

Automated Assessment Generation and Student Performance Prediction In E-Learning System

Vedashree Lahoti^{1*}, Dr. Anjali Raut²

^{1,2}Department of Computer Science and Engineering, HVPM College of Engineering and Technology, Amravati, India

ABSTRACT

Article Info

Volume 8, Issue 2

Page Number : 58-65

Publication Issue :

March-April-2022

Article History

Accepted: 05 March 2022

Published: 20 March 2022

In this project we proposed an automatic student performance and assessment generation models in e-learning using machine learning algorithms. Our proposed model will find out Students performance by using their behavior. Behavioral data like study material searching, video accessing time, and submission dates, assignment marks, question asking behavior etc. will be tracked into the database. In e-learning assessment generation is a very important and time-consuming activity for teachers. Therefore, to solve this issue we proposed automatic assessment generation model which will use Formal Concept Analysis algorithm. Formal Concept Analysis algorithm will be used to extract knowledge from the question-answers. A Learning Management System (LMS) is an application software that plays a vital role in educational technology. Such software can be designed to augment and facilitate instructional activities including registration and management of education courses, analyzing skill gaps, reporting, and delivery of electronic courses simultaneously.

Desired Implication

- To develop an e learning system
- To implement Automatic Assessment model
- To implement Decision Tree algorithm for student performance prediction
- To implement Formal Concept Analysis algorithm for question

Keywords: Learning Management System, Study Material Searching, Video Accessing Time

I. INTRODUCTION

The advancement in information technology has enhanced the effectiveness of web-based education (e-learning) system. The e-learning system allows students from anywhere and at any time to have access and carry out different learning activities such as reading, downloading and uploading documents, presentations and assignments. The activities normally take place in a platform called Learning Management System (LMS), which utilizes different technology mostly on the internet which facilitate the student's communication between students and tutors or among the students. The LMS provides a huge volume of varied data for usage by the students and the instructors as well, it also contains data about the student's personal data, learning styles, practices, behaviors and usability preferences addition, Massive open online courses (MOOCs) empower students around the world to gain from top instructive substance with ease. In e learning students refer study materials, solve assignments, and submit their subjective assignments or practical within specified submission date. We will track these activities to calculate student performance.

To calculate performance we proposed decision tree algorithm.

Decision Tree Algorithm

This algorithm will be used to calculate student's performance. Decision Tree algorithm belongs to the family of supervised learning algorithms. decision tree algorithm can be used for solving **regression and classification problems** too. The motive of using Decision Tree is to create a training model which can use to predict class or value of target variables by **learning decision rules** inferred from training data. The understanding level of Decision Trees algorithm is so easy compared with other classification algorithms. The decision tree

algorithm tries to solve the problem, by using tree representation. Each **internal node** of the tree match to an attribute, and each **leaf node** match to a class label.

The assessments required to find out performance of the students should be prepared in such a way that all the subtopics will be covered. In this project we proposed automatic assessment generation module. This model will make use of FCA algorithm to explore the knowledge about the questions which are stored in database.

Formal concept analysis (FCA) is a principled way of deriving a *concept hierarchy* or formal ontology from a collection of objects and their properties. Each concept in the hierarchy represents the objects sharing some set of properties; and each sub-concept in the hierarchy represents a subset of the objects (as well as a superset of the properties) in the concepts above it

II. LITERATURE REVIEW & RELATED WORK

In six types of features were extracted from click-stream logs within an MOOC and used to predict students' grades of next assessments. Brinton et al. [10] developed frameworks to extract event-based and position based sequences from student video-watching click streams in MOOC. Their experiments demonstrated that video-watching behavior can help improve student performance prediction. Due to declining participation over time in MOOC [11], Jiang et al. [12] utilized students' assessment performance and social data in week 1 to predict students' certificate obtaining.

In addition to MOOC, there is research on the prediction of student performance in SPOC. The work [13] developed a linear regression model and a deep learning model to predict student performance in SPOC, and it stated that the model can be generalized for MOOC or other online

learning. Wan et al. [14] used logical regression to predict the weekly test pass of students in SPOC. Marcos et al. [16] used features related to platform visiting and interactions with videos and quizzes to predict whether a student can pass the admission test, providing enlightenment for student performance prediction in SPOC-based blended learning.

Some studies predicted student performance in blended learning by using data directly from the learning management system (LMS). Raga and Raga [6] developed a deep neural network model for early prediction of student performance in blended learning.[7] developed linear and nonlinear prediction models based on the pedagogical types of blended learning. [8] used the number of online course access, coverage of digitally provided learning material, and the differences in student sequences with their golden standard to predict the online score.

Existing Question Paper Generation Systems

The research paper “Framework for Automatic Examination Paper Generation System” has provided a thorough insight into the process of automated paper generation [3]. As the manual generation of a balanced question paper by an individual is complex, the blending of technology into teaching and learning process is unavoidable. A simple way for an examination paper generation is provided. Generation of Examination Papers is governed by the Syllabus Engine, Pattern Composer and Question Aggregator.

The generated question paper is based on the pattern of the course. Another component called Bank Management takes care of User Rights and Privilege assignment. Questions are entered through the Question Aggregator. The attributes related to questions are type, marks and complexity. All these attributes are efficiently used during Question Paper Generation [3].

The paper generator selects a question according to the pattern and complexity. This engine also introduces marking systems wherein any selected question is marked so that it might not be selected again and then generated papers are stored as pdfs [3].

The paper on “Automatic Question Paper Generation System using Randomization Algorithm” describes a system which uses a shuffling algorithm (existing algorithm) as a randomization technique [4]. The system defines several modules like user administration, subject selection, difficulty level specification, question entry, question management, paper generation, and paper management. Also, mathematical formulae and diagrams is integrated in questions. The system includes a dual interface viz., web-based and desktop-based [4]. The system introduces a really efficient shuffling algorithm which uses an array to store randomly generated numbers. The questions are then selected against these array elements, so ensuring completely random generated question papers. However, this methods fails to utilize the highly efficient marking system. So, questions once selected is also repeated in previous papers. This is often limitation of this method[4].

The “Question Paper Generator System” has provided a able to use built-in question bank [1]. The paper aptly describes CQZ (Cloze Question Generation) putting more emphasis on the particular style of the questions [1].

Another paper on “Automatic Test Paper Generation supported Ant Colony Algorithm” has implemented a fancy but highly efficient Ant Colony Algorithm [5].It requires building of a mathematical model of constraint according to the per wants of the paper. This paper provides an efficient solution with their algorithm [5]. The paper on “An Integrated Automated Paperless Academic Module for Education Institutes” has stated the importance and

dealing of switching from Paper based systems to Paperless Systems [2]. The importance of automation is extremely well documented within the context of Task Engineering [2]. The paper defines the importance of data and Communication Technology (ICT) in academics and academic organizations [2].

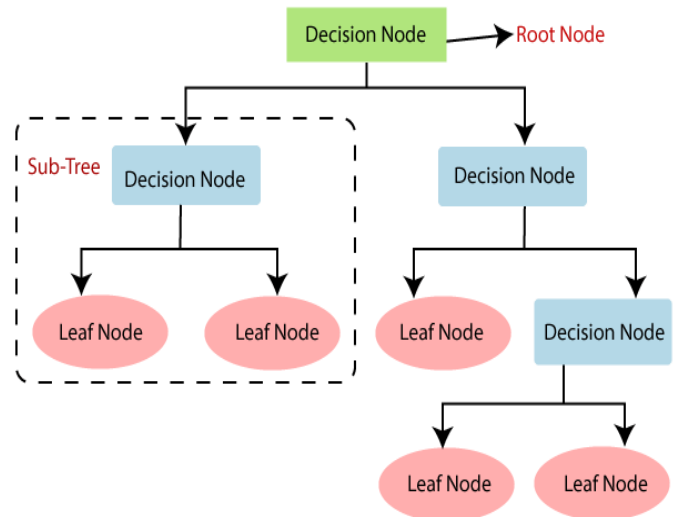
II. PROPOSED WORK

In this project we proposed e-learning system. In this e-learning our system will track student's activities like attendance, study material accessing time, total no of tests and attempted tests, test marks, assignment submission date etc. Depending upon these attributes our performance detection module will detect the student's performance. We proposed decision tree algorithm for student performance prediction. Along with this we proposed automatic assessment generation model. Staff members will register topics wise questions. System will detect key phrases of registered questions and their answers using RAKE algorithm. Extracted key phrases will be given as input to Formal Concept Analysis algorithm to detect concept hierarchy of questions. Output of FCA module will be sent to Automatic Assessment Generation Model.

Decision Tree Algorithm

- Decision Tree is a Supervised learning technique which will be used for both classification and Regression problems, but it is preferred for solving Classification problems. It's a tree-structured classifier, where internal nodes consider the features of a dataset, branches consider the decision rules and every leaf node consider the end result.
- In a Decision tree, total two nodes, which are the Decision Node and Leaf Node. Decision nodes are wont to develop any decision and have multiple branches, whereas Leaf nodes are the output of these decisions and don't contain any extra branches.

- The decisions or the test are performed on the idea of features of the given dataset.
- It could be a graphical representation for getting all the possible solutions to a problem/decision supported given conditions.
- It's a name decision tree because, same as a tree, it starts with the root node, which expands on more branches and constructs a tree-like structure.



Why use Decision Trees?

There are various algorithms in Machine learning, so choosing the most effective algorithm for the given dataset and problem is that the main point to remember while creating a machine learning model. Below are the 2 reasons for using the Decision tree:

- Decision Trees usually mimic human thinking ability while making a decision, so it's easy to grasp.
- The logic behind the decision tree are often easily understood because it shows a tree-like structure.

Decision Tree Terminologies

- **Root Node:** Root node is from where the decision tree starts. It represents the complete dataset, which further gets divided into two or more homogeneous sets.
- **Leaf Node:** Leaf nodes are the ultimate output node, and also the tree can not be segregated further after getting a leaf node.
- **Splitting:** Splitting is that the process of dividing the decision node/root node into sub-nodes

according to the given conditions.

- **Branch/Sub Tree:** A tree formed by splitting the tree.
- **Pruning:** Pruning is that the process of removing the unwanted branches from the tree.
- **Parent/Child node:** The root node of the tree is termed the parent node, and other nodes are called the child nodes.

How does the Decision Tree algorithm Work?

In a decision tree, for predicting the category of the given dataset, the algorithm starts from the root node of the tree. This algorithm compares the values of root attribute with the record (real dataset) attribute and, supported th comparison, follows the branch and jumps to the following node.

Python Implementation of Decision Tree

Now we'll implement the Decision tree using Python. For this, we are going to use the dataset "user_data.csv," which we've got employed in previous classification models. By using the identical dataset, we are able to compare the Decision tree classifier with other classification models like KNNSVM, Logistic Regression, etc.

Steps will remain the identical, which are given below:

- Data Pre-processing step
- Fitting a Decision-Tree algorithm to the Training set
- Predicting the test result
- Test accuracy of the result(Creation of Confusion matrix)
- Visualizing the test set result.

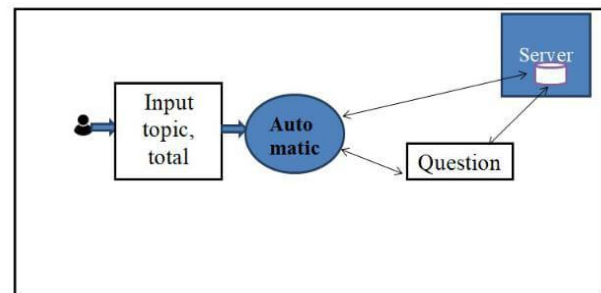
Formal Concept Analysis

Formal Concept Analysis (FCA) could be a method mainly used for the analysis of knowledge, i.e. for deriving implicit relationships between objects described through a group of attributes on the one

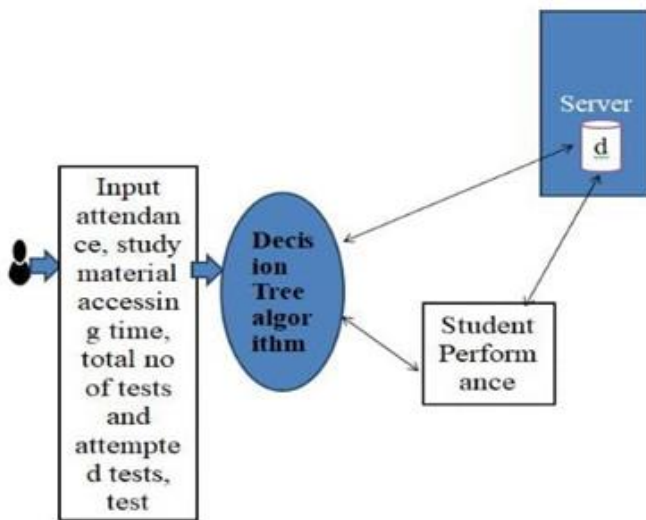
hand and these attributes on the opposite. The info are structured into units which are formal abstractions of concepts of human thought, allowing meaningful comprehensible interpretation. Thus, FCA may be seen as a conceptual clustering technique because it also provides intentional descriptions for the abstract concepts or data units it produces. Central to FCA is that the notion of a proper formal context.

We use FCA for automatic assessment generation. Each questionnaire consists of inquiries to be answered by respondents mostly selectively from possible answers. From the purpose of view of FCA, we are able to take the set of respondents because the set of objects and single questions as attributes.

The questions need't be yes/no questions. Rather, some questions like those concerning age and education are many-valued. Correspondingly, a filled-in questionnaire will be represented by a many-valued context. Such context will be transformed to a standard formal context $\langle X, Y, I \rangle$ via conceptual scaling.



Student Performance Prediction Working



Modules

- **Admin Panel**
 - Login
 - Create staff login
 - View staff details
 - Approve student account
 - View student performance
 - Register courses and course wise subjects
 - Allocate subjects to staff
- **Staff Panel**
 - Login
 - View allotted subjects
 - Register topics
 - Register questions
 - Register new sessions and upload videos
 - view generated assessments
 - upload study material
- **Student Panel**
 - Registration
 - Login
 - View course syllabus
 - View sessions
 - View study materials
 - Solve assignments online
 - View own performance
 - View marks
- **Automatic Assessment Generation model**

1. Generate concept hierarchy store in db using FCA
 2. Input topic, total marks, question wise marks distribution details, subtopics
 3. Input n =total no of questions
 4. Fetch questions related to the specified topics and subtopics by using concept hierarchy
 5. Check whether the fetched question is already present in question paper or not
 6. if question already present then fetch another question else add question in question paper
 7. decrement n ie $n--$
 8. repeat steps 3 to 5 till $n > 0$
 9. check whether all the subtopics are covered or not
 10. if all topics covered then stop otherwise revised the question paper
- **Student Performance Prediction**
 - Track students activities
 - Generate input as per the dataset
 - Predict student performance using decision tree algorithm

III. CONCLUSION

Our system has deployed an efficient algorithm which is totally randomized and avoids repetition of questions in consequent question papers, making it impossible to derive any pattern within the papers. The resultant automated system for Question Paper Generation provides improvement in terms of controlled access to the resources, random generation of question papers and a secure platform. Predicting Students Performance in e-learning and MOOCs could be a vital angle which helps instructors in enhancing the educational and educating process. The bulk of the researchers have utilized records got from log documents of various e-learning and MOOCs flat forms, While for prediction techniques, the classification method is employed in educational data processing area, the classification techniques, Naïve Bayes, Decision Tree

Nearest Neighbor and Neural Network, Multiple Linear Regression, Support Vector Machine are methods that are commonly used by the researchers for predicting students' performance, as they have been utilized by [15] in predicting student performance, [12] also in predicting student performance in virtual environment, [10] in predicting students' performance as well.

III. REFERENCES

- [1]. Surbhi Choudhary, Abdul Rais Abdul Waheed, Shrutika Gawandi and Kavita Joshi, "Question Paper Generator System," *International Journal of Computer Science Trends and Technology*, vol. 3, issue 5, Sept – Oct 2015.
- [2]. Prita Patil and Kavita Shirsat, "An Integrated Automated Paperless Academic Module for Education Institutes," *International Journal of Engineering Science Invention Research and Development*, vol. I, issue IX, March 2015.
- [3]. Ashok Immanuel and Tulasi.B, "Framework for Automatic Examination Paper Generation System," *International Journal of Computer Science Trends and Technology*, vol. 6, issue 1, Jan- March 2015
- [4]. Kapil Naik, Shreyas Sule, Shruti Jadhav and Surya Pandey, "Automatic Question Paper Generation using Randomization Algorithm," *International Journal of Engineering and Technical Research*, vol. 2, issue 12, December 2014.
- [5]. Dan Liu, Jianmin Wang and Lijuan Zheng, "Automatic Test Paper Generation Based on Ant Colony Algorithm," *Journal of Software*, vol. 8, no. 10, October 2013
- [6]. R. Raga and J. Raga, "Early prediction of student performance in blended learning courses using deep neural networks," in *Proc. IEEE Int. Symp. Educ. Technol. (ISET)*, Hradec Kralove, Czechia, 2019, pp. 39–43.
- [7]. J. H. Kim, S. Seodaemun-gu, Y. Park, J. Song, and I.-H. Jo, "Predicting students' learning performance by using online behavior patterns in blended learning environments: Comparison of two cases on linear and non-linear model," *Korea*, vol. 120, p. 750, 2014.
- [8]. A. Figueira, "Predicting grades by principal component analysis: A data mining approach to learning analytics," in *Proc. 16th Int. Conf. Adv. Learn. Technol. (ICALT)*, Austin, TX, USA, 2016, pp. 465–467.
- [9]. Violante, M. G., & Vezzetti, E. (2015a). Design and implementation of 3D Web-based interactive medical devices for educational purposes. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 1-14.
- [10]. C. G. Brinton, S. Buccapatnam, M. Chiang, and H. V. Poor, "Mining MOOC clickstreams: Video-watching behavior vs. in-video quiz performance," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3677–3692, Jul. 2016.
- [11]. S. Jiang, A. Williams, K. Schenke, M. Warschauer, and D. O'Dowd, "Predicting MOOC performance with week 1 behavior," in *Proc. 7th Int. Conf. Educ. Data Mining*, London, U.K., Jul. 2014, pp. 273–275.
- [12]. C. Yu et al., "SPOC-MFLP: A multi-feature learning prediction model for SPOC students using machine learning," *J. Appl. Sci. Eng.*, vol. 21, no. 11, pp. 279–290, 2018. [Online]. Available: <https://doi.org/10.1787/f8d7880d-en>
- [13]. H. Wan, J. Ding, X. Gao, Q. Yu, and K. Liu, "Predicting performance in a small private online course," in *Proc. Educ. Data Mining*, 2017, pp. 384–385.
- [14]. R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, "Predicting student performance from LMS data: A comparison of 17 blended courses using moodle LMS," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 17–29, Jan.–Mar. 2017.

- [15]. P. M. Moreno-Marcos et al., "Predicting admission test success using SPOC interactions," in Proc. 9th Int. Conf. Learn. Anal. Knowl., Tempe, AZ, USA, 2019, pp. 4–8.
- [16]. walia, N.; Kumar, M.; Nayar, N.; Mehta, G. Student's Academic Performance Prediction in Academic using Data Mining Techniques. In Proceedings of the International Conference on Innovative Computing & Communications (ICICC); Springer: Berlin/Heidelberg, Germany, 2020.
- [17]. OuahiMariame, S.K. Feature Engineering, Mining for Predicting Student Success based on Interaction with the Virtual Learning Environment using Artificial

Cite this article as :

Vedashree Lahoti, Dr. Anjali Raut, "Automated Assessment Generation and Student Performance Prediction In E-Learning System", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 2, pp. 58-65, March-April 2022. Available at doi : <https://doi.org/10.32628/CSEIT228213>
Journal URL : <https://ijsrcseit.com/CSEIT228213>