Finance Research Letters 39 (2021) 101655

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

A reality check on trading rule performance in the cryptocurrency market: Machine learning vs. technical analysis

Dan-Gabriel ANGHEL

Department of Money and Banking, The Bucharest University of Economic Studies, Romania

A R T I C L E I N F O

JEL Classification:
G11
G14

Keywords:
Reality check
Data snooping
Cryptocurrency market
Trading rules
Machine learning
Technical analysis

A B S T R A C T

This paper performs a reality check for the superior predictive ability of Machine Learning and Technical Analysis trading rules in the cryptocurrency market. After controlling for data snooping and various market frictions, we find that statistically significant positive excess returns are rarely achieved, independent of the data sampling frequency, type of trading position, or test significance level. Also, cross-sectional performance is correlated with risk factors such as beta and idiosyncratic volatility, implying that trading rules mostly capture market risk premiums. Overall, trading rules do not seem to provide additional benefits in cryptocurrency markets compared to traditional financial markets.

1. Introduction

The cryptocurrency market is rapidly expanding and has become a hot research topic in recent years. Among others, the study of trading rule performance is of great importance for investors looking for the best practical approaches to trade and diversify their portfolios in this relatively young financial market. This topic is also important for the theoretical dispute between the Efficient Market Hypothesis–EMH (Fama, 1970) and the Adaptive Market Hypothesis–AMH (Lo, 2004), the latter recently gaining some support (e.g., Khuntia and Pattanayak, 2018).

Papers examining cryptocurrencies generally find that they are "not efficient" (e.g., Zhang et al., 2018), meaning that returns are predictable to some statistically significant extent. Also, cryptoassets are less efficient compared to other asset classes (Al-Yahyaee et al., 2018; Sensoy, 2019). Even though efficiency improved in recent years (Khuntia and Pattanayak, 2018), economically significant profits seem to be attainable by trading using prediction models inspired from Technical Analysis, TA (Corbet et al., 2019; Grobys et al., 2020) or Machine Learning (ML) algorithms (Fischer et al., 2019; Chen et al., 2020; Sun et al., 2020). Such results contradict the modern interpretation of the EMH (Timmermann and Granger, 2004) for the cryptocurrency market.

Even though our understanding of cryptoassets is improving, the literature still presents an important limitation that undermines the economic relevance of reported results. Specifically, papers that analyze trading rule performance employ statistical tests that do not usually account for data snooping bias, despite the discussion on the dangers of data snooping that has gained visibility in the financial economics literature (Harvey, 2017). In this regard, the study of cryptocurrencies lags behind the one on more traditional

E-mail address: dan.anghel@fin.ase.ro.

https://doi.org/10.1016/j.frl.2020.101655
Received 3 April 2020; Received in revised form 25 May 2020; Accepted 14 June 2020
Available online 20 June 2020
1544-6123/

Table 1
Overview of ML trading rule universe.

Name	Python Package	Implementation			
SVM	sklearn	SGDClassifier(alpha=0.0001, average=False, class_weight=None, early_stopping=False, epsilon=0.1, eta0=0.0, fit_intercept=True, l1_ratio=0.15, learning_rate='optimal', loss='hinge', max_iter=1000, n_iter_no_change=5, n_jobs=None, penalty='l2', power_t = 0.5, random_state=7, shuffle=True, tol=0.001, validation_fraction=0.1, verbose=0, warm_start=False)			
Logistic Regression	sklearn	LogisticRegression(C = 1.0, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, l1_ratio=None, max_iter=1000, multi_class='auto', n_jobs=None, penalty='l2', random_state=None, solver='lbfgs', tol=0.0001, verbose=0, warm_start=False)			
Random Forest	sklearn	RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None, criterion='gini', max_depth=None, max_features='auto', max_leaf_nodes=None, max_samples=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=2000, n_jobs=None, oob_score=False, random_state=7, verbose=0, warm_start=False)			
RNN	keras	Layer (type)	Output Shape	Param #	
		simple_rnn (SimpleRNN)	*(None, None, 4)*	*36*	
		simple_rnn_1 (SimpleRNN)	*(None, 4)*	*36*	
		dense_5 (Dense)	*(None, 1)*	*5*	
WDNN	keras	Layer (type)	Output Shape	Param #	Connected to
		input_1 (InputLayer)	*[(None, 34)]*	*0*	
		dropout (Dropout)	*(None, 34)*	*0*	*input_1[0][0]*
		dense (Dense)	*(None, 32)*	*1120*	*dropout[0][0]*
		dropout_1 (Dropout)	*(None, 32)*	*0*	*dense[0][0]*
		dense_1 (Dense)	*(None, 32)*	*1056*	*dropout_1[0][0]*
		dropout_2 (Dropout)	*(None, 32)*	*0*	*dense_1[0][0]*
		dense_2 (Dense)	*(None, 32)*	*1056*	*dropout_2[0][0]*
		dropout_3 (Dropout)	*(None, 32)*	*0*	*dense_2[0][0]*
		dense_3 (Dense)	*(None, 32)*	*1056*	*dropout_3[0][0]*
		concatenate (Concatenate)	*(None, 66)*	*0*	*input_1[0][0]*
					dense_3[0][0]
		dense_4 (Dense)	*(None, 1)*	*67*	*concatenate[0][0]*

This may be Funding Agency. Please check.financial assets.[1] Moreover, TA and ML trading rules have not been analyzed together, in a multiple-hypothesis test that controls for the combined data snooping efforts of all researchers and investors.

This paper fills this gap by performing such an analysis. Our contribution is threefold. First, we use the Reality Check test (White, 2000) to make a data-snooping-free evaluation of the superior predictive ability of forecasting models inspired from ML and TA in the cryptocurrency market. Second, we test ML and TA together, effectively evaluating which of the two is better suited for predicting cryptoasset prices. Third, we account for market frictions (trading costs, restrictions on short trades) and risk, and evaluate the robustness of our results to changes in data sampling frequency (daily vs. intraday). Besides providing new insights that can help investors better navigate cryptocurrency markets, our results enable us to draw inferences about the weak-form efficiency of cryptoassets.

2. Data and methodology

2.1. Data sample

We analyze the entire cryptocurrency market with minimum filtering to assure a proper generalization of the results. Specifically, we retrieve all daily trading histories for cryptocurrencies tracked by www.coinmarketcap.com on February 10, 2020. For all trading days, we collect information on Open, High, Low, and Close prices (vs. the US dollar) and dollar traded volumes. Filtering out assets that have less than 730 days of available data–which we consider a minimum requirement to apply ML algorithms[2]–leaves a sample of 1000,036 daily observations for 861 cryptocurrencies. A data summary is provided as an online appendix.

Because previous results hint that price predictability may be more pronounced when analyzing higher data sampling frequencies (Sensoy, 2019; Chen et al., 2020), we perform additional tests using intraday data for Bitcoin, Ethereum, Litecoin, and Zcash. The sample is collected from the Gemini Exchange via https://www.cryptodatadownload.com/ and consists of prices and trading volumes sampled every 1 h, from as early as October 8, 2015 (for Bitcoin) through March 13, 2020. Tests on shorter non-overlapping samples of

[1] Park and Irwin (2007) provide an extensive review of the early literature. Recent contributions include Shynkevich (2012) for the stock market, Coakley et al. (2016) for the foreign exchange market, Shynkevich (2016) for the bond market, or Han, Hu and Yang (2016) for the commodity futures market.

[2] This assures a good balance between having sufficient time-series observations to avoid training suboptimal ML rules, having sufficient out-of-sample predictions to estimate ML performance, and keeping a sufficient number of cryptocurrencies in the sample as to achieve the goals of the paper.

Table 2
Overview of TA trading rule universe.

Rule Type	Strategy Type	Buy rule: ($Signal_{t+1} = 1$ when)	Sell rule: ($Signal_{t+1} = 0 or - 1$ when)	Parameterization settings	Total rules
F	TS	$L_t > (1+x)\underset{0<k<t}{\text{argmax}}\{L_k \mid L_k < L_{k-1}\}$	$L_t < (1-x)\underset{0<k<t}{\text{argmax}}\{L_k \mid L_k \rangle L_{k-1}\}$	$x \in [1\%: 1\%: 50\%]$;	50
TRB	TS	$L_t > \underset{1\le k\le n_1}{\text{argmax}}\{L_{t-k}\}$	$L_t < \underset{1\le k\le n_1}{\text{argmin}}\{L_{t-k}\}$	$n_1,n_2 \in [2: 3: 143]$;	2304
MA	TS	$MA_t(n_1) > MA_t(n_2)$	$MA_t(n_1) \le MA_t(n_2)$	$n_1,n_2 \in [2: 3: 143]$; $n_1 < n_2$;	1128
MACD	TS	$MACDp_t(n_1,n_2) > x$	$MACDp_t(n_1,n_2) \le x$	$n_1,n_2 \in [2: 3: 143]$; $n_1 < n_2$; $x \in [-16\%: 2\%: 16\%]$; $x \ne 0$	18,048
	TR	$MACD_t(n_1,n_2) > S_t(n_3)$	$MACD_t(n_1,n_2) \le S_t(n_3)$	$n_1,n_2 \in [2: 3: 143]$; $n_1 < n_2$; $n_3 \in [2: 3: 14]$;	5640
	TR	$MACD_t(n_1,n_2) > D_t(n_3)$	$MACD_t(n_1,n_2) \le D_t(n_3)$	$n_1,n_2 \in [2: 3: 143]$; $n_1 < n_2$; $n_3 \in [2: 3: 5]$;	2256
RSI	TS	$RSI_t(n) > x$	$RSI_t(n) \le x$	$n \in [2: 3: 143]$; $x \in [2: 2: 98]$;	2352
	TR	$RSI_t(n_1) > S_t(n_2)$	$RSI_t(n_1) \le S_t(n_2)$	$n_1 \in [2: 3: 143]$; $n_2 \in [2: 3: 14]$;	240
	TR	$RSI_t(n_1) > D_t(n_2)$	$RSI_t(n_1) \le D_t(n_2)$	$n_1 \in [2: 3: 143]$; $n_2 \in [2: 3: 5]$;	96
	TR-OS	$RSI_t(n) \ddots \uparrow x_1$	$RSI_t(n) > x_2$	$n \in [2: 3: 143]$; $x_1,x_2 \in [5: 5: 95]$;	17,328
	TR-OS	$RSI_t(n) \ddots \uparrow x_1$	$RSI_t(n) \therefore \downarrow x_2$	$n \in [2: 3: 143]$; $x_1,x_2 \in [5: 5: 95]$;	17,328
OBV	TS	$OBV_t(n_1) > S_t(n_2)$	$OBV_t(n_1) \le S_t(n_2)$	$n_1 \in [2: 2: 142]$; $n_2 \in [2: 2: 14]$;	497
	TS	$OBV_t(n_1) > D_t(n_2)$	$OBV_t(n_1) \le D_t(n_2)$	$n_1 \in [2: 2: 142]$; $n_2 \in [2: 2: 6]$;	213

Note. TS = Trend seeking strategy. TR = Trend reversal strategy. TR-OS = Trend reversal based on oversold signals. $\ddots \uparrow$ *is the operator denoting the indicator on the left is lower than and then rises above the value on the right.* $\therefore \downarrow$ *is the operator denoting the indicator on the left is higher than and then falls below the value on the right. MACDp denotes MACD expressed as a percentage of the last price.* $S(n)$ *denotes a signal band representing an n-period moving average of the indicator.* $D(n)$ *denotes a signal band representing an n-period delayed value of the indicator. The notation* $x \in [min: i: max]$ *is used to denote* $x \in \{min, min + i, min + 2i, min + 3i, \ldots, max\}$.

1 year, 1 month, and 1 week are also performed on the intraday data, thus enabling the analysis of possible time-variation in trading rule performance.

2.2. Trading rules

For testing if ML and TA have superior predictive ability, we frame the problem as a classification task. Specifically, we construct models that try to predict the 1-period-ahead direction of price movements, up or down, and not the exact expected price. Thus, the predictions made by the various models can be associated with trading "signals", i.e. the positions that investors should take in the market (long, short, or stay out of the market).

A set of prediction models is typically labeled as a trading rule "universe". The "ML" universe contains 5 trading rules and it is based on both traditional and recently introduced classification models: Support Vector Machines–SVM (Cortes and Vapnik, 1995), Logistic Regression, Random Forest (Ho, 1998), Recurrent Neural Networks–RNN (Rumelhart et al., 1986), and Wide and Deep Neural Networks–WDNN (Cheng et al., 2016). The trading rules are constructed and implemented in *Python*, with the help of the *Scikit-Learn, Keras*, and *TensorFlow* libraries. For brevity, the implementation for each model is listed in Table 1, while the reader is referred to Géron (2019) for details regarding hyperparameters and other properties.

ML predictions are obtained as follows. First, log-returns are computed from last prices, $r_t = \ln(L_t/L_{t-1})$. Second, a binary variable is constructed, taking 1 if the next period return is positive and 0 otherwise. The ML algorithms use this as the TARGET variable in the training phase, and try to predict it in the testing phase. Third, the set of features needed by the ML classifiers is constructed. Following similar approaches (e.g., Fischer et al., 2019), we use the average m-period return as a proxy for price momentum, $R_{t-m+1,t} = \sum_{k=1}^{m} r_{t-k+1}/m$, setting $m \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 17, 20, 25, 30, 40, 50, 70, 90, 120\}$. However, we add proxies for volatility ($v_t = r_t^2$), skewness ($s_t = r_t^3$), and liquidity ($l_t = V_t$, where V_t denotes trading volume), averaged over the same intervals as momentum. In total, 84 features are used to train the classifiers, except RNNs for which only 4 features (with $m = 1$) are used (RNNs explicitly handle time-series data). Fourth, we select an initial subsample consisting of the first year of trading data (365 days or 8760 h) and train each model on 70% of the subsample, cross-validating it on the remaining 30%. This 2-fold cross-validation tries to avoid overfitting by also employing early stopping when the prediction error on the validation set starts to diverge from the one on the training set. Finally, the fitted models are used to make one-period-ahead predictions on the remaining sample, updating the features at each step and retraining rules every 28 calendar days.[3]

The rival "TA" universe is defined using classes of TA indicators that are popular among investors and have also been widely employed in the literature (see, e.g., Park and Irwin, 2007). Specifically, we employ filters (F), trading range breakouts (TRB), moving average crossovers (MA), the Moving Average Convergence Divergence (MACD) indicator, the Relative Strength Index (RSI) indicator, and the On Balance Volume (OBV) indicator.[4] This selection assures that both trend-seeking and trend-reversal rules are considered. Actual trading rules are constructed by defining trading strategies and setting values to their parameters. Table 2 presents an overview

[3] ML rules can be retrained more often but this would not bring any significant benefits because adding few observations to the training sample would not materially change the estimated parameters and, thus, trading performance but it would be significantly more computationally expensive.
[4] Colby (2002) provides details on the characteristics of these indicators and how they can be implemented.

of this process. The resulting TA universe has a total of 67,480 trading rules. The size of the TA universe may seem excessive but incorporating all relevant alternatives used by researchers and investors assures a proper control for data snooping (White, 2000). The universe may even be unrepresentative compared to what others employ. For example, Coakley et al. (2016) use 113,148 rules when analyzing the FX market.[5]

2.3. Testing methodology

We evaluate the superior predictive ability of ML and TA rules by using their predictions to simulate trading, and then measure and test their empirical performance. We handle look-ahead bias by performing all trades on the Open price from the day following each prediction. To control for data snooping bias, we test the results using the Reality Check (RC) methodology (see White, 2000, for details), which evaluates the null hypothesis that *the best performing prediction model* in the universe is not capable of outperforming a benchmark model. The naïve buy and hold strategy (which buys the asset on day 1 and holds it until the end of the sample) is set as the benchmark, while the average excess return over the benchmark is set as the performance measure.

We vary test options and perform several runs to achieve our research goals. First, to evaluate the impact of data snooping and to directly compare ML with TA, we test 3 increasingly larger universes: ML, ML & MA, and ML & TA; where MA is a subset of TA (Table 2). Second, we evaluate performance both in a cost-free environment and in an environment with a fixed trading fee of 0.1% per trade.[6] Liquidity cost and price impact costs are also considered by substituting the Open price with the High price when buying and the Low price when selling. Third, we test the influence of short trading restrictions[7] by evaluating performance with and without them. Finally, we note that our implementation of the RC uses the stationary bootstrap of Politis and Romano (1994) with $B = 10{,}000$ iterations when the ML and ML&MA universes are tested, and $B_1 = 100$ otherwise.[8] In this latter case, we further bootstrap the best rule $B_2 = 10{,}000$ times for better reliability.

3. Results

Table 3 summarizes the results obtained by *the best trading rules* (the ones that maximize trading performance on each cryptocurrency) in 8 relevant test runs performed on daily data. Several interesting findings are worth noting. First, ML models have significant trading performance in cryptocurrency markets but underperform their TA counterparts when the two are considered together. Interestingly, the superior performance of ML and MA rules is approximately equal. Second, eliminating short positions significantly worsens trading performance compared to just staying out of the market. This becomes especially evident when strictly controlling for data snooping and trading costs in the final test runs. Third, and most importantly, trading rule excess performance increases with the size of the rule universe but the proportion of statistically significant results goes in the opposite direction. This constitutes evidence that data snooping bias significantly impacts test results when it is not handled. In the final test runs, only 4.41% RC null hypotheses are rejected at the 5% level (1.74% at the 1% level) when long positions are used. This slightly increases to 9.87% (4.53% at the 1% level) when both long and short positions are considered.

Fig. 1 reports the results in run 8, aggregated for 10×10 baskets of cryptocurrencies sorted by size and liquidity. Among others, it shows that ML does outperform TA on small, illiquid cryptocurrencies; and that RC null rejections mainly occur on liquid cryptocurrencies.

We elaborate on this result by performing several regressions that evaluate the relationship between cross-sectional variation in trading rule overperformance and possible cryptocurrency risk factors (see, e.g., Liu et al., 2020), such as their contribution to systematic risk (beta, estimated by regressing their return series on the equal-weighted index of all 861 cryptocurrencies in our sample), size (market capitalization at the end of the sample), liquidity (daily dollar trading volume averaged over the entire sample), idiosyncratic volatility (standard deviation of errors obtained from the beta regression) and conditional skewness (see, Harvey and Siddique, 2000). The results (Table 4) show that the cross-sectional overperformance of trading rules is correlated with some factors, in particular market beta and idiosyncratic volatility. Size and liquidity also explain superior performance but only in tests that use smaller universes or ignore trading costs (available at request). This implies that ML and TA rules are useful in capturing cryptocurrency risk premiums and that RC null rejections do not necessarily contradict modern-EMH (Timmermann and Granger, 2004).

Overall, the results show that deviations from market efficiency are relatively scarce and, when they do occur, can be associated with higher market frictions (trading costs, restrictions on short trades) and risk. This implies that the cryptocurrency markets are mostly weak-form informationally efficient. Robustness tests performed on intraday data support this conclusion. Table 5 reports a

[5] The universe may also seem unbalanced in favor of TA rules but this is not the case because the parameters for the 5 ML rules are selected (optimized) from a much larger parameter space.

[6] Actual trading fees vary by exchange, cryptoasset, portfolio value, trading activity, etc. but generally range between 0% and 0.2% (e.g., Binance, https://www.binance.com/en/fee/schedule). We report break-even transaction costs as an alternative way to evaluate trading performance given investor heterogeneous constraints.

[7] Some exchanges do not allow short trades, while others require some amount of the sold cryptoasset as collateral. This use of margin trading generates additional trading costs: opening fees, rollover fees (e.g., Kraken, https://www.kraken.com/en-us/features/fee-schedule). Moreover, similar to other financial assets, no lenders exist for some cryptocurrencies that are illiquid or in short supply.

[8] This aims to reduce computational demand. A robustness check performed on a limited subsample of cryptoassets using B_1=1000 yields that results do not materially change.

Table 3
Summary of results obtained ML and TA trading rules on daily data.

Panel A. Cross-sectional distribution of annualized excess returns

Rule universe,	ML,		ML & MA,		ML & TA		ML & TA	
Trading costs	0%		0%		0%		0.1%+liquidity	
Trading positions	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short
Minimum	−546.93%	−222.20%	−37.85%	1.86%	9.19%	14.22%	2.92%	3.81%
1st Decile	−8.76%	2.60%	87.40%	142.46%	98.50%	152.68%	35.31%	82.43%
2nd Decile	30.63%	50.04%	127.50%	202.23%	130.95%	205.81%	50.49%	114.98%
3rd Decile	59.06%	94.48%	160.78%	255.57%	170.86%	265.62%	65.56%	140.49%
4th Decile	92.24%	147.85%	197.56%	306.53%	207.49%	318.68%	78.31%	169.04%
Median	136.74%	205.41%	238.37%	366.75%	242.40%	364.57%	89.78%	198.71%
6th Decile	183.45%	282.75%	274.60%	429.76%	288.83%	431.87%	106.34%	231.33%
7th Decile	270.32%	420.37%	330.06%	504.01%	353.55%	511.77%	126.16%	262.66%
8th Decile	408.93%	561.89%	435.15%	620.44%	451.49%	627.77%	146.42%	313.72%
9th Decile	614.19%	866.10%	634.75%	909.84%	640.23%	879.53%	185.14%	381.38%
Maximum	2688.61%	3994.87%	2688.61%	3994.87%	2688.61%	3994.87%	364.65%	1421.20%
Average	244.16%	361.62%	321.16%	477.42%	333.83%	484.63%	103.08%	223.90%
Std. Dev.	340.75%	481.43%	308.48%	438.86%	310.06%	437.97%	60.86%	141.18%
Skewness	2.73	2.99	3.14	3.32	3.09	3.35	1.07	2.37
Kurtosis	11.54	13.57	14.94	16.97	14.38	17.02	1.33	12.73

Panel B. Cross-sectional distribution of Excess Break-Even Transaction Costs

Rule universe,	ML,		ML & MA,		ML & TA		ML & TA	
Trading costs	0%		0%		0%		0.1%+liquidity	
Trading positions	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short
Minimum	−16.91%	−25.32%	−0.83%	0.09%	0.13%	0.22%	1.27%	1.79%
1st Decile	−0.38%	0.05%	1.32%	2.27%	0.85%	2.00%	5.23%	8.19%
2nd Decile	0.59%	1.05%	1.88%	3.42%	1.24%	3.32%	8.08%	11.38%
3rd Decile	1.09%	1.68%	2.43%	4.25%	1.83%	4.53%	10.70%	15.57%
4th Decile	1.49%	2.41%	2.93%	5.11%	2.54%	5.63%	13.71%	19.91%
Median	2.08%	3.24%	3.65%	6.12%	3.42%	7.02%	17.69%	25.68%
6th Decile	2.79%	4.36%	4.48%	7.42%	4.29%	8.95%	23.50%	33.22%
7th Decile	3.67%	5.59%	5.63%	8.89%	5.73%	12.37%	32.87%	43.15%
8th Decile	5.12%	7.22%	6.82%	11.25%	7.62%	16.35%	46.74%	63.71%
9th Decile	7.72%	10.94%	9.29%	14.83%	11.85%	26.79%	92.02%	120.97%
Maximum	135.60%	135.60%	135.60%	135.60%	76.98%	708.25%	692.47%	916.08%
Average	3.09%	4.82%	5.00%	8.16%	5.54%	13.66%	39.68%	59.89%
Std. Dev.	6.32%	7.53%	6.40%	8.37%	7.61%	37.10%	63.12%	110.59%
Skewness	11.19	7.39	11.38	6.44	4.66	15.33	4.00	4.20
Kurtosis	225.58	110.66	208.84	75.39	29.91	276.75	22.60	19.85

Panel C. Cross-sectional distribution of Reality Check p-values

Rule universe,	ML,		ML & MA,		ML & TA		ML & TA	
Trading costs	0%		0%		0%		0.1%+liquidity	
Trading positions	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1st Decile	0.0020	0.0011	0.0011	0.0008	0.0000	0.0000	0.3400	0.0600
2nd Decile	0.0081	0.0060	0.0039	0.0029	0.0100	0.0100	0.6400	0.2000
3rd Decile	0.0217	0.0184	0.0079	0.0073	0.0300	0.0200	0.8300	0.4000
4th Decile	0.0453	0.0425	0.0151	0.0140	0.0700	0.0400	0.9200	0.6100
Median	0.0882	0.0833	0.0257	0.0228	0.1200	0.0700	0.9600	0.7600
6th Decile	0.1649	0.1512	0.0448	0.0405	0.2000	0.1400	0.9900	0.8600
7th Decile	0.2614	0.2604	0.0788	0.0818	0.2900	0.2600	1.0000	0.9300
8th Decile	0.4660	0.4243	0.1383	0.1465	0.4400	0.4100	1.0000	0.9800
9th Decile	0.8244	0.7252	0.3081	0.2855	0.6900	0.6200	1.0000	1.0000
Maximum	1.0000	1.0000	1.0000	0.9701	1.0000	0.9900	1.0000	1.0000
Average	0.2418	0.2211	0.1053	0.1003	0.2332	0.2038	0.8188	0.6289
Std. Dev.	0.3124	0.2826	0.1924	0.1801	0.2659	0.2587	0.2763	0.3542
Skewness	1.3382	1.3231	2.8942	2.8248	1.1750	1.3655	−1.6690	−0.5430
Kurtosis	0.4243	0.5168	8.5330	8.2218	0.2659	0.8078	0.0000	0.0000

Panel D. Summary of Reality Check test results

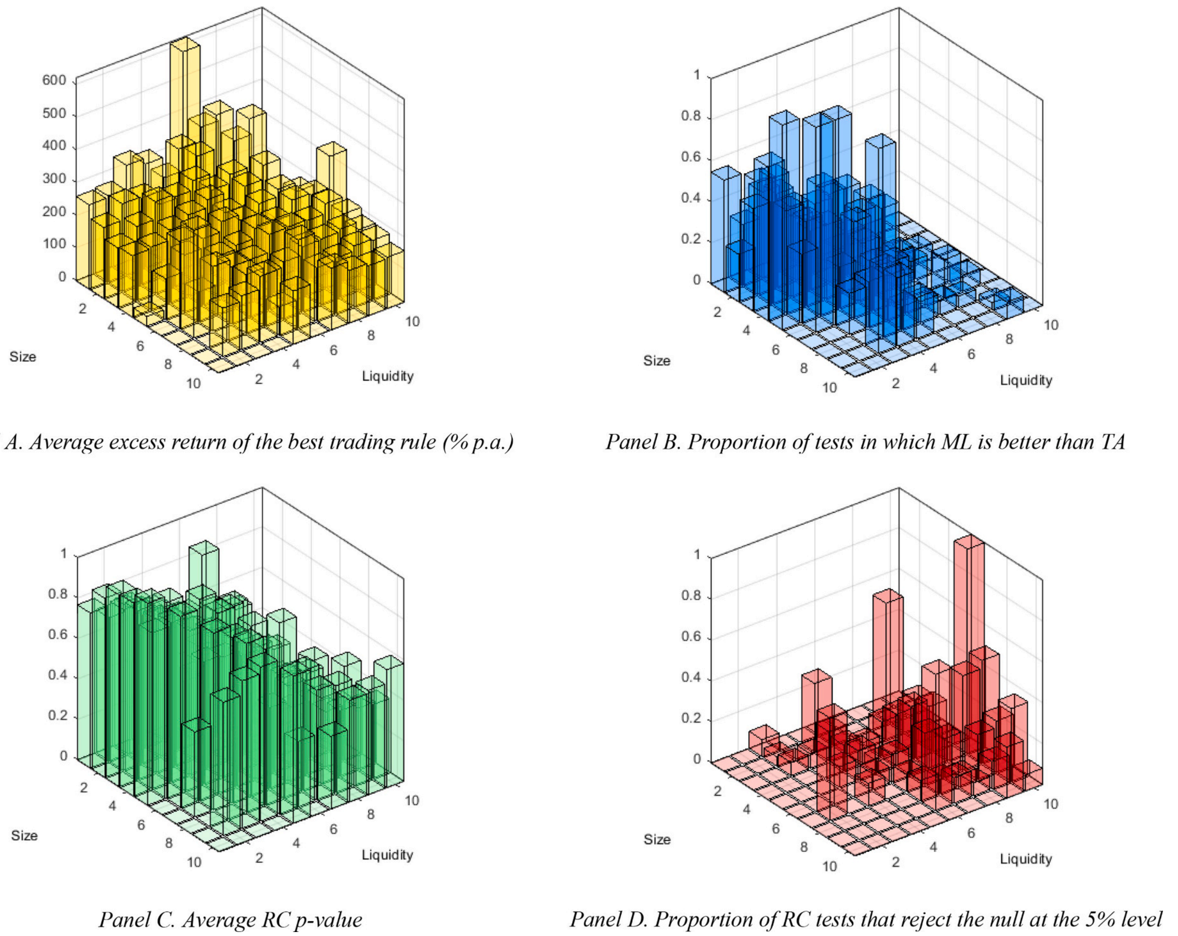
Rule universe,	ML,		ML & MA,		ML & TA		ML & TA	
Trading costs	0%		0%		0%		0.1%+liquidity	
Trading positions	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short
Total RC tests	861	861	861	861	861	861	861	861

(continued on next page)

Table 3 (*continued*)

Panel D. Summary of Reality Check test results								
Rule universe,	ML,		ML & MA,		ML & TA		ML & TA	
Trading costs	0%		0%		0%		0.1%+liquidity	
ML better in	*100.00%*	100.00%	38.10%	39.49%	30.55%	33.33%	0.70%	3.60%
Null rejections, 10%	451	458	633	632	406	483	38	117
of which ML	*100.00%*	100.00%	42.97%	42.56%	42.36%	40.99%	0.00%	0.85%
Null rejections, 5%	355	369	531	546	326	388	32	85
of which ML	*100.00%*	100.00%	45.20%	43.96%	41.41%	41.49%	0.00%	0.00%
Null rejections, 1%	185	216	293	303	186	233	15	39
of which ML	*100.00%*	100.00%	53.24%	53.14%	45.70%	46.78%	0.00%	0.00%

Note. This table reports the results obtained by the best trading rules (in each tested rule universe) on daily data for 861 cryptocurrencies. Panel A reports the cross-sectional distribution of the average (annualized) difference in returns between the best trading rule and the benchmark buy and hold rule. Panel B reports the cross-sectional distribution of break-even transaction costs, estimated as the total excess return earned by the best rule divided by its total number of trades. Panel C reports the cross-sectional distribution of p-values obtained when evaluating the statistical significance of the best average excess returns using the Reality Check test (White, 2000). Panel D summarizes the results of all tests and provides an overview regarding the relative performance of Machine Learning (ML) and Technical Analysis (TA) rules.



Panel A. Average excess return of the best trading rule (% p.a.)

Panel B. Proportion of tests in which ML is better than TA

Panel C. Average RC p-value

Panel D. Proportion of RC tests that reject the null at the 5% level

Fig. 1. Trading rule performance vs. size and liquidity. *Panel A. Average excess return of the best trading rule (% p.a.). Panel B. Proportion of tests in which ML is better than TA, Panel C. Average RC p-value, Panel D. Proportion of RC tests that reject the null at the 5% level, Note. This figure reports the performance of the best trading rules in the ML&TA universe aggregated for 10 × 10 baskets of cryptocurrencies sorted by size and liquidity.*

summary of the results obtained when testing the ML&TA universe while fully subtracting trading fees and liquidity costs. Among others, no RC null hypothesis is rejected at the 1% level and only 5 are rejected at the 5% level. Moreover, a simple time series analysis of weekly trading rule performance (Table 6) reveals no significant linear auto-dependency of statistically significant results, implying that null rejections occur randomly. Interestingly, RC p-values for Bitcoin record a statistically significant negative tendency. Contrary

Table 4
Cryptocurrency risk factors and the cross-section of trading rule performance.

Dependent variable (Y_t):	Excess returns		Break-even costs		RC p-values	
Trading costs:	No	Yes	No	Yes	No	Yes
δ	−0.4516	1.5303	0.0381	0.4987	0.2608	0.5731
	[−0.80]	[6.49]***	[0.55]	[2.48]**	[5.56]***	[9.96]***
δ_{BETA}	1.0819	0.2731	0.0851	−0.1502	−0.0699	−0.1888
	[2.01]**	[1.28]	[1.36]	[−0.83]	[−1.65]*	[−3.64]***
δ_{SIZE}	−1.29e-11	−1.93e-12	4.22e-13	−5.88e-13	4.83e-12	1.17e-12
	[−0.43]	[−0.16]	[0.12]	[−0.05]	[2.06]**	[0.40]
δ_{LIQ}	7.75e-14	−1.38e-13	−6.97e-15	−5.45e-14	−9.41e-15	3.11e-14
	[0.22]	[−1.03]	[−0.17]	[−0.47]	[−0.35]	[0.95]
δ_{VOL}	26.0850	2.4720	0.0793	1.6733	0.0812	1.5869
	[21.09]***	[5.05]***	[0.55]	[4.02]***	[0.83]	[13.31]***
δ_{COSKEW}	−1.1495	−0.3884	0.0431	−0.2682	−0.0381	0.0347
	[−1.09]	[−0.93]	[0.35]	[−0.75]	[−0.46]	[0.34]
$F-stat$	93.60***	6.48***	0.53	3.53***	2.71**	40.81***
R^2	0.3682	0.0388	0.0033	0.0215	0.0166	0.2026

Note. This table shows the results of the regression $Y_k = \delta + \delta_{BETA}Beta_k + \delta_{SIZE}Size_k + \delta_{LIQ}Liquidity_k + \delta_{VOL}Volatility_k + \delta_{COSKEW}Coskewness_k + \varepsilon_k$. *Total included observations: 806. T-statistics are reported in square parenthesis. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.*

Table 5
Summary of results obtained on intraday data.

Sample length	Full Sample		1 Year		1 Month		1 Week	
Trading positions	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short	Long only	Long & Short
Total RC tests	4	4	13	13	94	94	391	391
ML better in	*0%*	*0%*	0%	*0%*	0%	*0%*	0%	*0%*
Null rejections, 10%	0	0	0	1	0	4	0	10
of which ML	*N/A*	*N/A*	*N/A*	*0%*	*N/A*	0%	*N/A*	0%
Null rejections, 5%	0	0	0	0	0	1	0	4
of which ML	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	0%	*N/A*	0%
Null rejections, 1%	0	0	0	0	0	0	0	0
of which ML	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	N/A	*N/A*	N/A

Table 6
Time-varying performance in the cryptocurrency market.

Dependent variable (Y_t):	Bitcoin ($n = 178$)		Ethereum ($n = 149$)		Litecoin ($n = 21$)		Zcash ($n = 42$)	
	Excess return	RC p-value	Excess return	RC p-value	Excess return	RC p-value	Excess return	RC p-value
α	7.6972	0.7552	15.3519	0.7672	−3.4581	0.5868	5.6660	0.5483
	[4.26]***	[10.21]***	[5.06]***	[9.86]***	[−0.37]	[2.37]**	[0.93]	[3.82]***
β	0.0037	−0.0008	−0.0280	−0.0006	1.3731	−0.0082	0.2672	0.0006
	[0.24]	[−2.03]**	[−1.01]	[−1.10]	[1.95]*	[−0.62]	[1.39]	[0.16]
γ	0.2509	0.0657	0.1492	−0.0383	0.3549	0.2178	0.3389	0.1288
	[3.10]***	[0.86]	[1.49]	[−0.45]	[0.83]	[0.88]	[1.34]	[0.78]
$F-stat$	4.81***	2.79*	2.06	0.69	3.07*	0.63	1.89	0.32
R^2	0.0521	0.0310	0.0276	0.0094	0.2657	0.0691	0.0905	0.0170

Note. This table shows the results of the regression $Y_t = \alpha + \beta t + \gamma Y_{t-1} + \varepsilon_t$. *T-statistics are reported in square parenthesis. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.*

to previous results (e.g. Khuntia and Pattanayak, 2018), this directly questions the Adaptive Market Hypothesis in our sample.

4. Conclusions

We investigate the economic relevance of Machine Learning (ML) and Technical Analysis (TA) in the cryptocurrency market and find that their trading performance is not that different from more traditional financial markets. Specifically, ML and TA are capable to successfully forecast future price movements and capture risk premiums but they are not generally capable of earning “abnormal” returns after controlling for data snooping, market frictions, and risk. This implies that cryptoassets are more efficient than previously recognized (e.g., Corbet et al., 2019; Grobys et al., 2020).

We also find that our selection of ML rules generally underperforms simple alternatives derived from TA, especially after incorporating trading costs. Our ML rules trade more often but with no additional benefits, hinting that implementing them in practice is not worth the effort, as cheaper and easier to use TA strategies are sufficiently equipped to help investors navigate cryptocurrency markets. As an important exception, ML does outperform TA on small and less liquid cryptoassets but with seemingly no economic benefits (excess returns on these assets are not significant after controlling for data snooping). These surprising findings motivate future research. Why is this happening? How does it relate to different groups of cryptocurrencies and their exchanges? Would the results change if implementing hyperparameter optimization or more advanced ML architectures that may be used by investors (this is a limitation of our analysis)? In the end, many other forecasting models exist and they may also be considered. Adding relevant rules derived from ML, TA, etc. would further help to mitigate data snooping and improve the robustness of reported results.

CRediT authorship contribution statement

Dan-Gabriel ANGHEL: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition.

Acknowledgements

This paper was co-financed from the Human Capital Operational Program 2014–2020, project number POCU / 380/6/13/125245 no. 36482 / 23.05.2019 “Excellence in interdisciplinary PhD and post-PhD research, career alternatives through entrepreneurial initiative (EXCIA)”, coordinator The Bucharest University of Economic Studies.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.frl.2020.101655.

References

Al-Yahyaee, K.H., Mensi, W., Yoon, S.M., 2018. Efficiency, multifractality, and the long-memory property of the Bitcoin market: a comparative analysis with stock, currency, and gold markets. Finance Research Letters 27, 228–234.
Chen, Z., Li, C., Sun, W., 2020. Bitcoin price prediction using machine learning: an approach to sample dimension engineering. J Comput Appl Math 365, 112395.
Cheng, H.T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M. and Anil, R., 2016, September. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop On Deep Learning For Recommender Systems(7–10).
Coakley, J., Marzano, M., Nankervis, J., 2016. How profitable are FX technical trading rules? International Review of Financial Analysis 45, 273–282.
Colby, R.W., 2002. The Encyclopedia of Technical Market Indicators (2nd ed.). New York: McGraw-Hill.
Corbet, S., Eraslan, V., Lucey, B., Sensoy, A., 2019. The effectiveness of technical trading rules in cryptocurrency markets. Finance Research Letters 31, 32–37.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach Learn 20 (3), 273–297.
Fama, E.F., 1970. Efficient capital markets: a review of theory and empirical work. J Finance 25 (2), 383–417.
Fischer, T.G., Krauss, C., Deinert, A., 2019. Statistical arbitrage in cryptocurrency markets. Journal of Risk and Financial Management 12 (1), 31.
Géron, A., 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media.
Grobys, K., Ahmed, S., Sapkota, N., 2020. Technical trading rules in the cryptocurrency market. Finance Research Letters 32, 101396.
Han, Y., Hu, T., Yang, J., 2016. Are there exploitable trends in commodity futures prices? J Bank Financ 70, 214–234.
Harvey, C.R., 2017. Presidential address: the scientific outlook in financial economics. J Finance 72 (4), 1399–1440.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J Finance 55 (3), 1263–1295.
Ho, T.K., 1998. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20 (8), 832–844.
Khuntia, S., Pattanayak, J.K., 2018. Adaptive market hypothesis and evolving predictability of bitcoin. Econ Lett 167, 26–28.
Liu, W., Liang, X., Cui, G., 2020. Common risk factors in the returns on cryptocurrencies. Econ Model 86, 299–305.
Lo, A.W., 2004. The adaptive markets hypothesis. The Journal of Portfolio Management 30 (5), 15–29.
Park, C.H., Irwin, S.H., 2007. What do we know about the profitability of technical analysis? J Econ Surv 21 (4), 786–826.
Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J Am Stat Assoc 89 (428), 1303–1313.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323 (6088), 533–536.
Sensoy, A., 2019. The inefficiency of Bitcoin revisited: a high-frequency analysis with alternative currencies. Finance Research Letters 28, 68–73.
Sun, X., Liu, M., Sima, Z., 2020. A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Research Letters 32, 101084.
Shynkevich, A., 2012. Performance of technical analysis in growth and small cap segments of the US equity market. J Bank Financ 36 (1), 193–208.
Shynkevich, A., 2016. Predictability in bond returns using technical trading rules. J Bank Financ 70, 55–69.
Timmermann, A., Granger, C.W., 2004. Efficient market hypothesis and forecasting. Int J Forecast 20 (1), 15–27.
White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.
Zhang, W., Wang, P., Li, X., Shen, D., 2018. The inefficiency of cryptocurrency and its cross-correlation with Dow Jones Industrial Average. Physica A: Statistical Mechanics and its Applications 510, 658–670.