IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 5, 2022, accepted May 20, 2022, date of publication May 25, 2022, date of current version June 1, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3177888

A Deep Learning-Based Cryptocurrency Price Prediction Model That Uses On-Chain Data

GYEONGHO KIM[1], DONG-HYUN SHIN[2], JAE GYEONG CHOI[1], AND SUNGHOON LIM[1,3]

[1]Department of Industrial Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, South Korea
[2]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, South Korea
[3]Institute for the 4th Industrial Revolution, Ulsan National Institute of Science and Technology, Ulsan 44919, South Korea

Corresponding author: Sunghoon Lim (sunghoonlim@unist.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant through the Korea Government (MSIT) under Grant 2021R1F1A1046416, and in part by the AI Collaboration Project Fund of Ulsan National Institute of Science and Technology (UNIST) under Grant 1.220083.

ABSTRACT Cryptocurrency has recently attracted substantial interest from investors due to its underlying philosophy of decentralization and transparency. Considering cryptocurrency's volatility and unique characteristics, accurate price prediction is essential for developing successful investment strategies. To this end, the authors of this work propose a novel framework that predicts the price of Bitcoin (BTC), a dominant cryptocurrency. For stable prediction performance in unseen price range, the change point detection technique is employed. In particular, it is used to segment time-series data so that normalization can be separately conducted based on segmentation. In addition, on-chain data, the unique records listed on the blockchain that are inherent in cryptocurrencies, are collected and utilized as input variables to predict prices. Furthermore, this work proposes self-attention-based multiple long short-term memory (SAM-LSTM), which consists of multiple LSTM modules for on-chain variable groups and the attention mechanism, for the prediction model. Experiments with real-world BTC price data and various method setups have proven the proposed framework's effectiveness in BTC price prediction. The results are promising, with the highest MAE, RMSE, MSE, and MAPE values of 0.3462, 0.5035, 0.2536, and 1.3251, respectively.

INDEX TERMS Blockchain, cryptocurrency, Bitcoin, deep learning, prediction methods, change detection algorithms.

I. INTRODUCTION

With the advent of blockchain technology, there has been significant change in the form of currency as well as transactions. From its emergence to the present, currency's core role has been a means of payment as a medium of value delivery. This function relies on trust in the currency that is guaranteed and stabilized by a central agency (e.g., government, bank). However, central authorities have a critical shortcoming; there is the possibility of depravity that could risk transaction reliability. The blockchain, an open, anti-counterfeiting, and tamper-proof ledger, has created a currency called cryptocurrency. Based on blockchain technology, cryptocurrency can be trusted without the guarantee of a central authority, thus breaking away from the traditional relationship. Cryptocurrency that guarantees decentralization and transparency presents the possibility of a monetary system that relieves the risks of fraud and protects privacy [2]. The dominant cryptocurrency, Bitcoin (BTC), is an exemplary cryptocurrency in terms of its difference from existing traditional currencies. BTC is limited to 21 million issuances, resulting in practically no inflation that is caused by a central government's currency printing [3]. This strengthens the meaning of decentralization, leading cryptocurrency to function not only as a method of payment but also as a means of value storage. In fact, in addition to traditional investment vehicles, investing in cryptocurrencies is currently deemed one of the most effective ways to increase asset value.

With market capitalization at all-time high, BTC has been ranked in the top 10 of entire assets, which shows people's overwhelming interest in cryptocurrency. However, cryptocurrency differs from traditional financial products in several aspects. First, it could simultaneously function as value storage, method of payment, and platform. In addition, several types of cryptocurrencies, including BTC, do not necessarily correlate with traditional assets, such as gold and

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

crude oil, and even exhibit dissimilar behavior compared to stocks [4]. Unlike traditional assets (e.g., gold, stock, fiat money, etc.), cryptocurrency is intangible, volatile, and lacks an entity of ownership (i.e., corporation). Another distinctive characteristic of cryptocurrency is the existence of on-chain data, which contain information acquired from the blockchain [6]. On-chain data consist of valuable information regarding the blockchain network, including transactions, block size, and mining difficulty. Thus, existing traditional asset classification criteria and indicators cannot be directly applied to cryptocurrency. Considering the aforementioned points, a novel approach that reflects cryptocurrency's distinct characteristics is imperative for successful applications.

Accurately predicting cryptocurrency price forms the basis of successful investment, as it helps to establish risk management strategies and optimization methods that consider uncertainty and the possibility of loss. Furthermore, in order to achieve stability and maximum profits, investment strategies like portfolio management could also be reinforced by accurate price prediction [7]. However, since the market has a dynamic structure dominated by complex factors, it is difficult for investors and stakeholders to successfully make profits. Therefore, developing an accurate price prediction model is essential for establishing profitable investment strategies.

Recently, due to their ability to model non-stationarity in time-series data (unlike traditional approaches), machine learning methods have been widely used in predicting prices for financial products [8]. However, this work has found that there exist two issues in the literature. The first issue is due to a recent upsurge and plummet in cryptocurrency prices. As shown in Fig. 1, since the price moves in an unexpected range that has been previously unseen, constructed machine learning-based models are not able to predict future prices accurately. This problem does not apply only to certain prediction algorithms but could affect practically every prediction model constructed based on price data within a moderate range. This work therefore proposes a novel method to address the aforementioned problem using a change point detection (CPD) technique. In particular, during training, input data are segmented with CPD so that each segmented data has its own statistical characteristics. Based on segmentations, data are normalized separately to effectively reflect severe fluctuations. This has proven to be a practical solution to the first issue by the experiments in this work. The second issue that this work addresses for improvement of the cryptocurrency price prediction literature is that many existing works utilize only trite variables, such as historical prices and social media data. This work suggests using extensive blockchain-associated variables to enhance the ability of price prediction approaches. The proposed framework uses on-chain data, which are the most important factors for cryptocurrency price prediction, as independent variables. The contributions of this work to the literature are as follows:

1) Extensive blockchain-associated on-chain data are collected and used for cryptocurrency (i.e., BTC) price prediction.
2) Input variables are categorized into multiple groups based on domain knowledge and are utilized separately within a proposed prediction model.
3) For stable price prediction performance using non-stationary time-series data, the CPD technique is employed.
4) A novel prediction method dubbed SAM-LSTM is proposed; this method uses multiple LSTM modules for different input variable groups and a self-attention mechanism for extracting rich information immanent in on-chain data.

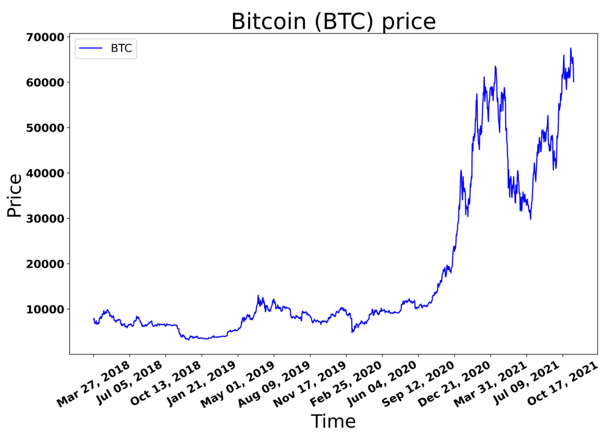

FIGURE 1. Price of Bitcoin (BTC).

An overview of the proposed framework for predicting the price of cryptocurrencies is illustrated in Fig. 2. The remainder of the paper is organized as follows. Section II discusses existing literature related to this work. Extensive data collection and selection, as well as detailed descriptions of methods used in the proposed framework, are illustrated in Section III. In addition, experiments with various setups based on real-world data, including on-chain data, and the corresponding results are detailed. The conclusion is drawn, and future works are illustrated in Section V. A summary of acronyms used in this work is presented in Table 1.

II. PRELIMINARIES

This section discusses preliminary materials, including existing works related to blockchain technology, cryptocurrency, and price prediction. In particular, existing cryptocurrency price prediction approaches based on machine learning and deep learning methods are discussed in depth. Additionally, applications for the CPD technique, an essential component of this work's proposed framework, are illustrated.

A. CRYPTOCURRENCY

The first blockchain technology-based cryptocurrency that functions as an electronic peer-to-peer transaction system without a trusted third party was allegedly proposed by S. Nakamoto in 2009 [9]. The underlying philosophies of

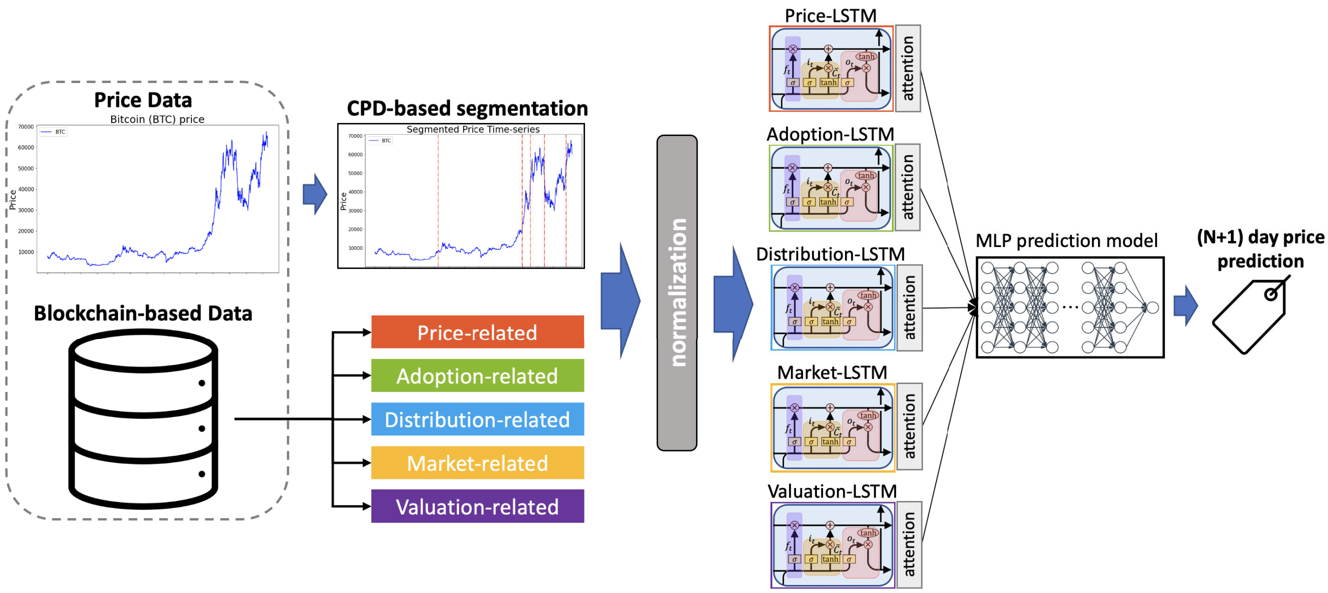

FIGURE 2. An overview of the proposed price prediction framework.

TABLE 1. Summary of acronyms.

Acronym	Full form
BTC	Bitcoin
CPD	change point detection
PELT	pruned exact linear time
CCF	cross-correlation function
DL	deep learning
RL	reinforcement learning
MLP	multi-layer perceptron
RNN	recurrent neural network
CNN	convolutional neural network
LSTM	long short-term memory
SAM-LSTM	self-attention-based multiple long short-term memory

cryptocurrency are threefold: safety, decentralization, and blockchain transparency. In the blockchain, each block contains transaction information that is distributively stored using its own hash value to prevent fraud or infringement [10]. Nodes, which are participants in the blockchain network, act as financial institutions so that the existence of a central agency is no longer necessary. The information recorded in each block is transparently disclosed. These are called on-chain data and are applied as indicators. Due to these innovative characteristics, cryptocurrency has gained much attention from the public, and many cryptocurrencies have been created.

Many cryptocurrencies that exist today have various purposes and usages depending on type. In all, BTC is the most popular cryptocurrency in terms of market capitalization. Although originally intended as an alternative payment method, BTC is now considered a speculative asset due to its high volatility and limited issuance, which make it unsuitable for use as a real currency [11]. Similar to BTC, Litecoin (LTC) can be used for transactions and also functions as a means of value storage due to limited issuance. In contrast, cryptocurrencies that control volatility to faithfully perform their role as a means of payment, called stable coin, exist [12]. The most famous stable coin with the largest market capitalization is Tether (USDT). It is issued one USDT per dollar and is stabilized through an arbitrage mechanism [13]. In addition to these cryptocurrencies, Ethereum (ETH) provides an ecosystem where decentralized applications (Dapp) can be implemented based on the blockchain [14]. Ripple (XRP) has a dissimilar purpose in the sense that it acts as a bridge between cryptocurrencies, whose transactions are not active, in a centralized manner [15]. Although these cryptocurrencies (i.e., BTC, LTC, USDT, ETH, XRP) have different inherent characteristics, they share coherent relationships that affect each other's prices. In particular, all cryptocurrencies affect the price of BTC in both direct and indirect manner.

B. PRICE PREDICTION

Macroeconomic variables, such as GDP, interest rate, and inflation rate, are considered important factors that influence the prices of conventional assets, such as gold, stock, and fiat money [5], [16], [17]. These indicators are commonly used in asset pricing and price prediction tasks. Another variable widely used in predicting conventional asset value is technical indicators, such as historical value, candles, trading volume, and moving average [18]–[20]. This work also utilizes some of the aforementioned factors, such as transaction sizes, transfer volumes, and market capitalization, which are associated with the price of cryptocurrency that exists in the on-chain data.

However, there are unique indicators for each asset based on distinctive characteristics used for price prediction. For instance, in the case of gold, scarcity, profitability, and political anxiety are considered in price prediction, since these factors affect demand and supply, which determine market equilibrium [21]. Stocks also have exclusive factors, such as income statements, financial statements, and cash flow statements, that are used for price prediction. In addition, factors like insider ownership and investor protection are also exploited [22]. Exogenous factors derived from social media (e.g., Twitter) data have also been utilized for stock price prediction [23]. In the case of the valuation of fiat money, the country's financial situation, monetary policy, and economic power are price determinants [24]. As shown in the conventional asset examples, to predict the price of cryptocurrency, a unique set of influencing factors should be taken into consideration. To this end, this work utilizes on-chain data that consist of factors that comprehensively reflect the ecosystem, market, and participation of cryptocurrency.

C. CRYPTOCURRENCY PRICE PREDICTION

Several recent works on cryptocurrency price prediction, especially BTC, have used various methods. In order to deal with cryptocurrency prices that have enormous fluctuations and non-stationarity, a dominant branch of research is based on machine learning methods [25], [26]. In particular, not only conventional machine learning algorithms but also more sophisticated methods, such as reinforcement learning (RL) and deep learning (DL)-based approaches, have been popularly utilized for handling the volatility of cryptocurrency prices.

Traditional machine learning algorithms, such as linear regression (LR) and support vector machine (SVM), use the sliding window data technique to predict daily BTC prices [27]. Chen *et al.* also suggest employing machine learning algorithms that use high-dimensional features to classify daily price movements [28]. Rebane *et al.* compare the ability of conventional auto-regressive integrated moving average (ARIMA) models and recurrent neural network (RNN)-based seq2seq models to predict prices and prove the superior efficacy of neural network-based models [29]. Multi-layer perceptron (MLP) and long short-term memory (LSTM) are also used to predict daily cryptocurrency prices [1]. Sin *et al.* employ an ensemble approach based on multiple MLP modules to classify BTC price movements [30]. However, MLPs used in existing works do not fully utilize the sequentiality of time-series data and thus are limited in the sense that only a fixed number of the data's days can be used (e.g, 7 days) to predict prices. For predicting BTC prices, neural networks that handle time-series data well, such as RNN and a convolutional neural network (CNN), have been most frequently adopted. Lahmiri and Bekiros address the chaotic dynamics of cryptocurrency markets and show that LSTM has greater predictability than standard neural networks (e.g., GRNNs) [31]. BTC price prediction has been conducted by machine learning algorithms and RNN-based methods that only use historical price data, such as open price, close price, and volume [32]. Patel *et al.* exploit LSTM and gated recurrent unit (GRU)-based prediction models in a hybrid scheme that especially focuses on predicting the price of LTC and Monero [33]. Awoke *et al.* also use LSTM and GRU to predict BTC price [34].

In addition to only using price data, other exogenous factors that directly and indirectly influence cryptocurrency prices have also been utilized for prediction. Biswas *et al.* use an LSTM-based prediction model that uses exogenous factors, including stock market capitalization, volume, distribution, and high-end delivery, to predict BTC prices [35]. Bai *et al.* suggest using other cryptocurrencies' price data to classify BTC price movement [36]. In addition, Alessandretti *et al.* introduce a machine learning-based prediction framework based on the daily price data of multiple cryptocurrencies [37]. Using a single-layer feedforward network (SLFN), Kurbucz utilizes transaction network data for BTC price movement prediction [38]. Like the above-mentioned works that take the joint interdependence of cryptocurrencies into account, this work also considers a wide variety of exogenous factors for price prediction. In particular, this work's proposed framework utilizes various on-chain data, which are associated with the blockchain technology that underlies cryptocurrencies, in a novel way.

Other works utilize social media data for price prediction. Mohapatra *et al.* introduce a real-time cryptocurrency price prediction platform with Twitter data that use sentiment analysis techniques based on natural language processing (NLP) algorithms [39]. Wolk proposes using various machine learning algorithms, such as support vector machine (SVM), decision tree (DT), and gradient boosting, for social media sentiment analysis-based cryptocurrency price prediction [40]. Based on Twitter data and Google trends, his work discusses social media's impacts on major cryptocurrency prices. The same user-based data types have been used with various deep learning methods to predict short-term price fluctuation [41]. Aggarwal *et al.* compare several neural network-based models, such as GRU, LSTM, and CNN, with varying hyperparameters in price prediction results [42]. Socioeconomic factors, including Twitter and gold price, are also utilized.

Shahbazi and Byun propose an RL-based price prediction for cryptocurrencies (e.g., LTC, Monero) [43]. In addition, inverse RL with agent-based modeling (ABM) predicts BTC price movements [44]. In their work, rather than modeling the relationship between BTC price and input variables, the synthetic behavior data of agents in a simulated market is modeled. Jiang and Liang use a CNN with the RL algorithm for optimal portfolio management [45]. Given the historical price data of cryptocurrency, the proposed model outputs weights for the portfolio given feasible sets. Betancourt and Chen also suggest using deep RL for optimal portfolio management with a dynamic number of cryptocurrency assets [46]. Schnaubelt uses RL methods to optimize cryptocurrency order placement so that the model learns optimal placement strategies that resemble realistic policies in established markets [47].

D. CHANGE POINT DETECTION (CPD)

CPD is a technique used to detect abrupt and significant changes in the behavior of time-series data [48]. There is a growing need to find ways to segment and detect edge, event, and anomaly in various application domains, such as finance [49], [50], genomics [51], speech recognition [52], and monitoring systems [53]. As the need for recognizing notable patterns in large time-series data increases, multiple works on CPD that consider computational complexity (i.e., time, cost) for practical applications have been conducted. For instance, binary segmentation (BS) [54] and variations of exact search algorithms based on dynamic programming (DP), such as the segment neighborhood method [55] and optimal partitioning algorithm [56], are designed to minimize cost function over the numbers and locations of change points. Among several existing CPD algorithms, pruned exact linear time (PELT) is the most widely used algorithm [57]. PELT not only efficiently reduces computational costs by adopting a pruning step that removes ineffective potential change point values but also achieves higher segmentation accuracy than other CPD methods [57], [58]. PELT's basic objective function is provided in (1).

$$\sum_{i=1}^{m+1}[\mathcal{C}(x_{(\tau_{i-1})+1:\tau_i})] + \beta f(m). \tag{1}$$

where:

m: number of change points
$\tau_{1:m} = (\tau_1, \ldots, \tau_m)$: timestamps (i.e., positions) of change points
$\mathcal{C}$: cost function for a segment
f: regularization term
β: regularization coefficient

Several works have applied CPD methods on economic and financial data for trend analysis and anomaly prediction. Fryzlewicz applies a variation of the BS algorithm called wild binary segmentation (WBS) on narrow time intervals to detect trends in the S&P 500 index [49]. Fu *et al.* propose a specialized binary tree representation for financial time-series segmentation [50]. Pepelyshev and Polunchenko suggest a real-time financial surveillance monitoring system that uses CPD methods to detect multi-cycling sequential changes and anomalies [59]. Zhu *et al.* utilize CPD to monitor and identify changes in risk dependence structure in banking data [60]. In addition, CPD uses stock return data to detect change points in a company's status [61]. Existing works on CPD applications have been focused on simply detecting change points in time-series data. However, in this work, CPD technique is further utilized for data normalization based on CPD segmentation results. To the best of our knowledge, there is no prior work that applies CPD to handle the data normalization of non-stationary time-series data (e.g., BTC prices).

III. PROPOSED PRICE PREDICTION FRAMEWORK

This section illustrates the overall framework proposed in this work. The proposed framework consists of five phases. First, extensive variable sets have been collected from on-chain data. Second, some variables are selected from the acquired dataset based on significant statistical correlations. Third, the segmentation of input data based on the CPD technique (PELT) is conducted. Fourth, the proposed price prediction model, composed of LSTM and the attention mechanism, is illustrated. Finally, an experimental setup, including data preprocessing, evaluation metrics, and implementation details, is provided. The complete algorithm for the proposed price prediction framework is presented in Algorithm 1.

A. EXTENSIVE VARIABLE COLLECTION

In this work, BTC price is selected as a prediction target due to several reasons. First, BTC not only has the largest number of transactions among existing cryptocurrencies but also has the largest number of owners (i.e., address/wallets) [62]. In addition, a large number of on-chain data has direct relationships with BTC and so are deemed effective in price prediction. For these reasons, this work, as well as many existing works in the literature, has focused on BTC price prediction using various data types. Daily BTC price data and other associated data that will be detailed later are collected from March 27, 2018 to November 16, 2021, having 1,331 time stamps (i.e., data points) in total. The BTC price trends used in this work are shown in Fig. 1.

Taking the distinct characteristics of blockchain-based cryptocurrencies (e.g., BTC) into account, diverse variables related to price are selected and used in this work. However, other exogenous variables that have been used in existing works (e.g., Twitter, Google Trends) in addition to on-chain data are not utilized, since one of this work's main contributions is to validate the efficacy of on-chain data-based cryptocurrency price prediction. To this end, an extensive variable collection is first conducted. On-chain data consisting of 254 variables are collected and used. Detailed information on all variables is shown in Table 8 in the Appendix A.

B. VARIABLE SELECTION

Since using all 254 variables collected is not only computationally but analytically excessive, a variable selection is performed. First, variables in available data with shorter periods and different time units (e.g., non-daily) are abandoned. Next, for a primary comparison of variable importance, a cross-correlation function (CCF) [63] is employed. According to (2), CCF values between the BTC price and each on-chain data variable, are computed using time lags from -3 to 3. Based on the computed CCF values, 42 variables with the highest CCF values, which are strongly related to BTC price,

Algorithm 1 The Proposed Price Prediction Framework That Takes Multivariate Data D Consisting of Price Data as Well as on-Chain Data From March 27, 2018 to November 16, 2021

Input: raw data $D = \{v_1, v_2, \ldots, v_{254}\}$ ▹ Containing 254 variables, where v_1 is BTC price
CCFs ← {}
for $i = 2$ **to** 254 **do** ▹ Compute CCF between BTC price and on-chain variables
compute $\rho_{v_1 v_i}$; CCFs.append($\rho_{v_1 v_i}$)
end for
sorted ← GetSordtedIndices(CCFs); topSet ← sorted[1:42]
$D_{selected}$ ← D[topSet] ▹ Select 42 variables with the highest CCFs
group $D_{selected}$ ▹ Divide variables into five groups
compute $\mathcal{S} \leftarrow PELT(D_{selected})$ ▹ Segment data using CPD
for all *segment* $s \in \mathcal{S}$ **do** ▹ Separately normalize data within each segmentation
normalize $D_{selected}[s]$
end for
transform $\{(X_b, y_b);\ b \in (1, \ldots, N)\} \leftarrow$ SidingWindow($D_{selected}$) ▹ Apply a sliding window method
define SAM-LSTM ▹ Employ the proposed prediction model
Train and evaluate SAM-LSTM.

are selected.

$$\rho_{yz}(s,t) = \frac{E[(y_s - \mu_{ys})(z_t - \mu_{zt})]}{\sqrt{E[(y_s - \mu_s)^2]E[(z_t - \mu_t)^2]}}. \tag{2}$$

where:

y, z: time-series data
s, t: time points
μ: mean value

After variable selection, 42 variables related to BTC price are divided into five groups: 1) price, 2) adoption, 3) distribution, 4) market, and 5) valuation. A detailed description of the variables that belong to each group is provided in Table 2. The criteria for grouping used in this work are as follows. The daily price of BTC, as well as other major cryptocurrencies (i.e., ETH, XRP, LTC), belongs to the group 'price'. Considering that the cryptocurrency market is unlike the stock market, 00:00 UTC is considered the price of the day. The variables that denote how much people participate in the cryptocurrency ecosystem and the fundamental information of the current blockchain phase belong to the group 'adoption'. Variables associated with the cryptocurrency distribution, including individual addresses, exchanges, and miners, fall under the group 'distribution'. Those related to total supply and demand and exchange supply and demand, which practically determine the real-time cryptocurrency prices in the market, belong to the group 'market'. This group also includes the data tracking activities of addresses with a certain balance amount and investors with distinct characteristics, such as long-term and short-term holders. For the last group 'valuation', variables used to check the fairness of cryptocurrency prices are chosen. The descriptive statistics of the selected variables that correspond to each group are shown in Table 3.

C. SEGMENTATION

As previously mentioned, the CPD technique is used in this work to segment the time-series data. In particular, PELT [57] is used as a component of the proposed framework. Using the algorithm, the BTC price time-series data are segmented into 6 groups that are deemed to have different statistical characteristics. Segmentation results, including a detailed time range, are provided in Table 4 and are visualized in Fig. 3. In particular, each segmented time-series data has a different mean value. For instance, the mean value of the sixth segment is more than ten times higher than that of the first segment, as shown in Table 4. In this work, the CPD-based segmentation results are utilized for data normalization. Each segmented time-series data is normalized (e.g., standardized) individually using one's own statistics. This has two benefits that help solve the aforementioned problem of machine learning prediction models not performing well on unseen price ranges. First, as data is normalized according to the statistics of each segmentation, the scale difference problem could be resolved. Second, as all data fall into certain ranges after normalization, price fluctuation trends and their implicit dynamics are better captured and reflected in the data. These two benefits of using the CPD technique resolve the scale problem and significantly increase the prediction performance of neural network-based models [64].

D. PRICE PREDICTION MODEL

1) LONG SHORT-TERM MEMORY (LSTM)

Analogous to many existing works on cryptocurrency prediction that use deep learning-based approaches, the proposed prediction model also employs LSTM [65], one of the most powerful RNN-based algorithms. Commonly used in sequence modeling tasks, such as natural language processing (NLP) and time-series forecasting, LSTM can handle data expressed in a sequential order [66]. Thus, LSTM is used as a primary feature extractor that directly takes time-series data as input to extract meaningful temporal information in the proposed model. LSTM recurrently updates hidden states using the output from the previous time step

TABLE 2. Description of selected variables used in the proposed framework.

Group (category)	Variable	Description
Price	BTC	Daily price of Bitcoin
	ETH	Daily price of Ethereum
	XRP	Daily price of Ripple
	LTC	Daily price of Litecoin
Adoption	BTC: Total addresses	Entire number of unique addresses identified through transaction of Bitcoin
	BTC: Block height	Total amount of blocks in Bitcoin blockchain
	BTC: Unspent transaction outputs value created (median)	Value of novel unspent transaction outputs (UTXOs) generated
	BTC: Unspent transaction outputs value spent (median)	Value of spent transaction outputs
	BTC: Difficulty	Estimated required number of hashes for mining one block of Bitcoin, which indicates the difficulty of mining
	USDT: Total addresses	Total number of unique addresses identified through transaction of USDT
	USDT: Transfer volume	Total number of USDT transferred successfully
Distribution	BTC: Thermocap	Cumulative amount of Bitcoin paid to miners converted to USD
	BTC: Exchange balance (stacked)	Amount of Bitcoin held at the address of each exchange converted (Only Bitcoin held at Binance)
	BTC: Addresses with non-zero balance	Number of addresses with more than 0 BTC
	BTC: Addresses with balance ≥ 0.01	Number of addresses with more than or equal to 0.01 BTC
	BTC: Addresses with balance ≥ 0.1	Number of addresses with more than or equal to 0.1 BTC
	BTC: Addresses with balance ≥ 10k	Number of addresses with more than or equal to 10k BTC
	USDT: Exchange balance	Total amount of USDT held at the exchange address
	USDT: Exchange balance (stacked)	Amount of USDT held at the address of each exchange (Only USDT held at Binance used in experiment)
	USDT: In-house exchange volume	Transaction of USDT within the same exchange
	USDT: Addresses with non-zero balance	Number of addresses with more than 0 USDT
Market	BTC: Market cap	Market capitalization of Bitcoin
	BTC: Delta cap	Gap between Realized Cap and Average Cap of Bitcoin
	BTC: Realized cap	Sum of each UTXO times price at the last point of movement of Bitcoin
	BTC: Investor capitalization	Gap between Realized Cap and Thermocap of Bitcoin
	BTC: Balanced price	Gap between Realized Price and Transfer Price of Bitcoin
	BTC: Realized price	Realized Cap divided by current number of Bitcoin issued
	BTC: Cumulative value days destroyed	Ratio of accumulated value of CDD and length of market existence of Bitcoin
	BTC: Stock-to-flow ratio	Ratio of current amount of Bitcoin circulating and new issuance
	BTC: Relative unrealized profit	Total amount of Bitcoin realized with lower price than the current price divided by the market cap
	BTC: Price drawdown from ATH (all time high)	Percentage drop in current price compared to high point of Bitcoin
	BTC: Unspent transaction outputs in profit	Amount of UTXOs of Bitcoin generated with lower price than the current price
	BTC: Supply last active 3+ years ago	The proportion of Bitcoin that hasn't been transferred for more than 3 years
	BTC: Supply last active 3y-5y	The proportion of Bitcoin that hasn't been transferred for 3 years to 5 years
	BTC: Supply last active 5y-7y	The proportion of Bitcoin that hasn't been transferred for 5 years to 7 years
	USDT: Market cap	Market capitalization of USDT
	USDT: Exchange inflow volume	Amount of USDT that flew in to exchange addresses
	USDT: Exchange outflow volume	Amount of USDT that flew out from exchange addresses
	USDT: Circulating supply	Total number of USDT issued
	USDT: Supply in smart contracts	Share of USDT in smart contracts
Valuation	BTC: Market value to realized value Z-score	Bitcoin's market cap subtracted by realized cap and divided by standard deviation of market cap
	BTC: Market value to realized value ratio	market cap and realized cap ratio of Bitcoin

TABLE 3. Descriptive statistics for selected variables.

Variable	N	Mean	Standard deviation	Minimum	Maximum
BTC price	1,331	17,714.9057	17,404.8761	3,228.7000	67,527.9000
ETH price	1,331	831.8834	1,100.8680	83.8100	4,808.3800
LTC price	1,331	96.1077	62.5293	23.1240	386.8200
XRP price	1,331	0.4655	0.3129	0.1360	1.8362
Addresses with balance $\geq$ 0.1	1,331	2,862,078.0691	295,347.7217	2,350,740.0000	3,259,148.0000
Addresses with balance $\geq$ 0.01	1,331	7,767,059.2344	1,010,615.0459	6,158,733.0000	9,203,146.0000
Addresses with balance $\geq$ 10k	1,331	104.6821	11.2237	81.0000	126.0000
Addresses with non-zero balance (BTC)	1,331	28,960,035.4808	5,703,196.8996	21,415,708.0000	38,704,373.0000
Addresses with non-zero balance (USDT)	1,331	1,238,117.6033	1,301,473.8884	262.0000	4,086,721.0000
Balanced price	1,331	6,999.6572	5,070.5749	3,329.4189	20,520.9904
Block height	1,331	613,322.9248	56,159.5391	515,462.0000	710,052.0000
Circulating supply	1,331	17,104,228,969.2182	22,233,387,666.2657	1,706,421,736.0000	73,864,417,288.5211
Cumulative value days destroyed	1,331	4,827.3506	2,610.8689	2,332.1223	11,883.0511
Delta cap	1,331	112,252,992,711.9373	88,180,431,583.1644	4,974,482,038.0194	342,161,722,954.9807
Difficulty	1,331	5.5884	2.5908	1.4871	1.0757
Exchange balance (stacked) (BTC)	1,331	299,242.5315	108,226.2316	170,436.5106	570,838.0560
Exchange balance (stacked) (USDT)	1,331	774,264,980.0331	1,053,516,454.5351	1,170.0000	4,191,935,324.0250
Exchange balance	1,331	1,789,225,502.8947	2,078,406,229.4885	2,954,185.6958	7,324,592,968.865815
Exchange inflow volume	1,331	348,221,961.7426	453,202,347.4026	290.0000	3,688,467,641.3318
Exchange outflow volume	1,331	309,126,772.4217	414,926,368.2728	0.0000	4,002,385,372.8736
In-house exchange volume	1,331	340,756,767.9706	492,807,550.4803	0.0000	5,208,459,205.7859
Investor capitalization	1,331	141,083,777,113.4224	105,980,122,307.4554	65,808,914,014.6945	425,251,898,231.3604
Market cap (BTC)	1,331	326,824,735,650.8270	328,789,045,791.6564	56,333,812,750.1928	1,274,451,962,953.7795
Market cap (USDT)	1,331	17,117,586,244.4681	22,240,984,075.4490	1,675,066,781.2009	73,890,294,466.9732
Market value to realized value ratio	1,331	1.7739	0.6674	0.6983	3.9537
Market value to realized value Z-score	1,331	1.6096	1.5932	-0.5086	7.6304
Price drawdown from ATH	1,331	-0.4897	0.2242	-0.8366	0.0000
Realized cap	1,331	157,383,002,485.2837	112,141,509,617.3931	76,517,682,919.4828	458,638,799,881.7129
Realized price	1,331	8,582.6425	5,865.0716	4,342.4790	24,298.3179
Relative unrealized profit	1,331	0.4825	0.1062	0.3059	0.7488
Stock-to-flow ratio	1,331	28,760.6863	36,165.1569	5,695.6180	111,085.8028
Supply in smart contracts	1,331	0.0670	0.0639	0.0042	0.2079
Supply last active 3y-5y	1,331	1,637,039.8374	554,751.7236	832,182.9024	2,527,782.2552
Supply last active 5y-7y	1,331	979,044.2697	217,153.7991	554,106.4684	1,303,287.5808
Supply last active 3+ years ago	1,331	0.3007	0.0316	0.2621	0.3538
Thermocap	1,331	16,299,225,371.8614	6,806,221,655.7357	6,671,516,975.3491	33,386,901,650.3526
Total addresses (BTC)	1,331	620,585,495.1307	154,836,111.6801	383,180,570.0000	903,090,061.0000
Total addresses (USDT)	1,331	5,757,287.7783	6,544,219.3044	705.0000	19,672,721.0000
Transfer volume	1,331	2,259,728,157.4901	3,021,315,774.5870	40,598.2903	24,517,329,457.5227
Unspent transaction outputs in profit	1,331	77,439,171.2945	3,1059,571.3369	28,789,464.0000	125,430,945.0000
Unspent transaction outputs value created	1,331	0.0078	0.0030	0.0019	0.0200
Unspent transaction outputs value spent	1,331	0.0096	0.0042	0.0002	0.0257

as a current input to learn sequentiality within time-series data. In addition, by using a cell state, LSTM not only learns long-term dependencies in sequences but also prevents gradient vanishing and exploding problems from happening. The cell state, the so-called memory cell, retains important sequential information during training. In addition, the

TABLE 4. Segmentation results using PELT.

Segment No.	Start date	End date	Mean	Standard deviation
1	March 27, 2018	May 26, 2019	5920.7724	1700.0463
2	May 27, 2019	December 16, 2020	9985.8247	2727.3034
3	December 17, 2020	February 9, 2021	32388.3254	5763.3584
4	February 10, 2021	May 15, 2021	54502.9821	4550.8680
5	May 16, 2021	October 7, 2021	40691.4296	6140.0379
6	October 8, 2021	November 16, 2021	61241.2853	3405.6998

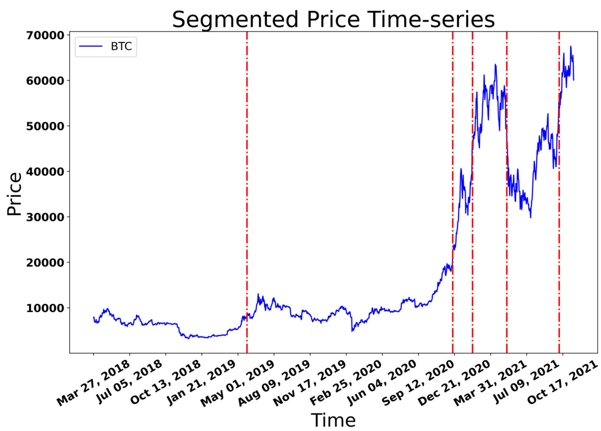

FIGURE 3. Segmentation of time-series data using CPD.

distinct internal mechanism of LSTM contains three gates (i.e., input, output, forget gates) that control the amount of information to be forgotten or updated [67]. The detailed architecture of LSTM is shown in Fig. 4. The computation of LSTM at time step t is defined as follows:

$$f_t = \sigma(W_f x_t + U_f h_{t-1}). \tag{3}$$
$$\bar{C}_t = tanh(W_c x_t + U_c h_{t-1}). \tag{4}$$
$$i_t = \sigma(W_i x_t + U_i h_{t-1}). \tag{5}$$
$$o_t = \sigma(W_o x_t + U_o h_{t-1}). \tag{6}$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \bar{C}_t. \tag{7}$$
$$h_t = o_t \odot tanh(C_t). \tag{8}$$

where:

$W_f, W_c, W_i, W_o, U_f, U_c, U_i, U_o$: trainable weight matrices containing bias terms
$\odot$: Hadamard product

2) ATTENTION MECHANISM

The attention mechanism, designed to help a model learn where to attend to during training, is an additional component that can be used within a neural network architecture [68]. Recently, attention has become a basis of modern seq2seq

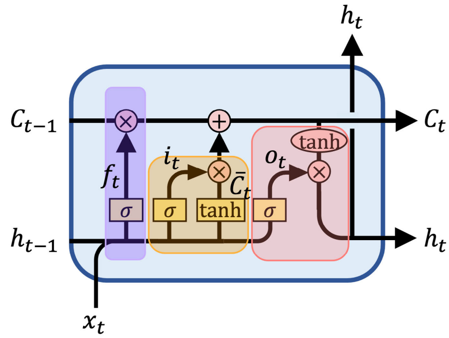

FIGURE 4. An internal architecture of long short-term memory (LSTM).

methods (e.g., BERT, ELMo, transformer) in sequence modeling for machine translation and document classification [69]–[71]. The attention mechanism is especially useful in situations where a large amount of information is provided to the model so that the model must selectively exploit important parts. In this regard, the proposed model of this work employs an attention mechanism in order to learn to attend to various input variables that more accurately predict BTC price. Multiple experiments in the next section further validate the effectiveness of using the attention mechanism in the proposed model. The attention mechanism also provides model interpretability to some degree [72] by using the computed attention scores that show where and how much the trained network has learned to attend to the training data.

There are several forms of the attention mechanism with varying internal computations, placement within neural networks, and softness [73]. Because LSTM uses multivariate time-series data for prediction in this work, the attention mechanism is attached to the end of LSTM module. A self-attention, the so-called Bahdanau attention, which uses representations (i.e., annotations as in [68]) generated from the LSTM module, is employed in the proposed model as shown in Fig. 5. The computation of the attention mechanism is shown as follows:

$$e_j = a(h_T, h_j). \tag{9}$$
$$\alpha_j = \frac{exp(e_j)}{\sum_{k=1}^{T} exp(e_k)}. \tag{10}$$
$$c = \sum_{j=1}^{T} \alpha_j h_j. \tag{11}$$

where:

j: timestamp within time-series data for $j = 1 \cdots T$
a: alignment module parametrized as a feedforward neural network with a tanh activation

3) PROPOSED ARCHITECTURE

As the selected variables are separated into five groups (i.e., price, adoption, distribution, market, valuation), the LSTM module is constructed for each variable group that has hidden units proportional to the number of variables,

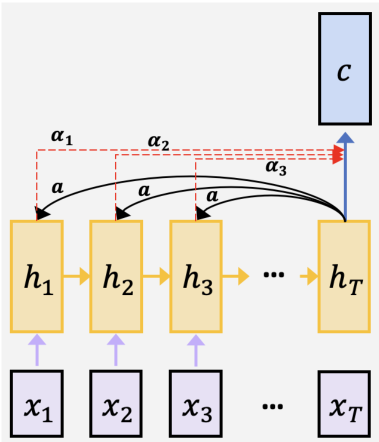

FIGURE 5. A diagram of the attention mechanism.

as shown in Fig. 2. In particular, each LSTM module for five variable groups has hidden units with size 32, 18, 22, 42, and 8, respectively. Furthermore, for each LSTM module, the attention mechanism is employed so that the final output is computed using representations with the weights of the computed attention scores. Using multiple LSTM for different variable groups associated with BTC prices should have several benefits. First, each variable group, divided according to domain knowledge, is used as a separate input for multiple LSTM modules, thus encouraging the model to learn optimal price dynamics. This use of LSTM with separate variable groups can also be seen as one type of secondary supervision for network training. Furthermore, different attention mechanisms employed for the LSTM modules learn to attend differently to extract richer and diverse timely information from the multivariate input data.

As mentioned above, multiple LSTM modules with attention mechanisms extract meaningful feature representations from input data. In order to utilize the extracted representations to predict future prices, MLP is employed in the proposed model so as to aggregate them and make price prediction. MLP is a standard neural network architecture that contains multiple fully connected layers and nonlinear activation functions between layers. This architecture is not only flexible to manipulate but also proven to model very complex underlying functions [74], which is a desirable characteristic for cryptocurrency prediction. An ordinary form of a single layer computation of MLP is given in (12). In the proposed model, an MLP-based prediction module is used on top of multiple LSTM module outputs.

$$h_{i+1} = g(W_i h_i). \tag{12}$$

where:

h_i: i^{th} representation
W_i: weights
g: activation function

The entire architecture, including multiple LSTM modules with the attention mechanism and the MLP-based prediction model, is dubbed SAM-LSTM, which stands for self-attention-based multiple LSTM. To reiterate, SAM-LSTM consists of two parts. First is a primary feature extractors using multiple LSTM modules for different variable groups with distinct self-attention mechanisms. Second is an MLP-based module that aggregates extracted representations from multivariate time-series inputs to predict the final price.

E. EXPERIMENTAL SETUP

This section illustrates data preprocessing, evaluation metrics for prediction performance, and implementation details. Data preprocessing includes 1) data normalization based on segmentation results using PELT [57] and 2) data transformation using a sliding window method. The experiments conducted in this work consist of two parts. First, LSTM-based models are compared in terms of BTC price prediction performance. Using only historical BTC price data, three methods are compared in an ablative manner: 1) LSTM without CPD (i.e., naïve LSTM); 2) LSTM with CPD; and 3) LSTM with CPD and the attention mechanism. Second, using the selected variables that are separately grouped as inputs, the proposed method (i.e., SAM-LSTM) is employed to predict BTC prices. Extensive variable group combinations are tested within the proposed SAM-LSTM.

1) DATA PREPROCESSING

To validate the effectiveness of the proposed SAM-LSTM to predict BTC prices, a series of data preprocessing have been conducted. First, each input variable is normalized via standardization according to (13). When CPD (i.e., PELT [57]) is used, each segmentation of the entire time-series data is separately normalized using its own statistics (i.e., mean and standard deviation). In this way, intense fluctuations are better reflected in a moderate manner after data normalization. Figure 6 visualizes the effects of the proposed CPD-based normalization, in which data show fluctuations in a modest range, while still maintaining distinctive patterns.

$$x_{scaled} = \frac{x - \bar{x}}{\sigma}. \tag{13}$$

where:

x: original independent variable
$\bar{x}$: mean value
σ: standard deviation value
x_{scaled}: scaled independent variable

In addition, like many other existing works on the price prediction, a sliding window method is employed for data preparation in this work [75]. Since the existing BTC price data, as well as multivariate time-series data, are lengthy (length=1,331), the sliding window is applied to crop the given data into multiple samples. Considering weekly BTC price trends, the sliding window method with n=7, following [1], is selected. This selection makes the prediction model take a 7-day long multivariate time-series data as input and outputs the next day's future price (i.e., price of day 8). The

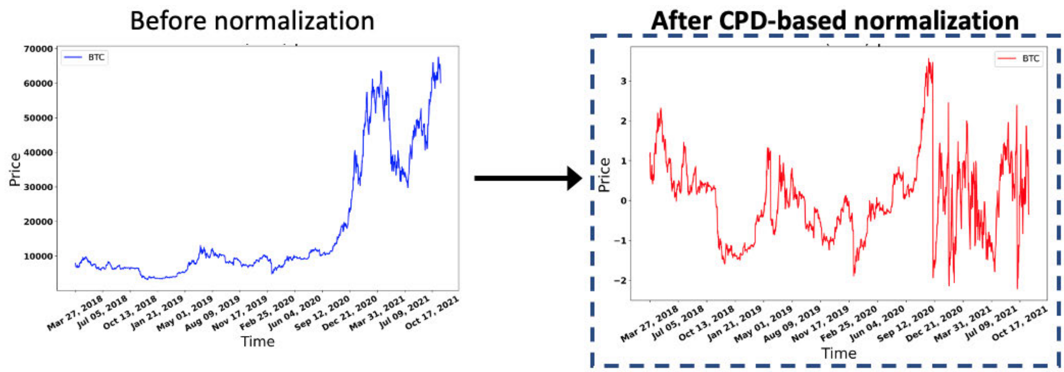

FIGURE 6. Effects of CPD-based normalization preprocessing.

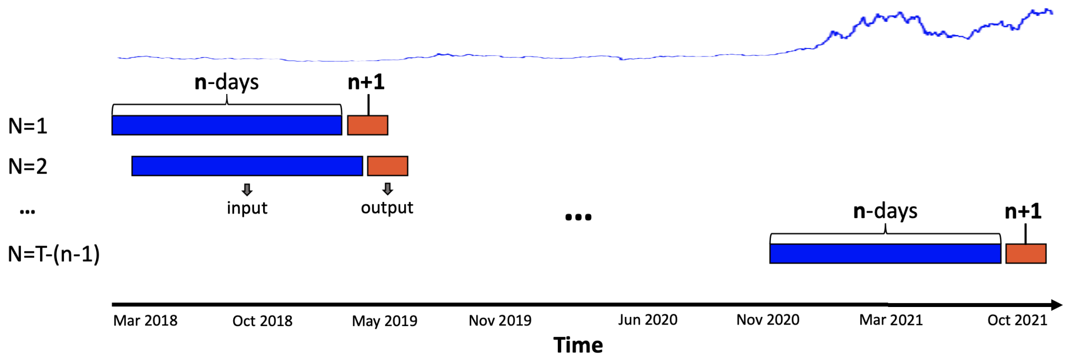

FIGURE 7. Sliding window for time-series data preparation.

sliding window method applied in this work is illustrated in Fig. 7.

Since the prediction uses sequential multivariate time-series data, the holdout sets are divided with a ratio of 8:2 so that the first 80% of the data are used for training, while the remaining 20% are used for evaluation.

2) EVALUATION METRICS

Since the target is a price, which is a continuous variable, several evaluation metrics for regression tasks are selected to validate the experimental results. Mean squared error (MSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean squared error (RMSE) are used. In particular, MSE is used as a loss function for network training. The equations for the metrics are defined as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|x_t - \hat{x}_t|. \tag{14}$$

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x}_t)^2. \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x}_t)^2}. \tag{16}$$

$$MAPE = \frac{100}{N}\sum_{t=1}^{N}\frac{|x_t - \hat{x}_t|}{|x_t|}. \tag{17}$$

TABLE 5. Model performance comparison using univariate price data.

Model	MAE	RMSE	MSE	MAPE
LSTM	2.2502 (±0.5956)	2.3346 (±0.5735)	5.7794 (±0.7096)	**1.3189** (±0.3968)
LSTM + CPD	0.4473 (±0.2039)	0.6069 (±0.1953)	0.4065 (±0.2940)	1.7209 (±0.2347)
LSTM + CPD + Attention **(proposed)**	**0.4270** (±0.025)	**0.5944** (±0.0064)	**0.3533** (±0.0063)	1.5381 (±0.1421)

where:

x_t: actual value
$\hat{x}_t$: predicted value
N: number of data

3) IMPLEMENTATION

The experiments are conducted on a graphical processing unit (GPU) Tesla V100 with an open source library TensorFlow [76]. During model training, Adam [77] with a learning rate of 0.01 with a decay rate of 0.0005, β_1 of 0.91, and β_2 of 0.98, is used as an optimization algorithm with a batch size of 32. Parameter optimization for the LSTM modules in the proposed prediction model utilizes backpropagation through time (BPTT) algorithm, which is one of the simplest and intuitive ways to train recurrent models based on gradient based optimization technique, as detailed in Appendix B. In addition, layer normalization [78] is used inside the network architecture to stabilize the training process. In order to prevent model overfitting, early stopping, which terminates

TABLE 6. Performance comparison of the proposed method with different configurations.

Model	Used variable group					MAE	RMSE	MSE	MAPE
	Price	Adoption	Distribution	Market	Valuation				
Attention-LSTM + CPD	○	×	×	×	×	0.4142 ±0.0169	0.5907 ±0.0212	0.3493 ±0.0250	1.7522 ±0.1964
	○	○	×	×	×	0.3731 ±0.0083	0.5188 ±0.0066	0.2962 ±0.0069	1.7672 ±0.1202
	○	×	○	×	×	0.4094 ±0.0279	0.5778 ±0.0260	0.3345 ±0.0302	1.7408 ±0.1139
	○	×	×	○	×	0.3996 ±0.0261	0.5563 ±0.0440	0.3114 ±0.0498	1.7606 ±0.1725
	○	×	×	×	○	0.3695 ±0.0027	0.5134 ±0.0041	0.2636 ±0.0042	1.6648 ±0.1189
	○	○	○	×	×	0.3756 ±0.0034	0.5196 ±0.0026	0.2699 ±0.0027	1.6591 ±0.0606
	○	○	×	○	×	0.3748 ±0.0051	0.5186 ±0.0083	0.2690 ±0.0087	1.5379 ±0.0627
	○	○	×	×	○	0.3749 ±0.0047	0.5203 ±0.0044	0.2708 ±0.0046	1.6666 ±0.0590
	○	×	○	○	×	0.3889 ±0.0142	0.5360 ±0.0161	0.2875 ±0.0171	1.7155 ±0.1791
	○	×	○	×	○	0.3996 ±0.0261	0.5563 ±0.0440	0.3114 ±0.0498	1.7606 ±0.1725
	○	×	×	○	○	0.3833 ± 0.0071	0.5180 ±0.0040	0.2684 ±0.0041	1.5314 ±0.0590
	○	○	○	○	×	0.3858 ±0.0044	0.5312 ±0.0063	0.2822 ±0.0067	1.6397 ±0.1967
	○	○	○	×	○	0.3751 ±0.0034	0.5176 ±0.0039	0.2680 ±0.0040	1.6692 ±0.1061
	○	○	×	○	○	0.3837 ±0.0135	0.5327 ±0.0194	0.2841 ±0.0210	1.5889 ±0.1249
	○	×	○	○	○	0.3761 ±0.0040	0.5188 ±0.0025	0.2691 ±0.0026	1.6677 ±0.1492
	○	○	○	○	○	0.3853 ±0.0101	0.5249 ±0.0107	0.2757 ±0.0111	1.5478 ±0.1774
SAM-LSTM + CPD **(proposed)**	○	×	×	×	×	0.4089 ±0.0536	0.5720 ±0.0585	0.3306 ±0.0674	2.1788 ±0.4426
	○	○	×	×	×	0.3767 ±0.0345	0.5444 ±0.0442	0.2983 ±0.0483	1.8589 ±0.3083
	○	×	○	×	×	0.3794 ±0.0232	0.5433 ±0.0178	0.2955 ±0.0193	1.4779 ±0.2768
	○	×	×	○	×	0.3803 ±0.0035	0.5450 ±0.0145	0.2973 ±0.0160	1.8273 ±0.3047
	○	×	×	×	○	0.3547 ±0.0120	0.5233 ±0.0182	0.2742 ±0.0193	1.7531 ±0.3858
	○	○	○	×	×	0.3601 ±0.0146	0.5234 ±0.0162	0.2472 ±0.0171	1.6720 ±0.1881
	○	○	×	○	×	**0.3462** ±0.0066	**0.5035** ±0.0085	**0.2536** ±0.0085	1.6897 ±0.0895
	○	○	×	×	○	0.3728 ±0.0329	0.5346 ±0.0284	0.2867 ±0.0308	1.6687 ±0.0890
	○	×	○	○	×	0.3475 ±0.0050	0.5072 ±0.0105	0.2574 ±0.0106	1.6191 ±0.1592
	○	×	○	×	○	0.3680 ±0.0151	0.5347 ±0.0239	0.2864 ±0.0259	1.4662 ±0.0638
	○	×	×	○	○	0.3816 ± 0.0163	0.5281 ±0.0247	0.2795 ±0.0262	**1.3251** ±0.1324
	○	○	○	○	×	0.3531 ±0.0099	0.5122 ±0.0079	0.2624 ±0.0081	1.6093 ±0.1804
	○	○	○	×	○	0.3601 ±0.0080	0.5170 ±0.0155	0.2674 ±0.0119	1.5900 ±0.1635
	○	○	×	○	○	0.3532 ±0.0174	0.5160 ±0.0201	0.2666 ±0.0210	1.4432 ±0.0497
	○	×	○	○	○	0.3704 ±0.0130	0.5342 ±0.0164	0.2856 ±0.0175	1.6465 ±0.1601
	○	○	○	○	○	0.3561 ±0.0093	0.5214 ±0.0090	0.2719 ±0.0093	1.4471 ±0.1430

the training procedure when the validation loss seems to be saturated, is employed. During the experiments, ten trials that use distinct random seeds are conducted for each configuration to obtain the error bars (i.e., standard deviations) for the results.

IV. RESULTS AND DISCUSSION

This section illustrates experimental results based on two parts. First, the effects of CPD and the attention mechanism utilized within the proposed SAM-LSTM method are validated through the task of predicting univariate BTC prices. After confirming the efficacy of the two techniques, the price prediction performance of SAM-LSTM is verified.

A. CPD AND ATTENTION MECHANISM

The first set of experiments compares the methods with and without two techniques: CPD (i.e., PELT [57]) and the attention mechanism. In order to first verify their efficacy, only historical BTC price data are used as an input for the models. A comparison of the three methods' price prediction performances is shown in Table 5. When comparing LSTM without CPD to LSTM with CPD, the latter shows significantly improved prediction performance in terms of three metrics: MAE, RMSE, and MSE. In fact, as shown in Fig. 8, LSTM without CPD is unable to learn intense fluctuations, especially during the price upsurge phase. This presents two inherent problems with using LSTM in a naïve

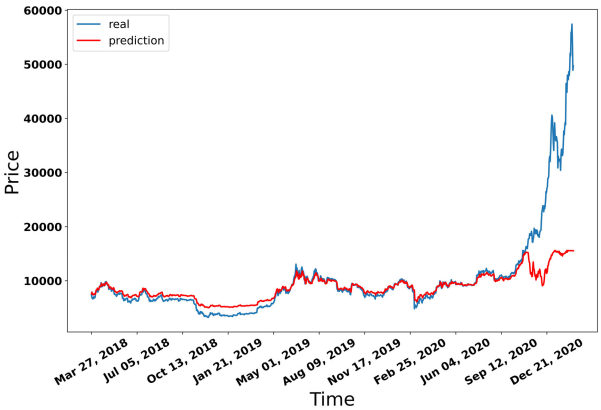

FIGURE 8. **The wrong prediction results from LSTM without CPD.**

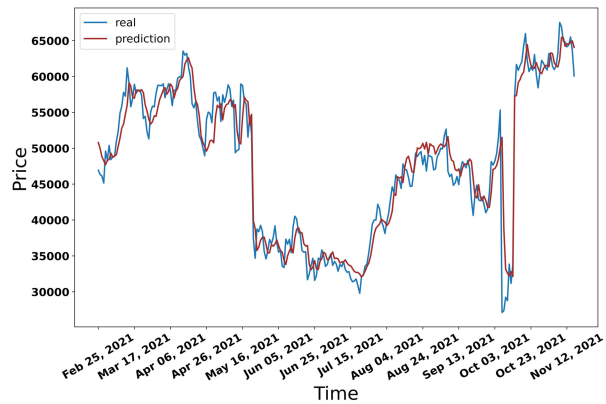

FIGURE 9. **Prediction results.**

manner. First, since the time-series data are normalized using all-time statistics, including that of price upsurge and decline, the fluctuation trends are not well reflected in the data, so the model cannot learn well. In addition, especially during the test stage, the trained model cannot effectively predict unseen data, which are outside of the price range of training data. The use of CPD (i.e., PELT), one of the techniques proposed in this work, has shown to handle the aforementioned problems well. In addition to using CPD for data normalization, the attention mechanism is attached to LSTM, and the performance of the model with CPD and the attention mechanism outperforms the model without the attention mechanism, as shown in Table 5. In fact, the proposed model using both CPD and attention mechanism shows better scores in terms of MAE, RMSE, and MSE, except for MAPE. MAPE of the proposed model that is slightly worse than a naïve LSTM might be due to the fact that MAPE uses ground truth value as a denominator for calculation, which is different for each model configuration. Thus, based on the aforementioned results, the efficacy of CPD and the attention mechanism are validated.

B. SAM-LSTM

As the use of CPD and the attention mechanism have shown to be effective in price prediction, the proposed method, which employs both, is used for price prediction. In particular, multivariate on-chain data are used as an input for the models. The proposed SAM-LSTM employs multiple LSTM modules with the attention mechanism and an MLP-based prediction module. Each LSTM module takes each group's input variables separately, and the MLP-based module aggregates the extracted representation and outputs a final price prediction result. A comparison of the two models, 1) LSTM with CPD and the attention mechanism that uses on-chain data and 2) SAM-LSTM that uses on-chain data, is detailed in Table 6. While the number of used variable groups as inputs is the same for the two models, internal architectures are different. The input variables are fed into the model as a whole for the first model whereas in the second model (SAM-LSTM), multiple LSTM modules take corresponding variables in each group separately. Using various variable group configurations according to Table 2, BTC price prediction tasks are conducted. In Table 6 the cells with a gray background denote a configuration in which the proposed SAM-LSTM outperforms LSTM without CPD. The numbers in *bold* denote the best-performing configuration and model. Based on the results, all configurations that use various on-chain input data better perform than those that only use univariate price data (compared with Table 5). This indicates that using rich, expressive multivariate on-chain data as an input is effective for price prediction. Furthermore, SAM-LSTM outperforms all other input variable group configurations in 15 out of 16 configurations in terms of MAE. In terms of RMSE and

TABLE 7. **Attention score distribution.**

SAM-LSTM module	Day	Mean	Standard deviation
Price	1	0.0634	0.0026
	2	0.0633	0.0028
	3	0.0624	0.0029
	4	0.0621	0.0029
	5	0.0620	0.0030
	6	0.2328	0.0373
	7	0.4548	0.0226
Adoption	1	0.0775	0.0078
	2	0.0809	0.0100
	3	0.0748	0.0079
	4	0.0737	0.0067
	5	0.1363	0.0221
	6	0.1260	0.0224
	7	0.3864	0.0375
Market	1	0.0670	0.0035
	2	0.0642	0.0031
	3	0.0625	0.0033
	4	0.0619	0.0034
	5	0.0643	0.0023
	6	0.2404	0.0453
	7	0.4393	0.0310

TABLE 8. Description of all variables (bolded variables are used for the proposed method).

Category	Variable
Price	**Bitcoin (BTC)**, **Ethereum (ETH)**, **Ripple (XRP)**, **Litecoin (LTC)**, USDT, Price OHLC (BTC), Price OHLC (ETH), Price OHLC (XRP), Price OHLC (LTC), Price OHLC (USDT)
Adoption	BTC: Active addresses, Sending addresses, Receiving addresses, New addresses, **Total addresses**, Average spent output lifespan, Median spent output lifespan, Spent output age bands, Spent outputs < 1h, Spent outputs 1h-24h, Spent outputs 1d-1w, Spent outputs 1w-1m, Spent outputs 1m-3m, Spent outputs 3m-6m, Spent outputs 6m-12m, Spent outputs 1y-2y, Spent outputs 2y-3y, Spent outputs 3y-5y, Spent outputs 5y-7y, Spent outputs 7y-10y, Spent outputs 10y >, Spent volume age bands, Spent volume < 1h, Spent volume 1h-24h, Spent volume 1d-1w, Spent volume 1w-1m, Spent volume 1m-3m, Spent volume 3m-6m, Spent volume 6m-12m, Spent volume 1y-2y, Spent volume 2y-3y, Spent volume 3y-5y, Spent volume 5y-7y, Spent volume 7y-10y, Spent volume 10y >, **Block height**, Block mined, Block interval (mean), Block interval (median), Block size, Block size (mean), **Difficulty**, Hash rate, Transaction count, Transaction rate, Transaction size, Transaction size (mean), Transfer volume, Transfer volume (median), Transfer volume (mean), Change-adjusted volume, Change-adjusted volume (mean), Change-adjusted volume (median), Unspent transaction output, Unspent transaction outputs created, Unspent transaction outputs spent, Unspent transaction outputs value created, Unspent transaction outputs value created (mean), **Unspent transaction outputs value created (median)**, Unspent transaction outputs value spent, Unspent transaction outputs value spent (mean), **Unspent transaction outputs value spent (median)**, Fee ratio multiple, Difficulty ribbon, Difficulty ribbon compression, Hash ribbon USDT: Active addresses, Sending addresses, Receiving addresses, New addresses, **Total addresses**, Transaction count, Transaction rate, **Transfer volume**, Transfer volume (mean), Transfer volume (median), Uniswap transactions, Uniswap volume, Uniswap liquidity
Distribution	BTC: Addresses with non-zero balance, **Addresses with balance $\geq$ 0.01**, **Addresses with balance $\geq$ 0.1**, **Addresses with balance $\geq$ 1**, **Addresses with balance $\geq$ 10**, **Addresses with balance $\geq$ 100**, **Addresses with balance $\geq$ 1k**, **Addresses with balance $\geq$ 10k**, Exchange balance (percent), **Exchange balance (stacked)**, Exchange balance, In-house exchange volume, Inter-exchange transfer, Inter-exchange volume, Fees, Fees (mean), Fees (median), Miner revenue, Miner revenue (fees), Miner revenue (rewards), **Thermocap**, Marketcap-to-thermocap ratio USDT: Addresses with non-zero balance, Supply of top 1 percent addresses, Gini coefficient, Herfindahl index, Exchange balance (percent), **Exchange balance (stacked)**, **Exchange balance**, **In-house exchange volume**, Inter-exchange transfer, Inter-exchange volume
Market	BTC: Exchange net position change, Exchange inflow volume, Exchange inflow volume (mean), Exchange outflow volume, Exchange outflow volume (mean), Exchange netflow volume, Exchange deposits, Exchange withdrawals, Coin days destroyed, Supply adjusted coin days destroyed, Binary coin days destroyed, Reserve risk, **Cumulative value days destroyed**, Coin years destroyed, Supply adjusted coin years destroyed, 90D coin days destroyed, **Balanced price**, Dormancy, Supply adjusted dormancy, Entity adjusted dormancy flow, Liveliness, Supply last active 2+ years ago, **Supply last active 3+ years ago**, Supply last active 5+ years ago, Supply last active < 24h, Supply last active 1d-1w, Supply last active 1w-1m, Supply last active 1m-3m, Supply last active 3m-6m, Supply last active 6m-12m, Supply last active 1y-2y, Supply last active 2y-3y, **Supply last active 3y-5y**, **Supply last active 5y-7y**, Supply last active 7y-10y, Supply last active > 10y, Stablecoin supply ratio, Stablecoin supply ratio oscillator, **Stock-to-flow ratio**, Stock-to-flow deflection, Realized profit, Realized loss, Net realized profit/loss, Realized P/L ratio, Realized profits-to-value ratio, **Investor capitalization**, Net unrealized profit/loss, **Relative unrealized profit**, Relative unrealized loss, Issuance, Inflation rate, **Circulating supply**, Adjusted supply, **Market cap**, **Realized cap**, Average cap, **Delta cap**, **Realized price**, Percent supply in profit, Supply in profit, Supply in loss, Percent unspent transaction outputs in profit, **Unspent transaction outputs in profit**, Unspent transaction outputs in loss, **Price drawdown from ATH (all time high)**, HODL waves, Realized cap HODL waves, RHODL ratio, Seller exhaustion constant, Bitcoin volatility index USDT: **Supply in smart contracts**, Exchange net position change, **Exchange inflow volume**, Exchange inflow volume (mean), **Exchange outflow volume**, Exchange outflow volume (mean), Exchange netflow volume, Exchange deposits, Exchange withdrawals, Circulating supply, **Market cap**
Valuation	BTC: Puell multiple, **Market value to realized value ratio**, **Market value to realized value Z-score**, NVT ratio, NVT signal, Velocity, Spent output profit ratio, Adjusted spent output profit ratio USDT: NVT ratio, NVT signal, Velocity

MSE, SAM-LSTM outperforms the others using the same configuration in 9 out of 16 configurations. This indicates that when rich multivariate input data are used for price prediction, SAM-LSTM, which uses multiple LSTM modules for each variable group with the attention mechanism, better learns the underlying dynamics that influence BTC prices. The prediction results for the best performing SAM-LSTM with CPD are visualized in Fig. 9.

The variable groups used in the best performing SAM-LSTM model with CPD applied, which are 'price', 'adoption', and 'market', can be found in Table 2. Using the best performing SAM-LSTM after training, attention scores are computed using the test data. For each data sample, the attention scores for each LSTM module are computed according to (10). For each separate attention mechanism attached to multiple LSTM modules, the computed attention score distribution is shown in Table 7. The different SAM-LSTM modules are shown to have learned where (i.e., which time stamp of the given multivariate on-chain data) to attend to during training for price prediction. In particular, each module has learned different patterns within each dataset type, as shown in the attention score distributions in Table 7. For instance, the attention scores for variable group 'adoption' has higher values for days 5 and 6 compared to that for variable groups 'price' and 'market'. In addition, attention scores for the 'market' group differ from the 'price' group. It is important to note that during training, data closer to the prediction point have been shown to have more influence. Specifically, according to the attention scores, the input data of the two to three latest days have shown to most influence price prediction.

V. CONCLUSION AND FUTURE WORKS

The authors propose a novel approach that uses multivariate on-chain time-series data to predict cryptocurrency prices. BTC price prediction is conducted within the proposed approach. Unlike traditional machine learning-based models, a CPD-based normalization technique enables price prediction models to predict unseen price ranges. Various on-chain variables are selected, grouped according to their inherent characteristics, and used as input variables for price prediction. The proposed price prediction model (i.e., SAM-LSTM), which consists of multiple LSTM modules with separate attention mechanisms and an MLP-based aggregation module, extracts distinctive features from grouped on-chain variables.

This work has five main steps. First, extensive variable collection using on-chain data is conducted. Second, based on CCFs, significant on-chain variables are selected as input variables and grouped according to their characteristics. Third, time-series data are segmented and normalized within each segmentation using a CPD technique called PELT [57]. Fourth, SAM-LSTM, which uses multiple LSTM for distinct on-chain variable groups and attention mechanism for price prediction, is proposed. Finally, the effectiveness of using CPD and SAM-LSTM in BTC price prediction is verified through rigorous experiments.

One limitation of this work is lack of a performance comparison with existing cryptocurrency price prediction methods. In fact, there are several reasons for being unable to conduct comparative experiments. First, each work in the literature uses different input data in terms of time spans, input data types (e.g., social media data, Google Trends), preprocessing steps, etc. In particular, existing works that use price data before recent plummet are not guaranteed to yield similar prediction results. In a similar vein, comparison with recent studies that claim to have considerable price prediction performances, such as [79]–[81], will be conducted in the future. Developing a holistic framework for cryptocurrency price prediction remains one possible future work. In particular, a unified framework that uses variables associated with prices, including on-chain data and social media data, should not only be developed, but a comprehensive aggregation model should also be created to model the price dynamics of the cryptocurrency market [6], [82]. In addition, a real-time price prediction model that uses various input data to make more frequent predictions (like hourly or per-minute) is worth being developed.

APPENDIX A

See Table 8.

APPENDIX B

Parameters of LSTM modules (W_f and U_f) in the proposed price prediction framework is optimized based on backpropagation through time (BPTT) as shown below.

$$\frac{\partial \mathcal{L}}{\partial U_f} = \sum_t \sum_{k=1}^{t+1} \frac{\partial \mathcal{L}(t+1)}{\partial \hat{x}_{t+1}} \frac{\partial \hat{x}_{t+1}}{\partial h_{t+1}} \frac{\partial h_{t+1}}{\partial h_k} \frac{\partial h_k}{\partial U_f}. \tag{18}$$

$$\frac{\partial \mathcal{L}}{\partial W_f} = \sum_t \sum_{k=1}^{t+1} \frac{\partial \mathcal{L}(t+1)}{\partial \hat{x}_{t+1}} \frac{\partial \hat{x}_{t+1}}{\partial h_{t+1}} \frac{\partial h_{t+1}}{\partial h_k} \frac{\partial h_k}{\partial W_f}. \tag{19}$$

ACKNOWLEDGMENT

The data used in this work were provided by Glassnode and can be accessed at https://glassnode.com/. Full attribution of the data is given to Glassnode and the authors are very grateful for allowing the usage of the on-chain data for academic use.

(Gyeongho Kim and Dong-Hyun Shin are co-first authors.)

REFERENCES

[1] P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar, and M. Alazab, "Stochastic neural networks for cryptocurrency price prediction," *IEEE Access*, vol. 8, pp. 82804–82818, 2020, doi: 10.1109/ACCESS.2020.2990659.

[2] L. Herskind, P. Katsikouli, and N. Dragoni, "Privacy and cryptocurrencies—A systematic literature review," *IEEE Access*, vol. 8, pp. 54044–54059, 2020, doi: 10.1109/ACCESS.2020.2980950.

[3] M. Kubát, "Virtual currency bitcoin in the scope of money definition and store of value," *Proc. Econ. Finance*, vol. 30, pp. 409–416, Jan. 2015, doi: 10.1016/S2212-5671(15)01308-8.

[4] T. Klein, H. P. Thu, and T. Walther, "Bitcoin is not the new gold—A comparison of volatility, correlation, and portfolio performance," *Int. Rev. Financial Anal.*, vol. 59, pp. 105–116, Oct. 2018, doi: 10.1016/j.irfa.2018.07.010.

[5] S. Hashim, H. Ramlan, N. Razali, and N. Nordin, "Macroeconomic variables affecting the volatility of gold price," *J. Global Bus. Soc. Entrep.*, vol. 3, no. 5, pp. 97–106, 2017.

[6] N. Jagannath, T. Barbulescu, K. M. Sallam, I. Elgendi, B. Mcgrath, A. Jamalipour, M. Abdel-Basset, and K. Munasinghe, "An on-chain analysis-based approach to predict ethereum prices," *IEEE Access*, vol. 9, pp. 167972–167989, 2021, doi: 10.1109/ACCESS.2021.3135620.

[7] S. Nanayakkara, A. Wanniarachchi, and D. Vidanagama, "Adaptive stock market portfolio management and stock prices prediction platform for Colombo stock exchange of sri Lanka," in *Proc. 5th SLAAI Int. Conf. Artif. Intell. (SLAAI-ICAI)*, Dec. 2021, pp. 1–6, doi: 10.1109/SLAAI-ICAI54477.2021.9664735.

[8] Preeti, R. Bala, and R. P. Singh, "Financial and non-stationary time series forecasting using LSTM recurrent neural network for short and long horizon," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2019, pp. 1–7, doi: 10.1109/ICCCNT45670.2019.8944624.

[9] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Online]. Available: https://www.debr.io/article/21260.pdf

[10] M. Nofer, P. Gomber, O. Hinz, and D. Schiereck, "Blockchain," *Bus. Inf. Syst. Eng.*, vol. 59, no. 3, pp. 183–187, Mar. 2017, doi: 10.1007/s12599-017-0467-3.

[11] F. Glaser, K. Zimmermann, M. Haferkorn, M. Weber, and M. Siering, ''Bitcoin-asset or currency? Revealing users' hidden intentions,'' in *Proc. 22nd Eur. Conf. Inf. Syst.*, 2014, pp. 1–14.

[12] M. Mita, K. Ito, S. Ohsawa, and H. Tanaka, ''What is stablecoin?: A survey on price stabilization mechanisms for decentralized payment systems,'' in *Proc. 8th Int. Cong. Adv. Appl. Inf.*, 2019, pp. 60–66, doi: 10.1109/IIAI-AAI.2019.00023.

[13] R. Lyons and G. Viswanath-Natraj, ''What keeps stablecoins stable?'' Nat. Bur. Econ. Res., Cambridge, MA, USA, Tech. Rep. 27136, 2020. [Online]. Available: https://www.nber.org/papers/w27136

[14] S. Tikhomirov, ''Ethereum: State of knowledge and research perspectives,'' in *Proc. Int. Symp. Found. Pract. Secur.*, 2017, pp. 206–221, doi: 10.1007/978-3-319-75650-9_14.

[15] G. Hileman and M. Rauchs, ''Global cryptocurrency benchmarking study,'' Cambridge Centre Alternative Finance, Cambridge, U.K., Tech. Rep. 113, 2017, vol. 33.

[16] Y.-C. Liu and L. Sun, ''Analysis of cointegration between macroeconomic variables and stock index,'' in *Proc. 4th Int. Conf. Natural Comput.*, vol. 5, 2008, pp. 318–322, doi: 10.1109/ICNC.2008.689.

[17] S. Antwi, M. Issah, A. Patience, and S. Antwi, ''The effect of macroeconomic variables on exchange rate: Evidence from Ghana,'' *Cogent Econ. Finance*, vol. 8, no. 1, Jan. 2020, Art. no. 1821483, doi: 10.1080/23322039.2020.1821483.

[18] Y. Udagawa, ''Predicting stock price trend using candlestick chart blending technique,'' in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 4162–4168, doi: 10.1109/BigData.2018.8622402.

[19] M. N. Bandara, R. M. Ranasinghe, R. W. M. Arachchi, C. G. Somathilaka, S. Perera, and D. C. Wimalasuriya, ''A complex event processing toolkit for detecting technical chart patterns,'' in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshop*, May 2015, pp. 547–556, doi: 10.1109/IPDPSW.2015.83.

[20] C. Christina and R. F. Umbara, ''Gold price prediction using type-2 neuro-fuzzy modeling and ARIMA,'' in *Proc. 3rd Int. Conf. Inf. Commun. Technol. (ICoICT)*, May 2015, pp. 272–277, doi: 10.1109/ICoICT.2015.7231435.

[21] A. Kiohos and N. Sariannidis, ''Determinants of the asymmetric gold market,'' *Investment Manage. Financial Innov.*, vol. 7, no. 4, pp. 26–33, 2010.

[22] G. Hu, Y. Liu, J. W. Wang, G. Zhou, and X. Zhu, ''Insider ownership and stock price crash risk around the globe,'' *Pacific-Basin Finance J.*, vol. 72, Apr. 2022, Art. no. 101714, doi: 10.1016/j.pacfin.2022.101714.

[23] S. Lim and C. S. Tucker, ''Mining Twitter data for causal links between tweets and real-world outcomes,'' *Expert Syst. Appl., X*, vol. 3, Sep. 2019, Art. no. 100007, doi: 10.1016/j.eswax.2019.100007.

[24] Z. Chen, ''The impact of trade and financial expansion on volatility of real exchange rate,'' *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0262230, doi: 10.1371/journal.pone.0262230.

[25] Y. Mezquita, A. Gil-González, J. Prieto, and J. Corchado, ''Cryptocurrencies and price prediction: A survey,'' in *Proc. Int. Congr. Blockchain Appl.*, 2021, pp. 339–346, doi: 10.1007/978-3-030-86162-9_34.

[26] A. M. Khedr, I. Arif, P. V. P. Raj, M. El-Bannany, S. M. Alhashmi, and M. Sreedharan, ''Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey,'' *Intell. Syst. Accounting, Finance Manage.*, vol. 28, no. 1, pp. 3–34, Jan. 2021, doi: 10.1002/isaf.1488.

[27] M. Poongodi, A. Sharma, V. Vijayakumar, V. Bhardwaj, A. P. Sharma, R. Iqbal, and R. Kumar, ''Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system,'' *Comput. Electr. Eng.*, vol. 81, Jan. 2020, Art. no. 106527, doi: 10.1016/j.compeleceng.2019.106527.

[28] Z. Chen, C. Li, and W. Sun, ''Bitcoin price prediction using machine learning: An approach to sample dimension engineering,'' *J. Comput. Appl. Math.*, vol. 365, Feb. 2020, Art. no. 112395, doi: 10.1016/j.cam.2019.112395.

[29] J. Rebane, I. Karlsson, P. Papapetrou, and S. Denic, ''Seq2Seq RNNs and ARIMA models for cryptocurrency prediction: A comparative study,'' in *Proc. SIGKDD Workshop Fintech*, 2018, pp. 2–6.

[30] E. Sin and L. Wang, ''Bitcoin price prediction using ensembles of neural networks,'' in *Proc. 13th Int. Conf. Natural Comput., Fuzzy Syst. Knowl. Discovery (ICNC-FSKD)*, Jul. 2017, pp. 666–671, doi: 10.1109/FSKD.2017.8393351.

[31] S. Lahmiri and S. Bekiros, ''Cryptocurrency forecasting with deep learning chaotic neural networks,'' *Chaos, Solitons Fractals*, vol. 118, pp. 35–40, Jan. 2019, doi: 10.1016/j.chaos.2018.11.014.

[32] T. Phaladisailoed and T. Numnonda, ''Machine learning models comparison for bitcoin price prediction,'' in *Proc. 10th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Jul. 2018, pp. 506–511, doi: 10.1109/ICITEED.2018.8534911.

[33] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, ''A deep learning-based cryptocurrency price prediction scheme for financial institutions,'' *J. Inf. Secur. Appl.*, vol. 55, Dec. 2020, Art. no. 102583, doi: 10.1016/j.jisa.2020.102583.

[34] T. Awoke, M. Rout, L. Mohanty, and S. Satapathy, ''Bitcoin price prediction and analysis using deep learning models,'' in *Communication Software and Networks*, 2021, pp. 631–640, doi: 10.1007/978-981-15-5397-4_63.

[35] S. Biswas, M. Pawar, S. Badole, N. Galande, and S. Rathod, ''Cryptocurrency price prediction using neural networks and deep learning,'' in *Proc. 7th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2021, pp. 408–413, doi: 10.1109/ICACCS51430.2021.9441872.

[36] C. Bai, T. White, L. Xiao, V. S. Subrahmanian, and Z. Zhou, ''C2P2: A collective cryptocurrency up/down price prediction engine,'' in *Proc. IEEE Int. Conf. Blockchain (Blockchain)*, Jul. 2019, pp. 425–430, doi: 10.1109/Blockchain.2019.00065.

[37] L. Alessandretti, A. ElBahrawy, L. Aiello, and A. Baronchelli, ''Anticipating cryptocurrency prices using machine learning,'' *Complexity*, vol. 2018, Nov. 2018, Art. no. 8983590, doi: 10.1155/2018/8983590.

[38] M. T. Kurbucz, ''Predicting the price of bitcoin by the most frequent edges of its transaction network,'' *Econ. Lett.*, vol. 184, Nov. 2019, Art. no. 108655, doi: 10.1016/j.econlet.2019.108655.

[39] S. Mohapatra, N. Ahmed, and P. Alencar, ''KryptoOracle: A real-time cryptocurrency price prediction platform using Twitter sentiments,'' in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 5544–5551, doi: 10.1109/BigData47090.2019.9006554.

[40] K. Wołk, ''Advanced social media sentiment analysis for short-term cryptocurrency price prediction,'' *Expert Syst.*, vol. 37, no. 2, p. e12493, Apr. 2020, doi: 10.1111/exsy.12493.

[41] A. Mittal, V. Dhiman, A. Singh, and C. Prakash, ''Short-term bitcoin price fluctuation prediction using social media and web search data,'' in *Proc. 12th Int. Conf. Contemp. Comput. (IC)*, Aug. 2019, pp. 1–6, doi: 10.1109/IC3.2019.8844899.

[42] A. Aggarwal, I. Gupta, N. Garg, and A. Goel, ''Deep learning approach to determine the impact of socio economic factors on bitcoin price prediction,'' in *Proc. 12th Int. Conf. Contemp. Comput. (IC)*, Aug. 2019, pp. 1–5, doi: 10.1109/IC3.2019.8844928.

[43] Z. Shahbazi and Y.-C. Byun, ''Improving the cryptocurrency price prediction performance based on reinforcement learning,'' *IEEE Access*, vol. 9, pp. 162651–162659, 2021, doi: 10.1109/ACCESS.2021.3133937.

[44] K. Lee, S. Ulkuatam, P. Beling, and W. Scherer, ''Generating synthetic bitcoin transactions and predicting market price movement via inverse reinforcement learning and agent-based modeling,'' *J. Artif. Soc. Social Simul.*, vol. 21, no. 3, p. 5, 2018, doi: 10.18564/jasss.3733.

[45] Z. Jiang and J. Liang, ''Cryptocurrency portfolio management with deep reinforcement learning,'' in *Proc. Intell. Syst. Conf. (IntelliSys)*, Sep. 2017, pp. 905–913, doi: 10.1109/IntelliSys.2017.8324237.

[46] C. Betancourt and W.-H. Chen, ''Deep reinforcement learning for portfolio management of markets with a dynamic number of assets,'' *Expert Syst. Appl.*, vol. 164, Feb. 2021, Art. no. 114002, doi: 10.1016/j.eswa.2020.114002.

[47] M. Schnaubelt, ''Deep reinforcement learning for the optimal placement of cryptocurrency limit orders,'' *Eur. J. Oper. Res.*, vol. 296, no. 3, pp. 993–1006, Feb. 2022, doi: 10.1016/j.ejor.2021.04.050.

[48] S. Aminikhanghahi and D. J. Cook, ''A survey of methods for time series change point detection,'' *Knowl. Inf. Syst.*, vol. 51, no. 2, pp. 339–367, 2017, doi: doi.org/10.1007/s10115-016-0987-z.

[49] P. Fryzlewicz, ''Wild binary segmentation for multiple change-point detection,'' *Ann. Statist.*, vol. 42, no. 6, pp. 2243–2281, 2014, doi: 10.1214/14-AOS1245.

[50] T. Fu, F. Chung, and C. Ng, ''Financial time series segmentation based on specialized binary tree representation,'' in *Proc. Conf. Data Mining*, 2006, pp. 3–9.

[51] S. Jia and L. Shi, ''Efficient change-points detection for genomic sequences via cumulative segmented regression,'' *Bioinformatics*, vol. 38, no. 2, pp. 311–317, Jan. 2022, doi: 10.1093/bioinformatics/btab685.

[52] R. Tahmasbi and S. Rezaei, ''Change point detection in GARCH models for voice activity detection,'' *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 5, pp. 1038–1046, Jul. 2008, doi: 10.1109/TASL.2008.922468.

[53] P. Yang, G. Dumont, and J. M. Ansermino, "Adaptive change detection in heart rate trend monitoring in anesthetized children," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2211–2219, Nov. 2006, doi: 10.1109/TBME.2006.877107.
[54] J. A. Scott and M. Knott, "A cluster analysis method for grouping means in the analysis of variance," *Biometrics*, vol. 30, pp. 507–512, Sep. 1974, doi: 10.2307/2529204.
[55] I. Auger and C. Lawrence, "Algorithms for the optimal identification of segment neighborhoods," *Bull. Math. Biol.*, vol. 51, no. 1, pp. 39–54, 1989, doi: 10.1016/S0092-8240(89)80047-3.
[56] B. Jackson, J. D. Scargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan, and T. T. Tsai, "An algorithm for optimal partitioning of data on an interval," *IEEE Signal Process. Lett.*, vol. 12, no. 2, pp. 105–108, Feb. 2005, doi: 10.1109/LSP.2001.838216.
[57] R. Killick, P. Fearnhead, and I. A. Eckley, "Optimal detection of changepoints with a linear computational cost," *J. Amer. Statist. Assoc.*, vol. 107, no. 500, pp. 1590–1598, 2012, doi: 10.1080/01621459.2012.737745.
[58] G. Wambui, G. Waititu, and A. Wanjoya, "The power of the pruned exact linear time (PELT) test in multiple changepoint detection," *Amer. J. Theor. Appl. Statist.*, vol. 4, no. 6, p. 581, 2015.
[59] A. Pepelyshev and A. S. Polunchenko, "Real-time financial surveillance via quickest change-point detection methods," 2015, *arXiv:1509.01570*.
[60] X. Zhu, Y. Xie, J. Li, and D. Wu, "Change point detection for subprime crisis in American banking: From the perspective of risk dependence," *Int. Rev. Econ. Finance*, vol. 38, pp. 18–28, Jul. 2015, doi: 10.1016/j.iref.2014.12.011.
[61] D. S. Matteson and N. A. James, "A nonparametric approach for multiple change point analysis of multivariate data," *J. Amer. Stat. Assoc.*, vol. 109, no. 505, pp. 334–345, 2014, doi: 10.1080/01621459.2013.849605.
[62] A. Phillip, J. S. K. Chan, and S. Peiris, "A new look at cryptocurrencies," *Econ. Lett.*, vol. 163, pp. 6–9, Feb. 2018, doi: 10.1016/j.econlet.2017.11.020.
[63] R. Shumway, D. Stoffer, and D. Stoffer, *Time Series Analysis and Its Applications*. New York, NY, USA: Springer, 2000.
[64] N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Deep adaptive input normalization for time series forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3760–3765, Sep. 2020, doi: 10.1109/TNNLS.2019.2944933.
[65] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
[66] K. Greff, R. K. Srivastava, J. Koutnìk, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.
[67] S. Lim, S. J. Kim, Y. Park, and N. Kwon, "A deep learning-based time series model with missing value handling techniques to predict various types of liquid cargo traffic," *Expert Syst. Appl.*, vol. 184, Dec. 2021, Art. no. 115532, doi: 10.1016/j.eswa.2021.115532.
[68] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
[69] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 6000–6010, doi: 10.5555/3295222.3295349.
[70] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," 2018, *arXiv:1802.05365*.
[71] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.
[72] J. Heo, H. B. Lee, S. Kim, J. Lee, K. J. Kim, E. Yang, and S. J. Hwang, "Uncertainty-aware attention for reliable interpretation and prediction," in *Proc. 32nd Conf. Neural Inf. Process. Syst.*, 2018, pp. 917–926, doi: 10.5555/3326943.3327028.
[73] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021, doi: 10.1016/j.neucom.2021.03.091.
[74] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Comput.*, vol. 3, no. 2, pp. 246–257, Mar. 1991, doi: 10.1162/neco.1991.3.2.246.
[75] J.-S. Chou and T.-K. Nguyen, "Forward forecast of stock price using sliding-window Metaheuristic-optimized machine-learning regression," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3132–3142, Jul. 2018, doi: 10.1109/TII.2018.2794389.
[76] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," 2016, *arXiv:1603.04467*.
[77] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[78] J. Lei Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," 2016, *arXiv:1607.06450*.
[79] R. Parekh, N. P. Patel, N. Thakkar, R. Gupta, S. Tanwar, G. Sharma, I. E. Davidson, and R. Sharma, "DL-GuesS: Deep learning and sentiment analysis-based cryptocurrency price prediction," *IEEE Access*, vol. 10, pp. 35398–35409, 2022, doi: 10.1109/ACCESS.2022.3163305.
[80] A. Gupta and H. Nain, "Bitcoin price prediction using time series analysis and machine learning techniques," in *Machine Learning for Predictive Analysis*, vol. 141. Singapore: Springer, 2021, pp. 551–560, doi: 10.1007/978-981-15-7106-0_54.
[81] S. Tanwar, N. P. Patel, S. N. Patel, J. R. Patel, G. Sharma, and I. E. Davidson, "Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations," *IEEE Access*, vol. 9, pp. 138633–138646, 2021, doi: 10.1109/ACCESS.2021.3117848.
[82] S. Tuarob, P. Wettayakorn, P. Phetchai, S. Traivijitkhun, S. Lim, T. Noraset, and T. Thaipisutikul, "DAViS: A unified solution for data collection, analyzation, and visualization in real-time stock market prediction," *Financial Innov.*, vol. 7, no. 1, pp. 1–32, Dec. 2021, doi: 10.1186/s40854-021-00269-7.

GYEONGHO KIM received the B.S. degree in industrial engineering from Ulsan National Institute of Science and Technology (UNIST), South Korea, in 2021, where he is currently pursuing the combined M.S./Ph.D. degree in industrial engineering. His research interests include semi-supervised learning, anomaly detection, the application of machine learning in industries (industrial AI), and knowledge distillation.

DONG-HYUN SHIN is currently pursuing the undergraduate degree in biomedical engineering with Ulsan National Institute of Science and Technology (UNIST), South Korea. His research interests include genomics, bioinformatics, and machine learning.

JAE GYEONG CHOI received the B.S. degree in design and human engineering from Ulsan National Institute of Science and Technology (UNIST), South Korea, in 2019, where she is currently pursuing the combined M.S./Ph.D degree in industrial engineering. Her research interests include machine learning/deep learning, industrial artificial intelligence, and video/audio processing.

SUNGHOON LIM received the B.S. and M.S. degrees in industrial engineering from KAIST, South Korea, in 2005 and 2009, respectively, and the Ph.D. degree in industrial engineering from the Pennsylvania State University, University Park, PA, USA, in 2018. Since 2018, he has been an Assistant Professor with the Department of Industrial Engineering, Ulsan National Institute of Science and Technology (UNIST), South Korea. Since 2021, he has also been the Head of the Institute for the 4th Industrial Revolution, UNIST. His research interests include machine learning/deep learning, industrial artificial intelligence, and smart manufacturing.