Real-time Anomaly Prediction from Cryptocurrency Time Series

Antonio Pellicani[1,3], Gianvito Pio[1,3], Sašo Džeroski[2], and Michelangelo Ceci[1,2,3]

[1] Dept. of Computer Science, University of Bari Aldo Moro, 70125 Bari, Italy
[2] Dept. of Knowledge Technologies, Jožef Stefan Institute, 1000 Ljubljana, Slovenia
[3] Big Data Lab, National Interuniversity Consortium for Informatics, Rome, Italy
{antonio.pellicani, gianvito.pio, michelangelo.ceci}@uniba.it,
saso.dzeroski@ijs.si

Abstract. Cryptocurrencies have gained popularity for their decentralized nature and ability to facilitate secure cross-border transactions, but they are susceptible to market volatility and anomalies. This extended abstract introduces a new real-time anomaly prediction method for cryptocurrency time series, that exploits temporal correlation. The proposed approach analyzes the temporal correlation between similar cryptocurrencies to identify clusters exhibiting similar patterns that may provide insights about future anomalies. Subsequently, an online multi-target LSTM model is adopted to predict upcoming anomaly events.
Our preliminary experiments on 17 real-world cryptocurrency demonstrated the potential of the proposed approach for detecting anomalies and improving trading strategies in the cryptocurrency market.

Keywords: Multi-target Classification · Anomaly Prediction · Data Stream Mining.

1 Introduction

The cryptocurrency sector has shown remarkable expansion since its emergence, peaking at a market capitalization of $3 trillion in November 2021. Cryptocurrencies are digital currencies used for online transactions, and can eventually be traded for other digital or conventional (fiat) currencies. Their unique characteristic is the absence of oversight from any central administrative authority. Transactions are recorded on the *blockchain*, a ledger system characterized by decentralization. Although the blockchain technology was initially proposed in the 90s [6], it gained recognition after Satoshi Nakamoto introduced Bitcoin in 2008. Bitcoin was designed to offer a decentralized algorithmic alternative to the conventional economic framework, which predominantly depends on central, trusted third parties, such as banks, that authorize and carry out financial transactions. Besides Bitcoin, several other cryptocurrencies have been introduced over the years, each residing on a blockchain with different characteristics.

Due to its speculative nature, the sphere of the blockchain and cryptocurrencies has attracted a lot of attention. The substantial market fluctuations observed in many cryptocurrencies (for instance, Bitcoin price journeyed from mere cents in 2009 to approximately $60k in 2021) make this landscape highly attractive to private and institutional investors, enthusiasts, and academics [2]. In such a context, conducting accurate, real-time evaluations of financial time series may be fundamental for traders and institutional investors to increase their potential profit (or reduce losses), e.g., by uncovering patterns or manipulative actions or by identifying buy/sell signals.

While prominent studies rooted in the *Efficient Market Hypothesis* [5] suggest that financial markets operate like a random walk and are inherently unpredictable, the domain of stock market analysis and forecasting attracted Machine Learning (ML) researchers, whose findings indicate that stock (and cryptocurrency) trends may be foreseeable by analyzing historical data. However, unlike the traditional stock market, the cryptocurrency market is characterized by sudden and abrupt fluctuations, that intensify the complexity of predicting its behaviour. Further complications stem from the influence of external factors, such as global politics, market cycles, public opinion, and (also fake) news.

Existing state-of-the-art systems examine past market patterns and other variables impacting a single cryptocurrency [4, 10]. However, a group effect exists across cryptocurrencies, based on several shared characteristics, such as the technology and the potential for growth. To the best of our knowledge, there is currently no approach in the literature that can capture in real-time the potential interdependence between several cryptocurrencies, which shared external aspects may impact. As a result, in this paper, we close this gap by presenting an innovative machine learning approach that examines the cryptocurrency market in real-time and takes advantage of potential relationships between different cryptocurrencies to predict possible future anomalies.

Methodologically, we initially identify clusters of cryptocurrencies affected by similar factors, as shown by their similar historical trends in terms of price, statistical indicators, and market sentiment. Subsequently, we adopt a data stream mining approach to train a deep multi-target model [3] for each identified group. These models can predict upcoming anomalous situations, such as sudden price fluctuations, for each cryptocurrency within the same cluster. It is essential to note that the problem we aim to address differs from the traditional anomaly detection task. The latter generally strives to train a batch model that can categorize an already observed instance as either *normal* or *anomalous*. In contrast, our goal is to anticipate the future occurrence of such anomalies.

The rest of the paper is organized as follows: Section 2 details the proposed method; Section 3 presents the results of a preliminary experimentation; finally Section 4 draws some conclusions and outline potential future work.

2 Methodology

In the following, we describe how the proposed approach solves the task of real-time anomaly prediction in cryptocurrency markets. We start with a preliminary data preparation phase, during which we acquire, pre-process, and store a first

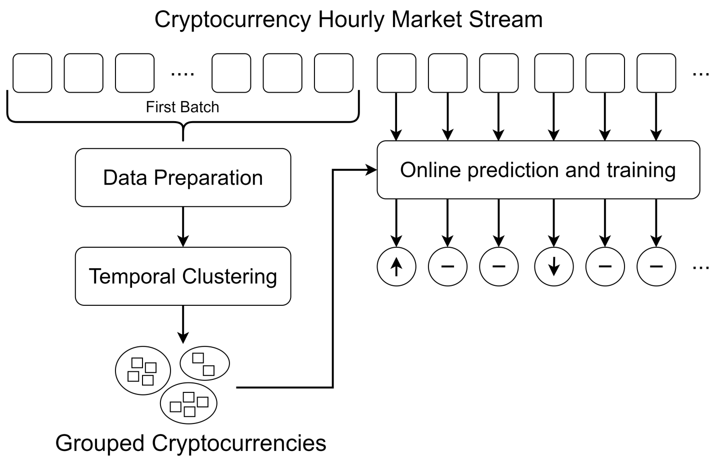

Fig. 1. General workflow of the proposed method.

batch of examples, that are used during the temporal clustering phase, which objective is to identify cryptocurrencies that exhibited a similar trend in the past. Finally, during the online prediction and training phase, we analyze the market data stream, labeling each instance and training/updating multiple deep multi-target models (one for each of the identified clusters). These models are then adopted to predict future market anomalies in real-time.

A graphical overview of the proposed approach is depicted in Figure 1.

2.1 Data preparation

The proposed method starts with an initial stage that retrieves and pre-processes data before performing the temporal clustering. We relied on the Yahoo! Finance API, which grants access to a wide range of financial information encompassing stock quotes, historical prices, company details, and trends within the crypto market. The API provides information for over 9000 unique coins. We have specifically chosen and processed hourly data for 17 of the most widely recognized cryptocurrencies based on their popularity index derived from Yahoo! Finance. The data consists of six key features, namely open price, close price, highest price, lowest price, adjusted close price, and exchange volume. These features have been consistently adopted in various studies that analyze financial data [1, 7]. Subsequently, we enhance the existing set of features by introducing 16 financial indicators with different parameter values, resulting in 54 novel features. An outline of the adopted financial indicators is provided in Table 1. Finally, in accordance with similar studies [11, 12], we incorporate a Fear and Greed indicator[4] into our dataset, that captures market sentiment by consider-

[4] www.alternative.me/crypto/fear-and-greed-index

Table 1. The considered financial indicators.

Feature	**Description**	**Parameters**
BBL_20	Bollinger Bands Lower	length = 20
BBM_20	Bollinger Bands Mid	length = 20
BBU_20	Bollinger Bands Upper	length = 20
CMO	Chande Momentum Oscillator	-
DPO	Detrend Price Oscillator	-
EMA_x	Exp. Moving Avg. on x days	$x \in \{5,12,14,20,21,23,26,30,50,100,200\}$
MACD	Moving Avg. Conv./Div.	fast = 12, slow = 26, signal = 9
MACDH	Moving Avg. Conv./Div. Hist.	fast = 12, slow = 26, signal = 9
MACDS	Moving Avg. Conv./Div. Signal	fast = 12, slow = 26, signal = 9
MOM	Momentum	-
RSI_x	Relative Strength Index on x days	$x \in \{5,12,14,20,21,23,26,30,50,100,200\}$
SMA_x	Simple Moving Average on x days	$x \in \{5,12,14,20,21,23,26,30,50,100,200\}$
STOCH_x	Slow Stochastic Oscillator on x days	$x \in \{3, 5\}$
STOCHF_x	Fast Stochastic Oscillator on x days	$x \in \{3, 14\}$
UO	Ultimate Oscillator	-
VWAP	Volume-weighted Average Price	-

ing several perspectives, including market volatility, market momentum, as well as social parameters, such as relevant hashtags and Google Trends indicators.

It is worth noting that the data preparation phase appears closely linked with the temporal clustering phase, since, at the beginning, the batch of instances that are pre-processed corresponds to the one that is subject to the clustering phase. However, data are extended with financial and sentiment indicators also in the online phase, even if according to the most recent examples of the stream.

2.2 Temporal clustering

Financial trends, especially those pertaining to cryptocurrencies, are subject to a variety of external factors, like business speculation, new product introductions, and political events. These factors can affect multiple assets simultaneously, and are often amplified by social medias. In order to capture such common influence on multiple cryptocurrencies, we group those showing analogous trends via a specific temporal clustering phase, based on the Dynamic Time Warping (DTW) distance measure [13] and the K-Medoids clustering algorithm [8].

Dynamic Time Warping (DTW) is a technique used for measuring the distance between two temporal sequences, also when they are not aligned, or have different lengths and speed. Initially introduced for speech recognition, DTW has been extensively used in time series classification.

In our approach, we first build a matrix D, by computing pair-wise DTW distances among all the cryptocurrencies, considering the first batch of data. Subsequently, we apply the k-medoids clustering algorithm on D to identify k clusters of cryptocurrencies that show similar trends. The identified clusters are then leveraged to learn multi-target prediction models, that possibly exploit dependencies among the cryptocurrencies to improve the prediction accuracy.

2.3 Online prediction and training

We analyze the continuous data stream from the cryptocurrency market, by training a different deep multi-target model for each previously identified cluster.

Although several models can be used for this complex task, we adopted Long Short-Term Memory (LSTM) networks. Indeed, this type of recurrent neural network architecture has shown to be effective when dealing with time-series data due to its ability to recognize long-term temporal dependencies. In particular, we use a multi-target variant of LSTM that concurrently learns from multiple time series and can simultaneously predict anomalies for all of them.

Given that we are dealing with a continuous data stream, we must refrain from resorting to the batch training/prediction setting, which involves learning the model from training data and using it to make predictions on new data (i.e., on the test set, during experimental evaluations). Instead, we employ the prequential method, a common practice in online learning models, where each instance in the stream is initially subject to the prediction of the current model, before the actual value is known. Subsequently, the same instance is used to update the model. This procedure is repeated for each instance in the stream.

In order to update the models, the actual label denoting if an example is normal or anomalous is required. Consequently, to label each instance, we rely on the definition of anomaly in financial time series, which corresponds to a sudden price fluctuation. Specifically, given a threshold s, we label the data according to three categories: *normal*, *upward anomaly*, and *downward anomaly*. Choosing the correct threshold is a challenging task because a low value leads to a large number of anomalies. In contrast, a high threshold may tend to identify an insufficient number of anomalous observations [9]. However, it is essential to note that setting a value for s is not a methodological decision, but rather it defines a different objective: if the goal is to learn a model that can detect minor changes, we can label the dataset with a low value for s. In contrast, if the purpose of the model is to predict significant changes, a high value for s would be more suitable. Formally, given an observation x_t, its label is defined as:

$$label(x_t) = \begin{cases} \text{an\_up} & \text{if } \Delta_t \geq s, \\ \text{an\_down} & \text{if } \Delta_t \leq -s, \\ \text{normal} & \text{otherwise,} \end{cases} \tag{1}$$

where Δ_t is the percentage of variation computed on the pair $\langle C(x_t), C(x_{t-1})\rangle$ with $C(x_t)$ and $C(x_{t-1})$ corresponding to the hourly closing price for observation x_t and x_{t-1}, respectively.

3 Experiments

To assess the performance of the proposed method, we gathered a real dataset using the Yahoo! Finance APIs, as described in Section 2. This dataset includes hourly observations of 17 of the most traded cryptocurrencies over a span of three years, from January 1, 2020, to December 31, 2022. Specifically, we considered

Table 2. Precision, Recall, F1-Score and Accuracy obtained by our approach. *Cluster* indicates the ID of the cluster in which the cryptocurrency was grouped into.

Crypto	**Cluster**	**Precision**	**Recall**	**F1 score**	**Accuracy**
BTC	3	0.612	0.566	0.587	0.681
BTS	5	0.652	0.620	0.635	0.620
DASH	4	0.626	0.589	0.607	0.594
DGB	5	0.630	0.604	0.621	0.609
DOGE	1	0.683	0.660	0.671	0.666
ETH	3	0.612	0.571	0.591	0.592
IOC	1	0.748	0.725	0.737	0.731
LTC	1	0.654	0.628	0.641	0.636
MAID	1	0.634	0.606	0.620	0.607
MONA	1	0.681	0.655	0.667	0.680
NAV	1	0.655	0.627	0.640	0.653
SYS	4	0.662	0.623	0.641	0.649
XCP	4	0.640	0.606	0.622	0.619
XLM	2	0.583	0.653	0.579	0.583
XMR	5	0.638	0.603	0.620	0.607
XRP	1	0.675	0.649	0.662	0.666
VTC	4	0.667	0.626	0.643	0.747
Average		**0.650**	**0.624**	**0.634**	**0.644**

Table 3. Average Precision, Recall, F1 Score and Accuracy for each cluster.

	Number of Cryptocurrencies	**Average Precision**	**Average Recall**	**Average F1 score**	**Average Accuracy**
Cluster 1	7	0.676	0.650	0.663	0.663
Cluster 2	1	0.583	0.653	0.579	0.583
Cluster 3	2	0.612	0.569	0.589	0.637
Cluster 4	4	0.649	0.611	0.628	0.652
Cluster 5	3	0.640	0.609	0.625	0.612

the first six months (from January 1, 2020 to June 30, 2020) as the first batch of data, i.e., to perform the initial data preparation and temporal clustering phases, as described in Sections 2.1 and 2.2. After some preliminary experiments, we set the number of clusters to $k = 5$, thereby deriving five groups of cryptocurrencies with similar behaviour during the considered period.

The remaining data (from July 1, 2020, to December 31, 2022) was adopted to simulate a continuous data stream for the online prediction and training phase. Each multi-target LSTM network was set to analyze sequences spanning 30 hours, with as many heads as the number of cryptocurrencies fallen in the corresponding cluster. The considered threshold to define an instance as anomalous (according to Eq. 1) was set to $s = 1\%$, i.e., an instance is considered an anomaly if there is a hourly variation of the closing price of at least 1%.

Table 2 presents the results obtained by our approach, in terms of precision, recall, F1 score and accuracy. We can observe that one of the cryptocurrencies under consideration, i.e., XLM, exhibited a distinctly different behaviour from

Table 4. F1 score and Accuracy obtained by the proposed approach and by a competitor batch single-target LSTM, for each cryptocurrency.

	Batch Single-target		**Online Multi-target**	
	F1 score	**Accuracy**	**F1 score**	**Accuracy**
BTC	0.423	0.581	**0.587**	**0.681**
BTS	0.404	0.485	**0.635**	**0.620**
DASH	0.556	0.585	**0.607**	**0.594**
DGB	0.366	0.417	**0.621**	**0.609**
DOGE	0.443	0.547	**0.671**	**0.666**
ETH	0.416	0.497	**0.591**	**0.592**
IOC	0.289	0.284	**0.737**	**0.731**
LTC	0.357	0.468	**0.641**	**0.636**
MAID	0.342	0.407	**0.620**	**0.607**
MONA	0.375	0.515	**0.667**	**0.680**
NAV	0.438	0.507	**0.640**	**0.653**
SYS	0.325	0.344	**0.641**	**0.649**
XCP	0.307	0.347	**0.622**	**0.619**
XLM	0.397	0.421	**0.579**	**0.583**
XMR	0.503	0.529	**0.620**	**0.607**
XRP	0.393	0.454	**0.662**	**0.666**
VTC	0.322	0.360	**0.643**	**0.747**
Average	0.391	0.456	**0.634**	**0.644**

all the others, leading to its placement in a singleton cluster, i.e., cluster 2. Looking at Table 3, we can also observe that the average results of the cluster 2 are the worst. Further highlighting the importance of capturing the relationships among cryptocurrencies, cluster 3, that is the second smallest cluster with only 2 cryptocurrencies, recorded the second-lowest F1 score. On the other hand, the other clusters, with an average size of 4, lead to better results.

In Table 4, we report the results of an additional analysis, in comparison with a single-target LSTM learned in a batch setting. As we can see from the results, the proposed real-time approach outperformed the competitor for all the cryptocurrencies. These results confirm that the proposed approach is effective in capturing the relationships in the trends and in exploiting them to improve the predictive accuracy of the learned models. Moreover, it is noteworthy that a positive contribution came also from the online update. This aspect is evident specifically for XLM, since our approach achieved better results then the batch single-target competitor, even if it also learned a single-target model for XLM.

4 Conclusions

Predicting future anomalies in the volatile cryptocurrency market is a challenging task. However, capturing temporal correlations among cryptocurrencies in their past trends can improve the predictive performance. In this paper, we introduced a new approach for real-time prediction of anomalies from cryptocurrency trends. The proposed method exploits clustering to group similar cryptocurrency time-series data, thereby highlighting their temporal correlation. Subsequently, a multi-target LSTM model is trained and updated for each group of cryptocurrencies using prequential evaluation. Experimental results emphasized that the proposed approach provides benefits in terms of predictive accuracy, also in comparison with a batch single-target LSTM model.

For future work, we plan to extend our experiments considering different clustering algorithms, different values of k and different values of the threshold s. Moreover, we plan to design an instance weighting scheme based on the strength of the anomaly, namely, on the actually observed percentage of variation.

Acknowledgements. This work was partially supported by the project FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI, under the NRRP MUR program funded by the NextGenerationEU.

References

1. Awoke, T., Rout, M., Mohanty, L., Satapathy, S.C.: Bitcoin price prediction and analysis using deep learning models. In: Communication software and networks, pp. 631–640. Springer (2021)
2. Bhutta, M.N.M., Khwaja, A.A., Nadeem, A., Ahmad, H.F., Khan, M.K., Hanif, M.A., Song, H., Alshamari, M., Cao, Y.: A survey on blockchain technology: Evolution, architecture and security. IEEE Access **9**, 61048–61073 (2021)
3. Ceci, M., Impedovo, A., Pellicani, A.: Leveraging multi-target regression for predicting the next parallel activities in event logs. In: ECML PKDD 2020 Workshops. pp. 237–248. Springer International Publishing, Cham (2020)
4. Chen, Z., Li, C., Sun, W.: Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics **365**, 112395 (2020)
5. Fama, E.F.: Efficient capital markets: A review of theory and empirical work. The journal of Finance **25**(2), 383–417 (1970)
6. Haber, S., Stornetta, W.S.: How to time-stamp a digital document. In: Conf. on the Theory and Application of Cryptography. pp. 437–455. Springer (1990)
7. Hasan, S.H., et. al: A novel cryptocurrency prediction method using optimum cnn. Computers, Materials & Continua **71**(1), 1051–1063 (2022)
8. Kaufman, L., Rousseeuw, P.J.R.: Partitioning Around Medoids (Program PAM), chap. 2, pp. 68–125. John Wiley & Sons, Ltd (1990)
9. Ko, J.U., Na, K., Oh, J.S., Kim, J., Youn, B.D.: A new auto-encoder-based dynamic threshold to reduce false alarm rate for anomaly detection of steam turbines. Expert Systems with Applications **189**, 116094 (2022)
10. Lahmiri, S., Bekiros, S.: Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons & Fractals **118**, 35–40 (2019)
11. de Oliveira Carosia, A.E., Coelho, G.P., da Silva, A.E.A.: Investment strategies applied to the brazilian stock market: A methodology based on sentiment analysis with deep learning. Expert Syst. Appl. **184**, 115470 (2021)
12. Valle-Cruz, D., et al.: Does Twitter Affect Stock Market Decisions? Financial Sentiment Analysis During Pandemics: A Comparative Study of the H1N1 and the COVID-19 Periods. Cogn. Comput. **14**(1), 372–387 (2022)
13. Vintsyuk, T.K.: Speech discrimination by dynamic programming. Cybernetics **4**(1), 52–57 (Jan 1968)