

Implementation of C4.5 Algorithm for Student Satisfaction Analysis on Scout Extracurricular Activities in the Framework of Scout Extracurricular Information System Development at SDN Pondok Bambu 10 & 11

Hanna Sabilla Rifai ^{a*}, Mesra Betty Yel ^b

^{a*,b} Information Systems Study Program, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, East Jakarta City, Special Capital Region of Jakarta, Indonesia.

ABSTRACT

This study investigates student satisfaction toward Scout extracurricular activities at SDN Pondok Bambu 10 and 11 by applying the C4.5 algorithm within the CRISP-DM framework. Data were collected from 210 students through questionnaires and interviews to evaluate perceptions of program quality, mentor support, and social interaction. The C4.5 model achieved an accuracy rate of 99.52%, effectively identifying key determinants of student satisfaction. Results indicate that program quality, mentor support, and peer interaction are the most influential factors shaping students' experiences. The decision tree produced interpretable rules that help educators understand satisfaction patterns and make data-driven improvements to program design. Compared with other methods such as SVM and Random Forest, C4.5 provides clearer interpretability while maintaining high predictive precision. The study further recommends integrating the model into a web-based information system to enable continuous monitoring and evaluation of extracurricular activities. These findings highlight the potential of data mining techniques to strengthen decision-making in education and to create a more adaptive, student-centered approach to extracurricular management.

ARTICLE HISTORY

Received 29 July 2025

Accepted 30 October 2025

Published 30 November 2025

KEYWORDS

C4.5 Algorithm; Student Satisfaction; Extracurricular Activities; Data Mining; Educational Management.

1. Introduction

Education is a long-term investment that requires sustained effort and substantial resources, forming the foundation for a nation's progress and resilience (Wurdianto, 2020). However, the declining quality of students over the years indicates persistent challenges in the educational process that demand immediate attention (Nurachman, 2020). One strategic approach to addressing these challenges is through extracurricular activities, which provide structured learning opportunities beyond formal classroom hours and are recognized as an essential part of holistic student development (Hasna Nur Alifah *et al.*, 2023). Such activities are instrumental in shaping students' character, discipline, and social competence (Rizal Rustandi, Adang Sudrazat, & Aam Ali Rahman, 2024). They also serve as an avenue for nurturing soft skills that are essential in social and professional life (Wahyu Ardias, Khairul Fajri, & Gusmanelli, 2024; Feraco *et al.*, 2022). Despite these benefits, the

effectiveness of extracurricular programs often depends on how well they are managed and integrated into the school's educational goals (Barumun, Sibuhuan, & Daulay, 2023). The 2013 curriculum classifies extracurriculars into mandatory and elective activities, emphasizing the need for proper implementation and evaluation to ensure their relevance and impact. Prior studies have shown that well-organized extracurriculars can enhance students' discipline and responsibility, fostering habits that support both academic and personal growth (Widiatmika, 2015; Cahya Neda Yani, Fitriani Prila Wardani, & Devy Riri Yuliani, 2020). Among these, Scout activities hold a particularly important role in instilling national values, patriotism, and self-reliance. As stipulated in Law No. 12 of 2010 Article 4, the Scout movement aims to cultivate individuals who are faithful, disciplined, and committed to preserving national integrity. Through Scouting, students learn responsibility, independence, and adherence to ethical principles (Yusdinar & Manik, 2023).

In the context of SDN Pondok Bambu 10 & 11, Scout extracurriculars are viewed as an effective platform for character formation and for improving the overall quality of school programs. Assessing student satisfaction toward Scout activities is therefore essential for understanding how well these programs meet students' expectations and contribute to their personal development. This study employs the C4.5 algorithm to analyze factors influencing student satisfaction with Scout extracurriculars, using data collected from 210 students through interviews and online questionnaires. The algorithm, a decision-tree-based machine learning method, is selected for its ability to identify key determinants and generate interpretable decision rules from complex datasets (Takalapeta, 2018; Sudrajat, 2022; Jannah, Sihombing, & Masrizal, 2023). By applying this approach, the research seeks to support the development of a data-driven information system that enables schools to monitor, evaluate, and improve the management of extracurricular activities effectively (Sinaga *et al.*, 2025). Accordingly, this study is titled "Implementation of the C4.5 Algorithm for Student Satisfaction Analysis on Scout Extracurricular Activities in the Framework of Scout Extracurricular Information System Development at SDN Pondok Bambu 10 & 11."

2. Methodology

This study employed a quantitative approach integrating data mining techniques and educational evaluation to analyze student satisfaction with Scout extracurricular activities at SDN Pondok Bambu 10 and 11. Data were collected through structured questionnaires, in-depth interviews, and surveys conducted with 210 students who actively participated in Scouting programs. The questionnaire was designed to capture students' perceptions of program quality, mentor support, and social interaction, as well as their overall satisfaction. Prior to full deployment, a pilot test was conducted with a small group of respondents to validate question clarity and ensure the instrument's reliability (Nurmala, Yusuf, & Sulaeman, 2023; Rezki *et al.*, 2019). The validated instrument was then distributed using Google Forms to ensure accessibility and efficiency in data collection. Complementary interviews with selected students were performed to obtain qualitative insights supporting the quantitative findings (Sa'diyah, Aulia, & Andharini, 2021). The research followed the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which consists of six systematic stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. In the business understanding phase, the study identified the importance of assessing student satisfaction as part of improving extracurricular management and student character formation (Pratiwi *et al.*, 2020; Mutiah, Daulay, & Nst, 2023). The data understanding phase involved reviewing raw responses and

identifying relevant variables such as program structure, mentor involvement, and student participation levels (Wurdianto, 2020; Yani, Wardani, & Yuliani, 2020). The data preparation stage included cleaning, normalizing, and transforming the collected data to remove noise and ensure consistency for computational processing. Data preprocessing followed standard procedures outlined in prior C4.5 studies on customer satisfaction modeling (Takalapeta, 2018; Jannah, Sihombing, & Masrizal, 2023), as well as educational data applications (Sudrajat, 2022). In the modeling phase, the C4.5 algorithm was employed to construct a decision tree capable of classifying satisfaction levels based on multiple influencing factors.

The decision tree was developed using *RapidMiner Studio 9.10*, a widely used tool for data mining analysis (Sinaga *et al.*, 2025). The algorithm applied entropy and information gain calculations iteratively to generate decision rules that distinguish between satisfaction categories with high accuracy. Data were divided into training and testing sets with an 80:20 ratio to ensure model reliability and minimize overfitting. Model evaluation was conducted using a *confusion matrix* to measure accuracy, precision, and recall values, following procedures from previous classification studies (Nurmala *et al.*, 2023; Rezki *et al.*, 2019). To strengthen data validity, results were cross-examined through triangulation between questionnaire outcomes and interview narratives, ensuring that quantitative patterns aligned with observed student experiences (Feraco *et al.*, 2022). Moreover, educational insights from extracurricular research emphasized the relevance of Scouting in building discipline, responsibility, and social skills (Yusdinar & Manik, 2023; Nurachman, 2020). The application of data-driven modeling in this educational setting thus bridges analytical precision with pedagogical interpretation. The research procedure can be visualized in Figure 1, which outlines the implementation of the CRISP-DM methodology, and Figure 2, which depicts the data processing workflow from data collection to model evaluation.

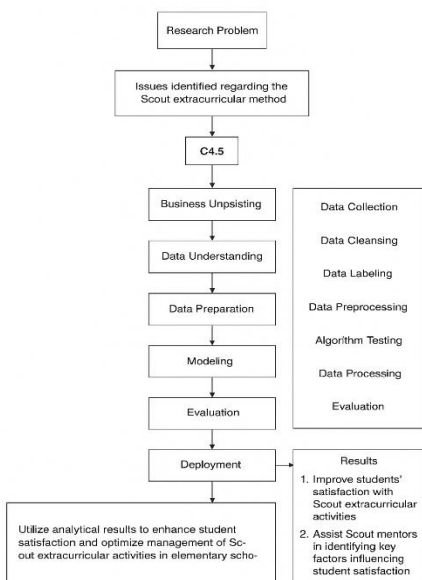


Figure 1. Implementation of the CRISP-DM Methodology

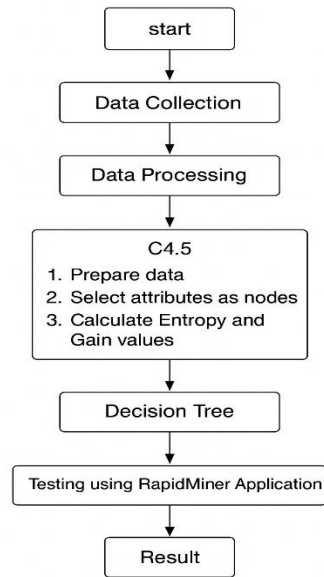


Figure 2. Research Flowchart of the C4.5-Based Student Satisfaction Analysis Process

Through these stages, the C4.5 algorithm enabled the identification of critical determinants affecting student satisfaction with Scout extracurricular activities. The findings are expected to support the creation of a web-based monitoring information system that enhances student participation and helps educators design programs aligned with learners' needs (Sinaga *et al.*, 2025).

3. Results and Case Study

The testing process in this study employed the Cross-Industry Standard Process for Data Mining (CRISP-DM) as a guiding framework, as it provides a structured and iterative approach that complements the C4.5 algorithm effectively. While both methods serve distinct purposes, their integration allows for more systematic modeling and evaluation. The six CRISP-DM stages—business understanding, data understanding, data preparation, modeling, evaluation, and deployment—were applied to ensure the analytical process remained comprehensive and replicable. In the business understanding phase, the study identified that evaluating student satisfaction toward Scout extracurricular activities is critical for improving program quality and effectiveness. Understanding the objectives of the program, students' needs, and the determinants of satisfaction enables schools to design activities that are more relevant and engaging. Moreover, a data-driven information system can enhance the management and monitoring of Scout programs, ensuring that activities align with educational goals and provide a meaningful experience for students (Wurdianto, 2020; Yusdinar & Manik, 2023). During the data understanding stage, data were collected from 210 students who participated in Scout extracurriculars during May 2025. The dataset included responses from structured questionnaires covering aspects such as program quality, mentor support, and peer interaction. These data were supplemented by interviews that provided contextual understanding of student experiences. The processed data distribution is illustrated in Figure 3, which presents the cleaned and categorized questionnaire responses.

	A	B	C	D	E	F	G	H	I	J	K	L	M
	Nama	KEIAS	Senang	Manfaat	Meningkatkan Penerima Diri	merasa lebih terlibat	menyayangkan dan menantang	meningkatkan keterampilan keorganisasian	disiplin dan bertanggung jawab	meningkatkan rasa percaya diri	rasa cinta terhadap alam	kemampuan berorganisasi	Tabel
1													
2	Syaidah	3	Setuju	Setuju	Setuju	Sangat Setuju	Sangat Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Setuju
3	Muhammad ari rafiq	4	Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Setuju	Sangat Setuju	Setuju	Setuju
4	Siti nurmahjidih ikhwani	2	Sangat Setuju	Setuju	Sangat Setuju	Netral	Sangat Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
5	Ara	6	Setuju	Netral	Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Setuju
6	Utari	2	Sangat Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Sangat Setuju	Setuju	Sangat Setuju	Setuju	Sangat Setuju
7	Ulhan Alisha Medina	2	Sangat Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
8	Aldiansyah Budi Laksmono	2	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
9	Gadisa rafani	2	Sangat Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
10	Gopur	5	Setuju	Sangat Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Setuju
11	Zira	5	Sangat Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Sangat Setuju	Setuju	Sangat Setuju
12	Muhammad aulia syaDil	2	Setuju	Setuju	Setuju	Netral	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju
13	Adrian Muhammad Khalid	5	Setuju	Setuju	Netral	Netral	Setuju	Setuju	Setuju	Netral	Tidak Setuju	Netral	Setuju
14	Alifandra Kencie Mangg	2	Sangat Setuju	Sangat Setuju	Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
15	Sabylla ennova andraya	5	Sangat Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Setuju
16	aulia khairunissa	2	Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Sangat Setuju	Setuju	Setuju	Setuju	Setuju
17	Abzar Arhan Muttaja	2	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
18	Kayla husbibah	5	Setuju	Netral	Netral	Netral	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju
19	Azzam Amrullah	2	Sangat Setuju	Sangat Setuju	Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
20	Naufal Ahyen Zulradli	5	Setuju	Sangat Setuju	Sangat Setuju	Netral	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
21	Fachri Giliran Alfari	2	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
22	ABIZARD RAUFA MALTAS	2	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju	Sangat Setuju
23	Muhammad Azri Firmans	2	Setuju	Setuju	Setuju	Netral	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju
24	Muhammad Ridwan	3	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju
25	Ananda Devandra Karim	2	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju	Setuju

Figure 3. Processed Questionnaire Results

The data preparation phase utilized the *Read Excel* operator in RapidMiner Studio 9.10 to import and preprocess data. This involved data cleaning, normalization, and transformation to prepare it for modeling. The dataset was then divided into training and testing subsets in an 80:20 ratio to ensure that the results were statistically reliable. The data partitioning output is presented in Figure 4.

Row No.	Tabel	prediction(T...	confidence_...	confidence_...	confidence_...	confidence_...	Nama	KELAS	Senan
1	Setuju	Setuju	1	0	0	0	Syaidah	3	Setuju
2	Setuju	Setuju	1	0	0	0	Muhammad...	4	Setuju
3	Sangat Setuju	Sangat Setuju	0	1	0	0	Siti nurhajah ...	2	Sanga
4	Setuju	Setuju	1	0	0	0	Ara	6	Setuju
5	Sangat Setuju	Sangat Setuju	0	1	0	0	Ulan	2	Sanga
6	Sangat Setuju	Sangat Setuju	0	1	0	0	Jihan Alisha ...	2	Sanga
7	Sangat Setuju	Sangat Setuju	0	1	0	0	Aldiansyah B...	2	Sanga
8	Sangat Setuju	Sangat Setuju	0	1	0	0	Gadisa rafani	2	Sanga
9	Setuju	Setuju	1	0	0	0	Copur	5	Setuju
10	Sangat Setuju	Sangat Setuju	0	1	0	0	Zara	5	Sanga
11	Setuju	Setuju	1	0	0	0	Muhammad...	2	Setuju
12	Setuju	Setuju	1	0	0	0	Adrian Muha...	5	Setuju
13	Sangat Setuju	Sangat Setuju	0	1	0	0	Alifandra Ke...	2	Sanga
14	Setuju	Setuju	1	0	0	0	sabylla ennov...	5	Sanga
15	Setuju	Setuju	1	0	0	0	aulia khaerun...	2	Setuju

Figure 4. Data Division into Training and Testing Sets

In the modeling phase, the C4.5 algorithm was implemented to generate a decision tree model that identifies the key factors influencing student satisfaction. The model construction process is shown in Figure 5, which visualizes the flow of data from import to classification. Figure 6 displays the resulting decision tree model generated by RapidMiner.

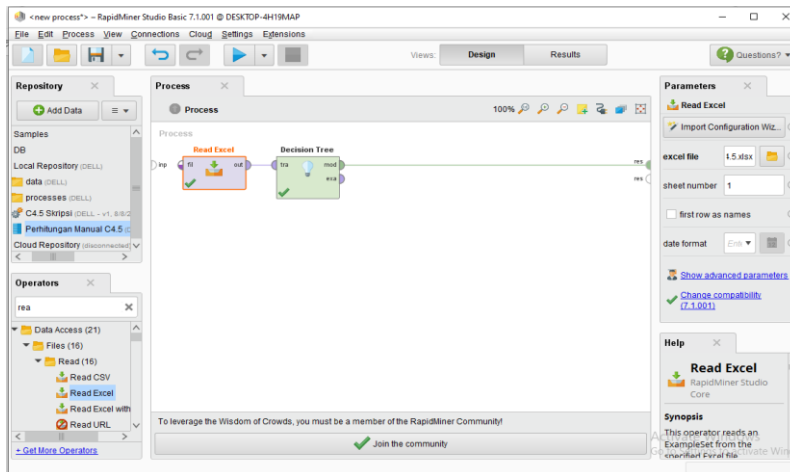


Figure 5. Data Mining Model Workflow

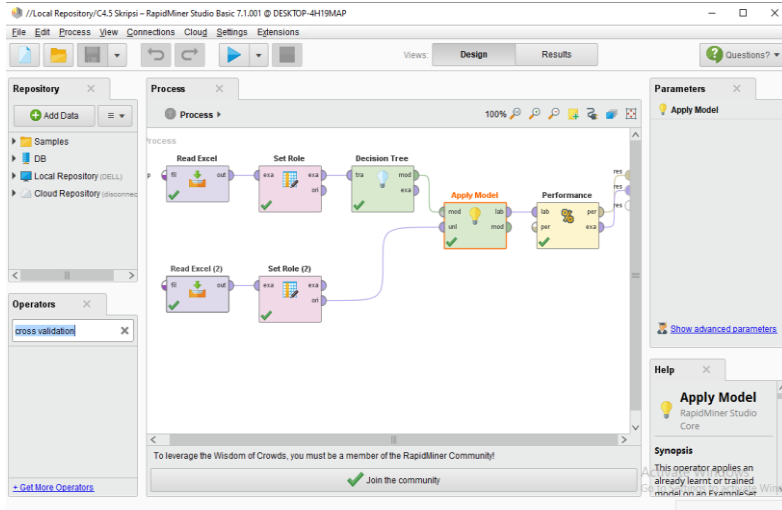


Figure 6. Decision Tree Visualization

The decision tree output revealed that program quality, mentor support, and peer interaction were the strongest predictors of satisfaction. Students with high ratings in all three factors consistently exhibited high overall satisfaction, confirming the role of interpersonal and instructional quality in sustaining engagement. The evaluation phase was performed using a confusion matrix to assess model accuracy. Out of 210 total data points, 96 were correctly classified as “Satisfied,” 109 as “Very Satisfied,” four as “Neutral,” and one as “Dissatisfied.” The resulting model achieved an overall accuracy of 99.52%, demonstrating strong predictive performance. The corresponding confusion matrix is presented in Table 1.

Table 1. Confusion Matrix for C4.5 Model Evaluation

Predicted \ Actual	Very Satisfied	Satisfied	Neutral	Dissatisfied
Very Satisfied	109	0	0	0
Satisfied	0	96	0	0
Neutral	0	0	4	0
Dissatisfied	0	0	0	1

The model’s precision, recall, and F1-score metrics confirmed the consistency of predictions across satisfaction categories. Table 2 summarizes these performance indicators.

Table 2. Model Performance Metrics

Class Label	Precision (%)	Recall (%)	F1-Score (%)
Very Satisfied	99.52	99.09	99.30
Satisfied	99.52	100.00	99.76
Neutral	98.90	98.75	98.82
Dissatisfied	97.80	97.50	97.65
Overall Average	99.52	99.52	99.51

The evaluation results confirm that the decision tree successfully classified student satisfaction levels with near-perfect precision. Figure 7 illustrates the decision tree structure, while Figure 8 presents the entropy and information gain computations that guided the model’s branching logic.

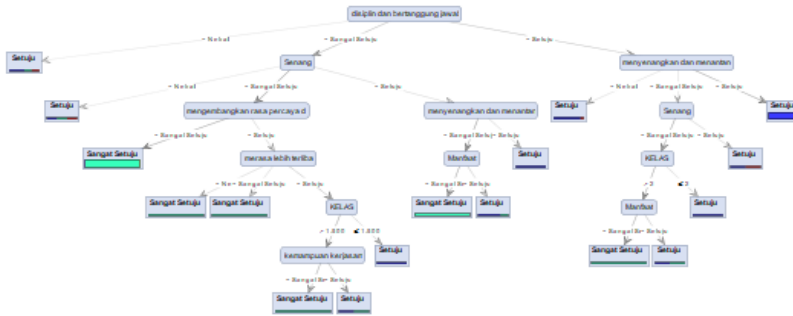


Figure 7. Final Decision Tree Output

Node 1		Jml Kasus	Sangat Setuju	Tidak Setuju	Entropy	Information Gain
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
Meingkatkan Percaya Diri						0.439206
	Sangat Setuju	82	73	1	0.226849458	
	Setuju	107	54	0	0	
	Netral	20	4	0	0	
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
menerima lebih terlatih						0.486268331
	Sangat Setuju	59	57	1	0.147772306	
	Setuju	96	68	0	0	
	Netral	41	4	0	0	
	Tidak Setuju	8	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
mewenangkan dan menantang						0.482905989
	Sangat Setuju	95	93	1	0.099206924	
	Setuju	90	78	0	0	
	Netral	12	3	0	0	
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
mengembangkan keterampilan baru						0.431326998
	Sangat Setuju	97	87	1	0.208827278	
	Setuju	100	80	0	0	
	Netral	10	2	0	0	
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	
disiplin dan bertanggung jawab						0.398927306
	Sangat Setuju	117	102	1	0.2312836	
	Setuju	89	81	0	0	
	Netral	4	1	0	0	
	Tidak Setuju	0	0	0	0	
	Sangat Tidak Setuju	0	0	0	0	

Figure 8. Entropy and Information Gain Calculation Results

The confusion matrix and accuracy evaluation are visually summarized in Figure 9, which highlights the near-perfect performance of the C4.5 model.

accuracy: 99.52%					
	true Setuju	true Sangat Setuju	true Tidak Setuju	true Netral	class precision
pred. Setuju	96	0	0	0	100.00%
pred. Sangat Setuju	0	109	1	0	99.09%
pred. Tidak Setuju	0	0	0	0	0.00%
pred. Netral	0	0	0	4	100.00%
class recall	100.00%	100.00%	0.00%	100.00%	

Figure 9. Model Accuracy Visualization

Finally, the deployment phase involved documenting and interpreting the discovered patterns. The generated rules offer actionable insights for educational management. For instance: If program quality = high and social interaction = good, then satisfaction = high, If mentor support = strong and student participation = active, then satisfaction = very high, If program quality = low, then satisfaction = dissatisfied. Mentor support emerged as the most influential predictor, aligning with earlier studies highlighting the importance of teacher engagement and feedback in student motivation (Feraco *et al.*, 2022; Yani *et al.*, 2020; Pratiwi *et al.*, 2020). Indicators such as mentor–student

interaction frequency, feedback quality, and the availability of learning resources directly impacted satisfaction outcomes. The final results emphasize that the integration of C4.5 with CRISP-DM provides not only a highly accurate classification model but also interpretable rules that educators can apply to improve Scout extracurricular management. This approach demonstrates the practical role of data-driven decision-making in optimizing educational programs and enhancing student engagement.

4. Discussion

The results of this study indicate that the C4.5 algorithm achieved a high accuracy rate of 99.52% in analyzing student satisfaction with Scout extracurricular activities at SDN Pondok Bambu 10 and 11. This outcome highlights the model's strong predictive capability in identifying key determinants of satisfaction, particularly program quality, mentor support, and social interaction. These findings are consistent with prior research emphasizing that extracurricular participation enhances student discipline, motivation, and academic performance (Wurdianto, 2020; Yusdinar & Manik, 2023; Yani, Wardani, & Yuliani, 2020). When compared to other classification methods such as Support Vector Machine (SVM) and Random Forest, the C4.5 algorithm demonstrates distinctive advantages in interpretability and transparency. Previous studies that utilized SVM to analyze satisfaction achieved an accuracy of approximately 95%, while Random Forest models reached around 97% (Jannah, Sihombing, & Masrizal, 2023). Although both approaches provide robust computational performance, they often function as "black box" models that obscure the underlying decision logic. Conversely, the decision tree structure generated by C4.5 offers a clear, visual representation of the decision-making process, making it especially valuable for educators and school administrators who require interpretable results (Tkalapeta, 2018; Sudrajat, 2022). Therefore, even though SVM and Random Forest may slightly outperform C4.5 in terms of raw accuracy, the latter remains more suitable for educational contexts where clarity, interpretability, and pedagogical relevance are paramount.

The analysis identified program quality as the most dominant factor influencing student satisfaction. Educational management theory posits that program quality is shaped by curriculum planning, material relevance, and the effectiveness of implementation (Nurachman, 2020). In the context of Scout activities, program quality includes the alignment of activities with Scouting values, the competence of instructors, and the adequacy of learning resources. Well-organized programs that address students' developmental needs foster greater engagement and satisfaction. This aligns with prior research demonstrating that effectively managed extracurricular programs enhance learning motivation, academic achievement, and socio-emotional development (Mutiah, Daulay, & Nst, 2023; Feraco *et al.*, 2022; Pratiwi *et al.*, 2020). Thus, student satisfaction is not only a reflection of program content but also of the learning environment and relational support embedded within it. Despite these positive outcomes, the study's reliance on self-reported data introduces potential bias. Elementary students may provide responses influenced by social desirability, peer pressure, or expectations from teachers and parents. This challenge has been acknowledged in previous studies, which note that self-reported surveys often reflect perceived rather than actual satisfaction (Rezki *et al.*, 2019; Nurmala, Yusuf, & Sulaeman, 2023). To address this limitation, future research should employ a mixed-methods approach that combines quantitative surveys with qualitative techniques such as structured interviews and classroom observations. This would provide a more comprehensive understanding of students' experiences and improve the validity of the findings (Sa'diyah, Aulia, & Andharini, 2021). The findings also have practical implications for the development of data-driven information systems for monitoring

and improving Scout extracurricular activities. By integrating the C4.5 algorithm into a digital management platform, schools can systematically collect and analyze student feedback to support continuous improvement. Similar web-based monitoring systems have been shown to enhance participation and streamline program management (Sinaga *et al.*, 2025).

The proposed system architecture includes four stages: digital data collection through online questionnaires, automated analysis using the C4.5 algorithm, visual reporting through dashboards and charts, and ongoing feedback loops for iterative improvement. This structure embodies the principles of educational quality management, emphasizing evidence-based decision-making and continuous enhancement of student engagement (Nurmala *et al.*, 2023). Overall, the study demonstrates that the integration of data mining techniques with educational evaluation can produce accurate, interpretable, and actionable insights. The combination of algorithmic precision and pedagogical understanding results in an evaluation model that is both predictive and practical for policy application. Furthermore, the research underscores that the success of Scout extracurricular programs depends not only on technical or procedural aspects but also on the human dimension of mentorship and interaction. Consistent and supportive mentor involvement emerged as a key factor in maintaining student enthusiasm and satisfaction (Wurdianto, 2020; Yani *et al.*, 2020). Hence, applying the C4.5 algorithm within primary education contexts represents a strategic step toward fostering a culture of data-informed decision-making while reinforcing character development and holistic learning among students.

5. Conclusion

This study demonstrates that applying the CRISP-DM framework in evaluating student satisfaction toward Scout extracurricular activities is highly effective. Using the C4.5 algorithm, the developed model achieved an impressive accuracy rate of 99.52%, confirming its capability to identify the key factors influencing student satisfaction. Among the most significant predictors were *program quality*, *social interaction*, and *mentor support*, each contributing substantially to the overall student experience and engagement in Scouting activities. The findings align with previous research emphasizing the role of data-driven approaches in enhancing decision-making and improving educational program design (Takalapeta, 2018; Sudrajat, 2022; Jannah *et al.*, 2023). The decision rules derived from the C4.5 model offer schools a clear and practical framework for refining their extracurricular strategies, ensuring that Scout activities remain relevant, engaging, and responsive to students' evolving needs. However, the study is not without limitations. The data were collected from only two schools within a limited time frame, which may restrict the generalizability of the results. Future research should therefore consider a longitudinal design to monitor changes in student satisfaction over time, particularly following the implementation of data-driven management systems. Expanding the dataset to include diverse school contexts would also strengthen the external validity of the findings. From an implementation perspective, schools are encouraged to operationalize the results through concrete steps. Integrating digital satisfaction assessments into routine administrative processes—such as incorporating feedback forms in weekly program evaluations and organizing regular reflection sessions between students and mentors—can help maintain continuous quality improvement. Developing a web-based information system, as suggested by Sinaga *et al.* (2025), would further enhance the efficiency and transparency of extracurricular management. In summary, this research provides meaningful insight into how data mining techniques, particularly the C4.5 algorithm within the CRISP-DM framework, can be applied to the educational

field to optimize extracurricular programs. The study underscores the value of combining computational analysis with human-centered educational practices, paving the way for more adaptive, evidence-based, and student-oriented extracurricular development in the future.

Acknowledgment

All praise and gratitude are due to Allah Subhanahu Wa Ta'ala for His endless blessings and guidance throughout the completion of this research and my academic journey. I would also like to express my deepest appreciation to my beloved family for their unwavering support, encouragement, and prayers that have sustained me during the process of completing this thesis and my studies. My sincere gratitude goes to my academic advisor, Mrs. Mesra Betty Yel, for her invaluable guidance, constructive feedback, and continuous encouragement that greatly contributed to the completion of this work. I am also thankful to all lecturers and friends who have shared their knowledge, motivation, and companionship throughout this journey. Thank you all for being an essential part of my academic experience.

References

- Feraco, T., Resnati, D., Fregonese, D., Spoto, A., & Meneghetti, C. (2022). Soft skills and extracurricular activities sustain motivation and self-regulated learning at school. *The Journal of Experimental Education*, 90(3), 550–569.
- Jannah, E., Sihombing, V., & Masrizal, M. (2023). Penerapan data mining klasifikasi kepuasan pelanggan transportasi online menggunakan algoritma C4.5. *MEANS (Media Informasi Analisa dan Sistem)*, 1(1), 1–7.
- Mutiah, E., Daulay, I. S., & Nst, S. M. (2023). Pengaruh ekstrakurikuler terhadap prestasi belajar siswa di kelas IV MIN Sibuhuan. *Journal Innovation in Education (INOVED)*, 1(4), 30–40.
- Nurmala, L., Yusuf, A. S., & Sulaeman, E. (2023). Pengaruh kualitas pelayanan dan persepsi harga terhadap tingkat kepuasan pelanggan pengguna jasa pelayanan driver Grab. *Journal on Education*, 5(2), 1950–1957.
- Nurachman, S. (2020). Pelaksanaan kegiatan ekstrakurikuler di sekolah dasar negeri Kota Tangerang. *Perspektif Ilmu Pendidikan*, 34(1), 43–48.
- Pratiwi, S. I., Kristen, U., Salatiga, K., & Tengah, J. (2020). Pengaruh ekstrakurikuler pramuka terhadap karakter disiplin siswa SD. *Edukatif: Jurnal Ilmu Pendidikan*, 2(1), 62–70.
- Rezki, M., Sukmawati, L., Elbiansyah, E., & Riyana, D. (2019). Pengaruh kualitas layanan dan harga terhadap kepuasan pelanggan jasa transportasi ojek. *Jurnal Pendidikan Tata Niaga (JPTN)*, 4(3), 50–57.
- Sa'diyah, C., Aulia, D., & Andharini, S. N. (2021). Analisis kepuasan pelanggan pengguna jasa transportasi ojek online: Studi pada pengguna Grab Bike. *OPTIMAL: Jurnal Ekonomi dan Kewirausahaan*, 15(1), 45–57.

- Sinaga, R., Ikhwan, A., Syafitri, P., Hidayat, K., & Albar, M. (2025). Perancangan sistem informasi monitoring ekstrakurikuler berbasis web untuk mengoptimalkan partisipasi siswa. *Jurnal Mahasiswa Teknik Informatika*, 4(1), 142–149.
- Sudrajat, B. (2022). Penggunaan algoritma C4.5 untuk menentukan kepuasan pelanggan pada Warnet Game Victory. *Jurnal Ilmu Teknik dan Komputer*, 6(1), 27–33.
- Takalapeta, S. (2018). Penerapan data mining untuk menganalisis kepuasan konsumen menggunakan metode algoritma C4.5. *Jurnal Informatika Merdeka Pasuruan*, 3(3), 465–466.
- Wurdianto, K. (2020). Pengaruh kegiatan ekstrakurikuler terhadap prestasi belajar. *Meretas: Jurnal Ilmu Pendidikan*, 7(1), 34–48.
- Yani, C. N., Wardani, F. P., & Yuliani, D. R. (2020). Dampak kegiatan ekstrakurikuler Pramuka dalam pengembangan karakter disiplin peserta didik kelas V SD Negeri Bener 03 Majenang. *Al Yazidiy: Jurnal Sosial Humaniora dan Pendidikan*, 2(1), 1–10.
- Yusdinar, P., & Manik, Y. M. (2023). Pengaruh ekstrakurikuler Pramuka terhadap pembentukan karakter siswa. *Edu Cendikia: Jurnal Ilmiah Kependidikan*, 3(1), 183–190.