

CLASSIFICATION OF STUDENT SATISFACTION WITH ONLINE LECTURE

Nanang Ruhyana^{1*)}, Tati Mardiana², Fachri Amsury³, Daning Nur Sulistyowati⁴

Sains Data^{1*2}, Sistem Informasi^{3,4}
Fakultas Teknologi dan Informasi, Universitas Nusa Mandiri
www.nusamandiri.ac.id
nanang.ngy@nusamandiri.ac.id, tati.ttm@nusamandiri.ac.id, fachri.fcy@nusamandiri.ac.id,
daningnur.dgs@nusamandiri.ac.id

(*) Corresponding Author

Abstrak

Covid-19 memberikan dampak yang signifikan dalam kehidupan masyarakat, mengakibatkan lumpuhnya hampir seluruh perekonomian dan pendidikan, terutama pada bidang pendidikan mengakibatkan banyak siswa dan mahasiswa tidak bisa melakukan aktifitas kegiatan belajar mengajar di sekolah atau universitas. Berdasarkan itu kementerian pendidikan dan kebudayaan memberikan himbauan untuk menghentikan kegiatan belajar mengajar secara tatap muka di sekolah maupun perguruan tinggi dan digantikan dengan melakukan pembelajaran jarak jauh atau daring. Mengakibatkan kegiatan belajar mengajar menjadi kurang maksimal untuk siswa ataupun mahasiswa, terdapat ketidakpuasan terhadap sistem pembelajaran jarak jauh atau daring, tujuan dari penelitian ini adalah mengukur tingkat kepuasan mahasiswa terhadap perkuliahan secara daring dengan menerapkan teknik data mining, melakukan klasifikasi tingkat kepuasan pembelajaran secara daring menggunakan pendekatan algoritma k-NN dan Decision Tree dengan 100 data kuesioner yang telah dikumpulkan dari mahasiswa aktif yang melaksanakan perkuliahan secara daring dengan tingkat akurasi dari algoritma k-NN sebesar 96.00% serta nilai presisi puas sebesar 95.51%, nilai recall puas sebesar 98.84% pada nilai presisi kelas tidak puas sebesar 90.91%, nilai recall kelas tidak puas 71.43%. Sedangkan hasil akurasi menggunakan pendekatan algoritma Decision Tree lebih rendah dengan akurasi sebesar 95.00%. berdasarkan hasil penelitian bahwa tingkat kepuasan mahasiswa terhadap pembelajaran jarak jauh atau daring cukup tinggi.

Kata kunci: covid 19, data mining, daring, k-NN, decision tree

Abstract

Abstra Covid-19 has had a significant impact on people's lives, resulting in the paralysis of almost the entire economy and education, especially in the education sector, resulting in many students being unable to carry out teaching and learning activities at schools or universities. Based on this, the Ministry of Education and Culture has issued an appeal to stop face-to-face teaching and learning activities at schools and universities and replace them with distance or online learning. Resulting in teaching and learning activities to be less than optimal for students or students, there is dissatisfaction with the distance or online learning system, the purpose of this study is to measure the level of student satisfaction with online lectures by applying data mining techniques, classifying the level of online learning satisfaction using an online learning approach. k-NN algorithm and Decision Tree with 100 questionnaire data that has been collected from active students who carry out online lectures with an accuracy rate of 96.00% from the k-NN algorithm and a satisfied precision value of 95.51%, a satisfied recall value of 98.84% on a precision value the dissatisfied class is 90.91%, the recall value of the dissatisfied class is 71.43%. While the accuracy results using the Decision Tree algorithm approach is lower with an accuracy of 95.00%. based on research results that the level of student satisfaction with distance learning or online is quite high.

Keywords: covid 19, data mining, online, k-NN, decision tree

INTRODUCTION

Covid-19 has an impact on all lines of economic, social and educational life (Damanik, Sumijan, & Nurcahyo, 2021). In the world of

education there has been a significant change in teaching and learning activities in connection with the policies of the ministry of education and culture give an appeal to stop teaching and learning activities face-to-face at school and



college and replaced by doing distance learning or online. So it requires carrying out. So that requires carrying out lectures at home by applying distance learning by utilizing the internet as well as virtual applications that can help teaching and learning activities, but sometimes there are obstacles in the process, both technically and non-technically (Mulyana, Rainanto, Astrini, & Puspitasari, 2020). Barriers to the online process are related to the readiness of the government and the community in the educational environment to quickly adapt to changes in the existing learning system (Putu et al., 2021). Online learning is the implementation of distance education in higher education aimed at equalizing access to quality learning (Napaswati, 2020).

Some universities have implemented an online learning system, but there are pros and cons in the process, one of which is experienced by students in online learning, there are things to complain about, mainly because of inadequate internet or unsupported devices, causing the teaching and learning process to be disrupted and ineffective. (Susilowati, Sabariah, & Gozali, 2015).

The effect of changing the learning model from face to face to online causes less effective teaching and learning activities due to the lack of direct interaction, it is necessary to re-adapt to find out the compatibility of learning media with learning methods (Musu, Simpen, & Samsie, 2021).

Related research to measure the level of student satisfaction with online lectures using the naive Bayes algorithmic approach (Natuzzuhriyyah & Nafisah, 2021). Determining the level of activeness of students participating in online training using the k-Mean algorithm (Rohman & Mujiyono, 2020). Prediction of online learning outcomes during the pandemic using the C4.5 . method (Fitriani, Defit, & Nurcahyo, 2021). Analysis of public opinion on policies for handling the covid-19 pandemic using the k-NN algorithm (Aulianita, 2016).

This study aims to classify student satisfaction with distance or online learning, by applying data mining techniques and the k-NN algorithm and decision tree approaches.

RESEARCH METHODS

Research using the concept of Knowledge Discovery in Database (Tsytzarau & Palpanas, 2012). To classify student satisfaction with online lectures at the University of Nusa Mandiri with the following stages:

Types of research

This study uses quantitative data in the form of questionnaires distributed to 6th and 7th semester students at Nusa Mandiri University.

Research Target / Subject

The population in this study were 6th and 7th semester students at Nusa Mandiri University by randomly distributing 100 questionnaires. The sample used is a questionnaire that has been filled out by students from the period 21 March 2021 - 14 June 2021 with 100 data. The sampling technique used is snowball sampling, which is a method to get a sample where the first selected sample is asked to choose and forward it to the group that matches the criteria to be used as a sample. (Susilowati et al., 2015).

Procedure

The following is a framework that is carried out based on the steps and procedures of this research process (Takdirillah, 2020).

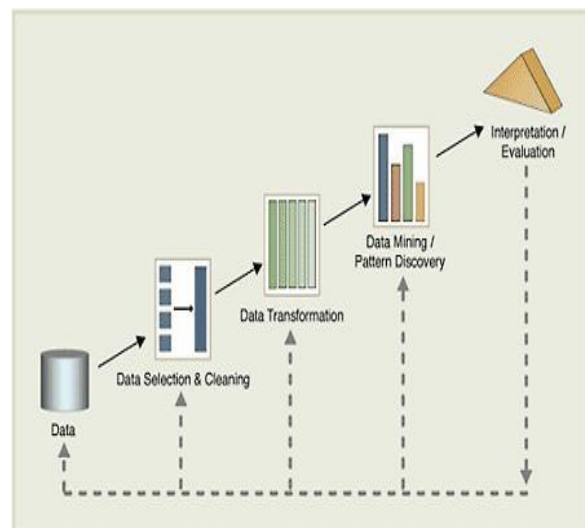


Figure 1. Knowledge Discovery in Database

Figure 1 describes the KDD process that is used as a way to find and identify a pattern in the data where in this study the data came from a questionnaire where the data was selected and cleaned of noise, then the selected data was transformed so that the data mining process could be carried out. the data mining process is carried out by searching for patterns or information from data using data mining methods or algorithms, finally at the interpretation/evaluation stage is the process of translating patterns that have been generated from data mining and presented in a form that is easily understood by interested parties.

1. Data
The data used comes from questionnaires that have been distributed to students of Nusa Mandiri University semesters 6 and 7, to classify satisfaction with the distance or online learning system.
2. Data Selection & Cleaning
Aims to process data derived from the results of questionnaires and select data that can be used in research.
3. Data Transformation
Data that has gone through the selection and cleaning process will be transformed into a separate file with the xls format so that the data is ready for the mining process.
4. Data mining
The transformed dataset will look for the right algorithm approach by comparing the two algorithms, namely k-NN and Decision Tree to see the level of accuracy generated from the two algorithms using RapidMiner Studio tools.
5. Interpretation / Evaluation
The data that has been modeled on the k-NN algorithm and decision tree will be visualized to make it easier to compare the best accuracy levels of the two algorithms.

Data, Instruments, and Data Collection Techniques

This study uses quantitative data in the form of questionnaires distributed to students of Nusa Mandiri University semester 6 and 7 from the period 21 March 2021 - 14 June 2021 as much as 100 data, in the process using Rapid Miner Studio version 9.8.001 tools.

Data analysis technique

This study uses a quantitative research approach because it uses a lot of numbers from the beginning of data collection, then transforms the data and performs data visualization (Palupi, 2021). This study uses quantitative data analysis techniques by conducting interviews with students who do distance or online learning and observing activities during learning and distributing questionnaires via google form to 6th and 7th semester students participating in online learning. There are 100 students who fill out the questionnaire and the resulting data is 100 data. Based on that, a study was designed to classify student satisfaction with distance or online learning systems.

RESULTS AND DISCUSSION

The following is a sample of the data used in the study.

Table 1. Student questionnaire data sample

Name	Jk	Smt	A	B	C	D	E
Andi	L	7	23	19	30	30	25
Muhamad Yusup							
Kandai	L	7	25	20	30	30	25
Andry	L	7	18	13	19	23	18
Stevanus P Rian	L	7	19	15	23	23	19
Drevianto Achyar Ibnu Hadjar	L	7	18	15	22	21	11
Santoso	L	7	25	29	27	30	25
Heny	P	6	21	19	24	25	20
Thesalonika Apresa	P	6	22	18	25	29	21
Yuliana dwi puspitasaki	P	7	21	15	23	29	21
Yoshua Kristiawan	L	7	19	16	26	28	22

Table 1 is a sample of research data consisting of name column, jk column is gender, smt column is semester, column A is tangibles value, column B is reliability, column C is responsiveness, column D is assurance value and column E is empathy value.

K-Nearest Neighbor

The basic concept of k-NN is the search for the shortest distance between the data to be evaluated with the k closest neighbors, Then the value of the distance between the test data and the training data is sorted from the lowest value, the sorting process is carried out to select a minimum distance of k (Nasution, Khotimah, & Chamidah, 2019). The value of k used in this study is k = 3. The equation for the calculation is as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \dots\dots\dots (1)$$

- Description::
 d : sample distance
 xi : knowledge sample data
 pj : data input var to i
 n : number of samples

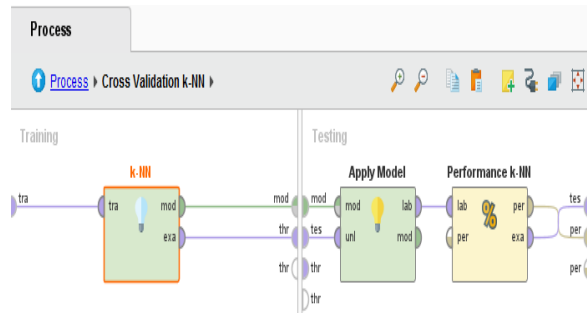


Figure 2. k-NN algorithm Algorithm Validation Process



Figure 2 is a validation process by applying the k-NN algorithm modeling to get an accuracy value, with a value of k=3.

accuracy: 96.00% +/- 5.16% (micro average: 96.00%)

	true Puas	true Tidak Puas	class precision
pred. Puas	86	4	95.56%
pred. Tidak Puas	0	10	100.00%
class recall	100.00%	71.43%	

Figure 3. The results of the accuracy of the k-NN algorithm

Figure 3 explains that the results of the accuracy of the k-NN algorithm are 96.00% with the results of class precision predictions being dissatisfied at 100.00%, class precision predictions being satisfied 95.56% and class recalling true being satisfied at 100.00%, class recalling being dissatisfying 71.43%. Based on the results of the questionnaire data, it shows that the classification for distance or online learning satisfaction is 86 data and the questionnaire data referred to in the dissatisfied prediction is 4 data. Questionnaire data satisfied with predictions of dissatisfied as much as 0 data and questionnaire data dissatisfied with predictions of dissatisfied as much as 10 data.

Decision Tree

Generally, the decision tree algorithm is used to build a decision tree by selecting an attribute that will be the root attribute and each branch will represent the result for the attribute value, break down the cases on the branch and repeat the process until all cases on the branch have received the same class. (Rosandy, 2016). The equation for the calculation is as follows:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \dots\dots\dots (2)$$

Description: S : case set; n : number of partitions S ; pi : the proportion of Si to S

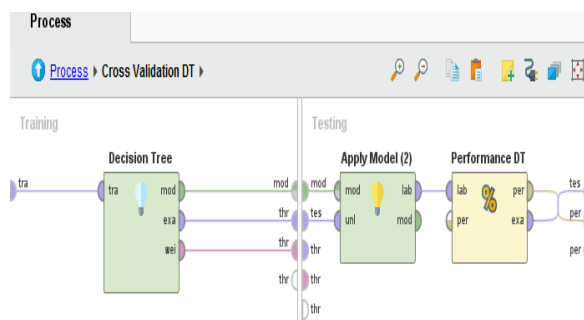


Figure 4. Decision tree algorithm validation process

Figure 4 is a validation process by applying decision tree algorithm modeling to get the accuracy value.

accuracy: 95.00% +/- 5.27% (micro average: 95.00%)

	true Puas	true Tidak Puas	class precision
pred. Puas	85	4	95.51%
pred. Tidak Puas	1	10	90.91%
class recall	98.84%	71.43%	

Figure 5. Decision tree algorithm accuracy results

Figure 5 explains that the results of the decision tree algorithm accuracy are 95.00% with class precision predictions satisfied 95.51%, class precision predictions dissatisfied 90.91% and class recall true results satisfied 98.84%, class recall true dissatisfied 71.43%. Based on the results of the questionnaire data, it shows that the classification for distance or online learning satisfaction is 85 data and the questionnaire data referred to in the dissatisfied prediction is 4 data. The questionnaire data is satisfied with the prediction of being dissatisfied with 1 data and the questionnaire data is not satisfied with the prediction of being dissatisfied with as much as 10 data.

CONCLUSIONS AND SUGGESTIONS

Conclusion

Qualification results from this study using the k-NN algorithm get an accuracy value of 96.00% with a class precision prediction of dissatisfied at 100.00%, class precision prediction is satisfied 95.56% and class recall true results are satisfied 100.00%, class recall true is not satisfied 71.43%. while the results of the accuracy of the decision tree algorithm are 95.00% with the results of class precision prediction satisfied 95.51%, class precision prediction dissatisfied 90.91% and class recall true results satisfied 98.84%, class recall true dissatisfied 71.43%. Based on the results obtained, the accuracy level using the k-NN algorithm has a higher level of accuracy compared to the accuracy results generated by the decision tree algorithm. It can be concluded that the k-NN algorithm with k=3 can be used as an appropriate approach model for predicting student satisfaction with distance or online lectures with a high level of accuracy, so that it can help universities to see the level of



student satisfaction with distance or online learning systems.

Suggestion

Future research is expected to further explore varied data and a larger number of samples, can also optimize for the level of accuracy by adding features and applying optimization techniques as well as conducting trials with different algorithmic approaches to see the advantages of the algorithm in the case of the classification of student satisfaction levels. towards distance or online learning.

REFERENCES

- Aulianita, R. (2016). Komparasi Metode K-Nearest Neighbors dan Support Vector Machine Pada Sentiment Analysis Review Kamera. *Journal Speed – Sentra Penelitian Engineering Dan Edukasi*, 8(3), 71–77. Retrieved from <http://ijns.org/journal/index.php/speed/article/view/995>
- Damanik, A. R., Sumijan, S., & Nurcahyo, G. W. (2021). Prediksi Tingkat Kepuasan dalam Pembelajaran Daring Menggunakan Algoritma Naïve Bayes. *Jurnal Sistim Informasi Dan Teknologi*, 3, 88–94. <https://doi.org/10.37034/jsisfotek.v3i3.137>
- Fitriani, Y., Defit, S., & Nurcahyo, G. W. (2021). Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4.5. *Jurnal Sistim Informasi Dan Teknologi*, 3, 149–157. <https://doi.org/10.37034/jsisfotek.v3i3.149>
- Mulyana, M., Rainanto, B. H., Astrini, D., & Puspitasari, R. (2020). Persepsi Mahasiswa Atas Penggunaan Aplikasi Perkuliahan Daring Saat Wabah Covid-19. *JAS-PT (Jurnal Analisis Sistem Pendidikan Tinggi Indonesia)*, 4(1), 47–54. <https://doi.org/10.36339/jaspt.v4i1.301>
- Musu, W., Simpen, W., & Samsie, I. (2021). Identifikasi Pola Pembelajaran Daring dimasa Pandemi Covid-19 menggunakan Teknik Data Mining *Identification of Online Learning Patterns during the Covid-19 Pandemic using Data Mining Techniques*. 11–20. <https://doi.org/10.30818/jpkm.2021.2060302>
- Napaswati. (2020). Analisis Situasi Pembelajaran IPA Fisika dengan Metode Daring di Tengah Wabah COVID-19 (Studi Kasus Peserta Didik MTS DDI Seppange Kabupaten Bone). *Karst Jurnal Pendidikan Fisika Dan Terapannya*, 3(1), 6–12. Retrieved from <http://ejournals.umma.ac.id/index.php/karts/article/view/546>
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and System Journal*, 4(1), 78–85. <https://doi.org/10.24114/cess.v4i1.11458>
- Natuzzuhriyyah, A., & Nafisah, N. (2021). Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Secara Daring Menggunakan Algoritma Naïve Bayes. *Techno Xplore : Jurnal Ilmu Komputer Dan Teknologi Informasi*, 6(2), 61–67. <https://doi.org/10.36805/technoxplore.v6i2.1377>
- Palupi, E. S. (2021). Prediction Of Android Handphone Sales During Pandemic Using Naïve Bayes And K-Nn Methods Based On Particle Swarm Optimization. *Jurnal Riset Informatika*, 4(1), 23–28. Retrieved from <http://ejournal.kresnamediapublisher.com/index.php/jri/article/view/279>
- Putu, N., Naraswati, G., Rosmilda, D. C., Desinta, D., Statistika, P. D., & Stis, P. S. (2021). Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification. *Jurnal SISTEMASI*, 10(1), 228–238. Retrieved from <http://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/1179>
- Rohman, A., & Mujiyono, S. (2020). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Keaktifan Peserta Pelatihan Pembuatan Dokumen Ilmiah Secara Daring. *Multimatrix*, 2(2), 32–35. Retrieved from <http://jurnal.unw.ac.id:1254/index.php/mm/article/view/947>
- Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier dengan Metode Decision Tree Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus : KSPPS / BMT AL-FADHILA). *Jurnal TIM Darmajaya*, 2(1), 52–62. Retrieved from <https://jurnal.darmajaya.ac.id/index.php/jtim/article/view/648>
- Susilowati, E., Sabariah, M. K., & Gozali, A. A. (2015). Impelentasi Metode Support Vector Machine Untuk Melakukan Klasifikasi Kemacetan Lalu Lintas Pada Twitter. *E-Proceeding of Engineering*, 2(1), 1478–1484. Retrieved from <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/2578>
- Takdirillah, R. (2020). Penerapan Data Mining



Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic: Jurnal Pendidikan Informatika*, 4(1), 37-46. <https://doi.org/10.29408/edumatic.v4i1.2081>

Tsytarau, M., & Palpanas, T. (2012). Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 24(3), 478-514. <https://doi.org/10.1007/s10618-011-0238-6>

