Student Perceptions and Achievements of Online Learning: Machine Learning Approaches

Hari Suparwito

Department of Informatics, Sanata Dharma University, Yogyakarta, Indonesia

shirsj@jesuits.net

Abstract. The Covid-19 pandemic is currently occurring affects almost all aspects of life, including education. *School From Home* (SFH) or online learning has been selected as one of the ways to prevent the spread of Covid-19. The face-to-face learning method in class turns into online learning using information technology facilities. Even though there are many barriers to implementing a class online, online learning provides a new perspective for student learning. The study's purpose was to analyze student perceptions of the online learning process. The research data were obtained from a student questionnaire. Students provided an assessment through a questionnaire about the online learning methods they experienced during the Covid-19 pandemic. The machine-learning algorithm was applied to examine the dataset. The study focused on the criteria (variable importance) that affect student perceptions of the online learning process. The results described that the student achievement in online learning is influenced by 1) technology to access online learning. 2) student efforts 3). Active and independent learning. The study contributes to improving the online learning method for the student.

INTRODUCTION

The Covid-19 pandemic has affected almost all aspects of life, such as health, economy, social, and education are no exception. Notably, in education, schools are closed, and an online learning process is implemented. Even though many schools and universities have used online learning in the daily learning process, this condition challenges and must be faced by lecturers, students, and all stakeholders. Some of the challenges faced include:

a) Changing face-to-face form in the classroom to digital form (e-learning). Lecturers are expected to deliver material online effectively so that students could understand well.
b) Online learning has never been imagined or anticipated for Lab classes.
c) The issues on the use of information technology are challenging.
d) It requires internet access if the learning is carried out through video conferencing. On the student side, this also has financial consequences.

Therefore, learning methods need to be continuously developed, either in-class or online, by analyzing student perceptions. A study to determine student satisfaction in online learning using a questionnaire has done by Wu et al. [1]. The interaction of lecturers and students is an essential point in the questionnaire. Thus, the person in charge of education can determine whether the entire learning process can help students develop themselves and obtain adequate knowledge [2].

Input from student perceptions became essential data and was used as helpful information to develop the learning process. Thus, a more comprehensive picture will be obtained to develop the learning process in class and online [3]. Therefore, it is necessary to analyze the collected data. This data is in the form of student track records and learning processes. Analyzing and evaluating large amounts of data requires tools, i.e., data mining and machine learning, to provide more accurate information [4]. The Data Mining and Machine Learning algorithm approaches can provide an overview of the learning process based on student questionnaire data. Several machine learning algorithms can be used, such as Random Forest, Support Vector Machine, and Artificial Neural Networks.

A study conducted by Ali [5] showed that data mining applied in the education sector could be used for various research, especially in prediction, classification, clustering, and associations associated with the learning process. Data were stored, processed, and analyzed to become information or knowledge, known as Data Mining [4]. Data analysis relating to the World of Education and its results to provide input to the learning process development is known as data mining in education [6]. Some researchers also mentioned the role of data mining and machine learning as a research tool in the field of education, for example, to predict applicants [7], study program selection [8], performance lecturers [9], and curriculum development [2]. Kushik et al. studied student performance in the online learning process. The research focused on online test problems. The machine learning approach method is used to predict student performance in the online learning process. The research results showed that choosing learning strategies can significantly increase student test scores [10].

The study focused on two things. The first is to determine which potential factors influence the student semester achievement index. The second is whether the machine learning method could predict student semester achievement index based on their perception of online learning. Random Forest (RF) and Support Vector Machine (SVM) algorithms are proposed to analyze the dataset from the student questionnaire. We examined three criteria: students' perceptions of online lectures, the roles of students and lecturers, and the relation to access and use of information technology.

By knowing student perceptions of the online learning process, the stakeholders could obtain helpful information for improving the online learning process. The results provide a new perspective on data mining and machine learning as tools in the education area. Therefore, to design a better learning process following lecturer and student expectations, we proposed that these study results could contribute to online learning development.

METHODS AND DATA COLLECTION

To evaluate whether the online learning process is running well, we conducted a survey to determine the challenges and difficulties of online learning. The survey includes five main elements considered necessary in online learning, i.e., self-management aspects, student efforts, technology utilization, student perceptions of self-role, and perceptions of lecturer role. Moreover, in this study, the five primary elements were examined further and focus more on three main student perceptions on their learning using an online method. The three student perceptions are 1) Student perceptions of online lectures. 2) Student and lecturer roles in online learning, such as student independence on learning, learning methods from lecturers, and time management. 3) Access to and use of information technology.

Data were collected from 8531 student surveys from the Quality Assurance and Internal Audit Institution (LPMAI) Sanata Dharma University Yogyakarta. From December 2020 to February 2021, students could provide their opinion and perception of the online learning method. The data consist of 33 questions on a Likert scale (1-4) and Yes-No questions. The 33 questions could be grouped into three main criteria, i.e., the student perceptions, the student and lecturer preparation on online learning, and how to access and use information technology on online learning. The respondents come from eight departments: Business, Pharmacy, Education, Postgraduate, Psychology, Science and Technology, Literature, and Theology. However, in this study, we only used questionnaires from Science and Technology students with 1931 respondents. As the target label, we added the student semester achievement index.

Data were processed following the steps in Knowledge Data Discovery. The raw data in the form of a questionnaire were modified and restructured so that Data Mining tools could read it. This study applied Random Forest and Support Vector Machine algorithms using R programming to analyze data. The data processing results were focused on student perceptions of the online learning process, which includes three main criteria in the questionnaire.

To analyze data, in the beginning, we have undertaken pre-processing data such as examining the missing values, scaling data, and lastly separating data into training and testing datasets in 70:30 portions. Furthermore, we performed two main processes. The first process is to find the variable importance. In this process, we applied the Random Forest algorithm to obtain three groups of variable importance, i.e., important, tentative, and unimportant attributes. The second process is to use the crucial attributes to predict the student semester achievement index. We have carried out Random Forest and Support Vector Machine regression. The RMSE (Root Mean Square Error) value predicts the testing dataset's dependent variable. A low RMSE value indicates a similarity between the predicted data and the actual data. Next, to find out variable importance from the validation dataset. Variable importance is needed to provide an overview of the independent variables that have the most influence in determining the RMSE value. In terms of this study, if we want to understand the online learning implementation better, then the importance of variables must focus on change.

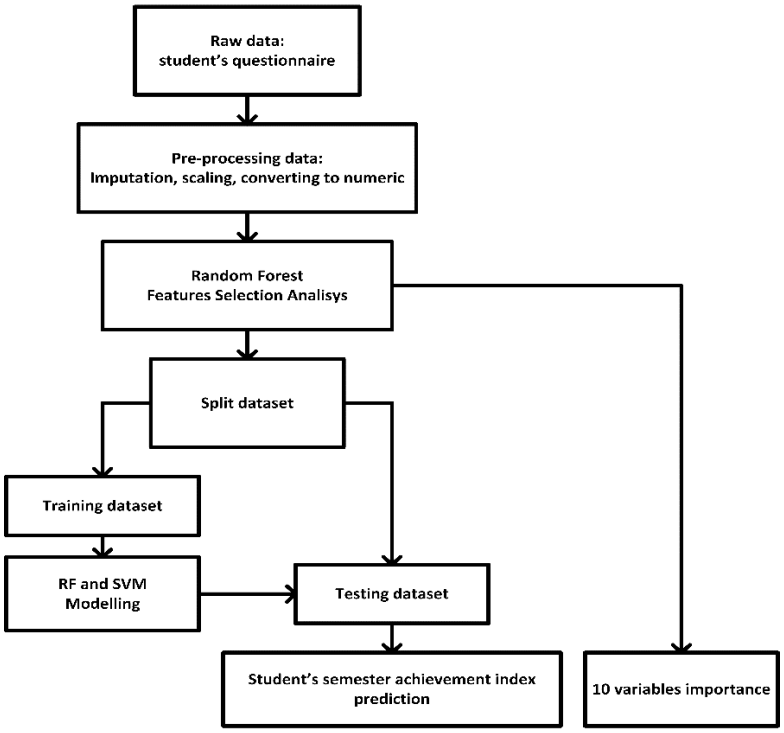

Figure 1. The methodology used in the research.

RESULTS AND DISCUSSIONS

Firstly, we performed variable importance analysis using the Random Forest algorithm. Figure 2 showed the 10-best attributes (green color). They have a significant factor in predicting the student semester achievement index.

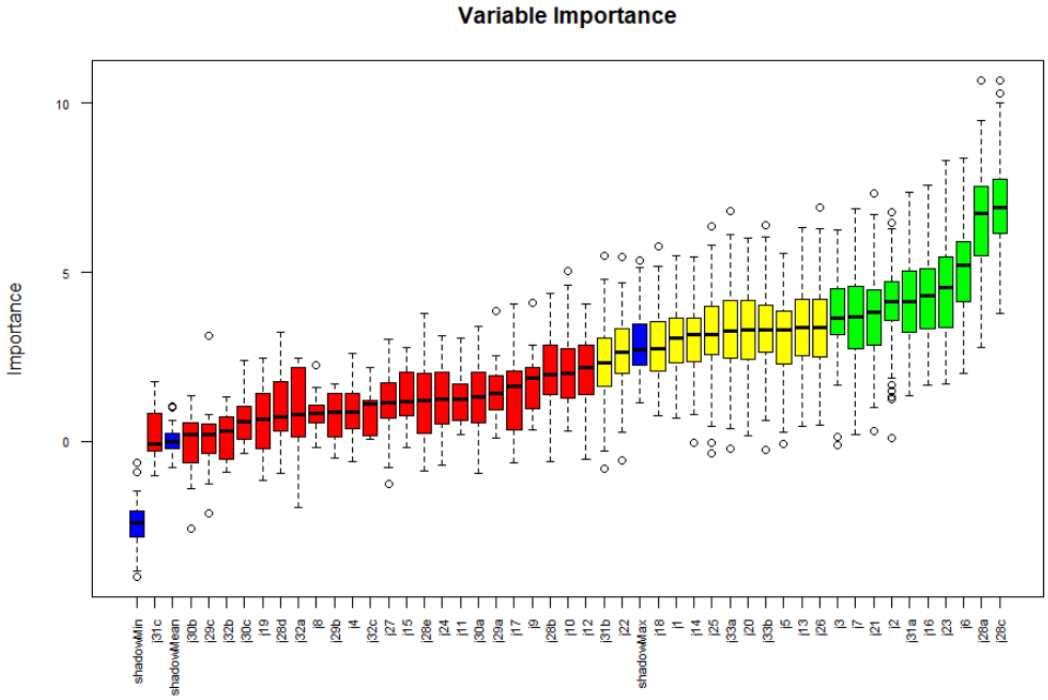

FIGURE 2. Variable importance. The Green color shows the significant attributes, the yellow color is a tentative variable, and the red color is unimportant. In this study, we used only the significant variable that is represented in green color.

Table 1 described what factors influence the student semester achievement index. We eliminated the tentative and unimportant attributes as a part of the feature selection process.

TABLE 1. The ten variables' importance.

Number	Variable importance
1	The recorded lecture on YouTube is the best tool for learning.
2	The Learning Management System is beneficial
3	Students need more time to learn
4	Students do not encounter any significant problems when they discuss or work in groups.
5	The features used in LMS make students easier to learn.
6	Students need a good internet access
7	Online learning makes students more independent
8	Online learning makes students more enthusiastic to learn
9	Students could explore a new knowledge easily
10	Students are enthusiastically attending the class even though it is online using Zoom

In general, we could do grouping these ten important attributes into three significant factors on student perceptions to make successful online learning, namely, the use of technology (1, 2, 5, and 6), student efforts (3, 4, and 9), and creating independent learning conditions (7, 8, and 10).

It is believed that technology is a crucial aspect of online learning. Therefore, respondents put this aspect into first and second vital things for successful online learning, followed by using various technological tools for online learning. Students also know how to use technology tools for online learning (e.g., handphone, computers, Google Classroom, Zoom, WA). This is not separated from the fact that information technology is becoming a mediator, which has an increasingly significant role in this online learning mechanism.

However, Engeström [11] claimed that tools refer to two levels: primary tools (cellphones, computers, internet, books, study rooms) and secondary tools (experiences, language skills, ideas, hopes, motives). Moreover, activity

systems in the context of online learning are mediated by five other primary elements, i.e., tools, rules, community, division of labor, and object. Technology becomes an integral part of the learning experience and an essential consideration for students from preparing student learning experiences. The vital role that technology plays in education allows teachers to design meaningful learning experiences that embed technology. This is not a new area for students; teachers should have considered the tools and resources that can best support learning activities for students. However, advances and accessibility of technologies have made the possibilities seem almost endless [12]. The lack of access to information technology due to unsupported equipment and the internet creates frustration, decreasing student motivation, and self-confidence [13]. Even though technology becomes the primary tool, however, online learning is the learning process revolution enabled by the new technologies that will present an effective and efficient learning process that does not exist before. Learning Management Systems (LMSs) are responsible for learning activities, while other systems are responsible for handling any supporting learning activities. All systems shall integrate and interoperate together to support educational institutions and online learning [14].

Considering that students as the main subject, we should not focus only on the technology itself. The variable importance also mentioned how student efforts to obtain the excellent semester achievement index. In Cultural-Historical Activity Theory (CHAT), the primary unit is students (as the main subject) who carry out various activities (objects) to achieve outcomes. Information technology in online learning mechanisms is a fundamental component in constructing knowledge based on human consciousness [15]. Learning performance between students who do well and those who ultimately fail has been a concern of researchers since the mid-1970s. What makes students successful compared to those who are not successful in learning? The main differences were in the way of learning.

Doing well in learning performance shows the ability to self-manage (self-regulation). Their metacognitive skills are much better because they could assess their strengths and weaknesses. They can find solutions and compensation that can overcome their weaknesses or shortcomings. The ability to self-regulate metacognitive skills refers to a series of interrelated activities, including the ability to set learning goals and orientation; determining the strategies that will be used to achieve the stated goals; management of various existing resources; the implementation; appropriate reactions to feedback from others; and the meaning of the resulting product. If a series of skills has never been possessed or developed, students will undoubtedly fail to develop regulation of metacognition skills [16].

The survey described that students could do more independently based on their interest in learning something through online learning. The student's high self-confidence is related to completing their studies in the online learning process. This factor is mainly determined by the availability of supporting tools in the form of technological equipment, and at the same time, the ability to use various existing technological tools. LMS features and good internet access have created student self-confidence. This finding differs from studies published that have confirmed the strength of feelings, isolation, and growing frustration because of this learning experience in the context of the Covid-19 Pandemic [13].

Independent learning is related to student self-confidence and the ability to manage the time to learning. The previous study on independent learning showed that developing online teaching approaches incorporating other techniques and tools is necessary [17, 18]. For that reason, in its implementation, lecturers provided complete access to communication and problem-solving. Thus, students do not feel separated from their lecturers because of this kind of emotional closeness. It becomes the basis of the formation of mutual trust. Three main elements to build this kind of well-being are communication skills, social tolerance, and creativity. The loss of encounters with lecturers and fellow students has erased the social structure in which students could share their burdens and care for one another. Data related to student preferences regarding the role of lecturers in adapting learning material to increase relevance to online learning reflects the tension that has been unconsciously built up so far.

On the one hand, this shows the lecturers' concern for their students, prioritizing well-being orientation or comfort [19]. However, at the same time, this also shows a strong indication of the existence of fundamental weakness in students, that is, the level of learning independence among students is still relatively low among students [20]. In other words, courses should be designed and taught with their specific strengths and needs in mind.

Based on this variable importance analysis, we inferred that the student's perception from the questionnaire could predict the student semester achievement index.

Moreover, we performed prediction analysis using ten attributes as an independent variable by applying Random Forest and Support Vector Machine algorithms. The target label (a dependent variable) is the student semester achievement index. We compared the prediction result using the default parameters and tuned parameters (a grid search technique) for Random Forest and Support Vector Machine algorithms.

TABLE 2. The default and tuned parameter on the prediction analysis.

Default Parameters	Values	Tuned Parameters	Values
SVM		SVM	
Gamma	0.1	Gamma	$2^{\wedge}(-5:5)$
Cost	1	Cost	$2^{\wedge}(-5:5)$
Epsilon	0.1	Epsilon	0.1; 0.01; 0.001
RF		RF	
Mtry	1	Mtry	1; 1.5; 2
Ntree	500	Ntree	500; 1000; 2000
Maxnodes	24	Maxnode	24
Maxsize	14	Maxsize	14

To determine the best parameter, we implemented a grid search technique. Table 3 expressed the best parameter for prediction analysis.

TABLE 3. The best parameter

Algorithms	Parameter values
SVM – tuned	
Gamma	0.03125
Cost	0.0625
Epsilon	0.1
RF – tuned	
Mtry	1
n-Tree	2000

Random Forest and Support Vector Machine algorithms were implemented, and the RMSE results are below:

TABLE 4. The RMSE results on the training and testing dataset.

Algorithms	Training dataset	Testing dataset
SVM – default	0.41	0.472
SVM – tuned	0.452	0.46
RF – default	0.367	0.465
RF – tuned	0.44	0.451

The results showed that we obtained the optimal predicted RMSE value using RF-tuned algorithms compared to other algorithms.

CONCLUSION

Machine learning approaches could be used as an alternative way to describe student perceptions of online learning. We predicted the student semester achievement index through the student questionnaire and the semester achievement index and obtained some inputs for successful online learning. The results showed that three things greatly influenced student perceptions in implementing online learning. Technology becomes the most crucial factor when the university wants to establish online learning. Students believed that information technology such as zoom technology, YouTube, and University Learning Management Systems are the main elements in supporting their study's success. Students hope that the learning material needs to be changed and adapted to the online learning process or method. In other words, lecturers are asked to change the format for delivering learning material.

The second and third aspects are related to students and study environments. Students acknowledged that they need more time to learn to catch new knowledge. They also recognized that they could do some discussion between students to enhance their knowledge. This happens through LMS, where they found that the LMS features are easily used and are helpful. In the online learning process, the lecturer-student relationship is a crucial factor that will significantly influence the learning process. Online learning is also conceded as having made students more enthusiastic in learning and completing their assignments.

Finally, the suggestions to the university stakeholders are necessary to re-examine what kinds of materials and technologies can be used to be more acceptable to students as a learning process to gain broader knowledge. It is also necessary to learn about how to prepare student personalities, especially concerning student independence in the learning process.

ACKNOWLEDGMENTS

The authors would like to thank LPMAI Sanata Dharma University Yogyakarta for providing the opportunity to research student perceptions and achievements of the online learning process during the Covid-19 pandemic.

REFERENCES

1. Wu, J.-H., R.D. Tennyson, and T.-L. Hsia, *A study of student satisfaction in a blended e-learning system environment.* Computers & Education, 2010. **55**(1): p. 155-164.
2. Hsia, T.-C., A.-J. Shie, and L.-C. Chen, *Course planning of extension education to meet market demand by using data mining techniques–an example of Chinkuo technology university in Taiwan.* Expert Systems with Applications, 2008. **34**(1): p. 596-602.
3. Owston, R., D. York, and S. Murtha, *Student perceptions and achievement in a university blended learning strategic initiative.* The Internet and Higher Education, 2013. **18**: p. 38-46.
4. Han, J., J. Pei, and M. Kamber, *Data mining: concepts and techniques*. 2011: Elsevier.
5. Ali, D.M.M., *Role of data mining in education sector.* International Journal of Computer Science and Mobile Computing, 2013. **2**(4): p. 374-383.
6. Cope, B. and M. Kalantzis, *Big data comes to school: Implications for learning, assessment, and research.* AERA Open, 2016. **2**(2): p. 2332858416641907.
7. Kovacic, Z., *Early prediction of student success: Mining students' enrolment data.* 2010.
8. Kardan, A.A., et al., *Prediction of student course selection in online higher education institutes using neural network.* Computers & Education, 2013. **65**: p. 1-11.
9. Mardikyan, S. and B. Badur, *Analyzing Teaching Performance of Instructors Using Data Mining Techniques.* Informatics in Education, 2011. **10**(2): p. 245-257.
10. Kushik, N., N. Yevtushenko, and T. Evtushenko, *Novel machine learning technique for predicting teaching strategy effectiveness.* International Journal of Information Management, 2016.
11. Engeström, Y., *Expansive learning at work: Toward an activity theoretical reconceptualization.* Journal of education and work, 2001. **14**(1): p. 133-156.
12. Eady, M. and L. Lockyer, *Tools for learning: Technology and teaching.* Learning to teach in the primary school, 2013. **71**.
13. Basilaia, G. and D. Kvavadze, *Transition to online education in schools during a SARS-CoV-2 coronavirus (COVID-19) pandemic in Georgia.* Pedagogical Research, 2020. **5**(4).
14. El-Bakry, H.M., et al. *E-learning and management information systems for E-universities*. in *WSEAS International Conference. Proceedings. Recent Advances in Computer Engineering*. 2009. WSEAS.
15. Postholm, M.B., *Methodologies in Cultural–Historical Activity Theory: The example of school-based development.* Educational Research, 2015. **57**(1): p. 43-58.
16. Nicol, D.J. and D. Macfarlane-Dick, *Formative assessment and self-regulated learning: A model and seven principles of good feedback practice.* Studies in higher education, 2006. **31**(2): p. 199-218.
17. Luke, B. and K. Hogarth, *Developing and enhancing independent learning skills.* Accounting Research Journal, 2011.
18. Hsu, J., K. Hamilton, and J. Wang, *Guided independent learning: A teaching and learning approach for adult learners.* International Journal of Innovation and Learning, 2015. **17**(1): p. 111-133.
19. Kislyakov, P.A., et al., *Monitoring of education environment according to the social-psychological safety criterion.* Asian Social Science, 2014. **10**(17): p. 285.
20. Little, D., *Language learner autonomy: Some fundamental considerations revisited.* International Journal of Innovation in Language Learning and Teaching, 2007. **1**(1): p. 14-29.