Available online at www.sciencedirect.com

Borsa İstanbul Review

Borsa İstanbul Review 23-S2 (2023) S38–S52

http://www.elsevier.com/journals/borsa-istanbul-review/2214-8450

Full length article

Big data–enabled sign prediction for Borsa Istanbul intraday equity prices

Abdurrahman Kılıç [a,*], Bülent Güloğlu [a], Atakan Yalçın [b], Alp Üstündağ [a]

[a] Istanbul Technical University, Turkiye

[b] Ozyegin University, Turkiye

Received 4 May 2023; revised 13 August 2023; accepted 14 August 2023

Available online 21 August 2023

Abstract

This paper employs a big data source, the Borsa Istanbul's "data analytics" information, to predict 5-min up, down, and steady signs drawn from closing price changes. Seven machine learning algorithms are compared with 2018 data for the entire year. Success levels for each method are reported for 26 liquid stocks in terms of macro-averaged F-measures. For the 5-min lagged data, nine equities are found to be statistically predictable. For lagged data over longer periods, equities remain predictable, decreasing gradually to zero as the markets absorb the data over time. Furthermore, economic gains for the nine equities are analyzed with algorithms where short selling is allowed or not allowed depending on these predictions. Four equities are found to yield more economic gains via machine learning–supported trading strategies than the equities' own price performances. Under the "efficient market hypothesis," the results imply a lack of "semistrong-form efficiency."

JEL classification: G14; G17; G4

Keywords: Borsa Istanbul; Data analytics; Intraday; Machine learning; Market efficiency; Sign prediction

1. Introduction

Predicting stock market returns is a key objective of market participants and academics alike. Market participants, such as traders, investors, and market makers, use technical and fundamental analyses to make predictions, and they build buy-or-sell strategies based on those estimates. Technical analysis uses past price levels and volume information to determine trends and momentum, while fundamental analysis focuses on macroeconomic and firm-specific financial data (Gunduz, Yaslan, & Cataltepe, 2017). On the other hand, academics in finance, and economics have been suspicious about the predictability of stock market returns. Despite contradictory theories and empirical results, recent studies have shown that the sign predictability of stock price returns is possible and economically viable from an investor's perspective (Chronopoulos, Papadimitriou, & Vlastakis, 2018). In addition to financial economics and time-series econometrics, computational finance has focused on using big financial data, especially intraday data, to predict stock price directions through machine learning and neural network algorithms (Gunduz et al., 2017).

Three objectives in this study aim to contribute to the literature. The first objective is understanding whether the "data analytics" commercially shared by the Borsa Istanbul can provide predictive information for sign prediction. To achieve this, the entire "data analytics" set is used as a predictive feature set for the first time in the literature.

The second objective is finding machine learning algorithms that work best for 5-min sign predictions for the Borsa Istanbul stock market using big data. In this regard, comparisons of seven different machine learning algorithms are reported, with the macro-averaged F-measure (MA F-measure) chosen as the comparison metric.

The third objective is to create a basis for further economic gain analysis by future researchers. There are few studies in

* Corresponding author. Graduate School, Maçka 34367, Istanbul, Turkiye.
E-mail address: kilicab16@itu.edu.tr (A. Kılıç).

Peer review under responsibility of Borsa İstanbul Anonim Şirketi.

https://doi.org/10.1016/j.bir.2023.08.005

this field, and the previous ones only consider "up" and "down" signs as two predicted classes to create a buy or sell strategy. Since researchers such as Chronopoulos et al. (2018) aim to predict daily price changes, they include daily riskless interest-earning financial tools in their benchmarking methods. However, for 5-min time intervals, this study suggests building a more viable buy, sell, or hold strategy by dividing predicted class labels into three categories, namely "up," "down," and "steady." These class labels are then used as trading strategy signals to assess results for financial earnings and to discuss the existence of market efficiency.

Within this framework, the next section summarizes previous work in financial economics, computational finance, and computer engineering. The complexity of the Borsa Istanbul's original big raw data and the method of extracting useful data from it are described in the third section. Dimension-reduction methods for the feature set, the machine learning algorithms compared, and the comparison methodology are explained in the fourth section. The empirical results and statistical performance of the relative successes of machine learning algorithms are reported in the fifth section. Depending on empirical results, economic gain analysis is described in the sixth section. Lastly, the interpretation of results in terms of the mentioned objectives and the importance of this study for further analysis are elaborated in the conclusions section.

2. Literature review

2.1. Predictability of stock market data

The "efficient market hypothesis" embodied by Fama (1970) suggests that if important data are available to everyone at the same time, they are fully reflected in market prices, and nobody can beat the market, which means one cannot consistently outperform the market. He proposes three forms of efficient markets. "Weak form efficiency" assumes that security prices reflect publicly accessible market information but not information that is not yet publicly accessible. This implies that historical price, volume, and return information cannot be used to estimate future prices. This form of the hypothesis suggests a "random walk" for security prices, which is tested via linear and nonlinear unit root tests or by detecting market anomalies (Shiller & Perron, 1985). "Semi-strong form efficiency" adds to weak form efficiency by assuming that prices adjust quickly to new publicly accessible information. Thus, it rejects the benefits of technical and fundamental analysis. Lastly, "strong form efficiency" assumes that private data are already reflected in the current price of a security (Fama E. , 1970).

On the other hand, the literature on whether stock market returns are predictable contains numerous empirical studies with various financial and macroeconomic variables. Rozeff (1984) revealed that stock market returns are not a random walk, and dividend yields can be used to predict future earnings. Later, Pontiff et al. (1998) showed that in addition to the dividend yield and interest yield spread, the book-to-market ratio contains predictive information about future returns. To explain stock market returns, macroeconomic indicators such as consumption and aggregate wealth are added to the literature (Lettau & Ludvigson, 2001).

Torun and Kurt (2008) investigated the existence of weak and semistrong form efficiency on stock exchanges in European Monetary Union countries with panel data variables such as stock market price indexes and macroeconomic variables. Panel unit root tests showed weak form efficiency. However, panel cointegration and causality analysis indicated that some stock markets are not semistrong-form efficient. Jordan, Vivian, and Wohar (2014) suggested that macroeconomic and technical variables can enhance prediction accuracy and produce economic gains for investors. Their predictions for European markets are more powerful than those for US results. Also, the magnitude of the forecasting gains for European markets is comparably better than those for the United States and other G7 countries. Their results imply that market development is related to the forecast performance of macro variables. There is also some evidence that forecast performance is related to liquidity and market size. Khan and Ikram (2010) have shown that both Indian stock markets, the National Stock Exchange, and the Bombay Stock Exchange, have a significant correlation with monthly average foreign institutional investment, and regression analysis suggests an impact from these investments on Indian markets, implying semistrong efficiency.

Tsay (2010) elaborated on the research methods by writing a comprehensive book on the analysis of financial time series. His work provided a systematic introduction to recent financial econometric methods and examples of modeling and predicting financial time series data. Linear and nonlinear models, high-frequency data analysis, multivariate analysis, Markov chain, and Monte Carlo methods are all described in the book. He also illustrates real financial data applications for the models depicted in the book.

Later, Shang (2017) suggested using the functional principal component analysis (PCA) method to estimate 5-min earnings on the S&P 500 index and achieved a successful estimation result. He also reported sign predictions depending on index level predictions. He showed that the accuracy rates for functional PCA were around ninety percent for 5-min intervals during the later trading hours of the examined days.

On the other hand, researchers such as Bossaerts and Hillion (1999) and Goyal and Welch (2003) have shown that it is difficult to predict return levels, especially for US markets.

Some literature exists on the predictability and market efficiency of Turkish stock markets. Kılıç and Buğan (2016) have summarized previous analyses regarding Borsa Istanbul indexes, where 9 of 16 studies covering various periods between 1988 and 2012 reject the existence of weak-form-efficient markets. They also found that two different nonlinear unit root tests yielded different results for the BIST 100, 30, and 50 indexes for the 2003–2015, where one test rejected the existence of weak-form market efficiency. Coşkun and Seven (2016) used unit root tests for the BIST 100 index for the 1993–2015 period, including a structural break in 2003, and found no unit roots, which indicates a lack of weak-form market efficiency.

Akyildirim, Sensoy, Gulay, Corbet, and Salari (2021) revealed that order imbalance data for the Borsa Istanbul are useful in predicting time-series and cross-sectional intraday excess future returns against the benchmark market index. The study depends on the same data analytics as this article, whereas their dataset only focuses on imbalances between buy and sell orders, not the entire big data analytics set. They focus on 1-min interval predictions covering data for the last 5 months of 2016. Their predictive findings imply market inefficiency for BIST 30 stocks.

2.2. Sign predictability of stock market data

Despite contradictory results in the literature regarding whether stock return levels are predictable, researchers such as Breen, Glosten, and Jagannathan (1989) pioneered a new branch of research to predict the sign of returns. Pesaran and Timmermann (1995) examined the robustness and economic significance of stock return predictability. Later, Pesaran and Timmermann (2000) illustrated an extended and generalized version of their recursive modeling method to predict stock returns in the United Kingdom. Furthermore, Christoffersen and Diebold (2006) showed that the direction of returns can be estimated even when the level of means cannot. Additionally, these predictions are not dependent on the distribution of returns and can be used for investment strategies with economic gains.

In an interesting study, Chronopoulos et al. (2018) pointed out the effect of information demand on stock return predictability. They approximated information demand using the daily internet search volume index from Google. The researchers suggested that incorporating an information demand variable in various GARCH family models significantly improves volatility forecasts. Based on the theory that these volatility predictions can be used to predict the direction of stock market prices, they found that the sign of stock returns is predictable, contrary to the return levels. Additionally, they illustrated the economic value of sign predictability and revealed that investors could form profitable investment strategies using information demand. Therefore, in this study, we focus on sign predictability and its economic significance rather than price level estimations.

2.3. Machine learning algorithms in stock return predictions

In addition to econometric regression methods, machine learning algorithms have lately been employed to predict stock return performance. Support vector machines (SVMs) (Cortes & Vapnik, 1995) are used by various researchers (Sapankevych & Sankar, 2009) for different purposes, such as forecasting the Belgrade Stock Exchange index (Markovi'c, Stojanovi'c, Stankovi'c, & Boži'c, 2014) or finding the optimal subset of features to predict stock price trends on the Istanbul Stock Exchange (Pehlivanlı, Aşıkgil, & Gülay, 2016).

Ballings, Poel, Hespeels, and Gryp (2015) compared the sign prediction performance of various methods, including random forest, ada boost, kernel factory, neurol networks, logistic regression, SVMs, and K-nearest neighbors (KNN), using European stock market data. They used the area under the curve (AUC) and receiver operating characteristic (ROC) as performance indicators. They found that random forest was the best-performing method, with SVM coming in second. Bastı, Kuzey, and Delen (2015) analyzed the performance of initial public offerings on the Istanbul Stock Exchange using SVMs and decision trees. Gunduz and Cataltepe (2015) added text-mining methods for financial news in Turkish to their naïve bias model to predict the daily direction of the BIST 100 index and determine a feature selection method. In this study, and likewise in this paper, only significant direction changes were accepted as positive or negative, with the rest accepted as steady. As three signs were used, and due to the unbalanced dataset, the MA F-measure was used to compare feature selection methods.

A smaller branch of literature uses neural network algorithms to make sign predictions in stock markets, including the Borsa Istanbul (C.S. Vui, 2013, pp. 477–482; Persio & Honchar, 2016; Gunduz et al., 2017). While intraday studies, such as the hourly analysis of Gunduz et al. (2017), did not yield strong prediction estimates, in a recent study by Aksoy (2021), balance sheet analysis with macroeconomic variables for quarterly price direction predictions, depending on machine learning algorithms and artificial neural network methods, resulted in more than ninety percent accuracy.

Finally, in another study, Akyildirim et al. (2021) employed machine learning methods to assess the excess future returns of single stocks compared with the BIST-30 index. In this study, they focused the order imbalances in terms of number and quantity and 1-min interval performances. They made sign predictions and compared them via ROC curves, calculated the (average) AUC measure of classifier performance, and employed the k-fold cross-validation procedure.

In this regard, this study also investigates single stock-level market efficiency in intraday trading with comprehensive machine learning methods for the year 2018. It differs from previous literature, however, by employing all the information provided by the "data analytics" set of the Borsa Istanbul, by benchmarking prediction-based trading strategies to the price performance of single stocks, and by allowing three-dimensional confusion matrices to assess "up," "down," and "stable" signs.

3. Data

In this study, intraday 5-min stock market data from Borsa Istanbul for all of 2018 are employed. The data cover 26 stocks included in the BIST 30 index for 2018, except for VAKBN, as its raw data for independent variables could not be parsed.

The dependent variables are derived from the 5-min closing prices of the equities by differentiating from the previous work of Akyildirim et al. (2021) to allow greater price changes that can cover trading costs such as commissions and price tick levels to achieve a realistic economic gain analysis. None of the stock price levels were found to be stationary, but differences in prices were stationary, as shown in Appendix A. The proposed ARIMA models are listed in Appendix A for those curious about univariate price level analysis. As Christoffersen, Diebold, Mariano, Tay, and Tse (2007) mentioned, even if the distribution is asymmetric and the expected return is zero, the sign can be predictable. The price returns are labeled into three classes to achieve good sign prediction. It is assumed that the commission for the brokerage house, plus other costs such as price tick levels, is approximately 0.025 percent. By multiplying this by 6, if the price increases more than 0.15 percent (i.e., 15 pips) in 5-min intervals, it is accepted that the sign is up (i.e., positive). If the price decreases by the same percentage, it is accepted as down (i.e., negative), and the price is steady for the in-between percentages. Unlike previous work related to intraday sign prediction of equities in the Borsa Istanbul, three sign classes are determined to make the research more consistent with real practices in the trading industry and make the results more useful for determining buy, sell, or hold strategies in assessing economic significance studies. Fig. 1 describes the labeling algorithm, where P_t represents the 5-min closing price for an equity taken from the data.

The features used to predict stock return signs are all collected from the "data analytics" information disseminated by the Borsa Istanbul through data vendors. The data analytics are calculated and updated for each BIST 100 index stock every second. There are 39 analytics related to the price, number, and quantity of buy or sell orders and canceled orders and some combining ratios regarding the last 1- and 5-min data from the beginning of the day. The same statistics are disseminated by the Deutsche Börse for German stocks. Appendix B includes the whole list of data analytics statistics and their explanations.

The raw data consist of 251 daily files totaling nearly 58 gigabytes and are highly complicated. By assigning Python software codes, the data for only the 26 stocks examined in the study were extracted. As the study focused on 5-min intervals, data for the last seconds of the 5-min intervals for each of the 39 analytics were picked up. However, there was no data disseminated for the second that does not include any change in an analytic from the previous second. Therefore, if there is no data for the last second of the targeted 5-min interval for a specific feature, we picked up the data by searching the last 30 s for the 1-min interval analytics, searching the last 150 s for the 5-min interval analytics, and searching until the beginning of the day for daily cumulative analytics.

There are 84 5-min trading intervals for a total of 7 h each day on the Borsa Istanbul, and there is a closing session. Data for the end of the first 5-min interval are used to predict the sign of the price change for the second 5-min interval, data for the end of the first session are used to predict the sign of the price change for the first 5-min interval of the second session, and the data for the end of the last interval are used to predict the sign of the price change for the closing session. Thus, there are 84 predictions each day.

There are 251 trading days, with some including only one session due to holidays. Additionally, stocks such as ASELS are closed to continuous trading on some days. Within this framework, at most, 20,889 data points are included in the data for each stock.

Test data are chosen as the last 5 percent of the data without randomization. Thus, the training data and test data refer to the first 95 percent and last 5 percent of the 5-min trading periods of the whole year, respectively.

4. Methodology

4.1. Standardization and dimension reduction

Data standardization methods enhance the symmetry of data and bring the analyzed data closer to a normal distribution. They reduce skewness and scale raw data into a narrower range to properly study relationships across different data series (Sree, Bindu, & Shoba, 1998). In this sense, the explanatory variables set is scaled using the standard scaler in the scikit-learn library of the Python software program. The standard scaler adjusts variables by setting their means to zero and standard deviations to one, as shown in Eq. (1):

$$z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

where μ is the mean of the elements of feature vector x, σ is the standard deviation for the elements of x, and z_i is the scaled value for the element x_i.

Since the feature-based model is high-dimensional, a common obstacle for time-series classification models is the substantial effort required to obtain essential attributes from the data. Determining only relevant features can take a long time, and it is necessary to use domain expertise to eliminate the irrelevant ones (Amjady & Daraeepour, 2009). However, the feature set in this study provided by Borsa Istanbul is newly used in the literature. Therefore, to cope with complexity, PCA is chosen as a rough-cut eigenvalue-based dimension-reduction model. Feature reduction in the dependent variables can provide some advantages, such as reducing noise and enhancing classification performance (Seijo-Pardo, Porto-Díaz, Bolón-Canedo, & Alonso-Betanzos, 2017).

The number of dimensions for explanatory variables is reduced to four, five, or six for the stocks via PCA to ensure

$$Y_t = \begin{cases} \text{Positive} & \text{if } (P_t - P_{t-1}) / P_{t-1} > 0.0015 \\ \text{Steady} & \text{if } -0.0015 < (P_t - P_{t-1}) / P_{t-1} < 0.0015 \\ \text{Negative} & \text{if } (P_t - P_{t-1}) / P_{t-1} < 0.0015 \end{cases}$$

Fig. 1. Labeling of dependent variables.

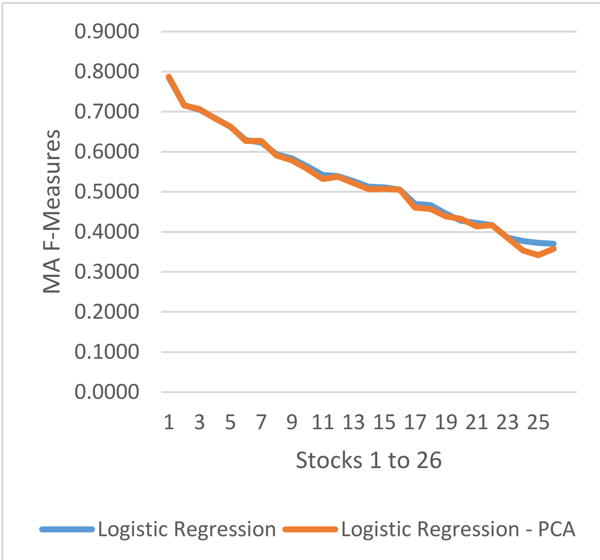

Fig. 2. Depiction of classification performances: LR vs. LR with PCA.

that those dimensions represent at least 60% of the total volatility of the 39 analytics. Since PCA projects high-dimensional data onto a reduced dimension by combining the original explanatory variables, the analysis results obtained from the generated principal components cannot be interpreted (Dunteman, 1989).

Reducing the dimensions of explanatory variables via PCA can cause some information loss (Björklund, 2019), but in this study, it causes a negligible effect on results. Following the example of Gunduz et al. (2017), to assess analytical performance with and without PCA, we trained a linear regression classifier for each of the 26 stocks' sign predictions with and without PCA. Figs. 2 and 3 depict and compare the MA F-measure results for the LR classification. LR classification with PCA has an average MA F-measure of 0.527 for 26 stocks,

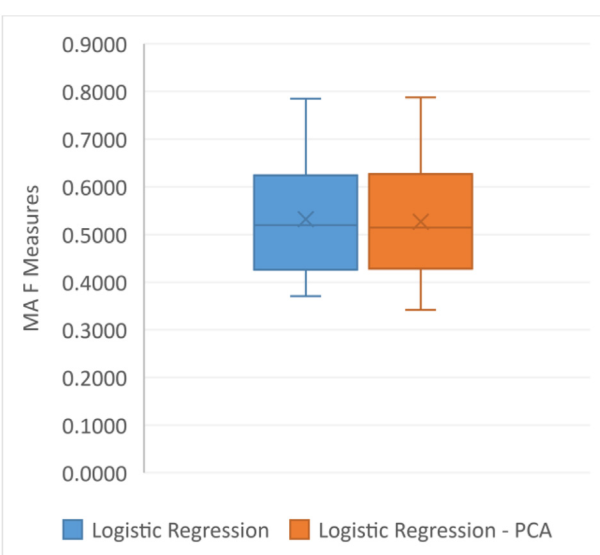

Fig. 3. Comparison of classification performances: LR vs. LR with PCA.

whereas the LR without PCA is 0.532. A one-tailed T-test comparison of the two averages results in no statistically significant difference. Therefore, we conducted further analysis with PCA dimension reduction.

4.2. Comparison of machine learning algorithms

Nine different machine learning algorithms were initially employed for sign prediction in this study: logistic regression, SVMs with linear, radial basis function (i.e., rbf), sigmoid basis kernels, naïve Bayes, decision tree, and random forest with 50 and 500 trees. SVMs with different kernel bases were used to examine the performance differentiation between SVM models. Additionally, random forest algorithms with different numbers of trees were included in the study.

For each stock, 5-min closing price data are labeled "positive," "steady," or "negative" as described in the "Data" section, and the number of explanatory variables is reduced via PCA.

Classifiers for each methodology are trained with the first 95 percent of the data. Their ability to predict each sign is tested using the last 5 percent of the data for the year 2018. To compare the resulting confusion matrices, F-measures are calculated for each class.

In this study, we have three classes instead of two to introduce a better economic gain analysis depending on trading strategies that are more relevant to algo-trading practices in the market. Thus, our study uses three-by-three confusion matrices, as illustrated in Table 1.

$$PrecisionA = T_{AA} \Big/ (T_{AA+FAB+FAC}) \tag{2}$$

$$RecallA = T_{AA} \Big/ (T_{AA+FBA+FCA}) \tag{3}$$

$$F-MeasureA = \frac{2*PrecisionA*RecallA}{PrecisionA + RecallA} \tag{4}$$

Likewise, in Eqs. (2)–(4), F-measures are calculated for each class, and the three F-measures for the three classes are simply averaged to obtain MA F-measures for comparison. On the other hand, micro-averaged F-measures (i.e., MiA F-measures) yield the overall accuracy, and weighted-averaged F-measures can be calculated by weighting F-measures for each class proportional to their actual presence rates within the observed samples.

Since an equal importance weight is assessed to the prediction results of each class, the resulting confusion matrices

Table 1
Three-by-three confusion matrix.

	Predicted			
Actual		ClassA	ClassB	ClassC
	ClassA	T_{AA}	F_{BA}	F_{CA}
	ClassB	F_{AB}	T_{BB}	E_{CA}
	ClassC	F_{AC}	F_{BC}	T_{CC}

-T and F stand for true and false predictions, respectively. Predicted and actual classes are represented by the first and second subscripts in order.

are used to determine MA F-measures by following the inspiring footsteps of Gunduz et al. (2017) instead of micro- or weighted-averaged F-measures.

To determine the statistical significance of the F-measure results of sign predictions, their standard deviations must first be known, and a theoretical benchmark for comparison is needed. With those tools, it is possible to determine whether the deployed machine learning algorithms yield predictive estimators.

In this sense, the article “Confidence interval for micro-averaged F1 and macro-averaged F1 scores” (Takahashi, Yamamoto, Kuchiba, & Koyama, 2022) provided an essential method to assess the significance of sign prediction methods. Since Takashi et al. use only confusion matrices to assess the standard deviations of MiA and MA F-measures, their method is particularly useful to determine whether an F-measure obtained from a machine learning algorithm is statistically superior to others.

As the MiA F-measure yields accuracy for a given test and training data, it is better to benchmark MiA F-measures with a dummy strategy that predicts every data point as the most common sign, “steady,” in the test data. This strategy yields a higher MiA F-measure for unbalanced data, as in our case. None of the machine learning algorithms for any equity can perform significantly better in MiA F-measure than this default accuracy-enhancing strategy of predicting all the signs as 0 by default. That is why the results for the MiA F-measures are not reported in the next section.

For the MA F-measure, the default benchmark strategy has to yield good performance for all signs. Therefore, the benchmark estimator for the MA F-measure is the random sign generator that predicts each data point according to the proportions of each sign in the training data.

For the confusion matrix of this random prediction method, from the same notation in Takashi et al. (2022), it can be inferred that.

r is the number of classes.

i is the indicator of a predicted class, and j stands for the actual class.

p_{ij} is the probability of i'th column and j'th row of confusion matrix.

$Ptrain_i$ is the proportion of the i'th class in the training data, and $Ptest_j$ is the proportion of the j'th class in the test data:

$$p_{ij} = Ptrain_i \text{ x } Ptest_j \; (i,j = 1 \ldots\ldots r) \tag{5}$$

As we assume there a multinomial distribution for the data, the expected values of cells in the confusion matrix for the random sign generator are

$$E[n_{ij}] = n_{test} x \, p_{ij}$$

where n_{test} stands for number of observations in the test data.

After obtaining MA F-measures for each equity using the aforementioned machine learning algorithms, their performances were statistically compared using their 95% confidence intervals and the random sign generator.

Reference MA F-measures from the random sign generator were calculated using p_ij obtained from Eq. (5). They were then embedded into Eq. (4) for each class to obtain F-measures of each class and simply averaged to find the reference MA F-measure of the benchmark strategy. Theoretically, the reference MA F-measure is 1/r if the proportions of incidences are the same for the training and test data for each class. It will be lower than that, as training, and test data differ. This method yielded MA F-measures less than one-third for the three-class model, since the proportions of classes are not the same in the test and training data. The MA F-measure scores of this random predictor are reported as a base level for benchmarking the MA F-measures obtained from the machine learning algorithms.

5. Results for sign predictability

The MA F-measure results for the 26 stock price direction predictions using 5-min lagged data are presented in Table 2 along with their standard deviations and significance relative to the reference levels.

The KNN algorithm yields the best performance for 11 stocks, while the decision tree algorithm outperforms the others for 10 of the 26 stocks. Additionally, random forest with 50 trees provides the best estimation performance for 4 stocks, and naïve Bayes does so for 1.

With 5-min-lagged data, nine of the stocks can be predicted with an MA F-measure that is significantly above the random benchmark predictor. AKBNK, GARAN, and THYAO can be predicted by decision tree and EREGL and KOZAL by KNN. The price direction of HALKB, KOZAA, KRDMD, and TKFEN can be predicted by multiple methods.

In total, 9 of 26 equities are predictable via machine learning algorithms, meaning the results are statistically significantly better than the random benchmark estimator. This result implies that the information provided by Borsa İstanbul can be used to predict equity price changes with a 5-min-lagged data, which could also provide evidence for the abundance of the “semistrong form of market efficiency” for those nine equities if it can be justified financially.

Compared with the random benchmark strategy's results, the significantly better MA F-measures of the machine learning algorithms are 9–28 percent higher.

Note that the MA F-measures without standard deviations do not differ much from the benchmarks, as their confusion matrix diagonals are zero, which implies no true prediction for some classes. In this sense, it can be observed from the tables that the existence of MA F-measure results without standard deviations did not affect the results.

The existence of market efficiency is also tested for 10-min-lagged data to understand how long it takes for the market to exploit this data or how long market inefficiency lasts. In this regard, the number of predictions for each day is reduced from 84 to 83 for the 10-min-lagged analysis. The number of determined dimensions for the PCAs for each equity did not change.

The MA F-measure results for the nine stock price direction predictions using 10-min-lagged data are presented in Table 3, along with their standard deviations and significance relative to reference levels.

Table 2
Macro-averaged F-measure performance of classification methods using 5-min-lagged data.

Equity	Logistic Reg.	KNN	SVM - rbf	SVM - sigmoid	Naive Bayes	Decision Three	Random Forest (50 trees)	Ref. Ma-F
AKBNK	0.2749	0.3548	0.2823	0.3621	0.2894	0.3701 *	0.3345	0.3322
		0.0151	0.0085			0.0154	0.0148	
ARCLK	0.3125	0.3237	0.2872	0.2775	0.3182	0.3243	0.3077	0.3297
	0.0117	0.0123				0.0141	0.0110	
ASELS	0.2745	0.3216	0.2676	0.2619	0.2898	0.3317	0.3116	0.3327
		0.0147		0.0122	0.0105	0.0149	0.0138	
BIMAS	0.2827	0.3484	0.2833	0.3098	0.3131	0.3343	0.2965	0.3332
		0.0156		0.0141	0.0125	0.0147	0.0098	
DOHOL	0.2354	0.3127	0.2383	0.2905	0.2426	0.3456	0.2914	0.3332
		0.0138			0.0067	0.0151	0.0127	
EKGYO	0.2427	0.3090	0.2425	0.3380	0.2668	0.3442	0.2778	0.3327
		0.0135		0.0146	0.0094	0.0150	0.0113	
EREGL	0.2955	0.3540 *	0.2950	0.1608	0.3398	0.3072	0.3442	0.3144
		0.0164		0.0113		0.0129	0.0160	
GARAN	0.2920	0.3218	0.2920	0.3138	0.3075	0.3631 *	0.3079	0.3224
		0.0133		0.0138	0.0104	0.0158	0.0104	
HALKB	0.3077	0.3812 *	0.2974	0.3364	0.3208	0.3685	0.3788	0.3286
	0.0115	0.0170		0.0145	0.0134	0.0158	0.0169	
ISCTR	0.2380	0.3259	0.2257	0.3074	0.2469	0.3140	0.3330	0.3331
	0.0082	0.0145	0.0060	0.0135	0.0092	0.0143	0.0148	
KCHOL	0.3027	0.3335	0.2967	0.2652	0.3123	0.3314	0.3071	0.3276
	0.0139	0.0139				0.0142	0.0090	
KOZAA	0.2933	0.3400 *	0.2937	0.3843	0.3463	0.3279	0.3492	0.2997
	0.0116	0.0144	0.0120	0.0158	0.0156	0.0146	0.0155	
KOZAL	0.2617	0.3588 *	0.2691	0.2712	0.2850	0.3305	0.3433	0.3280
		0.0155		0.0140	0.0097	0.0146	0.0151	
KRDMD	0.2931	0.3530 *	0.2808	0.3312	0.3433	0.3317	0.3455	0.3064
	0.0096	0.0148			0.0148	0.0148	0.0155	
PETKM	0.2611	0.3253	0.2681	0.3198	0.2881	0.3020	0.3371	0.3323
	0.0075	0.0143	0.0087	0.0118	0.0115	0.0140	0.0145	
PGSUS	0.2737	0.3341	0.2670	0.2329	0.3055	0.3340	0.3237	0.3295
		0.0146	0.0062	0.0125	0.0122	0.0147	0.0143	
SAHOL	0.3163	0.3376	0.3069	0.3165	0.3309	0.3354	0.3276	0.3247
		0.0147		0.0121		0.0147	0.0134	
SISE	0.2982	0.3348	0.3082	0.3126	0.3039	0.3229	0.3518	0.3320
	0.0132	0.0147	0.0138	0.0135	0.0138	0.0145	0.0150	
TAVHL	0.2622	0.3343	0.2389	0.2862	0.2667	0.3236	0.3213	0.3329
	0.0096	0.0143			0.0140	0.0145	0.0141	
TCELL	0.2608	0.3571	0.2520	0.3078	0.2663	0.3365	0.3100	0.3332
	0.0081	0.0147			0.0092	0.0146	0.0135	
THYAO	0.2854	0.3592	0.2894	0.2640	0.3049	0.3640 *	0.3250	0.3296
		0.0159	0.0082			0.0155	0.0139	
TKFEN	0.3028	0.3246	0.2858	0.2106	0.3462	0.3676 *	0.3879	0.3319
	0.0128	0.0147	0.0111		0.0150	0.0150	0.0157	
TOASO	0.2656	0.3300	0.2703	0.2434	0.3170	0.3440	0.3254	0.3322
		0.0147		0.0135	0.0131	0.0149	0.0143	
TTKOM	0.2416	0.2883	0.2516	0.3164	0.2646	0.3446	0.3000	0.3314
		0.0135	0.0096	0.0147	0.0109	0.0149	0.0136	
TUPRS	0.2710	0.3416	0.2711	0.2922	0.2962	0.3382	0.3185	0.3321
		0.0143			0.0104	0.0146	0.0135	
YKBNK	0.2514	0.3081	0.2449	0.2352	0.2554	0.3445	0.2794	0.3329
	0.0061	0.0133				0.0149	0.0114	

- The numbers in the first row of each cell represent the MA F-measures of the relevant methods for the corresponding equities.
- The numbers in the second row of each cell represent the standard deviations for the MA F-measures.
- MA F-measures that are higher than the reference levels, within their 95% confidence interval, are marked with an asterisk (*).
- Equities that have sign predictions with MA F-measures having 95% confidence intervals higher than the Reference MA F-measures are bolded.
- Since the standard deviation formula of Takashi et al. requires nonzero diagonals, confusion matrices with zero diagonals provided no standard deviations; therefore, their places are left empty.

Table 3
Macro-averaged F-measure performance of classification methods using 10-min-lagged data.

Equity	Logistic Reg.	KNN	SVM - rbf	SVM - sigmoid	Naive Bayes	Decision Three	Random Forest (50 trees)	Ref. Ma-F
AKBNK	0.2667	0.3564	0.3564	0.3552	0.2807	0.3581	0.3077	0.3321
		0.0147		0.0125	0.0085	0.0151	0.0122	
EREGL	0.2956	0.3568 *	0.3007	0.1717	0.3239	0.3096	0.3288	0.3147
		0.0172		0.0118	0.0125	0.0134	0.0134	
GARAN	0.2926	0.3411	0.2986	0.3171	0.3055	0.3420	0.3104	0.3224
		0.0155		0.0118		0.0150	0.0108	
HALKB	0.3051	0.3445	0.2987	0.2722	0.3186	0.3850 *	0.3459	0.3287
	0.0109	0.0147				0.0164	0.0150	
KOZAA	0.2951	0.3290 *	0.3026	0.2919	0.2980	0.3243	0.3191	0.2994
	0.0111	0.0142	0.0128		0.0156	0.0144	0.0139	
KOZAL	0.2661	0.3235	0.2771	0.2661	0.2761	0.3392	0.3495	0.3280
		0.0140		0.0140	0.0081	0.0149	0.0155	
KRDMD	0.2864	0.3162	0.2826	0.1674	0.3170	0.3353 *	0.3294	0.3063
	0.0081	0.0127	0.0070	0.0112	0.0129	0.0147	0.0144	
THYAO	0.2767	0.3338	0.2767	0.3408	0.2851	0.3533	0.3343	0.3294
		0.0144		0.0137		0.0152	0.0147	
TKFEN	0.2682	0.3598	0.2575	0.3081	0.3114	0.3578	0.3443	0.3319
	0.0092	0.0152	0.0111	0.0136	0.0136	0.0150	0.0147	

- The numbers in the first row of each cell represent the MA F-measures of the relevant methods for the corresponding equities.
- The numbers in the second row of each cell represent the standard deviations for the MA F-measures.
- MA F-measures that are higher than the reference levels, within their 95% confidence interval, are marked with an asterisk (*).
- Equities that have sign predictions with MA F-measures having 95% confidence intervals higher than the Reference MA F-measures are bolded.
- Since the standard deviation formula of Takashi et al. requires nonzero diagonals, confusion matrices with zero diagonals provided no standard deviations; therefore, their places are left empty.

Only four equity markets remain predictable with 10-min-lagged data. As all the tests for other equities yield insignificant results compared with the benchmark, for simplicity, they are not reported separately.

All the equities were also tested with 15- and 30-min-lagged data, and the number of equities that remained predictable decreased. For simplicity, we share the results for the 1-h-lagged data, where no equity was found to be predictable. To illustrate, only the four equities that remained predictable with 10-min-lagged data are reported in Table 4, with the results for the 1-h-lagged data summarized below.

We find a significant sign prediction for nine equities with 5-min-lagged data, but this number reduces to four equities with 10-min-lagged data and 0 when the time lag reaches 1 h. The results in Tables 2–4 suggest that the predictive power of machine learning algorithms disappears over time as the market gradually absorbs the data.

6. Economic gain analysis

6.1. Framework of economic gain analysis

Having prediction methods that yield statistically significant predictions does not always mean a profitable investment strategy (Diebold & Lopez, 1996). Therefore, in this section, we deploy our sign forecasts to the tested data to investigate economic gains.

As we explore the economic gains of the machine learning method predictions, we assume that the investor's position for each equity is holding it at the beginning, and a passive buy

Table 4
Macro-averaged F-measure performance of classification methods using 1-h-lagged data.

Equity	Logistic Reg.	KNN	SVM - rbf	SVM - sigmoid	Naive Bayes	Decision Three	Random Forest (50 trees)	Ref. Ma-F
EREGL	0.2992	0.3215	0.2985	0.2691	0.3041	0.2887	0.3272	0.3133
		0.0132	0.0125			0.0133	0.0148	
HALKB	0.2831	0.3346	0.2827	0.3334	0.3071	0.3435	0.3622	0.3293
		0.0155		0.0157		0.0159	0.0175	
KOZAA	0.2865	0.3176	0.2969	0.2770	0.2970	0.3217	0.3243	0.2981
	0.0107	0.0150	0.0121	0.0120		0.0154	0.0149	
KRDMD	0.2802	0.3192	0.2744	0.1746	0.3141	0.3331	0.3271	0.3073
		0.0143			0.0129	0.0157	0.0149	

- The numbers in the first row of each cell represent the MA F-measures of the relevant methods for the corresponding equities.
- The numbers in the second row of each cell represent the standard deviations for the MA F-measures.
- MA F-measures that are higher than the reference levels, within their 95% confidence interval, are marked with an asterisk (*).
- Equities that have sign predictions with MA F-measures having 95% confidence intervals higher than the Reference MA F-measures are bolded.
- Since the standard deviation formula of Takashi et al. requires nonzero diagonals, confusion matrices with zero diagonals provided no standard deviations; therefore, their places are left empty.

Table 5
Interpretation of trading strategies depending on predicted signs.

Short Sell Not Allowed			Short Sell Allowed		
SIGN\POSITION	Not Holding	Holding	SIGN\POSITION	Short Sold	Holding
Negative	Keep The Position	Sell	**Negative**	Keep The Position	Short Sell
Stable	Keep The Position	Keep The Position	**Stable**	Keep The Position	Keep The Position
Positive	Buy	Keep The Position	**Positive**	Buy	Keep The Position

and hold (i.e., B&H) strategy that follows the performance of the equity price is compared with strategies using the predicted signs. Since there is no risk-free financial tool for intraday trading, the 5-min-interval decision mechanisms could not be compared with interest-earning benchmarks. To provide a more detailed analysis of financial gains, we specified two trading environments where short sales are allowed and not allowed, similar to Chronopoulos et al. (2018).

Since we have three classes of predictions (positive, stable, and negative), we accepted the positive signal as a buy signal, stable as a keep-the-position signal, and negative as a sell signal. Since we determined the positive or negative move parameters as 15 pips, which is more than enough to cover transaction costs and commissions, we could describe an investor's positions under the different environments as shown in Table 5.

To detect true economic gains, transaction costs should be considered. There are different approaches in the literature to describe transaction costs, such as the including floor trader cost, which was used by Fama and Blume (1966) and later by Bekiros (2010). Another way is to use brokerage commissions and bid–ask spreads (Bhardwaj, 1992). In addition to those costs, short sales costs, and price impacts can also be considered. The literature regarding S&P analysis accepts those costs at approximately 2.4 basis points (bps) (Chronopoulos et al., 2018). Commission fees for the Borsa Istanbul can be reduced by up to 1 bps, but considering other trading costs, we rounded up the one-way transaction cost to 2.5 bps.

6.2. Results for economic gain analysis

Economic gains for the passive buy and hold (i.e., B&H.) strategy and trading strategies depending on machine learning algorithms were found to have statistically significant sign estimates according to their MA F-measures in the previous section, as listed in Table 6.

Among the nine equities that were statistically predictable, basic trading strategies that depended on machine learning algorithms outperformed the passive strategy—which depends on equity price performance only—for just three equities:

Table 6
Economic gain results for machine learning–supported strategies without short selling.

Equity	Parameter	B&H	KNN	SVM sigmoid	Naive Byes	Decision Tree	Random Forest
ABANK	Earning	7.355%				−0.290%	
	Average Return	0.007%				−0.001%	
	Std. Dev.	0.00205				0.00138	
EREGL	Earning	−4.875%	−5.473%				
	Average Return	−0.005%					
	Std. Dev.	0.00203	0.00113				
GARAN	Earning	4.599%				−3.737%	
	Average Return	0.004%				−0.004%	
	Std. Dev.	0.00189				0.00135	
HALKB	Earning	3.846%	−1.489%			0.874%	1.921%
	Average Return	0.003%				0.000%	0.001%
	Std. Dev.	0.00212	0.00120			0.00150	0.00157
KOZAA[a]	Earning	−7.198%	−3.421%	−4.742%	0.917%		5.624%
	Average Return	−0.007%	−0.003%	−0.005%	0.000%		0.004%
	Std. Dev.	0.00308	0.00163	0.00275	0.00253		0.00201
KOZAL[a]	Earning	−3.571%	−2.277%				
	Average Return	−0.003%	−0.002%				
	Std. Dev.	0.00220	0.00114				
KRDMD[a]	Earning	−0.909%	3.636%		−1.889%		10.034%
	Average Return	−0.001%	0.003%		−0.002%		0.009%
	Std. Dev.	0.00333	0.00159		0.00301		0.00239
THYAO	Earning	4.815%				−3.466%	
	Average Return	0.004%				−0.003%	
	Std. Dev.	0.00206				0.00146	
TKFEN	Earning	−0.956%				−8.367%	−1.361%
	Average Return	−0.001%				−0.008%	−0.001%
	Std. Dev.	0.00258				0.00181	0.00195

[a] Equites having superior economic gain results with machine learning–supported predictions compared to passive buy and hold strategy.

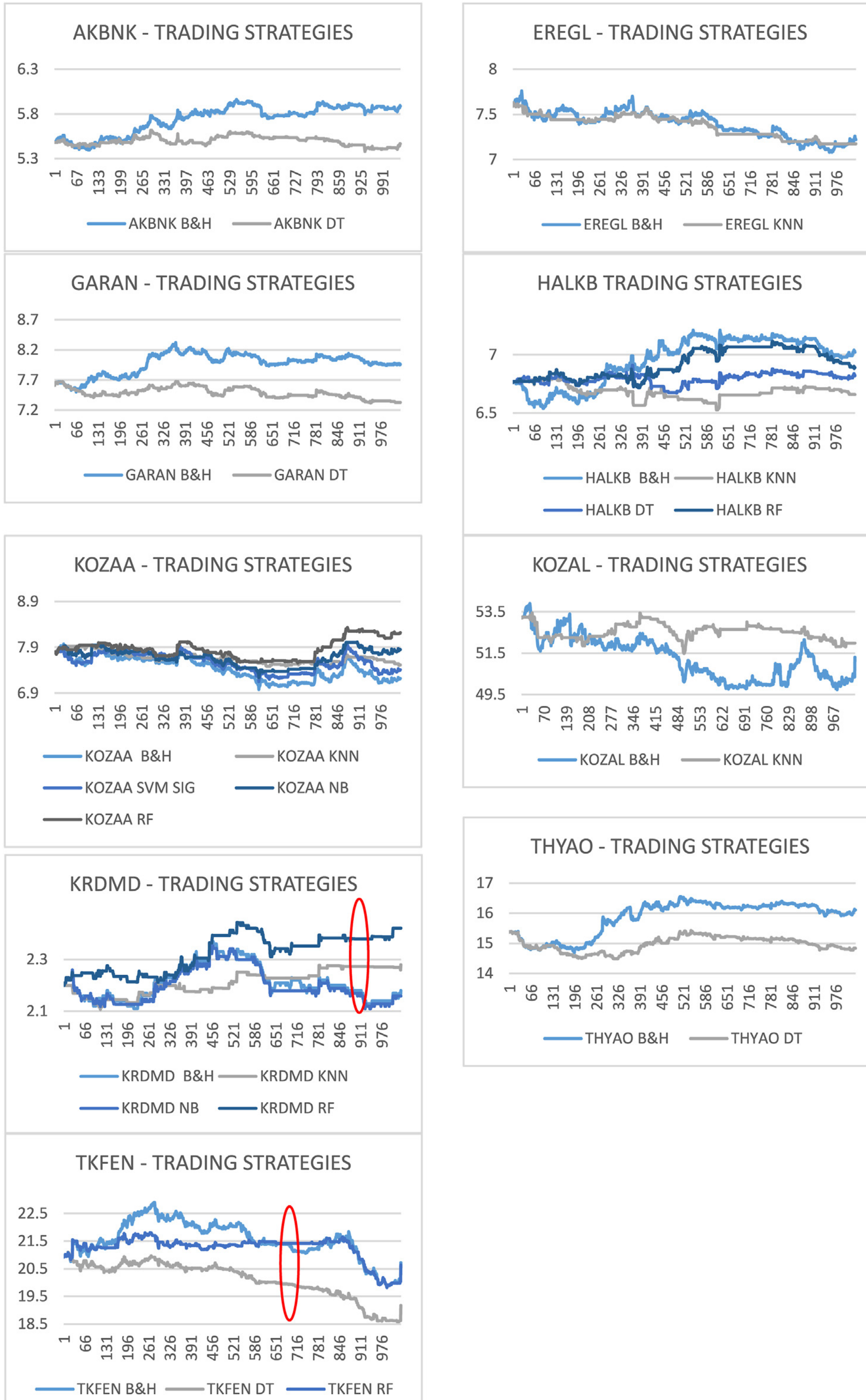

Fig. 4. Graphical illustrations of three trading-signal strategies without short selling.

Table 7
Economic gain results for machine learning–supported strategies—short selling allowed.

		B&H	KNN	SVM sigmoid	Naive Byes	Decision Tree	Random Forest
ABANK	Earning	7.355%				−3.683%	
		0.006%				−0.004%	
	Std. Dev.	0.00205				0.00205	
EREGL	Earning	−4.875%	−5.336%				
	Av. Excess Return						
	Std. Dev.	0.00203	0.00203				
GARAN	Earning	4.599%				−7.815%	
	Av. Excess Return	0.004%					
	Std. Dev.	0.00189				0.00189	
HALKB	Earning	3.846%	−4.207%			2.754%	2.386%
	Av. Excess Return	0.003%				0.002%	0.001%
	Std. Dev.	0.00212	0.00211			0.00211	0.00211
KOZAA[a]	Earning	−7.198%	4.118%	0.859%	11.065%		22.954%
	Av. Excess Return				0.009%		0.019%
	Std. Dev.	0.00308	0.00307	0.00308	0.00307		0.00307
KOZAL[a]	Earning	−3.571%	2.230%				
	Av. Excess Return						
	Std. Dev.	0.00220	0.00219				
KRDMD[a]	Earning	−0.909%	10.304%		−2.397%		24.001%
	Av. Excess Return		0.009%				0.020%
	Std. Dev.	0.00333	0.00333		0.00333		0.00332
THYAO	Earning	4.815%				−6.137%	
	Av. Excess Return	0.004%					
	Std. Dev.	0.00206				0.00206	
TKFEN[a]	Earning	−0.956%				−9.394%	2.296%
	Av. Excess Return						
	Std. Dev.	0.00258				0.00257	0.00258

[a] Equites having superior economic gain results with machine learning–supported predictions compared to passive buy and hold strategy.

KOZAA, KOZAL, and KRDMD. Earnings were calculated as a percentage of gains from the beginning to the end of the test periods. Random forest predictions yielded gains of 5.6% and 10%, while equity prices were reduced by 7.2% and 0.9%, respectively, during the test period for KOZAA and KRDMD. Likewise, KNN predictions achieved only 2.3% losses, as KOZAL lost 3.6% of its value in the same period.

Average returns were calculated as compounded returns of 5-min gains. The standard deviations of 5-min interval average returns are also described in the table. Since we described small changes as "Stable" signals and did not act to buy or sell in the trading strategy supported by machine learning predictions, we had some periods where we sold equities and did not buy them until a "Positive" signal occurred. Therefore, we kept our monetary value stable and had less volatility (i.e., standard deviations) in gains or losses. Some of those periods are indicated by red circles in Fig. 4.

Since the three classes allow for three different signals to determine trading actions, there may be periods when an equity is not held. The economically advantageous results can be explained by sell signals that prevent losses during price-decreasing periods. To address this concern, we created a trading strategy that allows for short selling, where the trader can have a hold or short-sell position for each equity.

When short selling is allowed, trading strategies depending on the machine learning algorithms resulted in better performance than the passive buy and hold strategy for four equities, namely KOZAA, KOZAL, KRDMD, and TKFEN (Table 7). Earnings are calculated as a percentage of gains from the beginning to the end of the test periods. The trading strategy depending on random forest predictions yielded 5.6%, 2.3%, and 10% gains, while equity prices were reduced by 7.2%, 0.9%, and 1%, respectively, during the test period for KOZAA, KRDMD, and TKFEN. Additionally, KNN predictions achieved 2.2% gains, as KOZAL's price was reduced by 3.6% in the same period. Notably, three of those four equities are the same for the trading algorithms not allowing short selling.

Since we allow only hold or short positions, the standard deviations of gains of the buy-and-hold and machine learning–supported trading strategies are very close. For detailed illustration, some holding, and short-sell periods are marked with blue and red circles, respectively, in Fig. 5.

The results indicate that real economic gains can be created using publicly available data and machine learning algorithms while including transaction costs for 4 of the 26 equities analyzed. From the point of view of market theory, those achievements imply a lack of semistrong market efficiency for KOZAA, KOZAL, KRDMD, and TKFEN.

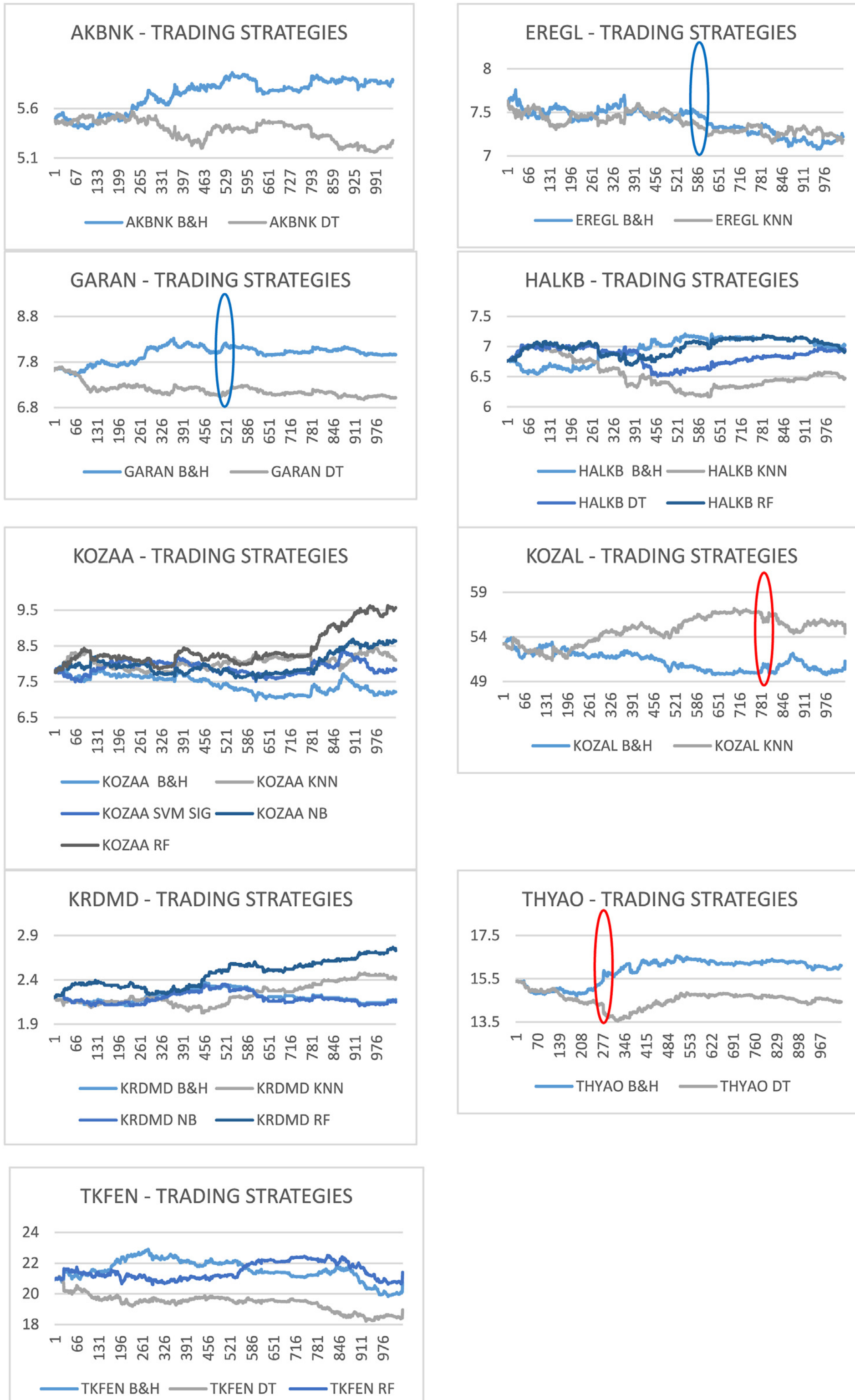

Fig. 5. Graphical illustrations of three signal trading strategies with short selling.

7. Conclusions

In this paper, the directions of 5-min-interval closing price movements of 26 highly liquid stocks are estimated via various machine learning algorithms. Unlike previous work in the literature, the direction of change is determined as three classes: up, down, and steady. By differentiating from the previous literature, the entire Borsa Istanbul's “data analytics” set is included in the feature set to understand the estimation performance of these data currently disseminated to real investors. All performance comparisons are made using MA F-measures obtained from the confusion matrices.

In the initial setup of the empirical study, the LR classifier is trained with and without PCA dimension reduction for all stocks, and it is concluded that there is no significant difference in the estimation performance. Thus, all the classifier comparisons are conducted by first applying PCA to the raw data.

Second, LR, SVM-rbf, SVM-sigmoid, and naive bayes, random forest, KNN, and decision tree classifiers are trained with the not-shuffled time series data. Chronologically, the last five percent of the data for each stock is employed in the test set. With the help of the study of Takashi et al. (2022), we include confidence intervals of MA-F measures into the analysis, and the existence of strong sign prediction methods indicates that 9 of 26 BIST 30 index equities are statistically predictable. This MA F-measure analysis provides a statistical comparison of machine learning algorithm performances for intraday sign prediction on the Borsa Istanbul. For the 10-min-lagged data, only four equities stayed predictable, whereas none of the equities are predictable with 1-h-lagged "data analytics."

However, statistical success cannot be regarded as the key to economic gain (Leitch & Tanner, 1991). Therefore, as a next step, we analyzed the economic gains by the inspiring methods from the study of Chronopoulos et al. (2018). The literature on sign prediction and its economic significance is a very new and small field of study, and studies in this field typically make predictions for two classes, namely "up" or "down," for stocks while including transaction costs. The mentioned study examines daily changes and proposes investing in an index or in riskless interest-earning assets. However, there are no intraday riskless interest-earning assets in practice. Additionally, only buy, and sell signs can create excessive commission costs for intraday algorithms. Therefore, this study proposes three classes, namely "up," "down," or "steady," for buy, sell, and hold signs, and thus could create a better basis for analyzing economic earnings via intraday sign predictions. Only the nine equities with statistically strong sign prediction results were allowed in the economic gain analysis. Both short-selling-allowed and short-selling-not-allowed strategies were included. The results indicated that better economic gains could be achieved for three of the nine equities when short selling is not allowed. In addition, four of the nine equities yield greater financial gains when short selling is allowed with the help of using estimations of machine learning algorithms.

This study shows that the "data analytics" set provided by the Borsa Istanbul is not being efficiently used by market players and that valuable information is not reflected in prices in a 5-min interval. According to the efficient market hypothesis, these results allow us to assess market efficiency and imply that the information provided by the Borsa Istanbul every second is not absorbed by the market, meaning that the "semistrong form of market efficiency" is not valid for at least four equities.

For further analysis from this study, it would be interesting to explore why some equities have inefficient markets and what factors contribute to these inefficiencies.

This study is expected to draw the attention of future researchers and data scientists engaged in algorithmic trading to the "data analytics" tools disseminated by the Borsa Istanbul. By demonstrating the estimation power of this big data set, new researchers can be encouraged to use commercially available data. Furthermore, this study reveals that, besides the Deutsche Börse, and Borsa Istanbul, other exchanges could provide big data analytics to traders and scientists to enhance market efficiency.

Declaration of competing interest

I hereby declare that the disclosed information is correct and that no other situation of real, potential or apparent conflict of interest is known to me.

Appendix A. Stationarity tests for stock prices and price changes and proposed ARIMA models

	ADF statistics for Price	PP Z values for Price	ADF statistics for Price Differences	PP Z values for Price Differences	Proposed Auto ARIMA Model
AKBNK	−1.5894 (0,7527)	−5.5658 (0,7997)	−26.845 (<0.01)	−23,058 (<0.01)	ARIMA (0,1,1)
ARCLK	−1.3543 (0,8524)	−4.095 (0,8815)	−27.272 (<0.01)	−22,701 (<0.01)	ARIMA (3,1,1)
ASELS	−1.9967 (0,5799)	−7.8092 (0,6748)	−27.088 (<0.01)	−19,645 (<0.01)	ARIMA (2,1,2)
BIMAS	−2.0127 (0.5731)	−8.0016 (0.6642)	−28.894 (<0.01)	−21,907 (<0.01)	ARIMA (1,1,1)
DOHOL	−1.9358 (0.6057)	−10.828 (0.5069)	−28.704 (<0.01)	−25,109 (<0.01)	ARIMA (1,1,1)
EKGYO	2.4398 (0.3919)	−13.336 (0.3674)	−28.25 (<0.01)	−25,961 (<0.01)	ARIMA (0,1,1)
EREĞL	−1.2927 (0.8785)	−5.3208 (0.8133)	−28.694 (<0.01)	−21,948 (<0.01)	ARIMA (0,1,4)
GARAN	−1.154 (0.9142)	−3.2008 (0.9253)	−26.172 (<0.01)	−23,015 (<0.01)	ARIMA (0,1,5)
HALKB	−2.497 (0.3676)	−9.7707 (0.5657)	−25.854 (<0.01)	−24,110 (<0.01)	ARIMA (5,1,2)
ISCTR	−1.6584 (0.7234)	−6.5233 (0.7464)	−26.8 (<0.01)	−23,445 (<0.01)	ARIMA (1,1,1)
KCHOL	−1.8403 (0.6462)	−5.6733 (0.7937)	−27.794 (<0.01)	−22,226 (<0.01)	ARIMA (1,1,1)
KOZAA	−2.2488 (0.4729)	−10.164 (0.5439)	−27.348 (<0.01)	−22,840 (<0.01)	ARIMA (1,1,1)
KOZAL	−2.7456 (0.2622)	−16.129 (0.222)	−27.865 (<0.01)	−21,991 (<0.01)	ARIMA (2,1,0)
KRDMD	−1.0467 (0.9315)	−2.8629 (0.9405)	−27.721 (<0.01)	−22,179 (<0.01)	ARIMA (5,2,0)
PETKM	−1.1384 (0.9167)	−2.9613 (0.9361)	−27.723 (<0.01)	−23,377 (<0.01)	ARIMA (1,1,0)
PGSUS	−2.444 (0.3901)	−11.575 (0.4654)	−26.665 (<0.01)	−22,728 (<0.01)	ARIMA (1,1,0)

(*continued on next page*)

(*continued*)

	ADF statistics for Price	PP Z values for Price	ADF statistics for Price Differences	PP Z values for Price Differences	Proposed Auto ARIMA Model
SAHOL	−2.3547 (0.428)	−11.569 (0.4657)	−27.161 (<0.01)	−22,418 (<0.01)	ARIMA (2,1,2)
SISE	−1.9538 (0.5981)	−7.9789 (0.6654)	−27.475 (<0.01)	−22,516 (<0.01)	ARIMA (1,1,2)
TAVHL	−1.8989 (0.6214)	−7.2713 (0.7048)	−27.024 (<0.01)	−23,220 (<0.01)	ARIMA (3,1,2)
TCELL	−2.0741 (0.547)	−8.3248 (0.6462)	−26.198 (<0.01)	−22,036 (<0.01)	ARIMA (4,1,4)
THYAO	−2.3087 (0.4475)	−10.269 (0.538)	−27.074 (<0.01)	−24,018 (<0.01)	ARIMA (2,1,2)
TKFEN	−2.5387 (0.3499)	−12.759 (0.3995)	−27.329 (<0.01)	−22,886 (<0.01)	ARIMA (2,1,1)
TOASO	−3.2446 (0.0799)	−21.377 (0.0527)	−28.081 (<0.01)	−22,617 (<0.01)	ARIMA (3,1,0)
TTKOM	−1.92 (0.6124)	−7.7113 (0.6803)	−26.904 (<0.01)	−23,302 (<0.01)	ARIMA (2,1,1)
TUPRS	−2.3931 (0.4117)	−9.4936 (0.5812)	−26.621 (<0.01)	−21,262 (<0.01)	ARIMA (0,1,2)
VAKBN	−1.1824 (0.9096)	−4.1943 (0.876)	−26.878 (<0.01)	−24,156 (<0.01)	ARIMA (2,1,0)
YKBNK	−2.8361 (0.2238)	−15.133 (0.2674)	27.303 (<0.01)	−25,192 (<0.01)	ARIMA (0,1,1)

- Numbers in parentheses near Augmented Dickey-Fuller Test and Phillips-Perron Unit Root Test Statistics represent probability (P) values for the statistics.

Appendix B. List of data analytics by Borsa Istanbul

Analytic	Definition
Number of arrived orders	Number of orders arriving per 60 s
Total number of arrived orders	Total number of arrived orders up to that time
Quantity of arrived orders	Quantity of arrived orders per 60 s
Total Quantity of arrived orders	Accumulated quantity of arrived orders up to that time
Number of arrived Buy orders	Number of Buy orders arriving per 60 s
Number of arrived Sell orders	Number of Sell orders arriving per 60 s
Quantity of arrived Buy orders	Quantity of Buy arriving per 60 s
Quantity of arrived Sell orders	Quantity of Sell arriving per 60 s
Number of Immediate-or-Cancel orders	Number of Immediate or Cancel Orders arriving per 60 s
Number of cancelled orders	Number of cancelled orders per 60 s interval
Quantity of cancelled orders	Quantity of cancelled orders per 60 s interval
Number of cancelled Buy orders	Number of cancelled buy orders per 60 s
Number of cancelled Sell orders	Number of cancelled sell orders per 60 s
Quantity of cancelled Buy orders	Quantity of cancelled buy orders per 60 s
Quantity of cancelled Sell orders	Quantity of cancelled sell orders per 60 s
Total number of cancelled orders	Total number of cancelled orders up to that time
VWAP of cancelled orders	Volume weighted average price of cancelled orders up to that time
VWAP of cancelled Buy orders	Volume weighted average price of cancelled buy orders up to that time
VWAP of cancelled Sell orders	Volume weighted average price of cancelled sell orders up to that time
Cancel/Order Ratio 1	The ratio of the total number of cancelled orders to the total number of arrived orders per 60 s
Cancel/Order Ratio 2	The ratio of the accumulated quantity of cancelled orders to the accumulated quantity of arrived orders per 60 s
Cumulative Cancel/Order Ratio 1	The ratio of the total number of cancelled orders to the total number of arrived orders up to that time
Cumulative Cancel/Order Ratio 2	The ratio of the accumulated quantity of cancelled orders to the accumulated quantity of arrived orders up to that time
Average quantity of arrived Buy orders	Average of buy order quantity in the last 5 min
Average quantity of arrived Sell orders	Average of sell order quantity in the last 5 min
Volatility of arrived Buy order quantities	Volatility of buy order quantity in the last 5 min
Volatility of arrived Sell order quantities	Volatility of sell order quantity in the last 5 min
VWAP of trades in the last 5 min	Volume weighted average price (VWAP) of trades per 5-minutesinterval
VWAP of all trades	Volume weighted average price (VWAP) of all trades up to that time
VWAP of Buyer initiated trades	Volume weighted average price (VWAP) of buyer-initiated trades per 5-minutesinterval
VWAP of Seller initiated trades	Volume weighted average price (VWAP) of seller-initiated trades per 5-minutesinterval
Number of Buyer initiated trades	Number of buyer-initiated trades per 60 s
Number of Seller initiated trades	Number of seller-initiated trades per 60 s
Quantity of Buyer initiated trades	Quantity of buyer-initiated trades per 60 s
Quantity of Seller initiated trades	Quantity of seller-initiated trades per 60 s
Buyer/Seller Ratio 1	The ratio of the number of buyer-initiated trades to the number of seller-initiated trades per 60 s
Buyer/Seller Ratio 2	The ratio of the quantity of buyer-initiated trades to the quantity of seller-initiated trades per 60 s
Cumulative Buyer/Seller Ratio 1	The ratio of the total number of buyer-initiated trades to the total number of seller-initiated trades up to that time
Cumulative Buyer/Seller Ratio 2	The ratio of the accumulated quantity of buyer-initiated trades to the accumulated quantity of seller-initiated trades up to that time

References

Aksoy, B. (2021). Predicting direction of stock price using machine learning. *Business and Economics Research Journal, 12*(1), 89–110.

Akyildirim, E., Sensoy, A., Gulay, G., Corbet, S., & Salari. (2021). Big data analytics, order imbalance and the predictability of stock returns. *Journal of Multinational Financial Management, 62*.

Amjady, N., & Daraeepour, A. (2009). Design of Input vector for day-Ahead price forecast- Ing of electricity markets. *Expert Systems with Applications, 36*(10), 12281–12294.

Ballings, M., Poel, D. V., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications, 42*(20), 7046–7056.

Bastı, E., Kuzey, C., & Delen, D. (2015). Analyzing initial public offerings' short-term performance using decision trees and SVMs. *Decision Support Systems, 73*, 15–27.

Bekiros, S. (2010). Heterogeneous trading strategies with Adaptive fuzzy Actor-critic reinforcement learning: A behavioural Approach. *Journal of Economic Dynamics and Control, 34*, 1153–1170.

Bhardwaj, R. B. (1992). The January anomaly: Effects of low share price. *Transaction Costs, and Bid-Ask Bias, 47*, 553–574.

Björklund, M. (2019). Be careful with your principal components. Evolution. *International Journal of Organic Evolution, 73*(10), 2151–2158.

Bossaerts, P., & Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies, 12*(2), 405–428.

Breen, W., Glosten, L. R., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of Finance, 44*(5), 1177–1189.

Christoffersen, P. F., & Diebold, F. X. (2006). Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science, 52*(8), 1273–1287.

Christoffersen, P., Diebold, F., Mariano, R., Tay, A., & Tse, Y. (2007). Direction-of-Change forecasts based on conditional variance, skewness and Kurtosis dynamics: International evidence. *Journal of Financial Forecasting, 1*, 3–24.

Chronopoulos, D. K., Papadimitriou, F. I., & Vlastakis, N. (2018). Information demand and stock return predictability. *Journal of International Money and Finance, 80*(C), 59–74.

Cortes, C., & Vapnik, V. N. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297.

Coşkun, Y., & Seven, Ü. (2016). Efficiency of financial markets. In A. Gündogdu (Ed.), *Financial markets and institutions: Theory and practice in Turkey* (pp. 288–312). Ankara: Seçkin Yayıncılık.

Diebold, F. X., & Lopez, J. A. (1996). 8 forecast evaluation and combination. *Handbook of Statistics, 14*, 241–268.

Dunteman, G. H. (1989). *Principal components analysis*. Thousand Oaks: SAGE Publications.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical works. *The Journal of Finance, 25*(2), 383–417.

Fama, E. F., & Blume, M. E. (1966). Filter rules and stock market trading. *Journal of Business, 39*(S1), 226–241.

Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science, 49*(5), 639–654.

Gunduz, H., & Cataltepe, Z. (2015). Borsa Istanbul (BIST) daily prediction using financial news and balanced feature selection. *Expert Systems with Applications, 42*(22), 9001–9011.

Gunduz, H., Yaslan, Y., & Cataltepe, Z. (2017). Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations. *Knowledge-Based Systems, 137*, 138–148.

Jordan, S. J., Vivian, A. J., & Wohar, M. E. (2014). Forecasting returns: New European evidence. *Journal of Empirical Finance, 26*, 76–95.

Khan, A., & Ikram, S. (2010). Testing semi-strong form of efficient market hypothesis in relation to the impact of foreign institutional investors (FII's) investments on Indian capital market. International journal of trade. *Journal of Economics and Finance, 1*, 373–379.

Kılıç, Y., & Buğan, M. F. (2016). The efficient market hypothesis: Evidence from Turkey. *International Journal of Academic Research in Business and Social Sciences, 6*, 262–272.

Leitch, G., & Tanner, J. (1991). Econometric forecast evaluation: Profits versus the conventional error measures. *The American Economic Review, 81*, 580–590.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth and expected stock returns. *The Journal of Finance, 56*(3), 815–849.

Markovíc, I., Stojanovíc, M., Stankovíc, J., & Božíc, M. (2014). Stock market trend prediction using support vector machines. *Facta Universitatis – Series: Automatic Control and Robotics, 13*(3), 147–158.

Pehlivanlı, A.Ç., Aşıkgil, B. A., & Gülay, G. (2016). Indicator selection with committee decision of filter methods for stock market price trend in ISE. *Applied Soft Computing, 49*(2), 792–800.

Persio, L. D., & Honchar, O. (2016). Artificial neural networks Architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing, 10*, 410–413.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance, 50*(4), 1201–1228.

Pesaran, M., & Timmermann, A. (2000). A recursive modelling Approach to predicting UK stock returns. *Economic Journal, 110*, 159–191.

Pontiff, J. S.-t.-m., & Pontiff, J. (1998). Book-to-market ratios as predictors of market returns. *Journal of Financial Economics, 49*(2), 141–160.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio Management, 11*(1), 68–75.

Sapankevych, N., & Sankar, R. (2009). Time series prediction using support vector ma-chines: A survey. *IEEE Computational Intelligence Magazine, 4*(2), 24–38.

Seijo-Pardo, B., Porto-Díaz, I., Bolón-Canedo, V., & Alonso-Betanzos, A. (2017). Ensemble feature selection: Homogeneous and heterogeneous approaches. *Knowledge-Based Systems, 118*, 124–139.

Shang, H. L. (2017). Forecasitng intraday S&P 500 index returns: A functional time series Approach. *Journal of Forecasting, 36*(5), 741–755.

Shiller, R. J., & Perron, P. (1985). Testing the random walk hypothesis. *Economics Letters, 18*(4), 381–386.

Sree, K. D., Bindu, & Shoba, C. (1998). Data analytics: Why data normalization. *International Journal of Engineering and Technology, 7*, 209–2013.

Takahashi, K., Yamamoto, K., Kuchiba, A., & Koyama, T. (2022). Confidence interval for micro-averaged F1 and macro-averaged F1 scores. *Applied Intelligence, 52*(5), 4961–4972.

Torun, M., & Kurt, S. (2008). Testing weak and semi strong form efficiency of stock exchanges in European mnetary union, countries: Panel data causality and Co-integration analysis. *International Journal of Economic and Administrative Studies, 1*, 67–82.

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). John Wiley & Sons.

Vui, C. S., & G S. (2013). *A review of stock market prediction with artificial neural network (ANN),. IEEE international conference on control system. Computers and industrial engineering (ICCSCE)*. IEEE.