DESIGN ANALYSIS AND MODELING TECHNIQUE PROJECT REPORT (2238-CSE-5301-004)

Project Name: Future Prediction and Analysis of Bitcoin Price.	**Champion: NA**
Business or Process Owner: NA	Project Leader: Dhruvin Sureshbhai Dholariya
Problem Statement: The primary issue with bitcoin is price volatility. Since there's a rapid growth with crypto currency's, the transaction of bitcoin over the network is slower and also very expensive to use it. Therefore, it is important to predict and analyze the bitcoin value.	Project Goal: To elevate the R-Squared value obtained from our comprehensive reference. This improvement is intended to enable more precise predictions and analyses of Bitcoin prices within a specific time frame. The resulting dataset will be a valuable asset for future trading activities, with the ultimate goal of achieving favorable results.
Business Case: To reduce the losses caused by Bitcoin trading and increase profits using a predictive model. Hence, it provides a better understanding of Bitcoin price fluctuations and creates opportunities for greater profit through trading.	Project Scope: The project is undertaken with the collected bitcoin prices to predict and analyze using linear regression prediction methods and further the results are analyzed based on the predictions.
Team Members: Dhruvin Sureshbhai Dholariya (1002156569) Srestha Somala (1002166735) Doppalapudi sai samhitha (1002165766) Varshith Konduru (1002132051) Sriya Sadineni (1002162910)	Benefits: Model effectively predicts Bitcoin price in the future which can be used in future investment. The model can be used to cut the underlying losses of the investors and it helps to predict when it is the good time to sell and buy Bitcoin.
	Timeline: 1. Proposal (Problem Identification-) – Sept 10th 2. Model Development - Oct 10th 3. Analysis – Nov 10th 4. Final report- Dec 5th

Future Prediction and Analysis of Bitcoin Price

Abstract

In this analysis, we accurately predict Bitcoin prices by considering various influencing parameters. In the initial stage of our prediction, our objective is to comprehend and recognize daily trends within the Bitcoin market, concurrently acquiring insights into the most effective features associated with Bitcoin prices. The dataset we employ encompasses a variety of features pertaining to Bitcoin prices and the payment network, recorded daily over a nine-year period. Subsequently, in the second phase of our investigation, we intend to utilize the gathered information to predict the direction of daily price changes with the utmost accuracy possible.

Introduction

The cryptocurrency known as Bitcoin is utilized all around the world for investing as well as digital payments. Since Bitcoin is decentralized, no one owns it. A digital wallet, which is essentially a virtual bank account, is where bitcoins are kept. Blockchain is the location where all transaction records and timestamp data are kept. A block is the name given to each record on a blockchain. Every block has a reference to the data block before it. Blockchain data is encrypted. The user's wallet ID is the only information made public during transactions; their identity is kept private.

Like stocks, the value of bitcoin fluctuates, but in a different way. For price prediction, a variety of algorithms are used to store stock market data. The factors influencing Bitcoins, however, differ. As a result, forecasting the value of Bitcoin is essential for making wise investment choices.

Since 2014, it is almost certain that there has been fraudulent manipulation of the price of bitcoin at some time. It is possible to conclude that there was 95% confidence in 2012 bitcoin manipulation, 95% confidence in 2017 bitcoin manipulation, and 98% confidence in 2019 bitcoin manipulation.

Dataset Data Source

The most important step was gathering past data on changes in Bitcoin values over the last 10 years. A nine-year archive of organized historical data is obtained from Yahoo finance.com.

The dataset includes bitcoin prices from September 17, 2014 to October 14, 2023 and it is .csv file. The dataset includes features with date, open, high, low, close, adj close, and volume.

Date	Open	High	Low	Close	Adj Close	Volume
9/17/2014	465.864	468.174	452.422	457.334	457.334	21056800
9/18/2014	456.86	456.86	413.104	424.44	424.44	34483200
9/19/2014	424.103	427.835	384.532	394.796	394.796	37919700
9/20/2014	394.673	423.296	389.883	408.904	408.904	36863600
9/21/2014	408.085	412.426	393.181	398.821	398.821	26580100
9/22/2014	399.1	406.916	397.13	402.152	402.152	24127600
9/23/2014	402.092	441.557	396.197	435.791	435.791	45099500
9/24/2014	435.751	436.112	421.132	423.205	423.205	30627700
9/25/2014	423.156	423.52	409.468	411.574	411.574	26814400
9/26/2014	411.429	414.938	400.009	404.425	404.425	21460800
9/27/2014	403.556	406.623	397.372	399.52	399.52	15029300
9/28/2014	399.471	401.017	374.332	377.181	377.181	23613300
9/29/2014	376.928	385.211	372.24	375.467	375.467	32497700
9/30/2014	376.088	390.977	373.443	386.944	386.944	34707300
10/1/2014	387.427	391.379	380.78	383.615	383.615	26229400
10/2/2014	383.988	385.497	372.946	375.072	375.072	21777700
10/3/2014	375.181	377.695	357.859	359.512	359.512	30901200
10/4/2014	359.892	364.487	325.886	328.866	328.866	47236500

Table 1. Dataset retrieved from Yahoo finace.com

Fig.1. Bitcoin price trends from September 17, 2014 to October 14, 2023.

Problem Statement

The primary issue with bitcoin is price volatility. Since there's a rapid growth with crypto currency's, the transaction of bitcoin over the network is slower and also very expensive to use it. Therefore, it is important to predict and analyze the bitcoin value.

Project Goal

To elevate the R-Squared value obtained from our comprehensive reference. This improvement is intended to enable more precise predictions and analyses of Bitcoin prices within a specific time frame. The resulting dataset will be a valuable asset for future trading activities, with the ultimate goal of achieving favorable results.

Business Case

To reduce the losses caused by Bitcoin trading and increase profits using a predictive model. Hence, it provides a better understanding of Bitcoin price fluctuations and creates opportunities for greater profit through trading.

Benefits

Model effectively predicts Bitcoin price in the future which can be used in future investment. The model can be used to cut the underlying losses of the investors and it helps to predict when it is the good time to sell and buy Bitcoin.

Prediction Model

Describing phrases that are used in this project the most frequently. The method used in predictive modeling is called linear regression. Fitting a linear equation (= a straight line) to the observed data is the method used in linear regression to try and predict the connection between two variables. Your income, for example, is an explanatory variable, while your spending, on the other hand, are regarded as dependent variables.

For linear regression models, the goodness-of-fit metric is called R-squared. This statistic shows the proportion of the dependent variable's variation that the independent variables account for together. R-squared provides a straightforward 0–100% scale for measuring the strength of the association between your model and the dependent variable.

Now that certain key words have been clarified, let's start with the prediction's steps.Data is transformed into the desired form in the first stage. The training dataset is put through a linear regression model after it has the proper format in order to identify the predictors and for each predictor's R-squared value is calculated.

Analyzing the Results

In our analysis, we used a simple method called linear regression to guess Bitcoin's future prices by looking at past prices. We checked how the highest and lowest prices of Bitcoin can help us predict its closing price.

Our findings showed that both the highest and lowest prices are good at helping us guess the closing price, as the numbers we got were significant. However, the starting point of our prediction, when we don't consider the high and low prices, doesn't really affect the closing price.

When we looked at the difference between what our model predicted and the actual prices, we found that most of these differences were small, which means our model did a good job. The graph comparing what we predicted and the actual prices showed a straight line, which is a good sign for our model.

But, there was a small issue. When we looked at the differences between the predicted and actual prices at different levels, we noticed that these differences weren't the same across the board. This could mean our model isn't perfect for all price ranges.

Even with this small hiccup, our model did a really good job of following the general trend of Bitcoin prices. In the future, we could make our predictions even better by looking at more factors or trying different methods to fix the small issue we found.

Screenshot of Results

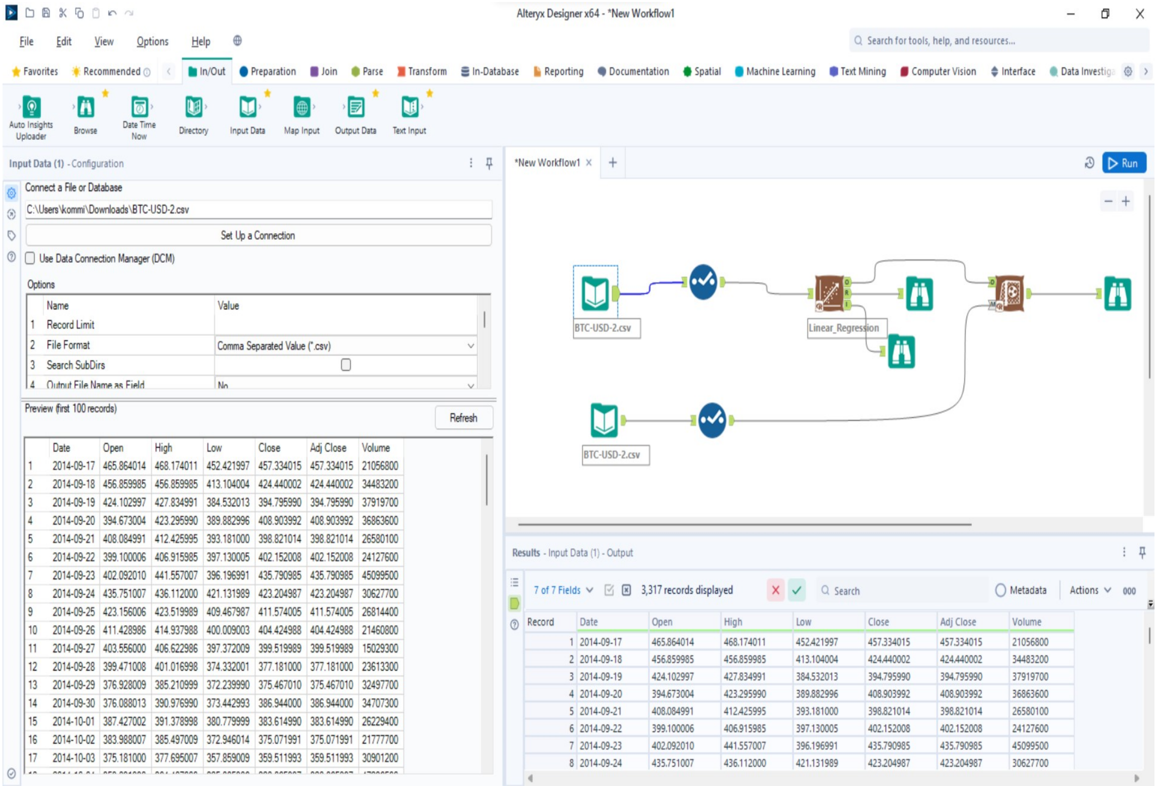


Fig.2. Load the dataset into Alteryx Designer

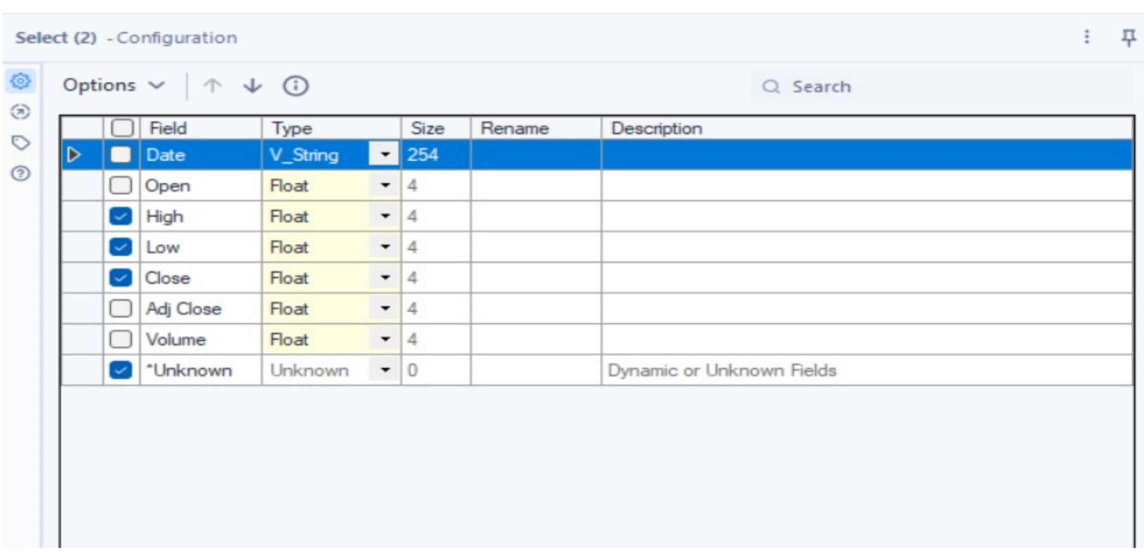


Fig.3. Configure the data type of the variables

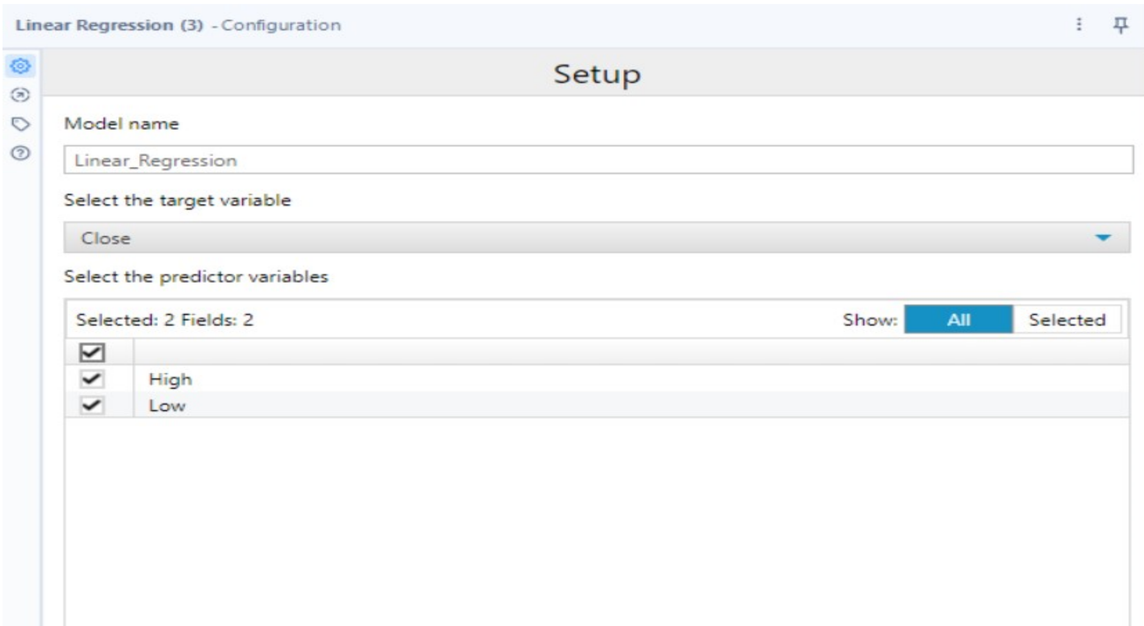


Fig.4. Setup the Linear Regression model

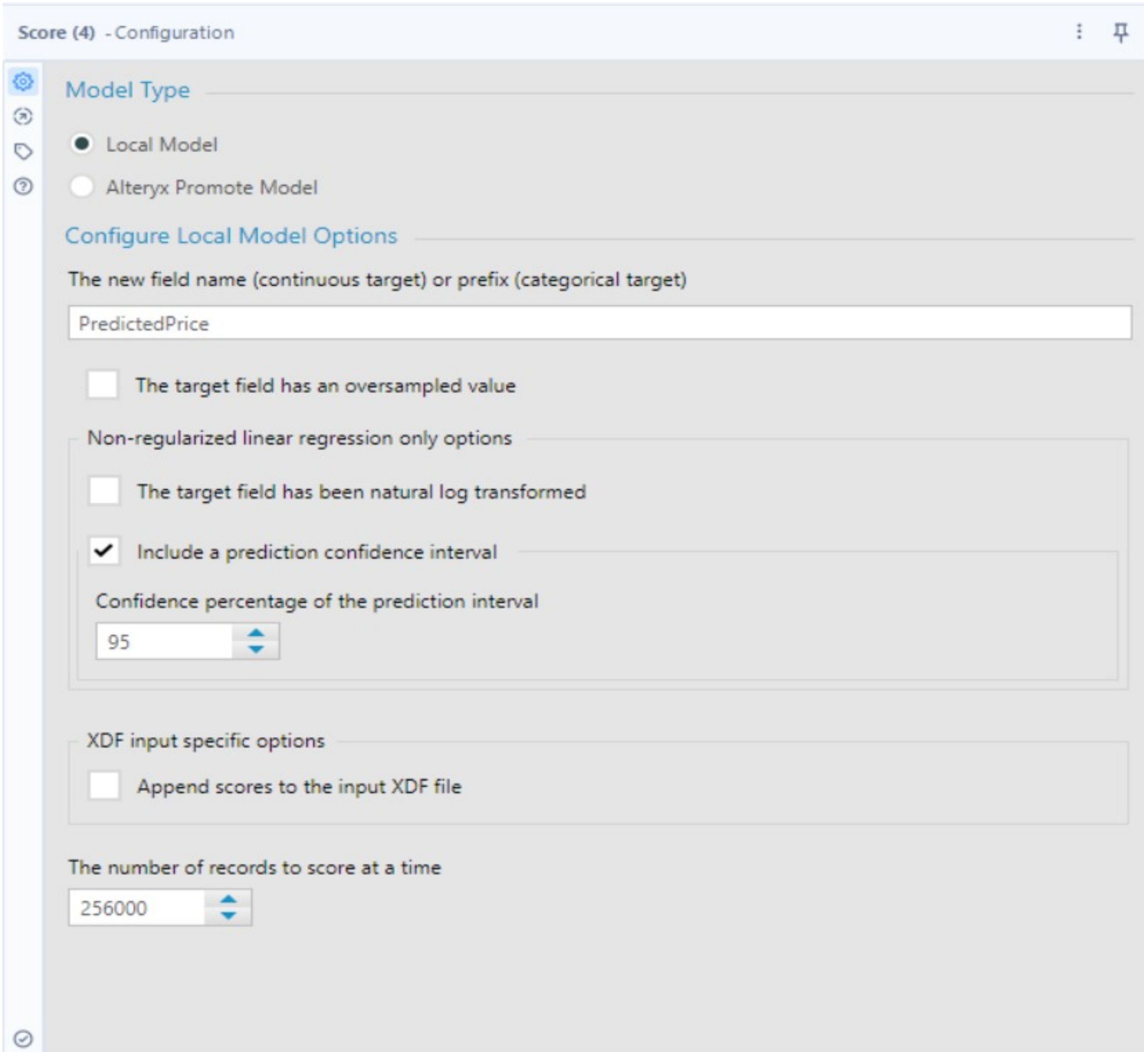


Fig.5. Configure Confidence percentage of prediction interval as 95 percent

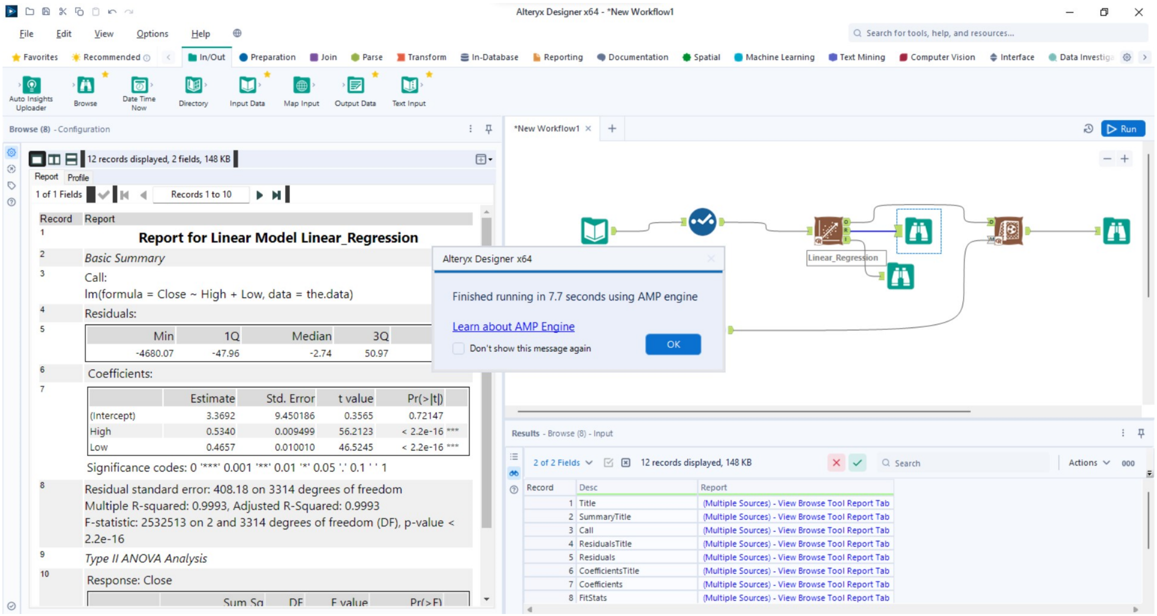


Fig.6. Shows Run Time is 7.7 seconds

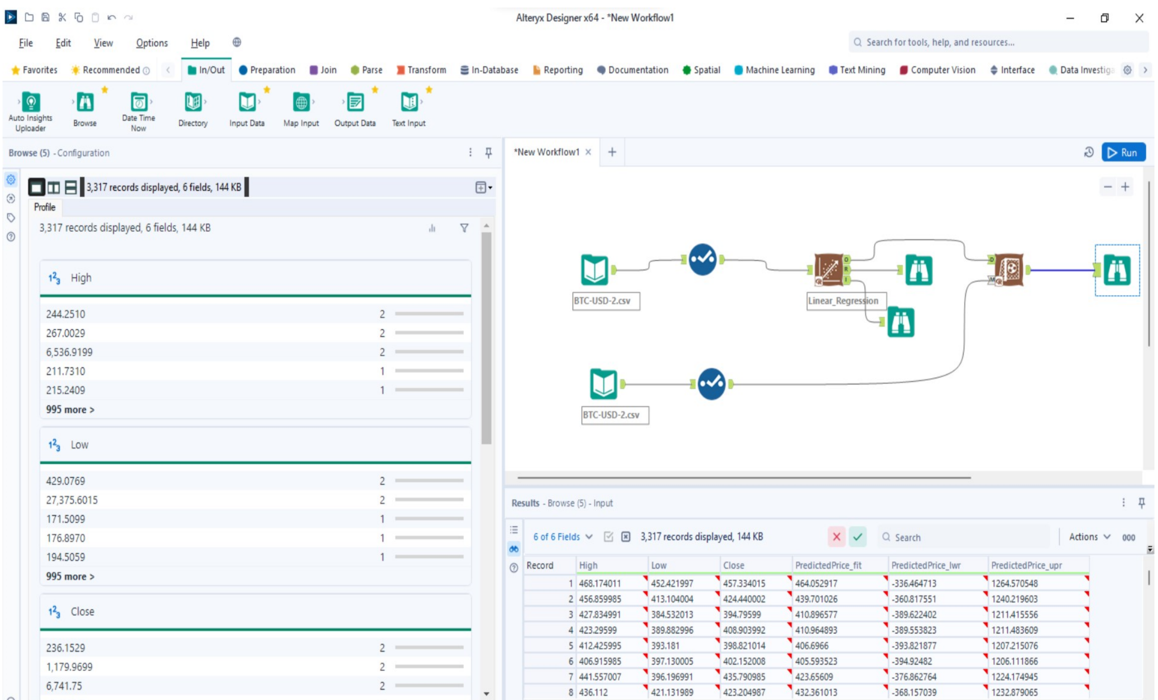


Fig.7. Displayed the results

Conclusion and Summary

In conclusion, our research underscores the dynamic evolution of Bitcoin as a compelling asset class within the alternative investment space, drawing considerable attention from investors. Through the construction of a predictive model and an investigation into the adherence of Bitcoin to market efficiency theories, we contribute valuable insights to both researchers and traders. The extensive dataset spanning September 17, 2014, to October 14, 2023, serves as the foundation for our analysis.

Utilizing linear regression and the r-squared method, we discern patterns and behaviors in Bitcoin movements. Whether Bitcoin aligns with an efficient market hypothesis or exhibits characteristics of a random walk is a crucial aspect of our inquiry. This study not only enhances our understanding of Bitcoin dynamics but also provides practical implications for investment strategies and market expectations in the rapidly evolving cryptocurrency landscape.

Record	Report
1	**Report for Linear Model Linear_Regression**
2	*Basic Summary*
3	Call: lm(formula = Close ~ High + Low, data = the.data)
4	Residuals:

5

Min	1Q	Median	3Q	Max
-4680.07	-47.96	-2.74	50.97	3789.48

6 Coefficients:

7

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	3.3692	9.450186	0.3565	0.72147	
High	0.5340	0.009499	56.2123	< 2.2e-16	***
Low	0.4657	0.010010	46.5245	< 2.2e-16	***

8 Residual standard error: 408.18 on 3314 degrees of freedom
Multiple R-squared: 0.9993, Adjusted R-Squared: 0.9993
F-statistic: 2532513 on 2 and 3314 degrees of freedom (DF), p-value < 2.2e-16

9 *Type II ANOVA Analysis*

10 Response: Close

	Sum Sq	DF	F value	Pr(>F)	
High	526459448.84	1	3159.82	< 2.2e-16	***
Low	360633586.44	1	2164.53	< 2.2e-16	***
Residuals	552147113.66	3314			

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fig.8. Report for Linear Regression Model

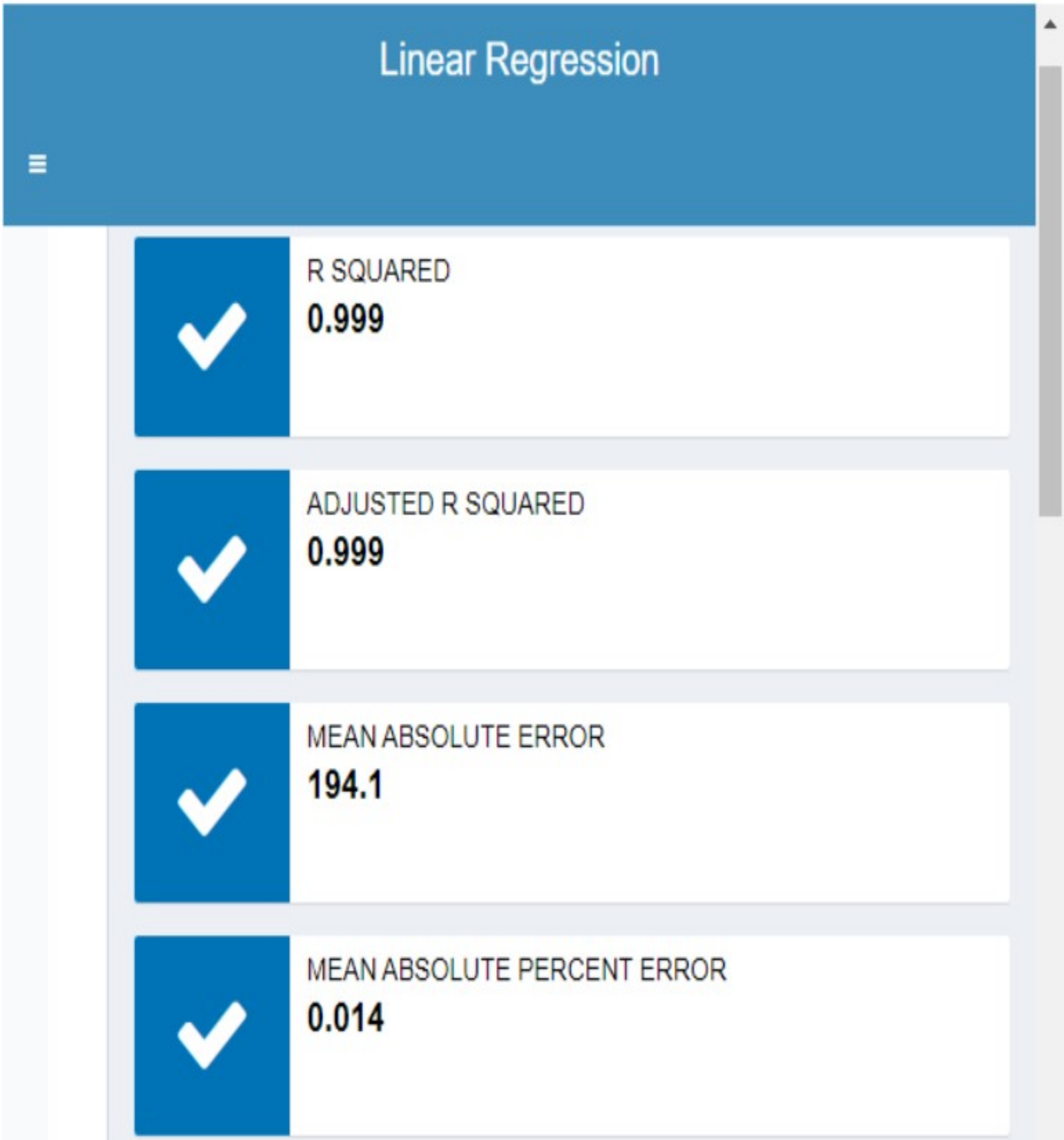


Fig.9. Shows r-squared value, adjusted r-squared, mean absolute error, mean absolute percent error

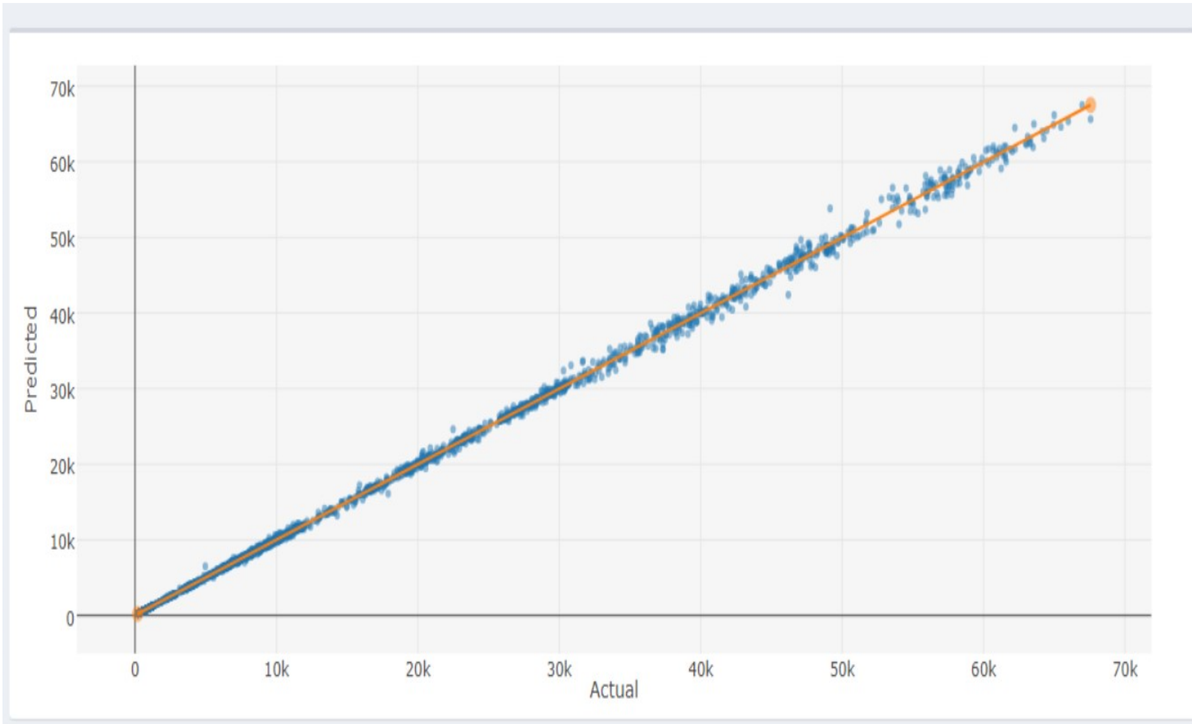
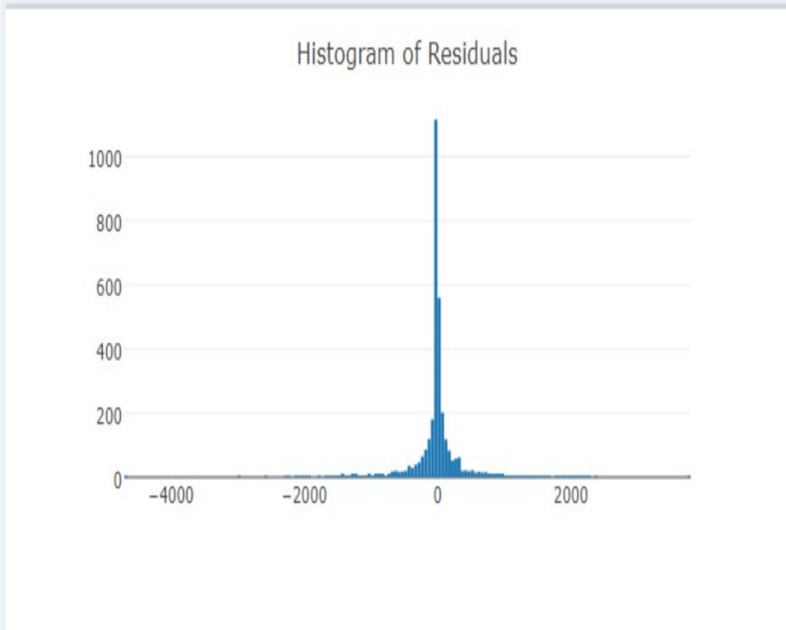


Fig.10. Shows plot of actual values vs predicted values



Statistic	Value
Minimum	-4680.07
1st Quartile	-47.96
Median	-2.74
Mean	0.00
3rd Quartile	50.97
Maximum	3789.49

Fig.11. Shows Histogram of Residuals

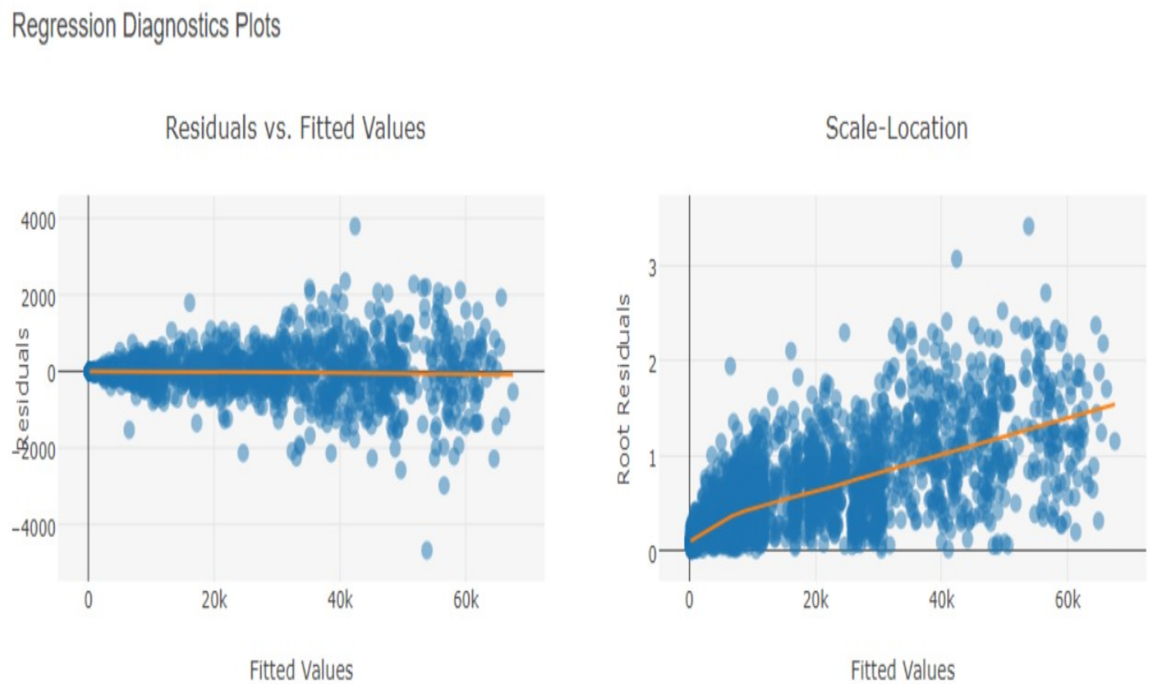


Fig.12. Shows plot of fitted values vs residuals and root residuals.

Fig.13. Shows plot of theoretical quantiles and leverage vs Standard Residuals.

References

D. Shah and K. Zhang, "Bayesian regression and Bitcoin", *52nd Annual Allerton Conference on Communication Control and Computing (Allerton)*, pp. 409-415, 2015.

Huisu Jang and Jaewook Lee, "An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information", *IEEE Early Access Articles*, vol. 99, pp. 1-1, 2017.

F. Andrade, de Oliveira, L. Enrique ZÃ!rate, M. de Azevedo Reis and C. Neri Nobre, "The use of artificial neural networks in the analysis and prediction of stock prices", *IEEE International Conference on Systems Man and Cybernetics*, pp. 2151-2155, 2011.

Aggarwal, Apoorva, Isha Gupta, Novesh Garg, and Anurag Goel. 2019. Deep Learning Approach to Determine the Impact of Socio Economic Factors on Bitcoin Price Prediction. Paper presented at 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, August 8–10.

M. Daniela and A. BUTOI, "Data mining on Romanian stock market using neural networks for price prediction", *informatica Economica*, vol. 17, 2013.

Akyildirim, Erdinc, Oguzhan Cepni, Shaen Corbet, and Gazi Salah Uddin. 2021. Forecasting mid-price movement of Bitcoin futures using machine learning. *Annals of Operations Research*, 1–32.

Awoke, Temesgen, Minakhi Rout, Lipika Mohanty, and Suresh Chandra Satapathy. 2021. Bitcoin Price Prediction and Analysis Using Deep Learning Models. In *Communication Software and Networks*. Singapore: Springer, pp. 631–40.

Basak, Suryoday, Saibal Kar, Snehanshu Saha, Luckyson Khaidem, and Sudeepa Roy Dey. 2019. Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance* 47: 552–67.

Baur, Dirk G., and Lai Hoang. 2021. The Bitcoin gold correlation puzzle. *Journal of Behavioral and Experimental Finance* 32: 100561.

Carbó, José Manuel, and Sergio Gorjón. 2022. Application of Machine Learning Models and Interpretability Techniques to Identify the Determinants of the Price of Bitcoin. Banco de Espana Working Paper No. 2215. Available online: **https://ssrn.com/abstract=4087481**

aur, Dirk G., and Thomas Dimpfl. 2021. The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics* 61: 2663–83.

García-Medina, Andrés, and Toan Luu Duc Huynh. 2021. What Drives Bitcoin? An Approach from Continuous Local Transfer Entropy and Deep Learning Classification Models. *Entropy* 23: 1582.

Kim, Alisa, Y. Yang, Stefan Lessmann, Tiejun Ma, M.-C. Sung, and Johnnie E. V. Johnson. 2020a. Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research* 283: 217–34.