



RESEARCH ARTICLE

ROLE OF DATA MINING IN EDUCATION SECTOR

Dr. Mohd Maqsood Ali¹

¹Asst. Professor and Head of Marketing Department, Jazan Community College,
Jazan University, Jazan Kingdom of Saudi Arabia

¹ maqsoodphd@gmail.com

Abstract— Universities either public or private and its colleges enroll thousands of students into various courses or programs every year. They collect information from students at the time of admissions and store the same in computers. Understanding the benefits of data is essential from business point of view. Data can be used for classifying and predicting the students' behaviour, performance, dropouts as well as teachers' performance. Therefore, this paper examines the role of data mining in an education sector. In addition, lays emphasis on application of data mining that contribute to offer competitive courses and improve their business.

Key Terms: - Data Mining; Data Mining Process; Data Mining Tasks; Data Mining Techniques; Data Mining Application

I. INTRODUCTION

Data mining is an essential tool, which is used to predict and classify the data collected from the customers. Today, many industries use data mining to make decisions and to prepare marketing strategies towards target segmented customers for the achievement of its objectives. But many universities neglected practicing data mining techniques. Application of data mining in education sector is an emerging trend in the global competitive business. Understanding the data mining terms, tasks, techniques and application are foundation of developing data mining in education sector. Therefore, there is a need to study the role of data mining in education sector.

This paper consists of six sections. Section 1 explains the conceptual framework of data mining and process of data mining is highlighted in section 2. Section 3 presents the benefits of data mining. In section 4 and 5, data mining tasks and techniques are discussed. Section 6 examines the application of data mining in education sector.

II. DATA MINING CONCEPTUAL FRAMEWORK

Many of us know that information is ubiquitous. Data mining is an exploratory data analysis (Olson and Delen, 2008), a data analysis methodology (Tiwari et al., 2013), a process of obtaining knowledge (Hsia et al., 2008), a process and methodology for applying tools and techniques (Berry and Linoff, 2004), an art of extracting information from data (Tuffery, 2011), a process of discovering patterns in data (Witten, 2005), and a task of discovering meaningful data from big data (Rajshree et al., 2010) with the aim of obtaining clear and useful results (Giudici 2003).

According to Pujari (2001) data mining is "the non-trivial process of identifying valid novel, potentially useful, and ultimately understandable patterns in data." Taniar and Rhaya (2002) define data mining as "a process on non-trivial extraction of implicit previously unknown and potential useful information from data in

database.” A defining characteristic of data mining is ‘Big data’ (Weiss, 1998). Luan (2003) defines data mining as “the process of discovering “hidden message,” patterns and knowledge within large amounts of data and of making predictions for outcomes or behaviour.” ‘Pattern’ is a single record that consists of input and output (Kalogirou, 2004).

The goal of data mining is to reduce large data into useful knowledge (Baritchi, 2004), to find useful patterns from large data (Ramageri and Desai, 2010-20110), to identify important trends and previously unknown behaviour pattern (Mark Last, 2008), to find patterns in historical data (Berry and Linoff, 2004), and to uncover previously unknown patterns (Riyazuddin et al., 2013). In brief, the definitions emphasise the importance of viewing data mining as an essential tool to build models by extracting raw data into usable information. Consequently education sector may gain usable information that helps to offer profitable courses to the students’ community.

III. DATA MINING PROCESS

Many authors develop different steps of data mining. According to Giudici (2003) data mining is “the process of selection, exploration, and modelling of large quantities of data to discover regulations or relations that are at first unknown with the aim of obtaining clear and useful results for owner of the database.” Table I presents the data mining process, which were outlined by different authors and also help us to compare these processes with each other.

Table I
Summary of Data Mining Process

| Authors’ Name | No. of Stages | Data Mining Stages |
|-------------------------|---------------|---|
| Hsu, Hui-Huang (2006) | Four | 1 Data Collection 2 Data Pre-processing 3 Data Mining 4 Information interpretation and Visualisation |
| Hsu, Jeffrey (2004) | Six | 1 Selection 2 Pre-Processing 3 Transformation 4 Data Mining 5 Interpretation 6 Evaluation |
| Berry and Linoff (2004) | Eleven | 1 Translate the business problem 2 Select appropriate data 3 Get to know the data 4 Create a model set 5 Fix problems with the data 6 Transform data to being information to the surface 7 Builds models 8 Assess models 9 Deploy models 10 Assess results 11 Begin again |
| Kantardzic (2004) | Five | 1 State the Problem 2 Collect the data 3 Pre-Process the data 4 Estimate the model 5 Interpret model and draw conclusions |
| Giudici (2003) | Four | 1 Strategic 2 Training 3 Creation 4 Migration |

The standard process mostly used by industries is CRISP-DM (Cross-Industry Standard Process Data Mining). Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment are the steps of CRISP-DM Process.

1. Business Understanding

In this step, first determine objectives of business, assess current situation, establish data mining goals and determine project plans keeping in view of business objectives.

2. Data Understanding

After determining business objectives, the next step is data collection and data description. It is important to collect data from various sources and describe it. Then, examine data quality carefully to know problems if any, or to know whether information is left out or not.

3. Data Preparation

In this step, data that was collected is processed through cleaning unwanted data. Cleaned data is transfer for modelling and constructing new data. Then data is kept ready for subsequent analysis. Moreover, formation of data can be done if required and prepared it for analysis.

4. Modelling

There are several data mining techniques such as classification, regression, cluster and association rules that can be applied to test the validity of data. Bayesian, decision tree, logistic regression and support vector machine are the tools of classification. Regression tools are linear regression, robust regression and neural network. Hierarchical and k-means are clustering tools while association tool is 'A Priori'. At this stage, models are evaluated to design the test.

5. Evaluation

Soon after deciding tools and techniques, the results of models are evaluated and analyse the data keeping in view of the objectives of business. The impact of model on business objectives is necessary to understand and review the data before deployment so that the data, which was not considered earlier, can be prepared once again.

6. Deployment

In this stage, reports and graphs are generated and monitored. Furthermore, decisions are made based on reports and graphs of data.

The above steps are practiced by many industries. Universities can also apply CRISP-DM model to achieve its objectives and attain profitability.

IV. DATA MINING BENEFITS

Data mining was used in financial services, banking, retail, healthcare, manufacturing, telecommunication, food industries, engineering, motor industries, forecasting water, road traffic, pharmaceutical, drug discovery, job seekers and insurance sectors but not used in an education sector. The following may be the benefits of data mining in an education sector:

- identifying students' needs and preferences towards course choices, and selection of specialisation
- identifying students' pattern trends
- predicting students' knowledge, grades, and final results
- supporting automatic exploration of data
- constructing students' profiles become easy
- helping management to understand business

V. DATA MINING TASK

There are many tasks of data mining practicing by various industries. Prediction, classification, association and clustering are the most important tasks of data mining.

1. Prediction

Sen et al. (2012) built models to predict secondary education placements-test using sensitive analysis on predicting models (Decision Tree Algorithm, Support Vector Machine, Neural Network and Logistic Regression). They have identified the following important predictors of placement-test: previous test experience, student has a scholarship or not, students' number of sibling and previous years' grade point average (GPA). They indicate that decision tree analysis is the best predictor, followed by support vector machine, and neural network. Logistic regression is the least predictor.

2. Classification

Classifying data into a fixed number of groups (Soman et al., 2006) and using it for categorical variables (Nisbet, 2009) is known as classification. Classification can be classified into two types: Supervised and Unsupervised. When the objects or cases are known in advance is called supervised classification whereas

unsupervised classification means the objects or cases are not known in advance. The following algorithm can be used for classification model (Gorunescu, 2011; Aggarwal et al., 1999).

- Decision/Classification tree
- K-nearest neighbour classifier
- Rule-based methods,
- Statistical analysis, genetic algorithms
- Bayesian classification
- Neural Networks
- Memory-based reasoning
- Support vector machines

In brief, data mining can be applied in classifying students based on academic achievements, knowledge, gender, age, employment (if any), semester-wise grades and less motivated students.

3. Association

Larose (2005) states that association task finds which attributes “go together”. According to Gopalan and Sivaselvan (2009) association rule is “the process of discovering interesting association or relationship among data items.” It summarises the entire data. Priori and GRI algorithm are used in association rules. Romero et al. (2008) indicates that association rule can be applied to discover relationship between the characteristics of the students and helped to find relationship perfectly (Aggarwal et al., 1999). According to Baker (2010) association rule is ‘discovering of if-then rule’ which means if the value of one variable is found, the value of another variable will have specific value. For example, students who select marketing can also select finance as a specialisation. In addition, students who select M.Com course will also select MBA course.

In summary, association rule can be used for opening new colleges, offering new courses and specialisation based on certain rules. Association rules are “derived from patterns in a dataset that correspondent to a particular situation” (Rajamani et al., 1999).

4. Clustering

Clustering is grouping similar objects. Rajshree et al. (2010) defines clustering as a process of grouping a set of physical or abstract object into a class of similar objects. According to Larose (2005) cluster does not classify, estimate or predict the value of target variables but segment the entire data into homogeneous subgroups. Heterogeneous population is classified into number of homogenous subgroups or clusters are referred as clustering (Berry and Linoff, 2004). Furthermore, clustering task is an unsupervised classification. For example, students can be targeted after segmenting heterogeneous students into similar groups.

Furthermore, clustering task in education sector can be based on enrolments, transfer, readmission, course selections, specialisation, age, gender and behaviour of students. Romero et al. (2008) concludes that clustering shows the characteristics of students in each group.

VI. DATA MINING TECHNIQUES

In this section, the most common data mining techniques are discussed to understand the theory without going into details. According to Chen et al. (2005) data mining brings various techniques together to discover pattern and to construct models from database. Ngai et al. (2009) indicates seven types of data mining model: Association, Classification, Clustering, Forecasting, Regression, Sequence Discovery and Visualisation. Huang et al. (2012) concludes that data mining technique is used to gain useful information or interesting knowledge. Perceived usefulness and perceived ease of use are the factors that affect an individual intention to use data mining tools. Figure 1 shows the techniques of data mining.

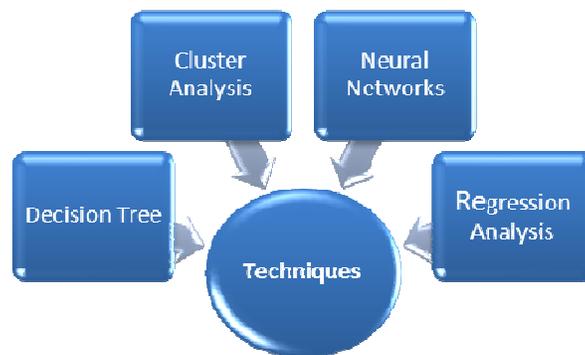


Fig. 1: Data Mining Techniques

1. Neural Networks

It is techniques which can be used for classification of large complex data. It can be used to study course selection by students, student course satisfaction, and specialisation selection. Rokach (2010) state the neural network represents each cluster by a neuron or “prototype”. The input data is also represents by neurons which are connected to the prototype neurons. Each such connection has a weight, which is learned adaptability during learning.

2. Decisions Tree

Decision tree is a data mining technique that can be used for classification and prediction of large data. Decision tree is used for profiling customers. Decision tree is also called rule induction technique (Luan, 2003). According to Rokach (2010) the data, in decision tree, is represented by “a hierarchical tree where each leaf refers to a concept and contains a probability description of the concept.”

Furthermore, Yin et al. (2011) defines decision tree as a “systematic method that uses a tree diagram. It consists of nodes and branches, nodes are connected by branches, time flows from left to right, each ranch represents a decision or a possible event.” In addition, decision tree make classification easy and understandable and also result-oriented. Many industries use to classify and predict customers’ behaviour, acquit ion, retention, and growth. Similarly, universities can use this technique to classify students’ performance, behaviour, and expectations.

3. Regression Analysis

Bryman and Duncan (2005) stated that regression is “a powerful tool for summering the nature of the relationship between variables and for making predictions of likely values of the dependent variables.” It is used as a continuous variable (Nisbet, 2009). Regression analysis can be applied in data mining to predict students’ GPA, expectations with regard to specialisation selection, complaints and satisfaction levels based on demographic, geographic, psychographic and behavioural variables. Furthermore, simple linear regression and multiple linear regressions are techniques used in regression analysis. These techniques can be done by using SPSS software.

4. Cluster Analysis

Customer analysis is an unsupervised learning technique (Tsai et al., 2011). Customer analysis refers to identifying groups of customers with similar characteristics (Ahn and Sohn, 2009), splitting the full data set into a set of clusters (Baker, 2010) where categories are not known in advance. Han and Kamber (2006) indicate that cluster analysis can be used to generate labels. The objects are clustered or grouped based on the principles of maximizing the intra-class similarity and minimizing the interclass similarity. Clustering is also known as segmentation (Sinha et al., 2010). Segmentation can be done based on demographic variables, (e.g., gender, income, age, qualification, religion, occupation) geographic variables, (e.g., city, country, state, zip code, region) psychographic variables (e.g., lifestyle, personality and values) and behavioural variables (e.g., benefits, and occasion)

In brief, classification and regression models are supervised model and cluster analysis is an unsupervised model (Rokach and Maimon, 2010). Both models are used to predict and classify the large data into usable information. Consequently, universities can use cluster analysis to examine similarities and differences between colleges, students, teachers, administrative staff, courses, and examinations.

VII. APPLICATION OF DATA MINING IN EDUCATION SECTOR

Baker (2010) identifies four areas of application of data mining in education: improve students’ model, discovering models of the knowledge structure, studying the pedagogical support provided by learning software and scientific discovery about learning and learners. He, Wu (2013) examines students’ participation and

learning behaviour by using LVS (Live Video Streaming) of data mining and text mining techniques. They found that students use chat messages to communicate positive emotions, negative emotions, and expressions of social support. In addition, there is no positive correlation between the number of chat messages and final grades. Maqsood (2013) states that data mining can be used to report and analyse the data that help in preparing marketing strategies for targeted students.

Samira et al. (2012) discusses the interface between data mining research in higher education and legal implications of data. They stated that there is 'hardly any research that focuses on applications for improving higher education using data in a holistic way. They suggest that teachers can use data mining to find innovative way to help improving and teaching as well as develop assessment procedures. Cios et al. (2007) indicates that data mining allows for automated discovery of patterns and trends in the data.

Tsai et al. (2011) examined computer proficiency test using data mining technique (cluster analysis, decision tree) to assess students' computer literacy before admitting into higher education. They concluded that data mining techniques assist universities to identify a number of groups who need reinforcement training and promote their computer proficiency more efficiently.

Furthermore, Pena et al. (2009) studied application of data mining in the web based education field. Data mining applied in student modelling, tutoring, content, and assessment areas. They conclude that education data mining help in understanding learning outcomes, identify students; behaviour, and characterise groups of students. Data mining can be applied in the following functional areas of an education sector.

1. Students' Enrolment Prediction

Aksenova et al. (2006) build predictive model for fresher, existing and returned students at both graduate and undergraduate levels. This model is built based on population, unemployment rates in the region, institutional tuition fees, household income, enrolment past data of institutions. Data is mined with the help of Cubist tool. They conclude that data mining has an enormous application in higher education.

Kovacic, J. Zlatko (2010) predicted students' success based on enrolment data (socio-demographic and environment variables) using data mining techniques such as CART (Classification and Regression Technique). They concluded that classifying students based on pre-enrolment information helps to identify students 'at-risk' of dropping the course and suggest using orientation, advising and mentoring programs to make them success.

2. Predicting Students' Profiling

Romdhane et al. (2010) indicate that data mining allows building customer models each describing the specific habits, need and behaviour of group of customers. It classifies new customers and predicts their special need. Consequently, data mining can help management to identify the demographic, geographic and psychographic characteristics of students based on information provided by the students at the time of admission. Profiles are often based on demographic and geographic variables (Berry and Linoff, 2004).

According to Chen et al. (2005) data mining can be applied to describe behaviour of customers. Furthermore, surveys are one common method of building customer profiles. Neural networking technique can be used to identify different types of students. In addition, Discriminant analysis can also be used to identify patterns. Regression analysis, decision tree and Bayesian classification can be applied. Consequently, cluster analysis can be done to students' profiling and separate marketing strategies can be prepared to target segmented students. Cluster analysis is also called data segmentation (Sinha et al., 2010).

3. Curriculum Development

Hsia et al. (2008) study course preferences, completion rates and profession of enrollees by using data mining algorithm such as decision tree, link analysis and decision forest. They found the correlation between course category and enrollee professions. They lay emphasis on importance of data mining in building curriculum and marketing in the field of higher education.

4. Students' complaints

Chen et al. (2012) proposes 'PARA' (P=Primary Diagnosis, A=Advanced Diagnosis, R=Review and A=Action) model of service failure based on customer complaints. Data mining technique is used to identify categories of complaints and develop strategy to improve services.

5. Course Completion

Universities can cluster students into groups based on students' satisfaction, students' loyalty and degree of complaints to understand students' patterns towards course completion.

6. Students' Targeting

Woo et al. (2005) defines customer targeting as "a process of building strategy towards specific customers." They indicate that customer map is the visualisation method for customer targeting. Customer map helps in building customer-oriented strategy. It is a "novel technique to find right target customers who are homogenous with characteristics, values and needs. It is organised with three dimensions of customer targeting: Customer Value (usage and behaviour), customer characteristics (demographic and Psychographic), and customers' needs (complaints and satisfaction). Target customers are detected and targeting strategies can be derived from a customer map.

7. Library facility

Data mining techniques can be used in library to explore students' reports in relation to loan accounts, books selections, and books shelves to gain information. In addition, clustering analysis can be applied to understand, for example, books selection and ordering system based on gender, age, and grades of students.

8. Students' course selection

Kardan et al. (2013) determine various factors influencing student course selection using neural networks such as characteristics, students' workload, course grades, course type, course duration, and number of time conflicts, final examination time and students' demand. These factors are used as input of neural network modelling. Furthermore, Guo (2010) analyse and predict student course satisfaction using neural networks. He found that number of students enrolled to a course and high distinction rate in final grading are the two most influential factors to student course satisfaction.

9. Teachers' teaching performance

Mardikyan and Badur (2012) identified factors that affect instructors' teaching performance in university by using stepwise regression and decision tree of data mining techniques. Instructor attitude, employee status, student attendance, and students' feedback affect teaching performance.

10. Students' Performance

Kumar and Uma (2009) studied students' performance in the course using data mining techniques, particularly classification techniques such as Naïve Bayes and Decision tree based on students ID and marks scored in course. Furthermore, they suggest that data mining process can be done to the teachers for classifying performance which helps in improving higher education system. Data mining methods helps students and teachers to improve students' performance.

11. Students' Dropouts

Massa and Puliafito (1999) studied the problem of university dropout by using Markov Chains, a new data mining technique. They conclude that behaviour of homogenous group can be studied through Markov chain. It can be used to define clusters of students associated with different dropout risk degree.

12. Students' relationship management

Analysing students' data by using data mining techniques is crucial for students' relationship management. Data mining techniques can be used to acquire, maintain and retain the students to achieve the profitability of an organisation. Maqsood (2013) defines CRM as "an integration of a business strategy, organisational process supported by technology for acquiring most profitable and unprofitable customers by providing value-added services and maintain better relationship by offering customer loyalty programs."

VIII. CONCLUSIONS

This paper examines application of data mining in education section. Data mining is used in industries for a variety of applications such as predicting and classifying customers and clustering customer's characteristics and prepares marketing strategies to segmented customers for the achievement of profitability. Similarly, universities can also apply data mining for predicting enrolment of students into various courses. Data mining can be applied for classifying and clustering students characteristics based on demographic, psychographic and behavioural variables. Data mining can also be applied by using if-then rule. In addition, it can describe the profile of successful and unsuccessful students based of GPA achieved during the semesters. It can also be used for dropout student, students' academic performance, teachers' performance, and students' complaints.

REFERENCES

- [1] Aggarwal, C. Charu and Yu, S. Philip. "Data Mining Techniques for Associations, Clustering and Classification." in Zhong, Ning and Zhou, Lizhu (Eds.) methodologies for knowledge discovery and data mining, third pacific Asia Conference, PAKDD, Beijing, China, April 26-28, 1999 proceedings, Springer, New York
- [2] Ahn, Jin. Sook, and Sohn, S. Young. "Customer pattern search for after sales services in manufacturing." Expert System with Applications, 6 (2009): 5371-5375
- [3] Aksenova, S. Sretlana., Zhang, Du and Lu, Meilin. "Enrolment prediction through Data Mining" IEEE Conference Publications (2006): 510-515
- [4] Baker, R.S.J.D., Data Mining for Education, in B. McGraw, P. Peterson, and E. Baker (Eds.), international Encyclopaedia of Education, Third Edition, Vol. 7, 2010, 112-118
- [5] Baritchi, Andi. Data Mining and Knowledge Discovery, in (Eds.) Business Intelligence in the Digital Economy, Raisinghani, S. Mahesh, 2004, Idea Group Inc, 2004,
- [6] Berry, J.A. Michael and Linoff S. Gordon., "Data Mining Techniques for Marketing, Sales, and Customer Relationship Management." Second Edition, Wiley Publishing, 2004
- [7] Bryman, Alan and Cramer, Duncan. Quantitative Data Analysis with SPSS 12 and 13: A Guide for Social Scientists, 2005, Routledge
- [8] Cios, K.J., Swiniarski, W.R., Pedrycz, W. and Kurgan, A. L., Data Mining: A Knowledge Discovery Approach, 2007, Springer, New York
- [9] Chen, Chi-Kuang., Shie, An-Jin., and Yu, Chang-Hsi. "A Customer-oriented Organisational diagnosis model based on data mining of customer complaints databases" Expert System with Applications, 39 (2012): 786-792
- [10] Chen, Mu-Chen, Chiu, Ai-Lun and Chang, Hsu-Hwa. "Mining changes in customer behaviour in retail marketing." Expert System with Applications, 28 (2005): 773-781
- [11] Delen, Dursun. "Analysis of Cancer Data: A Data Mining Approach." Expert Systems, vol.26 No.1 (2009)
- [12] Giudici, Paolo. Applied Data Mining: Statistical Methods for Business and Industry, John Wiley and Sons Ltd., 2013
- [13] Gopalan, N.P and Sivaselvan, B. Data Mining: Techniques and Trends, PHI, New Delhi, 2009
- [14] Gorunescu, Florin. Data Mining: Concepts, Models and Techniques, Springer-Verlag Berlin Heidelberg, 2011
- [15] Guo, W. William, "Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction." Expert System with Applications, 37 (2010) 3358-3365
- [16] Han, Jiawei and Kamber, Micheline. Data Mining: Concepts and Techniques, Second Edition, Morgan Kauffmann Publishers.
- [17] He, Wu. "Examining students' online interactions in alive video streaming environment using data mining and text mining" Computers in Human Behaviour, 29 (2013): 90-102.
- [18] Hsu, Hui-Huang. Introduction to Data Mining in Bioinformatics, in Hsu, Hui-Huang Reichgett, (Eds.), Advanced Data Mining Technologies in Bioinformatics, Idea Group Inc, 2006, 1-12
- [19] Hsu, Jeffrey. Data Mining and Business Intelligence: Tools, Technologies and Applications, in (Eds.) Business Intelligence in the Digital Economy, Raisinghani, S. Mahesh, 2004, Idea Group Inc, 2006, 141-190
- [20] Huang, T. Cheng-Kui, Lin, Chuang-Chun and Chang, D Cheng. "An Empirical Investigation of factors influencing the adoption of Data Mining Tools." International Journal of Information Management, 32 (2012): 257-270
- [21] Hsia, Tai-Chang., Shie, An-Jin and Chen, Li-Chen., "Course Planning of extension education to meet market demand by using data mining technique: an example of Chinkuo Technical University in Taiwan." Expert System with Applications, 34 (2008): 596-602
- [22] Kalogirou, A. Soteris. "Neural Network Modelling of Energy Systems" Encyclopaedia of Energy, Vol. 4 (2004): 291-299
- [23] Kantardzic, Mehmed. Data Mining: Concepts, Models, Methods and Algorithm, Second Edition, John Wiley and Sons, New Jersey, 2011
- [24] Kardan, A. Ahmad., Sadegh, Hamid, Ghidary, Shiry. Saeed., and Sani, Reza Fani., Mohammad. "Prediction of student course selection in online higher education institutions" 65 (2013): 1-11
- [25] Kumar, A.N.V. and Uma, G.V. "Improving academic performance of students by applying data mining techniques." European Journal of Scientific Research, No. 4 (2009): 526-534
- [26] Kovacic, J. Zlatko, "Early prediction of student success: Mining students' enrolment data, in proceedings of informing science and IT education Conferences (InSITE) 2010, 647-665.
- [27] Larose, T. Daniel. Discovering knowledge in data: An Introduction to Data Mining Techniques, John

- Wiley and Sons, New Jersey, 2005
- [28] Luan, Jing. Data mining and its applications in higher education, new directions for institutional research, No. 113, 2002, Springer
- [29] Nisbet, Robert and Elder, John and Miner, Gary., Handbooks of Statistical analysis and Data Mining Applications, Academic Press Publications, 2009
- [30] Ngai, E.W.T, Xiu, Li and Chau, D.C.K “Application of Data mining techniques in customer relationship management: A literature review and classification.” Expert System with Applications, 36 (2009): 2592-2602
- [31] Massa, S and Puliafito, P.P. “An application of data mining to the problem of the university students’ dropout using Markov chains.” in Zytknow, J.M and Rauch, J (Eds.), Principles of Data mining and Knowledge Discovery, third European conference, PKDD 99, Prague, Czech Republic, September 15-18, 1999, proceedings, Springer, vol. 1784, 51-60
- [32] Maqsood, A. Mohammed, “Customer Relationship Management in B-Schools: An Overview.” International Journal of Computer Sciences and Management Research, Vol. 2 issue 4 (April 2013):
- [33] Mardikyan, Sona and Badur, Bertan. “Analysing teaching performance of instructors using data mining techniques.” informatics in Education, Vol.10, No.2, (2011): 245-257
- [34] Mark, Last. Data Mining, in Cyber Warfare and Cyber Terrorism, Janczewski, Lech and Colarik, Andrew., Information Science Reforms, IGI Global, (2007): 358-365
- [35] Olson L. David and Delen, Dursun. Advanced Data Mining Technique, 2008, Springer
- [36] Pena, Alejandro. Dominguez, R., and Medd, J., “Education Data Mining: A sample of review and study case.” Academic World and Education and Research Center, 2 (2009):118-139
- [37] Pujari, K. Arun. Data Mining Techniques, Universities Press India Pvt. Ltd., Hyderabad, 2001
- [38] Rajamani, Karthick., Sung, Sam., and Lox, Alan. “Extending the applicability of association rules.” the proceedings of Third Pacific Asia Conference, PAKDD 99, Beijing, China, April 26-28, 1999
- [39] Ramageri, M. Bharti “Data Mining Techniques and Applications.” Indian Journal of Computer sciences and Engineering, Vol.1 No.4, (Dec-Jan 2010-2011): 301-305
- [40] Riyazuddin, Qureshi., AlManna, Yahiya Mohammed and Deshmukh, Anwar Pasha. “Big Data: Growing Pressure on global Storage by Data created on Social Networking Sites.” International Journal of Computer Sciences and Management Research, Vol. 2 issue 3 (March 2013): 1876-1881
- [41] Rokach, Lior. “A survey of clustering algorithm,” in O.Maimon, L. Rokach (Eds.), Data Mining and Knowledge Discovery Handbook, 2nd Edition, Springer, 2010
- [42] Rokach, Lior and Maimon, Oded. “Supervised Learning,” in Maimon, L. Rokach (Eds.), Data Mining and Knowledge Handbooks, 2nd Edition, Springer, 2010
- [43] Romdhae, L.B., Fadhel, N. and Ayeb, B. “An Efficient approach for building customer profiles from business data.” Expert System with Applications, 37 (2010): 1573-1585
- [44] Romero, Cristobal., Ventura, Sebastian and Garcia, Enrique. ‘Data Mining in Course Management Systems: Moodle case study and tutorial.’ Computers and Education 51 (2008): 368-384
- [45] Samira, AlEtia, Ipperciel, Donald and Hammad, Ahmed. “Implications and challenges to suing data mining in educational research in the Canadian context.” Canadian Journal of Education, 35, 2 (2012): 101-119
- [46] Sen, Baha., Ucar, Emine., Delen, Dursun. “Predicting and analysing secondary education placement-test scores: A data mining approach.” Expert System with Applications, 39 (2012): 9468-9476
- [47] Sinha, Nidhi., Vandana, Ahuja., and Medury, Y. “Effective Brand Management through customer profiling using clustering, IUP Journal of Brand Management, 2010, 21-33
- [48] Soman, K.P., Diwakar, Shyam and Ajay, V. “Insights into Data Mining: Theory and Practice, PHI, New Delhi, 2006
- [49] Taniar, David and Rahaya, T. Wenny. Parallel Data Mining in (ed.,) Data Mining: A Heuristic approach by Abbass, Hussein, Ruhul Sarker and Charles Newton, Idea Group Inc, 2002, 261-289
- [50] Tiwari, Mahendra., Randhir Singh and Neeraj Vimal., “An Empirical Study of Application of Data Mining Techniques for Predicting Student Performance in Higher Education.” International Journal of Computer Sciences and mobile Computing, Vol.2 Issue 2 (Feb 2013): 53-57
- [51] Tsai, C.F., Tsai, C.T., Hung, C.S., Hwang, P.S., “Data Mining Techniques for identifying students at risk of failing a computer proficiency test requires for graduation.” Australian Journal of Educational Technology, 2011, 27(3): 481-498
- [52] Tuffery, Stephane. Data Mining and Statistics for Decision making, John Wiley and Sons Ltd., 2011
- [53] Weiss, M. Sholom and Indurkha, Nitin. Predictive Data Mining: A practical guide, Morgan Kauffmann Publishers, USA, 1998
- [54] Woo Ji-young., Bae, Sung Min and Park, Sang Chan. “Visualisation method for customer targeting using customer map.” Expert System with Applications, 28 (2005): 763-772

- [55] Witten, H. Ian and Frank, Eibe. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition, Morgan Kaufmann, 2005
- [56] Yin, Yong. Kaku, Iku., Tang, Jiafu., and Zhu, Jian Ming. Data Mining: Concepts, Methods and Applications in Management and Engineering design, Springer, New York, 2011