Strategic Management of Sales Assisted by Machine Learning: A Case Study in a Large Food Business

Marcus Vinícius Silva[1], Luiz Melk de Carvalho[1], Diva de Souza e Silva Rodrigues[1], Flávio Henrique Batista de Souza[1]

[1]*Centro Universitário de Belo Horizonte – UNIBH, Brazil*

Abstract.

This paper implemented Machine Learning techniques to predict sales success of a large Brazilian food company. The strategy explores the use of recorded data to apply a pattern recognition process based in Multilayer Perceptrons (MLP) algorithms and thus proposes an efficient management of business proposals. As a research methodology, a database with information collected over 4 years (23,093 records) in an Enterprise Resource Planning (ERP) system was evaluated and standardized. Then sales success prediction experiments were performed, using MLPs of the following types: Standard Backpropagation, Backpropagation Momentum, Resilient Propagation, Backpropagation with Weight Decay and Quick Propagation. Resampling by test and validation sets was used, with the proportion variation of 65-35% and 75-25%. The number of neurons in the hidden layer were {3,5,7,10} and the periods of interaction were {50;100;1,000;5,000}. The Receiver Operating Characteristics (ROC curve) and its Area Under the Curve (AUC) were employed to evaluate the performance, considering the normalization of the measured AUCs between 0 and 1. The experiments are focused on the annual and monthly prediction of sales success, according to the profile of the customer. The main achievements of this work were: a methodology for collecting and standardizing the data of the company; an analysis of the best MLP configuration, according to the profile of the customer; and a structure that has an AUC between 0.975 and 0.983, with a processing time between 4.2 and 4.8 seconds, capable of supporting the decision taking of the company and reducing the loss of potential contracts (which reached the order of more than $180,000.00 Brazilian reais). Besides, the implemented approach allows the prediction of renegotiations and success cases.

Keywords: Machine Learning, Pattern Recognition, Multilayer Perceptron, Strategic Sales Management, Food Sector.

1. Introduction

By the use of pattern recognition algorithms, it is possible to predict relevant aspects about the operations in the sales of products that a commercial department offers to the customers, in order to determine a more precise recommendation according to the purchase profile of each customer (Ramochaeva & Klionskiy, 2019). According to Coccia (2019), a new technology is not necessarily a new technology lauchedon the market, but it can be, according to the interest of the analysis to be carried out, something that is new for the company bussiness.

The world crisis scenario, caused by the pandemic due to COVID-19, accentuates the need for accurate technologies to conduct the business, although the complexity of the negotiations has not decreased (Landier & Thesmar, 2020).

For processes that use complex control systems, it is possible to apply the concept of Machine learning, using pattern recognition algorithms. One of the main techniques used is Artificial Neural Networks (ANN), which in addition to the ability of modelling complex linear and non-linear systems, also present learning mechanisms (Wu et al., 2019). ANNs start from a premise, developed in 1958, by Frank Rosenblatt, who proposed a network topology called Percéptron, in which these are the neurons present in the network (which has the input layer, the hidden layers and the output layer), which made it possible to increase the work related to neural networks until 1969, generating the Multilayer Percentron (MLP) (Rosenblatt, 1958).

The application of these algorithms to real problems requires ways of assessing the accuracy of their predictions. Thus, some comparison tools started to be analyzed, where it was possible to measure and specify the performance problems of algorithms, such as the ROC Curve (Receiver Operating Characteristic) and AUC (Area Under the Curve). As the area under the ROC curve identifies a fraction of the area of a square, its value is always between 0 and 1. Based on the results generated by the algorithms, it is possible to generate graphic analysis and obtain a measure of evaluation of models in machine learning with less deficiency in face of a rate of classification error (Nhu et al., 2020).

A structure based on MLP algorithms, with assessments via AUC, are still considered important in the literature, in several segments of the industry, such as in health, education, supply chain and other segments (Souza et al., 2019; Rifat et al., 2019 ; Lima-Junior & Carpinetti, 2019).

Accordding to the presented technological context, it was taken as a reference for studies and experiments, a Brazilian company with more than 50 years on the market of food-related business. It is a traditional company, with thousands of direct and indirect employees. After a study carried out at the company, it was verified that the data generated during its commercial transactions were not treated correctly. There was no purpose for the recorded data after the transactions. Unfortunately, it was a finding that agrees with one fact: the BI technique used for this particular situation had not been widely adhered by the company. It had being only addressed in scientific research in the production department, without aiming at the way the sale and distribution of their products could be made. Currently, there are several types of systems capable of performing statistical calculations for recommendations of better products for certain customers, but they do not have any analysis and processing of this data that proposes a system learning feature.

In order to propose a strategic management, using these data, based on the demonstrated technological prerogative, some of these questions guided the work: what are the characteristics of customers and products during the sales transaction process that can compose a database for pattern recognition of successful transactions? what would be the correctness of MLP, a structure well founded in the literature and still with recent researches for this type of business? what is the performance of MLPs according to the size of the database obtained? what would be the benefits for the company and the customer in face of such an implementation?

Faced with such questions, the objective of this work is to seek, through an analysis and treatment of the database of the commercial sector of the company in question, to apply the pattern recognition technique with MLP-type algorithms. The main purpose of this work is that the structure developed has the ability to predict to the manager, according to the profile of the customer, whether the sales will be successful or a renegotiation will be necessary.

As specific objectives, this research aims to demonstrate: the analysis of the company database in question, looking for the relevant characteristics of the products and their customers from 2015 to 2018; the development of a sequence of experiments with MLPs to evaluate the assertiveness and performance of the structure; and demonstrating the strategic feasibility achieved. In addition to the proposed objectives, it is intended that the procedure will become standard for negotiations during and after the pandemic period, for subsistence and strengthening of the company.

2. Methods

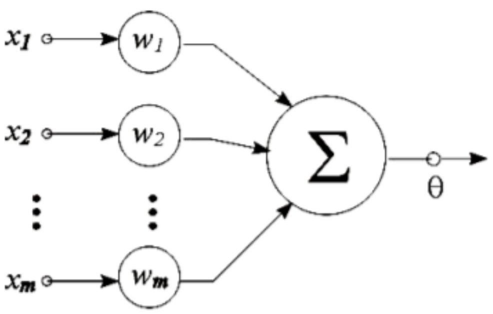
2.1 Analysis of Machine Learning and Pattern Recognition Technology

The Machine Learning perspective allows computer systems to employ the concept of learning, through the knowledge of previous data in a context. Regarding learning, one of the concepts used is pattern recognition. Pattern recognition can be employed by algorithms, such as artificial neural networks (Emmert-Streib & Dehmer, 2019). In this research, Multilayer Perceptron algorithms are used to predict sales success. During its use, a resampling process based on a test and validation set is performed. Finally, an assessment through AUC is carried out.

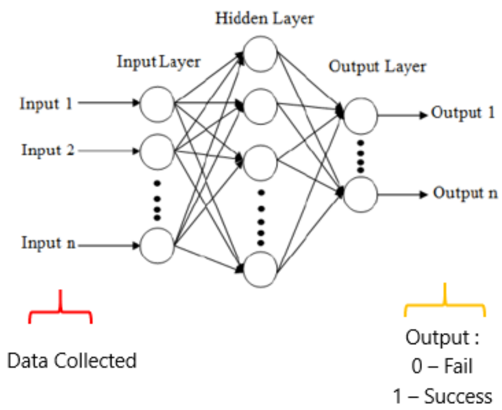
2.1.1 Multilayer Perceptron

The Multilayer Perceptron (MLP), is a type of artificial neural network (ANN), in which it is applied in problems of classification of patterns and predictions, with structures as described in Figure 1.

Figure 1: Perceptron model versus MLP model



(a)



(b)

Source: adapted from (Rosenblatt, 1958)

The single-layer perceptron is only able to classify linearly separable patterns. Its neuron shape (see Figure 1a) can be explained by placing x_i as inputs of the network and w_i as the synaptic weights. The synapse is known to make the connection between neurons, the same is similar to the synaptic weights, which in the connections are assigned values due to the connections.

In practice, the problem to be addressed does not allow an exact linear separation, making it necessary to use a multilayer perceptron, as shown in Figure 1b (Mohammed et al., 2016).

An MLP network is subdivided into layers: data input layer, intermediate or hidden layers (where perceptrons are interconnected) and output layer. Through the result of this calculation, it is possible to compare the calculated data with data generated by the database, finding possible errors, which is propagated in the opposite direction to the execution to adjust the network weights (Mohammed et al., 2016).

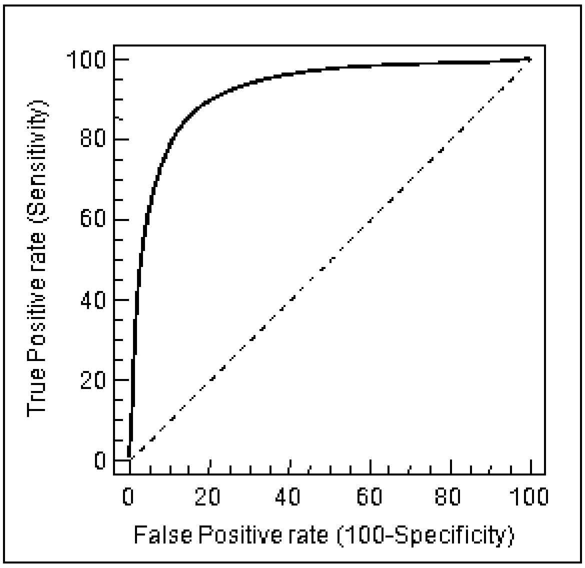
2.1.2 Resampling Methods

Resampling techniques are used to obtain classifier performance statistics. These techniques can be used to split a set of data into different subsets, working with training, validation and testing of the model (Berrar, 2019). The first subset is used to estimate parameters for the model definition and to validate the performance evaluation of the model for the set that has not been presented. In the second subset, single validation is used to perform tests on the model.

2.1.3 ROC Curve and AUC

The ROC Curve (Receiver Operating Characteristic) is widely used to visualize and analyze the behavior of diagnostic systems (Kannan; Vasanthi, 2019). Figure 2 shows the true positive rate (sensitivity) as a function of the false positive rate (specialty).

Figure 2: ROC Curve



Source: (Schoonjans, 2018)

Schoonjans (2018) defines AUC (Area Under the Curve) as the area under the ROC curve that identifies a fraction of the area of a square and its value is always between 0 and 1.

2.2 Experimental Methods

The work aims to demonstrate the development of an experimental research. An internal research was carried out in the company, identifying the departments that most impacted the profit, the processes and the satisfaction of the customers. Two main processes were identified during the development of the analyzes and one of them were taken as the basis for the development of this work.

The first one was the tax department, responsible for opening new branches and distribution centers, but despite its great influence within the company, the data provided did not have a necessarily robust history to apply the pattern recognition method and artificial neural networks .

The second process and the one developed throughout this work impacts the commercial area of the company. This department is responsible to present the product sales metrics and to identify the preferences of the curstomer through the profile of each customer present in the database provided by the company. The data were collected from the SAP system, that is the ERP system used by the company. Through the transactions of the Controllership module (CO), an ALV report (Advanced Business Application Programming List Viewer - ABAP List Viewer) was developed containing characteristics such as: sale period (month / year) of the product, quantity, weight and price of each sale. It was also necessary to collect the order number of the customer for identification of the customer and of the branch in which they operate, among other details that determine the purchase value of the products.

The Resampling technique used was the Validation Set, in order that the system can recognize the characteristics of the database and validate what was its performance after the learning process. Through this technique, the validation was applied to the MLP network algorithm, verifying the ability of the system to learn. The evaluators of the algorithms were the AUC curve and the processing time, which aims to enable a comparison and analysis between them, in order to define which was the algorithm that presented the best behavior for the proposal and the ability to treat through the score obtained by the curve, the main characteristic of the customer and the probability of the sale that may or may not be completed.

3. Results and Discussion

3.1 Data Proccessing

Data collection started through the development of a system in the ABAP language. The system ALV's Grid reports data of the years 2015 (November and December), 2016 and 2017 (January to December) were generated.) and 2018 (January to October). Among the variables gathered only those that added relevance to the neural network were kept, such as: Month / Year, Service Center, Material Code, City of Sale, Order Region, Production Center, Basic Unit of Measure (UM), Sales code (sale made equal to 1 or sale not made) in which the sale was obtained, which was successful or not (0 or 1), in a certain period of the year. Table 1 presents the amount of data generated through the previous-mentioned years and the number of months analyzed.

Table 1 – Annual Data.

Year	Months	Samples
2015	2	1000
2016	12	5524
2017	12	5524
2018	10	5524
Total	-	23093

Source: (Authors, 2020)

3.2 Accuracy Experiments

The database processing was carried out using R Language, responsible for generating the results of the statistical analysis.

Through the structure used, the data were processed, generating the output data, such as: the AUC Curve and the processing time of each experiment.

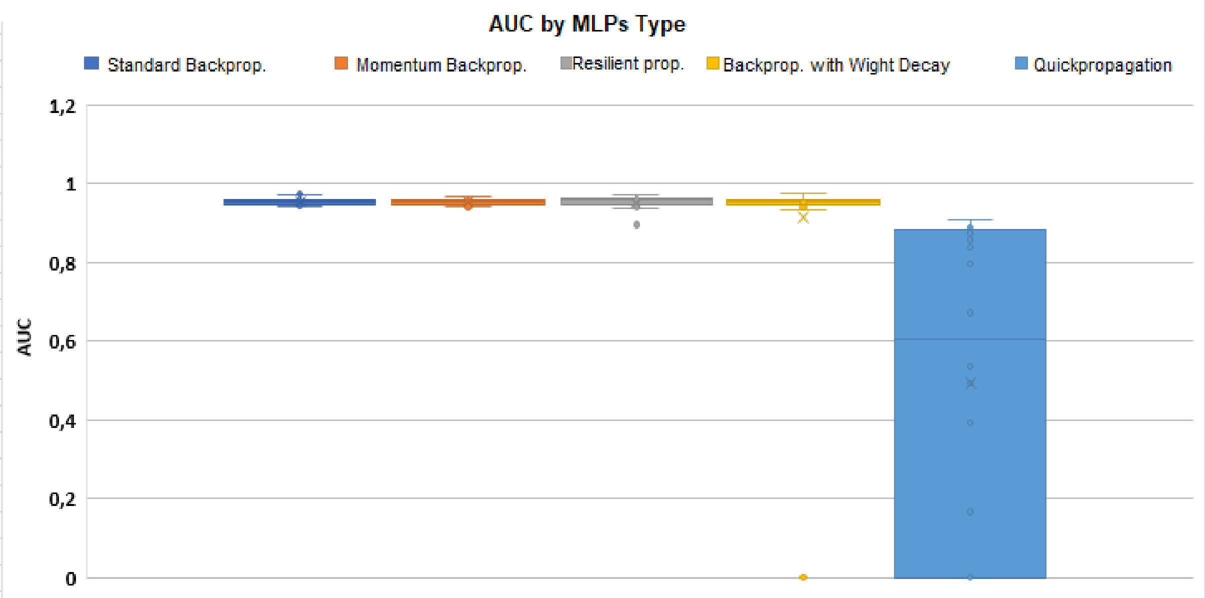
Variations of: five types of MLPs (Standard Backpropagation, Backpropagation Momentum, Resilent propagation, Backpropagation with Weight Decay and Quickpropagation) were used; numbers of neurons in the hidden layer (3, 5 and 10); two resampling rates for validation (25% and 35%) and training (75% and 65%); maximum number of seasons (50, 100, 1000 and 5000). In this way, scenarios were generated in which the company can present a higher number of sales in a certain period of the year.

3.1.1 Global Sales Scenario

The first experiment was carried out on the complete sales basis of the commercial department, varying the five types of MLPs selected.

Thus, the variation of the AUC curve and the processing time was obtained through the variations that were filtered. In the graph of Figure 3, it can be seen that when changing the type of MLP in the complete data base, there is no significant variation in the data, except in Quickprop in which the AUC values reach values close to 0.

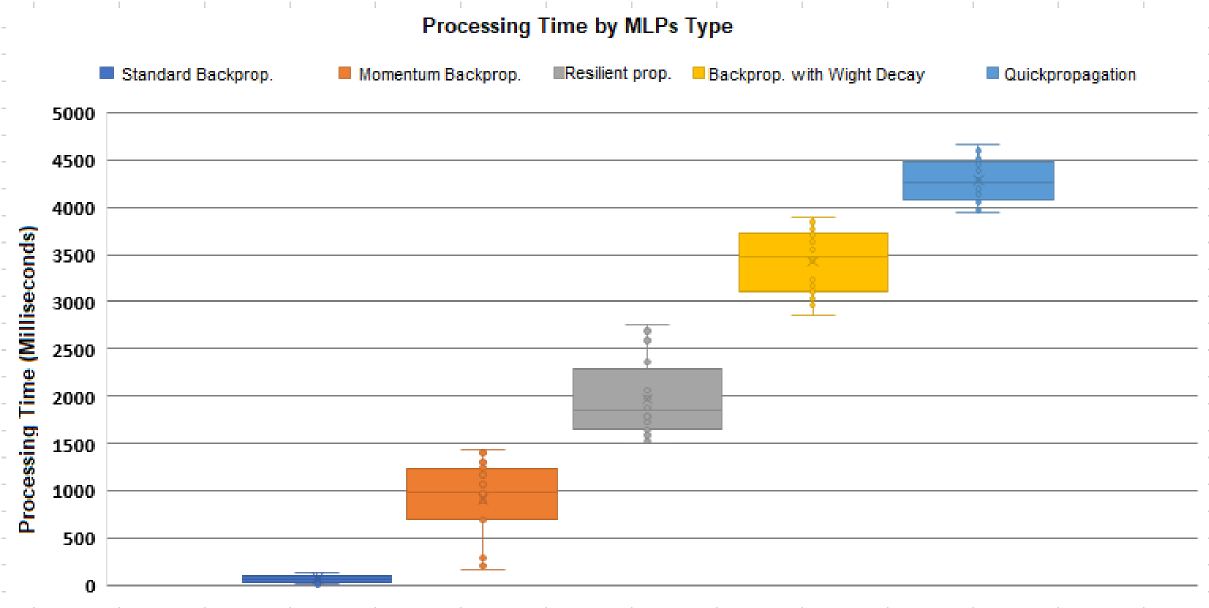
Figure 3: Evaluation of AUC by MLPs in global base



Source: (Authors, 2020)

In the graph of Figure 4, the processing time grows proportionally with the change of the types of MPLs, in which the one with the best time was the Standard Backpropagation. Therefore, the best result was found through Backpropagation, in which the AUC was closer to 1 and the processing time was between 0 and 500 milliseconds.

Figure 4: Analysis of MLP processing time on the global basis

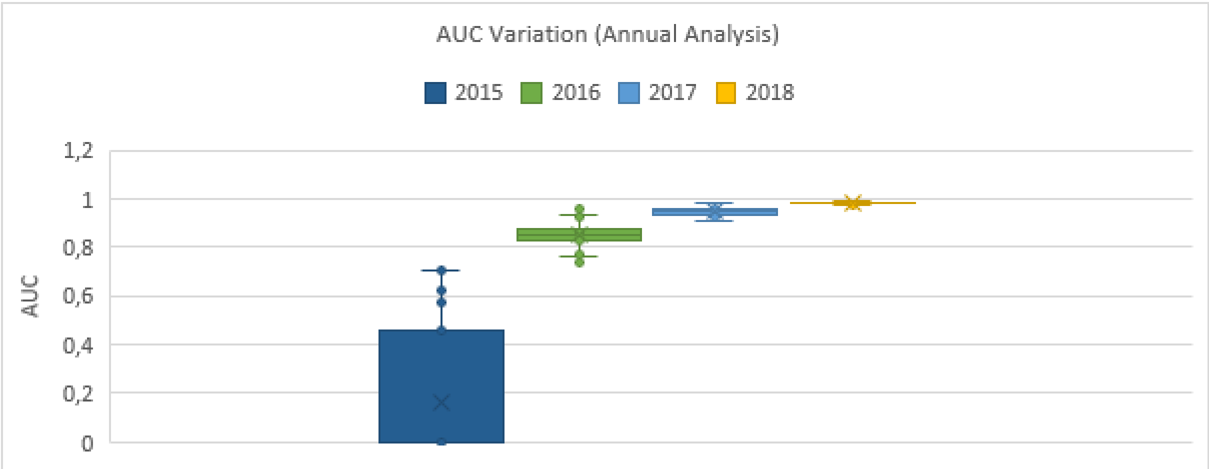


Source: (Authors, 2020)

3.1.2 Annual Experimentation

The second experiment was carried out on the database of the years 2015, 2016, 2017 and 2018, using the same variables of the first experiment, with the difference of using only one experimental model: the Standard Backpropagation. In Figure 5, it can be seen that the year 2015 was the year with the lowest seasonality rate and the highest sampling rate in relation to the other years. As it had the lowest data collection rate, it affected the analysis. Because of that, the algorithm did not generate an accurate analysis of that dataset.

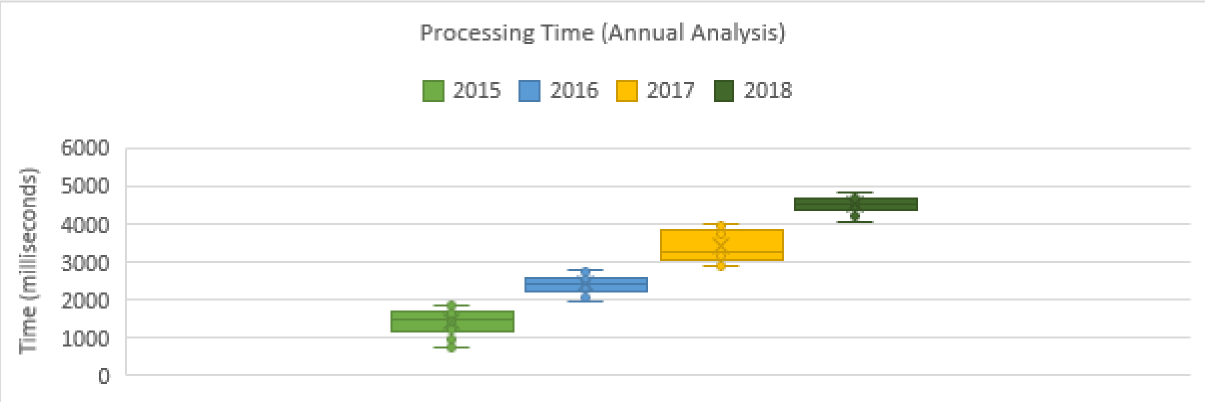
Figure 5: AUC in Annual Analysis



Source: (Authors, 2020)

The years 2016 and 2017 were stable, indicating that there was not a very significant change in these sales years. As can be seen in Figure 6, 2018 is the year with the worst processing time.

Figure 6: Processing Time in Annual Analysis



Source: (Authors, 2020)

The year 2015 is the one with the best processing time in terms of the amount of data, while the years of 2016 and 2017 have a high processing time, but are still below the year of 2018. Thus, the best result that could be achieved on the experiments of the complete annual base was in 2018, in which although the base does not have the last two months of the year, its AUC was closer to 1 and despite high processing time, it did not suffer great variations over the months, meaning that it was a stable year in the sales of products with less returns of goods.

3.1.3 Monthly Experimentation

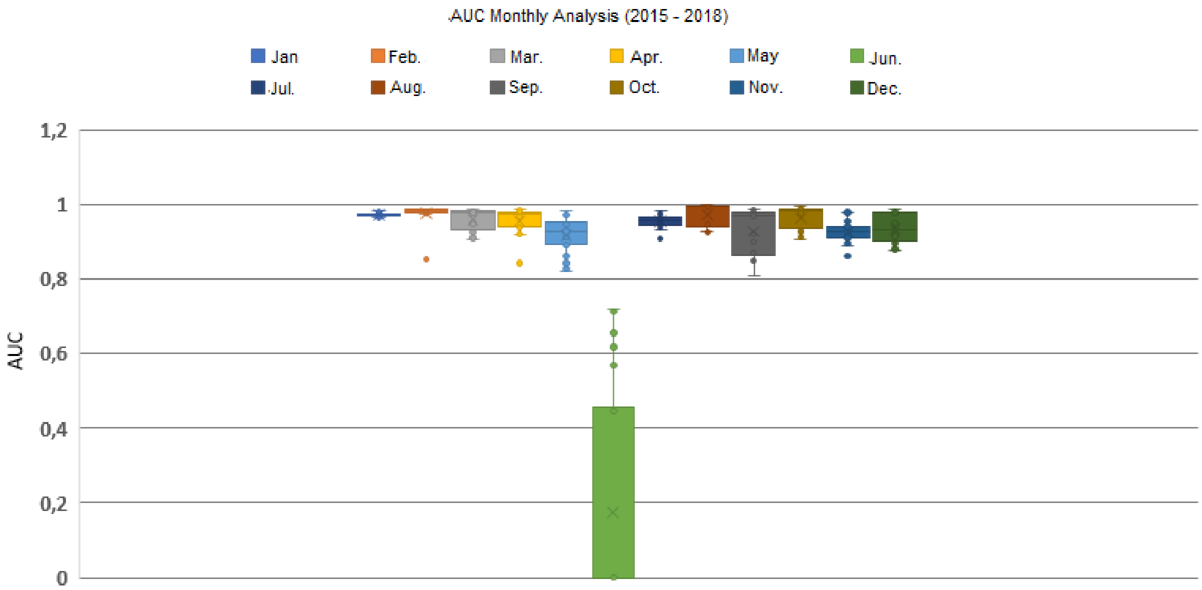
The third set of experiments carried out was the monthly analysis. All annual data from 2015 to 2018 were grouped in the months from January to December using the same variables of the first experiment and only one experimental model: the Standard Backpropagation.

It was observed that the month with the highest seasonality index was June, as this was the month with the highest number of AUC curves equal to or closer to 0. The months in which the best results were obtained with the least variations were January and February, in which the curve was closer to 1, as shown in the graph in Figure 7. The processing time is shown in Figure 8.

In Figure 8, it can be seen that the month with the longest processing time, was the month of May, in which it was the one that had the largest number of variations of experimental data over the years. The one with the shortest time was June, which has the closest AUC curve to 0. The largest variation in processing time was from July in which the processing reached from 0 to almost 4,000 milliseconds. In the graph of Figure 9, it can be seen a significant increase on the number of epochs, given the complexity of the actual data and there was no drastic interference in the processing time, despite the more prevalent complexity in processing the higher number of data.

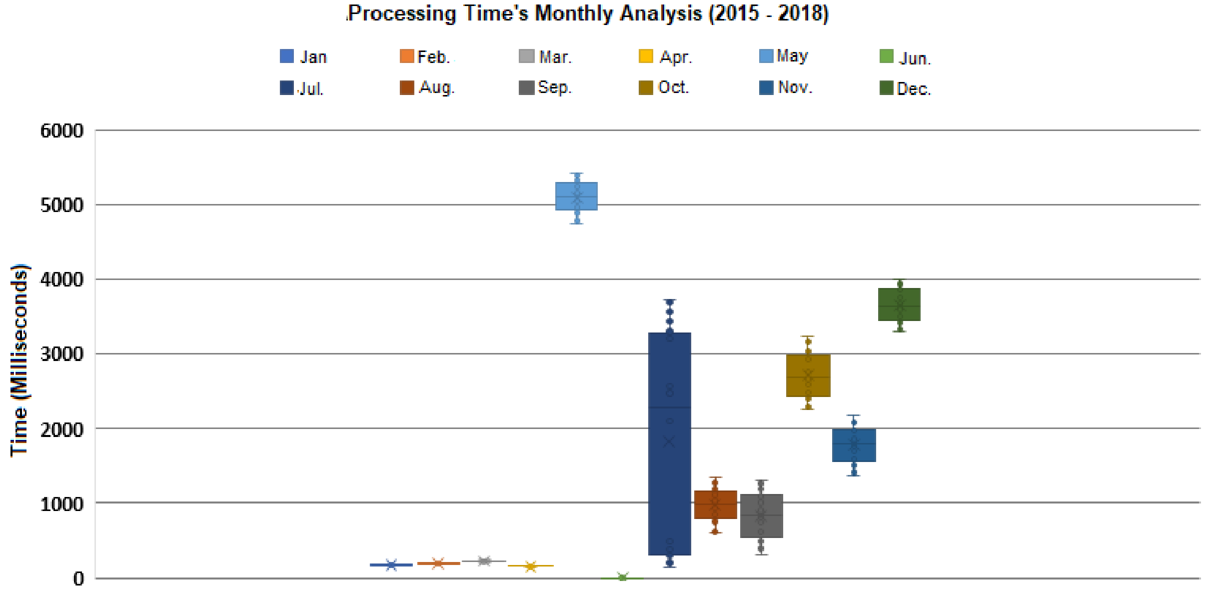
Thus, variations in processing time are due not only to the number of periods, but also to the set of neuron variables and test samples. The 5,000-epoch option, with 10 neurons and 0.35 resampling (35% of samples for validation), which has a time span of approximately 4.816 ms, does not get as relevant an analysis as a 50-epoch option, with 10 neurons and resampling 0.35 (elapsed time of approximately 4.736 ms).

Figure 7: AUC in Monthly Analysis



Source: (Authors, 2020)

Figure 8: Processing Time in Monthly Analysis

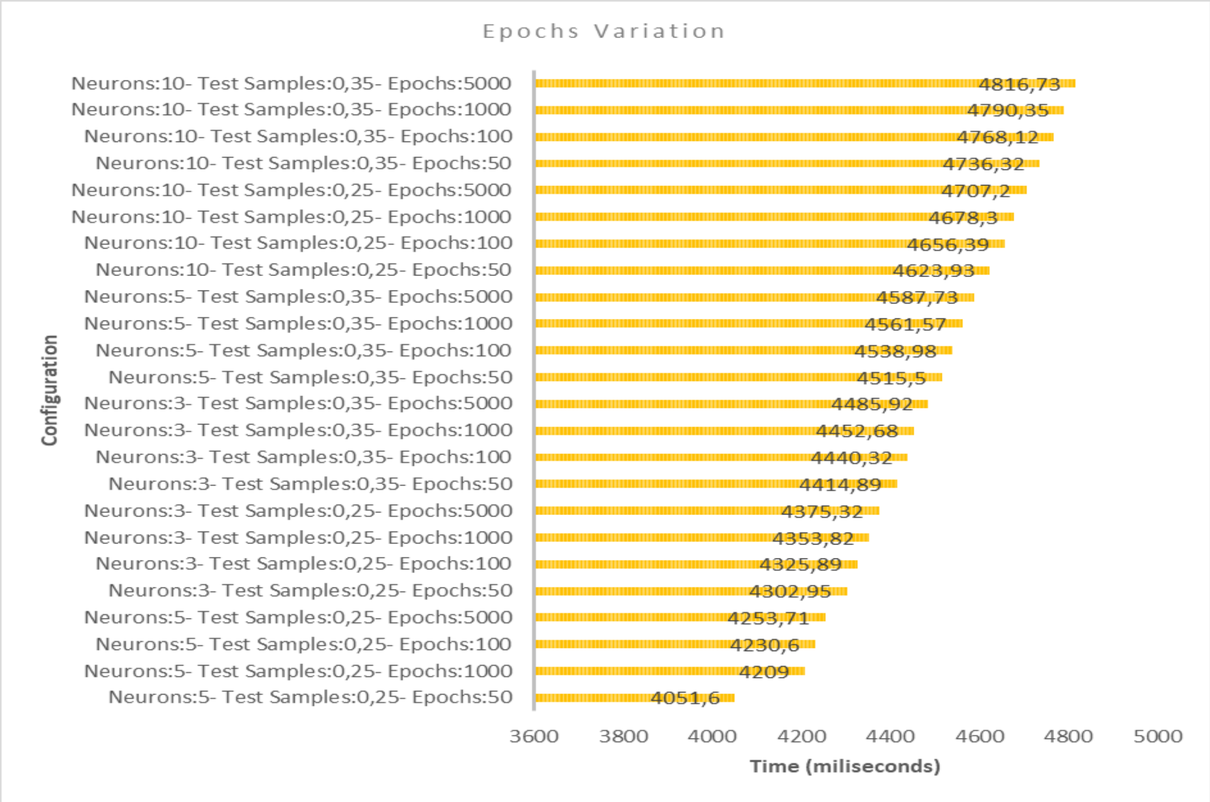


Source: (Authors, 2020)

When analyzing Figure 10, it can be seen that the number of epochs does not decisively influence the AUC curve. The results thus are also dependent on neuron variables and test samples. Based on that, the 5000-epochs option AUC curve, with 3 neurons and 0.35 resampling (with AUC equal to 0.9932927) does not obtain an AUC as relevant as a 50-time option, with 10 neurons and 0 resampling, 35 (with AUC equal to 0.9881057).

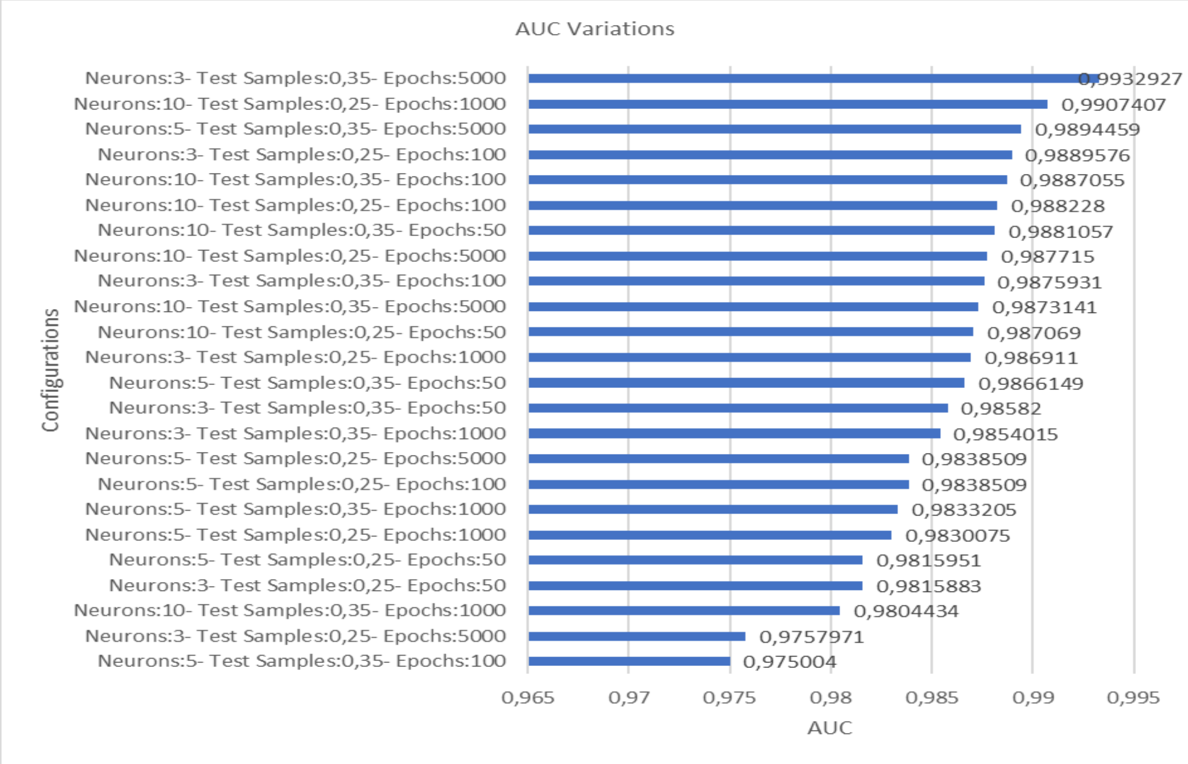
In the fourth and last test, it was performed for period variations, in which it is noticed that the period size does not directly influence the value of the AUC Curve, being necessary to analyze the number of neurons and resampling.

Figure 9: Processing Time in Monthly Analysis Ranking



Source: (Authors, 2020)

Figure 10: Neurons in Hidden Layer Analysis



Source: (Authors, 2020)

3.3 Impacts on Strategic Management

Through the developed study, it was possible to determine a standardization for the collection of the data to be used, for a management guided by pattern recognition algorithms.

As a result, as the user enters the characteristics used by the structure to train and validate MLP, it is possible to predict whether the proposed sale will be successful. An explicit case can be demonstrated in a sale proposal in the amount of R$ 186,184.32 in May of 2017, which was not closed. The structure predicts such a situation as 0 (there will be no sale).

According to this alert, the manager could change the strategy of the supplier center, the date or the number of units to be directed for sale, being able, for example, to offer a negotiation and a leaner quantity, with a value of R$ 140,000.00 that could have closed the sale, even with approximately R$ 46,000.00 less than the inittial amount.

4. Conclusion

Based on the aspects observed during the development of the work, it was concluded that it is possible to implement a Pattern Recognition algorithm for food sales management, seeking the assertiveness of the best sale of products at a certain time of the year for a given customer company and in the future for new customers. Through the generated database, relevant results of the analysis of the attributes of the experiments were presented.

As shown in Figure 10, results were obtained from the AUC curve between 0.993 and 0.975, which determines the company's selling and purchasing power, and the processing time, demonstrating that despite the high processing time of the data generated in a complete annual basis, a more in-depth and relevant content for research can be obtained from the AUC.

One of the contributions of this work is demonstration through experimentation and exhaustive tests of varied options, algorithms and resampling techniques for recognizing patterns of product sales.

Another contribution of the work is to demonstrate to the customer new possibilities of purchase at certain periods of the year, so that it is possible to diversify the portfolio catalog of the company and focus on products that are more available at certain periods of the year and that are below the market price.

The future possibilities of the work are the integration of tools already existing in the company, attributing the algorithm already developed in the SAP® Leonard Machine Learning system, integrating to the ERP system that is already in the company, in order to obtain the complete analysis of the data already available and future data collected. And also, in the development of an application in the SAP® Analytics Cloud capable of demonstrating the results achieved from sales in the company in real time.

References

[1] Coccia, M. (2019). Why do nations produce science advances and new technology? *Technology in society*, 59, 101124.

[2] Razmochaeva, N. V. and Klionskiy, D. M. (2019). Data Presentation and Application of Machine Learning Methods for Automating Retail Sales Management Processes. In 2019 *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering* (EIConRus) (pp. 1444-1448). IEEE.

[3] Lima-Junior, F. R. and Carpinetti, L. C. R. (2019). Predicting supply chain performance based on SCOR® metrics and multilayer perceptron neural networks. *International Journal of Production Economics*, 212, 19-38.

[4] Rifat, M. R. I., Al Imran, A. and Badrudduza, A. S. M. (2019). Educational performance analytics of undergraduate business students. *International Journal of Modern Education and Computer Science*, 11(7), 44.

[5] Emmert-Streib, F. and Dehmer, M. (2019). A machine learning perspective on Personalized Medicine: an automized, comprehensive knowledge base with ontology for pattern recognition. *Machine Learning and Knowledge Extraction*, 1(1), 149-156.

[6] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.

[7] Berrar, D. (2019). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*, 1, 542-545.

[8] Kannan, R. and Vasanthi, V. (2019). Machine learning algorithms with ROC curve for predicting and diagnosing the heart disease. *Soft Computing and Medical Bioinformatics* (pp. 63-72). Springer, Singapore.

[9] Schoonjans, F. (2018) ROC curve analysis with MedCalc. *MedCalc*.

[10] Landier, A. and Thesmar, D. (2020). Earnings expectations in the covid crisis (No. w27160). *National Bureau of Economic Research*.

[11] Wu, C. C., Yeh, W. C., Hsu, W. D., Islam, M. M., Nguyen, P. A. A., Poly, T. N. and Li, Y. C. J. (2019). Prediction of fatty liver disease using machine learning algorithms. *Computer methods and programs in biomedicine*, 170, 23-29.

[12] Nhu, V. H., Shirzadi, A., Shahabi, H., Singh, S. K., Al-Ansari, N., Clague, J. J. and Luu, C. (2020). Shallow Landslide Susceptibility Mapping: A Comparison between Logistic Model Tree, Logistic Regression, Naïve Bayes Tree, Artificial Neural Network, and Support Vector Machine Algorithms. *International Journal of Environmental Research and Public Health*, 17(8), 2749.

[13] Souza, F. H. B. D., Couto, B. R. G. M., Braga, G. M., Teixeira, J. A., Santos, R. C., Martins, J. M. C., Sousa, K. S. D., Souza, D. N. and ALVES, G. B. (2019). Medicine Allied to Technology: The Use of Artificial Neural Networks in The Prediction of Surgical Site Infection for General Surgery Service. In: 5th International Conference on Prevention & Infection Control (ICPIC 2019), Genebra. *Antimicrobial Resistance and Infection Control*, 2019. v. 8. p. 34-34.

[14] Mohammed, M., Khan, M. B. and Bashier, E. B. M. (2016). *Machine learning: algorithms and applications*. Crc Press.