Capstone Proposal

**Weekly Bitcoin Price prediction using
Sequence to Sequence model**

Hamed Niakan
Nov 2020

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.

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.

3. 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.

4. solution statement

To improve the performance of traditional time series models like ARMA, ARIMA, SARIMA or statistical learning, we used deep

learning platforms. Since the recurrent models has the capability of learn the pattern in the past and has a temporal memory, it can be the good candidate for this set up if the problem.

5. Benchmark Model

• Other statistical methods like ARIMA has accuracy between 50−55%,
[1].

6. Evolution Metrics

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

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

7. Project Design

Step 1: Cloning the data and preprocess to create a time frame by iterating over the selected window and split it into train, validation and test dataset.
Step 2: developing the structure of the stack LSTM model and programing and preparing the scripts.
Step 3: Using AWS SageMaker creating a package with all utility function and creating train and predict scripts.
Step 4: Train and test the model (Stack LSTM and Seq2Seq).
Step 5: Compare them in terms of the mean absolute percentage error (MAPE).

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.