

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem using Regression Machine Learning: Post-Experiment Analysis

Ivan Kristianto Singgih

Study Program of Industrial Engineering, University of Surabaya, Surabaya, Indonesia
Institut Supply Chain dan Logistik Indonesia, Indonesia
Asosiasi Peneliti Indonesia Korea (APIK), Republic of Korea
Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial (KORIKA), Jakarta, Indonesia
Indonesia Artificial Intelligence Society (IAIS)

Aditya Rio Prabowo

Department of Mechanical Engineering, Universitas Sebelas Maret, Surakarta 57126, Indonesia

Prita Meilanitasari*

Department of Logistics Engineering, Universitas Internasional Semen Indonesia, Gresik 61122, Indonesia

(Received: November 11, 2025 / Revised: February 11, 2026; April 18, 2026 / Accepted: April 27, 2026)

ABSTRACT

Traveling salesman problem (TSP) is an important problem because it is the base of many real-world transportation problems. One of the most effective and frequently used optimization methods for solving TSP is the genetic algorithm. Generally, genetic algorithms employ two types of operators: crossover and mutation. Most researchers apply the crossover operator to genetic algorithms to find better solutions, without analyzing the improvement in solution quality each time the crossover is applied. Such an approach is commonly used with the aim of simply finding the best route, for example, by comparing the quality of final solutions produced by the genetic algorithm and other methods. Unlike previous studies, this study examines the characteristics of the updated solution within the genetic algorithm run, which is modified by crossover, specifically, the improvement in solution quality. To the best knowledge of the authors, this study is the first to evaluate the effectiveness of crossover operators using regression machine learning methods. The numerical experiments show that the most important feature to predict the change in the best objective values is the distance between customers outside the cutting points, followed by the distance between customers within the cutting points, then the number of customers (outside then between the cutting points). This study provides insights for researchers to develop more effective solution methods.

Keywords: Travelling Salesman Problem, Genetic Algorithm, Regression Machine Learning, Crossover

* Corresponding Author, E-mail: prita.meilanitasari@uisi.ac.id

1. INTRODUCTION

To solve real optimization problems to ensure business profitability, e.g., travelling salesman problem (TSP) and vehicle routing problem, various optimization (operations research) methods are developed. Some methods guarantee the optimality of the solutions (e.g., exact methods (Riaventin *et al.*, 2021), while some other methods show that even with a much shorter computational time, close-to-optimal solutions could still be produced (e.g., metaheuristics (Meilanitasari and Shin, 2021), algorithms (Singgih and Kim, 2020)). One of the most used metaheuristics for solving routing problem is the genetic algorithm (GA) (Garside *et al.*, 2024; Ojeda Rios and Xavier, 2025; Ostermeier *et al.*, 2021; Shen *et al.*, 2022).

Most studies have a purpose to either produce the best solution quality that is usually defined as the best-known solution (regardless of the computational time) or reduce the computational time while still producing the best possible solution quality within the given time (Lee *et al.*, 2024; Xie *et al.*, 2025; Zhang and Woensel, 2023). Such an approach somehow fails to evaluate how well each part of the algorithm performs within the algorithm run because the method ignores how well the solution quality is improved after any operator is applied. Most studies focus on assessing their proposed method by observing the quality of the final solution and contrasting it with state-of-the-art methods. Some studies would then try to perform good crossover operations, e.g., undirected bipartite graph-based crossover operator (Nazif and Lee, 2012), adjusted order crossover that ensured copying more related nodes into the new sequence (Hvattum, 2022), an updated version of the well-known partially mapped crossover operator (Koohestani, 2020), two-layer chromosome-based crossover that captured more amount of necessary information simultaneously (Karakatič, 2021), adaptive crossover rate that adjusted the number of applied crossover operators based on the current quality of the solutions (Li *et al.*, 2024), modified crossover operators that applied arithmetic operations to allow better convergence (Naskar *et al.*, 2025), and multiple crossover operators which would aim to benefit from the advantage of each type of operator (Zhang *et al.*, 2025). It was common for previous studies to test several crossover operators and identify the ones producing the best solutions after the whole GA run, e.g., Kaelo and Ali (2007), Kaya (2011), Li *et al.* (2023), and Derouiche *et al.* (2025).

All previous studies above still do not answer fundamental questions about (1) why the crossover operator performs well, (2) how well the crossover operator performs, and (3) what values of parameters must be set for the crossover operators to perform well. The studies above only focused on attempting to produce good crossover operators that might work well and proved their effectiveness after a complete run. Without answering

such questions, it would be difficult to understand why and how crossover operators perform well in improving the solution quality, e.g., in a GA run, and to produce an effective operator.

The purpose of this study is to provide answers to the research questions above. As the initial study, this study focuses on the TSP, which has no vehicle capacity restriction and allows an easy understanding of the relationships between the nodes in the chromosome (solution representation). To the best knowledge of the authors, this study is the first to evaluate the effectiveness of GA's crossover operators using regression machine learning methods. This study is important to ensure understanding important parts of the crossover operators that must be considered for designing good operators.

When contrasted with statistical correlation analysis that only shows the relationship between features, implementing regression machine learning methods together with operations research approaches enables online learning, in which the operators in the optimization methods could be assisted by machine learning-based prediction to ensure obtaining better solutions (Szénási and Légrádi, 2024). The inclusion of machine learning methods to apply better operators in the GA initiates a whole new research field that could properly deal with the high uncertainty during the search process. Such an effort to utilize machine learning techniques to improve the quality of metaheuristics has been attempted by recent studies in several fields, e.g., autonomous ridesharing (Bongiovanni *et al.*, 2022), distribution system design (Xiao *et al.*, 2022), and global optimization for engineering problems (Farahmand-Tabar and Ashtari, 2024).

In this study, important parameters of the crossover operator are defined. Every time a crossover operator is applied, all the listed parameter values are recorded together with the amount of improvement in the objective value (the difference between the best new objective value and the best objective value before the crossover operation was applied). Those collected input data are then used to predict the target data by using regression machine learning models. As the result, this study successfully identifies the proper values of the parameters, which would improve the objective values the most.

The structure of this study is as follows. Section 2 describes the method that is solved using GA and proposes the prediction framework using regression machine learning models. Section 3 presents the numerical experiment results. Finally, Section 4 concludes the study.

2. METHOD

To evaluate the performance of crossover operators, this study discusses a Traveling Salesman Problem (TSP) case study. It focuses on route determination and consid-

ers the number of demands that must be met, while ignoring vehicle capacity. The suggested Crossover Analysis framework model is shown in Figure 1. The crossover and mutation operators are applied within the GA. Then, the crossover operator is assessed. Data analytics uses feature collection for analysis and regression machine learning-based crossover operator evaluation.

2.1 Genetic Algorithm (GA) for Traveling Salesman Problem (TSP)

The GA procedure used to solve the TSP is illustrated in Figure 2. The route starts from a depot and ends

at the same depot. In the considered TSP, the total distance is minimized, e.g., as shown in Figure 2. Initially, chromosomes (as many as the *population_size*) are generated randomly. To produce potentially better new solutions, crossover and mutation operators are applied, and the GA procedure is continued for *num_of_population* iterations (populations).

At the start of each population, n solutions with the best objective values are simply selected and stored as candidates for crossover operations. In the next step, crossover operators are applied by considering the initially selected n best solutions as parent chromosomes. Considering that chromosomes with least total distance must be given a

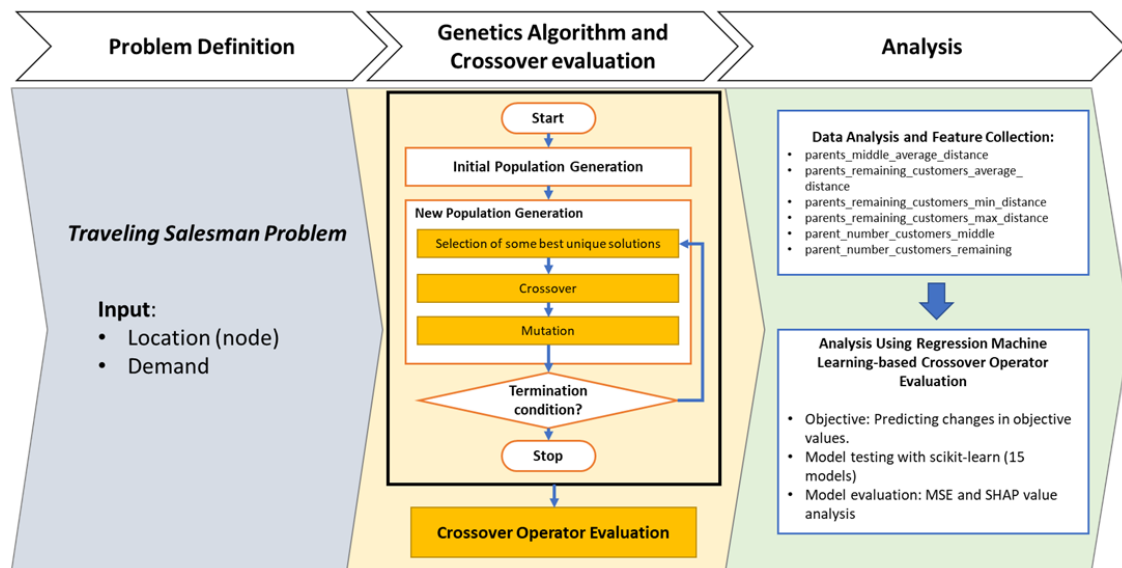


Figure 1. A framework of the crossover analysis research architecture.

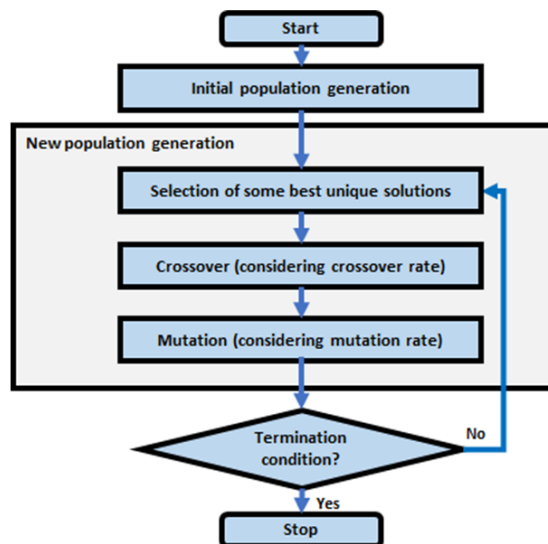


Figure 2. TSP-GA model framework.

higher selection probability, an *inverted_objective_value* for each parent chromosome is calculated in Equation (1), then the value is used to calculate the selection probability for that chromosome, as calculated in Equation (2).

$$iov = \frac{1}{ov} \quad (1)$$

$$sp = \frac{iov}{\text{total } iov \text{ of all best chromosomes}} \quad (2)$$

where:

iov : *inverted_objective_value*
ov : *objective_value*
sp : *selection_probability*

A random number between 0 and 1 is generated. If its value is less than the crossover rate, the crossover operation is applied. Otherwise, the parent chromosomes are set as the result of the crossover operation (stored as the new child chromosomes). If the two-point crossover operator is applied, two cutting points are randomly selected. The parts of the parents within the crossover points are copied to the child chromosomes. The remaining unfilled parts of the new child chromosomes are then obtained from other parent chromosomes (Figure 3). The crossover operator is applied until the number of chromosomes in the current population equals to *population_size*.

The mutation operator is applied when the random number is less than the mutation rate. The mutation is performed by randomly selecting two values within a parent chromosome and exchanging the positions of both values to produce a child chromosome (Figure 3).

2.2 Regression Machine Learning-based Crossover Operator Evaluation

The objective value of the TSP-GA (the new best total distance minus the previous best total distance) is predicted in this study using regression machine learning

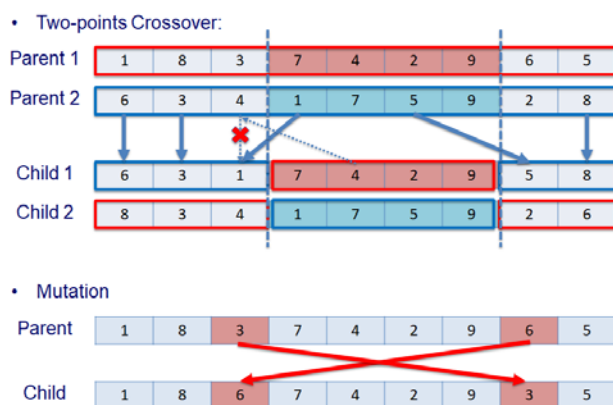


Figure 3. Crossover and mutation operators for this study.

models.

The input data represent the characteristics of the solution before the crossover is performed. Each input data is presented as follows.

- *parents_middle_average_distance_value*: the average distance of subsequent customers within two cutting points in the two-point crossover;
- *parents_middle_average_distance_percentage*: the value of *parents_middle_average_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_average_distance_value*: the average distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_average_distance_percentage*: the value of *parents_remaining_customers_average_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_min_distance_value*: the minimum distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_min_distance_percentage*: the value of *parents_remaining_customers_min_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_max_distance_value*: the maximum distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_max_distance_percentage*: the value of *parents_remaining_customers_max_distance_value* divided by the length of the area for the data set;
- *parent_number_customers_middle_value*: the number of customers in the cutting area of the chromosome;
- *parent_number_customers_middle_percentage*: the value of *parent_number_customers_middle_value* divided by the number of customer nodes in the data set;
- *parent_number_customers_remaining_value*: the number of customers outside the cutting area in the chromosome.
- *parent_number_customers_remaining_percentage*: the value of *parent_number_customers_remaining_value* divided by the number of customer nodes in the data set.

When different datasets are considered, the number of customers is added as an input feature. The predicted output is the total distance of the route.

The considered features are expected to highly affect the quality of the two-points crossover. Two important parts of the chromosomes are the part in the middle of both cutting points and the remaining area outside the area between the cutting points. The main idea of the crossover is to maintain the good part from parent chromosomes, which refers to the middle part between the cutting points. It is expected that through the random cutting and combination process in the crossover, the good middle parts would remain in the child chromosomes and eventually lead the search to a promising direction. Another way to evaluate the crossover would be understanding the effect of the distances between nodes outside the middle part of the chromosome on the objective value. A good crossover would select middle parts with good partial objective values, while having outer parts with bad partial objective values. In contrast, the crossover would produce worse results when the middle parts contain bad sequences, but have the good outer parts (with short travel times). The number of customers in each part could have a high impact on the crossover results as well. All these related features are considered in this study.

Several regression machine learning models are tested, and then the best models are reported in this section. The predictions were developed with the following scikit-learn (Scikit-learn, 2022) regression machine learning models such as: (1) random forest regression (Kim *et al.*, 2020), (2) linear regression (Abid Almubaidin *et al.*, 2023), (3) Ridge (Scikit-learn, 2023; Wang *et al.*, 2023), (4) ElasticNet (Scikit-learn, 2023; Yu *et al.*, 2021), (5) Lars (Scikit-learn, 2023; Lee and Jun, 2018), (6) LassoCV (Scikit-learn, 2023; Kramlinger *et al.*, 2023), (7) LassoLars (Scikit-learn, 2023), (8) OrthogonalMatchingPursuit (Scikit-learn, 2023; Polat and Kayhan, 2018), (9) ARDRegression (Sandhu *et al.*, 2017), (10) BayesianRidge (da Silva *et al.*, 2021), (11) HuberRegressor (Scikit-learn, 2023), (12) RANSACRegressor (Scikit-learn, 2023; Zhu *et al.*, 2023), (13) TheilSe-

nRegressor (Scikit-learn, 2023), (14) TweedieRegressor (Scikit-learn, 2023), (15) PassiveAggressiveRegressor (Scikit-learn, 2023; Jorge and Paredes, 2018), ExtraTreeRegressor (Ahmad *et al.*, 2018), GradientBoostingRegressor (Moghadaspoor *et al.*, 2026), HistGradientBoostingRegressor (Alawi *et al.*, 2024), AdaBoostRegressor (Kumari *et al.*, 2025), support vector regressor with radial basis function (Taiwo *et al.*, 2024), KNeighborsRegressor, DecisionTreeRegressor (Zade *et al.*, 2023), MLPRegressor (Zade *et al.*, 2023), and XGBoostRegressor (Moghadaspoor *et al.*, 2026). The linear and non-linear methods are implemented to properly find the best method for the data sets. At the time of writing, scikit-learn has more than 120,000 citations based on Google Scholar. Its minimal dependencies and ease of use allow a high reproducibility rate in many studies.

Data is generated as follows. The GA is run following the procedure explained in Section 2. As a result, 1,781 data records are recorded and used for the machine learning analysis. In the analysis, the models are compared based on the mean squared error (MSE), and the SHAP value analysis (SHAP, 2025a, 2025b) is conducted to identify the most important input feature for the prediction.

3. NUMERICAL EXPERIMENTS

Different data set sizes are generated randomly, as shown in Table 1. To produce the data, the parameters for the GA are listed in Table 2. Experiments using the machine learning models are conducted separately for each data set size, then for the whole data sets combined together to allow generalizing the findings. In the experiments 10-fold cross validation and MinMaxScaler are applied. Performance metrics obtained when dealing with data set 1, 2, 3, and all of them are presented in Tables 3, 4, 5, and 6, respectively.

Table 1. Data set sizes for the numerical experiments

Data set	Number of nodes	Area on which the nodes are placed	Number of instances
1	20	20 x 20	538
2	200	100 x 100	614
3	1,000	400 x 400	629

Table 2. Parameters of the genetic algorithm

Parameter	Value
Number of populations	20
Population size	100
Number of initial chromosomes	20% * population size
Crossover rate	0.8
Mutation rate	0.2

Table 3. Performance metrics for regression machine learning models on data set 1

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	2.22	6.07	8.36	59.32	2.89	7.70	87.70	11.81
LinearRegression	6.61	6.71	64.05	66.23	8.00	8.14	5.71	0.93
Ridge	6.62	6.71	64.08	65.99	8.01	8.12	5.67	1.28
ElasticNet	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
Lars	6.89	7.01	72.71	75.80	8.53	8.71	-6.76	-17.60
Lasso	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
LassoLars	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
OrthogonalMatchingPursuit	6.80	6.82	66.95	67.34	8.18	8.21	1.43	-1.03
ARDRegression	6.64	6.75	64.24	66.51	8.01	8.16	5.43	0.45
BayesianRidge	6.64	6.72	64.20	65.94	8.01	8.12	5.49	1.34
HuberRegressor	6.53	6.65	64.92	67.36	8.06	8.21	4.43	-0.34
RANSACRegressor	7.18	7.34	93.21	92.68	9.65	9.63	-37.24	-37.25
TheilSenRegressor	6.61	6.72	64.21	66.51	8.01	8.16	5.47	0.52
TweedieRegressor	6.86	6.88	67.64	68.05	8.22	8.25	0.43	-2.29
PassiveAggressiveRegressor	7.52	7.54	86.42	87.89	9.30	9.38	-27.27	-31.09
ExtraTreesRegressor	0.10	6.15	0.44	64.45	0.66	8.03	99.36	3.56
GradientBoostingRegressor	4.13	6.01	26.85	59.85	5.18	7.74	60.46	10.96
HistGradientBoostingRegressor	3.22	6.23	16.94	63.47	4.12	7.97	75.06	4.56
AdaBoostRegressor	6.27	6.64	52.85	62.90	7.27	7.93	22.19	4.86
SVR_RBF	6.53	6.65	67.31	69.20	8.20	8.32	0.90	-2.97
KNeighborsRegressor	5.22	6.72	44.64	72.82	6.68	8.53	34.28	-7.81
DecisionTreeRegressor	0.10	7.24	0.44	100.37	0.66	10.02	99.36	-52.27
MLPRegressor	6.61	6.74	62.75	65.68	7.92	8.10	7.61	1.42
XGBRegressor	1.44	6.16	3.85	63.97	1.96	8.00	94.34	4.63

Table 4. Performance metrics for regression machine learning models on data set 2

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	23.5	65.8	993.2	7991.8	31.5	89.4	86.8	-9.9
LinearRegression	62.0	62.9	6907.5	7117.4	83.1	84.4	8.0	-1.7
Ridge	62.0	62.8	6910.8	7103.4	83.1	84.3	8.0	-1.3
ElasticNet	63.8	64.0	7357.3	7418.9	85.8	86.1	2.1	-3.0
Lars	62.2	62.5	6962.6	7070.0	83.4	84.1	7.3	-0.6
Lasso	63.1	63.5	7175.6	7304.2	84.7	85.5	4.5	-2.7
LassoLars	63.1	63.5	7187.4	7306.6	84.8	85.5	4.3	-2.7
OrthogonalMatchingPursuit	63.4	63.8	7212.2	7327.4	84.9	85.6	4.0	-4.3
ARDRegression	62.0	62.7	6915.5	7085.3	83.2	84.2	7.9	-0.8
BayesianRidge	62.0	62.7	6919.2	7109.5	83.2	84.3	7.9	-1.1
HuberRegressor	61.4	62.4	7006.7	7194.8	83.7	84.8	6.7	-1.1
RANSACRegressor	64.6	64.5	8121.7	7792.3	90.1	88.3	-7.8	-11.3
TheilSenRegressor	61.9	62.7	6925.4	7125.9	83.2	84.4	7.8	-1.3
TweedieRegressor	64.0	64.2	7400.0	7451.5	86.0	86.3	1.5	-3.3
PassiveAggressiveRegressor	62.4	62.5	7161.7	7072.6	84.6	84.1	4.6	-0.1
ExtraTreesRegressor	0.0	67.8	0.0	8514.4	0.0	92.3	100.0	-18.2
GradientBoostingRegressor	41.1	66.3	2975.9	8659.1	54.6	93.1	60.4	-12.3
HistGradientBoostingRegressor	33.1	67.1	1973.8	8656.5	44.4	93.0	73.7	-18.3
AdaBoostRegressor	60.2	66.6	5732.3	7998.3	75.7	89.4	23.6	-13.0
SVR_RBF	62.2	62.7	7475.6	7564.9	86.5	87.0	0.5	-3.2
KNeighborsRegressor	53.3	66.7	5406.8	8358.1	73.5	91.4	28.0	-17.6
DecisionTreeRegressor	0.0	91.2	0.0	16358.5	0.0	127.9	100.0	-121.2
MLPRegressor	61.5	62.9	6794.3	7167.0	82.4	84.7	9.5	-2.9
XGBRegressor	12.9	67.4	319.7	8830.6	17.9	94.0	95.7	-18.7

Table 5. Performance metrics for regression machine learning models on data set 3

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	249.11	721.31	151441.08	1275915.07	389.15	1129.56	88.56	-26.38
LinearRegression	742.27	760.60	1200368.72	1257147.92	1095.61	1121.23	9.51	-53.95
Ridge	741.11	759.49	1202529.30	1256671.97	1096.60	1121.01	9.35	-53.51
ElasticNet	777.36	792.73	1301802.23	1346299.48	1140.97	1160.30	1.88	-65.68
Lars	744.94	767.94	1208269.22	1276741.75	1099.21	1129.93	8.89	-60.23
Lasso	742.06	761.89	1201521.10	1260090.02	1096.14	1122.54	9.42	-54.45
LassoLars	744.95	768.14	1209326.87	1276120.10	1099.69	1129.65	8.80	-59.46
OrthogonalMatchingPursuit	763.86	780.11	1280113.67	1330966.94	1131.42	1153.68	3.51	-77.06
ARDRegression	743.86	764.21	1203485.08	1264248.95	1097.03	1124.39	9.28	-54.98
BayesianRidge	741.14	759.96	1201820.72	1257419.14	1096.28	1121.35	9.40	-53.85
HuberRegressor	699.27	714.46	1276669.46	1318829.37	1129.90	1148.40	3.75	-20.11
RANSACRegressor	773.51	776.35	1784477.92	1567055.92	1335.84	1251.82	-33.80	-84.10
TheilSenRegressor	43 (10 ⁹)	41 (10 ⁹)	19 (10 ²¹)	18 (10 ²¹)	141 (10 ⁹)	135 (10 ⁹)	-157 (10 ¹⁶)	-423 (10 ¹⁶)
TweedieRegressor	781.42	796.47	1311187.56	1354636.92	1145.07	1163.89	1.18	-66.00
PassiveAggressiveRegressor	707.41	714.61	1380041.85	1394199.76	1174.75	1180.76	-3.91	-18.73
ExtraTreesRegressor	0.00	757.95	0.00	1437740.56	0.00	1199.06	100.00	-36.36
GradientBoostingRegressor	435.34	751.54	380870.44	1421742.25	617.15	1192.37	71.29	-48.96
HistGradientBoostingRegressor	346.25	765.08	273954.94	1426434.43	523.41	1194.33	79.32	-55.71
AdaBoostRegressor	734.09	866.56	878118.11	1517805.48	937.08	1231.99	34.10	-130.99
SVR_RBF	717.11	721.31	1434919.21	1446023.09	1197.88	1202.51	-8.06	-16.98
KNeighborsRegressor	590.70	733.18	827013.10	1258896.83	909.40	1122.01	37.64	-54.87
DecisionTreeRegressor	0.00	972.59	0.00	2350582.54	0.00	1533.16	100.00	-250.05
MLPRegressor	734.49	760.41	1161207.30	1230558.99	1077.59	1109.31	12.42	-75.35
XGBRegressor	112.23	763.05	25508.40	1448972.31	159.71	1203.73	98.08	-50.74

Table 6. Performance metrics for regression machine learning models on all data sets

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	98.9	291.7	55842.7	526353.5	236.3	725.5	89.7	-13.8
LinearRegression	310.9	339.9	434600.8	486185.8	659.2	697.3	20.1	-6826.6
Ridge	297.5	314.4	442188.9	486585.7	665.0	697.6	18.7	-485.5
ElasticNet	321.8	332.2	481246.4	508580.2	693.7	713.1	11.6	-2692.2
Lars	338494.0	416814.2	221 (10 ¹⁰)	420 (10 ¹⁰)	148 (10 ⁴)	205 (10 ⁴)	-399 (10 ⁶)	-26 (10 ⁸)
Lasso	298.2	314.7	443031.9	488952.9	665.6	699.3	18.5	-835.1
LassoLars	298.1	314.7	443044.7	489015.9	665.6	699.3	18.5	-832.4
OrthogonalMatchingPursuit	312.6	329.6	471069.2	507923.4	686.3	712.7	13.5	-680.1
ARDRegression	297.3	316.3	437940.2	485443.4	661.8	696.7	19.5	-283.5
BayesianRidge	297.6	317.0	437709.4	485186.7	661.6	696.6	19.5	-489.3
HuberRegressor	272.7	281.7	483311.0	505395.3	695.2	710.9	11.2	-43.3
RANSACRegressor	498.2	773.1	1085469.1	1722986.0	1041.9	1312.6	-99.8	-26899.8
TheilSenRegressor	302.4	333.8	442784.3	487451.3	665.4	698.2	18.6	-13253.4
TweedieRegressor	338.1	349.0	495021.2	520156.3	703.6	721.2	9.1	-7114.8
PassiveAggressiveRegressor	275.0	281.7	497727.7	518314.4	705.5	719.9	8.6	-17.3
ExtraTreesRegressor	0.0	297.0	0.1	541539.2	0.4	735.9	100.0	-17.6
GradientBoostingRegressor	183.4	314.7	146422.2	632543.6	382.7	795.3	73.4	-128.5
HistGradientBoostingRegressor	146.1	290.1	107817.6	500860.3	328.4	707.7	79.8	-42.4
AdaBoostRegressor	304.7	364.4	322521.8	587055.0	567.9	766.2	41.7	-11379.8
SVR_RBF	280.7	282.7	542527.7	551146.8	736.6	742.4	0.5	-22.3
KNeighborsRegressor	233.1	298.3	290406.1	486843.6	538.9	697.7	46.4	-23.1
DecisionTreeRegressor	0.0	389.0	0.1	897974.0	0.4	947.6	100.0	-112.6
MLPRegressor	306.6	337.2	451621.1	519803.3	672.0	721.0	16.9	-239.8
XGBRegressor	58.8	311.2	10420.8	608625.1	102.1	780.1	98.1	-41.0

To initially assess the experiment results, the R^2 score obtained using the test data in each data set is observed. The most promising R^2 score using the test data (11.81%) is obtained in data set 1, in which 20 customer nodes are considered (Table 3), when the RandomForestRegressor is applied. When the value is contrasted with the R^2 score for the train data, the large gap between the scores shows a high overfitting of the model. However, a positive R^2 score shows a promising research result, considering this study is the first to attempt observing the quality of the crossover operators in the GA. A low R^2 score is acceptable in some fields, e.g., social science and psychology, which is very affected by the complexity of the data (Gupta *et al.*, 2024). Considering that R^2 score refers to the degree to which the variation of the target variable is influenced by the variation in the input features, the metric is strongly influenced by the variation in the input features. However, the low R^2 score does not indicate a weak relationship between the variables (Hamilton *et al.*, 2015). Given that the attempt to understand new input features of the crossover operator in the GA is a completely new problem, a lot of effort might be necessary for studies to understand the relationships between variables properly. To ensure a better prediction, it is ne-

cessary to (1) identify more related input features, which might be specifically designed to understand hidden relationships that better predict the total distance, and (2) group the data based on specific characteristics to allow better prediction using more similar input features. Correlations between variables when using this data set are presented in Figure 4.

When the predictions are conducted in other data sets (Tables 4 and 5), the models fail to find good predictions, which is reflected by the negative R^2 values, which represent poor prediction capabilities (Zhang *et al.*, 2026). The negative R^2 values appear when the models fail to capture the nonlinear behavior of the data or fail to explain the data variations in reality (Mittlböck and Waldhör, 2000) because the performance of the prediction models is worse than when the average values of the target variables are considered (Chicco *et al.*, 2021). A similar result is obtained when all data sets are considered simultaneously (Table 6), that concludes that the generalization could not yet be conducted properly and needs further observations.

Figure 5 shows the SHAP analysis result. In Figure 5, the low value of each feature is represented in blue, and its high value is represented in red. The negative values in

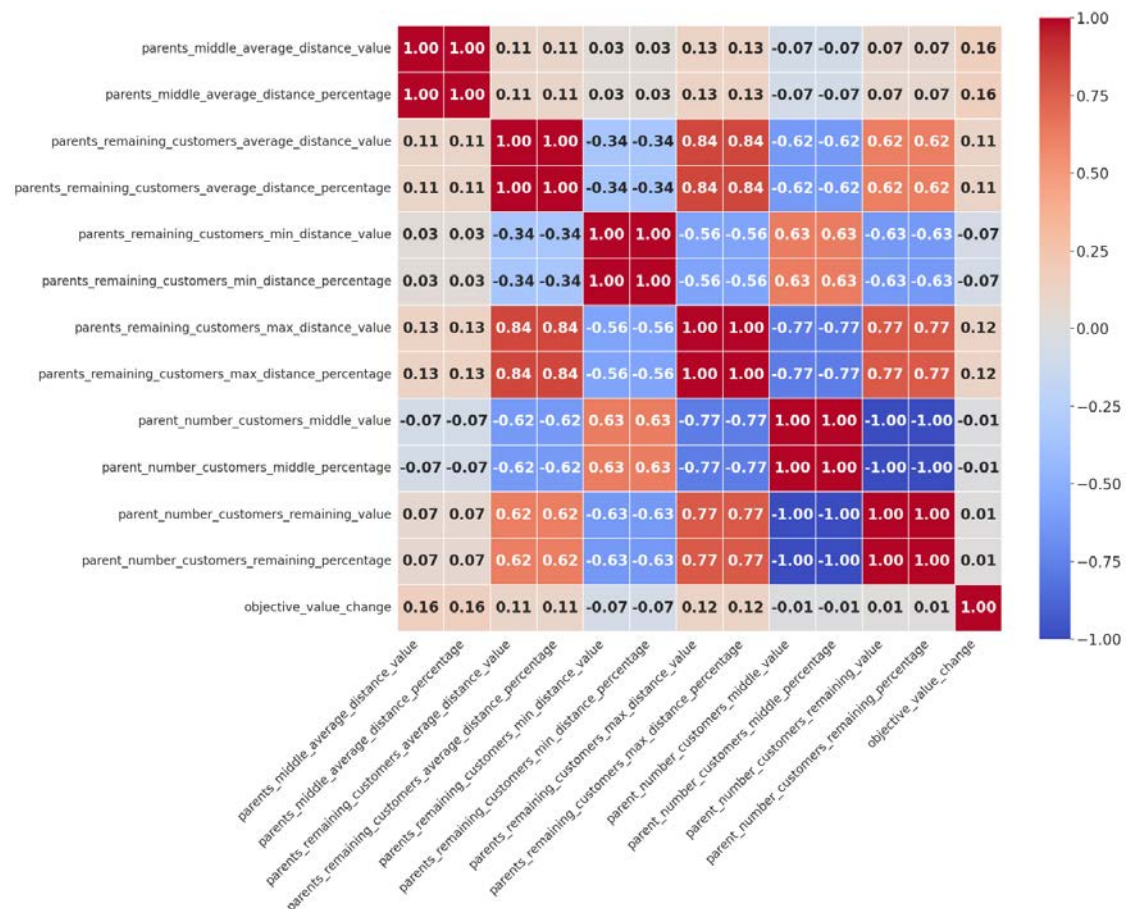


Figure 4. Correlations between variables.

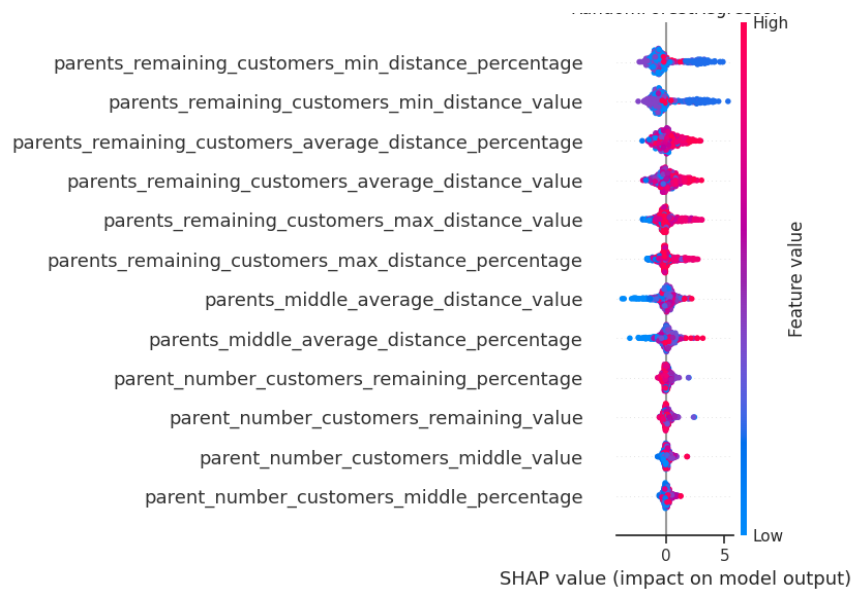


Figure 5. Feature importance analysis using SHAP value.

the x-axis refer to the reduction in prediction quality (the increase in MSE), while the positive values refer to the increase in prediction quality (the decrease in MSE). Based on Figure 5, it is shown that the most important input features to predict the change in the total distance are (1) minimum, average, then maximum distances between each pair of remaining customers outside the cutting points, (2) average distance between customers within the cutting points, then (3) number of customers (outside then between the cutting points).

The prioritized features are further explained as follows. First, the distance between each pair of remaining customers is important because its value shows the opportunity to improve outside parts of the new chromosomes, e.g., when customers with short distances are placed subsequently. Such a conclusion aligns with the order of importance for the minimum, average, then maximum values. Second, the total distance measured between subsequent customers within the cutting points affects the quality of the new chromosomes by ensuring that subsequent customers with short distances are maintained in the next chromosomes. The findings are interesting. The remaining customers feature that is more important than the customers within the cutting points indicates that when applying the crossover operators, it is necessary to sufficiently consider including customers with short distances into the area between the cutting points. Third, the number of customers (either inside or outside the cutting points) has the least impact when predicting the total distances. Such an observation is acceptable considering that the distance measure itself should be observed better than the number of customers because the objective (total distances) is strongly related to the distance values. As a conclusion, such important features must be considered

properly when applying the two-point crossover operators in the GA to produce generally good solutions in the long run, while reducing the number of required populations.

4. CONCLUSIONS

This study observes the potential of understanding the behavior of the crossover operator in the genetic algorithm for solving the travelling salesman problem. The numerical experiments show that the distance between customers outside the cutting points of the chromosome is the most important feature to predict the change in the best objective values before and after the crossover operator is applied, followed by the distance between customers within the cutting points, then the number of customers. The limitation of this study is only considering the two-point crossover operator. For further research topics, it would be necessary to consider more variants of crossover operators and more objectives, design more features that could significantly affect the prediction quality, design better regression machine learning methods with better generalization (e.g., by introducing more features related to the problem sizes), and discuss how effective genetic algorithms can be designed based on the insights found in this study.

REFERENCES

Abid Almubaidin, M. A., Latif, S. D., Balan, K., Ahmed, A. N., and El-Shafie, A. (2023), Enhancing sediment transport predictions through machine learning-based multi-scenario regression models, *Results in*

- Engineering*, **20**, 101585.
- Ahmad, M. W., Reynolds, J., and Rezgui, Y. (2018), Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees, *Journal of Cleaner Production*, **203**, 810-821.
- Alawi, O. A., Kamar, H. M., Salih, S. Q., Abba, S. I., Ahmed, W., Homod, R. Z., Jamei, M., Shafik, S. S., and Yaseen, Z. M. (2024), Development of optimized machine learning models for predicting flat plate solar collectors thermal efficiency associated with Al₂O₃-water nanofluids, *Engineering Applications of Artificial Intelligence*, **133**(B), 108158.
- Bongiovanni, C., Kaspi, M., Cordeau, J.-F., and Geroliminis, N. (2022), A machine learning-driven two-phase metaheuristic for autonomous ridesharing operations, *Transportation Research Part E*, **165**, 102835.
- Chicco, D., Warrens, M. J., and Jurman, G. (2021), The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, *PeerJ Computer Science*, **7**, e623.
- da Silva, F. A., Viana, A. P., Correa, C. C. G., Santos, E. A., de Oliveira, J. A. V. S., Andrade, J. D. G., Ribeiro, R. M., and Glória, L. S. (2021), Bayesian ridge regression shows the best fit for SSR markers in *Psidium guajava* among Bayesian models, *Scientific Reports*, **11**(1), 13639.
- Derouiche, H., Elarbi, M., and Bechikh, S. (2025), Deep crossover schemes for genetic algorithms: Investigations on the travel salesman problem, *Swarm and Evolutionary Computation*, **98**, 102094.
- Farahmand-Tabar, S. and Ashtari, P. (2024), Intelligent cross-entropy optimizer: A novel machine learning-based meta-heuristic for global optimization, *Swarm and Evolutionary Computation*, **91**, 101739.
- Garside, A. K., Ahmad, R., and Muhtazaruddin, M. N. B. (2024), A recent review of solution approaches for green vehicle routing problem and its variants, *Operations Research Perspectives*, **12**, 100303.
- Hamilton, D. F., Ghert, M., and Simpson, A. H. R. W. (2015), Interpreting regression models in clinical outcome studies, *Bone & Joint Research*, **4**(9), 152-153.
- Hvattum, L. M. (2022), Adjusting the order crossover operator for capacitated vehicle routing problems, *Computers & Operations Research*, **148**, 105986.
- Jorge, J. and Paredes, R. (2018), Passive-aggressive online learning with nonlinear embeddings, *Pattern Recognition*, **79**, 162-171.
- Kaelo, P. and Ali, M. M. (2007), Integrated crossover rules in real coded genetic algorithms, *European Journal of Operational Research*, **176**(1), 60-76.
- Karakatić, S. (2021), Optimizing nonlinear charging times of electric vehicle routing with genetic algorithm, *Expert Systems with Applications*, **164**, 114039.
- Kaya, M. (2011), The effects of two new crossover operators on genetic algorithm performance, *Applied Soft Computing*, **11**(1), 881-890.
- Kim, S. W., Lee, Y. G., Tama, B. A., and Lee, S. (2020), Reliability-enhanced camera lens module classification using semi-supervised regression method, *Applied Sciences*, **10**(11), 3832.
- Kramlinger, P., Schneider, U., and Krivobokova, T. (2023), Uniformly valid inference based on the Lasso in linear mixed models, *Journal of Multivariate Analysis*, **198**, 105230.
- Koohestani, B. (2020), A crossover operator for improving the efficiency of permutation-based genetic algorithms, *Expert Systems with Applications*, **151**, 113381.
- Kumari, S., Jha, R., Ray, A., Jena, J. J., Gourisaria, M. K., and Bandyopadhyay, A. (2025), An explainable AI approach for potato plant disease detection using enhanced feature engineering, *Procedia Computer Science*, **258**, 3570-3579.
- Lee, S. and Jun, C.-H. (2018), Fast incremental learning of logistic model tree using least angle regression, *Expert Systems with Applications*, **97**, 137-145.
- Lee, Y., Thyagarajan, K., Pinto, J. M., Charitopoulos, V. M., and Papageorgiou, L. G. (2024), Towards efficient solutions for vehicle routing problems for oxygen supply chains, *Computers & Chemical Engineering*, **191**, 108827.
- Li, D., Yang, C., and Zou, D. (2023), A nondominated sorting genetic algorithm III with three crossover strategies for the combined heat and power dynamic economic emission dispatch with or without prohibited operating zones, *Engineering Applications of Artificial Intelligence*, **123**(C), 106443.
- Li, J., Liu, R., and Wang, R. (2024), Handling dynamic capacitated vehicle routing problems based on adaptive genetic algorithm with elastic strategy, *Swarm and Evolutionary Computation*, **86**, 101529.
- Meilanitasari, P. and Shin, S.-J. (2021), A review of prediction and optimization for sequence-driven scheduling in job shop flexible manufacturing systems, *Processes*, **9**(8), 1391.
- Mittlböck, M. and Waldhör, T. (2000), Adjustments for R²-measures for Poisson regression models, *Computational Statistics & Data Analysis*, **34**(4), 461-472.
- Moghadaspoor, B., Bozorgi-Amiri, A., Tavakkoli-Moghaddam, R., and Bashiri, M. (2026), An ensemble learning method to predict a hydrogen production rate in the viable wastewater-hydrogen supply chain network design, *Energy*, **345**, 139950.
- Naskar, A., Ghosh, S., Kundu, M., and Sarkar, R. (2025), Feature selection using guided population based ge-

- netic algorithm with modified crossover and parent selection, *Applied Soft Computing*, **172**, 112872.
- Nazif, H. and Lee, L. S. (2012), Optimised crossover genetic algorithm for capacitated vehicle routing problem, *Applied Mathematical Modelling*, **36**(5), 2110-2117.
- Ojeda Rios, B. H. and Xavier, E. C. (2025), Metaheuristic approaches for the stochastic capacitated multi-depot vehicle routing problem with pickup and delivery, *Expert Systems with Applications*, **290**, 128258.
- Ostermeier, M., Henke, T., Hübner, A., and Wäscher, G. (2021), Multi-compartment vehicle routing problems: State-of-the-art, modeling framework and future directions, *European Journal of Operational Research*, **292**(3), 799-817.
- Gupta, A., Stead, T. S., and Ganti, L. (2024), Determining a meaningful R-squared value in clinical medicine, *Academic Medicine & Surgery*.
- Polat, Ö. and Kayhan, S. K. (2018), High-speed FPGA implementation of orthogonal matching pursuit for compressive sensing signal reconstruction, *Computers & Electrical Engineering*, **71**, 173-190.
- Riaventin, V. N., Cahyo, S. D., and Singgih, I. K. (2021), A model for developing existing ports considering economic impact and network connectivity, *Sustainability*, **13**(7), 3705.
- Sandhu, R., Pettit, C., Khalil, M., Poirel, D., and Sarkar, A. (2017), Bayesian model selection using automatic relevance determination for nonlinear dynamical systems, *Computer Methods in Applied Mechanics and Engineering*, **320**, 237-260.
- Scikit-learn (2022), 1. Supervised learning. Scikit-Learn. Available from: https://scikit-learn.org/stable/supervised_learning.html.
- Scikit-learn (2023), API Reference, cited 2023 Dec 2, Available from: scikit-learn.org.
- SHAP (2025a, October 12), Shap.KernelExplainer. Available from: <https://shap.readthedocs.io/en/latest/generated/shap.KernelExplainer.html>
- SHAP (2025b, October 12), Shap.LinearExplainer. Available from: <https://shap.readthedocs.io/en/latest/generated/shap.LinearExplainer.html>
- Shen, J., Liu, K., Ma, C., Zhao, Y., and Shi, C. (2022), Bibliometric analysis and system review of vehicle routing optimization for emergency material distribution, *Journal of Traffic and Transportation Engineering* (English Edition), **9**(6), 893-911.
- Singgih, I. K. and Kim, B. I. (2020), Multi-type electric vehicle relocation problem considering required battery-charging time, *European J. of Industrial Engineering*, **14**(3), 335-368.
- Szénási, S. and Légrádi, G. (2024), Machine learning aided metaheuristics: A comprehensive review of hybrid local search methods, *Expert Systems with Applications*, **258**, 125192.
- Taiwo, B. O., Fissaha, Y., Hosseini, S., Khishe, M., Karaman, E., Adebayo, B., Sazid, M., Adesida, P. A., Famobuwa, O. V., Faluyi, J. O., and Akinlabi, A. A. (2024), Machine learning based prediction of flyrock distance in rock blasting: A safe and sustainable mining approach, *Green and Smart Mining Engineering*, **1**(3), 346-361.
- Wang, X., Wang, X., Ma, B., Li, Q., Wang, C., and Shi, Y. (2023), High-performance reversible data hiding based on ridge regression prediction algorithm, *Signal Processing*, **204**, 108818.
- Xiao, Z., Zhi, J., and Keskin, B. B. (2022), Towards a machine learning-aided metaheuristic framework for a production/distribution system design problem, *Computers and Operations Research*, **146**, 105897.
- Xie, Y., Chu, K.-F., Lam, A. Y. S., and Iida, F. (2025), A multi-objective evolutionary algorithm with constraint-compliant initialization for energy transport and urban logistics in Electric Vehicle Routing, *Applied Soft Computing*, **183**, 113624.
- Yu, B., Chen, C., Wang, X., Yu, Z., Ma, A., and Liu, B. (2021), Prediction of protein-protein interactions based on elastic net and deep forest, *Expert Systems with Applications*, **176**, 114876.
- Zade, F. A., Ghafurian, M. M., Mesgarpour, M., and Niazmand, H. (2023), Predictive machine learning models for optimization of direct solar steam generation, *Journal of Water Process Engineering*, **56**, 104304.
- Zhang, J. and Woensel, T. V. (2023), Dynamic vehicle routing with random requests: A literature review, *International Journal of Production Economics*, **256**, 108751.
- Zhang, X., Hao, Y., Zhang, L., and Yuan, X. (2025), Application of improved genetic algorithm to vehicle routing problem considering the environmental self-regulation of the freight companies, *Expert Systems with Applications*, **274**, 127010.
- Zhang, P.-F., Zhang, D., Zhao, Q., Tuerxunmaimaiti, Y., Yang, X., Zhao, X., and Iqbal, M. (2026), Machine learning-guided interpretable modelling of code-based equation for bond strength of composite bars in ultra-high-performance concrete, *Advanced Engineering Informatics*, **73**, 104542.
- Zhu, F., Li, H., Li, J., Zhu, B., and Lei, S. (2023), Unmanned aerial vehicle remote sensing image registration based on an improved oriented FAST and rotated BRIEF- random sample consensus algorithm, *Engineering Applications of Artificial Intelligence*, **126**(B), 106944.

Ivan Kristianto Singgih is the Head of Systems Engineering Laboratory in the Study Program of Industrial Engineering at University of Surabaya (Ubaya), Indonesia. He received his Ph.D. degree in Industrial Engi-

neering from Pusan National University, Busan, Korea, in 2017. He has been a Research Assistant Professor under the Department of Industrial and Management Engineering at Pohang University of Science and Technology (POSTECH), Korea Advanced Institute of Science and Technology (KAIST) and Korea University. His research interests include operations research, smart systems, machine learning, and simulation.

Dr. Eng. Ir. Aditya Rio Prabowo is an Associate Professor in the Department of Mechanical Engineering at Universitas Sebelas Maret (UNS) in Surakarta, Central Java, where he also serves as Head of Laboratory. His research primarily investigates structural crashworthiness and accident-related phenomena in marine environments—such as collisions, groundings, and explosions—with a particular focus on engineering operations in extreme regions, such as the Arctic. He is interested in structural analysis and probabilistic estimation of large systems using numerical methods, including the Finite Ele-

ment and Monte Carlo methods. With an award as one of the top two percent of world researchers and a substantial publication record in impact engineering and marine structures, he is recognized for his technical expertise in offshore-marine structural safety.

Prita Meilanitasari a lecturer in the Logistics Engineering Department in one of Private University in Indonesia. She earned her Ph.D. from Hanyang University Graduate School of Technology & Innovation Management and has a research background in advanced manufacturing and supply chain management. Prita has worked as a research and teaching assistant at Hanyang University and Pukyong National University. Her areas of expertise include advanced manufacturing, logistics systems, data analytics, and supply chain optimization. Her academic and research experience includes developing data-driven models for manufacturing system optimization, logistics management, and supply chain analytics.

Industrial Engineering & Management Systems



[HOME](#)
[Contact Us](#)

On-line Submission

Industrial Engineering
& Management Systems

[About Us +](#)
[Journal Information](#)
[Introduction](#)
[Aims and Scope](#)
[Editorial Board +](#)
[Editor-in-Chief](#)
[Area Editors](#)
[Advisory Board](#)
[Board Members of
APIEMS](#)
[For Contributors +](#)
[Instruction For Authors](#)
[Ethical Guideline](#)
[Crossmark Policy](#)
[SUBMISSION CHECKLIST](#)
[Journal Search +](#)
[Most read Article](#)
[Most Read Articles in
the Recent 5 Years](#)
[Current Issue](#)
[Archive](#)

Introduction

IEMS has been launched in 2002 by the APIEMS (Asian Pacific Industrial Engineering and Management Society) and now it is an official journal of APIEMS and KIIE (Korean Institute of Industrial Engineers). IEMS is intended for reporting original quality works on industrial engineering and management systems. The journal publishes original works that have significant theoretical and academic values, novel industrial applications, and case studies. Every manuscript is strictly reviewed by two or more anonymous referees and final acceptance decision is made based on the review reports. The journal aims at speedy review and processing within 3 months and it will publish four issues per year.

Contact Us

Inquiry about submission and review:

Managing Editor

Dong-Ho Lee

E-mail: leman@hanyang.ac.kr

Inquiry about subscription and payment:

The Korean Institute of Industrial Engineers (KIIE)

victoria Bldg. Room 509, 327, Teheran-ro, Gangnam-gu, Seoul, Korea 06150

Tel: +82-2-527-3096, Fax: +82-2-527-3162, Email: kiie_admin@hate.com



ORCID
Open Researcher and Contributor ID





Copyright © The Korean Institute of Industrial Engineers. All rights reserved.
2006, Jungang Royal, 13 Seoun-ro, Seocho-gu, Seoul, Republic of Korea (06732)
Tel: +82-2-523-3095, Fax: +82-2-523-3090, Email: admin@kiiie.org

Industrial Engineering & Management Systems



[HOME](#)
[Contact Us](#)

On-line Submission

Industrial Engineering
& Management Systems

[About Us +](#)
[Journal Information](#)
[Introduction](#)
[Aims and Scope](#)
[Editorial Board +](#)
[Editor-in-Chief](#)
[Area Editors](#)
[Advisory Board](#)
[Board Members of
APIEMS](#)
[For Contributors +](#)
[Instruction For Authors](#)
[Ethical Guideline](#)
[Crossmark Policy](#)
[SUBMISSION CHECKLIST](#)
[Journal Search +](#)
[Most read Article](#)
[Most Read Articles in
the Recent 5 Years](#)
[Current Issue](#)
[Archive](#)

Aims and Scope

Industrial Engineering and Management Systems (IEMS)

covers all areas of industrial engineering and management sciences including but not limited to, applied statistics & data mining, business & information systems, computational intelligence & optimization, environment & energy, ergonomics & human factors, logistics & transportation, manufacturing systems, planning & scheduling, quality & reliability, supply chain management & inventory systems.



ORCID
Open Researcher and Contributor ID





Copyright © The Korean Institute of Industrial Engineers. All rights reserved.
2006, Jungang Royal, 13 Seoun-ro, Seocho-gu, Seoul, Republic of Korea (06732)
Tel: +82-2-523-3095, Fax: +82-2-523-3090, Email: admin@kiiie.org

Industrial Engineering & Management Systems


[HOME](#)
[Contact Us](#)

On-line Submission

 Industrial Engineering
& Management Systems

[About Us +](#)
[Journal Information](#)
[Introduction](#)
[Aims and Scope](#)
[Editorial Board +](#)
[Editor-in-Chief](#)
[Area Editors](#)
[Advisory Board](#)
[Board Members of
APIEMS](#)
[For Contributors +](#)
[Instruction For Authors](#)
[Ethical Guideline](#)
[Crossmark Policy](#)
[SUBMISSION CHECKLIST](#)
[Journal Search +](#)
[Most read Article](#)
[Most Read Articles in
the Recent 5 Years](#)
[Current Issue](#)
[Archive](#)

Editor-in-Chief

Editor-in-Chief

Chi-Hyuck Jun

Department of Industrial & Management Engineering

POSTECH

77 Cheongam-Ro, Nam-Gu, Pohang, Gyeongbuk, Korea 37673

Tel: +82-54-279-2197

Fax: +82-54-279-2870

E-mail: chjun@postech.ac.kr

Deputy Editor

Hirokazu Kono

Graduate School of Business Administration

Keio University

4-1-1 Hiyoshi, Kohoku-ku,

Yokohama 223-8526, Japan

Tel: +81-45-564-2019

Fax: +81-45-562-3502

Email: kono@kbs.keio.ac.jp

Managing Editor

Dong-Ho Lee

Department of Industrial Engineering

HANYANG University

222 Wangsimni-ro, Seongdong-gu, Seoul, 04763, KOREA

Tel: +82-2-2220-0475

E-mail: leman@hanyang.ac.kr

url: <http://pli.hanyang.ac.kr>

Manuscript Editor

Youngdoo Son

Department of Industrial&Systems Engineering dongguk University

30, Pildong-ro 1-gil, Jung-gu, Seoul, 04620, Republic of Korea

Tel: +82-2-2260-3840

Fax: +82-2-2269-2212

E-mail: iems.manuscript@gmail.com





Copyright © The Korean Institute of Industrial Engineers. All rights reserved.
2006, Jungang Royal, 13 Seoun-ro, Seocho-gu, Seoul, Republic of Korea (06732)
Tel: +82-2-523-3095, Fax: +82-2-523-3090, Email: admin@kiiie.org

Industrial Engineering & Management Systems



[HOME](#)
[Contact Us](#)

On-line Submission

Industrial Engineering
& Management Systems

[About Us +](#)
[Journal Information](#)
[Introduction](#)
[Aims and Scope](#)
[Editorial Board +](#)
[Editor-in-Chief](#)
[Area Editors](#)
[Advisory Board](#)
[Board Members of
APIEMS](#)
[For Contributors +](#)
[Instruction For Authors](#)
[Ethical Guideline](#)
[Crossmark Policy](#)
[SUBMISSION CHECKLIST](#)
[Journal Search +](#)
[Most read Article](#)
[Most Read Articles in
the Recent 5 Years](#)
[Current Issue](#)
[Archive](#)

Area Editors

Applied Statistics & Data Mining

Chien-Chih Wang

Ming Chi University of Technology, Taiwan

Jong-Seok Lee

KAIST, Korea

Weonyoung Joo

Ewha Woman' University, Korea

Kyoungok Kim

Seoul National University of Science and Technology, Korea

Business & Information Systems

Hyunseok Hwang

Hallym University, Korea

Bashar Almansour

Muscat University, Muscat, Oman

Hussein Kadhim Sharaf

Bilad Alrafidain University College, Iraq

Computational Intelligence & Optimization

Sang Jin Kweon

UNIST, Korea

Lixing Yang

Beijing Jiaotong University, China

Environment and Energy

Ming Lang Tseng

Asia University, Taiwan

Dong Gu Choi

POSTECH, Korea

Ergonomics & Human Factors

Jaehyun Park

Konkuk University

Sangeun Jin

Pusan National University

Logistics & Transportation

Byung-In Kim

POSTECH, Korea

Ching-Jung Ting

Yuan Ze University, Taiwan

Manufacturing Systems

Funlade Sunmola

University of Hertfordshire, U.K

Jei-Zheng Wu

Soochow University, Taiwan

Planning & Scheduling

Kazuko Morizawa

Osaka Prefecture University, Japan

Jun-Hee Han

Pusan National University, Korea

Sungbum Jun

Dongguk University, Korea

Quality & Reliability

Kwai-Sang Chin

City University of Hong Kong, China

Yi-Kuei Lin

National Yang Ming Chiao Tung University, Taiwan

Seong-joon Kim

Chosun University, Korea

SCM & Inventory Systems

Huynh Trung Luong

Asian Institute of Technology, Thailand

Kun Soo Park

Seoul National University, Korea

Bhaskar Bhandarkar

Indian Institution of Industrial Engineering, India



ORCID
Open Researcher and Contributor ID





Copyright © The Korean Institute of Industrial Engineers. All rights reserved.
2006, Jungang Royal, 13 Seoun-ro, Seocho-gu, Seoul, Republic of Korea (06732)
Tel: +82-2-523-3095, Fax: +82-2-523-3090, Email: admin@kiiie.org

Industrial Engineering & Management Systems



[HOME](#)[Contact Us](#)

On-line Submission

Industrial Engineering
& Management Systems

[About Us +](#)[Journal Information](#)[Introduction](#)[Aims and Scope](#)[Editorial Board +](#)[Editor-in-Chief](#)[Area Editors](#)[Advisory Board](#)[Board Members of
APIEMS](#)[For Contributors +](#)[Instruction For Authors](#)[Ethical Guideline](#)[Crossmark Policy](#)[SUBMISSION CHECKLIST](#)[Journal Search +](#)[Most read Article](#)[Most Read Articles in
the Recent 5 Years](#)[Current Issue](#)[Archive](#)

Advisory Board

Min K. Chung

Department of Industrial & Management Engineering
POSTECH

San 31 Hyoja-dong, Pohang 790-784, Korea

E-mail: mkc@postech.ac.kr

Augustine Esogbue

School of Industrial and System Engineering Georgia Institute of
Technology Groseclose 0205, Room 442

765 Ferst Drive, Atlanta, GA 30332-0205, USA

E-mail: augustine.esogbue@sye.gatech.edu

Hark Hwang

Department of Industrial Engineering Korea Advanced Institute of
Science & Technology Taejeon, Korea

E-mail: harkhwang@kaist.ac.kr

Kazuyoshi Ishii

Department of Industrial & Information Systems Engineering
Kanazawa Institute of Technology

Kanazawa, Japan

E-mail: ishiik@heptune.kanazawa-it.ac.jp

Erhan Kozan

Senior Lecturer, School of Management Science, Queensland
University of Technology GPO Box 2434 Brisbane, Q 4001,
Australia

E-mail: e.kozan@qut.edu.au

Mitsuru Kuroda

Department of Industrial and Systems Engineering
Aoyama Gakuin University JAPAN

E-mail: kuroda@se.aoyama.ac.jp

Kin Keung Lai

Department of Management Sciences City University of Hong Kong CIRRELT and Canada Research Chair in Distribution Management
Tat Chee Avenue, Kowloon, Hong Kong

E-mail: mskklai@cityu.edu.hk

Gilbert Laporte

CIRRELT and Canada Research Chair in Distribution Management
HEC Montréal

3000, chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T
2A7

E-mail: Gilbert.Laporte@cirrelt.ca

Tsong-Ming Lin

National Yunlin University of Science & Technology
123 University Road, Section 3, Doulin, Yunlin 64002, Taiwan
Email: tmlin@yuntech.edu.tw

Baoding Liu

Department of Mathematical Sciences
Tsinghua University
Beijing 100084, China

E-mail: liu@tsinghua.edu.cn

Masayuki Matsui

Department of Systems Engineering
The University of Electro-Communications
1-5-1, Chofugaoka, Chofu, Tokyo, 182-8585, Japan

E-mail: matsui@se.uec.ac.jp

Suebsak Nanthavanij

Department of Industrial Engineering
Thammasat University
P.O.Box 22, Thammasat-Rangsit Post Office

Klong Luang, Pathumthani 12121, Thailand

E-mail: suebsak@siit.tu.ac.th

Barry L. Nelson

Department of Industrial Engineering & Management Sciences
Northwestern University

McCormick School of Engineering and Applied Science
2145 Sheridan Road, Room C250, Evanston, IL 60208-3119, USA

E-mail: nelsonb@northwestern.edu

Yang-Byung Park

Department of Industrial Engineering,
Kyung Hee University

Yongin-si, Gyeonggi-do, Korea

E-mail: ybpark@khu.ac.kr



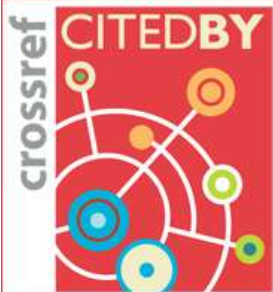
ORCID

Open Researcher and Contributor ID



Powered by iThenticate



*Linyan Sun*

School of Management
Xi'an Jiaotong University
Xi'an, 710049, China
E-mail: lysun@mail.xjtu.edu.cn

Mario Tabucanon

School of Advanced Technologies,
Asian Institute of Technology
P.O. Box 4, Klong Luang, Pathumthani 12120, Thailand
E-mail: mtt@ait.ac.th

James M. Tien

Decision Science & Engineering Systems
Rensselaer Polytechnic Institute
CII 5023 110 8th St. Troy, NY 12180-3590, USA
E-mail: tienj@rpi.edu

Dingwei Wang

Department of System Engineering
Northeastern University,
Shenyang 110006, China

Kuo-Ming Wang

Department of Industrial Engineering and Engineering Management
National Tsing Hua University
101, Kuang Fu Road, Sec. 2, Hsinchu 300, Taiwan
E-mail: kmwang@mx.nthu.edu.tw

Copyright © The Korean Institute of Industrial Engineers. All rights reserved.
2006, Jungang Royal, 13 Seoun-ro, Seocho-gu, Seoul, Republic of Korea (06732)
Tel: +82-2-523-3095, Fax: +82-2-523-3090, Email: admin@kiiie.org

Industrial Engineering & Management Systems



[HOME](#)
[Contact Us](#)

On-line Submission

Industrial Engineering & Management Systems

[About Us +](#)
[Journal Information](#)
[Introduction](#)
[Aims and Scope](#)
[Editorial Board +](#)
[Editor-in-Chief](#)
[Area Editors](#)
[Advisory Board](#)
[Board Members of APIEMS](#)
[For Contributors +](#)
[Instruction For Authors](#)
[Ethical Guideline](#)
[Crossmark Policy](#)
[SUBMISSION CHECKLIST](#)
[Journal Search +](#)
[Most read Article](#)
[Most Read Articles in the Recent 5 Years](#)
[Current Issue](#)
[Archive](#)


Industrial Engineering & Management Systems

Journal Abbreviation : Ind. Eng. Manag. Syst.

Frequency : Quarterly

DOI Prefix : 10.7232/iems

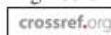
ISSN : 1598-7248 (Print)

ISSN : 2234-6473 (Online)

Year of Launching : 2002

Publisher : The Korean Institute of Industrial Engineers

Indexed/Tracked/Covered By :



Volume.25 No.3 September 2026

Bias Reduction in Statistical Modeling: Improved Estimation for Inverse Power Maxwell Survival Functions and Waring Regression Models View count **59**

Nazik J. Sadik, Haifa Taha Abd, Zakariya Yahya Algarni

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.542-553

DOI: <https://doi.org/10.7232/iems.2026.25.3.542>

[Open abstract](#)

[Download pdf](#)
[Export Citation](#)

Universal Design for the Bus Carousel in the Philippines View count **119**

Min K. Chung, Alma Maria Jennifer Gutierrez, Betina Ching, Zian Sol, Jose Sta. Cruz, Charles Tan

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.554-570

DOI: <https://doi.org/10.7232/iems.2026.25.3.554>

[Open abstract](#)

[Download pdf](#)
[Export Citation](#)

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem using Regression Machine Learning: Post-Experiment Analysis View count **38**

Ivan Kristianto Singgih, Aditya Rio Prabowo, Prita Meilanitasari

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.571-582

DOI: <https://doi.org/10.7232/iems.2026.25.3.571>

[Open abstract](#)

[Download pdf](#)
[Export Citation](#)

Intelligent Course Scheduling with Genetic Algorithm on a Web-Based System View count **57**

Rengganis Wahyuaning Putri, Imam Fahrurrozi, Muhammad Fakhurrifqi, Ronald Adrian, Irkham Huda, Harnan Malik Abdullah, Aulia Faqih Rifai, Muhammad Syafrudin, Ganjar Alfian

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.583-597

DOI: <https://doi.org/10.7232/iems.2026.25.3.583>

[Open abstract](#)

[Download pdf](#)
[Export Citation](#)

Attribute Quick Switch Sampling Inspection Plan for Time-Truncated Life Tests Using the Quasi Aradhana Distribution View count **31**

Qurratulain Hanif, Nadia Saeed, Mohammed Albassam, Muhammad Aslam

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.598-608

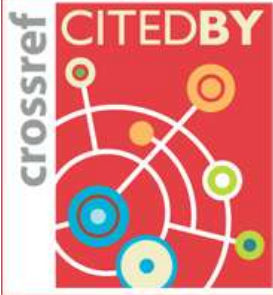
DOI: <https://doi.org/10.7232/iems.2026.25.3.598>

[Open abstract](#)

[Download pdf](#)
[Export Citation](#)

Sociodemographic Characteristics as Predictors of Human Error: A Case Study in Indonesian Secondary Hospital View count **30**





Tutik Farihah, Jamasri, Rini Dharmastiti

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.609–620

DOI: <https://doi.org/10.7232/iems.2026.25.3.609>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Robust DMAIC-Driven Optimization of Expansion Index in Cereal Extrusion Using DOE and SHAP-Based Factor Screening View count 107

Angel Geovanny Guamán Lozano, Julio Cesar Moyano Alulema, Jhofre Ramiro Guamán Quishpi

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.621–636

DOI: <https://doi.org/10.7232/iems.2026.25.3.621>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Evaluating Green Supply Chain Management Practices through GRI Framework-Based Sustainability Reporting in Indonesian Manufacturing View count 30

Elisa Kusrini, Vembri Noor Helia, Putri Dwi Annisa, Zelania In Haryanto, Didin Dwi Novianto

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.637–651

DOI: <https://doi.org/10.7232/iems.2026.25.3.637>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

A Simulation-Based Study on the Impact of Redundancy and Repair Capacity on Drone Availability View count 30

Seungjong Song, Seongam Moon, Kyunghwan Choi

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.652–667

DOI: <https://doi.org/10.7232/iems.2026.25.3.652>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Vehicle Parking Problem with Uncertain Arrival Times and Parking Durations Considering Time-Sensitive Users View count 49

Aditya Saputra, Ivan Kristianto Singgih

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.668–673

DOI: <https://doi.org/10.7232/iems.2026.25.3.668>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Design and Effectiveness Evaluation of an ML-Based Decision Support System (DSS) for Fraud Detection in Auto Insurance Companies View count 30

Theodore Jun

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.674–683

DOI: <https://doi.org/10.7232/iems.2026.25.3.674>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Using Upper Trapezius sEMG to Inform Rest Break Timing in Office Workers: A Preliminary Feasibility Study View count 67

Ika Fitri Wulan Dhari, Susy Purnawati, I Putu Gede Adiatmika, I Made Ady Wirawan

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.684–695

DOI: <https://doi.org/10.7232/iems.2026.25.3.684>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Optimizing Lean Six Sigma Project Selection in Higher Education Institutions: A Fuzzy MCDM Approach View count 31

Ig. Jaka Mulyana, Dian Retno Sari Dewi, Ivan Gunawan, Moses Laksono Singgih, Sri Gunani Partiw

Industrial Engineering & Management Systems :: Vol.25 No.3 pp.696–708

DOI: <https://doi.org/10.7232/iems.2026.25.3.696>

[Open abstract](#)

[Download pdf](#)

[Export Citation](#)

Sumake

We are global provider of air
screwdriver and electric tools,
assembly systems

sumake.com

Open

Industrial Engineering and Management Systems

South Korea | Universities and research institutions | Media Ranking

Country	Subject Area and Category	Publisher
South Korea	Economics, Econometrics and Finance	Korean Institute of Industrial Engineers
 SCImago Institutions Rankings	Economics, Econometrics and Finance (miscellaneous)	
 SCImago Media Rankings	Social Sciences Social Sciences (miscellaneous)	

Sumake

We are global provider of air screwdriver and electric tools, assembly systems

0.198

Q3

22

Publication type

Journals

ISSN

15987248, 22346473

Coverage

2014-2025

Information

Home

How to publish in this journal

chjun@postech.ac.kr

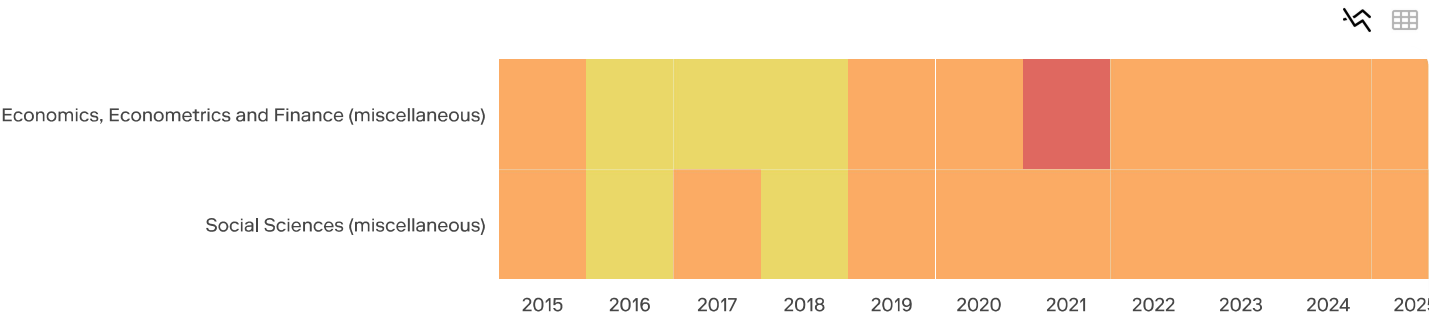
CircuitHub - Quick-turn PCB assembly

Rapidly iterate on PCB prototypes with reduced upfront design time. CircuitHub

Scope

Industrial Engineering and Management Systems (IEMS) covers all areas of industrial engineering and management sciences including but not limited to, applied statistics & data mining, business & information systems, computational intelligence & optimization, environment & energy, ergonomics & human factors, logistics & transportation, manufacturing systems, planning & scheduling, quality & reliability, supply chain management & inventory systems.

Quartiles



	Soekarno-Hatta Jakarta ✈️ Jakarta ✈️ Padang mulai Rp2.124.823 Lihat	Jakarta ✈️ Padang mulai Rp2.124.823 Lihat	Jakarta ✈️ Edinburgh mulai Rp13.107.523 Lihat	Soekarno-Hatta ✈️ Kuala Namu M mulai Rp2.423. Lihat
--	---	---	---	---

Find similar journals

All quartilesAll countriesAll subject categoriesClear filters

Download

Only Open Access Journals

1 - International Journal of Industrial Engineering and Production Research



63%

similarity

2 - Journal of Optimization in Industrial Engineering

61%

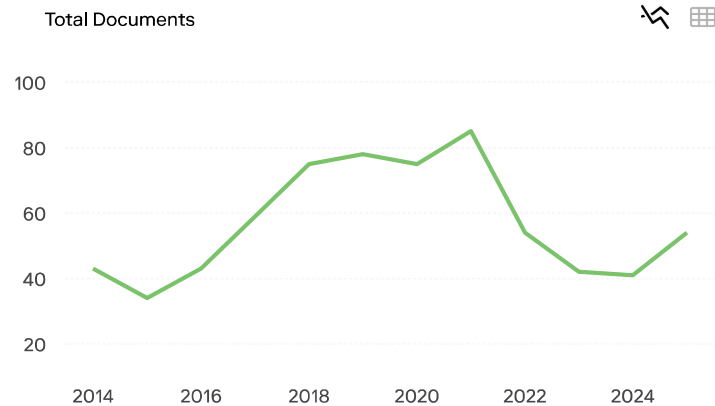
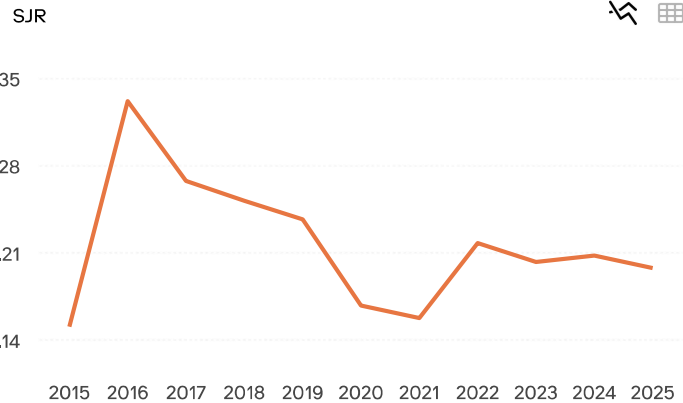
similarity

3 - OPSEARCH

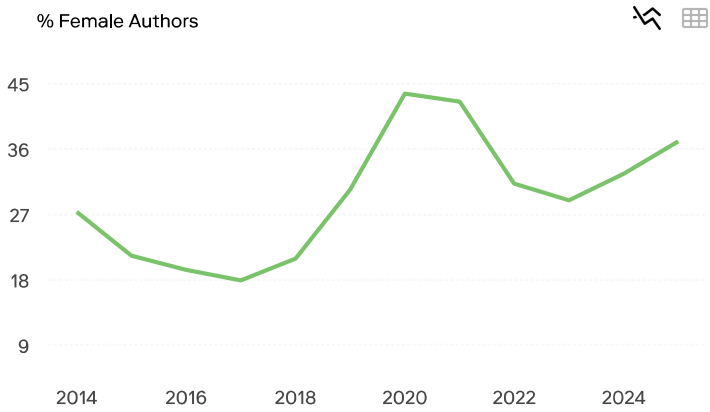
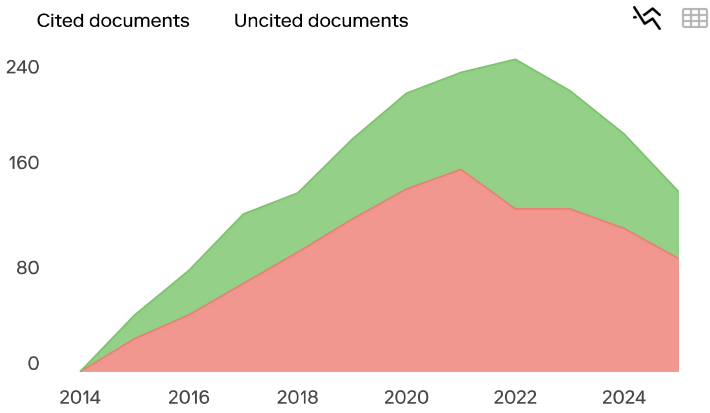
60%

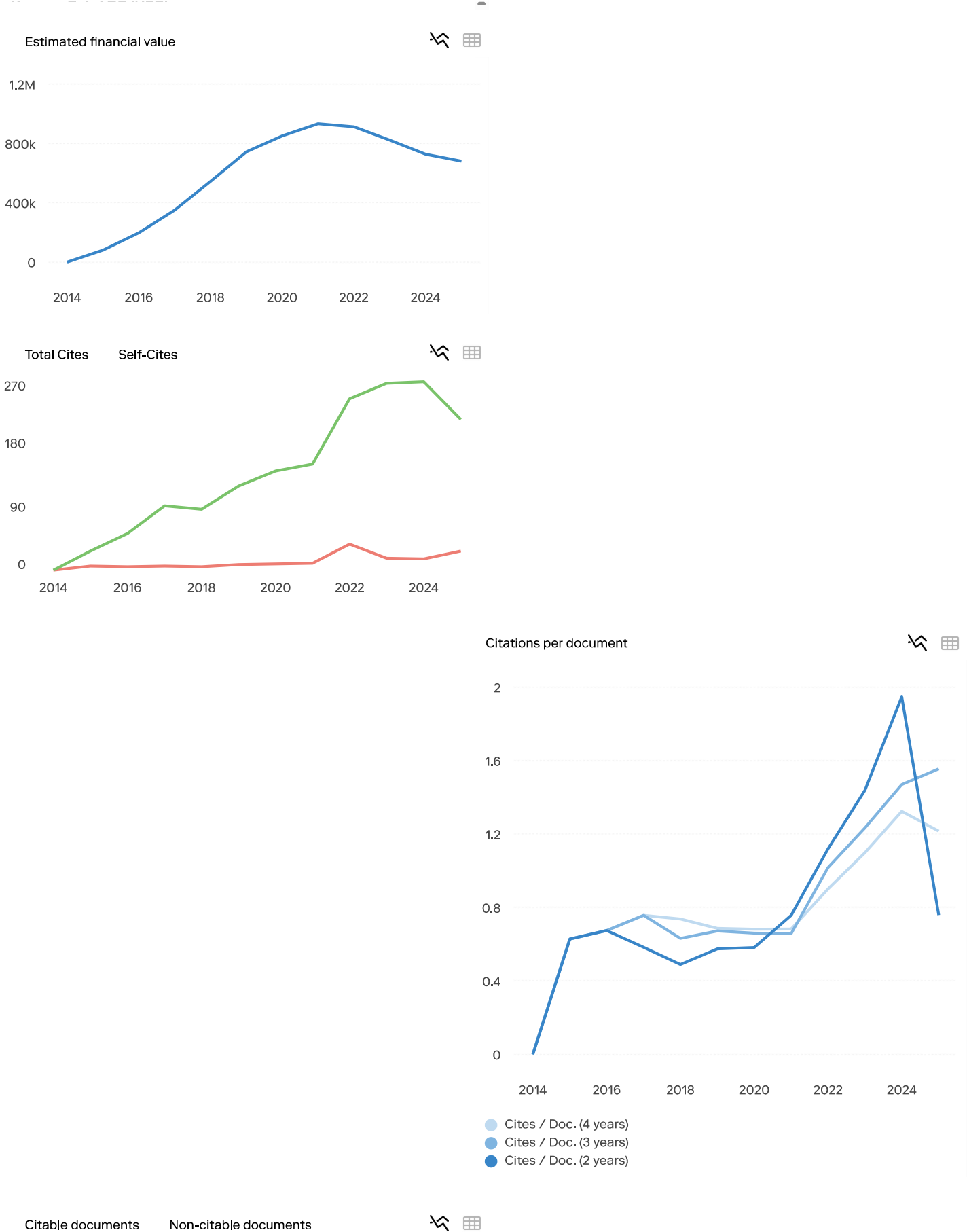
similarity

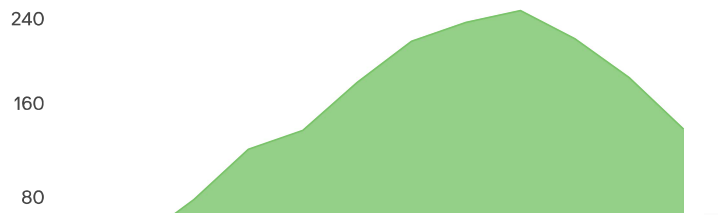
< >



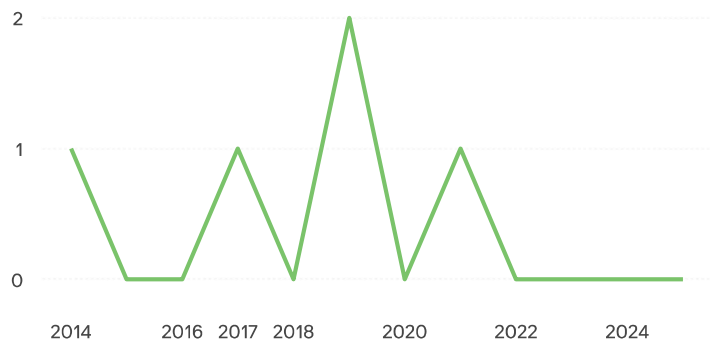
External Cites per DocCites per Doc







Documents cited by public policy (Overton)



Documents related to SDGs (UN)



Coming soon



← Show this widget in your own website

Just copy the code below and paste within your html code:

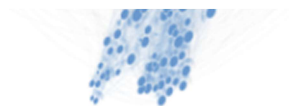
```
<a href="https://www.scimago
```

SCImago Graphica

Explore, visually communicate and make sense of data



with our **new data**
visualization tool.



Metrics based on Scopus® data as of March 2026

A

ASHIDA AP

1 month ago

is it in the Q3 list now? it shows as till 2025

reply



Melania Ortiz - SCImago Team

1 month ago

Dear Ashida, Thank you very much for your request. You can consult this information above. Best Regards, SCImago Team

L

La Gubu

2 years ago

We, La Gubu, et al have submitted an article in Industrial Engineering

reply



Melanie Ortiz

2 years ago

Dear La Gubu,

Thank you for contacting us. Could you please expand a little bit on your request so we can assist you better?

Best Regards, SCImago Team

L

La Gubu

3 years ago

Dear Editors,

How to submit a manuscript. Is it emailed directly to catenoid@kaist.edu? Or is there an ad

Warm regards

La Gubu

reply



Melanie Ortiz

3 years ago

Dear La Gubu,

Thank you for contacting us.

We suggest you visit the journal's homepage (See submission/author guidelines) or contact the journal's editorial staff , so they could inform you more deeply.

Best Regards, SCImago Team

K

Karnawi Kamar

4 years ago

Thank you very much.

reply

**Hafed M. Motair**

6 years ago

Is the (Industrial Engineering and Management Systems) journal still in scopus because I see it's coverage 2014-2019 and not writing 2014-ongoing.
Thanks

reply

**Melanie Ortiz**

6 years ago

Dear Hafed, thank you very much for your comment. We suggest you consult the Scopus database directly. Keep in mind that the SJR is a static image (the update is made one time per year) of a database (Scopus) which is changing every day.
Best Regards, SCImago Team

G

gehad kamel

8 years ago

how can i publish a paper in this journal ?
what is the terms to publish?
thank you

reply

**Elena Corera**

8 years ago

Thank you very much for participating!
Please, check out our FAQs <https://www.scimagojr.com/help.php?q=FAQ>

Name

Email

Message

☐ I'm not a robot

reCAPTCHA

Send

The users of Scimago Journal & Country Rank have the possibility to dialogue through comments linked to a specific journal. The purpose is to have a forum in which general doubts about the processes of publication in the journal, experiences and other issues derived from the publication of papers are resolved. For topics on particular articles, maintain the dialogue through the usual channels with your editor.



Services

Journal Rankings
Country Rankings
Journal Value

About us

Who we are
Contact us

SCImago's products

SCImago Journal Country & Rank
SCImago Institutions Ranking
SCImago Media Ranking
SCImago Iber
SCImago Arab
SCImago Graphica
Ediciones Profesionales de la Información

All the information shown in the SCImago Journal Country & Rank website can be used for non-commercial purposes as long as it is cited.

How to cite:

SCImago, (n.d.). SJR-SCImago Journal Country & Rank [Portal]. Retrieved (Date), from <https://www.scimagojr.com>

Powered by:
Scopus®



Source details

Industrial Engineering and Management Systems

Years currently covered by Scopus: from 2014 to 2026

Publisher: Korean Institute of Industrial Engineers

ISSN: 1598-7248 E-ISSN: 2234-6473

Subject area: Social Sciences: General Social Sciences

Economics, Econometrics and Finance: General Economics, Econometrics and Finance

Source type: Journal

- View all documents >
- Set document alert
- Save to source list

CiteScore 2025
3.0 ⓘ

SJR 2025
0.198 ⓘ

SNIP 2025
0.335 ⓘ

CiteScore CiteScore rank & trend Scopus content coverage

CiteScore 2025

3.0 = $\frac{571 \text{ Citations } 2022 - 2025}{191 \text{ Documents } 2022 - 2025}$

Calculated on 05 May, 2026

CiteScoreTracker 2026 ⓘ

1.2 = $\frac{209 \text{ Citations to date}}{172 \text{ Documents to date}}$

Last updated on 05 October, 2026 • Updated monthly

CiteScore rank 2025 ⓘ

Category	Rank	Percentile
Social Sciences		
General Social Sciences	#85/286	70th
Economics, Econometrics and Finance		
General	#92/289	68th

View CiteScore methodology > CiteScore FAQ > Add CiteScore to your site

About Scopus

[What is Scopus](#)

[Content coverage](#)

[Scopus blog](#)

[Scopus API](#)

[Privacy matters](#)

Language

[日本語版を表示する](#)

[查看简体中文版本](#)

[查看繁體中文版本](#)

[Просмотр версии на русском языке](#)

Customer Service

[Help](#)

[Tutorials](#)

[Contact us](#)

ELSEVIER

[Terms and conditions](#) ↗ [Privacy policy](#) ↗ [Cookies settings](#)

All content on this site: Copyright © 2026 Elsevier B.V. ↗, its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the relevant licensing terms apply.



IEMS_TSP_ML

by Ivan Kristianto Singgih

Submission date: 22-Oct-2026 02:17PM (UTC+0700)

Submission ID: 3021597586

File name: 1_Paper.pdf (726.23K)

Word count: 7589

Character count: 39012

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem using Regression Machine Learning: Post-Experiment Analysis

1

Ivan Kristianto Singgih

Study Program of Industrial Engineering, University of Surabaya, Surabaya, Indonesia
Institut Supply Chain dan Logistik Indonesia, Indonesia
Asosiasi Peneliti Indonesia Korea (APIK), Republic of Korea
Kolaborasi Riset dan Inovasi Industri Kecerdasan Artifisial (KORIKA), Jakarta, Indonesia
Indonesia Artificial Intelligence Society (IAIS)

9

Aditya Rio Prabowo

Department of Mechanical Engineering, Universitas Sebelas Maret, Surakarta 57126, Indonesia

Prita Meilanitasari*

Department of Logistics Engineering, Universitas Internasional Semen Indonesia, Gresik 61122, Indonesia

12

(Received: November 11, 2025 / Revised: February 11, 2026; April 18, 2026 / Accepted: April 27, 2026)

ABSTRACT

Traveling salesman problem (TSP) is an important problem because it is the base of many real-world transportation problems. One of the most effective and frequently used optimization methods for solving TSP is the genetic algorithm. Generally, genetic algorithms employ two types of operators: crossover and mutation. Most researchers apply the crossover operator to genetic algorithms to find better solutions, without analyzing the improvement in solution quality each time the crossover is applied. Such an approach is commonly used with the aim of simply finding the best route, for example, by comparing the quality of final solutions produced by the genetic algorithm and other methods. Unlike previous studies, this study examines the characteristics of the updated solution within the genetic algorithm run, which is modified by crossover, specifically, the improvement in solution quality. To the best knowledge of the authors, this study is the first to evaluate the effectiveness of crossover operators using regression machine learning methods. The numerical experiments show that the most important feature to predict the change in the best objective values is the distance between customers outside the cutting points, followed by the distance between customers within the cutting points, then the number of customers (outside then between the cutting points). This study provides insights for researchers to develop more effective solution methods.

Keywords: Travelling Salesman Problem, Genetic Algorithm, Regression Machine Learning, Crossover

* Corresponding Author, E-mail: prita.meilanitasari@uisi.ac.id

1. INTRODUCTION

To solve real optimization problems to ensure business profitability, e.g., travelling salesman problem (TSP) and vehicle routing problem, various optimization (operations research) methods are developed. Some methods guarantee the optimality of the solutions (e.g., exact methods (Riaventin *et al.*, 2021), while some other methods show that even with a much shorter computational time, close-to-optimal solutions could still be produced (e.g., metaheuristics (Meilanitasari and Shin, 2021), algorithms (Singgih and Kim, 2020)). One of the most used metaheuristics for solving routing problem is the genetic algorithm (GA) (Gars *et al.*, 2024; Ojeda Rios and Xavier, 2025; Ostermeier *et al.*, 2021; Shen *et al.*, 2022).

Most studies have a purpose to either produce the best solution quality that is usually defined as the best-known solution (regardless of the computational time) or reduce the computational time while still producing the best possible solution quality within the given time (Lee *et al.*, 2024; Xie *et al.*, 2025; Zhang and Woensel, 2023). Such an approach somehow fails to evaluate how well each part of the algorithm performs within the algorithm run because the method ignores how well the solution quality is improved after any operator is applied. Most studies focus on assessing their proposed method by observing the quality of the final solution and contrasting it with state-of-the-art methods. Some studies would then try to perform good crossover operations, e.g., undirected bipartite graph-based crossover operator (Nazif and Lee, 2012), adjusted order crossover that ensured copying more related nodes into the new sequence (Hvattum, 2022), an updated version of the well-known partially mapped crossover operator (Koohestani, 2020), two-layer chromosome-based crossover that captured more amount of necessary information simultaneously (Karakatič, 2021), adaptive crossover rate that adjusted the number of applied crossover operators based on the current quality of the solutions (Li *et al.*, 2024), modified crossover operators that applied arithmetic operations to allow better convergence (Naskar *et al.*, 2025), and multiple crossover operators which would aim to benefit from the advantage of each type of operator (Zhang *et al.*, 2025). It was common for previous studies to test several crossover operators and identify the ones producing the best solutions after the whole GA run, e.g., Kaelo and Ali (2007), Kaya (2011), Li *et al.* (2023), and Derouiche *et al.* (2025).

All previous studies above still do not answer fundamental questions about (1) why the crossover operator performs well, (2) how well the crossover operator performs, and (3) what values of parameters must be set for the crossover operators to perform well. The studies above only focused on attempting to produce good crossover operators that might work well and proved their effectiveness after a complete run. Without answering

such questions, it would be difficult to understand why and how crossover operators perform well in improving the solution quality, e.g., in a GA run, and to produce an effective operator.

The purpose of this study is to provide answers to the research questions above. As the initial study, this study focuses on the TSP, which has no vehicle capacity restriction and allows an easy understanding of the relationships between the nodes in the chromosome (solution representation). To the best knowledge of the authors, this study is the first to evaluate the effectiveness of GA's crossover operators using regression machine learning methods. This study is important to ensure understanding important parts of the crossover operators that must be considered for designing good operators.

When contrasted with statistical correlation analysis that only shows the relationship between features, implementing regression machine learning methods together with operations research approaches enables online learning, in which the operators in the optimization methods could be assisted by machine learning-based prediction to ensure obtaining better solutions (Szénási and Légrádi, 2024). The inclusion of machine learning methods to apply better operators in the GA initiates a whole new research field that could properly deal with the high uncertainty during the search process. Such an effort to utilize machine learning techniques to improve the quality of metaheuristics has been attempted by recent studies in several fields, e.g., autonomous ridesharing (Bongiovanni *et al.*, 2022), distribution system design (Xiao *et al.*, 2022), and global optimization for engineering problems (Farahmand-Tabar and Ashtari, 2024).

In this study, important parameters of the crossover operator are defined. Every time a crossover operator is applied, all the listed parameter values are recorded together with the amount of improvement in the objective value (the difference between the best new objective value and the best objective value before the crossover operation was applied). Those collected input data are then used to predict the target data by using regression machine learning models. As the result, this study successfully identifies the proper values of the parameters, which would improve the objective values the most.

The structure of this study is as follows. Section 2 describes the method that is solved using GA and proposes the prediction framework using regression machine learning models. Section 3 presents the numerical experiment results. Finally, Section 4 concludes the study.

2. METHOD

To evaluate the performance of crossover operators, this study discusses a Traveling Salesman Problem (TSP) case study. It focuses on route determination and consid-

ers the number of demands that must be met, while ignoring vehicle capacity. The suggested Cross²³r Analysis framework model is shown in Figure 1. The crossover and mutation operators are applied within the GA. Then, the crossover operator is assessed. Data analytics uses feature collection for analysis and regression machine learning-based crossover operator evaluation.

2.1 Genetic Algorithm (GA) for Traveling Salesman Problem (TSP)

¹ The GA procedure²⁰ used to solve the TSP is illustrated in Figure 2. The route starts from a depot and ends

at the same depot. In the considered TSP, the total distance is minimized, e.g., as shown in Figure 2. Initially, chromosomes (as many as the *population_size*) are generated randomly. To produce potentially better new solutions, crossover and mutation operators are applied, and the GA procedure is continued for *num_of_population* iterations (populations).

At the start of each population, n solutions with the best objective values are simply selected and stored as candidates for crossover operations. In the next step, crossover operators are applied by considering the initially selected n best solutions as parent chromosomes. Considering that chromosomes with least total distance must be given a

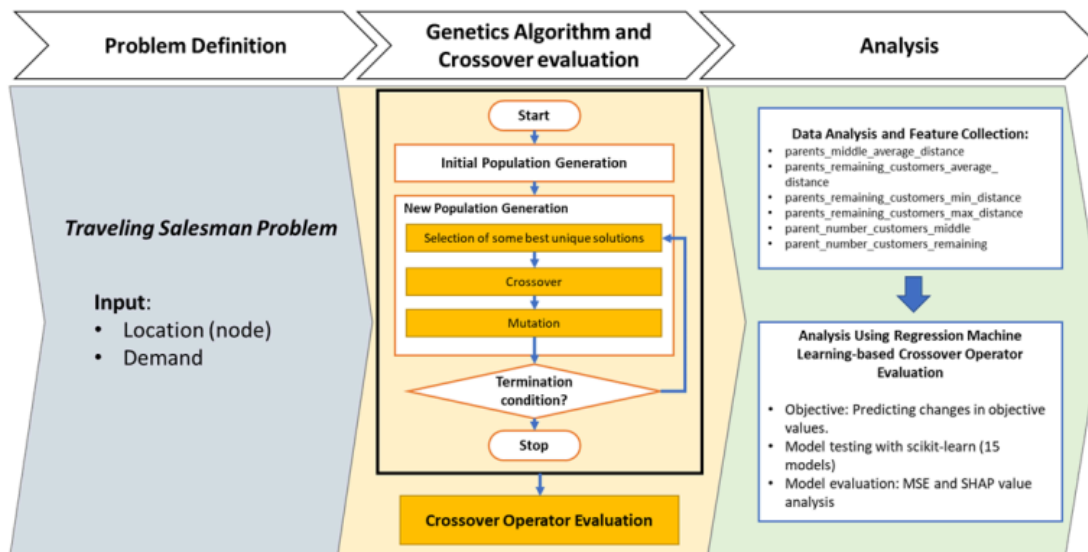


Figure 1. A framework of the crossover analysis research architecture.

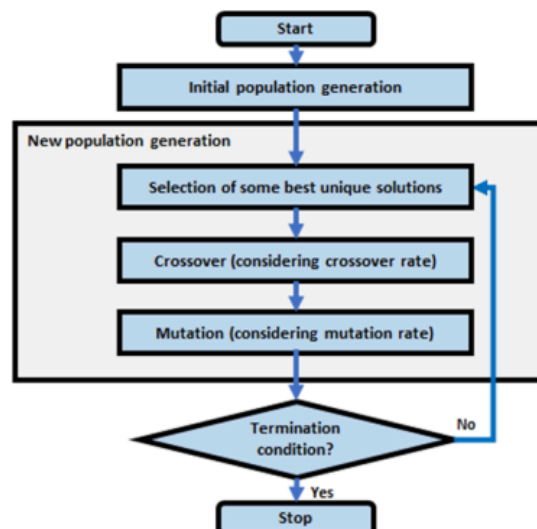


Figure 2. TSP-GA model framework.

higher selection probability, an *inverted objective value* for each parent chromosome is calculated in Equation (1), then the value is used to calculate the selection probability for that chromosome, as calculated in Equation (2).

$$iov = \frac{1}{ov} \quad (1)$$

$$sp = \frac{iov}{\text{total } iov \text{ of all best chromosomes}} \quad (2)$$

where:

iov : *inverted objective value*
ov : *objective value*
sp : *selection probability*

A random number between 0 and 1 is generated. If its value is less than the crossover rate, the crossover operation is applied. Otherwise, the parent chromosomes are set as the result of the crossover operation (stored as the new child chromosomes). If the two-point crossover operator is applied, two cutting points are random selected. The parts of the parents within the crossover points are copied to the child chromosomes. The remaining unfilled parts of the new child chromosomes are then obtained from other parent chromosomes (Figure 3). The crossover operator is applied until the number of chromosomes in the current population equals to *population_size*.

The mutation operator is applied when the random number is less than the mutation rate. The mutation is performed by randomly selecting two values within a parent chromosome and exchanging the positions of both values to produce a child chromosome (Figure 3).

2.2 Regression Machine Learning-based Crossover Operator Evaluation

The objective value of the TSP-GA (the new best total distance minus the previous best total distance) is predicted in this study using regression machine learning

model.

The input data represent the characteristics of the solution before the crossover is performed. Each input data is presented as follows.

- *parents_middle_average_distance_value*: the average distance of subsequent customers within two cutting points in the two-point crossover;
- *parents_middle_average_distance_percentage*: the value of *parents_middle_average_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_average_distance_value*: the average distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_average_distance_percentage*: the value of *parents_remaining_customers_average_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_min_distance_value*: the minimum distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_min_distance_percentage*: the value of *parents_remaining_customers_min_distance_value* divided by the length of the area for the data set;
- *parents_remaining_customers_max_distance_value*: the maximum distance between each possible pair of customers outside the cutting point area (customers at the left of the first cutting point and customers at the right of the second cutting point);
- *parents_remaining_customers_max_distance_percentage*: the value of *parents_remaining_customers_max_distance_value* divided by the length of the area for the data set;
- *parent_number_customers_middle_value*: the number of customers in the cutting area of the chromosome;
- *parent_number_customers_middle_percentage*: the value of *parent_number_customers_middle_value* divided by the number of customer nodes in the data set;
- *parent_number_customers_remaining_value*: the number of customers outside the cutting area in the chromosome.
- *parent_number_customers_remaining_percentage*: the value of *parent_number_customers_remaining_value* divided by the number of customer nodes in the data set.

When different datasets are considered, the number of customers is added as an input feature. The predicted output is the total distance of the route.

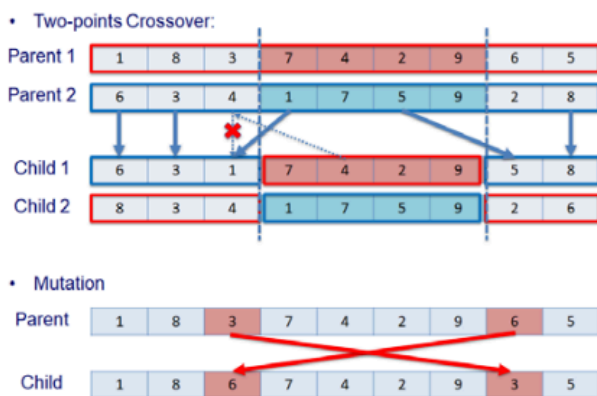


Figure 3. Crossover and mutation operators for this study.

The considered features are expected to highly affect the quality of the two-points crossover. Two important parts of the chromosomes are the part in the middle of both cutting points and the remaining area outside the area between the cutting points. The main idea of the crossover is to maintain the good part from parent chromosomes, which refers to the middle part between the cutting points. It is expected that through the random cutting and combination process in the crossover, the good middle parts would remain in the child chromosomes and eventually lead the search to a promising direction. Another way to evaluate the crossover would be understanding the effect of the distances between nodes outside the middle part of the chromosome on the objective value. A good crossover would select middle parts with good partial objective values, while having outer parts with bad partial objective values. In contrast, the crossover would produce worse results when the middle parts contain bad sequences, but have the good outer parts (with short travel times). The number of customers in each part could have a high impact on the crossover results as well. All these related features are considered in this study.

Several regression machine learning models are tested, and then the best models are reported in this section. The predictions were developed with the following scikit-learn (Scikit-learn 17022) regression machine learning models such as: (1) random forest regression (Kim *et al.*, 2020), (2) linear regression (Abid Almubaidin *et al.*, 2023), (3) Ridge (Scikit-learn, 2023; Wang *et al.*, 2023), (4) ElasticNet (Scikit-learn, 2023; Yu *et al.*, 2021), (5) Lars (Scikit-learn, 2023; Lee and Jun, 2018), (6) LassoCV (Scikit-learn, 2023; Kramlinger *et al.*, 2023), (7) LassoLars (Scikit-learn, 2023), (8) OrthogonalMatchingPursuit (Scikit-learn, 2023; Polat and Kayhan, 2018), (9) ARDRegression (Sandhu *et al.*, 2017), (10) BayesianRidge (da Silva *et al.*, 2021), (11) HuberRegressor (Scikit-learn, 2023), (12) RANSACRegressor (Scikit-learn, 2023; Zhu *et al.*, 2023), (13) TheilSe-

nRegressor (Scikit-learn, 2023), (14) TweedieRegressor (Scikit-learn, 2023), (15) PassiveAggressiveRegressor (Scikit-learn, 2023; Jorge and Paredes, 2018), ExtraTreeRegressor (Ahmad *et al.*, 2018), GradientBoostingRegressor (Moghadaspoor *et al.*, 2026), HistGradientBoostingRegressor (Alawi *et al.*, 2024), AdaBoostRegressor (Kumari *et al.*, 2025), support vector regressor with radial basis function (Taiwo *et al.*, 2024), KNeighborsRegressor, DecisionTreeRegressor (Zade *et al.*, 2023), MLPRegressor (Zade *et al.*, 2023), and XGBoostRegressor (Moghadaspoor *et al.*, 2026). The linear and non-linear methods are implemented to properly find the best method for the data sets. At the time of writing, scikit-learn has more than 120,000 citations based on Google Scholar. Its minimal dependencies and ease of use allow a high reproducibility rate in many studies.

Data is generated as follows. The GA is run following the procedure explained in Section 2. As a result, 1,781 data records are recorded and used for the machine learning analysis. In the analysis, the models are compared based on the mean squared error (MSE), and the SHAP value analysis (SHAP, 2025a, 2025b) is conducted to identify the most important input feature for the prediction.

3. NUMERICAL EXPERIMENTS

Different data set sizes are generated randomly, as shown in Table 1. To produce the data, the parameters for the GA are listed in Table 2. Experiments using the machine learning models are conducted separately for each data set size, then for the whole data sets combined together to allow generalizing the findings. In the experiments 10-fold cross validation and MinMaxScaler are applied. Performance metrics obtained when dealing with data set 1, 2, 3, and all of them are presented in Tables 3, 4, 5, and 6, respectively.

Table 1. Data set sizes for the numerical experiments

Data set	Number of nodes	Area on which the nodes are placed	Number of instances
1	20	20 x 20	538
2	200	100 x 100	614
3	1,000	400 x 400	629

Table 2. Parameters of the genetic algorithm

Parameter	Value
Number of populations	20
Population size	100
Number of initial chromosomes	20% * population size
Crossover rate	0.8
Mutation rate	0.2

Table 3. Performance metrics for regression machine learning models on data set 1

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	2.22	6.07	8.36	59.32	2.89	7.70	87.70	11.81
LinearRegression	6.61	6.71	64.05	66.23	8.00	8.14	5.71	0.93
Ridge	6.62	6.71	64.08	65.99	8.01	8.12	5.67	1.28
ElasticNet	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
Lars	6.89	7.01	72.71	75.80	8.53	8.71	-6.76	-17.60
Lasso	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
LassoLars	6.88	6.89	67.92	68.22	8.24	8.26	0.00	-2.59
OrthogonalMatchingPursuit	6.80	6.82	66.95	67.34	8.18	8.21	1.43	-1.03
ARDRegression	6.64	6.75	64.24	66.51	8.01	8.16	5.43	0.45
BayesianRidge	6.64	6.72	64.20	65.94	8.01	8.12	5.49	1.34
HuberRegressor	6.53	6.65	64.92	67.36	8.06	8.21	4.43	-0.34
RANSACRegressor	7.18	7.34	93.21	92.68	9.65	9.63	-37.24	-37.25
TheilSenRegressor	6.61	6.72	64.21	66.51	8.01	8.16	5.47	0.52
TweedieRegressor	6.86	6.88	67.64	68.05	8.22	8.25	0.43	-2.29
PassiveAggressiveRegressor	7.52	7.54	86.42	87.89	9.30	9.38	-27.27	-31.09
ExtraTreesRegressor	0.10	6.15	0.44	64.45	0.66	8.03	99.36	3.56
GradientBoostingRegressor	4.13	6.01	26.85	59.85	5.18	7.74	60.46	10.96
HistGradientBoostingRegressor	3.22	6.23	16.94	63.47	4.12	7.97	75.06	4.56
AdaBoostRegressor	6.27	6.64	52.85	62.90	7.27	7.93	22.19	4.86
SVR_RBF	6.53	6.65	67.31	69.20	8.20	8.32	0.90	-2.97
KNeighborsRegressor	5.22	6.72	44.64	72.82	6.68	8.53	34.28	-7.81
DecisionTreeRegressor	0.10	7.24	0.44	100.37	0.66	10.02	99.36	-52.27
MLPRegressor	6.61	6.74	62.75	65.68	7.92	8.10	7.61	1.42
XGBRegressor	1.44	6.16	3.85	63.97	1.96	8.00	94.34	4.63

Table 4. Performance metrics for regression machine learning models on data set 2

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	23.5	65.8	993.2	7991.8	31.5	89.4	86.8	-9.9
LinearRegression	62.0	62.9	6907.5	7117.4	83.1	84.4	8.0	-1.7
Ridge	62.0	62.8	6910.8	7103.4	83.1	84.3	8.0	-1.3
ElasticNet	63.8	64.0	7357.3	7418.9	85.8	86.1	2.1	-3.0
Lars	62.2	62.5	6962.6	7070.0	83.4	84.1	7.3	-0.6
Lasso	63.1	63.5	7175.6	7304.2	84.7	85.5	4.5	-2.7
LassoLars	63.1	63.5	7187.4	7306.6	84.8	85.5	4.3	-2.7
OrthogonalMatchingPursuit	63.4	63.8	7212.2	7327.4	84.9	85.6	4.0	-4.3
ARDRegression	62.0	62.7	6915.5	7085.3	83.2	84.2	7.9	-0.8
BayesianRidge	62.0	62.7	6919.2	7109.5	83.2	84.3	7.9	-1.1
HuberRegressor	61.4	62.4	7006.7	7194.8	83.7	84.8	6.7	-1.1
RANSACRegressor	64.6	64.5	8121.7	7792.3	90.1	88.3	-7.8	-11.3
TheilSenRegressor	61.9	62.7	6925.4	7125.9	83.2	84.4	7.8	-1.3
TweedieRegressor	64.0	64.2	7400.0	7451.5	86.0	86.3	1.5	-3.3
PassiveAggressiveRegressor	62.4	62.5	7161.7	7072.6	84.6	84.1	4.6	-0.1
ExtraTreesRegressor	0.0	67.8	0.0	8514.4	0.0	92.3	100.0	-18.2
GradientBoostingRegressor	41.1	66.3	2975.9	8659.1	54.6	93.1	60.4	-12.3
HistGradientBoostingRegressor	33.1	67.1	1973.8	8656.5	44.4	93.0	73.7	-18.3
AdaBoostRegressor	60.2	66.6	5732.3	7998.3	75.7	89.4	23.6	-13.0
SVR_RBF	62.2	62.7	7475.6	7564.9	86.5	87.0	0.5	-3.2
KNeighborsRegressor	53.3	66.7	5406.8	8358.1	73.5	91.4	28.0	-17.6
DecisionTreeRegressor	0.0	91.2	0.0	16358.5	0.0	127.9	100.0	-121.2
MLPRegressor	61.5	62.9	6794.3	7167.0	82.4	84.7	9.5	-2.9
XGBRegressor	12.9	67.4	319.7	8830.6	17.9	94.0	95.7	-18.7

Table 5. Performance metrics for regression machine learning models on data set 3

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	249.11	721.31	151441.08	1275915.07	389.15	1129.56	88.56	-26.38
LinearRegression	742.27	760.60	1200368.72	1257147.92	1095.61	1121.23	9.51	-53.95
Ridge	741.11	759.49	1202529.30	1256671.97	1096.60	1121.01	9.35	-53.51
ElasticNet	777.36	792.73	1301802.23	1346299.48	1140.97	1160.30	1.88	-65.68
Lars	744.94	767.94	1208269.22	1276741.75	1099.21	1129.93	8.89	-60.23
Lasso	742.06	761.89	1201521.10	1260090.02	1096.14	1122.54	9.42	-54.45
LassoLars	744.95	768.14	1209326.87	1276120.10	1099.69	1129.65	8.80	-59.46
OrthogonalMatchingPursuit	763.86	780.11	1280113.67	1330966.94	1131.42	1153.68	3.51	-77.06
ARDRegression	743.86	764.21	1203485.08	1264248.95	1097.03	1124.39	9.28	-54.98
BayesianRidge	741.14	759.96	1201820.72	1257419.14	1096.28	1121.35	9.40	-53.85
HuberRegressor	699.27	714.46	1276669.46	1318829.37	1129.90	1148.40	3.75	-20.11
RANSACRegressor	773.51	776.35	1784477.92	1567055.92	1335.84	1251.82	-33.80	-84.10
TheilSenRegressor	43 (10 ⁹)	41 (10 ⁹)	19 (10 ²¹)	18 (10 ²¹)	141 (10 ⁹)	135 (10 ⁹)	-1.57 (10 ¹⁶)	-423 (10 ¹⁶)
TweedieRegressor	781.42	796.47	1311187.56	1354636.92	1145.07	1163.89	1.18	-66.00
PassiveAggressiveRegressor	707.41	714.61	1380041.85	1394199.76	1174.75	1180.76	-3.91	-18.73
ExtraTreesRegressor	0.00	757.95	0.00	1437740.56	0.00	1199.06	100.00	-36.36
GradientBoostingRegressor	435.34	751.54	380870.44	1421742.25	617.15	1192.37	71.29	-48.96
HistGradientBoostingRegressor	346.25	765.08	273954.94	1426434.43	523.41	1194.33	79.32	-55.71
AdaBoostRegressor	734.09	866.56	878118.11	1517805.48	937.08	1231.99	34.10	-130.99
SVR RBF	717.11	721.31	1434919.21	1446023.09	1197.88	1202.51	-8.06	-16.98
KNeighborsRegressor	590.70	733.18	827013.10	1258896.83	909.40	1122.01	37.64	-54.87
DecisionTreeRegressor	0.00	972.59	0.00	2350582.54	0.00	1533.16	100.00	-250.05
MLPRegressor	734.49	760.41	1161207.30	1230558.99	1077.59	1109.31	12.42	-75.35
XGBRegressor	112.23	763.05	25508.40	1448972.31	159.71	1203.73	98.08	-50.74

Table 6. Performance metrics for regression machine learning models on all data sets

Model	Train MAE	Test MAE	Train MSE	Test MSE	Train RMSE	Test RMSE	Train R ² (%)	Test R ² (%)
RandomForestRegressor	98.9	291.7	55842.7	526353.5	236.3	725.5	89.7	-13.8
LinearRegression	310.9	339.9	434600.8	486185.8	659.2	697.3	20.1	-6826.6
Ridge	297.5	314.4	442188.9	486585.7	665.0	697.6	18.7	-485.5
ElasticNet	321.8	332.2	481246.4	508580.2	693.7	713.1	11.6	-2692.2
Lars	338494.0	416814.2	221 (10 ¹⁰)	420 (10 ¹⁰)	148 (10 ⁴)	205 (10 ⁴)	-399 (10 ⁶)	-26 (10 ⁸)
Lasso	298.2	314.7	443031.9	488952.9	665.6	699.3	18.5	-835.1
LassoLars	298.1	314.7	443044.7	489015.9	665.6	699.3	18.5	-832.4
OrthogonalMatchingPursuit	312.6	329.6	471069.2	507923.4	686.3	712.7	13.5	-680.1
ARDRegression	297.3	316.3	437940.2	485443.4	661.8	696.7	19.5	-283.5
BayesianRidge	297.6	317.0	437709.4	485186.7	661.6	696.6	19.5	-489.3
HuberRegressor	272.7	281.7	483311.0	505395.3	695.2	710.9	11.2	-43.3
RANSACRegressor	498.2	773.1	1085469.1	1722986.0	1041.9	1312.6	-99.8	-26899.8
TheilSenRegressor	302.4	333.8	442784.3	487451.3	665.4	698.2	18.6	-13253.4
TweedieRegressor	338.1	349.0	495021.2	520156.3	703.6	721.2	9.1	-7114.8
PassiveAggressiveRegressor	275.0	281.7	497727.7	518314.4	705.5	719.9	8.6	-17.3
ExtraTreesRegressor	0.0	297.0	0.1	541539.2	0.4	735.9	100.0	-17.6
GradientBoostingRegressor	183.4	314.7	146422.2	632543.6	382.7	795.3	73.4	-128.5
HistGradientBoostingRegressor	146.1	290.1	107817.6	500860.3	328.4	707.7	79.8	-42.4
AdaBoostRegressor	304.7	364.4	322521.8	587055.0	567.9	766.2	41.7	-11379.8
SVR RBF	280.7	282.7	542527.7	551146.8	736.6	742.4	0.5	-22.3
KNeighborsRegressor	233.1	298.3	290406.1	486843.6	538.9	697.7	46.4	-23.1
DecisionTreeRegressor	0.0	389.0	0.1	897974.0	0.4	947.6	100.0	-112.6
MLPRegressor	306.6	337.2	451621.1	519803.3	672.0	721.0	16.9	-239.8
XGBRegressor	58.8	311.2	10420.8	608625.1	102.1	780.1	98.1	-41.0

To initially assess the experiment results, the R^2 score obtained using the test data in each data set is observed. The most promising R^2 score using the test data (11.81%) is obtained in data set 1, in which 20 customer nodes are considered (Table 3), when the RandomForestRegressor is applied. When the value is contrasted with the R^2 score for the train data, the large gap between the scores shows a high overfitting of the model. However, a positive R^2 score shows a promising research result, considering this study is the first to attempt observing the quality of the crossover operators in the GA. A low R^2 score is acceptable in some fields, e.g., social science and psychology, which is very affected by the complexity of the data (Gupta *et al.*, 2024). Considering that R^2 score refers to the degree to which the variation of the target variable is influenced by the variation in the input features, the metric is strongly influenced by the variation in the input features. However, the low R^2 score does not indicate a weak relationship between the variables (Hamilton *et al.*, 2015). Given that the attempt to understand new input features of the crossover operator in the GA is a completely new problem, a lot of effort might be necessary for studies to understand the relationships between variables properly. To ensure a better prediction, it is ne-

cessary to (1) identify more related input features, which might be specifically designed to understand hidden relationships that better predict the total distance, and (2) group the data based on specific characteristics to allow better prediction using more similar input features. Correlations between variables when using this data set are presented in Figure 4.

When the predictions are conducted in other data sets (Tables 4 and 5), the models fail to find good predictions, which is reflected by the negative R^2 values, which represent poor prediction capabilities (Zhang *et al.*, 2026). The negative R^2 values appear when the models fail to capture the nonlinear behavior of the data or fail to explain the data variations in reality (Mittlböck and Walldör, 2000) because the performance of the prediction models is worse than when the average values of the target variables are considered (Chicco *et al.*, 2021). A similar result is obtained when all data sets are considered simultaneously (Table 6), that concludes that the generalization could not yet be conducted properly and needs further observations.

Figure 5 shows the SHAP analysis result. In Figure 5, the low value of each feature is represented in blue, and its high value is represented in red. The negative values in

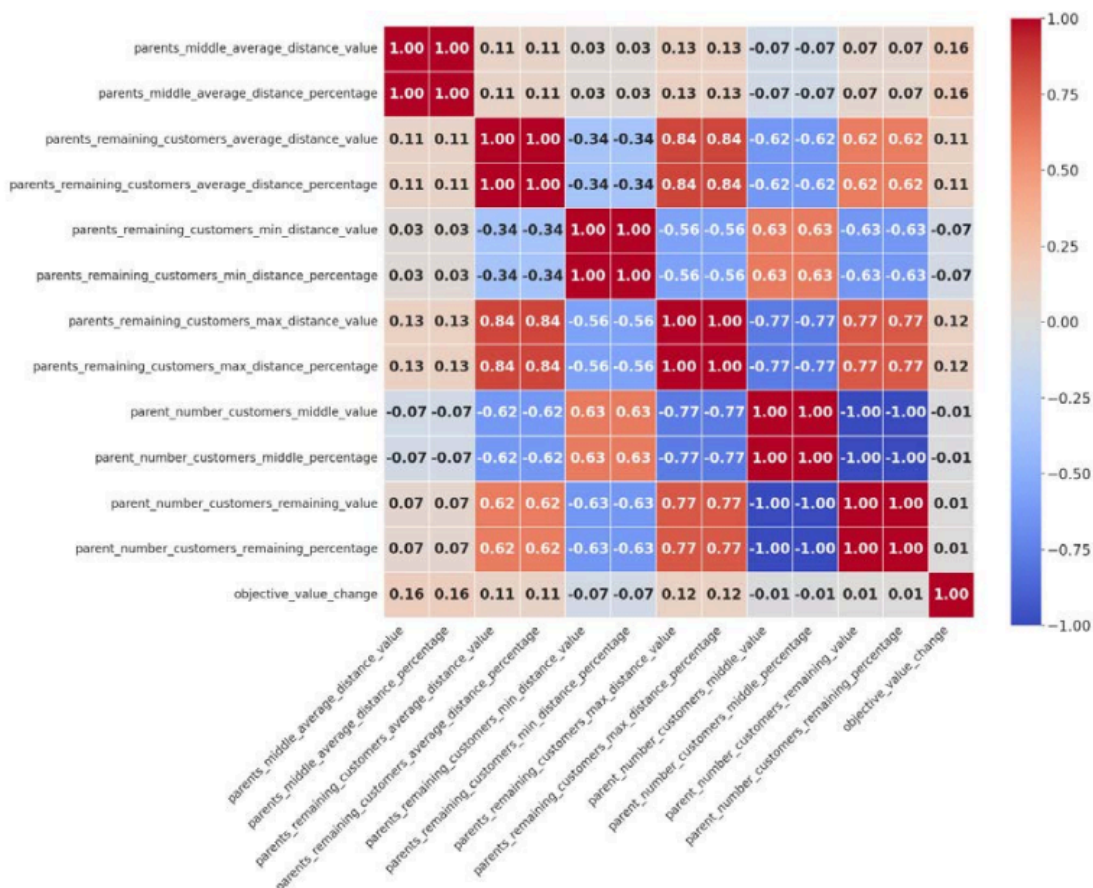


Figure 4. Correlations between variables.

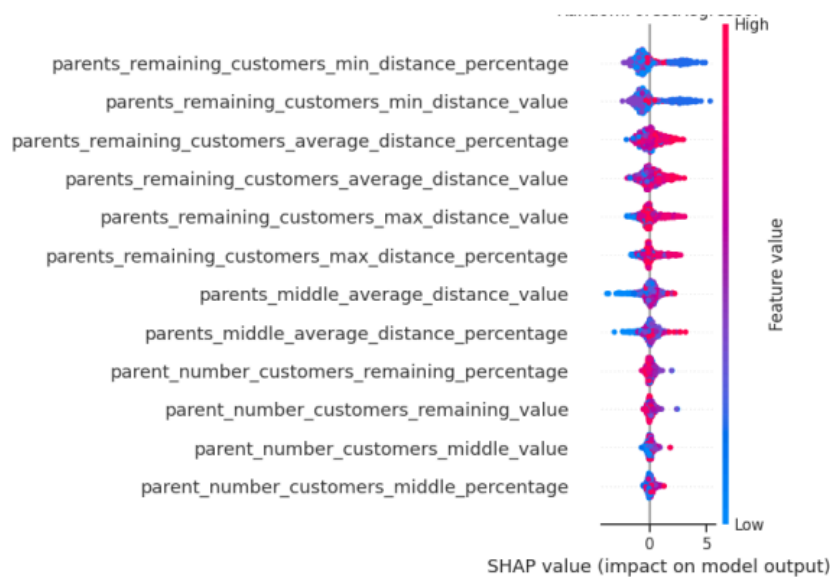


Figure 5. Feature importance analysis using SHAP value.

the x-axis refer to the reduction in prediction quality (the increase in MSE), while the positive values refer to the increase in prediction quality (the decrease in MSE). Based on Figure 5, it is shown that the most important input features to predict the change in the total distance are (1) minimum, average, then maximum distances between each pair of remaining customers outside the cutting points, (2) average distance between customers within the cutting points, then (3) number of customers (outside then between the cutting points).

The prioritized features are further explained as follows. First, the distance between each pair of remaining customers is important because its value shows the opportunity to improve outside parts of the new chromosomes, e.g., when customers with short distances are placed subsequently. Such a conclusion aligns with the order of importance for the minimum, average, then maximum values. Second, the total distance measured between subsequent customers within the cutting points affects the quality of the new chromosomes by ensuring that subsequent customers with short distances are maintained in the next chromosomes. The findings are interesting. The remaining customers feature that is more important than the customers within the cutting points indicates that when applying the crossover operators, it is necessary to sufficiently consider including customers with short distances into the area between the cutting points. Third, the number of customers (either inside or outside the cutting points) has the least impact when predicting the total distances. Such an observation is acceptable considering that the distance measure itself should be observed better than the number of customers because the objective (total distances) is strongly related to the distance values. As a conclusion, such important features must be considered

properly when applying the two-point crossover operators in the GA to produce generally good solutions in the long run, while reducing the number of required populations.

4. CONCLUSIONS

This study observes the potential of understanding the behavior of the crossover operator in the genetic algorithm for solving the travelling salesman problem. The numerical experiments show that the distance between customers outside the cutting points of the chromosome is the most important feature to predict the change in the best objective values before and after the crossover operator is applied, followed by the distance between customers within the cutting points, then the number of customers. The limitation of this study is only considering the two-point crossover operator. For further research topics, it would be necessary to consider more variants of crossover operators and more objectives, design more features that could significantly affect the prediction quality, design better regression machine learning methods with better generalization (e.g., by introducing more features related to the problem sizes), and discuss how effective genetic algorithms can be designed based on the insights found in this study.

REFERENCES

- Abid Almubaidin, M. A., Latif, S. D., Balan, K., Ahmed, A. N., and El-Shafie, A. (2023), Enhancing sediment transport predictions through machine learning-based multi-scenario regression models, *Results in*

- Engineering*, **20**, 101585.
- Ahmad, M. W., Reynolds, J., and Rezgui, Y. (2018), Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees, *Journal of Cleaner Production*, **203**, 810-821.
- Alawi, O. A., Kamar, H. M., Salih, S. Q., Abba, S. I., Ahmed, W., Homod, R. Z., Jamei, M., Shafik, S. S., and Yaseen, Z. M. (2024), Development of optimized machine learning models for predicting flat plate solar collectors thermal efficiency associated with Al₂O₃-water nanofluids, *Engineering Applications of Artificial Intelligence*, **133**(B), 108158.
- Bongiovanni, C., Kaspi, M., Cordeau, J.-F., and Geroliminis, N. (2022), A machine learning-driven two-phase metaheuristic for autonomous ridesharing operations, *Transportation Research Part E*, **165**, 102835.
- Chicco, D., Warrens, M. J., and Jurman, G. (2021), The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, *PeerJ Computer Science*, **7**, e623.
- da Silva, F. A., Viana, A. P., Correa, C. C. G., Santos, E. A., de Oliveira, J. A. V. S., Andrade, J. D. G., Ribeiro, R. M., and Glória, L. S. (2021), Bayesian ridge regression shows the best fit for SSR markers in *Psidium guajava* among Bayesian models, *Scientific Reports*, **11**(1), 13639.
- Derouiche, H., Elarbi, M., and Bechikh, S. (2025), Deep crossover schemes for genetic algorithms: Investigations on the travel salesman problem, *Swarm and Evolutionary Computation*, **98**, 102094.
- Farahmand-Tabar, S. and Ashtari, P. (2024), Intelligent cross-entropy optimizer: A novel machine learning-based meta-heuristic for global optimization, *Swarm and Evolutionary Computation*, **91**, 101739.
- Garside, A. K., Ahmad, R., and Muhtazaruddin, M. N. B. (2024), A recent review of solution approaches for green vehicle routing problem and its variants, *Operations Research Perspectives*, **12**, 100303.
- Hamilton, D. F., Ghert, M., and Simpson, A. H. R. W. (2015), Interpreting regression models in clinical outcome studies, *Bone & Joint Research*, **4**(9), 152-153.
- Hvattum, L. M. (2022), Adjusting the order crossover operator for capacitated vehicle routing problems, *Computers & Operations Research*, **148**, 105986.
- Jorge, J. and Paredes, R. (2018), Passive-aggressive online learning with nonlinear embeddings, *Pattern Recognition*, **79**, 162-171.
- Kaelo, P. and Ali, M. M. (2007), Integrated crossover rules in real coded genetic algorithms, *European Journal of Operational Research*, **176**(1), 60-76.
- Karakatič, S. (2021), Optimizing nonlinear charging times of electric vehicle routing with genetic algorithm, *Expert Systems with Applications*, **164**, 114039.
- Kaya, M. (2011), The effects of two new crossover operators on genetic algorithm performance, *Applied Soft Computing*, **11**(1), 881-890.
- Kim, S. W., Lee, Y. G., Tama, B. A., and Lee, S. (2020), Reliability-enhanced camera lens module classification using semi-supervised regression method, *Applied Sciences*, **10**(11), 3832.
- Kramlinger, P., Schneider, U., and Krivobokova, T. (2023), Uniformly valid inference based on the Lasso in linear mixed models, *Journal of Multivariate Analysis*, **198**, 105230.
- Koohestani, B. (2020), A crossover operator for improving the efficiency of permutation-based genetic algorithms, *Expert Systems with Applications*, **151**, 113381.
- Kumari, S., Jha, R., Ray, A., Jena, J. J., Gourisaria, M. K., and Bandyopadhyay, A. (2025), An explainable AI approach for potato plant disease detection using enhanced feature engineering, *Procedia Computer Science*, **258**, 3570-3579.
- Lee, S. and Jun, C.-H. (2018), Fast incremental learning of logistic model tree using least angle regression, *Expert Systems with Applications*, **97**, 137-145.
- Lee, Y., Thyagarajan, K., Pinto, J. M., Charitopoulos, V. M., and Papageorgiou, L. G. (2024), Towards efficient solutions for vehicle routing problems for oxygen supply chains, *Computers & Chemical Engineering*, **191**, 108827.
- Li, D., Yang, C., and Zou, D. (2023), A nondominated sorting genetic algorithm III with three crossover strategies for the combined heat and power dynamic economic emission dispatch with or without prohibited operating zones, *Engineering Applications of Artificial Intelligence*, **123**(C), 106443.
- Li, J., Liu, R., and Wang, R. (2024), Handling dynamic capacitated vehicle routing problems based on adaptive genetic algorithm with elastic strategy, *Swarm and Evolutionary Computation*, **86**, 101529.
- Meilanitasari, P. and Shin, S.-J. (2021), A review of prediction and optimization for sequence-driven scheduling in job shop flexible manufacturing systems, *Processes*, **9**(8), 1391.
- Mittlböck, M. and Waldhör, T. (2000), Adjustments for R²-measures for Poisson regression models, *Computational Statistics & Data Analysis*, **34**(4), 461-472.
- Moghadaspoor, B., Bozorgi-Amiri, A., Tavakkoli-Moghaddam, R., and Bashiri, M. (2026), An ensemble learning method to predict a hydrogen production rate in the viable wastewater-hydrogen supply chain network design, *Energy*, **345**, 139950.
- Naskar, A., Ghosh, S., Kundu, M., and Sarkar, R. (2025), Feature selection using guided population based ge-

- netic algorithm with modified crossover and parent selection, *Applied Soft Computing*, **172**, 112872.
- Nazif, H. and Lee, L. S. (2012), Optimised crossover genetic algorithm for capacitated vehicle routing problem, *Applied Mathematical Modelling*, **36**(5), 2110-2117.
- Ojeda Rios, B. H. and Xavier, E. C. (2025), Metaheuristic approaches for the stochastic capacitated multi-depot vehicle routing problem with pickup and delivery, *Expert Systems with Applications*, **290**, 128258.
- Ostermeier, M., Henke, T., Hübner, A., and Wäscher, G. (2021), Multi-compartment vehicle routing problems: State-of-the-art, modeling framework and future directions, *European Journal of Operational Research*, **292**(3), 799-817.
- Gupta, A., Stead, T. S., and Ganti, L. (2024), Determining a meaningful R-squared value in clinical medicine, *Academic Medicine & Surgery*.
- Polat, Ö. and Kayhan, S. K. (2018), High-speed FPGA implementation of orthogonal matching pursuit for compressive sensing signal reconstruction, *Computers & Electrical Engineering*, **71**, 173-190.
- Riaventin, V. N., Cahyo, S. D., and Singgih, I. K. (2021), A model for developing existing ports considering economic impact and network connectivity, *Sustainability*, **13**(7), 3705.
- Sandhu, R., Pettit, C., Khalil, M., Poirer, D., and Sarkar, A. (2017), Bayesian model selection using automatic relevance determination for nonlinear dynamical systems, *Computer Methods in Applied Mechanics and Engineering*, **320**, 237-260.
- Scikit-learn (2022), 1. Supervised learning. Scikit-Learn. Available from: https://scikit-learn.org/stable/supervised_learning.html.
- Scikit-learn (2023), API Reference, cited 2023 Dec 2, Available from: scikit-learn.org.
- SHAP (2025a, October 12), Shap.KernelExplainer. Available from: <https://shap.readthedocs.io/en/latest/generated/shap.KernelExplainer.html>
- SHAP (2025b, October 12), Shap.LinearExplainer. Available from: <https://shap.readthedocs.io/en/latest/generated/shap.LinearExplainer.html>
- Shen, J., Liu, K., Ma, C., Zhao, Y., and Shi, C. (2022), Bibliometric analysis and system review of vehicle routing optimization for emergency material distribution, *Journal of Traffic and Transportation Engineering* (English Edition), **9**(6), 893-911.
- Singgih, I. K. and Kim, B. I. (2020), Multi-type electric vehicle relocation problem considering required battery-charging time, *European J. of Industrial Engineering*, **14**(3), 335-368.
- Szénási, S. and Légrádi, G. (2024), Machine learning aided metaheuristics: A comprehensive review of hybrid local search methods, *Expert Systems with Applications*, **258**, 125192.
- Taiwo, B. O., Fissaha, Y., Hosseini, S., Khishe, M., Karaman, E., Adebayo, B., Sazid, M., Adesida, P. A., Famobuwa, O. V., Faluyi, J. O., and Akinlabi, A. A. (2024), Machine learning based prediction of flyrock distance in rock blasting: A safe and sustainable mining approach, *Green and Smart Mining Engineering*, **1**(3), 346-361.
- Wang, X., Wang, X., Ma, B., Li, Q., Wang, C., and Shi, Y. (2023), High-performance reversible data hiding based on ridge regression prediction algorithm, *Signal Processing*, **204**, 108818.
- Xiao, Z., Zhi, J., and Keskin, B. B. (2022), Towards a machine learning-aided metaheuristic framework for a production/distribution system design problem, *Computers and Operations Research*, **146**, 105897.
- Xie, Y., Chu, K.-F., Lam, A. Y. S., and Iida, F. (2025), A multi-objective evolutionary algorithm with constraint-compliant initialization for energy transport and urban logistics in Electric Vehicle Routing, *Applied Soft Computing*, **183**, 113624.
- Yu, B., Chen, C., Wang, X., Yu, Z., Ma, A., and Liu, B. (2021), Prediction of protein-protein interactions based on elastic net and deep forest, *Expert Systems with Applications*, **176**, 114876.
- Zade, F. A., Ghafurian, M. M., Mesgarpour, M., and Niazmand, H. (2023), Predictive machine learning models for optimization of direct solar steam generation, *Journal of Water Process Engineering*, **56**, 104304.
- Zhang, J. and Woensel, T. V. (2023), Dynamic vehicle routing with random requests: A literature review, *International Journal of Production Economics*, **256**, 108751.
- Zhang, X., Hao, Y., Zhang, L., and Yuan, X. (2025), Application of improved genetic algorithm to vehicle routing problem considering the environmental self-regulation of the freight companies, *Expert Systems with Applications*, **274**, 127010.
- Zhang, P.-F., Zhang, D., Zhao, Q., Tuerxunmaiti, Y., Yang, X., Zhao, X., and Iqbal, M. (2026), Machine learning-guided interpretable modelling of code-based equation for bond strength of composite bars in ultra-high-performance concrete, *Advanced Engineering Informatics*, **73**, 104542.
- Zhu, F., Li, H., Li, J., Zhu, B., and Lei, S. (2023), Unmanned aerial vehicle remote sensing image registration based on an improved oriented FAST and rotated BRIEF- random sample consensus algorithm, *Engineering Applications of Artificial Intelligence*, **126**(B), 106944.

Ivan Kristianto Singgih is the Head of Systems Engineering Laboratory in the Study Program of Industrial Engineering at University of Surabaya (Ubaya), Indonesia. He received his Ph.D. degree in Industrial Engi-

neering from Pusan National University, Busan, Korea, in 2017. He has been a Research Assistant Professor under the Department of Industrial and Management Engineering at Pohang University of Science and Technology (POSTECH), Korea Advanced Institute of Science and Technology (KAIST) and Korea University. His research interests include operations research, smart systems, machine learning, and simulation.

Dr. Eng. Ir. Aditya Rio Prabowo is an Associate Professor in the Department of Mechanical Engineering at Universitas Sebelas Maret (UNS) in Surakarta, Central Java, where he also serves as Head of Laboratory. His research primarily investigates structural crashworthiness and accident-related phenomena in marine environments—such as collisions, groundings, and explosions—with a particular focus on engineering operations in extreme regions, such as the Arctic. He is interested in structural analysis and probabilistic estimation of large systems using numerical methods, including the Finite Ele-

ment and Monte Carlo methods. With an award as one of the top two percent of world researchers and a substantial publication record in impact engineering and marine structures, he is recognized for his technical expertise in offshore-marine structural safety.

Prita Meilanitasari a lecturer in the Logistics Engineering Department in one of Private University in Indonesia. She earned her Ph.D. from Hanyang University Graduate School of Technology & Innovation Management and has a research background in advanced manufacturing and supply chain management. Prita has worked as a research and teaching assistant at Hanyang University and Pukyong National University. Her areas of expertise include advanced manufacturing, logistics systems, data analytics, and supply chain optimization. Her academic and research experience includes developing data-driven models for manufacturing system optimization, logistics management, and supply chain analytics.

ORIGINALITY REPORT

10%

SIMILARITY INDEX

7%

INTERNET SOURCES

8%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	3%
2	Ivan Kristianto Singgih, Moses Laksono Singgih. "Regression Machine Learning Models for the Short-Time Prediction of Genetic Algorithm Results in a Vehicle Routing Problem", World Electric Vehicle Journal, 2024 Publication	1%
3	dokumen.pub Internet Source	1%
4	kpfu.ru Internet Source	1%
5	Submitted to Sydney Institute of Higher Education Student Paper	<1%
6	Ivan Kristianto Singgih, Byung In Kim. "Multi-type electric vehicle relocation problem considering required battery-charging time", European J. of Industrial Engineering, 2020 Publication	<1%
7	Submitted to Telkom University Student Paper	<1%
8	Submitted to Sheffield Hallam University Student Paper	<1%
9	ijtech.eng.ui.ac.id Internet Source	<1%
10	Egba Anwaitu Fraser, Obikwelu R. Okonkwo. "GENETIC ALGORITHM APPROACH FOR OPTIMAL CYCLIC TOUR ROUND THE STATE CAPITALS IN NIGERIA'S NIGER DELTA REGION", International Journal of Research -GRANTHAALAYAH, 2021 Publication	<1%
11	hal.science Internet Source	<1%
12	Wen-Zhao Liu, Cong-Ying Han, Tian-De Guo. "Neural Re-Insertion Improvement Algorithm for Large-Scale TSP", Journal of the Operations Research Society of China, 2026 Publication	<1%

13	link.springer.com Internet Source	<1 %
14	"2nd EAI International Conference on Renewable Energy and Sustainable Manufacturing", Springer Science and Business Media LLC, 2025 Publication	<1 %
15	Lecture Notes in Computer Science, 2015. Publication	<1 %
16	boneandjoint.org.uk Internet Source	<1 %
17	simulation.su Internet Source	<1 %
18	Hoang Nguyen, Tran Dinh Bao, Xuan-Nam Bui, Van-Viet Pham et al. "Measuring and Predicting Blast-Induced Flyrock Using Unmanned Aerial Vehicles and Lévy Flight Technique-Based Jaya Optimization Algorithm Integrated with Adaptive Neuro-Fuzzy Inference System", Natural Resources Research, 2025 Publication	<1 %
19	Yong Wang, Mengyuan Gou, Siyu Luo, Lu Zhen, Xiuling Hei. "Collaborative multi-depot vehicle routing problem with time windows, split delivery, and uncertain demands", Transportation Research Part E: Logistics and Transportation Review, 2026 Publication	<1 %
20	export.arxiv.org Internet Source	<1 %
21	mdpi-res.com Internet Source	<1 %
22	minerva.usc.es Internet Source	<1 %
23	Amin Riazi. "Genetic algorithm and a double-chromosome implementation to the traveling salesman problem", SN Applied Sciences, 2019 Publication	<1 %

Exclude quotes

Off

Exclude matches

Off

Exclude bibliography

On



IVAN KRISTIAN TO SINGGIH _ <ivanksinggi h@staff.ubaya.ac.id>

Industrial Engineering & Management Systems - Manuscript ID IEMS-2025-0128

2 messages

Dong Ho Lee <onbehalfof@manuscriptcentral.com>

Tue, Nov 11, 2025 at 11:10 AM

Reply-To: leman@hanyang.ac.kr

To: ivanksinggi h@staff.ubaya.ac.id

11-Nov-2025

Dear Dr. Singgi h:

Your manuscript entitled "Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis" has been successfully submitted online and is presently being given full consideration for publication in the Industrial Engineering & Management Systems.

Your manuscript ID is IEMS-2025-0128.

Please mention the above manuscript ID in all future correspondence or when calling the office for questions. If there are any changes in your street address or e-mail address, please log in to ScholarOne Manuscripts at <https://mc03.manuscriptcentral.com/kiie-iems> and edit your user information as appropriate.

You can also view the status of your manuscript at any time by checking your Author Center after logging in to <https://mc03.manuscriptcentral.com/kiie-iems>.

Thank you for submitting your manuscript to the Industrial Engineering & Management Systems.

Sincerely,
Industrial Engineering & Management Systems Editorial Office

IVAN KRISTIAN TO SINGGIH _ <ivanksinggi h@staff.ubaya.ac.id>

Tue, Nov 11, 2025 at 11:11 AM

To: Aditya Rio Prabowo <aditya@ft.uns.ac.id>, prita.meilanitasari@uisi.ac.id

[Quoted text hidden]



IEMS_crossover_v2.docx
408K



IVAN KRISTIAN TO SINGGIH _ <ivanksinggih@staff.ubaya.ac.id>

Industrial Engineering & Management Systems - Decision on Manuscript ID IEMS-2025-01281 message

Dong Ho Lee <onbehalfof@manuscriptcentral.com>
Reply-To: leman@hanyang.ac.kr
To: ivanksinggih@staff.ubaya.ac.id

Wed, Feb 11, 2026 at 6:56 AM

11-Feb-2026

Dear Dr. Singgih:

Manuscript ID IEMS-2025-0128 entitled "Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis" which you submitted to the Industrial Engineering & Management Systems, has been reviewed. The comments of the reviewer(s) are included at the bottom of this letter.

The reviewer(s) have recommended publication, but also suggest some revisions to your manuscript. Therefore, I invite you to respond to the reviewer(s)' comments and revise your manuscript. While revising the paper, please prepare a separate response letter that shows how you address each of the comments.

To revise your manuscript, log into <https://mc03.manuscriptcentral.com/kiie-iems> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You may also click the below link to start the revision process (or continue the process if you have already started your revision) for your manuscript. If you use the below link you will not be required to login to ScholarOne Manuscripts.

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc03.manuscriptcentral.com/kiie-iems?URL_MASK=64a6d9b925d54b50a6136515771370fa

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Please also highlight the changes to your manuscript within the document by using the track changes mode in MS Word or by using bold or colored text.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.

When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided. You can use this space to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Because we are trying to facilitate timely publication of manuscripts submitted to the Industrial Engineering & Management Systems, your revised manuscript should be submitted by 12-May-2026. If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.

Once again, thank you for submitting your manuscript to the Industrial Engineering & Management Systems and I look forward to receiving your revision.

Sincerely,
Dr. Dong Ho Lee
Managing Editor, Industrial Engineering & Management Systems
leman@hanyang.ac.kr

Associate Editor Comments to Author:

Area Editor

Comments to the Author:

Dear Dr. Ivan Singgih (Corresponding Author of IEMS-2025-0128),

Thank you for submitting your manuscript entitled
"Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using
Regression Machine Learning: Post-Experiment Analysis"
to Industrial Engineering & Management Systems.

I have now completed my evaluation of your manuscript based on the peer-review process. According to the reviewers' comments, the editorial decision is "Major Revision".

Please carefully address all comments and suggestions provided by the reviewer(s) in a revised version of your manuscript and resubmit it within the specified deadline. Along with the revised manuscript, please include a detailed response-to-reviewers letter clearly explaining how each comment has been addressed, with specific references to page and line numbers in the revised manuscript.

The reviewer's report is attached for your reference.

Thank you again for considering Industrial Engineering & Management Systems as an outlet for your work. We appreciate your interest in and support of the journal.

We look forward to receiving your revised manuscript.

Best regards,

Sang Jin Kweon, Ph.D.

Associate Editor, Industrial Engineering & Management Systems.

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

This study analyzes the impact of cross-over operators using regression techniques. While the idea is interesting and relevant, the paper's presentation and experimental setup are rather straightforward. Several improvements are strongly recommended, as outlined below:

- In the Introduction, please discuss the importance and urgency of predicting or analyzing the quality of cross-over operators in genetic algorithms.
- Please include a Related Works section that reviews recent cross-over techniques in genetic algorithms. And some works which analyzing the impact of cross-over techniques to the GA objectives.
- The authors should justify the selection of the following features for analyzing the impact of cross-over operators: parents_middle_average_distance, parents_remaining_customers_average_distance, parents_remaining_customers_min_distance, parents_remaining_customers_max_distance, parent_number_customers_middle, and parent_number_customers_remaining. In particular, is this feature set sufficiently general?
- The current experimental design is overly simplistic. Please include additional experiments using different cross-over operator techniques.
- The TSP settings should be more diverse, ranging from simple to complex problems. For example, experiments could vary from instances with a small number of cities to those with a large number of cities, as well as from simple objectives or constraints to more complex ones.
- The XAI analysis is also quite basic. The authors are encouraged to employ statistical correlation analyses to better support their claims.

Reviewer: 2

Comments to the Author

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using
Regression Machine Learning: Post-Experiment Analysis

The authors investigate the influence of genetic algorithm (GA) parameters on the performance of the Traveling Salesman Problem (TSP) and employ a regression model to estimate changes in objective values. While the study demonstrates

promising results, several aspects require further clarification and improvement.

1. The abstract should clearly mention the use of regression techniques and briefly summarize the key experimental findings.
2. The novelty of the proposed work should be explicitly stated, particularly in the introduction section.
3. The purpose of employing a regression model in this study needs to be explained in greater detail. The manuscript appears to focus on identifying influential GA parameters that affect performance. However, similar insights could potentially be obtained using feature importance analysis or correlation measures without relying on regression. The authors should therefore clarify the role and purpose of the regression model in this study.
4. The dataset used to generate the 272 records for the TSP experiments should be described more explicitly. It is unclear whether the TSP instances are derived from prior studies or randomly generated. Providing a screenshot of the dataset along with descriptive statistics would improve reader understanding.
5. Furthermore, the dataset size (272 records) is relatively small for a machine learning-based approach, which raises concerns regarding potential overfitting. The authors are encouraged to report model performance on both training and testing data to demonstrate whether overfitting occurs.
6. Please clarify the rationale behind the selection of features used in the regression model.
7. To better assess regression performance, additional evaluation metrics such as the correlation coefficient (R) or coefficient of determination (R^2) should be reported alongside error-based metrics. The validation strategy should also be clearly described, including whether a hold-out method or cross-validation was used.
8. Finally, the issue of generalization requires further discussion. Since the regression model is trained on a specific TSP instance, the authors should clarify how and to what extent the model can be expected to perform well on different problem instances or domains.



IVAN KRISTIAN TO SINGGIH _ <ivanksinggih@staff.ubaya.ac.id>

Industrial Engineering & Management Systems - Decision on Manuscript ID IEMS-2025-0128.R1

Dong Ho Lee <onbehalf@manuscriptcentral.com>

Sat, Apr 18, 2026 at 7:14 AM

Reply-To: leman@hanyang.ac.kr

To: ivanksinggih@staff.ubaya.ac.id

18-Apr-2026

Dear Dr. SinggiH:

Manuscript ID IEMS-2025-0128.R1 entitled "Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis" which you submitted to the Industrial Engineering & Management Systems, has been reviewed. The comments of the reviewer(s) are included at the bottom of this letter.

The reviewer(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the reviewer(s)' comments and revise your manuscript. While revising the paper, please prepare a separate response letter that shows how you address each of the comments.

To revise your manuscript, log into <https://mc03.manuscriptcentral.com/kiie-iems> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You may also click the below link to start the revision process (or continue the process if you have already started your revision) for your manuscript. If you use the below link you will not be required to login to ScholarOne Manuscripts.

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc03.manuscriptcentral.com/kiie-iems?URL_MASK=b8db4455104c47f89ea746fd1462bfde

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Please also highlight the changes to your manuscript within the document by using the track changes mode in MS Word or by using bold or colored text.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.

When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided. You can use this space to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Because we are trying to facilitate timely publication of manuscripts submitted to the Industrial Engineering & Management Systems, your revised manuscript should be submitted by 17-Jul-2026. If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.

Once again, thank you for submitting your manuscript to the Industrial Engineering & Management Systems and I look forward to receiving your revision.

Sincerely,
Dr. Dong Ho Lee
Managing Editor, Industrial Engineering & Management Systems
leman@hanyang.ac.kr

Associate Editor Comments to Author:

Area Editor

Comments to the Author:

Below, please find the review letter by the reviewer:

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis

The authors examine how genetic algorithm (GA) parameters affect the performance of the Traveling Salesman Problem (TSP) and apply regression models to estimate variations in objective values.

The authors did not include a response-to-reviewers document, which made it difficult to track how reviewer comments were addressed in the revised manuscript. Nevertheless, they made a clear effort to incorporate my previous suggestions. They also expanded the study by adding experiments using multiple regression methods across three datasets, which increases the overall complexity of the work.

Overall, the revised manuscript shows improved results and reflects revisions made in accordance with my suggestions. Only my additional revision as follows:

- Please verify the calculation of the R^2 (coefficient of determination), as it appears to be incorrect. Typically, R^2 values range between 0 and 1.

Reviewer(s)' Comments to Author:

Reviewer: 2

Comments to the Author

Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis

The authors examine how genetic algorithm (GA) parameters affect the performance of the Traveling Salesman Problem (TSP) and apply regression models to estimate variations in objective values.

The authors did not include a response-to-reviewers document, which made it difficult to track how reviewer comments were addressed in the revised manuscript. Nevertheless, they made a clear effort to incorporate my previous suggestions. They also expanded the study by adding experiments using multiple regression methods across three datasets, which increases the overall complexity of the work.

Overall, the revised manuscript shows improved results and reflects revisions made in accordance with my suggestions. Only my additional revision as follows:

- Please verify the calculation of the R^2 (coefficient of determination), as it appears to be incorrect. Typically, R^2 values range between 0 and 1.



IVAN KRISTIAN TO SINGGIH _ <ivanksinggih@staff.ubaya.ac.id>

Industrial Engineering & Management Systems - Decision on Manuscript ID IEMS-2025-0128.R2

2 messages

Dong Ho Lee <onbehalfof@manuscriptcentral.com>
Reply-To: leman@hanyang.ac.kr
To: ivanksinggih@staff.ubaya.ac.id

Mon, Apr 27, 2026 at 8:48 AM

27-Apr-2026

Dear Dr. Singgih:

It is a pleasure to accept your manuscript entitled "Evaluating Effectiveness of Crossover Operator in Genetic Algorithm for Travelling Salesman Problem Case using Regression Machine Learning: Post-Experiment Analysis" in its current form for publication in the Industrial Engineering & Management Systems. The comments of the reviewer(s) who reviewed your manuscript are included at the foot of this letter.

Please send me the final version in MS Word format through e-mail. In the final version, please include bio data for each author at the end of the manuscript.

Thank you for your fine contribution. On behalf of the Editors of the Industrial Engineering & Management Systems, we look forward to your continued contributions to the Journal.

Sincerely,
Dr. Dong Ho Lee
Managing Editor, Industrial Engineering & Management Systems
leman@hanyang.ac.kr

Associate Editor Comments to Author:

Area Editor
Comments to the Author:
(There are no comments.)

Reviewer(s)' Comments to Author:

Reviewer: 2

Comments to the Author
(There are no comments.)

IVAN KRISTIAN TO SINGGIH _ <ivanksinggih@staff.ubaya.ac.id>
To: leman@hanyang.ac.kr
Cc: prita.meilanitasari@uisi.ac.id, Aditya Rio Prabowo <aditya@ft.uns.ac.id>

Wed, Apr 29, 2026 at 11:06 PM

Dear Professor Dong Ho Lee,

Thank you for allowing us to contribute to your journal.
Please find the final Word file you requested attached to this email.

Best regards,
Ivan K. Singgih
(on behalf of the corresponding author)

Ivan Kristianto Singgih (Ph.D.)
URL: https://www.researchgate.net/profile/Ivan_Singgih

Head of Systems Engineering Laboratory and Assistant Professor

Study Program of Industrial Engineering
University of Surabaya
Jalan Raya Kalirungkut, Surabaya, 60293
Indonesia

[Quoted text hidden]



IEMS_crossover.docx
1188K