

DCRL: Approach for Pattern Recognition in Price Time Series using Directional Change and Reinforcement Learning

Nora Alkhamees¹, Monira Aloud²

Management Information Systems Department
College of Business Administration, King Saud University, Riyadh, Saudi Arabia

Abstract—Developing an intelligent pattern recognition model for electronic markets has been a vital research direction in the field. Ongoing research continues for intelligent learning algorithms capable of recognizing and classifying price patterns and hence providing investors and market analysts with better insights into price time-series. In this paper, an adaptive intelligent Directional Change (DC) pattern recognition model with Reinforcement Learning (RL) is proposed, so called DCRL model. Compared with traditional analytical approaches that uses fixed time interval and specified features of the market, the DCRL is an alternative intelligent approach that samples price time-series using an event-based time interval and RL. In this model, the environment's behavior is incorporated into the RL process to automate the identification of directional price changes. The DCRL learns the price time-series representation by adaptively selecting different price features depending on the current state. DCRL is evaluated using Saudi stock market data with different price trends. A series of analyses demonstrate the effective analytical performance in detecting price changes and the extensive applicability of the DCRL model.

Keywords—Machine learning; reinforcement learning; directional-change event; pattern recognition; stock market

I. INTRODUCTION

Pattern recognition in financial markets has been widely studied in the fields of finance, economics, computer science, engineering, modern physics, and mathematics [30,31,37,38,48,51]. Furthermore, artificial intelligence and Machine Learning (ML) have been widely used for financial market forecasting, pattern recognition, and event detection to provide decision support in various financial market segments [19,28,32,40,42].

In the financial literature, most developed ML algorithms and methods are based on physical time, for which prices are sampled at fixed time intervals (such as daily, hourly, etc.) [26,27]. To avoid the discontinuous nature of the price time-series, the Directional Change (DC) event approach provides an alternative method for sampling time-series data [6,27]. A price point is sampled when a significant price change in the price trend is observed. Therefore, the DC event approach represents a time-series as downtrend or uptrend events based on the magnitude of price changes. Several studies have been developed based on the DC event approach for pattern recognition [26], profiling price time-series [10,46], regime change detection [47], event detection [2], time-series analysis

[7,33], forecasting models [15,16], and designing trading strategies [3,4,8-14, 29,50].

Reinforcement Learning (RL) is a learning method used for sequential decision-making problems [44]. RL is one of the three basic ML methods, along with supervised and unsupervised learning. In RL, the learning agent interacts and adapts from environmental interactions by exploitation or exploration. RL achieves performance improvements through continuous evaluations of and interactions with the environment [45]. RL has the advantages of self-learning and adapting to the environment towards decision making but lacks, to a certain extent, the environment's awareness capability. Despite the effectiveness of the RL approach, event detection and pattern recognition remain challenging in real-world time-series analysis for three reasons. First, using a physical time interval makes the price time-series discontinuous, given that prices are transacted at irregular times. Second, RL can be designed with a complex structure and a large number of parameters, which can interrupt the analysis. Lastly, the learning process of the dynamic continuous market environment's state representation and the associated learning strategy affect the RL model's interruption and converge.

In this work, an intelligent intrinsic time-driven model for automatic event detection in a price time-series - the Directional Change Reinforcement Learning (DCRL) - is developed. The DCRL is presented in two sequential phases: the RL phase and the DC event analysis phase. In particular, in the RL phase, the RL algorithm learns the environmental states and features to find the most applicable dynamic threshold for the DC event analysis. The aim is to find the best dynamic threshold definition method using the RL agent, which is subsequently used for DC event detection. We used the dynamic threshold introduced in [2], which replaces the DC given fixed threshold. For the DC event analysis phase, the generated threshold from the former phase is used to detect DC events in the price time-series. The proposed model is evaluated using the Saudi stock market (Tadawul). Stocks with different price trends and series patterns are selected to evaluate the model's performance. The experimental results demonstrate that the model is adaptable to various market conditions and might be used for designing algorithmic trading.

We are interested in developing an intelligent event detection (i.e. significant price movement) algorithm from a price time-series. This algorithm will allow investors or even artificial software agents to detect price movements in the market to capture investment opportunities. Hence, our motivation is that the DCRL can provide decision support methods for analysts and investors and can facilitate the automation of event detection for sampling price time-series. Therefore, a novel method for financial event detection and time-series sampling has been proposed.

The remainder of this paper is organized as follows. Section II reviews several related works in the financial literature. Section III introduces the proposed DCRL model. Section IV describes the datasets, presents the empirical evidence of learning and identification of events, and evaluates the effectiveness and robustness of the DCRL. The last section concludes the paper and presents some future directions.

II. RELATED WORK

In the financial literature, several studies including but not limited to the following have been using the RL method for financial signal representation [32], designing trading strategies and algorithmic trading [17, 20, 32, 41], portfolio management [5], optimizing trade execution [39], trading systems [49], Foreign Exchange (FX) asset allocations [22], cryptocurrency market trading [18], and changes in market regimes [5].

Supervised learning methods have been used to forecast stock prices and the direction of price trend movements [21]. Several studies used deep ML methods to forecast a stock price using historical numerical and textual data [1,23]. The authors in [23] used deep learning for event-driven stock predictions. The events are extracted from news text and formulated as dense vectors that are trained using a neural tensor network. The deep convolutional neural network is used to model the events' impact on price time-series movements. The results showed that the proposed model could obtain an approximate 6% improvement in S&P 500 index forecasting. Nonetheless, the proposed method is challenging if attempting to achieve adaptable learning and simultaneously lacks the quick response to new dynamic market conditions given the high cost of retraining [32]. Thus, when designing event detection algorithms and algorithmic trading, the inherent characteristics and evolution of market fundamentals should be considered.

The RL method might be an alternative solution for event detection and algorithmic trading, given that it is more applicable for continuous decision making in financial market trading [32]. Bertsimas and Lo in [17] examined the application of RL for trading large blocks of equity over a specific period to minimize the expected cost of executing trades. Their results demonstrated that the RL trading strategy saved between 25% and 40% in execution costs relative to the naïve strategy. Experimental results in [20,41,43] also show that the adaptive event detection mechanism and algorithmic trading with RL methods achieve more stable returns. The experimental results by [18] confirmed the effectiveness of deep RL methods on a dataset of one of the largest

cryptocurrency markets in the world, achieving average daily returns of over 24%.

Studies on algorithmic trading using the RL method can be categorized into two main groups: policy-based methods and value-based function methods. Work in [31] has designed an on-policy (policy-based) and an off-policy (Q-learning) discrete state and action RL agents for an individual retirement portfolio. Their study found that using the trading algorithm design results in the on-policy algorithm maintaining better evaluation and adaptation to the environment than Q-learning. Their study also found that the on-policy method's drawback is that it continuously remains to explore in the environment even when the best solution is learned. The works in [22,36] has demonstrated that the benefit of the policy-based model is that it has better results than the value-based function model.

The authors in [32] studied the representation of the stock market environmental state and developed a trading strategy using historical stock price and trading volume data. They developed a time-driven, feature-aware model jointly with a deep reinforcement learning model (TFJ-DRL) that had two parts - deep learning perception and RL decision making - to improve financial signal representation learning and, hence, decision making in algorithmic trading. The results showed that the TFJ-DRL model outperformed the state-of-the-art methods in the literature. A similar study by [24] introduced a decision support algorithm to filter trading signals based on RL and neural networks. The study aims to detect seasonality events of the basic strategy to improve the reward to risk ratios.

Maringer and Ramtohl in [34,35] introduced a regime-switching to the Recurrent RL (RRL), where regime-switching captures the different price trend movements over a time series. The results highlighted that the regime-switching RRL outperforms the traditional RRL when the price time series exhibits noticeably different regime characteristics. The RRL model in [36] is a policy-based model that offers the action of the previous time's trading with the current environmental state to direct RL, hence, create a trading action. This model's main obstacle is the direct input of all of the environmental features to the RL model without awareness and representation of the current environment's status. The study by [25] combined features based on Japanese candlesticks, a technical analysis technique, with RRL to produce a high-frequency algorithmic trading system for the E-mini S&P 500 index futures market. The results demonstrated a significant increase in both return and Sharpe ratio compared to relevant benchmarks, suggesting the capability of RRL to detect events in a high-frequency equity index futures trading environment.

Overall, the RL method has been recognized as being effective and efficient in forecasting asset prices in financial markets and, hence, make trading decisions. Previous studies used RL based on physical time, which is characterized by a fixed time interval, whereas the price time series is irregularly spaced in time. Therefore, to develop an adaptive RL algorithm for event detection, the DC event approach is used to represent and study the price time series. In this work, we use the RL to enhance the dynamic threshold definition

method presented in [2]. We want to improve the dynamic threshold definition method so that we can set the dynamic threshold without the need for an additional source of data (such as news).

III. METHODOLOGY

In this section, we introduce the DCRL model which aims to identify financial events from stock market price time series and, hence, represent periodic patterns of the price time series. First, the DC event approach which constructs a price time series of continuous DC events is described. Then, the process of defining the DC dynamic threshold is explained. Finally, the DCRL model is introduced as a dynamic adaptive process to select the optimal equation for the DC dynamic threshold. In other words, a DCRL model is developed to identify DC events in a price time series using the different dynamic threshold equations (actions). Therefore, the goal is to improve the reward function under different states.

A. DC Event Approach

Using the DC event approach, price time series data are sampled at irregular time intervals using a given size threshold (λ), which is defined by the observer (fixed value) and is typically expressed as a percentage [6]. Thus, the DC event approach transforms the discrete nature of the price time series into continuous DC events independent of the notion of fixed physical timescales. Under the DC event approach, the price time series is summarized into alternating uptrend and downtrend DC events.

A DC event is identified as a confirmed price change that is larger than, or equal to, a predefined threshold (λ) [6]. A DC event can be either a downturn or an upturn DC event. The time interval between an upturn DC event and the next downturn DC event is called an upward run, whereas a downward run is the time interval between a downturn DC event and the next upturn DC event. During an upward run, the last high price (p_h) is continuously updated to the maximum value between the current asset price $p(t)$ and the last high price (p_h). In a downward run, the last low price (p_l) is continuously updated to the minimum value between the current market price $p(t)$ and the last low price (p_l). At the beginning of a data sequence, the last low price (p_l) and last high price (p_h) are set to the initial asset price $p(t_0)$ at time t_0 . An upturn DC event is detected during a downward run and, in particular, when the current asset price $p(t)$ exceeds the last low price (p_l) by a given threshold (λ); refer to Formula (1). In contrast, a downturn DC event is detected during an upward run when the current asset price $p(t)$ is lower than the last high price (p_h) by a given threshold (λ); refer to Formula (2).

$$p_t \geq p_l \times (1 - \lambda) \quad (1)$$

$$p_t \leq p_h \times (1 - \lambda) \quad (2)$$

The DC event approach captures the short-term dynamics of the price time series by detecting significant events and a clear picture of the time series behavior on the basis of the observer's needs. Most importantly, this approach reduces the complexity of the financial market price time series, given the defined dataset of periodic price points to study and evaluate. The selected threshold value controls the magnitude of the DC

price events in a time series. Therefore, choosing a substantial threshold results in fewer detected DC price events, whereas a small threshold maps a series of insignificant patterns. The authors in [6] described the core mechanism of the DC event approach to study the financial price time series. In this work, a price time series is formulated using the DC event approach. Given a size threshold (λ), the mission is to detect events at the DC confirmation point regardless of whether or not the direction of the price trend changes at a certain point.

B. DC Dynamic Threshold

In this section, we describe the dynamic threshold definition method which replaces the DC fixed given threshold value [2]. The dynamic threshold definition method is suitable for markets that operate during specific opening and closing times (such as stock markets). The dynamic threshold is a flexible value and brings with it the advantage of allowing the identification of price changes (i.e., DC events) of different magnitudes in continuously changing environments.

In [2], significant price fluctuations were considered as an event occurrence indicator. Thus, the dynamic threshold definition method depends on the previous day's price behavior (short-term price history). The daily dynamic threshold value can be set in three possible ways, choosing the most appropriate one was not straightforward. They depend on an alternative source of data (news outlets) to facilitate the definition of the dynamic threshold. A suitable dynamic threshold definition method can be selected depending on the investigated asset news and market conditions. In this work, the best method for defining the dynamic threshold value without an alternative source of data is determined using RL. Hence, an agent is developed to select the most effective dynamic threshold definition method (i.e., the one that detects DC events at the right time).

Basically, the dynamic threshold can be set using one of the three equations (Eq. (3), Eq. (4), and Eq. (5)) as follow:

$$DT_{\text{Overnight}} = \text{Up/Downward_ROC} + \text{Overnight_ROC} \quad (3)$$

$$DT_{\text{PreviousDay}} = \text{Up/Downward_ROC} + \text{PreviousDay_ROC} \quad (4)$$

$$DT = \text{Up/Downward_ROC} + \text{PreviousDay_ROC} + \text{Overnight_ROC} \quad (5)$$

DC dynamic threshold depends on the price Rate Of Change (ROC) between the DC p_h/p_l (refer to Section III.A) and the high/low prices (depending on the examined trend) for the current day. In addition, it finds the price ROC for the previous day (between the previous day's opening and closing prices), and the price ROC that occurred overnight (between the previous day's closing price and the current day's opening price). The dynamic threshold is defined by the sum of the aforementioned metrics, as shown in Eq. (5). However, in some circumstances in which something has happened the previous day or overnight, the shortened version of the dynamic threshold definition method (Eq. 3 or Eq. 4) is used to ensure a reduced threshold value that certainly increases the chance of identifying a DC event (either an upturn or downturn event). Also, to be mentioned is that if a defined

threshold value by Eq. (5) was found to be less than 0.01, then we use the previous day's defined threshold instead. This approach is taken because we are almost certain that nothing has happened (experiencing a stable situation as no significant price changes occurred on the previous day or overnight), and an exceptionally low threshold value may detect a spurious or insignificant event.

C. DCRL

In this section, we introduce the DCRL model, which can identify financial events from a time series. The DCRL is based on the RL approach, which directs the dynamic thresholds definition method, and the DC approach, which is responsible for detecting the occurring DC events on the basis of the given threshold value from the RL phase.

RL is a learning approach through which an intelligence algorithm represented by an agent is designed to learn from interactions with the environment. Therefore, RL mimics human learning and, hence, appears well suited to processing the price time series. The goal is to train the RL agent based on a sequence of interactions to learn an optimal policy from the interaction to maximize the total cumulative reward obtained. In this section, the RL approach's key elements are introduced, and the approach is tailored to the goals of this study.

RL can be generally categorized into two types: the policy-based and value-based function methods [32]. Policy-based RL explicitly and directly builds a representation of a policy from the environment and, hence, creates continuous decisions from the policy. The established policy is stored in memory during the learning phase. The DCRL policy-based method is as follows: if the price ROC from the previous day or overnight is greater than a five-day price change moving average, then the first two equations (Eq. 3 or Eq. 4) from section III.B are used; otherwise, Eq. (5) is used.

The RL approach consists of the environment, agent, state, action, and reward. Considering discrete times $t = 0, 1, 2, 3, \dots$, at each time t , a RL agent receives some representations of a state in the environment, denoted by $s_t \in S$, where S is the set of all possible states. Based on the current state s_t and the previously obtained information, the agent takes action $a_t \in A(s_t)$, where $A(s_t)$ is the set of actions available in state s_t . The space of actions in DCRL consists of the three equations for defining the DC dynamic threshold, as described in Section 3.2. The RL agent chooses an action on the basis of its policy π_t , which is a mapping from each state to the probabilities of deciding on each possible action. Therefore, $\pi_t(s)$ denotes the chosen action when $S_t = s$ based on π_t . At the next time point $t+1$, the agent receives a numerical reward from the environment, denoted by $r_t \in R$, because of its action a_t , and moves to a new state s_{t+1} . Based on the earned reward, the RL agent learns to adapt its actions on the basis of the market condition to maximize its future rewards.

The DCRL agent interaction with its environment is depicted in Fig. 1. As an input, we have the price time series, and as an output, we get the optimal chosen action a_t (the best dynamic threshold definition method) and the assigned

reward r_t . The agent interaction with the environment is shown in the stage between input and output. Table I provides the set of all possible states in the environment and the set of actions presented for each state s_t , along with the associated rewards for each pair of state and action. In Table I, the DCRL approach takes the appropriate state-action policy $\pi_t(s_t, a_t)$, which indicates the expected reward r_t for each possible action a_t . For this purpose, the DCRL agent starts with random initial values of $\pi_t(s_t, a_t)$ for $s_t \in S$ and $a_t \in A(s_t)$. The DCRL agent then proceeds with the aforementioned interaction learning steps: (1) observes the current state s_t of the price time series, (2) executes action a_t , and (3) receives reward r_t and observes the next state s_{t+1} .

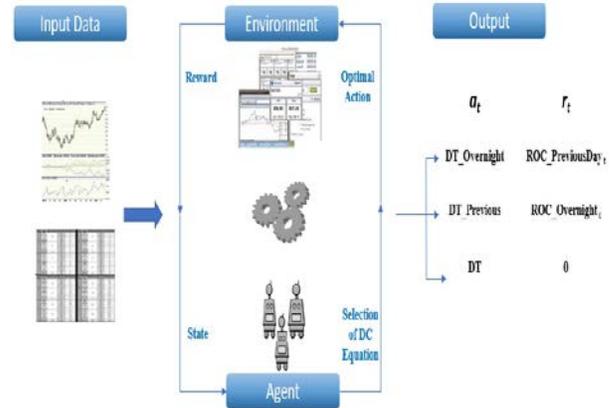


Fig. 1. DCRL Model Interaction with its Environment.

In each iteration, the DCRL agent observes the current state of the environment using the following state variables: a five-day price change moving average, the previous day opening and closing prices, and the previous day closing price and current opening price. This specification has established a learning architecture whereby the previous action at time $t - 1$ is considered. In this study, we choose the previous day ROC (Ext_Previous_t), overnight ROC (Ext_Overnight_t), and Neutral (Neutral_t) state to represent the set of possible states S . Following the observation of the current state s_t , the RL agent chooses action a_t from three possibilities. (1) Equation 3 (DT_Overnight) is used to define the DC dynamic threshold considering that an overnight event has occurred. (2) Equation 4 (DT_PreviousDay) is used to define the DC dynamic threshold considering that an event has occurred during the previous day. These two possible actions are associated with the two states Ext_Previous_t and Ext_Overnight_t. Note that an action that offers a lower threshold value is selected because it will increase the chance of detecting an event. For the Neutral_t state, only one possible action exists, which is (3) using Equation 5 assuming that no extreme price changes have occurred. Hence, the following set of possible actions is obtained: $a_t = \{DT_Overnight, DT_PreviousDay, DT\}$.

The agent receives a reward on the basis of the selected action. The reward is the maximum of either ROC_PreviousDay or ROC_Overnight when actions DT_Overnight or DT_PreviousDay are chosen. Alternatively, no reward is assigned (reward = 0) when action DT is taken because it is always taken whenever action DT_Overnight or

DT_PreviousDay cannot be taken. More specifically, if action DT_Overnight was executed, then the assigned reward is ROC_PreviousDay; because the action was based on the lowest threshold value and was DT_Overnight (it led to the lowest threshold value), then the reward is ROC_PreviousDay because it is of greater value than ROC_Overnight. The same applies for action DT_PreviousDay: if action DT_PreviousDay was executed, then the reward assigned is ROC_Overnight.

TABLE I. SET OF ALL POSSIBLE STATES IN THE ENVIRONMENT AND SET OF ACTIONS PRESENTED FOR EACH STATE S_t , ALONG WITH ASSOCIATED REWARDS FOR EACH PAIR OF STATE AND ACTION

S_t	S_{t+1}	a_t	P	R
Ext_Overnight _t	Ext_Overnight _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
	Ext_Previous _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
	N _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
Ext_Overnight _t	Ext_Previous _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
	Ext_Overnight _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
	N _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
Ext_Previous _t	Ext_Previous _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
	Ext_Overnight _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
	N _{t+1}	DT_PreviousDay	0.3 3	ROC_Overnight _t
Ext_Previous _t	Ext_Overnight _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
	Ext_Previous _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
	N _{t+1}	DT_Overnight	0.3 3	ROC_PreviousDay _t
Neutral _t	N _{t+1}	DT	0.3 3	0
	Ext_Overnight _{t+1}	DT	0.3 3	0
	Ext_Previous _{t+1}	DT	0.3 3	0

IV. DATA AND EMPIRICAL RESULTS

To verify the effectiveness and robustness of the proposed DCRL model, a series of experiments were conducted using four price datasets for stock exchange indices. The DCRL utilizes a policy-based model that learns the policy from historical prices and defines a variety of continuous actions according to the learned policy. A descriptive analysis of the identified events is presented in Section IVB, it shows a discussion of the identified events along with a statistical description of the associated DC dynamic threshold values. The last section presents the evaluation results of the effectiveness and accuracy of the proposed DCRL model.

A. Data

Our empirical study relies on data from the Saudi stock market (Tadawul¹) for the period from March 2015 to March 2020, the total number of investigated days is approximately 1285 days. We used four stock indices for the following two financial sectors: Al Rajhi, Alinma, and SABB banks (Sector: Financials Industry, Group: Banks), and STC (Sector: Telecommunication & Information Technology). These selected stock indices are well known in the Saudi financial market. The price time series for these four stock exchange indices are sourced from Yahoo finance². Each row data includes the date along with opening, low, high, and closing prices. The choice of these four stock indices is based on the strength of their economic and financial factors. The distribution of each dataset composes of a variety of price trends and a series of patterns, which will contribute to the effectiveness of evaluating the DCRL model under different situations. Fig. 2 shows the price time series for the four stock indices during the five investigated years (2015- 2020).

Table II presents the basic annualized average summary descriptive statistical analysis of the datasets for the four stock exchange indexes during the period from March 17, 2015, to March 13, 2020. The mean (μ), standard deviations (σ), Skewness, Kurtosis, minimum, and maximum price values are reported for each stock index.

B. Results

In Table III, we report the statistical analysis results of the identified DC events using the DCRL model. Table III provides the average annualized of the following quantities: number of identified DC events, number of times (days) the previous day's defined threshold was also used for the current day, number of times an event was identified as the ROC taking place overnight was significant, and the number of times an event was detected as the ROC taking place the previous day was significant. When *DT_Overnight* and *DT_PreviousDay* are used more often to define the DC dynamic threshold, this use could mean that the price changes occurring overnight (between the previous day's and the current day's opening prices) or occurring on the previous day (between previous day opening and closing prices) are considerably high. Therefore, the identified DC events using the dynamic threshold definition method can capture the sensitivity of the market changes and, hence, the identification of potential events.

TABLE II. DESCRIPTIVE STATISTICS OF FOUR STOCK INDICES: ALRAJHI, ALINMA, SABB, AND STC. THE PERIOD SPANS FROM MARCH 17, 2015, TO MARCH 13, 2020, AND THE TOTAL NUMBER OF INVESTIGATED DAYS IS APPROXIMATELY 1285 DAYS

Index	μ	σ	Skewness	Kurtosis	Min.	Max.
Alrajhi	48.25	3.64	0.38	-0.37	40.20	56.39
Alinma	18.99	2.12	-0.17	-0.22	13.92	22.64
SABB	28.87	3.34	-0.059	-0.53	21.06	34.78
STC	77.95	5.38	0.16	-0.09	65.53	90.07

¹ <https://www.tadawul.com.sa>

² <https://finance.yahoo.com/>

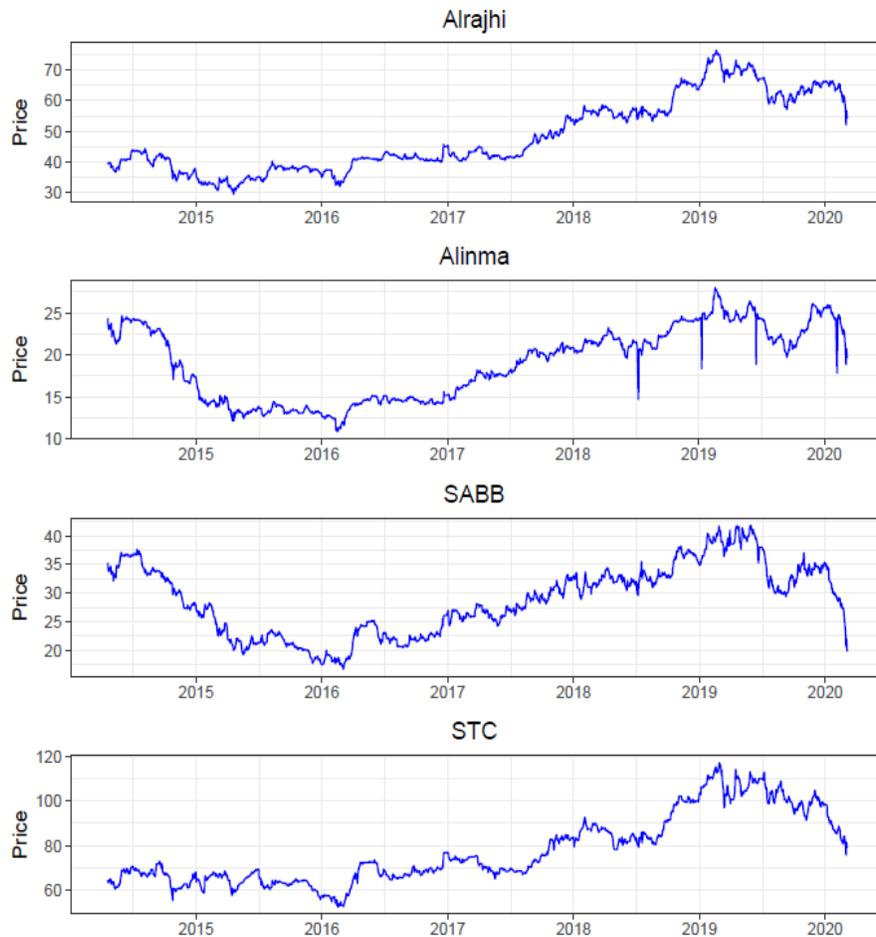


Fig. 2. Price Time Series for the Four Stock Indices: Alrajhi, Alinma, SABB, and STC over the Period from 17/3/ 2015-13/7/202 (x-axis Represents Price and y-axis Represents Year).

For all of the investigated stocks during the 5 examined years, more than half of the detected DC events from the dynamic threshold values were found using both equations DT_Overnight and DT_PreviousDay. Specifically, these equations have detected more than 60% of the DC events in Alinma and STC, and 70% of the DC events in Alrajhi and SABB, refer to Table III for a summary of the annualized average statistical analysis results. In addition, SABB had the fewest number of days on which the previous day defined threshold was also used for the current day to identify DC events (if any), on average, only 14% of the investigated days the previous day threshold was also used for the current day. In other words, on average each year in 221 days out of the 259 days, a new dynamic threshold was set each day to detect DC events, if any. Therefore, SABB may have been exhibiting a number of price variations; refer to SABB price time series in Fig. 2.

Also, the high number of identified DC events (an average of 54 DC event each year) confirms this phenomenon. In contrast, Alrajhi, Alinma and STC had a higher number of days for which the previous day threshold was again used for the current day (at least 25% of days); refer to Fig. 2 for

Alrajhi, STC and Alinma price time series, which also maintains a number of price stability trends.

In order to have a deeper and closer look, in Fig. 3, we illustrate in more details the identified DC events using DCRL over the period from March 2019 to March 2020 for Alrajhi, Alinma, SABB, and STC price time series. The X-axis represents the date, and the Y-axis represents the daily closing price. In the chart, the square-shape event represents a downturn DC event, and the x-shape event represents an upturn DC event. Forty-nine DC events were identified from Alrajhi and STC. In addition, Alinma had 45 identified DC events, and 57 DC events were detected in SABB (refer to Fig. 3).

TABLE III. STATISTICAL ANALYSIS RESULTS OF AVERAGE ANNUALIZED IDENTIFIED DC EVENTS USING DCRL MODEL

Index	N_{DC}	$N_{\lambda(-)}$	$N_{\lambda(0)}$	$N(DT_Overnight)$	$N(DT_Previous Day)$
Alrajhi	51.4	66.4	192.2	17.4	18
Alinma	49.6	77.8	180.4	16.2	16
SABB	53.8	36	220.6	20.4	16.4
STC	49.8	63.8	192.6	19.8	14.4

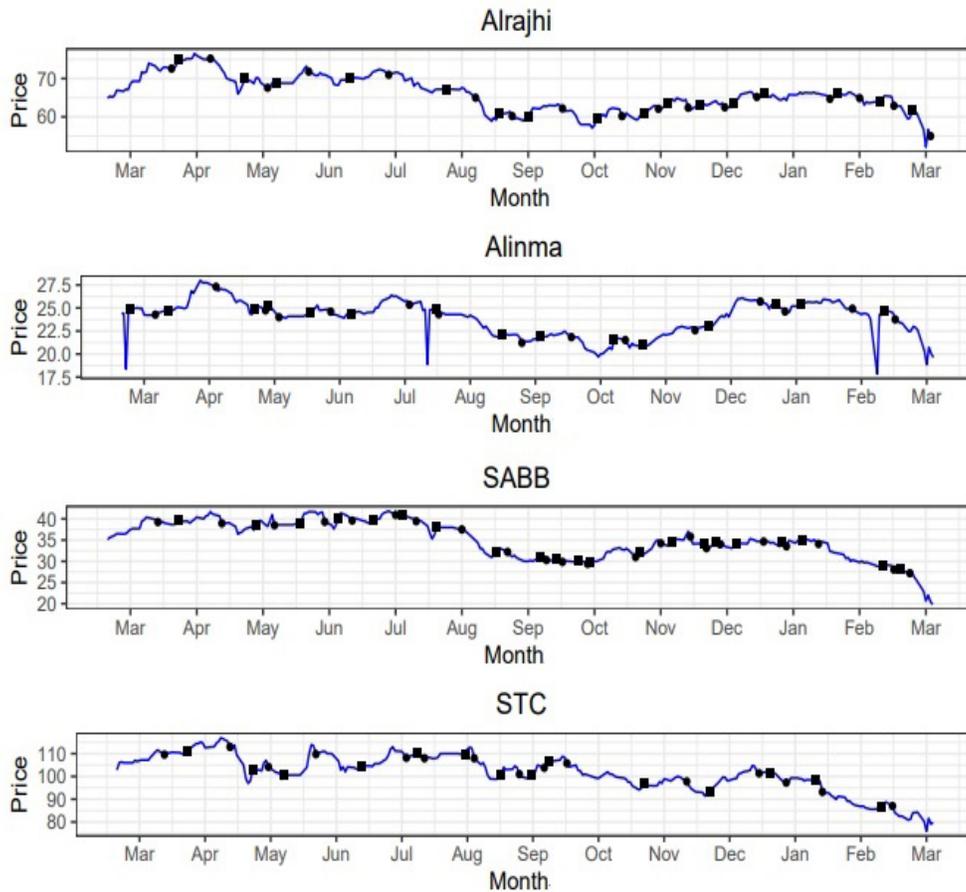


Fig. 3. Price Time Series for Four Stock Indices: Alrajhi, Alinma, SABB, and STC over the Period from March 17, 2019, to March 13, 2020. Identified DC Events (upward and downward) using the DCRL Model are shown for the Final Year in the Sampled Period.

Physical time (e.g., daily prices) fails to recognize the pattern flow of price movement, giving that the variety of price changes depends only on that considered time. Moreover, using daily or intraday prices to detect price patterns maps a range of patterns with different sizes, resulting in discontinuous pattern flow of price movements. On the other hand, the DC events reduces the complexity of a price time series giving that it detects periodic patterns in contrast to those detected by physical time.

Table IV reports an analysis of the defined DC dynamic threshold values and presents the mean value of the dynamic DC threshold values during the investigated period, the minimum and maximum DC dynamic threshold values, and, finally, the standard deviation values. Table IV clearly demonstrates that Alinma had a high standard deviation ($\sigma = 0.032$) relative to other stocks. This finding indicates that the defined threshold values are spread out with relatively high variations and are far from the mean. Additionally, Alinma has the highest maximum threshold (0.208), whereas all other stocks' maximum values were between 0.08 and 0.1 (Alinma's maximum value is at least two times higher than that of all of the other stocks); refer to Fig. 3 for the Alinma price time series to observe the significant price jumps encountered. For illustration, on March 18, 2019, the closing price was 24.4, which slipped to 18.3 on the next day. Subsequently, on March 20, 2019, the price bounced back to

24.4. During the investigated period, this significant jump occurred at least three times (in March and July 2019, and in February 2020). For the rest of the stocks, the standard deviation was between 0.012 and 0.016, indicating a smaller value. Thus, in these cases, the data points are close to the mean.

TABLE IV. DESCRIPTIVE STATISTICS OF THE DC DYNAMIC THRESHOLD VALUES DEFINED BY DCRL MODEL. FOUR DESCRIPTIVE VALUES ARE PRESENTED FOR DC DYNAMIC THRESHOLD: MEAN (μ), MINIMUM, MAXIMUM, AND STANDARD DEVIATION (σ)

Index	μ	Min	Max	σ
Alrajhi	0.022	0.046	0.0893	0.012
Alinma	0.027	0.029	0.208	0.032
SABB	0.027	0.0102	0.109	0.017
STC	0.023	0.0109	0.085	0.013

C. Evaluation

To verify the effectiveness and robustness of the proposed DCRL model, we evaluate the results using the (i) length of the price-curve coastline, and (ii) accumulated reward value from the DCRL model. The length of the price-curve coastline offers an indicator of the usefulness of sampling the price time series, whereas the accumulated reward value evaluates the efficiency of the learning process in the DCRL model.

1) *Price-curve coastline*: The authors in [26] uncovered the scaling laws used to estimate the length of the price-curve coastline on the basis of the intrinsic time, which turns out to be long. A price-curve coastline can capture the price variations and, hence, the potential profit [26]. In this section, we measure the length of the price-curve coastline using two different models: DCRL (intrinsic time) and physical time (fixed time intervals). The goal is to evaluate their performance by summarizing the price movements and, thus, improves the understanding of the dynamic behavior of the price time series in a simplified manner.

The length of a price-curve coastline is defined by the sum of all price changes during a defined period T . Under intrinsic time, the length of the price-curve coastline during period T is the average of the price changes between the identified DC events [6]. The length of the price-curve coastline under the DCRL model $c(\lambda)$ is defined by:

$$c(\lambda) = \frac{1}{N_{DC}} \sum_{i=1}^{N_{DC}} |p_i - p_{i+1}| \quad (6)$$

where N_{DC} is the number of identified events determined by the DC dynamic threshold (λ), p_i is the price of the i -th DC turning point, and p_{i+1} is the consequential DC turning point.

Under fixed physical time intervals, the length of the price-curve coastline during period T is the average of the price changes between the fixed points at which the time distance

between all fixed points are equivalents [6]. The length of the price-curve coastline under physical time $c(t)$ is defined by:

$$c(t) = \frac{1}{n} \sum_{i=1}^n |p_i - p_{i+1}| \quad (7)$$

where p_i is the price at point i (refer to the table to observe the length of PTI for all investigates stocks), and n refers to the total number of fixed points, which equals the number of identified DC events (to ensure fairness in comparison).

It is essential to being aware with how well the established DCRL and the physical time price-curves fit the real price time series to evaluate their performance and effectiveness of sampling price changes in a time series. For instance, Fig. 4 shows the price time series for the STC index over the March 17, 2019, to March 13, 2020 (the blue solid line) and a zoom-in in the lower chart from November 24, 2019, to January 5th, 2020. The chart compares the price changes defined under the DCRL model (the red dashed line) and the physical fixed time interval model (the green dashed line) using the same number of points (49 DC events in DCRL and 49 points in the physical time scale). The DCRL model can identify price movements (DC events) as they occurred, while the physical time scale model fails to do so. Hence, it is clear that the DCRL can better sample the price changes than the physical time scale.

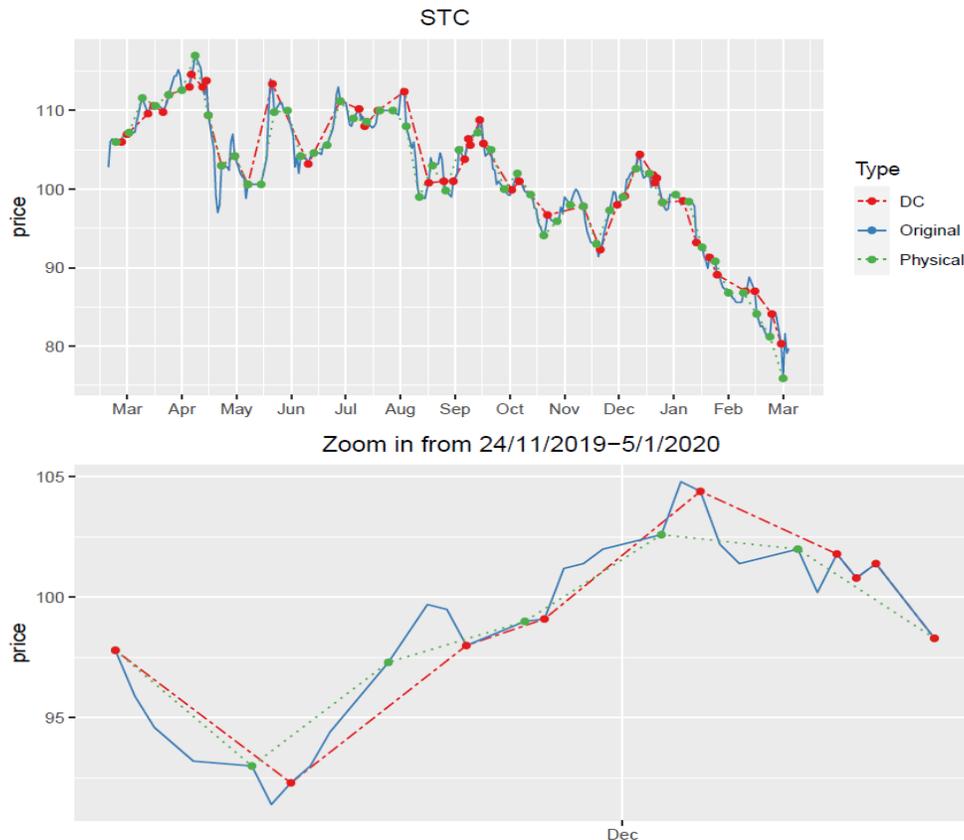


Fig. 4. The Price Curve Lines for STC Index using DC and Physical Time Compared to the Original Price Curve.

Table V compares the price-curve coastline of the events identified by the DCRL model using dynamic thresholds (intrinsic time intervals) against physical time changes (fixed time intervals) over the period March 17, 2019, to March 13, 2020. Table V clearly shows that the DCRL price-curve coastline $c(\lambda)$ is longer for all investigated stock indices relative to the physical time price-curve coastline $c(t)$. The coastline of Alinma under the DCRL model (i.e., DC intrinsic time) is more than three times longer than the physical time coastline (0.70 is the length of the DCRL coastline, and 0.22 is the length of the physical coastline). This difference can be the result of the time series evolution being unstable with significant price transitions occurring more frequently (refer to Fig. 2 for Alinma price time series). In contrast, SABB price-curve coastline using the DCRL model is slightly longer than its physical coastline but was closest to the physical coastline when compared with other investigated stock indices (0.89 was the length of the DCRL coastline, and 0.49 was for the physical time coastline). This finding can be the result of the often-recurring price transition but with insignificant price transitions (refer to Fig. 2 for SABB price time series). The DCRL price-curve coastline for the other stock indices (Alrajhi and STC) is at least two times longer than the physical coastline.

To summarize, the DCRL model-identified events using the intrinsic time outperforms the identified price transitions using the physical time for all investigated stock indices.

The natural fluctuation in the price time series suggests the need for diversification of the analytical scope of identifying financial events in the price time series. The DCRL mitigates the discontinuous price flow of prices in a time series and captures the periodic price changes.

2) *Random-Based DC model*: In this section, we further investigate the role and accuracy of the developed DCRL model in improving the decision-making process for the most appropriate dynamic threshold definition method. Therefore, we developed a random-based DC model that randomly selects a dynamic threshold definition method (randomly decide on one of the three dynamic DC equations: DT_Overnight, DT_Previous, or DT). The developed random-based DC model replaces the role of RL in selecting the DC dynamic threshold definition method. In contrast, DCRL finds the most appropriate dynamic threshold definition method using the DCRL policy (π).

Table VI provides a comparison for the accumulated reward value gained by the DCRL model and the random-based DC model over the period of March 17, 2019, to March 13, 2020. Evidently, the DCRL model outperformed the random-based DC model for all investigated stock indices, leading to the conclusion that the learning process of price movements (that is, upward and downward DC events) matters in the estimation of financial events in stock markets. The DCRL had proven to be effective in maximizing the accumulated value of the reward and, hence, the profitability during a sequence of learning steps for identifying events in different stock indices.

TABLE V. PRICE-CURVE COASTLINE LENGTH DEFINED BY: (I) THE DCRL MODEL C(λ) AGAINST (II) PHYSICAL TIME CHANGES C(T) FROM 17/3/2019-13/3/2020. THE NUMBER OF DC EVENTS IS DENOTED BY NDC, AND PTI REPRESENTS THE PHYSICAL TIME FIXED INTERVAL

Index	N _{DC}	c(λ)	PTI	c(t)
Alrajhi	49	1.47	5.3 days	0.69
Alinma	45	0.70	5.8 days	0.22
SABB	57	0.89	4.5 days	0.49
STC	49	3.06	5.3 days	1.35

TABLE VI. ACCUMULATED REWARD VALUE GAINED BY THE DCRL AND RANDOM-BASED DC MODELS

Index	DCLR	Random
Alrajhi	2.21	1.82
Alinma	3.90	3.25
SABB	3.44	2.78
STC	2.64	2.09

V. CONCLUSION

In this paper, the DC event and RL approaches were used for automated pattern recognition from price time series. We proposed an intelligent intrinsic time-driven DCRL joint model, which can (1) adaptively set the DC dynamic threshold and conduct an event-based time series analysis using the RL approach, hence, improving the effectiveness, adaptability, and interpretability of the identified financial events; (2) jointly construct a price time series using the DC event approach, thus acquiring periodic continuous price events and improving the accuracy of the price time series representation. The DCRL is suitable for markets that operate during specific opening and closing times and can identify financial events without the need for an additional data source.

The effectiveness of the DCRL model is validated on the Saudi stock market with different price trends and patterns. The experimental results demonstrate that the DCRL model outperforms other physical time-based analyses and the random-based DC model with higher rewards and a more reliable representation of the price curves.

This work can be further extended and improved in future research directions. One direction can be conducting experiments on large-scale data, such as high frequency time series data, to confirm the effectiveness of and further enhance the DCRL model. Another promising research direction is to further apply the DCRL model to emerging markets, such as the cryptocurrency market. In addition, algorithmic trading can be developed using the DCRL model to trade one asset at a time and then can be improved and expanded to manage the portfolios of several assets. Finally, some financial features could be introduced to enhance the DCRL model; for example, trade volume could provide significant information for selecting the dynamic threshold.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their useful comments and suggestions. The authors thank the

Deanship of Scientific Research and RSSU at King Saud University for their technical support.

REFERENCES

- [1] Akita, R., Yoshihara, A., Matsubara, T., and Uehara, K. (2016). Deep learning for stock prediction using numerical and textual information. Proceedings of the 15th International Conference on Computer and Information Science (ICIS) (pp. 1–6).
- [2] Alkhamees, N., and Fasli, M. (2017). Event detection from time-series streams using directional change and dynamic thresholds. IEEE International Conference on Big Data (Big Data), Boston, MA (pp. 1882–1891).
- [3] Alkhamees, N., and Fasli, M. (2017). An exploration of the Directional Change Based Trading strategy with dynamic thresholds on variable frequency data streams. 2017 International Conference on the Frontiers and Advances in Data Science (FADS), 108–113 p. Xi'an.
- [4] Alkhamees, N., and Fasli, M. (2017). A Directional Change Based Trading strategy with dynamic thresholds IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, 2017 (pp. 283–292).
- [5] Almahdi, S., & Yang, S. Y. (2017). An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown. Expert Systems with Applications, 87, 267-279.
- [6] Aloud, M., Tsang, E., Olsen, R., and Dupuis, A. (2012). A directional-change events approach for studying financial time series. Economics: The open-access, open-assessment e-journal 2012-36, 1–17.
- [7] Aloud, M., Fasli, M., Tsang, E., Dupuis, A., and Olsen, R. (2013). Stylized facts of trading activity in the high frequency FX market: An empirical study. Journal of Finance and Investment Analysis 2(4), 145–183.
- [8] Aloud, M. (2015). Directional-Change Event Trading strategy: Profit-maximizing learning strategy, in the Seventh International Conference on Advanced Cognitive Technologies and Applications, Nice, France (pp. 123–129), ISBN: 978-1-61208-390-2.
- [9] Aloud, M. (2016). Profitability of Directional Change Based Trading strategies: The case of Saudi stock market. International Journal of Economics and Financial Issues 6(1), 87–95.
- [10] Aloud, M. (2016). Time series analysis indicators under directional changes: The case of Saudi stock market. International Journal of Economics and Financial Issues 6(1), 55–64.
- [11] Aloud, M., and Fasli, M. (2017). Exploring trading strategies and their effects in the FX market. Computational Intelligence 33(2), 280–307.
- [12] Aloud, M. (2017). Investment opportunities forecasting: GP-Based Dynamic Portfolio Trading system under directional-change framework. Journal of Computational Finance 22(1), 1–35.
- [13] Ao, H., and Tsang, E. (2019). Trading algorithms built with direction changes. IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFER) Conference, Shenzhen, China.
- [14] Bakhach, A., Chinthapati, V., Tsang, E., and El Sayed, A. (2018). Intelligent dynamic backlash agent: A trading strategy based on the directional change framework. Algorithms 11(11), 171.
- [15] Bakhach, A., Tsang, E., and Jalalian, H. (2016). Forecasting directional changes in FX markets. IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (IEEE CIFER'16), Athens, Greece.
- [16] Bakhach, A., Tsang, E., and Raju Chinthapati, V. (2018). TSFDC: A trading strategy based on forecasting directional change. Intelligent Systems in Accounting, Finance and Management 25(3), 105–123.
- [17] Bertsimas, D., and Lo, A. W. (1998). Optimal control of execution costs. Journal of Financial Markets 1(1), 1–50.
- [18] Betancourt, A., Chen, W. (2021). Deep reinforcement learning for portfolio management of markets with a dynamic number of assets, Expert Systems with Applications, Volume 164, 114002.
- [19] Bustos, O., and Pomares-Quimbaya, A. (2020). Stock market movement forecast: A Systematic review. Expert Systems with Applications 156, 113464.
- [20] Carapuço, J., Neves, R., Horta, N. (2018). Reinforcement learning applied to Forex trading, Applied Soft Computing, 73, 783-794.
- [21] Chong, E., Han, C., and Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications 83, 187–205.
- [22] Dempster, M. A. H., and Leemans, V. (2006). An automated FX trading system using adaptive reinforcement learning. Expert Systems with Applications 30(3), 543–552.
- [23] Ding, X., Zhang, Y., Liu, T., and Duan, J. (2015). Deep learning for event-driven stock prediction. Proceedings of the IJCAI (pp. 2327–2333).
- [24] Eilers, D., Dunis, C., Mettenheim, H., Breitner, M. (2014). Intelligent trading of seasonal effects: A decision support algorithm based on reinforcement learning, Decision Support Systems, 64, 100-108, 0167-9236.
- [25] Gabriellsson, P., and Johansson, U. (2015). High-Frequency Equity Index Futures Trading Using Recurrent Reinforcement Learning with Candlesticks, 2015 IEEE Symposium Series on Computational Intelligence, Cape Town, pp. 734-741.
- [26] Glattfelder, J. B., Dupuis, A., and Olsen, R. B. (2011). Patterns in high-frequency FX data: Discovery of 12 empirical scaling laws. Quantitative Finance 11(4), 599–614.
- [27] Guillaume, D. M., Dacorogna, M. M., Davé, R. R., Müller, U. A., Olsen, R. B., and Pictet, O. V. (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. Finance and Stochastics 1(2), 95–129.
- [28] Henrique, B., Sobreiro, V., and Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications 124, 226-251.
- [29] Kampouridis, M., and Otero, F. E. B. (2017). Evolving trading strategies using directional changes. Expert Systems with Applications 73, 145–160.
- [30] Ko, B., Song, J., Chang, W. (2016). Simulation of financial market via nonlinear ising model, International Journal of Modern Physics C, 27(4), 1650038.
- [31] Ko, B., Song, J., Chang, W. (2018). Crash forecasting in the Korean stock market based on the log-periodic structure and pattern recognition, Physica A: Statistical Mechanics and its Applications, 492, 308-323.
- [32] Lei, K., Zhang, B., Li, Y., Yang, M., and Shen, Y. (2020). Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. Expert Systems with Applications 140.
- [33] Ma, J., Xiong, X., He, F., and Zhang, W. (2017). Volatility measurement with directional change in Chinese stock market: Statistical property and investment strategy. Physica A 471, 169–180.
- [34] Maringer, D., and Ramtohul, T. (2010). Threshold recurrent reinforcement learning model for automated trading. lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). LNCS (Part 2) 6025, 212–222.
- [35] Maringer, D., and Ramtohul, T. (2012). Regime-switching recurrent reinforcement learning for investment decision making. Computational Management Science 9(1), 89–107.
- [36] Moody, J., and Saffell, M. (2001). Learning to trade via direct reinforcement. IEEE Transactions on Neural Networks 12(4), 875–889.
- [37] Martins, T., Neves, R. (2020). Applying genetic algorithms with speciation for optimization of grid template pattern detection in financial markets, Expert Systems with Applications, 147, 113191.
- [38] Naranjo, R., Santos, M. (2019). A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition, Expert Systems with Applications, 133, 34-48.
- [39] Nevmyyaka, Y., Feng, Y., and Kearns, M. (2006). Reinforcement learning for optimized trade execution. Proceedings of the Twenty-third International Conference on Machine Learning, Pittsburgh (pp. 1–8).
- [40] Ozbayoglu, A., Gudelek, M., and Sezer, O. (2020). Deep learning for financial applications: A survey. Applied Soft Computing 93, 106384.

- [41] Pendharkar, P. C., and Cusatis, P. (2018). Trading financial indices with reinforcement learning agents. *Expert Systems with Applications* 103, 1–13.
- [42] Sezer, O., Gudelek, M., and Ozbayoglu, A. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing* 90, 106181.
- [43] Si, W., Li, J., Ding, P., Rao, R. (2017). A Multi-objective Deep Reinforcement Learning Approach for Stock Index Future's Intraday Trading. 2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, 2017, 431–436.
- [44] Sutton, R., and Barto, A. (1998). *Reinforcement Learning: An introduction*. Cambridge, MA: MIT Press.
- [45] Thrun, S., and Littman, M. (2005). Reinforcement learning: An introduction. *IEEE Transactions on Neural Networks* 16(1), 285–286.
- [46] Tsang, E. P. K., Tao, R., Serguieva, A., and Ma, S. (2017). Profiling high frequency equity price movements in directional changes. *Quantitative Finance* 17(2), 217–225.
- [47] Tsang, E., and Chen, J. (2018). Regime change detection using directional change indicators in the foreign exchange market to chart brexit. *IEEE Transactions on Emerging Topics in Computational Intelligence* 2(3), 185–193.
- [48] Vrontos, P., Galakis, J., Vrontos, I. (2020). Modeling and predicting U.S. recessions using machine learning techniques, *International Journal of Forecasting*.
- [49] Weng, L., Sun, X., Xia, M., Liu, J., and Xu, Y. (2020). Portfolio trading system of digital currencies: A deep reinforcement learning with multidimensional attention gating mechanism. *Neurocomputing* 402, 171–182.
- [50] Ye, A., Chinthapati, V., Serguieva, A., and Tsang, E. (2017). Developing Sustainable Trading strategies using directional changes with high frequency data. *IEEE International Conference on Big Data*, Boston (pp. 11–14).
- [51] Yoo, P. D., Kim, M. H., and Jan, T. (2005). Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. *Computational intelligence for modelling, control and automation and international conference on intelligent agents, web technologies and internet commerce* (pp. 835–841).