International Journal of Scientific Engineering and Technology
Volume No.4 Issue No.11, pp: 523-525
ISSN :2277-1581
01 Nov.2015

Moodle Data Retrieval for Educational Data Mining

Felermino M. D. A. Ali, S. C. Ng
School of Information Technology, SEGi University
No.9, Jalan Teknologi, Taman Sains Selangor Kota Damansara, PJU 5, 47810, Petaling Jaya, Selangor, Malaysia
felerminoali@gmail.com, ashleyng@segi.edu.my

***Abstract*—This study aims to incorporate a new service into existing Moodle Core Services for student's usage data retrieval. This service can ease the pre-processing in EDM and provide a generic mechanism for data portability. The proposed method expects to reduce the data pre-processing time and enhance the interoperability in Moodle.**

Keywords— Moodle, Web Services, E-learning, Educational Data Mining, Data retrieving.

I. Introduction

EDM recently became an important trends in education industry particularly for CBE settings which usually generates huge amounts of potential data for EDM [i][ii]. However, there is a need to have tools that efficiently extracts data from these systems. In this sense, this study adopts the basis of the source code available on Moodle 2.7 and adds a new plugin for student's usage data retrieval. The main aim of the study is to reduce the time and resources engaged in pre-processing tasks as well as provide an interoperability mechanism for EDM.

I.I Background

EDM is a new growing research field that concerns with exploration of educational data, and development of methods to apply in existing education systems to address educational issues. In CBE environment, EDM has found a wide use, especially to improve user's activities [ii], such as:

- Performance prediction for learners (e.g. final marks);
- Providing effective learning assistance by educators to their learners;
- Reducing the costs of educative personalization and adaptation process for educational administrators.

So far, Moodle (Modular Object Oriented Development Learning Environment) CBE is the most tested free learning platform that enables the creation of powerful, flexible and engaging online courses and experiences [iii]. Moodle is open-source, and can be extended through addition of new plugins or functionalities. In fact, Moodle enhancements comes from a global community of people who shares ideas, code, information and free support [iv].
Most of recent researches [v][vi][vii][vii][ix][x][xi][xiv] applies EDM over the Moodle student's usage data, which essentially consists of all information gathered from student's interaction with Moodle activities (e.g., forum, chat, email achieve, quizzes, and assignments) and the final mark in the course [viii] (see Table 1).

Table 1: Students' usage summary table on Moodle

Name	Description
id_student	Identification number of the student
id_course	Identification number of the course
num_sessions	Number of sessions
num_assignment	Number of assignments done
num_quiz	Number of quizzes taken
a_scr_quiz	Average score on quizzes
num_posts	Number of messages sent to the forum
num_read	Number of messages read on the forum
t_time	Total time spent on Moodle
t_assignment	Total time spent on assignments
t_quiz	Total time spent on quizzes
t_forum	Total time spent on forum
f_scr_course	Final score of the student obtained in a course

*Source from [ii]

The process of applying EDM on Moodle is similar to the Knowledge Discovery in Database (KDD or Data Mining), which consists of three tasks: data pre-processing, data mining and result interpretation. Data pre-processing phase is crucial for the effectiveness of the latter tasks. However, data pre-processing usually consumes over 60% of time, effort and resources employed in the whole EDM process as it involves data collection, data cleansing and data transformation into suitable format for mining [v]. To solve this problem, this paper proposes an approach to facilitate data pre-processing by using web services specifically on data collection and transformation.

I.II Related Work

So far there are two tools to facilitate data extraction in Moodle, namely MMT (Moodle Mining Tool) tool [xiii] and ADE [xiv] (Automatic data extraction). But both tools were adapted accordingly to a particular data mining tool. MMT has been implemented over the KEEL framework, whereas ADE has been implemented over RapidMiner framework. However, coupling Moodle and data mining frameworks usually sets boundaries in terms of data portability. For instance, to reuse the data extracted using ADE in MMT requires further steps of data transformation as dataset format depends heavily upon the data mining tool. The use of web services appears as a solution to the pointed limitations because it provides an intermediary data representation stage in XML. This method decouples and provides independence during data extraction on Moodle, as well as providing accessibility to wide range of platforms. Figure 1 illustrates the flow of the proposed web service execution. The service will transform Moodle relational data into a dataset by merging several tables into one table, and then return the table as XML format.

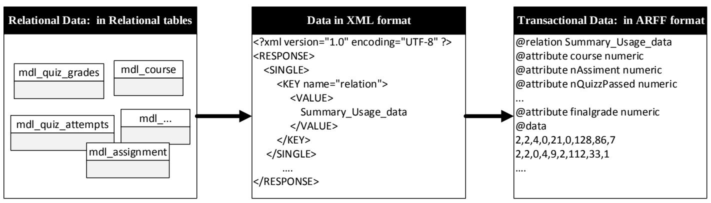

Figure 1: Proposed data transformation flow

II. Methodology

The core idea of this study is to provide a modular design to extract data for EDM through integration and layering of existing Moodle Core Services. The service will deliver student's usage data as this dataset remains consistent in most of recent studies, and it has been tested in most of the traditional data mining techniques. Figure 2 shows the proposed service implementation workflow. The process essentially consists of five steps, where the first four steps deals with the service implementation, whereas, the last step ensures the service accessibility and security.

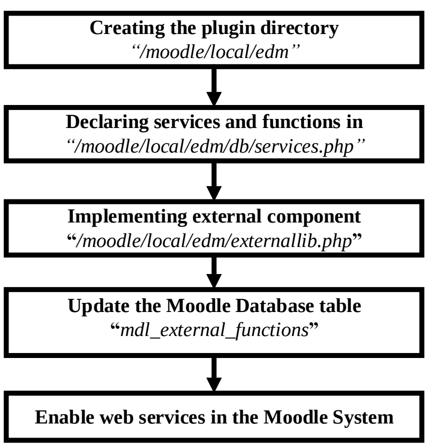

Figure 2: Service implementation workflow

The main implementation is in step 3 where three methods are created (see Figure 3). The first method implements the service itself. The second method describes the service parameters and the third one describes the service return. In addition, two other methods, *attributes_description* and *instances_description*, are also created to structure the XML in such way to represent a generic dataset so that the service response will be delivered in a Key-Value pair XML that contains two parts, known as header and data. Header contains the name of the relation, a list of the attributes, and attribute types. On the other hand, data contains the information. In this way, the data extracted using the web service will be portable and convertible to multiple types of dataset.

```
<?php
require_once($CFG->libdir . "/externallib.php");
class local_edm_external extends external_api {

 // formatting return Key and Value
 static function attributes_description() {...}
 static function instances_description() {...}

 // function implementation
 static function get_students_usage_data($param)
 {...}
 static function get_students_usage_data_parameters()
 {...}
 public static function get_students_usage_data_returns()
 {...}
}
```

Figure 3: Service externallib.php example

III. Results and Discussion

In order to test the service, a Simple Java Console Application (SJCA) was developed. This service delivers data from a simulated Moodle environment. The service response delivers a structured key-value pair XML. The role of data transformation from XML to a dataset is assigned to clients. In this particular experiment, SJCA consumes the service, and JAXB (Java Architecture for XML Binding) mechanism was able to successfully transform the returned XML into ARFF representation (see Figure 4).

```
@relation Summary_UsageDataNumerical
@attribute studentId numeric
@attribute course numeric
@attribute num_sessions numeric
@attribute num_assignment numeric
@attribute num_quiz numeric
@attribute a_scr_quiz numeric
@attribute num_posts numeric
@attribute num_read numeric
@attribute t_time numeric
@attribute t_assignment
@attribute t_quiz
@attribute t_forum
@attribute finalgrade numeric

@data
50,2,4,0,21,0,8,2,164, 128,86,7
90,2,0,4,9,2,28,121,164,112,33,1
98,3,1,3,24,6,35,42,111,76,115,6
```

Figure 4: Output service in ARFF

In summary using Moodle service for data pre-processing proved to bring the following advantages:

1. Portability - enables data conversion into multiple dataset types, moreover, the service can be accessible from a wide range of platform (e.g. desktop, web, android, etc.);
2. Modularity - from design perspective, using web services enables more flexibility in managing data from Moodle to data mining tools as services provides a layered approach to separate both tools according their purpose.
3. Reuse – developers will not need to explicitly implement algorithms for data retrieval because these algorithms are already available through services.

IV. Conclusion

This paper presented an approach for EDM data pre-processing through web services on Moodle environments specifically for data collection, and data transformation into dataset. The implementation results show that the proposed web service managed to reduce the time, effort and resources during the EDM pre-processing phase. This method also enables the portability, reuse, and modularity of EDM. As immediate future work, this method will be integrated into Moodle Core Service to improve CBE experience.

References

i. *Jormanainen I, Sutinen E. An open approach for learning educational data mining. Proc 13th Koli Call Int Conf Comput Educ Res - Koli Call '13. New York, New York, USA: ACM Press; 2013;203–4.*

ii. *. Bousbia N, Belamri I. Which Contribution Does EDM Provide to Computer-Based Learning Environments : Educational Data Mining. Peña-Ayala A, editor. Cham: Springer International Publishing; 2014;524:3–28.*

iii. *Moodle [Internet]. 2015. Available from: https://moodle.org*

iv. *Cole J, Foster H. Using Moodle. 2nd ed. O'Reilly Media, Inc; 2008.*

v. *Romero C, Romero JR, Ventura S, Data ÁDÁ. A Survey on pre-processing Educational Data Educational Data Mining. Peña-Ayala A, editor. Cham: Springer International Publishing; 2014.*

vi. *Zafra A, Romero C, Ventura S. Multiple instance learning for classifying students in learning management systems. Expert Syst Appl. Elsevier Ltd; 2011;38(12):15020–31. Romero C, Romero JR, Luna JM, Ventura S. Mining Rare Association Rules from e-Learning Data. :171–80.*

vii. *Romero C, Ventura S, Espejo PG, Hervás C. Data Mining Algorithms to Classify Students. In: International Conference of Educational Data Mining. Monteral, Canada; 2008.*

viii. *López MI, Luna JM, Romero C, Ventura S. Classification via clustering for predicting final marks based on student participation in forums. In: 5th International Conference on Educational Data Mining, EDM 2012. Chania, Greece; 2012. p. 148–51.*

ix. *Romero C, Ventura S, García E. Data mining in course management systems: Moodle case study and tutorial. Comput Educ. 2008 Aug;51(1):368–84.*

x. *García E, Romero C, Ventura S, Castro C De. An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. 2007;(Dm).*

xi. *Romero C, Espejo PG, Zafra A, Romero JR, Ventura S. Web usage mining for predicting final marks of students that use Moodle courses. Comput Appl Eng Educ [Internet]. 2010 Mar 18 [cited 2014 Oct 13];21(1):135–46. Available from:*

xii. *Romero C, Ventura S. A Java desktop tool for mining Moodle data. 2010;*

xiii. *Taylor P, Jovanovic M, Vukicevic M, Milovanovic M, Minovic M. Using data mining on student behavior and cognitive style data for improving e-learning systems : a case study. Int J Comput Intell Syst. 2012;(May 2012):37–41.*