

Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree

Melati Purba Bestari^{1)*} , Imam Yuadi²⁾ 

¹⁾²⁾Information and Library Science, Faculty of Social and Political Sciences, Airlangga University, Surabaya, Indonesia

¹⁾melati.purba.bestari-2024@fisip.unair.ac.id, ²⁾imam.yuadi@fisip.unair.ac.id

Abstract

Background: Social media has become an effective tool in promoting library services. This study uses the Decision Tree algorithm to predict student participation in the University of Muhammadiyah Malang library based on their social media activities.

Objective: The data used included features such as the type of social media (Instagram, TikTok, YouTube, Twitter), frequency of library visits, and level of social media engagement.

Methods: The decision tree model is processed using Orange Data Mining, resulting in a clear separation between participants and non-participants based on a combination of these features.

Results: The study's findings indicate that social media, particularly Instagram and TikTok, have a significant impact on student participation in the library. The accuracy of the obtained model is about 76.7%, indicating that decision trees are an effective method for predicting library participation.

Conclusion: This research provides valuable insights for designing strategies to increase library student engagement based on social media analysis

Keywords: Student participation, social media, Instagram, TikTok, decision tree, machine learning

INTRODUCTION

Library social media is increasingly recognized as a strategic tool to build brand loyalty, improve two-way communication, and increase library use, particularly in university settings (Okolo, 2020). As part of the knowledge infrastructure, academic libraries play a crucial role in disseminating collections, supporting information literacy, and facilitating the learning process (Rubin & Rubin, 2020). In the Indonesian higher education context, university libraries are integral to implementing the University's Tri Dharma program, serving students, lecturers, and staff through curated collections and facilities designed to meet evolving information needs (Harsana, 2009; Rosa & Storey, 2016)

The digital transformation has fundamentally altered how libraries engage with their academic communities. Social media platforms, such as Instagram, TikTok, Facebook, and WhatsApp, have emerged as powerful channels for communication, offering temporal and spatial flexibility (Aswad, 2018; Subono, 2017). For university libraries, these platforms offer opportunities to promote services, facilitate two-way interaction between librarians and users,

* Corresponding Author

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

and ultimately increase student participation in library activities (Parihar & Amte, 2024; Ritonga et al., 2022). Research has shown that students who are more active on platforms like Instagram and Twitter exhibit higher engagement in library activities, whereas those with limited social media access participate less frequently in library events (Smith et al., 2018). This digital engagement is particularly relevant in Indonesia, where most students belong to Generation Y and Generation Z—demographics that are inherently social media-savvy (Utomo, 2022).

Despite the recognized importance of social media in library promotion, existing research has primarily focused on descriptive analyses of social media as a promotional tool (Johnston & Lee, 2019; Parihar & Amte, 2024). While these studies provide valuable insights into platform usage and engagement strategies, few have employed predictive machine learning approaches to quantitatively measure and forecast the impact of social media on library participation. This gap is particularly pronounced in the Indonesian context, where academic libraries have yet to systematically leverage data-driven models to understand the relationship between social media engagement patterns and actual library usage.

Machine learning techniques, specifically decision tree algorithms, offer powerful capabilities for analyzing complex data patterns and making predictions based on multiple variables such as social media platform preference, engagement frequency, and library visit patterns (Roshini et al., 2019). Decision trees provide transparent decision flows and can handle both categorical and numerical data efficiently, making them ideal for classification tasks in educational contexts (Aglin et al., 2020). However, the application of such predictive models to measure the effectiveness of social media in driving student participation in Indonesian university libraries remains underexplored.

This study addresses the critical need to understand and predict the relationship between social media engagement and student participation in academic libraries using a data-driven approach. Focusing on the University of Muhammadiyah Malang library, this research employs a decision tree classification model to analyze survey data encompassing students' social media usage patterns (with emphasis on Instagram and TikTok), library visit frequency, and engagement levels with library-generated social media content.

The primary objectives are threefold: (1) to identify which social media platforms and engagement patterns most significantly influence student library participation; (2) to evaluate the performance of decision tree algorithms as a predictive tool for forecasting student engagement in library activities; and (3) to provide evidence-based recommendations for library managers to optimize social media strategies for enhancing student outreach and participation. By combining social media analytics with machine learning classification, this study aims to bridge the gap between descriptive promotional strategies and predictive, data-driven library management practices, offering a replicable framework for academic libraries in Indonesia and similar contexts to maximize the impact of their digital engagement efforts.

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

LITERATURE REVIEW

Social media has become an effective tool for promoting library services, particularly among millennials and Gen Z, who are accustomed to digital technology. A study on the influence of library promotion through Instagram social media on visitor interest reveals that promotion through Instagram can attract attention, arouse interest, foster desire, and encourage users to take action and visit the library. This underscores the importance of social media as an effective communication strategy to increase the use of libraries in the digital (Dari & Zulaikha, 2023). Social media has become an essential part of students' daily lives. According to research conducted by students who are active on social media platforms such as Instagram and Twitter, they tend to participate in more activities on campus, including library activities. Student engagement on social media improves their academic and extracurricular involvement by allowing them to connect and collaborate outside the classroom (Smith et al., 2018).

In addition, active engagement on social media is positively correlated with student participation in various academic activities, including events held by university libraries (Johnston & Lee, 2019). Library programs that use social media to promote their activities and services tend to attract more participants. Therefore, a significant relationship exists between social media use and student participation in library programs. This is especially true in cases where social media platforms are used to communicate and disseminate information related to library programs.

Machine learning (ML) is gaining popularity in education, and numerous studies focus on how it can predict student participation. Two techniques often used in research are decision trees and naïve Bayes to model and predict student behaviour based on existing data. Their study on decision trees to predict student participation in libraries found that decision trees can accurately distinguish participating students from those who do not. This model categorises students based on various things, such as social media, library visits, and other academic activities. Decision trees help identify decision-making paths based on variables such as satisfaction with website information, admin interactions, and relevant content (Tan & Walker, 2020). The decision tree method is used because of its ability to simplify the classification process into an easy-to-understand visual form (Kusumawardani et al., 2022).

In the context of libraries, Decision Trees provide many advantages regarding model interpretation because the classification results can be explained visually through the decision tree's structure. This makes it easier for library managers to understand what variables affect student participation in the activities held. Decision Trees are the perfect choice for diverse datasets because they effectively handle data involving numerical and categorical variables. In addition, research conducted by Smith shows that algorithms such as Random Forest and Gradient Boosting can also produce more stable predictions of student participation. Random Forest, an ensemble method of decision trees, can reduce the overfitting problem that often occurs in single decision trees by combining the results of multiple decision trees. (Smith et al., 2018)

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

With the right approach, social media can potentially improve the quality of library services and reach a wider audience, making it a key tool in developing modern libraries in the rapid technological developments (Amaliah & Sumarno, 2021). Students who are actively involved in these sites are more likely to visit the library and participate in activities held there, thereby creating a positive relationship between social media use and library visit frequency. This means that the student participation prediction model relies on data about the time spent on social media and the platforms used. College students who spend more time on platforms like Instagram and TikTok are more likely to engage in activities shared by libraries on those platforms, which increases their likelihood of participating in library activities (Gunter et al., 2019). Library visits are one of the main components that influence student participation. In this situation, social media demonstrates engagement, helps identify students more likely to participate in library programs, and encourages students to be active, interactive, and raise awareness to utilise library services. Libraries often utilise social media platforms like Instagram and TikTok to attract students and promote their events (Vovk & Malesa, 2024).

The Decision Tree algorithm can predict the relationship of certain variables through data mining. In predicting the influence of social media on students' choices in the Regional Elections, this algorithm can map factors such as gender, age, and type of social media platform (Facebook, Instagram, Twitter). The process involves data collection, preprocessing, model formation, and model accuracy testing with decision tree visualization that makes it easier to interpret the results. A decision tree is a handy tool for data classification, supporting pattern- and rule-based decision-making in datasets. Decision trees are the most commonly used machine learning techniques in classification and prediction. It works by dividing data according to specific standards to produce clear and easy-to-understand decisions. Based on research conducted by Mendonça et al. (2022) It was stated that the use of decision trees is appropriate for showing participation predictions because it is able to provide a visualization of the factors that influence student participation and identify key variables that influence participation. Tan & Walker say that Decision Trees are advantageous in visualizing results because they allow you to see the relationships between variables that influence decisions. Therefore, the decision tree used in this study was based on its ability to create a simple yet highly effective model for predicting student participation based on various features. (Tan & Walker, 2020)

METHODS

This study employs a quantitative predictive research design using a machine learning approach, specifically the Decision Tree classification algorithm, to model and forecast student participation in library activities based on social media engagement patterns and library usage behaviors. The machine learning framework enables pattern recognition in complex behavioral data and generates probabilistic predictions that can inform evidence-based library management strategies. The methodology of this study employs the Decision Tree model to predict student participation in library activities at the University of Muhammadiyah Malang, based on their social media activities. Data collection and processing, application of machine learning models, and evaluation of model performance are all parts of this process. Data collection and processing, application of machine learning models, and evaluation of model performance are

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

all parts of this process. Decision trees can be useful for handling large data sets and provide a clear visual representation of the decision-making process that aims to determine operational roles and strategies to achieve goals (Beay & Sarimole, 2024). The use of machine learning algorithms in this decision tree involves various sets of rules that serve as considerations for decision-making, structured in a tree-like format that enables the prediction of results, resources, and potential risks or consequences. The presentation of this decision tree is depicted in branches that can represent steps for decision-making. (Mardiani et al., 2023)

This methodology uses Orange Data Mining to apply machine learning techniques. The process is described in detail below.

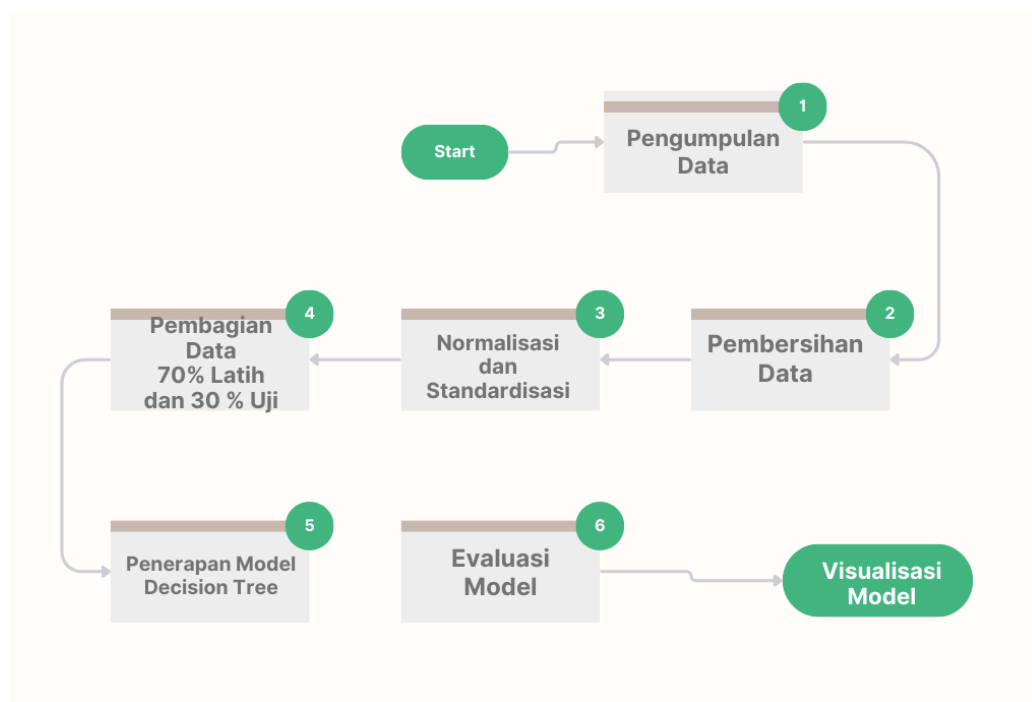


Figure 1. Data processing flowchart using a decision tree
Source: by the Author

Figure 1 describes this quantitative study, which uses machine learning methods to predict student participation in the library program of the University of Muhammadiyah Malang based on the data collected, namely the results of a survey conducted by the library related to the use of student social media to obtain information about the UMM library. The variables measured, including the type of social media used, the frequency of visits to the library, and the level of engagement on social media, are used in a decision tree model, or decision tree, to predict whether students will participate in library activities. The data for this study were collected through a survey conducted on students at a university, which included variables, namely the type of social media. Students were asked to choose social media platforms such as Instagram, TikTok, Twitter, and YouTube, which they used most often. Second, regarding the frequency of library visits, information about how much time students usually spend visiting the library every week is needed. Third, the level of engagement on social media is measured

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

by looking at how much time is spent on social media and how people interact with relevant library content. Fourth, participation in the library (Target), which is a target variable that determines whether students take part in library activities, is classified as "Yes" or "No". This survey involved one hundred active students from various departments of the University of Muhammadiyah Malang.

The sampling technique employed was purposive sampling, where participants were selected based on specific criteria: (1) active enrollment status at the University of Muhammadiyah Malang during the survey period, (2) regular use of at least one social media platform, and (3) awareness of the university library's social media presence. This non-probability sampling approach was chosen to ensure that respondents possessed the necessary characteristics to provide meaningful data about the relationship between social media engagement and library participation, though it limits generalizability to the broader student population. Data Cleanup is the process of deleting incomplete or irrelevant data. We use the imputation method to replace missing values based on averages or feature modes if there are missing values. Conversion of categorical variables. Some variables, such as participation in libraries and types of social media, are categorical. Therefore, these variables are transformed into forms acceptable to machine teaching models. For example, the nominal variable for the type of social media is changed to this form. Normalization and standardization ensure that numerical variables such as frequency of visits and time spent on social media are on the same scale and that the model is unaffected by scale differences between features.

Seventy percent of the data is used to train the model and build a decision tree. Thirty percent of this data is used to test the model that has been taught, which allows us to examine the model's performance on data that has never been seen before. Orange Data Mining uses the Data Sampler widget to split the data. Application of the Decision Tree Model: This model is applied to predict student participation in the library after the data is prepared. Orange Data Mining is the tool used to implement this model, which offers a variety of configurations for decision trees, i.e., the most expansive tree depth limit to prevent the tree from being overcrowded. Minimum amount of data in a leaf: To determine how much data should be in each decision tree leaf, see this section. To ensure that each leaf has enough data to make an informed decision, this value is set to 2. Pruning To prevent the model from becoming too complex, the decision tree model also uses pruning. The parameter used is the Induce binary tree to ensure that the decision tree only divides the data into two categories: "Yes" (participation) and "No" (non-participation).

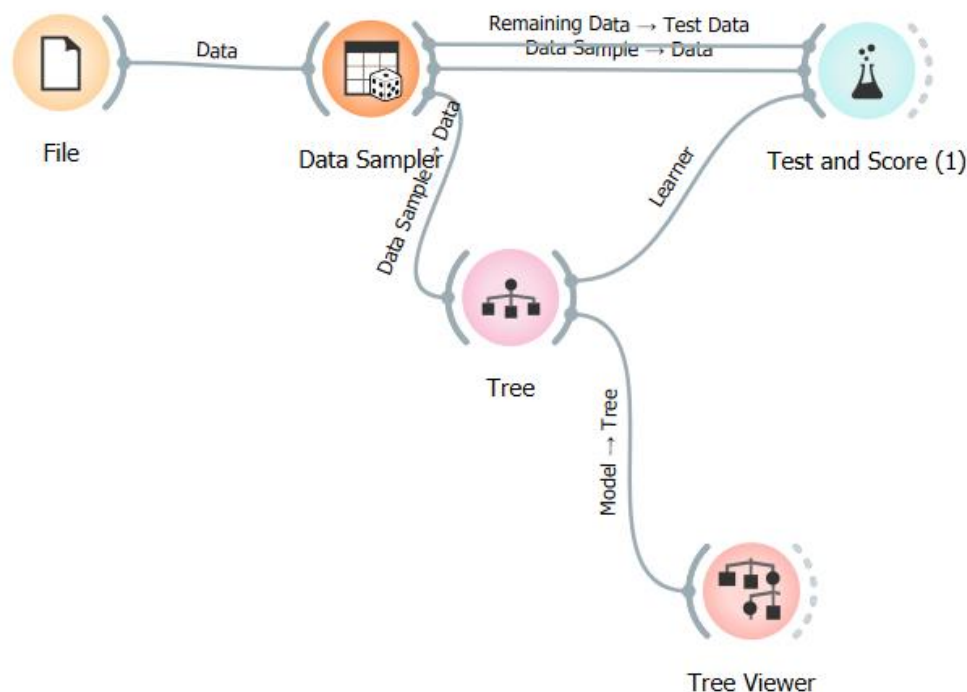


Figure 2. Data Analysis Workflow with Decision Tree Model
Source: by Authors from Orange Data Mining Process

Figure 2 describes model Evaluation; once the decision tree model is trained, the model's performance is evaluated using the Test & Score widget in Orange Data Mining. The primary metrics for assessing model performance are accuracy, precision, repeatability, and F1 scores. Additionally, the chaos matrix calculates the accuracy of predictions and determines whether the model tends to predict more of the majority or minority class. Test & Score compares the results of the decision tree model with test data and shows how well the model predicts student participation based on its features. This study uses the Tree Viewer at Orange Data Mining to visualize the decision tree structure and learn more about the model's decisions. This visualization shows how the model divides data based on specific features, such as social media and the frequency of visits to the library, to make decisions. The methodology used in this study integrates machine learning with data analysis to predict student participation in the library. Using decision tree models and other machine learning techniques, this study aims to provide better insights into the factors influencing student participation in libraries and develop data-driven strategies to enhance student engagement in libraries.

FINDINGS

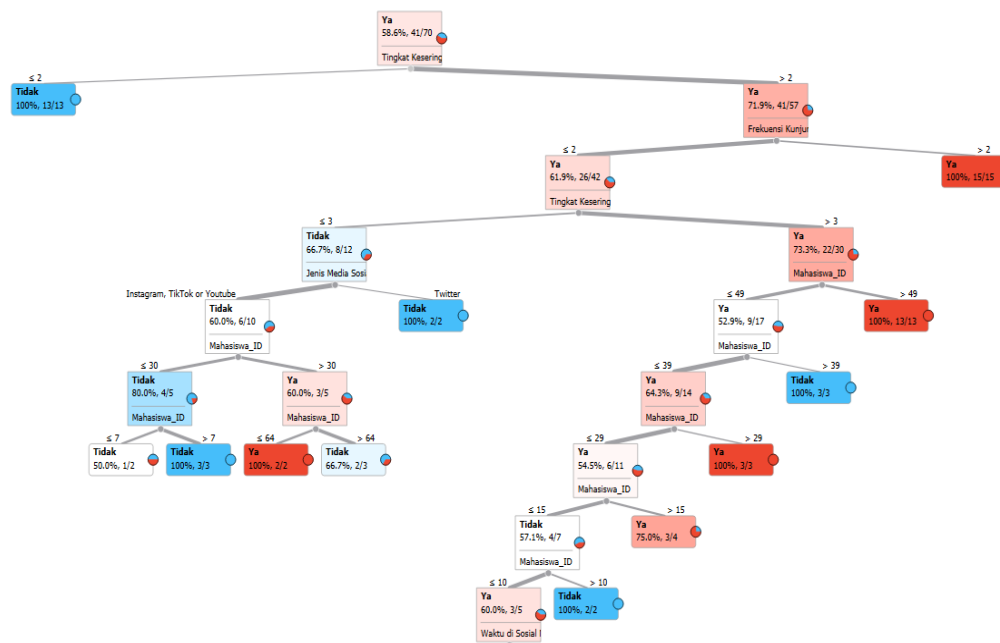


Figure 3. Decision Tree for Student Participation Based on Social Media Activity
Source: by Authors from Orange Data Mining Process

The results of the decision tree analysis from Figure 3 show several branches that are very clear in classifying student participation in the library based on various variables used. Here are the details of the resulting decision tree structure:

Decision Tree Structure

The results of the decision tree analysis reveal several distinct branches that effectively classify student participation in the library based on various variables. Here are the details of the resulting decision tree structure. Frequency Rate of Watching Library Content ≤ 2 : Students with a Frequency Rate of Watching Library Content ≤ 2 are predicted to be 100% not participating in library activities. In this group, all 13 cases were not involved. This 100% non-participation rate reveals a critical engagement threshold that functions as a gateway to library activities—students below this threshold cannot be converted through platform optimization or event design alone, requiring alternative outreach mechanisms such as classroom partnerships or peer referral programs. The finding suggests a fundamental "awareness gap" rather than an "engagement gap," where students are not consuming library content at all, necessitating different intervention strategies than those used for already-engaged students. This suggests that students with low engagement on social media or those who lack sufficient interest in participating in library activities are less likely to participate. Frequency of Watching Library Content > 2 : Students with a male-to-female ratio of 2 or More are more engaged in social media and are likely to participate in library activities. The next branch divides students based on the type of social media they use.

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

Types of Social Media and Participation

Instagram, TikTok, or YouTube: Students who actively use these platforms will likely participate in library activities. For example, students with a *Mahasiswa_ID* ≤ 30 are less likely to participate (80% of the predictions are "No"). However, students with a *Mahasiswa_ID* of 30 or more who use platforms like Instagram or TikTok are predicted to participate in library activities at a rate of 100% (2/2 cases). This indicates that the type of platform used influences student engagement in the library, with visual platforms, such as Instagram and TikTok, being more commonly used to promote library events. Twitter students who are more active on Twitter tend to be 100% non-participating in library activities, with 2/2 of the cases classified as non-participating. Despite its popularity, the use of Twitter does not significantly impact student participation in library activities. This could be due to Twitter's more open and fast-paced content, unlike platforms that focus more on visual content, such as Instagram and TikTok. The complete failure of Twitter (0/2 participation) versus the high effectiveness of Instagram/TikTok (100% for $ID > 30$) demonstrates a critical platform-content fit mismatch, suggesting that library promotional strategies must align with platform-specific user behaviors and content formats rather than pursuing universal social media presence. This finding challenges the assumption that mere platform popularity translates to effective engagement, indicating that libraries should strategically reallocate resources toward visual platforms that demonstrably drive participation, rather than maintaining a presence across all channels.

This platform-specific effectiveness can be contextualised through the nature of library promotional content, which is inherently visual, featuring book displays, facility tours, event announcements, and student testimonials that are optimally conveyed through images and short-form videos rather than text-based posts. Instagram and TikTok's algorithmic prioritisation of visual storytelling and their user bases' expectations for engaging, scroll-stopping content align naturally with library marketing needs, whereas Twitter's rapid-fire, text-heavy discourse format creates friction for promotional messages that require sustained attention and visual appeal. Furthermore, the demographics of each platform's dominant user base matter: Generation Z students, who constitute the majority of the university population, demonstrate higher engagement rates with visual and video content on Instagram and TikTok, while Twitter usage among this cohort has declined in favour of more visually-oriented platforms.

Frequency of Library Visits and Participation

The frequency of student visits to the library is a very important factor in predicting whether they will participate in library activities. The frequency of visits > 2 shows that students who visit the library more than twice a week have a high probability of participating. They are more likely to participate in activities held by the library, resulting in a higher participation rate. For example, college students with a visit frequency of more than 2 and a *Mahasiswa_ID* of ≤ 49 had a 73.3% chance of participating in 22 out of 30 cases. The 26.7% non-participation rate among frequent library visitors (8 out of 30) represents a significant missed opportunity—these students have already overcome barriers of physical access and library comfort yet remain uninvolved in activities. This "passive user" segment may be the lowest-hanging fruit for conversion, requiring only targeted in-library recruitment (posters, staff invitations, strategic event scheduling) rather than fundamental behavior change.

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

Model Accuracy and Evaluation

The decision tree model applied to the data yielded an accuracy of 76.7%, indicating that it correctly classified 76.7% of the data tested. This indicates that, based on the features provided, the decision tree can effectively differentiate between participating and non-participating students in the library. The model's performance metrics show that Precision (for "Yes" or participation classes) in 77.8% of the models is good enough to predict that students who are expected to participate are indeed participating in the library. Recall (for "Yes" or participation classes) 82.4% of the model successfully identified most of the participating students with a reasonably high recall rate. F1-Score 80% of F1-Score scores show a balance between precision and recall, which means the model works well in handling participation predictions.

Confusion Matrix

Based on the confusion matrix, the distribution of participation predictions is as follows:

TABLE 1.
Confusion Matrix Participation Prediction

	Predicted: Tidak	Predicted: Ya
Actual: No	9 (True Negatives)	4 (False Positives)
Current: Already	3 (False Negatives)	14 True Positives)

Source: by Authors

True Positives (TP): 14 (Students who participated and were predicted to participate). True Negatives (TN): 9 (Students who do not participate and are predicted not to participate). False Positives (FP): 4 (Students who did not participate but were predicted to participate). False Negatives (FN): 3 (Students who participated but were predicted not to participate). The asymmetric error patterns provide actionable insights: four false positives suggest that some students exhibit all behavioral markers of participants yet do not attend, warranting qualitative research to identify final conversion barriers such as scheduling conflicts or event-topic misalignment. Three false negatives indicate that alternative participation pathways exist outside the model—potentially peer invitation networks or course requirements—suggesting that libraries should diversify their recruitment strategies beyond dependence on social media. The model's 76.7% accuracy implies that approximately 23% of the variance in participation remains unexplained by social media and visit frequency, pointing to unmeasured factors such as academic workload, event characteristics, or individual motivation.

Analysis of Factors Influencing Participation

The results of this study indicate that social media factors and the frequency of library visits are the primary factors influencing student participation in the library. Students who are active on Instagram and TikTok participate more frequently in library activities, whereas students who are more active on Twitter tend to be less engaged. The frequency of visits to the library more than 2 times per week is directly related to student participation in library activities. This study demonstrates that the decision tree is an effective model for predicting student participation in the library, based on social media activities and the frequency of library

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

visits. With an accuracy of 76.7%, this model offers valuable insights for libraries to develop effective strategies for increasing student participation.

DISCUSSION

The study reveals that social media, particularly Instagram and TikTok, as well as the frequency of library visits, are the primary factors influencing student participation in the library. In this case, according to what Castells revealed, technology can affect the behavior of the information society (Castells, 2010). On the other hand, the study's results also show that students who are active on visual platforms such as Instagram and TikTok are more likely to be involved in events organized by the library.

Active engagement on social media is positively correlated with student participation in various academic activities, including events hosted by university libraries (Johnston & Lee, 2019). This study reveals a positive correlation between student activity and library social media participation, which can influence student participation in library visits and activities. This is predicted because the type of content promoted on these platforms is more visual and accessible, and directly related to academic and extracurricular activities on campus. Meanwhile, students who visit the library more often are more likely to participate in library programs. This indicates that physical access to the library plays a crucial role in enhancing student involvement in library activities.

CONCLUSIONS

This study uses the Decision Tree model to predict student participation in the library based on social media activities and the frequency of visits. Based on the results obtained from data analysis with Orange Data Mining. The results of the decision tree analysis revealed that the types of social media used by students, particularly Instagram and TikTok, as well as the frequency of library visits, played a significant role in predicting student participation in the library. Students who are more active on Instagram and TikTok tend to participate more in activities organized by the library. The decision tree model used in this study has an accuracy of 76.7%. With a precision of 77.8%, a recall of 82.4%, and an F1 score of 80%, this model can predict student participation in the library with a reasonably high success rate. This metric indicates that the model effectively classifies students who participate in and those who do not participate in the library. Students who visit the library more than twice a week are more likely to participate in library activities. This indicates that the frequency of library visits is a significant factor in determining student participation.

Based on the confusion matrix, most True Positives (TPs) (students who participated and were correctly predicted to do so) were successfully identified. However, there are False Positives (FPs) (students who did not participate but were predicted to participate) and False Negatives (FN) (students who participated but were predicted not to participate), indicating room for improvement in the model. Some of the disadvantages of the model include the possibility of overfitting, especially on tree branches that are too deep. The use of more features or more complex models can improve prediction results and reduce errors.

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

Although the results of this study are promising, several areas need improvement, including mitigating the risk of overfitting the model, especially when the decision tree is too deep. For this reason, settings such as pruning trees or using ensemble models (e.g., Random Forest or Gradient Boosting) can help improve the model's results. Adding other features, such as the type of activities students participate in or the level of satisfaction with library services, can improve the model's accuracy. More extensive and diverse datasets can help improve model generalization and reduce bias in the data.

AUTHOR CONTRIBUTIONS

[Melati Purba Bestari]: Conceptualization, methodology, writing the original draft, review and editing, supervision. [Imam Yuadi]: Software development, investigation, data curation, and writing the original draft.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no specific grant from any funding agency.

ACKNOWLEDGMENTS

Thank you to the University of Muhammadiyah Malang library has become a place for researchers to develop this journal research. Hopefully, this research can make a major contribution to the advancement of social media to increase user participation at the library and using machine learning for analysis to help make decisions or develop a strategy for developing the library in the future. Hopefully, this research can make a major contribution to the advancement of technology in Indonesia

REFERENCES

- Aglin, G. A., Nijssen, S., & Schaus, P. (2020, June 16). PyDL8.5: a Library for Learning Optimal Decision Trees. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20) Demonstrations Track*. <http://arxiv.org/abs/1606.05386>
- Amaliah, E., & Sumarno. (2021). Intensifikasi Penggunaan Media Sosial untuk Mewujudkan Perpustakaan Modern Di Era Digital. *Jurnal El-Pustaka*, 2(1). <https://doi.org/10.23887/XXXXXX-XX-0000-00>
- Aswad, E. (2018). The Future of Freedom of Expression Online. *Duke Law & Technology Review*, 17(1).
- Beay, R., & Sarimole, F. M. (2024). Application of Decision Tree Method for Sales Prediction at PT. Cipta Naga Semesta (Mayora Group) North Jakarta for 2023. *International Journal*

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

Software Engineering and Computer Science (IJSECS), 4(3), 943–952.
<https://doi.org/10.35870/ijsecs.v4i3.2999>

Castells, M. (2010). *The rise of the network society* (2nc ed.). Wiley-Blackwell.

Dari, S. W., & Zulaikha, S. R. (2023). Pengaruh Promosi Perpustakaan melalui Media Sosial Instagram terhadap Minat Kunjung Pemustaka di Perpustakaan Universitas Mercu Buana Yogyakarta. *Al-Kuttab: Jurnal Kajian Perpustakaan, Informasi Dan Kearsipan*, 5(1).

Gunter, W., Davis, P., & Hamil, S. (2019). Social Media Engagement and Library Participation: A Quantitative Analysis. *Library Analytics Review*, 31(4).

Harsana, L. (2009). *Kamus Kepustakawanan Indonesia*. Pustaka Book Publisher.

Johnston, K. M., & Lee, L. (2019). The impact of social media on student engagement in library services. *Journal of Library & Information Science*, 45(2).

Kusumawardani, H. H., Rosyadi, I., Artanto, F. A., Arzha, F. I., & Rachmayani, N. A. R. (2022). Analisis Decision Tree dalam Pengaruh Digital Marketing terhadap Penerimaan Siswa Baru. *Remik: Riset Dan E-Jurnal Manajemen Informatika Kompute*, 6(2).

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 Social Science Research*, 3(2).

Mendonça, Y. V. S., Naranjo, P. G. V., & Pinto, D. C. (2022). The Role of Technology in the Learning Process: A Decision Tree-Based Model Using Machine Learning. *Emerging Science Journal*, 6(Special Issue), 280–295. <https://doi.org/10.28991/ESJ-2022-SIED-020>

Okolo, E. S. (2020). Marketing the Library and It Services with the Use of Social Media: An Overview. *Information and Knowledge Management*, 10(2).
<https://doi.org/10.7176/ikm/10-2-04>

Parihar, M., & Amte, S. (2024). Marketing of Library Services Through Social Media. *GAP Bodhi Taru a Global Journal of Humanities*, 7(1). <https://www.gapbodhitaru.org/>

Ritonga, A. R., Dalimunthe, M. A., Veronica, A., Ginting, D. C. U., & Nur'aini. (2022). The Effectiveness of Social Media as a Promotional Medium of The University of Sumatera Utara (USU) Library The Effectiveness of Social Media as a Promotional Medium of The University of Sumatera Utara (USU) Library. *TALENTA Conference Series: Local Wisdom, Social, and Arts*. <https://doi.org/10.32734/lwsa.v6i3.1763>

Rosa, K., & Storey, T. (2016). American libraries in 2016: Creating their future by connecting, collaborating and building community. *Sage Journals*, 42(2).

Cite this article: Bestari, M. P., & Yuadi, I. (2025). Prediction of Student Participation in the Library of the University of Muhammadiyah Malang Based on Social Media Activities Using Decision Tree. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi dan Kearsipan*, 11(2), 29-42. <http://doi.org/10.14710/lenpust.v11i2.74771>

- Roshini, T., Sireesha, P. V., Parasa, D., & Bano, S. (2019, September 28). Social media survey using decision tree and naive Bayes classification. *Proceedings of the 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT 2019)*.
- Rubin, R. E., & Rubin, R. G. (2020). *Foundations of Library and Information Science* (5th ed.). American Library Association.
- Smith, A., Jenkins, M., & Young, R. (2018). Social media and university student engagement: A study of Instagram and Twitter. *Social Media Studies*, 22(1).
- Subono, N. I. (2017). *Dari Adat ke Politik: Transformasi Gerakan Sosial di Amerika Latin*. Marjin Kiri.
- Susilowati, G. A. (2019). Pengaruh Sosial Media Dalam Membentuk Persepsi Pemuda Terhadap Partisipasi Politik. *Ikon --Jurnal Ilmiah Ilmu Komunikasi*, 23(1).
- Tan, J. M., & Walker, J. L. (2020). Using machine learning for predicting library participation in higher education: A case study of decision trees. *Library Analytics Review*, 31(4).
- Utomo, T. P. (2022). Optimalisasi Media Sosial untuk Pemasaran Perpustakaan Perguruan Tinggi. *Buletin Perpustakaan Universitas Islam Indonesia*, 5(1), 99–133.
- Vovk, N., & Malesa, R. (2024). Assessing the users' awareness level: social media and university libraries in the context of modern educational environment. *SCIA-2024: 3rd International Workshop on Social Communication and Information Activity in Digital Humanities*.