An Enhanced Learning with Cognitive Skills and Analysis using Data Mining

D.Magdalene Delighta Angeline [1], Dr.P.Ramasubramanian[2], I.Samuel Peter James[3]

Assistant Professor, Dept. of . CSE., Dr.G.U.Pope College of Engineering, Thoothukudi, Tamilnadu,India [1]

Principal, Annai Vailankani College of Engineering, Kanyakumari, Tamilnadu, India [2]

Research Scholor, Dept. of . CS., Bharathiyar University,Coimbatore, Tamilnadu, India [3]

ABSTRACT: Data Mining is a tool that investigates huge databases in different perspective and generates new information to facilitate understanding the samples extracted with the set of rules. Also, data mining develop relationships among variables stored in large data set. Data mining find out hidden information in the data and affords a better solution for a specific problem. The main objective of this paper is to analyze the data mining algorithms applied in the learners learning environment. The educational sector gives much more importance on learners' results. The teaching methodology, listening ability of the learner and the outcome of the learner is the key element for the success in gaining knowledge in their concerned course. The key elements enlarge knowledge of the learner. This is achieved by scrutinizing the cognitive skills of the learner and also by predicting the performance of the learner using various data mining algorithms. This paper observes the variations among different learner and guidance is given based on the observed skills that impart the quality of education.

KEYWORDS: Blended Learning, Class Association Rule, Cognition, Cognitive Skills, Decision Tree, Naive Bayes

I. INTRODUCTION

Learning plays a major role in every part of human life style. Each human has their own style of learning. Generally, learning is performed by observing the environment, visualizing the content, understanding the concept and applying the concept to the existing problem. From birth till death learning is continually done. Whatever style or method each person adopts for learning, the final outcome is the essential to measure the individual performance. The learning environment inside the classroom by sitting in classes listening to teachers, memorizing assignments and spitting out answers do not make learner to learn much. The learner must talk about the information related to the topic, make a note on the topic, compare the topic with the previous knowledge and implement it in everyday lives [8].

The knowledge opening in an education structure can be linked using data mining methodology. An improved learning can be achieved with newer learner models. The assessment of a learner is to be done providing good prediction that leads to superior learning. Data prediction and description is the fundamental aim of data [2]. The performance prediction in the student database aid to forecast the unidentified values. The knowledge discovery in databases is a technique to extract significant pattern from a large set of data in the database.   Mining in educational environment is called Educational Data Mining. New methods are developed to find out knowledge from educational databases [11],[10], that aids in investigating students trends and behaviors toward education[2].

The prediction of student outcome is guaranteed with new methods with new type of learning structure. Achieving quality objectives is prevented due to the deficiency in required knowledge in higher educational system. This knowledge gap in higher education system can be filled and better quality can be achieved using data mining methods. The learning analytics helps in monitoring and predicting students' learning performance which helps to predict student at risk of failing a course. [9] [15].

To increase the success ratio of students', the student at risk of failing is identified. To implement this several learning analytics models have been developed. The examples of learning analytical methods comprises Purdue University's Course Signals system [4] and the Moodog system being used at the course level at the University of California, Santa Barbara, and at the institutional level at the University of Alabama [9]. The most important techniques of data mining is association rule. It is possibly the most common form of discovering local-pattern in unsupervised learning systems. The technique is expected to be very realistic in applications. The rules like If X then Y where X and

Y can be particular items ,values, words, etc. , or conjunctions of values, items, words, etc., is discovered with association rule . Association rules mining was proposed by Agarwal. Association rule mining is the technique to discover all possible rules that exceed the condition of minimum support and minimum confidence. The mining of rules can be done by computing support and confidence value for all of the rules and it is then compared with the threshold value to prune the rules with low values. Genetic algorithm, SaM algorithm, Apriori, Eclat and FP-Growth algorithms are used to mine frequent item sets. The problem occurs in mining frequent item sets while mining association rules [3]. The   frequent item set plays an important role when trying to find interesting patterns from databases.

II. RELATED WORK

In recent research works the role of Data mining is enormous. Lots of techniques and approaches are used in different applications like weather forecasting, genome data, performance analysis, predicting future outcomes etc. The fore most technique used in data mining is the KDD to find out the hidden patterns in the data. Hipp [14] used this technique in his research paper to extract patterns in the association rule. . Hijazi and Naqvi [13] conducted study to frame out the performance of a student related to family income, education and age. The study was conducted with a sample of 300 student from a group of colleges affiliated to Punjab university of Pakistan. Khan [16] carried out a performance study on 400 students including 200 boys and 200 girls chosen from the senior secondary school of Aligarh Muslim University, Aligarh, India. This study helps to know the success ratio predicting the value of different measures of cognition. The findings are the girls with high socio-economic status had comparatively higher academic achievement in science stream than boys with same situation.

Also study conducted by Baker to find the probability of a student slipped or guessed using machine learning approach. The outcome of this study increases the correctness of the prediction of student slipped or guessed up to 48 percent [5]. Bhardwaj and Pal [6] conducted study with Bayesian classification method on 17 attributes and identified the factors that affect the performance of a student. The factors like family income, parental education, living loaction, family status, medium of teaching highly influence the performance ratio. Brijesh and Saurabh [7] conducted study to identify the students who needed special consideration to reduce the failure ratio. This study conducted with 50 students from VBS Purvanchal University, Jaunpur. According to Moore [18] requiring students to work in groups and discussing their works orally and in writing, various diagnostic tools to analyze data, ands computer -intensive statistical proactive facilitates student learning.

In [11], gave a case study that use students data to analyze their learning behavior to predict the results and to warn students at risk before their final exams. [9] used educational data mining to identify and enhance educational process which can improve their decision making process**.** [11] applied a data mining technique based on association rules to find weak tertiary school students (n= 264) of Singapore for remedial classes.

Han and Kamber [12] explained data mining software that allow the users to analyze data from different dimensions, categorize it and summarize the relationships which are identified during the mining process. Galit [11] gave a case study about the learners learning behavior that helps to predict the results and to warn students at risk before their final exams. Al-Radaideh et al. [1], applied the data mining techniques, particularly classification to help in improving the quality of the educational system by evaluating learners data. Al-Radaideh, et al [1] applied a decision tree model to predict the final grade of students who studied the C++ course in Yarmouk University, Jordan in the year 2005. Three different classification methods namely ID3, C4.5, and the Naive Bayes were used. The outcome of their results indicated that Decision Tree model had better prediction than other models.

III. DATA MINING ALGORITHMS

There are various data mining algorithms such as decision tree, Naïve Bayes, Rule induction, Supervised Learning, Apriori algorithm, Association rule analysis are available. This paper scrutinizes the effect of different algorithm in producing students with best performance and to match their placement patterns with the industry. The algorithms discussed include Apriori mend, Apriori and class association rules.

A. *APRIORI MEND ALGORITHM:*
Apriori Mend algorithm is used to identify the frequent item sets derived association rules. Apriori Mend algorithm mainly composed of five steps:

Step 1:   Generate frequent-2 itemset using perfect hash function in the database specifying the minimum support.
Step 2:   Prune the frequent-2 item set in the database in order to create a compressed database.
Step 3:   Form a group, finding all the frequent item sets, and infer all the association rules.
Step 4:   A candidate implementation procedure must be set by the frequent-2 item set that have a candidate K. The algorithm is used to generate candidates' itemset C_K for K-item collection.
Step 5:   Calculate the support for the item.
Step 6:   Determine whether the degree of support can immediately become a frequent item sets by setting the minimum support.
Step 7:   When L_K is generated after the calculating $|L_K|$ and if $|L_K| < k+1$, then it is the end of the algorithm. Repeat this process until generating the frequent itemset, or else return to step 4.

B. *APRIORI ALGORITHM:*

Table 1 gives the Apriori algorithm. The first pass of the algorithm simply counts item occurrences to determine the large 1-itemsets. A successive pass k contains two phases: Initially, the large item sets L_{k-1} found in the (k-1)[th] pass are used to generate the candidate item sets C_k. Then, the database is scanned and the support of candidates in C_k is counted. It is necessary to determine the candidates in C_k for quick counting that are contained in a given transaction t.

```
L₁= {frequent items};
for (k= 2; Lₖ₋₁ !=∅; k++) do begin
Cₖ= candidates generated from Lₖ₋₁
for each transaction t in database do
The count that are enclosed in t of all candidates  in
Cₖ
  is to be incremented
Lₖ = candidates in Cₖ with min_sup
end
return     ₖ Lₖ;
```

Table 1. Apriori algorithm

C. CLASS ASSOCIATION RULE:

Normal association rule mining does not have any target where CAR finds all possible rules that exist in data, i.e., any item can appear as a consequent or a condition of a rule. A Class Association Rule (CAR) is a special type of Association Rule (AR) that describes an implicative co-occurring relationship between a set of items and a pre-defined class, expressed in the form of an rule.

"IF antecedent (input-attributes) THEN consequent (class)"

Let *T* be a transaction data set consisting of n transactions. Each transaction is also labeled with a class y. Let I be the set of all items in T, Y be the set of all class labels and $I \cap Y = \varnothing$. A class association rule (CAR) is an implication of the form

$$X \to y, \text{ where } X \subseteq I, \text{ and } y \in Y.$$

The key operation is to find all rule items that have support above minsup. A rule item is of the form:

$$(C, y)$$

where C is a set of items from I (i.e., $C \subseteq I$), and $y \in Y$ is a class label. Each rule item basically represents a rule:

$$C \to y,$$

The multiple minimum support idea can also be applied here. The user can specify different minimum supports to different classes, which effectively assign a different minimum support to rules of each class. For example, a data set with two classes, Yes and No is assumed.

- Rules of class "Yes" to have the minimum support of 5%
- Rules of class "No" to have the minimum support of 10%.

IV. BLENDED LEARNING THROUGH COGNITIVE APPROACH

Cognitive approach is a specialized learning method that the learners use in order to learn more effectively. It includes recurrence, organising new language, reviewing meaning, speculating meaning from context, using imagery for memorisation. All of these approaches involve purposeful manipulation of language to improve learning. Cognitive

approach is useful tool in supporting learner with learning problems. It helps learners to make use of the mind (cognition) to solve a problem or complete a task. The use of cognitive approaches can increase the efficiency with which the learner moves toward a learning task. Cognitive style refers to the way a person thinks and processes information. It is very essential for a learner to improve his/her cognitive skills in order to produce good outcome. The learner should possess

- Ability to sustain concentration on a particular object, action, or thought.
- Ability to manipulate objects.
- Ability to visualize images and scenarios.
- Abilities that enable goal-oriented behavior, such as the ability to plan, and execute a goal.
- The ability to withstand distraction, and internal urges.
- The ability to identify and manage one's own emotions for good performance.

   The faculty accomplishes an essential role, bridging the gap between learner and content to be learned. This role needs an understanding of the task to be completed and knowledge of an approach to the task that the faculty can communicate to the learner.

   The enhanced learning is done with the help of various methods such as Project-Based Approach, Interactive Learning Approach, Presentation Learning Approach, Team-Based Work and Creative Empowerment Approach. Each approach identifies the learner with different cognitive ability and the learning is made with the respective identified approach. Identifying the cognitive skills of each learner and providing learning based on the skill is a talented task of a faculty. The measurement of a learners learning outcome is predicted and enhancement is made to the performance of each learner using data mining algorithm.

A. *PROJECT-BASED APPROACH:*

   In project-based method, the learners are grouped into seven teams with four members in each team. All the team members are instructed to select a topic from microprocessor and microcontroller subject. The teams were given one month time to ready with their own topics. Each team used different materials to explain their concepts. The materials used by the teams are chart, newspaper content, information from internet source, cardboard work, real-life examples to explain the microprocessor working along with the instructions and instruction set etc. The cognitive skills observed at the end of the work from each team are Synthesis, Intelligence, Reasoning, Evaluation and Application. The percentage of skills observed from each team is analysed and grouped under a grade. If-then rules were formed as follows:

If skill percentage>=90 then Grade='A'
If skill percentage>=80 then Grade='B'
If skill percentage>=70 then Grade='C'
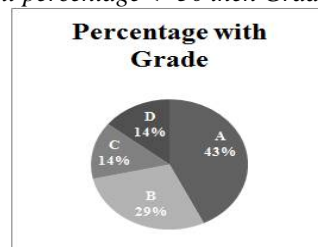*If skill percentage<=50 then Grade='D'*

Fig.1 Cognitive skill percentange for grade

   The grade 'A' has the highest percentage and the maximum team members came out with the good well explanation on their topics. The viewers of this approach learned well and able to grasp the basic concepts.

B. *TEAM-BASED WORK:*

   In team-based work, the learners are grouped and given different topics for each team in software testing subject. The time duration for preparing the topic is 45 minutes and the presentation is 15 minutes to each team. Totally, there were ten teams with two or three members in each team. The learners enjoyed in preparing, discussing among them, working together preparing materials. The presentation of each team was more creative and realistic. The outcomes from the students are

- Increased interest
- Cleared doubts
- Equally shared in doing

The cognitive skills are observed and analysis was made for each team.

C. *CREATIVE EMPOWERMENT APPROACH:*

In creative empowerment approach, the students were grouped into four teams. Each team were provided with a specialized topic on "how a final product undergoes various testing before releasing it". The learners were very much enthusiastic in proceeding with their work. They themselves implemented their own ideas in various levels of testing. The upshot of this method was

- Increased Creative Skills
- Came out with their own ideas
- Different implementation method
- Diverse styles in presenting
- Sound understanding in the technological approach.

D. *INTERACTIVE LEARNING APPROACH:*

In this approach, the learners were made to act, listen, view and ask question immediately. This approach was applied to learners for object oriented and programming and computer networks subject. The findings after this approach was

- Active participation
- Improved listening capability
- Knowing the models effortlessly
- Questioning ability enhanced
- Better communication between learner and facilitator

E. *PRESENTATION LEARNING APPROACH*

In presentation learning approach, the learner as well as the facilitator were made to present content using power point slides or multimedia presentation softwares. The learner were able to study about the working principles and their simulation in a real way. The simulation and presentation for mobile communication and programming paradigm subjects. The learners individual cognitive skills were analysed. The learners was educated to be aware of the concept in detail.

The learner's cognitive processes influence the nature of what is learned. People learn new information more easily when they can relate it to something they already know.

V. SIMULATION RESULTS

The learning in the classroom is made with the blended learning through cognitive approach. Initially, 28 learners were selected for this study. Each learner possesses different abilities, different observing skills and different background. In the learning environment, the learners' were allowed to observe the content of a topic, share their own ideas, clear doubt, and create new application and implementation of the content in practice. The cognitive skills such as intelligence, reasoning, attention, synthesis, knowledge, evaluation and processing speed were observed. The observed cognitive skills from blended learning were analysed using apriori algorithm, apriori mend algorithm and class association rule. The prediction of each learner was made with the above mentioned data mining algorithm.

The learners' observation details helps to get the predicted result that notify about the performance of each learner. The weekly assessment was made at the end of every week and periodical assessment was made three times before the final exam. This help to predict the performance of the learners under three categories: Good, Average and Poor. The rules were generated with the help of data mining techniques that assist in acquiring specific information depending on what the faculty wants to know about students' behaviour patterns. The rules were evaluated based on the confidence measure.

Fig.2 shows the analysis chart of various approaches for different cognitive skills. The creative empowerment approach produces good effect among learners with 92.2%.

	Project-Based Approach	Team-Based Work	Creative Empowerment Approach	Interactive Learning Approach	Presentation Learning Approach
■ Intelligence	90	85	92	88	77
■ Synthesis	87	72	93	75	79
■ Reasoning	89	85	91	89	65
■ Evaluation	91	93	90	92	86
■ Application	86	92	95	83	80
■ Percentage	88.6	85.4	92.2	85.4	77.4
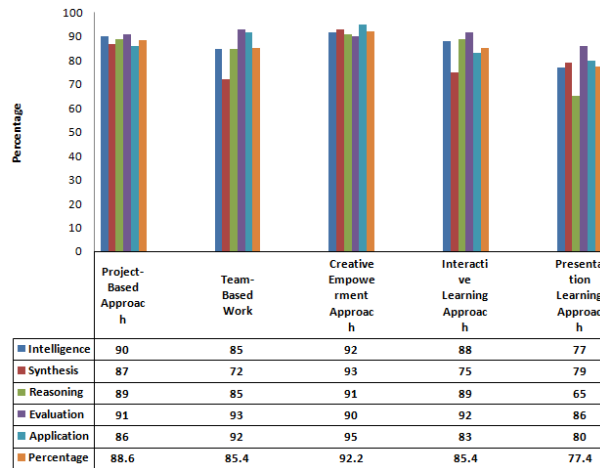
Fig.2 Analysis of blended learning with various approaches

It was found that the cognitive skill named evaluation has 90.4% in fig.4 which was the highest when compared with other skills. The analysis was made with the different subjects for the same group of learner in each approach which modernizes and enhances the learning of the learners. Intelligence encompasses a number of mental abilities such as reasoning, planning and problem-solving. Cognitive theory also highlights student-centered instruction, supportive learning groups, multiple presentations of key thoughts and energetic, investigative learning.
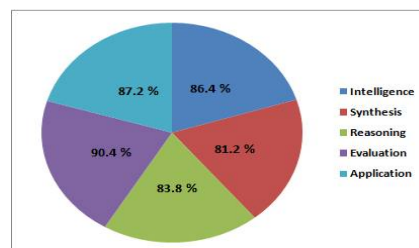

Fig.3 Overall Percentage of cognitive skills when applied with various approaches

Fig.4 shows the number of rules generated in each algorithm along with their confidence. This implementation is conducted to enhance the learners' outcome by analysing the pattern extracted from the rules.

After analysing the cognitive skills, the learners were taught using their own skills.The learners were monitored and evaluation were done based on their performance in the Internal Assessment Tests, Attendance, Degree aspiration, Parental education and discipline. The collected data from their performance were applied in various data mining algorithms and the learners drop out ratio were calculated.

The data mining algorithm discussed in section IV were implemented and rules were generated by which the drop out ratio were calculated. The strength of a rule is mesured by confidence. The maximum number of rules generated were 127 with apriori algorithm and 100% confidence.
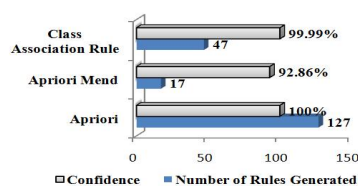

Fig.4 Rules with confidence value

The pattern extracted from the rules smooth the progress of the learner and the outcome of each learner get upgraded in every assessment.

The enhanced learning through cognitive skills along with their individual assessment yields a good outcome. Fig.5 shows the periodical assessment outcome along with the final outcome of the learner. Initially, the outcome of a learner is poor in the PA1. Then it gradually increases after applying the enhanced learning approaches. Finally, the learners' outcome reached to a maximum level stating that the learning approaches with a continuous analysis with data mining algorithms and training works good for each learner.
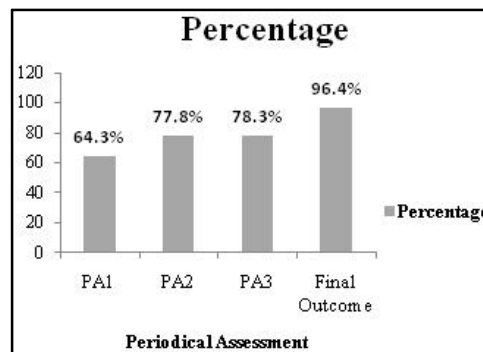

Fig.5 Comparison of final outcome with periodical assessment

From the implementation result, it is identified that the learning with cognitive skills produce good outcome with 96.4%. The drop out ratio of the learner gets reduced with this enhanced learning approach. Similarly, the apriori, apriori mend and class association rule algorithm plays a major role in the performance evaluation of learners'. Among them apriori algorithm evaluate the performance of learner with 127 rules and 100% confidence.

VI. CONCLUSION AND FUTURE WORK

The enhanced learning approaches help every learner to learn something new in their own style without any stress. The learner enjoyed learning along with some activity and they improve their questioning ability to create new things. The learner fall under poor category scored high marks when compared to other learners. The enhanced learning with cognitive approach proven that the outcome of the learner has a dramatic increase and the drop out ratio is reduced to a maximum. In future, this approach can be applied with clustering technique for faster evaluation.

REFERENCES

1. Q.A.AI-Radaideh, E.M. AI-Shawakfa, and M.I. AI-Najjar, "Mining Student Data using Decision Trees", International Arab Conference on Information Technology(ACIT'2006), Yarmouk University, Jordan, 2006.
2. R. Agrawal, Christos Faloutsos & Arun N. Swami, "Efficient Similarity Search in Sequence Databases", Proceedings of the 4th International Conference of Foundations of Data Organization and Algorithms, Pp. 69–84, 1994.
3. R. Agrawal, T. Imielienski, and A. Swami. Mining Association Rules between Sets of Items in Large Databases. Proc. Conf. on Management of Data, 207–216. ACM Press, New York, NY, USA 1993
4. Arnold, K. E, "Signals: Applying Academic Analytics. EDUCAUSE Quarterly33 (1). http://www.educause.edu/EDUCAUSE +Quarterly/EDUCAUSEQuarterlyMagazineVolum/Signals ApplyingAc ademic Analyti/19 9385, 2010
5. Baker, R. S. J. d., A.T. Corbett, and V. Aleven, "More Accurate Student Modeling Through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing." In Proceedings of the 9thInternational Conference on Intelligent Tutoring Systems. Berlin, Heidelberg: Springer-Verlag, 406–415, 2008.
6. B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.
7. Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational Data to Analyze Students' Performance", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.
8. Chickering, A. W., & Erhmann, S. C, "Implementing the seven principles: Technology as lever". AAHE Bulletin, 49(2), 3 – 6, 1996.
9. EDUCAUSE. "Next Generation Learning Challenges: Learner Analytics" Premises.http://www.educause.edu/Resources/Next Generation Learning Chal leng/215028, 2010.

10. Erdogan and Timor, "A data mining application in a student database. Journal of Aeronautic and Space Technologies" Vol. 2, No 2, pp.53-57, July 2005.
11. Galit.et.al ,"Examining online learning processes based on log files analysis: a case study", Research, Refelection and Innovations in Integrating ICT in Education, 2007.
12. J.Han, and M.Kamber, "Data Mining: Concepts and Techniques", 2nd edition. The Morgan Kaufmann Series in Data Management Systems, Jim Gray, Series Editor, 2006.
13. S. T. Hijazi, and R. S. M. M. Naqvi, "Factors affecting student's performance: A Case of Private Colleges", Bangladesh e-Journal of Sociology, Vol. 3, No. 1, 2006.
14. J. Hipp, U. Guntzer & N. Gholamreza, "Algorithm for Association Rule Mining: A General Survey and Comparison", ACM SIGKDD Explorations Newsletter, Vol. 2, No. 1, Pp. 58–64, 2000.
15. Johnson, L., R. Smith,H. Willis, A. Levine, and K.Haywood. The 2011 Horizon Report. Austin, TX: The New Media Consortium. http://net.educause.edu/ir/library/pdf/HR2011.pdf, 2011.
16. Z. N. Khan, "Scholastic achievement of higher secondary students in science stream", Journal of Social Sciences, Vol. 1, No. 2, pp. 84-87, 2005.
17. Y. Ma, B. Liu, C.K. Wong, P.S. Yu, and S.M. Lee, ―Targeting the Right Students Using Data Mining‖, Proceedings of KDD, International Conference on Knowledge discovery and Data Mining, Boston, USA, 2000, pp. 457-464.
18. Moore, D. S,." New Pedagogy and New Content: The Case of Statistics", International Statistical Review , 65, 123-165, 1997.
19. D. Magdalene Delighta Angeline and I. Samuel Peter James (2012), "Association Rule Generation using Apriori Mend Algorithm for Learner's Placement", International Journal of Emerging Sciences, Vol. 2, No. 1, Pp. 78–86.
20. D. Magdalene Delighta Angeline, "Association Rule Generation for Learner Performance Analysis using Apriori Algorithm", The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), Vol. 1, No. 1, March-April 2013, Pp. 12-16.
21. D. Magdalene Delighta Angeline, "Learner's prognostic Analysis Using Class Association Rule", International Journal of Advanced Research in Computer Science and Software Engineering 5(9),September- 2015, pp. 314-318

BIOGRAPHY

D. Magdalene Delighta Angeline is working as a Assistant Professor in the Department of Computer Science and Engineering in Dr.G.U.Pope College of Engineering, Sawyerpuram, Tamilnadu, India. She obtained her Bachelor degree in Information Technology from Anna University, Chennai in the year 2007 and she obtained her Master degree in Computer and Information Technology in Manonmaniam Sundaranar University, Tirunelveli. She is also a research scholar of Bharathiar University, Coimbatore. She has over 9.3 years of Teaching Experience and published nine papers in national conference, five papers in International conferences and also published thirteen papers in various international journals. She also published four books. She is a reviewer for various international journals. Her current area of research includes Image Processing, Neural Networks, and Data Mining.

Prof. Dr. P. Ramasubramanian, presently working as a Principal in Annai Vailankani College of Engineering, Kanyakumari Dist. Tamilnadu, India. He obtained his Bachelor, Master degree in Computer Science and Engineering from Madurai Kamaraj University, Madurai in the year 1989 and 1996 respectively. He also obtained his Ph.D degree from Madurai Kamaraj University in the year 2012. He has more than 24 years of Teaching Experience and authored 16 books and 39 research papers in International, National Journals and Conferences. His current area of research includes Data Mining, Data Ware housing, Neural Networks, Fuzzy and Rough Set logic. Prof. Dr. P.Ramasubramanian has been chosen for inclusion in MARQUIS Who's Who in Science and Engineering, 11th Edition, for the year 2011 – 2012. He is a member of various professional societies like Indian Society for Technical Education (ISTE), International Association of Engineers, Computer Science Teachers Association, International association of Computer Science and Information Technology, Multi-Criteria Decision Making and Fellow in Institution of Engineers (India). He is a reviewer and editor for various international conferences and journals.

I. Samuel Peter James is a research scholar of Bharathiar University, Coimbatore, Tamilnadu, India. He obtained his Bachelor degree in Computer Science and Engineering from Anna University, Chennai in the year 2009 and he obtained his Master degree in Computer Science and Engineering and also M.B.A. in Manonmaniam Sundaranar University, Tirunelveli in the year 2012. He has over 5 years of Teaching Experience and published one paper in national conference, five papers in International conferences and also published twelve papers in various international journals. He also published four books. He is a reviewer for various international journals. His current area of research includes Data Mining, Image Processing and Neural Networks.