Project Proposal

Zaiyi Liu zl2729
Lin Ye ly2458
Akshat Mittal am5022

October 9, 2019

1 Description

Goal The goal for this project is to use statistical approach to predict bitcoin price movement using one-minute binned bitcoin price. With Efficient Market Theory, predicting price movement of an asset class has been challenging. We are hoping to use a Bayesian approach to capture the latent relationship that has not yet been discovered by the general public.

Data To create a trading strategy of Bitcoin, we gather the open/close/high/low price of Bitcoin. The data set contains time series of bitcoin prices from two exchanges, namely Coinbase and Bitstamp, from Jan 2012 to August 2019. The dataset contains minute to minute updates of OHLC (Open, High, Low, Close), Volume in BTC and indicated currency, and weighted bitcoin price.

2 Proposal

Modeling With Hidden Markov Models, we will use current state price to predict the next state price. In our approach, we would like to consider both price change from open to next open, and high/low-to-open price during the one-minute interval to capture additional information.

The first step is to construct the features from raw data. We define n features, and each feature is a function of our observed variables, i.e. $F_{it} = f_i(o_{1:t}, h_{1:t}, l_{1:t}, c_{1:t})$, where $i = 1, 2, \ldots, n$. Here are some examples of the formula of our features.

Return at time t $\log(\frac{c_t}{c_{t-1}})$
High-to-open return at time t $\log(\frac{h_t}{o_t})$
Low-to-open return at time t $\log(\frac{l_t}{o_t})$

Now we begin to formulate our latent variables. At each time t, we assume that there are total m different states, and the state $s_t = Cat(g(s_{t-1}, F_t))$, where

$F_t = (F_{1t}, F_{2t}, ..., F_{nt})$ served as exogenous variables. g is a non-linear function which could be a neural network. The observed variables Y_t are generated as the Mixture Model.

$$Y_t \sim N(\mu_{s_t}, \Sigma_{s_t})$$

Where μ is a multi-variate gaussian distribution with $\mu = \mu_0$ and $\Sigma = \Sigma_0$. The prior of μ and $Sigma$ could be discrete distribution whose parameter is generated by by Dirichlet distribution. For simplicity, we might set Σ as a diagonal matrix.

Inference Our target is to find optimal g, μ, and Σ given observed variable Y.

$$\sum_{i=1}^{n} log(P(g, \mu, \Sigma | Y_t)$$

. We will use Black-Box variational inference algorithms.

Criticism Mean Square Error is a candidate for our loss criteria. (Not quite sure about Criticism).