Forecasting Bitcoin Price using Machine Learning Algorithms.

Smt.Abha Marathe

*Vishwakarma Institute of Technology* Pune, India

Vyankatesh Kulkarni  
*Computer Engineering*

*Vishwakarma Institute of Technology* Pune, India

Kush Agarwal  
*Computer Engineering*

*Vishwakarma Institute of Technology* Pune, India

Shrihari Mangle  
*Computer Engineering*

*Vishwakarma Institute of Technology* Pune, India

*Abstract*—Today’s growing world has totally became the great example for the use of digital technology. Over the recent years, everything is going online and internet has become the keen provider for the required resources. Considering, these development in our mind the most important domain is the rapid and sudden development of crypto currency during these times is the most controversial development in the global economy. Adding to it, today’s market is most critical that even good practitioner cannot predict the things since, the fluctuation and the volatile market of this Bitcoin currency led the confusion among the investors. So, this paper mainly focusses on the various approaches to predict and hence to forecast the fluctuating price of the Bitcoin. The various machine learning based approaches are discussed here with good accuracy so that there should not be any blind decision taken regarding investments particularly in the Bitcoin trading. The results of this paper verifies the applicability of model and give a direction to investors on how these machine learning techniques are used in decision making.

Keywords—Bitcoin, Machine Learning, crypto currency, multiple linear regression, Random Forest, SVR.

# *Introduction*

Transfer of direct money from a person to another is hectic and also there is a chance of loss of money in many ways so the digital currency which is named Crypto currency is introduced. Bitcoin is the worlds’ most valuable cryptocurrency and is traded on over 40 exchanges worldwide accepting over 30 different currencies. It is one of the most popular cryptocurrencies around the world. However, it is not the first cryptocurrency that appears in the world. Satoshi Nakamoto first introduced Bitcoin and the concept of Blockchain in one of his papers in 2008[luo].It has a current market capitalization of 9 billion USD according to https://www.blockchain.info/ and sees over 250,000 transactions taking place per day.[mcnally] Digital currency is a method for trade which is web based and utilizes cryptographical capacities to perform money related exchanges. . The principle highlight of cryptographic money is that it isn't constrained by any fundamental power: the circulated pith of the blockchain makes digital forms of money hypothetically invulnerable to the old methods of government control and obstruction. Digital currencies support blockchain innovation to pick up decentralization, permanence, and straightforwardness. Bitcoin is an advanced installment that uses cryptographic money and distributed (P2P) mechanization to produce and oversee financial exchanges instead of a focal power. Time arrangement estimating or expectation is a notable issue [mahendra2020]. With a market capitalization of around 170 billion US dollar (September 2020), bitcoin represents about 58% of the cryptocurrency market.[1s2.0]

As Bitcoin's price has very high volatility, many people are attracted to the research of predicting the price trend of Bitcoin because investors have a chance to gain high profit from the price change. As more and more researches focus on the topic, the methods used for prediction also expand from only time series models to machine learning models and deep learning models. The input feature sets also expand from historical price and exchange volume to text data from social media, news, internet search, and more.[luo2020].

In this paper , we are going to predict the closing price and will be showing the price trends for the same. These predicted prices will be based on the previously observed and studied data form a dataset taken for building the project. The main input parameters which will be taken into consideration is that the opening price of the Bitcoin , how much it has gone high during the span of the day and the low price it has acquired. On this basis, the price at which it is going to close at the end of the day is the key prediction we have done while framing the various machine learning models. We have used the Multiple Linear Regression[MLR], Random Forest Algorithm and Support Vector Regression[SVR] for prediction purpose. Also we have plotted various graphs which will depict the actual and the predicted value trends for clear visualization purpose.

# LITERATURE REVIEW

The studies which are carried out for the prediction of Bitcoin prices are quite good to study since they are very few by implementing through machine learning algorithm concepts. But there are also some relevant studies carried out which gives the total review of how should we carry out our implementation towards this domain of Bitcoin price prediction.

[1](mcnally2018) performed the study on the dataset where they have used various deep learning algorithms and finally came to the conclusion that deep learning models such as the RNN and LSTM are evidently effective for Bitcoin prediction with the LSTM more capable for recognizing longer-term dependencies. Their results mainly says that , the LSTM achieved the highest classification accuracy of 52% and a RMSE of 8%. The popular ARIMA model for time series forecasting is implemented as a comparison to the deep learning models. As expected, the non-linear deep learning methods outperform the ARIMA forecast which performs poorly.

[2](mahendra2020) mainly depicted the top deep learning approaches for the price prediction. As only predicting the price does not mean that we are done, so they have prepared the Python based GUI application (i.e API) for the real time calculation of the price based on the given input parameters. The most common deep learning approaches Long Short Term Memory (LSTM) , Gated Recurrent Unit (GRU), Gated Recurrent Unit (GRU) and Support Vector Machine are used for model building and evaluation. They tried to predict the value using deep learning and want to compare it with the machine learning model (SVM) but they got negative approach at that point. So it can be inferred that deep learning models take long time to train. Prediction of bitcoin price was a complicated task as it is based on large dataset.

[3](luo2020) In this research, they used historical Bitcoin transaction data, Twitter data, and COVID-19 data. They collected data from 1st January 2020 to 31st July 2020 from online open sources. Features of the transaction data includes Open, High, Low, Close prices, and Trading Volume. They formulated the data to four input feature sets, including: (1) Historical Bitcoin exchange data; (2) Historical Bitcoin exchange data + COVID-19 data (recovery, confirmed, death); (3) Historical Bitcoin exchange data + Twitter data; (4) Historical Bitcoin exchange data + COVID-19 data (recovery, confirmed, death) + Twitter data. And implemented four machine learning models including (1) Random forest; (2) Decision tree; (3) AdaBoost; (4) Support vector machine. Finally, they concluded that (1) Twitter data improve the performance of models. (2) People consider the information within 5 days when they make decisions on investments. (3) Support vector machine does not perform well in predicting Bitcoin return or price trend. (4) COVID-19 data does not help improve the prediction.

[4](A research on) predicted Bitcoin price by using various machine learning algorithms including Linear Regression Model, Random Forest, K-nearest Neighbour Algorithm, Naïve Bayes, etc .They performed the analysis by mentioning the how they carried out the steps to build the models. Also they mentioned the various techniques to perform such predictions with the advantages and disadvantages to get proper idea and concluded that survey report will be just introducing modules of Bitcoin price prediction and machine algorithms. The Comparison table of ML algorithm model accuracy which tells that the Linear regression model will have most accuracy then the other algorithms. In that paper they concluded that the linear regression algorithm is more efficient then the other algorithms.

#############################################

Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

## Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

## Units

* Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
* Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
* Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
* Use a zero before decimal points: “0.25”, not “.25”. Use “cm3”, not “cc”. (*bullet list*)

## Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

*a**b* 

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.