recognition, Vol. 30, No. 6, pp. 1145–1159, 1997.

Authors Profile

Pranshu Rupesh Dave is pursuing Bachelor of Technology in Computer Engineering from Birla Vishvakarma Mahavidyalya Engineering College, India. He is in the fourth year of his studies. His research interest include deep learning, artificial intelligence and finance. He has been fascinated by computers since he was a child and aspires to contribute to research in the field.

Priti Srinivas Sajja has been working at the Post Graduate Department of Computer Science, Sardar Patel University, India since 1994 and presently holds the post of Professor. She specializes in Artificial Intelligence and Systems Analysis & Design especially in knowledge-based systems, soft computing and multi-agent systems. She is author of Essence of Systems Analysis and Design (Springer, 2017) published at Singapore and co-author of Intelligent Techniques for Data Science (Springer, 2016 and Knowledge-Based Systems (J&B, 2009) published at Switzerland and USA, and four books published in India. She is supervising work of a few doctoral research scholars while eight candidates have completed their Ph.D research under her guidance. She has served as Principal Investigator of a major research project funded by the University Grants Commission, India.