Capstone Proposal

Weekly Bitcoin Price prediction using Sequence to Sequence model

Hamed Niakan
Nov 2020

1 Definition

1.1 Domain Background

This study us to forecast the bitcoin values based on historical data which consists of intrinsic patterns. In order to come up with a model to capture the behavior of coin values, a model to define the important features of the time series pattern and explain how the past affects the future or how two-time series can "interact". In [3], different neural network methods to forecast the Bitcoin future prices mentioned. He explored a Recurrent Neural Network with LSTM cells to predict the future price.

In this research, we introduced a Sequence to Sequence model (Seq2Seq) inspired from [4] as a learning method for natural language processing for the purpose of this study.

1.2 problem statement

Due to high volatility state of bitcoin price in the market, we are to predict the weekly future price of Bitcoin by observing the Bitcoin's past price data. We explored a Seq2Seq and Stack LSTM to get to the goal and compared the performance of each one.

The result shows that Seq2Seq has a better performance compared to the mentioned methods in this study and represent an average MAPE of almost 1.

2 Analysis

2.1 Datasets and Inputs

"Coindesk" api provides the researcher with historical data of USD and EUR bitcoin value. For this research we clone the recent 2500 daily price of bitcoin and split it into 80% trainset, 10 % validate set and test set 10%. For the preprocessing we select a window to split data as a observed sequence and future sequence which we have to predict the values.

2.2 Algorithm and Techniques

Natural Language processing was the first application explored the Seq2Seq approach [3] as their learning algorithm. In essence, sentence structure is a sequence of a words that can be treated as an observation over time series because it has a sequential model and it also has sequential input-output pairs. The methodology they introduced was to exploit multi-layered LSTM to map the input sequence to a vector with a fixed hidden state and then the hidden state of the fixed dimension goes to another deep LSTM to decode the target sequence from that vector. The sequence to sequence model is created of two blocks of LSTMs, which form an encoder and decoder block as shown in Fig 2.

Sequence to sequence NLP application [3] results showed that for translating from English to French based on the WMT-14 dataset was outperforming the other methods for the same application. Other than translation, Seq2Seq models can be extensively used for the dynamic, spatial-temporal characteristic of multivariate time series date.

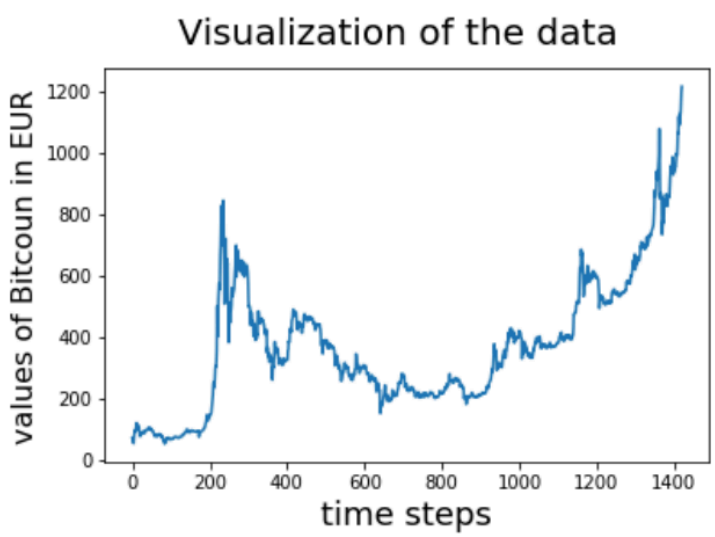

Figure 1 Visualization of the data, Bitcoin value in terms of USD

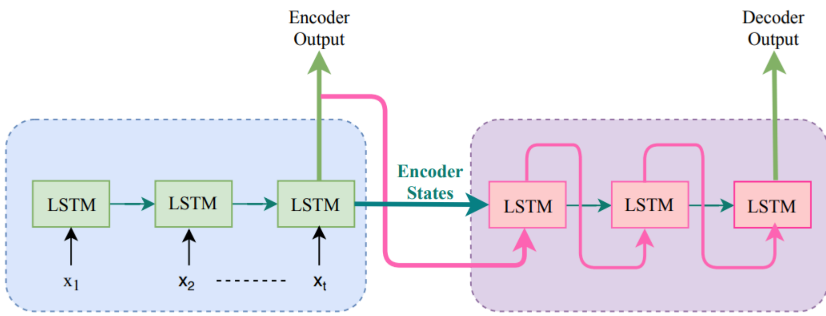

Figure 2 Sequence to Sequence model: This mode uses ”encoder-decoder” and maps input sequence with ”encoder states” and use it as initial state of the decoder.

3 Methodology

3.1 Pre-processing

Providing three-dimensional dataset matched with different deep learning platform is a first and tedious part of the learning. Other than there are a couple of steps like normalization and encoding. We split the data to three parts: Training, Validation and Testing and creating the labels for each of sets. The labels are actually the sequence of the sentence we put for the prediction period to make it as a supervised learning problem.

Furthermore, using sliding window we recreate datasets so that to have a supervised learning set up. We put earlier time stamps as input feature and the next stamps for target labels.

By doing this sliding window, we select a window of 7 for input features and 7 for the prediction. If one of them are shorter using the padding, we can pad it with zero. The total size of dataset is 1407, and we train the Stack LSTM and Seq2Seq models using these processed data.

3.2 Implementation

In this research, we tried two different approaches using time series prediction. In this study Stack sq2seq is implemented using GRU cells using keras Library to implement it and we constructed a data generator for training the model.

Next implementation was Seq2Seq model. Instead of Keras we used TensorFlow to have a better control over the network in order to tune hyper parameters however all the hyper param tuning was done manually to tune this model.

This model consists of two parts, encoder and decoder both of these blocks have block of LSTM/GRU, which I used GRU (it is really similar to LSTM).

3.3 Refinement

There are lots of parameters in neural network and in specific for seq to seq or time series prediction that can be tuned and some of them are really necessary. These hyper parameter are the length of window for training size, hidden layers, learning rate, batch size learning rate and so on.

Based on my results the best value for past window size = 7, number of stack LSTM in Seq2Seq is 2 and number of hidden neurons = 100, learning rate = 0.007, and batch size = 100.

4 Results

I improved the prediction result for both method and I improved the result from stack lstm and got better with Seq2Seq model. In Fig. 4, you can see the prediction from Stack LSTM. In Fig. 4 we can see the prediction with sequence to sequence model.

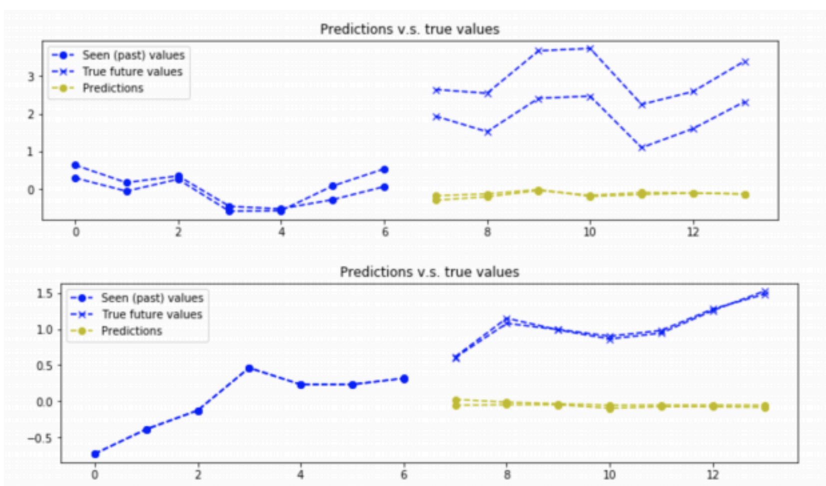

Figure 3 Prediction using stack lstm for Bitcoin value in terms of both USD and EUR.

5 Evolution Metrics

For this problem we use the mean absolute percentage error (MAPE) metric
like [2] which is defined as:

$$\frac{1}{N}\sum_{n=1}^{N} \left| \frac{y_{pred} - y_{actual}}{y_{actual}} \right| \quad (1)$$

I used this matrix to compare the result of these two methods, I compete the MAPE for value of batch data and then I report the average of this value by computing the mean of MAPE for each batch size. As a result the MAPE from Seq2Seq was the half of MAPE from stack LSTM. Finally, I got 0.9% as the MAPE in training process with Seq2seq and the average of MAPE in test process is 1.0

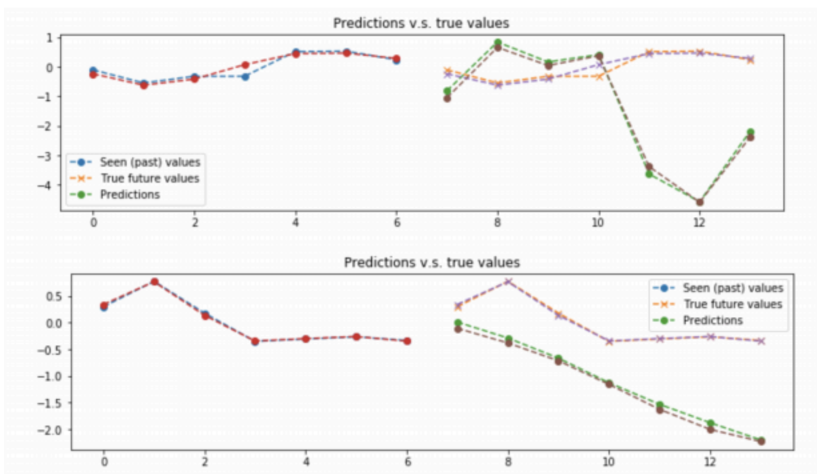

Figure 4 : Prediction using Seq2Seq model for Bitcoin value in terms of both

References

[1] Zheshi Chen, Chunhong Li, and Wenjun Sun. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365:112395, 2020.

[2] Suhwan Ji, Jongmin Kim, and Hyeonseung Im. A comparative study of bitcoin price prediction using deep learning. Mathematics, 7(10):898, 2019.

[3] Marco Santos. Predicting bitcoin prices with deep learning, 2019.

[4] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In Advances in neural information processing systems, pages 3104–3112, 2014.