PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

Ministry of Higher Education and Scientific Research

UNIVERSITY OF ECHAHID HAMMA LAKHDAR - EL OUED

FACULTY OF EXACT SCIENCES

Computer Science Department

Thesis

submitted in partial fulfilment of the requirements for the degree of

ACADEMIC MASTER

Field: **Mathematics and Computer Science**

Option: **Computer Science**

Specialty: **Distributed Systems and Artificial Intelligence**

Presented by : **Zekkour Ferhat Ali**

Moumen Messaoud Salem

Title

Using Artificial Intelligence Techniques for Analyzing and Predicting Cryptocurrencies Price

Defended on June 07 th 2023

Jury Commitee :

Mr.	Zaiz Faouzi	MCA. Univ El-Oued	Chair
Mme.	Bellila Khaoula	MAA. Univ El-Oued	Examiner
Mr.	Boucherit Ammar	MCA. Univ El-Oued	Supervisor

Academic year: 2022-2023

Dedication

This work is dedicated to the candles of my life, my lovely mother, the source of sympathy and love, I wish Mum that I had realized your dreams, Thank you great deal. And my thoughtful father for his sacrifices and mental support, My appreciation and many thanks to my supervisor Dr. Ammar Boucherit for his help, guidance and encouragement.

Acknowledgement

First and foremost, I owe thanks and gratitude to Almighty Allah for enlightening my path with knowledge and for granting me the ability to accomplish this work. Special thanks, all people who made this research work possible to all the teachers and students who participated in this research work for their cooperation and contribution.

Abstract

The development of financial technology has brought about a new type of asset called cryptocurrency, which presents an exciting opportunity for investment and we can see the trend of cryptocurrency is in continuous increasing. However, predicting the price of cryptocurrency is challenging due to its volatility and dynamic nature. Therefore, and in order to ensure a correct prediction with good accuracy, we have performed in this project a deep analysis on an existing dataset of the most 4 known cryptocurrencies such as Bitcoin, Binance coin, Ethereum, and Ripple to understand the market behavior by applying different machine learning algorithms such as LSTM, and Prophet to predict their short and long term price behavior. Our experimental result reaches a good level of accuracy.

Key words: Cryptocurrencies, Prediction, Machine Learning, Deep Learning, Time Series.

Résumé

Le développement de la technologie financière a engendré un nouveau type d'actif appelé crypto-monnaie, qui présente une opportunité d'investissement passionnante et nous pouvons voir que la tendance de la crypto-monnaie est en augmentation continue. Cependant, prédire le prix de la crypto-monnaie est difficile en raison de sa volatilité et de sa nature dynamique.

Par conséquent, et afin d'assurer une prédiction correcte avec une bonne précision, nous avons effectué dans ce projet une analyse approfondie sur un ensemble de données existant des 4 crypto-monnaies les plus connues telles que Bitcoin, Binance coin, Ethereum et Ripple pour comprendre le comportement du marché en utilisant différents algorithmes d'apprentissage automatique comme LSTM et Prophet pour prédire le prix à court et à long terme. Notre résultat expérimental atteint un bon niveau de précision.

Mots clés : crypto-monnaie, prédiction, machine learning, deep learning, séries temporelles.

الملخص

أدى تطور التكنولوجيا المالية إلى ظهور نوع جديد من الأصول يسمى العملات المشفرة، والذي يقدم فرصة مثيرة للاستثمار ويمكننا أن نرى اتجاه العملة المشفرة في ازدياد مستمر. ومع ذلك ، فإن التنبؤ بسعر العملة المشفرة يمثل تحديًا نظرًا لتقلبه وطبيعته الديناميكية.

وعليه، أصبح التنبؤ الصحيح بأسعار العملات المشفرة أكثر أهمية لكسب الأرباح. ولغرض تحديد التنبؤ الصحيح بدقة جيدة، أجرينا في مشروع بحثنا الحالي تحليلًا عميقًا لمجموعة البيانات السابقة لأكثر 4 عملات رقمية تداولا مثل **Bitcoin**، **Binance coin**، **Ethereum** و**Ripple** لفهم سلوك السوق باستخدام خوارزميات مختلفة للتعلم الآلي مثل **LSTM** و**Prophet** وذلك لغرض التنبؤ بالسعر على المدى القريب والبعيد. وقد توصلنا بعد التطبيق إلى نتائج تجريبية ذات دقة معتبرة.

الكلمات المفتاحية: العملات المشفرة، التنبؤ، التعلم الآلي، التعلم العميق، السلاسل الزمنية.

General Introduction

Digital currency or Crypto-currency is an encrypted virtual currency that enables its owners to purchase goods and services online. It uses a web-based digital ledger with strong encryption to secure the online financial transactions.

The cryptocurrency, which was only understood by a relatively fringe community of investors, is becoming a household name that cannot be ignored due to its long rising use by normal peoples.

In fact, the increasing interest in these virtual currencies comes as a result of what is known as "digital currency trading", in which these currencies are bought and sold to achieve real financial profits from them. Additionally, analysts estimate that the global cryptocurrency market will more than triple by 2030.

On the other hand, cryptocurrencies have suffered a severe beating in the past few months, i.e. the most popular cryptocurrencies such as Bitcoin, Ethereum, Binance coin and Ripple losing value by up to two-thirds. Therefore, deeply analyzing their Price turbulence/volatility using artificial intelligence techniques is essential for understanding the overall price of cryptocurrencies – and predicting their future values – especially in a highly uncertain cryptocurrency future.

Research problem

The price volatility is one of the main problems that characterize Bitcoin, Ethereum, and other cryptocurrencies. Where the crypto arena/market lacks a clear model that can predict its next movement, that is why the opening price point is often strange and always tempting for investors, especially those in the uptrend, and in fact anyone holding a cryptocurrency at 0.01$ can swing and make their investor get 10,000$.

However, the problem is that many predictions are made without evidence or analysis to support them. Due to the significant price swings (volatility) of cryptocurrencies, it is crucial to examine the viability of price prediction in order to facilitate the creation of precise investment plans for traders and to protect them from significant price swings.

To do this, it would be highly helpful to employ machine learning models to forecast cryptocurrencies prices using some of their features like price, volume and market cap across a number of years.

Through careful monitoring of the prices of the cryptocurrencies chosen for this study, we will address an important issue represented in the application of artificial intelligence techniques to predict cryptocurrency prices and give traders the opportunity to create investment plans for the near future.

Therefore, we can summarize our problem into the following questions:

- Is it possible to predict the price of cryptocurrencies such as Bitcoin, Ethereum?
- To what extent are we able to predict with high accuracy such prices?
- What is the best machine learning model that can predict the price of the most known cryptocurrencies?

Research goals

We aim to achieve a number of goals through our research on this subject. The most crucial of these goals is to:

- Studying the possibility of making use of machine learning models to predict the prices of cryptocurrencies, especially in light of the large volatility in their prices.
- Find the most accurate model to predict the price of cryptocurrencies.
- Provide a trained model to traders to predict the price of cryptocurrencies in a specific period/time.

Document structure

The present work is composed of four chapters, where we have devoted two chapters for the theoretical part and two chapters for the practical part. The first chapter discusses the concepts related to the problem of cryptocurrencies prices. In the second chapter, we presented the most important concepts of machine learning, how their models work and their importance in the case of our study. In the third chapter, we presented some existing works in the literature in order to benefit of their advantages.

Finally, in the last chapter we focused on the presentation and analysis of results obtained from our approach with the machine learning models used for the prediction of the price of cryptocurrencies.

Chapter 1

Overview on Cryptocurrencies and Price Prediction

1. Introduction

Cryptocurrencies have grown significantly since Nakamoto (2008) introduced Bitcoin. Thanks to blockchain technology, cryptocurrencies do not involve intermediaries and are therefore free from sovereign risk. Today, Cryptocurrency is an investment asset and despite its speculative nature and high volatility can be traded on specialized exchanges (Polasik et al. 2015) (Cheah).

In addition, identifying the suitable asset class for cryptocurrencies is a recurring subject in the literature; some studies compare cryptocurrencies to equities (Symitsi& Chalvatzis, 2019), bonds (Corbet et al., 2018), gold (Klein et al., 2018), or other precious metals (Mensi et al., 2019), among other things. Cryptocurrency may be most beneficial as an investment asset, despite the fact that researchers have not yet defined it and despite its appeal as a medium of trade. The literature that is currently available also depicts Bitcoin as a valuable portfolio diversifier, as it offers protection from sovereign risk and weakness (Bouri, Jalkh, et al., 2017; Bouri, Molnar, et al., 2017; Corbet et al., 2018; Dyhrberg, 2016); additionally, it is resilient during times of distress, confirming its potential role as a hedge and safe haven against the growing global uncertainty.

The cryptocurrency's extreme price volatility has prompted more literary research on its price development and the potential for an asset bubble. Due to its distinct characteristics that influence its price dynamic, such as cryptocurrency supply and demand (Bouoiyour&Selmi, 2015), investor attractiveness (Bouoiyour et al., 2016), user anonymity (Ober et al., 2013), and computer programming enthusiasts (Wilson &Yelowitz), the majority of researchers have argued that cryptocurrencies are isolated from the economic and financial fundamentals. It's interesting to note that vanWijk (2013) and Ciaian et al. (2016) found that macro financial development influences Bitcoin price fluctuations, as well as macroeconomic news (Alkhazali et al., 2018), economic policy uncertainty (Demir et al., 2018; Fang et al., 2019), and the risk in the US stock market (Conrad) [01].

2. Cryptocurrencies (what, When, Why, How many)

2.1 What is Cryptocurrency?

Cryptocurrency refers to a type of digital asset that uses distributed ledger, or blockchain, technology to enable a secure transaction. In other words, Cryptocurrency is the name given to a system that uses cryptography to allow the secure transfer and exchange of digital tokens in a distributed and decentralized manner. These tokens can be traded at market rates for fiat currencies. Although the technology is widely misunderstood, many central banks are considering launching their own national cryptocurrency [02, 21]. The following figure Fig.1.1 illustrates the symbols of some cryptocurrencies.

Fig.1.1. Symbols of some Cryptocurrencies.

2.2 When?

As used to be until now mentioned, most researches agree that the first cryptocurrency brought was once bitcoin. It was once created in 2008 and in accordance to Devries (2016) to make certain rarity, there is a distinctive number of this currency that can be generated by way of users. An uncommon element of blockchain finance is explored by using the creator who points out that, unlike traditional currencies, cryptocurrencies exist due to perceived fee that peoples are accepting them as repayments location on this technology. For instance, a seller receiving bitcoins has to agree with that this forex has fee when you consider that no institution can grant support for it. Due to this reason, one can argue that the blockchain finance market is greater intricate when in contrast to usual ones and requires an extra advanced technological know-how such as AI for acceptable analysis and predictions [03].

2.3 Why?

The emergence of cryptocurrencies has been driven by a combination of technological advancements, economic factors, and societal motivations. This section highlights some of the key reasons behind the rise of cryptocurrencies.

1. Decentralization and Trust:

One of the primary motivations for the development of cryptocurrencies was to create a decentralized form of digital currency that operates without the need for intermediaries like banks or government authorities. This concept of trustless transactions, facilitated by blockchain technology, appealed to individuals seeking greater financial autonomy and control (Nakamoto, 2008).

2. Financial Inclusion:

Cryptocurrencies have the potential to provide financial services to the unbanked and underbanked populations worldwide. By leveraging blockchain technology, individuals can access financial systems and conduct transactions without traditional banking infrastructure, reducing barriers to entry and promoting financial inclusion (Swan, 2015).

3. Privacy and Security:

The desire for increased privacy and security in financial transactions has also contributed to the emergence of cryptocurrencies. The cryptographic techniques employed in cryptocurrencies offer enhanced privacy features, allowing users to maintain anonymity in their transactions while ensuring the security and integrity of the blockchain network (Antonopoulos, 2017).

4. Speculation and Investment Opportunities:

The potential for high returns on investment and speculative trading has attracted many individuals to cryptocurrencies. The volatile nature of the cryptocurrency market, coupled with the increasing value of certain cryptocurrencies over time, has created opportunities for traders and investors seeking substantial profits (Griffin & Shams, 2018) [04].

4.4 How many?

Since the launch of Bitcoin in 2008, several hundred different ‘cryptocurrencies’ have been developed and become accepted for a wide variety of transactions in leading online commercial marketplaces. We cite here some of the most important ones.

4.4.1 Bitcoin:

January 2008, Satoshi Nakamoto developed the first cryptocurrency as Bitcoin, the first manufacturing of which took vicinity on January 2009. Bitcoin is expressed as a digital forex used in trading transactions that can be carried out barring the need for a middleman at some stage in digital transactions. Any transaction that can be finished with bodily cash can be performed in purchases made the use of this system circulation in the duration up to today is approximately 18,436 thousand. (Eğilmez, 2013). The figure Fig.1.2. gives an overview about it.

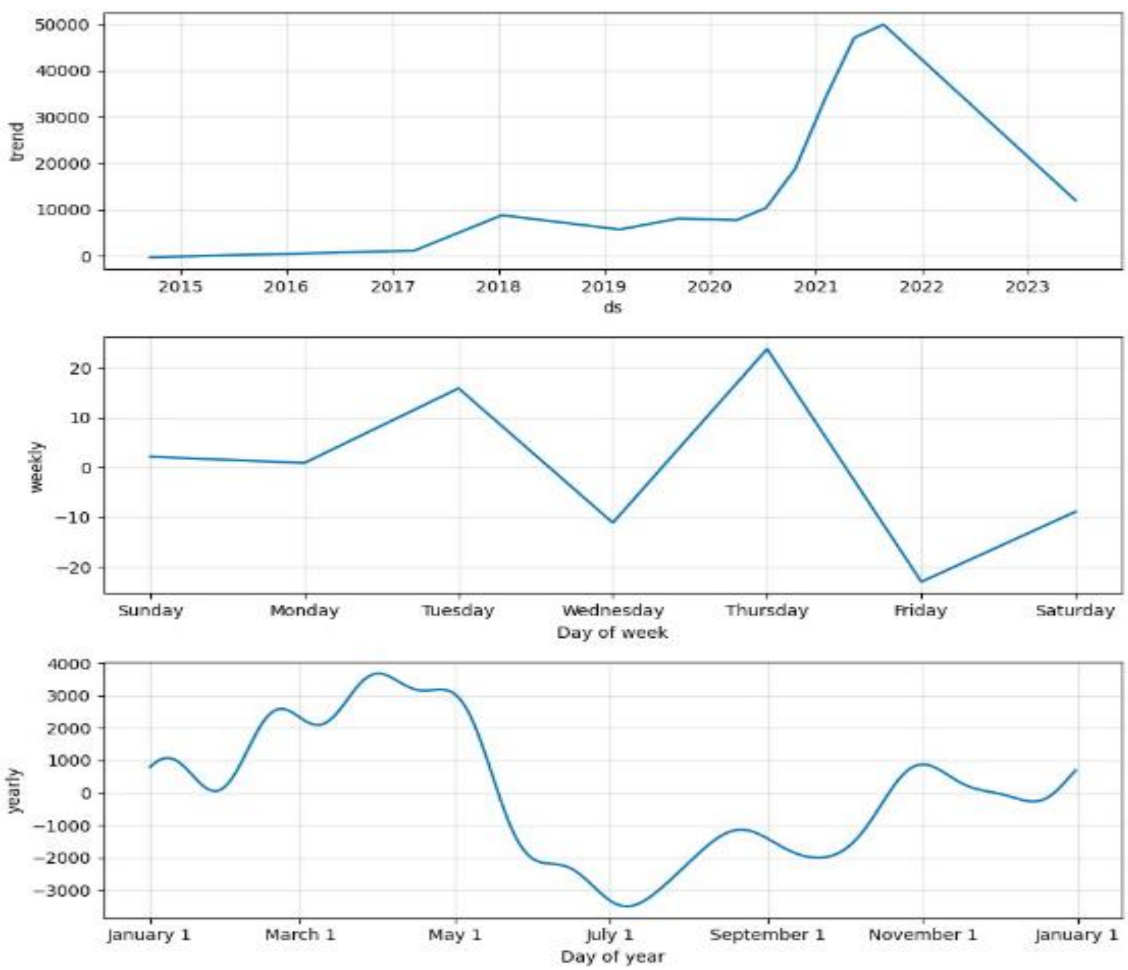

Fig.1.2. Bitcoin trend and seasonality.

4.4.2 Ethereum:

Ethereum is the second largest cryptocurrency by market capitalization behind bitcoin. Its current market cap stands at 290.69 billion USD, representing approximately 14% of the total cryptocurrency market1.

It was imagined in 2013, and created in 2015, the blockchain network Ethereum has grown in innovation and utility. Different from Bitcoin's primary function as a peer-to-peer electronic cash system, Ethereum invents a new world of peer-to-peer applications [22].

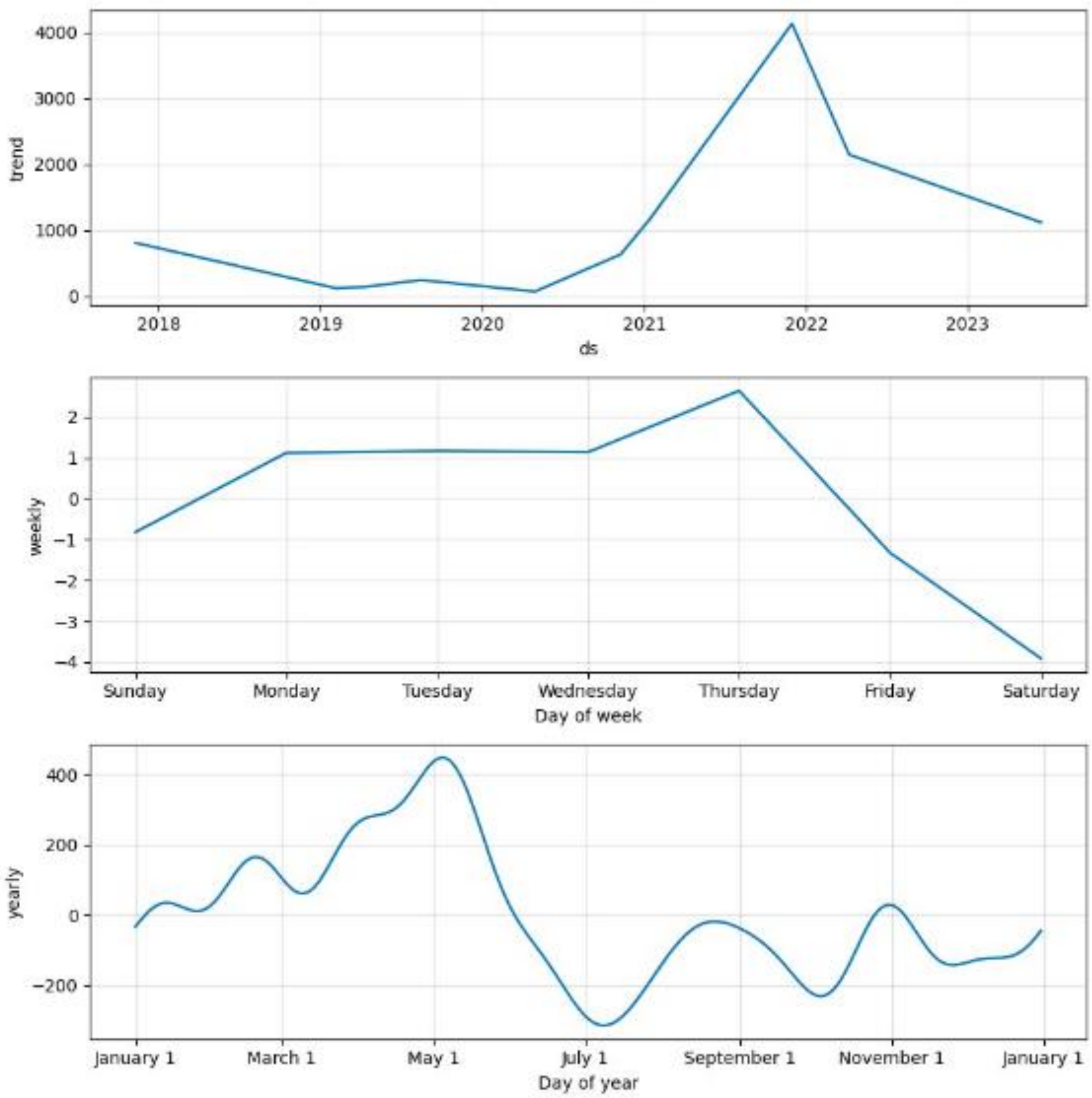

Fig.1.3. Ethereum trend and seasonality.

4.4.3 Ripple:

Ripple is a payment system and a digital currency which evolved completely independently of Bitcoin. The Ripple holds the second highest market cap after Bitcoin.

Although Ripple utilizes XRP and the XRP Ledger in its product offerings, XRP is independent of Ripple. The XRP Ledger is decentralized, open-source, and operates on what is known as a "consensus" protocol. While there are well over a hundred known use cases for XRP and the XRP Ledger, Ripple leverages XRP for use in its product suite because of XRP's suitability for cross-border payments [05]. The figure Fig.1.4. illustrates Ripple trend and seasonality.

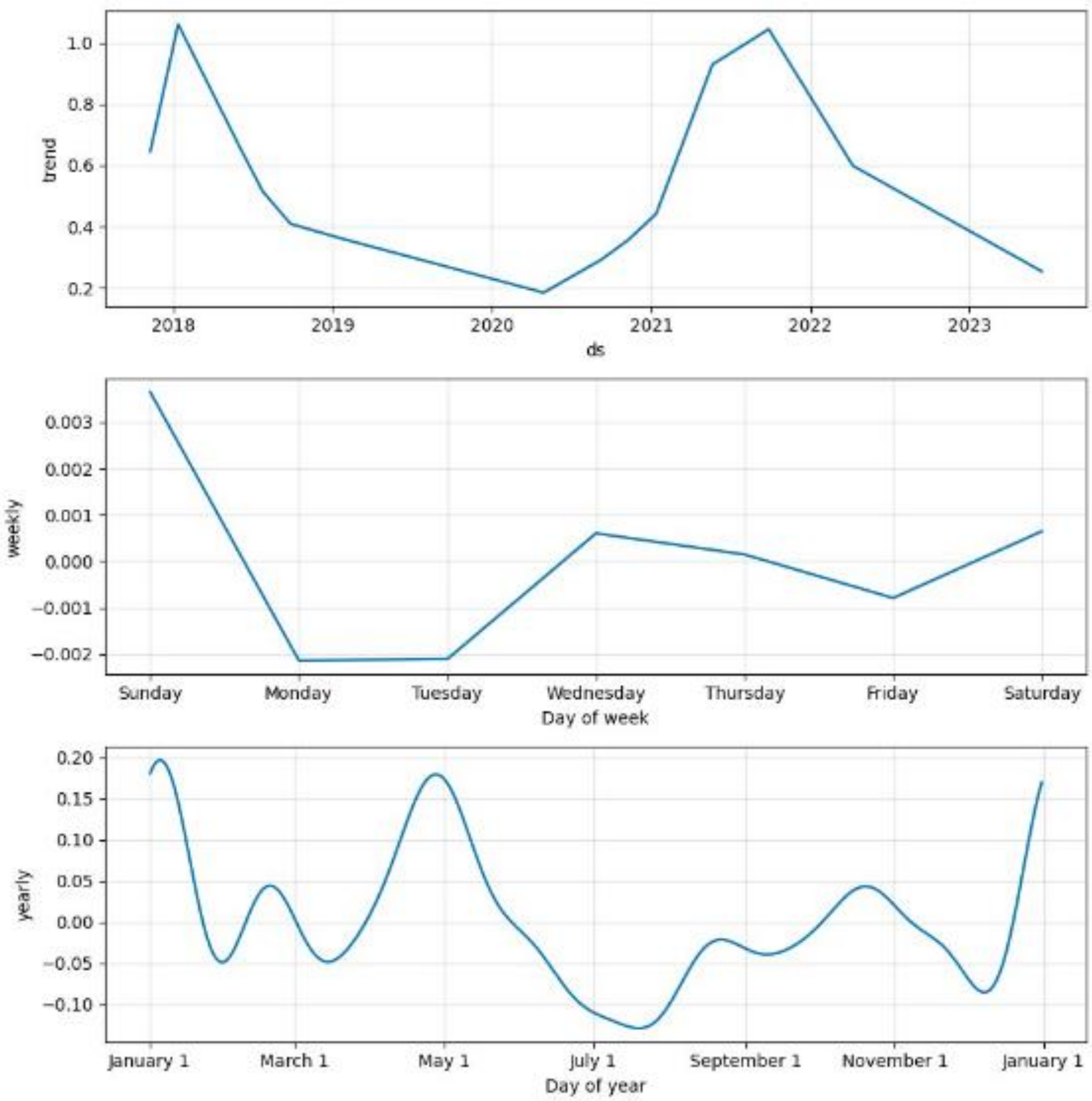

Fig.1.4. Ripple trend and seasonality.

4.4.4 Binance Coin (BNB):

Binance Coin (BNB) is the native cryptocurrency of the Binance exchange platform, one of the largest and most popular cryptocurrency exchanges in the world. BNB was launched in 2017 as a token on the Binance Chain blockchain [06].

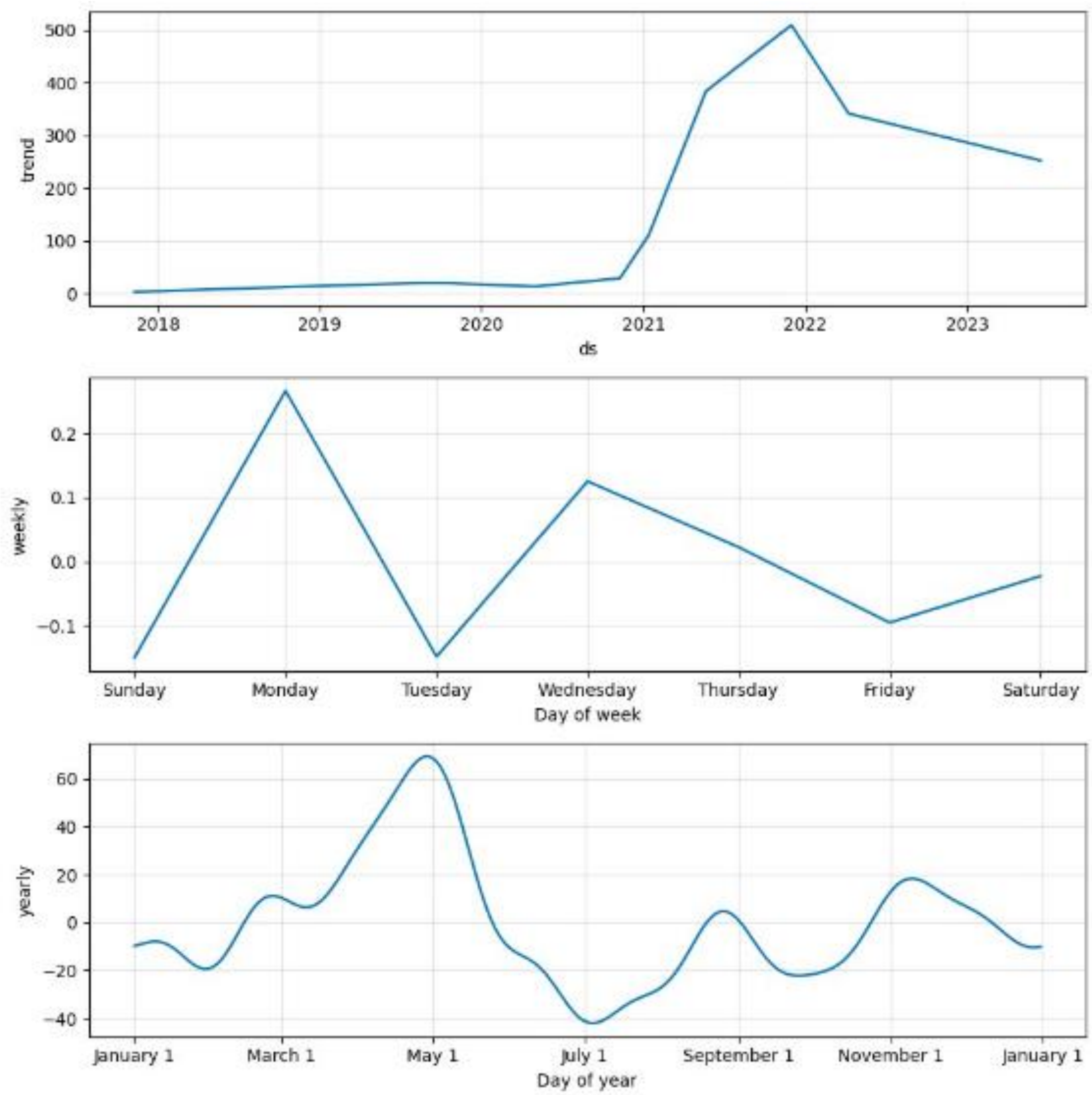

Fig.1.5. Binance trend and seasonality.

3. Purpose and significance of predicting cryptocurrency prices:

The purpose of predicting cryptocurrency prices and its importance can be summarized as follows:

3.1. Investment Decision Making:

Predicting cryptocurrency prices plays a crucial role in making informed investment decisions. By analyzing market trends, historical data, and various indicators, investors can gain insights into the potential profitability of different cryptocurrencies. This helps them identify opportunities for buying or selling assets at optimal prices, maximizing their returns and minimizing risks.

3.2. Risk Mitigation:

Cryptocurrency markets are highly volatile, which poses significant risks for investors. Predicting cryptocurrency prices allows individuals and organizations to assess and manage these risks more effectively. By understanding potential price fluctuations, investors can implement risk management strategies such as setting stop-loss orders, diversifying their portfolios, or adjusting their trading positions accordingly.

3.3. Trading Strategies:

Predicting cryptocurrency prices is essential for developing effective trading strategies. Traders utilize price predictions to identify patterns, trends, and market sentiment that can guide their trading decisions. Technical analysis tools and algorithms are often employed to forecast price movements and generate buy or sell signals. This enables traders to enter or exit positions at opportune moments, enhancing their chances of profitability.

3.4. Market Research and Analysis:

Cryptocurrency price predictions contribute to market research and analysis. Researchers and analysts use these predictions to gain insights into the behavior of different cryptocurrencies, assess their potential value, and identify factors that drive price movements. This knowledge aids in evaluating the market's overall health, identifying emerging trends, and informing investment strategies for individuals and institutions.

3.5. Economic and Financial Planning:

Cryptocurrencies have become a significant part of the global financial landscape. Predicting their prices is important for economic and financial planning. Governments, central banks, and financial institutions need accurate price forecasts to assess the impact of cryptocurrencies on the economy, determine regulatory measures, and develop strategies for incorporating cryptocurrencies into existing financial systems [07].

4. Conclusion

In this chapter, we have presented the most basic concepts about cryptocurrencies that have emerged as a revolutionary technology that aimed to change the traditional financial system since they offer a decentralized, secure, and transparent way of conducting transactions without the need for intermediaries such as banks. Bitcoin, the first cryptocurrency, paved the way for the development of numerous other cryptocurrencies, each with its own unique features and value propositions.

Despite their potential benefits, cryptocurrencies have also faced significant challenges, including regulatory hurdles, volatility in prices, and security concerns.

As cryptocurrencies continue to evolve and gain wider acceptance, it remains very important to look for the best technics to study the volatility of their price.

Chapter 2

Overview on Machine Learning

1. Introduction:

Machine learning is a rapidly growing field that has the potential to revolutionize the way we approach complex problems and make decisions. At its core, machine learning is the process of teaching computers to learn and improve from experience, without being explicitly programmed. This allows machines to identify patterns and make predictions based on data, leading to more accurate and efficient solutions to problems that were previously unsolvable.

Machine learning has already made significant contributions across a wide range of industries, including healthcare, finance, transportation, and entertainment. It has been used to develop personalized medical treatments, predict financial market trends, improve transportation systems, and create more immersive gaming experiences, among many other applications.

In this chapter, we will provide an overview of machine learning, including the different types of machine learning algorithms, the various applications of machine learning, and the potential benefits and challenges of using machine learning in different contexts.

2. Types of models and algorithms:

Machine Learning algorithms are mainly divided into four categories: Supervised learning, Unsupervised learning, Semi-supervised learning, and Reinforcement learning [10].

2.1 Supervised learning:

Supervised getting to know is normally the venture of desktop studying to examine a feature that maps an enter to an output based totally on sample input-output pairs . It makes use of labeled education data and a collection of coaching examples to infer a function. Supervised mastering is carried out when certain goals are recognized to be achieved from a sure set of inputs, i.e., a task-driven approach. The most common supervised tasks are “classification” that separates the data, and “regression” that fits the data. For instance, predicting the category label or sentiment of a piece of text, like a tweet or a product review, i.e., text classification, is an example of supervised mastering.

2.2 Unsupervised learning:

Unsupervised gaining knowledge of analyzes unlabeled datasets without the need for human interference, i.e., a data-driven procedure. This is widely used for extracting generative features, identifying meaningful trends and structures, groupings in results, and exploratory purposes. The most frequent unsupervised studying duties are clustering, density estimation, characteristic learning, dimensionality reduction, finding affiliation rules, anomaly detection, etc.

2.3. Semi-supervised learning:

Semi-supervised learning can be described as a hybridization of the above-mentioned supervised and unsupervised methods, as it operates on each labeled and unlabeled data. Thus, it falls between learning “without supervision” and learning “with supervision”. In the real world, labeled statistics could be uncommon in numerous contexts, and unlabeled data are numerous, the place semi-supervised learning is beneficial. The ultimate intention of a semi-supervised studying mannequin is to grant a higher effect for prediction than that produced the usage of the labeled facts by myself from the model. Some utility areas the place semi-supervised studying is used consist of machine translation, fraud detection, labeling statistics and text classification.

2.4. Reinforcement learning:

Reinforcement learning is a type of machine learning algorithm that enables software agents and machines to automatically evaluate the optimal behavior in a particular context or environment to improve its efficiency, i.e., an environment-driven approach. This type of learning is based on reward or penalty, and its ultimate goal is to use insights obtained from environmental activists to take action to increase the reward or minimize the risk. It is a powerful tool for training AI models that can help increase automation or optimize the operational efficiency of sophisticated systems such as robotics, autonomous driving tasks, manufacturing and supply chain logistics, however, not preferable to use it for solving the basic or straightforward problems.

3. Applications of ML and DL:

Machine Learning (ML) and Deep Learning (DL) have a wide range of applications across various industries and domains. Here are some key applications of ML and DL:

Image and Object Recognition: ML and DL algorithms are widely used for image and object recognition tasks. They can analyze and classify images, detect objects in images or videos, and perform tasks like facial recognition and object tracking. These applications are used in fields such as security, healthcare, autonomous vehicles, and entertainment.

Natural Language Processing (NLP): ML and DL techniques are used for NLP tasks such as text classification, language translation, speech recognition, and chatbots. NLP has applications in customer support, virtual assistants, content analysis, and language understanding [11].

Fraud Detection: ML algorithms can detect fraudulent activities by analyzing patterns and anomalies in large datasets. They are used in financial institutions, e-commerce platforms, and cybersecurity to identify fraudulent transactions, account hacking, and suspicious behavior.

Predictive Analytics: ML and DL models are used for predictive analytics to forecast future trends, behaviors, Prices and events. They can be applied in various domains such as stock market prediction, demand forecasting, healthcare outcome prediction, and weather forecasting.

Autonomous Systems: ML and DL techniques are used in autonomous systems like self-driving cars, drones, and robotics. These algorithms enable these systems to perceive and interpret their environment, make decisions, and navigate autonomously [12].

4. Choosing suitable model/models:

As the problem studied in our project is related to time series, we will be limited in this section to such specific axis. On the other hand, time series forecasting refers to the process of predicting future values of a variable based on its past observations or values. It is a statistical modeling technique commonly used in various fields such as finance, economics, sales forecasting, weather prediction, and more.

The key characteristics of time series data include the temporal ordering of observations, autocorrelation (the dependence of current values on past values), and potentially other patterns such as trends, seasonality, and irregular fluctuations.

In general, time series forecasting is a valuable tool for making predictions based on historical data, allowing decision-makers to anticipate future trends, make informed decisions, and optimize resource allocation [13]. In the following, we will present the most important concepts related to the used models for the price prediction.

4.1 On the near term:

4.1.1 LSTM:

Cryptocurrencies and blockchain technology have revolutionized the financial and technological landscape. One of the key challenges in this field is predicting cryptocurrency prices, as investors and traders are keen on understanding future price trends to make informed investment decisions. In recent years, deep learning techniques have emerged as powerful tools for price prediction, and among them, LSTM (Long Short-Term Memory) models have proven to be effective and efficient in forecasting cryptocurrency prices over time [14].

LSTM, one of the most powerful RNN-based algorithms, is used in the suggested prediction model, which is similar to many current efforts on cryptocurrency prediction that use deep learning-based techniques. LSTM can handle data presented in a sequential order and is commonly used in sequence modeling applications such as natural language processing (NLP) and time-series forecasting. Thus, in the proposed model, LSTM is employed as a primary feature extractor that directly accepts time-series data as input to extract relevant temporal information. LSTM updates hidden states on a regular basis using the preceding time step's output [15].

4.1.2 Principle of LSTM Algorithm:

LSTM (Long Short-Term Memory) is made up of three layers: an input layer, a hidden layer, and an output layer, and it can effectively solve the problem of gradient disappearance or explosion due to the long time span of dependencies between information, learn and predict based on historical data, balance the temporal and nonlinear relationships of electricity consumption data, and obtain better prediction results. Next figure depicts the LSTM network structure.

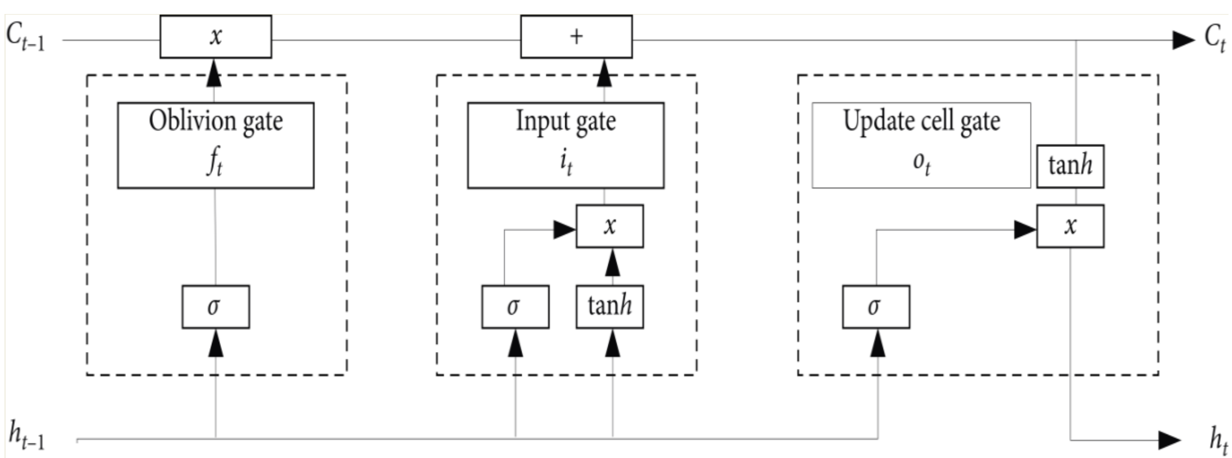

Fig.2.1 Network structure of LSTM.

- <u>Oblivion Gate :</u>

The forgetting gate ***ft*** how much information can be passed from the previous cell state to the current cell state, with an output of 1 indicating that all information from the previous moment is retained and an output of 0 indicating that all information from the previous moment is discarded.

$$f_t = \sigma\left(w_{xf}x_t + w_{hf}h_{t-1} + w_{cf}c_t + b_f\right)$$

where f_t is the oblivion gate, $\sigma$is the excitation function, taking values between [0,1], ω is the weight matrix, and b_f is the deviation vector.

- Input Gate :

By engaging the function $\sigma$to filter fresh input information and add it to the memory cell, the input gate $\mathbf{i_t}$ updates the candidate information in the memory cell.

$$i_t = \sigma\left(w_{xi}x_t + w_{hi}h_{t-1} + w_{ci}c_t + b_i\right)$$
$$g_t = \tan\ h\left(w_{xc}x_t + w_{hc}h_{t-1} + w_{ci}c_t + b_c\right)$$

where $\mathbf{i_t}$ and $\mathbf{g_t}$ are the input gate, x_t is the input vector, $b_i$andb_c are the deviation vectors.

- Update Cell State :

The output of the forgetting gatef_t and the old cell state $\mathbf{c_{t-1}}$ of the previous moment form part of the updated cell state ,$\mathbf{c_t}$ while the other part consists of the output of the input gate $\mathbf{i_t}$ and the memory cell candidate state.

$$c_t = fc_{t-1} + i_t\sigma_t\left(w_{xc}x_t + w_{hc}h_{t-1} + b_c\right)$$

where $\mathbf{c_t}$ is the updated cell state and$\mathbf{c_{t-1}}$ is the old cell state at the previous moment.

- Input Gates and Hidden Layer Outputs:

The output gate O_t can be obtained by the excitation function σ. The output vector h_t can be obtained by compressing and updating the cell state $\mathbf{c_t}$.

$$o_t = \sigma\left(w_{xo}x_t + w_{ho}h_{t-1} + w_{co}c_t + b_o\right)$$
$$h_t = o_t \tan\ h\left(c_t\right)$$

where O_t is the output gate and h_t is the output vector.[16]

4.2 At the long term:

4.2.1 FBProphet:

FBProphet is an open-source library developed by Facebook for time series forecasting. It is specifically designed to handle the forecasting tasks with simplicity and efficiency. FBProphet incorporates several key features that make it suitable for predicting cryptocurrency prices:

4.2.2 Main characteristics of FBProphet:

1. Additive Modeling:

FBProphet uses an additive model, where different components of a time series (trend, seasonality, and holidays) are combined to make predictions. This allows for capturing various patterns and dynamics present in the data.

2. Flexible Trend Modeling:

FBProphet automatically detects and models the trend component in the data. It can handle both linear and non-linear trends, enabling it to capture the upward or downward movements in cryptocurrency prices.

3. Seasonality Modeling:

FBProphet can capture seasonal patterns in the data, such as daily, weekly, or monthly fluctuations. It automatically detects and models seasonality, allowing it to adjust the forecasts accordingly.

4. Handling Outliers and Anomalies:

FBProphet is robust to outliers and anomalies in the data. It can handle extreme values or sudden changes in cryptocurrency prices without significantly affecting the overall forecast.

5. Incorporating Domain Knowledge:

FBProphet allows users to include domain-specific knowledge by incorporating additional inputs such as holidays or events that may impact cryptocurrency prices. This flexibility enables the model to consider external factors that can influence the forecasts.

6. Automatic Changepoint Detection:

FBProphet automatically identifies changepoints in the data, representing shifts or changes in the underlying patterns. This helps in capturing trend changes or sudden shifts in cryptocurrency prices.

7. Uncertainty Estimation:

FBProphet provides uncertainty estimation for the forecasts by generating prediction intervals. These intervals give an indication of the potential range of future cryptocurrency prices, taking into account the inherent variability in the data.

It is important to note that while FBProphet is a powerful tool for time series forecasting, its performance may vary depending on the specific characteristics of the cryptocurrency data and the forecasting horizon. It is always recommended to validate and evaluate the forecasts using appropriate evaluation metrics and domain knowledge [17].

5. Main metrics for evaluating the performance of regression prediction models

For each algorithm used for the prediction, it is important to calculated R-square, mean absolute error, and the mean squared error to test the performance the trained model.

5.1 RMSE:

Root Mean Squared Error (RMSE) is the square root of the mean squared error between the predicted and actual values.

$$RMSE = \mathrm{sqrt}\left(\frac{\sum(\text{actual} - \text{prediction})^2}{\text{Number of observations}}\right)$$

5.2 MAE:

Mean Absolute Error MAE is the average absolute error between actual and predicted values.

$$MAE = \mathrm{sqrt}\left(\frac{\sum|\text{actual} - \text{prediction}|}{\text{Number of observations}}\right)$$

5.3 MSE:

Mean Squared Error (MSE) is the average squared error between actual and predicted values**.**

$$MSE = \left(\frac{\sum(\text{actual} - \text{prediction})^2}{\text{Number of observations}}\right)$$

5.4 R-Squared:

R-Squared is a metric for assessing the performance of regression machine learning models

$$R^2 = 1 - \frac{\sum(Y\mathbf{pred} - Y\mathbf{mean})^2}{\sum(Y\mathbf{actual} - Y\mathbf{mean})^2}$$

After finding the R-square, Mean absolute error and mean squared error of all Algorithms we perform comparison and evaluation To check which algorithm gives us the best result. This entire architecture has been applied to all four Cryptocurrencies and in every cryptocurrency.

6. Existing works:

In 2020, Bitcoin gained global recognition as several wealthy individuals such as Elon Musk and companies like Morgan Stanley began discussing it. Since then, Bitcoin has experienced a significant increase in both its price and the number of transactions conducted using it. Researchers have examined various factors, such as social media activity and past trends, to determine the predictability of Bitcoin prices. On the other hand, and because the price data of cryptocurrencies is nonlinear and highly volatile, traditional parametric models have difficulty in effectively tracking and predicting it and they show limited success in tracking and prediction. Therefore, many researchers have tried to study the problem of prediction of price in the crypto market by using of a variety of methods and algorithms.

In their paper, they have reviewed several machine-learning techniques for the estimation of bitcoin pricing. They also analyzed various parameters and its effect on the pricing patterns. As result, it is found out the corresponding cross validation and R-squared scores for various algorithms. At the end, they recommend that a study have to be conducted on sentiment analysis technique on social media posts regarding bitcoins and its effect on the price fluctuations [08].

The authors of [09] made an effort to predict the value of Bitcoin with precision by considering multiple factors that influence it. Their initial goal was to analyze daily trends in the Bitcoin market and determine the most effective features for predicting Bitcoin prices. They collected data on Bitcoin prices and payment networks over a period of five years, recording them on a daily basis. In the second phase of their study, they aimed to predict the direction of daily price changes with the greatest possible accuracy using the data they had gathered.

According to [23], the use of deep neural networks for forecasting cryptocurrency prices has shown promising results in low volatility scenarios. However, due to the nonlinear nature of cryptocurrency price data, traditional parametric models such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model have advantages as well. As a solution, the authors propose a hybrid approach that combines the strengths of both models. They combine the GARCH model's non-stationary parametric approach with the LSTM neural network's nonlinear modeling potential. The practical results of their study demonstrate that their hybrid model has similar predictive performance in terms of MSE, MAE, and RMSE but outperforms in precision, accuracy, and F1 score under optimal hyperparameters. This study suggests that combining parametric models like GARCH with deep neural networks may yield better results in forecasting cryptocurrency prices, especially in highly volatile data or when short data sequences are available.

The authors of [24] proposed three different recurrent neural network (RNN) algorithms to predict the prices of three types of cryptocurrencies: Bitcoin (BTC), Litecoin (LTC), and Ethereum (ETH). They evaluated the models based on mean absolute percentage error (MAPE) and found that the gated recurrent unit (GRU) performed better than the long short-term memory (LSTM) and bidirectional LSTM (bi-LSTM) models for all three types of cryptocurrency. Therefore, GRU was deemed the best algorithm. Specifically, the GRU provided the most accurate prediction for LTC with MAPE percentages of 0.2454%, 0.8267%, and 0.2116% for BTC, ETH, and LTC, respectively. On the other hand, the bi-LSTM algorithm had the lowest prediction results compared to the other two algorithms with MAPE percentages of 5.990%, 6.85%, and 2.332% for BTC, ETH, and LTC, respectively. Overall, the prediction models in this paper produced accurate results that closely align with the actual prices of cryptocurrencies. These models can be valuable for investors and traders in determining cryptocurrency sales and purchases. As a suggestion for future work, the authors recommended investigating other factors that might affect cryptocurrency prices, such as social media, tweets, and trading volume.

The aim of authors in paper [18] is to investigate whether there is a cointegration between the daily closing price of Bitcoin and five other cryptocurrencies (Ethereum, Ripple, Bitcoin Cash, EOS, and Litecoin) during five different time periods that all ended on April 9, 2019. Two cointegration tests (Engle-Granger and Johansen) and a Vector Error Correction Model (VECM) are used to test the long-term relationship between Bitcoin and these cryptocurrencies. The results of both cointegration tests indicate that Bitcoin is cointegrated with Bitcoin Cash, Ethereum, Litecoin, and Ripple. However, the Johansen test and Engle-Granger method suggest that there is no cointegrating relationship between Bitcoin and EOS. Additionally, the VECM estimation shows that the price of Bitcoin has a statistically significant long-term impact on the prices of Bitcoin Cash, Ethereum, Litecoin, and Ripple.

To achieve accurate predictions through big data analysis, a systematic approach is necessary. In [19], the authors conducted a study that involved the entire data analysis process, from data collection to model evaluation, for cryptocurrency price prediction using the Long Short-Term Memory (LSTM) model. They collected data on three different coin prices directly from CoinMarketCap in near real-time using web scraping. The LSTM model was trained using this data, with the random seed or static seed parameters being varied to determine the optimal conditions for better accuracy. The accuracy of the LSTM model was evaluated using MAE, RMSE, and SMAPE indicators. The experiment found that most of the best candidate parameters were classified at the fixed seed trial in terms of RMSE for Bitcoin, Ethereum, and Litecoin.

7. Synthesis and Discussion:

The presented studies suggest that predicting cryptocurrency prices is challenging due to the nonlinear and volatile nature of the data. Machine learning techniques, such as deep neural networks and recurrent neural networks, have shown promising results in predicting cryptocurrency prices. Combining parametric models like GARCH with deep neural networks can yield better results, especially in highly volatile data or when short data sequences are available. The study of cointegration between Bitcoin and other cryptocurrencies reveals that Bitcoin is cointegrated with Bitcoin Cash, Ethereum, Litecoin, and Ripple.

However, no cointegrating relationship was found between Bitcoin and EOS. The importance of sentiment analysis on social media posts regarding Bitcoin and its effect on price fluctuations is also highlighted as an area for future exploration. Finally, a systematic approach to data analysis, such as the use of LSTM models trained with data collected from CoinMarketCap, can yield accurate predictions of cryptocurrency prices.

Finally, we rae agree to what authors presented in [20], the use of various machine learning algorithms and time series analysis techniques to predict cryptocurrency prices may enhance the prediction results. In fact, the used algorithms include classical machine learning models such as SVM and random forest, as well as deep learning models like LSTM. As result, the adopted that the combination of more technical and sentimental factors has been shown to improve prediction results.

8. Conclusion:

In conclusion, the use of machine learning techniques for predicting cryptocurrency prices has shown great potential in recent years. With the increasing popularity of cryptocurrencies and the volatility of their prices, accurate predictions are becoming increasingly important for investors and traders. Machine learning techniques, such as deep neural networks and recurrent neural networks, have been applied to cryptocurrency price prediction with promising results.

However, it is important to note that the nonlinear and volatile nature of cryptocurrency price data presents a significant challenge for accurate prediction. Therefore, it is essential to consider various factors that influence the prices of cryptocurrencies, such as social media sentiment, trading volume, and other market indicators, when developing machine learning models for cryptocurrency price prediction. Additionally, the combination of parametric models with deep neural networks has been shown to improve the accuracy of predictions, especially in highly volatile data or when short data sequences are available.

In summary, the use of machine learning for cryptocurrency price prediction can provide valuable insights for investors and traders. With further research and development, machine learning techniques can continue to improve the accuracy of cryptocurrency price predictions, ultimately leading to better decision-making in the cryptocurrency market.

Chapter 3

Price Prediction Approach

(Implementation and Results)

1. Introduction:

Cryptocurrency price prediction refers to the practice of using various techniques, models, and data analysis methods to forecast the future prices of cryptocurrencies such as Bitcoin, Ethereum, or any other digital assets in the cryptocurrency market. Cryptocurrency price prediction should be approached with caution and used as one of several tools for decision-making in the cryptocurrency market.

2. Cryptocurrencies price prediction Approach:

The main steps in the machine learning process include collecting and preparing data, selecting an appropriate model, training the model on the training data, evaluating the model's performance, and visualizing the results.

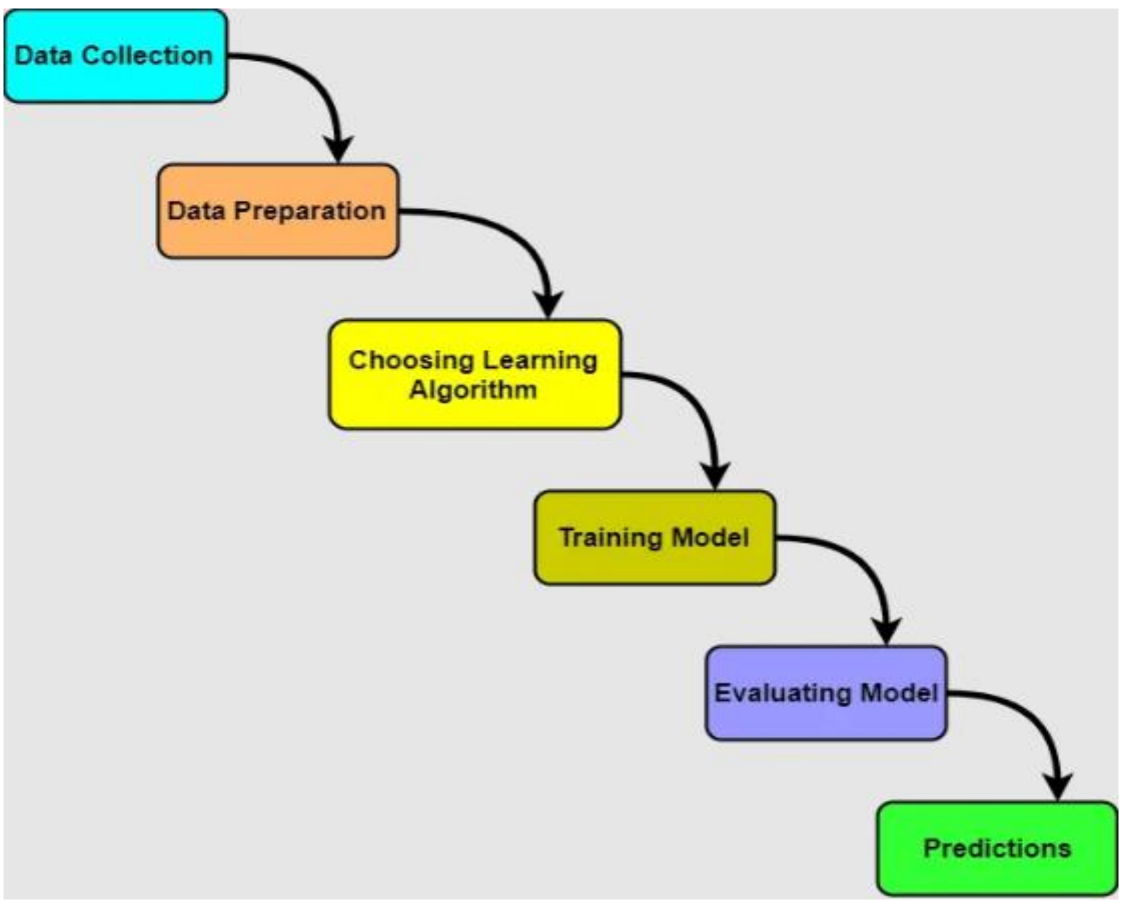

Fig.3.1 Steps price prediction approach.

3. Detailed steps:

3.1 Data collection and preprocessing:

The data set is taken from yahoo finance (yfinance.com) which contains the following features: date, open, close, high, low and volume from 2014 to the present day. However, the data about, market cap, volume and price from the date of 2013 to May 2023 is taken from coindgecko (coindgecko.com). Such historical data can be described as follows:

- Date: The day where recordings were taken.
- Price: actual coin price.
- High: The highest price on a given date.
- Low: The lowest price on a given date.
- Volume: The amount of cryptocurrency transactions in one day.
- Marketcap: The total value of the cryptocurrency market in one day.

The following table can illustrate the general description of our dataset:

Date	High	Low	Volume	Marketcap	Price
2014-09-17	468,174011230468	452,42199707031	21056800	6037835956,610002	454,6238000000001
….	….	….	….	….	….
2023-05-16	27299,3046875	26878,947265625	12732238816	527655229433,8867	27227,793422558792

Tab.3.1. Features description

✓ **Feature selection:**

In order to select the appropriate features, we have measured the Pearson coefficient between these features and the target. Results are given in the following table.

Feature	Target	Correlation
High	price	0.99
Low	price	0,99
Volume	price	0,71
Market cap	price	0,99

Tab.3.2. Features correlation

3.2 Model Selection and Training:

Once we finished pre-processing the data for each cryptocurrency, we first split the dataset into two parts: a training containing all the data minus 365, and a test containing the last 365 days of data. The second step is to apply machine learning algorithms to train the ML cryptocurrency model. Therefore, we applied two different types of Fb Prophet machine learning, Long Term Memory (LSTM).

In the case of LSTM, we used the rolling window; we took 30 days data as an input and the price of the next day as an output (target). However, in the Prophet case we take the date as input and price as output.

3.3 Model evaluation and Results visualizing:

After, the training and tests of the selected machine learning models, we have obtained the following results.

- ✓ **<u>Bitcoin</u>**

Algorithm	**R-square**	**MSE**	**RMSE**	**MAE**
Prophet	**0,91**	**22526066,79**	**4746,16**	**3358,68**
LSTM	**0,98**	**996232,31**	**998,11**	**714,17**

Tab.3.3. Bitcoin models evaluation metrics

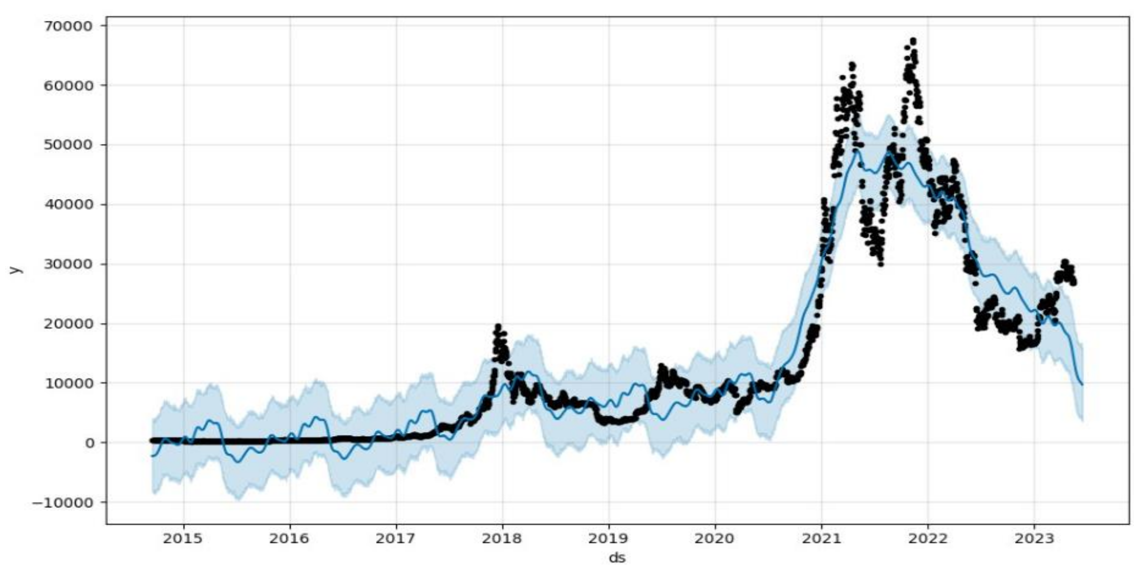

Fig.3.2. Real and predicted price of Bitcoin using prophet

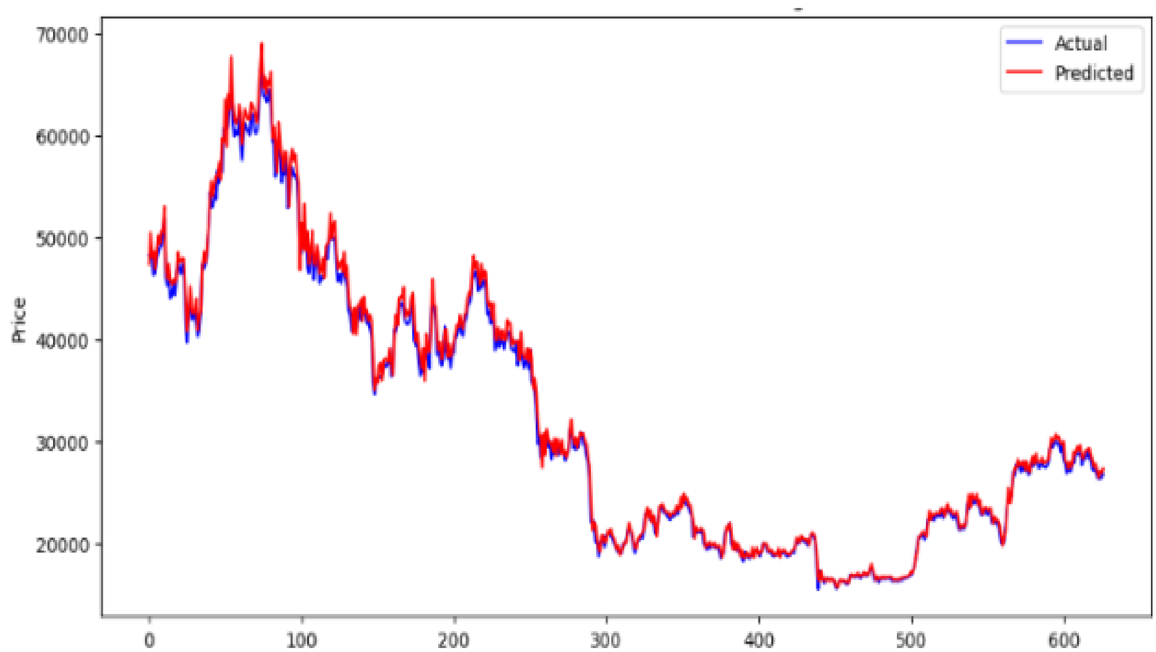

Fig.3.3. Real and predicted price of Bitcoin using LSTM

✓ **Ethereum**

Algorithm	R-square	MSE	RMSE	MAE
prophet	**0,94**	**76164,01**	**275,97**	**202,74**
LSTM	**0,98**	**1758,38**	**41,93**	**29,95**

Tab.3.4. Ethereum models evaluation metrics

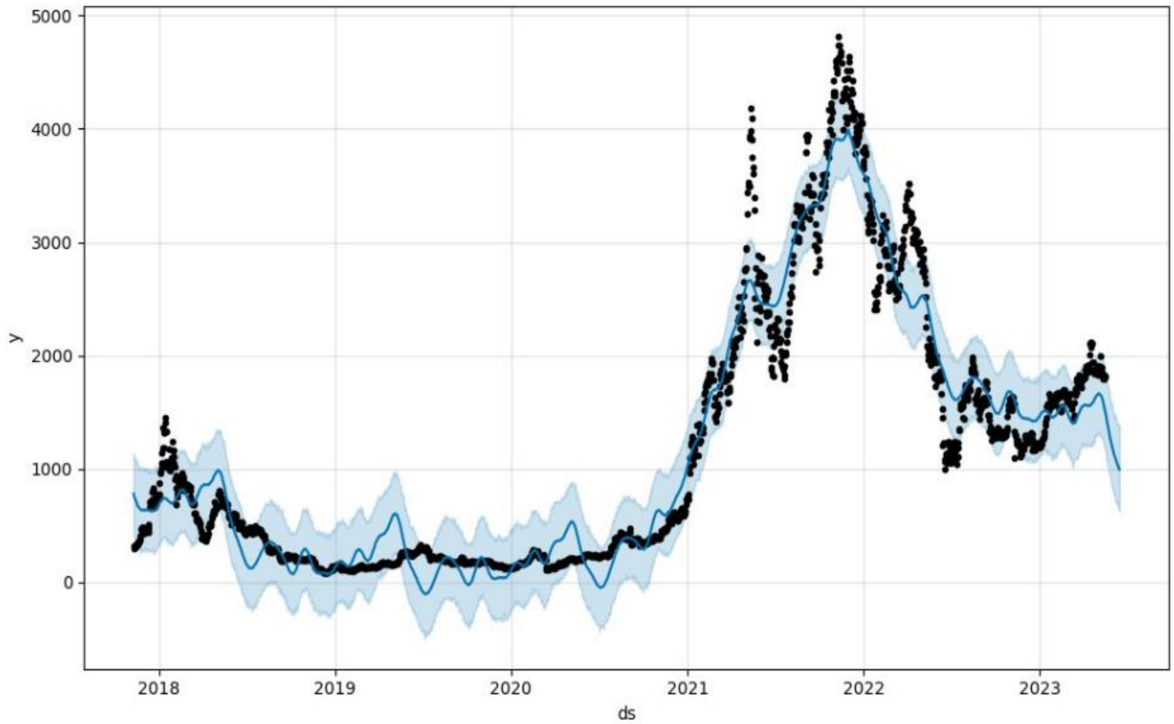

Fig.3.4. Real and predicted price of Ethereum using prophet.

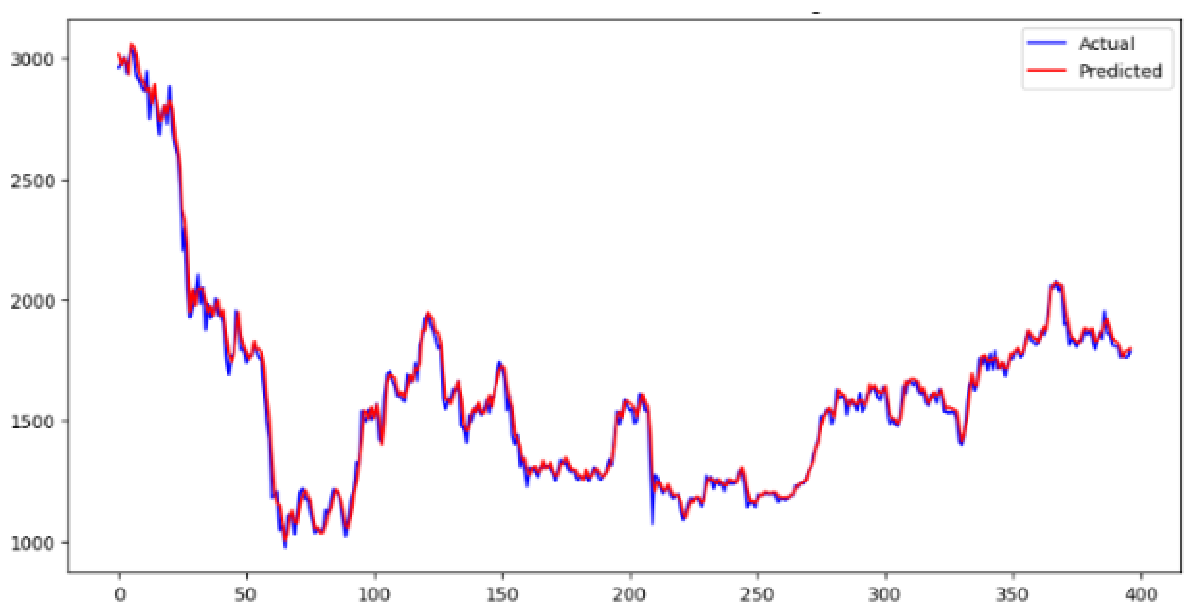

Fig.3.5. Real and predicted price of Ethereum using LSTM

✓ **Binance BNB**

Algorithm	R-square	MSE	RMSE	MAE
prophet	**0,93**	**1978,05**	**44,47**	**30,74**
LSTM	**0,96**	**43,55**	**6,59**	**4,89**

Tab.3.5. Binance BNB models evaluation metrics

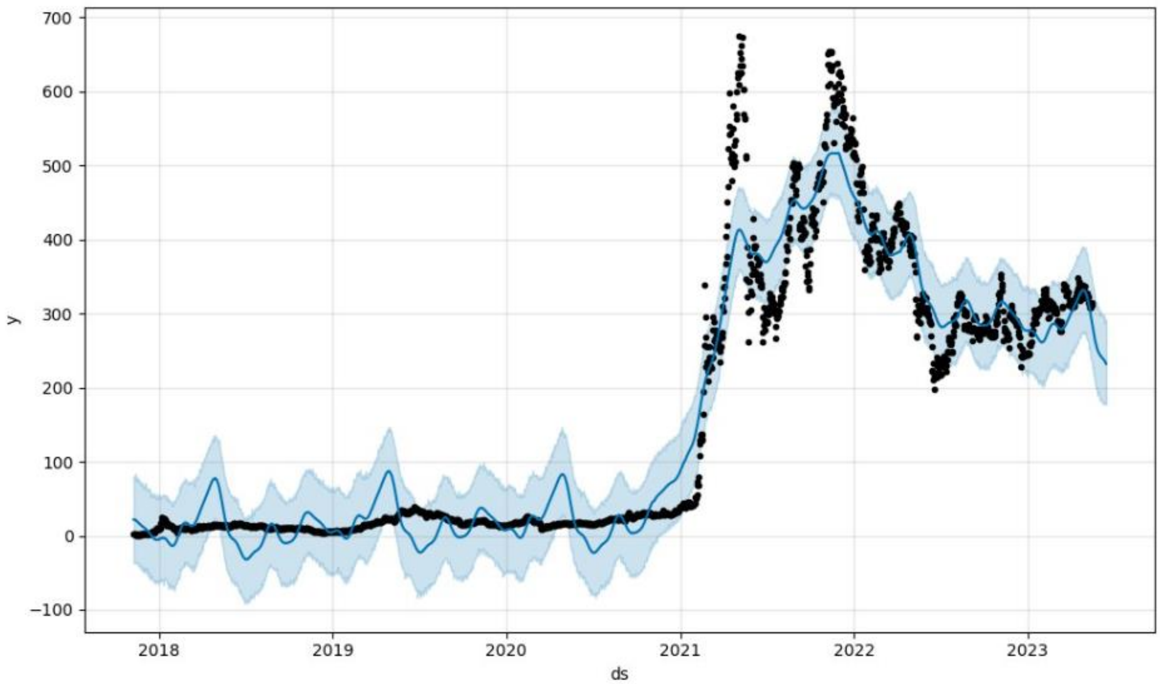

Fig.3.6. Real and predicted price of binance using prophet.

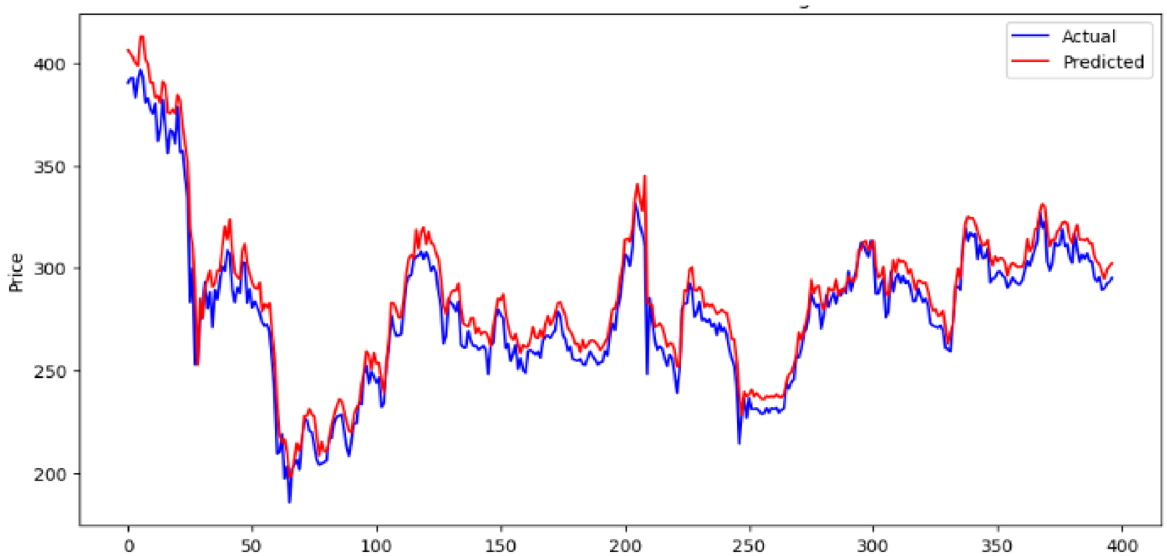

Fig.3.7. Real and predicted price of Binance BNB using LSTM.

✓ **Ripple XRP**

Algorithm	R-square	MSE	RMSE	MAE
Prophet	**0,68**	**0,04008**	**0,2002**	**0,1232**
LSTM	**0,93**	**0,0004**	**0,203**	**0,0196**

Tab.3.6. Ripple models evaluation metrics

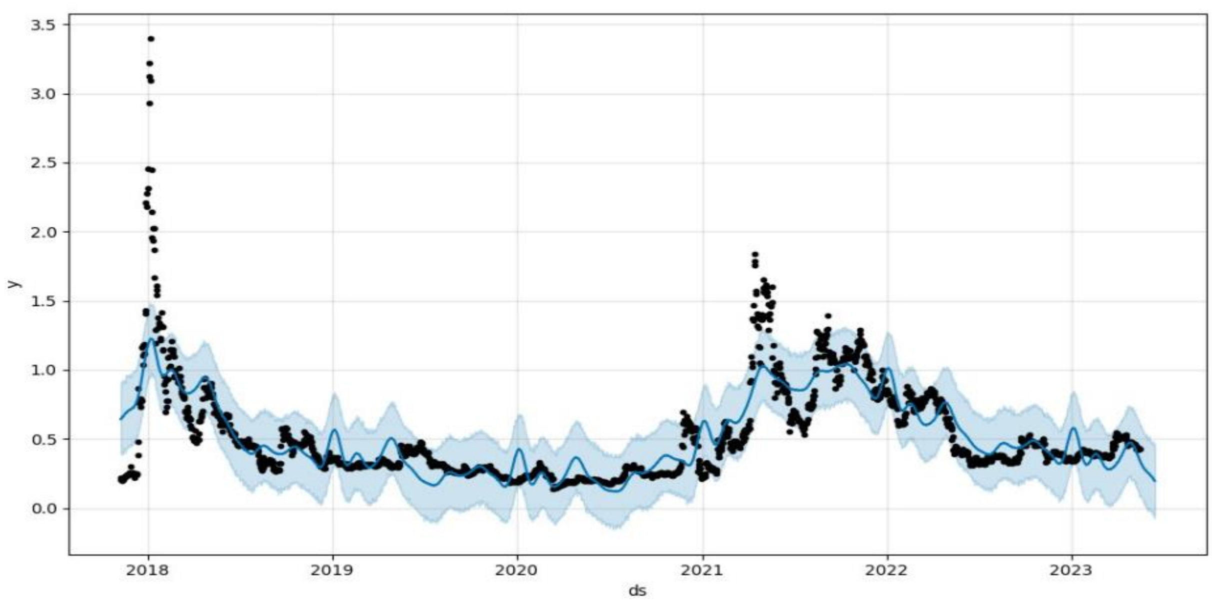

Fig.3.8. Real and predicted price of Ripple using prophet.

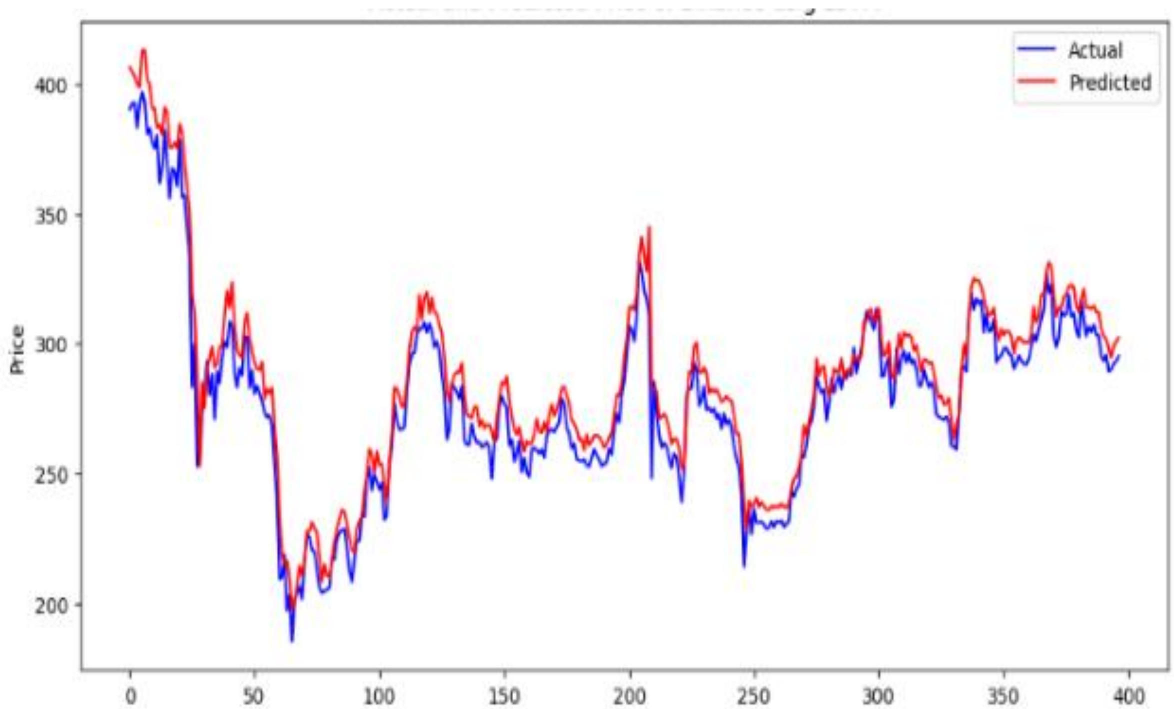

Fig.3.9. Real and predicted price of ripple using LSTM.

After that, we added features as regressors to the Prophet algorithm and we got the following results:

Cryptocurrencies	R-square		MSE		RMSE		MAE	
	Uni-var	With regressors	Uni-var	With regressors	Uni-var	With regressors	Uni-var	With regressors
Bitcoin	0,91	0,98	22526066	299199.21	4746,1	546.99	3358,6	518.51
Ethereum	0,94	0,80	76164,01	13943.32	275,97	118.08	202,74	112.53
Binance	0,93	0,76	1978,05	1441.06	44,47	37.96	30,74	31.44
Ripple	0,68	0,51	0,04	0.0015	0,2002	0.0388	0,1232	0.0350

Tab.3.7. Evaluation metrics of prophet models

4. Results and Discussion

Machine learning-based cryptocurrency price prediction has gained popularity as a research topic in recent years. Despite some encouraging results, it is crucial to know that cryptocurrency prices are notoriously unstable and can be impacted by a variety of variables, such as developments in the global economy, changes in regulations, and technological improvements.

In this context and although we have obtained a good results for cryptocurrencies price prediction in both short and long term by using the corresponding machine learning models, we still think that many other features must be studied and used for this purposes like: Sentiment analysis, Blockchain analytics, etc.

Finally, the application of machine learning techniques to cryptocurrency price forecasting is still in its early stages, but there is significant potential for future growth and development in this area.

5. Conclusion

In conclusion, predicting the prices of cryptocurrencies is challenging due to their volatility, the many factors that affect their value, the limitations of technical and fundamental analysis, the speculative nature of the market, and the lack of regulation. It is important to approach price predictions with caution and consider multiple sources of information before making investment decisions.

Conclusion

Conclusion

In this study, we have used two popular machine learning algorithms, namely LSTM and Prophet, to predict the short and long-term prices of several cryptocurrencies, including Bitcoin, Ethereum, Binance BNB, and Ripple XRP. Our technical analysis was based on a dataset dating back to 2015 for most of the cryptocurrencies studied. The results of our application showed a prediction accuracy of up to 98% in training.

To assess the reliability of the algorithms, we used well-known metrics for regression problems, such as MAE, RMSE, MSE, and R-square. The results were encouraging and showed that LSTM outperformed Prophet in predicting the price of cryptocurrencies with very high accuracy. Finally, this study highlights the potential of machine learning algorithms in predicting cryptocurrency prices and the importance of utilizing appropriate metrics to evaluate their reliability.

References

[01]: Yin, L., Nie, J., & Han, L. (2021). Understanding cryptocurrency volatility: The role of oil market shocks. *International Review of Economics & Finance*, *72*, 233-253.

[02]: Dourado, E., & Brito, J. (2014). Cryptocurrency. The New Palgrave Dictionary of Economics. Online Edition..

[03]: AlRayhi, N. (2020). Cryptocurrency Exchange Market Prediction and Analysis Using Data Mining and Artificial Intelligence (Doctoral dissertation, The British University in Dubai (BUiD)).

[04]: Antonopoulos, A. M. (2014). Mastering Bitcoin: unlocking digital cryptocurrencies. " O'Reilly Media, Inc.".

[05]: Deniz, A., &Teker, D. (2020). Determinants of Cryptocurrency Market: An Analysis for Bitcoin, Ethereum and Ripple. International Journal of Business and Social Science, 11(11), 36-45.

[06]: McNally, R. C. (2018). Cryptocurrency Price Predictions Using Machine Learning: An Empirical Study. International Journal of Forecasting, 34(5), 854-865. doi: 10.1016/j.ijforecast.2018.04.003.

[07]: Tanwar, S., Patel, N. P., Patel, S. N., Patel, J. R., Sharma, G., & Davidson, I. E. (2021). Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. IEEE Access, 9, 138633-138646.

[08]: Parikh, R., Sangle, Y., Kulkarni, V., & Sutaria, J. (2021, December). A Review on Machine-Learning Techniques for Estimation of Bitcoin Pricing. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 1493-1495). IEEE.

[09]: 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.

[10]: Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2(3), 160.

[11]: Yarali, A. (2022). Applications of Artificial Intelligence, ML, and DL.

[12]: LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[13]: Wei, W. W. S. (2006). Time series analysis: univariate and multivariate methods. PearsonEducation.

[14]: Hochreiter, S., &Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[15]: Kim, G., Shin, D. H., Choi, J. G., & Lim, S. (2022). A Deep Learning-Based Cryptocurrency Price Prediction Model That Uses On-Chain Data. IEEE Access, 10, 56232-56248.

[16]: Zhu, G., Peng, S., Lao, Y., Su, Q., & Sun, Q. (2021). Short-term electricity consumption forecasting based on the EMD-fbprophet-LSTM method. Mathematical Problems in Engineering, 2021, 1-9.

[17]: Taylor, S. J., &Letham, B. (2017). Forecasting at scale. The American Statistician, 72(1), 37-45.

[18]: Göttfert, J. (2019). Cointegration among cryptocurrencies: A cointegration analysis of Bitcoin, Bitcoin Cash, EOS, Ethereum, Litecoin and Ripple.

[19]: Kim, J., Wimmer, H., Liu, H., & Kim, S. (2021, September). A Streaming Data Collection and Analysis for Cryptocurrency Price Prediction using LSTM. In 2021 IEEE/ACIS 6th International Conference on Big Data, Cloud Computing, and Data Science (BCD) (pp. 45-52). IEEE.

[20]: MALLICK, S. K. (2020). Causal relationship between Crypto currencies: An Analytical Study between Bitcoin and Binance Coin. Journal of Contemporary Issues in Business and Government| Vol, 26(2), 2172

[21]: Härdle, W. K., Harvey, C. R., & Reule, R. C. (2020). Understanding cryptocurrencies. Journal of Financial Econometrics, 18(2), 181-208.

[22]: Vujičić, D., Jagodić, D., & Ranđić, S. (2018, March). Blockchain technology, bitcoin, and Ethereum: A brief overview. In 2018 17th international symposium infoteh-jahorina (infoteh) (pp. 1-6). IEEE.

[23]: Gao, Z., He, Y., & Kuruoglu, E. E. (2021, October). A Hybrid Model Integrating LSTM and Garch for Bitcoin Price Prediction. In 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). IEEE.

[24] : Hamayel, M. J., & Owda, A. Y. (2021). A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. AI, 2(4), 477-496.

Content

List of Figures

List of Tables