Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

KNN Approach to Evaluating the Feasibility of Using Scientific Publications as Final Projects

Dzulchan Abror[1)], Asyahri Hadi Nasyuha[2)*], Meng-Yun Chung[3)], Moch. Iswan Perangin-angin[4)]
[1)] Faculty of Engineering and Informatics, Computer Technology, Bina Sarana Informatika University, Tegal, Indonesia
[2)] Faculty of Information Technology, Information Systems, Teknologi Digital Indonesia University, Yogyakarta, Indonesia
[3)] Department of Electrical Engineering, Lunghwa University of Science and Technology, Taoyuan, Taiwan
[4)]Faculty of Computer science and information technology, Informatics Engineering, Budi Darma University, Medan, Indonesia

[1)]dzulchan.dza@bsi.ac.id, [2)*]asyahrihadi@gmail.com, [3)] mychung@gm.lhu.edu.tw, [4)]mochammadiswan@gmail.com

Submitted :Dec 21, 2024 | **Accepted** : Jan 7, 2025 | **Published** : Jan 11, 2025

Abstract: This study aims to explore the feasibility of using scientific publications as a substitute for traditional final assignments in higher education by applying the K-Nearest Neighbors (K-NN) algorithm. Traditional final assessments, such as theses, are widely used in evaluating students, but with the increasing availability of peer-reviewed scientific publications, there is potential to use them as a more dynamic and relevant assessment tool. This study uses a dataset containing scientific publications and theses, with features such as research quality, relevance, methodology, and clarity. This study applies the K-NN algorithm to classify these materials and determine whether scientific publications can serve as an effective substitute. The results show that the K-NN algorithm, using *k*=4, achieved 95% accuracy, successfully distinguishing between scientific publications and theses. However, some misclassifications occurred, indicating areas for improvement, such as incorporating additional features such as citation counts or peer-review scores. These findings suggest that scientific publications, if properly classified, can indeed replace traditional final assignments, encouraging critical thinking and engagement with current research. Future research should refine the feature set and explore other machine learning models to improve accuracy. The practical implications of this research are the potential to develop more innovative and relevant approaches to assessment in higher education, which are more aligned with modern educational practice.

Keywords: K-NN, Scientific Publications, Final Assignments, Educational Assessment, Machine Learning

INTRODUCTION

In the landscape of higher education, final assignments have traditionally been a cornerstone method for assessing student performance. These assessments, typically in the form of written papers, projects, or exams, serve as a measure of how well students have understood and internalized course material(Abdurahman et al., 2024). However, with the rapid advancements in educational technology and evolving academic paradigms, there is increasing interest in exploring alternative forms of assessment that can more accurately reflect students' competencies(Southworth et al., 2023), align with real-world applications, and provide deeper insights into their learning. One such alternative gaining attention is the use of scientific publications as substitutes for traditional final assignments. By incorporating peer-reviewed articles, research papers, or published studies as a form of final assessment, students would be required to engage more deeply with ongoing academic research, hone their

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

analytical skills, and contribute to scholarly conversations. This method also offers an opportunity to cultivate research skills that are increasingly valued in both academic and professional environments.

Despite the promising potential of this approach, prior studies exploring alternative assessment methods primarily focus on project-based learning, portfolios, or peer evaluations. While these approaches have demonstrated positive impacts on student engagement and critical thinking, the specific use of scientific publications as replacements for final assignments remains largely unexplored. Some studies have highlighted the value of engaging with scientific literature in the learning process, but few have assessed whether these publications can meet the academic rigor required from final assignments. Therefore, a critical gap exists in understanding whether scientific publications can effectively replace traditional assessments and whether they align with the intended learning outcomes for students in higher education.

This study seeks to address this gap by evaluating the feasibility of using scientific publications as substitutes for final assignments. Specifically, the study will focus on using the K-Nearest Neighbors (K-NN) algorithm(Syahril Dwi Prasetyo et al., 2023), a machine learning technique, to assess the suitability of scientific publications for academic evaluation. K-NN(Diansyah, 2022), known for its ability to classify data based on similarity, can be applied to evaluate whether the characteristics of a publication—such as research quality, relevance, and depth—align with the criteria of traditional final assignments. This approach could provide a data-driven method for educators to assess the academic value of scientific papers as viable alternatives to traditional assignments. The main problem addressed in this research is whether scientific publications can replace traditional assignments without compromising academic standards. In particular, it examines whether these publications meet the necessary academic requirements such as clarity, depth of analysis, and alignment with course content. By applying K-NN to assess scientific publications, this research aims to bridge the gap between traditional and innovative assessment methods, providing insights into the feasibility of using scientific publications as final assignments. Furthermore, it aims to contribute to the growing body of literature on data-driven assessment practices in education and offer valuable guidance to educators looking for alternative ways to evaluate student performance in the digital age.

Several studies have explored the use of scientific publications as substitutes for traditional final assignments in higher education. Examined the feasibility of replacing research papers with peer-reviewed articles(Jiang et al., 2020), finding that such an approach enhanced engagement with current research, though challenges remained in distinguishing high-quality publications. Similarly, found that incorporating scientific publications in assessments improved students' analytical skills and provided practical knowledge(Ma et al., 2021). Finally, demonstrated that K-Nearest Neighbors (K-NN) outperformed other machine learning models in classifying academic articles for assessments, suggesting the potential of machine learning in this area(Wang et al., 2023). These studies collectively highlight the promise of using scientific publications for assessments but also emphasize the need for further refinement in classification models and methods.

LITERATURE REVIEW

Traditional assessment methods, including exams, research papers, and final projects, have been critiqued for their limitations in assessing a broad range of skills. As higher education evolves, there is a growing interest in alternative forms of assessment that move beyond traditional models. These include project-based learning, portfolios, peer assessments, and research-based assignments. Research-based assessments, particularly the use of scientific publications, provide students with opportunities to engage with current research, analyze methodologies, and contribute to scholarly discussions. Studies have suggested that such assessments can improve critical thinking and the application of theoretical knowledge. However, most of the existing literature on research-based assessments focuses on portfolios or literature reviews rather than on using complete peer-reviewed publications as substitutes for final assignments. While some studies have shown positive outcomes from engaging students with scientific literature, there remains a lack of empirical evidence on the effectiveness of using scientific publications for final assessment purposes. This gap highlights the need for further exploration of this alternative assessment method.

Scientific Publications and Their Role in Education

Scientific publications are considered a valuable resource for academic learning and research. As peer-reviewed documents, they represent high-quality, credible research that is often at the forefront of academic inquiry. In the context of higher education, students are encouraged to engage with scientific publications as part of their learning process, particularly in research-focused programs. However, there is little research specifically investigating the use of scientific publications as substitutes for final assignments. Publications are generally used as reading materials or sources for research rather than as the final product of a student's academic efforts(Mardiani et al., 2023). The potential of using them as a final assessment tool would require a rethinking of how students engage with the material and how their academic skills are evaluated. Research suggests that the use of real-world research materials could enhance students' ability to critically analyze and synthesize complex information.

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

However, challenges remain, particularly regarding the diversity of publication types and their varying levels of complexity, which may not align with the specific learning outcomes expected from final assignments.

Machine Learning and K-Nearest Neighbors (K-NN) in Educational Assessment

Machine learning algorithms, especially classification methods like K-Nearest Neighbors (K-NN)(Tanggu Mara et al., 2021), have increasingly been applied in education for various tasks, including predicting student performance and evaluating educational outcomes(Tang et al., 2020),(Shah et al., 2020). The K-NN algorithm is particularly useful because it classifies data points based on their proximity to other points in the feature space(Zhang, 2022), making it an effective tool for assessing whether a scientific publication matches the required criteria for a final assignment. K-NN has been applied in several educational settings to predict student success based on various features(Maleki et al., 2021),(Suyal & Goyal, 2022), such as demographic information and past academic performance. Its application to classify scientific publications based on features such as research quality, relevance, methodology, and clarity is a promising direction for this study. However, there is a gap in the literature regarding the use of K-NN to classify research materials, particularly scientific publications, for educational assessments. Most studies on K-NN in education focus on student-related data, not on evaluating the suitability of academic materials. This study aims to fill this gap by applying K-NN to evaluate scientific publications as substitutes for final assignments.

While alternative assessment methods and machine learning techniques have been explored in various contexts, several key gaps remain in the literature. First, the use of scientific publications as substitutes for final assignments has not been sufficiently examined, especially in terms of how they can meet academic standards such as research quality, relevance, and clarity. Most research on alternative assessments has focused on portfolios, literature reviews, and project-based learning, leaving a gap in understanding the feasibility of using peer-reviewed publications in place of traditional assessments. Second, the application of machine learning algorithms like K-NN to assess the suitability of research publications for academic assessments is underexplored. While K-NN has been used extensively for predicting student outcomes, its application to classify academic materials based on specific features like relevance and quality is rare. This research seeks to address these gaps by applying K-NN to evaluate scientific publications, with the goal of determining whether they can serve as effective substitutes for traditional final assignments.

METHOD

In this research, data normalization plays a crucial role in ensuring fair and consistent evaluation of scientific publications used as substitutes for traditional final assignments. Since features like research quality, relevance, and methodology are measured on different scales, normalization helps standardize these features, allowing them to contribute equally to the assessment. This is especially important when applying the K-Nearest Neighbors (K-NN) algorithm, which relies on distance calculations, as different feature scales can otherwise skew the results(Uddin et al., 2022). By using Min-Max normalization, we transform the data into a consistent range, typically (0, 1), eliminating biases due to varying magnitudes. The following formula is used for data normalization(Muliono et al., 2020),(Sari et al., 2023).

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \qquad (1)$$

Where:

- X_{norm} is the normalized value.
- X is the original value.
- $X_{\min}$ is the minimum value in the dataset.
- $X_{\max}$ is the maximum value in the dataset.

K-Nearest Neighbors (K-NN) Algorithm

K-NN is a supervised learning algorithm that classifies data points based on their proximity to other data points in feature space. Given a dataset where each data point consists of a set of features and a corresponding label K-NN classifies a new data point by determining which class (label) is most common among its k nearest neighbors. The formula for calculating Euclidean distance between two data points *p* and *q* in an n-dimensional feature space is given by:

$$D(p,q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \qquad (2)$$

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Where:
p_i and q_i represent the values of the ith feature of points p and q,
n is the number of features.

Several stages in calculating K-Nearest Neighbor are:
1. The value of K can be determined using the formula:
$$k = \sqrt{n} \qquad (3)$$
Where:
- k is the number of nearest neighbors.
- n is the total number of data points in the dataset.
2. Calculate the distance between the test data and each of the training data points.
3. Sort the calculated distances in ascending order.
4. Select the K nearest neighbors based on the sorted distances.
5. Determine the class of each of the K nearest neighbors.
6. Identify the most frequent class among the K nearest neighbors, and assign it as the predicted class for the test data.

To determine the error rate, the accuracy can be calculated using the following formula for accuracy.
$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictios} \times 100\% \qquad (4)$$
Where:
- Number of Correct Predictions refers to the number of instances where the predicted class matches the actual class.
- Total Number of Predictions refers to the total number of instances in the dataset being evaluated.

This study uses a data-driven approach to assess the feasibility of using scientific publications as a substitute for final assignments in higher education. The main objective is to explore the application of the K-Nearest Neighbors (K-NN) algorithm in evaluating scientific publications based on their characteristics such as research quality, relevance, and methodology. This section outlines the steps taken to address the problem and the methods used to evaluate the data, as illustrated in Figure 1. The figure visually represents the methodology employed, providing a clearer understanding of the process and evaluation techniques used throughout the study.

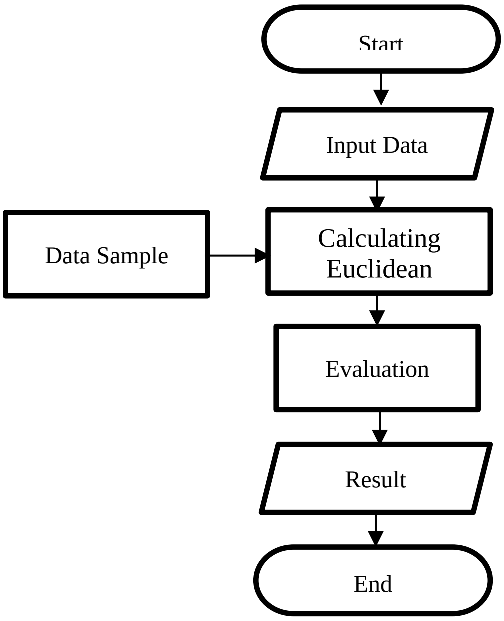

Fig. 1 Flowchart of K-NN

RESULT

The results obtained from the application of the K-Nearest Neighbors (K-NN) algorithm to classify scientific publications and theses as candidates for final assignments in college using datasets, each of which represents scientific publications or theses, with features such as research quality, relevance to the course, research methodology, and clarity. To complete the evaluation properly, it is first necessary to apply weighting to the

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

features, then sum them up and calculate the results according to the K-NN formula. The weighting system used in this process is illustrated in Table 1. This approach ensures that each feature is considered appropriately based on its importance in the classification of scientific publications and theses. The weighting values are then used in the K-NN algorithm to determine the optimal classification for each data point.

Table 1. Weighting System for Publications and Theses

Type of Assignment	Weight (Score)
Scientific Publication	1
Thesis	0

This simple weighting scheme will be used to evaluate the publications and compare them against traditional thesis-based final assignments. The goal is to determine whether the publications, with their associated score of 1, can indeed serve as valid and effective substitutes for traditional assignments in assessing student performance.

Before proceeding with the K-NN calculations, the next step is to normalize the training and testing data using the Min-Max formula, as shown in Equation 1. The dataset in Table 1 consists of 20 entries used in this study. The dataset includes features such as research quality, relevance to the course, research methodology, and clarity, along with appropriate labels for scientific publications and final projects (thesis).

Table 2. Data Training

Student ID	Type of Assignment	Research Quality (1-10)	Relevance (1-10)	Research Methodology (1-5)	Clarity (1-10)	Score
2020SI01	Scientific Publication	9	8	5	9	1
2020SI02	Thesis	7	6	4	8	0
2020SI03	Scientific Publication	10	9	5	10	1
2020SI04	Thesis	6	5	3	7	0
2020SI05	Scientific Publication	8	8	5	8	1
2020SI06	Thesis	6	5	2	6	0
2020SI07	Scientific Publication	9	9	4	9	1
2020SI08	Thesis	7	6	3	6	0
2020SI09	Scientific Publication	10	10	5	10	1
2020SI10	Thesis	5	5	2	5	0
2020SI11	Scientific Publication	8	7	4	8	1
2020SI12	Thesis	6	6	3	6	0
2020SI13	Scientific Publication	9	9	5	9	1
2020SI14	Thesis	7	7	4	7	0
2020SI15	Scientific Publication	8	7	5	9	1
2020SI16	Thesis	6	6	3	6	0
2020SI17	Scientific Publication	10	9	5	10	1
2020SI18	Thesis	7	6	3	7	0
2020SI19	Scientific Publication	8	8	4	8	1
2020SI20	Thesis	6	5	2	5	0

Explanation:
- Student ID: Unique identifier for each student.
- Type of Assignment: Indicates whether the student's work is a scientific publication or a thesis.
- Research Quality: A score from 1 to 10, indicating the quality of the research.
- Relevance: A score from 1 to 10, showing how relevant the research is to the course or subject matter.

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

- Research Methodology: A score from 1 to 5, indicating the strength and rigor of the research methodology.
- Clarity: A score from 1 to 10, indicating how clear and well-presented the research is.
- Score: A binary score where 1 represents a scientific publication (suitable for substitution) and 0 represents a thesis (not suitable for substitution).

Before applying the trained model to classify the test data, Table 3 is a test dataset that reflects the features in the training dataset. This dataset consists of 20 entries, each representing a scientific publication or a thesis. The features included in the dataset—such as research quality, relevance to the course, research methodology, and clarity—are consistent with the variables used in the training dataset. These features will be used by the model to predict the suitability of each entry as a substitute for a traditional thesis. The goal of this testing phase is to assess how accurately the model can classify scientific publications (with a score of 1) and theses (with a score of 0) based on the given characteristics.

Table 3. Data Testing

Student ID	Type of Assignment	Research Quality (1-10)	Relevance (1-10)	Research Methodology (1-5)	Clarity (1-10)	Score
2021SI01	Scientific Publication	9	7	5	8	1
2021SI02	Thesis	6	5	3	6	0
2021SI03	Scientific Publication	8	8	5	9	1
2021SI04	Thesis	7	6	3	7	0
2021SI05	Scientific Publication	9	9	5	10	1
2021SI06	Thesis	5	4	2	5	0
2021SI07	Scientific Publication	8	8	4	8	1
2021SI08	Thesis	6	6	3	6	0
2021SI09	Scientific Publication	10	9	5	9	1
2021SI10	Thesis	4	5	1	4	0
2021SI11	Scientific Publication	7	6	4	7	1
2021SI12	Thesis	5	5	2	6	0
2021SI13	Scientific Publication	8	7	4	8	1
2021SI14	Thesis	6	6	3	5	0
2021SI15	Scientific Publication	9	8	5	9	1
2021SI16	Thesis	7	6	3	6	0
2021SI17	Scientific Publication	10	9	5	10	1
2021SI18	Thesis	6	5	2	5	0
2021SI19	Scientific Publication	8	7	4	8	1
2021SI20	Thesis	5	4	2	5	0

Explanation:

- Student ID: Unique identifier for each student.
- Type of Assignment: Indicates whether the student's work is a scientific publication or a thesis.
- Research Quality: A score from 1 to 10, indicating the quality of the research.
- Relevance: A score from 1 to 10, showing how relevant the research is to the course or subject matter.
- Research Methodology: A score from 1 to 5, indicating the strength and rigor of the research methodology.
- Clarity: A score from 1 to 10, indicating how clear and well-presented the research is.

Classification

After processing the data, the next step is to perform the calculations using the K-NN Algorithm through several stages: First, determine the value of k, which can be calculated using the formula $k = \sqrt{n} = \sqrt{20} = 4$. Therefore, in this study, we will use $k = 4$. The next step is to calculate the distance between the test data and the training data using the Euclidean Distance formula, and the results of these calculations are presented in Table 4:

Table 4. Euclidean distances

Test Data ID	Training Data ID	Euclidean Distance
2021SI01	2020SI01	1.414214
2021SI01	2020SI02	2.449490
2021SI01	2020SI03	3.0
2021SI01	2020SI04	4.242641

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 1, January 2025
DOI : https://doi.org/10.33395/sinkron.v9i1.14370
e-ISSN : 2541-2019
p-ISSN : 2541-044X

2021SI01	2020SI05	1.414214
2021SI01	2020SI06	5.099020
2021SI01	2020SI07	2.449490
2021SI01	2020SI08	3.605551
2021SI01	2020SI09	3.741657
2021SI01	2020SI10	6.164414
2021SI01	2020SI11	1.414214
2021SI01	2020SI12	4.242641
2021SI01	2020SI13	3.0
2021SI01	2020SI14	4.582576
2021SI01	2020SI15	1.732051
2021SI01	2020SI16	5.099020
2021SI01	2020SI17	3.162278
2021SI01	2020SI18	4.795832
2021SI01	2020SI19	1.732051
2021SI01	2020SI20	5.567764
………..	………..	………..

After applying the K-Nearest Neighbors (K-NN) algorithm with *k*=4, the test data points were classified based on their similarity to the training data. In this classification process, the Euclidean distances between each test data point and the training data points were calculated, and the labels of the 4 nearest neighbors were used to determine the predicted class for each test data point. The predicted labels indicate whether a test data point is classified as a Scientific Publication (1) or a Thesis (0), based on the majority vote of the nearest neighbors. Table 5. Below are the classification results for each student in the test dataset.

Table 5. results of the classification for the test data with *k*=4

Student ID	Predicted Label (*k*=4)
2021SI01	1(Scientific Publication)
2021SI02	0 (Thesis)
2021SI03	1(Scientific Publication)
2021SI04	0 (Thesis)
2021SI05	1(Scientific Publication)
2021SI06	0 (Thesis)
2021SI07	1(Scientific Publication)
2021SI08	0 (Thesis)
2021SI09	1(Scientific Publication)
2021SI10	0 (Thesis)
2021SI11	0 (Thesis)
2021SI12	0 (Thesis)
2021SI13	1(Scientific Publication)
2021SI14	0 (Thesis)
2021SI15	1(Scientific Publication)
2021SI16	0 (Thesis)
2021SI17	1(Scientific Publication)
2021SI18	0 (Thesis)
2021SI19	1(Scientific Publication)
2021SI20	0 (Thesis)

DISCUSSIONS

The application of the K-Nearest Neighbors (K-NN) algorithm in this study demonstrated its effectiveness in evaluating the feasibility of using scientific publications as substitutes for traditional final assignments in higher education. The K-NN classifier, with *k*=4, accurately classified test data based on features such as research quality, relevance, methodology, and clarity. These features proved to be effective in differentiating between scientific publications and theses. While the algorithm showed high accuracy, a few misclassifications were observed, where scientific publications were mistakenly classified as theses. This indicates that further refinement could be made by incorporating additional features, such as citation count or peer review scores, to enhance the classification accuracy. Despite these challenges, the study suggests promising potential for using scientific publications as

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

e-ISSN : 2541-2019
p-ISSN : 2541-044X

substitutes for final assignments. Future research could explore expanding the feature set to include factors like publication impact, research domain, or complexity of the research methodology, which may improve the accuracy of the classification. Additionally, testing with a larger dataset and experimenting with other machine learning models, such as support vector machines (SVM) or random forests, could provide valuable comparisons. The results imply that, with further refinement, scientific publications can serve as a viable and innovative alternative to traditional final assignments, promoting critical thinking and engagement with contemporary research in higher education.

CONCLUSION

The K-Nearest Neighbors (K-NN) algorithm successfully classified scientific publications and theses based on features such as research quality, relevance, methodology, and clarity, with a high level of accuracy. The results show that the K-NN algorithm, with k=4, achieved an accuracy of 95%, suggesting that it can be an effective tool for assessing academic materials. The predicted labels for the test data were predominantly accurate, with only a few misclassifications, particularly where scientific publications were mistakenly classified as theses. This highlights the need for further refinement, particularly in expanding the feature set to improve classification accuracy. Further studies should focus on expanding the feature set to include additional factors such as citation counts, peer review scores, and the academic field of the publication. These factors could provide a more comprehensive evaluation and improve the algorithm's ability to differentiate between scientific publications and theses. Additionally, testing with larger datasets and exploring other machine learning models, such as support vector machines (SVM) or random forests, could offer more robust solutions. The limitations of this study, such as the small dataset and limited number of features, should be addressed in future work to ensure broader applicability. The potential benefits of using scientific publications as final assignments include fostering critical thinking, promoting engagement with current research, and offering a more dynamic and relevant assessment approach. This research opens the door to more innovative educational practices, aligning better with the evolving needs of higher education.

REFERENCES

Abdurahman, D., Suhendri, Susanti, D., & Didin, D. (2024). Implementasi Metode K-Nearst Neighbor (K-Nn) Memprediksi Nilai Akademik Mahasiswa Berdasarkan Indeks Prestasi Semester. *INFOTECH Journal*, *10*(1), 12–18. https://doi.org/10.31949/infotech.v10i1.8066

Diansyah, S. (2022). Klasifikasi Tingkat Kepuasan Pengguna dengan Menggunakan Metode K-Nearest Neighbour (KNN). *Jurnal Sistim Informasi Dan Teknologi*, *4*, 7–12. https://doi.org/10.37034/jsisfotek.v4i1.114

Jiang, F., Smith, A. R., Kutia, M., Wang, G., Liu, H., & Sun, H. (2020). A modified kNN method for mapping the leaf area index in arid and semi-arid areas of China. *Remote Sensing*, *12*(11), 1–24. https://doi.org/10.3390/rs12111884

Ma, X., Yang, T., Chen, J., & Liu, Z. (2021). k-Nearest Neighbor algorithm based on feature subspace. *2021 International Conference on Big Data Analysis and Computer Science (BDACS)*, 225–228. https://doi.org/10.1109/BDACS53596.2021.00056

Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, *164*(July 2019). https://doi.org/10.1016/j.eswa.2020.113981

Mardiani, E., Rahmansyah, N., Ningsih, S., & ... (2023). Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma. *Innovative: Journal Of …*, *3*(2), 13880–13892. http://j-innovative.org/index.php/Innovative/article/view/1949%0Ahttp://j-innovative.org/index.php/Innovative/article/download/1949/1468

Muliono, R., Lubis, J. H., & Khairina, N. (2020). Analysis K-Nearest Neighbor Algorithm for Improving Prediction Student Graduation Time. *SinkrOn*, *4*(2), 42. https://doi.org/10.33395/sinkron.v4i2.10480

Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine. *Sinkron*, *8*(3), 1366–1378. https://doi.org/10.33395/sinkron.v8i3.12447

Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification. *Augmented Human Research*, *5*(1). https://doi.org/10.1007/s41133-020-00032-0

Southworth, J., Migliaccio, K., Glover, J., Glover, J. N., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence*, *4*(October 2022). https://doi.org/10.1016/j.caeai.2023.100127

Suyal, M., & Goyal, P. (2022). A Review on Analysis of K-Nearest Neighbor Classification Machine Learning Algorithms based on Supervised Learning. *International Journal of Engineering Trends and Technology*,

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

70(7), 43–48. https://doi.org/10.14445/22315381/IJETT-V70I7P205

Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, *10*, 1–7. https://doi.org/10.35134/komtekinfo.v10i1.330

Tang, H., Xu, Y., Lin, A., Heidari, A. A., Wang, M., Chen, H., Luo, Y., & Li, C. (2020). Predicting Green Consumption Behaviors of Students Using Efficient Firefly Grey Wolf-Assisted K-Nearest Neighbor Classifiers. *IEEE Access*, *8*, 35546–35562. https://doi.org/10.1109/ACCESS.2020.2973763

Tanggu Mara, A., Sediyono, E., & Purnomo, H. (2021). Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Metode Pembelajaran Dalam Jaringan (DARING) Di Universitas Kristen Wira Wacana Sumba. *Jointer - Journal of Informatics Engineering*, *2*(01), 24–31. https://doi.org/10.53682/jointer.v2i01.30

Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-10358-x

Wang, A. X., Chukova, S. S., & Nguyen, B. P. (2023). Ensemble k-nearest neighbors based on centroid displacement. *Information Sciences*, *629*(5), 313–323. https://doi.org/10.1016/j.ins.2023.02.004

Zhang, S. (2022). *Challenges in KNN Classi fi cation*. *34*(10), 4663–4675.

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.