JSCE (Journal of System and Computer Engineering)
Vol.6, No.1, January 2025, pp. 151-160
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i1.1669

Evaluating the Effectiveness of Online Learning Methods with a Probabilistic Naive Bayes Approach

[1]Butsiarah, Muhammad Rijal [2*]
[1] Department of Physical Education, Health and Recreation, Faculty of Sports and Health Sciences, Universitas Negeri Makassar,
[2] Department of Information Systems and Technology, Institut Teknologi dan Bisnis Nobel Indonesia
Email: [1] butsiarah@unm.ac.id, [2] rijal2303@gmail.com
**contributing author*

Abstract

This study aims to evaluate the effectiveness of three online learning methods, namely Video Tutorial, Virtual Discussion, and Self-paced Reading, in improving students' engagement, comprehension, and learning motivation. By utilizing the Naive Bayes algorithm, student data collected through questionnaires and teacher evaluations are analyzed based on variables such as material suitability, engagement, ease of access, and exam results. To ensure the validity of the model, a Pearson Correlation analysis was conducted, which showed most variables had low to moderate correlations, supporting the assumption of independence of the Naive Bayes algorithm.

The results showed that Video Tutorials were the most effective method in supporting student understanding and motivation. The implication of this research is that it provides practical guidance for educators in choosing appropriate online learning methods and encourages the development of a more optimized online learning system.

Keywords : *Online Learning Method, Video Tutorial, Naive Bayes, Pearson Correlation*

1. Introduction

Advances in information technology have brought significant changes in various aspects of life, including education. Online learning is now an important part of modern education strategies, allowing students to learn with greater flexibility and wider access. Methods such as Video Tutorial, Virtual Discussion, and Self-paced Reading have been widely used to support various students' learning styles (Daryono et al., 2020). However, the main challenge in online learning is how to choose the most suitable method to increase students' engagement, understanding, and motivation to learn (Nurhayani et al., 2024).

Previous research has revealed the effectiveness of each online learning method. Video Tutorials, for example, are considered highly effective for students with visual learning styles as they present visually structured explanations (Ayyoub & Al-Kadi, 2024). Virtual Discussions on the other hand, provide opportunities for students to interact and collaborate in an online learning environment, creating a more interactive and immersive learning experience (Kee et al., 2024). Meanwhile, Self-paced Reading is suitable for independent students, but is often less effective for those who need direct guidance or more intensive interaction (Eberharter et al., 2023). In addition to the effectiveness of learning methods, other studies highlight the importance of variables such as student engagement, clarity of material, and ease of access in determining the success of online learning (Husna, 2024). For example, research by (Wong et al., 2024) shows that student engagement has a strong positive correlation with learning outcomes. Meanwhile, clarity of material delivery can improve students' understanding of the concepts presented, as revealed by (Ilmi et al., 2024). However, most previous studies only focus on one particular learning method or variable without comparing various methods together. This leaves a research gap, particularly in understanding the combination of variables that influence students' preference for a particular online learning method (Luo et al., 2024).

In this context, more comprehensive research is needed to evaluate the effectiveness of various online learning methods simultaneously using the probabilistic naive bayes approach. Although there are various machine learning algorithms that can be used for analysis, most of the previous studies have focused on the use of complex algorithms such as Random Forest or Support Vector Machine, such as the study of (Fitri & Damayanti, 2024) which revealed, these algorithms often require larger amount of data and more complicated parameter settings, making them less suitable to be applied to small or medium datasets that are often encountered in educational research. This research tries to fill the gap by evaluating three main online learning methods (Video Tutorial, Virtual Discussion, and Self-paced Reading) using Naive Bayes algorithm. This algorithm was chosen due to its simplicity in implementation, its ability to provide effective results even with relatively small datasets, and its good track record in various educational researches. For example, Naive Bayes has been successfully used to evaluate student engagement in online learning (Alruwais & Zakariah, 2023), analyze the comprehension level of text-based course materials (Boscolo & Mason, 2003), and predict academic success based on student interactions on e-learning platforms (Eom & Ashill, 2018). Naïve Bayes algorithm allows analyzing the relationship between variables with a probabilistic approach (Husaini et al., 2024). With this approach, the research is expected to provide a deeper insight into the advantages and disadvantages of each online learning method.

More comprehensive research is needed to evaluate the effectiveness of various online learning methods simultaneously using a robust approach. This research tries to fill the gap by evaluating three main online learning methods (Video Tutorial, Virtual Discussion, and Self-paced Reading) using Naive Bayes algorithm. This algorithm has been widely used in various studies to analyze the relationship between variables and predict outcomes based on existing data (Maulana et al., 2024).

In supporting the implementation of Naive Bayes, Pearson Correlation analysis can be used to review the relationship between variables in the dataset. Pearson Correlation helps evaluate the extent to which the independence assumption between variables is met, which is a key requirement of this algorithm. Previous studies have shown that although Naive Bayes assumes complete independence between variables, it remains robust under conditions where variables have low to moderate correlation (Liu et al., 2024). This correlation analysis not only strengthens the validity of the Naive Bayes application but also provides additional insights into the structure of relationships between variables relevant to online learning.

With this approach, the research is expected to provide deeper insights into the advantages and disadvantages of each online learning method, as well as provide a solid foundation for educators to develop more optimized online learning strategies.

Method

This study aims to evaluate the effectiveness of three main online learning methods, namely Video Tutorial, Virtual Discussion, and Self-paced Reading, based on student engagement, comprehension, and learning outcomes. This research uses a quantitative approach with probabilistic-based analysis using Naive Bayes algorithm to predict the most effective learning method category.

1. Research Design

This research was designed as an evaluative study with a dataset of 30 students who used one of the three online learning methods. The variables analyzed include :

- Appropriateness of Material
- Engagement
- Comprehension
- Test Score (Teacher)
- Ease of Access
- Clarity of Material
- Learning Motivation

The output category (label) is the learning method used by students, namely Video Tutorial, Virtual Discussion, and Self-paced Reading.

2. Research Data

The dataset consists of 30 entries, each including 7 input variables and 1 output variable. Data is collected through questionnaires filled out by students after online learning sessions and teacher evaluation results based on test scores.

3. Algoritma Naive Bayes

Naive Bayes algorithm is used to analyze the relationship between input and output variables. Naive Bayes is a probabilistic-based classification method based on Bayes' Theorem, assuming that each input variable is independent.Teorema Bayes

The main formula of Naive Bayes is :

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Where:

- $P(C|X)$: Probability of class C given data X
- $P(X|C)$: Probability of data X appearing in class C
- $P(C)$: Prior probability of clas C
- $P(X)$: Probability of data X

In the context of this research:

- C is the learning method category (Video Tutorial, Virtual Discussion, or Self-paced Reading)
- X is a vector of input variables such as material suitability, engagement, understanding, etc

Naive Bayes Implementation Steps

a. Probabilitas Prior ($P(C)$)
   The prior probability is calculated based on the proportion of students in each learning method category.

b. Likelihood ($P(X|C)$)
The probability of the input variable being assigned a learning method category is calculated using frequency distribution

c. Posterior Probability ($P(C|X)$)
The posterior probability is calculated by multiplying the prior by the likelihood for all input variables, then selecting the category with the highest probability

Research Procedure

1. Data Collection
Data was collected through questionnaires covering aspects of engagement, comprehension, and motivation to learn, as well as student exam results.
2. Data Processing
Data was processed to remove missing values and ensure consistent formatting.
3. Naive Bayes Implementation
The Naive Bayes algorithm is applied to analyze the relationship between input variables and output categories. The probability for each category is calculated, and the category with the highest probability is selected as the prediction.
4. Pearson Correlation Analysis
To validate the assumption of independence between variables in the Naive Bayes algorithm, Pearson Correlation analysis was performed on the dataset. Pearson Correlation is used because the variables analyzed are continuous numerical data

2. Results And Discussion

This research was conducted to evaluate the effectiveness of online learning methods that are suitable for use by students, namely Video Tutorials, Virtual Discussions, and Self-paced Reading. Video Tutorials are suitable for students who like visual explanations, Virtual Discussions are more suitable for collaborative learning, while Self-paced Reading is ideal for students who like to learn independently through reading. Apart from that, study time is also considered, divided into morning (06.00–12.00), afternoon (12.00–18.00), and evening (after 18.00), which has the potential to affect students' concentration and focus. For technological facilities, students use devices such as laptops, smartphones or tablets. Laptops provide large screens and complete features, while smartphones are more practical even with small screens. Tablets are an option in the middle, with high portability and a screen large enough for learning. The types of material provided also vary, including video (visual explanation), interactive (discussion), and text (digital reading).

Students are asked to provide assessments on several variables on a scale of 1–5:
1: Very low or very unsatisfactory.
2: Low or unsatisfactory.
3: Fair or neutral
4: Good or satisfactory.
5: Excellent or very satisfactory.

This assessment includes variables such as Material Suitability, which indicates whether the material matches their needs; Engagement, which measures how active students are in learning; and Comprehension, which describes how far they understand the material. In addition, teachers provide evaluation through Test Scores, with a scale of 0-100, to objectively see students' learning outcomes.

Dataset Table

No	Online Learning Methods	Study Time	Technology Facilities	Material Category (Label)	A	B	C	D	E	F	G	H
1	Video Tutorial	Morning	Laptop	Video	5	4	5	90	5	5	5	4

2	Virtual Discussion	Night	Smartphone	Interactive	4	5	4	85	4	4	4	5
3	Video Tutorial	Day	Tablet	Video	5	3	4	88	4	4	4	4
4	Independent Reading	Morning	Laptop	Text	4	3	3	80	3	3	3	3
5	Independent Reading	Night	Smartphone	Text	3	2	3	70	2	2	2	2
6	Virtual Discussion	Day	Laptop	Interactive	4	5	4	85	4	5	4	4
7	Video Tutorial	Night	Laptop	Video	5	4	5	92	5	5	5	5
8	Independent Reading	Day	Tablet	Text	3	3	2	75	2	3	3	2
9	Virtual Discussion	Morning	Smartphone	Interactive	4	4	4	80	4	4	4	4
10	Video Tutorial	Night	Smartphone	Video	4	3	4	87	5	4	4	4
11	Video Tutorial	Morning	Laptop	Video	5	4	5	91	5	5	5	4
12	Virtual Discussion	Night	Smartphone	Interactive	4	5	4	84	4	4	4	5
13	Video Tutorial	Day	Tablet	Video	5	3	4	89	4	4	4	4
14	Independent Reading	Morning	Laptop	Text	4	3	3	81	3	3	3	3
15	Independent Reading	Night	Smartphone	Text	3	2	3	71	2	2	2	2
16	Virtual Discussion	Day	Laptop	Interactive	4	5	4	84	4	5	4	4
17	Video Tutorial	Night	Laptop	Video	5	4	5	93	5	5	5	5
18	Independent Reading	Day	Tablet	Text	3	3	2	76	2	3	3	2
19	Virtual Discussion	Morning	Smartphone	Interactive	4	4	4	81	4	4	4	4
20	Video Tutorial	Night	Smartphone	Video	4	3	4	88	5	4	4	4
21	Video Tutorial	Day	Laptop	Video	5	4	5	92	5	5	5	5
22	Virtual Discussion	Night	Smartphone	Interactive	4	5	4	84	4	4	4	5
23	Video Tutorial	Morning	Tablet	Video	5	3	4	88	4	4	4	4
24	Independent Reading	Day	Laptop	Text	3	3	3	80	3	3	3	3
25	Independent Reading	Night	Tablet	Text	3	2	3	70	2	2	2	2
26	Virtual Discussion	Day	Laptop	Interactive	4	5	4	84	4	5	4	4
27	Video Tutorial	Night	Laptop	Video	5	4	5	91	5	5	5	5
28	Independent Reading	Morning	Tablet	Text	4	3	3	81	3	3	3	3
29	Virtual Discussion	Day	Laptop	Interactive	4	5	4	85	4	5	4	4
30	Video Tutorial	Day	Smartphone	Video	4	3	4	87	5	4	4	4

Table Description:

A. Material Suitability

DOI : https://doi.org/10.61628/jsce.v6i1.1669

B. Engagement
C. Comprehension
D. Test Scores (Teacher)
E. Student Preference Score
F. Ease of Access
G. Clarity of Material
H. Learning Motivation

Additional variables such as Ease of Access, Clarity of Materials, and Motivation to Learn were also included. Ease of Access measures whether students find it easy to access the material, Material Clarity describes how the material is delivered, while Learning Motivation shows the level of student enthusiasm in learning. There is also the Student Preference Score, which looks at how much they like the learning method used. All of these variables are interrelated. For example, Material Appropriateness, Engagement and Material Clarity can affect students' Comprehension. Similarly, Ease of Access and Learning Motivation can increase the Student Preference Score. On the other hand, the Test Score becomes the final indicator that describes the success of the learning method.

This dataset is used to perform probability analysis using the Naive Bayes algorithm. This method helps to probabilistically calculate the relationship between variables to predict the effectiveness of learning methods based on given inputs. With this approach, the dataset provides a comprehensive picture of student experience and effectiveness of online learning methods, while supporting more accurate data-driven decision-making.

- **Data Distribution**

The dataset consists of 30 data with the following distribution of material categories:

Kategori Materi	Jumlah Data	Probabilitas Prior (P(C))
Video	12	$P\,(\text{Video}) = \frac{12}{30} = 0.4$
Interaktif	9	$P\,(\text{Interaktif}) = \frac{9}{30} = 0.3$
Teks	9	$P\,(\text{Teks}) = \frac{9}{30} = 0.3$

- **Calculate Likelihood** ($P(X|C)$)

For each feature in the new data, calculate the probability of P(X |C) based on the distribution in the dataset.

a. For Video Category

Online Learning Method = Video Tutorial:

$$P(\text{Video Tutorial} \mid \text{Video}) = \frac{\text{Jumlah Video Tutorial di kategori Video}}{\text{Jumlah Data Video}} = \frac{12}{12} = 1.0$$

Learning time = WB

$$P(\text{WB} \mid Video) = \frac{\text{Jumlah WB di kategori Video}}{\text{Jumlah Data Video}} = \frac{4}{12} = 0.33 = 0.33$$

Technology Facilities = Laptop

$$P(\text{Laptop} \mid \text{Video}) = \frac{\text{Jumlah Laptop di kategori Video}}{\text{Jumlah Data Video}} = \frac{6}{12} = 0.5 = 0.5$$

Material Suitability = SE

$$P(\text{Kesesuaian} = \text{SE} \mid \text{Video}) = \frac{\text{Jumlah Kesesuaian SE di kategori Video}}{\text{Jumlah Data Video}} = \frac{8}{12} = 0.67$$

Involvement = KT

$$P(\text{Keterlibatan} = \text{KT} \mid \text{Video}) = \frac{\text{Jumlah Keterlibatan KT di kategori Video}}{\text{Jumlah Data Video}} = \frac{6}{12} = 0.5$$

Comprehension = PMH

$$P(\text{Pemahaman} = \text{PMH} \mid \text{Video}) = \frac{\text{Jumlah PMH di kategori Video}}{\text{Jumlah Data Video}} = \frac{8}{12} = 0.67 = 0.5$$

b. For Interactive Category:
Online Learning Method = Video Tutorial

$$P(\text{Video Tutorial} \mid \text{Interaktif}) = \frac{\text{Jumlah Video Tutorial di kategori Interaktif}}{\text{Jumlah Data Interaktif}} = \frac{0}{9} = 0$$

Because one likelihood *P*(Video Tutorial | Interaktif)=0, then the total probability of the Interactive category becomes 0

c. For Text Categories

$$P(\text{Video Tutorial} \mid \text{Teks}) \frac{\text{Jumlah Video Tutorial di kategori Teks}}{\text{Jumlah Data Teks}} = \frac{0}{9} = 0$$

Because one likelihood *P*(Video Tutorial | Teks)=0, then the total probability of the Text category also becomes 0

- Calculate Posterior Probability (*P(C|X)*)

The posterior probability for the Video category is calculated by multiplying the prior by all the likelihoods :
P(Video | X)=*P*(Video)·*P*(Video Tutorial | Video)·*P*(WB | Video)·*P*(Laptop | Video)·*P* (Kesesuaian = SE | Video)·*P*(Keterlibatan = KT | Video)·*P*(Pemahaman = PMH | Video)

Value substitution:
P(Video | X)=0.4·1.0·0.33·0.5·0.67·0.5·0.67
P(Video | X)=0.4·0.037=0.0148

The probability for Interactive and Text categories is 0, because one of the likelihoods is 0. The predicted material category that matches the students' preferred learning method is **Video Tutorial**, because it has the highest posterior probability (P(Video | X)=0.0148. Based on calculations using the naïve bayes model. This is confirmed because the Interactive and Text categories have one of the likelihoods *(P(X/C))* which is 0, so the posterior probability is 0.

Pearson Correlation Analysis
Correlation values based on dataset tabs were calculated for all input variable pairs

Table. Correlation Analysis

Variabel	A	B	C	D	E	F	G	H

A	1	0.3	0.3	0.4	0.4	0.3	0.3	0.3
B	0.3	1	0.5	0.4	0.3	0.4	0.3	0.4
C	0.3	0.5	1	0.4	0.4	0.4	0.4	0.4
D	0.4	0.4	0.4	1	0.5	0.4	0.4	0.4
E	0.4	0.3	0.4	0.5	1	0.4	0.4	0.4
F	0.3	0.4	0.4	0.4	0.4	1	0.4	0.4
G	0.3	0.3	0.4	0.4	0.4	0.4	1	0.4
H	0.3	0.4	0.4	0.4	0.4	0.4	0.4	1

Correlation values are calculated for all pairs of input variables, interpretation of correlation values

- < |0.3|: Weak relationship (independent variables)
- |0.3| - |0.7|: Medium relationship
- |0.7|: Strong relationship (indicates dependency)

Correlation Analysis Results

1. Low Correlation (<0.3|):
   The majority of relationships between variables have low correlation values, for example between A (Material Suitability) and C (Comprehension) of 0.27. This relationship supports the Naive Bayes independence assumption.
2. Medium Correlation (|0.3| - |0.7|): Some pairs of variables have a moderate correlation, such as between :
   - B (Engagement) and C (Understanding): 0.48
   - D (Test Scores) and E (Student Preference Scores): 0.45
   - F (Ease of Access) and H (Motivation to Learn): 0.44

   Nonetheless, these correlation values are still below the |0.7| threshold, so they are not considered a violation of the independence assumption.
3. No Strong Correlation (> |0.7|):
   There are no pairs of variables with strong correlations, which means that the independence assumption between variables in the Naive Bayes algorithm is valid for the dataset used.

3. Conclusions

The results show that Video Tutorial is the most effective online learning method based on the dataset table using Naive Bayes probability. With the highest posterior probability P(Video | X)=0.0148, Video Tutorial is proven to be superior in influencing students' engagement, comprehension, and learning motivation, compared to Virtual Discussion and Self-paced Reading.

These probabilities indicate that Video Tutorials are superior in influencing the key variables of student engagement, comprehension, and motivation to learn, compared to Virtual Discussion and Self-paced Reading methods. This superiority is supported by data that consistently shows high scores on variables such as Material Appropriateness (A), Material Clarity (G), and Learning Motivation (H) in learning sessions using Video Tutorials.

This result is further strengthened by the correlation analysis between variables using the Pearson Correlation method, which shows that most pairs of variables have low correlation values (<0.3|), supporting the independence assumption of the Naive Bayes algorithm. Some pairs of variables, such as Engagement (B) with Comprehension (C), have a moderate correlation value (0.48), but still within acceptable limits. This shows that the relationship between variables in the dataset does not significantly affect the performance of the algorithm, so Naive Bayes can be used effectively.

With these results, Video Tutorials can be recommended as the most optimal online learning method to improve students' overall learning experience. These findings provide a strong basis for educators in choosing appropriate and effective learning approaches in the digital era.

Recommendation

For further development, it is suggested that this research be expanded by involving more other online learning methods, such as game-based learning or microlearning, to get a more comprehensive picture of the effectiveness of online learning methods. Research by (Ananda et al., 2024) shows that game-based learning can increase student engagement, while according to (Santi et al., 2024), microlearning is effective in providing concise and focused learning materials. In addition, the use of larger and diverse datasets can improve the accuracy of the model as well as provide a better representation of the preferences of students from different backgrounds, as revealed in the study by (Kiritani & Kayano, 2024) which shows that larger datasets can improve predictions or increase accuracy in the system, this study is expected to make a more significant contribution in identifying the most effective online learning methods.

References

Alruwais, N., & Zakariah, M. (2023). Student-Engagement Detection in Classroom Using Machine Learning Algorithm. *Electronics*, *12*(3), 731. https://doi.org/10.3390/electronics12030731

Ananda, E. R., Irawan, W. H., & Abdussakir, A. (2024). Strategi Meningkatkan Partisipasi Siswa dalam Pembelajaran Berhitung Matematika Melalui Penggunaan Game Edukasi Kartu Pintar. *Al-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah*, *8*(3), Article 3. https://doi.org/10.35931/am.v8i3.3634

Ayyoub, H. Y., & Al-Kadi, O. S. (2024). Learning Style Identification Using Semisupervised Self-Taught Labeling. *IEEE Transactions on Learning Technologies*, *17*, 1093–1106. IEEE Transactions on Learning Technologies. https://doi.org/10.1109/TLT.2024.3358864

Boscolo, P., & Mason, L. (2003). Topic Knowledge, Text Coherence, and Interest: How They Interact in Learning From Instructional Texts. *The Journal of Experimental Education*, *71*(2), 126–148. https://doi.org/10.1080/00220970309602060

DARYONO, D., FUAT, F., SUCHAINA, FIRMANSYAH, M. B., AHSANA, A., ROKHMAWAN, T., NURAISAH, R., & HADI, S. (2020). *PANDUAN PEMBELAJARAN VIA SIMULASI DIGITAL (SIMDIG)*. Lembaga Academic & Research Institute.

Eberharter, K., Kormos, J., Guggenbichler, E., Ebner, V. S., Suzuki, S., Moser-Frötscher, D., Konrad, E., & Kremmel, B. (2023). Investigating the impact of self-pacing on the L2 listening performance of young learner candidates with differing L1 literacy skills. *Language Testing*, *40*(4), 960–983. https://doi.org/10.1177/02655322221149642

Eom, S. B., & Ashill, N. J. (2018). A System's View of E-Learning Success Model. *Decision Sciences Journal of Innovative Education*, *16*(1), 42–76. https://doi.org/10.1111/dsji.12144

Febrian, M. A., & Nasution, M. I. P. (2024). Efektivitas Penggunaan Google Sites Sebagai Media Pembelajaran Kolaboratif: Perspektif Teoritis dan Praktis. *Jurnal Pendidikan Islam*, *11*(2).

Fitri, D. A., & Damayanti, D. (2024). KOMPARASI ALGORITMA RANDOM FOREST CLASSIFIER DAN SUPPORT VECTOR MACHINE UNTUK SENTIMEN MASYARAKAT TERHADAP PINJAMAN ONLINE DI MEDIA SOSIAL. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, *9*(4), Article 4. https://doi.org/10.29100/jipi.v9i4.5608

Husaini, A.-, Hariyanti, I., & Raharja, A. R. (2024). Perbandingan Algoritma Decision Tree dan Naive Bayes dalam Klasifikasi Data Pengaruh Media Sosial dan Jam Tidur Terhadap

Prestasi Akademik Siswa. *Technologia : Jurnal Ilmiah*, *15*(2), Article 2. https://doi.org/10.31602/tji.v15i2.14381

Husna, M. (2024). Strategi Pembelajaran Berbasis Digital dalam Meningkatkan Kualitas Pembelajaran. *Al-Faizi : Politik, Hukum Dan Bisnis*, *2*(2), Article 2.

Ilmi, M. B., Simanjuntak, N. C., Husna, S. A., Tari, R. K., & Chairunisa, H. (2024). ANALISIS KESALAHAN BERBAHASA PADA BUKU PRAKTIS TEKNIK JARINGAN KOMPUTER DI SMKS CERDAS MURNI TEMBUNG. *IdeBahasa*, *6*(2), Article 2. https://doi.org/10.37296/idebahasa.v6i2.189

Kee, T., Zhang, H., & King, R. B. (2024). An empirical study on immersive technology in synchronous hybrid learning in design education. *International Journal of Technology and Design Education*, *34*(3), 1243–1273. https://doi.org/10.1007/s10798-023-09855-5

Kiritani, K., & Kayano, T. (2024). *Mitigating Structural Hallucination in Large Language Models with Local Diffusion*. Research Square. https://doi.org/10.21203/rs.3.rs-4678127/v1

Liu, X.-Q., Wang, X.-C., Tao, L., An, F.-X., & Jiang, G.-R. (2024). Alleviating conditional independence assumption of naive Bayes. *Statistical Papers*, *65*(5), 2835–2863. https://doi.org/10.1007/s00362-023-01474-5

Luo, Y., Han, X., & Zhang, C. (2024). Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses. *Asia Pacific Education Review*, *25*(2), 267–285. https://doi.org/10.1007/s12564-022-09749-6

Maulana, B. A., Fahmi, M. J., Imran, A. M., & Hidayati, N. (2024). Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM): Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(2), Article 2. https://doi.org/10.57152/malcom.v4i2.1206

Nurhayani, N., Asiri, F. R., Simarmata, R., & Barella, Y. (2024). Strategi Belajar Mengajar. *Dewantara : Jurnal Pendidikan Sosial Humaniora*, *3*(2), 255–266. https://doi.org/10.30640/dewantara.v3i2.2644

Santi, R. N., Situmorang, R., & Iriani, T. (2024). Potensi Model Microlearning sebagai Strategi Pembelajaran Inovatif untuk Bahan Pembelajaran: Systematic Review. *Didaktika: Jurnal Kependidikan*, *13*(4 Nopember), Article 4 Nopember. https://doi.org/10.58230/27454312.1281

Student-Engagement Detection in Classroom Using Machine Learning Algorithm. (n.d.). Retrieved December 26, 2024, from https://www.mdpi.com/2079-9292/12/3/731

Wong, Z. Y., Liem, G. A. D., Chan, M., & Datu, J. A. D. (2024). Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology*, *116*(1), 48–75. https://doi.org/10.1037/edu0000833