Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 7, No. 2, June 2019, pp. 165~181
ISSN: 2089-3272, DOI: 10.11591/ijeei.v7i2.1047

A modified genetic algorithm with a new crossover mating scheme

Allemar Jhone P. Delima[1], **Ariel M. Sison**[2], **Ruji P. Medina**[3]
[1,3]Graduate Programs, Technological Institute of the Philippines, Philippines
[2]Emilio Aguinaldo College, Philippines

Article Info

Article historys:

Received Feb 20, 2019
Revised Apr 11, 2019
Accepted Apr 24, 2019

Keywords:

Average crossover
Genetic algorithm
IBAX operator
Modified average crossover
Modified genetic algorithm
Novel crossover

ABSTRACT

This study introduced the Inversed Bi-segmented Average Crossover (IBAX), a novel crossover operator that enhanced the offspring generation of the genetic algorithm (GA) for variable minimization and numerical optimization problems. An attempt to come up with a new mating scheme in generating new offspring under the crossover function through the novel IBAX operator has paved the way to a more efficient and optimized solution for variable minimization particularly on premature convergence problem using GA. A total of 597 records of student-respondents in the evaluation of the faculty instructional performance, represented by 30 variables, from the four State Universities and Colleges (SUC) in Caraga Region, Philippines was used as the dataset. The simulation results showed that the proposed modification on the Average Crossover (AX) of the genetic algorithm outperformed the genetic algorithm with the original AX operator. The GA with IBAX operator combined with rank-based selection function had removed 20 or 66.66% of the variables while 13 or 43.33% of the variables were removed when GA with AX operator and roulette wheel selection function was used.

Corresponding Author:

Allemar Jhone P. Delima,
Graduate Programs,
Technological Institute of the Philippines,
938 Aurora Blvd., Cubao, Quezon City, Philippines.
Email: allemarjpd@ssct.edu.ph

1. INTRODUCTION

Data preprocessing [1-3] which is an imperative stride and considered to be one of the prime methods that is useful in data mining (DM), have led to the enhancement on the quality of data that positively contributes improvement to the precision and accuracy level as well as the mining efficiency of a prediction model [4, 5].

Data reduction, as an important data preprocessing technique in DM, is achieved through the selection and removal of unnecessary attributes and or variables in the dataset [6]. It is well known that in some cases, reducing original training set or variables by selecting the most representative information is advisable, yet obtaining nearly the same result or data-driven output [7-9]. Minimizing the size of the dataset aids in increasing the ability of generalization properties of the model. It also helped in lessening the space and computational time as well as minimizing the size of formulas used by the algorithm on the execution process [10]. Maximized accuracy through the reduced number of attributes [11, 6] and better understandability and interpretability of results are among the many benefits perceived in data reduction [12].

One of the competent data reduction, feature selection and global optimization algorithm that is widely used in related studies is the Genetic Algorithm [13-15]. Genetic Algorithms (GA), which was introduced by J.H. Holland in the 1970s, represents wide-ranging search method based on evolution and population genetics where its major executory mechanism relies on the crossover operator [16]. The unique integration of selection, crossover, and mutation operators serves as the driving force behind the successful implementation of GA. According to [17], the most widely-known problem in GAs is premature convergence. It occurs when genetic operators converge in an early stage after a few generations have been made and get stagnated there (local optima). Premature convergence occurs when the genetic operators cannot produce offspring that are a better representative of their parents whilst it is associated due to the loss of diversity in the population. According to [18], one of the technique to prevent premature convergence is to design an efficient crossover operator; thus, this study.

The activity that lies behind crossover is the creation of offspring that is achieved by combining information of the two parent chromosomes [19, 20]. For real encoding problems using the arithmetic function, the average crossover (AX) [21] is modified in this study. The simplicity of the average crossover has opened an avenue for improvements for better genetic algorithm performance. The modification will solve the weakness of the GA since a new method of pairing genes from the chromosomes will be observed, and those other researchers may use it for their experimental parameters setting.

There is an appeal in the literature that encourages the enhancement of crossover operators for more effective optimization schemes of evolutionary algorithms. The influence of crossover operators is vital to the whole genetic algorithm process in the quest for optimal search space [22, 23].

Therefore, this study proposed a novel crossover operator as an enhancement to the average crossover of the genetic algorithm. The novel crossover is called Inversed Bi-segmented Average Crossover (IBAX) that alters the offspring generation of parents that are instrumental for the next generation. The rest of the paper is arranged as follows. Section 2 discusses the literature review of Genetic Algorithm. Section 3 includes the design and methodology used in the study. Section 4 discusses the results and discussions while Section 5 highlights the conclusion and recommendation.

2. LITERATURE REVIEW

2.1. Genetic Algorithm

Genetic algorithms as defined by [24], is one of the many evolutionary algorithms based on the rules of biological evolution for global optimization solution.

GA is known as one of the most competent and widely held techniques that are used to search the best or ideal solution for problems with a huge search space especially in combinatorial problems where the search space is of factorial order. GA produce and controls some individuals through the integration of various suitable generic operators to look for optimal solutions. The bottleneck for an optimal genetic algorithm implementation relies on its three fundamental operations after creating the initial population viz., selection, crossover, and mutation functions. Figure 1 shows the flowchart of the genetic algorithm.

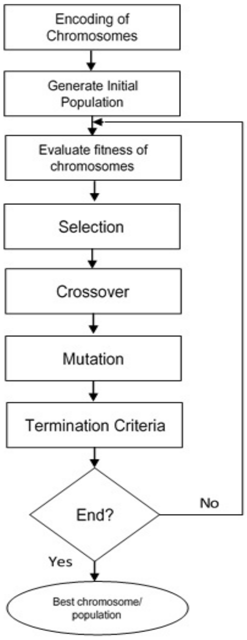

Figure 1. Genetic algorithm flowchart

2.1.1. Initialization / Evaluation of Fitness Function

Fitness function serves as the backbone of the evaluation process of fitted values; hence, a vital step in GA execution. This serves as a performance determinant for relevant judgment [25].

2.1.2. Selection

This stage of the genetic algorithm is where the members in the population are selected to enter into the mating pool for the next function which is the crossover stage. The selection of an optimal operator for this stage is vital to ensure that members of the population who have higher fitness values can have a bigger chance of being selected for mating. Although, members with lower fitness function do still have a slim chance of being selected for reproduction. It is important to select the best members of the population to ensure that the search process is global and does not simply meet the nearest local optimum [26]. Selection is one of the important aspects of the GA process, and there are several ways for the selection as to wit: Binary Tournament Selection, Stochastic Universal Sampling (SUS), Roulette Wheel Selection (RWS), Elitism Selection, and Rank-based Selection. For the detailed explanation of the abovementioned selection schemes, the study of [27] is recommended. Below are the following selection functions used in this study:

- Roulette Wheel Selection (RWS) Function

According to [28], Roulette selection is one of the simplest traditional GA selection technique. To execute, all the chromosomes in the population are placed on the roulette wheel according to their fitness value. A segment is assigned as representation to each individual commensurate to their fitness value; hence, the bigger the fitness value is, the larger the segment. Then, the virtual roulette wheel is spun. The individual corresponding to the segment on which roulette wheel stops are then selected. The process is repeated until the desired number of individuals is selected. Individuals with higher fitness have more probability of selection.

- Rank-based Selection Function

The rank-based selection function according to [29] can be assigned depending on the distribution of chromosomes according to their fitness values. This can be executed through positioning the chromosomes in decreasing order according to their fitness values. Next is to allocate a rank value on every chromosome

That corresponds to its arrangement in the set and then calculate the new fitness value for every chromosome using (1):

$$F = max - (max - \min) * \frac{rank-1}{Npop-1} \tag{1}$$

where $1<max<=2$ & $min = 2-max$

2.1.3. Crossover

The Crossover is identified to be the most important operator in genetic algorithms. It is responsible for generating new offspring that will be used for the next generation by combining features of two parent chromosomes [21].

A recent study was conducted by [22] that enhanced the Average Crossover (AX) operator of the genetic algorithm. The proposed operator is called Cross Average Crossover (CAX). The use of the modified genetic algorithm with CAX operator and rank-based selection function yielded to more decreased variables than the traditional genetic algorithm, but a degradation phenomenon [30] was depicted. The CAX operator with rank-based selection function eliminated those individuals with higher fitness values due to the structure of its mating scheme.

According to [19], there are two categories of crossover development. They are called parent-centric and mean centric operators. The parent-centric approach generates offspring within the vicinity of each of the parent chromosomes while the mean centric generates offspring solutions by identifying the central tendencies of the parents involved. The Average Crossover which is a well-known crossover operator for real encoding problems found in the study of [21], that is modified this study, is outlined below along with the CAX operator:

- Average Crossover (AX)

Part or all of the genes are averages of the same alleles in both the parents. Select two parallel parents and compute its average to create offspring.

- Cross Average Crossover (CAX)

A modified version of Average Crossover (AX) where the first gene in the first chromosome and the last gene of the second chromosome are averaged and get its value. The resulting average values are considered as offspring. Repeat the steps until genes from chromosomes have crossed in creating offspring.

2.1.4. Mutation

Studies have been carried out on the varieties of mutation techniques to improve GAs performance over the years. The purpose of mutation operation is to change the genes of the offspring and to increase the diversity of the population. This process enables GAs to jump out of local or suboptimal solutions to avoid premature convergence [19].

3. METHODOLOGY

In this study, the average crossover which is one of the crossover operators in the genetic algorithm is modified. The use of the roulette wheel and rank-based selection function were observed. The variables who obtained the lowest fitness function in each generation for ten generations were removed. Instead of pairing the parallel genes from chromosomes x and y and compute its average to produce offspring z as shown in Figure 2, it is suggested to segment the chromosomes (x and y) into two and inversely compute the average of genes within each segment created as depicted in Figure 3. The modified crossover will be called Inversed Bi-segmented Average Crossover (IBAX).

3.1. Existing Traditional Average Crossover

The average crossover is simple and can be implemented through the following steps:
Step 1: Take two parents from the selection pool.
Step 2: Create offspring Z from two parallel parents X and Y
Step 3: Use the formula

$$z = [x + y]/\,2 \qquad (2)$$

Step 4: For i = 1 to *n* do formula (2)
Step 5: End do

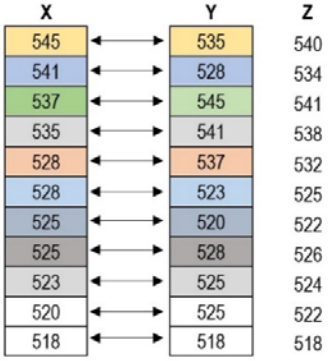

Figure 2. Average crossover with roulette wheel selection function

3.2. Modified Average Crossover

For the IBAX operator to be realized, the following steps must be executed:
Step 1: Take the parents from the selection pool.
Step 2: Count the number of genes found in the chromosomes. Identify if the dataset is in odd or even numbers.
Step 3: Segment the chromosomes (x and y) by dividing the total number of genes in the chromosomes into two and make sure that both first and second segments must contain an equal number of genes in an even count.
Step 4: On the first segment, create offspring Z for each gene by inversely pairing the first gene from chromosome X to the last gene on chromosome Y. Repeat until the last gene of the chromosome X and the first gene of the chromosome Y have inversely mated and have produced an offspring using formula (2).
Step 5: Execute the same process on the second segment until genes from all segments have produced offspring. In the case of odd datasets, the last genes of the chromosomes will not be combined in the second segment and will automatically be mated with each other to produce offspring.

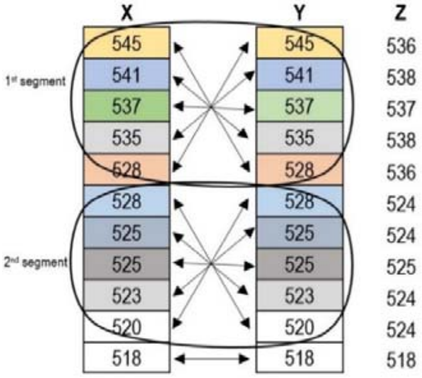

Figure 3. Inversed Bi-segmented Average Crossover with the rank-based selection function

3.3. Datasets

In this study, a total of 597 records of student-respondents in the evaluation of the faculty instructional performance from the four State Universities and Colleges (SUC) in Caraga Region, Philippines were used as the datasets. There were thirty (30) variables that represent the faculty instructional performance having divided into six (6) parts viz., methodology, classroom management, student discipline, assessment of learning, student-teacher relationship, and peer relationship. Each category has five items as shown in Table 1.

Table 1. Variables used in the study

Category	Reasons	Variable	Possible Value
Methodology	Utilizes varied designs/ techniques/ activities suited to the different types of learners.	M1	{1,2,3}
	Explains learning goals and instructional procedures to the students.	M2	{1,2,3}
	Uses real-life examples in the class to sustain student's interest in learning.	M3	{1,2,3}
	Creates a situation that encourages students to use critical thinking.	M4	{1,2,3}
	Delivers accurate/relevant/updated content knowledge.	M5	{1,2,3}
Classroom Management	Establishes routines to maximize instructional time.	C1	{1,2,3}
	Organizes and assign the daily cleaners.	C2	{1,2,3}
	Employs an effective system of classroom set-up.	C3	{1,2,3}
	Employs strategies to maximize the use of resources in learning activities.	C4	{1,2,3}
	Implements rules/policies inside the classroom.	C5	{1,2,3}
Student Discipline	Handles behavior problem concerning the student's rights.	SD1	{1,2,3}
	Imposes disciplinary sanction(s) to the misbehaving student(s).	SD2	{1,2,3}
	Encourages students to submit requirements on time.	SD3	{1,2,3}
	Motivates students to respect each other.	SD4	{1,2,3}
	Allows students to exercise their creativity.	SD5	{1,2,3}
Assessment of Learning	Constructs valid and reliable formative and summative tests.	A1	{1,2,3}
	Uses appropriate non-traditional assessment techniques and tools (i.e. portfolio, journals, rubric, etc)	A2	{1,2,3}
	Interprets and use test results to improve teaching and learning.	A3	{1,2,3}
	Uses tools for assessing authentic learning.	A4	{1,2,3}
	Provides timely and accurate feedback to students.	A5	{1,2,3}
Student-teacher relationship	Encourages students to participate in class/school activities actively.	ST1	{1,2,3}
	Allows students to communicate directly to him/her.	ST2	{1,2,3}
	Provides equal opportunities for all students.	ST3	{1,2,3}
	Promotes teamwork among students.	ST4	{1,2,3}
	Makes him/herself available to students.	ST5	{1,2,3}
Peer relationship	Demonstrates appropriate behavior in dealing with students/peers/superiors.	P1	{1,2,3}
	Manifests flexibility when deemed necessary.	P2	{1,2,3}
	Exhibits collegiality with colleagues.	P3	{1,2,3}
	Observes professionalism at all times.	P4	{1,2,3}
	Empathizes other needs and concern.	P5	{1,2,3}

4. RESULTS AND DISCUSSION

4.1. Simulation result for GA with AX operator and RWS Function

The simulation on the genetic algorithm was done for ten generations utilizing the existing traditional average crossover and roulette wheel selection function. The 597 records of random student-respondents in the evaluation of the faculty instructional performance (IP) from the four State Universities and Colleges (SUC) in Caraga Region, Philippines were instrumental in this study.

First Generation: Variable C2 is removed from the chromosome since it obtained the lowest fitness value of 171396 as evident in Table 2.

Second Generation: Variables M5 and A2 were removed from the chromosome since both obtained the lowest fitness value of 263169 as evident in Table 3.

Table 2. G1 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
M5	552	304704	24	474	C3	513	263169	
C1	490	240100	3	546	A3	518	268324	
C2	354	125316	1	474	C3	414	171396	Remove
C3	474	224676	2	556	ST1	515	265225	
C4	542	293764	18	490	C1	516	266256	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A2	500	250000	4	526	P3	513	263169	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
A5	516	266256	7	516	A5	516	266256	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Table 3. G2 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
M5	552	304704	24	474	C3	513	263169	Remove
C1	490	240100	3	546	A3	518	268324	
C3	474	224676	2	556	ST1	515	265225	
C4	542	293764	18	490	C1	516	266256	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A2	500	250000	4	526	P3	513	263169	Remove
A3	546	298116	20	513	A1	529.5	280370.3	

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
A4	518	268324	8	565	M2	541.5	293222.3	
A5	516	266256	7	516	A5	516	266256	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Third Generation: Variable C3 is removed from the chromosome since it obtained the lowest fitness value of 265225 as evident in Table 4.

Fourth Generation: Variables C4 and A5 were removed from the chromosome since it obtained the lowest fitness value of 266256 as evident in Table 5.

Table 4. G3 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C1	490	240100	3	546	A3	518	268324	
C3	474	224676	2	556	ST1	515	265225	Remove
C4	542	293764	18	490	C1	516	266256	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
A5	516	266256	7	516	A5	516	266256	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Table 5. G4 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
1	546	98116	2	52	5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C1	490	240100	3	546	A3	518	268324	
C4	542	293764	18	490	C1	516	266256	Remove
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A3	546	298116	20	513	A1	529.5	280370.3	

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
A4	518	268324	8	565	M2	541.5	293222.3	
A5	516	266256	7	516	A5	516	266256	Remove
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Fifth Generation: Variable C1 is removed from the chromosome since it obtained the lowest fitness value of 268324 as evident in Table 6.

Sixth Generation: Variable P3 is removed from the chromosome since it obtained the lowest fitness value of 272484 as evident in Table 7.

Table 6. G5 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C1	490	240100	3	546	A3	518	268324	Remove
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Table 7. G6 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P3	526	276676	10	518	A4	522	272484	Remove
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Seventh Generation: Variable A1 is removed from the chromosome since it obtained the lowest fitness value of 274052.3 as evident in Table 8.

Table 8. G7 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A1	513	263169	6	534	ST2	523.5	274052.3	Remove
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Eight Generation: Variable SD1 is removed from the chromosome since it obtained the lowest fitness value of 274576 as evident in Table 9.

Table 9. G8 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD1	548	300304	23	500	A2	524	274576	Remove
SD2	512	262144	5	542	C4	527	277729	
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Ninth Generation: Variable SD2 is removed from the chromosome since it obtained the lowest fitness value of 277729 as evident in Table 10.

Table 10. G9 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C5	528	278784	12	531	ST3	529.5	280370.3	
SD2	512	262144	5	542	C4	527	277729	Remove
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A3	546	298116	20	513	A1	529.5	280370.3	
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

Tenth Generation: Variables C5 and A3 were removed from the chromosome since it obtained the lowest fitness value of 280370.3 as evident in Table 11.

Table 11. G10 using an average crossover with RWS function

IP	X	Fitness	Rank	Pool		Off-spring	Fitness	Decision
				Y	IP			
M1	546	298116	22	552	M5	549	301401	
M2	565	319225	30	565	M2	565	319225	
M3	558	311364	27	548	SD1	553	305809	
M4	559	312481	28	546	SD5	552.5	305256.3	
C5	528	278784	12	531	ST3	529.5	280370.3	Remove
SD3	565	319225	29	546	A3	555.5	308580.3	
SD4	556	309136	26	558	M3	557	310249	
SD5	546	298116	21	528	C5	537	288369	
A3	546	298116	20	513	A1	529.5	280370.3	Remove
A4	518	268324	8	565	M2	541.5	293222.3	
ST1	556	309136	25	556	ST1	556	309136	
ST2	534	285156	16	559	M4	546.5	298662.3	
ST3	531	281961	14	546	M1	538.5	289982.3	
ST4	541	292681	17	556	SD4	548.5	300852.3	
ST5	527	277729	11	552	M5	539.5	291060.3	
P1	531	281961	13	565	SD3	548	300304	
P2	533	284089	15	541	ST4	537	288369	
P4	526	276676	9	565	SD3	545.5	297570.3	
P5	544	295936	19	565	SD3	554.5	307470.3	

4.2. Simulation result for GA with IBAX operator and rank-based selection function

The simulation on the genetic algorithm was done utilizing the novel Inversed Bi-segmented Average Crossover (IBAX) operator and rank-based selection function on the same datasets and number of generations.

First Generation: Variable C2 was removed from the list of variables after applying the rank-based selection. The variable C2 obtained the lowest fitness value in the rank-based selection. Hence, it does not have any chance to be selected. Moreover, after applying the inversed bi-segmented average crossover (IBAX) operator and obtained the fitness value of the offspring, variable C3 was removed from the chromosomes since it obtained the lowest fitness value of 224676 that will not warrant for the next generation. Thus, in the first generation, there were two variables removed from the list as shown in Table 12.

Table 12. G1 using IBAX with the rank-based selection function

		Rank-based			IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	541	553	305809
SD3	565	319225	29	3504670.8	565	542	553.5	306362.3
M4	559	312481	28	3379354.8	559	544	551.5	304152.3
M3	558	311364	27	3254038.8	558	546	552	304704
SD4	556	309136	26	3128722.8	556	546	551	303601
ST1	556	309136	25	3003406.8	556	546	551	303601
M5	552	304704	24	2878090.8	552	548	550	302500
SD1	548	300304	23	2752774.8	548	552	550	302500
M1	546	298116	22	2627458.8	546	556	551	303601
SD5	546	298116	21	2502142.8	546	556	551	303601
A3	546	298116	20	2376826.8	546	558	552	304704
P5	544	295936	19	2251510.8	544	559	551.5	304152.3
C4	542	293764	18	2126194.8	542	565	553.5	306362.3
ST4	541	292681	17	2000878.8	541	565	553	305809
ST2	534	285156	16	1875562.8	534	490	512	262144
P2	533	284089	15	1750246.8	533	500	516.5	266772.3
ST3	531	281961	14	1624930.8	531	512	521.5	271962.3
P1	531	281961	13	1499614.8	531	513	522	272484
C5	528	278784	12	1374298.8	528	516	522	272484
ST5	527	277729	11	1248982.8	527	518	522.5	273006.3
P3	526	276676	10	1123666.8	526	526	526	276676
P4	526	276676	9	998350.8	526	526	526	276676
A4	518	268324	8	873034.8	518	527	522.5	273006.3
A5	516	266256	7	747718.8	516	528	522	272484
A1	513	263169	6	622402.8	513	531	522	272484
SD2	512	262144	5	497086.8	512	531	521.5	271962.3
A2	500	250000	4	371770.8	500	533	516.5	266772.3
C1	490	240100	3	246454.8	490	534	512	262144
C3	474	224676	2	121138.8	474	474	474	224676
C2	354	125316	1	-4177.2				

Second Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables ST2 and C1 were removed from the chromosomes since they obtained the lowest fitness value of 262144. In the second generation, there were two variables removed from the list as shown in Table 13.

Table 13. G2 using IBAX with the rank-based selection function

		Rank-based			IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	541	553	305809
SD3	565	319225	29	3504670.8	565	542	553.5	306362.3
M4	559	312481	28	3379354.8	559	544	551.5	304152.3
M3	558	311364	27	3254038.8	558	546	552	304704
SD4	556	309136	26	3128722.8	556	546	551	303601
ST1	556	309136	25	3003406.8	556	546	551	303601
M5	552	304704	24	2878090.8	552	548	550	302500
SD1	548	300304	23	2752774.8	548	552	550	302500
M1	546	298116	22	2627458.8	546	556	551	303601
SD5	546	298116	21	2502142.8	546	556	551	303601
A3	546	298116	20	2376826.8	546	558	552	304704
P5	544	295936	19	2251510.8	544	559	551.5	304152.3
C4	542	293764	18	2126194.8	542	565	553.5	306362.3
ST4	541	292681	17	2000878.8	541	565	553	305809
ST2	534	285156	16	1875562.8	534	490	512	262144
P2	533	284089	15	1750246.8	533	500	516.5	266772.3
ST3	531	281961	14	1624930.8	531	512	521.5	271962.3
P1	531	281961	13	1499614.8	531	513	522	272484
C5	528	278784	12	1374298.8	528	516	522	272484
ST5	527	277729	11	1248982.8	527	518	522.5	273006.3
P3	526	276676	10	1123666.8	526	526	526	276676
P4	526	276676	9	998350.8	526	526	526	276676
A4	518	268324	8	873034.8	518	527	522.5	273006.3
A5	516	266256	7	747718.8	516	528	522	272484
A1	513	263169	6	622402.8	513	531	522	272484
SD2	512	262144	5	497086.8	512	531	521.5	271962.3
A2	500	250000	4	371770.8	500	533	516.5	266772.3
C1	490	240100	3	246454.8	490	534	512	262144

Third Generation: After applying the inversed bi-segmented average crossover (IBAX) operator and obtained the fitness value of the offspring, variables ST4 and A2 were removed from the chromosomes since both obtained the lowest fitness value of 270920.3. In the third generation, there were two variables removed from the list as shown in Table 14.

Table 14. G3 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	542	553.5	306362.3
SD3	565	319225	29	3504670.8	565	544	554.5	307470.3
M4	559	312481	28	3379354.8	559	546	552.5	305256.3
M3	558	311364	27	3254038.8	558	546	552	304704
SD4	556	309136	26	3128722.8	556	546	551	303601
ST1	556	309136	25	3003406.8	556	548	552	304704
M5	552	304704	24	2878090.8	552	552	552	304704
SD1	548	300304	23	2752774.8	548	556	552	304704
M1	546	298116	22	2627458.8	546	556	551	303601
SD5	546	298116	21	2502142.8	546	558	552	304704
A3	546	298116	20	2376826.8	546	559	552.5	305256.3
P5	544	295936	19	2251510.8	544	565	554.5	307470.3
C4	542	293764	18	2126194.8	542	565	553.5	306362.3
ST4	541	292681	17	2000878.8	541	500	520.5	270920.3
P2	533	284089	15	1750246.8	533	512	522.5	273006.3
ST3	531	281961	14	1624930.8	531	513	522	272484
P1	531	281961	13	1499614.8	531	516	523.5	274052.3
C5	528	278784	12	1374298.8	528	518	523	273529
ST5	527	277729	11	1248982.8	527	526	526.5	277202.3
P3	526	276676	10	1123666.8	526	526	526	276676
P4	526	276676	9	998350.8	526	527	526.5	277202.3
A4	518	268324	8	873034.8	518	528	523	273529
A5	516	266256	7	747718.8	516	531	523.5	274052.3
A1	513	263169	6	622402.8	513	531	522	272484
SD2	512	262144	5	497086.8	512	533	522.5	273006.3
A2	500	250000	4	371770.8	500	541	520.5	270920.3

Fourth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables P2 and A1 were removed from the chromosomes since both obtained the lowest fitness value of 273529. In the fourth generation, there were two variables removed from the list as shown in Table 15.

Table 15. G4 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	544	554.5	307470.3
SD3	565	319225	29	3504670.8	565	546	555.5	308580.3
M4	559	312481	28	3379354.8	559	546	552.5	305256.3
M3	558	311364	27	3254038.8	558	546	552	304704
SD4	556	309136	26	3128722.8	556	548	552	304704
ST1	556	309136	25	3003406.8	556	552	554	306916
M5	552	304704	24	2878090.8	552	556	554	306916
SD1	548	300304	23	2752774.8	548	556	552	304704
M1	546	298116	22	2627458.8	546	558	552	304704
SD5	546	298116	21	2502142.8	546	559	552.5	305256.3
A3	546	298116	20	2376826.8	546	565	555.5	308580.3
P5	544	295936	19	2251510.8	544	565	554.5	307470.3
C4	542	293764	18	2126194.8	542	512	527	277729
P2	533	284089	15	1750246.8	533	513	523	273529
ST3	531	281961	14	1624930.8	531	516	523.5	274052.3
P1	531	281961	13	1499614.8	531	518	524.5	275100.3
C5	528	278784	12	1374298.8	528	526	527	277729
ST5	527	277729	11	1248982.8	527	526	526.5	277202.3
P3	526	276676	10	1123666.8	526	527	526.5	277202.3
P4	526	276676	9	998350.8	526	528	527	277729
A4	518	268324	8	873034.8	518	531	524.5	275100.3
A5	516	266256	7	747718.8	516	531	523.5	274052.3
A1	513	263169	6	622402.8	513	533	523	273529
SD2	512	262144	5	497086.8	512	542	527	277729

Fifth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables ST3 and A4 were removed from the chromosomes since both obtained the lowest fitness value of 275100.3. In the fifth generation, there were two variables removed from the list as shown in Table 16.

Table 16. G5 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	546	555.5	308580.3
SD3	565	319225	29	3504670.8	565	546	555.5	308580.3
M4	559	312481	28	3379354.8	559	546	552.5	305256.3
M3	558	311364	27	3254038.8	558	548	553	305809
SD4	556	309136	26	3128722.8	556	552	554	306916
ST1	556	309136	25	3003406.8	556	556	556	309136
M5	552	304704	24	2878090.8	552	556	554	306916
SD1	548	300304	23	2752774.8	548	558	553	305809
M1	546	298116	22	2627458.8	546	559	552.5	305256.3
SD5	546	298116	21	2502142.8	546	565	555.5	308580.3
A3	546	298116	20	2376826.8	546	565	555.5	308580.3
P5	544	295936	19	2251510.8	544	512	528	278784
C4	542	293764	18	2126194.8	542	516	529	279841
ST3	531	281961	14	1624930.8	531	518	524.5	275100.3
P1	531	281961	13	1499614.8	531	526	528.5	279312.3
C5	528	278784	12	1374298.8	528	526	527	277729
ST5	527	277729	11	1248982.8	527	527	527	277729
P3	526	276676	10	1123666.8	526	528	527	277729
P4	526	276676	9	998350.8	526	531	528.5	279312.3
A4	518	268324	8	873034.8	518	531	524.5	275100.3
A5	516	266256	7	747718.8	516	542	529	279841
SD2	512	262144	5	497086.8	512	544	528	278784

Sixth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables C5 and ST5 were removed from the chromosomes since both obtained the lowest fitness value of 278256.3. In the sixth generation, there were two variables removed from the list as shown in Table 17.

Table 17. G6 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	546	555.5	308580.3
SD3	565	319225	29	3504670.8	565	546	555.5	308580.3
M4	559	312481	28	3379354.8	559	548	553.5	306362.3
M3	558	311364	27	3254038.8	558	552	555	308025
SD4	556	309136	26	3128722.8	556	556	556	309136
ST1	556	309136	25	3003406.8	556	556	556	309136
M5	552	304704	24	2878090.8	552	558	555	308025
SD1	548	300304	23	2752774.8	548	559	553.5	306362.3
M1	546	298116	22	2627458.8	546	565	555.5	308580.3
SD5	546	298116	21	2502142.8	546	565	555.5	308580.3
A3	546	298116	20	2376826.8	546	512	529	279841
P5	544	295936	19	2251510.8	544	516	530	280900
C4	542	293764	18	2126194.8	542	526	534	285156
P1	531	281961	13	1499614.8	531	526	528.5	279312.3
C5	528	278784	12	1374298.8	528	527	527.5	278256.3
ST5	527	277729	11	1248982.8	527	528	527.5	278256.3
P3	526	276676	10	1123666.8	526	531	528.5	279312.3
P4	526	276676	9	998350.8	526	542	534	285156
A5	516	266256	7	747718.8	516	544	530	280900
SD2	512	262144	5	497086.8	512	546	529	279841

Seventh Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables SD5 and SD2 were removed from the chromosomes since both obtained the lowest fitness value of 279841. In the seventh generation, there were two variables removed from the list as shown in Table 18.

Table 18. G7 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	546	555.5	308580.3
SD3	565	319225	29	3504670.8	565	548	556.5	309692.3
M4	559	312481	28	3379354.8	559	552	555.5	308580.3
M3	558	311364	27	3254038.8	558	556	557	310249
SD4	556	309136	26	3128722.8	556	556	556	309136
ST1	556	309136	25	3003406.8	556	558	557	310249
M5	552	304704	24	2878090.8	552	559	555.5	308580.3
SD1	548	300304	23	2752774.8	548	565	556.5	309692.3
M1	546	298116	22	2627458.8	546	565	555.5	308580.3
SD5	546	298116	21	2502142.8	546	512	529	279841
A3	546	298116	20	2376826.8	546	516	531	281961
P5	544	295936	19	2251510.8	544	526	535	286225
C4	542	293764	18	2126194.8	542	526	534	285156
P1	531	281961	13	1499614.8	531	531	531	281961
P3	526	276676	10	1123666.8	526	542	534	285156
P4	526	276676	9	998350.8	526	544	535	286225
A5	516	266256	7	747718.8	516	546	531	281961
SD2	512	262144	5	497086.8	512	546	529	279841

Eighth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables M1 and A5 were removed from the chromosomes since both obtained the lowest fitness value of 281961. In the eight generations, there were two variables removed from the list as shown in Table 19.

Ninth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables A3 and P3 were removed from the chromosomes since both obtained the lowest fitness value of 287296. In the ninth generation, there were two variables removed from the list as shown in Table 20.

Table 19. G8 using IBAX with rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	548	556.5	309692.3
SD3	565	319225	29	3504670.8	565	552	558.5	311922.3
M4	559	312481	28	3379354.8	559	556	557.5	310806.3
M3	558	311364	27	3254038.8	558	556	557	310249
SD4	556	309136	26	3128722.8	556	558	557	310249
ST1	556	309136	25	3003406.8	556	559	557.5	310806.3
M5	552	304704	24	2878090.8	552	565	558.5	311922.3
SD1	548	300304	23	2752774.8	548	565	556.5	309692.3
M1	546	298116	22	2627458.8	546	516	531	281961
A3	546	298116	20	2376826.8	546	526	536	287296
P5	544	295936	19	2251510.8	544	526	535	286225
C4	542	293764	18	2126194.8	542	531	536.5	287832.3
P1	531	281961	13	1499614.8	531	542	536.5	287832.3
P3	526	276676	10	1123666.8	526	544	535	286225
P4	526	276676	9	998350.8	526	546	536	287296
A5	516	266256	7	747718.8	516	546	531	281961

Table 20. G9 using IBAX with the rank-based selection function

Rank-based					IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	552	558.5	311922.3
SD3	565	319225	29	3504670.8	565	556	560.5	314160.3
M4	559	312481	28	3379354.8	559	556	557.5	310806.3
M3	558	311364	27	3254038.8	558	558	558	311364
SD4	556	309136	26	3128722.8	556	559	557.5	310806.3
ST1	556	309136	25	3003406.8	556	565	560.5	314160.3
M5	552	304704	24	2878090.8	552	565	558.5	311922.3
SD1	548	300304	23	2752774.8	548	526	537	288369
A3	546	298116	20	2376826.8	546	526	536	287296
P5	544	295936	19	2251510.8	544	531	537.5	288906.3
C4	542	293764	18	2126194.8	542	542	542	293764
P1	531	281961	13	1499614.8	531	544	537.5	288906.3
P3	526	276676	10	1123666.8	526	546	536	287296
P4	526	276676	9	998350.8	526	548	537	288369

Tenth Generation: After applying the inversed bi-segmented average crossover (IBAX) operator, variables M5 and P4 were removed from the chromosomes since both obtained the lowest fitness value of 290521. In the tenth generation, there were two variables removed from the list as shown in Table 21.

Table 21. G10 using IBAX with the rank-based selection function

		Rank-based			IBAX			
IP	X	Fitness	Rank	New Fitness	Parent 1	Parent 2	Offspring	Fitness
M2	565	319225	30	3629986.8	565	556	560.5	314160.3
SD3	565	319225	29	3504670.8	565	556	560.5	314160.3
M4	559	312481	28	3379354.8	559	558	558.5	311922.3
M3	558	311364	27	3254038.8	558	559	558.5	311922.3
SD4	556	309136	26	3128722.8	556	565	560.5	314160.3
ST1	556	309136	25	3003406.8	556	565	560.5	314160.3
M5	552	304704	24	2878090.8	552	526	539	290521
SD1	548	300304	23	2752774.8	548	531	539.5	291060.3
P5	544	295936	19	2251510.8	544	542	543	294849
C4	542	293764	18	2126194.8	542	544	543	294849
P1	531	281961	13	1499614.8	531	548	539.5	291060.3
P4	526	276676	9	998350.8	526	552	539	290521

4.3. Evaluation of the efficacy and reduction rate using GA with AX and IBAX operators

The variable minimization process using the genetic algorithm with average crossover operator and roulette wheel selection function has depicted a decrease after the ten generations. From the 30 variables, the numbers were minimized to 17 variables. 43% of variables were removed as depicted in Table 22.

Meanwhile, the variable minimization process using the genetic algorithm with the proposed novel mating scheme called inversed bi-segmented average crossover operator, and rank-based selection function has depicted a noticeable decrease after the ten generations. From the 30 variables, the numbers were minimized to 10 variables after the generations. A total of 66.66% of variables were removed as depicted in Table 23. Since the amount of reduction varies according to the genetic algorithms used, removing 66.66% of the variables in the dataset is good enough as the notion of dropping one or more variables should help reduce dimensionality is certain. The ratio of feature reduction of more than 60% is acceptable just like in the work of [31].

Table 22. Simulation result for GA with AX operator and RWS function

Number of Generations	Number of Variables Left	Number of Variables Removed	Variables Removed	Percentage
1	30	1	C2	3.33%
2	29	2	M5, A2	6.66%
3	27	1	C3	3.33%
4	26	2	C4, A5	6.66%
5	24	1	C1	3.33%
6	23	1	P3	3.33%
7	22	1	A1	3.33%
8	21	1	SD1	3.33%
9	20	1	SD2	3.33%
10	19	2	C5, A3	6.66%
10	17	-	-	-
Total Percentage of Variables Removed				43.33%

Table 23. Simulation result for GA with IBAX operator and rank-based selection function

Number of Generations	Number of Variables Left	Number of Variables Removed	Variables Removed	Percentage
1	30	2	C3, C2	6.66%
2	28	2	ST2, C1	6.66%
3	26	2	ST4, A2	6.66%
4	24	2	P2, A1	6.66%
5	22	2	ST3, A4	6.66%
6	20	2	C5, ST5	6.66%
7	18	2	SD5, SD2	6.66%
8	16	2	M1, A5	6.66%
9	14	2	A3, P3	6.66%
10	12	2	M5, P4	6.66%
10	10	-	-	-
Total Percentage of Variables Removed				66.66%

To have a further evaluation on the efficacy and reduction rate of the proposed novel crossover, the GA with the IBAX operator was compared along with the other real encoding-based crossover mechanism of the GA such as geometrical crossover [32] and cross average crossover (CAX) [22] aside from the AX operator. The simulation result showed that the genetic algorithm with a new crossover mating scheme outperformed the other existing real encoding-based crossover operators of genetic algorithm in reducing variables as depicted in Table 24.

Table 24. Comparative result for variable minimization using genetic algorithms

Genetic Algorithms	Number of variables	Number of variables left	Number of variables removed	Percentage of variables removed
GA with AX operator and RWS function	30	17	13	43.33%
GA with Geometrical crossover operator with RWS function	30	20	10	33.33%
GA with CAX operator with RWS function	30	13	17	56.66%
GA with IBAX operator and Rank-based Selection function	30	10	20	66.66%

5. CONCLUSION AND RECOMMENDATION

Through the study, a novel approach for the optimization process using another crossover operator of genetic algorithm was introduced and added to the body of knowledge. The proposed modification on the genetic algorithm with inversed bi-segmented average crossover (IBAX) has paved the way to the enhancement of GA's average crossover mating scheme that affects GA's optimization performance in general. It is evident that the IBAX operator performed the minimization process way better than the average crossover since there were 10 and 17 variables left, respectively, after the tenth generations.

For future works, it is suggested to use the novel IBAX operator in the different type of datasets and incorporate the modified genetic algorithm in various data mining techniques and approaches that need variable minimization or feature reduction process such as in prediction.

REFERENCES

[1] E. S. K. Tiu, Y. F. Huang, and L. Ling, "Improving the performance of streamflow forecasting model using data-preprocessing technique in Dungun River Basin," *E3S Web Conf.*, vol. 34, no. 02014, 2018.

[2] M. Munk and L. Benko, "Using Entropy in Web Usage Data Preprocessing," *Entropy*, vol. 20, no. 67, pp. 1–15, 2018.

[3] A. Roy, R. M. O. Cruz, R. Sabourin, and G. D. C. Cavalcanti, "A study on combining dynamic selection and data preprocessing for imbalance learning," *Neurocomputing*, vol. 286, pp. 179–192, 2018.

[4] P. Vrushali, M. Trupti, G. Pratiksha, and G. Arti, "Crime Rate Prediction using KNN," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 6, no. 1, pp. 124–127, 2018.

[5] S. García, J. Luengo, and F. Herrera, "Tutorial on practical tips of the most influential data preprocessing algorithms in data mining," *Knowledge-Based Syst.*, vol. 98, pp. 1–29, 2016.

[6] A. Baldominos, P. Isasi, and U. C. I. I. I. De Madrid, "Feature Set Optimization for Physical Activity Recognition Using Genetic Algorithms," *Proc. Companion Publ. 2015 Genet. Evol. Comput. Conf. - GECCO Companion '15*, pp. 1311–1318, 2015.

[7] C.-F. J. Kuo, C.-H. Lin, and M.-H. Lee, "Analyze the energy consumption characteristics and affecting factors of Taiwan's convenience stores-using the big data mining approach," *Energy Build.*, vol. 168, pp. 120–136, 2018.

[8] I. M. El-hasnony, H. M. El Bakry, and A. A. Saleh, "Comparative Study among Data Reduction Techniques over Classification Accuracy," *Int. J. Comput. Appl.*, vol. 122, no. 2, pp. 9–15, 2015.

[9] F. Herrera and S. Garc, "Prototype Selection for Nearest Neighbor Classification Taxonomy and Empirical Study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 417–435, 2012.

[10] J. R. Cano, F. Herrera, and M. Lozano, "On the combination of evolutionary algorithms and stratified strategies for training set selection in data mining," *Appl. Soft Comput. J.*, vol. 6, no. 3, pp. 323–332, 2006.

[11] Y. Cheng, K. Chen, H. Sun, Y. Zhang, and F. Tao, "Data and knowledge mining with big data towards smart production," *J. Ind. Inf. Integr.*, vol. 9, pp. 1–13, 2018.

[12] S. Ramírez-Gallego, B. Krawczyk, S. García, M. Woźniak, and F. Herrera, "A survey on data preprocessing for data stream mining: Current status and future directions," *Neurocomputing*, vol. 239, pp. 39–57, 2017.

[13] M. Mafarja, I. Aljarah, A. A. Heidari, A. I. Hammouri, H. Faris, and A. M. Al-zoubi, "Evolutionary Population Dynamics and Grasshopper Optimization Approaches for Feature Selection Problems," *Knowledge-Based Syst.*, vol. 154, no. 7, pp. 25–45, 2017.

[14] S. R. A. Archana and M. S. Thanabal, “Optimization Algorithms for Feature Selection in Classification : A Survey,” *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 4, no. 2, pp. 1124–1127, 2016.

[15] C. Tsai, W. Eberle, and C. Chu, “Knowledge-Based Systems Genetic algorithms in feature and instance selection,” *Knowledge-Based Syst.*, vol. 39, pp. 240–24, 2013.

[16] P. Horton, M. Jaboyedoff, and C. Obled, “Using genetic algorithms to optimize the analogue method for precipitation prediction in the Swiss Alps,” *J. Hydrol.*, vol. 556, pp. 1220–1231, 2018.

[17] S. F. Galan, O. J. Mengshoel, and R. Pinter, “A novel mating approach for genetic algorithms,” *Evol. Comput.*, vol. 21, no. 2, pp. 197–229, 2013.

[18] V. K. Dabhi and S. Chaudhary, “Empirical modeling using genetic programming: a survey of issues and approaches,” *Nat. Comput.*, vol. 14, no. 2, pp. 303–330, 2015.

[19] S. M. Lim, A. B. Sultan, N. Sulaiman, A. Mustapha, and K. Y. Leong, “Crossover and Mutation Operators of Genetic Algorithms,” *Int. J. Mach. Learn. Comput.*, vol. 7, no. 1, pp. 10–13, 2017.

[20] S. Picek, “Comparison of a Crossover Operator in Binary-coded Genetic Algorithms,” *WSEAS Trans. Comput.*, vol. 9, no. 9, pp. 1064–1073, 2010.

[21] G. Pavai and T. V Geetha, “A Survey on Crossover Operators,” *ACM Comput. Surv.*, vol. 49, no. 4, pp. 1–43, 2016.

[22] M. Y. Orong, A. M. Sison, and R. P. Medina, “A new crossover mechanism for genetic algorithm with rank-based selection method,” in *Proceedings of 2018 5th International Conference on Business and Industrial Research: Smart Technology for Next Generation of Information, Engineering, Business and Social Science, ICBIR*, pp. 83–88, 2018.

[23] P. Błażej, M. Wnętrzak, and P. Mackiewicz, “The role of crossover operator in evolutionary-based approach to the problem of genetic code optimization,” *BioSystems*, vol. 150, pp. 61–72, 2016.

[24] A. B. Hassanat, V. B. S. Prasath, M. A. Abbadi, S. A. Abu-Qdari, and H. Faris, “An improved Genetic Algorithm with a new initialization mechanism based on Regression techniques,” *Inf.*, vol. 9, no. 7, 2018.

[25] S. Renjith and C. Anjali, “Fitness Function in Genetic Algorithm based Information Filtering - A Survey,” in *International Conference on Mobility in Computing- ICMiC13*, 2013, pp. 80–86.

[26] O. Al Jadaan, L. Rajamani, and C. R. Rao, “Improved selection operator for GA,” *J. Theor. Appl. Inf. Technol.*, vol. 4, no. 4, pp. 269–277, 2008.

[27] F. Alabsi and R. Naoum, “Comparison of Selection Methods and Crossover Operations using Steady State Genetic Based Intrusion Detection System,” *J. Emerg. Trends Comput. Inf. Sci.*, vol. 3, no. 7, pp. 1053–1058, 2012.

[28] S. L. Yadav and A. Sohal, “Comparative Study of Different Selection Techniques in Genetic Algorithm,” *Int. J. Sci. Math.*, vol. 6, no. 3, 2017.

[29] F. Alabsi and R. Naoum, “Comparison of Selection Methods and Crossover Operations using Steady State Genetic Based Intrusion Detection System,” *J. Emerg. Trends Comput. Inf. Sci.*, vol. 3, no. 7, pp. 1053–1058, 2012.

[30] Y. Fan and D. Chen, “Application of Improved Adaptive Genetic Algorithm in Train Energy Saving,” in *Proceedings of the Fourth International Forum on Decision Sciences, Uncertainty and Operations Research*, 2017, pp. 723–736.

[31] H. Rao *et al.*, “Feature selection based on artificial bee colony and gradient boosting decision tree,” *Appl. Soft Comput. J.*, vol. 74, pp. 634–642, 2018.

[32] Z. Michalewicz, G. Nazhiyath, and M. Michalewicz, “A note on the usefulness of geometric crossover for numerical optimization problems,” *Proc. 5th Annu. Conf. Evol. Program.*, pp. 305–312, 1996.

BIOGRAPHY OF AUTHORS

Allemar Jhone P. Delima was born in Surigao City, Philippines. He is a faculty under the College of Engineering and Information Technology of Surigao State College of Technology. He received his B.S. degree from Surigao State College of Technology (SSCT), Surigao City, in 2014 and the master's degree from the same institution in 2016, both in Information Technology. He is currently pursuing his Doctor in Information Technology (DIT) degree in the Technological Institute of the Philippines – Quezon City. His research interest is in Data Mining, Data Analytics, and Machine Learning.

Ariel M. Sison earned his Doctor in Information Technology (DIT) degree at the Technological Institute of the Philippines Quezon City in 2013 and graduated with highest honors. He took up his master’s degree in computer science at De La Salle University, Manila in 2006 and obtained his BS in Computer Science at Emilio Aguinaldo College, Manila in 1994. He is currently connected at the School of Computer Studies, Emilio Aguinaldo College in Manila, Philippines. His research interests include data mining and data security.

Ruji P. Medina is the dean of the graduate programs of the Technological Institute of the Philippines in Quezon City. He holds a Ph.D. in environmental engineering from the University of the Philippines with a sandwich program at the University of Houston, Texas where he worked on the synthesis of nanocomposite materials.