Designing Data Mining Applications with Rough Set Algorithm for Provision of Recommendations in the Selection of Training Topics on Online Learning

Hotler Manurung1, Erianto Ongko1, Astro Julida Harahap1, Hamzah Ahmad1, D. A. Abdullah3, Cut Ita Erliana1, S Sriadhi1, Aditya Halim Putra1, P. Kurnia Putra1, Muslim7, Hamzah Ahmad1, Ricardo Freedon7, Nanuru8, A Saleh A9, Asmara Indahingwati10, Citra Kurniawan11, Ida Bagus W. Indrarta12, A B. Totrozzaq Hasibuan13, Eni Wuryani14, Wiwien Hadikurniawan15, and Edy Winarno15

1Department of Informatics, Akademi Teknologi Industri Immanuel, Indonesia
2Department of Computer Science, STMIK IBBI, Medan, Indonesia
3Department of Informatics, Universitas Malikussaleh, Aceh, Indonesia
4Department of Industrial Engineering, Universitas Malikussaleh, Indonesia
5Department of Electrical Engineering, Universitas Negeri Medan, Indonesia
6Department of Management, STMIK Bina Darma, Bangka, Indonesia
7Department of Accounting, Universitas Islam Sumatera Utara, Medan, Indonesia
8Universitas Kristen Indonesia, Malukutuban, Indonesia
9Statistics, Financial, & Social Sciences Research Group, Indonesia
10Manajemen School of Economic & Management, "STIESIA" Surabaya, Indonesia
11Department of Electrical Engineering, Sekolah Tinggi Teknik Malang, Malang, Indonesia
12Informatics Department, STMIK STIKOM, Indonesia
13Faculty of Engineering, Universitas Islam Sumatera Utara, Medan, Indonesia
14Department of Accounting, Universitas Negeri Surabaya, Surabaya, Indonesia
15Faculty of Information Technology, Universitas Stikubank, Semarang, Indonesia
*dahlan@unimal.ac.id

Abstract

Online Learning is a learning model that utilizes internet network. As one of the institutions providing vocational education, ATI Immanuel is planned to implement online learning for improving the quality of education. Help learners in choosing learning topics. The concept of data mining itself is a process of analysis of the database of students learning ATI Immanuel which has been collected and will be obtained patterns of student behavior in learning. The results of the study are expected to help online learning participants in this case ATI Immanuel students so they can choose the best training topics. Data mining applications with this rough set algorithm will be designed using PHP Programming Language and using System Development Life Cycle (SDLC) design method.

1. Introduction

Online learning is a type of teaching and learning process that allows the delivery of teaching materials to students by using internet media, intranet or other computer network media[1]. In online
learning, learning participants need to be more responsible for the learning outcomes that take place independently through a good planning and implementation process[2]. The online learning process needs to be developed to ensure equality of access to quality education to ensure a sense of justice[3]. Data mining applications with rough sets algorithm are expected to help the learning participants in choosing the training topics. Data mining is the process of finding useful knowledge in a large-scale database[4]. In the development of online learning applications need to pay attention to the quality of learning. Therefore, a benchmarking process is required[5]. The benchmarking process is also expected to ensure efficient implementation of online learning[6].

The rough set algorithm itself is a data mining algorithm that represents a set of data in the form of a table, in which rows in a table represent objects and columns representing attributes of objects. The attributes of the object itself can be distinguished into attribute conditions and decision attributes[7]. The implementation of this rough set application can make it easier for students to take decisions on the courses taken[8]. In the process of determining these subjects it is necessary to group subjects of existing subjects with a good classification method of the learning data in the database ensuring that all relevant information can be collected[9]. This is necessary to increase data diversity[10] and data sensitivity[11].

2. Related Works

Liu et al. suggested strengthening online learning to predict emotions using physiological signals[12]. Asir has proposed a method to identify students’ readiness for online learning, to test their merits and perceptions and to measure the quality of online tutorials[13]. Blended learning provides added pedagogical value compared to on-line learning systems of teaching college nursing skills in clinical supervision[14]. Bousbahi and Chorfi used the Case Based Reasoning (CBR) method to recommend appropriate training topics in the Massive Open Online Course (MOOC) lesson[15]. Aher and Lobo use a combination of data mining algorithms K-Means Clustering and Apriori Association Rule to generate recommendations for the selection of training topics on learning MOOC. The combination of these algorithms will result in recommendations of whether or not someone may take a training topic[16].

3. Research Methodology

The research methodology can be seen in Figure 1. Through figure 1 can be seen that in the design of data mining applications with rough set algorithm can be seen that what needs to be done is a database analysis to gain knowledge from past learning. Then based on the analysis of this database rough set algorithm will generate learning recommendations when students enter in the application of online learning.

4. Results and Discussion

4.1. Data Mining System

The data mining system can be seen in Figure 2. In Figure 2 it can be seen that the data mining system has a user interface that interacts with the user. The process of data mining is basically to analyze the pattern or interesting knowledge of a large-scale database using data mining engine that uses a method. One of the methods that can be used is the rough set method.

4.2. Rough Set Algorithm

The rough set algorithm was developed by Z. Pawlak in 1982 as the set of theories of an intelligent system. The rough set algorithm starts from the assumption that each object is associated with a knowledge. Two objects are expressed as "indistinguishable" if they provide the same information[17]. The Algorithm of Rough Set can be described as follows.

1. Data Selection (Selection of resources to be used)
2. Establishment of Decision System which contains attribute condition and decision attribute.
3. Establishment of Equivalence Class, ie by eliminating repetitive data.
4. The formation of Discernibility Matrix Modulo D, ie a matrix that contains comparison between different data attribute conditions and decision attribute.
4.3. SDLC Method
SDLC Method can be seen in Figure 3.

5. Conclusion
The conclusions that can be obtained from the results of this study are as follows. First, Online Learning is needed to improve the quality of learning. Second, Data mining can be used to analyze the learning patterns that exist in the database. Third, using the rough set algorithm can be recommended to students on learning topics that match their abilities.

References
[1] Hartley D E 2001 Selling E-Learning Association for Talent Development
[2] You J W 2016 Identifying significant indicators using LMS data to predict course achievement in online learning The Internet and Higher Education 29 23-30
[3] Alesyanti A, Ramlan R, Hartono H and Rahim R 2018 Ethical decision support system based on hermeneutic view focus on social justice International Journal of Engineering & Technology 7 (2.9) 74-77
[4] Han J, Kamber M and Pei J 2011 *Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems)* Morgan Kaufmann

[5] Abdullah D, Tulus, Suwilo S, Effendi S and Hartono 2018 DEA Optimization with Neural Network in Benchmarking Process *IOP Conference Series: Materials Science and Engineering* **288** 012041

[6] Abdullah D, Tulus, Suwilo S, Effendi S, Hartono and Erliana C I 2018 A Slack-Based Measures for Improving the Efficiency Performance of Departments in Universitas Malikussaleh *International Journal of Engineering & Technology* **7**(2) 491-494

[7] Chen L F and Tsai C T 2016 Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain *Tourism Management* **50** 197-206

[8] Simanihuruk T *et al.* 2018 Hesitant Fuzzy Linguistic Term Sets with Fuzzy Grid Partition in Determining the Best Lecturer *International Journal of Engineering & Technology* **7**(3) 35-39

[9] Hartono, Sitompul O S, Tulus and Nababan E B 2018 Optimization Model of K-Means Clustering Using Artificial Neural Networks to Handle Class Imbalance Problem *IOP Conference Series: Materials Science and Engineering* **288** 012075

[10] Hartono, Sitompul O S, Nababan E B, Tulus, Abdullah D and Hartono A S 2018 A New Diversity Technique for Imbalance Learning Ensembles *International Journal of Engineering & Technology* **7**(2) 478-483

[11] Hartono, Sitompul O S, Tulus T and Nababan E B 2018 Biased support vector machine and weighted-smote in handling class imbalance problem *International Journal of Advances in Intelligent Informatics* **4**(1) 21-27

[12] Liu W, Zhang L, Tao D and Cheng J 2018 Reinforcement online learning for emotion prediction by using physiological signals *Pattern Recognition Letters* **125** 125-130

[13] Asiry M A 2017 Dental students’ perceptions of online learning *The Saudi Dental Journal* **29**(4) 167-170

[14] McCutcheon K, Halloran P O and Lohan M 2018 Online learning versus blended learning of clinical supervisee skills with pre-registered nursing students: A randomised controlled trial *International Journal of Nursing Studies* **82** 1-9

[15] Bousbahi F and Chorfi H 2015 MOOC-Rec: A Case Based Recommender System for MOOCs *Procedia - Social and Behavioral Sciences* **195** 1813-1822

[16] Aher S B and Lobo L M R J 2013 Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data *Knowledge-Based Systems* **51** 1-7

[17] Huang C C, Seng T T and Tang C Y 2016 Feature extraction using rough set theory in service sector application from a managerial perspective *Computers & Industrial Engineering* **91** 30-41