

ARTIFICIAL INTELLIGENCE AND FINANCIAL ASSET PRICE FORECASTING: A SYSTEMATIC REVIEW

Ewerton Alex Avelar

ewertonaavelar@gmail.com
Minas Gerais Federal University –
UFMG, Belo Horizonte, MG, Brazil.

Octávio Valente Campos

octaviovc@yahoo.com.br
Minas Gerais Federal University –
UFMG, Belo Horizonte, MG, Brazil.

Jacqueline Braga Paiva Orefici

j.orefici@gmail.com
Leonardo da Vinci University
Center – UNIASSELVI, Brazil.

Sergio Louro Borges

sergio.borges@ufjf.edu.br
Juiz de Fora Federal University –
UFJF, Juiz de Fora, MG, Brazil.

Antônio Artur de Souza

artur@face.ufmg.br
Minas Gerais Federal University –
UFMG, Belo Horizonte, MG, Brazil.

ABSTRACT

Highlights: An efficient market is one in which prices always fully reflect the available information (Fama, 1970). However several agents have focused on potential inefficiencies for abnormal returns over time. Artificial intelligence (AI) algorithms have also been used to predict the prices of financial assets. This systematic literature review presented a series of contributions to using AI algorithms to forecast asset prices in the financial market. **Purpose:** To develop a systematic review of AI algorithm applications for forecasting asset prices in the financial market. **Methodology:** The systematic review focused on the guidelines from *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA). It addressed two main journal databases: Web of Science and Scopus. We conducted a qualitative analysis of the studies and based our conclusions on pm content analysis. The analysis employed categories based on the results reported in the literature. We also used descriptive statistics and the chi-squared test in the analysis. **Findings:** This study presents some relevant contributions: (i) the identification of the main features of the developed models based on artificial intelligence (AI) and the algorithms used to forecast asset prices in the financial market; (ii) the review of features and applications of the main algorithms used in forecasting; and (iii) the gaps in previous studies, as well as tendencies and perspectives for further analysis. **Limitations:** This study focuses only on papers publicly available on two major databases. Moreover, some subjects' issues might influence the categorization process. **Practical implications:** This paper presents the main features of AI algorithmics used to forecast asset prices in the financial market. These results can support market agents in improving their investment models. **Originality/value:** This paper addressed many relevant theoretical and practical issues. It also enhanced the importance of understanding the efficient market hypothesis (EMH) under the comprehensive automation of processes and the use of AI. Lastly, it systematically reviewed the different features among the analyzed studies.

Keywords: Artificial Intelligence; Price Forecast; Financial Market; Systematic Review.

INTRODUCTION

Cao, Lin, Li, and Zhang (2019) highlight that knowing behavior patterns and being able to make predictions about asset prices in the financial market is an important issue to be addressed in scientific research. In this sense, Moon, Jun, and Kim (2018) state that forecasting the prices of financial assets is a relevant theme in finance, as such forecasts enable, for example, economic agents to make their profits and protect themselves from market risks.

Ding and Qin (2020) consider that this type of research has always been relevant for economic agents, and several different methods have been used to forecast asset prices. Such methods range from widespread statistical techniques to the latest advances in artificial intelligence (AI). Regarding AI algorithms, Rundo, Trenta, Stallo, and Battiano (2019) emphasize that such use is in the context of the progressive automation of processes in different fields, and the financial market has not been an exception. According to these authors, many researchers have proven that AI algorithms make it possible to fast analyze a large volume of data with great accuracy and effectiveness.

However, according to the Efficient Market Hypothesis (EMH), forecasting asset prices for economic agents to obtain abnormal profits in the financial market is not theoretically possible. In his classic work, Fama (1970) defines an efficient market as one in which prices always fully reflect the availability of information. Thus, the EMH implies that, on average, an investor could not achieve an abnormal return (Ross, Westerfield, Jaffe, & Lamb, 2015).

Contrary to the hypothesis mentioned above, several studies that used AI algorithms to forecast asset prices have presented models with high performance in terms of predictive power, maximizing abnormal returns (*e.g.*, Cao *et al.*, 2019; Qian & Rasheed, 2007; Shynkevich, McGinnity, Coleman, Belatreche & Li, 2017). It should be noted that such studies have highlighted significant differences, from different perspectives, of the main algorithms used, such as *Artificial Neural Networks* (ANN), *Decision Tree/Random Forest* (DTRF), *k-Nearest Neighbors* (KNN), *Naive Bayes* (NB) and *Support Vector Machine* (SVM) (Shynkevich *et al.* 2017; Cao *et al.*, 2019; Ding & Qin, 2020).

Recognizing and exploring this research gap, the study presented in this article aims to answer the following research question: How has the application of AI algorithms to forecast asset prices in the financial market been addressed in the literature? Thus, the objective of the research was to carry out a systematic review (mapping the state of the art) on the application of AI algorithms to

forecast asset prices in the financial market. Therefore, the review was developed using the Web of Science and Scopus bibliographic databases, focusing on the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidelines (Page *et al.*, 2021a; Page *et al.*, 2021b).

The research developed can be justified from different perspectives. Firstly, there is the great importance of the topic, both from a theoretical and a practical point of view, for the Academy and the different market agents, such as investors, companies, and regulators (Cao *et al.*, 2019; Ding & Qin, 2020; Moon *et al.*, 2018). Furthermore, the importance of understanding EMH is highlighted in a new environment in which processes are extensively automated, and AI algorithms have been used to obtain higher-than-expected returns (Rundo *et al.*, 2019). Finally, the importance of presenting the different characteristics of the studies that focused on asset price forecasting and the algorithms used for that purpose is highlighted since they present different levels of performance in different contexts (Shynkevich *et al.*, 2017; Cao *et al.*, 2019).

THEORETICAL FOUNDATION

This section discusses critical aspects of the systematic review presented in this work. Initially, in subsection 2.1, the importance of forecasting asset prices in the financial market is discussed in the context of the EMH. Sequentially, the main AI algorithms used for such activity are highlighted in subsection 2.2. Finally, aspects of the development of models that employ such algorithms are highlighted in subsection 2.3.

Forecasting asset prices in the financial market

According to Ding and Qin (2020), the increased or decreased price of assets in the financial market is influenced by many factors, such as political, economic, social, and market-based. Additionally, according to these authors, these movements in asset prices, in turn, directly influence the returns obtained by investors, who can benefit from the correct prediction of these movements, which is, however, a very complex activity.

Such complexity can be related to EMH. According to Fama (1970), an efficient market is one in which prices always fully reflect the available information. It should be noted that efficiency varies in each form (weak, semi-strong, or strong), related to the speed with which the market assimilates information (Fama, 1970). Thus, as Ross *et al.* (2015) explained, the EMH implies that, on average, an investor could not achieve an abnormal re-

turn. However, the conditions listed by Fama (1970) for efficiency are ideal, allowing for abnormal returns from potential inefficiencies. Thus, over time, several agents have focused on this possibility.

According to Rundo *et al.* (2019), in recent decades, researchers have proposed a series of models based on statistical methods to forecast the prices of these assets, such as the *autoregressive integrated moving average* (ARIMA) and the *exponential smoothing model*. However, the authors point out that these models need help in this task due to their low performance when dealing with a large volume of intrinsically complex data, such as these assets' prices. These approaches also seem they need to be more suitable for understanding hidden relationships (dependencies) between data (Rundo *et al.*, 2019).

Ding and Qin (2020) emphasize that, in addition to statistical techniques, AI algorithms have also been used to predict the prices of financial assets. Among such algorithms, some stand out in the literature on the subject: ANN, DTRF, KNN, NB, and SVM (Shynkevich *et al.*, 2017; Cao *et al.*, 2019; Ding & Qin, 2020). Rundo *et al.* (2019) point out that its use is related to the effects of the progressive automation of certain processes in different fields, including those in the financial area. It is important to highlight that, according to Faceli, Lorena, Gama, Almeida, and Carvalho (2021), the previously mentioned algorithms can be used to solve both regression problems referring to value estimation in an infinite and ordered set (*e.g.*, share price in a given context) and classification problems estimating values from a discrete and unordered set, that is, a class (*e.g.*, whether the price of a stock will rise or fall). The following subsection details each of these algorithms.

AI algorithms for asset price prediction

In the review presented in this work, the algorithms mentioned in the previous subsection are focused on: ANN, DTRF, KNN, NB, and SVM. Faceli *et al.* (2021) argues that ANNs are inspired by abstract models of how the human brain is believed to work. These authors claim that such networks are composed of simple processing units responsible for implementing mathematical functions that simulate the functions performed by neurons. Such units can connect to many other connections, simulating synapses, which allows ANNs to solve complex problems.

As far as *decision trees* are concerned, Moon *et al.* (2018) highlight that they can be used to create a model that predicts the value of a target variable based on several input variables using recursive partitioning. A variable that best divides the sample set is chosen at each

step. Different impurity measures or splitting criteria can be used in binary trees, such as Gini impurity, information entropy, or misclassification. It should be noted that the *random forest algorithm* can be considered a development regarding *decision trees*. The idea is to combine several trees to determine the final result rather than relying on individual trees, reducing the model's variance (Vijh, Chandola, Tikkiwal, & Kumar 2020). Thus, for the study presented in this article, both algorithms (*decision trees* and *random forests*) are considered to fall into the same analysis category as DTRF.

The KNN algorithm, as highlighted by Moon *et al.* (2018), chooses the class label of the new data point by majority vote among its "k" nearest neighbors. The chosen distance metric determines these nearest neighbors. KNN is simple to implement, but it is sensitive to the local structure of the data and the computational complexity to classify new samples, which grows linearly with the number of samples in the training set. The parameter k can be chosen depending on the data, and generally, larger values of k reduce the effect of noise on classification but make the boundaries between classes less distinct (Moon *et al.*, 2018).

In turn, according to Faceli *et al.* (2021), NB computes all probabilities (priors and conditional) of the training data. According to these authors, the term "naive" is related to the hypothesis that the attribute values of an example are independent of its class. Finally, Rundo *et al.* (2019) emphasize that the SVM algorithm finds a decision function that maximizes the margin between classes. The algorithm performs a mathematical optimization based on the labeled data during the training step. Training examples that limit the maximum margin defined by the SVM during training are called "support vectors." The following subsection describes AI models' development by employing algorithms such as those mentioned above.

AI Models

Regardless of the AI algorithm used for asset price prediction, Ferreira, Gandomi, and Cardoso (2021) present a flowchart of the process usually used by studies for such activity (**Figure 1**). These authors include five main steps: (1) input data acquisition; (2) data transformation and selection; (3) model training; (4) parameter optimization; and (5) evaluation of the predictor's performance.

There are several input data usually verified in studies that aim to predict asset prices in the financial market, such as (a) trading history - closing, opening, maximum, and minimum prices - and trading volume (Chun & Ko, 2020; Gu, Shibukawa, Kondo, Nagao, & Kamijo, 2020;

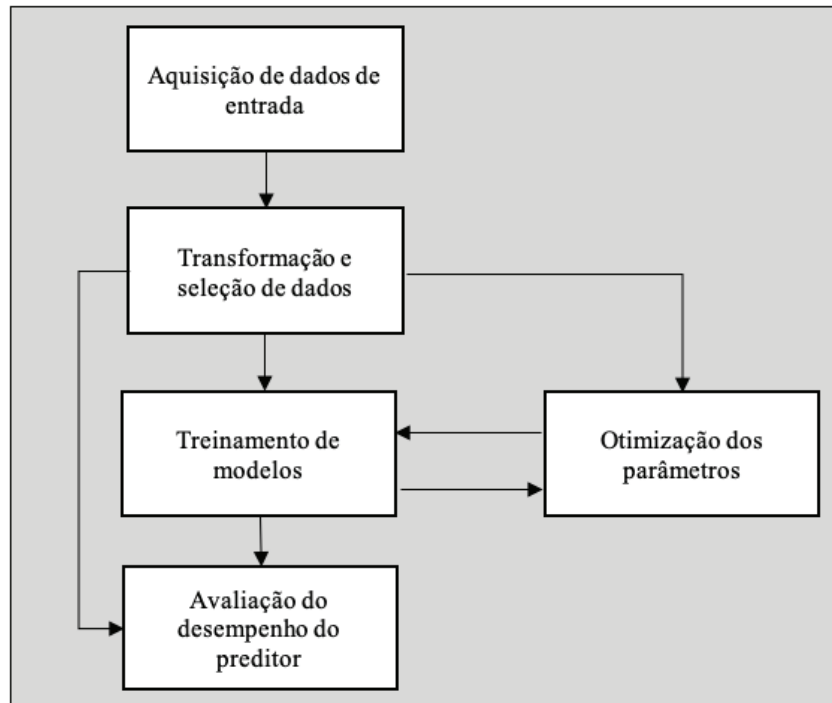


Figure 1. Process flowchart usually employed by studies using AI algorithms to forecast financial assets' prices.

Source: Adapted from Ferreira, Gandomi, and Cardoso (2021, p. 30,906)

Awan, Rahim, Nobanee, Munawar, Yasin, & Zain, 2021); (b) technical analysis indicators (Rundo *et al.*, 2019); (c) financial indicators (Janková, Jana, & Dostál, 2021); and (d) unstructured data for behavior analysis, employing natural language processing (NLP) (Almehmadi, 2021; Awan *et al.*, 2021).

In addition, different types of assets have been the focus of price prediction via AI algorithms, such as market indices (Cavdar & Aydin, 2020; Ding & Qin, 2020; Shynkevich *et al.*, 2017); share prices (Colliri & Zhao, 2019; Awan *et al.*, 2021); and prices of other financial assets, such as options (Sheu & Wei, 2011).

After selecting the algorithm to forecast prices, it is necessary to train it based on the collected data. That needs to be done with data related to the input variables and with data related to the estimated prices. At this stage, most of the data is used for training the algorithm, while the remainder is used for testing it (Moon, Jun, & Kim, 2018; Shynkevich *et al.*, 2017). It is also necessary optimize parameters according to the algorithms, such as parameter k in the case of KNN, *kernels* in the case of SVM, and adjustment of the weights in the case of ANN (Faceli *et al.*, 2021).

Finally, several performance evaluation metrics of the algorithms that can be used are highlighted, such as Accu-

racy (ACU), *Mean Square Error* (MSE), *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), and *Mean Absolute Percentage Error* (MAPE) (Ecer *et al.*, 2020; Awan *et al.*, 2021; Vijh *et al.*, 2020). It is important to note that while the ACU is more suitable for evaluating algorithms aimed at classification, the other metrics are more suitable for evaluating those aimed at performing regression analysis (Faceli *et al.*, 2021).

METHODOLOGY

The systematic review presented in this article was developed, focusing on PRISMA guidelines (Page *et al.*, 2021a; Page *et al.*, 2021b). Before proceeding with the review, a search was carried out in the *Open Science Framework* (OSF) records, and no developments were found in this regard (OSF Home, 2021), indicating the unprecedented nature of the study. The literature review was carried out in two journal databases: Web of Science and Scopus. Chadegani *et al.* (2013) highlight the importance of both databases for the scientific community. These authors emphasize that the Web of Science (Thomson Reuters) could be considered the main scientific reference for several areas until the launch of Scopus (Elsevier Science). The latter started to compete directly with the

former, being constituted in similar amplitude and scale (Chadegani *et al.*, 2013).

For the selection of articles, each of the databases was accessed in the last week of May 2021, and a *Boolean search* was performed with the following search query: [(“Machine Learning” OR “Artificial Intelligence”) AND (“stock market” OR “stock return” OR “stock price” OR “share market” OR “share return” OR “share price”)]. After performing these procedures, 840 documents were initially selected. Then, to refine the search, selection filters were employed, restricting the search to documents classified as “articles,” which resulted in 392 records. After this refinement, all titles and abstracts were read and analyzed to verify whether the texts referred to the phenomenon of using algorithms for forecasting financial assets (research focus), and 268 articles were selected.

Subsequently, the articles were downloaded and read in full. It should be noted that when the full text was not available in the selected database, it was searched directly on Google[®] Scholar. However, 68 articles were not found with the use of such procedures, being excluded from the sample. Finally, of the remaining 200 articles, it was verified, during the full reading, that 12 of them did not refer to the research focus, 2 were duplicates, and 51 approached AI algorithms different from those highlighted in subsection 2.2 (especially hybrids), generating the final sample of 135 articles. In this sense, **Figure 2** presents this article selection process based on the flowchart proposed by PRISMA.

After reading the selected articles in full, they were qualitatively analyzed and classified according to different analysis categories. These categories mainly focused on the AI algorithms highlighted in subsection 2.1 and some of the steps in the model by Ferreira *et al.* (2021) shown in **Figure 1**: (a) world region; (b) type of anticipated asset; (c) AI algorithm employed; (d) input data for training the algorithm; and (e) measures of algorithm performance. Finally, a categorization of the main conclusions of the articles was also carried out.

It should be noted that at least two reviewers with a Ph.D. in Administration or Accounting with experience in research in finance carried out the entire selection process of articles. The researchers who led the systematic review development analyzed and discussed the studies. Disagreements were resolved by consensus among reviewers. This procedure was applied at all stages.

For the data presentation and analysis, the techniques of descriptive statistics and the chi-square test were used, as recommended by Maroco (2010). This test was employed to evaluate statistically significant associations

between the different AI algorithms analyzed in the research concerning the other categories developed for this research. In this case, a statistical significance level of 10% was considered. The *Statistical Package for the Social Sciences* (SPSS) and MS Excel were used to operationalize the analyses.

RESULTS

This section presents the results derived from the systematic review of the literature. Three subsections compose this section. Firstly, the next subsection highlights the results referring to the following categories: world region; anticipated asset type; data used for training the algorithm; and metrics to measure the performance of the algorithm. Then, the results related to each of the analyzed algorithms are presented: ANN, DTRF, KNN, NB, and SVM. Finally, the main conclusions of the analyzed studies are discussed.

General analysis

Figure 3 presents the number of articles published per year. In total, 122 articles on the topic were identified. There is a strong trend of growth in the theme throughout the studied period. It is important to stress that more than 58% of the articles were published in the last three years. This result demonstrates the recent attention given to the topic at the Academy.

In turn, **Table 1** presents the number of articles published by region of the world. It was possible to observe that the first studies were carried out mainly in developed countries. Some others were carried out simultaneously in several countries. Only in 2014 were studies carried out exclusively in emerging countries consistently recorded. Since then, the number of studies in these countries has been greater than in other countries for several years, corresponding to a total of 54 studies against 51 carried out in developed countries. The initial preference for developed countries can be related to their more advanced capital markets compared to those of emerging countries.

Table 2 shows the frequency of the types of assets for which prices were predicted in each study. Until 2012, studies aimed at predicting the values of capital market indices predominated. From that year on, surveys focusing on stock prices became more numerous, surpassing those related to indices in some periods. However, contrary to what was observed in **Table 1**, all the studies predicting index prices remained the most frequent (51.5%) compared to those focusing on stocks (46.4%). It is noteworthy that only some studies are focusing on other types of as-

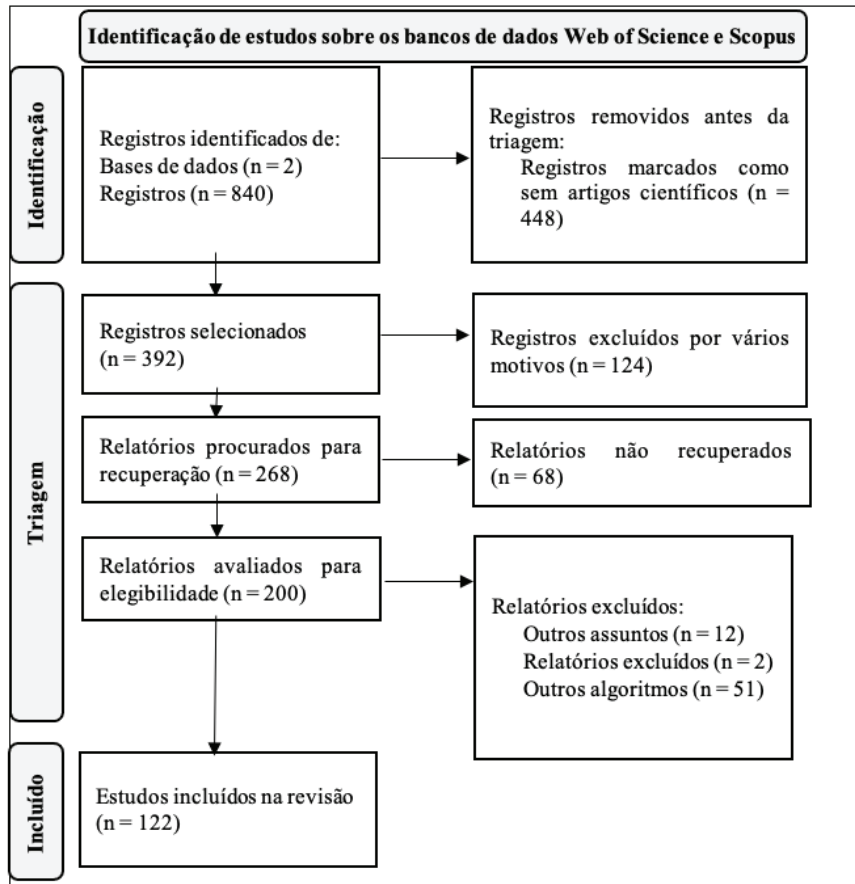


Figure 2. Article selection process.

Source: Adapted based on the model proposed by PRISMA (2021)

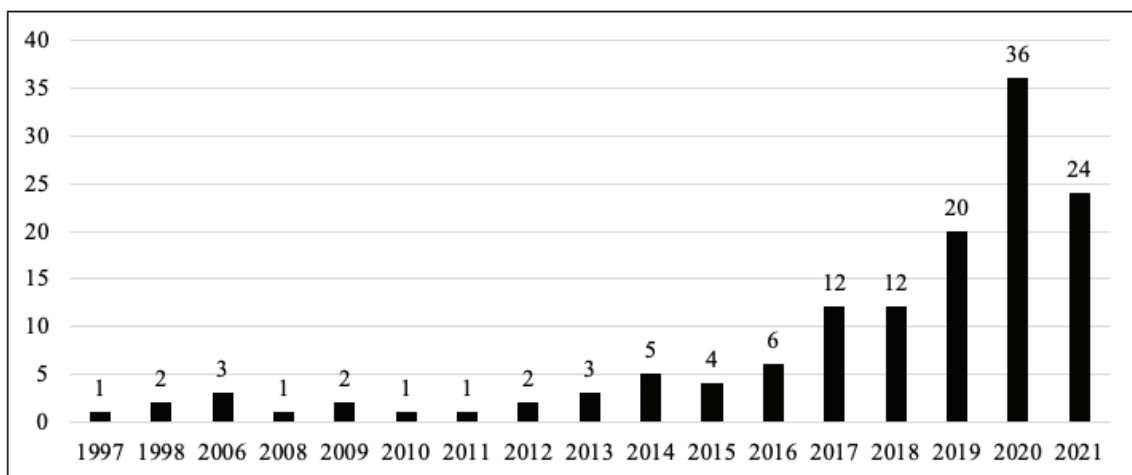


Figure 3. Number of articles published per year.

Source: Prepared by the authors.

Table 1 - Number of articles published by world region.

Year	Frequency				Relative (%)			
	General	Developed	Emerg.	Total	General	Developed	Emerg.	Total
1997	0	0	1	1	-	-	0.74	0.74
1998	0	1	1	2	-	0.74	0.74	1.48
2006	0	3	0	3	-	2.22	-	2.22
2008	1	0	0	1	0.74	-	-	0.74
2009	0	2	0	2	-	1.48	-	1.48
2010	1	0	0	1	0.74	-	-	0.74
2011	0	1	0	1	-	0.74	-	0.74
2012	1	0	1	2	0.74	-	0.74	1.48
2013	2	1	0	3	1.48	0.74	-	2.22
2014	0	2	3	5	-	1.48	2.22	3.70
2015	0	1	3	4	-	0.74	2.22	2.96
2016	2	2	2	6	1.48	1.48	1.48	4.44
2017	2	6	4	12	1.48	4.44	2.96	8.89
2018	3	5	4	12	2.22	3.70	2.96	8.89
2019	6	8	6	20	4.44	5.93	4.44	14.81
2020	7	11	18	36	5.19	8.15	13.33	26.67
2021	5	8	11	24	3.70	5.93	8.15	17.78
Total	30	51	54	135	22.22	37.78	40.00	100.00

Source: Search results.

sets, such as options. Thus, one can observe a gap in the literature that can be exploited in new research.

Table 3 presents the evolution of the number of studies, considering the different input data used in the training models. Initially, 215 different types of training data were observed, resulting in 1.6 types of input data per article. The most common input data refers to historical asset prices (opening, closing, maximum, and minimum), identified in 46.1% of the works. Another common input data type refers to technical indicators (moving average), which are present in 27.9% of the articles. Sentiment analysis is one type of input data that has become more frequent in recent years of analysis (83.3% of research using this *input* was published since 2017). The financial indicators were observed in only 7.9% of the analyzed studies. These are interesting data points for studies that employ fundamental analysis.

Finally, **Table 4** presents the different algorithm performance metrics employed in the studies. It is important to highlight that 232 different types of algorithm performance measures were mentioned in the articles, meaning that 1.7 measures were used per article. The most

common metric to measure algorithm performance was the ACU, present in 32.8% of the cases. Next, the RMSE metric is used by 15.1% of the studies. Other measures worth mentioning were MSE, MAE, and MAPE. It is also relevant to mention that 30.6% of the articles present other performance metrics and that these became more diverse throughout the studied period.

Analysis of AI algorithms

This subsection presents some information about the algorithms used in the analyzed studies. Initially, on average, 1.8 algorithms were presented in each study, showing that studies tend to employ more than one algorithm for price prediction. Sequentially, **Figure 4** shows the number of studies that used ANN as an asset price prediction algorithm. This is the most used algorithm for this task, observed in 33.8% of the analyzed articles. There was a statistically significant association between the use of the ANN algorithm and performance metrics other than the ACU ($\chi^2 = 4.1$, significant at less than 10.0%). This case shows that this algorithm tends to be used for regression and not classification purposes.

Table 2 - Frequency of the type of asset predicted in each study.

Year	Indices	Frequency			Indices	Relative (%)		
		Absolute	Stocks	Other		Total	Indices	Stocks
1997	1	0	0	1	0.72	-	-	0.72
1998	2	0	0	2	1.45	-	-	1.45
2006	2	1	0	3	1.45	0.72	-	2.17
2008	1	0	0	1	0.72	-	-	0.72
2009	2	0	0	2	1.45	-	-	1.45
2010	1	0	0	1	0.72	-	-	0.72
2011	1	0	0	1	0.72	-	-	0.72
2012	2	0	0	2	1.45	-	-	1.45
2013	1	3	0	4	0.72	2.17	-	2.90
2014	1	4	0	5	0.72	2.90	-	3.62
2015	4	1	0	5	2.90	0.72	-	3.62
2016	4	2	0	6	2.90	1.45	-	4.35
2017	6	6	0	12	4.35	4.35	-	8.70
2018	10	1	1	12	7.25	0.72	0.72	8.70
2019	14	6	1	21	10.14	4.35	0.72	15.22
2020	10	25	0	35	7.25	18.12	-	25.36
2021	9	15	1	25	6.52	10.87	0.72	18.12
Total	71	64	3	138	51.45	46.38	2.17	100.00

Source: Search results.

In turn, **Figure 5** presents the articles that present the SVM algorithm to predict the prices of assets. This is the second-most-used algorithm, observed in 27.6% of the analyzed articles. Interestingly, the chi-square test indicates a statistically significant association between the use of the SVM algorithm and studies based on emerging markets ($\chi^2 = 6.4$, significant at less than 1.0%). As studies in such markets have grown more than proportionally over the last decade, the general predominance of SVM can also be explained, as there seems to be a preference for such an algorithm in research in these regions.

There was also a strong association between the use of SVM and sentiment analysis as input data ($\chi^2 = 6.7$, significant at less than 5.0%). It is important to consider that SVM is a very complex algorithm for working with unstructured data using NLP, which is essential for dealing with this type of data. Furthermore, there was a statistically significant association between the use of the SVM algorithm and the use of ACU performance metrics ($\chi^2 = 4.5$, significant at less than 5.0%) and measures other than MAPE ($\chi^2 = 3.8$, significant at less than 10.0%). In

this case, it is possible to infer that this algorithm is used more for classification purposes than regression.

The third most frequent algorithm observed in the studies refers to DTRF (mentioned in 23.1% of the studies), whose observed frequency is presented in **Figure 6**. Regarding the input data, a statistically significant association was found between the use of these algorithms and financial indicators ($\chi^2 = 5.6$, significant at less than 5.0%). Thus, studies that use such algorithms tend to use these indicators as a basis for training. There was also a statistically significant association between the use of other performance metrics (alternatives) and the use of the DTRF algorithm ($\chi^2 = 7.4$, significant at less than 1.0%). It should be stressed that, in all studies that used this algorithm, measures different from those presented in subsection 2.3 were verified.

In turn, **Figure 7** presents the frequency of articles that employed the NB algorithm (used in 8.0% of the cases). It is interesting to note that its use was only observed from 2014 onward. Concerning the predicted asset, sta-

Tabela 3. Diferentes dados de entrada usados em modelos para treinamento.

Ano	Frequência											
	Absoluta						Relativa (%)					
	Preços	TI	FI	SA	Outros	Total	Preços	TI	FI	SA	Outros	Total
1997	1	0	0	0	0	1	0,47	-	-	-	-	0,47
1998	1	0	2	0	0	3	0,47	-	0,93	-	-	1,40
2006	1	2	0	0	0	3	0,47	0,93	-	-	-	1,40
2008	1	1	0	0	0	2	0,47	0,47	-	-	-	0,93
2009	2	0	0	2	0	4	0,93	-	-	0,93	-	1,86
2010	1	0	0	0	0	1	0,47	-	-	-	-	0,47
2011	1	1	0	0	0	2	0,47	0,47	-	-	-	0,93
2012	2	1	0	0	0	3	0,93	0,47	-	-	-	1,40
2013	1	1	1	1	0	4	0,47	0,47	0,47	0,47	-	1,86
2014	2	3	0	2	0	7	0,93	1,40	-	0,93	-	3,26
2015	2	2	1	0	0	5	0,93	0,93	0,47	-	-	2,33
2016	5	4	0	0	1	10	2,33	1,86	-	-	0,47	4,65
2017	12	7	1	2	1	23	5,58	3,26	0,47	0,93	0,47	10,70
2018	11	4	3	5	0	23	5,12	1,86	1,40	2,33	-	10,70
2019	16	6	4	5	0	31	7,44	2,79	1,86	2,33	-	14,42
2020	29	15	4	8	2	58	13,49	6,98	1,86	3,72	0,93	26,98
2021	11	13	1	5	5	35	5,12	6,05	0,47	2,33	2,33	16,28
Total	99	60	17	30	9	215	46,05	27,91	7,91	13,95	4,19	100,00

Fonte: Resultados da pesquisa.

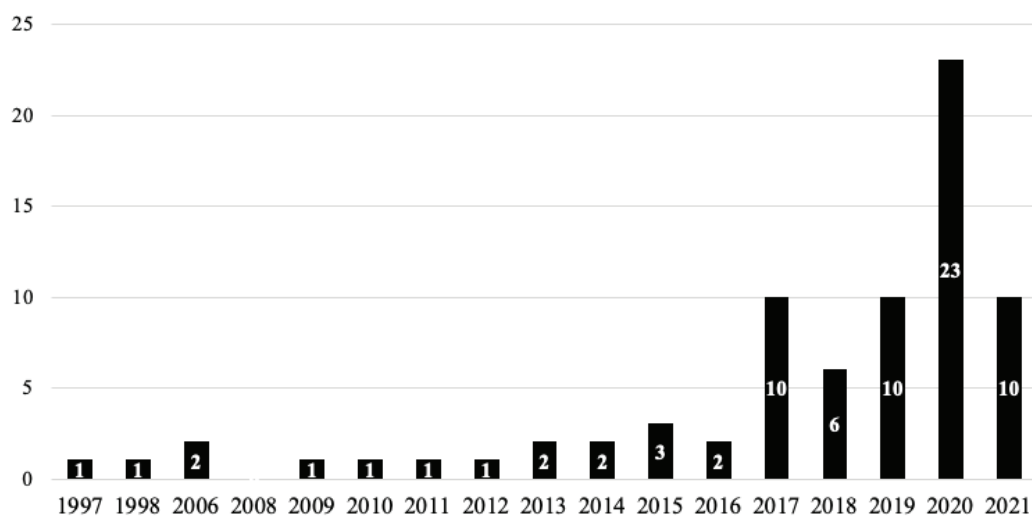


Figure 4. Number of studies that used the ANN algorithm annually

Source: Search results.

Table 4 - Performance metrics of the algorithms used by the analyzed studies.

Year	Absolute							Frequency						
								Relative (%)						
	ACU	MSE	MAE	MAPE	RMSE	Other	Total	ACU	MSE	MAE	MAPE	RMSE	Other	Total
1997	1	0	0	0	0	0	1	0.43	-	-	-	-	-	0.43
1998	2	0	0	0	0	0	2	0.86	-	-	-	-	-	0.86
2006	2	0	0	0	0	2	4	0.86	-	-	-	-	0.86	1.72
2008	1	0	0	0	0	0	1	0.43	-	-	-	-	-	0.43
2009	1	1	0	0	0	1	3	0.43	0.43	-	-	-	0.43	1.29
2010	0	0	0	0	0	1	1	-	-	-	-	-	0.43	0.43
2011	1	0	0	0	0	0	1	0.43	-	-	-	-	-	0.43
2012	0	0	0	1	1	1	3	-	-	-	0.43	0.43	0.43	1.29
2013	1	0	0	1	0	2	4	0.43	-	-	0.43	-	0.86	1.72
2014	4	0	0	1	1	2	8	1.72	-	-	0.43	0.43	0.86	3.45
2015	2	1	1	1	1	1	7	0.86	0.43	0.43	0.43	0.43	0.43	3.02
2016	4	1	0	0	1	2	8	1.72	0.43	-	-	0.43	0.86	3.45
2017	7	2	3	1	5	8	26	3.02	0.86	1.29	0.43	2.16	3.45	11.21
2018	7	0	1	0	3	10	21	3.02	-	0.43	-	1.29	4.31	9.05
2019	12	1	1	1	4	8	27	5.17	0.43	0.43	0.43	1.72	3.45	11.64
2020	16	7	6	3	10	23	65	6.90	3.02	2.59	1.29	4.31	9.91	28.02
2021	15	6	5	5	9	10	50	6.47	2.59	2.16	2.16	3.88	4.31	21.55
Total	76	19	17	14	35	71	232	32.7	8.19	7.33	6.03	15.1	30.6	100.0

Source: Search results.

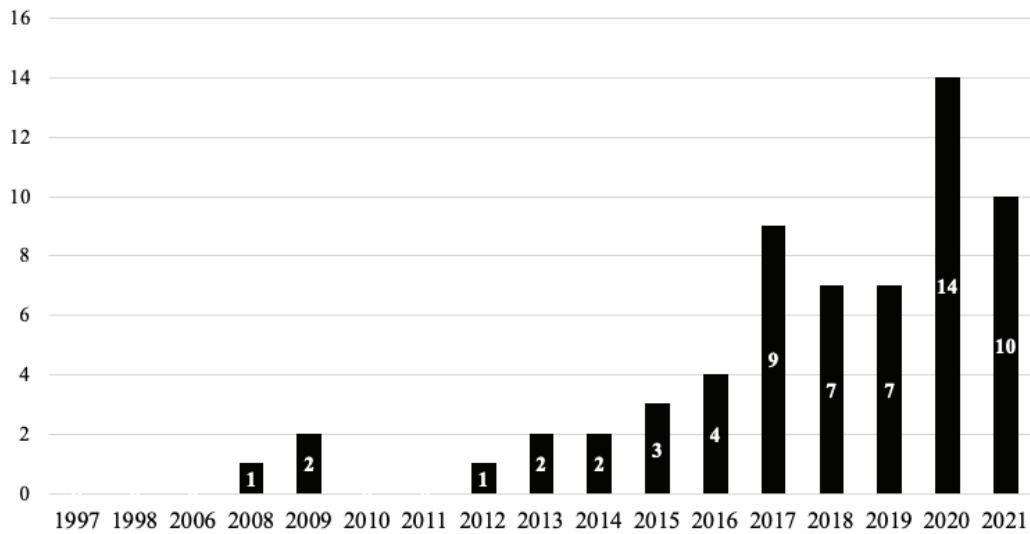


Figure 5. Number of studies that used the SVM algorithm annually

Source: Search results.

tistically significant associations were found regarding the use of the algorithm to predict stock prices ($\chi^2 = 14.33$, significant at less than 1.0%) as well as to forecast assets other than market indices ($\chi^2 = 10.8$, significant at less than 1.0%). In this case, it was found that there is a tendency to use models based on NB for forecasting stock prices to the detriment of using them for forecasting market indices.

As for the input data, there were statistically significant associations between the use of the NB algorithm and the use of sentiment analysis ($\chi^2 = 9.2$, significant at less than 1.0%) and with other data that were not historical prices ($\chi^2 = 3.4$, significant at less than 10.0%). In this case, there is a tendency for models that employ such an algorithm to use sentiment analysis as input data but not historical price data for this purpose. According to the recent advance in sentiment analysis in the area, the increase in the use of NB can be understood as a possible consequence. There was also a statistically significant association between using the ACU performance metric and the NB algorithm ($\chi^2 = 6.2$, significant at less than 5.0%), which shows the greater use of this algorithm for classification purposes

Finally, the frequency of studies that used the KNN is presented in **Figure 8**. It should be stressed that, despite being observed since 2006, this algorithm was the least used one in the studies (7.6%). It should also be noted that no study using this algorithm was recorded between 2009 and 2016. Interestingly, the chi-square test indicated a statistically significant association between the use

of the KNN algorithm and studies based on developed markets ($\chi^2 = 3.7$, significant at less than 10.0%).

Analysis of the main findings

Finally, this subsection presents a categorical analysis of the main conclusions of the analyzed articles. **Table 5** presents the frequency of such categories. It appears that 60.0% of the articles indicated that the results obtained through a particular AI algorithm were superior to traditional statistical techniques or to previous versions of other algorithms.

In 18 articles (13.3% of the sample), the authors argued that the AI algorithms generated good results but did not report a significant superiority to other techniques or algorithms. In turn, the third most frequent category of results indicates that hybrid algorithms, or the common use of several AI algorithms, provide better results than AI algorithms individually. Thus, in the vast majority of the analyzed articles (71.9%), the authors report having obtained results superior to those previously obtained.

However, four studies showed similar performance between AI algorithms and traditional statistical techniques (e.g., Parray *et al.*, 2020; Jaggi *et al.*, 2021). In this case, the authors did not observe significant advantages in using AI algorithms compared to traditional statistical techniques. On the other hand, two studies found that the performance of these algorithms would be even lower than that of traditional techniques (Pyo *et al.*, 2017;

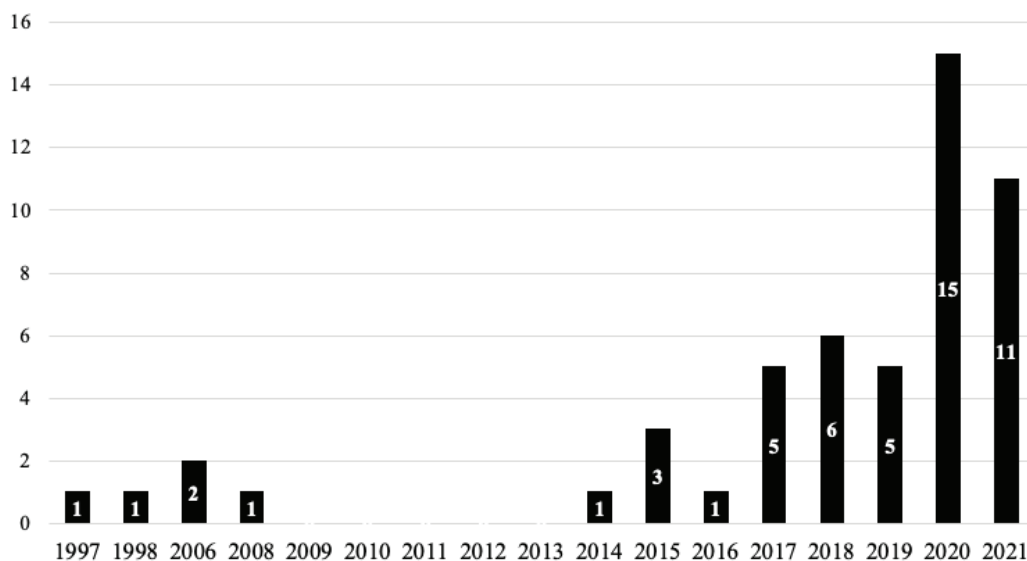


Figure 6. Number of studies that used DTRF algorithms annually
 Source: Search results.

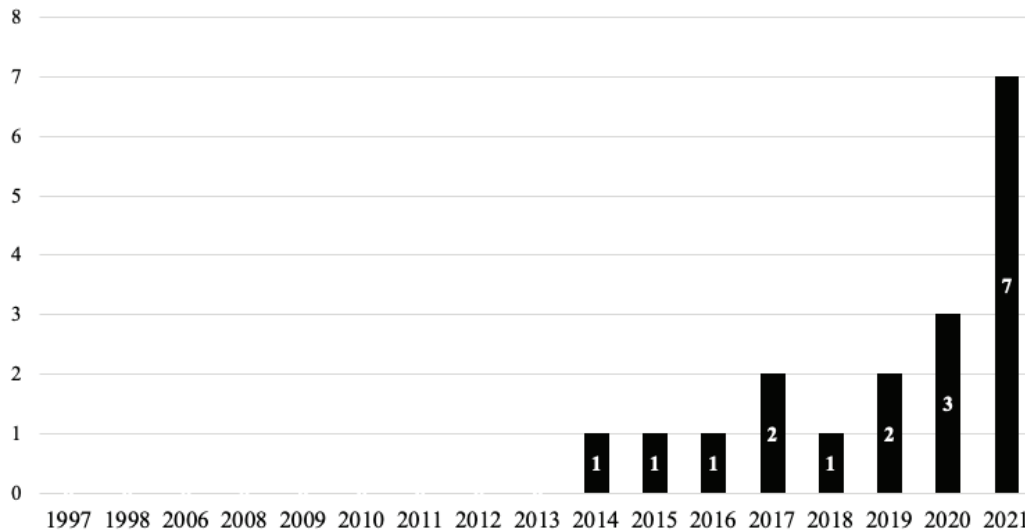


Figure 7. Number of studies that used NB algorithms annually

Source: Search results.



Figure 8. Number of studies that used the KNN algorithm annually

Source: Search results.

Jang & Lee, 2019). It is important to highlight that such studies correspond to only 4.4% of all analyzed articles.

Given the above, the rapid technological development from which AI algorithms benefit and their application in the financial market have opened up new possibilities in predicting asset prices, showing performances superior to traditional techniques and previous algorithms. On the other hand, an epistemological question may be raised, given the objectives outlined in the articles.

According to the survey, the authors intended to obtain favorable results with the proposed models. This

intention, by itself, can deprive some of the robustness of the results they have produced since independence in data management is lost since researchers would already start with the premise that the proposed models would be superior. That said, as they are models that need human analysis, the work ends up being directed, even if unconsciously, to overestimate the results of some models compared to others. Thus, epistemologically, it seems unclear how to discriminate between the part of the results due to the proposed models' actual superiority and the part due to the researchers' purposeful action to maximize these models.

Table 5 - Summarization of the main conclusions of the analyzed studies.

Main Conclusions	Frequency	
	Absolute	Relative (%)
The use of a particular AI algorithm is superior to traditional statistical techniques or to previous versions of other algorithms.	81	60.00
AI algorithms generated good results.	18	13.33
Hybrid algorithms or the joint use of several AI algorithms provide better results than AI algorithms used individually.	16	11.85
Specific variables are important in AI algorithms.	4	2.96
AI algorithms and traditional statistical techniques have comparable performance.	4	2.96
The performance of AI algorithms would be even lower than that of traditional techniques.	2	1.48
Others	10	7.41
Total	135	100.00

Source: Search results.

A possible way to solve part of this epistemological question is by analyzing the “*skin in the game*” (Taleb, 2020) of these researchers. Considering that there have been advances reported in the literature so that the current models provide superior returns compared to other models used for decades in the market, it must be assumed that researchers have differentiated returns on their investments. Thus, it would be advisable to observe whether these researchers allocate their own resources according to the indicated forecasts. If they do not allocate, it would be a strong indicator that they do not trust their published forecasts. Subsequently, if they use such information to allocate their own capital, it must be observed that the yield they obtain matches the returns indicated in the articles. A survey applied to this target audience could provide enlightening results.

CONCLUSIONS

This article presented a systematic literature review (mapping the state of the art) on applying AI algorithms for forecasting asset prices in the financial market. The review was developed using the Web of Science and Scopus databases, focusing on the PRISMA guidelines. The final sample corresponded to 135 articles, which were analyzed based on categories previously developed from the themes approached in the specific literature, with a special focus on the AI algorithms used.

Initially, it is important to highlight that there has been an evolution in the number of articles published on the subject since 2012, with a significant increase in the last three years. It is important to stress that many studies were observed that analyzed the markets of emerging countries, although studies from the first decade of the 2000s focused particularly on developed countries.

It was also observed that there is a preference in the studies for predicting market indices rather than other assets. Such indices can be traded as *Exchange-Traded Funds* (ETFs), allowing such studies to have theoretical and practical contributions. However, the number of articles that study stock price prediction has increased considerably in recent years, which may become a trend in the future. Such assets may also be more easily analyzed using financial indicators underused in the analyzed studies. In this sense, the little-used financial indicators indicate a gap to be explored by further studies, even in a context where there is an increase in the number of studies that focus on specific company actions. Furthermore, it is important to emphasize that only a few studies have dedicated themselves to developing models for forecasting other types of assets traded in the financial market, such as options, which represents another opportunity to be explored by future research.

Regarding the input data used as a basis for training the AI models, there was wide use of historical trading prices and technical indicators were widely used. Thus, there is a tendency for studies to explore the weak form of the EMH, according to Fama (1970). On the other hand, many recent studies have focused on sentiment analysis, which demands unstructured data and using NLP to treat it. Thus, it seems to be a new trend in the studies in the area since they can use both historical and actual data for forecasting, which increases the performance of the developed algorithms.

Regarding the algorithms used in research, there is a tendency to use more than one algorithm for forecasting prices, with the ANN being the most commonly used in general. It was found that it is usually exploited for regression purposes and is not associated with sentiment analysis, despite its potential to be used for NLP purposes.

es. The second most used algorithm in emerging countries was the SVM. These algorithms became the focus of studies in the second decade of this century. Unlike the ANN, there is a trend toward using the ANN for classification purposes.

Other algorithms employed in the studies were the DTRFs, which showed a trend to use financial indicators as training bases. In turn, studies that used the NB algorithm tended to focus on predicting stock prices rather than market indices (with a focus on ranking), as well as on the use of sentiment analysis data. The input data and the forecasted assets have shown a strong growth trend in the last few years of analysis. This may indicate the strengthening of the NB algorithm as a basis for forecasting asset prices in the next decade. Finally, the KNN algorithm was the least used as a basis in the articles. Such algorithms were more commonly used in developed countries for classification purposes.

Based on the above, the systematic literature review presented in this article features a series of contributions to the study on the use of AI algorithms to forecast asset prices in the financial market: (i) the main characteristics were identified (*e.g.*, input data, expected assets, performance measures) of models developed for this purpose; (ii) the characteristics and uses of the main algorithms used for forecasting were highlighted; (iii) the identification of associations between the different categories of analysis; and (iv) gaps in the studies were presented, indicating trends and proposing new and broad research perspectives on the phenomenon.

REFERENCES

- Almehmadi, A. (2021), "COVID-19 Pandemic Data Predict the Stock Market", *Computer Systems Science and Engineering*, Vol. 36, No. 3, pp. 451–460.
- Awan, M. J.; Rahim, M. S. M.; Nobanee, H.; Munawar, A.; Yasin, A.; Zain, A. M. (2021), "Social Media and Stock Market Prediction: A Big Data Approach", *Computers, Materials & Continua*, Vol. 67, No. 2, pp. 2569–2583.
- Cao, H.; Lin, T.; Li, Y.; Zhang, H. (2019), "Stock Price Pattern Prediction Based on Complex Network and Machine Learning", *Complexity*, pp. 1–12.
- Cavdar, S. C.; Aydin, A. D. (2020), "Hybrid Model Approach to the Complexity of Stock Trading Decisions in Turkey", *The Journal of Asian Finance, Economics and Business*, Vol. 7, No. 10, pp. 9–21.
- Chadegani, A. A.; Salehi, H.; Yunus, M. M.; Farhadi, H.; Fooladi, M.; Farhadi, M.; Ebrahim, N. A. (2013), "A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases", *Asian Social Science*, Vol. 9, No. 5.
- Chun, S. H.; Ko, Y. W. (2020), "Geometric Case Based Reasoning for Stock Market Prediction", *Sustainability*, Vol. 12, No. 17, pp. 7124.
- Colliri, T.; Zhao, L. (2019), "A Network-Based Model for Optimizing Returns in the Stock Market". *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*, 645–650. <https://doi.org/10.1109/BRACIS.2019.00118>
- Ding, G.; Qin, L. (2020), "Study on the prediction of stock price based on the associated network model of LSTM", *Int. J. Mach. Learn. & Cyber.* Vol. 11, pp. 1307–1317.
- Ecer, F.; Ardabili, S.; Band, S. S.; Mosavi, A. (2020), "Training Multilayer Perceptron with Genetic Algorithms and Particle Swarm Optimization for Modeling Stock Price Index Prediction", *Entropy*, Vol. 22, No. 11, pp. 1239.
- Faceli, K.; Lorena, A. C.; Gama, J.; Almeida, T. A. de; Carvalho, A. C. P. L. F. de., (2021), *Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina*, 2nd ed., LTC.
- Fama, E. F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work". *The Journal of Finance*, Vol. 25, No. 2, pp. 383.
- Ferreira, F. G. D. C.; Gandomi, A. H.; & Cardoso, R. T. N. (2021), "Artificial Intelligence Applied to Stock Market Trading: A Review", *IEEE Access*, Vol. 9, pp. 30898–30917.
- Gu, Y.; Shibukawa, T.; Kondo, Y.; Nagao, S.; Kamijo, S. (2020), "Prediction of Stock Performance Using Deep Neural Networks", *Applied Sciences*, Vol. 10, No. 22, pp. 8142.
- Janková, Z.; Jana, D. K.; Dostál, P. (2021), "Investment Decision Support Based on Interval Type-2 Fuzzy Expert System", *Engineering Economics*, Vol. 32, No. 2, p. 118–129.
- Maroco, J., (2010), *Análise Estatística: com utilização do SPSS*, 3rd ed., Sílabo.
- Moon, K.S.; Jun, S., Kim, H. (2018), "Speed up of the Majority Voting Ensemble Method for the Prediction of Stock Price Directions", *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 52, No. 1, pp. 215–228.
- OSF Home, (2021), *Open Science Framework (OSF)*, disponível em: <https://osf.io/?goodbye=true> (acesso em 01 maio 2021).
- Page, M. J. et al., (2021a), "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews", *BMJ*, n. 71.
- Page, M. J. et al., (2021b), "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews", *BMJ*, n. 160.

- Qian, B.; Rasheed, K., (2007), "Stock market prediction with multiple classifiers", *Applied Intelligence*, Vol. 26, n. 1, pp. 25–33.
- Ross, S. A.; Westerfield, R. W.; Jaffe, J.; Lamb, R., (2015), *Administração Financeira*, 10th ed., AMGH Editora.
- Rundo, F.; Trenta, F.; Stallo, A. L. di; Battiato, S., (2019), "Machine Learning for Quantitative Finance Applications: A Survey", *Applied Sciences*, Vol. 9, n. 24, pp. 5574.
- Sheu, H.-J.; Wei, Y.-C., (2011), "Effective options trading strategies based on volatility forecasting recruiting investor sentiment", *Expert Systems with Applications*, Vol. 38, n. 1, pp. 585–596.
- 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*, Vol. 264, pp. 71–88.
- Taleb, N. N., (2020), "Skin in the game: Hidden asymmetries in daily life". *Random House Trade Paperbacks*.
- Vijh, M.; Chandola, D.; Tikkiwal, V. A.; Kumar, A., (2020), "Stock Closing Price Prediction using Machine Learning Techniques", *Procedia Computer Science*, Vol. 167, pp. 599–606.

Received: Ago. 6, 2022

Approved: Dec. 8, 2022

DOI: 10.20985/1980-5160.2022.v17n3.1807

How to cite: Avelar, E.A., Campos, O.V., Orefici, J.B.P., Borges, S.L., Souza, A.A. (2022). Artificial intelligence and financial asset price forecasting: a systematic review. *Revista S&G* 17, 3. <https://revistasg.emnuvens.com.br/sg/article/view/1807>