A MODEL FOR PREDICTING ETHEREUM FUTURE PRICE RATES USING MACHINE LEARNING

BY

Shiam Bin Azad
ID: 201-15-13758

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

Md Assaduzzaman
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Research based project titled **"A model for predicting Ethereum future price rates using machine learning"**, submitted by Shiam Bin Azad, ID No: 201-15-13758 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *24.01.2024*.

BOARD OF EXAMINERS

Chairman

Narayan Ranjan Chakraborty (NRC)
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Md. Sazzadur Ahamed (SZ)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Amatul Bushra Akhi (ABA)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Md. Zulfiker Mahmud (ZM)
Associate Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

I hereby declare that, this research-based project has been done by us under the supervision of **Shah Md Tanvir Siddiquee, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

20.01.24

Md Assaduzzaman
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Submitted by:

Shiam Bin Azad

Shiam Bin Azad
ID: 201-15-13758
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

I really grateful and wish our profound our indebtedness to **Shah Md Tanvir Siddiquee**, **Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to Dr. Sheak Rashed Haider Noori, Professor & Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of our parents.

ABSTRUCT

Considered by many to be the forerunner of smart contract technology, Ethereum has become a widely used blockchain platform. Beyond being a cryptocurrency, Ethereum's main purpose is to offer a decentralized platform for smart contract execution. There are several benefits to using machine learning to predict Ethereum prices in the ever-changing cryptocurrency markets. Machine learning algorithms are able to identify complex patterns and trends that may be difficult for humans to analyze by examining large amounts of historical data. A thorough analysis of Ethereum's price fluctuations is made possible by this data-driven strategy, which considers market sentiment, technical indications, and outside events. In this study, I examine historical Ethereum data that I gathered from the Yahoo Stock Market between 2017 to the present. I used the regression model and the neural network model as my two sorts of algorithms. I utilize the Huber Regressor, Least Angle Regression, Linear Regression, Orthogonal Matching Pursuit, and Lasso Least Angle Regression for my regression models, and I use LSTM for my neural network model. According to my research, the Huber Regressor performs best when taking R2 score into account, having the greatest R2 score of 0.9906. LSTM, however, handles errors less frequently. 35.9866 is the Mean Absolute Error (MAE), which is less than that of any other model. High accuracy is correlated with low error. The Accuracy of LSTM is 98.3582% which is higher when taking MAE into account. According to this research, the best model for high-volume timestamp data is an LSTM.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Blockchain is a powerful technology that can access any banking or financial institution. Because of its diversity, it has drawn attention from fields including computer science, encryption, and the expansion of communications. Blockchain technology can offer a platform for money exchange without the need for a middleman like the government or other financial institutions [6]. Blockchain includes cryptocurrency as an integral component. A cryptocurrency is a kind of virtual or digital money that may be used to transfer and exchange digital assets. Most scholars and economists tend to think about cryptocurrencies as speculative financial assets meant mainly for short-term investment horizons [8]. This is why it would be essential to create suitable forecasting tools in order to make decisions in the cryptocurrency market. Bitcoin is the first and most well-known cryptocurrency example. Using cryptocurrencies, secure, anonymous online transactions are possible [5]. The Ethereum is another cryptocurrency that is becoming more and more popular every day. The two most representative cryptocurrencies, Bitcoin and Ethereum, have the highest market capitalizations and trade volumes on record. According to CoinMarketCap (2021) as of April 2021, Ethereum has attained a trading volume of USD 38 billion and a market capitalization of USD 265 billion, while Bitcoin has recorded a trading volume of USD 64 billion and a market value of USD 1,304 billion [7]. The Ethereum coin is used by the decentralized blockchain network Ethereum. It is a distributed software platform that is open-source and makes use of blockchain technology. Vitalik Buterin proposed it towards the end of 2013, and on July 30, 2015, it became operational. A number of changes have been made to the Ethereum network to improve its usefulness, security, and scalability. Byzantium, Constantinople, Istanbul, and Ethereum 2.0—also referred to as ETH 2.0—are a few of the significant improvements. The latter seeks to increase scalability and sustainability by switching from a proof-of-work (PoW) to a proof-of-stake (PoS) consensus mechanism. Ethereum is still a major participant in the blockchain and cryptocurrency industry, providing the framework for decentralized apps and the larger ecosystem of decentralized finance (DeFi). Anyone can use Ethereum to develop any kind of secure digital technology. If approved, users can use the token to pay for actual goods and services in addition to the work done to support the blockchain. Ethereum is intended to be decentralized, secure, programmable, and scalable. For developers and businesses building technology on top of it to transform several industries and our way of life, this is the preferred blockchain. Smart contracts, a key component of decentralized apps, are natively supported. Smart contracts and blockchain technology are used in many decentralized finance (DeFi) and other applications. There is a great deal of uncertainty when predicting the future price of Ethereum (ETH) or any other cryptocurrency because there are many variables that can affect pricing, such as macroeconomic trends, market sentiment, legislative changes, technological improvements, and more. While machine learning (ML) models are useful for examining historical data and spotting trends, it's crucial to remember that cryptocurrency markets are quite unpredictable and that past success does not always guarantee future outcomes. Within the Ethereum ecosystem, Ether (ETH), the native coin of Ethereum, is utilized for a number of applications. Ether is a sort of currency used for transactions and value exchange within the Ethereum network. It is also used to pay miners, or participants that do computations and validate transactions. The Ethereum blockchain is based on a dispersed, decentralized network of computers called nodes that cooperate to keep the network safe and secure. Ethereum has had a number of updates over the years to address issues with scalability and security. Ethereum 2.0 is used to switch from a proof-of-work (PoW) to a proof-of-stake (PoS) consensus mechanism. Ethereum, which offers a framework for the development of decentralized applications, decentralized finance (DeFi) initiatives, non-fungible tokens (NFTs), and more, has been essential in the growth of the blockchain and cryptocurrency field. Because of its adaptability and programmability, it has become a fundamental component of blockchain research and development. Among cryptocurrencies, Ethereum distinguishes out for being the first to provide decentralized apps (DApps) and smart contracts. Smart contracts are a feature of Ethereum that enables developers to design sophisticated and programmable applications, in contrast to traditional cryptocurrencies like Bitcoin, which concentrate on peer-to-peer transactions. Its programming language, which is Turing-complete, offers versatility and allows for the execution of numerous computations. The native cryptocurrency of Ethereum, called ether (ETH), is used as fuel for smart contract execution in addition to acting as a medium of exchange inside the network. The platform's drive to continuous improvement is demonstrated by its frequent upgrades, such as Ethereum 2.0's switch to proof-of-stake. With its emergence as a center for decentralized finance (DeFi) initiatives, Ethereum has demonstrated how important it is to blockchain innovation and the development of a wide range of applications that go beyond basic financial transactions. To begin using Ethereum (ETH) as money, people must first get a safe Ethereum wallet in which to save your virtual assets. Get Ethereum at a cryptocurrency exchange, then move it to your wallet to increase security. People can send and receive transactions by using the recipient's Ethereum address once you have ETH in your wallet. Remember how crucial it is to verify the gas fees that are being paid to miners at this time. Ethereum is supported for transactions by certain retailers, websites, and decentralized finance (DeFi) services, despite its lower level of acceptance than traditional currencies. Remain up to date with Ethereum ecosystem advancements and investigate new use cases that go beyond conventional money transactions, such DeFi platforms that provide chances for lending, borrowing, and earning. Ethereum's function as a flexible digital asset keeps growing as the cryptocurrency market develops, giving users access to a variety of options beyond straightforward transactions.

1.2 Problem Statement

There are risks and difficulties associated with investing in cryptocurrency. The first issue is the high degree of volatility present in cryptocurrency markets, where price fluctuations can occur suddenly and without warning and result in significant gains or losses. Another issue is regulatory ambiguity because there are regional variations in the legal status and adoption of cryptocurrencies. Due to exchanges' and wallets' vulnerability to fraud and hacking, security is a major concern, underscoring the significance of safe storage procedures. Investors in the cryptocurrency area should proceed with care and do extensive study because there are currently no comprehensive rules or safeguards for investors. In addition to limited adoption in the real world, market emotions and speculation can cause inflated price movements and affect the overall value proposition of cryptocurrencies. The intricacy and dynamic character of the cryptocurrency ecosystem are further highlighted by technological hazards, liquidity issues, and the possibility of market manipulation. Given the dynamic and inventive nature of the cryptocurrency market, it is imperative for investors to exercise caution and develop safe investment methods, considering their level of risk tolerance.

1.3 Research Objectives

- To apply a machine learning algorithm on the Ethereum historical data to find out the future price of Ethereum cryptocurrency.
- To find out which model is best fit for the Ethereum historical dataset.
- To investigate which model is not suitable for timestamp dataset.

1.4 Research Questions

- How can I Apply algorithms on the timestamp dataset?
- How can I find out the flows between other models?

CHAPTER 2
LITERATURE REVIEW

2.1 Related works

The authors of this study combine classic logit econometric models with tree-based machine learning classifiers. The research yields a number of significant conclusions like compared to logit models, random forests are more accurate at predicting the future prices of gold and bitcoin. The accuracy of the prediction ranges from 75 to 80% for a 5-day forecast. The acuity increases along with the day. The accuracy of bagging and random forests is more than 85% for predictions made for 10 and 20 days. The volatility of oil prices is also significant for projecting the prices of Bitcoin and gold, suggesting that Bitcoin can be used as a gold alternative to diversify this kind of volatility [1].

The authors of this study cross-validate a few writers' hypotheses regarding how social media affects bitcoin values. To forecast prices, sentiment scores from news feeds and tweets are considered in addition to past prices and volume. Authors used RNN, LSTM, random forest for the model. Authors finds random forest's MAE score 2.7526 and RMSE 13.7033 for first day. The results of the experiment indicate that sentiment scores have no effect unless they are biased towards a specific class [2].

The authors of this study have conducted research on the forecast of bitcoin prices using both statistical methods and machine learning (ML). ML approaches generally perform better in comparison studies. This review is a thorough investigation into how combining earlier research can improve our ability to forecast bitcoin prices. The distinction from previous review research is made clear by the group presentations of studies on Bitcoin price prediction. These include causality, ML and statistical approaches, ML-ML, frequency effect of chosen time, impact of social media and online search engine, and hyperparameter optimization techniques. This study will employ seven distinct models, including MLP, SVM, GRNN, RNN, LSTM, GRU, and CNN, on six distinct datasets. ARIMA model also perform well in this situation [3].

The goal of this study is to make highly accurate predictions about the changes of Bitcoin values. Four distinct machine learning (ML) algorithms are used to achieve this goal: support vector machines (SVM), artificial neural networks (ANNs), in addition to logistic regression (LR), the Naïve Bayes (NB), the Random Forest (RF), and the (ANN) as a standard model. The NB has the lowest forecasting performance in the continuous dataset, according to the authors, while the RF has the best. ANN Accuracy 0.843, SVM accuracy is 0.808, NB Accuracy is 0.717, LR 0.781 and RF accuracy is 0.884 [4].

An extensive overview of earlier research in the area of bitcoin price prediction from 2010 to 2020 is given in this article. This article's discussion will assist scholars in bridging the gap in upcoming research and acquire additional understanding. This paper's primary contribution is a current analysis that applies models utilizing conventional statistical and machine learning techniques to examine and compile the publications published in the field of bitcoin price prediction. Data was taken from publications that were released between 2010 and 2020. From the analysis, It is advised that researchers examine how LSTM models—such as CNN LSTMs, bidirectional LSTMs, and encoder-decoder LSTMs—are used in future research and compare the outcomes to gain useful future understanding and enhance price prediction outcomes. Because of the fact that the intraday behavior of the intraday variables varies over time, researchers should use caution while studying any element of cryptocurrencies [5].

Three distinct models—Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTMs), and Bi-directional Long Short-Term Memory (Bi-LSTMs)—are compared in this research. The dataset, which is used to forecast both short-term (30 days) and long-term (90 days) Ethereum prices, is made up of the closing price for the last 2000 days. These prices, which are updated daily, are retrieved from an API in JSON format. The suggested model evaluates several price prediction models and concludes that, when compared to RNN, LSTM, and Bi-LSTM, bidirectional LSTM is the most effective model for predicting Ethereum prices [6].

This study looks into the connection between Ethereum pricing and information found on the Ethereum Blockchain. Moreover, we look into the relationship between Ethereum pricing and Blockchain data about other coins that are available for purchase. Our main conclusions show that the macroeconomic variables, Ethereum-specific Blockchain information, and other cryptocurrency's Blockchain information all have a significant impact on predicting Ethereum pricing. Researchers discover that macroeconomic considerations greatly enhance the accuracy of Ethereum price prediction. Additionally, they discover that there is a strong correlation between Ethereum prices and the variables in Ethereum-specific Blockchain information, such as the uncle block, gas price, gas usage, and gas limit [7].

In this research, supervised machine learning (ML) technique to short-term forecasting of bitcoin time series is discussed. We used the most potent ensemble techniques, Random Forests (RF) and Stochastic Gradient Boosting Machine (SGBM), to achieve this aim. The three most valuable coins in the dataset were Bitcoin (BTC), Ethereum (ETH), and Ripple (XRP). The dataset was gathered from the daily close prices of these coins, and we employed historical price data and technical indicators (moving average) as features. We used the one step ahead technique to create an out-of-sample forecast for a subset of time series in order to assess the efficacy of these models. The anticipated prices' accuracy rate using RF and GBM was computed. The outcomes confirm that the ML ensembles approach may be applied to forecast bitcoin prices. The three most capitalized cryptocurrencies (BTC, ETH, and XRP) have out-of-sample accuracy of short-term prediction daily close prices derived by the SGBM and RF in terms of Mean Absolute Percentage Error (MAPE) within 0.92-2.61 % [8].

First, the researchers group the price of Bitcoin according to its daily and high-frequency values. For predicting the daily price of Bitcoin using high-dimensional characteristics, statistical techniques like logistic regression and linear discriminant analysis beat more complex machine learning algorithms, achieving an accuracy of 66%. They achieve greater performance when compared to benchmark results for daily price prediction, with the maximum accuracies of machine learning algorithms and statistical approaches being 65.3% and 66%, respectively. With an accuracy of 67.2%, machine learning models such as Random Forest, XGBoost, Quadratic Discriminant Analysis, Support Vector Machine, and Long Short-term Memory outperform statistical methods for predicting the price of bitcoin during a 5-minute interval. They might view our analysis of Bitcoin price prediction as a first examination into the significance of sample dimension in machine learning methods [9].

This research study teaches how to create a model prediction for the bitcoin stock market using Long Short-Term Memory (LSTM). Similar to RNN, LSTM is made up of modules with recurrent consistency and is another type of module provided for RNN that was later developed and popularized by numerous researchers. The approach we used for this study, along with the methods and resources used by Yahoo Finance to forecast Bitcoin prices on the stock market, can forecast a price above $12,500 USD for the days that follow the projection. In the final section, we draw conclusions and talk about our plans for future research. The outcome isn't good enough in terms of RMSE; it may be in the hundreds or International Conference on Electrical Engineering and Computer Science (ICECOS) 2019 209 almost 50 Score RMSE, which is the drawback [10].

2.2 Scope of the Problem

The unique features of the cryptocurrency market present a number of difficulties for machine learning models attempting to predict Ethereum's price. Like other digital assets, Ethereum's price fluctuates rapidly and erratically. This makes it extremely volatile. Because cryptocurrencies lack traditional fundamentals, unlike traditional financial markets, it is difficult for machine learning models to rely solely on fundamental analysis. Since cryptocurrencies are relatively new compared to traditional assets, the lack of comprehensive historical data makes it much more difficult to train reliable models. Another level of complexity is added by market sentiment, which is impacted by news, legislative changes, and social media trends. As a result, unstructured data must be included into machine learning models. Prediction models may also be confused by the bitcoin market's vulnerability to manipulation and the changing regulatory environment. Cryptocurrency values are significantly shaped by global factors, including geopolitical events and economic situations, therefore machine learning models must take a wide range of external variables into account. Challenges also include the non-linear nature of price swings and the dynamic adaptation needed for quickly shifting market conditions. In the end, accurate forecasting in the cryptocurrency arena necessitates a sophisticated comprehension of market dynamics, cautious assessment of the constraints of the data at hand, and the capacity of machine learning models to operate in an environment where conventional financial rules frequently do not apply.

2.3 Challenges

There are some challenges faced at the time of predicting timestamp data like Ethereum historical data. Those challenges are the following:

High Volatility: Ethereum and other cryptocurrency markets are notorious for having extremely volatile prices, which are marked by abrupt and erratic changes in value.

Lack of Traditional Fundamentals: It is difficult for ML models to rely on these conventional indications since, in contrast to traditional financial markets, cryptocurrencies like Ethereum lack traditional fundamental data (such as earnings reports).

Limited Historical Data: There is a dearth of comprehensive historical data due to the comparatively recent existence of cryptocurrencies, such as Ethereum. To find trustworthy patterns and trends, machine learning algorithms frequently need a large amount of historical data.

Impact of News and mood: Social media trends, news, and market mood all have a big impact on bitcoin pricing. The prediction problem becomes more challenging when these unstructured data sources are included in machine learning models.

Market manipulation: Pump and dump operations and acts by major holders (whales) are two examples of how cryptocurrency markets can be manipulated. ML models have difficulty distinguishing between real market movements and manipulation.

Regulatory Uncertainty: Prices may be impacted by the constantly changing regulatory environment pertaining to cryptocurrencies. For machine learning models, predicting the impact of regulatory developments on the market is a major difficulty.

Global variables: Events, economic situations, and geopolitical variables all have an impact on cryptocurrency pricing. For machine learning models to yield reliable predictions, a variety of external factors must be considered.

Non-linear interactions: The price movements of cryptocurrencies may not conform to linear patterns, and there may be intricate interactions between different variables. Non-linear correlations must be taken into consideration in machine learning models, particularly in those that use deep learning.

CHAPTER 3

MATERIALS AND METHODS

3.1 Working Process

The complete task can be completed in four stages. The following are these:

- Collected dataset
- Data pre-processing with timestamp
- Model Selection
- Predicted result analysis

The total process from collecting dataset to prediction result are shown in the figure.

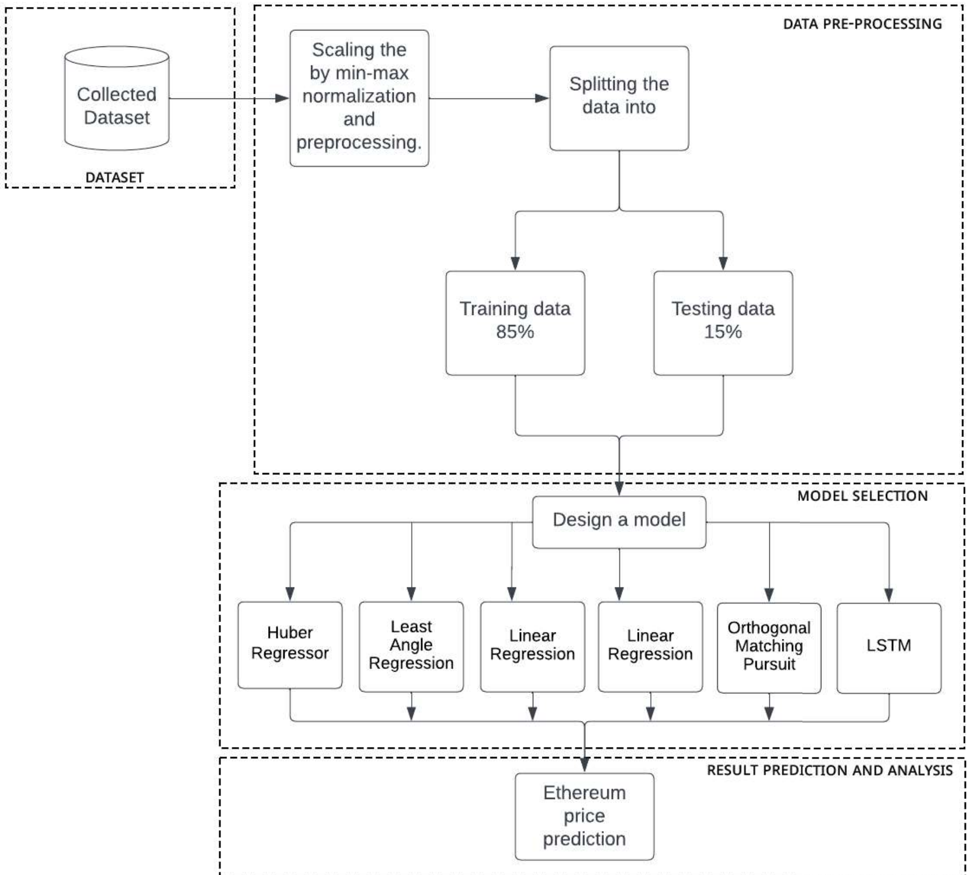


Figure.3.1.1 A basic diagram of the prediction model

3.2 Dataset Preparation

From 2017 to 2023, I gathered 2192 data points from Yahoo stock prices. The dataset has seven attributes. Six properties are present in every date: Open, High, Low, Close, Adj Close, and Volume. The day's high and low prices for the Ethereum cryptocurrency are respectively the highest and lowest values. The first price of the day is known as the open, and the last price is known as the close. A cryptocurrency's adjusted closing price is applied to the closing price to represent its value after any corporate actions. It is commonly used for examining historical returns or doing a comprehensive analysis of prior outcomes. The quantity of Ethereum exchanged on cryptocurrency exchanges is referred to as volume in an Ethereum price dataset.

Table3.2.1 Number of data and type of the dataset utilized in this research.

Name	Number of data	Type
Data	2192	Mixed
Attribute	7	Text
Timestamp	2192	Date
Open	2192	float
High	2192	float
Low	2192	float
Close	2192	float
Adj Close	2192	float
Volume	2192	float

Table3.2.2 overview of first 20 days of the dataset.

Date	Open	High	Low	Close	Adj Close	Volume
11/10/2017	320.671	324.718	294.542	299.253	299.253	885985984
11/11/2017	298.586	319.453	298.192	314.681	314.681	842300992
11/12/2017	314.69	319.153	298.513	307.908	307.908	1613479936
11/13/2017	307.025	328.415	307.025	316.716	316.716	1041889984
11/14/2017	316.763	340.177	316.763	337.631	337.631	1069680000
11/15/2017	337.964	340.912	329.813	333.357	333.357	722665984
11/16/2017	333.443	336.159	323.606	330.924	330.924	797254016
11/17/2017	330.167	334.964	327.523	332.394	332.394	621732992
11/18/2017	331.98	349.616	327.687	347.612	347.612	649638976
11/19/2017	347.401	371.291	344.74	354.386	354.386	1181529984
11/20/2017	354.094	372.137	353.289	366.73	366.73	807027008
11/21/2017	367.443	372.47	350.693	360.401	360.401	949912000
11/22/2017	360.312	381.42	360.147	380.652	380.652	800819008
11/23/2017	381.439	425.548	376.088	410.166	410.166	1845680000
11/24/2017	412.501	480.973	402.758	474.911	474.911	2292829952
11/25/2017	475.676	485.192	461.053	466.276	466.276	1422080000
11/26/2017	465.974	472.723	451.606	471.33	471.33	1197779968
11/27/2017	471.531	493.405	468.485	480.355	480.355	1396480000
11/28/2017	480.518	482.48	466.347	472.902	472.902	1346499968
11/29/2017	473.281	522.307	425.071	427.523	427.523	2675940096
11/30/2017	431.215	465.497	401.243	447.114	447.114	1903040000

3.3 Data Pre-processing

An important step in this Ethereum price prediction is data pre-processing. I start the process by gathering historical pricing data from dependable sources, including financial data providers or bitcoin exchanges. After obtaining, the data is cleaned to deal with outliers, duplicates, and missing values. The relevant variables influencing Bitcoin prices are then found using feature selection, with a focus on historical prices, trade volumes, and technical indicators. Time series decomposition is frequently used to identify residuals, seasonality, and underlying trends. To bring features to a consistent magnitude, normalization or scaling is necessary. Other feature engineering techniques could be generating lag features or calculating daily returns. The dataset is divided into training and testing sets, considering the temporal sequence of the time series data, and categorical data is suitably encoded. Particular attention is given to unbalanced data and time series-specific elements like frequency and temporal order. The data is ready for precise and trustworthy Ethereum price forecasts thanks to this thorough pre-processing.

3.3.1 Preparing data for prediction

First, I take A variable for predicting one day out into the future. Then I create a new Column shifted ‘n’ units up. The new column’s name is ‘future_price’.

Table3.3.1.1 overview of first 20 days of the dataset.

	Close	Future_Price
0	385.199707	370.671722
1	370.671722	386.295166
2	386.295166	389.875488
3	389.875488	401.590576
4	401.590576	394.961945
...	...	...
1091	1874.744873	1881.068848
1092	1881.068848	1861.643799
1093	1861.643799	1856.162354
1094	1856.162354	1834.165039
1095	1834.165039	NaN

1096 rows × 2 columns

3.4 Proposed Model

3.4.1 Huber Regressor

A kind of resilient regression model called the Huber Regressor was created to address how sensitive classic least squares regression was to data outliers. Biased estimates can result from outliers' excessive influence on the model parameters in ordinary least squares (OLS) regression. By utilizing a hybrid loss function, the Huber Regressor aims to achieve a compromise between the effectiveness of OLS and the resilience of techniques like the median. When there are potential outliers or high leverage points in the dataset that could unnecessarily affect the model's fit, the Huber Regressor is especially helpful.

3.4.2 Least Angle Regression

An algorithm for linear regression called Least Angle Regression (LARS) was developed as a substitute for conventional least squares techniques. The basic principle of LARS is to add predictors to the model one at a time, traveling in the direction of the predictor that has the highest correlation with the current residual at each step, so progressively building the regression model. Since the regularization parameter determines how much shrinkage is applied to the coefficients, LARS has a path wise optimization property, which means it computes the whole path of solutions for all possible values of the regularization parameter. When a sparse solution—one in which only a portion of the predictors will have non-zero coefficients in the final model—is anticipated or when there are more predictors than observations, LARS is frequently used.

3.4.3 Linear Regression

Linear Regression is one of the popular regression model. A basic statistical technique for simulating and comprehending the relationship between a dependent variable and one or more independent variables is called linear regression. The fundamental premise of linear regression is that a linear equation may adequately describe this relationship. One dependent variable and one independent variable make up a simple linear regression model; two or more independent variables make up a multiple linear regression model. The goal of the model is to identify the coefficients that reduce the discrepancy between the dependent variable's observed values and those predicted by the linear equation. The model incorporates an error term to account for variables that are not observed and assumes that the relationship is additive and linear. The direction and strength of the associations are indicated by the coefficients in the equation, which also represent the intercept and slopes of the regression line. Many different fields use linear regression extensively for objectives including making predictions, comprehending correlations, and determining the importance of variables. But it's important to understand the underlying presumptions of linear regression and to consider any potential drawbacks, particularly when working with real-world data that might not exactly match these assumptions.

3.4.4 Orthogonal Matching Pursuit

An iterative technique called Orthogonal Matching Pursuit (OMP) is frequently employed in machine learning and signal processing for sparse signal recovery and feature selection. When there are significantly more features than observations, OMP effectively finds and chooses a selection of pertinent features to roughly represent a target signal. By determining which feature is most correlated with the current residual and changing the solution, the algorithm iteratively moves forward. One characteristic that sets OMP apart is that it preserves orthogonality in the features that are chosen at every stage, which adds to the findings' stability and interpretability. The algorithm gradually adds characteristics that minimize the residual error in the most efficient way as it goes along. OMP finds use in a variety of fields, such as machine learning tasks like sparse regression, which aim to find and utilize the most informative features in a high-dimensional dataset, and compressive sensing, where it helps reconstruct signals from sparse observations.

3.4.5 Lasso Least Angle Regression

L1 regularization, also referred to as the Lasso (Least Absolute Shrinkage and Selection Operator) penalty, is included into the Least Angle Regression (LARS) technique in the Lasso Least Angle Regression (LARS Lasso) algorithm. When there are a lot of predictors (features) and simultaneous variable selection and regularization are required, LARS Lasso is especially helpful. It constructs the regression model step-by-step, adding predictors one at a time, just as LARS. But the LARS Lasso encourages sparsity in the coefficient estimations by adding the L1 penalty to the least squares objective function. In other words, LARS Lasso efficiently does feature selection by tending to generate models with a subset of predictors having non-zero coefficients. A parameter, commonly represented by α, governs the degree of regularization by weighing the trade-off between accurately fitting the data and maintaining a straightforward model. For applications like linear regression with feature selection, LARS Lasso is frequently utilized in machine learning and statistics. It helps identify the most important variables while reducing the possibility of overfitting in high-dimensional datasets.

3.4.6 LSTM

Recurrent neural network (RNN) architecture with Long Short-Term Memory (LSTM) was created to overcome the difficulties associated with learning long-term dependencies in sequential input. LSTMs were created as a solution to the drawbacks of conventional RNNs, and they work especially well for applications like time-series analysis, speech recognition, and natural language processing. The primary breakthrough of LSTMs is its capacity to selectively store and retrieve data across long sequences, hence reducing the vanishing gradient issue that conventional RNNs suffer from. To control information flow, LSTMs make use of an output gate, forget gate, input gate, and memory cell. Long-term information retention is achieved by the memory cell, while the input gate controls the amount of fresh information that enters the cell, the forget gate selects which information to remove, and the output gate sets the output depending on the state of the cell at that moment. LSTMs are well-suited for jobs that require modeling dependencies over extended temporal contexts because of their architecture, which enables them to recognize and retain pertinent patterns in sequential data. Because LSTMs are so good at handling both short- and long-term dependencies, they have become a mainstay in deep learning applications requiring sequential data.

3.4.7 Best Model

From those models above the Huber Regressor perform the best in this situation with the higher R2 score and low MAE score. Also, LSTM works best for long time prediction and for larger dataset.

3.5 Statistical Analysis

3.5.1 MSE

Mean Squared Error, or MSE for short, is a widely used metric to calculate the average squared difference between a set of data's actual and projected values. Regression analysis and machine learning employ the Mean Squared Error (MSE) extensively to assess a model's effectiveness.

The equation for Mean Squared Error (MSE) is as follows:

$$MSE = \frac{1}{n}\sum_{i-1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

Where:

- n is the number of data points.
- Y_i is the actual (observed) value for the i th data point.
- $\hat{Y}_i$ is the predicted value for the i th data point.

The mean squared error (MSE) is the sum of the squared deviations between the observed and expected values. Larger errors are penalized more severely than smaller ones when the discrepancies are squared. Better agreement between the actual and anticipated values is indicated by a lower MSE value, whilst larger disparities are suggested by a higher MSE value.

3.5.2 MAE

Another popular metric for calculating the average absolute difference between the actual and anticipated values in a set of data is MAE, or mean absolute error. Similar to MSE, MAE is often used to assess a model's performance in regression analysis and machine learning.

The equation for Mean Absolute Error (MAE) is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

- n is the number of data points.
- Y_i is the actual (observed) value for the i th data point.
- $\hat{Y}_i$ is the predicted value for the i th data point.
- $|\cdot|$ represents the absolute value.

The mean absolute difference (MAE) between the expected and actual values is computed. Since MAE does not square the differences like MSE does, every error adds the same amount to the metric. Because MAE is less susceptible to outliers than MSE, it is a good option when you want to highlight the significance of each individual error without exaggerating the effects of larger errors.

3.5.3 RMSE

Another metric for calculating the average magnitude of errors between expected and actual values is called Root Mean Squared Error, or RMSE. Like MSE and MAE, it is very well-liked in regression analysis and machine learning evaluation.

The equation for Root Mean Squared Error (RMSE) is derived from the MSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}$$

- n is the number of data points.
- Y_i is the actual (observed) value for the i th data point.
- $\hat{Y}_i$ is the predicted value for the i th data point.

Since RMSE is essentially the MSE's square root, its units are the same as those of the actual and projected values. To improve the RMSE's interpretability and match its scale to the original data, the square root is used. Like MSE, RMSE shows how well the actual and anticipated values match, with a lower RMSE number indicating greater agreement and a higher RMSE suggesting greater disparities. Because RMSE is sensitive to big errors, it can be used as a useful tool to highlight and penalize larger prediction errors.

3.5.4 R2

The percentage of the dependent variable's variance that can be predicted from the independent variable(s) is represented by the statistical measure known as R-squared (R2). Put differently, it offers a measure of the degree to which the independent variables in a regression model account for the variability of the dependent variable. The dependent variable's variability can be fully explained by the model when the R2 value is 1, meaning that zero indicates that no variability in the dependent variable is explained by the model. An increased R2 value indicates that the model fits the data more accurately.

The R2 value is often calculated using the following formula:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

- SS_{res} is the sum of squared residuals (the differences between the observed and predicted values)
- SS_{tot} is the total sum of squares, which represents the total variability in the dependent variable.

The square of the correlation coefficient (r) between the independent and dependent variables can also be used to define the formula for R2 in a straightforward linear regression with a single independent variable:

$$R^2 = r^2$$

R2 is a helpful metric, but it should be used cautiously and other criteria, including the suitability of the model and the context of the data, should also be considered. A regression model's validity cannot be determined by R2 alone.

3.5.5 Huber Regressor analysis

Huber Regressor works by minimizing a combination of the mean squared error and mean absolute error. The method minimizes the Huber loss function, which varies depending on the properties of the data, by iteratively changing the model parameters in the context of timestamp data. The model computes the Huber loss, a piecewise function, at each timestamp by evaluating the prediction error for each data point. It emphasizes precision by using the mean squared error for data points that are near to the prediction. On the other hand, it changes to the mean absolute error for points with greater variances, highlighting robustness against outliers.

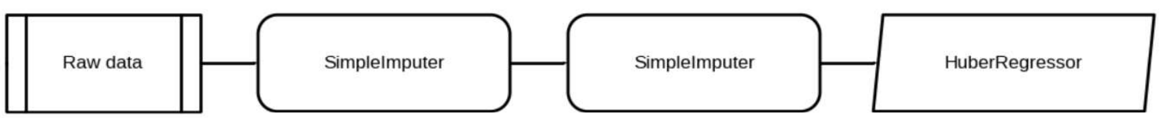


Figure.3.5.5.1 A basic Plot diagram of Huber Regressor.

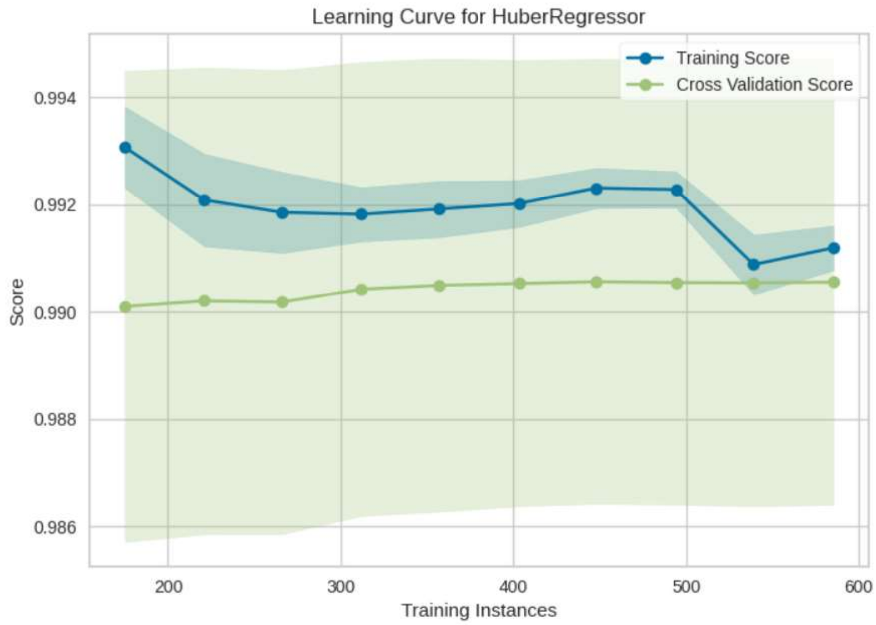


Figure.3.5.5.2 Learning Curve of Huber Regressor.

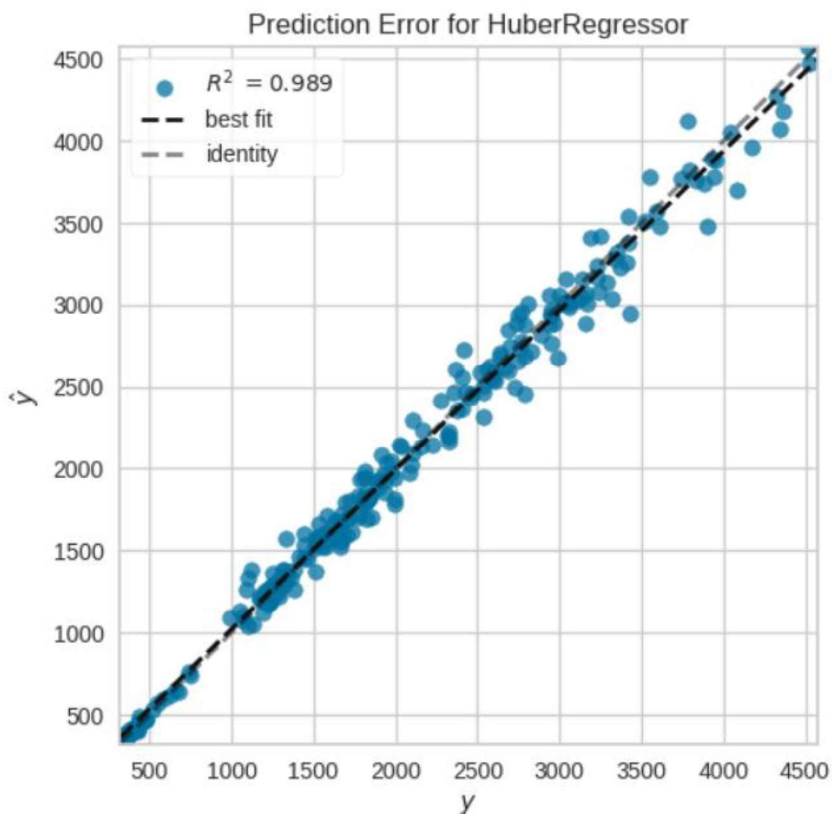


Figure.3.5.5.3 Prediction Error of Huber Regressor.

3.5.6 long short-term memory analysis

The input, forget, and output gates (input, output, and memory cells) that make up the fundamental architecture of LSTMs let the network to process and interpret sequential data in an efficient manner. Considering the current input data, the input gate determines at each timestamp which data should be stored in the memory cell. The forget gate lets the network handle long-term dependencies by simultaneously deciding what data from the prior cell state should be deleted. The LSTM can then adjust to changing patterns over time by updating the cell state in response to input and forget gate decisions. The output gate then selects the pertinent data to be delivered as the output at the current timestamp from the modified cell state. By using backpropagation to adjust their parameters during training, LSTMs maximize their capacity to identify patterns and temporal dependencies in the timestamp dataset.

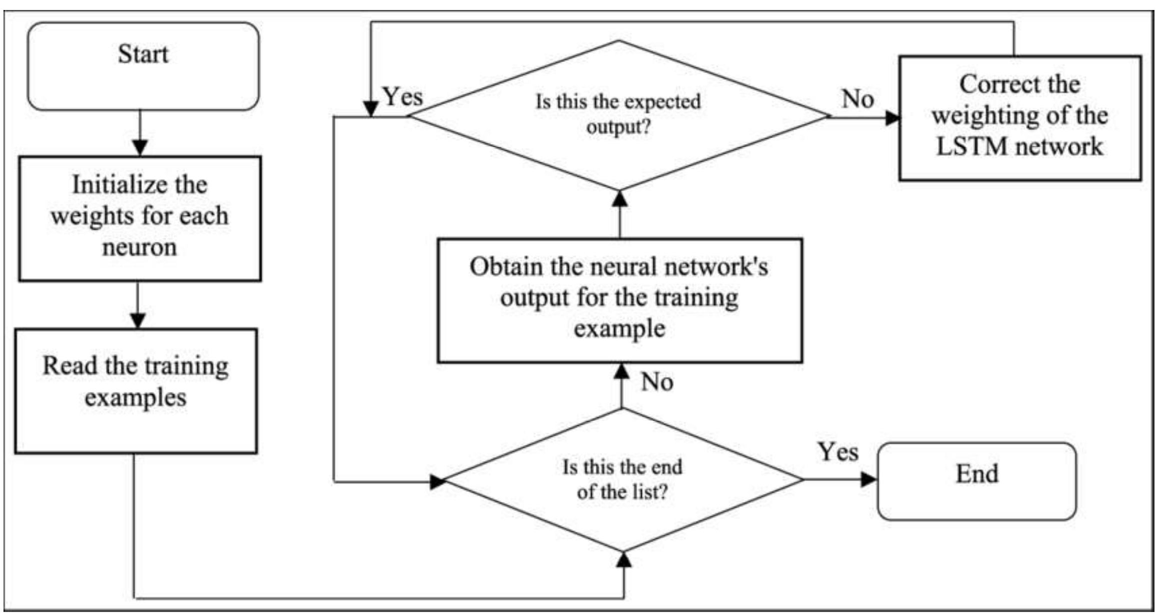


Figure.3.5.6.1 working procedure of LSTM training stage.

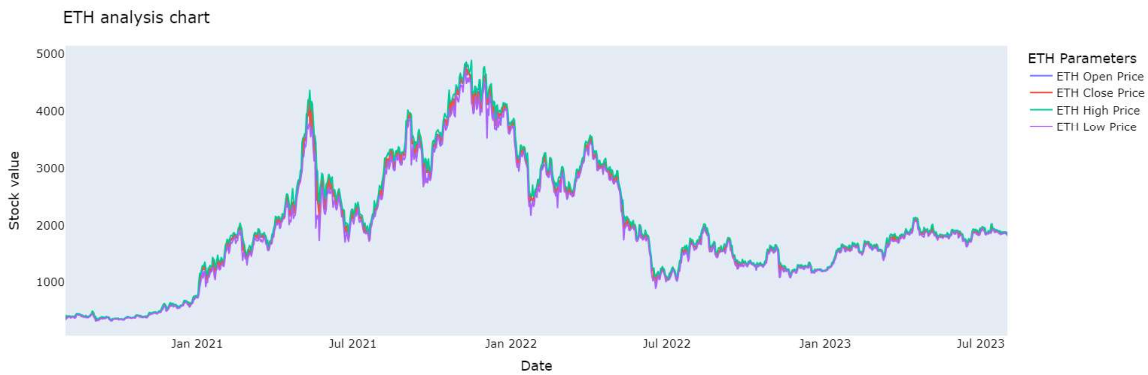


Figure.3.5.6.2 Ethereum data analysis chart for LSTM model.

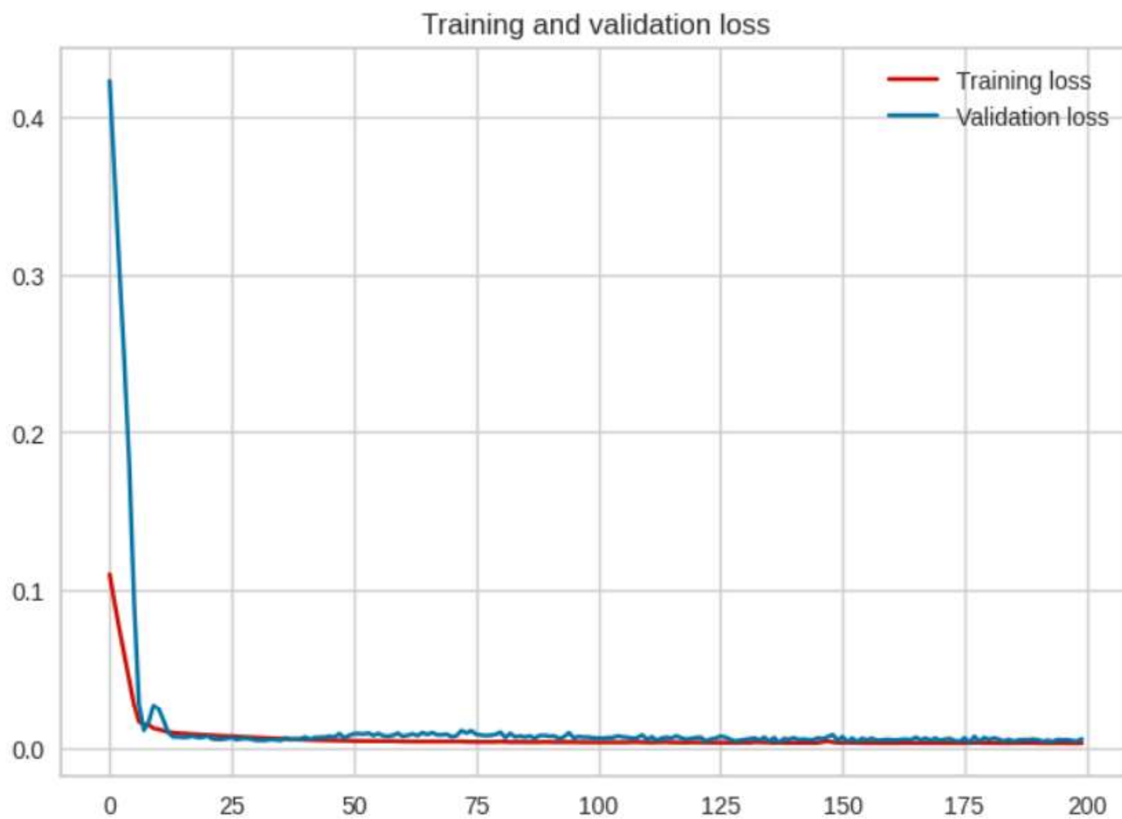


Figure.3.5.6.3 Training and validation loss curve of LSTM model.

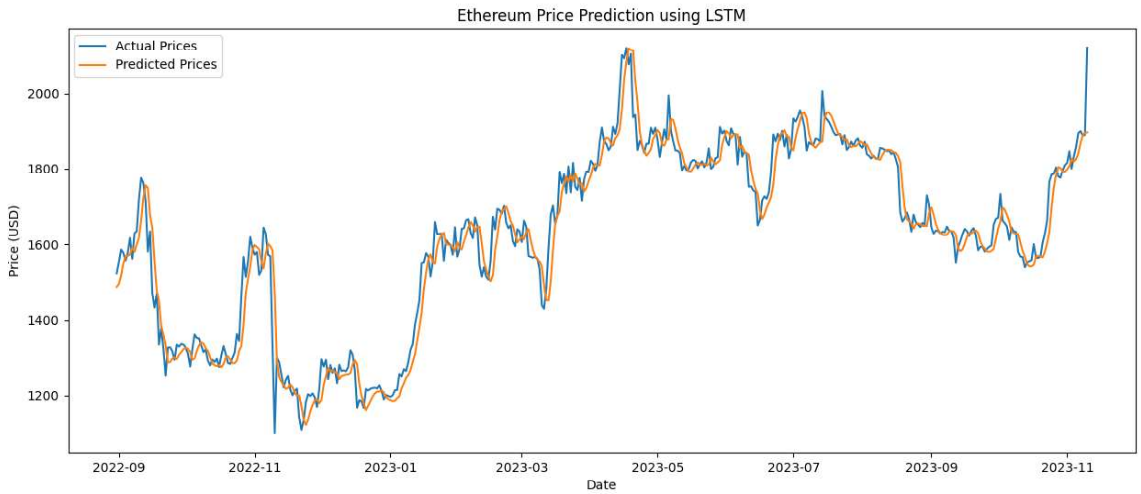


Figure.3.5.6.4 Full prediction price and actual price curve of Ethereum by using LSTM model.

3.6 Implementation Requirements

Developing and training the LSTM architecture requires a deep learning framework like TensorFlow or PyTorch. High-level abstractions and tools are provided by these frameworks to enable the effective construction of recurrent neural networks, including LSTMs. Furthermore, Python is a widely used programming language because of its large library and strong community, particularly in the fields of deep learning and machine learning.

The Huber Regressor can only be implemented with a good machine learning package like scikit-learn and an appropriate programming language and environment like Python. As part of its regression module, Scikit-learn offers a version of the Huber Regressor with an intuitive interface and interoperability with other data processing and visualization tools.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Results & Analysis

4.1.1 Accuracy:

$$\text{Accuracy based on MAE} = 100 - \frac{Absolute\ Error}{Total\ Data} 100\%$$

The results from different models are as follows:

Table.4.1.2 MAE, MSE, RMSE, R2, MAPE value for 6 models.

Model	MAE	MSE	RMSE	R2	Accuracy based on MAE
LSTM	35.9866	2863.48	53.5115	0.9486	98.358%
Huber Regressor	71.9439	12427.82	109.9026	0.9906	96.719%
Least Angle Regression	72.5203	12469.65	110.0593	0.9905	96.693%
Linear Regression	72.5203	12469.66	110.0593	0.9905	96.693%
Orthogonal Matching Pursuit	72.5203	12469.65	110.0593	0.9905	96.693%
Lasso Least Angle Regression	72.5204	12459.65	110.0593	0.9905	96.693%

From this research LSTM Mean Absolute Error (MAE) for predicting Ethereum is 35.9866 means that the magnitude of difference between the prediction of an observation and the true value of that observation is 35.9866. The Mean Squared Error (MSE) is 2863.48 means that the average of the squared differences between the actual and estimated values is 2863.48. Root Mean Squared Error (RMSE) measures the average difference between the values predicted by a model and the actual values. R squared (R2) value in machine learning is referred to as the coefficient of determination or the coefficient of multiple determination in case of multiple regression.

4.2 Discussion

According to the analysis, Huber Regressor has highest R2 score. However, because this is a regression model, the accuracy of the forecast will decrease with increased data. In that case, the large timestamp data set will get the best results using LSTM (long short-term memory). Because Timestamp datasets are a good fit for Long Short-Term Memory (LSTM) networks because of their special architecture and capabilities. In this research LSTM has the lower Mean Absolute Error with 35.9866. that means it has less error and higher accuracy. Long-term connections within sequences are well captured by LSTMs, which makes them useful for applications where the temporal order of events is important. Long-short-term memory (LSTM) networks are designed to address the vanishing gradient issue, which prevents them from learning and remembering patterns as long as a regular recurrent neural network (RNN) would. LSTMs' gating mechanisms improve their ability to adapt to multiple temporal patterns by allowing them to selectively recall or forget information at different timestamps. Additionally, LSTMs process sequences of varying durations with efficiency and excel in handling temporal delays.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND ETHICAL ASPECTS

5.1 Impact on Society

Predictions about the price of Ethereum have a wide-ranging effect on several industries and stakeholders in society. Predictions about the price of Ethereum have a big impact on personal investing decisions, investment plans, and portfolio management in general. Positive projections could draw in additional capital, encouraging a rise in the use of cryptocurrencies and possibly assisting in the general public's acceptance of digital assets as reliable financial tools. However, extreme volatility or gloomy forecasts could cause market ambiguity, which could affect investor confidence and interest in the bitcoin field. This might have wider economic ramifications, particularly if a sizable percentage of people or organizations have substantial Ethereum investments. Predictions about the price of Ethereum have an impact on the community that goes beyond simple investors and includes companies who are part of the Ethereum ecosystem, blockchain developers, and entrepreneurs. On the other hand, negative forecasts could cause the development community to become cautious, which would slow down the rate at which Ethereum-based applications gain traction and attract funding. This might impact future developments in technology as well as the expansion of the ecosystem as a whole. Ethereum's price projections may also have an impact on how the general public and regulatory bodies view cryptocurrencies. Anticipating favorable outcomes could lead regulatory agencies to investigate favorable policies, acknowledging the possible financial advantages and technological advancements linked to Ethereum and blockchain technology. At the same time, poor price projections could result in more regulatory scrutiny, which would hinder the advancement and use of decentralized technology.

5.2 Impact on Environment

The effect that Ethereum price forecasts have on the environment is directly related to the proof-of-work (PoW) consensus process that Ethereum currently uses. In order to solve challenging mathematical puzzles, Ethereum miners need a lot of processing power, which results in considerable energy usage. An increase in mining activity could result from a rise in Ethereum's price, which would amplify the environmental impact of PoW-based cryptocurrencies. Positive expectations for the price of Ethereum could draw additional miners looking to make big gains, which would increase energy usage. Negative price projections might force a reassessment of the existing PoW model and its environmental effects, while positive forecasts might highlight how urgent it is to switch to PoS or investigate other environmentally favorable options. In conclusion, the community's decisions on adopting more environmentally friendly consensus techniques may be influenced by Ethereum's price projections. They can also have an impact on the environmental conversation around blockchain technologies.

5.3 Ethical Aspects

Concerns about investor safety, responsible communication, and market behavior are some of the ethical issues underlying Ethereum price prediction. Like any financial estimate, price predictions for Ethereum have the ability to affect market dynamics. When there is a chance of purposefully manipulating prices for one's own benefit, taking advantage of information imbalances, or using dishonest tactics, ethical considerations surface. Predicting Ethereum prices ethically entails upholding openness, giving precise information, safeguarding investors, and encouraging responsible communication. Respecting these guidelines is essential to fostering trust in the cryptocurrency community and making sure that price forecasts advance the growth and development of the industry. In order to maintain transparency and honesty in their forecasts, analysts and financial specialists need to abide by ethical standards and steer clear of conflicts of interest that can jeopardize the accuracy of their studies.

CHAPTER 6

CONCLUTION AND FUTURE WORK

6.1 Conclusion

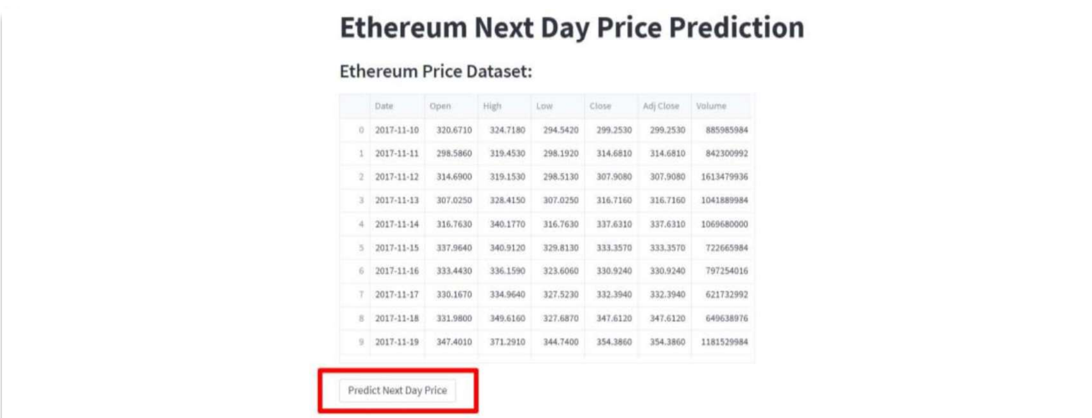
In the quickly changing cryptocurrency markets, there are both benefits and drawbacks to using machine learning (ML) to predict Ethereum prices. To improve prediction accuracy, these models can combine a range of characteristics, such as sentiment analysis from social media, market indicators, and historical price data. The quality and applicability of the input data, in addition to the meticulous selection and fine-tuning of model parameters, are critical factors that determine how well machine learning (ML) models predict Ethereum prices. Unexpected obstacles may be introduced by changes in the market, laws, and technology, which could affect how reliable forecasts are. Notwithstanding these difficulties, ML-driven Ethereum price prediction is a dynamic and cutting-edge method for comprehending and navigating the cryptocurrency market. ML has the ability to provide insightful information as technology and methods advance, helping analysts, investors, and the larger cryptocurrency community make wise decisions within the developing Ethereum environment.

6.2 Limitations and Future Work

Ethereum was created in 2013 and became publicly accessible in 2015. However, the available historical data only goes back to 2017. Since there are only about 3000 data points on the price of Ethereum, the forecast is not as precise as it would be with a larger dataset. For this reason, LSTM's performance in this circumstance was poor. But for this kind of dataset, LSTM is the ideal model. Also, I only use one neural network model, however this timestamp dataset can also be used with other neural network models, such as CNN and ANN which I will use in my future work.

6.3 Simple Web Application Implementation

I created a web application to visualize the output of my model. It will analyze the dataset and show the next day's close price. This predicting web application will train the dataset using LSTM model and save it for the future. By clicking the predict button it will use those train data and give the next one-day price as an output. The primary goal of this prediction web app is to give people an idea of how the price will be of ether cryptocurrency. In the future, I want to add some more models and give idea how the next 30 or more days price will be.



Ethereum Next Day Price Prediction

Ethereum Price Dataset:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2017-11-10	320.6710	324.7180	294.5420	299.2530	299.2530	885985984
1	2017-11-11	298.5860	319.4530	298.1920	314.6810	314.6810	842300992
2	2017-11-12	314.6900	319.1530	298.5130	307.9080	307.9080	1613479936
3	2017-11-13	307.0250	328.4150	307.0250	316.7160	316.7160	1041889984
4	2017-11-14	316.7630	340.1770	316.7630	337.6310	337.6310	1069680000
5	2017-11-15	337.9640	340.9120	329.8130	333.3570	333.3570	722665984
6	2017-11-16	333.4430	336.1590	323.6060	330.9240	330.9240	797254016
7	2017-11-17	330.1670	334.9640	327.5230	332.3940	332.3940	621732992
8	2017-11-18	331.9800	349.6160	327.6870	347.6120	347.6120	649638976
9	2017-11-19	347.4010	371.2910	344.7400	354.3860	354.3860	1181529984

Predict Next Day Price

Figure.6.2.1 Web app for predicting next day's price of Ethereum.

Ethereum Next Day Price Prediction

Ethereum Price Dataset:

	Date	Open	High	Low	Close	Adj Close	Volume
2182	2023-11-01	1,815.8721	1,858.3129	1,786.4095	1,847.0897	1,847.0897	10628825648
2183	2023-11-02	1,847.2599	1,873.8850	1,790.1147	1,800.6210	1,800.6210	9004197724
2184	2023-11-03	1,800.9165	1,835.0725	1,779.6355	1,832.7952	1,832.7952	7622864055
2185	2023-11-04	1,833.3689	1,867.2614	1,825.7366	1,857.6986	1,857.6986	4845080427
2186	2023-11-05	1,857.3956	1,911.6086	1,848.6201	1,894.1577	1,894.1577	8867152645
2187	2023-11-06	1,894.0271	1,914.5826	1,871.9730	1,899.8374	1,899.8374	8104122602
2188	2023-11-07	1,900.5997	1,907.5614	1,852.7909	1,888.1243	1,888.1243	9203228152
2189	2023-11-08	1,887.1190	1,904.5463	1,874.5457	1,889.3224	1,889.3224	6751627017
2190	2023-11-09	1,888.9403	2,130.8853	1,884.2043	2,120.5610	2,120.5610	24709695029
2191	2023-11-10	2,121.0674	2,132.9697	2,080.1399	2,090.4531	2,090.4531	27245359104

Predict Next Day Price

Predicted Next Day Price: 2070.8223

Figure.6.2.2 Result show in Web app for predicting next day's price of Ethereum.

REFERENCE

[1] Basher, S. A., & Sadorsky, P. (2022). Forecasting Bitcoin price direction with random forests: How important are interest rates, inflation, and market volatility?. *Machine Learning with Applications*, *9*, 100355.

[2] Inamdar, A., Bhagtani, A., Bhatt, S., & Shetty, P. M. (2019, May). Predicting cryptocurrency value using sentiment analysis. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 932-934). IEEE.

[3] Sibel Kervanci, I., & Fatih, A. K. A. Y. (2020). Review on bitcoin price prediction using machine learning and statistical methods. Sakarya University Journal of Computer and Information Sciences, 3(3), 272-282.

[4] Pabuçcu, H., Ongan, S., & Ongan, A. (2023). Forecasting the movements of Bitcoin prices: an application of machine learning algorithms. arXiv preprint arXiv:2303.04642.

[5] Khedr, A. M., Arif, I., El-Bannany, M., Alhashmi, S. M., & Sreedharan, M. (2021). Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. Intelligent Systems in Accounting, Finance and Management, 28(1), 3-34.

[6] Monish, S., Mohta, M., & Rangaswamy, S. (2022). Ethereum Price Prediction Using Machine Learning Techniques–A Comparative Study. International Journal of Engineering Applied Sciences and Technology, 7, 137-142.

[7] Kim, H. M., Bock, G. W., & Lee, G. (2021). Predicting Ethereum prices with machine learning based on Blockchain information. Expert Systems with Applications, 184, 115480.

[8] Derbentsev, V., Babenko, V., Khrustalev, K. I. R. I. L. L., Obruch, H., & Khrustalova, S. O. F. I. I. A. (2021). Comparative performance of machine learning ensemble algorithms for forecasting cryptocurrency prices. International Journal of Engineering, 34(1), 140-148.

[9] Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365, 112395.

[10] Ferdiansyah, F., Othman, S. H., Radzi, R. Z. R. M., Stiawan, D., Sazaki, Y., & Ependi, U. (2019, October). A lstm-method for bitcoin price prediction: A case study yahoo finance stock market. In 2019 international conference on electrical engineering and computer science (ICECOS) (pp. 206-210). IEEE.

[11] Ghosh, A., Bose, S., Maji, G., Debnath, N., & Sen, S. (2019, September). Stock price prediction using LSTM on Indian Share Market. In Proceedings of 32nd international conference on (Vol. 63, pp. 101-110).

[12] Rathan, K., Sai, S. V., & Manikanta, T. S. (2019, April). Crypto-currency price prediction using decision tree and regression techniques. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 190-194). IEEE.

[13] Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural computing and applications, 1-15.

[14] Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Research Letters, 32, 101084.

[15] Velankar, S., Valecha, S., & Maji, S. (2018, February). Bitcoin price prediction using machine learning. In 2018 20th International Conference on Advanced Communication Technology (ICACT) (pp. 144-147). IEEE.

siyam

ORIGINALITY REPORT

24%	21%	13%	17%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	Submitted to Daffodil International University Student Paper	2%
3	Submitted to Liverpool John Moores University Student Paper	1%
4	Submitted to University of Hertfordshire Student Paper	1%
5	www.researchgate.net Internet Source	1%
6	Submitted to Coventry University Student Paper	1%
7	Submitted to University of Bradford Student Paper	1%
8	Han-Min Kim, Gee-Woo Bock, Gunwoong Lee. "Predicting Ethereum Prices with Machine Learning based on Blockchain Information", Expert Systems with Applications, 2021 Publication	1%