International Journal of Early Childhood Special Education (INT-JECSE)
DOI: 10.48047/INTJECSE/V14I7.407 ISSN: 1308-5581 VOL 14, ISSUE 07 2022

Individual Micro Analysis and Prediction of Rural School Students Performance using Efficient Data Mining Techniques

Suresh Solomon. G[1], Nancy Jasmine Goldena[2],
[1]Research Scholar, Manonmaniam Sundaranar University, Tirunelveli,Tamil Nadu
Email: sureshsolomong@rediffmail.com
[2]Associate Professor, Department of Computer Applications, Sarah Tucker College, Tirunelveli, Tamil Nadu

Abstract
The student performance analysis and prediction is helpful for thestudent'scommunity for the entire education domain components. The process of analysing the educational domain components using data mining techniques is an interesting way with lot of stepping stones.The researches on educational data mining play the vital role in thefield of education towards the development of the country. This paper presents the individual student microeducation performance analysis. It also performs the prediction of rural school student's performance using efficient data mining techniques. The individual student micro education performance analysis contains 4 phases they are, deep phase, surface phase, evolution phase and strategy phase.The objective of this research article focuses on the implementing of data mining approaches to handle the 4-phases individual student data analysis for the rural educational data collections towards the prediction of performance in various aspects of education environment. In near future this paper will be extended with the implementation of soft computing techniques using data mining approaches.

Keywords—Data Mining, Rural, Prediction, Education, Performance

I. INTRODUCTION

Educational data mining (EDM) is a research field concerned with the application of data mining, machine learning and statistics to information generated from educational settings (e.g., universities and intelligent tutoring systems)[1,2].

At a high level, the field seeks to develop and improve methods for exploring this data, which often has multiple levels of meaningful hierarchy, in order to discover new insights about how people learn in the context of such settings [4]. In doing so, EDM has contributed to theories of learning investigated by researchers in educational psychology and the learning sciences. The field is closely tied to that of learning analytics, and the two have been compared and contrasted [3].

Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students [5], and the settings which they learn as in figure-1.

Whether educational data is taken from students' use of interactive learning environments, computer-supported collaborative learning, or administrative data from schools and universities, it often has multiple levels of meaningful hierarchy[6], which often need to be determined by properties of the data itself, rather than in advance. Issues of time, sequence, and context also play important roles in the study of educational data [7].

Basic values for Student performance:

1. Personal Values –These values make a person good for himself. Examples being ambition, cleanliness, discipline etc.

2. Family Values – Family as a social institution is based on certain universally defined value system which are nurtured and cultivated within a family system.
3. Social Values – These values are good for the society and form the basis of the relationship of an individual with other people in society. Examples are courtesy, charity, civic duty etc.
4. Moral Values – These values constitute attitude and behaviour that a society consider essential for co-existence. Example is fairness, justice, human dignity etc.
5. Spiritual Values – it refers to the process of reflecting on non-material dimensions of life and acquiring insights into personal experiences.

DOI: 10.48047/INTJECSE/V14I7.407 ISSN: 1308-5581 VOL 14, ISSUE 07 2022

6. Cultural Values – It gives importance to preserve cultural practices, ceremonies, traditions etc. which might be threatened by the materialistic culture of modern times.

Prediction:

A prediction, or forecast, is a statement about a future event or data. They are often, but not always, based upon experience or knowledge. There is no universal agreement about the exact difference from "estimation"; different authors and disciplines ascribe different connotations.

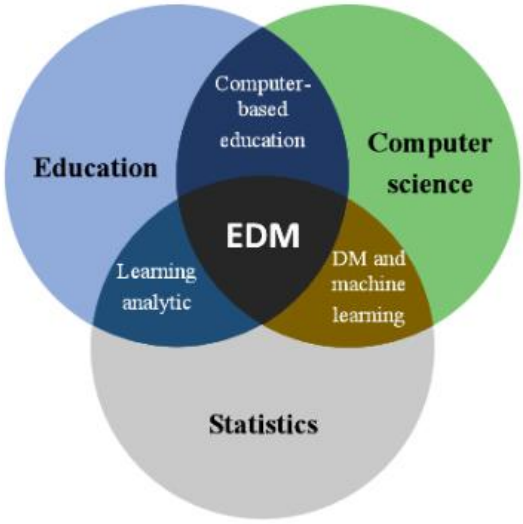

Fig-1: Education Data mining architecture

II. LITERATURE REVIEW

The literature review for this research article is represented in the following table-1.

Table-1: Literature review for Educational Data Mining

Sl.No	Author	Paper title	Publication	Concept	Findings
1	Khribi, M. K., Jemni, M., Nasraoui, O.	Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval	IEEE International Conference on Advanced Learning Technologies,	Investigate different Information Technology (IT) tools for E-learning development. Firstly, the issue of information retrieval was identified and the various web usage mining techniques were presented. Then, the most important IT tools used in information retrieval development were discussed. Finally, some strengths and weaknesses of described tools	This development has to be based on various IT tools. Traditional IT tools for E-learning focused on acquisition, analysis and discovering new knowledge from Internet, social media, and big data requiring more advanced IT tools.

				were demonstrated.	
2	A. S. Arunachalam and T. Velmurugan	A Survey on Educational Data Mining Tools and Techniques	International journal of data mining technology and applications-2016	To improve the understanding of the educational data mining tools and techniques.	The evolution and implementation structure for different versions of data mining strategies. The data mining implementation were discussed.
3	P. Bachhal, S. Ahuja and S. Gargrish	Educational Data Mining: A Review	ICMAI-Mar-2021	This research article review data are extracted for analysis. Use of data derived from computational modeling for epidemiology and the application analysis for the various technical implementations.	The complicated issue of data-driven science's limitations and connection to automated data retrievals are not explained properly.
4	P. Shabrina, B. Mostafavi, S. D. Tithi, M. Chi, and T. Barnes	Learning problem decomposition-recomposition with data-driven chunky parsons problem within an intelligent logic tutor	July 2023. International Educational Data Mining Society	The data driven graph-mining-based method to decompose historical student solutions of logic-proof problems into Chunks discusses learning problem decomposition	Gaps in recomposition with data-driven chunky parsons problem within an intelligent logic tutor.
5	Mr. Pradeep Nayak , Mohammed Sufiyan , Mohan Raju. V , Monisha. N. S. , Moollya Gautami Bhaskar	Review Paper on Educational Data Mining	International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)-Jan 2023	It contains the educational data mining existing research article results with different domain result comparisons.	Description of result variation is not described elaborately but the comparisons are widely explained.

III.METHODOLOGY

The proposed methodology deals with 4-phase architecture towards the micro analysis and prediction of student's performance in rural education area.

a. The problem and solution structure:

The mathematical problem formulation for education process analysis-mini is as follows,

i. EPA-Education Process Analysis Score (micro); 0<=EPA<=1 depends on the following fundamental components.

EPA-micro(X) =Pv(X) +Bs(X) +FSr(X) +Se(X) where X=rural student sample space,

Pv(X) =Personal Values Analysis impact on rural student educational performance computation based on Genetic algorithm cross over operation with data mining techniques such that 0<=Pv(X) <=0.25.
Bs(X)=Behavioural & psychological state Analysis application on rural student educational performance computation based on Sequential patterns with data mining techniques such that 0<=Bs(X)<=0.25.
FSr(X) =Family and Society relationship analysis implementation on rural student educational performance computation by classification clustering approach of data mining techniques such that 0<=FSr(X) <=0.25.
Se(X) =School environment analysis effect on rural student educational performance computation based on clustering approach of data mining techniques such that 0<=Se(X) <=0.25.

ii.Student Performance Prediction:

The student performance prediction is based on the education performance analysis score-micro with the following computation approach for each student i.

Based on the EPA-micro score value, the following interpretations are available for prediction as in table-2.

Table-2: Students Performance Prediction interpretation

Sl.No	Criterion	Prediction interpretation
1	Pv(X)>=0.2	Student Personal value is good
2	0.2> Pv (X)>=0.1	Student Personal value is average
3	Pv (X)<0.1	Student Personal value is not ok
4	Bs(X)>=0.2	Student behaviour state is good
5	0.2> Bs (X)>=0.1	Student behaviour state is average
6	Bs (X)<0.1	Student behaviour state is not ok
7	FSr(X)>=0.2	Family/Society relation is good
8	0.2> FSr (X)>=0.1	Family/Society relation is average
9	FSr (X)<0.1	Family/Society relation is not ok
10	Se(X)>=0.2	School environment is good
11	0.2>Se (X)>=0.1	School environment is average
12	Se (X)<0.1	School environment is not ok
13	EPA(X)-micro>=0.5	Student performance progress is good

14	EPA(X)-micro<0.5	Poor student performance in future.

The student performance prediction will be completed through data mining linear threshold based approach.

b. Proposed methodology architecture design:

The proposed methodology architectural design is as follows in figure-2.

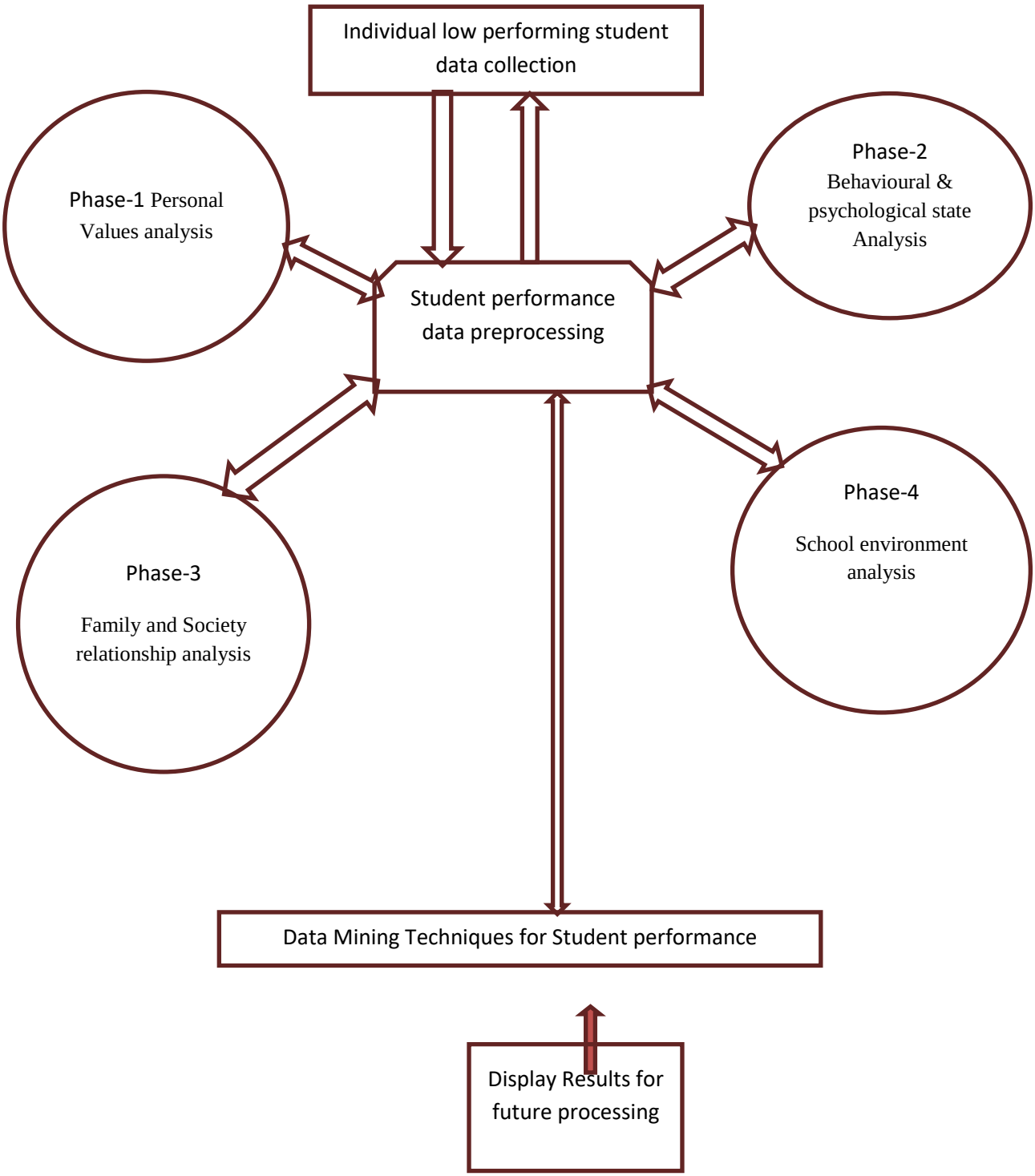

Fig-2:Proposed methodology architectural design for education performance prediction

The proposed methodology consists of 4-phases approach for student performance analysis and prediction such as personnel value analysis, behaviour state analysis, family& society relation analysis and school environment analysis attributes. The data mining techniques are used for both analysis and prediction for the effectiveness in education data domain analysis and handling. Each individual component incorporates data mining techniques for the entire computation.

IV. IMPLEMENTATION

The implementation consists of 5 stages, the initial stage focuses on personnel value analysis on rural student educational performance computation based on fuzzy membershipapproach of data mining technique followed by the behaviour state analysis on rural student educational performance computation through classification and clustering analysis of data mining technique. The third stage concentrates on the family and society relationship analysis implementation on rural student educational performance computation by pattern analysis approach of data mining techniques. The fourth stage concentrates on the school environment analysis implementation on rural student educational performance computation by genetic algorithm cross over approach of data mining techniques. The final stage performs the prediction of student performance analysis using regression technique.

a. Fuzzy membership association based Personnel value analysis-Pv(X)

The membership values associated with each factors insist the methodology for focusing on the individual student personnel value assessment, The following Fuzzy membership assignment as in table-3 shows the personnel value components impact,

Table-3: Fuzzy membership assignment for personnel values

Sl.No	**Personnel value component**	**Fuzzy membership assignment for positive response-Fm(X)**
1	Motivation	0.1
2	Physical disability	0.1
3	With/Without family	0.1
4	Comparisons	0.1
5	School distance	0.1
6	Available notes/books etc.	0.1
7	Teacher rapport	0.1
8	Mutual friends	0.1
9	Emotional intelligence	0.1
10	Discipline	0.1
Total		1.0

The personnel value $Pv(X)=Fm(X)/4$ such that $0<=Pv(X)<=0.25$

b. Data mining Classification and clustering based behaviour state computation- Bs(X)

The behaviour state of the student observed in the school based on 10 or more events over a period of 45 days with frequent monitoring produces the following classification exhibits along with the state values as in table-4.

Table-4: Student Behaviour type table

Sl.No	Behaviour type	Observation count
1	Overt	B1
2	Covert	B2
3	Conscious	B3
4	Unconscious	B4
5	Rational	B5

6	Irrational	B6
7	Voluntary	B7
8	Involuntary	B8
9	Target	B9
10	Exceptional	B10

There are 3 main clusters in student performance behaviour state assessment such that

Cluster1=C1=b1+b3+b5+b7
Cluster2==C2=b2+b4+b6+b8
Cluster3=C3=b9+b10

If C1>5 then
Cluster1 activated
Else if C2>5 then
Cluster2 activated
Else if C3>5 then
Cluster3 activated
Else
Combination cluster activated

The Data mining Classification and clustering based behaviour state computation - Bs(X) are shown in the following table-5.

Table-5: Behaviour state computation table

Sl.No	Cluster value	Cluster Label	Behaviour state computation - Bs(X)
1	C1>7	Cluster1	0.25
2	7>C1>=5	Cluster1	0.2
3	0<C1<5	Cluster1	0.1
4	C2>7	Cluster2	0.05
5	7>C2>=5	Cluster2	0.1
6	0<C2<5	Cluster2	0.15
7	C3>=5	Cluster3	0.05
8	0<C3<5	Cluster3	0.1

c. Data mining Pattern analysisbased Family & society value analysis-FSr(X)

The student education performance analysis using family and society values depends on two factors ,they are the society value Sv(X) and the family support pattern value Pn(X) such that, the Family state relation value FSr(X) is determine by the following formula,

$$FSr(X) = (Sv(X) + Pn(X))/2$$

The societal relationship states are as follow in table-6

Table-6: Social relationship values-Sv(X) table

Sl.No	**Social Relation type**	**Social relationship values-Sv(X)**
1	Exchange	0.25
2	Competition	0.2
3	Conflict	0.15

4	Cooperation	0.2
5	Accommodation	0.2
6	Isolation	0.1

The family support analysis incorporates the pattern matching schema for the student performance analysis. The family support is stored as different patterns for isomorph the target student pattern for the associated computation. The patterns are represented in the following figure-3.

Pattern-1	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Any Relation support	Internal guardian support	External guardian support

Pattern-2	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Any Relation support	Internal guardian support
	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Any Relation support	External guardian support
	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Internal guardian support	External guardian support
	Parent-Father/Mother/both support	Any Relation support	Internal guardian support	External guardian support
	Sibling-Brother/Sister/both support	Any Relation support	Internal guardian support	External guardian support

Pattern-3	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Any Relation support
	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	Internal guardian support
	Parent-Father/Mother/both support	Sibling-Brother/Sister/both support	External guardian support
	Parent-Father/Mother/both support	Any Relation support	Internal guardian support
	Parent-Father/Mother/both support	Any Relation support	External guardian support
	Parent-Father/Mother/both support	Internal guardian support	External guardian support
	Sibling-Brother/Sister/both support	Any Relation support	Internal guardian support
	Sibling-Brother/Sister/both support	Any Relation support	External guardian support
	Sibling-	Internal guardian	External

	Brother/Sister/both support	support	guardian support
	Any Relation support	Internal guardian support	External guardian support

Pattern-4	Parent-Father/Mother/both support	Sibling-Brother/Sister/both Support
	Parent-Father/Mother/both support	Any Relation support
	Parent-Father/Mother/both support	Internal guardian support
	Parent-Father/Mother/both support	External guardian support
	Sibling-Brother/Sister/both Support	Any Relation support
	Sibling-Brother/Sister/both Support	Internal guardian support
	Sibling-Brother/Sister/both Support	External guardian support
	Any Relation support	Internal guardian support
	Any Relation support	External guardian support
	Internal guardian support	External guardian support

Pattern-5	Parent-Father/Mother/both support Sibling-Brother/Sister/both Support Any Relation support Internal guardian support External guardian support

Pattern-6	No-Support

Fig-3:Stored patterns for target student analysis.

DOI: 10.48047/INTJECSE/V14I7.407 ISSN: 1308-5581 VOL 14, ISSUE 07 2022

The family support for the rural student education performance analysis corresponding pattern values are represented in table-7.

Table-7: Family support for performance analysis score-Pn(X)

Sl.No	**Family support Pattern type**	**Pattern value Pn(X)**
1	Pattern-1	0.25
2	Pattern-2	0.2
3	Pattern-3	0.15
4	Pattern-4	0.1
5	Pattern-5	0.05
6	Pattern-6	0.0

d.Genetic algorithmic crossover based school environment value analysis Se(X):

Consider total number of students in a class=N
Number of students failed in initial assessment test=m
Number of groups to be formed in the class=G(X) such that G(X) <=m; most probably G(X) =m, but in exceptional cases such that more number of failures > (N/2)
Number of candidates in each group=3; exceptional cases=2.

Select the toppers with higher Pv(X), Bs(X), and FSr(X) and perform the crossover with the low performing student mi based on the total group member's|N/m|

If |N/m|=4 then
It flows on crossover-1;
Else if |N/m|=3 then
It flows on crossover-2;
Else if |N/m|=2 then
It flows on crossover-3;
Else No group

Classgroup-crossover-1

$\mathbf{m_i}$	Pv(X)>0.2
Bs(X)>0.2	FSr(X)>0.2

Classgroup-crossover-2

$\mathbf{m_i}$	
Bs(X)>0.2	Pv(X)>0.2

Classgroup-crossover-2

$\mathbf{m_i}$	
Pv(X)>0.2	FSr(X)>0.2

Classgroup-crossover-2

$\mathbf{m_i}$	
Bs(X)>0.2	FSr(X)>0.2

Classgroup-crossover-3

m_i	Pv(X)>0.2

Classgroup-crossover-3

m_i	Bs(X)>0.2

Classgroup-crossover-3

m_i	FSr(X)>0.2

The class group cross over value Cc(X) is defined as follows in table-8

Table-8: Cc(X) computation

Sl.No	Class group cross over type	Class group cross over value Cc(X)
1	Classgroup-crossover-1	0.15
2	Classgroup-crossover-2	0.12
3	Classgroup-crossover-3	0.09

The school environment component support factor-Sc(X) is defined in table-9

Table-9: Sc(X) computation

Sl.No	School environment component	Component support factor
1	Management-M	0.02
2	Principal-P	0.02
3	Class Teacher-C	0.02
4	Other Teachers-OT	0.02
5	Other Friends-OF	0.02
Total Sc(X)		0.1

School environment component support factor-Sc(X) =M(X) +P(X) +C(X) +OT(X) +OF(X)

The School Environment Support Analysis Se(X) =Cc(X) +Sc(X) such that 0<=Se(X) <=0.25

Student performance prediction

The rural students education performance analysis score EPA-micro(X) is computed as follows,

EPA-micro(X) =Pv(X) +Bs(X) +FSr(X) +Se(X) where X=rural student sample space, the prediction score interpretation is as follows as in table-10,

Table-10: Performance prediction interpretation

Sl.No	Prediction score	Prediction Interpretation
1	EPA(X)-micro>=0.5	Student performance progress is good
2	EPA(X)-micro<0.5	Poor student performance in future.

V. RESULTS AND DISCUSSION

The data collection from the results of 12rural educational institution in and around Tenkasi and Tuticorin district,Tamilnadu, India are collected through proper monitoring support from the respective schools as in table-11,

DOI: 10.48047/INTJECSE/V14I7.407 ISSN: 1308-5581 VOL 14, ISSUE 07 2022

Table-11: Educational institution data collection

School /student data of Educational institution type	**Count**
Private	3
Aided	4
Government	5

Thetotal of 4 low performing students is taken into consideration for the implementation of the proposed methodology computation.

Government school students count=4

Aided school students count=0

Private school students count=0

Student performance prediction score computation consists of 4 sets of sub-computations such as,
1. Personnel value analysis-Pv(X)
2, Behaviour state computation - Bs(X)
3. Family & society value analysis-FSr(X)
4. School environment vale analysis Se(X):

1. Personnel value analysis-Pv(X)
The personnel value Pv(X) =Fm(X)/4 such that 0<=Pv(X) <=0.25; the obtained results are shown in the following table-12.

Table-12: Pv(X) Computation

Sl.No	Student-ID	Fuzzy membership assignment for positive response-Fm(X)	Personnel value analysis-Pv(X)
1	G1S1	0.4	0.1
2	G1S2	0.5	0.125
3	G2S1	0.5	0.125
4	G2S2	0.6	0.15

2, Behaviour state computation - Bs(X)
The behaviour state analysis value Bs(X) based on classification and clustering technique implementations is represented in the following table-13

Table-13: Bs(X) Computation

Sl.No	Student-ID	Cluster value	Cluster Label	Behaviour state computation - Bs(X)
1	G1S1	C1>7	Cluster1	0.25
2	G1S2	C2>7	Cluster2	0.05
3	G2S1	0<C2<5	Cluster2	0.15
4	G2S2	0<C2<5	Cluster2	0.15

DOI: 10.48047/INTJECSE/V14I7.407 ISSN: 1308-5581 VOL 14, ISSUE 07 2022

3. Family & society value analysis-FSr(X)

The student education performance analysis using family and society values depends on two factors ,they are the society value Sv(X) and the family support pattern value Pn(X) such that, the Family state relation value FSr(X) is determine by the following formula,

$FSr(X) = (Sv(X) + Pn(X))/2$

The Family & society value analysis-FSr(X) based pattern analysis approach is represented in the following table-14

Table-14: FSr(X) Computation

Sl.No	Student-ID	Society value Sv(X)	Pattern value Pn(X)	Family & society value analysis-FSr(X)
1	G1S1	0.15	0.2	0.175
2	G1S2	0.1	0.05	0.075
3	G2S1	0.2	0.15	0.175
4	G2S2	0.2	0.2	0.2

4. School environment value analysis Se(X):

The School Environment Support Analysis Se(X) =Cc(X) +Sc(X)

School environment component support factor-Sc(X) =M(X) +P(X) +C(X) +OT(X) +OF(X) computation is represented in table-15.

Table-15: Se(X) Computation

Sl.No	**Student-ID**	**Cc(X)**	**M(X)**	**P(X)**	**C(X)**	**OT(X)**	**OF(X)**	**Sc(X)**	**Se(X)**
1	G1S1	0.15	0	0	0.02	0.02	0.02	0.06	0.21
2	G1S2	0.12	0	0	0	0	0.02	0.02	0.14
3	G2S1	0.12	0.02	0.02	0.02	0.02	0.02	0.1	0.22
4	G2S2	0.12	0	0	0.02	0.02	0.02	0.04	0.18

The student performance microanalysis prediction score value computation is represented in the following table-16.

Table-16: Student performance micro prediction score computation

No	**Student ID**	**Personnel value analysis-Pv(X)**	**Behaviour state computation - Bs(X)**	**Family & society value analysis-FSr(X)**	**School environment value analysis-Se(X)**	**Students Education Performance Microanalysis score EPA-micro(X)=Pv(X)+Bs(X)+FSr(X)+Se(X)**

1	G1S1	0.1	0.25	0.175	0.21	0.735
2	**G1S2**	**0.125**	**0.05**	**0.075**	**0.14**	**0.39**
3	G2S1	0.125	0.15	0.175	0.22	0.67
4	G2S2	0.15	0.15	0.2	0.18	0.68

The performance analysis and prediction of student education performance shows that 3out of 4 students will perform well in their current education state. The suggestive feedback and enforcement towards the student G1S2 through proper guidance and counselling, hostel accommodation,the specific area improvement possibility in the order of personnel value, school environment modification by upgraded crossover and support, encouraging the family engagement and commitment, internal and external guardians arrangements . The final results are as follows,

2	**G1S2**	**0.15**	**0.1**	**0.1**	**0.18**	**0.53**

After the half yearly exams the performance assessment also improved with all pass categories as a successful candidate (average).
While comparing the proposed methodology results with the quarterly and half yearly performance analysis the final results shows that the proposed methodology prediction results are accurate with 4 out of 4 students.

The proposed methodology produces 100% of success when implemented directly twice with an improvement test scenario in the real time school environment in and around the rural schools of Tenkasi and Tuticorin districts.

The following figure-4 shows the final results for the proposed methodology implementation in real-time environment.

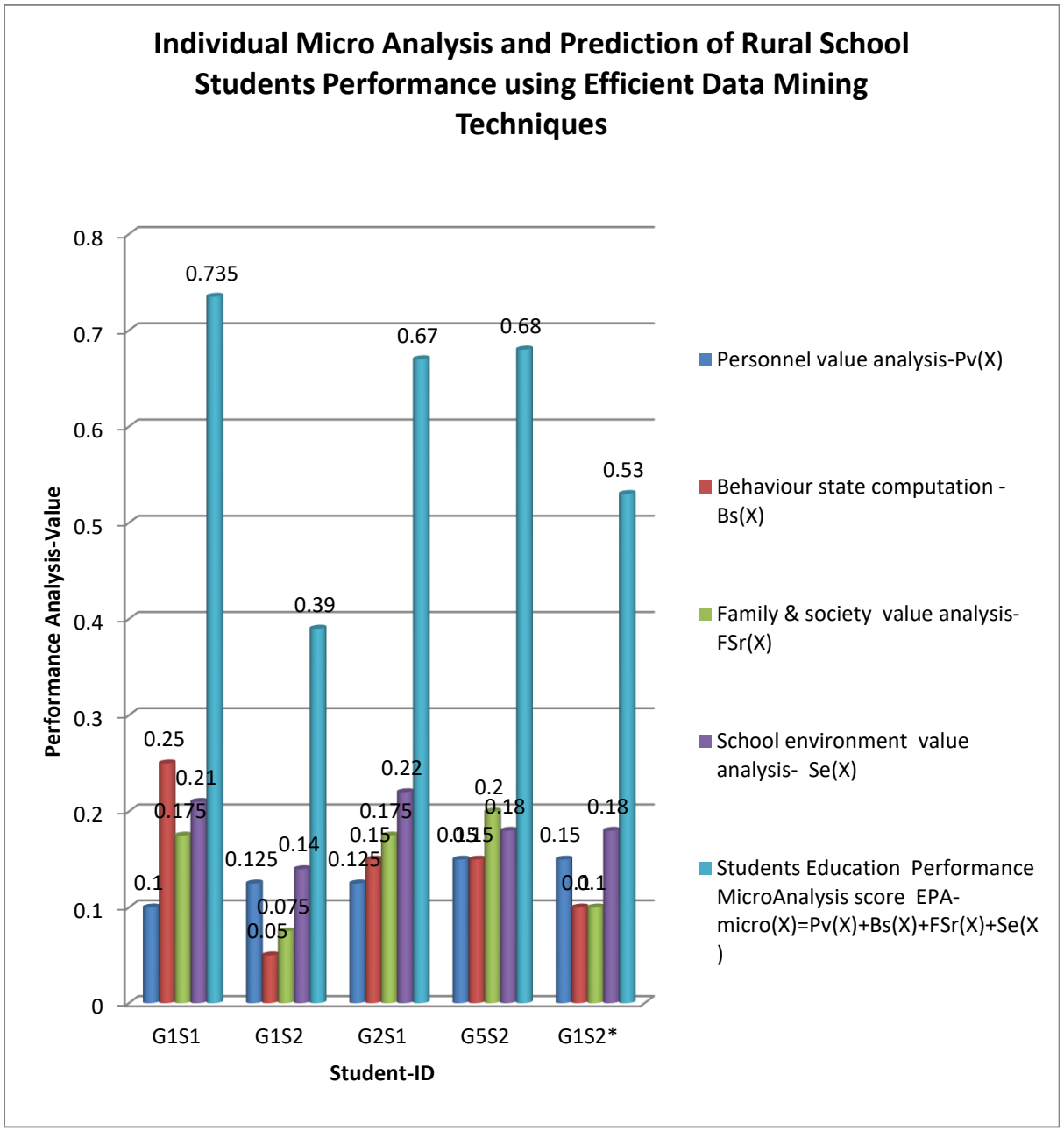

Fig-4: Student performance prediction score computation result

The comparison of the proposed methodology with the existing approaches with a sample space of 45 students along with the computation of accuracy,recall,precision and F1 score is as follows in table-17,

Table-17: Proposed methodology comparison using accuracy,precision,recall and F1 score

Sl.No	Methodology	Accuracy	Precision= tp/(tp+fp)	Recall= tp/(tp+fn)	F1 score=2*tp/(2*tp+fp+fn)
1	Node detection method	52%	0.53	0.54	0.56
2	Trend extrapolation	63%	0.64	0.65	0.67
3	Back casting	69%	0.68	0.71	0.69

4	Proposed Individual micro analysis and prediction of rural school students performance using efficient data mining techniques	98%	0.98	0.97	0.96

VI.CONCLUSION

The educational data mining application in the process of analysing each individual student performance and predict their future educational results are useful in the process of implementing the proper guidance and support to modify the progress in the upward direction. The student education data are collected and handled in a proper way with providing the options for injecting or vaccinating the each individual affecting component for the student welfare.

This research article initially collecting and pre-processing the educational data for students' performance, followed by the 4-phase proposed methodology for the impact on individual microanalysis and prediction of rural school student's performance using efficient data mining techniques. The first phase focuses on the personnel value analysis phase; the second phase concentrates on the student behaviour analysis. The third phase leads the analysis for the family and society support in student education performance improvement. The fourth phase looks into the school environment support analysis.

Finally this research article performs the threshold based prediction for the student education performance improvement assessments. The practical implementation of this proposed research methodology produces exact results for 4 out of 4low performing students education performance prediction score with a success rate of 100%.while analysing the class room performance with the proposed methodologywith 45 students, the proposed methodology produces 44 out of 45 students performance with correct prediction results which is 98% of success rate.

In future we will implement machine learning approach based methodology for faster and accurate results in this educational data mining domain.

REFERENCES

[1]L. Tang and H. Liu, "Toward predicting collective behavior via social dimension extraction," IEEE Intelligent Systems, vol. 25, pp. 19–25, 2013.

[2] M. Newman, "Finding community structure in networks using the eigenvectors of matrices," Physical Review E (Statistical, Nonlinear, and Soft Matter Physics), vol. 74, no. 3, 2006. [Online].

[3] L. Tang and H. Liu, "Scalable learning of collective behavior based on sparse social dimensions," in CIKM '09: Proceeding of the 18th ACM conference on Information and knowledge management. New York, NY, USA: ACM, 2015, pp. 1107–1116.

[4] Baker, Ryan. "Data mining for education", oxford, UK: Elsevier. Retrieved 9 February 2014.

[5]. C. Romero, S. Ventura. Educational Data Mining: A Review of the State-of-the-Art. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 40(6), 601-618, 2016.

[6] Baradwaj, B.K. and Pal, S., 2011. Mining Educational Data to Analyze Students' Performance. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.

[7] Ahmed, A.B.E.D. and Elaraby, I.S., 2014. Data Mining: A prediction for Student's Performance Using Classification Method. World Journal of Computer Application and Technology, 2(2), pp.43-47.