A Gated Recurrent Unit Approach to Bitcoin Price Prediction

SoReZ

https://www.zerotheorem.com

October 11, 2021

Highlights

- Implement a framework that predict BitCoin pricing effectively by using set of exogenous and endogenous factors that outperforms the traditional approaches.
- Understand the working of seasonal auto regressive integrated moving average model that proves to be the best architecture.

Background

With the advancement in technology, cryptocurrency garnered a lot of attention. Hence, to predict BitCoin price and volatility, different machine learning methods have been implemented. However, traditional neural networks have several shortcomings in effectively using prior information for future predictions of Bitcoin price. Therefore, there is a need to investigate a framework with a set of advanced machine learning algorithms that consists of set of exogenous along with endogenous factors to predict daily prices of BitCoin.

Introduction

To overcome the loopholes related to traditional approaches, sequence models with robust feature engineering are required that predict future pricing. In this context, Dutta, Aniruddha and Kumar, Saket and Basu, Meheli (2020) focus on two aspects such as exogenous and endogenous factors to predict BitCoin price along with introducing the set of advanced machine learning forecasting methods to predict daily BitCoin prices. Moreover, Dutta, Aniruddha and Kumar, Saket and Basu, Meheli (2020) compared the recurrent neural network (RNN) models with traditional machine learning models by proposing a gated recurring unit (GRU) model as the best architecture to predict Bitcoin price.

Proposed Methodology

For this purpose, Dutta, Aniruddha and Kumar, Saket and Basu, Meheli (2020) uses neural network models for the Bitcoin price prediction which are simple neural network (NN), Long Short-Term Memory (LSTM) and GRU. In this regard, the neural networks are initially trained with optimum hyper-parameters before being tested on the test set. Following that, RNN with an LSTM and GRU with dropout and repeated dropouts are learned and deployed. Finally, a seasonal auto regressive integrated moving average (SARIMA) is used to make a comparison. This model is basically used to investigate seasonality where time series models fail to capture long term dependencies in the presence of high volatility, which is an inherent characteristic of a cryptocurrency market. To have a better understanding of the proposed methodology, let's have a look at the details.

Details of Proposed Methodology

In this regard, a single LSTM unit used that is composed of a cell, an input gate, a forget gate (a tanh layer and sigmoid layer) and an output gate. This means that the flow of information which is in and out of the LSTM cell is controlled by the gates. Furthermore, the GRU has similar properties to LSTM but it combines the input and the forget gates of the LSTM to form a single update gate which depicts that a GRU unit consists of a reset gate, an update gate and current memory content. Moreover, to check the robustness of the proposed architecture, model is compared with tradational neural network and LSTM where the update gate is represented by z_{t}, where x_{t} is the input and $h_{\text{t-1}}$ is the output which are represented in such a way $z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$, $r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$, and $h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot h_t$ respectively.

Moreover, for the simple NN, 2 dense layers are used, each with hidden nodes 1 and 25. The simple GRU as well as the GRU with recurrent dropout architecture comprised of one dense layer with 1 node and one GRU layer (50 nodes). The final GRU architecture is tuned with two GRU layers (10 nodes and 50 nodes) with a dropout and recurrent dropout of 0.1. The LSTM layer is modelled with one dense layer (1 node) and one LSTM layer (50 nodes). The optimized batch size for the RNN models and the neural network are determined to be 100 and 125 respectively. A higher batch size led to validation loss and higher training during the learning process.

Results and Discussion

In order to evaluate the proposed model, all the neural network models are tested on the test data together with the SARIMA model. For this, the RMSE for all the models on the train and test data is evaluated which shows that the LSTM architecture have better performance than the simple NN architecture due to memory retention capabilities. As shown in Table 1, the GRU model with a recurrent dropout generates an RMSE of 0.017 on the test set and 0.014 on the training set. RNN-GRU performs better than LSTM and a possible reason is the fact that GRUs are computationally faster with lesser tensor operations and number of gates. These results indicate that the GRU with recurrent dropout is the best performing model for predicting Bitcoin prices.

Table 1: Train and Test RMSE of 30 days lookback period

Model	RMSE Train	RMSE Test
Neural Network	0.020	0.031
LSTM	0.010	0.024
GRU	0.010	0.019
GRU-Dropout	0.014	0.017
GRU-Dropout-GRU	0.012	0.034
SARIMA	0.034	0.041

Conclusion

This led to the conclusion that recurrent neural network models such as LSTM and GRU out-performs conventional time-series approaches like SARIMA for price prediction. From the results it is clear that, With limited data, neural networks like GRU and LSTM can regulate past information to learn effectively from non-linear patterns. Moreover, adding a recurrent dropout will improve the performance of the GRU architecture, but still there are room for improvements that need to be carried out to investigate the dropout phenomenon in GRU architectures.

References

Dutta, Aniruddha and Kumar, Saket and Basu, Meheli (2020). A gated recurrent unit approach to bitcoin price prediction. *Journal of Risk and Financial Management*, 13(2):23.