Trend Prediction Classification for High Frequency Bitcoin Time Series with Deep Learning

SoReZ

https://www.zerotheorem.com

October 11, 2021

Highlights

- Applied combine deep learning approach for trend prediction classification for bitcoin time-series data.
- Understanding the working of Random Sampling Method (RSM) using non-stationary time series data.

Background

With the advancement in deep learning algorithms, machine learning is applied to time series data for prediction. Hence, this study focuses on how systematic ML will work on high-frequency trading data for one minute. Implementation of Deep learning requires large-scale data set for learning while applying one-minute trade data set is not acceptable and it removes non-stationary data from the training set while learning. This became challenging in recent times when cryptocurrency is in huge demand. However, the existing approach used an order-driven dataset for deep learning and trained recurrent neural networks with variations in data that might be able to understand non-stationary patterns. They had no promising results as time series data continuously changes in training and testing which introduces noise in the prediction phase. Therefore, there is a need to develop a method that works on time-series data that would be able to understand non-stationary patterns and produce better outcomes.

Introduction

In this context, Shintate and Pichl (2019) introduced a metric learning approach such as Random Sampling Method that aims to address limitations by observing accurate samples for prediction. For this purpose, recent data is considered that is more accurate in bitcoin for predicting trend that restricts data on current samples rather than existing records. This approach in time-series is more beneficial to target highly non-stationary patterns and reduce class variance issues.

Conceptual Framework

In this framework, Shintate and Pichl (2019) used three types of feasible events for price concerning time series. Sample belonging to the same type will categorize to one class such as label up patterns will be more similar to label down or static patterns. Starting with collecting all recent data with the hypothesis that data is non-stationary. The collection of samples is based on the closed interval $[t-k-l, t-k]$ where l determines the size of the interval, k is the window size and t is time.

After data collection, preprocessing takes place to remove outliers and some threshold alpha is adjusted to eliminate sample sequence whose value is greater than the threshold. In this regard, cosine similarity is used to compute the hidden representation of input samples. Moreover, the encoder is attached to pass the sample sequence to the hidden network t^{th} refers to input sequence and samples are labeled as $\mathbf{x}_t^{(up)}, \mathbf{x}_t^{(\text{down})}$, and $\mathbf{x}_t^{(\text{static})}$. Encoder passes all these sequence samples independently to LSTMNet and converts t^{th} to $\mathbf{h}_t^{(\text{input})}$ using equation

$$\mathbf{h}_t = \text{LSTMNet}\left(\mathbf{x}_t\right)$$

In the next phase $\mathbf{h}_t^{(\text{input})}, \mathbf{h}_t^{(\text{up})}, \mathbf{h}_t^{(\text{down})}$, and $\mathbf{h}_t^{(\text{static})}$ are passed to bi-directional LSTM which produces hidden features of the sample by the equation

$$\hat{\mathbf{h}}_{ti} = \overrightarrow{\mathbf{h}}_{ti} + \overleftarrow{\mathbf{h}}_{ti} + \mathbf{h}_{ti}$$

At last, the ratio is computed among hidden features of input sequence which are produced using metric learning approach and sample with cosine similarity. The class which is more similar to the sample is declared as output.

Details of Conceptual Framework

A combination of Multi-Layer Perceptron and LSTM network is used for predicting trends. ReLU is used as an activation function in all layers with Adam optimizer. Learning rate is set to 0.3. Both networks have two layers with 32 hidden units in each. A probability measure is calculated for MLP, LSTM, and RSM that shows which class has maximum probability.

Results and Discussion

From the maximum probability achieved by all three models, metric scores of accuracy, recall, precision and F1 score evaluated to get accurate results. According to findings it is clear that RSM model outperforms the accuracy of LSTM and MLP model which depicts that proposed method can be applied to available financial factors. Below Table 1 shows representation of metric scores.

	Accuracy	Recall	Precision	F1 Score
MLP	0.4766	0.4570	0.4822	0.4511
LSTM	0.4788	0.48770	**0.5581**	0.4657
RSM	**0.5353**	**0.5182**	0.5458	**0.5092**

Table 1: Model evaluation scores on accuracy, recall, precision, and F1 measure.

Conclusion

This leads to the conclusion that a new metric learning approach known as the random sampling method performed better in comparison to MLP and LSTM. Its major concern was to target non-stationary sample patterns and removing noise in the prediction phase. In the future, this methodology can be applied to other financial factors affecting time series data.

References

Shintate, T. and Pichl, L. (2019). Trend prediction classification for high frequency bitcoin time series with deep learning. *Journal of risk and Financial management*, 12(1):17.