

Data Mining in The M-Learning Domain

Fadi R. Shahroury

Princess Sumaya University for Technology

Department of Electronics Eng.

Amman, Jordan

Email: fadi@psut.edu.jo

Abstract—Terabytes of data are generated everyday in many organizations all over the world; these huge volumes of data raised the need for evolutionary ways of managing and using data such as; Data mining which is capable of discovering hidden patterns and trends in organization's data assets. Data mining is a thriving research area however; the concept of data mining in academics is still at its early stages but appears promising; thus this paper aims to highlight the potential of data mining in the m-learning domain, and propose a model for applying data mining in m-learning and suggests that various data mining tools can be beneficial for m-learning from many perspectives.

Keywords-Data mining, M-learning, E-learning, M-web, Clustering, Association analysis.

I. INTRODUCTION

During the last decade the mobile industry have changed the world, due to the mobile revolution that introduced mobile devices with advanced features that can perform many enormous functions in different aspects, mobile technology has made a massive contribution to the human life not limited to communication and connectivity but it's contribution extended to education; making the learning experience widely accessible for remote learners all over the world and more flexible and suitable against various constraints such as work, family or any other commitments.

Learning nowadays has been transformed from "print" to "e-learning" to "m-learning"; Electronic learning referred to as E-learning is a learning experience in which total automation of teaching and learning processes is achieved through Learning Management Systems, and by utilizing web-based technologies for universities and other educational institutes, e-learning has been a great success in many institutions such UK Open University, Hong Kong Open University, Athabasca Open University, etc [1].

Mobile learning referred to as M-learning can be considered as an expansion of e-learning; M-learning is the ability to learn independently without any constraint of place and time, facilitated by a range of mobile devices, such as mobile phones, iPod, MP3 players, PDAs, pocket PCs, and blackberries. These devices allow learners to access bite-sized information like, course materials, online references materials, audio and video learning materials, exam schedules, online quizzes and exams, counseling sessions, and

assessments; however the mobile device should be enabled with compatible technology) [2].

Since almost half of the world's population are mobile phone holders; M-learning appears to be a promising field since it's empowering learners with the ability to learn any-time and anywhere, access learning resources, communicate and interact with other learners all over the globe, participate in creating learning resources, and access the World Wide Web; thus many Open Universities are coming up with a new concept of the web which is mobile web (m-web); which is internet for small screen that would be suitable and viewable on a mobile devices which allow quick access to web content; however websites with no mobile versions appear compressed on the small screen, and link and menus may appear overlapped; m-web can be characterized as following:

- Information access anytime and anywhere; m-web enable users with 24-7 internet access regardless of location and ability to retrieve and exchange information.
- Limitless access; m-web is not limited to websites with mobile versions but allow access to the whole web.
- Interactive capabilities; m-web allow learners to interact with other learners by creating assignments, uploading videos even taking by their camera phones, sharing comments and posts, starting conversations and video-conferencing.
- Location awareness; since almost all smart phones these days have a Global Positioning System (GPS) which can locate the mobile phone all the time; this feature can help learners to search for study points nearby, and retrieve directions for a specific group or destination [3].

II. DATA MINING

Data mining is collection of tools and methods that search for valuable and meaningful information in large volumes of data, which allow decision makers to learn from their data assets; data mining discover hidden relationships, patterns, and interdependencies in massive databases and data warehouses. Data mining is the process of extracting implicit, unknown and useful information such as knowledge rules and constraints using pattern recognition techniques and statistical and mathematical techniques; Data mining tools include various artificial intelligence methods such

as expert systems, fuzzy logic and others, decision trees, genetic algorithms, association rules, clustering tools, neural networks, and data visualization [4].

III. DATA MINING STEPS

There are three main steps in Data mining; preparing data, reducing data and finally searching for valuable information, however [5] defined four major steps in data mining:

- Predictive Modeling: the process of creating predictive models using inductive reasoning techniques and algorithms.
- Database Segmentation: the process of partitioning data into different clusters and segments using various clustering techniques.
- Link Analysis: the process of discovering valuable associations between data.
- Deviation Detection: the process of detecting and explaining the reason for excluding certain records from specific segments.

Further [6] explained that data mining passes through phases; at first selecting the dataset that is extracted from operational database or business transactions. In the second phase; cleaning and pre-processing is applied on the selected dataset to remove redundancy, inconsistency, and noisy data. In the third phase, the dataset is examined to discover relationships between data and thus identify hidden patterns by applying some data mining tools and algorithms, such as clustering, classification and prediction, genetic algorithms, expert system, decision trees, etc. In the fourth phase, the patterns will be evaluated by testing samples or other datasets. Finally the data mining phases are pursued iteratively until useful knowledge is extracted. [7] summarized data mining phases shown in Fig. 1

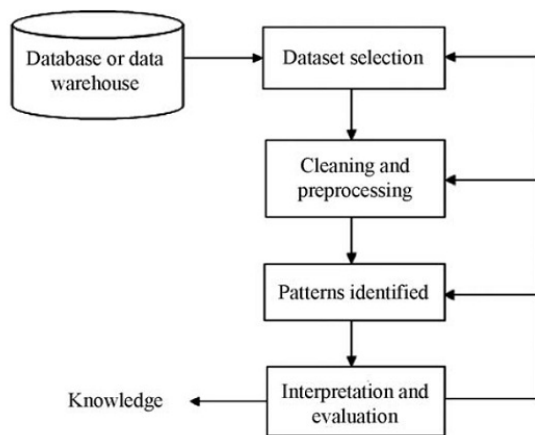


Figure 1. Data mining phases.

IV. WEB MINING

Applying data mining to the web can be valuable in different aspects; web mining means search for data related to the web, this data may actually be present in the web page or data that is related to web activity; web data can be classified into; content of web page, intra-page structure which is the linkage structure between web pages, usage data which is data about how users access web pages, and user profile that include registration information and demographics, see Fig. 2 [8].

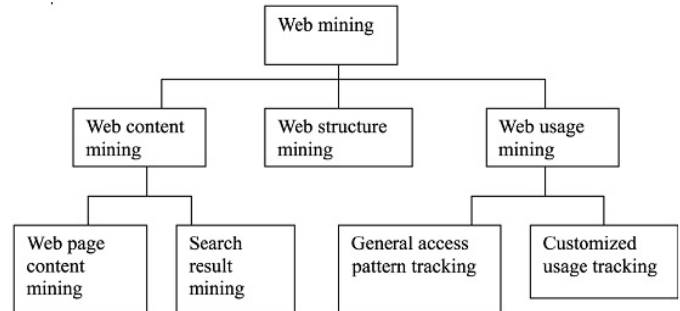


Figure 2. Web mining taxonomy.

V. WHY DATA MINING IN M-LEARNING DOMAIN?

The continuous generation of huge amounts of data every-day all over the world and the immediate need to manage and utilize this data has driven the need for data mining in all businesses and services industry; data mining is a natural result of the technology evolution in data collection and creation, data management and data analysis.

In m-learning there is a need to manage two groups of users; learners and learning providers, for learners; databases should store all personal information such as name, age, address, education details, qualification, moreover; information about work experience, income range, courses taken, courses of interest, this kind of information will be valuable to predict future behavior of different employed people, also personal interests and hobbies would be beneficial.

Regarding the learning providers; databases should keep track of name of the institute, offered courses, area of courses, duration and credits of courses, pre-requisites for each course, course content, course path, cost, and targeted learners; these information will be valuable to suggest relevant or complementary courses, data mining can allow learning providers to view learners and courses data from multiple angles and that would help in better decisions such as, investment in highly recommended courses. Data mining will help in recognizing patterns and future trends of learners, making more efficient decisions regarding courses and certificates, suggesting other courses from market point-of-view [8]

VI. APPLICABLE DATA MINING TOOLS IN M-LEARNING

- 1) Associations Analysis, this tool identify rules of affinities among data, application of association analysis include market basket analysis, attached mailing in direct marketing, fraud detection, department store etc. Association in m-learning can be used to track students activities related to discipline programs, specializations and courses. Association analysis can also be beneficial to find rules for example between age of learners and courses of interest [9].
- 2) Clustering, this tool is used to segment data into similar groups, Clustering techniques are ideal for classification and category prediction problems. According to [4] clustering starts with creating a set of random, normalized representatives. then follows an iterative process consisting of; randomly choosing a record from the pre-processed normalized database, order the representatives by their similarity to the random record, and then adjust the closest representative in the direction of the data record, then similarly adjust each of the other representatives, this process continues until a satisfying classification is made, or no significant movement in the representatives occurs for a long period of time, or a pre-specified number of iterations have occurred .
In the context of m-learning clustering can be used to group learners based on their educational backgrounds, specialization, course interest and other similarities.
- 3) Classification and Prediction, these two tools are used to define data classes and predict future data classes and trends, these tools can be valuable in m-learning in terms of classifying learners into groups based on their learning performance to excellent learners, average learners, or classify courses to attractive courses and non-attractive courses, and predicting learners behavior.
- 4) Genetic Algorithm, this method is based on the biological mechanism of natural selection and survival of the fittest, genetic algorithms evolve complex data structures for finding solution to complex optimization problems, it will run for hundreds of generations until an acceptable solution is found or after a pre-specified amount of time, and thus will generate optimal solutions to problems.
- 5) Expert System, a system that can imitate the reasoning of human experts, it consists of a knowledge base of rules, facts, and a logic based inference engine which creates new rules and facts based on previously accumulated knowledge and facts. However; the knowledge should be easy to capture, self-consistent, simple to explain, and straightforward to represent.

Data mining tools are highly recommended to enhance enrollment management. For example; discover which in-

quiries are most likely to turn into actual applications. Moreover; applying prediction on course enrollment can rate the success of a particular course. Data mining can forecasts the demand for specific existing or new courses, and thus enhance the admission process, moreover data mining can improve the learning process by identifying groups of students who may need special attention, or monitoring attendance or grades or any other indicators [9].

VII. THE PROPOSED MODEL

Fig. 3 shows the proposed model for applying data mining in m-learning domain, the model suggests that there is a potential to cluster similar m-learners into groups based on many factors such as learners backgrounds, gender, major, common interests; thus there is a possibility to identify m-learners behavioral patterns. This model includes five steps which will be explained as follows.

- 1) Select the m-learning database or target subset of the database on which the data mining tools will be applied; further it is necessary to identify the goal of the mining from different perspectives.
- 2) Cleaning and processing data which include handling errors and noisy data, remove redundancy, managing missing data, and preparing database for the mining process.
- 3) Clustering m-learners into similar groups based on factors that must be defined earlier depending on the learning provider point of view, these factors might be m-learners backgrounds, gender, major, common interests, etc.
- 4) Applying various data mining tools such as association rules, classification, genetic algorithms and other techniques to predict m-learners behavioral and thus identify m-learners patterns.
- 5) Recommend potential courses for m-learners based on the patterns identified earlier after applying data mining methods.

VIII. CONCLUSION

In this evolutionary era; mobility is truly coming of age and offering huge scope for innovative solutions. M-learning is becoming a pioneering solution for education problems regarding constraints of work, time limitation, etc. M-learning offers the potential to learn anytime and anywhere, interact with other learners globally, and access learning materials. However mining m-learning content is becoming a necessity to track, manage, and better understand volume of data, which will lead to more efficient decisions. This paper suggest that data mining in the m-learning domain is applicable to enhance enrollment management, admission process, segmenting learners based on different indicators, recognizing hidden learning behavior, and predicting future learning trends, further this paper propose a model to apply data mining in the m-learning domain, however; this model

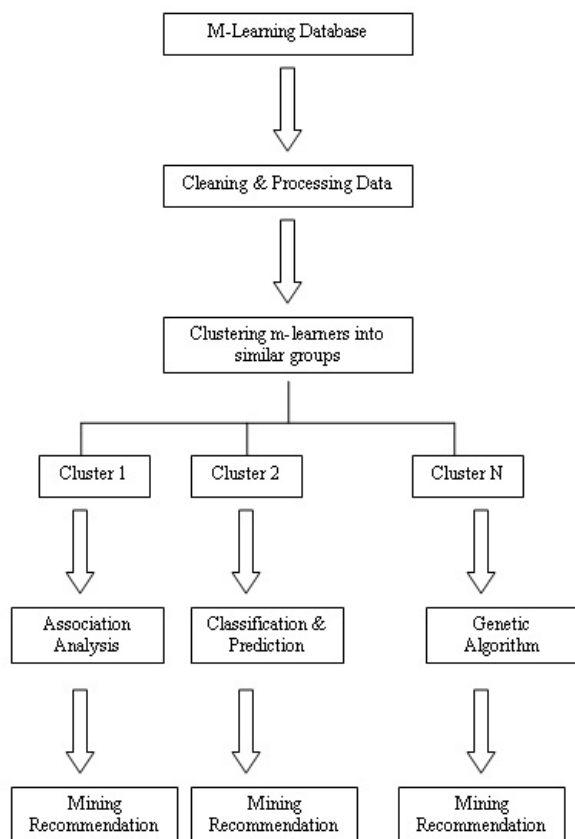


Figure 3. Proposed model for applying data mining in m-learning.

needs to be investigated on real case studies to suggest further recommendations. Moreover; the research about data mining in academics is still evolving and more investigation is highly recommended to address the implications and challenges of data mining in academics generally and in m-learning specifically.

REFERENCES

- [1] S. Chandhok and P. Babbar, "M-learning in distance education libraries: A case scenario of indira gandhi national open university," *The Electronic Library*, vol. 29, no. 5, 2011.
- [2] K. Donnelly, "Learning on the move: how m-learning could transform training and development," *Development and Learning in Organizations*, vol. 23, no. 4, 2009.
- [3] E. Kroski, "On the move with the mobile web: libraries and mobile technologies," *Library Technology Reports*, vol. 44, no. 5, 2008.
- [4] M. L. Gargano and B. G. Raggad, "Data mining - a powerful information creating tool," *OCLC Systems & Services*, vol. 15, no. 2, 1999.
- [5] S. J. Lee and K. Siau, "The effects of lean production on worker job stress," *Industrial Management & Data Systems*, vol. 101, no. 1, 2001.
- [6] W. H. J. L. P. H. Tai, D.W.S., "Effective e-learning recommendation system based on self-organizing maps and association mining," *The Electronic Library*, vol. 26, no. 3, 2008.
- [7] C. M. Huang, M.J. and S. Lee, "Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis," *Expert Systems with Applications*, vol. 32, no. 3, 2007.
- [8] M. Hanna, "Data mining in the e-learning domain," *Campus-Wide Information Systems*, vol. 21, no. 1, 2004.
- [9] J. Ranjan and K. Malik, "Effective educational process: a data-mining approach," *The journal of information and knowledge management systems*, vol. 37, no. 4, 2007.