REVISTA DE GESTÃO E SECRETARIADO
MANAGEMENT AND ADMINISTRATIVE PROFESSIONAL REVIEW
ISSN: 2178-9010

Revista GeSec
São Paulo, SP, Brasil v. 14, n. 7, p. 11294-11308, 2023

DOI: http://doi.org/10.7769/gesec.v14i7.1670

Sales forecasting using machine learning algorithms

Previsão de vendas utilizando algoritmos de aprendizagem automática

Emerson Martins[1]
Napoleão Verardi Galegale[2]

Abstract

Retail companies, as production systems, must use their resources efficiently and make strategic decisions to obtain growing and stable revenues, especially when market conditions are becoming more competitive and profit margins are increasingly pressured. Thus, sales forecasting is crucial to maintain competitiveness in the retail segment, but obtaining inaccurate forecasts can lead to stock shortages, causing delays in deliveries and generating customer dissatisfaction, as well as increasing inventory, increasing the cost of warehousing, forcing the "burn" of stock through promotional campaigns, directly affecting profitability. Forecasting the demand for products and services and adapting the supply chain by finding a balance has always been and will continue to be a challenge in the retail segment. This research aims to evaluate the main methods and identify the one with the greatest accuracy in sales prediction. Based on an integrative literature review (ILR), three main methods were evaluated: time series, artificial neural networks and machine learning algorithms. The results show that machine learning is more suitable in terms of accuracy, particularly when models contain exogenous and endogenous variables, in addition to allowing the identification of hidden patterns in demand that can be used to identify market trends. However, in markets

[1] Master's Degree in Management and Technology in Productive Systems from the Centro Estadual de Educação Tecnológica Paula Souza (CEETPS), Rua dos Bandeirantes, 169, Bom Retiro, São Paulo - SP, CEP: 01124-010. E-mail: emerson.martins@cpspos.sp.gov.br Orcid: https://orcid.org/0000-0001-5296-8900

[2] PhD in Controllership and Accounting from Faculdade de Economia, Administração, Contabilidade e Atuária da Universidade de São Paulo (FEA - USP), Rua dos Bandeirantes, 169, Bom Retiro, São Paulo - SP, CEP: 01124-010. E-mail: napoleao.galegale@cpspos.sp.gov.br Orcid: https://orcid.org/0000-0003-2228-9151

with constant demands and few external interferences, its use is not justified because, for these cases, the use of time series is simpler and less costly.

Keywords: Sales Forecast. Retail. Machine Learning. Time Series. Productive Systems.

Resumo

As empresas de varejo, como sistemas produtivos, devem utilizar seus recursos de forma eficiente e tomar decisões estratégicas para obter receitas crescentes e estáveis, principalmente quando as condições de mercado estão se tornando mais competitivas e as margens de lucro cada vez mais pressionadas. Assim, a previsão de vendas é crucial para manter a competitividade no segmento de varejo, mas a obtenção de previsões imprecisas pode levar à falta de estoque, causando atrasos nas entregas e gerando insatisfação do cliente, além de aumentar o estoque, elevando o custo de armazenagem, forçando a "queima" de estoque por meio de campanhas promocionais, afetando diretamente a lucratividade. Prever a demanda de produtos e serviços e adequar a cadeia de suprimentos buscando o equilíbrio sempre foi e continuará sendo um desafio no varejo. Esta pesquisa tem como objetivo avaliar os principais métodos e identificar aquele com maior precisão na previsão de vendas. Com base em uma revisão integrativa da literatura (ILR), três métodos principais foram avaliados: séries temporais, redes neurais artificiais e algoritmos de aprendizado de máquina. Os resultados mostram que o aprendizado de máquina é mais adequado em termos de precisão, principalmente quando os modelos contêm variáveis exógenas e endógenas, além de permitir a identificação de padrões ocultos na demanda que podem ser usados para identificar tendências de mercado. Entretanto, em mercados com demandas constantes e poucas interferências externas, seu uso não se justifica, pois, para esses casos, o uso de séries temporais é mais simples e menos oneroso.

Palavras-chave: Previsão de Vendas. Varejo. Aprendizado de Máquina. Séries Temporais. Sistemas Produtivos.

Introduction

In the last two decades, computing power has evolved considerably, in this context we can mention: large data storage, more robust processors, faster internet connection, among other examples. Problems that seemed to be extremely complex or costly to solve are now within our reach. New trends such as Big Data, Cybersecurity, Internet of Things (IoT) and blockchain have emerged, jointly exploiting the technological advances mentioned above. The IoT, which aims to use embedded systems, including sensors and actuators, together with the internet, to allow control and immediate access to information in real time (Atzori, 2010; Cecchinel, 2014), represents one of these challenges, as the report of "Juniper Research", informs that by 2024 we will have more than 83 billion devices and sensors connected (IoT). In addition, some of these devices will have the ability to generate significant amounts of data on the order of Zettabytes, information that can be valuable for a company's strategy, so sales forecasting cannot ignore these new trends; it must use it to support competitive advantage.

Currently, data processing through information systems to generate knowledge has become vital for decision makers, particularly in some important areas such as forecasting sales of retail products or services, in which external variables such as weather or global economy can affect people's consumption decisions (KRAWCZYK, 2016).

One of the recent and popular techniques aimed at tackling these new business challenges is Big Data Analytics (BDA). A precise definition of BDA is given in Hofmann (2018). In short, it is the alignment of Big Data and Machine Learning (ML) techniques to provide reliable insights for decision making. ML and Big Data benefit from each other as they can be coupled to create more complete models. Furthermore, the main purpose of the BDA is to transform information into useful knowledge.

Predictive analytics, in turn, encompasses methods that use information to create models and perform simulations that will provide insights into future events, allowing the most attentive executives to be able to predict strategic actions that improve their company's performance. By definition, the results obtained through these techniques are not 100% accurate, as no method can predict the future, so a good predictive analysis is one that provides the most accurate results in a reasonable time (CASTILHO *et al*, 2017).

One of the common uses of predictive analytics in business is sales forecasting – this point will be covered in section 3 of this paper – but there are also several other applications in domains such as: cost estimation, where (LOYER *et al*, 2016) applied techniques of ML to quickly estimate the cost of manufacturing aircraft engine components. In the performance evaluation, (FAN *et al*, 2013) used ML to estimate the performance of the supply chain based on the "5 Dimensional Balanced Scorecard" (5DBSC), in order to provide quick results and avoid biased performance evaluations by the managers.

The machine learning (ML) or machine learning approach can be described as the study of computer algorithms that automatically improve with experience. It is treated with an artificial intelligence (AI) subarea. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or impractical to develop conventional algorithms to perform the necessary tasks (RASCHKA *et al*, 2017).

Humans learn through experience, we use a process of trial and error to figure out what actions should be taken under certain circumstances. This allows us to make abstractions and build knowledge. ML is somewhat similar, it can be seen as algorithms that aim to improve a performance measure, automatically deriving their own rules and creating their own decision models based on certain information (RASCHKA *et al*, 2017) and Mitchell (1997).

In this context, the purpose of this paper is to provide an insight into the latest ML techniques applied to retail sales forecasting through an ILR. The question that guided the research can be posed as follows: Is ML more suitable than traditional methods of forecasting retail sales, in terms of accuracy, particularly when the models contain exogenous and endogenous variables?

Research Methodology

For this descriptive and qualitative research, an ILR was carried out on the use of ML to forecast sales in the retail segment. The PRISMA-P protocol was used, this protocol aims to support researchers to improve the reporting of systematic reviews and meta-analyses, filtering the number of publications with greater relevance to the research topic (MOHER *et al*, 2015)

In the identification stage, a search was performed for publications in the Google Scholar, Microsoft Academic and Scopus databases, with the following search string: ((“Sales Forecasting” OR “Sales Predictions” OR “Predictive”) AND (“Machine Learning” OR “Algorithm” OR “Big Data Analytics”) AND Retail”), in the period from 2015 to 08/2021. 1,022 publications were returned.

In the screening stage, 87 Books were removed, 159 publications classified as HTML and PDF, 13 Conference Papers and 7 publications classified as “Other”, totaling 266 records excluded, resulting in 756 records selected for the next stage. The inclusion and exclusion criteria of the studies are presented in Table 1.

Inclusion criteria
It helps to define what sales prediction is and the impact of exogenous and endogenous variables
Helps categorize types of ML algorithms

Presents accuracy metrics for evaluating ML algorithms
Exclusion Criteria
Duplicate paper
Other literatures, that is, they are not scientific papers (theses, books, dissertations, interviews, etc.)
Search outside the scope of interest

Table 1 - Criteria for inclusion and exclusion of studies.

In the eligibility stage, 193 papers without an H-Index were excluded, 43 papers whose access is not public, 488 papers whose title and keywords are not related to the objective of this research. After exploratory analysis of the remaining 32 papers, 23 papers with low adherence to the research topic were discarded, as the content of the papers did not present metrics used to assess the performance of the algorithms presented. Thus, 9 publications were selected for qualitative analysis as shown in Table 2.

	Title	**Author**	**Year**
01	Machine learning methods for demand estimation	P Bajari, D Nekipelov, SP Ryan, M Yang	2015
02	A machine learning framework for customer purchase prediction in the non-contractual setting	A Martínez, C Schmuck, S Pereverzyev Jr	2020
03	Retail forecasting: Research and practice	R Fildes, S Ma, S Kolassa	2019
04	A deep learning approach for the prediction of retail store sales	Y Kaneko, K Yada	2017
05	Sales forecasting by combining clustering and machine-learning techniques for computer retailing	IF Chen, CJ Lu	2017
06	Intelligent Sales Prediction Using Machine Learning Techniques	S Cheriyan, S Ibrahim, S Mohanan, S Treesa	2018
07	Applying computational intelligence methods for predicting the sales of newly published books in a real editorial business management environment	Castillo, P.A., et al	2017
08	Sales-forecasting of retail stores using machine learning techniques	A Krishna, V Akhilesh, A Aich	2018
09	Comparison of different machine learning algorithms for multiple regression on black friday sales data	CSM Wu, P Patil, S Gunaseelan	2018

Table 2 - Papers selected for this research.

Integrative Literature Review (ILR)

Traditional forecasting methods are based on time series, this means that they are applied under the assumption that past demand can statistically estimate future demand. Typically, these methods are easy to apply and perform well in markets where demand is mostly stable (CHOPRA *et al*, 2013). Unfortunately, we have many cases where this scenario cannot be applied, as demand often depends on exogenous factors that are not effectively represented by past values. For example, on-demand transport services such as UBER cannot estimate their demand just relying on time series, they must take into account other elements such as weather conditions, time of day, day of the week (KE *et al*, 2017).

To satisfy this need, other types of forecasting, known as causal modeling, propose methods that include exogenous elements, such as macroeconomic variables, weather conditions, marketing strategies, etc. (CHOPRA *et al*, 2013). These techniques allow us to face the limits found in time series models. In this sense, ML itself could be considered as a causal modeling provider because it can deal with time series, categorical variables, fuzzy variables, text analysis, images and other elements.

ML has been applied to sales forecasting since 1980 through methods such as Artificial Neural Networks (ANN) (HIPPERT *et al*, 2001). Over the past two decades, these methods have yielded interesting results and demonstrated some potential, but several ANN search applications were not valid due to validation or implementation issues (ADYA, *et al*, 1998). This problem was probably caused by the scarcity of data to effectively train ANNs, as this technique has a good generalizability, but it needs a lot of data, well distributed data and time for training. Previous limitations, such as limited storage capacity, low computing power and slow internet connections, could have influenced the reluctance to use ML in forecasting sales.

Nowadays, ML enjoys a good reputation, probably because most of the aforementioned restrictions have been overcome. Thanks to this, new applications were published leading to new trends and techniques. To identify these new trends and meet the objective of this research, a literature study was carried out, according to the steps described in section 2.

As a result, 9 publications presented in Table 2 were selected. For each study, the objective of the application, the type of resource found in the dataset, the method used for data pre-processing and the ML techniques used for sales prediction are presented. Some nomenclatures are used in this paper: (T) - Time series variables (past sales); (Ve) - Endogenous model variables (price, POS number, period of sale); (Vx) - Exogenous model variables (Weather, time of day, day of the week (weekend), spatial location, seasonality, unemployment rate, inflation rate, population size, average income).

In the paper [01], 1,510,563 sales transactions were analyzed in a supermarket chain with 3,149 stock keeping units or SKUs (Stock Keeping Units), the data was collected through IRI Marketing Research through an academic license from the university. of Chicago. 25% of the data was used as a validation sample, 15% for validation and 60% for training. The RMSE root mean square prediction error metric was used to assess the accuracy of the methods used.

As a result, the RF and SVM algorithms presented an accuracy of 65% and 15% respectively, while the traditional Linear Regression method presented 6% of precision.

[02] highlights that one of the main challenges of retail is to differentiate customers who made a punctual purchase and do not intend to make new purchases, from customers who are more likely to make regular purchases but are "paused" between a purchase and other. It is accepted by business wisdom and research literature that it costs five to ten times more to acquire a new customer than it does to retain an existing customer (Daly, 2002; Bhattacharya, 1998). In this study, a data set with more than 10,000 customers and 200,000 purchase transactions was used, and the GTB method achieved the best performance, achieving 89% accuracy and 0.95 (AUC - Area Under The Receiver Operating Characteristic Curve) in predicting monthly purchases in the test dataset.

In [03], the authors perform a benchmark through a literature review comparing the results of other researchers in various scenarios in the retail segment, confirming that time series models have been widely used to forecast aggregate retail sales, where Simple Exponential Smoothing and its extensions, together with ARIMA models, have been the most used time series models for sales forecasts, but due to dependence on limited data or use of inadequate evaluation metrics, such as sample adjustment, some researchers have found that standard time series models are sometimes inadequate to assess aggregate retail sales, identifying evidence of nonlinearity and volatility in the retail sales time series, for example (ALON *et al*, 2001; CHU & ZHANG , 2003; KUVULMAZ, USANMAZ, & ENGIN, 2005; ZHANG & QI, 2005), they resorted to nonlinear models, especially artificial neural networks. The results indicate that traditional time series models with a stochastic trend, such as Simple Exponential Smoothing and ARIMA, performed well when macroeconomic conditions are relatively stable, however, when economic conditions are volatile (with rapid changes in economic conditions), Artificial neural network ANNs have been claimed to outperform linear methods (ALON *et al*, 2001). As a result, Clustering techniques and DT algorithms showed the best results for forecasting sales for new items with limited historical data. In another scenario, 2 mass brands were evaluated in two different retail stores, with a history of three years of daily sales in promotional products, the SVM and ANNs methods showed better results.

In the study [04], three-year data from a point-of-sale (POS) collected between 2002 and 2004 from a supermarket located in the Kanto region of Japan were used, in order to predict the sales volume of the next day, applying Deep Learning (DP) methods in comparison with the LGR logistic regression model. Sales data were grouped into three categories according to product attributes: Category 1 (62 attributes), Category 2 (569 attributes) and Category 3 (3,312 attributes). The three years of data were divided so that 80% was used for learning and 20% for verification. As a result, the DL was superior to the LGR obtaining an accuracy of 86%.

In [05], actual sales data from 124 points of sale between Jan-2005 and Sept-2009 of three large retailers with domain in computer products in the Taiwan region were grouped. Three products were considered for the sales history: Computers (PC), Notebooks (NB) and Liquid Crystal Displays (LCD). In the study, three clustering data clustering techniques were compared: SOM, GHSOM and K-Means, together with two ML algorithms: SVM and ELM. The data sets PCs, NBs and LCDs were subjected to six cluster-based prediction models and also in isolation using two machine learning techniques (single SVR) and (single ELM), that is, without using a clustering algorithm . The first 88 points of sale (71% of the sample) are used as a training sample, while the remaining 36 points of sale (29% of the sample) are used to forecast sales. Two evaluation criteria were used to measure sales performance: MAPE and RMSPE.

The experimental result demonstrated that of the 8 models created the combination of the GHSOM clustering with the ELM machine learning algorithm provided superior performance for all 3 selected products.

In [06], it is highlighted that the precision in the sales forecast has a great impact on the business. Data mining techniques are very effective tools in extracting hidden knowledge from a huge dataset to increase prediction accuracy and efficiency. Organizations face serious challenges to identify a data mining technique and an effective presale strategy (MATHEW et al, 2015), due to the exponential growth in the volume of data used in e-commerce transactions. Traditional forecasting methods are difficult to handle a large amount of Big Data data, in this context the Data Mining data mining technique becomes a strong ally in sales prediction. At the organizational level, sales forecasts are essential inputs to support decision making in various business areas, such as operations, marketing, sales, production, logistics, inventory, finance (cash flow). With almost 85,000 records, the initial database considered in this research was reduced after pre-processing due to redundant records as well as irrelevant information for analysis. The dataset used for this paper is based on a fashion store with three consecutive years (2015 to 2017) of sales data. In this paper, three ML algorithms were compared: GTB, DT and GLM, which showed 98%, 71% and 64% accuracy respectively.

[07] mentions the problem of predicting sales for newly launched products on the market. For this study, sales data consisting of 6000 books from a Spanish publisher Trevenque Editorial S.L were collected, whose data were analyzed using the SOM classifier and two pre-processing techniques: Correlation-based feature selection and Relief for attribute estimation. This is a challenge for the industry, as printing a much larger number of volumes than those eventually sold will lead to losses, while printing an adequate number of copies will optimize the publisher's sales and profits. In addition, there are several difficulties inherent in forecasting new book sales, such as the limited amount of historical data or market variability (fads, seasonality) that make the task of predictive methods difficult. As a result, the DT and RF algorithms showed the highest precision and the best performance.

In the paper [08], the precision of different algorithms in the dataset extracted from the website https://www.kaggle.com/ was compared, where AdaBoost and GTB presented RMSE of 1,350 and 1,088 respectively, thus GTB presented lower error rate, therefore better precision in the evaluated dataset.

In the paper [09], a dataset with 550,000 sales transactions is collected from a retail company to train supervised ML algorithm to predict the amount of customer purchases, enabling the creation of personalized offers according to their consumption profile. customers during the Black Friday period. In the research, it is highlighted that in ML algorithms, the data set used must be balanced, that is, all classes must contain the same number of samples, otherwise, the prediction or classification will be biased towards that category of data in which the data is skewed, meaning a good ML algorithm is useless without the proper data. The accuracy of the prediction model depends on the reliability of the data it is built on, however, real world data is often confusing and needs cleaning, to aid in this task there are pre-processing techniques. As a conclusion it is highlighted that complex models such as neural networks are overkill to deal with simple problems such as regression, in this way we can use simpler models together with pre-processing to obtain better results. In the applied ML methods, XGBoost which internally uses SR and RR presented the best precision with 2400 RMSE. It is evident that ML can be applied to forecast sales in a wide range of different types of products.

[01,03,07] highlight that methods such as RF and DT provide an unparalleled level of interpretation, together with good accuracy and decreasing computation time.

Most studies include endogenous inputs and exogenous [03,07] in their models, which shows the good flexibility of ML to handle a wide range of inputs. Furthermore, [07] addressed the problem of implementing sales forecasting in new products for which historical sales data is not available. For this, historical data from other products were used, along with endogenous variables, such as the number of weeks on sale and the retail price.

In [06,07,08,09] data pre-processing techniques were effectively applied, which resulted in simpler models, lower computational cost and good precision. As we are in an era with a high volume of data being generated daily, this causes noisy and meaningless variables, so analysing the most relevant variables in real world applications is mandatory, this way we can differentiate whether a model will be useful or not.

3.1 Comparing Machine Learning and Traditional Forecasting Methods

Traditional models offer enormous advantages in terms of simplicity and accuracy, as they can forecast sales in a matter of seconds for multiple SKUs (YU *et al*, 2011). However, they need to be designed by an expert who can tailor it to the company's needs. Also, do not include exogenous variables. ML models partially solve this problem because they can include other types of data, such as endogenous and exogenous variables, allowing a better representation of reality. Furthermore, ML techniques that are correctly implemented outperform most traditional forecasting methods (YU *et al*, 2011). In order to assess whether this statement is valid in our study of the literature, the Table 3 mentions which were the traditional and ML models applied by the authors and which had better accuracy.

Paper	**Application**	**Data Set**	**Pre-processing technique**	**Statistical and/or classification methods**	**Evaluation metrics**	**ML**	**Model with better accuracy**
01	Grocery, 6-year database – 3,149 SKU and 1,510,563 sales transactions	T, Ve	-	Linear Regression	RMSE	SR, FSR, LASSO, SVM, Bag, RF	RF, SVM
02	10,000 customers and 200,000 purchase transactions	T, Ve	-	-	Precision, AUC	GTB, LASSO, ELM	GTB
03	482 textile items with 52 weeks of sales history.	T, Ve, Vx	-	Simple Exponential Smoothing, ARIMA	RMSE, MAPE, MdAPE	DT	DT, SVM, ANN
	2 pasta brands from 2 different retail stores, with 3-year historical data with daily sales			ARIMA, ETS e HW		SVM, ANN	

04	Three-year data from a point-of-sale (POS) collected between 2002 and 2004 from a supermarket located in the Kanto region of Japan	T, Ve	-	Logistic Regression	Accuracy, Precision, Recall, F-measure, AUC	-	DL
05	Actual sales data from 124 points of sale from Jan-2005 to Sept-2009 of three large IT-domain retailers in Taiwan region	T, Ve	-	SOM, GHSOM, K-Means	MAPE, RMSPE	SVR, ELM	GHSOM + ELM
06	Actual sales data collected from a fashion store for three consecutive years (2015 to 2017)	T, Ve	outlier detection	-	Accuracy, Error Rate, Precision, Recall, Kappa	GLM, DT, GTB	GTB
07	Sales forecast for new books	T, Ve, Vx	Correlation-based feature selection, Relief for atribute estimation	Multiple Linear Regression	MAE, RMSE, RAE, RRSE	ELM, KNN, DT, ANN, RF, SVM	DT, RF
08	Kaggle dataset, with 8,523 entries	T, Ve	Outlier detection, including missing values	Multiple Regression, Polynomial Regression, Ridge Regression	RMSE	AdaBoost, LASSO, GTB	GTB
09	Dataset with 550,000 transactions in the Back Friday period	T, Ve	Transforming categorical data into numerical data	Logistic, Polynomial, Stepwise, Ridge, Lasso and Elastic Regression	RMSE	LR, MLK, DL, DT, Bagging, XGBoost	XGBoost

Table 3 - Comparison of ML with traditional sales forecasting methods.

The fact that ML outperforms traditional models in the sales forecasting landscape does not necessarily mean that companies should change their forecasting tools. This leads us to consider an important question: when should companies invest in ML for sales forecasting?

Sales forecasting must be flexible and responsive, especially when it comes to short-term forecasting, which is needed in industries like retail. Forecast calculation should be agile as most companies have hundreds or even thousands of SKUs, all of which may need an immediate estimate.

Even though data pre-processing techniques can substantially reduce computing time, they are still complex and therefore costly to establish in a company, as they require skilled people with adequate equipment. This difficulty is especially present in small companies, where available resources are limited and employees rarely have advanced knowledge on the subject. In this case, if a company is positioned in a stable market and if historical demand is sufficient to achieve good sales forecast accuracy with traditional methods, this company should postpone the migration to ML techniques until it really identifies an added value in use them.

On the other hand, if a company sells products in a market subject to constant evolution, where it is mandatory to be at the forefront of trends for being competitive, ML models for sales forecasting will be a valuable asset. However, it is not an easy change, as companies must guarantee three fundamental aspects: data storage capacity, data processing capacity and employee qualification. Data storage is essential, as ML needs to deal with Big Data for good performance, and this requires large storage capacity, as well as data processing capacity can vary from simple models to more robust models (ZHOU *et al*, 2017); therefore, the company must consider its objectives versus the resources it needs.

In short, a company should choose to use ML over traditional sales forecasting methods when its economic environment really requires a digital transformation, and when the company is able to muster the necessary resources to take on the challenge of projecting future sales.

Conclusion

Dealing with a high volume of data is one of the most common challenges posed by modern business trends. As data has become one of the most valuable resources, sales managers are eager to extract relevant information that will lead to a competitive advantage. This paper conducted a literature study to investigate the application of ML in sales forecasting as a method to achieve this advantage. Nine recent research papers that apply ML to sales forecasting were selected and analyzed to identify new trends. One of the findings was that ML extends the reach of sales prediction, as it is capable of dealing with complex variables. More precisely, ANN approaches have shown excellent performance when dealing with imprecise data such as exogenous variables, while DT and RF offer incredible interpretability. Furthermore, data pre-processing techniques have been proven to substantially reduce the complexity of models, allowing both good accuracy and reasonable computation time.

As the process of adapting a company to new technologies often comes with doubts raised by uncertainty, a comparison between ML and traditional forecasting methods was made with the aim of providing insights to managers who are willing to implement ML in their processes. The results of this study show that ML is more suitable than traditional forecasting methods in terms of accuracy, particularly when models contain both exogenous and endogenous variables. In addition, it allows the identification of hidden patterns in demand that can be used as a baseline to identify new market trends.

In the findings, it is evident that the absence of a reliable training base, that is, one that accurately reflects real-world transactions, affects the algorithm's efficiency, containing biases that make its usefulness unfeasible. Furthermore, it is found that in markets with constant demands and few external interferences, the use of ML is not justified, because for these cases the use of traditional methods through time series is simpler and less costly.

For further research, a similar study should be carried out on data pre-processing techniques, as they offer significant advantages in terms of computational cost reduction in the application of ML models. In addition, the processing cost and time for each method must be evaluated. Finally, addressing these new trends for small businesses is vital as they are numerous but often lack the financial means or experience to implement disruptive technologies.

Acknowledgement

I first thank God, who allowed all the achievements in my life to materialize. Thanks to CPS, which gave me the knowledge I needed to complete this work. Thanks to my advisor for all the help and guidance. I thank my mother, wife and son for their understanding of my absence at various times.

References

Adya, M. and Collopy, F., 1998 - *How effective are neural networks at forecasting and prediction? A review and evaluation*. - Journal of Forecasting

Alon, I. et al, 2001 - *Forecasting aggregate retail sales: A comparison of artificial neural networks and traditional methods* - Journal of Retailing and Consumer Services

Atzori, L. et al, 2010 - *The internet of things: A survey* - Computer Networks.

Bhattacharya, C. B., 1998 - *When customers are members: Customer retention in paid membership contexts* - Journal of the academy of marketing science.

Cecchinel, C. et al, 2014 - *An architecture to support the collection of big data in the internet of things*. - IEEE World Congress on Services

Chopra, S. and Meindl, P., 2013 - *Supply Chain Management* - Pearson Education, Inc.

Chu, C. W. and Zhang, G. P., 2003 - *A comparative study of linear and nonlinear models for aggregate retail sales forecasting* - International Journal of Production Economics

Daly, J. L., 2002 - Pricing for profitability: activity-based pricing for competitive advantage – Vol. 11 - John Wiley & Sons.

Fan, X., et al, 2013 - An evaluation model of supply chain performances using 5DBSC and LMBP neural network algorithm - Journal of Bionic Engineering

Hippert, H. S., et al, 2001 - Neural networks for short-term load forecasting: a review and evaluation - IEEE Transactions on Power Systems

Hofmann, E. and Rutschmann, E., 2018 - Big data analytics and demand forecasting in supply chains: a conceptual analysis - International Journal of Logistics Management

Juniper Research, 2020. Available in: https://www.juniperresearch.com/press/iot-connections-to-reach-83-bn-by-2024?ch=IOT%20CONNECTIONS%20TO%20GROW

Ke, J., et al, 2017 - Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach - Transportation Research

Krawczyk, B. 2016 - Learning from imbalanced data open challenges and future directions - Progress in Artificial Intelligence – Springer

Kuvulmaz, J., et al, 2005 - Time-series forecasting by means of linear and nonlinear models - In A. Gelbukh, A. DeAlbornoz, & H. TerashimaMarin (Eds.), MICAI 2005: Advances in artificial intelligence

Loyer, J. L., et al, 2016 - Comparison of machine learning methods applied to the estimation of manufacturing cost of jet engine components. - International Journal of Production Economics

Mathew, N. M. and Jomo, K., 2015 - A Survey on the Clustering Algorithms in Sales Data Mining - International Journal of Computer Applications Technology and Research

Mayo, M., 2016 - The data science puzzle, explained. Available in: https://www.kdnuggets.com/2016/03/data-science-puzzle-explained.html/2

Mitchell, T. M., 1997 - Machine Learning - McGraw Hill

Moher, D., et al, 2015 - Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P)

Raschka, S. and Mirjalili, V., 2017 - Python Machine Learning, 2nd Ed.- Packt Publishing, Birmingham

Vahdani, B., et al, 2016 - A high performing meta-heuristic for training support vector regression in performance forecasting of supply chain - Neural Computing and Applications

Yu, Y., et al, 2011 - An intelligent fast sales forecasting model for fashion products - Expert Systems with Applications

Zhang, G. P. and Qi, M., 2005 - Neural network forecasting for seasonal and trend time series - European Journal of Operational Research

Zhou, L., et al, 2017 - Machine *learning on big data: Opportunities and challenges* – Neurocomputing

Submetido em: 19.06.2023

Aceito em: 18.07.2023