International Review of Financial Analysis 88 (2023) 102657

Contents lists available at ScienceDirect

International Review of Financial Analysis

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

Data vs. information: Using clustering techniques to enhance stock returns forecasting

Javier Vásquez Sáenz [a], Facundo Manuel Quiroga [a,b], Aurelio F. Bariviera [c,*]

[a] Facultad de Cs. Exactas, Universidad de Buenos Aires, Buenos Aires, Argentina
[b] Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata, 50 y 120, La Plata, 1900, Buenos Aires, Argentina
[c] Universitat Rovira i Virgili, Department of Business, ECO-SOS, Av. Universitat 1, Reus, 43204, Tarragona, Spain

ARTICLE INFO

Keywords:
Stock price forecast
Clustering
Financial Reports
Deep learning
Investment algorithms
Trading

ABSTRACT

This paper explores the use of clustering models of stocks to improve both (a) the prediction of stock prices and (b) the returns of trading algorithms.

We cluster stocks using k-means and several alternative distance metrics, using as features quarterly financial ratios, prices and daily returns. Then, for each cluster, we train ARIMA and LSTM forecasting models to predict the daily price of each stock in the cluster. Finally, we employ the clustering-empowered forecasting models to analyze the returns of different trading algorithms.

We obtain three key results: (i) LSTM models outperform ARIMA and benchmark models, obtaining positive investment returns in several scenarios; (ii) forecasting is improved by using the additional information provided by the clustering methods, therefore selecting relevant data is an important preprocessing task in the forecasting process; (iii) using information from the whole sample of stocks deteriorates the forecasting ability of LSTM models.

These results have been validated using data of 240 companies of the Russell 3000 index spanning 2017 to 2022, training and testing with different subperiods.

1. Introduction

There is currently a consensus that the stock market is a complex dynamic system, with some variables that are difficult to measure and have great sensitivity to unpredictable events that can cause disruptions. This complexity is increased by the emotional component of many market actors, which amplify underlying trends and can generate market runs. As a consequence, the traditional Efficient Market Hypothesis (EMH) began to be superseded by a more realistic Adaptive Market Hypothesis (AMH), proposed by Lo (2004). Under such framework, markets have (slight) time-varying departures from the theoretical random behavior that, if properly detected and exploited, can be used to generate profitable trading strategies.

In search of this advantage, numerous mechanisms have been generated to predict markets' behavior and consequently find investment strategies with a higher than average return (Batten et al., 2018). Thus, a plethora of quantitative and qualitative models that have been developed to predict market trends and prices, such as statistical, technical, morphological or graphic indicators. Some of these methods have a methodological basis, while others obey more to superstitions (Dyl & Maberly, 1988). Among the most popular methods are moving averages, exponential moving averages, temporal price structures, daily volume ratios or methods that determine, based on graphic patterns, the possibility of a rise or fall in the price of a given stock. More sophisticated methods use correlations between different economic indicators with groups of stocks, such as the price of oil, oil production, and energy stocks prices.

A definitive answer regarding market behavior has not yet been found, giving rise to the development of new forecasting methods, thus, an active line of research in quantitative finance. In particular, the use of machine learning techniques could bypass certain limitations of traditional econometric techniques. For example recently, Amini et al. (2021) detected that returns on major stock market indices exhibit non-linear dependencies. Therefore, using linear models (e.g. ARIMA) may produce flawed predictions due to biased coefficient estimates.

Machine learning techniques can also use various sources of information to predict the price of a stock, including information of the same stock, other stocks, and other market indicators. While there are many sources and types of information that can be added, a major concern in these approaches is that the information added should be relevant to avoid overfitting the prediction model, but also help to improve the performance of investment strategies that use the predictions of

* Corresponding author.
E-mail address: aurelio.fernandez@urv.cat (A.F. Bariviera).

https://doi.org/10.1016/j.irfa.2023.102657
Received 2 November 2022; Received in revised form 22 March 2023; Accepted 11 April 2023
Available online 22 April 2023

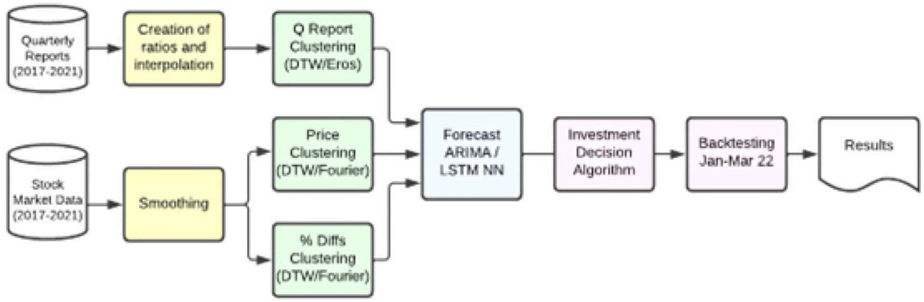

Fig. 1. Diagram of our combined clustering, prediction and trading model. We use data both from quarterly reports and stock prices time series to cluster stocks. These are clustered independently since they require specialized distance measures, using K-Means. Using the clustering information, we train several forecasting models based on ARIMA and LSTM. Finally, we use these predictions to back-test a trading algorithm.

the model to obtain a profit. This precludes the possibility of simply using all available information to train and use a stock price prediction model.

In this work we contribute a novel approach to modeling market behavior based on selecting relevant information to predict the price of stocks. Our approach explores the hypothesis that using information of similar stocks, determined via clustering, can improve the results of a prediction model, compared to a prediction model that does not take into account such cluster-derived data.

In our case, we cluster stocks according to financial ratios (built from the quarterly financial reports), to their prices or to their daily returns. Even though all these data are given as time series, they are different in nature. Consequently, we explore several non-trivial features and distance measures to perform the clustering. The different distance measures compared were based on Dynamic Time Warping (DTW), Fourier decomposition and Extended Frobenius Norm (EROS) (Yang & Shahabi, 2004).

With these clustering results, we train specific prediction models for each cluster to isolate the dynamics of specific "submarkets". In this way, we select relevant submarket information that can complement the information of a single stock, avoiding using either all the information available, or only the information of the stock itself. Fig. 1 shows a diagram of our proposed pipeline.

To validate our approach, we compare both statistical and deep learning models to predict prices and returns for a representative set of data from the Russell 3000.[1] We also compare the results against some naïve strategies (perfect forecast, buy-and-hold and random strategies) and the well-known Moving Average Convergence/Divergence oscillator (MACD). Our experiments show that this approach can be used in combination with an investment strategy to obtain positive returns in real world scenarios.

We contribute to the literature in the following aspects: (i) we propose clustering financial reports (by means of time series of financial ratios) in order to improve stock price and return forecasts; (ii) we provide evidence that deep learning methods offer an advantage over traditional linear forecasting models; (iii) we uncover that stock-related data, rather than market-wide data, is better at enhancing forecast accuracy and performance.

The article is organized as follows. Section 2, reviews previous approaches for representing, clustering and predicting stock information. Then, Section 3 describes the dataset and features used, as well as the clustering, prediction and investing methods tested. Section 4 presents the evaluation of the methods on a subset of 240 stocks from the Russell 3000 index, and discusses the results. Finally, Section 5 summarizes the conclusions of our work and delineates future research paths.

[1] https://www.ftserussell.com/products/indices/russell-us

2. Background

The behavior of the stock markets has been a subject of debate for more than 100 years. The seminal paper by Bachelier (1900) could be considered the first theoretical attempt to model bond prices. However, it was not until the middle of the 20th century that formal asset pricing models were developed. Osborne (1959, 1962) find evidence that price motion follows a simple random walk. Shortly after, Samuelson (1965) demonstrates that properly anticipated prices should follow a random walk. Such situation became crystallized in the Efficient Market Hypothesis (EMH), which, as defined by Fama (1970), describes a market where prices fully reflect all available information. Informational efficiency is classified into three broad categories: (i) weak efficiency corresponds to market where prices reflect the information contained in past prices; (ii) semi-strong efficiency reflects a market where prices incorporates all public information (beyond prices); (iii) strong efficiency affirms that prices reflect all public and private information. The EMH remained unchallenged until the 1980s, when data and computational power became more easily available. Then, empirical papers found inconsistencies between empirical regularities and the theoretical models.

In this section we will briefly review the literature around the three pillars on which this paper is based: clustering methods for stocks, general prediction models, and combined clustering/prediction models.

2.1. Market data clustering

In search of indirect sources of information that would enable better price predictions, some researchers have studied the clustering of accounting or sentiment data from various stock markets and the companies listed there.

The additional information explored includes fundamental data of each company, either as static data (Dhakal, 2019), as a time series (Basalto et al., 2005; Nair et al., 2017) or as quarterly financial reports (Lee et al., 2010). Other alternative data sources include sentiment inferred from forums and news regarding the companies (Li & Wu, 2021).

The first clustering attempts used different indicators by risk rating agencies, such as Standard & Poor's, Moody's or Fitch Rating. Even nowadays, these agencies classify companies listed on the stock exchanges according to different criteria, including risk rating, sector, type of company and other indicators that can be used to distinguish companies.

Among the works using static data for clustering, Dhakal (2019) utilizes balance sheet magnitudes to associate companies in order to study whether stock price growth is significant in any of the clusters found. Babu et al. (2012) have an interesting approach, converting the financial reports into feature vectors and clustering them to determine the short term variation of the stock after the financial report is released and examines the diverse properties of several clustering methods. In different research line, Gubareva and Borges (2022) use clustering to identify periods of the regime switching in interest rates.

Clustering of static data is a widely explored topic and there are very efficient dimensionality reduction and grouping techniques. However, clustering of time series is not as mature, especially with multivariate time series. The result of these clusterings indicates a favorable contribution in trend and price predictions, as analyzed by Lee et al. (2010), Nair et al. (2017) and Wang et al. (2021).

There are several methods for clustering time series, either by shape similarity, by analyzing principal components or by point-to-point comparison, as described in Bagnall et al. (2017) and Aghabozorgi et al. (2015). In particular it is worth to mention the use in the current paper of the Extended Frobenius Norm (EROS), proposed by Yang and Shahabi (2004), to cluster multivariate time series. This method employs Principal Component Analysis (PCA) to reduce the dimensionality of the multivariate time series, and then uses the Frobenius Norm to compare time series, achieving better results than simply comparing the original multivariate series with the Dynamic Time Warping (DTW) distance.

2.2. Forecasting models

A similar scenario is found for stock price or trend prediction models, either with time series or static data, which usually refer to the latest balance sheet data compared to the results observed in other similar companies. Among the models used are both univariate and multivariate statistical methods such as VARIMA (Vector Autoregressive moving average) or Prophet (Fang et al., 2019; Samal et al., 2019; Taylor & Letham, 2018; Yenidoğan et al., 2018) or methods based on gradient boosting, decision trees and neural networks (Krauss et al., 2017).

The use of neural networks to predict prices and trends of stocks is a topic with abundant literature, with techniques reaching different levels of accuracy. For example, the approach by Liu and Motani (2021) combines convolutional recurrent and generative adversarial networks, whereas Hadavandi et al. (2010) propose the use of genetic algorithms to refine predictions. In a similar vein, Lin et al. (2022) combines market and sentiment data for stock market forecasting.

There is debate as to whether statistical methods are superior to neural networks in terms of prediction accuracy. Some researchers, such as Yamak et al. (2019), forecasting Bitcoin, hold this position. Others refute it, such as Siami-Namini et al. (2018) working with several financial indexes, or Adebiyi et al. (2014) and Ma (2020) both using Dell's stock price.

2.3. Forecasting models combined with clustering

Combining clustering with various forecasting methods is rising as a separate research line. In this sense, Hadavandi et al. (2010) propose clustering the data of the time series of stocks using self-organized maps (SOM) and then feeding this data into a genetic algorithm, with good results in terms of the accuracy of the predictions *vis-à-vis* other methods. In a similar vein, other authors (Bandara et al., 2020; Li et al., 2022; Tadayon & Iwashita, 2020) adopt distance based clustering methods such as DTW or feature clustering methods to create features. Afterwards, they use characteristics of the time series such as entropy, stability, autocorrelation of fast Fourier transform coefficients or continuous wavelet transform coefficients, as well as the time series, to feed a neural network. In these cases, the authors find that the features added by clustering improve the predictions made by the neural networks, when compared to models without clustering.

Long et al. (2020) exploits the transaction records and public market information for each stock, using the transaction records to identify and cluster investment patterns and adding this information to the price and volume data to predict through neural networks based on CNN and Bidirectional LSTM to predict price movement, getting a very good prediction accuracy.

Finally, Wang et al. (2021) propose clustering the time series of prices using the distance of morphological similarity and k-means to find different types of series. Afterwards, they train different neural networks for each type of series. This method is the most similar to our own approach, but our work differs mainly from these approach in that we include financial information as part of the clustering scheme, and also experiment with various combinations of clustering, forecasting and investment models to fully characterize the impact of the clustering scheme.

3. Methods

In this section, we describe how we obtain the dataset used in our experiments, as well as the clustering, prediction and investment models we simulate in Section 4.

3.1. Obtaining and preparing the dataset

The dataset employed in this paper is based on the constituent list of the Russell 3000 index as of 01-Mar-2021, which contains the list of the 3000 companies with the highest market value.

Due to computational limitations, we randomly sampled 300 stocks from the Russell 3000. After verifying the availability of all quarterly reports and that stocks actually traded in the market during the period 01-Jan-2017 to 31-Mar-2022, our final dataset is composed by 240 stocks. Then, we gather the historical data of prices and reports for the indicated period from Yahoo Finance[2] and Alpha Vantage.[3]

The time series of prices and returns are simple to analyze since they are already adjusted for splits or consolidations of the shares. To avoid high frequency oscillations, we slightly smoothed the series with a moving average over a 5-days period. Fig. 2 shows how we employ the dataset to train and test the models. Note that the period used to train the clustering models does not overlap with any forecasting period (F) to avoid leaking information from the training set.

Quarterly balance sheet series values, however, differ markedly according to the company and the quarter, which made it difficult to compare them. Therefore, we sought to identify financial ratios that would standardize and more adequately represent the financial statements of companies, allowing for cross-comparisons. Based on the work of Chen and Shimerda (1981) who analyzed the usefulness of more than one hundred accounting and financial ratios, and the analysis of Wang and Lee (2008) who used the clustering of various financial ratios to identify the most representative ones, we chose ten ratios that met these conditions. Thus we generated quarterly time series of ratios for the five years between 01-Jan-2017 and 31-Dec-2021. However, given that the quarterly closings of the companies do not necessarily match, we transformed these series into monthly series, interpolating the missing points by means of cubic splines as can be seen in the example shown in Fig. 3.

3.2. Stocks clustering

We clustered the companies in our sample in clusters according to either prices, daily returns or financial ratios. The aim was to generate clusters of companies that share similar characteristics for each method.

Afterwards, we used these clusters to augment the data used for each stock to train a neural network or ARIMA model by adding the closing prices or daily returns of the rest of the stocks in the respective cluster.

[2] https://finance.yahoo.com/

[3] https://www.alphavantage.co/

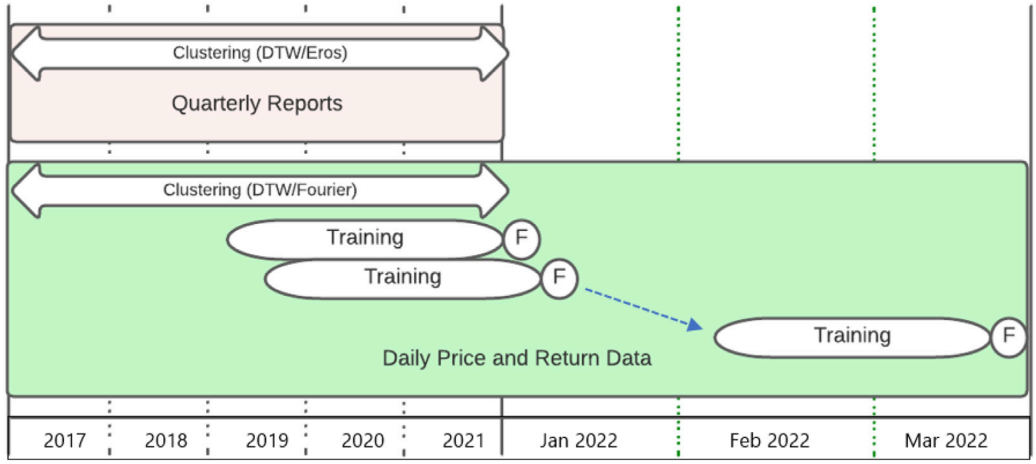

Fig. 2. Temporal distribution of our data that includes stock data from 2017 to 2022. We use a subset of data from 2017 to 2021 for clustering. Then, we use the remaining data to train daily models, so that for each day a window of N previous days was used to train the model, and the next day to test the forecast (F) of the daily model. The forecast of the value for the next day was performed with ARIMA and LSTM methods.

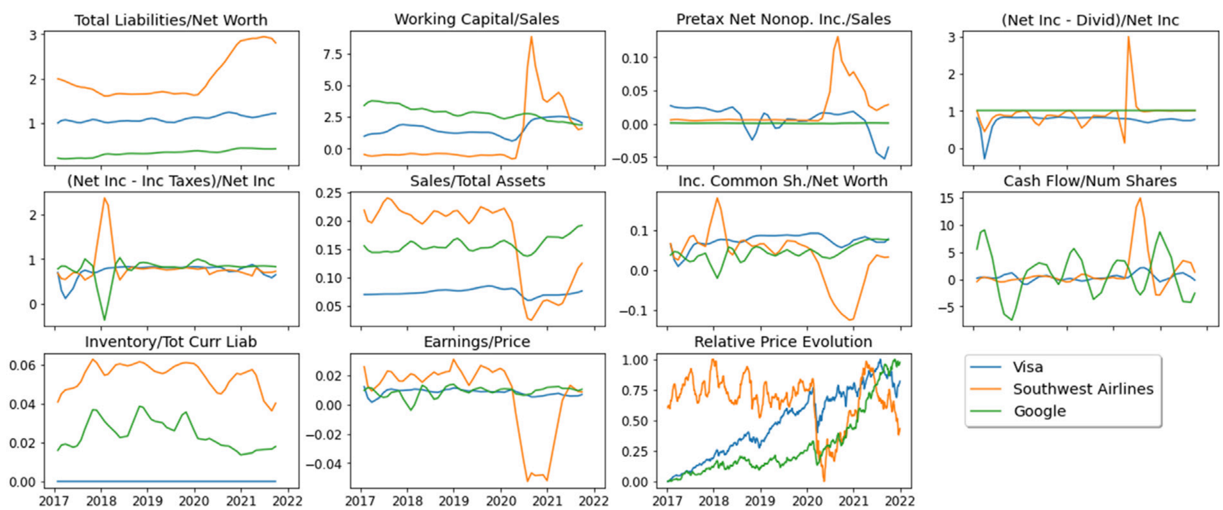

Fig. 3. Sample data from three stocks in the sample (Visa, Southwest Airlines and Google) during the 2017–2021 period. We show the values of the ten ratios used for clustering quarterly reports, and their relative price evolution.

In the case of prices and daily returns, we experimented with two different distance measures for the clustering. The first distance used is DTW, based on curve similarity and the other one is based on the euclidean distance between features of the series, represented as Fourier coefficients. In the case of the financial statements, due to the different nature of this data, we experimented with clustering using DTW but also EROS (Yang & Shahabi, 2004), which characterizes the relationship between the different features of the time series.

3.2.1. Price and daily return series

We used the daily closing stock prices for the period 2017–2021, obtaining 1250 data points for each stock. The clustering of this data is simpler since the clustering of univariate time series is well analyzed. In this case we explored the clustering of two different time series: closing prices and daily returns, which we also included as features since they can remove trends in the data. In Fig. 3 it is possible to observe an example of the normalized trend for an easier comparison.

The series of closing prices and daily returns were clustered using k-means with two different distance metrics: (a) DTW as distance between time series of prices and returns; and (b) the Euclidean distance on the 100 most relevant Fourier Transform real and imaginary coefficients.

In all cases, we optimized the number of clusters measuring Silhouette's coefficient for k between 2 and 30 and choosing a balance between the number of clusters and the value of the Silhouette coefficient. The goal was to set a number of clusters that provides a good differentiation of groups, but having a reasonably low value of k.

3.2.2. Financial ratios series

Note that although the clustering of time series based on prices has been attempted before in various occasions, the clustering of time series based on quarterly financial reports is not as frequent and was analyzed in a limited fashion by Tupe (2014) and Marvin (2015), among others. Therefore, we expanded this approach by using a set of more representative distance measures for these ratios.

The series of the ten financial ratios have 58 monthly points for the period 2017–2021. We explored not only the relationship between the equivalent series of each stock of interest (for example, Total Sales to Assets for each stock), but also the relationship among the different series of ratios for the same stock. In other words, our clustering is based in both inter-stock and intra-stock comparison of ratios. Therefore, we used two different clustering methods to capture different sets of information, either in the similarity of curves or in the relationship between the time series.

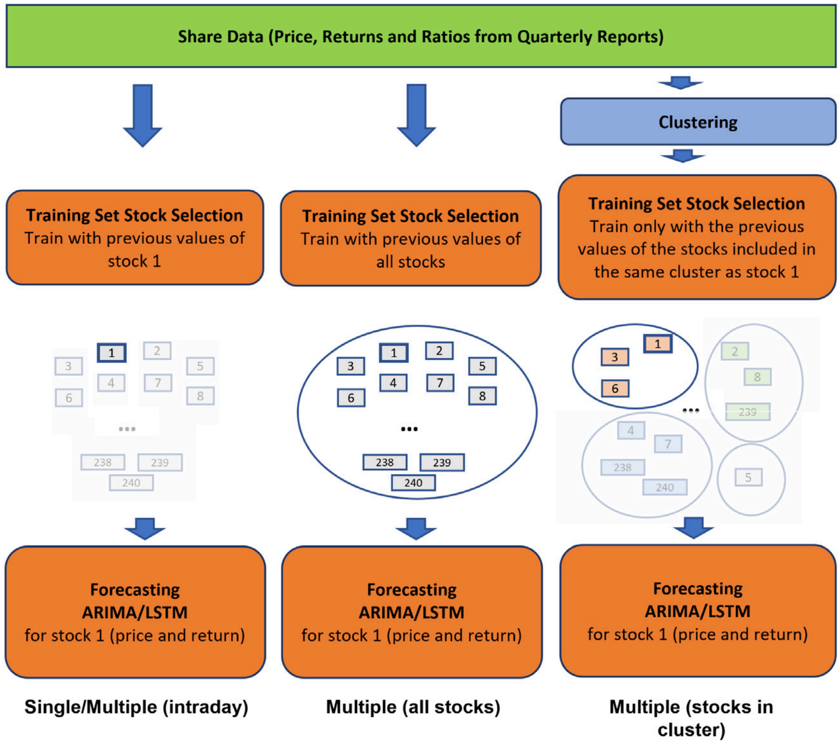

Fig. 4. Data usage of the different ways a forecasting model may use existing data, for a single stock named stock 1. In the first case, we train with only the previous values of the stock for previous timesteps. In the second case, we use the previous data of *all* stocks to train the forecast model. The third and last case, our proposed approach, uses the previous values of all stocks that are in the same cluster as Stock 1.

The first method, based on Dynamic Time Warping (DTW) (Berndt & Clifford, 1994), obtains a distance that stems from the correlation of the shapes of the curves, considering possible displacements or compressions in time. We used a combination of the DTW distances of each of the ten different financial ratios series, inducing a weighted distance between stocks. Using this distance, we clustered the stocks with k-means, optimizing the number of clusters via the Silhouette's coefficient. However, in some cases we detected that the combination of distances between the different series does not necessarily allow obtaining a good correlation between the series that can provide interesting information. As a consequence, we also used a second distance method. It is called Extended Frobenius Norm (EROS) and was proposed by Yang and Shahabi (2004). The measure is based on a combination of the results of the Principal Component Analysis and the Frobenius Norm to compare time series. This method allows taking the series of ten ratios and obtaining the distance between the set of series of each stock. As k-means does not support arbitrary distance/similarity measures, we employed k-medoids when using EROS, which gives similar results to k-means. In this way we could compare two forms of clustering that reflect different properties of the data.

3.2.3. Summary

The result of these tasks generated five different clustering schemes, two for the price series, two for the financial ratios and one for daily returns, as can be seen in Table 1 according to the type of distance used.

Table 1
Distances used for each type of time series. We use DTW in for all types of data. We also experiment with Eros given the large number of financial ratios features and the need to model their interdependence through time. We also use a Fourier-based distance to compare daily prices, modeling them in terms of their high and low oscillation components.

Time series	Distance		
	DTW	Eros	Fourier
Financial ratios	X	X	
Prices	X		X
Daily returns	X		

3.3. Forecasting methods

The clusters generated in the previous step were used to execute different forecasting methods applied to the prices and to the returns time series. We use the data from 01-Jul-2019 to 31-Dec-2021 as a training set to predict prices and returns for the 62 market trading sessions between 03-Jan-2022 to 31-Mar-2022. We run the scenarios using a sliding window so we added the closing or return data of the previous day to our historical set to predict the following day, as shown in Fig. 2.

We design several scenarios (depicted in Fig. 4), using unclustered and clustered data to predict the closing price of each stock. These scenarios describe the training data (stocks and features) used to train a model to predict the closing price or daily returns of a specific stock:

- Forecasting using unclustered data:
  - Single: Uses only the closing prices or daily returns of the same stock we predict.
  - Multiple (intraday): Uses the opening, closing, high, low and volume data or the daily changes of the same stock we predict.
  - Multiple (all stocks), all stocks: Uses the closing data or daily returns of all 240 stocks.
- Forecasting using clustered data:
  - Multiple (stocks in cluster): Uses the closing prices or daily returns of the stock we predict, and the closing prices or daily returns of the rest of the stocks of the same cluster.

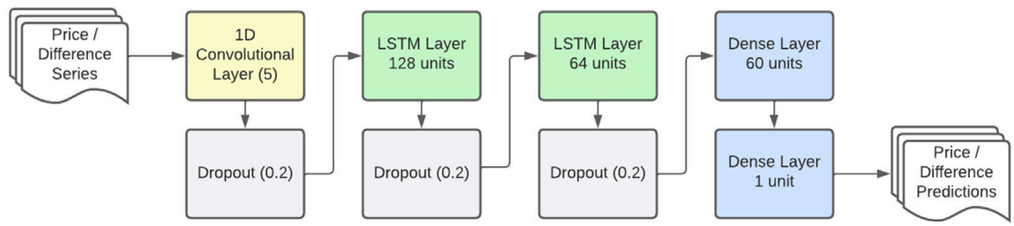

Fig. 5. Architecture of the CNN-LSTM neural network used for the forecasting of prices and returns. The network is a simple stack of 1D convolutional layers to compute 5 features from the input data, which are afterwards processed temporally by two LSTM layers and a Dense layer and finally the price prediction for the next day is performed by a Dense layer.

The first two scenarios were executed only with the data of each stock, either in Single, with only the closing prices or daily returns of the stock or Multiple, with the opening, closing high, low and volume data or daily changes. The third scenario consists in the closing prices or daily returns for all the stocks in the dataset. The fourth scenario is actually a set of different scenarios. They add, in addition to the closing price or daily return of the stock to forecast, the closing data or daily returns of the rest of the stocks in the same cluster according to each clustering method, which is the key proposal of the current paper. The rationale is to assess if the addition of data of those stocks (allegedly related according to the clustering process) enhances the forecasting capability.

We run these scenarios by means of ARIMA and Long Short-Term Memory neural network (LSTM) models for each of the different sets of features. The selection of the ARIMA model for each stock was selected using the Akaike information criteria. The architecture of the neural network is described in Fig. 5. We use a simple network to ensure that the results are driven mainly by the features and the clustering and, to a lesser extent, by the modeling capacities a complex network. The network has a typical design composed of a convolutional layer to extract features from the input data, two recurrent LSTM layers to encode the temporal relationship between the values in the time series, and two dense layers to generate the prediction of the price and return for the following day.

3.3.1. Evaluation with MASE and sMAPE

In order to compare results obtained by the different forecasting methods, it is common practice to use some metric, to assess how well each method fitted with the actual results. The mean squared error (MSE) and the mean absolute error (MAE) can be hard to interpret given their direct relationship to the scale of the values being predicted, and that different stocks can have vastly different price scales. Therefore, we use two metrics that overcome these limitations: (i) the Mean Absolute Scaled Error (MASE) and (ii) the Symmetric Mean Absolute Percentage Error (sMAPE). MASE, defined in Eq. (1) is a scale free error measure that compares the error with the average error. sMAPE, defined in Eq. (2), is a measure that present the errors in terms of percentages (Goodwin & Lawton, 1999).

$$MASE = \frac{1}{k}\frac{\sum_{t=1}^{k}|Y_t - \hat{Y}_t|}{\frac{1}{n-m}\sum_{t=m+1}^{n}|Y_t - Y_{t-m}|} \tag{1}$$

$$sMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{2 * |y_i - \hat{y}_i|}{|y_i| - |\hat{y}_i|} \tag{2}$$

MASE compares the error of the predicted series, against a naïve prediction. In our case, this prediction is simply the closing value of the previous day. This metric tells how much better or worse our prediction is compared to the naïve case, with values under 1 showing a better prediction than the naïve case. sMAPE, on the other hand, indicates the percentage deviation of the fitted value from the actual data, so the benchmark is not defined by the researcher as with MASE. Table 2 displays the results obtained with these two metrics.

3.4. Investment strategy

We used as the starting point an equally weighted portfolio, independent from each other, reinvesting the gains or losses for the next transaction and allowing fractional investments for simplicity. Therefore we tracked the results of 240 individual funds and averaged the returns of each one at the end of the period.

Our strategy produces buy, sell or do nothing signals for each stock on a daily basis, and the holding period spans from opening to close daily.

The trading algorithm triggers a signal for each market day t and each stock between 01-Jan-2022 to 31-Mar-2022 according to the following rules:

- Buy (Long buy) if $close_{t-1} < predicted_t$ and $open_t < predicted_t$.
- Sell (or short sale) if $close_{t-1} > predicted_t$ and $open_t > predicted_t$.
- Do nothing, when none of the above conditions is met.

After daily signals are computed, the trading algorithm executes one of the following transactions for each stock in the sample and calculates the return on each trade as follows:

- Buy (Long Buy): The return between the *closing and opening prices* for the stock is applied to the amount invested in given stock.
- Sell (or Short Sale): In this case the return between the *opening and closing prices* is applied to the amount invested in that stock.
- Do nothing signal: No trade is executed for that particular stock on that day. Consequently, there is no return.

We benchmarked these results with two common stock market trading strategies among small traders, and two baseline methods:

(a) Buy and Hold (B&H): This is the simplest strategy possible, where a trader buys some stocks at the beginning of the holding period and sells them at the end of the period, recording the profit or loss.
(b) Moving Average Convergence Divergence (MACD): This technical analysis method was devised by Gerald Appel in 1979 and is very popular among amateur traders. It is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price. It calculates the difference between an instrument's 26-day and 12-day exponential moving averages (EMA) and produces entry and exit signals for trades (Appel & Dobson, 2007).
(c) Random purchase: This strategy creates random buy and sell signals for daily operations (using a pseudorandom number generator), producing 25% long buys, 25% short sales and 50% of do nothing for each share, irrespective of the actual behavior of the stock. This method is used to find any bias that may be generated by the custom method when feeding these signals generated into the same trading execution algorithm.
(d) Perfect forecast: This is the return obtained when the trader has a perfect knowledge of the future. Thus, trade forecasts are 100% right. This strategy shows the maximum profit that may be attained in the period under study, with the same constraints and initial budget.

Table 2
Results of the forecasting methods applied on prices and on returns, showing the number of trades executed, the percentage of successful forecasts, and the average error (sMAPE, MASE) of each prediction generated with ARIMA or LSTM. The last 5 methods on each section of the table belong to Multiple (stocks in cluster) named as the clustering method, indicating the different data/distance combination used.

Method	Forecast							
	ARIMA (Prices)				ARIMA (Returns)			
	# Predictions	% Success	sMAPE	MASE	# Predictions	% Success	sMAPE	MASE
Single	6124	39.5%	1.994	1.009	14000	48.5%	11.726	6.518
Multiple (intraday)	8261	59.7%	1.864	0.939	13625	48.0%	12.393	7.199
Multiple (all stocks)	12861	49.0%	23.058	17.769	13968	48.4%	11.626	24.239
Eros Q.Report	8890	41.6%	3.671	1.936	14002	48.4%	11.818	6.554
DTW Q.Report	8466	39.4%	2.476	1.232	13990	48.5%	11.834	6.567
DTW Prices	8562	40.2%	2.642	1.340	13984	48.5%	11.736	6.442
Fourier Prices	8428	40.0%	2.236	1.121	13980	48.5%	11.805	6.561
DTW Returns	9197	43.0%	5.310	3.185	14008	48.4%	14.618	7.256
Method	Forecast							
	LSTM (Prices)				LSTM (Returns)			
	# Predictions	% Success	sMAPE	MASE	# Predictions	% Success	sMAPE	MASE
Single	13473	49.0%	4.725	2.651	14385	48.4%	12.191	6.623
Multiple (intraday)	13553	50.4%	5.502	3.166	N/A	N/A	N/A	N/A
Multiple (all stocks)	14343	48.8%	89.399	167.123	14387	49.2%	16.381	7.732
DTW Q.Report	13913	51.0%	9.397	5.18	14390	49.1%	14.901	7.373
Eros Q.Report	13858	50.8%	8.358	4.707	14387	49.1%	15.795	7.71
DTW Prices	13951	51.2%	9.469	5.372	14390	49.2%	14.877	8.161
Fourier Prices	13896	51.0%	9.269	5.225	14392	49.1%	14.71	7.348
DTW Returns	10680	44.3%	5.116	2.615	14389	49.3%	15.439	7.609

Finally, to simplify the evaluation, the calculation of the returns does not take into account transaction costs, which usually are small for long transactions, but can be substantial compared to the return obtained in the case of shorting stocks.

4. Results and discussion

The results of the different clustering methods for the quarterly reports are not as intuitive as the results of the clustering of the price time series. To validate that the clustering methods would return reasonable clusters in the five methods used, we identified pairs of stocks that *a priori* should be closely related to each other. For example, the two shares that Google has on NASDAQ (GOOG and GOOGL) (that trade in identical manner with the same financial returns), and Mastercard (MA) and Visa (V) (stocks that have very similar parameters) were assigned by all clustering methods to the same clusters, respectively. Such outcomes follow the financial intuition and give a good indication of the soundness of the clustering methods.

4.1. Results

We measure the goodness of the forecasts using two broad measures. One is the accuracy, i.e. the ratio of the correct transaction classification to the number of all transactions. The other is the performance, which is the return generated on the right classifications minus the return generated on the wrong classifications. We also provide the returns of the four baseline investment strategies.

Additionally we calculate the forecast accuracy of each method for those transactions effectively executed, either short or long. Table 2 displays the accuracy, sMAPE and MASE of the forecasting models.

With the sole exception of the Multiple (intraday) model for prices in ARIMA, the accuracy is not significantly higher than that of a random prediction, and in many cases substantially worse. The sMAPE and MASE values generally indicate a substantial deviation from the naïve case, where the price or return of the previous day is taken as a prediction for the following day, improving it only in the case of the Multiple (intraday) model for prices in ARIMA.

However, the situation changes substantially when we back test the trading recommendations (buy and sell orders) and measure them in terms of returns, whose results are summarized in Table 3. First, we observe that regardless of the forecasting method, results are very far from a perfect forecast. It means that market movements are mostly unpredictable, but they are not fully random. Second, the empirical simulations reflect that all forecasting methods (except ARIMA multiple intraday for prices) outperform all the baseline methods. Third, and more importantly, those methods that used targeted augmented information to predict the prices had an average return larger than the methods that used only the data available for each stock. On the contrary, if we use data from the 240 stocks in the sample (Multiple — all stocks in Table 3) results are worse than other methods (except for ARIMA forecasting with prices). In a certain way, this situation indicates that more data is not necessarily better. Therefore, including a clustering technique when preparing data may bring two benefits: to improve forecasting results and to reduce computing time.

In agreement with part of the literature, we find LSTM forecasting ability generally outperforms that of ARIMA. The reason for such outcome could be due to the better suitability of neural networks to deal with nonlinearities and thus, to extract additional information, as observed in average returns shown in Table 4. Finally, price forecasts generally give better results than return forecasts as it is shown in Table 3.

Fig. 6 shows the dynamic evolution of returns obtained by the different forecasting methods. On the left panel we select the best four methods (best two ARIMA and best two LSTM), and display also some baseline methods for comparison. The right panel exhibits the remaining simulations with ARIMA and LSTM as a reference. We can observe a dominance of the neural network methods over the ARIMA counterparts for the holding period. Note that most of the forecasting methods show a clear jump in the last week of February 2022. Around this date Russia invaded Ukraine, producing a major escalation in the Russia–Ukraine conflict. Therefore, the forecasting methods seem able to capture such market stress, and take advantage of it.

To have a better understanding of the behavior of our back-testing, we aggregated the results as regards all transactions executed considering the best two methods for each forecasting model, and compared it with the returns obtained through a random strategy. As a result, we find a small average return with low standard deviations around the mean return, with a greater average when running LSTM than ARIMA. However, considering the large amount of transactions (we obtain up to one trading signal for each stock each day of the sample period), they accumulate and create the variations observed in Fig. 6. It

Table 3
Three month returns for each forecasting method. The first two methods correspond to the perfect trading and a random trading. The second two, to popular trading method (B&H and MACD). The last eight methods use the forecasts we generated through ARIMA and LSTM, using either (i) each stock's own data (Single), (i) all stocks data (Multiple) (iii) or only the stocks in the same cluster as the stock to predict (our approach).

Baseline Methods	Return			
Perfect forecast	241.18%			
Random purchase	1.38%			
Buy and Hold	−2.04%			
MACD	−3.02%			
Forecasting methods	**ARIMA (prices)**	**ARIMA (returns)**	**LSTM(prices)**	**LSTM (returns)**
Single	3.33%	4.32%	3.17%	3.57%
Multiple (intraday)	−1.83%	3.13%	6.92%	N/A
Multiple (all stocks)	6.19%	4.21%	3.82%	4.09%
Cluster + DTW Q.Report	4.64%	4.42%	8.13%	3.35%
Cluster + Eros Q.Report	5.73%	4.75%	9.09%	3.31%
Cluster + DTW Prices	4.68%	4.51%	10.24%	3.51%
Cluster + Fourier Prices	4.04%	4.56%	8.37%	5.37%
Cluster + DTW Returns	5.89%	4.37%	3.99%	4.78%

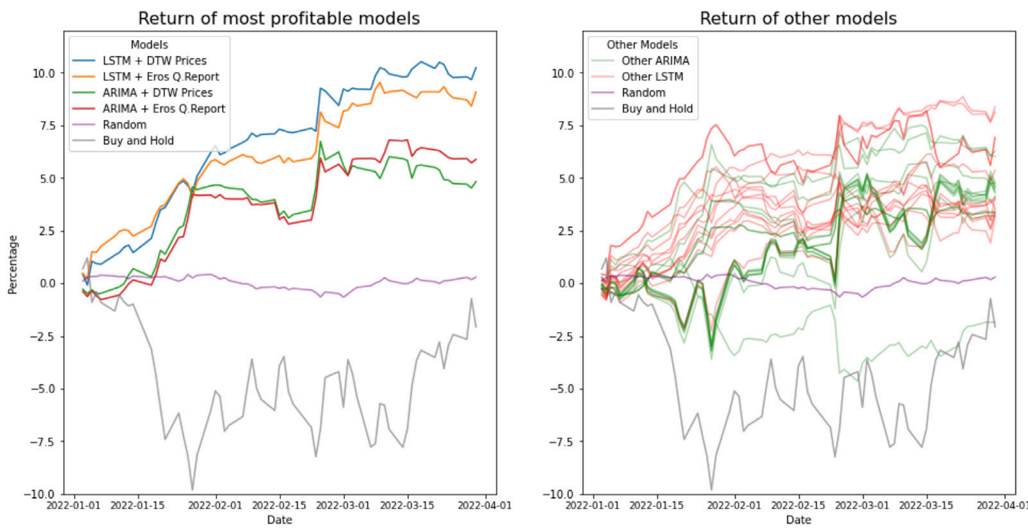

Fig. 6. Evolution of returns for the test set for the first quarter of 2022 for several methods. The left panel displays only the results of the most profitable methods for clarity. The right panel features the remaining methods colored according the forecasting method category. In both cases, random purchase and the buy and hold investment strategy is included as a reference.

Table 4
Descriptive statistics of the *returns* obtained using the two best performing methods for the best forecasting techniques. We also include the random investment strategy for comparison.

Methods		Descriptive statistics of selected methods			
Clustering	Forecast	Mean	Std. Dev.	Maximum	Minimum
DTW Prices	LSTM	0.15%	2.30%	21.13%	−26.74%
Eros Q.Report	LSTM	0.14%	2.29%	22.04%	−26.74%
DTW Prices	ARIMA	0.08%	1.78%	22.04%	−16.18%
Eros Q.Report	ARIMA	0.10%	1.81%	22.04%	−13.92%
–	Random	−0.01%	1.65%	22.04%	−17.45%

should be noted that LSTM has higher standard deviation than ARIMA. Additionally, the random strategy, as expected, has the lowest standard deviation of all methods and a mean very close to 0%.

4.2. Robustness checks

Although the methods used had found a clear positive return, a question that lingers is whether it was just luck or was effectively capturing useful information. Therefore we designed two additional experiments to check the robustness of the results under other market situations. The first experiment extends the testing period to the second quarter of 2022. This period is interesting because from 1/Apr/22 to 30/Jun/22 the stock market showed a declining tendency with a 15% fall. The second experiment goes back in time and tests the period from 1/Feb/20 to 30/Apr/20. The selection of this time window is because COVID-19 could be regarded as an unforeseeable, one-time-ever event with important economic consequences. The COVID-19 pandemic hit the markets, causing a 30% decline (between 20/Feb to 20/March), followed by a 15% recovery by the end of April.

In the first case, we used the same clusters as in Section 4.1 and we forecasted those three months and used the forecast to perform our investment strategy only with Neural Networks as the results with ARIMA were not as good. We also calculated the returns for the Buy and Hold and MACD strategies to have a benchmark as before. The results are displayed in Table 5.

Regarding the second case, the training period changes. Therefore, we run the process from the beginning, meaning that we have clustered financial data (quarterly reports and prices) for the period 1/Jan/17 to 31/Jan/20. Subsequently, we used these clusters to forecast prices for the period 1/Feb/20 to 30/Apr/20 (using both ARIMA and LSTM) and used the results to conduct our investment strategy. Besides, we also calculated the results of Buy and Hold and MACD for comparison purposes. The results can be seen in Table 6.

In both cases the returns achieved through our method with LSTM are better than Buy and Hold and (in the first case) better than MACD, being similar in the second case. Consequently, these results confirm

Table 5
Results and returns of the forecasting methods applied to prices showing the number of trades executed, the percentage of successful forecasts and the average error for each prediction generated with LSTM for the 2nd Quarter 2022.

Base Methods			Return		
Buy and Hold			−14.59%		
MACD			−10.63%		
Method	# Predictions	% Success	sMAPE	MASE	Returns
DTW Q.Report	13704	51.1%	10.107	5.233	7.41%
Eros Q.Report	13728	50.0%	10.145	5.382	2.23%
DTW Prices	13694	50.5%	10.035	5.246	5.59%
Fourier Prices	13722	50.9%	10.363	5.589	6.38%

Table 6
Results and returns of the forecasting methods applied to prices showing the number of trades executed, the percentage of successful forecasts and the average error for each prediction generated with ARIMA and LSTM for the period February–April 2020.

Base Methods			Return		
Buy and Hold			−15.99%		
MACD			6.25%		
		ARIMA			
Method	# Predictions	% Success	sMAPE	MASE	Return
DTW Q.Report	9251	32.8%	4.323	0.982	−1.63%
Eros Q.Report	9536	33.9%	5.404	1.266	−3.73%
DTW Prices	9053	32.9%	4.495	1.022	−0.85%
Fourier Prices	8941	32.9%	4.414	1.005	−3.19%
		LSTM			
Method	# Predictions	% Success	sMAPE	MASE	Return
DTW Q.Report	15383	47.2%	21.170	5.274	6.54%
Eros Q.Report	15422	47.0%	21.211	5.399	4.11%
DTW Prices	15381	47.3%	21.786	5.554	4.99%
Fourier Prices	15381	47.6%	21.827	5.637	5.40%

the robustness of our proposal, which performs very well, even under a "black swan" situation (Covid-19 pandemic) or a declining market due to uncertainty regarding the Russia–Ukraine war and US speeding inflation. One interesting fact is that ARIMA does not perform well in any sense in the second case, as ARIMA fails to adapt to the sharp fall caused by the lockdowns and supply chain disruptions due to COVID-19 and their impact in the stock markets, as it can be seen in the poor success rate and negative returns.

4.3. Discussion

Our results suggest some interesting implications. First, we find that too much information can be as harmful as the lack of it. The empirical simulations show that when the information for a given stock is supplemented with the information from the rest of the cluster, the forecasting results are better either in terms of accuracy or performance. In fact, they exceed the forecasting using only single stock information and full stock sample information. This indicates that the selection of similar stocks adds useful information to the forecasting methods. On the contrary, using information for all stocks adds more noise and contradictory information that harms the forecasting ability. In Table 3, we can observe that using all the data (prices/returns and quarterly reports) in a LSTM network to forecast the price or return does not reach a better result than that obtained through supplementing the data to forecast the stock prices.

This finding challenges the conventional wisdom that feeding a neural network with more data helps the forecasting. This bias is not as evident in the case of ARIMA, where the results are all similar, except in the Multiple (intraday) case (using high, low, opening and close prices of one stock) where the return is negative.

On the other hand, the good performance results (Table 3) are in apparent conflict with the accuracy results (Table 2). We offer the following explanation for this paradox: forecasting methods (especially LSTMs) are good at triggering the right signal for the higher return transactions, whereas those days with shallow returns, the predicted signal is more frequently wrong. As can be observed in Table 7, the mean return on profits is greater than the mean return on losses (significant at 1% level), except (as expected) for the random purchase.

An additional insight in the results shown in Table 7, is the number of transactions executed by each method. It can be observed that LSTM models induce more trades, whereas ARIMA models triggers less frequently buy or sell (short sell) signals. Then, using ARIMA forecasting, the trading algorithm decides do nothing (no trade) more frequently than LSTM. The statistically significant difference in positive mean returns and the greater number of trades explain the better performance of the neural network based models *vis-à-vis* their ARIMA counterparts.

5. Conclusions

This article provides three important results. First, deep learning techniques, such as LSTM, generally outperforms (especially in prices) traditional ARIMA models. This could be explained by the ability of neural networks to deal with nonlinearities in financial data, made evident in the verification done with the period February–April 2020. This is reflected in a greater number of trade signals triggered by neural network methods and the average profit obtained in each trade. Second, the different simulations indicate that not all data is useful in a forecasting method. Clustering of stock information provides relevant information that supplements the specific data of a given stock, improves the forecasting power and allows obtaining higher returns than using only stock-specific data. In this aspect, clustering both quarterly accounting reports data and prices (or alternatively returns) adds knowledge to the forecasting model. Third, adding information (both from financial reports and prices/returns) from all the stocks in the sample harms the prediction ability of the neural network. This means that more data is not necessarily followed by better forecasts. Probably, data from unrelated stocks adds noise and contradictory signs that lower the forecast accuracy and performance in the test set. In other words, stock-related information (provided by the cluster where it belongs to) is blurred by adding unrelated data. As such, the prediction is closer to the mean market behavior, rather than to a specific stock behavior.

As expected, by economic intuition and common sense, only related data provides useful information for the forecasting process. Therefore, the first point that should be highlighted is that some of this related data, invisible to the usual analysis, emerges through the clustering process, evidencing this method as a useful preprocessing step. The second point is that more data does not imply better results; quality rather than quantity is the key here. These points are good news, as a practical consequence, less data to process derives in a reduction of computational process times. The third point that should be mentioned is that the exploitation of data in financial reports through the creation of temporal series and subsequent clustering provides the means to reveal hidden associations.

Finally, the results show that our proposal performs well under different market situations. In particular, we tested the methods in the first and second quarter of 2022, as well as during February/2020–April/2020 (COVID-19 pandemic outbreak). In all three cases, the combination of clustering financial data and using LSTM neural networks provided better performance that the benchmarking methods.

We propose in future works to evaluate the sensitivity of the method with respect to stock sector or size. Moreover, clustering could be complemented using alternative data such as sentiment analysis, or could combine the data from the financial reports with the price evolution of commodities or macroeconomic variables to further refine the clustering.

Table 7
Descriptive statistics of the *profits and losses* obtained using the two best performing models (LSTM and ARIMA) and clustering method combinations. We also include the random investment strategy for comparison.

Methods		Profits			Losses		
Clustering	Forecast	Mean	Std. Dev.	# Trades	Mean	Std. Dev.	# Trades
DTW Prices	LSTM	1.97%	2.01%	7078	1.80%	1.82%	6498
Eros Q.Report	LSTM	1.96%	2.03%	7022	1.81%	1.82%	6470
DTW Prices	ARIMA	1.96%	2.11%	4301	1.81%	1.80%	4007
Eros Q.Report	ARIMA	1.98%	2.11%	4458	1.76%	1.74%	4165
–	Random	1.87%	1.90%	3532	1.88%	1.92%	3557

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, *2014*.

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering–a decade review. *Information Systems*, *53*, 16–38.

Amini, S., Hudson, R., Urquhart, A., & Wang, J. (2021). Nonlinearity everywhere: implications for empirical finance, technical analysis and value at risk. *The European Journal of Finance*, *27*, 1326–1349. http://dx.doi.org/10.1080/1351847X.2021.1900888.

Appel, G., & Dobson, E. (2007). *Understanding MACD, vol. 34*. Traders Press.

Babu, M. S., Geethanjali, N., & Satyanarayana, B. (2012). Clustering approach to stock market prediction. *International Journal of Advanced Networking and Applications*, *3*(4), 1281.

Bachelier, L. (1900). *ThÉorie de la SpÉculation* (p. 148). Paris: Annales scientifiques de l'École Normale Supérieure, ISBN: 2876471299.

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.

Bandara, K., Bergmeir, C., & Smyl, S. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Systems with Applications*, *140*, Article 112896.

Basalto, N., Bellotti, R., De Carlo, F., Facchi, P., & Pascazio, S. (2005). Clustering stock market companies via chaotic map synchronization. *Physica A: Statistical Mechanics and its Applications*, *345*(1–2), 196–206.

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. http://dx.doi.org/10.1016/j.intfin.2017.06.005.

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series.. *10*, In *KDD workshop* (16), (pp. 359–370). Seattle, WA, USA:.

Chen, K. H., & Shimerda, T. A. (1981). An empirical analysis of useful financial ratios. *Financial Management*, 51–60.

Dhakal, S. (2019). Reliability of clustering in forecasting stock prices of companies traded on the stock exchanges. *Merge*, *3*(1), 5.

Dyl, E. A., & Maberly, E. D. (1988). The anomaly that isn't there: A comment on friday the thirteenth. *The Journal of Finance*, *43*, 1285–1286. http://dx.doi.org/10.1111/j.1540-6261.1988.tb03971.x.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417. http://dx.doi.org/10.2307/2325486, URL https://www.jstor.org/stable/2325486?origin=crossref.

Fang, W.-X., Lan, P.-C., Lin, W.-R., Chang, H.-C., Chang, H.-Y., & Wang, Y.-H. (2019). Combine facebook prophet and LSTM with BPNN forecasting financial markets: the morgan Taiwan index. In *2019 International symposium on intelligent signal processing and communication systems* (pp. 1–2). IEEE.

Goodwin, P., & Lawton, R. (1999). On the asymmetry of the symmetric MAPE. *International Journal of Forecasting*, *15*(4), 405–408.

Gubareva, M., & Borges, M. R. (2022). Governed by the cycle: interest rate sensitivity of emerging market corporate debt. *Annals of Operations Research*, *313*(2), 991–1019. http://dx.doi.org/10.1007/s10479-021-03972-x.

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.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702.

Lee, A. J., Lin, M.-C., Kao, R.-T., & Chen, K.-T. (2010). An effective clustering approach to stock market prediction.. In *PACIS* (p. 54).

Li, X., & Wu, P. (2021). Stock price prediction incorporating market style clustering. *Cognitive Computation*, 1–18.

Li, M., Zhu, Y., Shen, Y., & Angelova, M. (2022). Clustering-enhanced stock price prediction using deep learning. *World Wide Web*, 1–26.

Lin, Y.-L., Lai, C.-J., & Pai, P.-F. (2022). Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis. *Electronics*, *11*(21), 3513. http://dx.doi.org/10.3390/electronics11213513.

Liu, S., & Motani, M. (2021). Towards better long-range time series forecasting using generative adversarial networks. arXiv preprint arXiv:2110.08770.

Lo, A. W. (2004). The adaptive markets hypothesis. *The Journal of Portfolio Management*, *30*, 15–29. http://dx.doi.org/10.3905/jpm.2004.442611.

Long, J., Chen, Z., He, W., Wu, T., & Ren, J. (2020). An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. *Applied Soft Computing*, *91*, Article 106205.

Ma, Q. (2020). Comparison of ARIMA, ANN and LSTM for stock price prediction. *218*, In *E3S web of conferences* (p. 01026). EDP Sciences.

Marvin, K. (2015). *Creating diversified portfolios using cluster analysis*. Princeton University.

Nair, B. B., Kumar, P. S., Sakthivel, N., & Vipin, U. (2017). Clustering stock price time series data to generate stock trading recommendations: An empirical study. *Expert Systems with Applications*, *70*, 20–36.

Osborne, M. F. M. (1959). Brownian motion in the stock market. *Operations Research*, *7*(2), 145–173.

Osborne, M. F. M. (1962). Periodic structure in the Brownian motion of stock prices. *Operations Research*, *10*(3), 345–379.

Samal, K. K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time series based air pollution forecasting using SARIMA and prophet model. In *Proceedings of the 2019 international conference on information technology and computer communications* (pp. 80–85).

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, *6*(2), 41–49.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE international conference on machine learning and applications* (pp. 1394–1401). IEEE.

Tadayon, M., & Iwashita, Y. (2020). A clustering approach to time series forecasting using neural networks: A comparative study on distance-based vs. feature-based clustering methods. arXiv preprint arXiv:2001.09547.

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, *72*(1), 37–45.

Tupe, T. (2014). Financial ratio analysis for stock price movement prediction using hybrid clustering.

Wang, Y.-J., & Lee, H.-S. (2008). A clustering method to identify representative financial ratios. *Information Sciences*, *178*(4), 1087–1097.

Wang, X., Yang, K., & Liu, T. (2021). Stock price prediction based on morphological similarity clustering and hierarchical temporal memory. *IEEE Access*, *9*, 67241–67248.

Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019). A comparison between arima, lstm, and gru for time series forecasting. In *Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence* (pp. 49–55).

Yang, K., & Shahabi, C. (2004). A PCA-based similarity measure for multivariate time series. In *Proceedings of the 2nd ACM international workshop on multimedia databases* (pp. 65–74).

Yenidoğan, I., Çayir, A., Kozan, O., Dağ, T., & Arslan, Ç. (2018). Bitcoin forecasting using ARIMA and PROPHET. In *2018 3rd international conference on computer science and engineering* (pp. 621–624). IEEE.