International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Stock Price Forecast, with Multi Layer Perceptron Artificial Neural Networks

Cristiane Orquisa Fantin[a]*, Eli Hadad Junior[b]

[a,b]*Social and Applied Sciences Center, Mackenzie Presbyterian University, Rua Da Consolação, 930, 01302-907 São Paulo, Brazil*

[a]*Email: cris.orquisa@hotmail.com* , [b]*Email: eli.hadad@mackenzie.br*

Abstract

This study maps and analyzes the academic literature regarding the stock price forecast with multi layer perceptron artificial neural networks, through bibliometric analysis and systematic review. The adoption of these methods requires the use of RStudio, VOSViewer and Rank Words software. In the bibliometric analysis, its main laws are verified [1,2,3]. As a result of the bibliometric analysis, the most frequent keywords are forecast, model(s), neural network and market, and most authors are associated with institutions located in China. Concerning the systematic review results, research on the different training methods, as well as the different data pre- and post-processing models are urgent, as they may reduce risks and maximize investors' returns. As for the directions for new research, further studies are suggested on the different models and architectures of multi-layer perceptron artificial neural networks, associations with other statistical and intelligent models, and research focusing on specific market segments, such as industry, energy and civil construction.

Keywords: Stock price forecast; Multi layer perceptron; Back propagation; Bibliometric analysis; Systematic review.

1. Introduction

The financial market is characterized by being a dynamic, complex and non-linear system, with intense data, noise, non-stationary, unstructured nature and with a high degree of uncertainties [4]. Because so many factors interact simultaneously, such as political events, macro and microeconomic conditions and investor´s expectations, predicting these movements is a very challenging task.

---

* Corresponding author.

The increasing role that the stock market plays in the world economy stimulates the development of research focused on the construction of theories involving the theme of stock price forecasting, and accurate methods are crucial for the investors' portfolio management. Assessing expected returns relative to total exposure assumes that portfolio managers understand the portfolio distribution. Experts can model the influence of tangible assets in relation to market value, but not of intangible assets, such as rights, experiences or brand value. An important ally in trying to minimize risk in relation to exposure, artificial intelligence and machine learning with their neural networks have taken a great leap in quality in recent decades, improving detection, diagnosis, prediction and problem solving [5]; that is because in this market, future events are at least partially dependent on past events and data [11], and not completely random.

Perhaps, the most common type of neural network [21] is the multi layer perceptron (MLP), which is a feed-forward neural network, formed by two or more layers of neurons. It is a type of artificial neural network (ANN) applied in several fields, having obtained successful results in predictions. The reason why this model is chosen for this study, among many others available, is that, according to the conclusions of recent studies [11], MLP is an ANN method providing flexible and consistent results. This being so, this study aims at mapping and analyzing the published academic literature on stock price forecasting, by using multi layer perceptron artificial neural networks. For such, a bibliometric analysis and a systematic review of the literature on the subject are carried out, during the period from January 1, 1945 to March 28, 2021, with a final sample of 26 articles. Bibliometric analysis refers to quantitative analysis, which is developed through counting frequencies and co-citations. The systematic review, a qualitative analysis, considers the correlation between more significant themes, although still little studied by the academy. The research base used comes from the Web of Science (WoS) database, and both the bibliometric analysis and the systematic review require the use of R, RStudio, Biblioshiny, VOSViewer and Rank Words software. In the bibliometric analysis, the verification of the main laws is adopted [1,2,3].

The literature review is presented in item 2, with the identification of theories and methods for stock price forecast with multi layer perceptron artificial neural networks, mentioned in the papers of final sample. The bibliometric analysis and systematic review methodologies are described in item 3 and, in item 4, the results of both methodologies are reported, with descriptive statistics of the most relevant characteristics of the papers in final sample and the knowledge gaps on this topic. Item 5 presents the conclusions, directions for future studies and limitations of this research.

2. Materials and methods

The environment in which this study focuses, the one of stock market, is quite noisy and non-linear, but with not totally random movements [30]. The fluctuation of the daily closing prices of a company's shares is strongly influenced by certain predictors recorded over time, such as the volume of transactions, the number of shares, the price range or the daily exchange rate, for example [31]. In this context, researches with ANN and learning algorithms are quite adequate, as they demonstrate the ability to retain previously learned information and to acquire additional knowledge when new data are introduced [4]. Neural networks mimic the way the human brain works. MLP is a supervised learning algorithm capable of learning from a training dataset. Given a set of

resources and a target output, MLP can learn a nonlinear function for regression. The difference between classic multiple linear regression and MLP for regression is the number of non-linear layers, also known as hidden layers [5].

MLP ANNs have been being widely applied to supervised learning problems [18], and in particular, they offer excellent performance in predicting stock market index values [19]. This is because they are algorithms whose most striking feature is extracting non-linear and hidden relationships from the data. MLP ANNs are neural networks that feed-forward with units connected to each other, called neurons. These sensory units are organized into input layers, an arbitrary number of hidden layers of processing units (perceptrons) and output layer [11]. The input layer is fed by external information and transfers them, through connections, to the units of the first hidden layer. The latter ones calculate their activations on the information and pass the results to the units (neurons) of the following layers [32]. The network has to be trained to produce the correct output with minimal error [23]. The training itself consists of two phases, the forward and backward calculation pass. In the forward pass, external information (the so-called training vectors) are introduced into the network and their responses are obtained. These responses are combined with the information provided (that is, the target vectors indicate the optimal response for each training vector), in order to define the network error in relation to a loss function. This error is then propagated back to the neural network and used in the reverse pass direction, to update the weights, the network parameters, that is, the cycle is repeated and the weights are adjusted continuously until the error is minimized [23], that is, reducing the mean square error between the network prediction output and the target output.

This is achieved by exploring the differentiable property of the neuron activation functions, following a descending gradient learning approach, called error back-propagation [21], with the accumulated information being propagated in the network up to the output layer. In short, the model is trained to learn the correlations between inputs and outputs and enabled to adjust parameters, weights and bases, from time to time, in order to minimize errors in the training process [18]. If the learning curve stabilizes, after many iterations, at an error level not considered as acceptable, it is the time to rethink on the network topology (more hidden layers or a totally different topology) or on the training procedure (another gradient search with more sophisticated techniques).

In this context, designing a good network architecture for the studied problem is not a simplistic task. The number of layers, of neurons per layer and the selection of activation functions directly affect the model performance. The refinement process may be time consuming [33], as there is no formal method to derive an MLP network configuration for a given classification task, nor a direct method to find the best suited final structure to the modeling process. This is necessarily done through the performance of iterative tests, with several parameters, different architectures, in which only the most assertive structure is kept. The information quality with which the network is fed, as much or more, is reflected in the accuracy of the output layer response. The selection of information that will be provided to the input layer is a key factor in an intelligent decision system design, because even the best model will have a poor performance if the resources are not well chosen. Specific methods shall be used in selecting the really relevant information [5].

2.1. Methodology

The purpose of this study is to answer the question – using multi layer perceptron artificial neural networks, can we obtain reliable stock price forecasts? To do so, the 7 steps described below are implemented. Steps 1 to 5 meet both the bibliometric analysis and systematic review methodologies, and steps 6 and 7 refer to the systematic review.

Step 1 – Choosing the database. Sample articles are sourced from WoS, the world's leading citation database. It contains records of papers published in high impact factor scientific journals, classified by the Journal Citation Reports (JCR). Step 2 - Usage of initial research parameters from WoS, for the period from January 1, 1945 to March 27, 2021. 86 papers are identified, based on the keyword variations, such as stock, market, multi layer perceptron, forecast, predictive, regression, supervised, learn and back propagation. Exclusions are performed by applying filters in WoS itself, resulting in an intermediate sample of 35 papers, as shown in Table 1.

Table 1: Sample evolution through WoS´ filters.

Sign	Description	Articles
(+)	Keywords like: "stock* market*" and "MLP" or "stock* market*" and "multi-layer perceptron" or "stock* market*" and "multi layer perceptron" or “forecas* stock*” and “MLP” or “forecas* stock*” and “multi-layer perceptron” or “forecas* stock*” and “multi layer perceptron” or “predictiv* regression*” and “MLP” or “predictiv* regression*” and “multi-layer perceptron” OR “predictiv* regression*” and “multi layer perceptron” OR “supervis* learn*”and “multi-layer perceptron*” and “backpropagation”	86
(-)	Type of document: other than “articles”	44
(-)	Language: other than “English”	1
(-)	WoS´categories: other than “computer science artificial intelligence”, “computer science theory methods”, “computer science information systems”, “computer science interdisciplinary applications”, “computer science hardware architecture”, “computer science software engineering”, “multidisciplinary sciences”, “economics”, “business finance”, “mathematics interdisciplinary applications”, “operations research management science”, “management”, “mathematics”, “social sciences mathematical methods”, “engineering multidisciplinary”	6
(=)	**Intermediate sample**	**35**

Step 3 - Exclusion of unrelated papers. After the initial reading of the papers´ abstract, introduction and conclusion - to verify if they are in accordance with the defined theme - 09 out of the 35 papers in the intermediate sample are excluded. The reasons for these exclusions are because they are not related to stock price forecast, but to forecast of companies´ bankruptcy (02), to Bitcoins price forecast (01), to exchange rate

forecast (01), to sigmoid calculation (01), to physical temporal analyses (01), to advanced supervised learning in multilayer perceptrons (01), or for not being available in the researched sources (02). Thus, the final sample consists of 26 papers [6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31]. Step 4 - Database creation and papers´ collection. The 26 articles in the final sample are obtained through the following academic research bases: Science Direct, Web of Science and Google Scholar. From this, the following information are collected to capture the paper´s general data: title, author's name, affiliated institution and authors/researchers´ country of origin, journal name, volume and issue number, initial and final page, year of publication, country of data source and number of years of sample data, keywords, Digital Object Identifier (DOI), Journal of Economic Literature (JEL) and number of citations of articles in the WoS´ database.

Step 5 – Bibliometric analysis. By using the R, RStudio, Biblioshiny and VOSviewer software, papers´data are analyzed for the concoction and analysis of tables and relationship/co-citation maps. The analyses performed by these tools are complemented by the verification of the main laws of bibliometrics, (i) Zipf 's law [1], which refers to the keyword frequency´s categorization and estimation, carried out with the help of Rank Words software for Goffman 's T point calculation [32], which is the transition point from low-frequency words to high-frequency words, a region that, theoretically, concentrates the words with high semantic load, (ii) Bradford's law [2], which verifies the journals producing many papers, as opposed to those producing few papers on a given topic, and (iii) Lotka's law [3], which identifies the researchers who have the highest frequency of production in a given knowledge area. Step 6 - Papers´ reading and coding. It deals with the identification of the objectives, sample, methods and contributions of the articles. In addition, they are classified and coded into structured categories and subcategories, as shown in Table 2. Each of the 10 categories has its own subcategories and a single article may be classified into more than one subcategory.

Table 2: Matrix of (sub) categorization.

Categories	Subcategories	Definition
1.Main topic / study focus	A – Stock price forecast with multi layer perceptron	A - Stock price forecasting with a Multi Layer Perceptron ANN, which is a neural network with one or more hidden layers and an undetermined number of neurons.
	B – Stock price forecast with other artificial neural networks and results compared to MLP ANN	B – Stock price forecast research with Other Artificial Neural Networks, whose resultsare compared to the results with MLP.
	C – Stock price forecast with combined neural networks, including MLP	C - Stock price forecasting with hybrid models among MLP artificial neural networks and other neural networks
	D – Others	D - Other topics not related to subcategories 1A to 1C.
2. Research methods	A – Artificial Neural Networks	A - Computational techniques presenting a mathematical model inspired in the neural structure of intelligent organisms. They are capable to perform both the machine learning and the pattern recognition.

Categories	Subcategories	Definition
2. Research methods	B - Statistical models for time series (ARMA, ARIMA, GARCH)	B - In time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both models are adjusted to time series data to better understand the data or to predict future points in the series. Generalized autoregressive conditional heteroskedasticity (GARCH) is a statistical model used in the analysis of time series data, in which the variance error is believed to be serially auto-correlated.
	C – Others	C - Other research methodologies not related to subcategories 2A or 2B.
3. Neural networks used in research	A – MLP	A - Multi Layer Perceptron is a neural network with one or more hidden layers and an undetermined number of neurons.
	B – LSTM	B - Long short-term memory (LSTM) is an artificial recurrent neural network, used in the deep learning field. It can not only process individual data points, but also entire data sequences.
	C – HONN	C - Higher Order Neural Networks (HONN) can simulate high frequency nonlinear businesses and deal with discontinuities in the training dataset input.
	D – CNN	D - The convolutional neural network (CNN or ConvNet) is a class of deep neural networks, most commonly applied to the analysis of visual images. It has applications in image and video recognitions, natural language processing, brain-computer interfaces, and financial time series.
	E – Others	E - Other methods not related to subcategories 3A to 3D.
4. Theories related to the objectives of the papers	A – Random walk dilemma theory	A - The forecasts generated by the models show a characteristic delay of one step in relation to the original time series data. This behavior has been seen as a dilemma in relation to the financial representation of time series, having been stated that the series follow a random walk model and, therefore, cannot be predicted. The main reason behind this theory is that researchers believe that stock prices were mainly affected by news and politics.
	B – Theories on stock market volatility	B - Several studies discuss about the price x volume relationship and its volatility.
	C – Non mentioned theories	C – Theories are not mentioned.
5 – Types of analysed data	A – Daily stock closing prices	A - This is the quotation of the last trade closed on the day, for a given share on the stock exchange trading floor.
	B – Daily stock opening and closing prices	B - The opening price quotation represents the price at which the first trade with a certain asset was made in the trading session in question. The closing price is the price at which the last trade was made on the trading session in question.

Categories	Subcategories	Definition
5 – Types of analysed data	C – Opening, closing, highest, lowest and trading volume prices	C – With respect to a given trading floor day, the opening price quotation represents the price at which the first trade was made, the closing price represents the price at which the last trade was carried out, the highest price is the maximum price at which the shares were traded, and the lowest, is the minimum price at which a deal was closed. In the concept of volume, it represents the total of transacted resources.
	D – Daily closing of Stock Price Index	D - This is the last quotation on the day, for a given stock price index.
	E – Other associations	E - Other methods not related to subcategories 5A to 5D.
6. Data source	A – United States and Canada	A - Data from operations in the United States and /or in Canada.
	B – Europe	B – Data from operations in Europe.
	C – Asia / Oceania	C - Data from operations in Asia and /or Oceania.
	D – Latin America	D - Data from operations in Latin America.
	E – Africa	E - Data from operations in Africa.
	F – Non mentioned	F - Data source is not mentioned
7. Analysis period	A – Up to 5 years	A - Data between 0 and 5 years.
	B - From 6 to 10 years	B – Data between 6 and 10 years.
	C – More than 10 years	C – Over 10 years.
	D - Not applicable/not informed	D - Studies that do not inform the analysis period.
8. Results	A - New perspectives or new conclusions	A - Presentation of a new theory or new conclusions on topics previously discussed.
	B - Similar conclusions to previously presented works	B - Studies that do not present new perspectives or new conclusions.
	C - Others	C - Other results not related to subcategories 8A or 8B.
9. Conclusions on study objectives	A – Confirmation of the main objective	A - Study confirms the main objective proposed by the author.
	B – Non confirmation of main objective	B - Study does not confirm the main objective proposed by the author.
	C – Inconclusive result regarding the main objective	C - Study presents inconclusive result regarding the main objective proposed by the author.
	D - Not applicable	D - Other conclusions not related to subcategories 9A to 9C.
10. Directions for future studies	A– Examining other models using MLP	A - Research using MLP, but with different patterns.
	B – Examining different artificial neural network models	B- Research using other artificial neural networks for this purpose.
	C – Evaluating other stock price forecast models	C - Research considering other stock price forecast models, in addition to artificial neural networks.
	D – Non mentioned	D - Author does not mention directions for future studies.

Step 7 – Systematic review. After coding the matrix of (sub) categorization in Table 2 from the 26 papers in the final sample, the frequency count of the subcategories is performed to make the identification of knowledge gaps easier. From there, these gaps are compared to the subcategories of Category 10 - Directions for future studies, to obtain gaps subject to new studies.

3. Analysis of results

The bibliometric analysis results are presented in item 3.1, mentioned in Step 5 of the Methodology. In item 3.2 are the results of the systematic review, with descriptions in Steps 6 and 7 of item 2.1.

3.1 Bibliometric Analysis

According to Figure 1, the final sample is composed of 26 articles, distributed between the years 2005 and 2021, obtained from the WoS database, and in this period, the publication of up to 5 papers per year, dealing with the stock price forecast theme using MLP ANN.

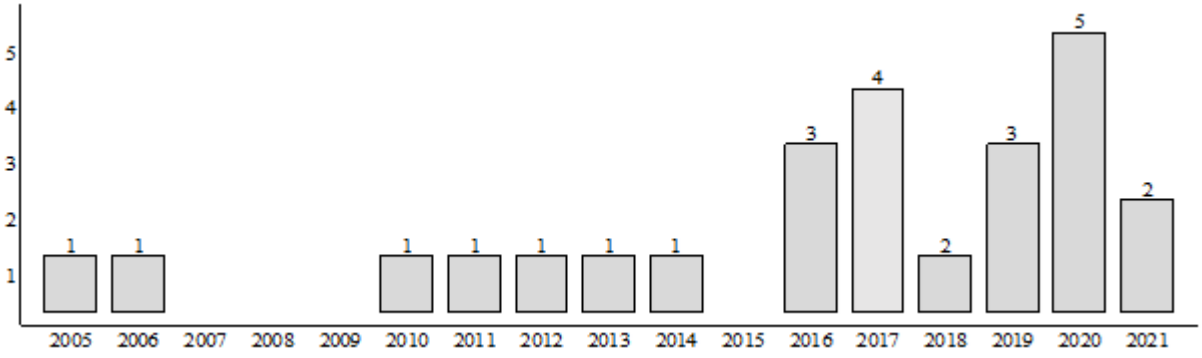

Figure 1: Annual distribution of articles

Among the 69 keywords identified in the 26 articles in the final sample of the study, prediction (10.1%), model / models (7.2%) and neural network (5.8%) stand out.

Based on the analysis of all articles in the final sample, Rank Words software provides a list of words in descending order of frequency. From this base, the words that are repeated only once are identified to calculate the Goffman´s T point [32]. Then, the words above the classification indicated by that point are located. In the 26 articles, this point varies between 43.75 [5] and 25.08 [22], with an average of 35.31. Once the Goffman´s T point [32] of each article has been calculated, to analyze the region, in which the most adherent words to the main theme of the text are located, is possible. Next, those words that are not relevant to the study are excluded, such as prepositions, definite and indefinite articles, pronouns and adverbs, and the words having the highest frequency are classified. In article [8], the word model has the highest frequency, that is, it is repeated 295 times in a text with 12,315 words. In addition, it appears that the word model is the most cited, in 06 of the 26 (23%) articles in the final sample. Words such as stock, data, forecasting, trading, regression, prediction, neural, networks, market, interval, index, efficiency and algorithm also have high frequency, in the 20 (77%) other

articles in the final sample. Also, 56 authors are identified for the final sample with 26 articles, and among them, only 5 publish individually. These 56 authors are associated with institutions from 16 countries, with 17 of them (30%) being associated with institutions located in China, 6 (11%) with Turkish institutions, 4 (7%) Brazilian, 4 (7%) Indian, and the rest, distributed among Egypt, the Czech Republic, Thailand, Denmark, Finland, South Africa, United Arab Emirates, Iran, Poland, Romania, Bahrain and Lebanon. Out of the 219 total citations, 77 (35%) are from articles written by authors associated with institutions located in Brazil. The other citations are from authors who are associated with the following countries: Bahrain 40 (18%), India 26 (12%), China 18 (8%), Turkey 18 (8%), Iran 16 (7%), Egypt 4 (2%), Lebanon 4 (2%), South Africa 3 (1.4%), United Arab Emirates 3 (1.4%), Czech Republic 2 (0.9%), Thailand 2 (0.9%), Finland 2 (0.9%), Poland 2 (0.9%), Denmark 1 (0.5%) and Romania 1 (0.5%). Bradford's law [2] verifies that there are few journals producing many papers and many journals producing few papers on a given topic. Thus, if journals are classified in descending productivity order, they may be distributed in zones having a variation in the ratio 1: n: n^2, and so on. These zones are formed by dividing the total of published articles by two. Zone A is identified as the core of the subjects, composed of a journal with two publications, in zone B there are journals with only one publication. 8% of the publications on the theme of stock price projection with multi layer perceptron neural networks are in zone A and this percentage is limited to only one academic journal, the Shams Engineering Journal, with an individual amount equal to 2. About the co-citation network among the journals in the final sample of 26 articles, the most cited are the journals Expert Systems With Applications with 20 co-citations; Econometrics and Neural Computing and Applications with 11 co-citations each; Physica A, Neurocomputing and Applied Soft Computing, with 10 co-citations each. This fact demonstrates that the journals publishing the most on a given topic are not, necessarily, the most cited, emphasizing the importance of the papers. The size of the nodes indicates the relevance of these words in the articles, the thickness of the lines is related to the connections strength between them, and the colors separate the groups. Among the papers on the topic of stock price forecast with multi layer perceptron neural networks and their citations, the article by Maia and Carvalho [20] stands out, with 51 (23%) of the citations and an annual average of 4.636. The article presents three approaches for forecasting time series, which are based on MLP neural networks, Holt´s exponential smoothing methods, and on a hybrid methodology combining the MLP and Holt models, respectively. The second and third most cited papers are those by Mostafa [23] and by Araujo and Ferreira [7], with 40 (18%) and 26 (12%) citations, respectively. Lotka 's law [3] states that a small number of authors produce many papers and that the production obtained by this small number of researchers is equal, in quantity, to the performance of the others. This theory is called the Inverse Square Law. It appears that the 26 papers in the final sample are produced by 56 authors, and one author publishes 4 papers, two authors publish 3 papers and the other authors publish a single article. 20% of the authors (including those who publish the most) are responsible for 29% of the publications, however, there is not a smaller number of researchers equaling, in quantity, the performance of the others, making Lotka 's Law [3] impossible to be confirmed.

3.2 Systematic review

The systematic literature review seeks to identify knowledge gaps related to the topic of this study – stock price forecast with multi layer perceptron neural networks. In Step 6 of Item 2.1 – Methodology, a matrix of (sub) categorization is defined – see Table 2. For each of the 26 papers in the final sample, there is the identification

of up to 3 subcategories per category. Thus, the frequency count is made in relation to the total of the subcategories, and not to the total of the 26 papers. Figure 2 highlights the subcategories with most and least frequency – those having the potential to be prioritized in future research. In category 1, the theme of stock price forecast with MLP ANN (A) is central to 23% of the papers analyzed, and these researches differ from each other in terms of their refinement methods and architectures. This is because there is no standard method to find the most suitable final structure for the modeling process, so, iterative tests are performed, with different amounts and qualities of input information, because even with the best architecture, performance will be compromised, if input features are not well selected. Furthermore, the number of layers and neurons per layer also vary, and all of this is directly reflected in the assertiveness of the output layers' responses. The stock price forecast with other ANNs and results compared to MLP ANN (B), is the theme of 35% of the papers, and there are also an additional 12% studying both the stock price forecast with MLP ANN and with other ANNs, and they compare these results (A+B). There is also the opportunity of studying stock price forecast through hybrid models, with MLP ANN and other neural networks, as shown by 19% of the studies. Category 2 indicates that 35% of the papers focus their research methods on ANN, which are computational techniques presenting a mathematical model inspired in the neural structure of intelligent organisms (A), another 35% combine, as a research method, ANN and statistical models for time series (B), such as ARMA, ARIMA, GARCH, and 23% use other research methods (C). There are also 8% that combine ANN, statistical models for time series and other research methods (A+B+C). As for the neural networks used in the research, category 3 also has combined subcategories. Thus, 58% of the papers use MLP artificial neural networks (A), 15% use, in addition to MLP ANN, the LSTM (A+B), which is an artificial recurrent neural network used in the field of deep learning, which can process not only individual data points, but also entire data sequences. 8% combine ANN and other neural networks (A+E), and 8% combine MLP, LSTM and CNN (A+B+D). Regarding the theories supporting stock price forecast with MLP ANN, category 4, the random walk dilemma theory (A), mentioned in 3 papers (12%) and its association with stock market volatility theories, mentioned in one paper (4%, A+B) stand out. The other papers do not mention the theories associated with their objectives or hypotheses. The random walk dilemma theory, or random walk hypothesis, mentioned in 3 articles, is a statistical phenomenon, according to which variables move randomly. The random walk theory applied to trading was introduced by Burton Gordon Malkiel, in his book A Random Walk Down Wall Street, in 1973 [34], and it states that stock prices move randomly, and therefore, any attempt to predict these prices is useless. The stock market volatility is studied, in particular the relationship between price and volume traded [46], which are the result of a single mechanism. Based on these results, to state that the supply and demand theory is useful in the stock market analysis [36] is questionable. In category 5, types of analyzed data, 62% of the papers fall into the subcategory using daily stock closing prices as the main data source for the studies, 15% use opening and closing prices, highest and lowest prices and traded volume, 8% use daily opening and closing prices of stocks and 4% use daily closing of stock price index. Variations in the types of analyzed data are directly related to the origin, objectives and depth of each of researches.

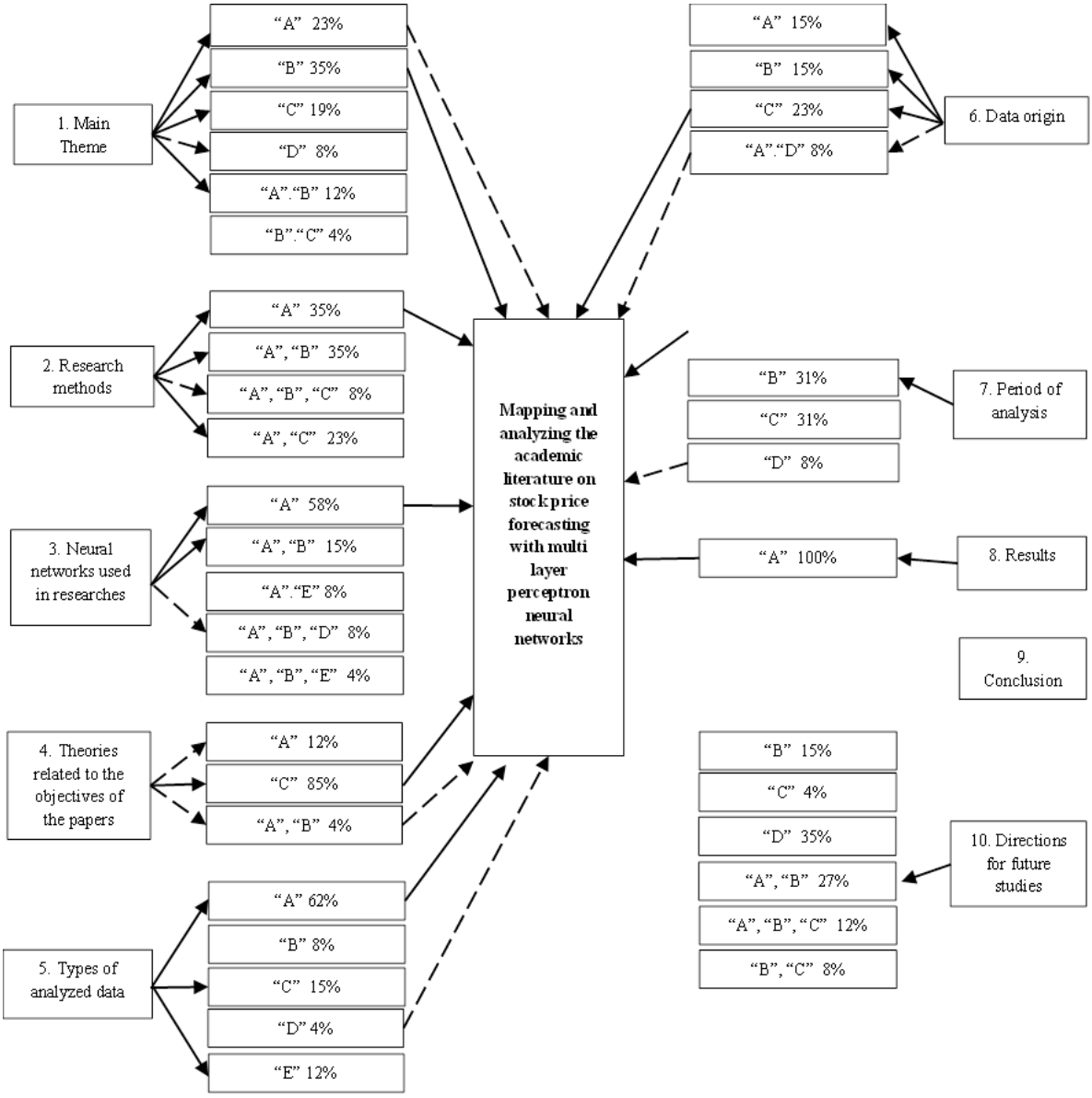

Figure 2: Analysis of (sub) categories to identify knowledge gap.

Notes: Subcategory with most frequency in each category ——————→

Subcategory to be prioritized in researches – – – – – – – →

As far as the data origin is concerned, category 6 demonstrates dispersion in its subcategories. 23% of the data originate from Asia/Oceania (C), 15% originate from the United States and Canada (A), and this percentage is repeated for Europe (B), and the United States, Canada, Europe, Asia and Oceania (A+B+C), demonstrating that the studies focus especially on developed countries. Thus, there is great opportunity for investigations with data from Latin America and Africa. Concerning the period of analysis, category 7, the subcategories up to 5 years (A), 06 to 10 years (B) and over 10 years (C) are distributed equally, with a frequency of 31% each. Category 8 concludes that 100% of the papers present new perspectives or new conclusions, no article presents similar conclusions to previously presented works. Furthermore, according to category 9, 100% of the papers confirm

the main objective of their studies. Finally, category 10 indicates the directions for future studies or knowledge gaps - according to the perspective by the authors of the 26 papers in the final sample. In 15% of the subcategories, the authors suggest carrying out studies on stock price forecast with other ANN models (B). 27% suggest carrying out tests with other MLP ANN models, added to other ANN models (A+B). 12% suggest adding to this, the assessment of other stock price forecasting models (A+B+C).

From the above, the main aspects to be investigated in future studies, related to the theme of stock price forecast with MLP ANN are highlighted: a) different architectures and processes of data selection in MLP, i.e., quality of input data, of the number of layers, of neurons per layer and of output layers different from those already researched; b) the sum of the MLP ANN results to other ANN models and/or to other stock price forecasting models; c) combine the study of the daily stock closing prices and the variations and the relationships among the opening, the highest and the lowest prices, the volumes traded and the stock price indices; and d) data from the Latin American and African markets. Regarding the main theme/focus of the study, subcategory A - Stock price forecast with multi layer perceptron - is present in 23% of the analyzed subcategories. Stock price forecast with other artificial neural networks, and results compared to MLP ANN, category B, present in 9 (34.6%), Stock price forecast with combined neural networks, including MLP, category C, present in 5 (19.2%) and other combinations of analyses, in 6 (23.1%) of the sample. As for the research methods, there is a predominance of subcategory A - Neural networks (34.6% or 9 occurrences) over the other research methods present in the sample articles. The other articles also present analyses based on neural networks, but combined with statistical analyses of time series (A, B: 34.6% or 9 occurrences), or even with other research methods. Among the neural networks used in the research, subcategory A – MLP is the most present, occurring separately in 15 (57.7%) of the samples, and jointly, in another 10 (38.5%). Regarding the theories related to the objectives of the articles, the data analyzed with greater concentration refer to subcategory A - Random Walk Dilemma Theory, with 3 isolated occurrences and 1 occurrence together with subcategory B - Combination theory. The remaining 22 articles in the sample do not mention the theories related to their objectives. The types of data analyzed with the highest concentration are those of subcategory A - Daily stock closing prices (61.5%), C - Opening and closing prices, highest and lowest prices and traded volume (15.4%). Subcategories B, D and E concentrate the other types of analyzed data. The data originate more concentrated in Asia/Oceania (23.1%), followed by data from the United States/Canada (15.4%) and from Europe (15.4%). However, noteworthy is that there are studies in a relevant volume (46.2%) including combined data from more than one of these regions. As regards the period of analysis of the studies´ data sample, 30.8% of the data belong to subcategory A - Up to 5 years, the same percentage for categories B - From 06 to 10 years and for category C - More than 10 years. Only 7.7% do not indicate the period of analyzed data. Directions for future studies: 38.5% of the papers exclusively recommend the examination of different ANN training methods, 11.5% recommend associating different ANN training methods with the evaluation of different models of pre and post data processing, 7.7% recommend associating different ANN training methods with other statistical and intelligent models, and 34.6% do not mention any direction for future studies.

4. Conclusion

The stock market is intense in dynamics, noise and uncertainties, an environment where companies, politics,

economy and investors interact. Since so many factors relate at the same time, stock price forecast is a very complex task. MLP is an ANN method providing flexible and consistent results, and above all, it is successful in forecasting. The complexity in this process is inserted in the identification of the most adequate architecture and variables, so that the learning curve stabilizes after the iterations at an error level considered as acceptable, otherwise, it is time to rethink about the network topology (more hidden layers or an entirely different architecture) or on the training procedure (another gradient research with more sophisticated techniques). Considering the scarcity of research on the subject, this study performs a bibliometric analysis and a systematic literature review, regarding stock price forecasting with MLP ANN. As a result, there is the identification of knowledge gaps to be filled by a proposal for a future research agenda related to this topic.

The initial sample consists of 86 papers, which after adopting the exclusion criteria mentioned in item 3, Methodology, is reduced to 26 final studies, obtained from the WoS database. Bibliometric analysis presents quantitative data through graphs, relationship maps and tables on the bibliometrics´ main dimensions and laws, namely: keywords – Zipf's law [1], journals - Bradford's law [2] and authors – Lotka's law [3]. These checks take place through RStudio, Biblioshiny, VOSviewer and Rank Words software. In turn, the systematic review identifies the frequency of (sub) categories defined for the final sample. Its verification allows the perception of what combinations of subcategories are feasible for future investigations.

As a result of the bibliometric analysis, the words forecast, model (s), neural network and market are the most used. 56 authors are responsible for the 26 articles analyzed, and 17 of them (30%) are associated with institutions based in China. The journal that stands out the most in the final sample is ain Shams Engineering Journal, and the most cited article is by Maia and Carvalho [20], with 51 citations and annual average of 4.636. In it, the authors empirically investigate the daily stock closing prices of 15 companies, traded on the B3 (Brazil) and Nasdaq (United States) stock exchanges, and their experiments suggest that MLP, Holt and a hybrid model may be successfully used to forecast financial time series. Finally, it appears that 20% of authors are responsible for 29% of publications, making Lotka's Law impossible to be confirmed. With regard to the systematic review, the analysis is performed by identifying the frequency of the subcategories described in Table 2. For stock price forecasting with MLP ANN, research on the different training methods and on the different models of data pre and post processing are urgent, as they can reduce risks and maximize returns for investors. Another concern refers to obtaining data for inputs in the evaluation models. The scarcity of reliable data outside the stock exchanges of developed countries leads to the restriction of studies beyond these markets. Regarding the paths for future research, the authors emphasize the recommendation for examining different MLP ANN training methods and their associations to other statistical and intelligent models.

As far as the limitations of this study are concerned, the obtained results are restricted to the selection criteria defined in item 3, Methodology, and to the articles made available by WoS. Thus, for the evolution of this research, more detailed studies on stock price forecasting with MLP ANN are suggested, under the scope of its different architectural models, associations with other statistical and intelligent models, and research focusing on market´s specific segments, such as industry, energy and civil construction.

References

[1] G.K. Zipf. Human behavior and the principle of least effort: an introduction to human ecology. *Cambridge, Mass: Addison-Wesley Press, 1949.*

[2] S.C. Bradford. "Sources of information on scientific subjects". Engineering: An Illustrated Weekly Journal, 137, 85-86, 1934. DOI: 10.1177/016555158501000407

[3] A.J. Lotka. "The frequency distribution of scientific productivity". Journal of the Washington Academy of Sciences, 16(12), 317-323, 1926.

[4] D. Lunga e T. Marwala. "Online forecasting of stock market movement direction using the improved incremental algorithm". Neural Information Processing, PT3, Proceedings, Lecture - Notes in Computer Science, 4234, 440-449, 2006

[5] S. Belciug e A. Sandita. "Business Intelligence: Statistics in predicting stock Market". Annals of the University of Craiova, Mathematics and Computer Science Series, 2017, 44(2), 292-298. ISSN: 1223-6934

[6] M. Al-Shboul e N. Alsharari. "The dynamic behavior of evolving efficiency: Evidence from the UAE stock markets". The Quarterly Review of Economics and Finance, 73, 119–135, 2019. DOI: 10.1016/j.qref.2018.05.007

[7] R.A. Araújo e T.A.E. Ferreira. "A Morphological-Rank-Linear evolutionary method for stock market prediction". Information Sciences, 237, 3–17, 2013. DOI: 10.1016/j.ins.2009.07.007

[8] M. Bildirici e O. Ersin. "Modeling Markov Switching ARMA-GARCH Neural Networks Models and an Application to Forecasting Stock Returns". Hindawi Publishing Corporation - The Scientific World Journal, Article ID 497941, 2014. DOI: 10.1155/2014/497941

[9] S.K. Chandar. "Grey Wolf optimization-Elman neural network model for stock price prediction". Soft Computing, 25, 649–658, 2021. DOI: 10.1007/s00500-020-05174-2

[10] Q. Chen, W. Zhang e Y. Lou. "Forecasting Stock Prices Using a Hybrid Deep Learning Model Integrating Attention Mechanism, Multi-Layer Perceptron, and Bidirectional Long-Short Term Memory Neural Network". IEEE Access, 8, 117365-117376, 2020. DOI 10.1109/ACCESS.2020.3004284

[11] U. Demirel, H. Cam e R. Unlu. "Predicting Stock Prices Using Machine Learning Methods and Deep Learning Algorithms: The Sample of the Istanbul Stock Exchange". Gazi University Journal of Science, 34(1), 63-82, 2021. DOI: 10.35378/gujs.679103

[12] G. Ding, G. L. Qin, L. "Study on the prediction of stock price based on the associated network model

of LSTM". International Journal of Machine Learning and Cybernetics, 11, 1307–1317, 2020. DOI: 10.1007/s13042-019-01041-1

[13] N. Eiamkanitchat, T. Moontuy, e S. Ramingwong S. "Fundamental analysis and technical analysis integrated system for stock filtration". Cluster Comput, 20, 883–894, 2017. DOI: 10.1007/s10586-016-0694-2

[14] I.M. Hamed, A.S. Hussein, M.F. Tolba (2012). "An Intelligent Model for Stock Market Prediction". International Journal of Computational Intelligence Systems, 5 – 4, 639-652. Atlantis Press, 2012.

[15] U. JuHyok, P. Lu, C. Kim, U. Ryu e K. Pak. "A new LSTM based reversal point prediction method using upward / downward reversal point feature sets". Chaos, Solitons and Fractals, 132 – 109559, 2020. DOI: 10.1016/j.chaos.2019.109559

[16] A. Karathanasopoulos. "Modelling and trading the English stock Market with new forecasting techniques". Economics and Business Letters, 5 (2), 50-57, 2016. DOI: 10.17811/ebl.5.2.2016.50-57

[17] M. Khashei e Z. Hajirahimi. "Performance evaluation of series and parallel strategies for financial time series forecasting". Financial Innovation, 2-24, 2017. DOI 10.1186/s40854-017-0074-9

[18] Z.X. Li e V. Tam. "A Machine Learning View on Momentum and Reversal Trading". Algorithms, 11(11), 170, 2018. DOI: 10.3390/a11110170

[19] W.J. Lu, J.Z. Li, Y.F. Li, A.J. Sun e J.Y. Wang. "A CNN-LSTM-Based Model to Forecast Stock Prices". Complexity, 6622927, 2020. DOI: 10.1155/2020/6622927

[20] A.L.S. Maia e F.D.T. de Carvalho. "Holt's exponential smoothing and neural network models for forecasting interval-valued time series". International Journal of Forecasting, 27(3), 740-759, 2011.

[21] Y. Makinen, J. Kanniainen, M. Gabbouj e A. Iosifidis. "Forecasting jump arrivals in stock prices: new attention-based network architecture using limit order book data". Quantitative Finance, 19(12), 2033-2050, 2019. DOI: 10.1080/14697688.2019.1634277

[22] K. Michalak e P. Lipinski. "Prediction of high increases in stock prices using neural networks". Neural network world, 15(4), 359-366, 2005.

[23] MM. Mostafa. "Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait". Expert Systems with Applications, 37(9), 6302-6309, 2010. DOI: 10.1016/j.eswa.2010.02.091

[24] S.C. Nayak. "Escalation of Forecasting Accuracy through Linear Combiners of Predictive Models". EAI Endorsed Transactions on Scalable Information Systems, 6(22), 2, 2019. DOI: 10.4108/eai.10-6-2019.159345

[25] S.C. Nayak, B.B. Misra e H.S. Behera. "An Adaptive Second Order Neural Network with Genetic-Algorithm-based Training (ASONN-GA) to Forecast the Closing Prices of the Stock Market". International Journal of Applied Metaheuristic computing, 7(2), 39-57, 2016. DOI: 10.4018/IJAMC.2016040103

[26] S.C. Nayak, B.B. Misra e H.S. Behera, H.S. "Artificial chemical reaction optimization of neural networks for efficient prediction of stock market índices". Ain Shams Engineering Journal, 8(3), 371-390, 2017.

[27] S.C. Nayak, B.B. Misra e H.S. Behera. "Artificial chemical reaction optimization based neural net for virtual data position exploration for efficient financial time series forecasting". Ain Shams Engineering Journal, 9(4), 1731-1744, 2018. DOI: 10.1016/j.asej.2016.10.009

[28] P. Tufekci. "Classification-based prediction models for stock price index movement". *Intelligent Data Analysis, 20(2),* 357-376, 2016. DOI: 10.3233/IDA-160809

[29] M. Vochozka, J. Horak e T. Krulicky. "Innovations in Management Forecast: Time Development of Stock Prices with Neural Networks". *Marketing and Management of Innovations, 2,* 324-339, 2020

[30] P.D. McNelis. "Neural Networks in Finance: Gaining the predictive edge in the market". Elsevier Academic Press 1,2005. ISBN-10: 0124859674, ISBN-13: 978-0124859678

[31] I. Pardoe, I. Applied Regression Modeling. A business approach. John Wiley e Sons, 2006. ISBN 9781119615903

[32] M. Riedmiller. "Advanced supervised learning in multi-layer perceptrons — From backpropagation to adaptive learning algorithms". *Computer Standards e Interfaces, 16(3),* 265-278,1 1994.

[32] W. Goffman. "A mathematical method for analyzing the growth of a scientific discipline". Journal of the Association for Computing Machinery, 18(2), 173-185, 1971. DOI: 10.1145/321637.321640

[33] P. Swicegood e J.A. Clark. "Off-site monitoring systems for prediction bank underperformance: A comparison of neural networks, discriminant analysis, and professional human judgment". *Intelligent Systems in Accounting Finance e Management, 10(3),* 169-186, 2001. DOI:10.1002/isaf.201

[34] B.G. Malkiel, B.G. A Random Walk Down Wall Street. W. W. Norton e Company, 2013.

[35] C.C. Ying. "Stock Market Prices and Volumes of Sales". *Econometrica, 34*, 676 – 686, 1966. DOI: 10.2307/1909776

[36] R.L. Crouch. "The Volume of Transactions and Price Changes on the New York Stock Exchange". Financial Analysts Journal, 26, 104-109, 1970.