

# Hybrid Agent Based Educational Data Mining Model for Student Performance Improvement

Falguni Ranadive, Akil Z. Surti

**Abstract**—After the successful applications in the business domain, the research interest is now diverting towards applying data mining techniques in the education sector. In this paper, we discuss how a hybrid agent based data mining model can help to improve an education system by enabling better and effective teacher-student interaction. The extra information can help the teachers to manage their classes better and to provide proactive feedback to the students. This paper proposes the methodology to predict students' performance based on their psychological characteristics, internal and final examination results. An intelligent system design is required to analyze varied personality and intellect dimensions of both students and well as teacher in order to provide effective and successful interaction. The educational systems currently face number of issues such as high dropout rates, identifying students in need, personalization of training and predicting the quality of student interactions. Agent Based System involving Educational Data mining and Psychological Testing provides a set of agents, which can help the educational system to overcome these issues.

**Index Terms**— Hybrid Agent Based System, Intelligent System, Educational Data Mining (EDM), Predicting student performance

## I. INTRODUCTION

This paper presents a novel model which combines strengths of Agents as well as Data Mining techniques. Though Data Mining has started its own vicinity in education – EDM, Agent based systems are not very commonly used in educational area. The model is focusing autonomy, openness and goal-oriented characteristics of agent based systems with analyzing capabilities of Data Mining techniques.

Agent based system can be said as the most attractive and prominent and technologies in Computer Science, now a days. Currently, the technologies, methods and theories of Agent and multi-agent system have their contribution to varied domains. Some of them include computer mediated collaboration, computer games, information retrieval, user interface design, smart environments, robotics, electronic commerce, education and training, ubiquitous computers, and social simulation. An agent is a computer system located in some environment, and is capable of autonomous action in this environment to meet its design objectives. Multi-agent systems are systems made up of multiple interacting agents.

**Manuscript received April 11, 2014**

Falguni Ranadive, Parul Institute of Computer Application, Vadodara, India.

Akil Surti, Parul Institute of Engineering & Technology, Vadodara, India.

Educational data mining (EDM) is an emerging interdisciplinary research area. This area deals with the development of methods to explore data originating in an educational context. EDM combines use of computational approaches and educational data in order to analyze and study educational questions. The competitive and dynamic environment and Educational opportunities demands for quality education. Students are exposed to more and better opportunities and obtain high quality education at institutions all over the world, due to education globalization. The educational institutions are brazen out with a ruthless competition among each other in order to try to make students capable enough to successfully complete the educational process while the retention ratio is reduced. Although educational data mining has not achieved greater success as compared to other domains such as e-commerce, some institutes are able to make use of data mining to discover knowledge and improve the learning experience of students as well as increase their profits. With the rising of learning management systems, one can expect if the data mining could integrate with learning management systems, then it would become possible to discover real-time knowledge and provide personalized learning experience to the students.

## II. HYBRID AGENT BASED MODEL

Agent-based models consist of agents that interact within an environment. Agents are either separate computer programs or, more commonly, distinct parts of a program that are used to represent social actors—individual people, organizations such as firms, or bodies. Agents are programmed to react to the computational environment in which they are located.

The environment is a model of the real environment in which the real actors operate. It is the virtual world in which the agents act. It may be an entirely neutral medium with little or no effect on the agents.

A crucial feature of agent-based models is that the agents can interact; they can pass informational messages to each other and act on the basis of what they learn from these messages. The messages may represent information flow.

## III. DATA MINING TECHNIQUES

Data mining, also popularly known as Knowledge Discovery in Database, refers to extracting or “mining” knowledge from large amounts of data. Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making. Various algorithms and techniques such as Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic

## Hybrid Agent Based Educational Data Mining Model for Student Performance Improvement

Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases.

### A. Classification

Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records at large. This approach frequently employs decision tree or neural network-based classification algorithms.

### B. Clustering

Clustering can be said as identification of similar classes of objects. By using clustering techniques we can further identify dense and sparse regions in object space and can discover overall distribution pattern and correlations among data attributes.

### C. Predication

Regression technique can be adapted for predication. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining independent variables are attributes already known and response variables are what we want to predict. Unfortunately, many real-world problems are not simply prediction. Therefore, more complex techniques (e.g., logistic regression, decision trees, or neural nets) may be necessary to forecast future values.

## IV. THE BIG FIVE PERSONALITY TEST

The Big Five personality characters are five broad dimensions of personality, which are used to depict human personality. Five Factor Model (FFM) is the theory in psychology, based on these five personality traits. The Big Five factors of FFM are :

- Openness
- Conscientiousness
- Extraversion
- Agreeableness
- Neuroticism

The big five personality traits are the best accepted and most commonly used scientific measure of personality and have been extensively researched.

## V. METHODOLOGY

Data Mining has been used to study the factors leading students to choose to engage in behaviors which reduce their learning and to understand the factors influencing university student retention.

In the proposed system, Student profiling is done by any one of the psychological methods - Big 5 methods or Psychometric profiling. This method can identify personality traits of each student.

As Data Mining and Psychometric profiling requires expertise, the role of agents comes in picture.

The proposed model has separate agents :

1. Data Preprocessing Agent – for normalization of data using data preprocessing techniques of Data Mining.
2. Personality Identification Agent-for psychometric test

3. EDM Agent – for employment of various techniques of Data Mining, such as classification, clustering and prediction, as well as the final allocation of the clusters identified from the clustering technique.
4. Student Performance Assessment Agent – for evaluation of students performance compared to recent past performance.

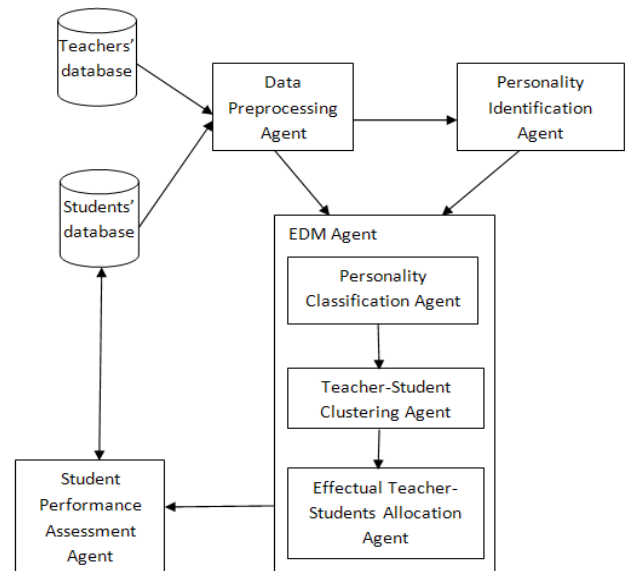


fig.1 Intelligent system model for student's performance Improvement

The model in fig.1 is inspired from a similar example of the social application of data mining is its use in expertise finding systems, whereby descriptors of human expertise are extracted, normalized and classified so as to facilitate the finding of experts, particularly in scientific and technical fields. In this way, data mining can facilitate Institutional Memory.

Expertise extraction for teacher as well as students is done using one of the various tried and tested psychological methods mentioned earlier.

Students are classified using decision tree induction method in three categories based on their performance in various parameters such as, continuous evaluation in class room, class tests, mid-semester exams:

1. 'Slow learner'
2. 'Medium learner' and
3. 'Rapid learner'

The expertise attributes are analyzed using k-means method for clustering and the teacher is allocated for the students lacking in particular expertise for academic as well as personal attribute improvement. Various visualization methods such as, bar graph, pie graph and line chart are used for output of analysis.

## VI. FUTURE SCOPE

The proposed model can be extended to accommodate the improvement agent by considering feedback of evaluation of the allocation agent and re-arranging the group allocation.

REFERENCES

- [1] Alaa el-Halees (2009) *Mining Students Data to Analyze e-Learning Behavior: A Case Study*
- [2] -----
- [3] Connolly T., C. Begg and A. Strachan (1999) *Database Systems: A Practical Approach to Design, Implementation, and Management (3rd Ed.)*. Harlow: Addison-Wesley.687
- [4] Erdogan and Timor (2005) *A data mining application in a student database. Journal of Aeronautic and Space Technologies* July 2005 Volume 2 Number 2 (53-57)
- [5] Galit.et.al (2007)*Examining online learning processes based on log files analysis: a case study Research, Refelection and Innovations in Integrating ICT in Education*
- [6] Henrik (2001) *Clustering as a Data Mining Method in a Web-based System for Thoracic Surgery: 2001*
- [7] Han,J. and Kamber, M., (2006) "*Data Mining: Concepts and Techniques*", 2nd edition. The Morgan Kaufmann Series in Data Management Systems, Jim Gray, Series Editor
- [8] Kifaya(2009) *Mining student evaluation using associative classification and clustering* Communications of the IBIMA vol. 11 ISSN 1943-7765
- [9] ZhaoHui. Maclennan.J, (2005). *Data Mining with SQL Server 2005* Wihely Publishing, Inc.
- [10] Shaeela Ayesha, Tasleem Mustafa, Ahsan Raza Sattar, M.Inayat Khan (2010) "*Data Mining Model for Higher Education System*", European Journal of Scientific Research, ISSN 1450-216X Vol.43 No.1 (2010), pp.24-29
- [11] Cristóbal Romero, Sebastián Ventura, (2010), "*Educational Data Mining : A Review of theState-of-the-Art*", IEEE Transactions on Systems, Man and Cybernetics-Part C : Applications and Reviews, Vol. 40, No. 6
- [12] Nigel Gilbert "*Agent-Based Models*", University of Surrey, Guildford, UK
- [13] Brijesh Kumar Baradwaj, Saurabh Pal, (2011), "*Mining Educational Data to Analyze Students' Performance*", IJACSA, Vol. 2, No. 6
- [14] Vikram Jamval, Sridhar Iyer. KR School of Information Technology IIT Bombay, Powai, <http://www.it.iitb.ac.in/~sri/papers/ma-ail01.pdf>
- [15] Adina Magda Florea, "*An Agent-based Collaborative Learning System, Advanced Research in Computers and Communications in Education*", G. Cumming et al. (Eds.), IOS Press
- [16] Sayyad Rasheeduddin, (2013), "*The Theoretical Framework of Data Mining & its Techniques*",IJSSIR,Vol2(1)
- [17] Floyd Garvey, Suresh Sankaranarayanan, (2012), "*Intelligent Agent based Flight Search and Booking System*", IJARAI, Vol. 1, No. 4
- [18] Longbing Cao, Vladimir Gorodetsky, Pericles A. Mitkas, (2009), "*Agent Mining : The Synergy of Agents and Data Mining*", IEEE Intelligent Systems, 1541-1672/09
- [19] "*Agent-Based Models*", Issue 153, By Nigel Gilbert



**Falguni Ranadive**, pursuing Ph.D. from Gujarat Technological University. Her area of interest includes Intelligent Systems, Artificial Intelligence, Data Mining, Image Processing.



**Akil Z Surti**, pursuing Ph.D. from Gujarat Technological University. His area of interest includes Web Technologies, Database Management System, Data Mining.