

A Web-based Intelligent Report e-learning System Using Data Mining Techniques

Shivchandra Bhusse

ME 2nd Year, Department of Computer Science,

Padmabhooshan Vasantdada patil Institute of Technology Pune, India.

Abstract- This paper presents a method of enhancing web-based intelligent reports of an e-learning system as an intelligent system through PDCA (Plan, Do, Search, Act). By predicting behaviour patterns of students and modifying the layout of these electronic courses, the emphasis is on developing e-learning systems. Using data mining techniques and increasing the usefulness of courses using personalised modules, an existing learning management system is enhanced. This study presents the system's architecture, implementation, and assessment.

Index Terms- e-Learning using PDCA, web-based Intelligent report of e-Learning, Data Mining Techniques.

1. INTRODUCTION

E-learning is the use of electronic technologies for the method of teaching and learning. It describes the use of electronic media in the process of teaching and learning. We define e-learning as any learning that involves use of internet or intranet. The term e-learning comprises a lot more than online learning, virtual learning, distributed learning, networked or web-based through networked or standalone computers and other electronic devices, as the letter "e" in e-learning stands for "electronic" E-learning should not be thought of as a one-way conversation. It is challenging for professors to consider student's viewpoints in a course where there is no face-to-face contact.

E-learning is not just concerned with providing easy access to learning resources, anytime, anywhere, via a repository of learning resources, but is also concerned with supporting such features as the personal definition of learning goals and the synchronous and asynchronous communication and collaboration, between learners and

between learners and instructors. The use of web-based e-learning system has resulted in a need for all participants to track and develop behavioural patterns. Allowing students to complete an a variety of motivational techniques. Following are the advantages of e-learning system.

- Student and faculty interaction will be improved.
- Providing opportunities for student-centred learning approaches.
- Providing 24/7 accessibility to course materials.
- Providing tools for measuring and reviewing student success that are just-in-time.
- Reducing the amount of paperwork associated with course management.
- Assist teachers in keeping track of their student's learning.
- E-learning is fast and effective. It makes it simple to understand and digest knowledge.

Flexibility, Time and Location independence, Efficient connectivity, Cost efficiency, and Ease of delivery of learning materials are some of the other major advantages.

The remainder of this paper has the following structure. Section 2 contains and Analysis for this proposed system. Concluding remarks are found in Section 5.

2. LITERATURE SURVEY

E-learning platforms have been applied in a variety of ways, each with its own set of goals. This section examines a various learning management system.

There are several open sources and proprietary e-Learning tools available that any educational institution may

use. These systems also allow for the incorporation of personalised modules to suit the needs of educational institutions.

A. Moodle

It's a web-based e-learning framework that's open source. Its major features include collaborative tools and activities, file management using cloud storage services, private and broadcast messaging, progress tracking, mass enrolment, multi lingual capability, bulk course creation, high interoperability, detailed reporting and logs, design and management of courses, content driven collaboration, peer and self-assessment and integrating multimedia and external resources.

Aim in Title	Work	Problem Statement	Year
Data Mining Concepts and its Techniques.	Classification technique applied.	Decision Tree if used, can address Classification and Regression needs.	2011
A Review on Predicting Student's Performance using Data Mining Techniques	Concluded that Decision Tree and Neural Networks both performed with highest accuracy	Did not address that Neural Networks are not good with psychometric factors.	2015
Predictive Analytics by applying a Data Mining Approach.	Application of Rule Based If Then Classification & Decision Tree algorithms.	Decision Tree working with Big Data could be explained.	2015
Using Big Data in the Academic Environment.	Unstructured data studied using the graphical Gephi tool.	Data Mining input in Big Data Academic environment was not included in this work	2015
A Web Based intelligent report e-learning system Author: Marija Blagojevic	Predict the Students performance using Decision Tree	Not comparing linear regression and decision tree	2015

B. Coursera

Coursera is an education platform that partners with top universities and organizations worldwide. It typically provides immediate feedback on points student didn't understand. Many institutions have incorporated courses offered on coursera in their curriculum to provide blended learning to students.

	edX	Coursera	Udacity
Used Technology to Build	Ubuntu, Docker, Django & React	JavaScript, React, AWS, Scala, Cassandra, S3 and MySQL.	AngularJS, Apache, Docker, JS, MongoDB, Python, R
Video with audio	Yes	Yes	Yes
Articles	Yes	Yes	No
Projects	No	No	Yes
Discussions	Yes	Yes	Yes
Assignments	Yes	Yes	Yes
Quiz Tests	Yes	Yes	Yes
Transcripts	Yes	No	No
Certificate	Yes	Yes	Yes
Peer Assessment	No	Yes	No
Adaptive Learning	No	No	No
Course joining timings	Scheduled & Anytime	Scheduled & Anytime	Scheduled & Anytime
Target Users	Anyone	Anyone	Professionals

Table 1. Summary of features supported by various e-learning platforms

3. PROPOSED SYSTEM

The proposed system presents the creation of web-based intelligent report e-learning system using the data mining techniques with PDCA (Plan, Do, Check & Act). This system has ability to track and develop behavioural patterns. Proposed system has three main roles as Admin, Institute/Teacher and Student.

Admin Login

1. Create, Retrieve, Update and Delete operations for Programmes, Categories, Classes and Courses.
2. The admin may create several language Courses, each with its own set of chapters and sub-chapters based on the levels (e.g., Beginner, Intermediator, Professional).
3. Course Material i.e., Video, Documents, etc. can be managed.
4. Admin can set the course fees, MCQ Test on each Chapter.
5. Manage Institute or Teachers.
6. Manage Profile.

Institute/Teacher Login

1. Institute or Teacher must have to participate for the course to access its videos and other materials.
2. Manage separate assignments for Chapters.
3. Check and give remarks on solved assignments by the students.
4. Institute or Teacher can check student's course progress report and other track history.
5. Get Course wise or Student wise reports.
6. Manage Profile (Name, Password, Payment Mode, etc).

Student Login

1. Student can separately participate to courses they want or can directly access the courses already enrolled by their Institute/Teacher.
2. Student can access course chapters, sub-chapters videos and related documents.
3. Course progress and certificates after its completion.

4. Submit Assignments.
5. Give Assessment (MCQ Tests, etc).
6. Manage Profile (Name, Password, Payment Mode, etc).

Our study presents the PDCA cycle, in creating web based intelligent report e-learning systems.

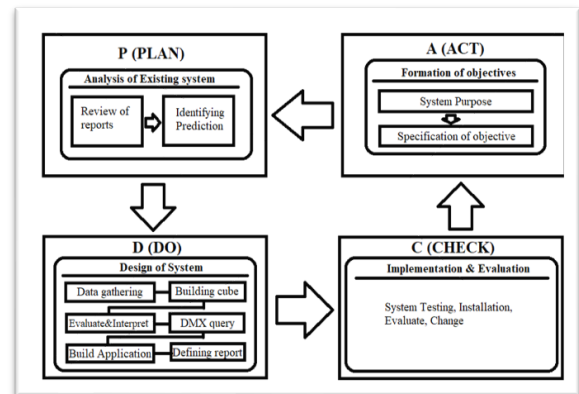


Fig.1 shows the reporting activity in a model system through four activities (Plan, Do, Check, and Act)

3.1 Architecture of Proposed System

The proposed system has following three parts.

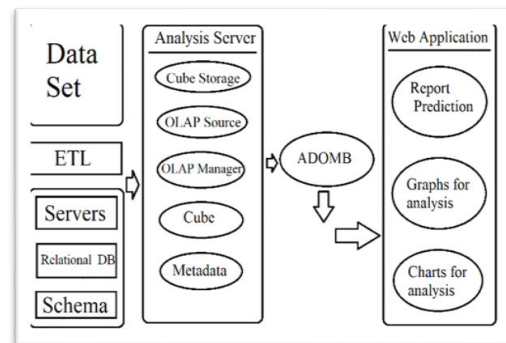


Fig. 2 shows architecture of the system.

3.1.1 SQL server and the Relational Database

A Data set is a set or collection of data. This set normally presents course information, total no. of. enrolled students and their activity data. Every data column describes a particular variable.

ETL is Extract, Transform and Load. It is a process that extracts the data from different source systems, then

transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system.

3.1.2 Analysis Server

Pre-processing

It is a phase that removing entries that contain errors. Here, anything that is irreverent to the specific research is considered as an error. In an addition, Time is presented in hours, days, weeks, months or years.

Creating Dimensions and the OLAP Cube

After Pre-processing, the OLAP Cube and dimensions are created. Microsoft visual studio is used to create a cube. The Data source and source view must be created before the OLAP cube creation. The data source view has two aims: It identifies the tables and views from the data source that will be used in the concrete Analysis Service Project, and It retrieves and stores the metadata about these objects, allowing for the building of cubes without having to maintain an active open connection to the data source.

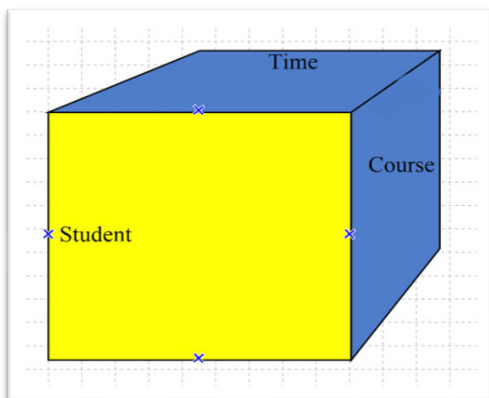


Fig. 3. OLAP cube

Creating Data Mining Models

Data Mining Model having two parts i.e., decision tree and neural network.

1. Decision Tree

A decision tree is used to predict the percentage of occurrence modules with selected input parameters. This is

done through the DMX queries. The input parameters are hour, day, month, and activity.

2. Linear Regression

We use the linear regression algorithm. This method allows us to summarize and study relationship between two continues variables. We predict the value of a variable based on the value of another variable. the variable we are using to predict is called the independent variable and the variable we want to predict is called the dependent variable.

Writing the DMX Query and Test Models

DMX queries is written for obtaining a prediction of behavioural patterns. DMX query language is used for model creation, training, prediction and content access.

```
Select t.*,Predict([Module]) as [Module]
vbalformat(PredictProbability([Module]),'Percent') as (Probability)
from
[dectree]
prediction join
openquery
'select Hour,Day,Month,Activity from table_t') as t
on[dectree].[Hour]=t.Hour and
[dectree].[Day]=t.Day and
[dectree].[Month]=t.Month and
[dectree].[Activity]=t.Activity
```

Fig. 4 shows the Example of DMX queries.

This DMX-in predictable variable is module whose probability is to show. The input variable in this case are hour, day, month and activity.

Neural Network is used to determine the values of attributes. The user provides information on the impact of all attribute values that relate to predictable state.

3.1.3 Web Application

The predictable analysis result is shown on web application. The link between the analysis and the web server applications is accomplished through ADOMD (ActiveX Data Objects Multi-dimensional). The Admin, Institute or Teacher and Students can now directly access all the course wise, time wise progress reports. Admin or Institute can check prediction before the actual implementation begins for any decisions which they will take.

3.2 Implementation of the proposed system prototype

The proposed system is implemented on a server for both the processing and pre-processing of the data, as well as the performing of data mining extensions. Given the number of records to be processed, which exceeds 100 and is constantly increasing, the server should have above-average performance for faster response time.

3.3 Implementation of the described cases

Using the numeric parameter i.e., the results showing both to the input data and the probability of the occurrence of modules within the e-course. The system also provides the possibility of expressing the result in percentages (from 0 to 100). By implementing this system through PDCA approach, an analysis can be conducted for n no. of years (i.e. times t1, t2, and t3, tn). PDCA is implemented in spirals of increasing knowledge of the system that converge on the goal, each cycle closer than the previous.

4. METHODOLOGY & RESULTS

4.1 Possible scenario of use

The following table is showing the possibilities of the proposed system used for various scenarios comparison with the existing system.

Possibilities	Existing System	Proposed System
User Activity Report	Yes	Yes
Real time data analysis	Yes	Yes
Visualisation of Results	Yes	Yes
Interactivity	Yes	Yes

Existence of web-based intelligence	No	Yes
Prediction of user behavioural patterns	No	Yes
Numerical result	No	Yes

Table 2. The possibilities of existing and proposed system

The Teacher enters the input parameter that are relevant to the analysis (e. g. ID, Name of course or student, and time). Based on the input data, the intelligence report system provides graphical representation of the result. The teacher receives display modules and activities associated with percentage probabilities. Then Teacher can decide on the selection module for students in the course depending on result.

Using this proposed system facilitates the process of policy-making and increase the opportunity for more efficient use.

Advantages of the proposed system relate primarily to the opportunities provided by the intelligent features, while a disadvantage is the need for above average performance of the server. The advantages are reflected in the existing web intelligence and the possibility of predicting user behavioural patterns. Furthermore, these advantages have a direct effect on quality assurance in e learning and on the improvement of the teaching process through the adaptation of content by predicting behavioural patterns.

4.2 Characteristics of proposed system

1. Simplicity

The architecture and use of the system are in general simple. The interface is user-friendly and requires no special preparation to use it.

2. Enabling more real time data analysis

Given the fact that the database is constantly updated with new records of user activity, the application automatically performs the same process, based on real time analysis

3. Cost effective

This system is cost effective. The only required costs are related to the server for data storage and processing. Considering the number of users on the system, as well as teachers who use the intelligent system, this cost is manageable.

The Teacher enters the input parameter that are relevant to the analysis (e. g. ID, Name of course or student, and time). Based on the input data, the intelligence report system provides graphical representation of the result. The teacher receives display modules and activities associated with percentage probabilities. Then Teacher can decide on the selection module for students in the course depending on result.

Using this proposed system facilitates the process of policy- making and increase the opportunity for more efficient use.

Advantages of the proposed system relate primarily to the opportunities provided by the intelligent features, while a disadvantage is the need for above average performance of the server. The advantages are reflected in the existing web intelligence and the possibility of predicting user behavioural patterns. Furthermore, these advantages have a direct effect on quality assurance in e learning and on the improvement of the teaching process through the adaptation of content by predicting behavioural patterns.

4.3 Characteristics of proposed system

4. Simplicity

The architecture and use of the system are in general simple. The interface is user-friendly and requires no special preparation to use it.

5. Enabling more real time data analysis

Given the fact that the database is constantly updated with new records of user activity, the application automatically performs the same process, based on real time analysis

6. Cost effective

This system is cost effective. The only required costs are related to the server for data storage and processing. Considering the number of users on the system, as well as teachers who use the intelligent system, this cost is manageable.

1. eb intelligence support features

The system includes a web intelligence related to prediction, which is the main purpose of the system (i.e., obtaining predictive behaviour patterns).

```
# Lets take a look at how well this model preforms
print("The Accuracy is: ", linear.score(X_test, y_test))
coeff = linear.coef_
intercept = linear.intercept_

for i in range(len(coeff)):
    print(maths.columns[i], ': ', coeff[i])
print('The intercept of our slope is: ', intercept)

The Accuracy is: 0.8937549722501006
famrel : 0.27076566247889483
studytime : -0.16957717494611743
failures : -0.33394723148131306
absences : 0.040069346651734214
G1 : 0.14518214242984612
G2 : 0.9975071361433688
The intercept of our slope is: -2.738132927531492
```

Fig. Accuracy using Linear Regression.


```
predictions = grade_classifier.predict(X_test)
predictions

array([[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0])
```

Finally, we can measure the accuracy of our classifier:

```
[ ] accuracy_score(y_true = y_test, y_pred = predictions)

0.7848101265822784
```

Fig. Accuracy using Decision Tree.

2. CONCLUSION

Proposed system is created with data mining techniques leads to a modern and intelligent way of reporting user activity. This system presents an improvement for an existing e-learning system since it predicts behavior patterns thus leading to the increased efficiency of the participants. In PDCA cycle of activities, the following can be concluded (in time t, PtDtCtAt):

- P: This includes the review of existing reports, and the need for identified reports that provide predictions of user behaviour in an LMS.
- D: This includes the design and implementation of the new system and it is consistent with the planned outcomes, adjustment for all activities, and successful functioning of an intelligent, web-based reporting system.
- C: This is the compulsory phase control, carried out through the testing activities, installation and evaluation systems, the execution of change, and the improvement system.
- A: This includes the possible improvement through the re use of system analysis tasks. Here, future work should be related to the creation and implementation of new modules in the given time

t Pt, Dt, Ct, At.

3. REFERENCES

- [1] Castro F, Vellido A, Nebot A, Mugica F. Applying data mining techniques to e-learning problems. <<http://www.sci2s.ugr.es/keellpdf/specific/capitulolApplingDataMiningTechniques.pdf>>[accessed 10.01.121.]
- [2] Zaiane O, Luo J. Web usage mining for a better web-based learning environment. In: Proceedings of conference on advanced technology for education. Banff, Alberta; 2001. p. 60-4
- [3] Romero C, Ventura S. Data mining in e-learning. Southampton, UK: Wit Press; 2006.
- [4] Klossgen W, Zitzkow J. Handbook of data mining and knowledge discovery. New York: Oxford University Press; 2002.
- [5] Pollard H, Jordin R, Abdallah C, Levine Blevins M. A new computer system for education. Comput Electr Eng 1991;17:261-6.
- [6] Agah A. Human interactions with intelligent systems: research taxonomy. Comput Electr Eng 2000; 27:71-1.