


Improvement in the purchase of imported goods through machine learning models for intelligent decision making

Mejoramiento en la adquisición de bienes importados mediante modelos de aprendizaje automático para la toma de decisiones inteligentes

Adolfo González^{1*}  <https://orcid.org/0009-0008-4452-1285>

Mauricio Arriagada^{1*}  <https://orcid.org/0000-0002-2953-663X>

Recibido 06 de abril de 2023, aceptado 06 de junio de 2023

Received: April 06, 2023 Accepted: June 06, 2023

ABSTRACT

Demand planning related to making purchases of SKUs to maintain the SLA given by the company's strategy and thus avoid stock breaks has an important role in the operation of the supply chain and the company's operation. Demand forecasts based on qualitative methods and manual methods based on historical data obtained impact on production planning, consequently, the fulfillment of the products the customer requires. The objective of defining and implementing a purchasing recommender has been raised based on the machine learning model that more effectively adapts to variations in demand for products classified under an ABC model. Scikit-learn libraries are used to implement demand prediction models trained with historical product information. The result is a proposed prediction model with a better confidence level than the company's current prediction model.

Keywords: Machine learning, demand forecasting, demand planning, purchasing recommendation, quantitative models, time series analysis.

RESUMEN

La planificación de la demanda relacionada con la realización de compras de SKU para mantener el SLA dado por la estrategia de la empresa y así evitar quiebres de stock, tiene un papel importante en la operación de la cadena de suministro y la operación de la empresa. Los pronósticos de demanda basados en métodos cualitativos y métodos manuales, basados en datos históricos obtenidos, tienen un impacto en la planificación de la producción, en consecuencia en el cumplimiento de los productos requeridos por el cliente. Se ha planteado el objetivo de definir e implementar un recomendador de compras, basado en el modelo de machine learning que se adecue de forma más efectiva a las variaciones de la demanda para productos que están clasificados bajo un modelo ABC. Se utilizan las librerías Scikit-learn, para implementar modelos de predicción de demanda que son entrenados con la información histórica de los productos. El resultado es un modelo de predicción propuesto con un mejor nivel de confianza que el modelo de predicción actual usado por la empresa.

Palabras clave: Machine learning, pronóstico de la demanda, planificación de la demanda, recomendador de compras, modelos cuantitativos, análisis series de tiempo.

¹ Departamento de Ingeniería en Computación e Informática, Facultad de Ingeniería, Universidad de Tarapacá. Arica, Chile.
E-Mail: adolfo.rgc@gmail.com; marriagada@academicos.uta.cl

* Autor de correspondencia: Adolfo González, Mauricio Arriagada.

INTRODUCTION

The purchase of products to satisfy the demand is a vital part of the company's business that participates in the industry of bolts and screws and is oriented to the competitive strategy of cost leadership. The strategy implies that purchases must be made with a forecast of at least nine months since it is considered that the stock is stored between three to six months with the estimated consideration of three months of transportation. The competitive environment drives the company to buy products effectively to avoid overstock or outages; this implies that the demand forecasting methods to make purchases must reach an optimal level of confidence.

This paper aims to propose a purchase recommender that responds to a prediction model using machine learning techniques with specialized libraries in the use of time series based on predictive analysis techniques. The recommender uses historical data that allows machine learning models to be trained and then to evaluate and establish the confidence level of each proposed model based on the model used. Each model is applied to four groups that are made based on the application of the ABC model of inventory classification.

Five machine learning models are applied to each SKU from each group, and the model that delivers the most correlated error concerning the historical demand for each SKU is selected.

The proposed recommender model exposes an improvement of 20% on average. Compared to the current model, which allows it to be implemented in applications used by purchase analysts.

The present work proposes a platform that allows adding demand prediction models, whether statistical

or neural networks, for demand prediction so that all the models are applied to the platform, and then the result of the model that best responds to each Stock Keeping Unit (SKU) Behavior. As a reference, this project uses the CRISP-DM methodology [1] [35] to apply machine learning models in this project. Therefore, following the framework shown in Figure 1, used in this research, the data extraction, cleaning, and processing phase is performed at the initial stage, followed by machine learning models that finally provide a trained model as a recommendation.

OVERVIEW

The company and its competitive strategy

Chile's Bolts and Screws Industry gravitate in a highly competitive environment or red ocean [2]. Therefore, it must differentiate itself from the competition by creating a blue ocean, to achieve more significant market shares. The company uses the competitive strategy of cost leadership [3] and seeks to differentiate itself from its competition, adding to its strategy, innovations that drive it to change the operation in its value chain demanding to be more flexible oriented to the needs of the client and/or changing conditions of the market and the environment.

The company focuses on the import and sale of bolts, fasteners, and screws in general, orienting itself to the customer through the mix of products and the availability of its products, where the guidelines are followed; "always have available the products sought by the customer" and "always have the products available when customers need them."

The guidelines guide the company's value chain, generating coordination of the areas based on the company's competitive strategy.

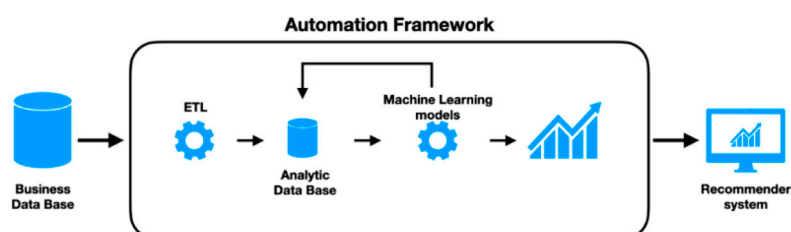


Figure 1. Data extraction by ETL.

The requirement to separate channels to reach groups of clients generates the need to convert the products purchased and imported in large volumes from abroad into packages with smaller volumes through the process called “maquilado,” which consists of packaging of products.

The company calls the products imported from abroad in large volumes “Parent SKUs” or “Parent Codes,” and the manufactured products, which correspond to the parent products grouped in various packaging formats, are called “Child SKUs” or “Codes Son.”

The inventory management model

In a company that must purchase products to satisfy customers’ product needs sales information is the source of information for determining demand, understanding that the information incorporates variations due to uncertainty due to external market conditions and internal operations of the company.

The company that uses an inventory management model based on competitive strategy [4] uses ABC classification models to organize internal operational management. In this context, inventory planning becomes important, given its impact on costs and economic results of the company based on compliance with the SLA, as shown in Figure 2.

The first step identifies or places the company’s strategy in a competitive environment concerning service levels, defined as a percentage of product delivery to the customer, given by equation (1) that the company sets at 98%.

$$N_s = \frac{P_F}{P_{NV}} \times 100 \quad (1)$$

Where:

N_s : Percentage of the Service Level, per sales note.

P_F : The SKU number, effectively delivered to the customer in a sales note.

P_{NV} : Number of SKUs committed to deliver to the Customer in a Sales Note.

The second step establishes a classification of the products according to criteria consistent with the company’s strategy. This classification establishes the relevance of products based on demand, such as where they are named in groups A. B. C and D, where:

Source: Inventory management based on competitive strategy [4].

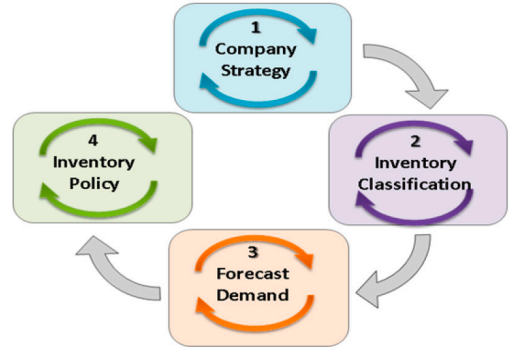


Figure 2. Management reference model.

- Classification A: represents 3% of the Child SKUs and 1% of the Parent SKUs that generate 80% of the turnover. These products require strict checkout, with periodic reviews regarding stock levels with attention to the records’ accuracy.
- Classification B: represents 4% of Child SKUs and 3% of Parent SKUs that generate 10% of turnover. These products require an intermediate checkout, with continuous revisions concerning stock levels with attention to data recording.
- Classification C: represents 6% of Child SKUs and 4% of Parent SKUs that generate 5% of turnover. These products require a low checkout, with continuous reviews regarding stock levels with attention to data recording.
- Classification D: represents 88% of Child SKUs and 92% of Parent SKUs that generate 5% of turnover. These products need to be reviewed to determine whether to remove or promote their sale.

The actions defined to be carried out for each classification are;

- Classification A. Carrying out strict control, with periodic reviews concerning stock levels with attention to the records’ accuracy.
- Classification B. Carrying out an intermediate control, with continuous revisions concerning stock levels with attention to data recording.
- Classification C. Implementation of less rigid control, with continuous reviews concerning stock levels with attention to data recording.
- Classification D. Carrying out an analysis of these products to determine, unsubscribe or promote their sale.

In the third step, a demand forecast is made for each Parent SKU. The forecast is made for each product by analyzing the historical behavior of 12 moving months from the month prior to the one in which it is located.

For products classified as A, B, and C, demand is forecast using the Simple Average Method, Simple Moving Average Method, Moving Weighted Average Method, Simple Exponential Smoothing, and Double Exponential Smoothing [5-6]. Classified products are only applied to Simple Average Method analysis if they are in demand.

After applying the statistical techniques for the analysis of demand, the most appropriate model for the historical behavior of the product is selected through the analysis of the absolute standard deviation, the accumulated forecast error, the mean absolute percentage error, and the signal tracking [7] and the use of the cumulative sum of the errors for the bias together with the mean absolute deviation for the dispersion.

In the analysis process, it is taken into consideration that the company assigns the same weight to bias errors and dispersion errors; consequently, for each Parent SKU, the forecast method that best represents the variability of the variation is used, demand and allow better adjustment of inventories based on the guidelines of the company's strategy [5], [8].

A criterion that the company defines as part of the company's competitive strategy concerning inventories is the existence of three months of inventory per monthly rotation for the Child SKUs and 9 months for the Parent SKUs, that is, the number of monthly turnovers multiplied by three months for the Child SKUs and the amount of monthly turnover multiplied by 9 months for the Parent SKUs [9].

The cycle for calculating the demand for each product is reviewed periodically based on the classification of the products, from classification A to C; therefore, the variability of the demand caused by commercial actions, market conditions, or changes in the company strategy cause changes in stocks that affect the level of service.

The fourth step is oriented to the selection of an inventory policy following the needs of the company's

competitive strategy; the analysis of the demand provides guidelines based on the variability of this and the seasonal nature. Based on this, coupled with the given classification, inventory policies are selected for classifications A, B, and C [10], [11].

Classification A. It is defined to use a periodic review policy since this policy contains a stricter control criterion in accordance with the definition of the product classification [12], [13].

For classification B and C; it is defined to use a continuous review policy since this policy contains less strict control criteria following the definition of the classification [10].

Problematic situation

The purchase of SKUs from foreign suppliers carried out by the company's import area requires a high degree of coordination with the suppliers to avoid falling into stock breaks due to delays in the dispatch and transportation processes of the SKUs acquired, mainly from China and India. Therefore, the increase in time dedicated to analyzing the quantities of Parent SKUs to be purchased, which include demand projections, added to the time it takes for suppliers to respond to quotes and send orders once the conditions have been commercially accepted become more complex as the number of SKUs increases.

There are diverse situations that make the analyst's job difficult. First of all, although the analyst is supported by tools that provide him with the data calculated with the mobile monthly demand and the classification, the associated regression analysis work that must be applied to the Parent SKUs, together with the increase in the demand in a sustained way, does not allow to review in its entirety the SKUs classified as C. Another problem is observed in the Parent SKUs classified as D since not all have zero demand, and they can also be products that are on the rise. These issues can turn into high-turnover products, and if they go bankrupt, even if the SLA is met, the loss of competitiveness is latent.

Lastly, the current market uncertainty, inflation, and the increase in currency exchange, among others, do not allow simple product purchase decisions without considering the company's historical data on purchases.

Given this context, the following questions arise: How can a support system for analysts be implemented to calculate demand, which considers statistical regression analyses automatically? Is it possible that the model delivers a function that estimates the demand and indicates whether the demand is rising or falling for the Parent SKUs of all the product master classifications? Is it possible to have a purchase recommender for Parent SKUs that reduce analysis times for purchase?

Availability of a medium-long-term demand projection. Machine Learning processes historical sales data to transform it into useful information.

It is possible to use empirical information to understand better the behavior of the products purchased in the company, train a model, and thus make better decisions based on data. This understanding is how the future prediction of the demand for different products is an area where Machine Learning contributes to reducing risk when purchasing. This risk reduction provides valuable tools for studying variables that will facilitate decision-making related to inventory management, by predicting an attribute of continuous value using the supervised learning perspective.

THEORETICAL FRAMEWORK

Related Works

The study of the demand prediction models' state of the art concludes, "It is necessary to evaluate and test different models to find the most suitable for the reality under analysis. A resource used in different realities is constructing its demand forecasting model using techniques, aspects, concepts, and characteristics of different methods and models. However, monitoring the adopted model by measuring the forecast error is essential to exercise control over the model and make corrections as quickly as possible." In this sense, apply prediction models and compare them based on their error coefficients to select the most suitable for prediction at a given time [14].

As described in the work of Qian and Rasheed [15], stock market prediction is a challenge where some machine learning models can contribute to increasing such prediction. The researchers Nayyeri, Jabani, and Mosavi [16] applied a similar approach, where ten

different machine learning algorithms were applied. Some other approaches have been included in the literature, such as the work presented by Singh and Srivastava [17], where deep learning improved the accuracy of stock market forecasts.

The article by Diaz [18] addresses the application of a Machine Learning forecast model to establish the demand for the main material used in oil well drilling by applying a model leaving aside other models, so that this work is necessary to evaluate more models that participate in the prediction and thus select the model that best responds in predicting the demand for each SKU in a given time.

Demand estimation through machine learning models or forecasting techniques based on quantitative and qualitative historical information [19]. The research work called "Comparative analysis of traditional and modern models for demand forecasting" [20] performs an analysis of demand forecasting methods through time series analysis, concluding that; "There is no forecasting model that can perfectly determine future demand," exposing the situation in which applying several methods for a future forecast in a given time, it will not be known which is the best until the result is at any given moment. The work exposes that the precision of the prediction will depend on the quantity, the quality, and the particular behavior of the time series data. This information indicates that the ABC grouping helps to obtain better results in the SKUs with higher rotation, therefore, greater quantity and quality of data. An approach to selecting the prediction model applied to anticipate a demand is addressed in the work of Porteiro, L. Hernández-Callejo, and S. Nesmachnow [21], where the model is selected through the error coefficients.

ABC classification

The ABC classification is used in companies of various industries to categorize products based on demand or a relationship between demand and costs, among other possible combinations. That is why this type of classification groups the products from the most significant to the least significant [22].

After classification, forecasting techniques are used to obtain a future value based on the historical information of the classified products. This estimate determines the pattern that allows exposing the

product's behavior in the future under similar conditions, both within and in the environment of the company, understanding that the conditions are a source of variability generation.

Forecast Methods

Forecasting methods are classified as qualitative and quantitative. The qualitative ones, in general, depend on the expert judgment of actors who have been studying the variables of the company's products for a long time and support future predictions. In this context, techniques such as the so-called Delphi are used, which consists of bringing together several experts and establishing a consensus forecast. In this context, techniques such as the so-called Delphi are used to bring together several experts and establish a consensus forecast. Curve fitting is also used, where the expert prepares a representative curve of the product's behavior and compares new products with characteristics similar to the one represented with this curve. These methods are currently used in the company under study to support our forms of estimation that have been incorporated.

Quantitative methods aim to find patterns in the products' historical data and establish future projections of behavior. These methods lead to the application of numerical analysis, such as time series methods that use historical data to estimate future results; where there are several methods for such an effect, such as the naive method, which assumes that the variable to be predicted has the same value as its measurement, previous; moving average, where the prediction is the result of the average of n previous observations; decomposition methods, which break down the time series into four components to describe each of them separately, thus managing to describe and predict the time series as a whole; exponential smoothing, seeks to smooth the series to have a better idea of the behavior of the trend, to obtain a more reliable forecast [23].

Predictive statistical analysis in multi-product environments guides the company to find ways to apply supervised models to determine the demand in specific periods of time. In this sense, Méndez and López [22] expose the application of various predictive models for estimating demand, where by analysis criteria, Exponential Smoothing (SE), Additive Exponential Smoothing (EA), Multiplicative

Exponential Smoothing (EM), Linear Additive Smoothing (LA), Linear Multiplicative Smoothing (LM), On the other hand, low-involvement techniques such as Linear Models (ML).

Regression Analysis

Regression analyzes are used to discover trends in data sets. An equation determines the prediction based on the input data entered into the equation. In this sense, a dependent variable must be defined that will be explained given one or more independent variables.

Nonlinear regression aims to find a function that describes the nonlinear relationship between a continuous response variable and one or more variables, predicting new observations. The aim is to find a polynomial function of a higher degree that minimizes the error. This method allows an almost perfect fit, obtaining an error equal to zero. With this comes the problem of considering the noise (outliers) as part of the valid data, which means that the function obtained does not correctly represent the behavior since it considers the noise data as part of the phenomenon. In the context of Machine Learning models, overfitting should be considered, called "overfitting," it is a problem that causes a model to be unable to generalize [24].

Machine Learning Models

Machine Learning Models use historical information to improve the performance of predictions through training processes. This condition makes the application of Machine Learning models allow the arrangement of classes to which a new entry belongs. The regressions, where it seeks to predict a value; ranking, where data is ordered based on some criteria, clustering, used for the segmentation of customers and products to guide commercial actions, reduction of dimensionality, where it is sought to transform a large number of dimensions into fewer dimensions, maintaining the properties of the original representation [25].

Machine learning is the discipline that guides the development and application of models, algorithms to learn from data. This learning refers to learning from data; this indicates that an algorithm is not developed to solve a specific problem; instead, a generic learning algorithm provides solutions and allows learning how to solve the problem [34].

Generalized linear models

Regression analysis belongs to the area of supervised Machine Learning to establish a model for the relationship between a specific number of characteristics and a target variable, seeking to obtain a quantitative answer.

Regarding a data set, the regression algorithm will establish a model that adjusts the dependency relationship between the independent variable and the value of the dependent variable; it is obtained by establishing an arbitrary line and the distance from the line to the given points. This distance, the vertical lines, are the “residuals” or the “prediction errors.” The regression algorithm recalculates and moves the line with each interaction until it reaches the line that fits the data points with the slightest error.

Simple linear Regression is a relationship between a dependent and an independent variable. Equation (2) exposes the relationship, called the equation of the line, where m is the slope and n is the intercept

$$Y = mX + n \quad (2)$$

When obtaining the equation of the line from a sample, the coefficients are statistical and not parametric. The equation is obtained by incorporating the error ϵ into the equation of the line, which is known as random error or noise and is defined as the difference between the real value and the approximate value $\epsilon = Y_R - T_A$, as shown in equation (3). This method allows obtaining a prediction of Y knowing the variable X , seeking to obtain an ϵ close to 0. Several methods can be applied to find the relationship between the variables x and y , each trying to mitigate the problems the other model does not address [36].

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \epsilon = \beta_0 + \sum_{j=1}^i \beta_j X_j + \epsilon \quad (3)$$

Least square approach is a numerical analysis method where, given a set of data, it is sought to determine the function that best represents the data employing a fit line. In its simplest form, this method seeks to minimize the sum of squares of the difference in ordinates between points generated by the function and the input data. This method is used to analyze data obtained from a study to express

its behavior in a linear way, minimizing the errors of the data set obtained, as shown in equation (4).

$$\sum_{i=1} (Y_R - Y_A)^2 \quad (4)$$

Bayesian Linear Regression is a regression model defined in probabilistic terms. The estimate is not made as a single value but is drawn from a probability distribution. The model for Bayesian linear regression with the response sampled from a normal distribution is given by equation (5).

$$y \sim N(\beta^T X \sigma^2 I) \quad (5)$$

The mean for the linear regression is the transpose of the weight matrix multiplied by the predictor matrix. The variance is the square of the standard deviation σ multiplied by the Identity matrix because this is a multidimensional model formulation. This regression model aims to determine the posterior distribution of the model parameters. The response is generated from a probability distribution, and the model parameters are assumed to come from a distribution. The posterior probability of the model parameters depends on the training inputs and outputs, as shown in equation (6), where $P(\beta | y, X)$ is the posterior probability distribution of the model parameters given the inputs and outputs [26].

$$P(\beta | y, X) = \frac{P(y | \beta, X) * P(\beta | X)}{P(y | X)} \quad (6)$$

Linear Lasso Model (Least Absolute Shrinkage and Selection Operator) combines a regression model with a procedure of contraction of some parameters towards zero and selection of variables, imposing a restriction or a penalty on the regression coefficients [27]. Robert Tibshirani introduced lasso in 1996 based on Leo Breiman's work on the Garrote [28] Non-Negative to improve the accuracy of the predictions of statistical regression models.

As shown in equation (7), it mathematically consists of a linear model with an added regularization term. Where the exact relationship between t and λ depends on the data.

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\} \quad (7)$$

This linear model estimates scattered coefficients. It is used in contexts where the tendency to prefer solutions with fewer non-zero coefficients effectively reduces the number of characteristics on which the given solution depends [29].

Polynomial regression is a form of linear regression where the relationship between the independent and dependent variables is modeled as a polynomial of order n . It is used to describe non-linear phenomena, where the linear model is extended by adding additional predictors obtained by raising each of the original predictors to power, as shown in equation (8).

$$y = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0 \quad (8)$$

Determination coefficient

The R-squared coefficient of determination (R^2) predicts outcomes or tests a hypothesis. It determines the quality of the model and the proportion of variation in the results. In simple linear regression, R^2 the square of the Pearson correlation coefficient indicates the percentage of the variation in one variable due to the variation in the other. If there are several results for a single variable, R^2 corresponds to the square of the coefficient of multiple determination. In both cases, the R^2 acquires values between 0 and 1 [30]. Table 1 shows a way to classify the results of R^2 , to have an interpretation of the results obtained:

Table 1. Interpretation criteria R2.

R^2 Value	Interpretation
0.00	There is no linear correlation between the variables
0.10	Weak positive correlation
0.50	Mean positive correlation
0.75	Considerable positive correlation
0.90	Very strong positive correlation
1.00	Perfect positive correlation

In this study, when applying a linear regression model to a set of values, the application delivers an R^2 , where the model that delivers a higher value will be the one with the highest correlation.

Automatic adaptation to change

Machine Learning trains a model from a data set. The selected model is evaluated for errors; if not found or within acceptability parameters, the model can be used. Otherwise, the errors are analyzed, and the algorithm is retrained [31]. Thus, a model of automatic adaptation to applied change is presented in this proposal, where the operational relationships between each stage are found, as can be seen in Figure 3.

Steps one (1) and two (2) are the business applications that are related to the business database where the transactions are recorded. Then, through the ETL (Extract, Transform, and Load), the data is processed (step 3) and transferred to the analysis database (step 4). Consecutively, in the stage of applying Machine Learning (step 5), the data is taken from the database, processed, and the results are delivered, which are entered again into the analysis database (step 4). Subsequently, it returns to the ETL stage (step 3), where the information is taken from the analysis base and deposited in the business base, remaining available for use in step 1. Special mention should be given to steps 3 and 5, which an Automator coordinates to synchronize the executions temporarily.

IMPLEMENTATION

Data identification

Given the monthly demand volumes, the company uses a four-stage inventory management model [4] to organize the company's sales resources and meet market needs.

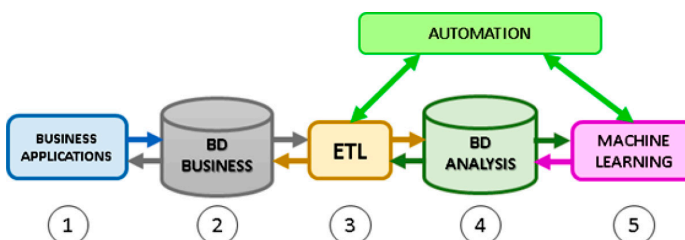


Figure 3. Automatic adaptation to change model.

As shown in Table 2, there are 20,003 SKUs in the company's product sample (10,525 + 9,478). In addition, the sum of the monthly demands of each parent SKU added to the demands of the child SKUs adds up to an average of 65,782,846 monthly units, as shown in Table 3.

Generally, there is a significant rotation of A-class products, with 80% of the total units, but it represents 1% of Parent SKUs of the same class. In comparison, a high percentage of parent SKUs are classified D (92%) but with a low turnover (5%).

There are 9,478 existing Parent SKUs in Table 3, which is the current number of Parent SKUs selected. However, this number varies from month to month depending on the variations and needs of both the business and the market; therefore, a set of Parent SKUs with data history was selected, eliminating the new Parent SKUs and leaving a stratification by classification as shown in Table 4, leaving 267 SKUs out of the analysis.

The company, based on its business definitions and considering that it makes purchases with a scope of 6, 9, and 12 months to avoid bankruptcy, defines to consider only products that are sold six months in a period of one year, leaving out the SKU

Parents that do not meet this condition, as shown in Table 5, where the largest number of products that will not be selected are classified D, mostly for sales projections between 11 and 12 months.

As a result of eliminating the Parent SKUs that do not meet the sales forecast condition in short periods (6, 9, and 12 months), the number of Parent SKUs is reduced by 45% of the data set. The distribution of Parent SKUs by classification is shown in Table 6.

Data subsets

The data set extracted from the company's database is divided into subsets, which comprise the training, validation, and testing data set. The data from the training subset is used to train the model. The results obtained are contrasted with the validation set to evaluate the results. Finally, the model is used with data from the test subset to generalize with the model on data yet to be used.

The subset selection is usually made with 80% for the training subset and 20% for the validation subset. The validation subset is selected with diverse samples and a sufficient number of samples to check the results once the model has been trained.

Table 2. Classification of products by rotation. (Author's elaboration).

Classification	N° Child SKU	% Child SKU	N° Parent SKU	% Parent SKU	N° Average monthly turnover	% Rotation
A	338	3%	142	1%	52,613,688	80%
B	391	4%	255	3%	6,594,101	10%
C	582	6%	381	4%	3,288,966	5%
D	9,214	88%	8,696	92%	3,286,091	5%
Total	10,525	100%	9,478	100%	65,782,846	100%

Table 3. Classification of Parent SKU products by rotation. (Author's elaboration).

Classification	N° Parent SKU	% Parent SKU	N° Average monthly turnover	% Rotation
A	142	1%	52,613,688	80%
B	255	3%	6,594,101	10%
C	381	4%	3,288,966	5%
D	8,696	92%	3,286,091	5%
Total	9,478	100%	65,782,846	100%

Table 4. Stratification of Parent SKUs for analysis. (Author's elaboration).

A	B	C	D	Total
142	255	385	8,429	9,211

Table 5. Number of Parent SKUs with more than six months without sale. (Author's elaboration).

Months without sales	A	B	C	D	Total
Between 11 and 12	0	0	0	3,342	3,342
Between 10 and 11	0	0	0	214	214
Between 9 and 10	0	0	0	194	194
Between 8 and 9	0	0	1	217	218
Between 7 and 8	0	0	0	207	207
Between 6 and 7	0	0	0	0	0
Total	0	0	1	4,174	4,175

Table 6. Quantity of Parent SKUs used by classification. (Author's elaboration)

A	B	C	D	Total
142	255	384	4,255	5,036

Data collection and organization

The data extraction from the business database (Base of Operations) is executed through the ETL process. The implementation is done with the Pentaho tool [32]. Finally, the data is deposited in an analysis database, as shown in Figure 4.

The data obtained from the operational system corresponds to the database tables that contain the following features:

- **Product Master:** Contains the information on the active codes of master products and their relationship with the child codes and

the packaging units to be able to carry out a normalization of demand units.

- **SKU classification:** The classification model that is applied to the child SKUs and the parent SKUs is registered in classification tables that are modified monthly. From these tables, the classification table of parent SKUs is obtained.
- **Movements:** This table registers all movements of units in and out for each SKU so that the monthly demand is the outputs minus the inputs.
- **Historical monthly demand:** This table is loaded based on the historical consolidation of the data obtained from the Product Master, SKU Classification, and Movements tables, it records the monthly demands for each parent SKU from 2017 to date.

In the analysis base, a table is created to organize the results that can be obtained from applying the models. This table collects the results of the application of the model to analyze and use the results of the training in the analysis period and to maintain the data of the applications of the models during the execution of the process. In this table, the column M1_P stores the slope information resulting from the application of model 1, while the field M1_B stores the intercept and M1_err stores the score. This is repeated iteratively depending on the number of models used at the same time; that is, if 4 models are applied, the columns go from M1 to M4.

In addition, an attribute called M_SELECT indicates which model (examples M1 to M4) best represents the demand based on the score obtained. In this sense, the SLOPE, and INTERCEPT attributes store the value of the selected model. Finally, the TREND attribute indicates, based on the slope sign, whether the forecast trend is upward, downward, or linear if it is zero.

Data selection

The selection of the data considers the performance of cross-validation as a statistical technique to test the performance of a machine learning model.

The monthly demand historical table data is segmented into groups of 12 moving months, which are used to train the model and validate. The segmentation is selected based on the definitions of the business, which must acquire the parent SKUs with a projection of 6, 9, and 12 months.

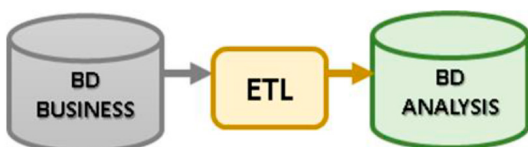


Figure 4. Data extraction by ETL.

The groups selected for training and testing are separated by a split at the time of training the model in 80% training and 20% testing. The groups called for validation are the data used to evaluate the prediction of the trained model, as shown in Table 7, Table 8, and Table 9. Here it has been considered that the percentage of data selected for validation is much higher than the data group for testing because it was necessary to ensure that the model entered values very close to those projected by the validation set. Thus, group 1 of Table 7 is used for the base configuration of the model, to then begin training and data validation with those of the validation group. The data segmentation structure is oriented to measure the variations of the prediction.

Machine Learning Models

Given the nature of the problem to be solved, regression models were selected to be able to make product demand predictions. These models are part of the supervised machine learning algorithms group since it will be necessary to train the models with historical data from the business. The libraries used to implement the models are obtained from Scikit-learn [33]. A brief description of the models used is shown below:

- Ordinary Least Squares (M1). This model is linearly fitted with coefficients $w = (w_1, \dots, w_p)$ to minimize the residual sum of squares between the targets observed in the data set and those predicted by the linear approximation.

Table 7. First data set. (Author's elaboration).

Item	Group 1 training set (months)	Group 2 test set (months)	Group 3 validation set (months)
1	Sep-09	Sep-10	Sep-11
2	Aug-09	Aug-10	Aug-11
3	Jul-09	Jul-10	Jul-11
4	Jun-09	Jun-10	Jun-11
5	May-09	May-10	May-11
6	Apr-09	Apr-10	Apr-11
7	Mar-09	Mar-10	Mar-11
8	Feb-09	Feb-10	Feb-11
9	Jan-09	Jan-10	Jan-11
10	Dec-08	Dec-09	Dec-10
11	Nov-08	Nov-09	Nov-10
12	Oct-08	Oct-09	Oct-10

Table 8. Second and third data sets. (Author's elaboration).

Item	Group 4 test & training set (months)	Group 5 validation set (months)	Group 6 test & training set (months)	Group 7 validation set (months)
1	Sep-12	Sep-13	Sep-14	Sep-15
2	Aug-12	Aug-13	Aug-14	Aug-15
3	Jul-12	Jul-13	Jul-14	Jul-15
4	Jun-12	Jun-13	Jun-14	Jun-15
5	May-12	May-13	May-14	May-15
6	Apr-12	Apr-13	Apr-14	Apr-15
7	Mar-12	Mar-13	Mar-14	Mar-15
8	Feb-12	Feb-13	Feb-14	Feb-15
9	Jan-12	Jan-13	Jan-14	Jan-15
10	Dec-11	Dec-12	Dec-13	Dec-14
11	Nov-11	Nov-12	Nov-13	Nov-14
12	Oct-11	Oct-12	Oct-13	Oct-14

Table 9. Fourth and fifth data sets. (Author's elaboration).

Item	Group 8 training & test set (months)	Group 9 validation set (months)	Group 10 training & test set (months)	Group 11 validation set (months)
1	Sep-16	Sep-17	Sep-18	Sep-19
2	Aug-16	Aug-17	Aug-18	Aug-19
3	Jul-16	Jul-17	Jul-18	Jul-19
4	Jun-16	Jun-17	Jun-18	Jun-19
5	May-16	May-17	May-18	May-19
6	Apr-16	Apr-17	Apr-18	Apr-19
7	Mar-16	Mar-17	Mar-18	Mar-19
8	Feb-16	Feb-17	Feb-18	Feb-19
9	Jan-16	Jan-17	Jan-18	Jan-19
10	Dec-15	Dec-16	Dec-17	Dec-18
11	Nov-15	Nov-16	Nov-17	Nov-18
12	Oct-15	Oct-16	Oct-17	Oct-18

- Bayesian Ridge Regression (M2). This model includes regularization parameters in the estimation procedure. The regularization parameter is not set strictly but is adjusted to the available data.
- LARS Lasso (M3) has been implemented based on the LARS algorithm. Unlike the implementation based on coordinate descent, this model provides a more exact solution, which is piecewise linear based on the norm of its coefficients.
- Polynomial Regression (M4), this model focuses on maintaining the generally fast performance of linear methods while allowing them to fit a large volume of data.

The following sequence of steps is applied to train the models (M1, M2, M3, and M4) for each proposed machine-learning technique:

1. The first model selected (M1) has the normalization parameters applied and is parameterized with 10% of the data to create the test data set.
2. The algorithm is executed, and the score is recorded for each SKU evaluated. Then the process is repeated for the same algorithm from the first step; however, the test data set is parameterized at 20% this time. Finally, the process is repeated, with the parameterization of the test data set at 30%.
3. The recorded score data of the three executions are evaluated, and the parameterization with

the best score is selected, registering the slope (M1_P), the intercept (M1_B), and the score (M1_err) for the model executed by each SKU of the model data set.

With the information obtained, the data for each SKU of the columns M1_err, M2_err, M3_err, and M4_err are evaluated, which represents the score of each applied model and is used to identify the model that best fits the prediction (with the best score) for each Evaluated SKU. For example, if the best score corresponds to the M3 model, the prediction for the SKU is made using the third model.

Data Processing

Through Python, the outputs of the 4 applied models are processed to consolidate the best SKUs. The script must pick the largest M_err of each SKU for the four models.

Subsequently, a column called Trend with the values <Down Linear, Rising> must be added to the consolidated file and stored in a company database. Finally, the developed application connects to the analysis database, extracting data and processing it to deliver the result of the application of the models, as shown in Figure 5.

The result of the processing is deposited in the analysis database, obtaining information in the data set for each SKU as shown in Table 10, Table 11, Table 12, and Table 13. The tables are highlighted

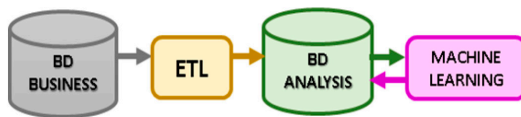


Figure 5. Processing in Python application.

in color, indicating which model is selected; the model with the highest correlation is selected, being used for the prediction highlighting in the column ending in “_P” the slope, for example, for the SKU 741020999900G, the M1_err is the prediction that is closest to a positive correlation; therefore, for the prediction made for the SKU 741020999900G with the given data set, for a given time the model M1_P is selected, and the slope is -1,589.17. This process is applied for each SKU used by the company’s product master every time the model is applied, and the data set changes as the month progresses.

Table 14 shows the selected model for each SKU that the models to be evaluated apply to.

EVALUATION OF RESULTS

Current model

The current demand forecasting model for Parent SKUs is grouped using the ABC classification. This classification is used to keep under strict control the products that represent 20% of the SKUs of the total but that represent 80% of the products that are most in demand. From this perspective, a period of 12 months is evaluated, obtaining the results shown in Table 15.

It is observed that the predictions at 12 months of group A have a confidence level of 60%; that is, the model predicts, on average, with 60% effectiveness, the future demand of the Parent SKUs classified in group A. For the other classifications, lower percentages of effectiveness are observed. The estimated general percentage of the currently applied model is 43%, lower than that obtained by the proposed model.

The details of the model’s behavior based on the classification and ranges of successes in the

Table 10. Subset that includes SKU applied to M1. (Author’s elaboration).

Parent SKU	Code	Class	M1_P	M1_B	M1_err
741020	7410209999900G	A	-1,589.17	202,723.61	0.0129
741018	7410189999900G	A	-2,153.33	73,333.33	0.2059
740220	7402209999900G	A	-2,736.67	186,483.33	0.0373
740218	7402189999900G	A	452.5	116,195.83	0.0114
740216	7402169999900G	A	-1,585.83	254,990.28	0.0169
740034	7400349999900G	A	-8,963.33	2,947,383.33	0.0136
601895	6018959999900G	B	-15,430	1,297,372.22	0.1325
592540	5925409999900G	B	-3,742.92	288,070.14	0.0303
592532	5925329999900G	B	204.76	10,728.57	0.0148
592524	5925249999900G	B	-1345	34,347.22	-0.4677
750340	7503409999900G	C	2,251.19	2,232.14	-1.5205
750315	7503159999900G	C	-915	27,463.89	0.0703
750252	7502529999900G	C	886.79	7,892.46	0.1256
750135	7501359999900G	C	1,175.83	13,898.61	0.0858
999476	9994769999900G	D	801.63	1,690.29	0.5277
999473	9994739999900G	D	-303.83	11,932.17	0.2318
999472	9994729999900G	D	14.29	-39.29	-4.4082
999471	9994719999900G	D	20	-55	0.4444
999470	9994709999900G	D	20.95	-56.79	0.3974
999469	9994699999900G	D	202.14	-549.64	0.8273

Table 11. Subset that includes SKU applied to M2. (Author's elaboration).

Parent SKU	Code	Class	M2_P	M2_B	M2_err
741020	7410209999900G	A	-9.235E-06	194,777.78	-0.0797
741018	7410189999900G	A	-0.0006809	62,566.67	-2.0191
740220	7402209999900G	A	-0.0002265	172,800	-0.0423
740218	7402189999900G	A	3.7526E-05	118,458.33	-0.0021
740216	7402169999900G	A	-2.5812E-05	247,061.11	-0.0002
740034	7400349999900G	A	-2.2626E-05	2,902,566.67	-0.0028
601895	6018959999900G	B	-7.3503E-05	1,220,222.22	-0.2719
592540	5925409999900G	B	-0.00008857	269,355.56	-0.0164
592532	5925329999900G	B	0.00014319	11,650	-0.0983
592524	5925249999900G	B	-0.00064136	27,622.23	-0.5439
750340	7503409999900G	C	0.00178984	12,362.49	-0.2945
750315	7503159999900G	C	-0.00278622	22,888.9	-1.1417
750252	7502529999900G	C	526.472251	9,513.87	0.0714
750135	7501359999900G	C	1,019.31189	14,681.22	0.0280
999476	9994769999900G	D	757.489396	1,888.92	0.5532
999473	9994739999900G	D	-0.00067588	10,413	-0.9773
999472	9994729999900G	D	12.2312606	-30.04	0.8967
999471	9994719999900G	D	18.2312877	-47.04	0.8402
999470	9994709999900G	D	19.310828	-49.4	0.6914
999469	9994699999900G	D	186.103665	-477.47	0.8737

Table 12. Subset that includes SKU applied to M3. (Author's elaboration).

Parent SKU	Code	Class	M3_P	M3_B	M3_err
741020	7410209999900G	A	-1,589.05	202,723.03	0.0128
741018	7410189999900G	A	-2,153.22	73,332.75	0.2060
740220	7402209999900G	A	-2,736.55	186,482.75	0.0374
740218	7402189999900G	A	452.38	116,196.41	0.0113
740216	7402169999900G	A	-1,585.72	254,989.7	0.0168
740034	7400349999900G	A	-8,963.22	2,947,382.75	0.0137
601895	6018959999900G	B	-15,429.88	1,297,371.64	0.1326
592540	5925409999900G	B	-3,742.8	288,069.56	0.0304
592532	5925329999900G	B	204.64	10,729.13	0.0147
592524	5925249999900G	B	-1,344.88	34,346.64	-0.4676
750340	7503409999900G	C	2,251.07	2,232.7	-1.5202
750315	7503159999900G	C	-914.88	27,463.31	0.0702
750252	7502529999900G	C	886.66	7,893.02	0.1258
750135	7501359999900G	C	1,175.72	13,899.19	0.0857
999476	9994769999900G	D	801.51	1,690.84	0.5278
999473	9994739999900G	D	-303.72	11,931.59	0.2316
999472	9994729999900G	D	14.16	-38.73	-3.8487
999471	9994719999900G	D	19.88	-54.44	0.4949
999470	9994709999900G	D	20.83	-56.23	0.4313
999469	9994699999900G	D	202.02	-549.09	0.8314

Table 13. Subset that incudes SKU applied to M4. (Author's elaboration)

Parent SKU	Code	Class	M4_P	M4_B	M4_err
741020	7410209999900G	A	2,833.67	-212,180.56	-8.2235
741018	7410189999900G	A	-136.11	73,197.62	-25.2816
740220	7402209999900G	A	-898.32	284,260.32	-1.9523
740218	7402189999900G	A	-1,198.32	160,326.39	-26.0303
740216	7402169999900G	A	-2,518.64	314,755.56	-100.0000
740034	7400349999900G	A	-3,173.23	2,848,614.29	-8.2950
601895	6018959999900G	B	548.99	948,190.48	-12.5805
592540	5925409999900G	B	591.79	294,094.64	-7.6742
592532	5925329999900G	B	110.1	-400	0.5136
592524	5925249999900G	B	187.96	4,368.25	0.7887
750340	7503409999900G	C	82.58	-13,150	0.2064
750315	7503159999900G	C	-60.52	36,444.44	0.2699
750252	7502529999900G	C	-137.83	7,518.07	-17.7305
750135	7501359999900G	C	61.91	13,275.79	-0.7874
999476	9994769999900G	D	0	0	-100.0000
999473	9994739999900G	D	0	0	-100.0000
999472	9994729999900G	D	0	0	-100.0000
999471	9994719999900G	D	0	0	-100.0000
999470	9994709999900G	D	0	0	-100.0000
999469	9994699999900G	D	0	0	-100.0000

Table 14. SKU Data Extract, model selection. (Author's elaboration).

Parent SKU	Code	Class	M_SELEC	Trend
741020	7410209999900G	A	RL	BAJA
741018	7410189999900G	A	RLS	BAJA
740220	7402209999900G	A	RLS	BAJA
740218	7402189999900G	A	RL	ALZA
740216	7402169999900G	A	RL	BAJA
740034	7400349999900G	A	RLS	BAJA
601895	6018959999900G	B	RLS	BAJA
592540	5925409999900G	B	RLS	BAJA
592532	5925329999900G	B	RP3	ALZA
592524	5925249999900G	B	RP3	ALZA
750340	7503409999900G	C	RP3	ALZA
750315	7503159999900G	C	RP3	ALZA
750252	7502529999900G	C	RLS	ALZA
750135	7501359999900G	C	RL	ALZA
999476	9994769999900G	D	RB	ALZA
999473	9994739999900G	D	RL	BAJA
999472	9994729999900G	D	RB	ALZA
999471	9994719999900G	D	RB	ALZA
999470	9994709999900G	D	RB	ALZA
999469	9994699999900G	D	RB	ALZA

Table 15. Current model evaluation. (Author's elaboration).

Classification	3 months	6 months	9 months	12 months	% Predictions
A	60%	59%	60%	60%	60%
B	47%	47%	47%	46%	47%
C	47%	47%	47%	46%	47%
D	20%	19%	19%	19%	19%
	43%				

predictions are shown in Table 16, Table 17, and Table 18. These tables report the quantity of SKUs for importing merchandise without ABC classification grouping, where the number of SKUs are segmented into groups of predictions from 1 to 12 months.

Table 16 shows that the 3-month forecast groups several 1,107 SKUs in a confidence range of between 60% and 100%, regardless of the ABC classification.

Figure 6, shows the behavior of the predictions of the current model, which shows predictions with

an accuracy level between 0 and 40% affectivity in the prediction.

Proposed model

The proposed demand prediction model for Parent SKUs is evaluated over a 12-month in the same way that the current model is evaluated. The results obtained are shown in Table 19.

The predictions at 12 months of group A are 80% reliable; the proposed model predicts future demand with 80% effectiveness. For the other classifications,

Table 16. Behavior between 1 to 4 months of the prediction for classification A, B, C, and D - Current Model. (Author's elaboration).

Intervals	Prediction 1 months	Prediction 2 months	Prediction 3 months	Prediction 4 months
0.6 and 1	1,443	1,217	1,107	1,301
0.4 and 0.6	1,011	957	825	994
0 and 0.4	2,582	2,862	3,104	2,741
Total SKU	5,036	5,036	5,036	5,036

Table 17. Behavior between 5 to 8 months of the prediction for classification A, B, C, and D - Current Model. (Author's elaboration).

Intervals	Prediction 5 months	Prediction 6 months	Prediction 7 months	Prediction 8 months
0.6 and 1	1,137	1,196	1,261	1,195
0.4 and 0.6	885	921	967	905
0 and 0.4	3,014	2,919	2,808	2,936
Total SKU	5,036	5,036	5,036	5,036

Table 18. Behavior between 9 to 12 months of the prediction for classification A, B, C, and D - Current Model. (Author's elaboration)

Intervals	Prediction 9 months	Prediction 10 months	Prediction 11 months	Prediction 12 months
0.6 and 1	1,179	1,158	1,210	1,099
0.4 and 0.6	936	926	924	847
0 and 0.4	2,921	2,952	2,902	3,090
Total SKU	5,036	5,036	5,036	5,036

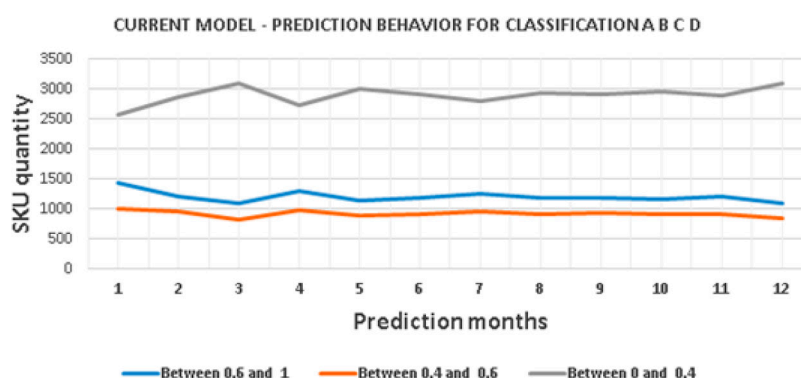


Figure 6. Behavior of the prediction for classification A, B, C, and D - Current Model.

Table 19. Proposed model evaluation. (Author's elaboration)

Classification	3 months	6 months	9 months	12 months	% Prediction
A	80%	80%	80%	79%	80%
B	70%	70%	70%	69%	70%
C	64%	65%	65%	65%	65%
D	25%	25%	25%	25%	25%
	60%				

Table 20. Behavior between 1 to 4 months of the prediction for classification A, B, C, and D - Proposed Model. (Author's elaboration).

Ranges	Prediction 1 months	Prediction 2 months	Prediction 3 months	Prediction 4 months
0.6 y 1	2,394	2,394	2,079	2,416
0.4 y 0.6	1,054	976	1,025	989
0 y 0.4	1,588	1,666	1,932	1,631
Total SKU	5,036	5,036	5,036	5,036

it is observed that the proposed model obtains 70% confidence for classification B, 65% confidence for classification C and 25% confidence for classification D. The general estimated result of the proposed model is 60%.

Table 20, Table 21, and Table 22 show the details of the behavior of the proposed model based on the classification ranges of success in the predictions.

Figure 6 shows the behavior of the predictions where the tendency to improve the predictions with an interval between 60% and 100% confidence in the prediction is observed.

The proposed model recommends the future demand trend for the Parent SKU (Table 14). This trend is expressed based on three criteria. The first is called **Rising**, which indicates that the trend is for an increase in demand relative to the demand when making the prediction.

The second trend is called **Linear**, which indicates that the trend does not vary concerning the demand when making the prediction.

The third trend is called **Down**, which indicates that the trend is downward concerning the demand when making the prediction.

Table 21. Behavior between 5 to 8 months of the prediction for classification A, B, C, and D - Proposed Model. (Author's elaboration).

Ranges	Prediction 5 months	Prediction 6 months	Prediction 7 months	Prediction 8 months
0.6 y 1	2,173	2,282	2,311	2,261
0.4 y 0.6	1,011	1,002	1,011	1,026
0 y 0.4	1,852	1,752	1,714	1,749
Total SKU	5,036	5,036	5,036	5,036

Table 22. Behavior between 9 to 12 months of the prediction for classification A, B, C, and D - Proposed Model. (Author's elaboration).

Ranges	Prediction 9 months	Prediction 10 months	Prediction 11 months	Prediction 12 months
0.6 y 1	2,231	2,246	2,273	2,157
0.4 y 0.6	1,063	964	985	1,044
0 y 0.4	1,742	1,826	1,778	1,835
Total SKU	5,036	5,036	5,036	5,036

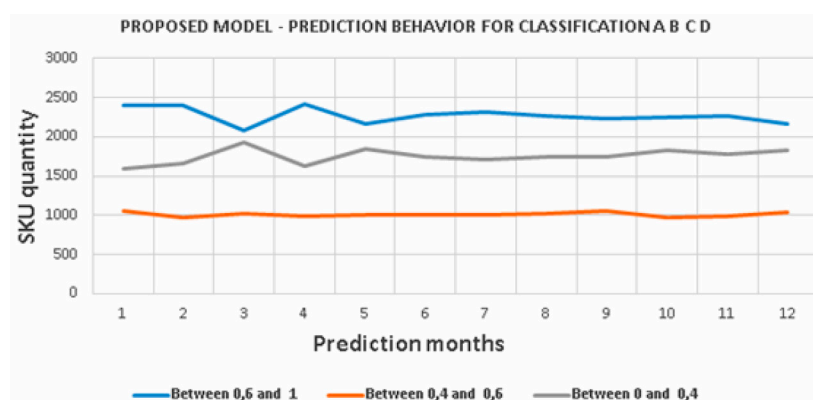


Figure 7. Behavior of the prediction for classification A, B, C, and D - Proposed Model.

Table 23, Table 24, and Table 25 show the result of the analysis of the trend predictions made at 6, 9, and 12 months, respectively, where classification groups also separate the analyses.

Comparison Current Model and Proposed Model

The reliability percentages of the proposed model are contrasted with the current model, where the current model exposes classification A with a confidence level of 60%, classification B with a confidence level of 47%, classification C with a confidence level of 47%, and classification D with a confidence level

Table 23. 6 Months trend prediction evaluation. (Author's elaboration).

Ranges	Classification SKU			
	A	B	C	D
Rising	26	48	64	737
Linear	0	0	0	3
Down	57	80	133	1,345
Total hits	83	143	199	1,371
Total SKU	142	255	384	4,255
	58%	56%	52%	32%

Table 24. 9 months trend prediction evaluation.
(Author's elaboration).

Ranges	Classification SKU			
	A	B	C	D
Rising	16	43	55	713
Linear	0	0	0	2
Down	67	83	138	1,328
Total hits	83	126	193	2,043
Total SKU	142	255	384	4,255
	58%	56%	52%	32%

Table 25. 12 months trend prediction evaluation.
(Author's elaboration).

Ranges	Classification SKU			
	A	B	C	D
Rising	15	31	47	658
Linear	0	0	0	6
Down	69	86	151	1,415
Total hits	84	117	198	2,079
Total SKU	142	255	384	4,255
	59%	46%	52%	49%

of 19%. These results validate the tendency of the proposed model to improve the prediction for each classification compared to the current model.

The results presented after applying the proposed model show a marked improvement. Now, from the perspective of an overall evaluation, the current model provides a confidence level of 43%, as shown in Table 15. The proposed model provides a confidence level of 60%, as shown in Table 19, exposing the tendency of the proposed model to improve the prediction.

Figure 7, shows the results when applying the proposed model to the Parent SKUs, where the confidence level is between 60% and 100% and groups the most significant number of Parent SKUs, unlike the current model shown in Figure 6, where most of the Parent SKUs are in the confidence intervals from 0% to 40%, this illustrates a tendency to improve the predictions made with the proposed model.

The monthly review of the results after the application of the current model and the proposed model for

the third month, as shown in Figure 6 and Figure 7, shows a decrease in the level of confidence for the predictions between the intervals of 60% to 100% and intervals from 0% to 40%, showing a low level of confidence in the application of the current model and the proposed model. The review of the Parent SKUs involved in this month exposes a problem related to out-of-stock in the indicated month, which impacts sales by not having products available; therefore, the prediction is affected.

From a general point of view, the increase in the level of confidence from 43% of the current model (Table 15) to 60% of the proposed model (Table 19) shows a greater precision in the demand predictions, marked in the products of classification A and B, where, due to the inventory management model followed by the company, they carry out significant control of stocks.

Applying machine learning models to each Parent SKU offers a prediction for the next twelve months. When reviewing the results of the proposed model, applied to each Parent SKU, a reliability level of 80% for classification A, 70% for classification B, 65% for Classification C, and 25% for classification D is illustrated.

The proposed model delivers three types of trends along with the predictions, which are; upward, downward, or unchanged trend of the Parent SKU, with an average confidence of 58% for rank A, 56% for rank B, 52% for rank C and 32% for rank D, for the projections from 6 to 9 months as shown in tables 22 and 23, information that the current model does not provide.

Recommender Proposal

The recommender system exposes the information of the predictions that the end user needs to support the purchase decision. The operation of the proposed recommender is shown in Figure 8.

The recommender is an application that integrates with the system modules used by analysts, as shown in Figure 9. The recommender application is connected to the business database, allowing synchronous access to information. In addition, it exposes the operational view of the recommender used by the purchase analyst, where the code of the SKU to be reviewed is entered, and the recommender provides information for the purchase decision.

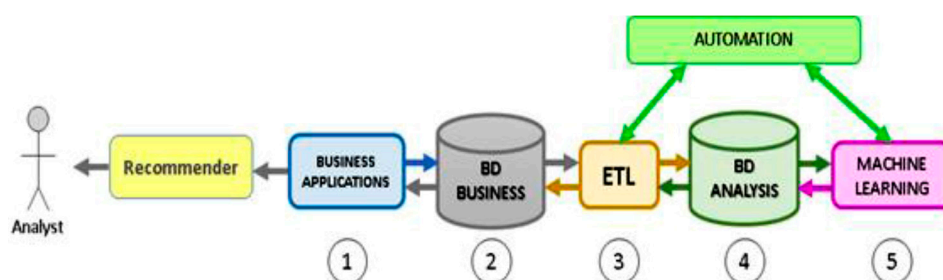


Figure 8. Scheme of the recommender system.

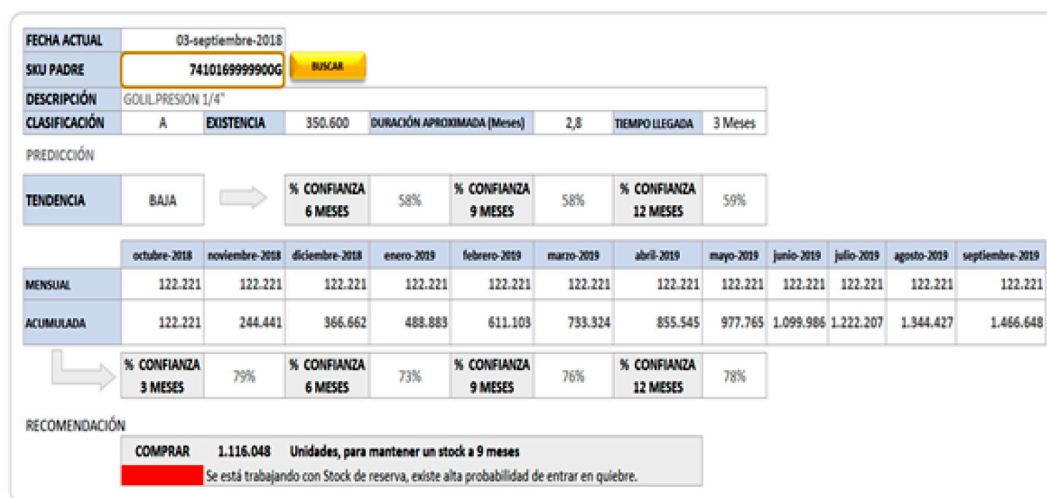


Figure 9. Operational view of the recommender.

CONCLUSIONS

Using a predictive inventory management model allows the company to make better decisions. The model, which in one of its stages uses the ABC classification to organize purchasing logistics and focus logistics models, helps the purchasing department in logistics to align with the ABC classification used in the organization, which generates groups with a priority of merchandise replacement. For this purpose, five machine learning models are applied to each parent SKU belonging to each classification group, where the application automatically selects the model that offers the best score concerning historical demand.

It is possible to deliver timely and relevant information to the company of those Parent SKU products that have a low correlation in the prediction. Such is the case of classification group D, which

according to the ABC classification that the company uses, these products are not highly supervised or controlled by the company's operations. Identified Parent SKUs are tagged as new Parent SKUs with no sales history or less than the defined six-month sales minimum to feed back into the machine learning model.

The reliability of the predictions obtained by applying the proposed model exceeds the reliability obtained by applying the current model. Confirmed that technology adapted to predict future demands in multi-product scenarios delivers greater speed and accuracy in prediction in orientation to the company's business model.

The decrease in the confidence level of the prediction when comparing the two models at the third month of execution, caused by stock breaks, indicates that the proposed model and the current model are equally

sensitive to a stock break; therefore, handling this cause requires consideration when handling the prediction according to the affected SKU.

The implemented operating model responds to the problematic situation exposed. It automatically calculates for each SKU, through a data management system and prediction models, if the demand rises or falls while recommending quantities to acquire and maintain stocks.

The prediction model selection process allows adding and assessing new demand prediction models. This process means that models or algorithms such as Support Vector Machine (SVM) and Recurrent Neural Networks, among others, are added to the current models of the platform.

The recommender operation model exposed in Figure 9 exposes a simplified view for the user of the processed data; therefore, it exposes in a simplified way the result of the application of the Machine Learning models used, delivering a tool with usability to User Level.

The proposed model allows the integration of new machine-learning prediction models. Future work should continue with the validation methodology used, seeking to improve the prediction with higher reliability levels.

REFERENCES

- [1] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, vol. 1, 2000, pp. 29-39.
- [2] W.C. Kim y R. Mauborgne, *La estrategia del océano azul: Crear nuevos espacios de mercado donde la competencia sea irrelevante*. Bogotá, Colombia: Grupo editorial Norma, 2015.
- [3] M. Porter, *Competitive advantage: Creating and sustaining superior performance*. New York, NY, USA: Free Press, 1998.
- [4] A. González Cruz, "Un modelo de gestión de inventarios basado en estrategia competitiva", *Ingeniare, Revista chilena de ingeniería*, vol. 27, no. 4, 2019, doi: 10.4067/S0718-33052020000100133.
- [5] D. Anderson and D. Sweeney, *Statistics for Business and Economics*. Mason, OH, USA: South-Western, Cengage Learning, 2008.
- [6] M. Valencia-Cárdenas, F.J. Díaz-Serna y J.C. Correa-Morales, "Planeación de inventarios con demanda dinámica. Una revisión del estado del arte", *Dyna*, vol. 82, no. 190, pp. 183-191, 2015.
- [7] B. Render, R.M. Stair y M.E. Hanna, *Métodos cuantitativos para los negocios*. 9na. ed. Madrid, España: Pearson Prentice Hall, 2006.
- [8] L.J. Krajewski and L.P. Ritzman, *Administración de operaciones, estrategia y análisis*. 5ta ed., Nacualpan de Juárez, México: Pearson educación, 2000.
- [9] L. Mora García, *Gestión logística en centros de distribución, bodegas y almacenes*, 1ra ed., Bogotá, Colombia: Ecoe ediciones, 2011.
- [10] V. Gutiérrez y C.J. Vidal, "Modelos de gestión de inventarios en cadenas de abastecimiento: Revisión de la literatura", *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 43, pp. 134-149, 2008.
- [11] A. Leal, "Modelo de planificación y control de inventarios para mantenimiento", Tesis de Magister, Univ. Zulia, Maracaibo, Venezuela. [En línea]. Disponible: http://tesis.luz.edu.ve/tde_arquivos/78/TDE-2011-06-08T14, vol. 52, pp. 08Z-1053, 2014
- [12] M.L. Naranjo de Lucca, "Manejo de Inventarios en una Cadena de Supermercado", Tesis de Magister, Universidad de Chile, Santiago, Chile, 2006.
- [13] E. Gutiérrez-González, O.V. Panteleeva, M.F. Hurtado-Ortiz y C. González-Navarrete, "Aplicación de un modelo de inventario con revisión periódica para la fabricación de transformadores de distribución", *Ingeniería, investigación y Tecnología*, vol. 14, no. 4, pp. 537-551, 2013.
- [14] A.E.F. Ackermann y M.A. Sellitto, "Métodos de previsão de demanda: uma revisão da literatura", *Innovar*, vol. 32, no. 85, pp. 83-99, 2022, doi: 10.15446/innovar.v32n85.100979.
- [15] B. Qian and K. Rasheed, "Stock market prediction with multiple classifiers," *Applied Intelligence*, vol. 26, pp. 25-33, 2007, doi: 10.1007/s10489-006-0001-7.
- [16] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, and E. Salwana, "Deep learning for

- stock market prediction,” *Entropy*, vol. 22, no. 8, p. 840, 2020, doi: 10.3390/e22080840.
- [17] R. Singh and S. Srivastava, “Stock prediction using deep learning,” *Multimedia Tools and Applications*, vol. 76, pp. 18569-18584, 2017, doi: 10.1007/s11042-016-4159-7.
- [18] C. Díaz Ramos y L. Orjuela Camargo, “Pronóstico de la demanda del biodiesel mediante la aplicación de técnicas de machine learning”, Fundación Universitaria Los Libertadores, 2021. [En línea]. Disponible: <http://hdl.handle.net/11371/4090>
- [19] J.A. Arango Marín, J.A. Giraldo García y O.D. Castrillón Gómez, “Gestión de compras e inventarios a partir de pronósticos Holt-Winters y diferenciación de nivel de servicio por clasificación ABC”, *Scientia et Technica*, vol. 18, no. 4, pp. 743-747, 2013.
- [20] C.A. Fierro Torres, V.H. Castillo Pérez y C.I. Torres Saucedo, “Comparative Analysis of Traditional and Modern Models for Forecasting Demand: Approaches and Features”, *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 12, no. 24, 2022.
- [21] R. Porteiro, L. Hernández-Callejo y S. Nesmachnow, “Electricity demand forecasting in industrial and residential facilities using ensemble machine learning”, *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 102, pp. 9-25, 2022.
- [22] G.A. Méndez Giraldo y E.R. López Sant, “Methodology to demand forecasting under multiproduct environments and high variability”, *Tecnura*, vol. 18, no. 40, pp. 89-102, 2014.
- [23] F. Pallares Cabrera, “Desarrollo de un Modelo Basado en Machine Learning para la predicción de la demanda de habitaciones y ocupación en el sector hotelero”, Maestría, Universidad Tecnológica de Bolívar, Cartagena, Colombia, 2015. [En línea]. Disponible: <https://repositorio.utb.edu.co/handle/20.500.12585/1020#page=1>
- [24] T. Hastie, R. Tibshirani, J.H. Friedman, and J.H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
- [25] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*, 2nd ed. London, UK: MIT Press, 2018.
- [26] A. Gelman, J. B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin, *Bayesian data analysis*, 3rd ed. Boca Raton, FL, USA: Chapman and Hall/CRC, 2013.
- [27] L. Ramos Castillo, “Regresión Lasso”, Trabajo Fin de Grado, Universidad de Sevilla, Sevilla, España, 2018. [En línea]. Disponible: <https://idus.us.es/handle/11441/77576>
- [28] L. Breiman, “Better subset regression using the nonnegative garrote,” *Technometrics*, vol. 37, no. 4, pp. 373-384, 1995.
- [29] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267-288, 1996.
- [30] R.G.D. Steel and J.H. Torrie, *Principles and procedures of statistics, a biometrical approach*. 2da ed. New York, NY, USA, McGraw-Hill, 1980.
- [31] L. Moreno González-Páramo, “Herramientas avanzadas de análisis de datos de aplicación en ingeniería civil: Machine Learning using Python”, Trabajo Fin de Master, Universidad de Cantabria, España, 2017. [En línea]. Disponible: <https://repositorio.unican.es/xmlui/handle/10902/12684>
- [32] R. Bouman, M. Casters, and J. Van Dongen, *Pentaho Kettle solutions: building open source ETL solutions with Pentaho Data Integration*, Indianapolis, IN, USA: Wiley, 2010.
- [33] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825-2830, 2011. [En línea]. Disponible: <http://jmlr.org/papers/v12/pedregosa11a.html>
- [34] X. Teng and Y. Gong, “Research on application of machine learning in data mining,” in *IOP Conference Series: Materials Science and Engineering*, 2018, vol. 392, no. 6, p. 062202.
- [35] J. Stirrup, “What’s wrong with CRISP-DM and is there an alternative?,” jenstirrup.com. [Online]. Available: <https://jenstirrup.com/2017/07/01/whats-wrong-with-crisp-dm-and-is-there-an-alternative/>
- [36] V. Roman, “Machine Learning Supervisado: Fundamentos de la Regresión Lineal,” medium.com. [En línea]. Disponible: <https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd>