Using Machine Learning for Predicting Arbitrage Occurrences in Cryptocurrency Exchanges

Author Name
Affilliation
Affiliation
City and Country
email

***Abstract*—Cryptocurrency arbitrage, a riskless trading strategy, can yield profits but requires swift execution due to volatile opportunities that vanish rapidly. Utilizing arbitrage bots for algorithmic trading is essential for immediate trade execution across exchanges like Binance and Bybit. This study implements such a system focusing on BTCUSDT and ETHUSDT pairs. Integrating Machine Learning (ML) aims to predict arbitrage occurrences in advance for faster trade execution, a tactic overlooked by many traders. Logistic Regression, Random Forest, Support Vector Machine, and Multilayer Perceptron models are implemented. The addition of Machine Learning principles required the collection of a dataset with historical prices of the observed cryptocurrency pairs for various time intervals, on which we trained the model. Afterward, the model was evaluated in a live-trading environment. Results show Random Forest predicting exploitable arbitrage intervals ahead for Ether, with ML models more effective during less volatile periods. However, careful consideration is needed as predictions may not always align with market realities, leading to mixed trading outcomes. Furthermore, the training led to a model that is able to predict the occurrence of arbitrage, however, classifying the calculations even more carefully than in reality, resulting in a partially profitable or partially lossy trading strategy depending on the time of day and the current stage of the market.**

***Index Terms*—arbitrage, trading, cryptocurrency, machine learning, exchange, Blockchain**

I. Introduction

As the popularity of cryptocurrency trading grows, the various approaches and trading strategies evolve as well. One of them is arbitrage, which is based on obtaining profit from price discrepancies between exchanges or cryptocurrency pairs. This way of trading is nothing new, but it occurs rather sporadically between fiat currencies. Due to the youth of the cryptocurrency market and the variety of pairs available, an exploitable arbitrage opportunity can develop more often, providing a higher percentage of potential profit [1]. In order to ensure that the profit is captured, the whole process has to be executed simultaneously and frequently [2] making this strategy unfeasible for humans and thus requiring an arbitrage bot to be implemented. This method of algorithmic trading extracts data from a chosen cryptocurrency exchange via its API and available endpoints, providing requested data in JSON or XML format [3].

To provide a more detailed explanation of the arbitrage procedure, it can be described as constantly selling or buying cryptocurrency assets of one exchange or as a part of one pair and simultaneously selling or buying it on the other exchange before the profit disappears [4].

As there are many arbitrage bots creating a competitive environment, an implementation of Machine Learning models can represent a valuable advantage when trying to predict the occurrence of profitable opportunities in advance. Prediction in the field of cryptocurrencies could possibly profit from the implementation of the following algorithms [5]:

- Logistic Regression - classification model that estimates a relationship between a binary outcome and a dataset consisting of any number of independent variables [6],
- Random Forest - algorithm suitable for classification, as well as regression problems based on the average or the most frequently occurring output of the multiple decision trees that the Random Forest consists of [7], [8],
- Support Vector Machine - classification model trying to fit an optimal line to separate observations belonging to different classes [9],
- Multilayer Perceptron - as one of the simplest neural network implementations, it is composed of layers of neurons where the information flows from input through hidden to output layer with corresponding weights to be learned [10].

Various ways of integrating the described Machine Learning algorithms were inspected to find the most profitable one by experimenting with the time interval of the obtained data and values of hyperparameters for each algorithm separately. The contribution of such implementation is particularly in the field of immediate transactions, to capture the potential profit before it disappears. That is exactly what can be the main asset of this study, speeding up the arbitrage search and resulting in a higher number of exploited opportunities in comparison with other trading bots. When Machine Learning algorithms can predict an exploitable arbitrage opportunity even before it happens, the purchase can be timed at the beginning of the considered time interval, being between the first ones to take full advantage of the profit.

II. Background

Arbitrage is known as the process of trading an asset in one market while concurrently making an opposite purchase in another market at a different price [4]. These price discrepancies are typical of highly volatile, unstable, and decentralized markets with low regulations. All conditions mentioned are met by cryptocurrency markets, supported by the emergence of multiple currencies and exchanges, providing a convenient way of trading with the possibility to algorithmize the whole process [1], [2]

The arbitrage opportunity can occur for various reasons that are more typical for cryptocurrencies than conventional currency markets. The reasons can be summarized as follows:

- Efficiency of the market, which is maturing over time but still providing a vast amount of profitable arbitrages [11],
- Price variations across exchanges cause the price to move away from the average price defined by market volume [11],
- The currency spread, or the difference between the bid and ask price, expresses the contrast in willingness to buy or sell a specific asset [12],

- Specific patterns, considering the impact of the time of the day, liquidity, or exchange heists and hacks [11].

There is no general approach to arbitrage spotting, as many aspects can play a significant role. It is advised that trading between exchanges, and specifically Bitcoin for other assets, may be more beneficial [13]. Some arbitrage opportunities might occur as a result of a market crash, but those are very rare [14]. The better principle seems to be reliance on the decentralized nature of Blockchain technology, which makes the fundamentals of cryptocurrencies feasible and supports the emergence of exploitable arbitrages [15].

A. *Bitcoin & Ether*

Cryptocurrencies bring a new dimension to trading. They are becoming more convenient and popular, mainly due to their decentralized and immutable nature [3]. As a result, many cryptocurrencies evolved, but only a small amount of them became popular worldwide and had high trading volume. The purpose of this thesis is to examine cryptocurrencies with high volatility and many executed transactions. Due to these requirements, Bitcoin and Ether are considered as the best possibilities, with USD Tether as their quote asset.

Bitcoin is currently the largest cryptocurrency measured by its market capitalization. It was introduced in 2009 by an anonymous developer known by his pseudonym Satoshi Nakamoto. The rise in price and popularity was dramatic, especially between 2016 and 2021, when the number of people possessing Bitcoin rose from 5 to over 220 million people. The value developed from the starting price of 1$ in February 2011 up to 69000$ in November 2021 [16]. Currently, the prices move between 25000 USDT and 30000 USDT, reflecting the current global economic situation. As another example, Ether was used as the second-largest cryptocurrency, still with significant market capitalization. Its prices are currently moving around 1800 USDT [17].

Both cryptocurrencies mentioned are traded against USDT which stands for USD Tether. Instead of fiat currency, Tether represents a stablecoin that, unlike traditional cryptocurrencies, offers more stability and lower volatility. This can be achieved by pegging to nonvolatile underlying assets such as the US Dollar and still relying on distributed ledger technology [18]. Thus bringing more stability into cryptocurrency trading while keeping trades only between cryptocurrencies, stablecoins are gaining popularity and are beneficial for the purpose of this thesis as well.

B. *Cryptocurrency exchanges*

Cryptocurrency exchanges, known as places where cryptocurrencies can be traded conveniently and quickly, represent a gateway between crypto and fiat currencies [19]. With the number of cryptocurrencies rising, the number of exchanges is rising as well, making it an extensive resource for arbitrages to occur [3].

To algorithmize the process of trading, connecting to an exchange via its API is necessary. This study uses Binance and Bybit exchanges to spot arbitrage possibilities between them. Binance is currently the world's leading Blockchain ecosystem, providing the core infrastructure services for organizing the largest digital asset exchange [20]. With similar goals but less widespread among users, Bybit is used in order to profit from price discrepancies between these two exchanges. Both exchanges are accessed via their API endpoints summarized in their corresponding documentation to provide implementation of algorithmic trading to the public. [21], [22].

III. Related Work

Various studies are dedicated to arbitrage bot trading on the cryptocurrency market. Only a small number of them are trying to promote the strategy by using Machine Learning algorithms to predict a profitable arbitrage opportunity even before it occurs.

According to T. G. Fischer et al. [5] implementation of Logistic Regression or Random Forest could be beneficial. The study highlights that the most significant impact is determined by the minute-binned price and volume obtained and the highest market capitalization of the considered cryptocurrencies to reflect the volatility of the market and meet the minimal liquidation criteria. The choice of Machine Learning algorithms is mainly about the comparison of transparent and black-box approaches. The results have shown that Random Forest outperforms Logistic Regression in every metric utilized for the measurement of profitability.

The study by Madan, I. et al. [23] is dedicated to Bitcoin price predictions and focuses on the relevance of available cryptocurrency features. Again, the choice of a Machine Learning algorithm does not play a crucial role. However, many different features are collected, considering mainly the number of transactions or blocks, transaction fees, or traded volume. For the prediction of the price movement, Binomial Logistic Regression, Random Forest, and Support Vector Machine were implemented and trained on price data for various time intervals. The outcomes favor Binomial Logistic Regression when working with daily data and Random Forest for intraday collection of price data.

Other relevant solutions include various arbitrage bots whose goal is to algorithmize trading strategies that are sensitive to the exact timing of transactions. Although they do not include any Machine Learning strategy, the procedure of arbitrage search and calculation of potential profit is identical and represents an inspiration and verification method for the bot implemented within this thesis. Specifically, there is FreqTrade, which represents a universal trading bot with the possibility to specify one's own trading strategy. Furthermore, this work also contains techniques of Artificial Intelligence summarized in the FreqAI framework, providing possibilities to test and compare various time-series forecasting methods on cryptocurrency price data that are most of the time chaotic and at first glance unpredictable [24]. One of many solutions for a simple arbitrage bot without any Machine Learning principles included provides a clear understanding of the execution process of an arbitrage bot when trading on the futures market and represents a valuable verification of the bot's functionality [25].

IV. Solution Design

To perform algorithmic arbitrage trading, an arbitrage bot has to be implemented. As summed up in III, there are many arbitrage bots that provide users with the possibility to algorithmize the whole process of an arbitrage strategy. In order to differentiate from the rivals waiting for the limited amount of exploitable arbitrages, Machine Learning models are trained and included in the arbitrage bot in order to predict the occurrence of a profitable arbitrage even before it arises. The bot is represented via a Python script that executes arbitrages between two exchanges, namely Binance and Bybit. In order to be able to obtain price data and perform necessary operations, the bot is connected to both exchanges via their APIs[1]. The bot is searching for exploitable arbitrage opportunities for BTCUSDT and ETHUSDT cryptocurrency pairs continuously until its execution is terminated. The search is performed on the futures market as a limit order, which protects the user even more in this mostly risk-less trading strategy

[1]Format of queries is inspired by the following repositories: https://github.com/CryptoFacilities/REST-v3-Python/blob/master/cfRestApiV3.py, https://github.com/bybit-exchange/api-usage-examples/blob/master/V3_demo/api_demo/contract/Encryption_HMAC.py

because the trade is not executed immediately; instead, it is waiting for the market to meet the desired price. The high-level architecture of the execution is provided in Figure 1 and provides two approaches when searching for arbitrage opportunities. One is composed of a bare calculation of the potential profit, and the other one takes advantage of Machine Learning integration. In both cases, the current price data are gathered from both exchanges; the difference is in the time at which an arbitrage is assumed to occur. When a chosen algorithm is included, the execution is postponed by the time interval specified by the used dataset. Otherwise, the corresponding trades are executed immediately in order to capture the profit.

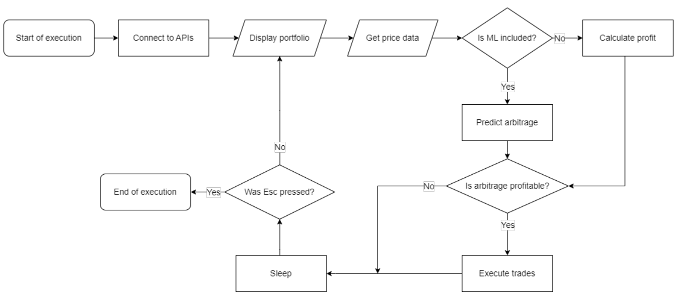

Fig. 1. Architecture

To be able to decide whether a purchase should be executed or not, the potential profit of the sequence of transactions has to be calculated. When considering profit in terms of arbitrage trading, the current stage of the order book and transaction fees on both exchanges have to be taken into consideration. The process of calculation consists of obtaining the traded amount of the base asset, getting the price difference defined by the bid-ask spread, and deducting transaction fees. In order to keep transaction fees as low as possible, all traded assets are held on both exchanges; thus, no transactions between exchanges are necessary, and naturally, no withdraw or deposit fees are included. The calculation proceeds as follows [25]:

Traded amount of base asset

$$quote = quote/((1 + fee) * price) \tag{1}$$

$$amount = min(bid, ask, base, quote) \tag{2}$$

where:

bid	= amount available on the bid side of the order book
ask	= amount available on the ask side of the order book
$base$	= amount of base assets available in the portfolio
$quote$	= amount of quote assets available in a portfolio
fee	= transaction fee of exchange where buying
$price$	= ask price from the order book

Transaction fees

$$fee = taker * amount * ask + taker * amount * bid \tag{3}$$

where:

$taker$	= transaction fee as a percentage of the trade
$amount$	= traded amount of a base asset
ask	= ask price
bid	= bid price

Potential profit

$$profit = |(ask - bid) * amount| - fee \tag{4}$$

$$percentage = profit/ask * 100 \tag{5}$$

where:

ask	= ask price
bid	= bid price
$amount$	= traded amount of a base asset
fee	= transaction fee of exchange where buying
$percentage$	= percentage expression of profit

Besides the described bot's execution, all steps of Machine Learning integration have to be performed. Specifically, consisting of gathering data from both exchanges, data cleaning and pre-processing, visualization of the datasets, data normalization, and building the models for all algorithms with evaluation and comparison of them based on the accuracy metrics.

Datasets used for the training of models consist of price data, specifically OHLCV data, consisting of open, high, low, close, and trading volume for the corresponding datetime. Various time intervals are tested, namely 1, 5, and 15 minutes, to find the interval that is the best reflection of the cryptocurrency's volatility. The data are gathered for the past half year to be able to spot current trends on the market uninfluenced by the historical development of the currency. In order to be able to simulate the potential arbitrages in the past, the datasets are gathered for Binance and Bybit separately.

As a part of the pre-processing, deletion of null values, handling of extreme outliers, and composition and appending of features representing change and arbitrage are executed. The change represents the percentage movement of the currency since the last collected observation. Arbitrage possibilities are expressed by 1 or 0, representing whether the observations from both exchanges for the specific datetime could form an exploitable arbitrage or not.

V. Implementation and Evaluation

The arbitrage bot was implemented as proposed, continuously searching for arbitrage opportunities at 1-minute intervals. In order to perform the required operations, connection to the APIs of both exchanges is established via an API and secret keys corresponding to the created accounts.

The financial profit obtained strongly depends on the current stage of the cryptocurrency markets and world economics as a whole, seasonality, and time of day. The profitability of the bot was measured from the day 29.04.2023 07:20 until 12.05.2023 07:00, where the price of Bitcoin on Binance fluctuated from 29201.1 to 27846 and on Bybit from 29392.5 to 26000 and the price of Ether on Binance from 1897 to 1860.9 and on Bybit from 1901.7 to 1769.65. The overall change in the portfolio is displayed in Table I, resulting in a 32.11% change in the total balance of USDT on both exchanges. This percentage change represents the rise (or potentially decline) of the total amount of USDT on both exchanges when compared to the stage of the portfolio before and after the run of the arbitrage bot. When expressed in monetary values, the total amount of USDT rose from 26364.721 to 34830.39. As a result of trading on the futures market, the total balance of BTC and ETH was not changed. As already mentioned and also proven by the visualized execution, the arbitrage trading strategy results in profit when the purchases are repeated for a longer period of time without any interruptions.

The inclusion of Machine Learning models required data gathering, data visualization, and data pre-processing. The obtained datasets consisted of OHLCV data with a corresponding timestamp and traded volume for a half year, specifically from 22.9.2022 until 23.3.2023. Datasets were obtained for each traded pair from both exchanges and for time intervals of 1, 5, and 15 minutes. As a part of pre-processing, datasets were cleaned from null values and extreme

TABLE I
PORTFOLIO WITHOUT MACHINE LEARNING

Asset	**Binance**	**Bybit**	**Total**	**Percentage change**
BTC	0.045	1.000	1.045	0.000%
ETH	0.637	0.542	1.179	0.000%
USDT	17867.600	16962.790	34830.390	32.11%

outliers determined by the values out of the range of 0.001 and 0.999 quantile. Later, datasets were merged in pursuit of a congruent traded pair and time interval. Furthermore, the percentage change between two open prices in a row and the identification of arbitrage between exchanges for open and close prices and combinations of high and low prices were appended to the datasets. As a part of the visualization, prices on both exchanges, overall traded volume, and change on exchanges were visualized. Figure 2 shows the development of open prices and signifies that most of the price fluctuations resulting in arbitrage opportunities originate on the Bybit exchange. The quantity of specified ranges of traded volume from both exchanges visualized in Figure 3 proves that in most of the intervals, the traded volume stays very low, and Figure 4 supports the price fluctuations visible from comparisons of open prices as well.

In order to enrich the arbitrage search, Machine Learning models were trained. Specifically, Logistic Regression, Random Forest, Support Vector Machine, and Multilayer Perceptron. The training of each model consisted of gradual testing of individual hyperparameters and subsequently training the model with the best parameters and determining which time interval suits the process of predicting arbitrage the best. Due to the higher number of hyperparameters, large datasets for small time intervals, and negligible differences in individual tests, the set of all possible values for hyperparameters was not searched exhaustively. All tested hyperparameters for each Machine Learning algorithm, with the chosen one highlighted, are summarized in Tables II, III, IV and V. The gathered results based on the proportion of accuracy and F1 score with the most suitable time interval for each cryptocurrency pair are summarized in Table VI.

Subsequently, due to better results, the model for the prediction of ETHUSDT was included in the execution of the arbitrage bot, and the profitability was measured again on the day 15.05.2023 11:56 until 20.05.2023 11:31, where the price of Ether from Binance fluctuated from 1860.9 to 1861.1 and on Bybit from 1828.37 to 1813.82. The overall change in the portfolio is displayed in Table VII, resulting in a 0.455% change in the total balance of USDT on both exchanges, with identical meaning to the percentage change in the run without Machine Learning inclusion. As can be seen from the obtained results, the profit gain is not that significant, indicating the carefulness of the trained model. However, a still profitable result was observed, which suggests a valuable resource for arbitrage search when the cryptocurrency market matures, and the occurrence of exploitable arbitrage becomes even rare with a higher number of trading bots fighting for limited profit. Although these results have to be compared carefully due to the different time ranges of the execution, namely 13 and 5 days, caused by a shortage of time, the results still provide sufficient evidence of the profitability of both modifications of the arbitrage bot.

VI. DISCUSSION

The results of the training of Machine Learning models signify a relationship between OHLCV price data of a cryptocurrency pair, traded volume, percentage change, and the occurrence of exploitable arbitrage. When considering the differences between individual Machine Learning models, the best results were obtained from Random Forest, but being aware of the potential overfitting and thus preferring a 5-minute time interval, as a 1-minute interval is prone to learning the dataset by heart. The main difference was identified between the considered cryptocurrency pairs. The more volatile ones are much harder to predict regardless of the considered time interval. As Figure 2 shows, the pair ETHUSDT is much more stable for the selected time range of the dataset, resulting in significantly better results for all considered metrics for the implemented models. On the other hand, less volatile cryptocurrencies create fewer arbitrage opportunities, resulting in lower available profit. When all combined, Random Forest for a 5-minute time interval and trained on the ETHUSDT cryptocurrency pair leads to the best evaluation with the accuracy of 0.985, precision of 0.635, recall of 0.836, and an F1 score of 0.722.

The study focused on the integration of Machine Learning algorithms into an arbitrage trading strategy executed between cryptocurrency exchanges. From all the models described above, Random Forest performed the best when used for a 5-minute time interval in the dataset. When integrated into a real trading strategy, the approaches with and without the inclusion of the Machine Learning model were compared, resulting in a profit of 32.11% without and 0.455% with the Machine Learning model. Although these results were obtained during different time ranges, both numbers support the profitable nature of the modifications to the implemented arbitrage bot.

TABLE II
HYPERPARAMETERS FOR LOGISTIC REGRESSION

Logistic Regression					
penalty	l1	l2			
C	0.01	0.1	1	10	100
max_iter	500	1000	1500		

TABLE III
HYPERPARAMETERS FOR RANDOM FOREST

Random Forest			
n_estimators	50	100	150
criterion	gini	entropy	
min_samples_split	2	10	50
max_features	sqrt	log2	None
bootstrap	True		

TABLE IV
HYPERPARAMETERS FOR SUPPORT VECTOR MACHINE

Support Vector Machine					
C	0.01	0.1	1	10	100
penalty	l1	l2			
max_iter	500	1000	1500		
kernel	linear				

TABLE V
HYPERPARAMETERS FOR MULTILAYER PERCEPTRON

Multilayer Perceptron			
activation	logistic	relu	tanh
solver	sgd	adam	
learning_rate_init	0,001	0.01	0.1
max_iter	100	150	200
hidden_layer_sizes	(5, 5)		

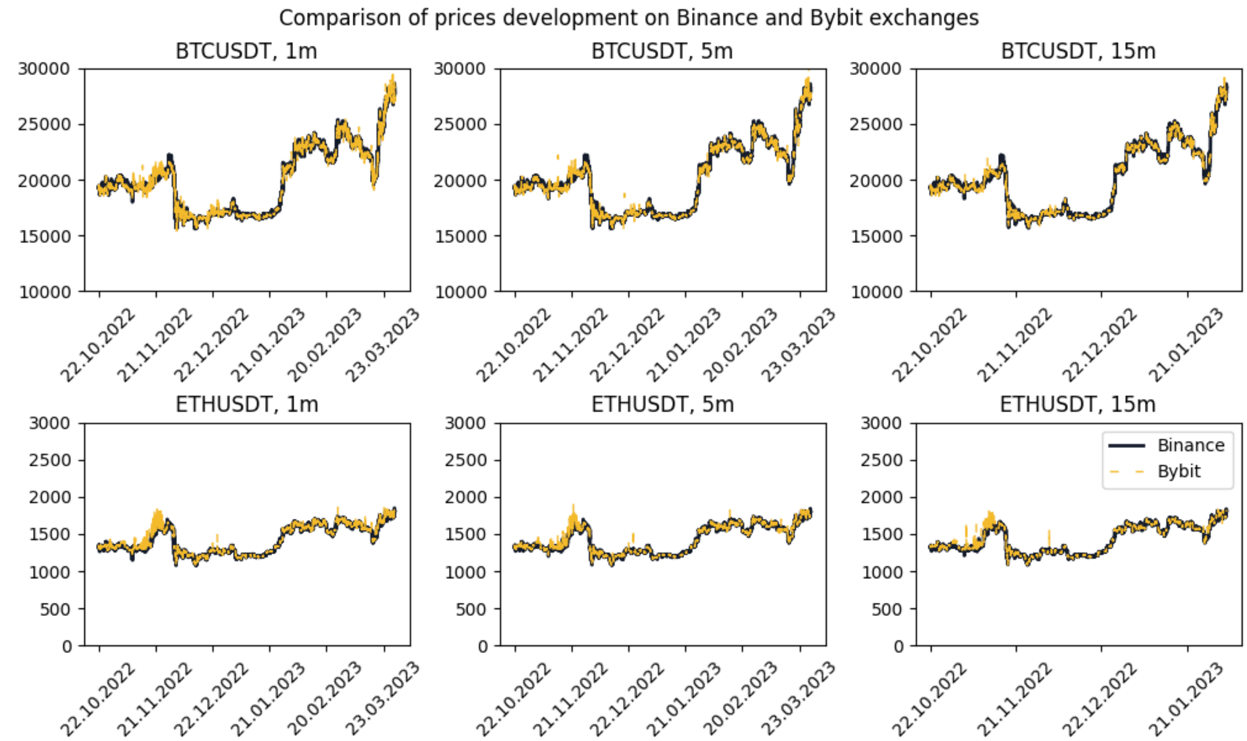

Fig. 2. Visualization of open prices

TABLE VI
MACHINE LEARNING MODELS

Machine Learning model	Cryptocurrency	Time interval	Accuracy	Precision	Recall	F1 score
Logistic Regression	BTCUSDT	1 minute	0.981	0.135	0.814	0.232
	ETHUSDT	1 minute	0.975	0.433	0.963	0.597
Random Forest	BTCUSDT	5 minutes	0.985	0.284	0.617	0.389
	ETHUSDT	5 minutes	0.985	0.635	0.836	0.722
Support Vector Machine	BTCUSDT	1 minute	0.986	0.179	0.793	0.292
	ETHUSDT	1 minute	0.974	0.426	0.965	0.591
Multilayer Perceptron	BTCUSDT	1 minute	0.984	0.175	0.899	0.292
	ETHUSDT	5 minutes	0.973	0.466	0.864	0.605

TABLE VII
PORTFOLIO WITH MACHINE LEARNING

Asset	Binance	Bybit	Total	Percentage change
BTC	0.045	1.000	1.045	0.000%
ETH	0.637	0.316	0.953	0.000%
USDT	17986.879	17002.060	34888.939	0.455%

In comparison with related work, the obtained results were similar in some aspects but different in others. Equally, to T. G. Fischer et al. [5], Random Forest outperformed Logistic Regression; however, with datasets consisting of nearly the same price data, it thrived better with a 5-minute interval in comparison to the 1-minute interval from the related work. Similarly, Madan et al. [23] considered Random Forest to be one of the more successful models together with Logistic Regression, whose results were not that influential in this study. This difference could be explained by the diametrically different features and time intervals considered. In general, it is evident that Random Forest could reflect the price features gathered in the datasets with the best results measured by various Machine Learning metrics.

VII. CONCLUSION AND FUTURE WORK

To sum up, it was demonstrated that a combination of Machine Learning algorithms and an arbitrage trading strategy between cryptocurrency exchanges can provide interesting outcomes. Multiple Machine Learning models with multiple parameters were tested. Random Forest with a 5-minute time interval provided the best result for both considered cryptocurrency pairs, namely BTCUSDT and ETHUSDT. Both approaches to arbitrage trading with and without the inclusion of Machine Learning algorithms were tested and provided profitable results. The key novelty of the study is precisely the demonstration of the profitability of the prediction of exploitable arbitrage one-time interval beforehand, making it a valuable resource to outrun the rivals represented by an increasing number of existing arbitrage bots.

Furthermore, to obtain the maximal relationships from the datasets, the possibility of arbitrage occurrence appended to the dataset could be replaced by an actual observation of whether trades leading to arbitrage were executed by obtaining a historical order book from the involved cryptocurrency exchanges. Such an analysis of data could lead to more precise results, even in volatile crypto markets. The inclusion of less-known, more unstable currencies could also

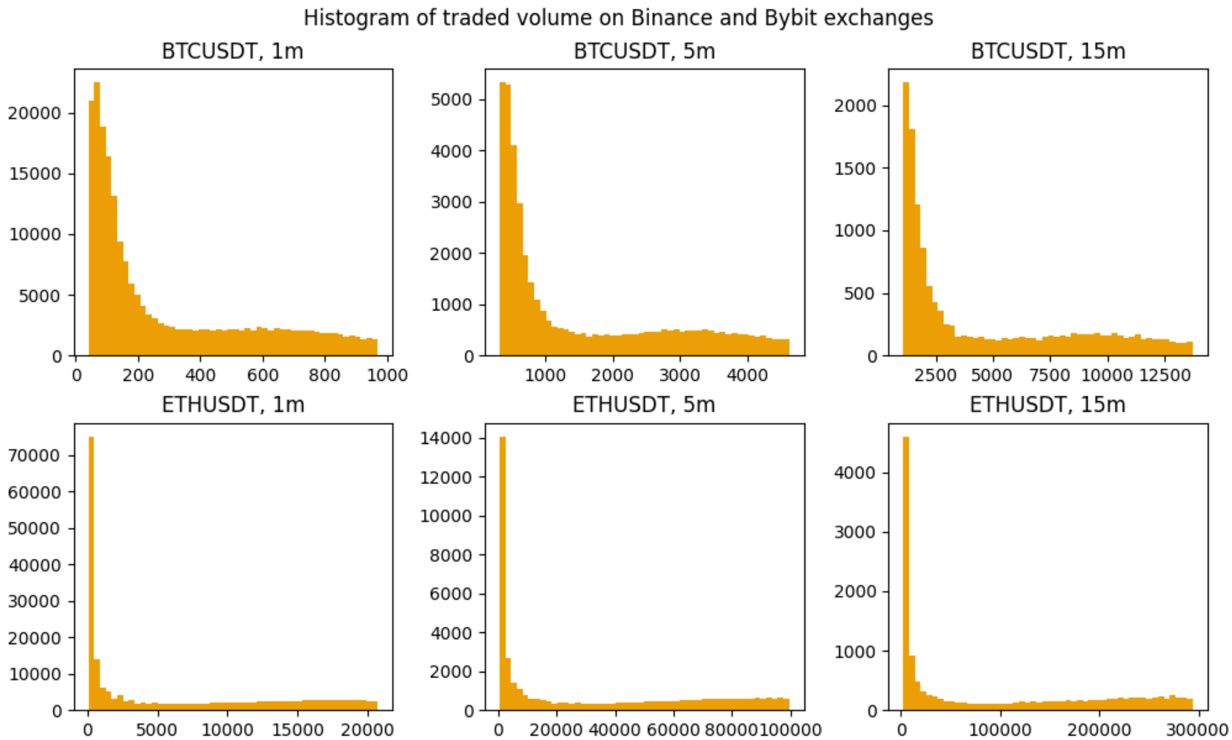

Fig. 3. Visualization of traded volume

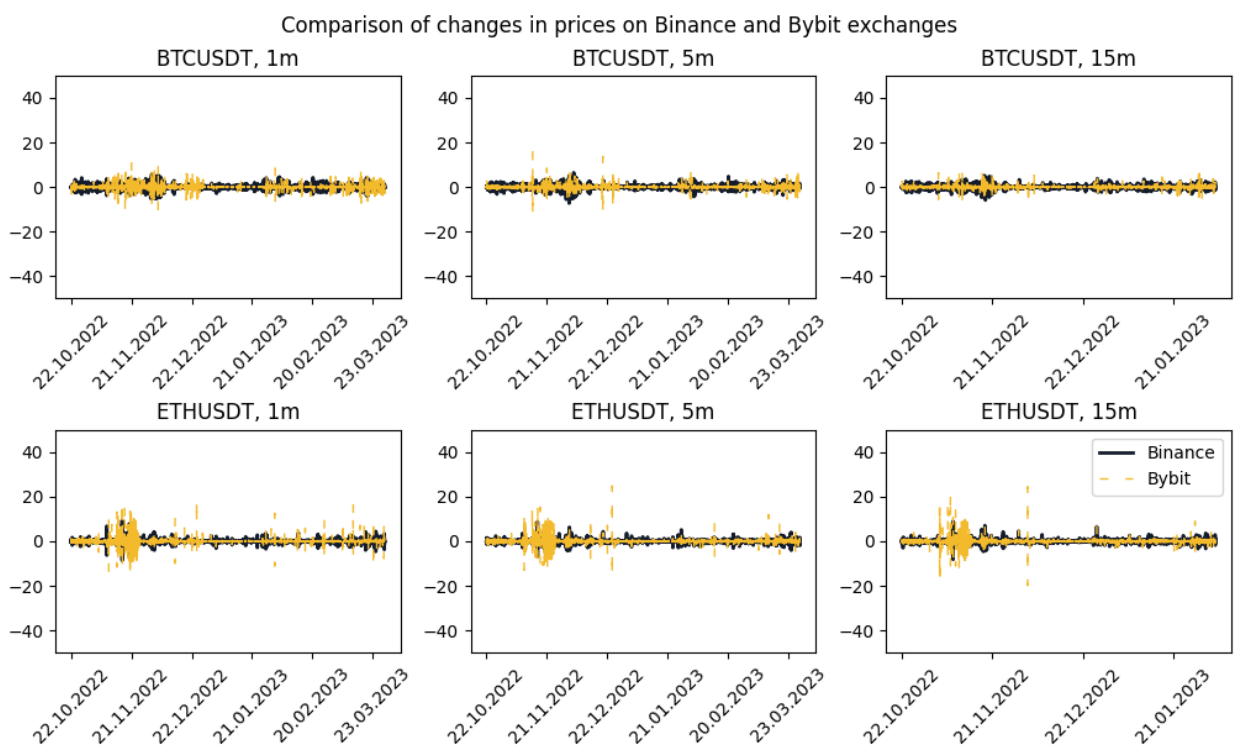

Fig. 4. Visualization of price change

influence the results due to the more frequent occurrence of arbitrage possibilities, however, making it harder to predict exploitable arbitrages. Lastly, the implementation of more complex neural networks could be beneficial for the revelation of the relationships between the features in the datasets.

References

[1] R. Levus, A. Berko, L. Chyrun, and V. Panasyuk, "Intelligent System for Arbitrage Situations Searching in the Cryptocurrency Market," p. 34.

[2] G. Peduzzi, J. James, and J. Xu, "Jack the Rippler: Arbitrage on the Decentralized Exchange of the XRP Ledger," in 2021 3rd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS), Sep. 2021, pp. 1–2. doi: 10.1109/BRAINS52497.2021.9569833.

[3] V. Lytvyn et al., "Development of the system to integrate and generate content considering the cryptocurrent needs of users," EEJET, vol. 1, no. 2, pp. 18–39, Jan. 2019, doi: 10.15587/1729-4061.2019.154709.

[4] L. Zhou, K. Qin, and A. Gervais, "A2MM: Mitigating Frontrunning, Transaction Reordering and Consensus Instability in Decentralized Exchanges." arXiv, Jun. 17, 2021. Accessed: Aug. 17, 2022. [Online]. Available: http://arxiv.org/abs/2106.07371

[5] T. G. Fischer, C. Krauss, and A. Deinert, "Statistical Arbitrage in Cryptocurrency Markets," Journal of Risk and Financial Management, vol. 12, no. 1, Art. no. 1, Mar. 2019, doi: 10.3390/jrfm12010031.

[6] P. Schober and T. R. Vetter, "Logistic Regression in Medical Research," Anesth Analg, vol. 132, no. 2, pp. 365–366, Feb. 2021, doi: 10.1213/ANE.0000000000005247.

[7] E. Akyildirim, A. Goncu, and A. Sensoy, "Prediction of cryptocurrency returns using machine learning," Ann Oper Res, vol. 297, no. 1, pp. 3–36, Feb. 2021, doi: 10.1007/s10479-020-03575-y.

[8] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," The Stata Journal, vol. 20, no. 1, pp. 3–29, Mar. 2020, doi: 10.1177/1536867X20909688.

[9] "Chapter 6 - Support vector machine — Elsevier Enhanced Reader." https://reader.elsevier.com/reader/sd/pii/B9780128157398000067?token=15264E02361C2B29E22D841989A0F94B16BB5A6BAEC5BF69FF0\C665906B85A75C8739C7A8992AF2253AE1C3A6A9AC6DD&originRegion=eu-west-1&originCreation=20221127170227 (accessed Nov. 27, 2022).

[10] H. Taud and J. F. Mas, "Multilayer Perceptron (MLP)," in Geomatic Approaches for Modeling Land Change Scenarios, M. T. Camacho Olmedo, M. Paegelow, J.-F. Mas, and F. Escobar, Eds., in Lecture Notes in Geoinformation and Cartography. Cham: Springer International Publishing, 2018, pp. 451–455. doi: 10.1007/978-3-319-60801-3_27.

[11] S. Krückeberg and P. Scholz, "Decentralized Efficiency? Arbitrage in Bitcoin Markets," p. 32.

[12] E. Han, "Intra-exchange Cryptocurrency Arbitrage Bot," Master of Science, San Jose State University, San Jose, CA, USA, 2018. doi: 10.31979/etd.6xze-y9xu.

[13] S. Bistarelli, A. Cretarola, G. Figà-Talamanca, and M. Patacca, "Model-based arbitrage in multi-exchange models for Bitcoin price dynamics," Digit Finance, vol. 1, no. 1–4, pp. 23–46, Nov. 2019, doi: 10.1007/s42521-019-00001-2.

[14] T. Hattori and R. Ishida, "The relationship between arbitrage in futures and spot markets and Bitcoin price movements: Evidence from the Bitcoin markets," J Futures Markets, vol. 41, no. 1, pp. 105–114, Jan. 2021, doi: 10.1002/fut.22171.

[15] A. Kabašinskas and K. Šutienė, "Key Roles of Crypto-Exchanges in Generating Arbitrage Opportunities," Entropy, vol. 23, no. 4, Art. no. 4, Apr. 2021, doi: 10.3390/e23040455.

[16] R. Auer et al., "Crypto Trading and Bitcoin Prices: Evidence from a New Database of Retail Adoption," SSRN Journal, 2023, doi: 10.2139/ssrn.4357559.

[17] Y. Tan, Z. Wu, J. Liu, J. Wu, Z. Zheng, and T. Chen, "Bubble or Not: Measurements, Analyses, and Findings on the Ethereum ERC721 and ERC1155 Non-fungible Token Ecosystem." arXiv, Jan. 05, 2023. Accessed: Mar. 20, 2023. [Online]. Available: http://arxiv.org/abs/2301.01991

[18] B. Łet, K. Sobański, W. Świder, and K. Włosik, "What drives the popularity of stablecoins? Measuring the frequency dynamics of connectedness between volatile and stable cryptocurrencies," Technological Forecasting and Social Change, vol. 189, p. 122318, Apr. 2023, doi: 10.1016/j.techfore.2023.122318.

[19] J. Chen, D. Lin, and J. Wu, "Do cryptocurrency exchanges fake trading volumes? An empirical analysis of wash trading based on data mining," Physica A: Statistical Mechanics and its Applications, vol. 586, p. 126405, Jan. 2022, doi: 10.1016/j.physa.2021.126405.

[20] "Binance - Cryptocurrency Exchange for Bitcoin, Ethereum & Altcoins," Binance. https://www.binance.com (accessed Mar. 20, 2023).

[21] "Change Log – Binance API Documentation." https://binance-docs.github.io/apidocs/futures/en/#change-log (accessed Mar. 28, 2023).

[22] "Introduction — Bybit API Documentation." https://your-docusaurus-test-site.com/docs/v5/intro (accessed Mar. 28, 2023).

[23] I. Madan, S. Saluja, and A. Zhao, "Automated Bitcoin Trading via Machine Learning Algorithms," p. 5.

[24] R. A. Caulk et al., "FreqAI: generalizing adaptive modeling for chaotictime-series market forecasts," JOSS, vol. 7, no. 80, p. 4864, Dec. 2022, doi: 10.21105/joss.04864.

[25] K. Au, "Crypto Arbitrage." Mar. 22, 2023. Accessed: Mar. 29, 2023. [Online]. Available: https://github.com/kelvinau/crypto-arbitrage/blob/2f8956fe37b62002985edfba84006f19490697de/engines/exchange_arbitrage.py