Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11431
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Applying genetic algorithm for optimization income value

Ayu Febri Siagian[1)*], Gomal Juni Yanris[2)], Sahat Parulian Sitorus[3)]
[1,2,3)]Universitas Labuhanbatu, Indonesia
[1)]ayufebsiagian@gmail.com, [2)]gomaljuniyanris@gmail.com, [3)]sahatulb@gmail.com

Submitted : May 8, 2022 | **Accepted** : May 13, 2022 | **Published** : May 13, 2022

Abstract: In this digital era, the use of information technology and internet technology cannot be separated from digital services. Starting from product promotion media, recording customer data, determining the amount of revenue from product sales, and optimizing the value of revenue. Sales of digital service products owned by PT. XYZ needs to be evaluated to find out which products are most in demand by customers from each product offering that has been made. Therefore we need a system to calculate revenue from the number of customers who use the product for further promotion. The object of this research focuses on optimizing the value of income at PT. XYZ of the products they market, the results of the object will be used as an evaluation to determine a new strategy in carrying out promotions for products that are less attractive to customers. The data used in this study is customer data for January 2017-December 2021. The method used in this study uses a genetic algorithm to determine the optimization of the revenue value. For the optimization results, the genetic algorithm went well, because it resulted in a smaller comparison of error values compared to values that were not optimized. The error value in January 2019 with a non-optimized value was 35,498.8 and the optimized value got an error value of 32,364.9. The results of this study are used as a sales evaluation to increase promotions on digital services that are less attractive to customers. In addition, the results of the application of this genetic algorithm method can provide a better solution to increase income in the next period.

Keywords: Genetic algorithm; Income; Optimization; Prediction; Regression.

INTRODUCTION

Advances in information technology today have changed the work pattern of a company or organization. Large amounts of data and have been stored for a long time can be processed into information that can help analyze the existence of a company or organization. Automatic analysis of large or complex data is carried out to find important patterns or trends (Ermatita, 2009). In addition, optimization problems are things that are often encountered in a company. This is inseparable from the desire of a company that wants to get big profits by minimizing losses. Optimization is an attempt to solve problems in the best way in conditions of limited resource allocation (Primadani, Palgunadi, & Harjito, 2015).

PT. XYZ as a private company located in Rantauprapat provides various information and communication facilities to meet the needs of the community in the media and entertainment sector. PT XYZ has digital service products in the Cable TV and Satellite TV industry. PT. XYZ requires a better income value than the previous period. To determine the value of the next income, it can be done by optimizing the value oriented to the value of income, and to each product that is marketed. To determine the optimization of the income value, the right method is needed, one of which is by using a Genetic Algorithm.

A number of previous studies have proven that genetic algorithms can be applied to optimization problems in a company. Optimization methodology with genetic algorithm technique has found the optimal design with the lowest cost for the concrete arch bridge. The results obtained after geometric improvement have a percentage reduction range from 30% to 35% compared to traditional designs (Abd Elrehim, Eid, & Sayed, 2019). Genetic algorithms are also effective in solving flexible job shop scheduling. The results of this study prove that genetic algorithms can overcome the shortcomings of traditional algorithms with better performance (Luo, Qian, & Fu, 2020). Other studies have also proven that genetic algorithms have been able to reduce investment risk for the stock market (Martins & Neves, 2020).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11431
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Based on the background of the problems and phenomena that have been described, this study aims to apply a genetic algorithm to optimize the value of income at PT. XYZ. Meanwhile, the formulation of the problem is how to compare the value of the number of customers and the number of products that are optimized with those that are not optimized for the value of revenue at PT XYZ.

LITERATURE REVIEW

Optimization is a method of solving problems in the best way to get ideal results (Primadani et al., 2015). The optimization objective is to determine the optimal operating parameters under various conditions and constraints of the variable cycle (Zhang et al., 2022). In addition to improving the quality of the parameters, optimization algorithms are used to determine new weights for learning efficiency (Khotimah, Syarief, Miswanto, & Suprajitno, 2021). Optimization techniques work by minimizing or maximizing the function of discrete variables (Garcia & Lôndero, 2021). Optimization algorithms can be used for general pattern search, direct search, stochastic (Alfarizi, Stanko, & Bikmukhametov, 2022), and model predictions (Song, Cai, Cen, Xu, & Zeng, 2022). Optimization techniques also aim to speed up the process by reducing the number of iterations (Al Rivan, Steven, & Tanzil, 2020).

Genetic Algorithm is a search algorithm based on natural selection (genetic) mechanisms that are suitable for complex optimization problems (Wicaksono, 2019). This algorithm is the most commonly used evolutionary algorithm in optimization problems (Khan et al., 2022). Genetic algorithms are able to find the best solutions in various optimization problems using a number of steps, namely, initialization, selection, crossover, mutation, and replacement (Sari, Alkaff, Wijaya, Soraya, & Kartikasari, 2019). Genetic Algorithm is a technique of finding the right solution from many available solutions without having to test them one by one (Istianto & 'Uyun, 2021). The ability, efficiency, and dynamic nature of genetic algorithms in solving optimization problems with the right tools have become the main reasons for a number of researchers to solve complex problems (Shehadeh, Alshboul, Tatari, Alzubaidi, & Hamed El-Sayed Salama, 2022).

METHOD

The problem analysis was conducted based on research at PT XYZ. Where there are several problems, which include marketing strategies, forecasting the value of income per month, and optimizing the value of the income. The focus of this research is only on the problem of optimizing the value of income. In this research, Genetic Algorithm is used to optimize the income value. The data analyzed is customer data obtained from company archives. Then from the data taken 2 attributes, namely the number of subscribers who subscribe, and the number of products. The data represents the value of revenue optimization. The flow of the stages of this research is shown in Figure 1.

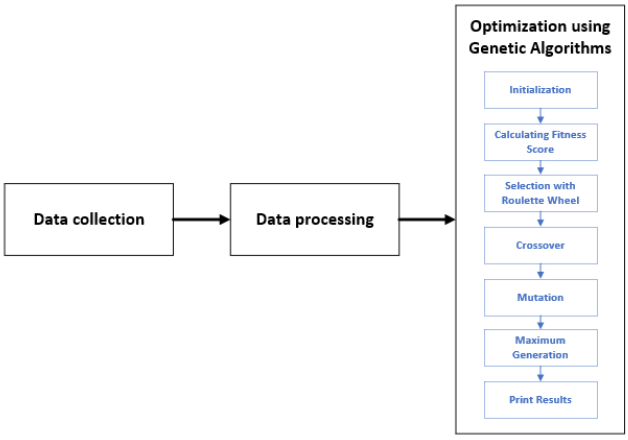

Fig 1. Research Framework

Data collection

Data collection is done by taking data from company archives in the form of customer data. The data contains information obtained by the company regarding customers while subscribing to the service from January 2017 to December 2021. The information includes subscriber numbers, addresses, registration dates, service dates can be used, subscription dates, Cable TV services used, and prices from each Cable TV service.

Data processing

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v7i2.11431 p-ISSN : 2541-044X

Data processing is done by multiple linear regression method. The process carried out is by adding up the total customers, adding up the total product, adding up the total revenue, raising the value, and determining the value of the determinant.

Optimization using Genetic Algorithms

At this stage the optimization process is carried out using a genetic algorithm through a number of stages as follows:

- **Initialization**
  The initialization process is carried out by giving the initial values of the genes with random values according to predetermined limits.

- **Calculating Fitness Score**
  The factors supporting the fitness value in this study are the number of customers in a certain month period and the number of products in a certain month period. In this selection process, it is done by making chromosomes that have a small evaluation value, getting a greater probability of being selected with a high probability value.

- **Selection with Roulette Wheel**
  Selection using the roulette wheel method is carried out in stages: calculating the fitness value of each individual, and calculating the probability value of each individual's selection.

- **Crossover**
  The crossover process is done by forming two new offspring (child chromosomes) from two parents. Only offspring that meet the requirements of the problem will be added to the existing population. The crossover process is done randomly.

- **Mutation**
  This mutation process is carried out by selecting the chromosome to be mutated at random, and then determining the mutation point on the chromosome at random. Mutations are applied with very little probability. To carry out the mutation process is done by calculating the total length of the gene. To select the position of the mutated gene, it is done by giving random numbers 1-12. Then the value of the gene at that position is replaced with a predetermined ID.

- **Maximum Generation**
  This maximum generation process will check the new population that has been formed by going through quite long stages. If the new population produced is the best, it will then repeat the stages in the formation of a new population such as evaluating the fitness value, selection process, crossover process, and mutation to form the next new population.

- **Print Results**
  The next process is to print the results of the new population formed with the optimal solution to solve the problem of optimizing the value of income in this study.

RESULT

This optimization technique is carried out because the error value resulting from this multiple regression method is quite high, so optimization is needed in the hope of reducing the error value. The following are the results of the steps that have been carried out in the optimization process using genetic algorithms.

In this initialization stage, the number of products and the number of customers are given an initialized identity or ID in the form of the number 1 to the amount of data to be predicted, so that an initialization of identity or ID in the form of the number 1 to the amount of data to be used is a gene. Each number is a gene from the chromosome that has been assigned an identity in lieu of the number of customers and the number of products. The number of genes has been determined as much as the number of values to be processed. Chromosomes are used to represent a possible solution of a problem to be solved using a genetic algorithm. The following are the chromosomes that are generated or that will be used in the optimization process.

- Chromosome[1] = [1 4]

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11431
e-ISSN : 2541-2019
p-ISSN : 2541-044X

- Chromosome[2] = [2 5]
- Chromosome[3] = [3 6]
- Chromosome[4] = [6 3]
- Chromosome[5] = [3 2]
- Chromosome[6] = [4 1]

Next, the objective function of the generated chromosome is calculated.

- Chromosome[1] = (27 + 8.15*1 + 17*4) = 103.15
- Chromosome[2] = (27 + 8.15*2 + 17*5) = 128.3
- Chromosome[3] = (27 + 8.15*3 + 17*6) = 153.45
- Chromosome[4] = (27 + 8.15*6 + 17*3) = 126.9
- Chromosome[5] = (27 + 8.15*3 + 17*2) = 85.45
- Chromosome[6] = (27 + 8.15*4 + 17*1) = 76.6

Average score of evaluation = (103.15 + 128.3 + 153.45 + 126.9 + 85.45 + 76.6) / 6 = 112.308

Next, the fitness value is calculated.

- Fitness[1] = 1/103.15 = 0.0097
- Fitness[2] = 1/128.3 = 0.0078
- Fitness[3] = 1/153.45 = 0.0065
- Fitness[4] = 1/126.9 = 0.0079
- Fitness[5] = 1/85.45 = 0.0117
- Fitness[6] = 1/76.6 = 0.0130

Total fitness score = 0.0097 + 0.0078 + 0.0065 + 0.0079 + 0.0117 + 0.0130 = 0.0566

After the crossover is done, the new chromosomes that are formed are as follows:

- Chromosome[1] = [1 6]
- Chromosome[2] = [2 5]
- Chromosome[3] = [3 6]
- Chromosome[4] = [3 2]
- Chromosome[5] = [2 1]
- Chromosome[6] = [4 1]

To select the position of the mutated gene, it is done by giving random numbers 1-12. The value of nm = 20% is determined, then the number of mutated genes is 0.2 * 12 = 2.4 = 2. Two gene positions that will be mutated, after being randomized are gene positions 4 and 7. Then the value of the gene in that position is replaced with a predetermined ID.

- Chromosome[1] = [1 6]
- Chromosome[2] = [2 5]
- Chromosome[3] = [3 6]
- Chromosome[4] = [1 2]
- Chromosome[5] = [2 1]
- Chromosome[6] = [4 1]

At this stage, the imported dataset will be calculated regarding the correlation between these variables.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11431
e-ISSN : 2541-2019
p-ISSN : 2541-044X

```
> results <- lm(pendapatan~jumlah_pelanggan+jumlah_produk,dataset)
> summary(results)

Call:
lm(formula = pendapatan ~ jumlah_pelanggan + jumlah_produk, data = dataset)

Residuals:
    Min      1Q  Median      3Q     Max
-5245.9 -2399.0  -459.8  2330.0  5431.8

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -3.685e+03  3.138e+03  -1.174    0.245
jumlah_pelanggan -4.035e+02  4.757e+02  -0.848    0.400
jumlah_produk    7.959e-02  6.626e-03  12.011   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2665 on 57 degrees of freedom
Multiple R-squared:  0.7663,    Adjusted R-squared:  0.7581
F-statistic: 93.48 on 2 and 57 DF,  p-value: < 2.2e-16
```

Fig 2. Variable Correlation

The results of the calculation of the correlation of the variables in Figure 2 are quite good for the correlation between the independent variable (number of customers and number of products) and the dependent variable (income). Because the resulting Multiple R-squared is 0.7663 (76%) and the Adjusted R-squared is 0.7581 (75%). These results indicate that there is a correlation between variables, but it is not strong enough because it still produces a fairly high standard error, namely the residual standard error: 2665 at 57 degrees of freedom.

In this study, the type of coding that will be used is real numbers because it is suitable for this research, namely for value optimization. Pcrossover serves to provide a probability value between Chromosome with a given value of 0.8. This fitness is useful for determining the best solution from the many solutions obtained from the optimization process. For lower and upper this is the lower and upper limit values in the optimization process. Popsize is a population that is raised, this study generates a population of 60 with a maximum iteration or repetition of 50 times. To avoid damage or loss of population, it is better to do iterations between 50-200 times, this is useful so that the population is not lost or damaged during iterations. After the optimization and iteration process is complete, the results obtained are shown in Table 1 below.

Table 1. Genetic Algorithm (GA) Optimization Results

GA settings					**GA results**		
Population size	Number of generations	Elitism	Crossover probability	Mutation probability	Iterations	Solution	
						X_1	**X_2**
60	50	3	0.8	0.1	**50**	**2.097333**	**1.00302**

In Table 1, it is shown that, from the iteration process 50 times, the solution values of X_1 = 2.097333 and X_2 = 1.00302 are obtained. For elitism, the best number of individuals or chromosomes that can survive in each generation or iteration. The results showed that the chromosomes that could survive 50 iterations were 3 chromosomes. After the optimization and iteration process is complete, the next plot of results is shown in Figure 4 below.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11431
e-ISSN : 2541-2019
p-ISSN : 2541-044X

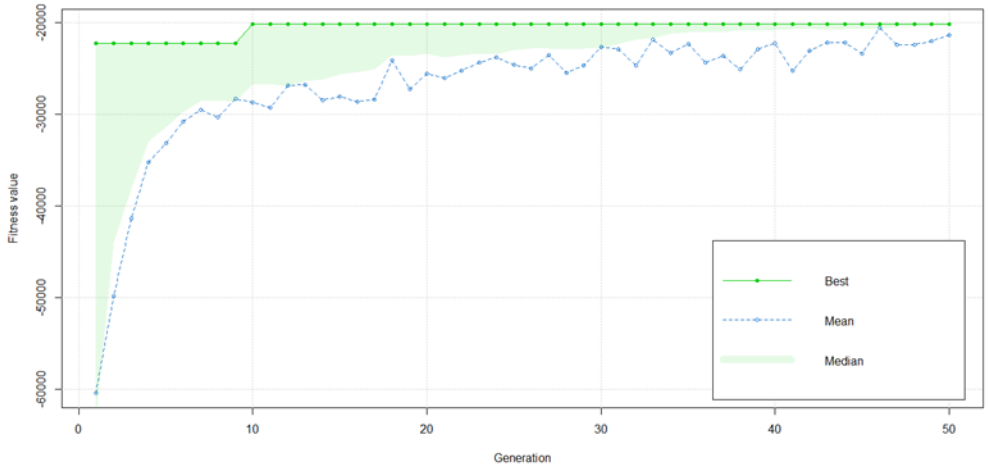

Fig 4. Plot Optimization Results

DISCUSSIONS

The results of predictions with optimized values with non-optimized values have negative and positive values, this has no effect because what is seen is the difference between the real value of income and the resulting prediction value, the smaller the difference, the better the predictions that can be displayed. From the testing process, it can be seen that the Model Summary listed is the result of data processing with Genetic Algorithms. So that the final result of this processing can be better and can be re-implemented with the hope that the value of the standard error estimate is smaller than the data that is not optimized. By using the Genetic Algorithm method for optimizing the value of income so that the prediction results can be better.

The research method used in this study focuses on the process of predicting the value of income, this prediction is needed to estimate the number of products to be marketed according to the specified variables, so as to provide maximum results in the next period. In addition, in this study, the Genetic Algorithm method was also used to optimize the supporting variables in order to provide a better solution or result with a smaller Standard Error value than the value that was not optimized. The results of these methods can be used as a reference for PT. XYZ to continue to make improvements to the products marketed so that in the next period it can generate better income and can reduce the value of the difference between the predicted results and the real income value.

CONCLUSION

Based on the research formulation, the comparison of the resulting error values between the values that are optimized and those that are not, can be implemented properly. For the predicted value in January 2021, the value that was not optimized was 316,388.8 with a difference or error value of 35,498.8 and the optimized value obtained a predicted value of 313,254.9 with a difference or error value of 32,364.9. This shows that the optimization has been successfully carried out and the results of this optimization can affect the error value with a smaller difference with the real value.

REFERENCES

Abd Elrehim, M. Z., Eid, M. A., & Sayed, M. G. (2019). Structural optimization of concrete arch bridges using Genetic Algorithms. *Ain Shams Engineering Journal*, *10*(3), 507–516. https://doi.org/https://doi.org/10.1016/j.asej.2019.01.005

Al Rivan, M. E., Steven, S., & Tanzil, W. (2020). Optimasi Fuzzy C-Means dan K-Means Menggunakan Algoritma Genetika untuk Pengklasteran Dataset Diabetic Retinopathy. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *7*(5), 993. https://doi.org/10.25126/jtiik.2020711872

Alfarizi, M. G., Stanko, M., & Bikmukhametov, T. (2022). Well control optimization in waterflooding using genetic algorithm coupled with Artificial Neural Networks. *Upstream Oil and Gas Technology*, *9*(March), 100071. https://doi.org/10.1016/j.upstre.2022.100071

Ermatita. (2009). Analisis Optimasi Query Pada Data Mining. *Jurnal Sistem Informasi (JSI)*, *1*(1), 47–54. Retrieved from http://ejournal.unsri.ac.id/index.php/jsi/index

Garcia, L. G., & Lôndero, V. (2021). A parameter optimizer based on genetic algorithm for the simulation of carbonate facies. *Intelligent Systems with Applications*, *12*, 200057.

*name of corresponding author

https://doi.org/10.1016/j.iswa.2021.200057

Istianto, Y., & 'Uyun, S. (2021). Klasifikasi Kebutuhan Jumlah Produk Makanan Customer menggunakan K-Means Clustering dengan Optimasi Pusat Awal Cluster Algoritma Genetika. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *8*(5), 861. https://doi.org/10.25126/jtiik.202182990

Khan, A. H., Hossain, S., Hasan, M., Islam, M. S., Rahman, M. M., & Kim, J. H. (2022). Development of an optimized thermodynamic model for VVER-1200 reactor-based nuclear power plants using genetic algorithm. *Alexandria Engineering Journal*, *61*(11), 9129–9148. https://doi.org/10.1016/j.aej.2022.02.052

Khotimah, B. K., Syarief, M., Miswanto, & Suprajitno, H. (2021). Optimasi Bobot K-Means Clustering Untuk Mengatasi Optimization Weight of K-Means Clustering. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *8*(4), 745. https://doi.org/10.25126/jtiik.202184912

Luo, X., Qian, Q., & Fu, Y. F. (2020). Improved Genetic Algorithm for Solving Flexible Job Shop Scheduling Problem. *Procedia Computer Science*, *166*, 480–485. https://doi.org/https://doi.org/10.1016/j.procs.2020.02.061

Martins, T. M., & Neves, R. F. (2020). Applying genetic algorithms with speciation for optimization of grid template pattern detection in financial markets. *Expert Systems with Applications*, *147*, 113191. https://doi.org/https://doi.org/10.1016/j.eswa.2020.113191

Primadani, L., Palgunadi, Y., & Harjito, B. (2015). Optimasi Produksi Menggunakan Algoritma Fuzzy Linear Programming (Studi Kasus : Produksi Tas UKM Cantik Souvenir). *Jurnal Teknologi & Informasi ITSmart*, *4*(2), 63. https://doi.org/10.20961/its.v4i2.1764

Sari, Y., Alkaff, M., Wijaya, E. S., Soraya, S., & Kartikasari, D. P. (2019). Optimasi Penjadwalan Mata Kuliah Menggunakan Metode Algoritma Genetika dengan Teknik Tournament Selection. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *6*(1), 85. https://doi.org/10.25126/jtiik.2019611262

Shehadeh, A., Alshboul, O., Tatari, O., Alzubaidi, M. A., & Hamed El-Sayed Salama, A. (2022). Selection of heavy machinery for earthwork activities: A multi-objective optimization approach using a genetic algorithm. *Alexandria Engineering Journal*, *61*(10), 7555–7569. https://doi.org/10.1016/j.aej.2022.01.010

Song, H., Cai, M., Cen, J., Xu, C., & Zeng, Q. (2022). Research on energy saving optimization method of electric refrigerated truck based on genetic algorithm. *International Journal of Refrigeration*, *137*(June 2021), 62–69. https://doi.org/10.1016/j.ijrefrig.2022.02.003

Wicaksono, S. A. (2019). Optimasi Sistem Penempatan Magang Menerapkan Algoritme Genetika. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *6*(1), 17. https://doi.org/10.25126/jtiik.201961950

Zhang, K., Ma, H., Li, Q., Wang, D., Song, Q., Wang, X., & Kong, X. (2022). Thermodynamic analysis and optimization of variable effect absorption refrigeration system using multi-island genetic algorithm. *Energy Reports*, *8*, 5443–5454. https://doi.org/10.1016/j.egyr.2022.04.004

*name of corresponding author