Received: 9 May 2022 | Accepted: 15 November 2022

DOI: 10.1002/for.2933

RESEARCH ARTICLE

WILEY

# Early prediction of Ibex 35 movements

I. Marta Miranda García[1] | María-Jesús Segovia-Vargas[2] | Usue Mori[3] | José A. Lozano[3]

[1]Department of Financial Economy and Accounting, Universidad Pablo de Olavide, Seville, Spain

[2]Department of Financial and Actuarial Economics and Statistics, Complutense University of Madrid, Campus de Somosaguas, Madrid, Spain

[3]Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, Leioa, Spain

**Correspondence**
I. Marta Miranda García, Department of Financial Economy and Accounting, Universidad Pablo de Olavide, Ctra. Utrera, Km. 1, 41013, Seville, Spain.
Email: immirgar@upo.es

**Funding information**
This research was supported by a grant from the Santander-UCM research project, call 2019, with reference PR87/19-22586 and by Spanish Ministry of Science and Innovation research project, with reference PID2020-115700RB-I00.

**Abstract**

In this paper, we examine the early predictability of the market's directional movement using intraday high-frequency data (695,764 observations) from an stock index (Ibex 35 Index) to provide, either private or institutional investors, an early warning system based on an "early indicator" of the financial market fluctuations with an optimal combination of the two more relevant variables for this strategy, accuracy, and earliness. A novel supervised machine learning early classification technique (Artificial Intelligence) has been applied, for the first time, to the high-frequency time series of both price and certain technical indicators.

The results obtained allow us to assert that the intraday movement of the Ibex 35 can be predicted with acceptable levels of accuracy 24 min after the start of the session and to establish certain informative intraday hourly patterns. Consequently, different indicators of precision and earliness in the session are generated, obtaining that, after a certain point in the session, no gains in precision are generated.

**KEYWORDS**

artificial intelligence, high-frequency data, intraday pattern, price discovery, stock price prediction, trading hours

**JEL CLASSIFICATION**

G17, G14, C55, C22, C38

## 1 | INTRODUCTION

One of the main functions of financial markets is to provide resources for the growth of companies and, therefore, of the economy and society as a whole, which is associated with sustainability and well-being.

Countries in which companies access and diversify their sources of financing through the capital markets have stronger and less vulnerable economies, as well as allow numerous investment alternatives in accordance with the level of risk they are willing to assume. In addition, the past financial crisis and the current crisis resulting from the COVID-19 pandemic have caused market turmoil affecting the main stock indexes, and, as a result, the systematic risk. Consequently, having tools and indicators that anticipate market movements, represented by the stock market index, is still a very topical issue.

Among the many strategies for trading in financial markets, there is the market timing strategy, for which market movements are particularly relevant. These types of active strategies aim to predict the direction of the markets and decide the timing of moving in and out of the markets or switching between different asset classes. This way, "adjusting the systematic risk level of a portfolio in anticipation of broad market price movements", it is possible to obtain returns in excess of those expected of an unmanaged portfolio (Pfeifer, 1985). Therefore, the attractiveness of market timing is a function of risk attitude and the ability to forecast better than the mean–variance approach.

Thus, predicting the movement of the stock indexes, as well as their volatility, is an issue that remains challenging today and could contribute to the sustainability of a financial system as they take part, in a great extent, to the market risk.

There is no theoretical or the empirical consensus as to whether the prices of financial assets or their movement are predictable or not. Although until the beginning of the 1980s the Efficient Market Hypothesis (Fama, 1970) and its complements, the Random Walk Theory (Malkiel, 1973), were widely accepted as a pillar of financial theory, the emergence of a new paradigm consisting of the transition from rational to behavioral economics provided new ways of trying to explain why financial markets are not efficient but predictable (Lo & MacKinlay, 1988), because to some extent prices behave according to predictable patterns, unlike the paradigm defended by the efficiency hypothesis.

There are anomalies caused by the biased and non-rational behavior of the human being that causes the appearance of such patterns (Sewell, 2007). These are repeated over time and turn the financial markets into a memory repository that can be used to obtain extraordinary returns, thereby exceeding those that would be obtained for that investment adjusted to its level of risk.

All this seems to suggest that investors are heterogeneous, and, if they are, not all have the same attitude toward risk (Chen & Li, 2016). These authors point out that the heterogeneity of investors could be associated with different factors such as portfolio objectives, modeling techniques, trading frequencies, investment strategies, speculative intention, herding behavior, and/or momentum (MTM) trading.

In this way, the market timing strategy argues that asset prices and their movement are to some extent predictable. So, these predictions can be used for optimal and profitable decision making, such as portfolio composition and rebalancing, risk management, and arbitrage. Therefore, it would be useful to incorporate the time variable together with accuracy to facilitate decision-making on market entry and exit.

Thus, we have selected a methodology that does not aim to improve the accuracy of other AI techniques, but rather to optimally combine timing and prediction. This strategy will allow us to obtain an instrument or tool that facilitates decision-making for agents with different risk profiles that use the "timing market" strategy, which tries to avoid the worst days and take advantage of the best days.

As a result, the objective of our research is to provide, either private or institutional investors, an early warning system based on an "early indicator" of the financial market fluctuations with an optimal combination of the two more relevant variables for this strategy, accuracy, and earliness. For this purpose, we will focus on predicting the movement of the stock indexes. In particular, we will focused on IBEX 35 index that is of the Spanish Stock Exchange, using time series of prices and technical indicators and trying to avoid more complex systems that use multiple factors to design early warning systems (Wang et al., 2020).

This analysis consists of a variety of solutions that find different trade-offs between the two strategies (accuracy and earliness), allowing the user to select the most appropriate solution a posteriori, depending on the accuracy and earliness requirements.

The solutions are obtained by applying, for the first time, a multi-objective early classification algorithm to the financial time series formed by the values of different technical indicators used in many studies (Atsalakis & Valavanis, 2009; Van den Poel et al., 2016). This method based on artificial intelligence will allow the incoming sequence to be classified as soon as possible, while maintaining the appropriate levels of accuracy in the predictions. Therefore, we can say that the problem of early classification is to optimize two objectives simultaneously: accuracy and earliness.

In contrast to the time series classification models used in a large number of papers (Gu et al., 2020), early classification of time series is a supervised learning problem whose objective is to build models that are capable of classifying as accurately and early as possible, when only a part of the series is available. It is logical to think that accuracy and earliness are contradictory objectives, because the more we wait, the more data points of the series are available, and as a result, it is easier to make accurate predictions of the class. On the contrary, if we want to make very early predictions, we will have great uncertainty about the time series and it will be more difficult to make accurate class predictions. Taking into account these facts, our research question can become a multi-objective optimization problem and has to be solved taking into account this especial feature (Mori et al., 2019).

In addition, we will focus on the prediction of the Spanish financial market using high-frequency intraday data, because agents and investors operating in these time periods need earliness in predictions, above all, together with accuracy (Tanaka-Yamawaki et al., 2009).

The rest of this paper is presented in the following order. In Section 2, a literature review is provided for both the conceptual and methodological levels. The data and methodology are presented in Section 3. The last sections include the findings of the study and some final observations.

## 2 | LITERATURE REVIEW

This section reviews the conceptual and methodological framework that has guided the most significant recent research in terms of predicting asset and index prices and/or their movement.

As pointed out in Section 1 of this paper, the prevailing financial theory, based on the assumption of investor rationality and its homogeneity, has remained the efficient market theory until almost the end of the 20th century (Fama, 1970).

The price reflects all relevant historical, public, and private information (Fama, 1970, p. 383) so that only new information, whose entry is unpredictable, would affect the movement of those values, and the past cannot be used to predict the future. Following this theoretical framework, prices could not be predicted nor the future events that guide the movements of those prices.

Later, Malkiel's Random Walk Theory (Malkiel, 1973) states that successive changes in asset prices are not dependent. Consequently, these changes behave as a random process, so that the changes that occur in a period of time are independent of others and are identically distributed (Rechenthin & Street, 2013).

If these assumptions are true, the prices of financial assets behave in a way that fluctuates randomly and are not predictable beyond mere chance. Therefore, the accuracy of the prediction would not exceed 50% (Bollen et al., 2011) when trying to forecast their movement, as well as their profitability, the future direction of the price of a financial asset being no more predictable than the pattern that follows a series of random numbers.

This has important implications, not only for researchers but also for investors, because the historical price series could not be used to predict future behavior and according to Malkiel (2003), statistical dependencies are extremely small and are unlikely to allow investors to obtain higher returns.

Despite the theoretically sound nature of the efficient market hypothesis and the huge empirical research confirming the efficiency of the financial market and its randomness, the evidence is not conclusive.

Studies of the last 30 years have also found evidence of a significant dependence between price movements, which would confirm that they are partially or temporarily predictable (Lo & MacKinlay, 1988; Shiller, 2000). Consequently, it could be possible to find patterns that repeat over time in a systematic way, and this fact will provide the possibility of beating the market. Therefore, the market would not be efficient and would have memory, being able to predict to some extent both long and short term (Abergel & Politi, 2013; Timmermann, 2008).

There are countless studies that come out against the efficiency of financial markets, some of them are collected in the recent review of Ying et al., 2019, especially because of the development of the advanced prediction systems, thanks to modern computing and artificial intelligence (Henrique et al., 2019).

The explanation for the lack of efficiency and the presence of these patterns of the financial markets has been explained by different theoretical contributions, the most relevant being Behavioral Finance, which is undoubtedly the emerging economic paradigm and whose beginning dates back to 1985 (De Bondt & Thaler, 1985).

This approach argues that the market presents patterns that repeat themselves because inefficiencies are caused by human behavior, where decisions are made that are not always rational but biased in an almost systematic way (Jegadeesh & Titman, 2001).

Thus, psychological factors affect investment behavior (Sewell, 2007), and this behavior affects price movements, reflecting goals, errors, and overconfidence, overreacting or underreacting to new events, etc.; consequently, they move away from their fundamental value and cause the inefficiency of markets. Since the 1990s, several studies have focused on examining the inefficiency of markets in an attempt to detect anomalies, particularly in relation to the predictability of profitability (Rossi, 2016).

The persistence of certain patterns associated with market movements clearly indicates that markets are somewhat predictable. Some of the most relevant patterns can be found in Jacobsen et al. (2005). They developed a taxonomy based on price, size, and timing anomalies, the most significant being the calendar effect related to seasonal effects. This anomaly is because of the market behaving differently according to the time of day, the day of the week, or in certain months or years (Rossi, 2016). These anomalies are the MTM anomaly (Jegadeesh & Titman, 1993; Vayanos & Woolley, 2013) and trend reversal anomaly or return reversal effect (Basdekidou, 2017). Jegadeesh and Titman (1993) described the fact that stocks that have risen in the last 6–12 months continued to perform well in the following

6–12 months (and vice versa). Basdekidou (2017) defined this effect as financial assets that performed well in the past tend to perform poorly in the future, and stocks that performed poorly in the past tend to perform better in the future.

In behavioral finance, the presence of the MTM effect is often attributed to cognitive biases of irrational investors such as investor herding, overreaction and underreaction by investors, and confirmation bias (Barberis et al., 1998).

Although MTM and trend-reversal anomalies have been analyzed especially in monthly, weekly, or daily frequency, the increase of technology has led to a substantial increase in high-frequency trading (HFT) and, therefore, in the interest in the presence of these anomalies in the intraday. The impact of HFT has changed the way of trading, the way markets are structured, and how price discovery arises (O'Hara, 2015). The identification of intraday patterns also constitutes a temporary market anomaly, as already identified by Wood, McInish, and Ord in Wood et al., 1985.

An important work on intraday patterns is that of Gao et al. (2018) who introduce a form of intraday MTM, where the return of the first half-hour calculated from the previous day's closing price predicts positively the return of the last half-hour before the market closes on the same trading day in the SPY ETF.

Caporale et al. (2016) also point out some of these intraday patterns and anomalies, such as half-day effects, abnormally low returns in the middle of a trading session, accompanied by a sharp drop in trading volumes; last-minute and first-minute effects (with the last trading hour being the best and the first hour being the worst in terms of performance); and the time of day anomaly (with stocks tending to rise in the first 45 and last 15 min of the trading day).

The presence of intraday patterns in price behavior is especially relevant for this research, and specifically in the IBEX 35, because this supports the hypothesis that certain predictions can be made about intraday price movements, following the idea suggested by Bildik (2001) that there are significant potential gains if investors use trading systems based on intraday seasonality in share returns.

In the case of the Spanish market, some of these patterns have also been detected, different effects and patterns on profitability and volatility being studied (García-Machado & Rybczyński, 2017; Ortiz et al., 2015).

Camino (1996) investigated the intraday effect by dividing the IBEX-35 index into 15-min periods and found that yields were statistically different in the first 4 h after the opening of transactions.

Among the most recent articles, we can mention the Miralles-Quirós et al. (2015)'s paper. They follow Chen's (2013) findings and identify intraday patterns by analyzing and evaluating the most profitable strategies depending on whether we enter at certain times of the day from the opening of the market until its closing, as well as the linkage with the American market. Several hypotheses are contrasted such as the strategy of entering the market long or short at 09:14:59 depending on how prices are between 09:04:59 and 09:00:01. They conclude that investors should wait until the second part of the trading day and evaluate the effects of the opening of the market in New York to decide whether to enter or not at the end of the session.

Considering the conflicting results provided by the empirical evidence, this research argues that the movement of the indexes can be predicted to some extent and could therefore be used to implement market timing strategies using as a complement an "early warning" indicator or signal of market movement. So, the following research question is proposed:

> **RQ1.** *To what extent is the capital market predictable by setting the time horizon at the end of a session and using only high-frequency intraday data in the form of historical prices and/or technical indicators based on those prices?*

On the other hand, it should not be forgotten that, although the emergence of new techniques and methodologies has made it possible to analyze a greater amount of data more accurately, the task of predicting the movement of financial markets is by no means easy (Esling & Agon, 2012). A large number of economic and social factors, political and psychological events, expectations, and movements in other markets, among others, affect the price of financial assets or indices (Kara et al., 2011). Thus, time series of a complex, dynamic, chaotic, nonparametric, nonlinear nature with high dimensionality and noise are generated (Hadavandi et al., 2010).

In addition, the dynamic changes in the relationship between independent and dependent variables often occur in financial time series, and identical statistical properties cannot be observed at every point in time (Hsu et al., 2009). Another factor that makes it difficult to predict future values is the high volatility in the price time series (Atsalakis & Valavanis, 2009). All these characteristics inherent in financial time series require more adaptable and flexible mechanisms to improve not only forecasting accuracy but also earliness.

From the market timing strategies, it is logical to think that the more accurate the predictions are, the less risk investors will take. Therefore, the techniques applied to financial time series have sought to make progress in improving that accuracy. This can reduce the risk of investing in financial markets in order to maximize capital gains and minimize losses (Qiu & Song, 2016).

Focusing on the methods, many economists prefer econometric techniques to machine learning (ML) and a gap is produced which, as Hsu et al. (2016) point out, leads to contradictory and disparate research results because of the use of different "modeling cultures". This paper also points out that, comparing the accuracy of the most commonly used econometric and ML methods, the best ML methods surpass the best econometric techniques, thus supporting the results obtained by numerous studies (Ballings et al., 2015).

Consequently, even though the number of algorithms is huge in the academic literature, we can mention some of the time series classifier algorithms that have been used more in previous research because one of our objectives is to try to predict the intraday movement of the Ibex 35 (up or down) and not its value. This means that we treat our research question as a time series classification problem, as has been done in numerous studies (Qian & Rasheed, 2007.

Nonlinear models used for financial forecasting that use supervised learning techniques to classify time series have experienced significant growth over the past two decades (Chourmouziadis & Chatzoglou, 2016).

Among the most widely used are artificial neural networks, particularly the multi-layer perceptron (MLP), support vector machines, (SVM), Bayesian networks, and fuzzy system models (Atsalakis & Valavanis, 2009; Cavalcante et al., 2016; Gerlein et al., 2016; Henrique et al., 2019).

Among these methods, the ones that have received the most attention in the field of financial market forecasting and the best results have been neural networks and SVM (Tkáč & Verner, 2016) as well as their many innovations from deep learning such as convolutional neural networks (CNNs).

One of the explanations for their popularity for time series forecasting is that they are better suited to handle continuous covariates (Lessmann et al., 2015), which often occur in financial time series forecasting (e.g., prices, price differences, or technical indicators).

As an alternative to all these classification proposals, in the past few years, there has been a great advance in artificial intelligence methods specifically designed to work with time series. One of the most prominent tasks within this framework is the classification of time series (Esling & Agon, 2012). The starting point for this task is a database of various time series, each of which belongs to a specific class or group. The objective of supervised classification of time series is to use this labeled dataset to learn a model that will be able to predict the class of new unlabeled time series.

Given the interest in this problem in many application domains, there has been an explosion of publications concerning this problem and its variants. A detailed review can be found in Bagnall et al. (2017). Also, given its usefulness, many variants of this problem have been proposed in the past decades. As an example, various authors have shown an interest in performing early classification of time series. In this case, the objective is to predict the classes of new unlabeled time series, but as early as possible, preferably, before the full time series is collected. As such, the objective is not only to obtain a high accuracy in the class predictions, but it is also required that these predictions be made early. Xing, Pei, and Yu (2011) defined this problem initially and proposed a method based on nearest neighbor classifiers. After this study, many others have been proposed in the literature in order to provide predictions of market movements. Compared to other ML methods that attempt to make time series forecasts, they require the full time series to achieve accurate results.

Early time series classification methods can be separated into two main groups. The first group of methods is based on finding shapelets (Ye & Keogh, 2009) that are useful to discriminate between the classes as early as possible. These shapelets are subsections of the time series that contain patterns which allow to discriminate between the classes as accurately as possible, and they have been widely used in the time series classification domain. (Ghalwash et al., 2014; He et al., 2015; Xing, Yu, & Wang, 2011) extended the concept of shapelet to allow early classification of time series. The second group of methods combines a set of classifiers built in different timestamps with a mechanism to decide whether the predictions issued by them are reliable or not. Different types of classifiers and specific methods for controlling their reliability give way to a large set of different methods based on the same principle (Anderson et al., 2012; Ghalwash et al., 2012; Hatami & Chira, 2013; Mori et al., 2017; Mori et al., 2018; Xing, Pei, & Yu, 2011).

As an improvement on all these methods and with the state-of-the-art performance, in Mori et al. (2019), a multi-objective method was proposed to deal with the problem of early classification. For the first time, the authors of this study dealt with the problem of early classification of time series from a multi-objective point of view, that is, they treated both objectives—earliness and accuracy—simultaneously. The latter study will be taken as a baseline for this paper.

Even if these methods have gained considerable popularity in many application domains, it is noticeable that there are hardly any studies that use this type of methods in the financial domain, none in the case of early classification. The benefits of these approaches are various, but the most important is their inherent consideration of the temporality of the data, which is obviated if classical classification methods are used. In view of the evident temporal nature of financial data, the consideration of this type of techniques is interesting in itself. Therefore, once the first research question is answered, we proposed a second research question:

> **RQ2.** *Is the capital market being able to give early signals as to what its movement will be at the end of the day and if so, what kind of indicators can it give?*

## 3 | DATA AND METHODOLOGY

### 3.1 | Data: The stock index Ibex 35, high-frequency and technical indicators

Because our objective is to make the most accurate and early predictions possible about the direction of the stock index Ibex 35's movement in the intraday market, we have selected high-frequency data, whose availability, although it has grown recently, has not yet generated a sufficient number of studies in the Spanish market, in line with what has happened in other countries (Rechenthin & Street, 2013; Tay & Cao, 2001).

This application is especially relevant for high-frequency trading (HFT), which is characterized by a short-term investment horizon and high speed, so the ability to process data in a short time is essential (Tran et al., 2018).

The data collected are from the Ibex 35 index. This index of the Spanish market is composed of the 35 most liquid securities listed on the Stock Exchange Interconnection System of the four Spanish stock exchanges and is therefore considered a key indicator of market performance globally, as it represents approximately 90% of the cash traded on the stock exchange. It is also used as an underlying in the trading of derivative products. Therefore, it is used as a national and international benchmark of market profitability and risk.

Stock market indices have been used in a large number of previous studies that use ML to predict the movement of the stock market (Atsalakis & Valavanis, 2009; Qian & Rasheed, 2007) as well as in the Spanish market (Dunis et al., 2013; Pérez-Rodríguez et al., 2005).

The importance of stock market indices is that they can be replicated, that is, to create and/or undo a basket with the values of the index in such a way that it is a method of channeling investment and describing the trend and behavior of the movements of an economy and a market (Dunis et al., 2013).

In addition, it constitutes an early warning system for investors, especially short-term investors, against sudden falls in the market (Wang, 2014) so that forecasts of its movement are relevant in order to establish effective trading strategies (Leung et al., 2000) and that investors can hedge against possible market risks and speculators and arbitrageurs detect opportunities to make profits by trading in the stock market index (Manish & Thenmozhi, 2005).

We follow in this paper a technical approach, which considers that previous market behavior patterns are repeated based on the decisions to buy and sell an asset on its previous price patterns, trading volumes, and, potentially, other public information (Jiang et al., 2019), and there are very significant studies that confirm its usefulness to beat the market (being the most used for this purpose in short-term investment horizons) (Evans et al., 2013). Hu et al. (2015) made a comprehensive review of the literature in 2015 and concluded that the most existing studies applying soft computing and neural network for financial forecasting are based on technical analysis.

The data used in this paper are high frequency and have been acquired from BME Market Data, belonging to Grupo Bolsas y Mercados Españoles, which is specialized in the processing, generation, and commercialization of the information coming from the different Regulated Markets and Multilateral Trading Systems of the BME Group. They refer to Ibex 35 prices obtained from January 2, 2015 to October 20, 2017. This means a total of 724 trading days, corresponding to 256 in 2015 and 2016, respectively, and 212 in 2017. Each of these trading days is a time series of 961 data points, corresponding to the Ibex 35 quotations from 9:00 am to the close of the market at 17:00 pm in 30-s intervals. This means a total of 695,764 observations, which will be used directly as raw data after having demonstrated their effectiveness in the paper "Early classification of time series using multi-objective optimization techniques" (Mori et al., 2019) or as a basis for the calculation of the various technical indicators evaluated and which will be explained later.

It is important to mention the challenge of using high-frequency information, pointing out that intraday models, because of their greater complexity, are the most unusual (Gao et al., 2018).

The use of high-frequency data may incorporate "more noise" into the model and make forecasting and accuracy more difficult. Nevertheless, some research suggests that it generates better predictions (Degiannakis & Potamia, 2017).

Schulmeister (2009) argue that the profitability of technical trading may have shifted from daily data to intraday data and thus as Buchanan (2015, p.161) points out, to do "trading at the speed of light: to minimize risks, we must learn more about how financial markets operate at ever faster rates".

Timmermann (2008) also pointed out that the prediction in the intraday market is important economically speaking because the shorter the interval of time, the more times a trading strategy could be applied to take advantage of any prediction and, therefore, the greater the potential of greater profitability will be.

In the tests, we will use raw data and technical indicators selected from among the most widely used by both academics and practitioners to design their trading strategies and systems (Batten et al., 2018).

Following Schulmeister (2009), technical analysts believe that the dynamics of an asset's price pattern is a sequence of trends, which repeat across different time scales, so these indicators or models can be used for any frequency from daily data to tick data. In this way and adjusting the calculations to intraday data, the formulas are presented below, generating for each trading day a number of observations of these indicators, according to the delay (lag) used, in 30-s intervals.

These technical indicators are the following:

*Simple Moving Average (MA) for 1 and 5 intervals of 30 s.*

*The MA is the arithmetic mean of T past prices Cli*

$$\mathrm{MA} = \frac{1}{T}\sum_{i=1}^{T} Cl_i \tag{1}$$

As the number of $n$ cases increases, the moving average becomes less sensitive to short-term fluctuations.

*Weighted Moving Average (WMA) for 5 intervals of 30 s.*

A variation of the moving average, known as the WMA used in this study, assigns greater weight to more recent prices. A WMA of $P$ periods is shown in Patel et al. (2015, p. 2164). In this study, $Cl_i$ is defined as daily closing prices or updated prices, depending on the frequency used in the analysis.

$$\mathrm{WMA} = \frac{PCl_i(P-1)Cl_{i-1}\ldots + Cl_{i-P}}{P + (P-1) + \ldots + 1} \tag{2}$$

*Exponential Moving Average (EMA) for 5 intervals of 30 s.*

Attempts to assign more weight to the most recent data make it less sensitive to price. Thus, the EMA is calculated in three stages: First, calculate the MA for the period ($n$); second, calculate the Smoothing Constant (Sm) using the formula:

$$Sm = (2/(n+1)) \tag{3}$$

and finally, calculate the EMA using the formula:

$$EMA = (\text{price} - \text{previous EMA}) \times \text{Sm} + \text{previous EMA} \tag{4}$$

EMA is a type of moving average where weights of past prices decrease exponentially.

*MTM for 1, 5, and 12 intervals of 30 s.*

MTM is a simple momentum indicator. The MTM oscillator measures the pace at which a trend is accelerating or decelerating. It does it by comparing the last price near the end of a given period in the past. Like the rate of change (ROC), this indicator is very sensitive to price changes. MTM is simply calculated by subtracting the previous price close from the latest price close using the formula:

$$MTM = \textit{latest close} - \textit{specified close} \tag{5}$$

where the specified period is any previous price close specified by the trader.

*ROC for 5 intervals.*

The ROC is computed as the percentage change of the current closing price at time $t$ to the closing price from $t - n$, where $n = 5$, which implies 150 s.

$$ROCn = (Ct - Ct - n)/Ct - n. \tag{6}$$

Equation 6 shows the relative difference between the closing price on the day of forecast and the closing price $n$ interval as previously, where $n$ is equal to the input window length.

## 3.2 | Methodology: Early time series classification using multi-objective optimization techniques

This study will attempt to predict whether a specific financial index will increase or decrease during a given trading session. However, we will try to make this prediction as soon as possible, preferably, in the early hours of the trading session.

This problem is denominated early classification of time series (Xing, Pei, & Yu, 2011) in the area of ML and can be defined more formally as follows. Suppose we have a set of historical time series $X$ = {(TS1, CL1), (TS2, CL2), ..., (TSn, CLn)}, each TSi is a time series that collects all the measurements of a specific daily session from the past. CLi represents its respective class label; thus, it takes a value of 0 if the index has decreases and a value of 1 if it has increases. The objective is to use this set of historical data to build a model that relates the time series to their class labels at different times of the session. This model will then be used to predict whether the index will increase or decrease for new daily sessions as early and as accurate as possible (Mori et al., 2017).

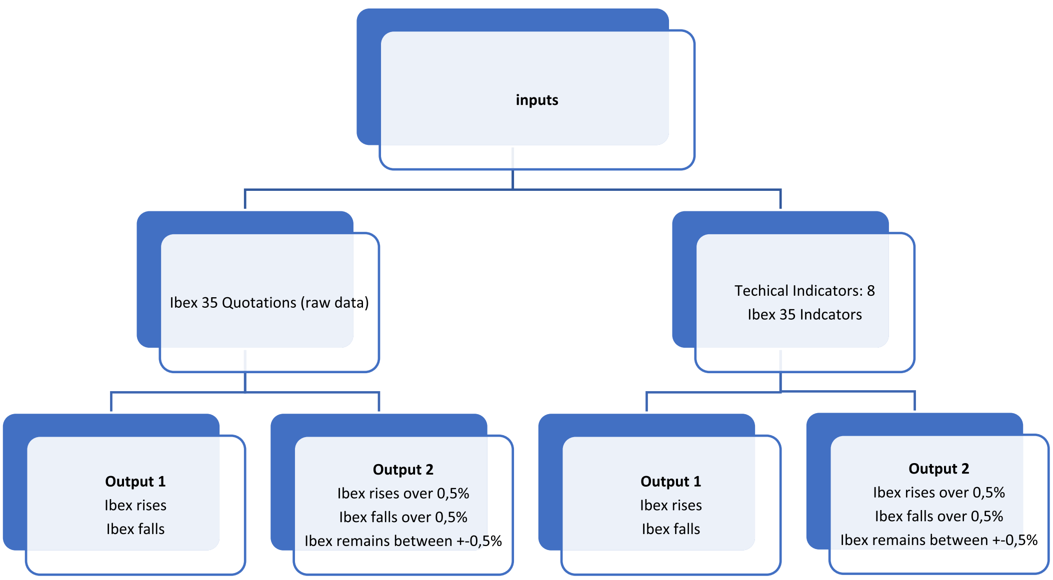
The methodology we will use to build this early classifier was initially proposed in (Mori et al., 2019) and is based on multi-objective optimization methods. The early classifier used in this study is obtained by combining two components which are explained briefly in the following paragraphs.

As a first component, our early classifier builds a set of classifiers (h1, h2, ..., hL) using Gaussian process classifiers for time series (Rasmussen & Williams, 2006). Each of these classifiers hi receives the available data of the daily session until time *i* and outputs the probability that the index will decrease (or increase). For example, classifier h11:00 will receive all the values of the selected index obtained until 11:00 and will predict, after analyzing the data, with what probability the index will decrease (or increase). These classifiers are able to predict the most probable class of a daily session at different times. However, these predictions are not guaranteed to be accurate. Possibly, at early time, there is not enough information to make good predictions. As such, we require some additional mechanism to decide whether this prediction is to be trusted or not.

In this context, the second component of the early classifier is a set of trigger functions obtained by using multi-objective optimization techniques. These trigger functions are linear rules that combine the probability value obtained by the h_t classifiers, and also the time in which the prediction is made (*t*). Based on this input information, each linear rule will tell us if the prediction provided by a classifier h_t is reliable or not. As the most innovative part of the methodology, the parameters of these linear functions are optimized using a multi-objective genetic algorithm (NSGA-II) (Deb et al., 2002), with the objective of finding the trigger functions that maximize accuracy and earliness simultaneously.

Multi-objective optimization deals with the simultaneous optimization of two or more conflicting objective functions, in this case accuracy and earliness. Typically, the sooner we want to make the predictions, the less information we have about the trading session, so it is more difficult to make accurate predictions (Mori et al., 2017). But how can we choose between a classifier that yields an accuracy of 60% at 11:00 and another that yields an accuracy of 80% at 16:00? The choice depends on the interests of the dealer, and the risks he/she is willing to take. In this context, we can say that accuracy and earliness are in conflict, and instead of searching for only one solution, we will try to find the set of Pareto solutions (Tan et al., 2005), which is the set of solutions that describe all the possible optimal trade-offs between the two objectives.

With all this in mind, note that as a direct result of using multi-objective optimization techniques, each trigger function will be different in nature: some will accept higher risks in order to provide earlier predictions, whereas other will tend to wait more in order to make




Source: Own elaboration.

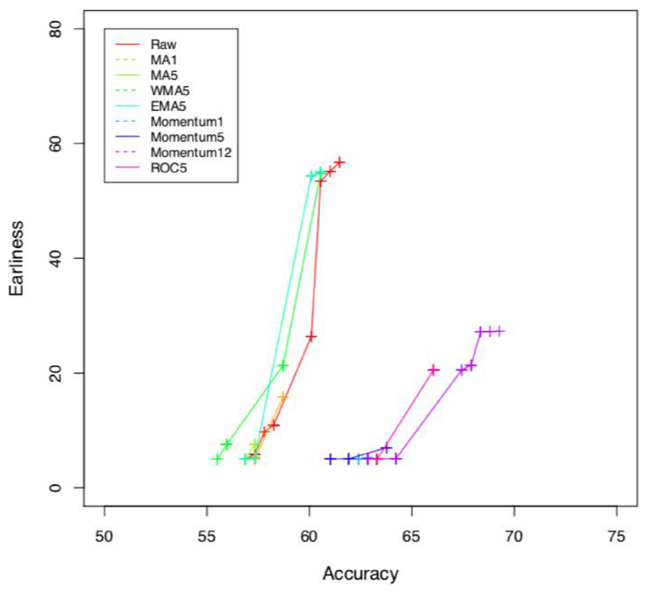
FIGURE 1 Architecture of the research Source: Own elaboration.

TABLE 1 Accuracy and earliness of Ibex 35 movement predictions by type of indicator used

| Earliness (%) | Accuracy (%) | Type of data |
|---|---|---|
| 5.00 | 56.88 | Raw |
| 5.00 | 56.88 | Ma1 |
| 5.00 | 56.88 | Ma5 |
| 5.00 | 55.50 | Wma5 |
| 5.00 | 56.88 | Ema5 |
| 5.00 | 61.93 | Momentum1 |
| 5.00 | 61.01 | Montmentum5 |
| 5.00 | 62.84 | Momentum12 |
| 5.00 | 63.30 | Roc5 |
| 5.02 | 62.39 | Momentum1 |
| 5.05 | 57.34 | Ma1 |
| 5.09 | 61.93 | Montmentum5 |
| 5.09 | 64.22 | Momentum12 |
| 5.18 | 62.84 | Momentum1 |
| 5.30 | 57.34 | Ema5 |
| 5.87 | 57.34 | Raw |
| 7.00 | 63.76 | Montmentum5 |
| 7.55 | 55.96 | Wma5 |
| 7.57 | 57.34 | Ma5 |
| 9.75 | 57.80 | Raw |
| 10.94 | 58.26 | Raw |
| 15.89 | 58.72 | Ma1 |
| 20.57 | 67.43 | Momentum12 |
| 20.57 | 66.06 | Roc5 |
| 20.57 | 66.06 | Roc5 |
| 21.33 | 58.72 | Wma5 |
| 21.38 | 67.89 | Momentum12 |
| 26.40 | 60.09 | Raw |
| 27.20 | 68.35 | Momentum12 |
| 27.27 | 68.81 | Momentum12 |
| 27.34 | 69.27 | Momentum12 |
| 53.44 | 60.55 | Raw |
| 54.36 | 60.09 | Ema5 |
| 54.72 | 60.55 | Wma5 |
| 55.05 | 60.55 | Ema5 |
| 55.14 | 61.01 | Raw |
| 56.72 | 61.47 | Raw |

Source: Own elaboration.

more accurate class predictions (Mori et al., 2019). However, an approximation of the results obtained by each trigger function will be provided to the user beforehand, and, as such, he/she will be able to choose the classifier that best suits their interests. This is especially interesting in the financial context, because different dealers could have different interests.




Source: Own elaboration.

FIGURE 2 Graphical representation of accuracy results as the session progresses (by type of indicator) Source: Own elaboration.

To summarize all this, we explain how our early classifier could be used to make predictions in real time. First, the user will be provided with a summary of the performances (accuracies and earliness values) of the different trigger functions obtained during the training phase of the method. Based on its interests, a specific trigger function will be selected. Then, at pre-specified times, the data from the daily session will be introduced into the framework. The classifier ht responsible for that timestamp will output a prediction. As the last step, the trigger function selected by the user will evaluate whether this prediction should be trusted, or, on the contrary, if the user should wait for more data to come. The procedure will end when the trigger function accepts a prediction, or when the session is finished.

## 4 | RESULTS AND DISCUSSION

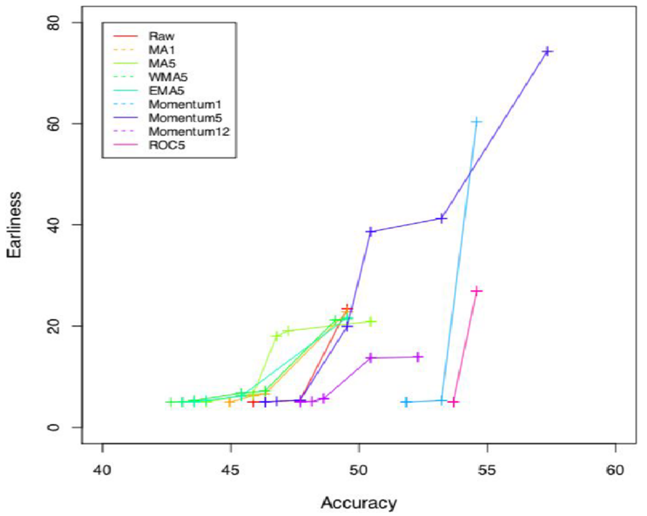
In order to answer the RQ1 and try to predict the daily movement with high-frequency intraday data, a classification problem has been designed. For this purpose, two types of experiments have been developed, which means two classifications to offer a more complete proposal to facilitate investors' decision-making.

First, a binary classification, consisting of predicting whether the market represented by Ibex 35 index, will rise (Class 0) or fall (Class 1) at the end of the daily session.

**TABLE 2** Accuracy and earliness of movement predictions ±0.5% of the Ibex 35 according to the type of indicator used

| Earliness (%) | Accuracy (%) | Type of data |
|---|---|---|
| 5.00 | 45.87 | Raw |
| 5.00 | 44.95 | Ma1 |
| 5.00 | 43.58 | Ma5 |
| 5.00 | 42.66 | Wma5 |
| 5.00 | 43.12 | Ema5 |
| 5.00 | 51.83 | Momentum1 |
| 5.00 | 46.33 | Momentum5 |
| 5.00 | 47.71 | Momuntum12 |
| 5.00 | 53.67 | Roc5 |
| 5.02 | 46.33 | Raw |
| 5.02 | 44.04 | Ma5 |
| 5.02 | 43.12 | Wma5 |
| 5.02 | 43.58 | Ema5 |
| 5.14 | 48.17 | Momuntum12 |
| 5.16 | 46.79 | Momentum5 |
| 5.32 | 53.21 | Momentum1 |
| 5.34 | 43.58 | Wma5 |
| 5.34 | 44.04 | Ema5 |
| 5.39 | 47.71 | Raw |
| 5.39 | 47.71 | Momentum5 |
| 5.71 | 48.62 | Momuntum12 |
| 6.22 | 45.41 | Ema5 |
| 6.28 | 45.87 | Ma1 |
| 6.31 | 45.41 | Ma5 |
| 6.56 | 45.87 | Ma5 |
| 6.58 | 46.33 | Ma1 |
| 6.79 | 45.41 | Wma5 |
| 7.20 | 46.33 | Wma5 |
| 13.72 | 50.46 | Momuntum12 |
| 13.88 | 52.29 | Momuntum12 |
| 18.00 | 46.79 | Ma5 |
| 19.08 | 47.25 | Ma5 |
| 19.93 | 49.54 | Momentum5 |
| 20.89 | 50.46 | Ma5 |
| 21.19 | 49.08 | Wma5 |
| 21.40 | 49.54 | Wma5 |
| 21.67 | 49.54 | Ema5 |
| 22.87 | 49.54 | Ma1 |
| 23.42 | 49.54 | Raw |
| 26.95 | 54.59 | Roc5 |
| 38.62 | 50.46 | Momentum5 |
| 41.24 | 53.21 | Momentum5 |
| 60.37 | 54.59 | Momentum1 |
| 74.31 | 57.34 | Momentum5 |

Source: Own elaboration.




Source: Own elaboration.

**FIGURE 3** Graphical representation of accuracy results as the session progresses of movement predictions ±0.5% of the Ibex 35 (by type of indicator) Source: Own elaboration.

Second, a triple ranking, which consists of predicting whether the market will go up (Class 0) more than a percentage set at 0.5%, a threshold that has been used in the preceding literature for similar problems (Shynkevich et al., 2017), whether it will go down more than that percentage (Class 1) or remain between that up-and-down range (Class 3).

A total of 18 experiments are carried out, 2 for each of the 9 inputs used.

The first two tests are performed on the "raw data" quotations of the Ibex 35 every 30 s. The remaining tests use technical indicators selected.

In this way, the architecture of the research is as follows (Figure 1):

In order to answer the RQ2, we will try to predict as soon as possible in the session whether the market represented by Ibex 35 index will rise (Class 0) or fall (Class 1) at the end of the daily session or whether the market will go up more than that percentage (Class 0) or it will go down more than that percentage (Class 1) or remain between that up-and-down range (Class 3) The aim of this classification is to obtain an early indicator.

Consequently, applying the novel methodology, the first result achieved allows us to identify how each of the selected inputs, prices (raw data), and technical indicators behave, in terms of accuracy and earliness (percentage of the session elapsed). Additionally, we find that the same indicator is not equally effective at any moment of time, when predicting the movement of the index at the end of the session.

Thus, in order to predict whether the Ibex 35 will rise or fall at the end of the session, we find the following results (see Table 1):

As it can be seen in the following graphical representation (Figure 2), the approach to the economic problem in terms of optimizing the earliness–accuracy binomial allows us to identify the MTM and ROC indicators (that technical analysis uses as trend indicators when measuring the "strength" of the trend), as those allow us to make earlier and more accurate forecasts in the session. In contrast, the indicators related to the moving average, although they are also trend indicators, obtain less accurate results and act with more delay in the session.

The results for the problem of predicting whether the index will rise or fall, or remain at the ±0.5% level, are as follows (Table 2):

Similar to the previous figure, for the more complex problem of predicting a 0.5% rise or fall in the Ibex 35, we again identify the MTM and ROC indicators (see

**TABLE 3** Time window for the IBEX 35 rise or fall at the end of the session

| Earliness (%) | Accuracy (%) | Type of data |
|---|---|---|
| 5.00 | 63.30 | Roc5 |
| 5.02 | 62.39 | Momentum1 |
| 5.05 | 57.34 | Ma1 |
| 5.09 | 64.22 | Momentum12 |
| 5.18 | 62.84 | Momentum1 |
| 5.30 | 57.34 | Ema5 |
| 7.00 | 63.76 | Montmentum5 |
| 7.55 | 55.96 | Wma5 |
| 7.57 | 57.34 | Ma5 |
| 9.75 | 57.80 | Raw |
| 10.94 | 58.26 | Raw |
| 15.89 | 58.72 | Ma1 |
| 20.57 | 67.43 | Momentum12 |
| 21.33 | 58.72 | Wma5 |
| 21.38 | 67.89 | Momentum12 |
| 26.40 | 60.09 | Raw |
| 27.20 | 68.35 | Momentum12 |
| 27.27 | 68.81 | Momentum12 |
| 27.34 | 69.27 | Momentum12 |
| 53.44 | 60.55 | Raw |
| 54.36 | 60.09 | Ema5 |
| 54.72 | 60.55 | Wma5 |
| 55.05 | 60.55 | Ema5 |
| 55.14 | 61.01 | Raw |
| 56.72 | 61.47 | Raw |

Source: Own elaboration.

**TABLE 4** Time window for the IBEX 35: to rise or fall, or remain at the ±0.5% level

| Earliness (%) | Accuracy (%) | Type of data |
|---|---|---|
| 5.00 | 53.67 | Roc5 |
| 5.02 | 46.33 | Raw |
| 5.14 | 48.17 | Momuntum12 |
| 5.16 | 46.79 | Momentum5 |
| 5.32 | 53.21 | Momentum1 |
| 5.34 | 44.04 | Ema5 |
| 5.39 | 47.71 | Raw |
| 5.39 | 47.71 | Momentum5 |
| 5.71 | 48.62 | Momuntum12 |
| 6.22 | 45.41 | Ema5 |
| 6.28 | 45.87 | Ma1 |
| 6.31 | 45.41 | Ma5 |
| 6.56 | 45.87 | Ma5 |
| 6.58 | 46.33 | Ma1 |
| 6.79 | 45.41 | Wma5 |
| 7.20 | 46.33 | Wma5 |
| 13.72 | 50.46 | Momentum12 |
| 13.88 | 52.29 | Momentum12 |
| 18.00 | 46.79 | Ma5 |
| 19.08 | 47.25 | Ma5 |
| 19.93 | 49.54 | Momentum5 |
| 20.89 | 50.46 | Ma5 |
| 21.19 | 49.08 | Wma5 |
| 21.40 | 49.54 | Wma5 |
| 21.67 | 49.54 | Ema5 |
| 22.87 | 49.54 | Ma1 |
| 23.42 | 49.54 | Raw |
| 26.95 | 54.59 | Roc5 |
| 38.62 | 50.46 | Momentum5 |
| 41.24 | 53.21 | Momentum5 |
| 60.37 | 54.59 | Momentum1 |
| 74.31 | 57.34 | Momentum5 |

Source: Own elaboration.

Figure 3), as those allow earlier and more accurate forecasts to be made, as opposed to indicators based on the moving average that are not as early and accurate as the previous ones in identifying the movement of the Ibex 35.

In the light of these findings, the following issues should be highlighted.

For all indicators and prices, the results are above 50% in the binary classification and 33.33% for the triple classification, which confirms that the Spanish market has a certain degree of intraday predictability and therefore could suggest that the intraday market is not efficient, at least for that period of time. This point would require further studies because in the event that the hypothesis of an efficient market is not satisfied, not only must there be predictability but also the market must be beaten, that is, profitable trading systems must be established (Malkiel, 2003).

The earliest time achieved for all the indicators in obtaining an accuracy of more than 50% or 33.33%, respectively, is 5% of the time spent in the session. Thus, it is only necessary to consume 24 min of the session so that the prices or technical indicators have sufficient informative value.

In the same way, all the indicators "stop working" at some point during the session. This finding is significant because we found that even if more time passes from the session contrary to what might be expected, it does not enhance the accuracy of the predictions.

This result reveals that not all the information contained in the prices or indicators has the same preeminence in predicting the movement of the market at the end of each session, and therefore, there is an informative pattern of prices.

In the first problem proposed (to predict as soon as possible if the IBEX will go up or down at the end of the session), the inputs used provide us with an informative time window that goes from 24 min after the beginning of the session, the earliest moment, to 272.25 min (4.53 h approximately from the beginning of the session), that is to say, until approximately 12:32 h.

In that time window, the informative value of prices in terms of the prediction is concentrated between approximately 9:24 and 11:12 h. Focusing on the results that are efficient (shown in bold in Table 3), a relevant finding is highlighted: although there is a time window in which different indicators can be used to make more or less accurate forecasts, there is one indicator (MTM 12) that shows the greatest efficiency in this double objective of accurately and early predicting the rise or fall of the Ibex 35.

In the second problem, the window (Table 4) is between the required minimum of 24 and 356.68 min (i.e., about 6 h after the opening of the session).

Selecting the combinations that can be efficient (shown in bold in Table 4), we can see that there are four key moments in the day. Once again, it can be seen that for this problem, the MTM and ROC oscillators obtain the best results in accuracy and earliness. This way, we can see that there are four key moments: 9:24, 11:09:36, 13:49:02, and 14:56:04 h.

Therefore, to the best of our knowledge, we found a new phenomenon in the Spanish market. We observed that the predictive power of prices and indicators occurs in a time window, and that the highest predictive value is concentrated in a narrow time range. In addition, the finding that, depending on our objective, the time window is modified as it has been shown is equally relevant.

In this way, performing a transversal analysis, we can see that the indicator that presents the best combinations of accuracy and earliness is the MTM indicator with a delay of 12 intervals for the first objective, except at 9:24 min, which is the ROC (5), whereas for the second problem, the most relevant is the ROC with a "lag" of 5 intervals, and the MTM indicators 1 and 5.

## 5 | CONCLUSIONS

The capital market can be considered as a betting market in which each trader bets on the outcome of a certain event (Page, 2012); our study would allow different strategies to be posed according to the different levels of risk tolerance. In this way, we introduce an added value by incorporating in the decision-making process not only the criterion of accuracy of the forecast (greater accuracy, less risk of suffering losses) but also earliness, so necessary in intraday trading and even more so in the case of high-frequency traders. Consequently, our results show that we have been able to limit and offer investors a certain time range where the indicators seem to show more information.

Thus, we answer the two research questions proposed. On the one hand, we are able to predict with acceptable percentages, the daily market movement using intraday data (RQ1) without using the entire time series of that session. On the other hand, as a consequence, we generate an alert system based on two evidences: the early time window in which it is possible to predict the movement of the Ibex 35 and the finding that not all information has the same utility, because after that window of time, the data do not generate any accuracy gain (RQ2).

Compared to the more conventional and ML methodologies, where the models obtain a result with their corresponding level of accuracy, this research provides a "map" of possible combinations of accuracy and earliness, which makes the decision-making for heterogeneous investors with different risk profiles easier.

Undoubtedly, investors not only need to obtain good accuracy results to make more profitable decisions but also to be fast, because the value of information is essential; every minute and every second count to establish possible profitable strategies.

It is not a question of being the earliest but of doing it in the right direction and taking into account the economic importance of earliness.

Investors are continually pursuing these information advantages, so many financial asset managers aim to generate these advantages and exploit them quickly (Kadan et al., 2018). These authors point out that having a small information advantage, even if the duration of this information is short, can generate significant advantages, and our analysis can help to achieve this.

The earliness of the classification is necessary to adapt to environments of stream data (as it is the one in our study) without impairing the performance as there is not enough learning time.

The time, therefore, is a variable to take into account especially when there are repeating patterns and anomalies (Thaler, 1987) that can be used to try to beat the market. In line with many studies that have identified intraday patterns (Admati & Pfleiderer, 1988) in different markets and scenarios, this research is able to identify the predictive power contained in certain trading hours.

The theoretical models that attempt to explain these anomalies are related, among other issues, to behavioral finance and the presence of different types of investors and their different profiles (Hasbrouck, 1991), including the behavior of liquidity traders and informed investors (Admati & Pfleiderer, 1988). The central point of the analysis of the market microstructure is the concept that in a market with asymmetrically informed agents, transactions transmit information and therefore cause a persistent price impact.

Previous studies on intraday patterns have focused on identifying the behavioral patterns of three variables in particular: the profitability, volume, and volatility of financial assets over the course of a session (Harris, 1986). As noted by Caporale et al. (2016), the so-called mid-day effects, where stock returns and trading volumes tend to be lower in the middle of a trading session, the last hour and first hour effects, meaning that stocks show lower returns in the first hour of trading and higher returns in the last hour, and the time of day anomaly, where stocks are higher in the first 45 and last 15 min of the trading day, are particularly important. Harris (1986) showed that prices and last trades tend to be higher during the first 45 min of trading sessions (every day except Monday). Wood et al. (1985) found that all the positive returns are obtained during the first 30 min and at the close of the market.

Camino's study of the Spanish stock market (Camino, 1996) found positive returns in the first hour of the session every day of the week except Monday and Wednesday, and a strong tendency for prices to rise in the first and last 15-min trading period.

Other intraday effects are, for example, the more intense trading volume at the beginning and end of the trading day combined with higher price volatility (Admati & Pfleiderer, 1988). Brooks et al. (2003) found higher trading volumes on the New York Stock Exchange at the beginning and end of the day, and that, on average, significantly positive returns are obtained during the first 30 min of trading and at the market closure.

The variation in intraday volumes has also been explained by the interactions between informed traders and liquidity traders. Liquidity traders need to meet certain liquidity needs, which are random in nature, whereas informed traders trade only in response to information. To the extent that much information is produced during non-trading hours, it could explain the high volume of market opening. Furthermore, the need to rebalance portfolios could be the cause of the increase in trading volume around the time of market closure (Admati & Pfleiderer, 1988).

There is also another intraday gap anomaly, that is, when the opening prices of certain assets of the market are different from the previous day's closing prices (Plastun et al., 2019). This fact has been used to explain such price movement because of changes in investor expectations, as well as unexpected events, among others.

Therefore, it seems reasonable to assume that the predictive power of prices concentrated on certain hours of the session, which has been detected with this research in the Spanish market, could be linked to other intraday anomalies, although the empirical analysis of this relationship will be the subject of a future study.

The results obtained with this innovative technique, which in the future can be developed, lead us to consider several interesting lines of future research: first, to deepen the knowledge of intraday anomalies and the relationship that may exist between the one detected in this paper with others found for this period such as the intraday pattern of profitability and risk; second, the use of this technique to try to anticipate the movement of individual financial assets, as well as to try to anticipate stock market shocks; and finally, the study and analysis of whether this information anomaly in prices (characterized by the presence of a certain regularity in which useful information is concentrated in certain time bands) is also detected for other time horizons, such as the weekly, monthly, or annual horizon.

## ACKNOWLEDGMENTS

U. Mori and Jose A. Lozano have been supported by the Basque Government under grant IT1504-22 and Spanish Ministry of Science PID2019-104933GB-10/AEI/10.13039/501100011033. Funding for open access publishing: Universidad Pablo de Olavide/CBUA.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

All authors have contributed to this study. All authors have read and agreed the final version of the manuscript.

## INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

## INFORMED CONSENT STATEMENT

Not applicable.

## DATA AVAILABILITY STATEMENT

Not applicable.

## ORCID

*I. Marta Miranda García* https://orcid.org/0000-0001-9731-496X

## REFERENCES

Abergel, F., & Politi, M. (2013). Optimizing a basket against the efficient market hypothesis. *Quantitative Finance*, *13*(1), 13–23.

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies*, *1*(1), 3–40. https://doi.org/10.1093/rfs/1.1.3

Anderson, H. S., Parrish, N., & Gupta, M. R. (2012). Early Time-Series Classification with Reliability Guarantee. Retrieved from http://prod.sandia.gov/techlib/access-control.cgi/2012/126961.pdf

Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*, *36*(3), 5932–5941. https://doi.org/10.1016/j.eswa.2008.07.006

Bagnall, A., Lines, J., Bostrom, A., Large, J., & Keogh, E. (2017). The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, *31*(3), 606–660. https://doi.org/10.1007/s10618-016-0483-9

Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, *42*(20), 7046–7056. https://doi.org/10.1016/j.eswa.2015.05.013

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307–343. https://doi.org/10.1016/S0304-405X(98)00027-0

Basdekidou, V. A. (2017). The momentum & trend-reversal as temporal market anomalies. *International Journal of Economics and Finance*, *9*(5), 1–19. https://doi.org/10.5539/ijef.v9n5p1

Batten, J. A., Lucey, B. M., McGroarty, F., Peat, M., & Urquhart, A. (2018). Does intraday technical trading have predictive power in precious metal markets? *Journal of International Financial Markets, Institutions and Money*, *52*, 102–113. https://doi.org/10.1016/j.intfin.2017.06.005

Bildik, R. (2001). Intra-day seasonalities on stock returns: Evidence from the Turkish stock market. *Emerging Markets Review*, *2*(4), 387–417. https://doi.org/10.1016/S1566-0141(01)00026-7

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Brooks, C., Hinich, M. J., & Patterson, D. M. (2003). *Intraday patterns in the returns, bid-ask spreads, and trading volume on stocks traded on the New York stock exchange*. Henley Business School, Reading University.

Buchanan, M. (2015). Physics in finance: Trading at the speed of light. *Nature*, *518*(7538), 161–163. https://doi.org/10.1038/518161a

Camino, D. (1996). The role of information and trading volume on intradaily and weekly returns patterns in the Spanish stock market. Working paper available at: https://www.researchgate.net/profile/D avid_Camino/publication/41599927.

Caporale, G. M., Gil-Alana, L., & Plastun, A. (2016). Intraday anomalies and market efficiency: A trading robot analysis. *Computational Economics*, *47*(2), 275–295. https://doi.org/10.1007/s10614-015-9484-9

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L. F., Nobrega, J. P., & Oliveira, A. L. I. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, *55*, 194–211. https://doi.org/10.1016/j.eswa.2016.02.006

Chen, C. (2013). The predictability of opening returns for the returns of the trading day: Evidence from Taiwan futures market. *International Review of Economics and Finance*, *25*, 272–281. https://doi.org/10.1016/j.iref.2012.07.012

Chen, W., & Li, H. (2016). Wavelet decomposition of heterogeneous investment horizon. *Journal of Economics and Finance*, *40*(4), 714–734. https://doi.org/10.1007/s12197-015-9321-y

Chourmouziadis, K., & Chatzoglou, P. D. (2016). An intelligent short term stock trading fuzzy system for assisting investors in portfolio management. *Expert Systems with Applications*, *43*, 298–311. https://doi.org/10.1016/j.eswa.2015.07.063

De Bondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), 793–805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197. https://doi.org/10.1109/4235.996017

Degiannakis, S., & Potamia, A. (2017). Multiple-days-ahead value-at-risk and expected shortfall forecasting for stock indices, commodities and exchange rates: Inter-day versus intra-day data. *International Review of Financial Analysis*, *49*, 176–190. https://doi.org/10.1016/j.irfa.2016.10.008

Dunis, C. L., Rosillo, R., de la Fuente, D., & Pino, R. (2013). Forecasting IBEX-35 moves using support vector machines.

Forecasting IBEX-35 moves using support vector machines. *Neural Computing and Applications*, *23*(1), 229–236. https://doi.org/10.1007/s00521-012-0821-9

Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys (CSUR)*, *45*(1), 1–34. https://doi.org/10.1145/2379776.2379788

Evans, C., Pappas, K., & Xhafa, F. (2013). Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. *Mathematical and Computer Modelling*, *58*(5–6), 1249–1266. https://doi.org/10.1016/j.mcm.2013.02.002

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Gao, L., Han, Y., Zhengzi Li, S., & Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics*, *129*(2), 394–414. https://doi.org/10.1016/j.jfineco.2018.05.009

García-Machado, J. J., & Rybczyński, J. (2017). How Spanish options market smiles in summer: An empirical analysis for options on IBEX-35. *The European Journal of Finance*, *23*(2), 153–169. https://doi.org/10.1080/1351847X.2015.1067634

Gerlein, E. A., McGinnity, M., Belatreche, A., & Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, *54*, 193–207. https://doi.org/10.1016/j.eswa.2016.01.018

Ghalwash, M. F., Radosavljevic, V., & Obradovic, Z. (2014). Utilizing temporal patterns for estimating uncertainty in interpretable early decision making. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD'14, 402–411. https://doi.org/10.1145/2623330.2623694

Ghalwash, M. F., Ramljak, D., & Obradovic, Z. (2012). Early classification of multivariate time series using a hybrid HMM/SVM model. IEEE International Conference on Bioinformatics and Biomedicine, 1–6.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hadavandi, E., Shavandi, H., & Ghanbari, A. (2010). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems*, *23*(8), 800–808. https://doi.org/10.1016/j.knosys.2010.05.004

Harris, L. (1986). A transacion data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics*, *16*(1), 99–117. https://doi.org/10.1016/0304-405X(86)90044-9

Hasbrouck, J. (1991). Measuring the information content of stock trades. The. *Journal of Finance*, *46*(1), 179–207. https://doi.org/10.1111/j.1540-6261.1991.tb03749.x

Hatami, N., & Chira, C. (2013). Classifiers With a Reject Option for Early Time-Series Classification. IEEE Symposium on Computational Intelligence and Ensemble Learning (CIEL), 9–16.

He, G., Duan, Y., Peng, R., Jing, X., Qian, T., & Wang, L. (2015). Early classification on multivariate time series. *Neurocom-Puting*, *149*, 777–787. https://doi.org/10.1016/j.neucom.2014.07.056

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, *124*, 226–251. https://doi.org/10.1016/j.eswa.2019.01.012

Hsu, M., Lessmann, S., Sung, M., Ma, T., & Johnson, J. E. V. (2016). Bridging the divide in financial market forecasting: Machine learners vs. financial economists. *Expert Systems with Applications*, *61*, 215–234. https://doi.org/10.1016/j.eswa.2016.05.033

Hsu, S., Hsieh, J. P., Chih, T., & Hsu, K. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications*, *36*(4), 7947–7951. https://doi.org/10.1016/J.Eswa.2008.10.065

Hu, Y., Liu, K., Zhang, X., Su, L., Ngai, E. W. T., & Liu, M. (2015). Application of evolutionary computation for rule discovery in stock algorithmic trading: A literature review. *Applied Soft Computing*, *36*, 534–551. https://doi.org/10.1016/j.asoc.2015.07.008

Jacobsen, B., Mamun, A., & Visaltanachoti, N. (2005). *Seasonal, size and value anomalies*. Dept. of Commerce, Massey University at Albany. Retrieved from. https://natlib-primo.hosted.exlibrisgroup.com/primoexplore/search?query=any,contains,9911767993502836&tab=catalogue&search_scope=NLNZ&vid=NLNZ&offset=0, https://doi.org/10.2139/ssrn.784186

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702

Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, *56*(2), 699–720. https://doi.org/10.1111/0022-1082.00342

Jiang, F., Tong, G., & Song, G. (2019). Technical analysis profitability without data snooping bias: Evidence from Chinese stock market. *International Review of Finance*, *19*(1), 191–206. https://doi.org/10.1111/irfi.12161

Kadan, O., Michaely, R., & Moulton, P. C. (2018). Trading in the presence of short-lived private information: Evidence from analyst recommendation changes. *Journal of Financial and Quantitative Analysis*, *53*(4), 1509–1546. https://doi.org/10.1017/S0022109018000212

Kara, Y., Acar Boyacioglu, M., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange. *Expert Systems with Applications*, *38*(5), 5311–5319. https://doi.org/10.1016/j.eswa.2010.10.027

Lessmann, S., Baesens, B., Seow, H., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Leung, M. T., Daouk, H., & Chen, A. (2000). Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting*, *16*(2), 173–190. https://doi.org/10.1016/S0169-2070(99)00048-5

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, *1*(1), 41–66. https://doi.org/10.1093/rfs/1.1.41

Malkiel, B. G. (1973). *A random walk down Wall Street [by] Burton G.* Norton.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, *17*(1), 59–82. https://doi.org/10.1257/089533003321164958

Manish, K., & Thenmozhi, M. (2005). *Forecasting stock index movement: A comparison of support vector machines and random forest. In Proceedings of ninth Indian institute of capital markets conference*. Mumbai.

Miralles-Quirós, J. L., Miralles-Quirós, M. d. M., & Daza-Izquierdo, J. (2015). Intraday patterns and trading strategies in the Spanish stock market. *Applied Economics*, *47*(1/3), 88–99. Retrieved from. http://www.econis.eu/PPNSET?PPN=815559151, https://doi.org/10.1080/00036846.2014.962224

Mori, U., Mendiburu, A., Dasgupta, S., & Lozano, J. A. (2018). Early classification of time series by simultaneously optimizing the accuracy and earliness. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(10), 4569–4578. https://doi.org/10.1109/TNNLS.2017.2764939

Mori, U., Mendiburu, A., Keogh, E., & Lozano, J. A. (2017). Reliable early classification of time series based on discriminating the classes over time. *Data Mining and Knowledge Discovery*, *31*(1), 233–263. https://doi.org/10.1007/s10618-016-0462-1

Mori, U., Miranda, I. M., Mendiburu, A., & Lozano, J. A. (2019). Early classification of time series using multi-objective optimization techniques. *Information Sciences*, *492*, 204–218. https://doi.org/10.1016/j.ins.2019.04.024

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, *116*(2), 257–270. https://doi.org/10.1016/j.jfineco.2015.01.003N/(SICI)1099-131X(199705)16:3<165::AID-FOR653>3.0.CO;2-Y

Ortiz, C., de Zárate, O., María, J., & Vicente, L. (2015). New evidence of quarterly return patterns in the Spanish stock market. *Applied Economics Letters*, *22*(13), 1025–1029. https://doi.org/10.1080/13504851.2014.995358

Page, L. (2012). 'It ain't over till it's over.' Yogi Berra bias on prediction markets. *Applied Economics*, *44*(1), 81–92. https://doi.org/10.1080/00036846.2010.498578

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, *42*(4), 2162–2172. https://doi.org/10.1016/j.eswa.2014.10.031

Pérez-Rodríguez, J. V., Torra, S., & Andrada-Félix, J. (2005). STAR and ANN models: Forecasting performance on the Spanish "Ibex-35" stock index. *Journal of Empirical Finance*, *12*(3), 490–509. https://doi.org/10.1016/j.jempfin.2004.03.001

Pfeifer, P. E. (1985). Market timing and risk reduction. *Journal of Financial and Quantitative Analysis*, *20*, 451–459. https://doi.org/10.2307/2330761

Plastun, A., Makarenko, I., Khomutenko, L., Shcherbak, S., & Tryfonova, O. (2019). Exploring price gap anomaly in the Ukrainian stock market. *Investment Management and Financial Innovations*, *16*(2), 150–158. https://doi.org/10.21511/imfi.16(2).2019.13

Qian, B., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence*, *26*(1), 25–33. https://doi.org/10.1007/s10489-006-0001-7

Qiu, M., & Song, Y. (2016). Predicting the direction of stock market index movement using an optimized artificial neural network model. *PLoS ONE*, *11*(5), e0155133. https://doi.org/10.1371/Journal.Pone.0155133

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. *Int J Neural Syst*, *14*(2), 69–106. https://doi.org/10.1142/S0129065704001899

Rechenthin, M., & Street, W. N. (2013). Using conditional probability to identify trends in intra-day high-frequency equity pricing. *Physica A: Statistical Mechanics and its Applications*, *392*(24), 6169–6188. https://doi.org/10.1016/j.physa.2013.08.003

Rossi, M. (2016). The capital asset pricing model: A critical literature review. *Global Business and Economics Review*, *18*(5), 604. https://doi.org/10.1504/GBER.2016.10000254

Schulmeister, S. (2009). Profitability of technical stock trading: Has it moved from daily to intraday data? *Review of Financial Economics*, *18*(4), 190–201. https://doi.org/10.1016/j.rfe.2008.10.001

Sewell, M. (2007). *Behavioural finance* (pp. 1–14). University of Cambridge.

Shiller, R. J. (2000). *Irrational exuberance* (2nd ed.). Princeton Univ. Press.

Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, *264*, 71–88. https://doi.org/10.1016/j.neucom.2016.11.095

Tan, K. C., Khor, E. F., & Lee, T. H. (2005). *Multiobjective evolutionary algorithms and applications (advanced information and knowledge processing)*. Springer-Verlag.

Tanaka-Yamawaki, M., Tokuoka, S., & Awaji, K. (2009). Short-term price prediction and the selection of indicators. *Progress of Theoretical Physics Supplement*, *179*, 17–25. https://doi.org/10.1143/PTPS.179.17

Tay, F. E. H., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega*, *29*(4), 309–317. https://doi.org/10.1016/S0305-0483(01)00026-3

Thaler, R. H. (1987). Anomalies: Weekend, holiday, turn of the month, and intraday effects. *The Journal of Economic Perspectives*, *1*(2), 169. Retrieved from–177. https://search.proquest.com/docview/1300221567, https://doi.org/10.1257/jep.1.2.169

Timmermann, A. (2008). Elusive return predictability. *International Journal of Forecasting*, *24*(1), 1–18. https://doi.org/10.1016/j.ijforecast.2007.07.008

Tkáč, M., & Verner, R. (2016). Artificial neural networks in business: Two decades of research. *Applied Soft Computing*, *38*, 788–804. https://doi.org/10.1016/j.asoc.2015.09.040

Tran, D. T., Iosifidis, A., Kanniainen, J., & Gabbouj, M. (2018). Temporal attention-augmented bilinear network for financial time-series data analysis. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(5), 1407–1418. https://doi.org/10.1109/TNNLS.2018.2869225

Van den Poel, D., Chesterman, C., Koppen, M., & Ballings, M. (2016). Equity price direction prediction for day trading: Ensemble classification using technical analysis indicators with interaction effects. In *In 2016 IEEE Congress on Evolutionary Computation (CEC)* (pp. 3455–3462). Ieee.

Vayanos, D., & Woolley, P. (2013). An institutional theory of momentum and reversal. *The Review of Financial Studies*, *26*(5), 1087–1145. https://doi.org/10.1093/rfs/hht014

Wang, P., Zong, L., & Ma, Y. (2020). An integrated early warning system for stock market turbulence. *Expert Systems with Applications*, *153*, 113463. https://doi.org/10.1016/j.eswa.2020.113463

Wang, Y. (2014). Stock price direction prediction by directly using prices data: An empirical study on the KOSPI and HSI. *International Journal of Business Intelligence and Data Mining*, *9*(2), 145–160. https://doi.org/10.1504/IJBIDM.2014.065091

Wood, R. A., McInish, T. H., & Ord, J. K. (1985). An investigation of transactions data for NYSE stocks. *The Journal of Finance*, *40*(3), 723–739. https://doi.org/10.1111/j.1540-6261.1985.tb04996.x

Xing, Z., Pei, J., & Yu, P. S. (2011). Early classification on time series. *Knowledge and Information Systems*, *31*(1), 105–127.

Xing, Z., Yu, P. S., & Wang, K. (2011). Extracting Interpretable Features for Early Classification on Time Series. Proceedings of the Eleventh {SIAM} International Conference on Data Mining, 247–258.

Ye, L., & Keogh, E. (2009). Time Series Shapelets: A New Primitive for Data Mining. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 947–949.

Ying, Q., Yousaf, T., Akhtar, Y., & Rasheed, M. S. (2019). Stock investment and excess returns: A critical review in the light of the efficient market hypothesis. *Journal of Risk and Financial Management*, *12*(2), 97. https://doi.org/10.3390/jrfm12020097

## AUTHOR BIOGRAPHIES

**I. Marta Miranda García** received the PhD in Economics and Business Administration from the Rey Juan Carlos University Degree in Economics and Business Administration. She is an associate professor at the Rey Juan Carlos University of Madrid (URJC) and Pablo de Olavide University of Sevilla Department of Financial Management, 2014 to present time. She was a professor of the Centre for Advanced Studies Felipe II (UCM), 2000–2016. She was the subdirector of the Cooperative School of the Complutense University of Madrid until 2010. Her main research areas include artificial intelligence applied to the internationalization of business and finance and Social Economy, in general the use of artificial intelligence techniques to solve economic and financial problems.

**María-Jesús Segovia-Vargas** received the BS and PhD degrees in Economics and Business Administration from Universidad Complutense de Madrid, Spain, in 1994 and 2003, respectively. Currently, she is an associate professor at the Department of Financial and Actuarial Economics and Statistics, at Universidad Complutense de Madrid. Her main research topic is focused on the application of operational research methods to analyze financial problems, especially banks and insurance companies' solvency, internationalization success, and social entrepreneurship. She has published several monographs and papers in internationally renowned journals. She has also led and participated in several research projects.

**Usue Mori** is a researcher and PhD at University of UPV/EHU. She received the MSc degree in mathematics and the PhD degree in computer science from the University of the Basque Country UPV/EHU, Spain, in 2010 and 2015, respectively, where she has been working as a lecturer with the Department of Computer Science and Artificial Intelligence since 2019. She has also participated in several research projects. She has published several monographs and articles in internationally renowned journals. Her main research interest includes time series classification, early classification of time series, time series clustering, time series distance measures, and traffic modeling.

**José Antonio Lozoano** received the PhD degree in computer Science from the University of the Basque Country in 1998. He is currently the scientific director with the Basque Center for Applied Mathematics and a full professor in computer science and artificial intelligence with the University UPV/EHU. His research interests include machine learning, heuristic optimization, and its application to different scenarios. He is a member on the editorial board of the main journals of his scientific fields, such as the IEEE Transactions on Neural Networks and Learning Systems and IEEE Transactions on Evolutionary Computation.

**How to cite this article:** Miranda García, I. M., Segovia-Vargas, M.-J., Mori, U., & Lozano, J. A. (2022). Early prediction of Ibex 35 movements. *Journal of Forecasting*, 1–17. https://doi.org/10.1002/for.2933