Implement of Students Result by Using Genetic Algorithm

Abhishek katiyar[1*], Anil pandey[2]

[1,2]Department of Computer Science, Invertis University Bareilly (U.P.), India

Corresponding Author: Abhishekkatiyar1003@gmail.com

Abstract: The artificial intelligence technique such as Genetic algorithm plays a significant role for handling in many fields such as artificial intelligence, engineering, robotic, etc. This is the technique to evaluate the new populations from natural population and provide the best result generation to generation. This is applied in students' quantitative data analysis to identify the most impact factor in their performance in their curriculum. The results will help the educational institutions to improve the quality of teaching after evaluating the marks achieved by the students' in academic career. This student analysis model considers the quantitative factors such as compiler, automata, data structure and other departmental marks to find the most impacting factor using genetic algorithm.

Keywords: students' performance, quantitative factors, genetic algorithm, influencing parameter, student's evaluation results.

I. INTRODUCTION

In present day's educational system, performance in career growth is determined by the assessment and examination achievements. The assessment is carried out by curricular activities such as class test, viva, seminar, assignments, general proficiency, attendance, lab work and finally semester exams. On considering the quantitative factors such as attendance, CGPA (cumulative), marks obtained in theoretical , technical's, elective and departmental papers, knowledge in learning, understanding capability, communication etc. thus the computation is done by real genetic algorithm is used to find the most influencing factor in those selected factors. In order to formulate the equation that contains several important subjects and its marks, their priority values plays the major role in the sensitivity analysis. It can be added to the equation along with their parametric variables and the output returned from the equation is considered as the overall performance value. Now according to the passing criteria as fixed, the performance is classified as good, average and poor categories. The real genetic algorithm is applied RGSPAT (Real Genetic Student Performance Analysis Tool) that helps us to know where the variations happened can be identified. From the result, the focus towards the potential problems in the course is gained. Educators can also use this information to guide their way of evaluation and implementing according to curriculum changes.The remaining section of this paper is organized as follows, section 2 describes the several techniques used in evaluation of student's performance, section 3 discuss about designing principles of various components in our student performance analysis model, section 4 describes about application of real genetic algorithm to find the most important quantitative feature, section 5 about working model of our Analysis tool   and section 6 concludes the paper with remarkable outcome.

II. LITERATURE REVIEW

P ramya and M, Mahesh kumar Discuss data mining techniques like classification algorithm such as naïve Bayesian classification algorithm and weight naïve Bayesian algorithm on the student data set which is collected from LBRCE IT dept., malavaram for building this model. Based on those result we have to overcome on suicide rate of student, who have to weak their performance [1].

Chew li sa and Emmy Dehlana Hossain discuss the data mining techniques, classification algorithm applied in this project to ensure the prediction of student performance in course "TMC 1013 system analysis and design". The main theme of this paper to assists the lecturers to give the permission to take the previous semester record throughout the semester to analyzed the performance of student who are failed in "TMC 1013 system analysis and design"[2].

Saddam Khan and Sunny Gupta discuss various techniques of data mining to predict and improve the quality and performance of the student in education sector. These techniques are clustering to group the similar data type, association rule

to find the relationship of the student in education, decision tree is better in terms of speed, ANN (Artificial Neural Network) in terms of accuracy and error rate but it is very slow and genetic algorithm to give the better performance by operators such as crossover and mutation[3].

Somnath mazi discuss about various types of crossover operation to improve the performance of students such as linear crossover, order based crossover, heuristic crossover etc. This is applied in students' quantitative data analysis to identify the most impact factor in their performance in their curriculum. The results will help the students as well as educational institutions to improve the quality of teaching after evaluating the marks achieved by the students' in academic career. So proposed genetic algorithm with different crossover with new student analysis model to find the most impacting factor in student's performance[4].

A. Martin and V. Prasanna Venkatesan discuss about various techniques that are used to evaluate and predict the performance of employees, staffs, teachers and students and implementation of real genetic algorithm in bank applications. They are as follows.The most impact banking features in bankruptcy model are analyzed and results are produced. Dakin's bankruptcy model and its various features such as net income, cash assets, current assets, sales and total assets are considered. On application of Genetic algorithm they predicted the most important feature in this Deakin model is cash assets/ total assets. This ratio is crucial for predicting the bankruptcy with 96% accuracy; similarly this approach is adapted here to find the most impacting parameter from the students' results[5].

Ramjeet Singh Yadav discuss fuzzy Expert System tested with 20 student's marks obtained by semester-1 and semester-2 examinations, both inputs had same Triangular Membership Functions. The proposed a Fuzzy Expert System (FES) for student academic performance evaluation based on fuzzy logic techniques uses a suitable fuzzy inference mechanism and associated rule[6].

O.K Chaudhari and P.G khot discussThe academic performances are taken and compare the results (performance) with existing statistical method. Similarly FES was developed for evaluation of teachers' performance in teaching activity is especially relevant for the academic institutions. It helps to define efficient plans to guarantee quality of teachers and the teaching learning process by considering students' feedback, result, students' attendance, teaching learning process and academic development of teachers[7].

V.O. Oladokun discuss The various factors in his research that may likely influence the performance of a student were identified such as examination scores, age on admission, parental background and gender etc., The Artificial neural network (ANN) shown the potential for enhancing the effectiveness of a university admission system. The model was developed based on some selected input variables from the pre admission data of five different sets of university graduates. It achieved an accuracy of over 74%, which shows the potential efficacy of Artificial Neural Network as a prediction tool and a selection criterion for candidates seeking admission into a university[8].

Osman Taylan and Bahattin Karagozog discuss The adaptive Neuro-fuzzy technique introduced for prediction of student performance evaluation. Here Assessment of students' achievements in the previous year to evaluate the learning process and performance of students is considered to be more meaningful than referring the marks individually. The previous and current student's performance data are kept in a computer system that is an excellent source to achieve satisfactory results in evaluation of the students' performance as a group, their learning level and evaluate the quality of education. It is illustrated with students' achievements, academic performances, linguistic grades granted. It also gives the outcomes of the SAP for statistical and neuro-fuzzy modeling approaches with their linguistic values[9].

Emmanuel N.Ogor discuss the methodology by the derivation of performance prediction indicators to deploying a simple student performance assessment and monitoring system focusing on performance monitoring of students' continuous assessment (tests) and examination scores in order to predict their final achievement status upon graduation. This is done based on various data mining techniques (DMT) and the application of machine learning processes, rules are derived that enable the classification of students in their predicted classes[10] .

From these literature surveys, few researches have compared for analyzing the performance of student quantitative data and there is no considerable work on finding any equation model for to find the best parameter that influence the performance of the students. This research also analyzes the impact of quantitative parameters in student's academic performance.

III. GENETIC APPROACH FOR STUDENTS PERFORMANCE ANALYSIS

Genetic algorithms are now widely applied in science and engineering as adaptive algorithms for solving practical problems. The general acceptance is that GA is particularly suited to multidimensional global search problems where the search space potentially contains multiple local minima. Unlike other search methods, correlation between the search variables is not generally a problem. Genetic algorithm works modeling the parameters of a problem as real strings. The following figure depicts the various components of RGSPAT model.
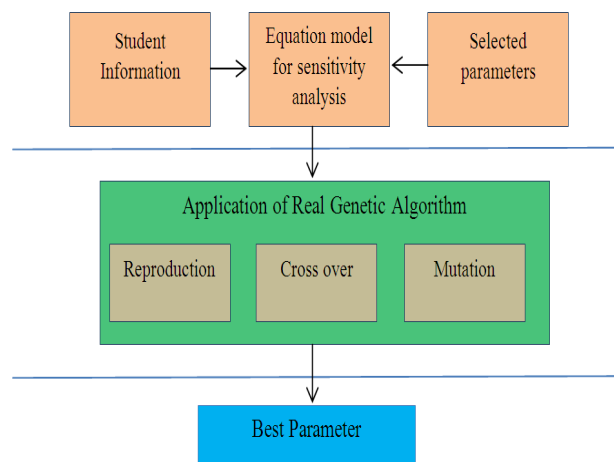

Figure 1. Design model of RGSPAT (Real genetic student performance analysis tool)

The proposed system consists of starting edge from the student analyzing data and ending edge by finding the most important parameter that is found by undergoing several processes as follows.

• Analyzing the performance information in educational systems that are quantifiable such as attendance, internal assessment marks, project marks, previous semester marks, seminar, general proficiency, papers that are considered important in that course etc.
• On collecting those parameters, the equation Construction that containing all those parameter values with weight values are designed.
• After designing the equation, the real-coded genetic algorithm is applied.
• The operations such as mutation and cross over are carried out.
• The best parameter is identified, once analyzing all the parameters through real genetic algorithm.

The quantitative factors that are considered into the Analysis of students' performance is given below.
• **PSM** Previous Semester Marks/Grade obtained in   course.
• **IA**   Internal Assessment marks.
• **SEM** Seminar Performance obtained. In each semester seminar are organized to check the performance of students.
• **ASS** Assignment marks. In each semester two   assignments are given to students by each teacher.
• **GP** General Proficiency performance. Like seminar, in each semester general proficiency tests are organized.
• **ATT** Attendance of Student. Minimum 70% attendance is compulsory to participate in End Semester Examination. But even through in special cases low attendance students also participate in End Semester Examination on genuine reason. Attendance   is divided.
• **PM** Project Work. Completion of the full project with report, presentation, system model.
• **C** Compiler subject marks obtained in semester.
• **A** Automata subject marks obtained in all semester.
• **E** Elective subject marks obtained in all semesters
• **DA** Data structure subject marks obtained in semester.
• **O** Other departmental subject marks in all semesters
• **PV** Performance value obtained from the equation as the result.

Here RGSPAT tool that is developed for predicting student performance requires the quantitative attributes of student (i.e.) measurable variables like internal mark, seminar mark, attendance, daily test marks etc. are listed below.

Table 1: **Range of Quantitative Attributes.**

Variable	Description	Possible range of values
PSM	Previous semester marks	0-100
IA	Internal Assessment	0-25
L	Lab	0-50
P	Project	0-75
ASS	Assignment	0-5
ATT	Attendance	0-10
R	Record	0-5
GP	GP viva	0-5
C	Compiler subject	0-100
DA	Data structure subject	0-100
E	Elective subject	0-100
A	Automata subject	0-100
O	Other dept. subject	0-100
PV	Overall Performance value	0-100

Of the students now the equation used to analyze and categorize the performance of the student consists of the input parameters such as attendance, internal assessments, Assignment, General proficiency, compiler subject, data structure subject, elective and automata subject marks as mentioned above. The parameter which shows more variation in the evaluation of the performance using the algorithm, it is chosen as the optimum parameter. On finding the result for each student, we can generally find the most optimizing parameter for the student management and taking necessary steps for improving the overall performance.

IV. MOST IMPORTANT PARAMETER USING RGA

This analysis is concerned with finding most important attribute that affects the performance of student. As the fore mentioned properties of RGA are highly advantageous, RGA for RGSPAT is designed using Crossover and Mutation process. For this experiment we have selected the quantitative factors among the student in College/school. The real-valued genetic algorithm (RGA) uses a real value as a parameter of the offspring in populations without performing coding and encoding process before calculates the fitness values of individuals. The performance analysis for RGSPAT model is given by the following equation.

Performance value (*PV)* = 0.1(IA) + 0.15(L) + 0.25(P) + 0.10(C) + 0.06(DA) +0.05(A) +   0.04 (E) + 0.05(O) + 0.20 (PSM)

The working of real genetic algorithm in performance analysis is as follows.
   a. [Start] Generate random population of attributes as   chromosomes
   b. [Fitness] Evaluate the fitness f(x) of each chromosome x in the population
   c. [New population] Create a new population of attributes by repeating following steps until the new population is complete

i. [Selection] Select two parent chromosomes from a population according to their fitness (which satisfies the fitness function)
ii. [Crossover] with a crossover probability Pc crossover the parents to form a new offspring (children). For real values, linear crossover is performed. ax+(1-a)y, where a is priority value in the equation and x, y are real values.
iii.[Mutation] with a mutation probability Pm mutate new offspring at each locus (position in chromosome), For real value mutation, add any random value as x + N(0,0.1).
iv [Accepting] Place new offspring in a new population
   d. [Replace] Use new generated population for a further run of algorithm.
   e. [Test] if the end conditions are satisfied, stop, and returns the best parameter in current population.
   f. [Loop] go to step b.

Initially a population of chromosomes, each of which represents a potential solution to the problem at hand, is generated randomly and each of them is evaluated by finding its fitness. The next generation of same size is created by selecting more fit individuals from this population and by applying genetic operators like crossover and mutation to them. Mutation is an operator which creates a new individual by making a random change in the old one, whereas crossover creates new individuals by combining parts from multiple individuals. Classic mutation randomly alters a single gene, while crossover exchanges genetic material between two or more parents. This completes one generation and after repeating this procedure for a number of generations, due to selective pressure the algorithm converges and it yields a better solution.

V. RESULTS AND DISCUSSION

Table 2 illustrates the sample iterative process of the genetic algorithm to find the best parameter from the population. The selected iterated values are given in the Table 2. The fitness value is chosen from these values. According to that the influencing parameter can be identified. In this application, we have chosen the performance fitness value as>2.

Table2. Iterative Process of Genetic Algorithm

Iteration No.	IA	L	P	A	C	DA	O	E	PSM	PV
S3	3.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.20
S6	6.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.50
S8	8.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.70
S10	10.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.90
S12	1.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.15
S14	1.00	4.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.45
S16	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.75
S18	1.00	8.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2.05
S20	1.00	10.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	2.35
S22	1.00	1.00	4.00	1.00	1.00	1.00	1.00	1.00	1.00	1.75
S24	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00	2.25
S26	1.00	1.00	8.00	1.00	1.00	1.00	1.00	1.00	1.00	2.75
S28	1.00	1.00	10.00	1.00	1.00	1.00	1.00	1.00	1.00	3.25
S30	1.00	1.00	1.00	2.00	1.00	1.00	1.00	1.00	1.00	1.40
S32	1.00	1.00	1.00	4.00	1.00	1.00	1.00	1.00	1.00	1.60
S34	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.06
S36	1.00	1.00	1.00	8.00	1.00	1.00	1.00	1.00	1.00	1.36
S38	1.00	1.00	1.00	10.00	1.00	1.00	1.00	1.00	1.00	1.54
S40	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	1.00	1.05
S42	1.00	1.00	1.00	1.00	4.00	1.00	1.00	1.00	1.00	1.15
S44	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.25
S46	1.00	1.00	1.00	1.00	8.00	1.00	1.00	1.00	1.00	1.35
S48	1.00	1.00	1.00	1.00	10.00	1.00	1.00	1.00	1.00	1.45
S50	1.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00	1.08
S52	1.00	1.00	1.00	1.00	1.00	4.00	1.00	1.00	1.00	1.20
S54	1.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.32
S56	1.00	1.00	1.00	1.00	1.00	8.00	1.00	1.00	1.00	1.05
S58	1.00	1.00	1.00	1.00	1.00	10.00	2.00	1.00	1.00	1.15
S60	1.00	1.00	1.00	1.00	1.00	1.00	4.00	1.00	1.00	1.25
S62	1.00	1.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.35
S64	1.00	1.00	1.00	1.00	1.00	1.00	8.00	1.00	1.00	1.45
S66	1.00	1.00	1.00	1.00	1.00	1.00	10.00	1.00	1.00	1.40
S68	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2.00	1.00	2.00
S70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	4.00	8.00	2.40
S72	1.00	1.00	1.00	1.00	1.00	1.00	1.00	6.00	9.00	2.60
S74	1.00	1.00	1.00	1.00	1.00	1.00	1.00	8.00	10.00	2.55

The following diagram illustrates the working of RGSPAT model for analyzing the student performance in automata, compiler and data structure. The respective subject marks are loaded and their corresponding influence on students is shown in the graph. In fig 2, the data structure subject Marks and compiler marks are taken whose maximum value is 90, given in y axis. The x axis represents the corresponding quantitative factors in terms of numbers. Thus each Subject marks are analyzed from their semester values and results are produced separately to find the most influencing parameter as given below. In the above Figure 2, we have selected the 50 students of IT departments and the results are used to analyze the most important parameter. This graph represents that parameter 2 and 4 have high impact on the performance of the students (2 indicates compiler subject and 4 is data structure subjects). Thus those two subjects have given high importance to reach the better performance. Thus this system can be implemented in the educational environment to evaluate the student results and to find the importance given to the subjects through analyzing their end results. Further research is to implement fuzzy clustering technique to group the students on the basis of their performance towards the selected parameters.

Table 3: Subject Marks of Previous Semester

Sr no.	Automata	Compiler	Data Structure
1	70	80	55
2	60	85	59
3	75	66	72
4	64	77	84
5	56	78	66
6	54	56	65
7	76	57	68
8	77	61	69
*			
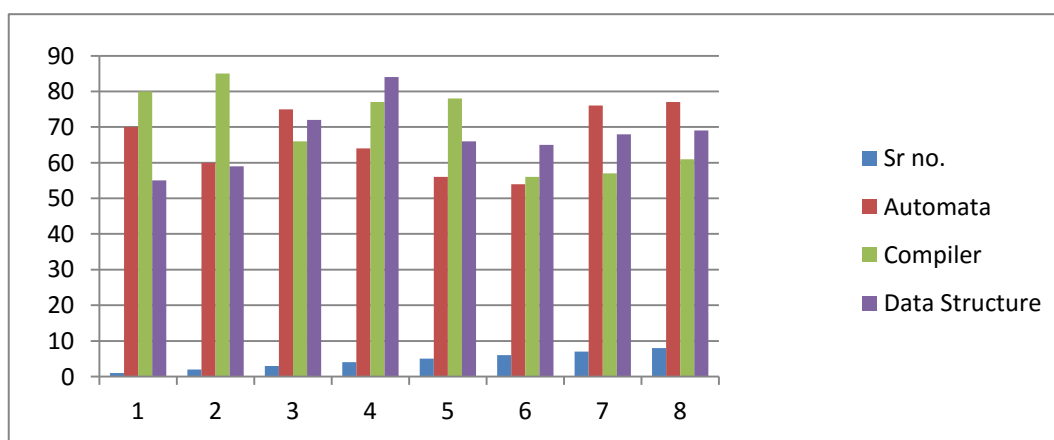
Figure 2. Working model of RGSPAT model.

VI. CONCLUSION

Thus real genetic process is successfully implemented to find the most influencing parameter that affects performance in the educational system. It will be a better solution for the classification, analyzing and evaluating the quantitative factors while course year and thereby predicting the students' overall performance at the end of the semester examination. In this genetic process when we made changes in single parent Characteristic we are able to find important ratios. This way of using Real Genetic Algorithm can be applied to any areas to find the most influencing ratios or parameters to predict the performance effectively. This research outlines the best features to develop a student performance model to analyze and can predict individual's performance efficiently. It will also act as a Self-assessment tool for the students to know their position and area where to get improve. Thus we will provide the successful tool to predict the student's performance using genetic algorithm.

REFERENCE

[1] **P** ramya, M, Mahesh kumar,"Student Performance Analysis using Educational Data Mining", International Journal of Computer science and Information Security (IJCSIS), ISSN NO 1947-5500, vol-14, pp:69-76, 2016.

[2] Chew li sa, Emmy Dehlana Hossain, "Student Performance Analysis System (SPAS)", JANUARY 2015.

[3] Saddam Khan, Sunny Gupta,"A Study of Data Mining Techniques and Genetic Algorithm in education sector", International Journal of Computer Science and Mobile Computing,ISSN 2320-088X, vol 4, march, 2015, pg 681-683.

[4] Somnath mazi,"genetic algorithm with different crossover for new analysis model of students performance"international journal of advanced research and development, ISSN: 2455-4030, march-2018, volume:3,pp:67-72.

[5] A. Martin, V. Prasanna Venkatesan et al, "To find the most impact financial features for bankruptcy model using genetic algorithm", International Conference on Advances in Engineering and Technology, (ICAET-2011), May 27-28, 2011.

[6] Ramjeet Singh Yadav, "Modeling Academic Performance Evaluation Using Soft Computing Techniques: A Fuzzy Logic Approach", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 2 Feb 2011.

[7] O.K Chaudhari, P.G khot, "Soft computing model for academic performance of teachers using fuzzy lo logic", British journal of applied science and technology, 2(2):213-226, 2012.

[8] V.O. Oladokun, "Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course", The Pacific Journal of Science and Technology, Volume9. Number 1. May-June 2008 (Springer).

[9] Osman Taylan, Bahattin Karagozog, "An adaptive neuro-fuzzy model for prediction of student's academic performance computers & Industrial Engineering", 57 (2009) 732-741.

[10] Emmanuel N. Ogor "Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques" published on Fourth Congress of Electronics, Robotics and Automotive Mechanics, 2009