International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

BITCOIN PRICE PREDICTION USING MACHINE LEARNING

Adithi Priya[1], Kunda Lakshmi Devi[2], Somaram Akanksha[3]

Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women, Telangana, India

Abstract:
lSTM(Long Short-Term Memory) is a kind of Recurrent Neural Network which is used in the field of Machine Learning. Traditional neural networks can't remember previous inputs But Recurrent Neural Networks enable us to learn from previous sequence input data. A
LSTM unit is composed of a cell, an input gate, an output gate and aforget gate.
this repository was written a Bitcoin Price Prediction project based on Google Trend keywords by using LSTM algorithm and Python 3.6 version. Here we tried to determine, " Does LSTM algorithm predict Bitcoin Close price by adding keywords volume from Google trends". Bitcoin price dataset was downloaded hourly using coinapi.io API and Google Trends keywords were downloaded hourly using python pytrend library. Finally chosen Bitcoin, BTC, Blockchain, Crypto currency LSTM algorithm predicted Bitcoin Close prices better than we expected by improving its learning in every epoch.

Keywords:LSTM, Bitcoin, Machine Learning, RNN, Google Trend Keywords.

Introduction

About ProjectBITCOIN
Bitcoin uses a peer-to-peer technology to operate with no central authority or banks. Bitcoin is open-source; its design is public, nobody owns or controls Bitcoin and everyone can take part. Digital currency bring into use as open source software in 2009 by pseudonymous creator Satoshi Nakamoto It is a cryptocurrency, so-called because it uses cryptography to control the creation and transfer of money. Users send payments by broadcasting digitally signed messages to the network. Participants known as miners verify and timestamp transactions into a shared public database called the blockchain, for which they are rewarded with transaction fees and newly minted bitcoins. Conventionally "Bitcoin" capitalized refers to the technology and network whereas "bitcoins" lowercase refers to the currency itself. Bit coins can be obtained by mining or in exchange for products, services, or other currencies.

PREDICTION

The Bitcoin's value varies just like any other stock. There are many algorithms used on stock market data for price forecast. However, the parameters affecting Bitcoin are different. Therefore it is necessary to foretelling the value of Bitcoin so that correct investment decisions can be made. The price of Bitcoin does not depend on the business events or intervening government authorities, unlike the stock market. Thus, to forecast the value we feel it is necessary to leverage machine learning technology to predict the price of Bitcoin

PYTHON

We chose to work on python programming language for the coding. Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data

structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

ANACONDA

Anaconda is a distribution of the Python and R programminglanguages for scientific computing (data science, machine
learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment.

JUPYTER

The Jupyter Notebook application allows you to create and edit documents that display the input and output of a Python or R language script. Once saved, youcan share these files with others.

Purpose

The purpose of the project to come up with a understandable, and Easy and simple visualization as well.Crypto is the world's fastest growing financial instrument. With us you can execute profitable trades with Investment Advice functionality that incorporates future prices generated by our AI: you will no longer need to spend hours analyzing price movements - get the Buy or Sell action in just one tap.

The advantage of this system is u need not be an expert for the usage of it rather this makes you an expert path. The system can be used by the professional who tend to teach on daily basis. It can be used for any future analysis routine as well. It can laterbe developed as an app as well.

Objectives of the Project

THE PROBLEM

The problem we aim to solve is the difficulty and lack of knowledge in crypto investment. We see many individuals who have interest in investment but lack the timeand the confidence to boldly invest in the future which is cryptocurrency.

People generally fear what they don’t understand. It makes sense, then, to learn all you can about the subject you are afraid of.

But now that crypto is rapidly growing this project would a vital help for so.
Cryptocurrency difficulty is a measure of how difficult it is to mine a block in
a blockchain for a particular cryptocurrency. A high cryptocurrency difficulty means ittakes additional computing power to verify transactions entered on a blockchain—a process called mining.

WHY BITCOIN

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Bitcoin is considered for the prediction. Not everyone gets to spend a solid time of the day for studying completely about investing. Project is supposed to make things easier.

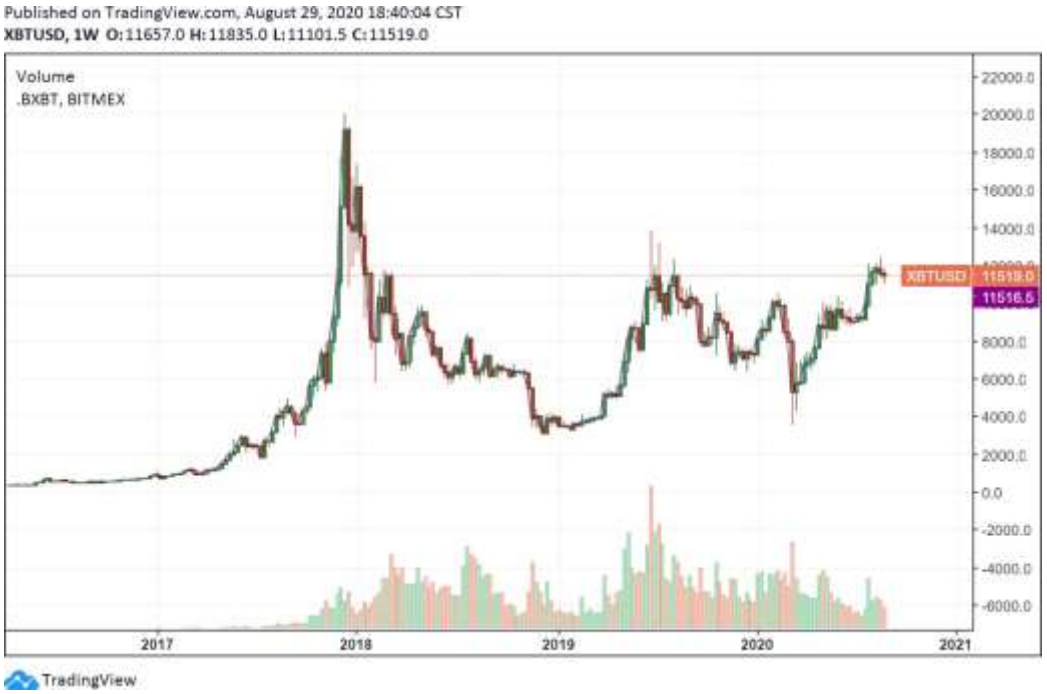


Figure 1. Bitcoin Statistics

Another reason for working on Bitcoin is Bitcoin's price appreciation is its growing adoption as a payment method. Recently, PayPal (PYPL) announced that it would soon allow its users and merchants to buy, sell, hold, and accept Bitcoin and other cryptocurrencies as a form of payment.

PREDICTION

Prediction is based on the LSTM algorithm which seem opt for the solution from the other machine Learning algorithms for its obvious advantages.

Predicting the future of sequential data like stocks using Long Short Term Memory (LSTM) networks. Forecasting is the process of predicting the future using current andprevious data. The major challenge is revising the patterns in the sequence of data and then using this pattern to analyze the future. If we were to hand-code the patterns, it would be tedious and changes for the next data. Deep Learning has proven to be betterin understanding the patterns in both structured and unstructured data.

To understand the patterns in a long sequence of data, we need networks to analyze patterns across time. Recurrent Networks is the one usually used for learning such data. They are capable of understanding long and short term dependencies or temporaldifferences.

REPRESENTATION

In Statistics, a graphical representation is a visual display of data in the form of a diagram or graph. A chart is a graphical representation of data, in which "the data is represented by symbols, such as bars in a bar chart, lines in a line chart, or slices in apie chart".

It represents the set of data in a meaningful way. It provides data where it helps to take decisions in a much better way. Some of the various types of graphical representation include – Line Graphs, Bar Graphs, Histograms, etc.

Graphical representation is selected for the visualization for obvious reasons which areGraphic visual representation of information is a crucial component in understanding and identifying patterns and trends in the ever increasing flow of data.

Graphical representation enables the quick analysis of large amounts of data at one time and can aid in making predictions and informed decisions. Data visualizations also make collaboration significantly more efficient by using familiar visual metaphors to illustrate relationships and highlight meaning, eliminating complex, and long-winded explanations of an otherwise chaotic-looking array of figures.

1.3 Scope of the ProjectTHE IDEA
The idea is to be able to reach out a common person to help him/her later on with herstudies on crypto investment. In which the concept of easy to understand the visualization is the main agenda.

PREDICTION

A common misconception is that predictive analytics and machine learning are the same thing, but it reality they only overlap.

At its core, predictive analytics encompasses a variety of statistical techniques including machine learning, predictive modeling and data mining and uses statistics to estimate, or 'predict' future outcomes. These outcomes might be behaviors a customeris likely to exhibit or possible changes in the market, for example.

Predictive analytics help us to understand possible future occurrences by analyzing thepast.

REASONS FOR SELECTING BITCOIN

Bitcoin broke a key resistance level making investors believe there is further upside.

Rising inflation and the potential for even more stimulus continues to push people tosafe-haven assets.

Increased adoption from payment applications like PayPal will give far more people easy access to cryptocurrency.
Publicly traded companies purchasing Bitcoin shows a high level of confidence in itsappreciation.

Bitcoin's historical trend of closely following its halving stock-to-flow model shows anambitious and extremely bullish outlook.

This makes it evident about the future scope of bitcoin which made us choose it.

Literature Survey

Existing System

they pointed to understand and identify daily changes in the Bitcoin market while obtaining insight into most appropriate features surrounding Bitcoin price. They willpredict the daily price change with highest possible accuracy[1]

The objective of this paper is to determine the predictable price direction of Bitcoin inUSD by machine learning techniques and sentiment analysis.We have applied sentiment analysis and supervised machine learning principles to the extracted tweets from Twitter and Reddit posts, and we analyze the correlation between bitcoin price movements and sentiments in tweets. [2]

This paper proposes three types of recurrent neural network (RNN) algorithms used to predict the prices of three types of cryptocurrencies, namely Bitcoin (BTC), Litecoin (LTC), and Ethereum (ETH). The models show excellent predictions depending on the mean absolute percentage error (MAPE).[3]

This project used LASSO finding of the results from a larger database is quick and fast and aim to understand and find daily trends in the Bitcoin market while gaining insight into optimal features surrounding Bitcoin price. [4]

In this paper, we attempt to predict the Bitcoin price accurately taking into consideration various parameters that affect the Bitcoin value. For the first phase of our survey, we aim to understand and identify daily trends in the Bitcoin market while gaining insight into optimal features surrounding Bitcoin price. For the second phase of our survey, using the available information, we will predict the sign of the daily price change with highest possible accuracy.[5]

Proposed SystemCRYPTO

A cryptocurrency is an encrypted data string that denoted a unit of currency. It is monitored and organized by a peer-to-peer network called a blockchain, which also serves as a secure ledger of transactions which accounts buying, selling, and transferring. Unlike physical money, crypto currencies are decentralized, which means they are not issued by governments or other financial institutions.

Bitcoin, Ether, Litecoin, and Monero are popular cryptocurrencies.

BITCOIN

Bitcoin was the first digital asset to beget the current ecosystem of crypto. For a while, it grew an underground following of investors who saw its future as a possible replacement to the physical monetary system. Now Bitcoin has become a household name as institutions and governments develop ways to serve their customers growing demand for exposure.

In 2021, El Salvador became the first country n 2021, El Salvador became the 1st country in the world to make Bitcoin a legal tender; Paraguay and other countries lookto follow suit. El Salvador is also the first and only country to have Bitcoin in its treasury. As of early 2022, El Salvador has 1,800 bitcoin. President Nayib Bukele has not been shy about announcing his purchases on Twitter.

Figure 2. QR Code of Bitcoin

As the traditional finance world realizes Bitcoin's potential for disruption, they must choose either to adopt cryptocurrencies or face irrelevance. The personal decision to invest in bitcoin comes down to your appetite for risk and your perspective on the future of humanity. For example, Russia has stated they are looking into cryptocurrencies to lower their dependence on the US dollar. Bitcoin has the potential to disrupt the US Dollar in a massive way, and it is simply too big to be ignored at thispoint.

If the question occurs why bitcoin is preferred than other variety of other crypto, It has obvious advantages. Such as

Transactions with Bitcoin don't include banking fees

Low transaction fees for Bitcoin international payments

These transactions are mobile and secure

Bitcoins transactions are also P2P and pseudonymous

In addition to all of the above mentioned, Bitcoin transactions are also pseudonymous. It means that they're not completely anonymous and that transactions are able to be identified only by using a blockchain address. Therefore, an individual is able to have multiple addresses and multiple passwords and usernames for just one account.

It's also important to know that the payment system of Bitcoin is purely peer-to-peer. It means that users are able to send and also receive payments worldwide from or to anyone on the network. The parties to a transaction are not required approval from authority or an external source unless they are receiving or sending bitcoin from the institution or regulated exchange. Forecasting is the process of predicting the future using current and previous data.

The major challenge is understanding the patterns in the sequence of data and then using this pattern to analyze the future. If we were to hand-core the patterns, it would be tedious and changes for the next data. Deep Learning has proven to be better in understanding the patterns in both structured and unstructured data. To understand thepatterns in a long sequence of data, we need networks to analyze patterns across time. Recurrent Networks is the one usually used for lrarning such data. They are capable ofunderstanding long ans short terms dependencies or temporal differences

The proposed system implements machine algorithm to predict the price of the bit-coin based on dataset. The bitcoin price prediction can be done using the LSTM. The tools used for project are anaconda-navigator. The procedure is as follows:

Collection of data set through Google keywords

Arrange the data into the data frame.

Extra columns are removed and stored in a CSV file.

data-preprocessing is performed.

Build the model for the data-set.

Test the predictions.

We preferred working with python 3.6 because Python 3.6 picks up where many of those improvements left off and nudges them into new realms.

Python 3.5 added syntax used by static type checking tools to ensure software quality.

Python 3.6 expands on that idea, which could eventually lead to high-speed statically compiled Python programs.

Proposed Architecture

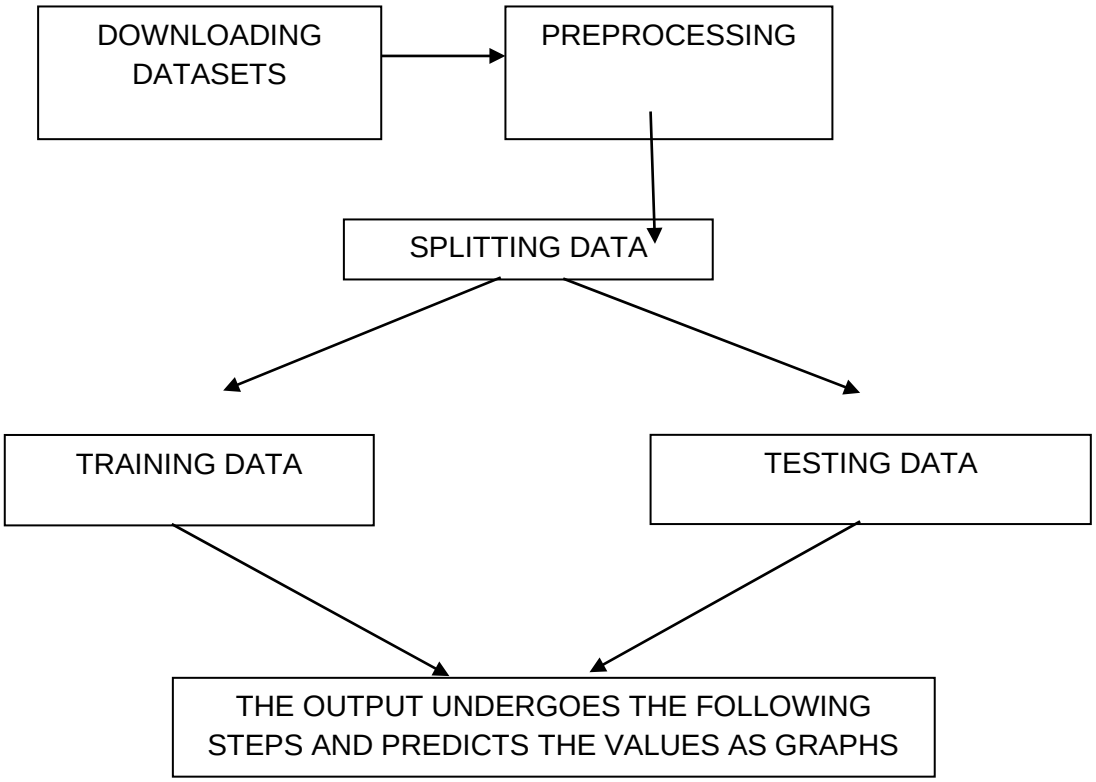
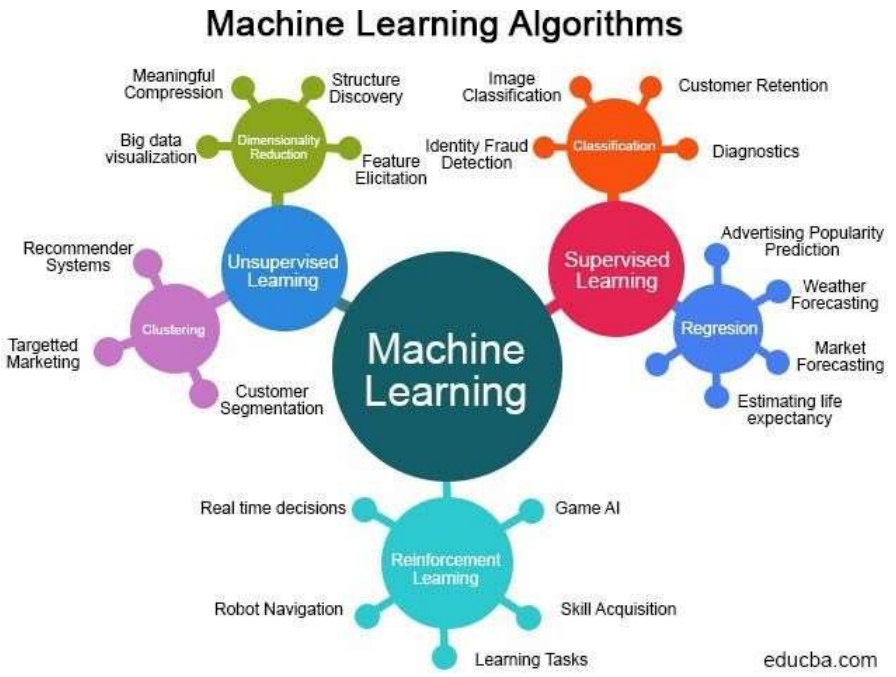


Figure 3: Architecture

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

System Methodology

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.



Classical **Machine Learning** is often categorized by how an algorithm learns tobecome more accurate in its predictions. The type of algorithm data scientists choose to use depends on what type of data they want to predict.

TYPES

Classical machine learning is often categorized by how an algorithm learns to become more accurate in its predictions. The type of algorithm data scientists choose to use depends on what type of data they want to predict.

Supervised learning: In this type of machine learning, data scientists supply algorithms with labeled training data and define the variables they want the algorithm to assess forcorrelations. Both the input and the output of the algorithm is specified

Unsupervised learning: This type of machine learning involves algorithms that train on unlabeled data. The algorithm scans through data sets looking for any meaningful connection. The data that algorithms train on as well as the predictions or recommendations they output are predetermined.

Semi-supervised learning: This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labeled training data, but the model is free to explore the data on its own and develop its ownunderstanding of the data set..

RNN

Based on our requirement we are choosing the supervised learning algorithm on Recurrent Neural Network

Recurrent neural networks (RNN) are the state of the art algorithm for sequential data and are used by Apple's Siri and Google's voice search. It is the first algorithm that remembers its input, due to an internal memory, which makes it perfectly suited for machine learning problems that involve sequential data. It is one of the algorithms behind the scenes of the amazing achievements seen in deep learning over the past few years. In this post, we'll cover the basic concepts of how recurrent neural networks work, what the biggest issues are and how to solve them.

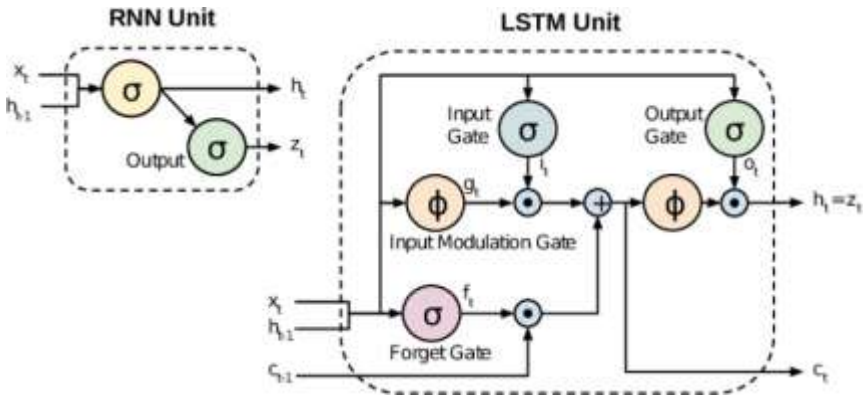
LSTM

LSTM is considered the best for our prediction because LSTM has advantage over RNN as LSTM can handle the information in memory for the long period of time as compare to RNN. But the question is what is different in LSTM than RNN by which LSTMs are capable of maintaining long term temporal dependencies (remembering information for long period of time).

A set of gates is used to control information within memory in general, such as when it enters the memory, how long and how much information may be kept, when it begins to provide output, and when it begins to decay or be forgotten.

Implementation

Algorithm



APPLICATIONS OF LSTM

LSTM networks find useful applications in the following areas:

Language modeling

Machine translation

Handwriting recognition

Image captioning

Image generation using attention models

Question answering

Video-to-text conversion

Polymorphic music modeling

Speech synthesis

Protein secondary structure prediction

This list does give an idea about the areas in which LSTM is employed but not how exactly it is used. Let's understand the types of sequence learning problems that LSTM networks are capable of addressing.

LSTM neural networks are capable of solving numerous tasks that are not solvable by previous learning algorithms like RNNs. Long-term temporal dependencies can be captured effectively by LSTM, without suffering much optimization hurdles. This is used to address the high- end problems.

Code Implementation

Step 1: import pandas as pd , numpy as np, matplotlib.pyplot as plt

Step 2: Bitcoin and Google datasets are loaded

Step3: Preprocessing: Loaded raw datasets are printed, Unused column are droppedand preprocessed

Step4: final results are printed for two datasets, Finally two datasets are concatenated, Final dataset exported

Step5: Load dataset by using Pandas library

Step6: Here was prepared column for visualizing, plot each column, convert series to supervised learning

Here is created input columns which are (t-n, ... t-1)

Here is created output/forecast column which are (t, t+1, ... t+n)

put it all together

drop rows with NaN values

Step 7: here checked values numeric format , Dataset values are normalized by using MinMax method, print(scaled)

Step8: Normalized values are converted for supervised learning, reframed.drop(reframed.columns[[9,10,11,12,13,14,15]], axis=1, inplace=True)

Step9: Dataset is splitted into two groups which are train and test sets

Splitted datasets are splitted to trainX, trainY, testX and testY

Train and Test datasets are reshaped in 3D size to be used in LSTM

LSTM model is created and adjusted neuron structure

Dataset is trained by using trainX and trainY

Loss values are calculated for every training epoch and are visualized Step10: Prediction process is performed for train dataset
Prediction process is performed for test dataset

Trains dataset inverts scaling for training

Test dataset inverts scaling for forecasting

invert scaling for actual

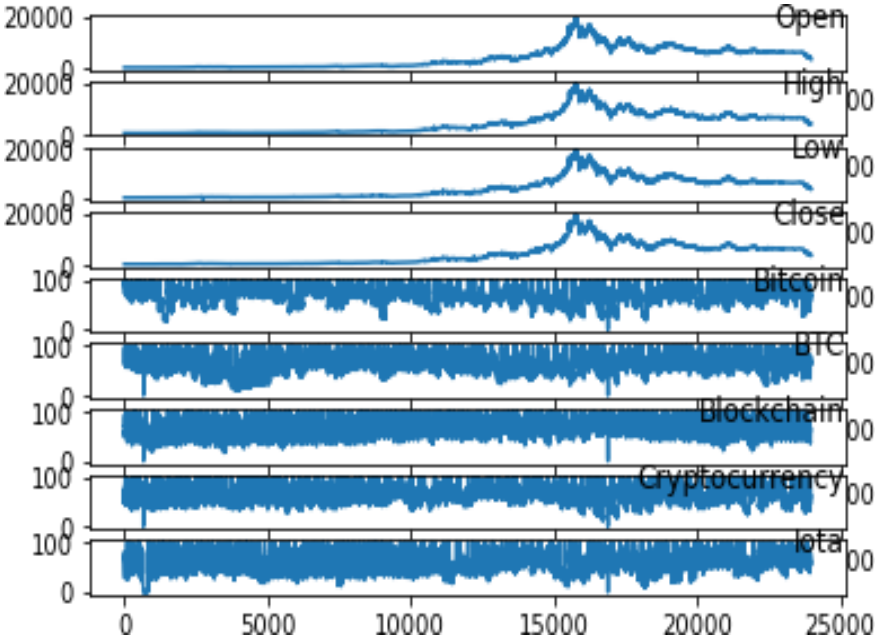
print('actual: ', len(inv_y))

Performance measure calculated by using mean_squared_error for train and test prediction, print(testPredict), print(type(trainPredict))

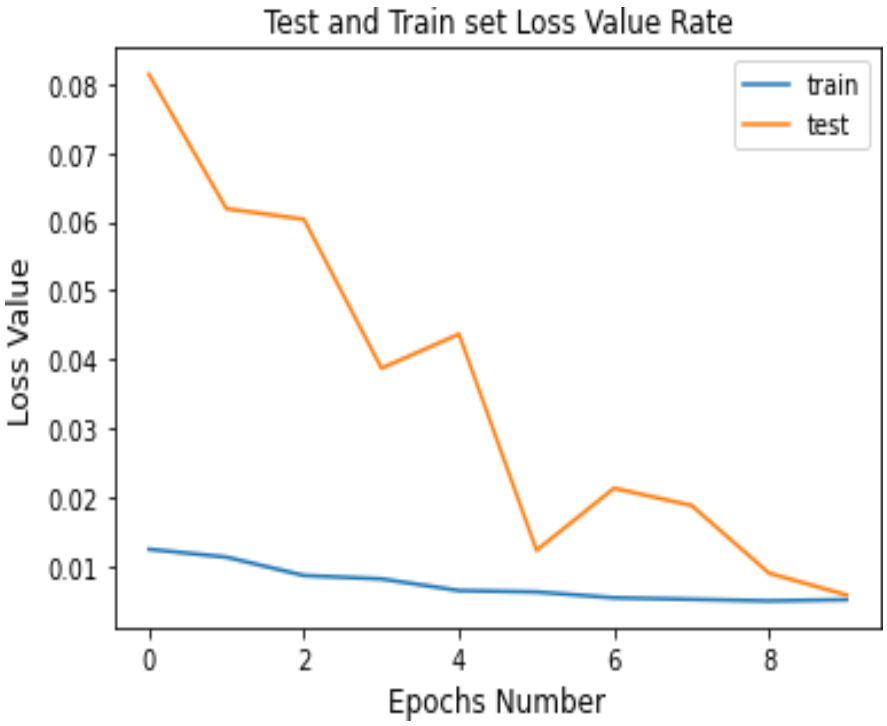
train and test prediction are concatenated, print(len(son)) Step 11: Finally training and prediction result are visualized

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

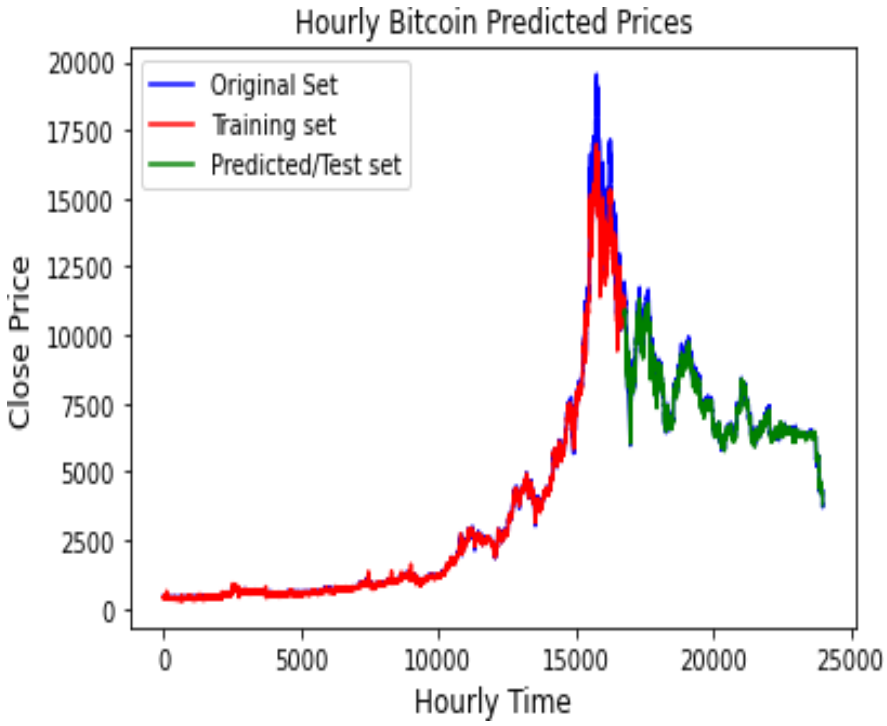
Result



Plotting each column of the dataset.



Test and Train sets loss value rates



Bitcoin predicted prices

The above is giving out clear explanation that the red line is the training set of data, theblue being the original set of data and the green being the Predicted/ Test Set. We understand the original is almost very close to the prediction of our system. Thisclearly conveys the consistency and approximation of the system.

Conclusion

PREDICTION

In conclusion with data like cryptocurrency such a bitcoin, Machine Learning Algorithm Long Short Term Memory is used to predict the data. There are obvious

reasons for selecting the algorithm LSTM for the project and the results are crystal clear as expected in accordance to visualization and information. The reason mentioned above of giving priority to LSTM algorithm rather than RNN had obvious advantage over the solution. The reason for selecting bitcoin prediction is the fact that it is growing rapidly than internet. For such a growing figure to estimate and predict values isn't a task that can be barred by a common person. The task being able to provide such interface is successfully done through the project.

BITCOIN

A collective insanity has sprouted around bitcoin over the last decade, it's hard to predict whether this cryptocurrency will become the global reserve currency or a store value as widely accepted as gold. The thrill of riches or ruin leaves some investors wary, but other want ti chase the chance for massive profits from investing in bitcoin. Bitcoin is certainly a revolutionary technology, and it's much less risky in 2022 than it was in 2012.

After becoming legal tender in EL Salvador in 2021, other countries will look to copy this move to attract innovation, while others may ban it entirely in attempt to save their fiat currency. Bitcoin has taken center stage in global geopolitical climate, and 2022 looks like it may be the year of massive adoption.

PREDICTION

Bitcoin Price Prediction using machine learning helps you discover the future value of coin and other financial assets traded on an exchange. The entire idea of predicting stock prices is to gain significant profits. Predicting how the bitcoin price will perform is a hard task to do. There are other factors involved in the prediction, such as physical and psychological factors, rational and irrational behavior, and so on. All these factors combine to make coin prices dynamic and volatile. This makes it very difficult to predict bitcoin prices with high accuracy.

The LSTM rectifies a huge issue that recurrent neural networks suffer from: short- memory. Using a series of 'gates,' each with its own RNN, the LSTM manages to keep, forget or ignore data points based on a probabilistic model.

LSTM

LSTMs also help solve exploding and vanishing gradient problems. In simple terms, these problems are a result of repeated weight adjustments as a neural network trains. With repeated epochs, gradients become larger or smaller, and with each adjustment, it

becomes easier for the network's gradients to compound in either direction. This compounding either makes the gradients way too large or way too small. Whileexploding and vanishing gradients are huge downsides of using traditional RNN's, LSTM architecture severely mitigates these issues.

After a prediction is made, it is fed back into the model to predict the next value in the sequence. With each prediction, some error is introduced into the model. To avoid exploding gradients, values are 'squashed' via (typically) sigmoid & tanh activation functions prior to gate entrance & output.

Future Scope

This is expected to later deal with other cryptocurrency as well and consider the crypto world in another level for analysis. The fact that the prediction by the system is so close gives hope and a standard sophistication to experiment in several other aspects. The successful prediction of a bitcoin's future price could yield significant profit. The efficient-market hypothesis suggests that coin prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable.

References

Lekkala Sreekanth Reddy, Dr.P. Sriramya A Research On Bitcoin Price Prediction Using Machine Learning Algorithms , 2020 [4]

Mohammad J. Hamayel and Amani Yousef Owda *A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms ,2020 [3]

Neha Mangla1 ,Akshay Bhat2 , Ganesh Avabratha3, Narayana Bhat4 Bitcoin Price Prediction Using Machine Learning , 2019 [1]

Shivam Pandey1, Mr.Anil Chavan2, Miss. Dhanashree Paraskar3, Prof. Sareen Deore Bitcoin Price Prediction using Machine Learning, 2021[5]

S M Rajua,∗ , Ali Mohammad Tarifb Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis , 2020[2]

SN Paidimarry, CKK Reddy, LG Lolla, Internet of things enabled smart baggage follower with theft prevention, Journal of Physics: Conference Series 1950 (1), 012002, 2021

PR Anisha, CKK Reddy, K Apoorva, CM Mangipudi, Early Diagnosis of Breast Cancer Prediction using Random Forest Classifier, IOP Conference Series: Materials Science and Engineering 1116 (1), 012187, 2021

RM Mohana, CKK Reddy, PR Anisha, BVR Murthy, Random forest algorithms for the classification of tree-based ensemble, Materials Today: Proceedings, 2021

Kishor Kumar Reddy C, Anisha P R, Shastry R, Ramana Murthy B V, "Comparative Study on Internet of Things: Enablers and Constraints", Advances in Intelligent Systems and Computing, 2021

Kishor Kumar Reddy C, Anisha P R, Apoorva K, "Early Prediction of Pneumonia using Convolutional Neural Network and X-Ray Images", Smart Innovation, Systems and Technologies, 2021

R Madana Mohana, Kishor Kumar Reddy C and Anisha P R, "A Study and Early Identification of Leaf Diseases in Plants using Convolutional Neural Network", Springer 4th Int Conference on Smart Computing and Informatics, 2020, India

Anisha P R, C Kishor Kumar Reddy and Nuzhat Yasmeen, "Predicting the Energy Output of Wind Turbine Based on Weather Condition", Springer 4th Int Conference on Smart Computing and Informatics, 2020, India

Kishor Kumar Reddy C, Apoorva K, Anisha P R, "Early Prediction of Pneumonia using Convolutional Neural Network and X-Ray Images", Springer 4th Int Conference on Smart Computing and Informatics, 2020, India

Viswanatha Reddy, Dr. Elango NM and Dr. C Kishor Kumar Reddy and Anisha P R, "Prediction of Diabetes using Internet of Things (IOT) and Decision Trees: SLDPS, Springer FICTA, January 2020, India.

P R Anisha, Dr. C Kishor Kumar Reddy, "Early Detection of Diabetes using Machine Learning Algorithms and Internet of Things: ADPA", Springer INDIA, 2019, India.

Dr B V Ramana Murthy , Kishor Kumar Reddy C, Anisha P R, Rajasekhar Shastry, "IOT Based Smart Stale Food Detector", Springer INDIA, 2019, India.

Kishor Kumar Reddy C, Anisha P R, Rajasekhar Shastry, Dr B V Ramana Murthy and Dr Vuppu Padmakar "Automated rainwater Harvesting System", IEEE ICCES, 2019.

Kishor Kumar Reddy C, Anisha P R, Rajasekhar Shastry, Dr B V Ramana "Comparative Study on Internet of Things: Enablers and Constraints", Springer-ICDECT, 2019.
Rajasekhar Shastry, Dr B V Ramana Murthy, Kishor Kumar Reddy C, Anisha P R " Automated Lighting Smart Parking Using Internet of Things", Springer-ICICCS, 2019
R Gangadhara Reddy, M Ramesh Babu, CH jaya Prakash Rao, Kishor Kumar Reddy C , PPSIC: Pulsated Power Supply Inverter Circuit", IEEE CICN, 2017