A PROPOSAL FOR PREDICTIVE ANALYSIS OF CRYPTO-CURRENCY LIKE RIPPLE

Utkarsh Singh, Abhyudaya P. Singh, and Pratik Parida, Student*, K.I.I.T*

***Abstract*— this paper scrutinizes the predictive power of Data Analytics in the field of Crypto-currency. We address the question of the future rise and fall in the price of Crypto-currency for the benefit of the investors. Data Analytics is undisputed in the field of prediction, helpful in various researches giving a better outcome of what to expect in the future. The rise in the popularity of crypto-currency and subsequently its investors has made it a compelling area in research. Recent rise and fall in the crypto-currency market has made it necessary for a better predictor so as to give a more accurate prediction to the investors as well as market analysts.**

**INTRODUCTION**

Crypto-currency is a digital asset designed to work as a medium of exchange that uses *cryptography* to secure its transactions, control creation of additional units, and to verify the transfer of assets. They are also referred to as Alternative Currencies, as most crypto-currencies are mediums of financial exchange that exists outside the bonds of State Monetary Policy. These currencies are marked by *Decentralized Control* i.e. supply and values are controlled by the activities of their users and highly complex protocols programmed into their governing codes. The decentralized control of each crypto-currency works through a *Blockchain*, which is a public transaction database, functioning as a distributed ledger.

According to *Jan Lansky*, a crypto-currency is a system that meets the following conditions: 1. The system doesn’t require a central authority, distributed achieve consensus on its state. 2. The system keeps an overview of crypto-currency units and their ownership. 3. The system defines whether new crypto-currency units can be created. 4. Ownership of crypto-currency units can be proved exclusively cryptographically. 5. The system allows transactions to be performed in which owner of cryptography is unchanged. 6. If two different instructions for changing the ownership of the same cryptographic units are simultaneously entered, the system performs at most one of them.

Early origin of crypto-currency dates back to the year *1983*, when American cryptographer *David Chaum* conceived anonymous cryptographic electronic money called *ecash*. In 1998, Wei Dai published a description of “b-money”, an anonymous, distributed electronic cash system, followed by Nick Szabo who created “bit gold”. The first decentralized crypto-currency, Bitcoin, was created in 2009 by pseudonymous developer Satoshi Nakamoto. Later other crypto-currencies such as *Namecoin* (2011), *Ripple* (2012) were created.

*Ripple* is a *real-time gross settlement system* (RTGS), currency exchange and remittance network created by the Ripple company. It is built upon a distributed open source Internet protocol, consensus ledger and native crypto-currency called XRP (ripples).Its’ original authors were *Arthur Britto, David Schwartz, Ryan Fugger .*Ripple purports to enable "secure, instantly and nearly free global financial transactions of any size with no chargebacks." It supports tokens representing fiat currency, crypto-currency, commodity or any other unit of value such as frequent flier miles or mobile minutes. Ripple has been increasingly adopted by banks and payment networks as settlement infrastructure technology, with American Banker explaining that "from banks' perspective, distributed ledgers like the *Ripple system have a number of advantages over crypto-currencies like bitcoin.* In India, Axis bank and Yes bank are among the pioneers and will start using ripple in future.

# I. RESEARCH METHOD

## Problem Formulation and Research Question

In 2009, after the introduction of Bitcoin only a handful of people had heard about the term “crypto-currency” and only a limited people knew of its working and use.

Crypto currency gained most of its popularity through bitcoin in the United States of America, where people who had invested 100 USD were now worth millions of dollars. The people showcased their fortune and thus the money loving and tech-savvy people got onto the trend of investing money in crypto. It further shot to fame when bitcoin was accepted as a mode of payment by many organisations. At the moment over a 100,000 merchants accept Bitcoin worldwide. Notable among those are Microsoft and Expedia, as well as online electronic retailer Newegg.

Crypto-currency is dependent on various factors for their overall price in the market. Some of the *major factors* include:

* **Demand and Supply:** A crypto-currency supply is inversely proportional to its demand. If a currency has a lot of supply but very less demand then it will see very less price movement whereas a currency with less supply but greater demand will see a significant price movement like Bitcoin.
* **Market Sentiment:** Positive or negative market news about a crypto-currency affects its price movement also. Positive news even during the price fall of a crypto-currency helps to retain its customers as well as the market share.
* **Government Regulation:** Many governments have not yet approved of the crypto-currency as it does not abide by their monetary policies. A ban on such a crypto-currency by a government in a particular country creates a series of crumbling events for the particular crypto-currency.
* **Stability of the particular Crypto-currency network:** Stability in the market price of a crypto-currency is necessary so as to garner more support from the people. It gives a measure of its reliability. The Bitcoin price drop in the month of January and February led to unrest amongst many supporters. In case of ripple not much fluctuation was to be seen and thus is very trusted amongst Bankers.
* **Wider Mainstream Acceptance:** It basically stands for acceptance by the society. Unless and until a particular thing is accepted by the society it is likely to thrive. Acceptance of the crypto-currency by the people as a secure medium for trading or going about the daily purchase activities is a herculean task.
* **Technological Changes and Innovations:** The use of technology for the usage of cryptocurrencies for daily usage is really of great importance. Integration of bitcoins in *PayPal’s* payment system has raised awareness and stimulated a lot of interest in digital currencies among most people.

Many investors aren’t interested to pay attention to the following factors. The only thing they are concerned with is the *price of the crypto-currency at present and in the future*.

Even the new investors are far more interested to know the future cost of the crypto-currency so as to invest the right amount of money at the right time.

Ripple has gained prominence over the years and is one of the most sought after cryptos in the world. With its *stable market cap* unlike the bitcoin, it has come out to be the top 5 contenders in the market.

**In this project we formulate a method to predict the nearby future closing price of Ripple so as to help the investors and shareholders to have a better idea of what to expect in the upcoming days and buy and sell accordingly.**

## B. Collection and Cleansing of Dataset

The dataset is directly being retrieved from [www.coinmarketcap.com](http://www.coinmarketcap.com) , a website that contains all the repositories of crypto-currencies in the market using the following python package: 1. *BeautifulSoup4* 2. *html5lib*. The data obtained is available up to a week before the data is extracted.

*BeautifulSoup4* is a python package that sits atop an HTML or XML parser, providing Pythonic idioms for iterating, searching, and modify the parser time. *Html5lib* is a Python package that implements the HTML5 parsing algorithm which is heavily influenced by current browsers and based on the WHATWG HTML5 specification is a Python package that implements the HTML5 parsing algorithm which is heavily influenced by current browsers and based on the *WHATWG HTML5 specification*.

Data extracted is then made free of null values. Missing data or Null values *reduces the representativeness* of the sample and can therefore *distort inferences* about the Data. The data after being cleansed is stored in an excel sheet where it is tested for duplicate data and if found, the duplicate data is deleted.

*C. Model Implemented on the Dataset*

ARIMA (Auto Regressive Integrated Moving Average) model was implemented above the dataset. ARIMA is basically used for forecasting a time series. It is a class of model that captures a suite of different standard temporal structures in time series data.

The “AR” of ARIMA illustrate the evolving variable of interest is regressed on its own antecedent values. The “MA” indicates that the regression error is a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The “I” indicates that the data values have been replaced with the difference between their values and the previous values, where the difference might have occurred more than once. The motive of each of these features is to make the model *fit the data* as well as possible.

The model later on was coupled with the python package *“pickle”* which implements a fundamental, but powerful algorithm for serializing and de-serializing a Python object structure. The process is mentioned as *pickling.*

# II. EXPERIMENTAL SETUP

For the following research, Integrated Development Environment (IDE) *SPYDER* was used with Python 3.6. The python packages used for the following were Pandas, Numpy, pip, html5lib, Scipy, matplotlib, statsmodels. Data was obtained from [www.coinmarketcap.com](http://www.coinmarketcap.com) , which has an updated data repository of all the crypto-currency in the present market. The data obtained has attributes 1.Date 2.Open 3.High 4.Low 5.Close 6.Volumes 7.Market Cap.

ARIMA (Auto Regressive Integrated Moving Average) model was used for the testing the dataset and ultimately Trained according to the prediction. Later on the following model was used for the future prediction of the closing price of the RIPPLE for the upcoming week.

# III. TERMINOLOGY

1. **Market Cap:** The total value held in a crypto-currency. It is calculated by multiplying the total supply of coins by the current price of an individual unit.
2. **Volume:**total amount of particular crypto-currency trades in particular platform within 24 hours.
3. **Open/Close:** Open generally refers to the price at 12:01 AM UTC of any given day and close generally refers to the price at 11:59 PM UTC of any given day. UTC stands for *Coordinated Universal Time.*
4. **High/Low:** Refers to the highest and lowest value of crypto-currency on a particular day.
5. **Blockchain:** A blockchain, originally block chain, is a continuously growing list of records, called blocks, which are linked and secured using cryptography.
6. **Cryptography:** The practice of securing information in a communication from adversaries.
7. **Integrated Development Environment:** An integrated development environment (IDE) is a software suite that consolidates the basic tools developers need to write and test software. Typically, an IDE contains a code editor, a compiler or interpreter and a debugger that the developer accesses through a single graphical user interface (GUI). An IDE may be a standalone application, or it may be included as part of one or more existing and compatible applications.
8. **Pickling:** process whereby a Python object hierarchy is converted into a byte stream and “unpickling” is the inverse operation, whereby a byte stream is converted back into an object hierarchy. Pickling (and unpickling) is alternatively known as “serialization”, “marshalling” or “flattening”.

# iv. Hardware specification

## The project was implemented on a 64 bit Windows 10 operating system having an Intel core i5 processor paired with 16.0 GB of RAM.

# v. RESULT AND ANALYSIS

Of the dataset available with us, 70% of it was used for training and the rest 30% was used as testing dataset. The dataset was available from 4 August 2013 to 26 March 2018.

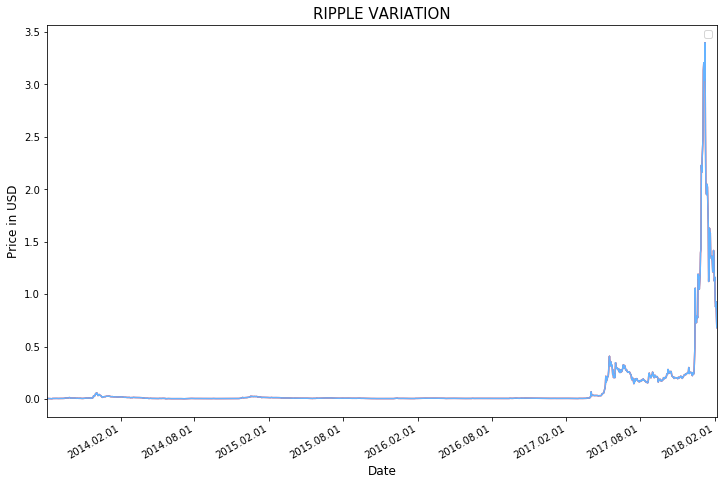
w

Fig 5.1 Actual Value Graph

The above graph depicts the original value of the dataset in US Dollars from the start of the ripple value.

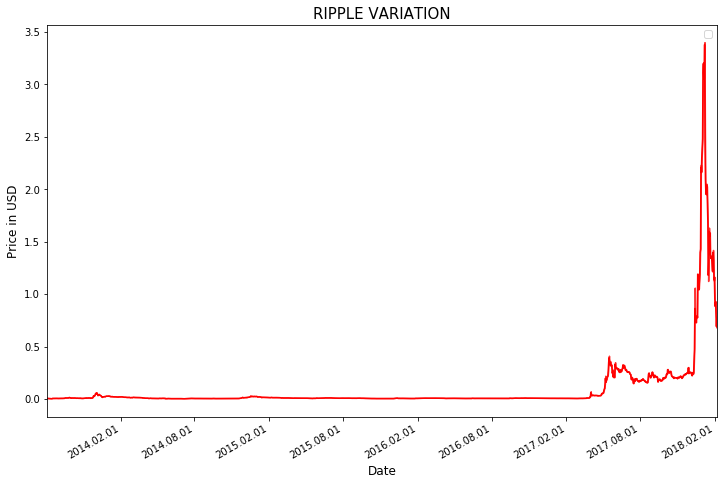


Fig 5.2 Prediction Graph

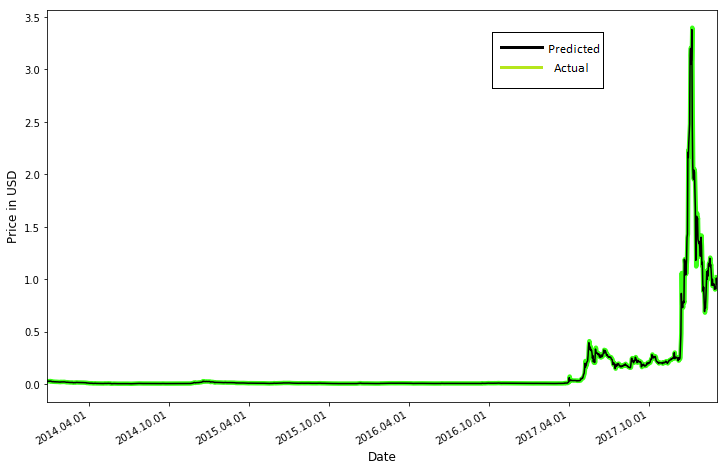


Fig 5.3 Actual vs Predicted

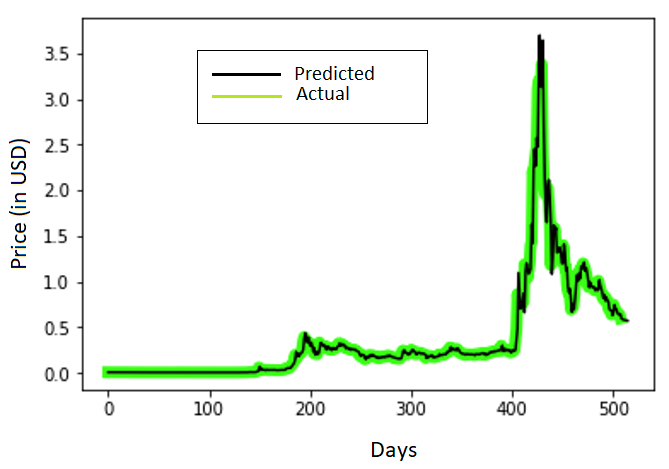
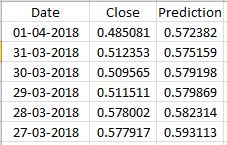


Fig 5.4 Graph with prediction of next 6 days



Actual closing price vs the predicted value

As observed the closing price and the predicted value for the next 6 days starting from 27 march is very near to the expected value and thus imposes the reliability of the model being used.

# vi. Conclusion

The ARIMA model used provided us with near accurate predictions for the given dataset. *Mean squared Error* of the prediction with respect to the testing dataset was found to be 0.01 and the *Variance score* is 0.97. With the same accuracy we can calculate the future closing price of Ripple for the next 7 days.

vii. Acknowledgment

The success and final outcome of the project required a lot of guidance and assistance and we are extremely privileged to have got this all along the completion of our project from many people.

We thank [Name Surname, title] for assistance with [particular technique, methodology], and [Name Surname, position, institution name] for comments that greatly improved the manuscript.

We would also like to show our gratitude to the (Name Surname, title, institution) for sharing their pearls of wisdom with us during the course of this research, and we thank 3 “anonymous” reviewers for their so-called insights. We are also immensely grateful to (List names and positions) for their comments on an earlier version of the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed persons.

viii. References

1. <https://www.kaggle.com/txtrouble/time-series-analysis-with-python>
2. <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>
3. <https://github.com/Ajaypal91/Predicting-Price-of-Cryptocurrency>
4. <https://www.kaggle.com/taniaj/cryptocurrency-predictions-with-arima/notebook>
5. <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>
6. <https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>
7. <https://dashee87.github.io/deep%20learning/python/predicting-cryptocurrency-prices-with-deep-learning/>
8. <http://statisticsbyjim.com/regression/choosing-regression-analysis/>
9. <https://blog.patricktriest.com/analyzing-cryptocurrencies-python/>
10. <http://ieeexplore.ieee.org/document/7028484/>
11. <http://ieeexplore.ieee.org/document/8125674/>