Trading Bitcoin and Online Time Series Prediction

Adel SOHBI
Seoul Artificial Intelligence Meetup
30th September 2017

Article Review

Trading Bitcoin and Online Time Series Prediction

Muhammad J Amjad MAMJAD@MIT.EDU
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

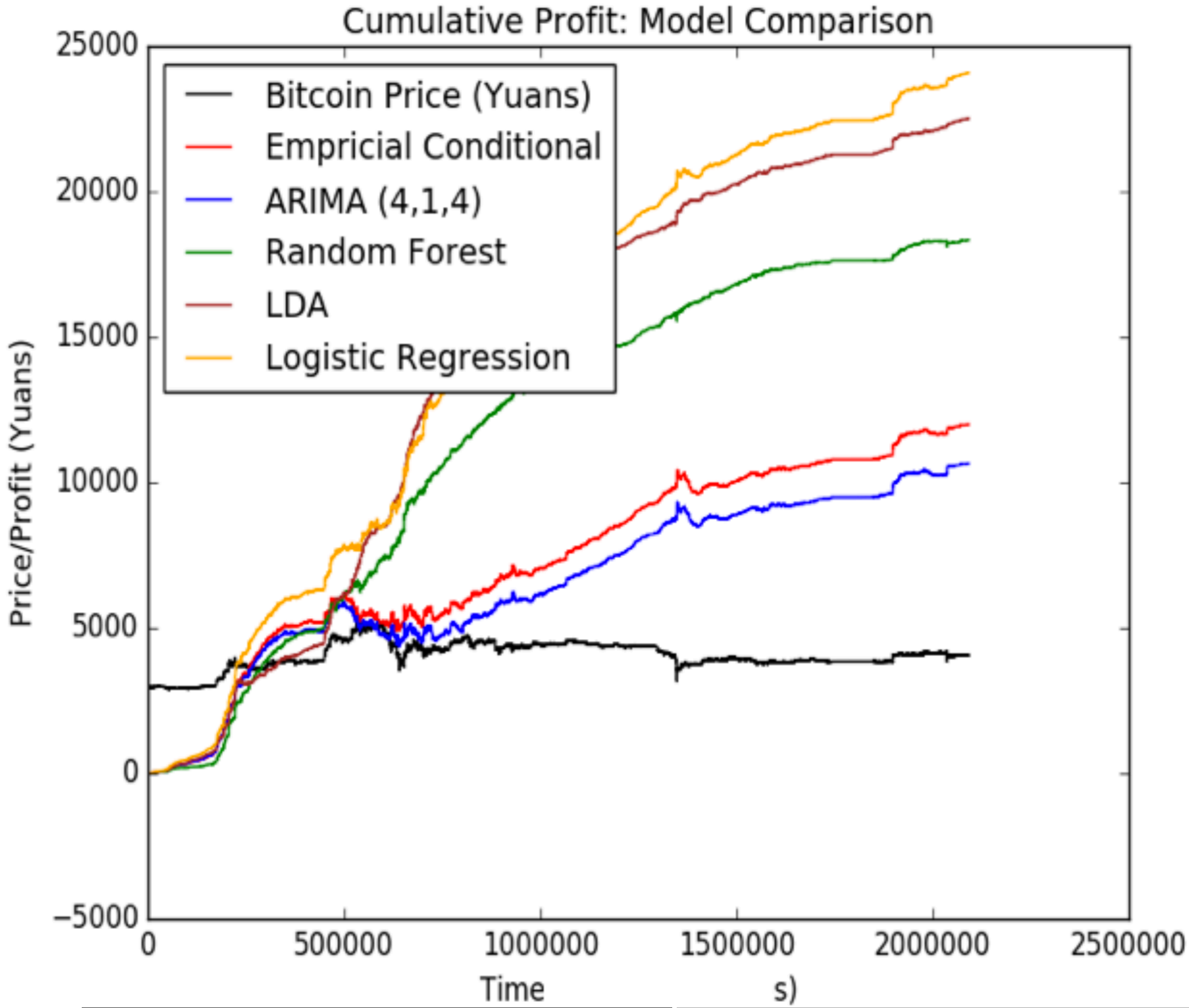
Devavrat Shah DEVAVRAT@MIT.EDU
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technolog
Cambridge, MA 02139, USA

Editor: Oren Anava, Marco Cuturi, Azadeh Khaleghi, Vitaly Kuznetsov, Alexander Rakhlin

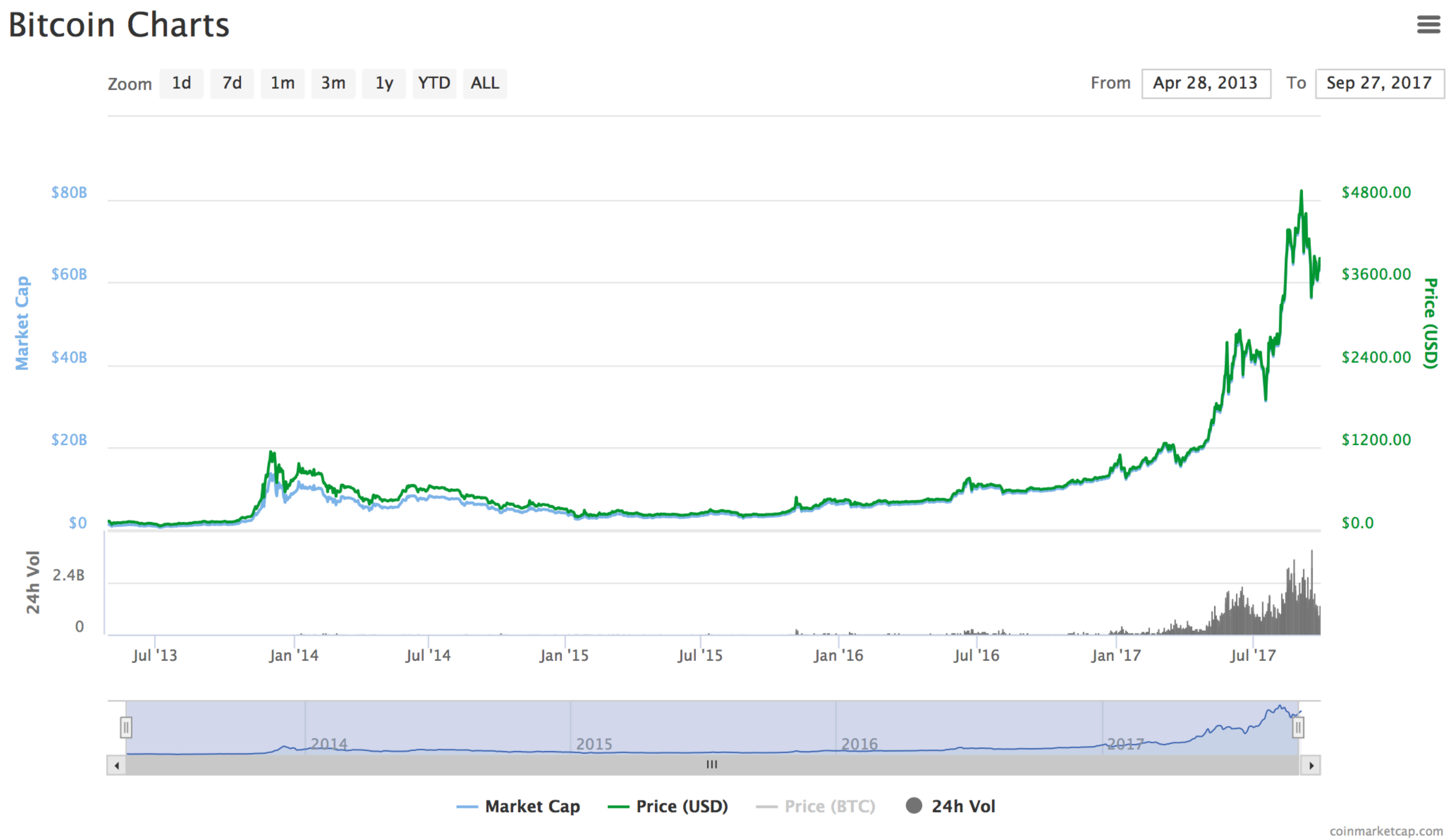
Abstract

Given live streaming Bitcoin activity, we aim to forecast future Bitcoin prices so as to execute profitable trades. We show that Bitcoin price data exhibit desirable properties such as stationarity and mixing. Even so, some classical time series prediction methods that exploit this behavior, such as ARIMA models, produce poor predictions and also lack a probabilistic interpretation. In light of these limitations, we make two contributions: first, we introduce a theoretical framework for predicting and trading ternary-state Bitcoin price changes, i.e. increase, decrease or no-change; and second, using the framework, we present simple, scalable and real-time algorithms that achieve a high return on average Bitcoin investment (e.g. 6-7x, 4-6x and 3-6x return on investments for tests in 2014, 2015 and 2016), while consistently maintaining a high prediction accuracy (> 60-70%) and respectable Sharpe Ratio (> 2.0). Furthermore, when trained on a period eight months earlier than the test period, our algorithms performed nearly as well as they did when trained on recent data! As an important contribution, we provide a justification for why it makes sense to use classification algorithms in settings where the underlying time series is stationary and mixing.

The Results


Cumulative Profit: Model Comparison
Bitcoin Price (Yuans)
Empricial Conditional
ARIMA (4,1,4)
Random Forest
LDA
Logistic Regression
Price/Profit (Yuans)
25000
20000
15000
10000
5000
0
−5000
0
500000
1000000
1500000
2000000
2500000
Time
s)

Why Trading Bitcoin?



Aim of the Article

- **Forecasting Bitcoin price changes for algorithmic trading.**

- Make **scalable** and **accurate** forecasts in **real-time**, given a live stream of time series data.

The Problem

- **Prediction:** For any time *t*, given the historical price time series up to time *t*, predict the price for future time instances, $s \geq t + 1$.

- **Trading:** For any time *t*, using current investment and predictions, decide whether to buy new Bitcoins or sell any of the Bitcoins that are in possession.

Simple Trading Strategy

- Focus **predictions accuracy**. The trading strategy is to demonstrate the **utility** of predictions.

- Trading Model:

$$d[t] = \begin{cases} \text{buy}, & \text{if } h[t] = 0 \text{ \& price is predicted to increase, with high confidence} \\ \text{sell}, & \text{if } h[t] = 1 \text{ \& price is predicted to decrease, with high confidence} \\ \text{hold}, & \text{otherwise.} \end{cases}$$

Where *h[t]* = 1, if we are in possession of a Bitcoin at time t and *h[t]* = 0, otherwise.

Article Contributions

- Theoretical framework for time series analysis based: **stationarity** and **mixing**.

- **Simple**, **scalable**, **real-time** algorithms for prediction and trading that yield **high prediction accuracy** and **highly profitable returns** on investment in Bitcoin.

Stationary and Mixing Time Series

- A time series is said to be:
  - **Stationary** if its joint probability distribution is time-invariant.
  - **Mixing** if the distribution at a specific time is primarily dependent on the recent past.
- Classical time series regression algorithm : **ARIMA** (**A**uto**R**egressive **I**ntegrated **M**oving **A**verage) has poor performances.
- **Overcome** this limitation with a new model.

Bitcoin Data

- OKCoin exchange using their APIs. All prices are reported in Chinese Yuans. The APIs return lists of bid[t] and ask[t] at the time t.
- Several months of data from the exchange in 2014, 2015 and 2016.
- **Estimate** of price:

$$p[t] = (\max(\mathsf{bid}[t]) + \min(\mathsf{ask}[t]))/2$$

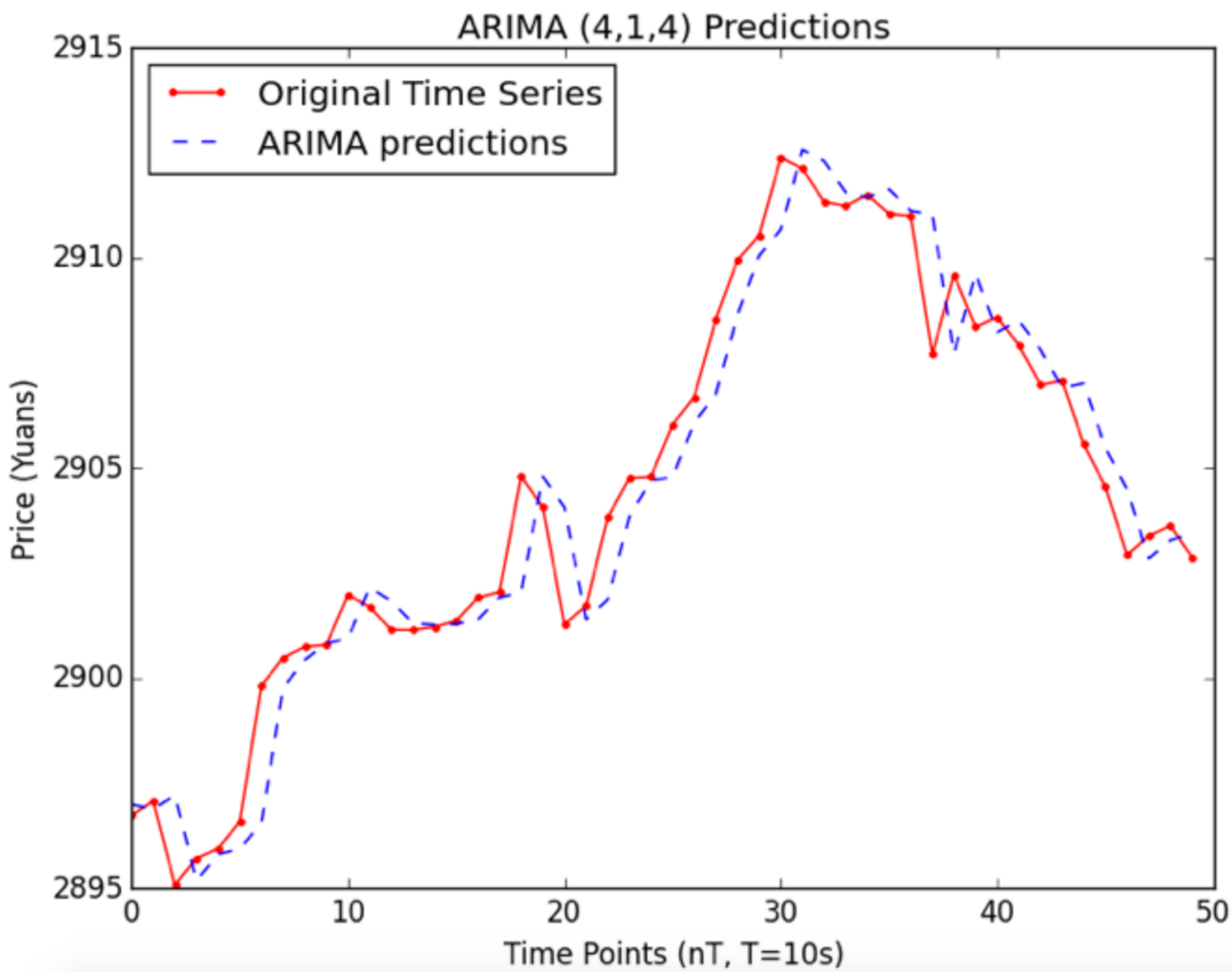
Stationarity and Mixing

- Tests reveal that the Bitcoin price time series is **not stationary**.
- Let y[t] be the time series produced by **the first-differences**:

$$y[t] = p[t] - p[t-1]$$

- The first-differences of the Bitcoin price time series is **stationary** and **mixing**.

Classical Modeling: ARIMA



Stationarity and Mixing

- **Stationary** and **mixing** suggest there exist a large enough d such that:

$$\mathbb{P}(y[t]|y[-\infty : t-1]) \approx \mathbb{P}(y[t]|y[t-d : t-1])$$

- **Stationary**:

$$\mathbb{P}(y[t]|y[t-d : t-1]) \approx \mathbb{P}(y[s]|y[s-d : s-1]), \forall s$$

$$F_\theta(y[t-d : t-1]) = \mathbb{P}(y[t] > \theta|y[t-d : t-1])$$

The Model

- The model:

$$x[t] = \begin{cases} -1, & \text{if } y[t] < -\theta \\ 1, & \text{if } y[t] > \theta \\ 0, & \text{otherwise.} \end{cases}$$

- Define a 3-dimensional probability vector:

$$\mathbb{Q}[t] = (\mathbb{P}(x[t] = \sigma | y[t-d : t-1]), \sigma \in \{-1, 0, 1\})$$

$$\mathbb{Q}[t] = F(y[t-d : t-1])$$

Method

- Interest is truly in this 3-dimensional.
- The problem effectively reduces to ternary-state classification
- Generate training data as follows:

$$(x[t], y[t-d:t-1])$$

Method

- The goal is to learn $F(\cdot) \in [0, 1]$, using a classification algorithm.

- Choice of **classification algorithms:**
  - Random Forest (RF),
  - Logistic Regression (LR) and
  - Linear Discriminant Analysis (LDA).

- Learn the transition probability distribution function $P(x[t]|x[t-d : t-1])$.

Method

- We require the model to **produce a probability** associated with each prediction to give us a sense of **confidence** in each individual prediction.

- Trading Model:

$$\hat{x}[t] = \begin{cases} \sigma^* \text{ if } \mathbb{P}^*(x[t] = \sigma^* | h_d[t-1]) \geq \gamma, \\ 0, \text{ otherwise,} \end{cases}$$

- Combining **multiple predictions:**

$$\hat{x}_w[t] = \sum_{d \in \mathbb{S}} \hat{x}_d[t] \times w_d$$

Experiments and Results

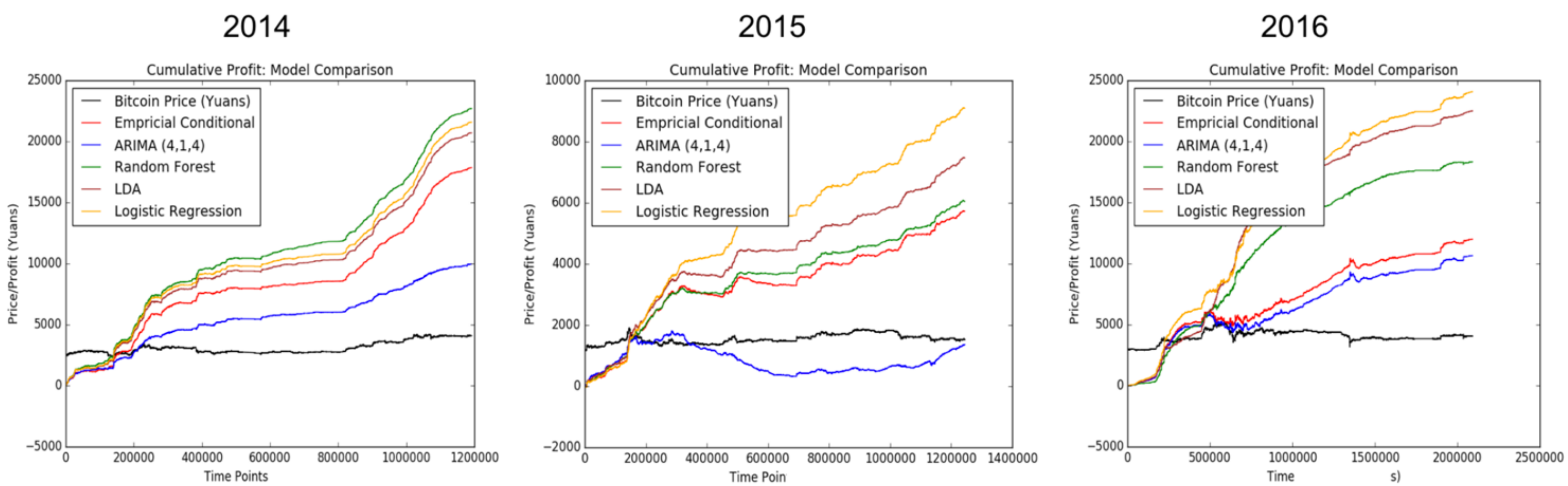


Figure 3: Cumulative Profit and Bitcoin Price in 2014, 2015, 2016 with $d \in 3, 4, 5$. γ selected via validation. Each time step represents 5s.

(Left): Training: 2/16/14 - 3/14/14, Validation: 3/15/14 - 3/31/14, Test: 4/1/14 - 6/11/14.

(Center): Training: 12/1/14 - 12/31/14, Validation: 1/1/15 - 1/15/15, Test: 1/16/15 - 3/31/15.

(Right): Training: 2/26/16 - 4/15/16, Validation: 4/16/16 - 5/15/16, Test: 5/16/16 - 9/15/16.

Quick Summary

- Learning framework based on properties such as **stationarity** and **mixing**.
- **Limitations** of **classical time series methods** like the ARIMA models.
- Prediction **accuracy** near **70%** with Classification Algorithms.
- **High return** on average Bitcoin investment (e.g. 6-7x, 4-6x and 3-6x return on investments for tests in 2014, 2015 and 2016),