

Machine learning algorithms used to predict product sales in the retail sector

Angel Alfonso Gonzales Molina
 Faculty of Systems Engineering and Informatics
 Universidad Nacional Mayor de San Marcos
 Lima, Perú
 10200156@unmsm.edu.pe

Igor Aguilar-Alonso
 Professional School of Systems Engineering
 Universidad Nacional Tecnológica de Lima Sur
 iaguilar@untels.edu.pe
 Faculty of Systems Engineering and Information
 Technology
 Universidad Nacional Mayor de San Marcos
 Lima, Perú
 iaguilara@unmsm.edu.pe

Abstract—Nowadays, due to the COVID-19 pandemic, the Retail sector has been impacted by the sale of technological equipment, which is used to carry out remote work and communicate with other people in different parts of the world. The objective of the research is to find out which machine learning algorithms are applied for the prediction of product sales, where the aim is to solve the problem faced by companies in the retail sector, which do not have prediction tools and the large companies take advantage of this. The methodology used for the research was the literature review. The result of the study addresses 11 machine learning algorithms that are used to predict product sales. Finally, it is concluded that the XGBoost algorithm is the most efficient, precise and tends to lower error for prediction cases, finally a model proposal is made for these specific cases of sales prediction.

Keywords—*machine learning, prediction, algorithms, retail*

I. INTRODUCTION

Due to the pandemic, the retail sector has been strongly impacted by the sale of technological equipment, as well as cell phones, tablets, laptops, desktop PCs and other accessories. Today digital communication has become widespread and therefore the need to resort to the use of these technological equipment is vital to carry out our daily activities. This research is dedicated to all those companies that are dedicated to the purchase and sale of technological equipment where sales prediction can be evidenced using machine learning applied to business intelligence.

The problem that these types of companies face in the retail sector is that they have a high-demand market, but they do not have the ability to know what, when to buy at a good price and then sell it and thus obtain a better profit margin, improve response time to customers and win more sales opportunities. Another event is that at the global level, small and medium-sized enterprises (SMEs) have played a very important role in the development of countries, despite technological changes, and, above all, the disadvantages they face with large companies. Business. Their prominence is due to the

importance they have in the economy of the countries and the high degree of job creation.

The help of information technologies is of vital importance for all businesses and having strategies aligned with business objectives is key to continuous improvement.[1] The set of models and architectures in the information technology domain describe their importance in business, as well as their strong impact on all business areas.[2]

The life cycle of technologies demands a key administration of all companies and thus satisfy the needs of the client to cover the objectives set.[3]

This article aims to make the literature available to inform what machine learning algorithms exist to predict product sales.

The motivation to carry out this research is that small and medium-sized companies can have an additional plus to compete with large companies.

The structure of this article consists of six sections: Starting with the introduction, going through describing the main concepts, and then detailing the review of the literature. In section four, the results obtained after the literature review will be known. Section five describes the results obtained and an analysis. Finally, we will mention the conclusions and recommendations in section six.

II. THEORETICAL FRAMEWORK

A. Machine Learning

The concept of machine learning, despite its growing popularity, has been misdefined. Many authors of the existing literature define it as a process through which a system interacts with its environment in such a way that the structure of the system changes and due to structural alterations, the interaction also changes the process.[4]

Goodfellow et al. They define machine learning as the ability of artificial intelligence (AI) systems to acquire knowledge by obtaining patterns from raw data sets.[5]

Lantz indicates that machine learning is related to different techniques to process and transform data into actionable intelligence.[6]

Mohammed et al. They describe machine learning as enabling machines to learn without programming. A key advantage of various machine learning techniques such as ANNs.[7]

Here is a list of the algorithms/techniques that are used in machine learning:

- 1) Decision trees
- 2) Naive Bayes Classification
- 3) Ordinary Least Squares Regression
- 4) Logistic Regression
- 5) Support Vector Machines
- 6) Ensemble Methods
- 7) Clustering Algorithms
- 8) Principal component analysis
- 9) Singular Value Decomposition
- 10) Independent Component Analysis

B. Sales Prediction

Prediction deals with events that occur in the future. The use of machine learning algorithms improves the intelligence of the system without manual intervention. Machine learning (ML) is used to optimize performance criteria using sample data or the past based on experience.[8]

III. METHODOLOGY

For the development of the present investigation, the guide established by Kitchenham and Charles was used, which consists of three important parts [9].

- **Revision planning.** In this phase it is necessary identify the need to carry out the review, considering the search protocol and review protocol..
- **Carrying out the review.** In this phase, the information from the main studies is methodically selected according to the inclusion and exclusion criteria.
- **Results of the review.** This phase presents the statistical results and analysis of the studies chosen for the review.

A. Review Planning

To develop this research, a literature review was carried out in several scientific articles with a time not exceeding 5 years, extracted from important bibliography databases such as Science Direct, Google Scholar, IEEE Explore, Scopus and Web of Science.

Based on the articles of the different authors of the research on sales prediction algorithms in the area of machine learning, the following question was asked:

¿What machine learning algorithms are used to predict product sales in the retail sector?

The search string applied in the bibliographic databases was the one detailed below:

MACHINE LEARNING ALGORITHMS TO PREDICT PRODUCT SALES

The inclusion and exclusion criteria considered are detailed in Table 1.

TABLE 1. INCLUSION AND EXCLUSION CRITERIA

Inclusion criteria	Exclusion criteria
Articles published between 2016 and 2021	Research that does not meet the inclusion criteria
Articles related to the research question	
Conference articles and journals	

B. Conducting the review

Based on the articles identified in the search process using the established string, the articles were filtered according to the inclusion and exclusion criteria in Table 1.

C. Results of the review

The result of the review search process was possible after developing the search chain where there was a total of 50 potential studies, 32 relevant articles and 21 selected articles based on machine learning algorithms applied for sales prediction.

Table 2 shows the number of articles selected for each source consulted.

TABLE 2. POTENTIAL, RELEVANT AND SELECT ARTICLES

Database consulted	Articles			%
	Potentials	Relevant	Selected	
Science Direct	13	8	4	19%
Google Scholar	15	10	6	29%
IEEE Explore	8	5	4	19%
Scopus	10	7	5	24%
Web Of Science	4	2	2	9%
TOTAL	50	32	21	100%

IV. DESCRIPTION OF THE ANNOTATED RESULTS

The results found during the literature review process is the product of the search and analysis of the 21 selected articles, 11 machine learning algorithms used for the prediction of product sales were found.

According to F. Haselbeck et al. [10] developed a comparison between different machine learning algorithms and classical prediction algorithms to predict the demand for vegetables, which are products with a short shelf life. In this research they show that the machine learning algorithms were superior to the others, specifically with the XGBoost gradient algorithm being the best in 14 of 15 comparisons and as the dataset grew the advantage over the other algorithms.

In the publication by D. Rohaan et al. [11], present a method to use sales prediction in B2B business types. They apply supervised machine learning algorithms and natural language techniques to learn. In the comparison of the Gradient Boosting Classifier, Random Forest and Logistic Regression algorithms, where the measurement base is the F1 - Score precision, it was found that the random forest algorithm had the highest precision.

According to the research of D. Raditya et al. [12] compare several algorithms to decide which one has the best performance in predicting shoe prices. The algorithms used are linear regression and random forest using a dataset of historical sales data. For both algorithms, a result with similar values is obtained, but carrying out an additional evaluation, it is concluded that the random forest algorithm has a better performance compared to linear regression for the prediction of the price of sneakers.

According to the research of M. Leippold et al. [13], use several machine learning algorithms for the construction and analysis of a comprehensive set of performance prediction factors in asset prices in a Chinese product store. They find that the effectiveness of the precision under the measure of R2 there is a variance with respect to the annual, monthly and daily dataset. Here models are implemented with the Ordinary Least Squares, Partial Least Squares, Lasso, Gradient Boosting Decision Tree, Random Forest and KNN algorithms, being a partial result for different critical factors.

According to the research of BM Pavlyshenko [14], he mentions the use of machine learning models to predict sales and aims at the uses of machine learning to predict sales. For a relatively small amount of historical data it uses the time series algorithm and with an approach of using several regression sets (Neural Network, Lasso, RandomForest, ARIMA, ExtraTrees) it gives a better result for predictive models. It is concluded that the use of regression in various approaches for sales prediction improves the results compared to time series methods.

According to the research of S. Cheriyan et al. [15], carry out the detailed study and analysis of various predictive models to improve sales predictions. They use several data mining techniques and have the reliability and accuracy as a measure of the prediction models, applying different algorithms in the predictive models, they find that the model

that shows the highest accuracy in sales prediction is the Gradient Boost algorithm. .

According to the article by I-Fei Chen et al. [16], proposes a clustering-based prediction model that combines clustering and machine learning methods. The clustering method was used to partition the training data, following similar patterns for each cluster based on a particular feature. Among these techniques we have K-means, GHSOM and SOM. And the machine learning algorithms it uses are Support Vector Regression or (SVR) and Extreme Learning Machine (ELM). The results showed that the combination between GHSOM and ELM showed superior prediction performance compared to the other models.

In the publication by Mohit Gurnani et al. [17], evaluate and compare several machine learning models based on ARIMA algorithms, autoregressive neural network, XGBoost, SVM and hybrid models that are combinations of the previously presented. These models are applied to predict the sales of a pharmacy company and the data set used contains historical sales and complementary information. MAE and RMSE were used to measure the models. As a result, it was found that they give better results when we combine techniques than when we use them individually.

According to the article by A. Krishna et al. [18], predict the sales of a retail store using different machine learning techniques and algorithms. In this way they determine which is the most suitable algorithm for their particular problem. They implemented normal regression techniques and boosting techniques in their approach, for which they finally found that boosting algorithms perform better than regular regression algorithms.

According to the article by J. Rincon-Patino et al. [19], present a machine learning approach to predict the amount Hass avocados can sell monthly and as well as knowing the total sales of Hass avocados in the United States territory. Historical sales records and weather data are the datasets they use to train the model. The 4 algorithms evaluated are Linear Regression, Multilayer Perceptron, Support Vector Machine and multivariate regression model. These last two are the ones that showed better precision.

According to the article by Yiyang Niu [20], he proposes a Walmart sales prediction model using XGBoost that combines with painstaking feature engineering processing. The model effectively extracts the different attributes of the dimensions and in this way makes good predictions. The results show that the method achieves superior performance over other machine learning algorithms. The metric used is RMSSE with which it is measured for all the models used. In addition, a set of data from Walmart supermarkets that were captured in Kaggle is used.

According to the publication of J. Chen et al. [21], propose a sales prediction model based on neural networks to predict Walmart sales. They evaluate this model based on the RMSE metric and compared with the linear regression algorithms and SVM algorithm, the neural network algorithm is more optimal. In addition to this, they use the SHAP technique to interpret the Neural Network (NN) model. The data set is obtained from

Kaggle, which is a portal where companies post databases to solve a complex problem, but if it is solved, a bonus is obtained.

In Gopalakrishnan T. et al. [22], analyze the sales of a store and predict its sales to help increase profits and make the brand more competitive. The technique used for sales prediction is the linear regression algorithm. To implement the model, they obtain historical sales data from 2 years ago and then also take data in real time. They compare forecast data to calculate accuracy and get very good results.

According to the research of H. Wang et al. [23], describe a product fraud detection model based on the decision tree algorithm, which combines feature engineering processing and algorithms to predict the sales problem of a certain product. The accuracy rate is higher with the decision tree algorithm than the logistic algorithm and the SVM algorithm.

According to the article by F. Barboza et al. [24], do research on bankruptcy using traditional statistical techniques and models such as discriminant analysis and logistic regression. As machine learning models they use SVM, random forest, bagging and boosting. They compare their performance with the results of discriminant analysis, logistic regression, and neural networks, using data from the years 1985 onwards. In this case study, the random forest algorithm had greater precision than the others.

In the publication Loureiro et al. [25], make use of a deep learning approach to forecast sales in the fashion industry, thus predicting sales of new products for future seasons. This study aimed to support a fashion retail company in its operations. The data set used is based on the company's historical sales data. With this study they compare sales predictions obtained with the deep learning approach based on a set of techniques such as decision trees, support vector regression, neural networks and linear regression. The algorithm with the best performance was random forest.

According to the research of Min Xia et al. [26], propose a hybrid method based on an extreme learning machine model with adaptive input metrics to improve the accuracy of sales prediction. The proposed algorithms are validated using real data from 3 fashion retailers. It was obtained as a result that the proposed model is practical for sales prediction and outperforms autoregression, artificial neural network and extreme learning machine models.

In the publication of the article by Massimo Aria et al. [27], are based on a study of the best performing and most used models in the framework of machine learning, the random forest algorithm. This is known as an efficient learning model as it ensures high predictive accuracy. But it is also considered a black box model due to the large number of deep decision trees that are produced in it. This study reviews methods and tools for ensemble tree interpretation.

According to the article by Hailong Cui et al. [28], mention that they are based on the development and comparison of models based on data to predict the volume of returns in the retail sector. The dataset is composed of detailed operations on

each product and the goal is to achieve good prediction accuracy. They found that the least absolute selection and contraction operator (LASSO) produces a predictive model that achieves the best prediction accuracy for this future return case.

According to the research of K. Saraswathi et al. [29] propose a prediction of Rossamann store sales using machine learning algorithms. They analyze customer buying behavior and forecast future sales demand with a dataset based on historical store sales data. The algorithms used for the prediction models are simple linear regression (SLR) and multiple linear regression (MLR), which produce a better prediction rate compared to other algorithms.

In the publication by K. Singh et al. [30], propose to build machine learning algorithms that are capable of predicting sales of an electronic commerce platform. Where in the end they compare all the precisions and the errors of the models in order to obtain the best model that has the lowest error and the highest precision to forecast sales. Using the random forest and gradient boosting algorithms, it was possible to specify that the best model is the one that used gradient boosting. And based on time series algorithms like ARIMA and SARIMA, ARIMA has the lowest error, but it is very high compared to gradient boosting prediction.

V. ANALYSIS OF RESULTS

This chapter describes the results obtained after reviewing the literature that corresponds to the research question. Table 3 reflects the results with 11 existing algorithms for product sales prediction.

Table 3 of existing algorithms for product sales prediction is presented below:

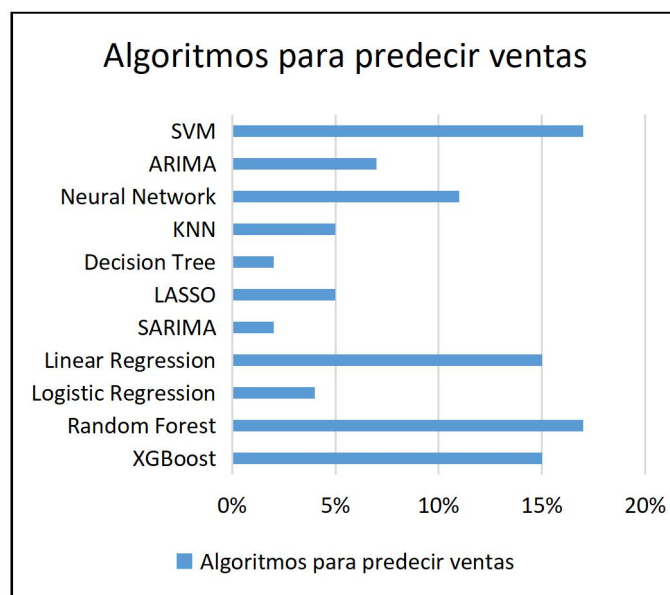


Figura 1. Muestra el número de algoritmos usados para predecir ventas en los artículos seleccionados

Tabla 3. Algorithms used to predict product sales

N°	Referencia	Algorithms used to predict product sales										
		XGBoost	Random Forest	Logistic Regression	Linear Regression	SARIMA	LASSO	Decision Tree	KNN	Neural Network	ARIMA	SVM
1	[10]	X										
2	[11]	X	X	X								
3	[12]		X		X							
4	[13]		X				X		X			
5	[14]		X							X	X	
6	[15]	X										
7	[16]								X			X
8	[17]	X								X	X	X
9	[18]				X							
10	[19]				X							X
11	[20]	X										
12	[21]				X					X		X
13	[22]				X							
14	[23]			X								X
15	[24]	X	X									X
16	[25]		X		X			X		X		X
17	[26]									X		X
18	[27]		X									
19	[28]						X					
20	[29]				X							
21	[30]	X	X			X					X	
Total		7	8	2	7	1	2	1	2	5	3	8
%		15	17	4	15	2	5	2	5	11	7	17

Results:

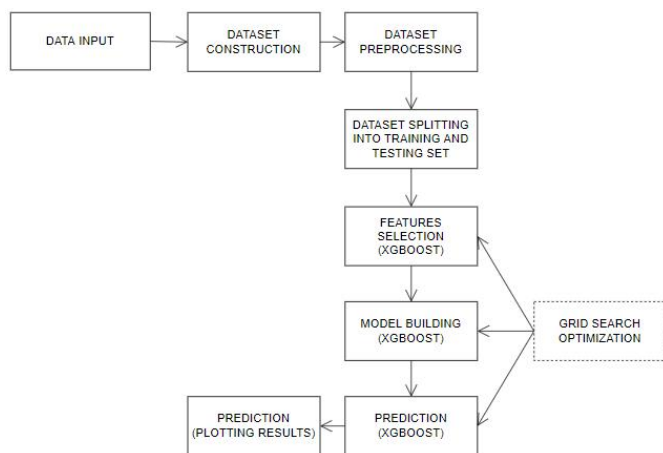
- In 6 articles analyzed, they use the XGBoost algorithm to predict sales, this represents 14% of the total articles in this investigation.
- In 4 articles analyzed, they use the Random Forest algorithm to predict sales, this represents 9% of the total articles in this investigation.
- In 2 articles analyzed, they use the Logistic Regression algorithm to predict sales, this represents 5% of the total articles in this investigation.

- In 7 articles analyzed, they use the Linear Regression algorithm to predict sales, this represents 16% of the total articles in this investigation.
- In 1 article analyzed, they use the SARIMA algorithm to predict sales, this represents 2% of the total articles in this investigation.
- In 3 articles analyzed, they use the LASSO algorithm to predict sales, this represents 7% of the total articles in this investigation.
- In 3 articles analyzed, they use the Decision Tree algorithm to predict sales, this represents 7% of the total articles in this investigation.
- In 1 article analyzed they use the KNN algorithm to predict sales, this represents 2% of the total articles in this investigation.
- In 5 articles analyzed they use the Neural Network algorithm to predict sales, this represents 11% of the total articles in this research.
- In 3 articles analyzed, they use the ARIMA algorithm to predict sales, this represents 7% of the total articles in this investigation.
- In 8 articles analyzed, they use the SVM algorithm to predict sales, this represents 18% of the total articles in this investigation.

In answer to the research question there are 11 most relevant algorithms to predict sales based on research and items found.

VI. ALGORITHM PROPOSAL

In this section, a proposal is made for a new prediction model based on the construction with the use of the XGBoost algorithm for the prediction of product sales, this having as precedents the results obtained after reviewing the literature. According to the literature reviewed and the results obtained, the XGBoost algorithm was found to be the most efficient when used as a training and data processing base for a set of historical sales data. This algorithm is useful for its efficiency, precision and low level of error for prediction and although the most common algorithm is SVM, it is not as efficient as the recommended algorithm for the case presented.



FLOWCHART OF THE XGBOOST ALGORITHM FOR SALES PREDICTION

The detail of each of the boxes belonging to the flowchart of the proposed model with the XGBoost algorithm for sales prediction is:

- **DATA INPUT:** The algorithm takes as input the configuration data such as what type of data is the field to predict, the number of fields in the dataset that will be processed and the dataset training time.

- **DATASET CONSTRUCTION:** The data segment assembled in a database or datawarehouse where the necessary fields are available to be able to predict the selected field is defined.

- **DATASET PREPROCESSING:** In the model there is a pre-processing to validate if the configuration entered in the Data Input and the DataSet Construction match and have what is necessary for the algorithm to process the data without problems.

- **DATASET SPLITTING INTO TRAINING AND TESTING SET:** In this process, the data set is divided into the training set and the test for the subsequent validation of the prediction.

- **FEATURES SELECTION:** In this process, the model is entered which date or range of dates to predict is needed and based on which date type variable will be taken as reference.

- **MODEL BUILDING:** The prediction model is built based on the XGBoost algorithm, the data splits (Training) are also taken and passed as parameters with the previously indicated characteristics, the prediction is started.

- **PREDICTION:** The prediction model already has a result of all the previous processes, but they are data that cannot be seen from the model itself, but later this result is processed by a system or set of graphs where the results are understood. Before giving a result, the model processes the result with the test data split and validates the precision and margin of error.

- **RESULTS:** The results produced by the prediction model are displayed based on graphs for the understanding of the result, based on the field to be predicted and related to the dates indicated in the selection of characteristics.

- **GRID SEARCH OPTIMIZATION:** The model from the selection of features to the evaluation of the prediction is optimized under another algorithm that is the grid search and whose main objective is to select the best parameters for the optimization problem from a list of parameters, automates the trial and error method.

VII. CONCLUSIONES

In this research, 50 articles were analyzed, of which 32 were relevant to the study and 21 articles were selected to provide support and information for the study. When carrying out a more detailed analysis of the selected articles, it was obtained that there are 11 machine learning algorithms that are used to predict sales based on datasets of historical data. As a result, it was possible to obtain the most used algorithm, which is the SVM, but the most efficient algorithm for these cases is XGBoost.

In the proposal, a new prediction model based on the construction with the use of the XGBoost algorithm is detailed and a flowchart is drawn where the algorithm is used from the selection of data to the presentation of results in the prediction.

REFERENCES

- [1] Alonso I.A., Verdún J.C., Caro E.T. (2010) Information Technology to Help Drive Business Innovation and Growth. In: Sobh T., Elleithy K. (eds) Innovations in Computing Sciences and Software Engineering. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-9112-3_90
- [2] I. A. Alonso, J. C. Verdún and E. T. Caro, "The Importance of IT Strategic Demand Management in Achieving the Objectives of the Strategic Business Planning," 2008 International Conference on Computer Science and Software Engineering, 2008, pp. 235-238, doi: 10.1109/CSSE.2008.1307
- [3] I. A. Alonso, J. C. Verdún and E. T. Caro, "The IT implicated within the enterprise architecture model: Analysis of architecture models and focus IT architecture domain," 2010 IEEE International Conference on Service-Oriented Computing and Applications (SOCA), 2010, pp. 1-5, doi: 10.1109/SOCA.2010.5707174.
- [4] Ryll, L., & Seidens, S. (2019). Evaluating the performance of machine learning algorithms in financial market forecasting: A comprehensive survey. <https://arxiv.org/abs/1906.07786>
- [5] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. <https://www.deeplearningbook.org/>
- [6] Lantz, B. (2019). Machine learning with R: Expert techniques for predictive modeling. Packt Publishing.
- [7] Mohammed, M., Khan, M. B., & Bashier, E. B. M. (2016). Machine learning: Algorithms and applications. CRC Press. <https://doi.org/10.1201/9781315371658>
- [8] Lytvynenko, T. I. (2016). Problem of data analysis and forecasting using decision trees method
- [9] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," Inf. Softw. Technol., vol. 51, no. 1, pp. 7-15, 2009, doi: 10.1016/j.infsof.2008.09.009.

- [10] F. Haselbeck, J. Killinger, K. Menrad, T. Hannus, D. G. Grimm, "Machine Learning Outperforms Classical Forecasting on Horticultural Sales", 2021. <https://doi.org/10.1016/j.mlwa.2021.100239>
- [11] D. Rohaan, E. Topan, C.G.m. Groothuis-Oudshoorn, "Using supervised machine learning for B2B sales forecasting: A case study of spare parts sales forecasting at an after-sales service provider", 13 October 2021. <https://doi.org/10.1016/j.eswa.2021.115925>
- [12] D. Raditya, N. Erlin, F. Amanda, N. Hanafiah, "Predicting Sneaker Resale Prices using Machine Learning". 2021. 10.1016/j.procs.2021.01.037
- [13] M. Leippold, Q. Wang, W. Zhou, "MAchine learning in the Chinese stock market". 28 October 2021. <https://doi.org/10.1016/j.jfineco.2021.08.017>
- [14] B. M. Pavlyshenko, "Machine-Learning Modeles for sales time series forecasting". 18 January 2019. <https://doi.org/10.3390/data4010015>
- [15] S. Cheriyan, S. Ibrahim, S. Mohanan, S. Treesa, "Intelligent Sales Prediction Using Machine Learning Techniques". Aug. 2018. <https://doi.org/10.1109/ICCECOME.2018.8659115>
- [16] I-Fei & Chi-Jie Lu, "Sales forecasting by combining clustering and machine learning techniques for computer retailing" 2016. <https://doi.org/10.1007/s00521-016-2215-x>
- [17] Mohit Gurnani, Y. Korke, P. Shah, S. Udmale, V. Sambhe, S. Bhirud, "Forecasting of sales by using fusion of machine learning techniques". 2017. DOI: 10.1109/ICDMAI.2017.8073492
- [18] A. Krishna, A. V, A. Aich and C. Hegde, "Sales-forecasting of Retail Stores using Machine Learning Techniques," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2018, pp. 160-166, doi: 10.1109/CSITSS.2018.8768765.
- [19] Juan Rincon-Patino, Emmanuel Lasso, Juan Carlos Corrales, "Estimating avocado sales using machine learning algorithms and weather data" 2018. <https://doi.org/10.3390/su10103498>
- [20] Y. Niu, "Walmart Sales Forecasting using XGBoost algorithm and Feature engineering," 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), 2020, pp. 458-461, doi: 10.1109/ICBASE51474.2020.00103.
- [21] J. Chen, W. Koju, S. Xu and Z. Liu, "Sales Forecasting Using Deep Neural Network And SHAP techniques," 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), 2021, pp. 135-138, doi: 10.1109/ICBAIE52039.2021.9389930.
- [22] G. T., R. Choudhary and S. Prasad, "Prediction of Sales Value in Online shopping using Linear Regression," 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018, pp. 1-6, doi: 10.1109/CCAA.2018.8777620.
- [23] H. Wang, F. Yang and S. Shen, "Supply Fraud Forecasting using Decision Tree Algorithm," 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), 2021, pp. 344-347, doi: 10.1109/ICCECE51280.2021.9342556.
- [24] F. Barboza, H. Kimura, E. Altman, "Machine learning models and bankruptcy prediction", 2017. <https://doi.org/10.1016/j.eswa.2017.04.006>
- [25] A.L.D. Loureiro, V.L. Miguéis, Lucas F. M. da Silva, "Exploring the use of deep neural networks for sales forecasting in fashion retail". 2018. <https://doi.org/10.1016/j.dss.2018.08.010>
- [26] Min Xia, Yingchao Zhang, Liguang Weng, Xiaoling Ye, "Fashion retailing forecasting based on extreme learning machine with adaptive metrics of inputs". 2012. <https://doi.org/10.1016/j.knosys.2012.07.002>
- [27] Massimo Aria, Corrado Cuccurullo, Agostino Gnasso, "A comparison among interpretative proposals for Random Forests". 2021. <https://doi.org/10.1016/j.mlwa.2021.100094>
- [28] Hailong Cui, Sampath Rajagopalan, Amy R. Ward, "Predicting product return volume using machine learning methods". 2020. <https://doi.org/10.1016/j.ejor.2019.05.046>
- [29] K. Saraswathi, N. T. Renukadevi, S. Nandhimidevi, S. Gayathridevi, P. Naveen, "Sales Prediction using machine learning approaches". 2021. <https://doi.org/10.1063/5.0068655>
- [30] Karandeep Singh, Booma P M, Umapathy Eaganatha. "E-commerce system for sale prediction using machine learning technique". 2020. doi:10.1088/1742-6596/1712/1/012042