Article

Enhancing Cryptocurrency Price Forecasting by Integrating Machine Learning with Social Media and Market Data

Loris Belcastro *[ID], **Domenico Carbone** [ID], **Cristian Cosentino** [ID], **Fabrizio Marozzo** *[ID] and **Paolo Trunfio** [ID]

Department of Informatics, Modeling, Electronics and Systems Engineering, University of Calabria, 87036 Rende, Italy; dcarbone@dimes.unical.it (D.C.); ccosentino@dimes.unical.it (C.C.); trunfio@dimes.unical.it (P.T.)
* Correspondence: lbelcastro@dimes.unical.it (L.B.); fmarozzo@dimes.unical.it (F.M.)

Abstract: Since the advent of Bitcoin, the cryptocurrency landscape has seen the emergence of several virtual currencies that have quickly established their presence in the global market. The dynamics of this market, influenced by a multitude of factors that are difficult to predict, pose a challenge to fully comprehend its underlying insights. This paper proposes a methodology for suggesting when it is appropriate to buy or sell cryptocurrencies, in order to maximize profits. Starting from large sets of market and social media data, our methodology combines different statistical, text analytics, and deep learning techniques to support a recommendation trading algorithm. In particular, we exploit additional information such as correlation between social media posts and price fluctuations, causal connection among prices, and the sentiment of social media users regarding cryptocurrencies. Several experiments were carried out on historical data to assess the effectiveness of the trading algorithm, achieving an overall average gain of 194% without transaction fees and 117% when considering fees. In particular, among the different types of cryptocurrencies considered (i.e., high capitalization, solid projects, and meme coins), the trading algorithm has proven to be very effective in predicting the price trends of influential meme coins, yielding considerably higher profits compared to other cryptocurrency types.

Keywords: cryptocurrency; social media data; price prediction; trading recommendation; machine learning; deep learning

1. Introduction

Following the pioneering launch of Bitcoin by Satoshi Nakamoto, a multitude of virtual currencies have seen notable increases in both their value and widespread acceptance. For instance, Ethereum and Litecoin emerged as prominent contenders, with Ethereum's smart contract capabilities driving its rapid ascent, while Litecoin focused on enhancing the transaction speed and scalability, making it a popular choice for everyday transactions. Today, the cryptocurrency landscape boasts thousands of coins, many of which are characterized by volatility and a lack of substantial projects [1]. Their value and volatility are influenced by factors such as popularity obtained through word of mouth on social media. For example, viral memes related to specific coins can drive a wave of interest, prompting people to invest based on the excitement generated by these online phenomena.

In recent years, the cryptocurrency market has witnessed extraordinary success, largely as a result of innovative marketing strategies adopted by exchange platforms [2]. These strategies, encompassing user-friendly interfaces and educational resources, have made cryptocurrency investments more approachable, leading to a surge in market growth. Moreover, influential people have wielded their substantial social media presence to endorse or critique cryptocurrencies through tweets and public statements, contributing to both market volatility and a surge in public interest [3]. Based on the dynamic context of the cryptocurrency market, this paper delves into the analysis of factors that influence price

movements, mainly focusing on market dynamics and content published on social media. Our methodology involves predicting the cryptocurrency price movements by analyzing two large datasets: one comprising market data like prices and trading volumes for specific cryptocurrencies, and the other containing social media posts discussing these coins. These datasets were properly merged and analyzed using a set of statistical, text analytics, and deep learning techniques. Several studies have demonstrated substantial economic benefits for investors by leveraging machine learning predictions, which allow one to predict stock [4], equity risk [5], or bond [6] premiums, compared to a risk-free investments. In particular, in many cases, the use of machine learning techniques surpasses leading regression-based strategies, even doubling their effectiveness in certain cases, especially using decision trees and neural networks [7,8]. The goal of this methodology is to develop a trading recommendation algorithm that can identify optimal moments for buying and selling cryptocurrencies in order to maximize profits.

The proposed methodology is composed of different phases: data collection, data preprocessing, data enrichment, training machine learning models, and trading recommendation. To enhance the predictive capabilities of our methodology, we augmented these datasets with three additional features: (*i*) correlation between social media activities and cryptocurrency price fluctuations; (*ii*) causal connection among prices of cryptocurrencies, so as to consider how they influence each other and the impact that the most popular cryptocurrencies have on the broader market; and (*iii*) the sentiment of users about cryptocurrencies, obtained through a textual analysis of posts published on social media. Such datasets were used to train an long short-term memory (LSTM) model for predicting the prices of cryptocurrencies, supporting a trading recommendation algorithm capable of identifying and exploiting the direction in which the price of a cryptocurrency is moving (e.g., upward or downward trend). Unlike other existing works, our research supports price forecasting for different types of cryptocurrencies (i.e., high capitalization, solid projects, and meme coins), providing excellent market coverage. Our methodology stands out for its comprehensive analysis, encompassing a wide range of features to enhance cryptocurrency price forecasting. This includes evaluating correlations and causal connections among coins, as well as the detection of bot activities that can influence price predictions. Additionally, our solution extensively utilizes text analysis techniques on social media data to extract information about the popularity, overall market perception, and users' sentiment regarding specific cryptocurrencies.

Several experiments were carried out on historical data to assess the effectiveness of the trading algorithm. By investing in a selected set of cryptocurrencies, the trading algorithm achieved an overall average profit of 194% when transaction fees were not taken into account and 117% when transaction fees were considered. Focusing only on influential meme coins, the algorithm resulted in a very high profit of 902% with fees and 1258% without fees.

The structure of this paper is as follows. Section 2 discusses related work in the cryptocurrency field and provides a comparison between our methodology and existing research. Section 3 describes the proposed methodology. Section 4 discusses the achieved results. Finally, Section 5 concludes the paper.

2. Related Work

The cryptocurrency market has become a sector of growing interest, characterized by significant volatility and a wide variety of speculative activities. In this context, the accurate prediction of cryptocurrency prices has played a crucial role for investors, traders, and stakeholders in the financial industry. The use of machine learning and deep learning techniques has emerged as a research avenue to support cryptocurrency price predictions. In this section, we provide an overview of the existing related work, highlighting the most used approaches and challenges in cryptocurrency price forecasting.

Nowadays, the use of machine learning and deep learning techniques for price prediction is widespread [9–11]. Several studies have explored the application of neural networks,

support vector machines, and other supervised learning models to predict cryptocurrency prices over time [12–14]. In particular, the use of methods such as long-short-term memory (LSTM) neural networks has shown particular promise in addressing time series challenges and identifying nonlinear patterns in price fluctuations [15–20]. Some works have exploited linear regression and random forest methods to analyze the model performance and profitability of trading strategies [21,22]. Other studies have proposed novel models based on recurrent neural network (RNN) models, such as GRU, LSTM, and bi-LSTM, to achieve precise cryptocurrency price forecasts [23,24]. Such approaches have provided accurate results and demonstrated their ability to detect complex dynamics, which can support trading decisions [25–27].

Other studies have explored the potential of employing text analysis on social media data to enhance the ability to predict cryptocurrency prices. Specifically, studies conducted by [28,29] have demonstrated that the sentiment analysis applied to social media posts can offer valuable insights into understanding the dynamics behind price fluctuations. The sentiment analysis of cryptocurrency-related tweets has mostly been conducted using tools such as VADER [30] and TextBlob [31], both of which have proven their effectiveness in several related works [3,27,32].

Kim et al. [33] demonstrated that the utilization of on-chain data (i.e., data directly from the blockchain) can provide valuable insights into the dynamics of cryptocurrency prices. Similarly, the analysis of on-chain data in conjunction with change point detection methodologies has contributed to enhancing the accuracy of cryptocurrency price forecasts and assisting investors in making more informed decisions [34,35].

Table 1 presents a comparison among the most recent works in the field of cryptocurrency price forecasting, which exploit different data and features. Specifically, each proposed algorithm or methodology has been evaluated based on the following characteristics:

- *Price Trend (PT)*: identifying and exploiting the direction in which the price of a cryptocurrency is moving (e.g., upward trend, downward trend).
- *Instant of Trading (IT)*: discovering the specific moment to buy/sell a cryptocurrency.
- *Publication Frequency (PF)*: exploiting social media data mentioning a particular cryptocurrency.
- *Impressions (I)*: evaluating the overall market sentiment or public perception about a specific cryptocurrency.
- *Sentiment Analysis (SA)*: analyzing users' sentiments and opinions to gauge the market's mood regarding a particular cryptocurrency.
- *Bot Analysis (BA)*: detecting the activities of bots on social media, as the posts they publish can influence the price prediction process.
- *Correlation among Coins (CC)*: measuring the relationship between the price movements of different cryptocurrencies, such as correlation and causality.
- *Trading Indicators (TI)*: exploiting market metrics and signals used by traders to make trading decisions, such as moving averages, relative strength index (RSI), moving average convergence divergence (MACD).
- *Deep Learning*: using deep learning approaches for price forecasting.
- *Type of Coins*: defining the type of cryptocurrencies considered for price prediction, which can belong to the four categories discussed in Section 3.1, i.e., Solid Project (SP), High Capitalization (HC), Influential Meme (IM), and Volatile Meme (VM) coins.

Our work stands out in the comparison due to its comprehensive analysis, encompassing almost all available features. Unlike other works, our research supports price forecasting for different types of cryptocurrencies (HC, SP, and IM), providing excellent market coverage. We excluded support for Volatile Meme (VM) coins due to their unpredictable and speculative nature, which makes them highly susceptible to market manipulation. In addition, we also discard posts with advertising content, often generated by bots for promoting online trading platforms, which can disturb the price prediction process.

Table 1. Comparison among existing related works and their features.

Related Work	PT	TI	Social Media Data				Market Data		Deep Learning	Type of Coins
			PF	I	SA	BA	CC	TI		
Hitam et al. [12]	x	-	-	-	-	-	x	x	x	HC, SP
Abraham et al. [10]	x	-	-	-	x	-	-	-	-	HC
Lahmiri et al. [16]	x	-	-	-	-	-	-	x	x	HC
Vo et al. [28]	x	-	-	-	x	-	-	-	x	HC
Rathan et al. [25]	x	-	-	-	-	-	-	-	-	HC
Valencia et al. [9]	x	-	-	-	x	-	-	x	x	HC
Wołk [27]	x	-	x	x	x	-	-	-	x	HC, SP
Patel et al. [15]	x	-	-	-	-	-	-	-	x	-
Ioannis et al. [24]	x	x	-	-	-	-	x	x	x	-
Khedr et al. [13]	x	-	-	-	-	-	-	x	x	HC, SP
Jay et al. [14]	x	-	-	-	-	-	-	x	x	HC, SP
Poongodi et al. [18]	x	-	-	-	-	-	-	x	x	-
Mirtaheri et al. [26]	-	-	x	x	x	x	-	-	-	HC
Hamayel et al. [23]	x	-	-	-	-	-	-	-	x	HC
Tanwar et al. [34]	x	-	-	-	-	-	x	-	x	HC
Shahbazi et al. [35]	x	-	-	-	-	-	-	x	x	HC, SP
Kim et al. [33]	x	-	-	-	-	-	-	-	x	HC
Van Tran et al. [21]	x	x	-	-	-	-	x	x	-	HC, SP, IM
Ammer et al. [17]	x	x	-	-	-	-	x	x	x	HC, SP, IM, VM
Fleischer et al. [20]	x	x	-	-	-	-	x	x	x	-
Sun et al. [22]	x	x	-	-	-	-	-	-	-	HC, SP, IM
Sebastião et al. [29]	x	x	-	-	-	-	-	x	-	HC, SP
Lamon et al. [19]	x	-	-	-	x	-	-	-	-	HC
Our work	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**HC, SP, IM**

3. Proposed Methodology

Our methodology aims to predict the price movements of cryptocurrencies through the examination of two large datasets: the first contains market data (e.g., prices and exchanged volumes) of a set of cryptocurrencies; and the second contains posts published on social media discussing such coins. These datasets have been combined and analyzed using different statistical analysis, text analysis, and deep learning techniques. The goal of the methodology is to define a trading algorithm capable of suggesting the moments in which to carry out sell-and-buy operations to maximize profits.

As illustrated in Figure 1, the methodology is composed of different phases: data collection, data preprocessing, data enrichment, training machine learning models, and trading recommendation. In the following, we discuss each phase in detail, highlighting the crucial and strategic decisions that characterize our methodology.
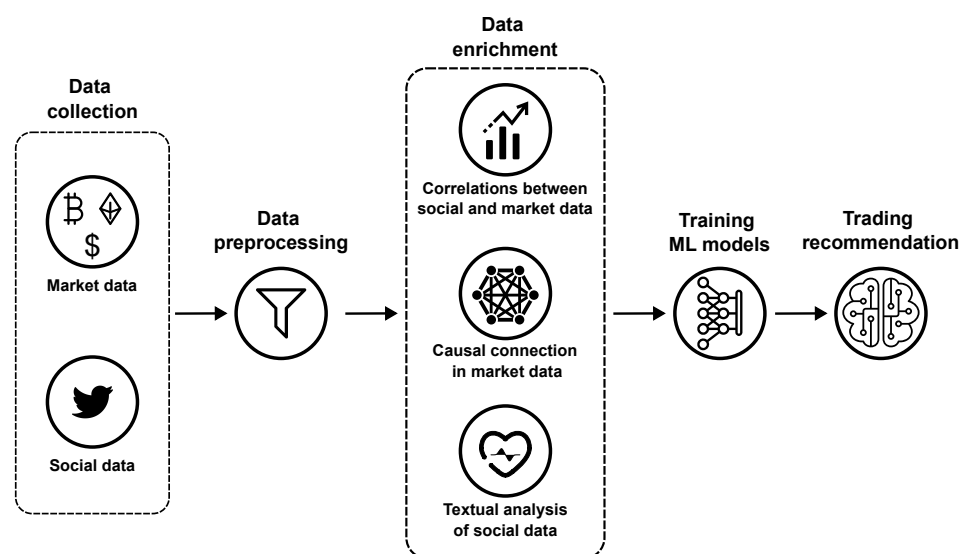
Figure 1. Execution flow of the proposed methodology.

To deal with such large and heterogeneous data, we leveraged Apache Spark to process them efficiently. The use of such a framework for Big Data is widely adopted in the realm of Big Data analytics, enabling faster and more scalable data processing [36].

3.1. Data Collection and Preprocessing

In the data collection phase, we gather market and social media data related to a selected group of cryptocurrencies. Our focus is on a heterogeneous set of representative cryptocurrencies chosen from the most popular ones. As shown in Table 2, these cryptocurrencies have been categorized into four distinct categories based on their characteristics:

- *High Capitalization (HC)*: this category includes cryptocurrencies such as Bitcoin and Ethereum, which are highly popular and have a significant impact on the world of cryptocurrencies.
- *Solid Project (SP)*: it includes cryptocurrencies backed by a robust project, although they may be less popular. Examples include Solana and Conflux, which form the foundation for various types of blockchains, as well as projects like The Sandbox, which is associated with the metaverse and NFT-related initiatives.
- *Influential Meme (IM)*: it includes coins that do not rely on solid projects (i.e., meme coins). Despite their lower capitalization and the absence of substantial projects, they have a significant influence on the world of cryptocurrencies due to their history and popularity on social media.
- *Volatile Meme (VM)*: this category comprises cryptocurrencies created purely for speculative purposes, characterized by high volatility and substantial price fluctuations within short time periods.

Table 2. List of cryptocurrencies used for the analysis.

Category	Acronym	Cryptocurrencies
High Capitalization	HC	Bitcoin (BTC), Ethereum (ETH), Polygon (MATIC), Polkadot (DOT), Solana (SOL), Cosmos (ATOM), Stellar (XLM), Avalanche (AVAX), Tron (TRX), Litecoin (LTC)
Solid Project	SP	Conflux (CFX), Stacks (STX), Fantom (FTM), Quant (QNT) Loopring (LRC), The sandbox (SAND), Gala (GALA), Lido Dao (LDO), Cronos (CRON), Zilliqa (ZIL), Chiliz (CHZ), Neo (NEO), Vethor Token (VTHO), Bancor (BNT), The Graph (GRT)
Influent Meme	IM	Dogecoin (DOGE), Shiba Inu (SHIB), Decentraland (MANA)
Volatile Meme	VM	Babydoge Coin (BabyDoge), Floki (Floki), Catecoin (CATE), Dogelon Mars (ELON), Volt Inu v2 (VOLT), Dejitaru Tsuka (TSUKA), Kishu Inu (KISHU), Shiba Predator (SHIBAP), Pitbull (PIT), Akita Inu (AKITA)

For each coin considered, we collected historical data on market performance from the CoinMarketCap website (https://coinmarketcap.com/, accessed on 1 November 2023), which provides information about price fluctuations (hourly, daily, weekly), market capitalization, daily traded amounts, and volumes of coins that are circulating in the market. Specifically, we gathered comprehensive market data for a selected group of cryptocurrencies spanning from January 2021 to March 2023. Cryptocurrency prices have been tracked over time in Tether (USDT), which is a stable coin designed to maintain a fixed 1:1 ratio with the US dollar. The analysis of market data was of great use in evaluating some key aspects of the relationships existing among the different cryptocurrencies, as discussed in the following.

Subsequently, the posts published by users talking about these coins were collected on social media platforms. In particular, we collected a large set of tweets published in the considered period, using the Twitter APIs with a set of keywords associated with the considered coins. Each collected tweet contains specific attributes, including the timestamp indicating when the tweet was posted, textual content, hashtags used, the author's username, the number of followers, and a flag indicating the user's account verification status,

enhancing the reliability of the collected information. Furthermore, we decided to discard tweets with advertising content, for example, those generated for promotional purposes by online trading platforms, often utilizing bots, which frequently mention multiple cryptocurrencies in their text. Then, each post in the dataset has been modified by applying some common preprocessing operations, such as removing usernames and mentions, special characters, URLs, and so on.

The final datasets include the hourly market information of the coins under analysis and approximately 133 million tweets, which are used to extract valuable insights about popularity trends, price fluctuations, and users' investment behaviors.

3.2. Data Enrichment

After the preprocessing phase, our final datasets comprise detailed hourly market data and an extensive collection of tweets related to the selected cryptocurrencies. To enhance the predictive capabilities of our methodology, we augmented these datasets with three additional information, as detailed in the following sections. In Section 3.2.1, we analyzed the correlation between social media activities and cryptocurrency price fluctuations. Section 3.2.2 investigates how cryptocurrency prices influence each other, with a particular focus on the impact of well-known cryptocurrencies on the broader market. Finally, in Section 3.2.3, we examined the textual content of posts published by social media users, aggregating their expressed opinions about cryptocurrencies, so as to identify the sentiment and utilize it to improve the prediction of cryptocurrency prices.

3.2.1. Correlation between Social Media and Market Data

At this step, we started from the intuition that the information extracted from social media has a strong correlation with the price fluctuations of cryptocurrencies. For each considered cryptocurrency, we calculated some social engagement metrics (number of tweets, followers, likes, and retweets) to determine their correlation with the daily closing prices. In particular, we used both Pearson and Spearman correlation tests, which are statistical measures used to assess the relationships between variables. The Pearson correlation evaluates linear relationships between normally distributed time series data, while Spearman evaluates monotonic relationships between variables, which may not necessarily be linear. Given the nature of cryptocurrency prices and social data, both correlation measures were considered to provide a comprehensive understanding of the relationship among time series. For example, Figure 2 provides a view of the social metrics and price fluctuations of Shiba Inu, where the values have been normalized to make the chart more understandable and comparable.
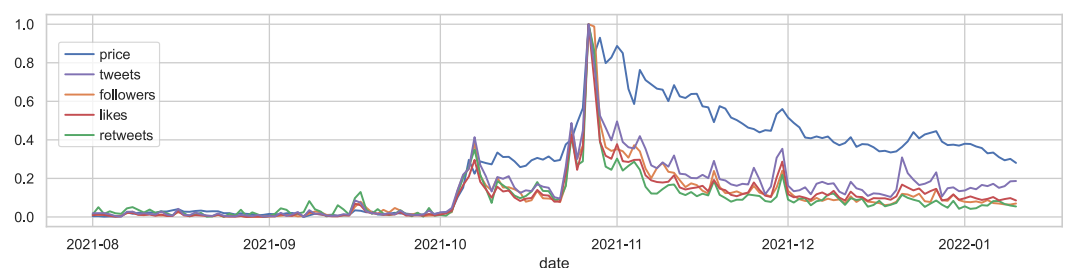

Figure 2. A five-month view of daily Twitter metrics and a closing price associated with Shiba Inu.

For all the cryptocurrencies considered, it has been verified that there is a strong correlation between the social metrics introduced above and the daily closing price. For example, Table 3 reports the Pearson and Spearman correlation coefficients of three meme coins (i.e., Shiba Inu, Floki, and CateCoin). As shown, the results indicate strong positive correlations between the tweet volume and cryptocurrency prices (ranging from 0.723 to 0.868 for Pearson and 0.841 to 0.909 for Spearman). The follower count exhibits moderate correlations, while likes and retweets show variable values, with Floki showing a strong correlation.

Table 3. Correlation coefficients between the price and social metrics for three meme coins (Shiba Inu, Floki, and CateCoin).

	Shiba Inu		*Floki*		*CateCoin*	
Category	*Pearson*	*Spearman*	*Pearson*	*Spearman*	*Pearson*	*Spearman*
Tweets	0.723	0.841	0.868	0.909	0.797	0.880
Followers	0.659	0.761	0.523	0.858	0.320	0.751
Likes	0.752	0.849	0.896	0.913	0.414	0.854
Retweets	0.796	0.850	0.889	0.913	0.674	0.831

Finally, these correlation values are added as features to the dataset in order to use them for training a machine learning model.

3.2.2. Causal Connection in Market Data

In this phase, it was evaluated how the prices of the different cryptocurrencies considered influence each other. The underlying idea behind this analysis is that fluctuations in the prices of popular cryptocurrencies, such as Ethereum and Bitcoin, may causally affect the prices of other coins. Following the approaches used in [37,38], the Granger causality test was employed to determine whether price variations in one cryptocurrency can be considered the cause of price changes in another coin.

The Granger causality test is used to determine whether a time variable X can be considered the cause of another time variable Y. During this test, the p-value, a widely used statistical concept for assessing evidence in support of or against a statistical hypothesis, is computed using the *F-test*. Such a test involves two main steps. In the first step, two regression models are built: the first uses only past values of Y to predict its current value, while the second model uses both past values of Y and X to predict the current value of Y. In the second step, the goodness of fit of the two models is compared using a statistical test. The test makes use of the following regression model:

$$y_t = \alpha_0 + \sum_{j=1}^{m} \alpha_j y_{t-j} + \sum_{j=1}^{m} \beta_j x_{t-j} + k_t$$

where: $alpha_0$ is the intercept term in the regression model; x_t represents the variable X at time t; y_t represents the variable Y at time t; β_j for $j = 1, \ldots, m$ are the regression coefficients; and k_t represents the error at time t. This test is based on the null hypothesis:

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_m = 0$$

when X causes Y, according to the Granger causality test, the null hypothesis is rejected.

Let us suppose that the null hypothesis represents the initial assumption that the price variations of one cryptocurrency do not influence the price variations of another currency. The alternative hypothesis represents the hypothesis we want to support if the data provide sufficient evidence against the null hypothesis. The p-value is compared to a predefined level of significance, denoted by *alpha*, which represents the maximum probability of making an error, incorrectly rejecting the null hypothesis when it is true. If the p-value is less than the level of significance *alpha* (typically 0.05), the null hypothesis is rejected, and it is concluded that the data provide statistical evidence in support of the alternative hypothesis. In other words, it is believed that there is a significant effect or relationship between the price variations of one cryptocurrency and that of another. Conversely, if the p-value is greater than *alpha*, there is not enough evidence to reject the null hypothesis, and therefore, it cannot be stated that there is a significant effect or relationship between the price variations of the two currencies. For example, considering the cryptocurrency Cosmos (ATOM), a cryptocurrency belonging to the solid project category, the Granger causality test indicated that its price is mostly influenced by Litecoin (LTC), as evidenced by a p-value of 0.0102.

Considering a 3-day time window, *p*-values for each pair of cryptocurrencies have been calculated using hourly prices. Subsequently, for each cryptocurrency, the hourly price variations and exponential moving average of the price for the three cryptocurrencies that most influence that coin (i.e., the ones with the lowest *p*-value and those below the *alpha* level) were added to the final dataset.

3.2.3. Textual Analysis of Social Data

At this stage of the methodology, a textual analysis was carried out on the large dataset collected from social media. As discussed before, the dataset is a rich repository of posts authored by users who explicitly mention at least one of the cryptocurrencies under analysis. In particular, we analyzed the textual content of each post for determining its sentiment, categorizing it as either negative, positive, or neutral. In such a way, it is possible to assess the collective sentiment of social media users at any given moment with respect to a particular cryptocurrency. This insight is very important for predicting cryptocurrency price movements: a positive sentiment often heralds a potential price increase, while conversely, a negative sentiment can foreshadow a decline in value. Specifically, the sentiment analysis of cryptocurrency-related posts has been carried out using two different tools, VADER [30] and TextBlob [31], which have been used in many other related work [3,27,32].

VADER is a lexical-based model using an annotated lexicon of English words with sentiment valence scores. It also considers negations, intensity modulators, and word order for precise sentiment analysis. In contrast, TextBlob is a Python library for natural language processing that assigns polarity scores on a scale from -1 (very negative) to 1 (very positive) and provides a subjectivity score. Following the same approach used in [39], to improve the VADER lexicon and better adapt it to the cryptocurrency context, the scores of some terms in VADER have been redefined. In fact, in the context of cryptocurrencies, specific terms are commonly used to identify phenomena of considerable importance, but VADER identifies them as common terms. For example, the terms *buy*, *moon*, and *rocket* suggest that the price of a cryptocurrency is going to see a huge increase, but they have a neutral score according to the original VADER lexicon.

The information on the collective sentiment about the different cryptocurrencies has then been added to the final dataset, in order to provide additional training data for the machine learning model used to predict prices.

3.3. Training Machine Learning Models

In this phase, a wide set of machine learning algorithms have been evaluated to choose the best solution for predicting prices of cryptocurrencies, including ensemble regressor and neural network algorithms. Concerning ensemble regressors, we used: *Random Forest* [40], which exploits a forest of decision trees; *XGBoost* [41], which provides a parallel tree boosting; and *CatBoost* [42], which utilizes a categorical feature-aware boosting algorithm. As for neural networks, we employed the following algorithms: *Conv1D* [43], which is a one-dimensional convolutional neural network (CNN) architecture suitable for sequence data; *GRU* (Gated Recurrent Unit) [44], which is a specialized recurrent neural network (RNN) designed for handling long-range dependencies in sequential data; and *LSTM* [44], which is an RNN that is particularly effective in modeling complex patterns and relationships over extended sequences. In particular, using the hyperparameter values shown in Table 4, these algorithms have been trained using data collected during the period from January 2021 to December 2021, related to the cryptocurrencies listed in Table 2. Subsequently, the different models obtained were tested on data collected during the period ranging from January 2022 to March 2023.

Table 4. Hyperparameter values used for the algorithms under comparison.

Model	Hyperparameters
Random forest	max_features: $\sqrt{n}$; min_samples_split: 5; estimators: 300
XGBoost	eta: 0.01; gamma: 150; n_estimators: 100; subsample: 1
CatBoost	depth: 6; iterations: 200; learning_rate: 0.1; l2_leaf_reg: 0.2
Conv1D	conv1d_layer: [units: 256; kernel_size: 2; activation: ReLU]; flatten_layer: yes; dense_layer_1: [units: 8; activation: ReLU]; dense_layer_2: [units: 1; activation: linear]; optimizer: Adam; learning_rate: 0.0001; epoch: 200
GRU	gru_layer_units: 256; dense_layer_1: [units: 8; activation: ReLU]; dense_layer_2: [units: 1; activation: linear]; optimizer: Adam; learning_rate: 0.0001; epoch: 200
LSTM	lstm_layer_units: 32; lstm_layer_2_units: 64; dense_layer_1: [units: 8; activation: ReLU]; dense_layer_2: [units: 1; activation: linear]; optimizer: Adam; learning_rate: 0.0001; epoch: 200

Table 5 shows a comparative overview of the performance obtained by the different machine learning models that have been tested. As shown, long short-term memory (LSTM) appears to be the best-performing model among the ones listed. It has the lowest RMSE (0.003), MAE (0.002), and MAPE (1.2%), indicating that it predicts cryptocurrency prices with the smallest errors compared to the other models. Additionally, it has the highest R^2 value (0.97), suggesting that it explains a larger portion of the variance in the data, making it a strong choice for predicting cryptocurrency prices in this context. The results obtained are in line with what we expected. In fact, the benefits of LSTMs for cryptocurrency price prediction have been confirmed in other studies [45,46], which have highlighted that LSTMs are the best model for short-term price prediction.

Table 5. Performance comparison of the different machine learning algorithms in predicting cryptocurrency prices.

Model	Category	RMSE	MAE	R^2	MAPE
Random forest	Tree-based	0.085	0.055	0.75	5.2%
XGBoost	Tree-based	0.110	0.070	0.68	6.8%
CatBoost	Tree-based	0.025	0.035	0.92	1.7%
Conv1D	CNN	0.005	0.003	0.95	1.4%
GRU	RNN	0.004	0.002	0.96	1.3%
LSTM	**RNN**	**0.003**	**0.002**	**0.97**	**1.2%**

Taking into account such results, we used an architecture consisting of two LSTM layers, followed by two densely connected layers. The first two LSTM layers capture long-term dependencies in the time sequence, while the subsequent dense layers handle data transformation and final prediction. Specifically, the first LSTM layer has been configured with 32 memory units, while the second one has been configured with 64 memory units. The next densely connected layers exploit a rectified linear unit (ReLU) activation function, which is commonly used to introduce nonlinearity into the neural network.

3.4. Trading Recommendation

The final phase of our methodology focused on defining a trading recommendation algorithm that exploits price predictions based on the LSTM model. The trading algorithm aims to suggest when is most appropriate to initiate trading operations (buy or sell) for a given cryptocurrency. To simplify the proposed heuristic, the algorithm invests the entire capital available in the account at each operation. However, in the future, less risky approaches could be studied, which involve better capital management in order to control losses. The algorithm takes into account some aspects:

- *Impact of commissions*: commission costs depend on the trading platform used, and thus, the algorithm is designed to take into account a certain percentage of the invested capital to be paid as transaction fees.
- *Identification of strong trends*: the algorithm implements a heuristic to limit the number of transactions, starting a new one only in the presence of a significant event. In this way, it is possible to avoid imprudent operations during phases of price uncertainty, with notable benefits in terms of profits.
- *Use of take-profit*: it leads the algorithm to close operations when the profit percentage exceeds a certain threshold.
- *Use of stop-loss*: it closes operations when the loss percentage exceeds a certain threshold.

It is worth noting that the algorithm is based on *future trading*, which allows traders to buy or sell cryptocurrencies at a predetermined price at a specified future date. This trading strategy enables traders to speculate on the price movements without owning the cryptocurrency. Specifically, traders can do two different types of trading operations: *shorting* a cryptocurrency or *going long* on a cryptocurrency. Shorting a cryptocurrency means betting that its price will decrease. Traders who short-sell borrow the cryptocurrency and sell it at the current price, hoping to buy it back later at a lower price, thus making a profit from the difference. On the contrary, going long on a cryptocurrency means betting that its value will rise, allowing traders to sell it later at a higher price and make a profit from the difference.

Algorithm 1 shows the pseudo-code of the *OpenTransaction* procedure, implementing the proposed trading algorithm that automates decisions on when to open and close operations based on the analysis of real and predicted prices. The algorithm receives the following parameters as input: a cryptocurrency *C*; a time *E* beyond which the execution of the trading algorithm ends; the loss percentage *LV* beyond which to activate the stop-loss procedure; the percentage gain *PV* beyond which to activate the take-profit procedure; a number *D* that indicates the time window (in days) for data used by the prediction model; a time *W* to wait before opening a trading operation; *DB*, the reference to the dataset containing aggregated social media and market information; *LSTM*, the trained neural network model for predicting prices; and a sleep time *S* between one completed operation and the next.

Algorithm 1 Pseudocode of the trading algorithm.

```
 1: procedure OPENTRANSACTION(C,E, LV, PV, D, W, DB, LSTM, S)
 2:     op ← null
 3:     while now() < E do
 4:         if op is null then
 5:             previousData ← DB.getLastDaysData(C, D)
 6:             P_p ← LSTM.predict(C, previousData)
 7:             P_r ← getRealPrice(C)
 8:             if intersect(P_r, P_p) then
 9:                 wait(W)                                          ▷ Safety waiting time
10:                 if P_r > P_p then
11:                     op ← openSellOperation(C)         ▷ Betting against (or shorting) cryptocurrency C.
12:                 else
13:                     op ← openBuyOperation(C)       ▷ Betting in favor of (or going long on) cryptocurrency C.
14:                 end if
15:             end if
16:         else
17:             if op.isSell() then
18:                 if hoursPredictedPriceGreaterThanReal(C) > W or checkStopLossOrTakeProfit(LV, PV) then
19:                     op.close()
20:                 end if
21:             else
22:                 if hoursRealPriceGreaterThanPredicted(C) > W or checkStopLossOrTakeProfit(LV, PV) then
23:                     op.close()
24:                 end if
25:             end if
26:             if p.isClosed() then
27:                 op.calculateProfit()
28:                 op ← null
29:             end if
30:         end if
31:     end while
32:     return
33: end procedure
```

Given a reference cryptocurrency C, the algorithm initializes a variable op as null, which will be used to represent the current trading operation (line 2). Then, it initiates a while loop that continues until the current time exceeds the specified duration E for trading (line 3). If the op variable is null, indicating no active trading operation, then the algorithm tries to start a new one.

To this end, it calculates the predicted and real prices for the cryptocurrency. In particular, the predicted price is given by the LSTM model based on the provided last D-days historical data (lines 5–6), while the real price is gathered from market coin APIs (line 7). The algorithm checks whether the predicted price and real price intersect (line 8), indicating a potential trading opportunity. After waiting a safety time W (line 9), aiming to avoid price retracements, the algorithm decides the type of trading operation to perform. Price retracements represent temporary and relatively short corrections within a growth or decline trend of a specific asset. They are common phenomena in financial markets, including cryptocurrency markets, and can offer trading and investment opportunities. Specifically, if the real price is greater than the predicted price, it starts a sell operation (lines 10–11), betting that the price will decrease (shorting). Otherwise, it starts a buy operation (lines 12–13), betting that the price will rise (going long). If op is defined, which means a trading operation is active, the algorithm starts monitoring the prices of the cryptocurrency to establish whether the current operation should be closed (lines 17–25). In particular, it checks whether the hours in which the predicted price exceeded the real price (or vice versa) have exceeded a given waiting time of W or if the stop-loss/take-profit condition is met. To close a sell operation, the algorithm checks whether, during the monitoring period (i.e., the period in which the operation is active), the predicted price P_r of the cryptocurrency always remained higher than the real one or if the stop-loss/take-profit condition is met (lines 17–20). Similarly, to close a buy operation, the algorithm checks whether the real price has always remained higher than the predicted one for a time greater than W or if the stop-loss/take-profit condition is met (lines 21–25). If op has been closed, the profit (or loss) obtained from the trading operation is calculated (line 27); afterward, the variable op is set to null (line 28) to allow the start of a new trading operation.

Figure 3 illustrates how the trading algorithm operates, using the cryptocurrency Shiba Inu as an example. As shown in Figure 3a, the first black vertical line (A) marks the point where the predicted price (P_p) intersects with the real price (P_r). Following this intersection, the real price consistently remains higher than the predicted price for the next time W (depicted in orange). As there are no subsequent intersections during W, indicating a stable downward trend, the algorithm suggests that a downward trend may be ongoing and starts a new shorting operation.

(**a**) Example of shorting operation.

(**b**) Operation opened in a period of two months for Shiba Inu.
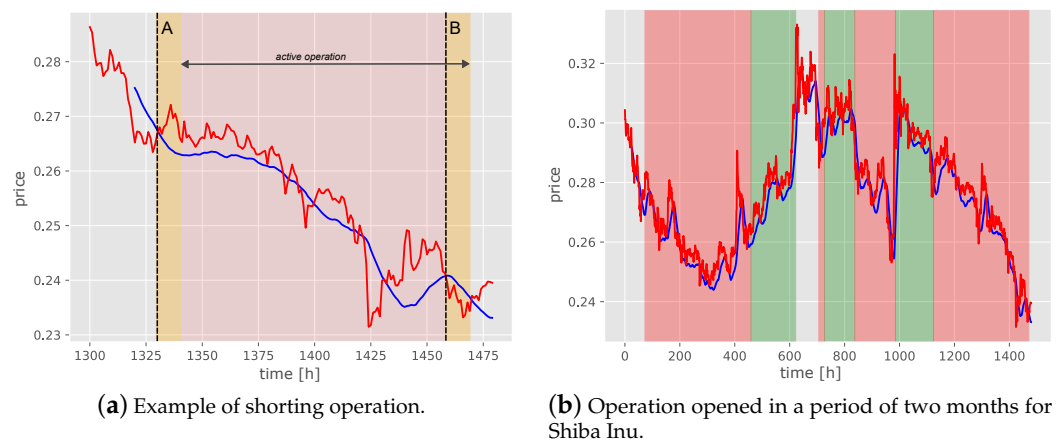
Figure 3. Example of how the trading algorithm works (the red line indicates the real price while the blue line indicates the predicted price).

At each subsequent intersection, the algorithm assesses whether it is advantageous to close the existing trading operation, observing a waiting period of W before making a decision. In the example, the second vertical black line (B) represents the intersection that leads to the closure of the operation. After B, the predicted price remains below the real price for a period W, suggesting a possible end of the downward trend. Consequently, the shorting operation is closed. The interval during which the algorithm kept the shorting operation open is represented by the red area.

Figure 3b illustrates the operations that have been initiated in two months of tests, using the cryptocurrency Shiba Inu. Specifically, the periods during which shorting operations were opened are highlighted in red, those in which long ongoing operations were opened and marked in green, while periods with no active operations were marked in gray.

4. Experimental Results

Several experiments were conducted on historical data to assess the effectiveness of the trading algorithm. Additionally, a phase of parameter tuning was carried out with the goal of identifying optimal values to maximize profits. The parameters assessed during this process included the take-profit (PV) and stop-loss (LV) percentages, as well as the duration of the safety interval before opening a trading operation (W), and the number of days for data used by the prediction model (D). In particular, after evaluating different values for such parameters, we identified the following optimal configurations: $PV = 12\%$, $LV = 8\%$, $W = 12$ h, and $D = 3$ days.

The training of the LSTM model has been carried out using a dataset spanning the period January 2021–December 2021, related to the cryptocurrencies listed in Table 2. Subsequently, we evaluated the obtained profits on data collected on a different period, ranging from January 2022 to March 2023. For our tests, among those listed in Table 2, we selected 28 coins in which to invest. In particular, we decided to invest exclusively in three categories of cryptocurrencies: high capitalization (HC), solid project (SP), and influential meme coin (IM). We decided not to invest in volatile meme coins (VM) due to their unpredictable nature, lack of fundamental value, and susceptibility to market manipulation, which can lead to significant financial losses. The evaluation was carried out starting from a virtual initial capital of USD 1000 for each cryptocurrency (USD 28,000 in total). Subsequently, the profits generated by the trading algorithm were evaluated, also taking into account the transaction fees of 1%. It is worth noting that fees are paid on each transaction, so the greater the number of open trading operations, the greater the amount paid.

The overall results of the algorithm are presented in Table 6, highlighting an overall gain of 194% when transaction fees are not taken into account, and of 117% when transaction fees are considered. Specifically, starting from an initial capital of USD 28,000, we obtained a final capital of USD 82,359 with no transaction fees and of USD 60,871 by considering fees. The trading algorithm proved to be extremely effective in predicting the price trend of influential meme (IM) coins, which appear to be significantly influenced by trends and popularity on social media. Specifically, trading on IM leads to a very high average profit of 902.48% with fees and 1257.96% without fees. However, it is worth noting that not all cryptocurrencies have produced profits, but some of them have experienced losses. In particular, fees have a significant impact on profits. As an example, in the case of BTC, the algorithm produced a loss of 14.91% with fees, while it produced a gain of 227.38% without fees. This phenomenon is due to the fact that the algorithm, in some situations, tends to open and close trading operations whose profit is not able to cover the cost of commissions. This is an aspect that will need to be better evaluated in the future, introducing additional operational constraints into the algorithm to address such situations. In some other cases (i.e., TRX and CHZ), the algorithm produced zero profits as no trading operations were carried out during the period considered. Finally, in some rare cases (e.g., SAND, GRT, and FTM), losses have occurred, both with and without fees. This behavior is most likely due to

the fact that data collected for training were limited and did not provide the LSTM model with sufficient predictive capabilities.

Table 6. Results obtained on selected cryptocurrencies categorized as high capitalization (HC), solid project (SP), and influential meme coin (IM).

Cryptocurrency	Acronym	Category	Profit % with Fees	Profit % without Fees
Bitcoin	BTC	HC	−14.91	+227.38
Ethereum	ETH	HC	+6.41	+16.35
Polygon	MATIC	HC	+164.55	+178.43
Polkadot	DOT	HC	−47.80	+28.13
Solana	SOL	HC	−33.56	−10.55
Cosmos	ATOM	HC	−42.27	+13.89
Stellar	XLM	HC	+40.01	+48.26
Avalanche	AVAX	HC	+23.26	+29.37
TRON	TRX	HC	0.00	0.00
Litecoin	LTC	HC	+10.58	+85.98
Conflux	CFX	SP	+112.46	+202.52
Stacks	STX	SP	+59.89	+109.21
Fantom	FTM	SP	−25.42	−22.88
Quant	QNT	SP	+74.17	+83.96
Loopring	LRC	SP	+3.73	+4.73
The Sandbox	SAND	SP	−80.04	−30.06
Gala	GALA	SP	+259.01	+297.28
Lido DAO	LDO	SP	−79.72	+85.95
Cronos	CRO	SP	−15.32	+5.68
Zilliqa	ZIL	SP	+38.22	+58.44
Chiliz	CHZ	SP	0.00	0.00
Neo	NEO	SP	+132.66	+240.80
VeThor Token	VTHO	SP	+26.57	+32.84
Bancor	BNT	SP	−7.44	−5.50
The Graph	GRT	SP	−25.37	−18.11
Dogecoin	DOGE	MCI	+27.68	+39.41
Shiba Inu	SHIB	MCI	+432.47	+771.74
Decentraland	MANA	MCI	+2247.29	+2962.72
Mean profit for HC coins			**+10.63**	**+61.72**
Mean profit for SP coins			**+31.56**	**+69.66**
Mean profit for IM coins			**+902.48**	**+1257.96**
Overall mean profit			**+117.40**	**+194.14**

5. Conclusions

In conclusion, the growing popularity and value of various cryptocurrencies, including the emergence of meme coins like Dogecoin and Shiba Inu, have been driven by a combination of technological innovation and marketing strategies. In particular, social media platforms and influential figures like Elon Musk have played key roles in shaping the cryptocurrency landscape. Our study has successfully identified the major factors influencing cryptocurrency price fluctuations, with a primary emphasis on social media data. By analyzing the correlation between tweet frequency, likes, retweets, and user popularity, we revealed the significant impact of social media on cryptocurrency prices. Additionally, we explored how high-cap cryptocurrencies can influence the broader market, especially meme coins, which are highly susceptible to external factors. Using a combination of different statistical analysis, text analysis, and deep learning techniques, we developed a methodology for predicting the price fluctuations of cryptocurrencies and suggesting the optimal moments for trading in order to maximize profits. In particular, we defined a trading recommendation algorithm that, exploiting price predictions provided by an LSTM model, leads to a total profit of 194% when transaction fees are not taken into account and of 117% when transaction fees are considered. Moreover, it proved highly effective in predicting the price trend of influential meme coins. Considering only such a category of coins, the algorithm resulted in a substantial average profit of 902% with fees and 1258% without fees. The proposed trading algorithm can serve as a powerful tool to optimize the trading strategies and maximize profits. Looking ahead, there is potential to further refine the algorithm by adopting a less risky capital management approach to better control

losses. This involves mitigating the risks associated with using the entire capital in each financial operation and minimizing the negative impact of trading commissions.

References

1. Härdle, W.K.; Harvey, C.R.; Reule, R.C. Understanding cryptocurrencies. *J. Financ. Econom.* **2020**, *18*, 181–208. [CrossRef]
2. Hashemi Joo, M.; Nishikawa, Y.; Dandapani, K. Cryptocurrency, a successful application of blockchain technology. *Manag. Financ.* **2020**, *46*, 715–733. [CrossRef]
3. Kraaijeveld, O.; De Smedt, J. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *J. Int. Financ. Mark. Inst. Money* **2020**, *65*, 101188. [CrossRef]
4. Cakici, N.; Fieberg, C.; Metko, D.; Zaremba, A. Do Anomalies Really Predict Market Returns? New Data and New Evidence. *Rev. Financ. Forthcom.* **2023**, rfad025. [CrossRef]
5. Gu, S.; Kelly, B.; Xiu, D. Empirical asset pricing via machine learning. *Rev. Financ. Stud.* **2020**, *33*, 2223–2273. [CrossRef]
6. Bianchi, D.; Büchner, M.; Tamoni, A. Bond risk premiums with machine learning. *Rev. Financ. Stud.* **2021**, *34*, 1046–1089. [CrossRef]
7. Bali, T.G.; Beckmeyer, H.; Moerke, M.; Weigert, F. Option return predictability with machine learning and big data. *Rev. Financ. Stud.* **2023**, *36*, 3548–3602. [CrossRef]
8. Zhou, X.; Zhou, H.; Long, H. Forecasting the equity premium: Do deep neural network models work? *Mod. Financ.* **2023**, *1*, 1–11. [CrossRef]
9. Valencia, F.; Gómez-Espinosa, A.; Valdés-Aguirre, B. Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy* **2019**, *21*, 589. [CrossRef]
10. Abraham, J.; Higdon, D.; Nelson, J.; Ibarra, J. Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Sci. Rev.* **2018**, *1*, 1.
11. Branda, F.; Marozzo, F.; Talia, D. Ticket Sales Prediction and Dynamic Pricing Strategies in Public Transport. *Big Data Cogn. Comput.* **2020**, *4*, 36. [CrossRef]
12. Hitam, N.A.; Ismail, A.R. Comparative performance of machine learning algorithms for cryptocurrency forecasting. *Ind. J. Electr. Eng. Comput. Sci* **2018**, *11*, 1121–1128. [CrossRef]
13. Khedr, A.M.; Arif, I.; El-Bannany, M.; Alhashmi, S.M.; Sreedharan, M. Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. *Intell. Syst. Account. Financ. Manag.* **2021**, *28*, 3–34. [CrossRef]
14. Jay, P.; Kalariya, V.; Parmar, P.; Tanwar, S.; Kumar, N.; Alazab, M. Stochastic neural networks for cryptocurrency price prediction. *IEEE Access* **2020**, *8*, 82804–82818. [CrossRef]
15. Patel, M.M.; Tanwar, S.; Gupta, R.; Kumar, N. A deep learning-based cryptocurrency price prediction scheme for financial institutions. *J. Inf. Secur. Appl.* **2020**, *55*, 102583. [CrossRef]
16. Lahmiri, S.; Bekiros, S. Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos Solitons Fractals* **2019**, *118*, 35–40. [CrossRef]
17. Ammer, M.A.; Aldhyani, T.H. Deep learning algorithm to predict cryptocurrency fluctuation prices: Increasing investment awareness. *Electronics* **2022**, *11*, 2349. [CrossRef]
18. Poongodi, M.; Nguyen, T.N.; Hamdi, M.; Cengiz, K. Global cryptocurrency trend prediction using social media. *Inf. Process. Manag.* **2021**, *58*, 102708.
19. Lamon, C.; Nielsen, E.; Redondo, E. Cryptocurrency price prediction using news and social media sentiment. *SMU Data Sci. Rev.* **2017**, *1*, 1–22.
20. Fleischer, J.P.; von Laszewski, G.; Theran, C.; Parra Bautista, Y.J. Time Series Analysis of Cryptocurrency Prices Using Long Short-Term Memory. *Algorithms* **2022**, *15*, 230. [CrossRef]

21. Van Tran, L.; Le, S.T.; Tran, H.M. Empirical Study of Cryptocurrency Prices Using Linear Regression Methods. In Proceedings of the 2022 RIVF International Conference on Computing and Communication Technologies (RIVF), Ho Chi Minh City, Vietnam, 20–22 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 701–706.

22. Sun, J.; Zhou, Y.; Lin, J. Using machine learning for cryptocurrency trading. In Proceedings of the 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS), Taipei, Taiwan, 6–9 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 647–652.

23. Hamayel, M.J.; Owda, A.Y. A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. *AI* **2021**, *2*, 477–496. [CrossRef]

24. Livieris, I.E.; Pintelas, E.; Stavroyiannis, S.; Pintelas, P. Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms* **2020**, *13*, 121. [CrossRef]

25. Rathan, K.; Sai, S.V.; Manikanta, T.S. Crypto-currency price prediction using decision tree and regression techniques. In Proceedings of the 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 23–25 April 2019; IEEE: Piscataway, NJ, USA, 2019, pp. 190–194.

26. Mirtaheri, M.; Abu-El-Haija, S.; Morstatter, F.; Ver Steeg, G.; Galstyan, A. Identifying and analyzing cryptocurrency manipulations in social media. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 607–617. [CrossRef]

27. Wołk, K. Advanced social media sentiment analysis for short-term cryptocurrency price prediction. *Expert Syst.* **2020**, *37*, e12493. [CrossRef]

28. Vo, A.D.; Nguyen, Q.P.; Ock, C.Y. Sentiment analysis of news for effective cryptocurrency price prediction. *Int. J. Knowl. Eng.* **2019**, *5*, 47–52. [CrossRef]

29. Sebastião, H.; Godinho, P. Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financ. Innov.* **2021**, *7*, 3. [CrossRef]

30. Hutto, C.; Gilbert, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the International AAAI Conference on Web and Social Media, Ann Arbor, MI, USA, 1–4 June 2014; Volume 8, pp. 216–225.

31. Loria, S. textblob Documentation. *Release 0.16* **2018**, *2*, 269.

32. Pano, T.; Kashef, R. A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. *Big Data Cogn. Comput.* **2020**, *4*, 33. [CrossRef]

33. Kim, G.; Shin, D.H.; Choi, J.G.; Lim, S. A deep learning-based cryptocurrency price prediction model that uses on-chain data. *IEEE Access* **2022**, *10*, 56232–56248. [CrossRef]

34. Tanwar, S.; Patel, N.P.; Patel, S.N.; Patel, J.R.; Sharma, G.; Davidson, I.E. Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. *IEEE Access* **2021**, *9*, 138633–138646. [CrossRef]

35. Shahbazi, Z.; Byun, Y.C. Improving the cryptocurrency price prediction performance based on reinforcement learning. *IEEE Access* **2021**, *9*, 162651–162659. [CrossRef]

36. Belcastro, L.; Cantini, R.; Marozzo, F.; Orsino, A.; Talia, D.; Trunfio, P. Programming big data analysis: Principles and solutions. *J. Big Data* **2022**, *9*, 1–50. [CrossRef]

37. Al Guindy, M. Cryptocurrency price volatility and investor attention. *Int. Rev. Econ. Financ.* **2021**, *76*, 556–570. [CrossRef]

38. Aslanidis, N.; Bariviera, A.F.; López, Ó.G. The link between cryptocurrencies and Google Trends attention. *Financ. Res. Lett.* **2022**, *47*, 102654. [CrossRef]

39. Mardjo, A.; Choksuchat, C. HyVADRF: Hybrid VADER–Random Forest and GWO for Bitcoin Tweet Sentiment Analysis. *IEEE Access* **2022**, *10*, 101889–101897. [CrossRef]

40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

41. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

42. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 3–8 December 2018; Volume 31.

43. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [CrossRef]

44. Hua, Y.; Zhao, Z.; Li, R.; Chen, X.; Liu, Z.; Zhang, H. Deep learning with long short-term memory for time series prediction. *IEEE Commun. Mag.* **2019**, *57*, 114–119. [CrossRef]

45. Moustafa, H.; Malli, M.; Hazimeh, H. Real-time Bitcoin price tendency awareness via social media content tracking. In Proceedings of the 2022 10th International Symposium on Digital Forensics and Security (ISDFS), Istanbul, Turkey, 6–7 June 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6.

46. Maqsood, U.; Khuhawar, F.Y.; Talpur, S.; Jaskani, F.H.; Memon, A.A. Twitter Mining based Forecasting of Cryptocurrency using Sentimental Analysis of Tweets. In Proceedings of the 2022 Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain, 14–17 February 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6.