A Proposed Genetic Algorithm Adaptation Based Model for Students' Tracks Prediction

Mourad Kaddes
University of Jeddah, College of Computing and Information Technology at Khulais, Department of Information Systems, Jeddah, Saudi Arabia
mkaddes@uj.edu.sa

Ayman E. Khedr
University of Jeddah, College of Computing and Information Technology at Khulais, Department of Information Systems, Jeddah, Saudi Arabia
aeelsayed@uj.edu.sa

Amira M. Idrees
Faculty of Computers and Information Technology, Future University in Egypt, Cairo, Egypt
amira.mohamed@fue.edu.eg

Abstract: *Evolutionary algorithms such as genetic algorithms have proved their effectiveness and reliability in optimization solutions. The genetic algorithm is one of the most powerful algorithms in optimizing solutions to various problems. However, such algorithms suffer from performance issues resulting from bottlenecks in their mechanisms. This research proposes an effective solution for raising the performance of a genetic algorithm with the idea of merging its mechanism with one of the swarm intelligence techniques. The proposed solution presents an effective model for the initialization task as well as minimizing the iterations while ensuring the optimized solution. The mimic concept for natural processes has leveraged the genetic algorithm computation to the optimized level. Linking genetic algorithms and particle swarm intelligence algorithm has proved their effectiveness through a set of experiments. Moreover, the proposed adapted algorithm has been applied to two experiments to prove the effectiveness compared with literature and in the education field in generating the most effective track for students targeting to enhance the student's performance which is considered one of the strategic targets in all economies.*

Keywords: *Genetic algorithms, optimization, particle swarm optimization, education, student tracks.*

Received August 01, 2024; accepted November 05, 2024
https://doi.org/10.34028/iajit/22/2/13

1. Introduction

Aligning sequences approaches in the bioinformatics field are commonly applied in identifying similarities and exploring relationships between sequences. Matching and exploring unknown sequences from known ones are employed [16]. Different types arise including pairs of sequences or multiple relationships. While the two types are employed, however, the most generic type is the multiple sequences alignment according to its ability to explore different families, simultaneously. One of the well-known methods for building pairs of trees is a genetic algorithm following the genetics nature [13, 20]. Genetic algorithm is utilized in different research in solving many problems with different natures which highlights the algorithm's effectiveness. However, the algorithm performance is still an open area of research. One of the algorithm challenges is the computational complexity which directly affects the time consumed to reach a solution [1]. The complexity of the genetic algorithm is determined to be equal to O(PC3) where P refers to the population and C refers to the selected items to contribute. Reducing the algorithm complexity can be efficiently performed by utilizing the concept of parallelism with an efficient selection method as will be proposed in the current research. Manipulating complexity can also be performed by adapting the selection process [12]. Applying the tournament concept to the selection process provides less computation complexity as will be proved in the current research. Fitness function could also play a vital role. Moving the fitness from the random conceptualization to a more systematic method could exclude solutions that need not be evaluated [17]. The exclusion task minimizes the required time execution and the complexity of the applied method. On the other hand, the main parameters of the genetic algorithm model especially the mutation rate also have an impact on the algorithm convergences and the performance [11, 22]. The search space reaches an inefficient process when the mutation rate becomes too high while failure to mitigate the premature convergence can result in low mutations. Therefore, efficient determination is required. This research focuses on enhancing genetic algorithm performance with the support of one of the vital optimization models which is particle swarm intelligence. The proposed enhancement ensures the optimized solution with the least complexity in time and operation.

The study applies the proposed enhanced model to one of the most vital fields, education. The idea of having good learning as a result of good teaching only has not become quite accurate. Learning in its traditional meaning is limited to the provided teaching topics quality. However, in case these topics are not compatible with the student's critical thinking and interestingness, then, it does not matter how high quality the courses are, it still will not provide the student with the prospective career level. As a consequence, the structural contradiction of the students has become deep, yet clear. Therefore, additional educational effectiveness factors have succeeded in gaining their own attention. One prominent factor is the fact that the student may have the required level of good learning, however, as it does not meet his qualifications or personality, then he may not gain the estimated learning outcomes [5]. Therefore, a vital research question has been raised, it is how could student preferences and personality contribute to students' track selection?

Continuously seeking educational quality has deepened the reformation mission to move beyond the courses' content, curriculum, instructors' experience, and exam equity [28]. All these objectives are still on focus; however, other vital approaches have proved their effectiveness in the quality-seeking mission [27]. Therefore, reforming the map of educational quality is the focus of this research. The educational society is currently paying more attention to the effectiveness of intervening in non-traditional, proven-effectiveness sources. Some of these sources are the data describing the student personality and preferences, other sources are the historical educational institutions' data, and more sources are the market scope, required qualifications, and interestingness. The current research aims to bind the goals of exploring, learning, thinking, and teaching, targeting to build a beneficial learning community. This goal is reached by the proposed framework which explores the students' abilities and preferences and then provides validated effective educational track recommendations. Ensuring the most suitable track that is oriented to the student leads directly to continuous career development and dramatically raises confidence in the educational system. The proposed framework aims to interpret the data sources using business intelligence methods as they require sophisticated solutions for the required targeted insights. Data mining methods contribute to the patterns discovery and students' tracks' classification according to the interpreted data. Applying analytical tools over the data sources supports the unification among data sources diminishes the loss of valuable data cohesion and avoids unrecognizing valuable relationships.

2. Research Problem

The optimization concept was first initiated in the 60s with the target of solving engineering problems [2, 4, 5, 8]. The main concept was to apply a set of bio-nature operators to nominate a set of solutions for the problem in which the most valued operators are the mutation and the strategy for solution selection [29]. One of the earliest developed algorithms is the genetic algorithm. One of the main features of genetic algorithms is their generic perspective as a biological abstraction methodology rather than a solution to a specified problem. As described by Mzili *et al*. [24], the foundation of genetic algorithms is to move a population of chromosomes representing the solution to another using a set of operations including mutation, inversion, crossover, and selection.

A vital parameter is the fitness function mechanism and its method of measuring the chromosomes' effectiveness in reaching the solution. Epistasis is one of the main issues that significantly affects the fitness function behavior [6, 18]. Epistasis is the fitness feature that measures the interaction level between the gene values while the chromosome resides in various locations. The direct relationship between the epistatic and the convergence to the local optimum results in the convergence delay of the method to reach a solution.

In general, the genetic algorithm could be illustrated in the following steps:

1. Identify a set of chromosomes in a random fashion with a fixed length.
2. Calculate the fitness function for all contributing chromosomes in the population.
3. If a solution is reached, stop, otherwise, continue.
4. Repeat from step 5 to step 8 until satisfying n offspring initiated.
5. Select step Identify two chromosomes on a random basis for the created offspring.
6. Combination step combine the two chromosomes and create two offspring.
7. Mutation step mutate the genes that are identified in a random fashion for all offspring.
8. Replacement step identify the offspring as a replacement of the identified chromosomes.

Following the previous steps, the contribution of the current research could be described in a set of points. First, in the selection step, the identification process is performed randomly. One of the main bottlenecks affecting the PSO performance is the randomness of the process in searching the data space. This randomness has been a main reason for performance degradation in one direction as well as lacking the ability to reach the local optima. In this research, a novel method is proposed to avoid the randomness selection. The proposed method is based on a more systematic approach for the selection task which minimizes the performance deviation. On the other hand, focusing on the fitness function performance, the repeating calculation of the fitness function should be also performed with the highest speed to avoid the calculation repeating on the algorithm performance. The current research also proposes an adapted fitness function for the genetic algorithm which is based on the evaluation metrics impact on the results accuracy. The proposed adaptation ensures reaching the solution with the highest possible speed which provides a higher positive impact on reducing the performance deviation.

Focusing on the fitness function, the simple definition of the fitness function can be stated as a function that evaluates a defined solution to a problem.

The fitness function of an algorithm is to keep evaluating the given problem solutions until reaching the optimum solution. The fitness function should be able to perform efficiently ensuring to reach the optimum solution. The performance of the fitness function has a direct impact on the algorithm's overall performance. In some situations, the fitness function is not able to reach the optimum solution, rather, it provides the approximate optimum possible solution which degrades the algorithm's accuracy. Therefore, focusing on enhancing the fitness function performance while ensuring reaching the optimum solution is one of the main goals of this research. One of the directions that supports increasing the fitness function speed in finding the required solution is vectorization. The mechanism of the genetic algorithm inference checks each point solely at a time and examines this defined point with the fitness function. While vectorization sets a set of points into a vector and produces a matrix in which the rows represent the points and the columns represent the evaluation of the solution given these points. Accordingly, the fitness function is obviously a problem dependent on one side and one of the main criteria for the algorithm's success on the other side.

3. Related Work

Genetic algorithm is one of the most popular heuristics strategies. Its ability to identify the highest beneficial elements for the optimal solution in the early stages is a unique advantage. However, the randomness, diversity of identification, and global optimum discovery are some of the bottlenecks in the algorithm that hinder its performance [7]. Some solutions are proposed. One solution is following the nearest neighbor heuristic selection. Elements that are identified to belong to the nearest group could be the seeds for subsequent search levels [2]. Another strategy was proposed by Bourahouat *et al.* [8] with a greedy strategy. In this research, the gene bank seeds were considered the population that is produced when the fitness reached the defined threshold with the shortest length. On the other hand, the research by Mzili *et al.* [24] highlighted the contribution of sorting the population for generating children's solutions. The research argued that the concept of sorting which was based on fitness value reveals better sub-solutions. Moreover, some survey research presented the earlier work performed for genetic algorithm enhancement and presented the performance of different approaches. The survey comparison revealed the outperformance of NN over other strategies. The performance of selective approaches and gene banks was also high performance.

The research by Valenzuela *et al.* [32] also performed a study for different enhancement strategies including NN, selective, sorting, and others. With the TSP experiment, the performance order was also confirmed by the research by measuring a set of evaluation measures including error and convergence rates. The concept of ordering was also followed in Alaoui *et al.* [3] which is applied to the transferring strategy of services. This research was one of the many research that confirmed the order strategy performance. However, the researchers focused on one direction of performance which is the deviation rates without considering the computation cost. Different algorithms contributed to the research for enhancing the genetic algorithm performance. For example, in [3], the k-means algorithm was applied for initiating the population using the grouping strategy by Li [21]. However, the randomness of the initial centroids of the k-means strategy was one of the drawbacks. Another direction in Soni and Kumar [30] followed the knowledge representation strategy. The research was based on identifying the significant features which were then used to explore the initial population which helped in minimizing the crossover in the following iterations. The splitting approach was also adopted in Korejo *et al.* [19] which proposed a regression model to identify the initial seeds and perform a series of clustering parental trees of four branches. It is not confirmed whether the distribution could reveal higher performance by more or less branch distribution. As presented in the literature, the initial population and number of iterations with the diversity existence were tackled in different research, however, the performance of the algorithm could still be enhanced with more efficient solutions.

Focusing on education, up to our knowledge, little work has been accomplished for the students' track recommendations. A study was conducted in Qaffas *et al.* [28] to identify the student courses. The research conducted a well-known classification algorithm by using students' degrees to classify students to a set of tracks with no further interpretation. Another research in Goswami and Dubey [15] proposed an approach that applies data mining for career recommendations. The research utilized graduate degree data for the forecasting task, however, the current research argues that marks should not be the only countable factor for accurate recommendations. Regarding the factors affecting the student's progress, research was conducted in Parhizkar *et al.* [26] to identify these factors in universities using only geographical data which are then analyzed by direct statistical methods. The study in Bujang *et al.* [9] applied an experiment using previously collected academic data for students to predict students' performance which is considered in the research as an early warning for students about their achievement. Another research in Qaffas *et al.* [28] has tackled the fact of the personal data contribution in the prediction task, however, the research was limited in the data and segment scope. Predicting career paths is also highlighted in Qaffas *et al.* [28] which used four data sources including the student academic data, surveys, and behavioral data. However, the research was based on applying the clustering technique for the prediction task with no further analysis which moved the research to collaborative recommendations.

4. Proposed PSO_GA Algorithm

This research focuses on two of the most efficient algorithms that are developed based on optimization of the selected population. Both algorithms act in the same manner. PSA is based on the identification of a child and evaluation of the identified child while GA is based on the integration which operates with the same process.

GA mechanism as described in the literature works by evaluating the presented solutions and identifies the most suitable one according to the given constraints at a defined iteration. The generation of children which represents new solutions for examination is applied through mutation. The reconstruction of new child solutions is applied by involving the pre-generated parents. The generated child originally emerged from the collaboration of the parents' information. On the other hand, another strategy for finding solutions is mutation. The solutions represented in a child, or more, are initiated from a member in each mutation cycle with the ability to adapt the parent through an optimization approach targeting the best-emerged child. At the end of the stages, a new population emerges which is usually validated with respect to the current solutions to explore the good solutions and then starts the new algorithm iteration with the explored population. Maintaining the best solutions for each iteration is an essential step and the iterations are finalized if one or more termination conditions are triggered, then, the best solution is finally identified (see Figure 1).

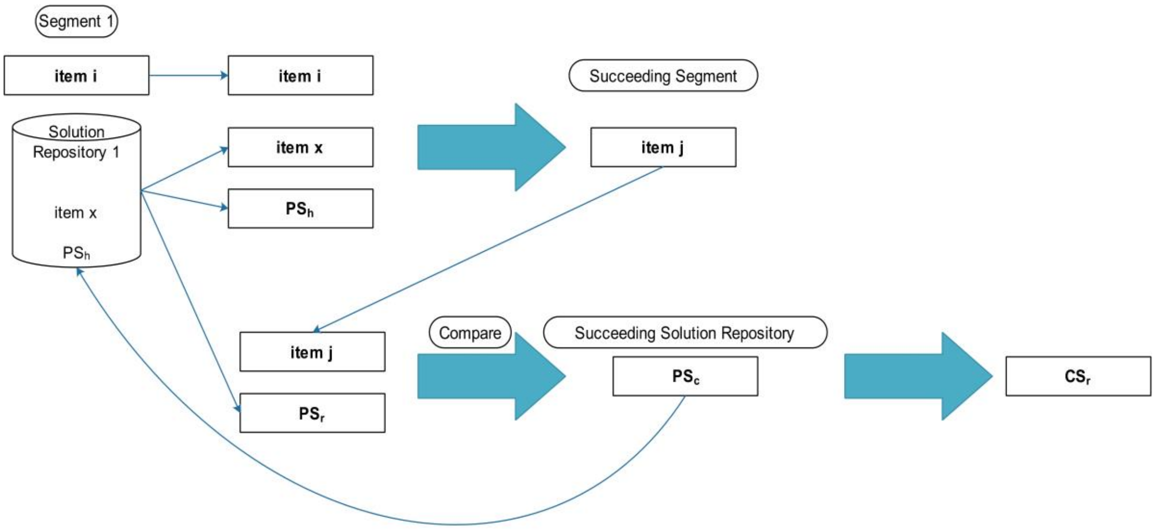

Figure 1. Genetic algorithm construction with three leveling stages.

Formally speaking, the following can represent the formal representation of the main genetic algorithm process players. *D* is the dataset, and *S* is the starting segment of the dataset representing the algorithm population which includes *i* as the number of contributing items in the iteration *t*. The items *ix* and *iy* are the item members of *D* and *S* respectively. The best global solution *ibs* in the iteration is identified when this solution is identified as the best solution in all iterations. In each iteration, three solutions are considered the parent solutions are *psh*, *psr*, *psc* which are the main dataset members, the starting iteration set member, and the current iteration set member. They contribute to generating the best child solution *csr*. The population for the following iteration is defined according to the best child solution which is generated in the iteration until termination conditions are met.

Three main operators contribute to the proposed algorithm (see Figure 2), they are selection, crossover, and mutation. The first operator selects the main players in the current stage while the remaining two operators are able to explore the best player as the best solution in the current stage.

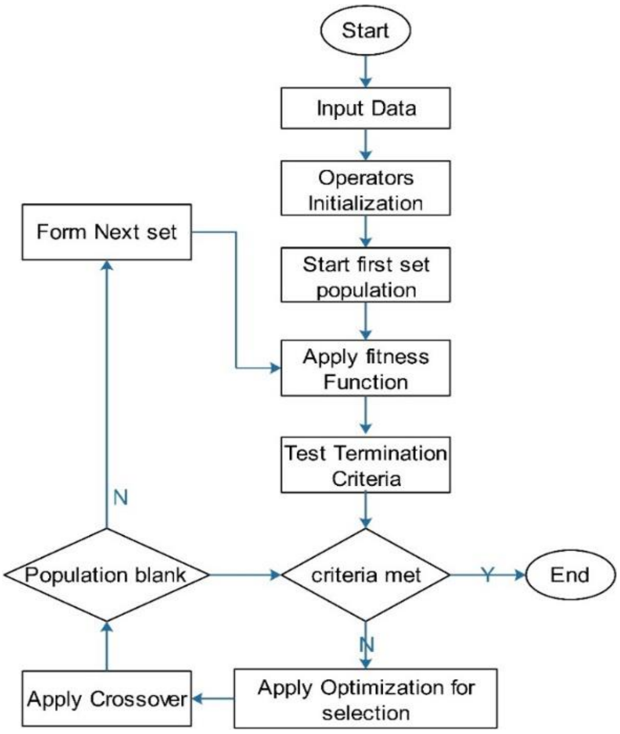

Figure 2. Main steps of proposed adaptation.

The main contribution of the current study is adapting the selection process which is usually performed randomly. Generating the best child solution can be formally represented as in equation 1 where v_x are identified as positive constants. The proposed equation for selection ensures the best fitness function for the selection. Moreover, the selection is applied for two portions of the dataset, each has its own. The selection process is based on two different equations [14, 15, 25] as in Equations (2) and (3) while the score that determines the penalty score is illustrated in Equation (4). In Equation (2), the equation determines the sum of $\vartheta\ (PS_x^t, PS_y^t)$ which refers to the score of each two consecutive players. Equation (3) identifies the accumulation of all players' attributes in a defined iteration while $v_1 and\ v_2$ represent the beginning and end of the gap scores in Equation (4).

$$\begin{aligned} cs_r^t = {} & v_1 * (ps_c^{t-1} - ps_c^{t-2}) \\ & + v_2 * \left(ps_h^{t-1} - ps_h^{t-2}\right) \\ & + v_3 * (ps_r^{t-1} - ps_r^{t-2}) \end{aligned} \tag{1}$$

$$f(P_1) = \sum\sum \vartheta\ (PS_x^t, PS_y^t) \tag{2}$$

$$f(P_1) = \sum_{att=1}^{x} ps_x^t \tag{3}$$

$$Score = v_1 + v_2(\sum_{i=1}^{n} |Score_i|) \tag{4}$$

The best child solution exploration process is illustrated in Figure 3.

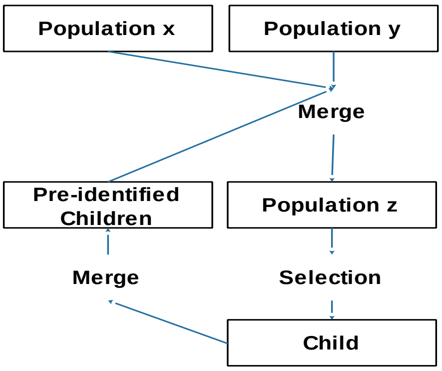

Figure 3. Child exploration process.

Figure 3 presents the mechanism of child solution creation by utilizing the three generations. The proposed adaptation of GA ensures an equal opportunity for all members to be selected as the best solution for all the generations that are involved in the selection process. Following the described formula, leads to minimizing the impact of the mutation operations as the three generations are equally involved in the process with no intention for random operation. According to the previous description, the similarity between the described process and the fundamental particle swarm optimization algorithm is noticed except for the velocity parameter. The proposed utilization of PSO in GA mechanism ensures an advancement over the traditional algorithm. Comparing the proposed mechanism in [14, 15, 25] which child creation is based on updating the probability distribution or a set of participating swarms, the proposed creation update ensures its ability to efficiently perform with different problems nature. The main advancement of the proposed adaptation is the ability to utilize previously detected solutions and enhance these solutions by creating an updated solution based on this utilization.

5. Experimental Study

This section discusses the performed experiments for the proposed algorithm evaluation. The contributed dataset is BAlisBASE [15]. The dataset and contributing evaluation measures are discussed, and the results are compared with the existing literature. The proposed adaptation effectiveness is finally confirmed by the non-parametric Wilcoxon test [25].

5.1. Dataset Description

BAliscore is utilized in different research to measure the conserved scores for sequences. The utilized version in this experiment is BAliscore 3.0. It consists of a total of two hundred and eighteen subsets representing alignments of multiple proteins. BAliscore is divided into a total of six subsets. Each subset presents a defined protein family [15]. The six subsets can be described in Table 1.

Table 1. Description of BAliscore dataset subsets.

	No of sets of sequence	Identity percentage range
Sub1	38	< 20
Sub2	44	20-40
Sub3	41	>40
Sub4	30	25-40
Sub5	49	>20
Sub6	16	>20

5.2. Experimental Results

The experiment setup includes different aspects. One of these aspects is the setting of the experiment parameters. Setting the parameters of the experiment considered the standards of GA algorithm settings [22]. The parameters include the dataset size (150), crossover and mutation probabilities (0.2-0.5), experiment counts per each dataset (20), and count of generations (100). The proposed model has been tested against the existing literature such as in [33]. Two evaluation metrics are used for the evaluation phase, they are Total Column Score (TS); which specifies the percentage of the true count of aligned columns to the total count of columns; and Quality (Q); which specifies the appropriate division percentage of pairs aligned [31]. In case the results resemble the value of the reference, then the percentage is zero, otherwise, the percentage is a fraction less than one. Table 2 illustrates the results of the proposed model by comparing these results with the literature. The experiment applied the compared models on the BALiBASE sets. Q and Tc metrics are calculated for the models in all subsets. According to the illustrated results (Table 2), the proposed model revealed the highest average scores in addition to the highest and consistent scores in most of the applied subsets. Moreover, Figures 4 and 5 illustrate a comparison for all iterations between the proposed model and the highest scores in the literature in terms of Q and TC measures respectively.

Table 2. Comparative results of the proposed model and the literature.

Model	Sub1		Sub2		Sub3	
	Q	TC	Q	TC	Q	TC
Proposed Approach	0.89	0.76	0.94	0.91	0.9	0.62
T-Coffee [10]	0.49	0.29	0.79	0.859	0.85	0.19
HMOAB C [28]	0.69	0.49	0.94	0.887	0.91	0.5
ClustalX [30]	0.6	0.39	0.82	0.794	0.9	0.49
BSAGA [10]	0.79	0.7	0.69	0.851	0.88	0.49
MSAProbs [33]	0.59	0.48	0.89	0.87	0.89	0.29
Model	**Sub4**		**Sub5**		**Sub6**	
	Q	**TC**	**Q**	**Q**	**TC**	**Q**
Proposed Approach	0.93	0.67	0.91	0.65	0.93	0.8
T-Coffee [10]	0.87	0.54	0.89	0.39	0.79	0.19
HMOABC [28]	0.88	0.65	0.92	0.66	0.75	0.61
ClustalX [30]	0.87	0.49	0.87	0.45	0.76	0.54
BSAGA [10]	0.79	0.69	0.9	0.7	0.79	0.59
MSAProbs [33]	0.65	0.56	0.89	0.61	0.9	0.61

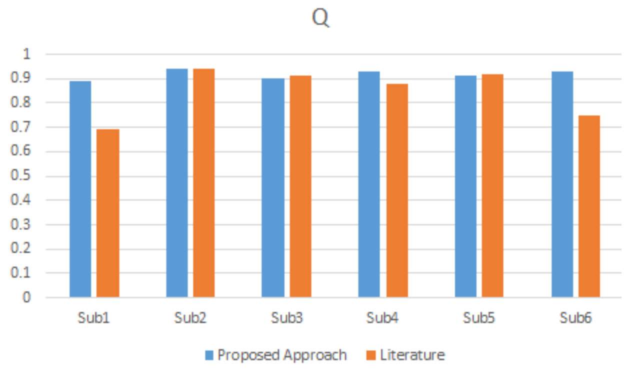

Figure 4. The performance comparison for Q measure between the proposed approach and the literature.

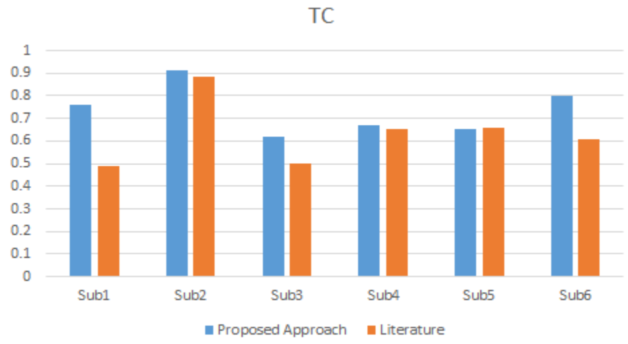

Figure 5. The performance comparison for TC measure between the proposed approach and the literature.

6. Applying the Proposed Model to Students' Dataset

The model has been applied to a students' dataset. The dataset included 30641 records described by 14 attributes. The attributes are distributed between main student data such as the gender and the parents' education, education attributes such as the number of studied hours, and scoring attributes such as math, reading, and writing scores. The target was to predict the expected student score. Following the same process and setting the same parameters, the data has been divided into six subsets and the model has been applied. The results present the high performance of the proposed model which confirms its applicability (see Table 3).

Table 3. Results for applying the proposed model to students' dataset.

Sub1		Sub2		Sub3		Avg.	
Q	TC	Q	TC	Q	TC	Q	TC
0.63	0.68	0.82	0.9	0.8	0.79	0.842	0.796
Sub4		**Sub5**		**Sub6**			
Q	**TC**	**Q**	**TC**	**Q**	**TC**		
0.9	0.8	0.9	0.78	0.97	0.83		

7. Conclusions

Evolutionary algorithms such as genetic algorithms have proved their effectiveness and reliability in optimization solutions. The genetic algorithm is one of the most powerful algorithms in optimizing solutions to various problems. However, such algorithms suffer from performance issues resulting from bottlenecks in their mechanisms. This research proposes an effective solution for raising the performance of a genetic algorithm with the idea of merging its mechanism with one of the swarm intelligence techniques. The proposed solution presents an effective model for the initialization task as well as minimizing the iterations while ensuring the optimized solution. The mimic concept for natural processes has leveraged the genetic algorithm computation to the optimized level. Linking genetic algorithms and particle swarm intelligent algorithms has proved their effectiveness through a set of experiments. Moreover, the proposed adapted algorithm has been applied to the education field in generating the most effective track for students targeting to enhance the students' performance which is considered one of the strategic targets in all economies

Acknowledgment

This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant no. (UJ-23-DR-8). Therefore, the authors thank the University of Jeddah for its technical and financial support.

References

[1] Abdelgwad M., Abed A., and Bahloul M., "Authenticated Diagnosing of COVID-19 Using Deep Learning-Based CT Image Encryption Approach," *Future Computing and Informatics Journal*, vol. 8, no. 2, pp. 31-58, 2022. https://digitalcommons.aaru.edu.jo/fcij/vol8/iss2/4

[2] Abo-Elnaga Y. and Nasr S., "Modified

Evolutionary Algorithm and Chaotic Search for Bilevel Programming Problems," *Symmetry*, vol. 12, no. 5, pp. 767, 2020. DOI:10.3390/sym12050767

[3] Alaoui N., Adamou-Mitiche A., and Mitiche L., "Effective Hybrid Genetic Algorithm for Removing Salt and Pepper Noise," *IET Image Process*, vol. 14, no. 2, pp. 289-296, 2020. https://doi.org/10.1049/iet-ipr.2019.0566

[4] Al-Oqaily A. and Shakah G., "Solving Non-Linear Optimization Problems Using Parallel Genetic Algorithm," *in Proceedings of the International Conference on Computer Science and Information Technology*, Amman, pp. 103-106, 2018. DOI:10.1109/CSIT.2018.8486176

[5] Alyousufi A., Naidu V., Jesrani K., and Dattana V., "Tracking Students' Progress using Big Data Analytics to Enhance Student's Employability: A Review," *in Proceedings of the International Conference on Teaching and Learning-Digital Transformation of Education and Employability*, Muscat, pp. 1-7, 2023. https://doi.org/10.1051/shsconf/202315607001

[6] Amorim A., Zafalon G., De Godoi Contessoto A., Valêncio C., and Sato L., "Metaheuristics for Multiple Sequence Alignment: A Systematic Review," *Computational Biology and Chemistry*, vol. 94, pp. 107563, 2021. https://doi.org/10.1016/j.compbiolchem.2021.107563

[7] Arkhipov D., Wu D., Wu T., and Regan A., "A Parallel Genetic Algorithm Framework for Transportation," *IEEE Access*, vol. 8, pp. 106506-106515, 2020. DOI:10.1109/ACCESS.2020.2997812

[8] Bourahouat G., Abourezq M., and Daoudi N., "Word Embedding as a Semantic Feature Extraction Technique in Arabic Natural Language Processing: An Overview," *The International Arab Journal of Information Technology*, vol. 21, no. 2, pp. 313-325, 2024. https://doi.org/10.34028/iajit/21/2/13

[9] Bujang S., Selamat A., Krejcar O., Mohamed F., Cheng L., and Chiu P., "Imbalanced Classification Methods for Student Grade Prediction: A Systematic Literature Review," *IEEE Access*, vol. 11, pp. 1970-1989, 2023. DOI:10.1109/ACCESS.2022.3225404

[10] Chowdhury B. and Garai G., "A Bi-Objective Function Optimization Approach for Multiple Sequence Alignment Using Genetic Algorithm," *Soft Computing*, vol. 24, pp. 5871-15888, 2020. https://doi.org/10.1007/s00500-020-04917-5

[11] De Jong K., *Evolutionary Computation: A Unified Approach*, MIT Press, 2006. https://ieeexplore.ieee.org/servlet/opac?bknumber=6267245

[12] DeRonne K. and Karypis G., "Optimal Pairwise Sequence Alignment*," IEEE/ACM Transactions of Computational Biological Bioinformtics*, vol. 10, no. 2, pp. 481-493, 2013. https://doi.org/10.1109/TCBB.2013.2

[13] Du X., Du C., Chen J., and Liu Y., "An Energy-Aware Resource Allocation Method for Avionics Systems Based on Improved Ant Colony Optimization Algorithm," *Computers and Electrical Engineering*, vol. 105, pp. 108515, 2023. https://doi.org/10.1016/j.compeleceng.2022.108515

[14] Ehrgott M., *Multicriteria Optimization*, Springer Science and Business Media, 2005. https://doi.org/10.1007/3-540-27659-9

[15] Goswami A. and Dubey K., "A Novel Population-Based Optimization for Multiple Sequence Alignment in Protein Sequencing," *Engineering Sciences*, vol. 21, no. 786, pp. 1-9, 2022. DOI:https://dx.doi.org/10.30919/es8d786

[16] Gupta S. and Sharma N., "Machine Learning Driven Threat Identification to Enhance FANET Security Using Genetic Algorithm," *The International Arab Journal of Information Technology*, vol. 21, no. 4, pp. 711-722, 2024. https://doi.org/10.34028/iajit/21/4/12

[17] Handl J. and Knowles D., "Multiobjective Optimization in Bioinformatics and Computational Biology," *IEEE/ACM Transactions of Computational Biological Bioinformtics*, vol. 4, pp. 279-292, 2007. DOI:10.1109/TCBB.2007.070203

[18] Hu X. and Chuang Y., "E-Commerce Warehouse Layout Optimization: Systematic Layout Planning Using a Genetic Algorithm," *Electronic Commerce*, vol. 23, pp. 97-114, 2023. https://doi.org/10.1007/s10660-021-09521-9

[19] Korejo I., Yang S., Brohi K., and Khuhro Z., "Multi-Population Methods with Adaptive Mutation for Multi-Modal Optimization Problems," *International Journal on Soft Computing, Artificial Intelligence and Applications*, vol. 2, no. 2, pp. 1-19, 2013. https://airccse.org/journal/ijscai/papers/0413scai01.pdf

[20] Kouka N., BenSaid F., Fdhila R., Fourati R., Hussain A., and Alimi A., "A Novel Approach of Many-Objective Particle Swarm Optimization with Cooperative Agents Based on an Inverted Generational Distance Indicator," *Information Sciences*, vol. 623, pp. 20-241, 2023. https://doi.org/10.1016/j.ins.2022.12.021

[21] Li B., "An Experiment of K-Means Initialization Strategies on Handwritten Digits Dataset," *Intelligent Information Management*, vol. 10, no. 2, pp. 43-48, 2018. DOI:10.4236/iim.2018.102003

[22] Maulik U. and Bandyopadhyay S., "Genetic

Algorithm-Based Clustering Technique," *Pattern Recognition*, vol. 33, no. 9, pp. 1455-1465, 2000. https://doi.org/10.1016/S0031-3203(99)00137-5

[23] Mousa M., Khedr A., and Idrees A., "Hierarchical Method for Automated Text Documents Classification," *The International Arab Journal of Information Technology*, vol. 22, no. 1, pp. 11-19, 2025. https://doi.org/10.34028/iajit/22/1/2

[24] Mzili T., Mzili I., Riffi M., and Dhiman G., "Hybrid Genetic and Spotted Hyena Optimizer for Flow Shop Scheduling Problem," *Algorithms*, vol. 16, no. 6, pp. 265, 2023. https://doi.org/10.3390/a16060265

[25] Naznin F., Sarker R., and Essam D., "Vertical Decomposition with Genetic Algorithm for Multiple Sequence Alignment," *BMC Bioinformatics*, vol. 12, no. 353, pp. 1-26, 2011. https://doi.org/10.1186/1471-2105-12-353

[26] Parhizkar A., Tejeddin G., and Khatibi T., "Student Performance Prediction Using Datamining Classification Algorithms: Evaluating Generalizability of Models from Geographical Aspect," *Education and Information Technologies*, vol. 28, no. 11, pp. 14167-14185, 2023. https://doi.org/10.1007/s10639-022-11560-0

[27] Perez A., Green J., Ball G., Amin M., Compton S., and Patterson S., "Behavioural Change as a Theme that Integrates Behavioural Sciences in Dental Education," *European Journal of Dental Education*, vol. 26, no. 3, pp. 453-458, 2022. DOI:10.1111/eje.12720

[28] Qaffas A., Alharbi I., Idrees A., and Kholeif S., "A Proposed Framework for Student's Skills-Driven Personalization of Cloud-Based Course Content," *International Journal of Software Engineering and Knowledge Engineering*, vol. 33, no. 4, pp. 603-617, 2023. https://doi.org/10.1142/S0218194023500092

[29] Rubio-Largo A., Vega-Rodríguez M., and González-Álvarez D., "A Hybrid Multiobjective Memetic Metaheuristic for Multiple Sequence Alignment," *IEEE Transactions of Evolutionary Computation*, vol. 20, pp. 499-514, 2015. DOI:10.1109/TEVC.2015.2469546

[30] Soni N. and Kumar T., "Study of Various Mutation Operators in Genetic Algorithms," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 4519-4521, 2014. https://www.ijcsit.com/docs/Volume%205/vol5issue03/ijcsit20140503404.pdf

[31] Thompson J., Koehl P., Ripp R., and Poch O., "BAliBASE 3.0: Latest Developments of the Multiple Sequence Alignment Benchmark," *Proteins Structural Function Bioinformatics*, vol. 61, pp. 127-136, 2005. DOI:10.1002/prot.20527

[32] Valenzuela O., Rojas F., Pomares H., Ortun F., Florido J., Urquiza J., and Rojas F., "Optimizing Multiple Sequence Alignments Using a Genetic Algorithm Based on Three Objectives: Structural Information, Non-Gaps Percentage and Totally Conserved Columns," *Bioinformatics*, vol. 29, no. 17, pp. 2112-2121, 2013. https://doi.org/10.1093/bioinformatics/btt360

[33] Wang J., Zhang F., Liu H., Ding J., and Gao C., "Interruptible Load Scheduling Model Based on an Improved Chicken Swarm Optimization Algorithm," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 232-240, 2021. DOI:10.17775/CSEEJPES.2020.01150

[34] Zhang G., Jia N., and Zhu N., "Robust Drone Selective Routing in Humanitarian Transportation Network Assessment," *European Journal of Operation Research*, vol. 305, no.1, pp. 400-428, 2023. https://doi.org/10.1016/j.ejor.2022.05.046

Mourad Kaddes received the Ph.D. Degree from University of le Havre, France in 2013. He is currently an Assistant Professor at the University of Jeddah. His research interests include Real-Time System, Real-Time Database Systems and Unmanned Aerial Vehicle.

Ayman E. Khedr currently a professor at the University of Jeddah. have been the Vice Dean of post-graduation and research and the Head of the Information Systems Department in the Faculty of Computers and Information Technology, at Future University in Egypt. He is A Professor in the Faculty of Computers and Information, at Helwan University in Egypt. have previously worked as the general manager of the Helwan E-Learning Center. His research focused on the themes (scientific) Data and Model Management, Data Science, Big Data, IoT, E-Learning, Data Mining, Bioinformatics, and Cloud Computing.

Amira M. Idrees a Professor in Information Systems. She has been the Head of Scientific Departments and the Vice Dean of Community Services and Environmental Development, at the Faculty of Computers and Information, at Fayoum University. A professor in the Faculty of Computers and Information Technology at Future University, the Head of IS Department, and the Head of the University Requirements Unit. Research interests include Knowledge Discovery, Text Mining, Opinion Mining, Cloud Computing, e-Learning, Software Engineering, Data Science, and Data Warehousing.