XVIII ICIEOM 2012
International Conference on Industrial Engineering and Operations Management
Guimarães - Portugal
July 9 - 11, 2012

Maximizing Profits with the Improvement in Product Composition, Using Genetic Algorithms and K-Means Application to a Company of the Printing Industry

María A. Guerrero, Rodrigo A. Batistelo and André F. H. Librantz

Industrial Engineering Post Graduation Program, Nove de Julho University (UNINOVE), São Paulo, Brazil

Email: malejandragh@gmail.com, rodrigo.batistelo@gmail.com, librantz@uninove.br

Abstract

In recent years, the printing industries have struggled to remain competitive due to increased concurrence and the appearance of new electronic media aimed at replacing the printed paper. In this context arises the need, as not only high technology, as optimization methods for such industries that can support decision making aiming cost reduction and profit improvements. In this context, it was proposed to maximize the profits of a Colombian printing industry by means of the best composition of their products to be sold, in order to improve its contribution margin (financial remainder of production of each product) and have a strategic vision for the sale of its many products. For this, the products were grouped by similarity, using clustering algorithm called K-Means and it was applied genetic algorithms technique for maximize the Contribution Margin. Results pointed that it was possible to achieve good increments in the revenues and contribution margin, as well.

Keywords: printing industry; optimization methods; genetic algorithm; k-means algorithm.

1 Introduction

1.1 Definition of the Problem

Over the last decade, the printing industry has been faced numerous problems, mostly because of the fierce competition, and the emergence of new communications, graphic and electronic media, which have replaced the printed product. However, lots of companies that are still working in this industry are day-to-day seeking to become highly competitive and efficient. Investments in technology are often not sufficient and do not solve entirely the problem, it is also necessary to develop and apply methodologies in decision making to better use the companies' productive resources. The high quality standards are not enough for competing. For most of the companies, their management boards have sophisticated quality control equipments that help to produce similar and high quality products making very difficult to obtain a remarkable differentiation from the competitors.

The best way to compete is by offering the best prices, even sacrificing utility. Sales budgets are usually not specific, considering only sales by customer or geographic location, no matter in which products the company has the best efficiency and productivity, or those which the company has the bigger contribution margin.

1.2 Company features

The experiments were performed on a printing company located in Colombia.

According to the processes classification defined by (Davis, Aquilano, Nicholas, & Chase, 2001), the type of process where this company can be classified is: Make to Order Production System, which focuses the production in highly customized products (the products are produced with the customer's specifications). According to the same author, these types of processes require more flexibility than the Make to Stock Production System and as a result, tend to be slower, inefficient, and consequently more expensive. Due

to the high diversity in customer orders and the long list of products generated / produced, the tasks of classifying the products, taking into consideration the characteristics of contribution margin and productivity, become very difficult if only traditional methods are adopted by the company, i.e. Excel.

The company has already made some classifications, based on the physical characteristics of the products; defining thirty-five possible product groups (clusters), but each one of these clusters could have important variations of productivity and contribution margins.

2 Scope

The goal of this study is to maximize the utility of the company and to determine which is the best composition of products, that provides the best contribution margin, and thus offer the Sales & Marketing departments, a solid basis for defining sales budgets and also guide them about the products to be sold, based on the best relation between costs and productivity impacting greatly the financial results.

2.1 Relevant Variables for Analysis

For establishing the best composition of products, are considered the two following variables:

- Contribution Margin ($ - %)

Contribution margin is equal to sales revenue minus variable expenses (both manufacturing and non-manufacturing). In other words, the contribution margin means: A measurement of the profitability of a product is the financial leftover of each product's production; it is used for the amortization of fixed costs. Its understanding allows discussions and actions focused on the achievement of the expected utility.

- Production Efficiency: Time Adding Value to the Product

It is measure of the efficiency used by the company. It is the amount of time that the company employs for adding value to the product.

3 Bibliographic Review

3.1 K-Means Algorithm

K-means is a well known clustering algorithm. It consists of a method for clustering analysis whose goal is to partition a set of samples into k clusters in which each sample belongs to the cluster with the nearest mean (Mitchell, 1997). Given a set of samples (x1, x2, …, xn), where each sample is a n-dimensional real vector, k-means algorithm aims to partition the set of samples into k sub-sets S = {S1, S2, …, Sk}, optimizing an objective function F that can be defined as:

$$F = \sum_{i=1}^{k} \sum_{x_j \in S_i} \left\| x_j \quad \mu_i \right\|$$

where μi is the mean of points in Si.

The most common algorithm uses an iterative technique that, in general, considers Euclidean distance as a similarity measure of vectors and variance as a measure of cluster scatter. Since the number of clusters *k* is an input parameter of the algorithm to define the *k* centroids, an inadequate choice of *k* can generate poor results. Thus, for using k-means, it is important to run diagnostic checks for determining the suitable number of clusters for the considered data set.

3.2 Genetic Algorithms

The Genetic Algorithms (GA) consists in a technique of optimization inspired by natural evolution theory. It has been used for solving optimization problems in several areas over the last decades, mainly because its efficiency in irregular search spaces (Goldberg, 1989; Haupt & Haupt, 1998; Mitchell, M., 1998). The GA generally uses binary strings called chromosome or individual to represent solutions of the problem. At the beginning, a population of individuals is randomly generated. The size of population depends on the problem to be solved (Haupt & Haupt, 1998). By means of competition, the most able chromosomes of the population are selected and crossed each other, to generating new chromosomes better than those ones of the previous population. So, at each generation the probability of one or more individual to be a solution of the problem is increased (Goldberg, 1989; Haupt & Haupt, 1998; Librantz, Coppini, Baptista, Araújo, & Rosa, 2011; Santana, Araújo, Librantz, & Tambourgi, 2010). A GA can find the global optimum solution in a complex multi-modal search space without requiring specific knowledge about the target problem that it was developed for solve. Besides, a GA operates over the population in parallel, yielding various solutions at a time. Hence, this method has found applications in engineering problems involving complex combinatorial optimization (Librantz, Coppini, Baptista, Araújo, & Rosa, 2011; Santana, Araújo, Librantz, & Tambourgi, 2010). GA procedure involves four main operations: Evaluation, selection, crossover and mutation. In the evaluation operation fitness function is used to measure the aptitude of the individuals of the population, providing information such as the number of new individuals each one can generate according with its aptitude. The selection consists of the choice of the best individuals for reproduction. Hence, the individuals with better aptitudes are selected while the other ones are discarded, that is, each individual has a probability to be selected according to its aptitude. In the crossover operation, the genetic material of the best chromosomes is mixed to generate the individuals of the next population. Finally, a random change in a small number of bits, with some small probability, is performed in order to preventing the population of chromosomes becoming too similar to each other or, in other words, to preserve the diversity of the population. This operation is essential to avoid the premature convergence.

It is important to remember that the suitable convergence of the GA depends on some parameters such as: size of chromosome, size of population, crossover points and mutation rate (Librantz, Coppini, Baptista, Araújo, & Rosa, 2011). Unfortunately, the determination of these parameters is a difficult task because and, in general, they are empirically defined (Pacheco, 1999).

4 Methodology

Here, it was described how we work with the data available, the application of optimization techniques, the objective function definition and the tools used in this work.

4.1 Data

In order to develop the experimental part of this project it was used historical data from 10 months of a company in the printing industry. These data were obtained from a corporate database ORACLE (Figure 1) and add a total of 2618 records with data on products, customers, vendors, business, revenue, product costs, contribution margin and time adding value to the product (sec / und). The classification of products was defined by the analysts of the product and classified into 35 groups (clusters), taking into account basically the shape of the product.

CLUSTER	PRODUCT	MONTH	REVENUES	SALES VOLUME (UNITS)	PRODUCTION COSTS	TOTAL COSTS	CONTRIBUTION MARGIN (%)	EGP	EFFICIENCY (SEC/UNIT)
26	LIBROS - INTERES GENERAL	1	4.871.988	1.000	2.679.593	3.410.392	0,45	61%	0,23
51	PAPELERIA - CUADERNO	1	2.479.997	200	1.363.998	1.735.998	0,45	53%	0,28
80	REVISTA - REVISTA	1	70.985.654	600.200	49.689.958	60.337.806	0,30	36%	0,38
65	POP - POP	1	974.400	12.000	623.616	769.776	0,36	63%	0,22
24	LIBROS - ENCICLOPEDIA/COLECCION	1	159.067.854	3.167	119.300.891	143.161.069	0,25	12%	0,53
82	REVISTA - CIRCULACION ABIERTA	1	12.377.200	19.400	8.787.812	10.644.392	0,29	37%	0,38
83	REVISTA - CIRCULACION CERRADA	1	4.455.104	69.611	3.341.328	4.009.594	0,25	32%	0,41
82	REVISTA - CIRCULACION ABIERTA	1	5.342.700	32.380	4.007.025	4.808.430	0,25	21%	0,47
82	REVISTA - CIRCULACION ABIERTA	1	29.200.000	40.000	21.024.000	25.404.000	0,28	37%	0,38
80	REVISTA - REVISTA	1	11.772.585	23.783	8.711.713	10.477.601	0,26	23%	0,46
82	REVISTA - CIRCULACION ABIERTA	1	24.038.055	247.815	17.067.019	20.672.727	0,29	38%	0,37

Figure 1: Sample of data from Oracle database (Colombian printing industry).

4.2 Tools and Procedures

Because of the wide variety of products and to better analyze the product portfolio, the project was divided in two stages:

- Grouping by similarity of products, using clustering algorithm K-Means, which allowed grouping the products in nine different clusters. For this, it was used function K-Means of the MATLAB software.
- Minimization of the objective function using Genetic Algorithms Toolbox of the MATLAB software.

4.2.1 Application of K-Means Algorithm

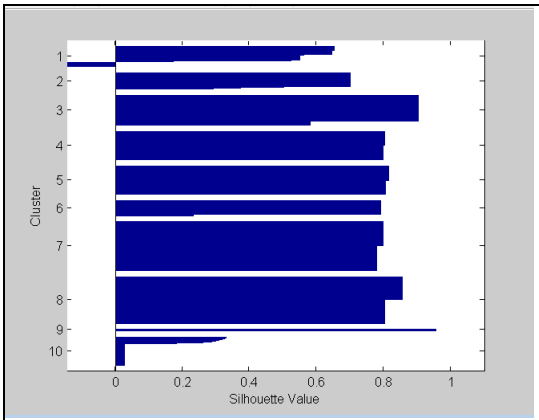
The number of clusters was defined and evaluated through graphics of silhouettes given by K-Means.

- **Graphics of Silhouettes**

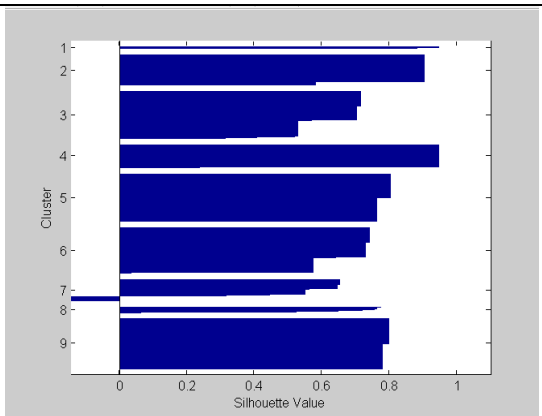
The graphic silhouette is a graphical representation of the quality of generated clusters. This chart can give an idea if the resulting number of clusters was set properly. These graphics are shown in Figure 2.

The silhouettes show a measure of how close is the point of a given cluster of points of neighboring clusters. This measure ranges from +1, indicating points which are distant neighboring clusters, through 0 indicating which are difficult to classify, according to their cluster, to -1 indicating points which area assigned to the cluster probably wrong.

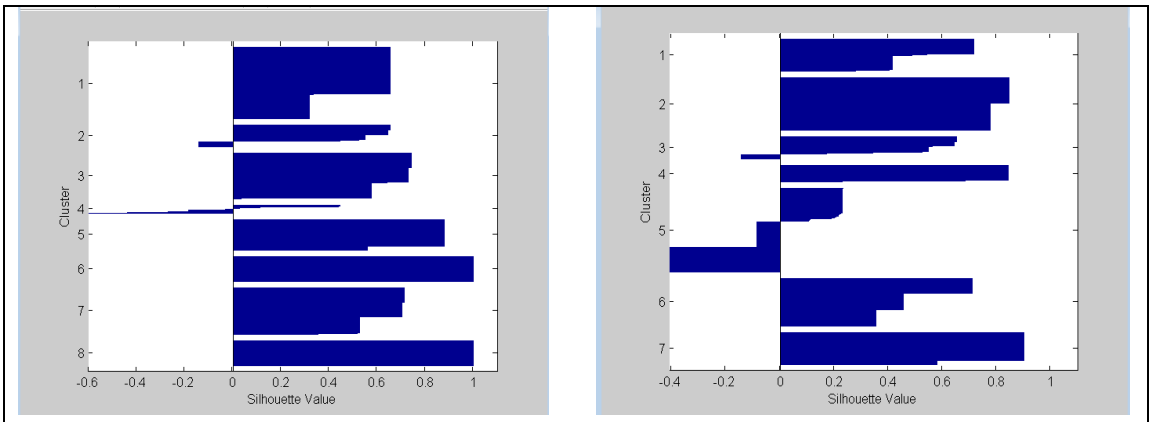
The database was analyzed to be clustered in 10, 9, 8 and 7 clusters.



Average Silhouettes for K-Means with 10 clusters = 0.6939



Average Silhouettes for K-Means with 9 clusters = 0.7303



Average Silhouettes for K-Means with 8 clusters = 0.6676 Average Silhouettes for K-Means with 7 clusters = 0.4527

Figure 2: The graphic silhouette for several cluster numbers.

- **Number of Clusters Definition**

Based on the results evaluated by the graphics of silhouettes for 7, 8, 9 and 10 clusters, it was obtained the following results (Table 1):

Table 1: Average silhouettes for several cluster numbers.

CLUSTER	7	8	9	10
AVERAGE SILHOUETTES	0,4527	0,6676	0,7303	0,6939

It was chosen the cluster nine, once it showed the best average silhouettes values.

- **MATLAB K-Means function**

Based on the results evaluated by the graphics of silhouettes for 7, 8, 9 and 10 clusters, it was executed in MATLAB environment the command lines:

1) X = textread ('C:/ICIEOM2012/Database_K-Means.txt');
2) v_options = statset ('Display', 'final');
3) [value, centroid] = kmeans (X, 9, 'Distance', 'city', 'Replicates', 10, 'Options', v_options)

After that, it was possible to obtain the clustering by K-means. In Table 2, it was shown the labels for the values of centroids for each cluster, returned after the execution of MATLAB commands listed above.

Table 2: Labels for the centroids.

CENTROID1	CENTROID2	CENTROID3	CENTROID4	CENTROID5
0.4500 0.3300	0.5500 0.2700	0.4900 0.3100	0.2500 0.4500	0.2900 0.4300
CENTROID6	CENTROID7	CENTROID8	CENTROID9	
0.8400 0.0900	0.2100 0.4700	0.4600 0.3200	0.5800 0.2500	

The K-Means ranks with numbers between 1 and 9, the data whose centroids were calculated. The chat below (Figure 3) shows the clustering obtained by K-Means.

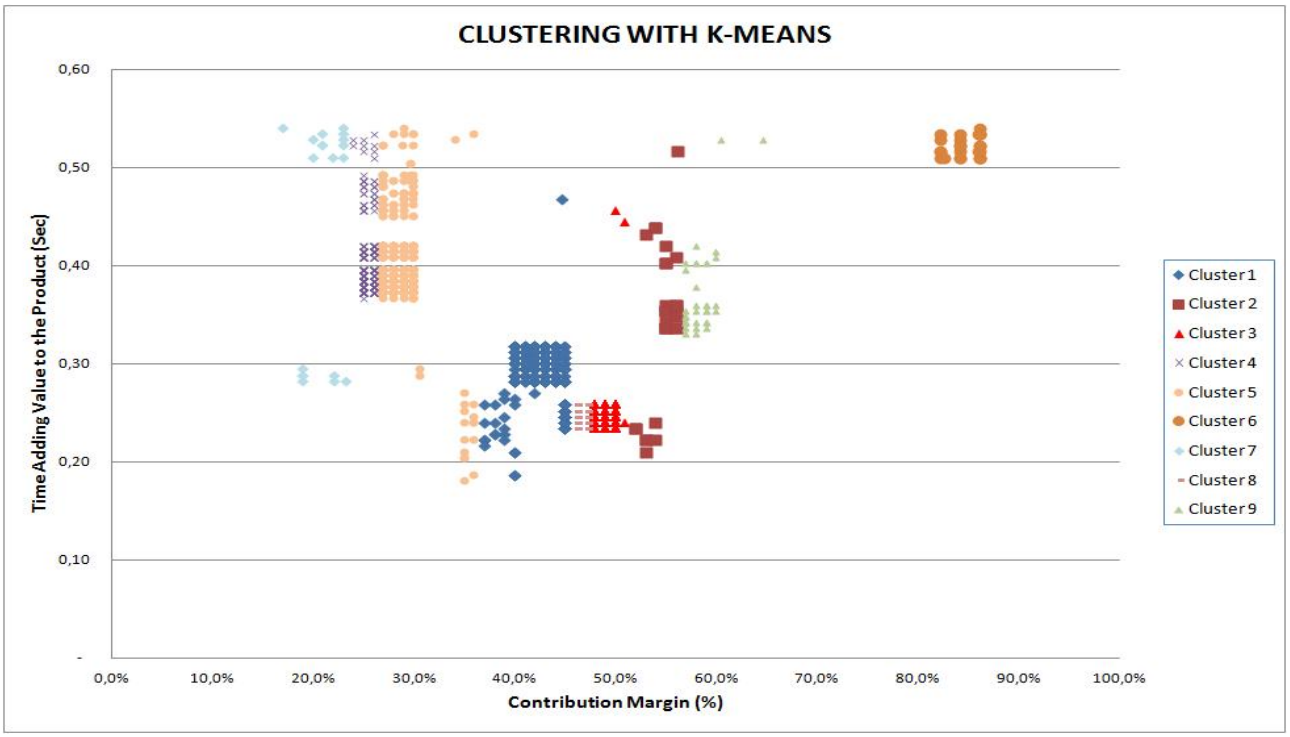


Figure 3: Clustering obtained by K-Means.

These numbers were taken back to the database, to label the data and have a good classification to continue with the phase of minimization with Genetic Algorithms (GA).

4.2.2 Application of Genetic Algorithms (GA)

The project´s final goal is to maximize the profit of a printing company doing the best composition of the products to be sold to improve the contribution margin of the company. As the Genetic Algorithms Toolbox of the MATLAB works only by minimizing functions, we seek to minimize costs so as to maximize the profit.

- **Definition of Fitness Function / Objective Function**

Taking into consideration the purpose of the present work, understanding the problem as much as the business, we chose the following variables (Table 3) to compose our fitness function:

Table 3: Variables chosen to compose the fitness function.

P	**PRICE ($/unit)**
M	**CONTRIBUTION MARGIN ($)**
R	**EFFICIENCY (sec/unit)**
V	VOLUME (unit)
I	**CLUSTER**

Then the fitness function that minimizes the costs of the company was defined as:

$$\sum_{i=1}^{i=9} \frac{P_i(1-M_i)}{R_i} \times V_i$$

The volume for each cluster (Vi) is the variable that will be found by the GA and which will provide an ideal composition of groups of products, in other words, it will inform the volume of each group of products to form the preferred portfolio sales. This result could guide the sales department at the moment of product offering.

- **Fitness Function in MATLAB**

Below is the function fitness.m, created in MATLAB for execution of its toolbox of genetic algorithms. It is that allows us to obtain, through a data file, the appropriate volumes of each of the groups (clusters) of product.

```
function result = fitness (V)
   X = textread ('C:/ICIEOM2012/Database_GA.txt');
   result = 0;
   tam = length (X(:,1));
   for i=1:tam
      R = X (i, 1);
      M = X (i, 2);
      P = X (i, 3);
      result = result + (((P*(1-M)) / R) * V(i));
   end
end
```

- **Input Data**

The input data of the file used in the fitness function, are prepared as follows (Table 4):

Table 4: Sample data used in applying the fitness function.

EFFICIENCY (sec/unit)	CONTRIBUTION MARGIN (%)	PRICE PER UNIT PRODUCED ($/unit)
0.26	0.464	2.29
0.24	0.441	1.00
0.38	0.253	0.44
0.25	0.493	1.66
0.24	0.482	2.26
0.42	0.562	5.16
0.43	0.271	0.58
0.38	0.287	0.44
0.30	0.421	2.61

The principal historical data are summarized in the Table 5 below:

Table 5: Principal historical data.

REVENUES (BRL)	PRODUCTION COSTS (BRL)	CONTRIBUTION MARGIN (%)	SALES VOLUME (units)
$ 90,194,579.1	**$ 51,012,304.6**	**43%**	**83,934,774**

- **Restrictions**

In this work we have established a single primary constraint, which is the sum of the volume of the nine clusters limited to 84 million units, which is the maximum production capacity of the factory according to the information of the production's analyst.

$$\sum_{t=1}^{t=9} V_t = 84.000.000$$

- **Definition of Chromosome**

The chromosome of the objective function is composed of 9 genes; each gene represents the volume of each cluster, which corresponds to the objective variables, those found by the genetic algorithm.
The way in how the genetic algorithms are presented is based on the binary system, the number of alleles of each gene is defined by the function: 2^k, where *k* is the size of the gene, making the conversion to the decimal system should contain the value of 84 million (restriction of number of production units described above). In this case, the k value for the gene studied is 26, so the number of alleles for the gene would be 27 (2^{26}, where $0 \leq k \leq 26$). All genes in the chromosome have the same structure as presented in Figure 4.

1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 4: Representation of the gene with k=26

In general it can be stated that all variables in this chromosome may represent a maximum of 84,000,000 units, meanwhile the other variables are 0.
The chromosome is composed of 9 genes (cluster), each with 27 alleles, totalizing 243 alleles, which can take values 1 or 0 of the binary system, respecting the constraint that the sum in decimal is less than or equal to 84,000,000.

4.3 Results and discussion

Some scenarios were analyzed by changing the values of the operators of the Genetic Algorithm. Each of the proposed scenarios is presented below. They show the improvements obtained with respect to the current scenario / position of the company.

4.4 Scenario One

For scenario one, the operators were chosen, as shown in Table 6, both; the value of fitness function and the best individual of Genetic Algorithms for this scenario are shown in Figure 5.

Table 6: Operators of genetic algorithms, referring to scenario one.

Population Size: 50, **Selection**: Tournament, **Crossover Fraction**: 0,8, **Crossover**: Two Point, **Mutation**: Use constraint dependent default

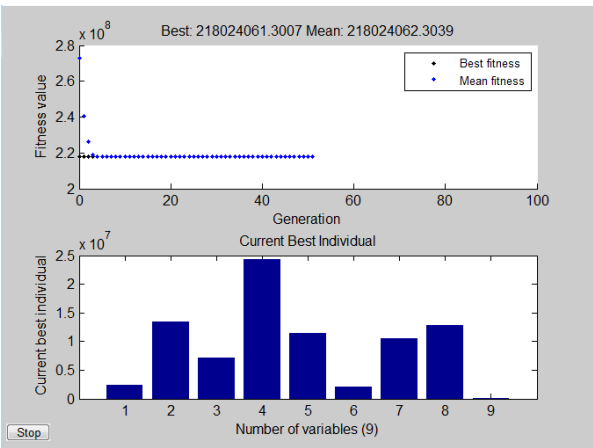


Figure 5: Value of fitness function and best individual obtained with genetic algorithms in scenario one.

Below (Figure 6), we are presenting a comparison between the company's current situation and the result obtained in scenario one.

	Current Scenario				Scenario One			
CLUSTER	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)
1	1.040.438	$ 2.386.111,18	$ 1.107.155,59	46,40%	2.445.932	$ 5.609.431,99	$ 2.602.776,45	46,40%
2	5.099.756	$ 5.111.750,75	$ 2.254.282,08	44,10%	13.432.721	$ 13.464.313,73	$ 5.937.762,35	44,10%
3	12.488.002	$ 5.505.058,63	$ 1.392.779,83	25,30%	7.173.454	$ 3.162.257,95	$ 800.051,26	25,30%
4	3.194.999	$ 5.295.899,88	$ 2.610.878,64	49,30%	24.285.046	$ 40.253.897,65	$ 19.845.171,54	49,30%
5	2.894.923	$ 6.545.844,40	$ 3.155.097,00	48,20%	11.385.257	$ 25.743.730,68	$ 12.408.478,19	48,20%
6	6.833.058	$ 35.235.058,03	$ 19.802.102,61	56,20%	2.012.081	$ 10.375.413,08	$ 5.830.982,15	56,20%
7	28.977.042	$ 16.914.815,88	$ 4.583.915,10	27,10%	10.524.630	$ 6.143.559,32	$ 1.664.904,57	27,10%
8	22.044.625	$ 9.643.743,09	$ 2.767.754,27	28,70%	12.740.879	$ 5.573.683,77	$ 1.599.647,24	28,70%
9	1.361.930	$ 3.556.297,31	$ 1.497.201,17	42,10%	0	$ 1,18	$ 0,50	42,10%
	83.934.774	**$ 90.194.579,14**	**$ 39.171.166,29**	**43,43%**	**84.000.000**	**$ 110.326.289**	**$ 50.689.774**	**45,95%**

Figure 6: Comparison between the current scenario and scenario one evaluated.

This scenario generated an improvement of $ 20.131.710BRL in billing and a contribution margin increase of $ 11.518.607BRL, which corresponds to an improvement of 5.48%.

4.5 Scenario Two

For scenario two, the operators were chosen, as shown in Table 7 and both; the value of fitness function and the best individual of Genetic Algorithms for this scenario are shown in Figure 7.

Table 7: Operators of genetic algorithms, referring to scenario two.

Population Size: 20, **Selection**: Stochastic uniform, **Crossover Fraction**: 0,8, **Crossover**: Scattered, **Mutation**: Use constraint dependent default

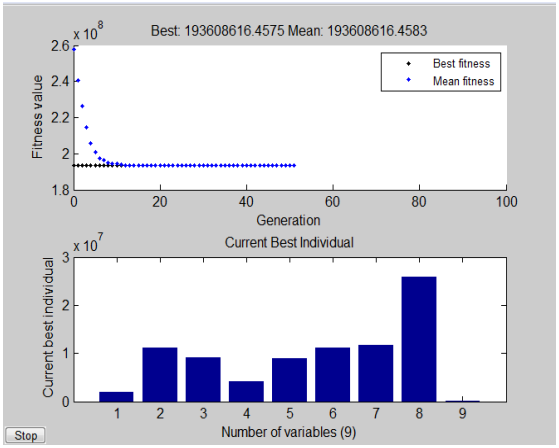


Figure 7: Value of fitness function and best individual obtained with genetic algorithms in scenario two.

Below (Figure 8), we are presenting a comparison between the company's current situation and the result obtained in scenario two.

	Current Scenario				Scenario Two			
CLUSTER	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)
1	1.040.438	$ 2.386.111,18	$ 1.107.155,59	46,40%	1.978.189	$ 4.536.723	$ 2.105.040	46,40%
2	5.099.756	$ 5.111.750,75	$ 2.254.282,08	44,10%	11.107.635	$ 11.133.760	$ 4.909.988	44,10%
3	12.488.002	$ 5.505.058,63	$ 1.392.779,83	25,30%	9.025.165	$ 3.978.544	$ 1.006.572	25,30%
4	3.194.999	$ 5.295.899,88	$ 2.610.878,64	49,30%	4.223.245	$ 7.000.279	$ 3.451.137	49,30%
5	2.894.923	$ 6.545.844,40	$ 3.155.097,00	48,20%	8.960.041	$ 20.259.963	$ 9.765.302	48,20%
6	6.833.058	$ 35.235.058,03	$ 19.802.102,61	56,20%	11.107.635	$ 57.277.161	$ 32.189.764	56,20%
7	28.977.042	$ 16.914.815,88	$ 4.583.915,10	27,10%	11.773.092	$ 6.872.326	$ 1.862.400	27,10%
8	22.044.625	$ 9.643.743,09	$ 2.767.754,27	28,70%	25.824.997	$ 11.297.522	$ 3.242.389	28,70%
9	1.361.930	$ 3.556.297,31	$ 1.497.201,17	42,10%	0	$ 0	$ 0	42,10%
	83.934.774	**$ 90.194.579,14**	**$ 39.171.166,29**	**43,43%**	**84.000.000**	**$ 122.356.278**	**$ 58.532.592**	**47,84%**

Figure 8: Comparison between the current scenario and scenario two evaluated.

This scenario generated an improvement of $ 32.161.698BRL in billing and a contribution margin increase of $ 19.361.426BRL, which corresponds to an improvement of 9.21%.

4.6 Scenario Three

The operators were chosen, as shown in Table 8, and both; the value of fitness function and the best individual of Genetic Algorithms for this scenario are shown in Figure 9.

Table 8: Operators of genetic algorithms, referring to scenario three.

Population Size: 20, **Selection**: Tournament, **Crossover Fraction**: 0,8, **Crossover**: Intermediate, **Mutation**: Use constraint dependent default

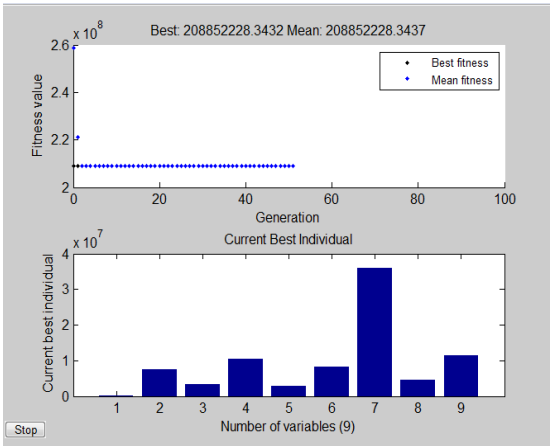


Figure 9: Value of fitness function and best individual obtained with genetic algorithms in scenario three.

Below (Figure 10), we present a comparison between the company's current situation and the result obtained in scenario three.

	Current Scenario				Scenario Three			
CLUSTER	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)	SALES VOLUME (UNITS)	REVENUES	CONTRIBUTION MARGIN ($-BRL)	CONTRIBUTION MARGIN (%)
1	1.040.438	$ 2.386.111,18	$ 1.107.155,59	46,40%	0	$ 0	$ 0	46,40%
2	5.099.756	$ 5.111.750,75	$ 2.254.282,08	44,10%	7.467.111	$ 7.484.674	$ 3.300.741	44,10%
3	12.488.002	$ 5.505.058,63	$ 1.392.779,83	25,30%	3.386.922	$ 1.493.050	$ 377.742	25,30%
4	3.194.999	$ 5.295.899,88	$ 2.610.878,64	49,30%	10.384.749	$ 17.213.336	$ 8.486.174	49,30%
5	2.894.923	$ 6.545.844,40	$ 3.155.097,00	48,20%	2.732.452	$ 6.178.473	$ 2.978.024	48,20%
6	6.833.058	$ 35.235.058,03	$ 19.802.102,61	56,20%	8.212.544	$ 42.348.455	$ 23.799.831	56,20%
7	28.977.042	$ 16.914.815,88	$ 4.583.915,10	27,10%	35.887.014	$ 20.948.385	$ 5.677.012	27,10%
8	22.044.625	$ 9.643.743,09	$ 2.767.754,27	28,70%	4.624.897	$ 2.023.229	$ 580.667	28,70%
9	1.361.930	$ 3.556.297,31	$ 1.497.201,17	42,10%	11.304.311	$ 29.518.034	$ 12.427.092	42,10%
	83.934.774	**$ 90.194.579,14**	**$ 39.171.166,29**	**43,43%**	**84.000.000**	**$ 127.207.634**	**$ 57.627.284**	**45,30%**

Figure 10: Comparison between the current scenario and scenario three evaluated.

This scenario generated an improvement of $ 37.013.055BRL in billing and a contribution margin increase of $ 18.456.117BRL, which corresponds to an improvement of 4.13%.

5 Conclusions

In this work it was proposed to maximize the profits of a Colombian printing industry by using k-means clustering algorithm combined to GA technique.

The printing industry is characterized by a high variety of products, because they are responding to customer specifications. Appraising the best mix of products and the quantity of units produced per product type is extremely important, but this analysis is technologically unfeasible and would result in a long list of objective variables to be be analyzed by the seeking algorithm, in this case, the genetic algorithm. The classification of these products requires that the patterns' identification is accurate and that leads to an effective as well as assertive classification to assist the decision taking.

The use of clustering techniques with K-Means proved to be a good alternative when it is necessary to group this type of "population" as well as reducing the number of variables to be used and of course, evaluated by the genetic algorithm. This genetic algorithm is a robust and flexible approach to evaluate alternatives that improve and optimize the performance of the company and also adapt them to market conditions.

These works contribute to improve the rentability of the companies, providing important information such as: What products to promote/sell considering, not only their contribution margin, but also, which products have the best performance in production, so the company can improve both; financial and productive indicators.

The results of this research demonstrated the great importance of the combination of these two techniques for supporting the decision taking. The experiments allowed offering different alternatives for the mixture of products that enable the company to improve its financial performance and adapt its production to the market needs.

References

Davis, M. M., Aquilano, N. J., & Chase, R. B. (2001). Fundamentos da Administração da Produção, 3ª Edição, Artmed Editora.

Goldberg, D.E. (1989). Genetic algorithms in search, optimization, and machine learning. Addison-Wesley: Massachusetts, 432 pp.

Haupt, R.L. & Haupt, S.E. (1998). The Binary Genetic Algorithm, In: Haupt, R.L; Haupt, S.E. Pratical Genetic Algorithms (1 ed.). Wiley-Interscience: New York, 276 pp.

Librantz, A. F. H., Coppini, N. L., Baptista, E. A., Araújo, S. A., & Rosa, A. F. C. (2011). Genetic Algorithm Applied to Investigate Cutting Process Parameters Influence on Workpiece Price Formation. Materials and Manufacturing Processes, v. 26, p. 550-557.

Mitchell, M. (1998). An introduction to genetic algorithms, First MIT Press paperback edition.

Mitchell, T. (1997). Machine Learning. McGraw-Hill, USA.

Pacheco MA. (1999). Algoritmos Genéticos: Princípios e Aplicações. In: V Congreso Internacional de Ingeniería Electrónica, Eléctrica y Sistemas, Lima, 11-16.

Santana, J. C. C., Araújo, S. A., Librantz, A. F. H., & Tambourgi, E. B. (2010). Optimization of Corn Malt Drying by Use of a Genetic Algorithm. Drying Technology, v. 28, p. 1236-1244.