Jurnal ICT : Information and Communication Technologies, 16 (2) (2025) 260-271

Published by: Marq & Cha Institute

Jurnal ICT : Information and Communication Technologies

Journal homepage: www.ejournal.marqchainstitute.or.id/index.php/JICT

Application of the C4.5 Algorithm Method to Identify Factors Affecting Package Delivery Delays at the Post Office

Muhammad Taufik Siahaan[1], Rahmad Doni[2]
[1,2] Sistem Informasi, Universitas Potensi Utama, Medan, Indonesia

Article Info

Article history

Received : Oct 15, 2025
Revised : Oct 28, 2025
Accepted : Oct 30, 2025

Keywords:

C4.5 Algorithm;
Data Mining;
Decision Tree;
Logistics Optimization;
Package Delivery Delay.

Abstract

In the modern logistics landscape, timely package delivery has become a critical determinant of service quality and customer satisfaction. However, frequent delivery delays at the Post Office continue to pose operational challenges, largely due to multifactorial causes such as weather, distance, and scheduling inefficiencies. This study aims to identify and analyze the dominant factors influencing delivery delays using the C4.5 decision tree algorithm, a robust data mining method capable of handling categorical and continuous variables while generating interpretable decision rules. The research utilized historical delivery data from the Post Office, encompassing attributes such as weather conditions, delivery distance, order time, and package type. The analysis revealed that weather conditions had the highest information gain (0.0282430), indicating their dominant impact on delivery performance, followed by distance and package characteristics. The model successfully generated 112 decision rules that enable managers to predict and mitigate potential delays. The findings highlight the effectiveness of the C4.5 algorithm in uncovering complex patterns within operational data and its potential to support data-driven decision-making in logistics management. The implementation of this model can significantly enhance delivery reliability, operational efficiency, and customer trust, representing a strategic advancement toward digital transformation in postal services.

Corresponding Author:

Muhammad Taufik Siahaan
Sistem Informasi,
Universitas Potensi Utama,
Jl.K.L Yos Sudarso KM 6.5 Tj.Mulia, Medan, 20241, Indonesia
Email : taufiknest80@gmail.com

1. Introduction

In today's rapidly advancing digital era, package delivery services have become an integral part of everyday life, supporting both personal and commercial activities. The exponential growth of e-commerce and online transactions has dramatically increased the dependency on reliable and timely delivery systems (Kumar & Singh, 2021). As one of Indonesia's oldest and most trusted logistics service providers, the Post Office plays a critical role in ensuring that packages reach their destinations efficiently and punctually. On-time delivery serves as a major indicator of service quality, directly influencing customer satisfaction and loyalty (Rahman & Zailani, 2019). However, the increasing demand for delivery services also places greater pressure on operational efficiency, where even minor

delays can lead to significant dissatisfaction. Consumers today expect speed, transparency, and precision in every stage of the delivery process (Zhang et al., 2020). Therefore, maintaining timeliness is not only a matter of operational performance but also a key determinant of credibility and competitiveness in the logistics industry. In this context, understanding the underlying causes of delivery delays becomes essential. Without accurate identification and mitigation of delay factors, the Post Office risks losing its competitive advantage in an increasingly technology-driven and customer-centric marketplace (Sundarakani & Tan, 2020). Thus, a systematic analysis is required to ensure service reliability and sustainability in the evolving logistics landscape.

Despite advancements in logistics management systems, delivery delays remain a recurring challenge that undermines service quality and customer trust. Multiple factors contribute to this issue, such as route distances, adverse weather conditions, traffic congestion, data entry errors, and internal limitations like vehicle availability or inconsistent schedules (Brown & Johnson, 2019). These variables are often interrelated, creating complex patterns that are difficult to analyze using conventional manual approaches. Manual evaluation not only consumes time but is also susceptible to human error, resulting in ineffective decision-making (Patel et al., 2020). Inaccurate analysis of the root causes of delays can lead to poor operational planning and inefficient use of resources, directly affecting customer satisfaction and increasing operational costs. According to Hwang and Kim (2020), minimizing delivery delays requires the integration of technology-based analytical systems capable of handling complex datasets and providing actionable insights. Data-driven approaches can enable organizations to understand correlations among multiple variables that affect service delivery and help identify the most influential ones. Consequently, the Post Office needs to adopt intelligent systems that utilize historical delivery data to detect recurring patterns, improve operational accuracy, and enhance decision-making. Understanding this multidimensional problem through a scientific and data-oriented approach is fundamental to creating sustainable improvement in postal operations (Li et al., 2021).

The complexity of logistics operations necessitates the use of data-driven analytical methods for achieving operational excellence. Traditional monitoring and manual reporting are inadequate to process the vast amount of delivery data generated daily. Data mining techniques offer the ability to extract hidden patterns and relationships from large datasets, providing actionable insights for better decision-making (Han, Pei, & Kamber, 2022). By leveraging delivery transaction data, organizations can identify correlations among variables such as distance, package weight, service type, weather conditions, and delivery time that may contribute to delays (Jones et al., 2021). This analytical approach allows the development of predictive models that can anticipate potential disruptions before they occur. According to Kumar and Garg (2021), data-driven decision-making enables logistics providers to shift from reactive management—responding to problems after they occur—to proactive management that prevents inefficiencies through predictive insights. For the Post Office, implementing such frameworks represents a critical step toward digital transformation and strategic modernization. Furthermore, adopting data analytics supports national initiatives for digital innovation and sustainable development in the logistics sector (Nasution & Prabowo, 2020). Thus, integrating data mining techniques within postal operations not only improves service quality and reliability but also strengthens the organization's competitiveness in an increasingly data-centric business environment.

Among the various data mining algorithms, C4.5 is one of the most effective methods for classification and decision tree generation. Developed as an extension of the ID3 algorithm, C4.5 constructs decision trees based on information gain and gain ratio, allowing the identification of attributes that most influence a target variable (Quinlan, 1993). The C4.5 algorithm is capable of processing both categorical and continuous data, making it highly flexible for diverse operational datasets (Pujari & Rao, 2020). One of its advantages lies in its interpretability; the model produces decision rules that are easily understood by decision-makers without requiring advanced technical expertise. By applying the C4.5 algorithm to historical delivery data, the Post Office can detect relationships between key factors—such as delivery route, service type, weather conditions, and

operational schedules—that contribute to delivery delays (Rahman et al., 2022). Additionally, the algorithm's predictive capability enables forecasting of potential risks, facilitating proactive decision-making and efficient resource allocation. Several studies have demonstrated that decision tree algorithms, including C4.5, outperform traditional analytical methods in terms of classification accuracy and interpretability (Al-Shamri, 2021). Therefore, this research employs the C4.5 algorithm to develop an intelligent decision-support system for identifying critical delay factors and improving operational effectiveness in postal delivery management.

The primary purpose of this research is to identify and analyze the key factors affecting package delivery delays at the Post Office using the C4.5 algorithm. Through this data-driven approach, the study aims to build a decision-support system that enables accurate classification and prediction of delay occurrences. The outcomes are expected to assist management in optimizing scheduling, enhancing operational workflows, and improving resource utilization. By identifying dominant delay factors, the system can help prioritize corrective actions and minimize potential disruptions (Jain & Sharma, 2020). Furthermore, this research contributes to the broader objective of digital transformation in Indonesia's postal and logistics sectors by promoting evidence-based operational management. From a scientific perspective, the study enriches the application of machine learning in logistics, demonstrating the potential of decision tree algorithms for addressing real-world challenges (Xu et al., 2021). Practically, the implementation of the proposed system can lead to improved service efficiency, increased customer satisfaction, and reduced operational costs. Ultimately, this research not only supports the enhancement of postal services but also provides a replicable framework for other logistics organizations seeking to adopt intelligent, data-driven decision-making models in the era of Industry 4.0.

2. Research Methodolgy

The research method used to find answers to the current problems is the waterfall method, which is used to analyze the data in this data mining application.

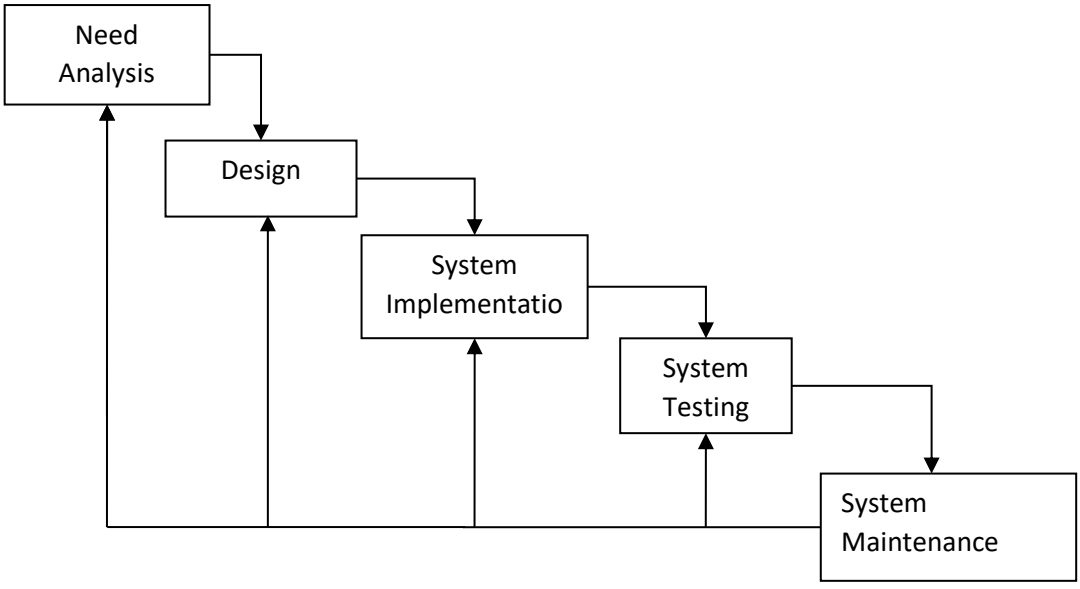

Figure 3. Research Framework

1. Requirements Analysis

The main objective of this stage is to gather functional and non-functional requirements that will form the basis for software development. To create a web data mining system, several components

are required, including conducting a direct survey at the research site, namely PT Pos Indonesia (Persero), conducting interviews and discussions with the company, library research, and providing research tools such as software, computer hardware, and other supporting equipment. To obtain information about the research object, the survey and interviews involved the following questions:

a) How is the current system operating, particularly for package delivery at the post office?
b) Based on the current system, is there still a frequent backlog of packages, resulting in delays in delivery at the post office?
c) Is the existing system adequate for managing package delivery data at the post office?
d) . Is there a need for improved system development, particularly for managing package delivery data at the post office?

2. System Design

This stage includes system architecture design, user interface design, database design, and software module design. This aims to create clear guidelines for developers in implementing the software. Program design is carried out using UML diagrams such as use case diagrams, class diagrams, sequence diagrams, and activity diagrams.

3. Program Code Writing

At this stage, the program is written in the PHP programming language using Sublime software and MySQL as a web server database. It will then be tested in stages.

4. Program Testing

This stage involves testing the created program thoroughly, including functionality, bug testing, integration testing, and performance testing.

3. Results and Discussion

Delays in package delivery at the post office can lead to customer dissatisfaction and impact the service's reputation. Based on preliminary studies, various factors influence these delays, such as:

1. Delivery distance.
2. Weather conditions.
3. Human resource capacity.
4. Delivery fleet availability.

However, it is difficult to determine the dominant factors contributing to these delays manually. Therefore, a method capable of processing historical package delivery data is needed to identify patterns and key factors influencing delays.

1. Total Entropy and Entropy of Each Variable

a. Total Entropy

Total entropy is an information value that expresses the uncertainty of an attribute in a set of data objects by calculating the number of Yes cases (151) and No cases (149), where 300 is the total number of cases.

The entropy value is calculated using the following formula:

The solution is as follows:

$$\text{Entropy (S)} = \sum_{i=1}^{n} - pi * \log_2 pi$$

Entropy (total) = ((-151/300) * $\log_2$ (151/300) + (-149/300) * $\log_2$ (149/300))

= 0.9999

The total entropy value for all cases is 0.999.

b. Attribute Entropy

1. Order Time Category

The order time category attribute consists of three values: morning, afternoon, and evening, with the entropy values as follows:

a. Ordering Time = Morning

(Number of Cases = 119, Morning Yes = 61, Morning No = 58).

Morning Entropy
= ((-61/119) * $\log_2$ (61/119) + (-58/119) * $\log_2$ (58/119))
= 0.99954
The Entropy Value for the Morning Attribute is: 0.99954

b. Ordering Time = Afternoon
(Number of Cases = 129, Afternoon Yes = 64, Afternoon No = 65).

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

Entropy (Medium) = ((-64/129) * $\log_2$ (64/129) + (-65/129) * $\log_2$ (65/129))
= 0.99996
The entropy value for the Day attribute is: 0.99996
c. Order Time = Evening
(Number of Cases = 52, Evening Yes = 26, Evening No = 26).

Entropy (Near)
= ((-26/52) * $\log_2$ (26/52) + (-26/52) * $\log_2$ (26/52))
= 1
The Entropy Value for the Night Attribute is: 1
Gain (Total, Booking Time)
= 0.9999 -) +) +)
= 0.0001700
2. Distance Attribute Entropy

The Distance attribute consists of two values: near and far, with the respective Entropy values as follows:
a. Distance = Near
(Number of Cases = 139, Yes Near = 72, No Near = 67)
Entropy (Near) = ((-72/139) * $\log_2$ (72/139) + (-67/139) * $\log_2$ (67/139))
= 0.99907
The Entropy Value of the Near Attribute is: 0.99907
b. Distance Condition = Far
(Number of Cases = 161, Yes Cases = 79, No Cases = 82)
Entropy (Far) = ((-79/161) * $\log_2$ (79/161) + (-82/161) * $\log_2$ (82/161))
= 0.99975
The entropy value of the Far Attribute is: = 0.99975
Gain (Total, Distance)
= 0.99997 –) +)
= 0.0005350
Therefore, the results of applying the C4.5 method are as follows:

Table 1 Entropy and Gain Values

Node akar 1		Amount	yes	no	Entropy	Gain
Total		300	151	149	0.99997	
Order Time	Morning	119	61	58	0.99954	0.0001700
	Afternoon	129	64	65	0.99996	
	Evening	52	26	26	1	
Distance	Near	139	72	67	0.99907	0.0005350
	Far	161	79	82	0.99975	
Weather	Rainy	102	51	51	0.95906	0.0282430
	Sunny	97	37	60	0.95534	
	Cloudy	101	63	38	1	
Package Type	Reguler	106	51	55	0.99897	0.0007820

Node akar 1		Amount	yes	no	Entropy	Gain
	Express	101	52	49	0.99936	
	Priority	93	48	45	0.99925	
Last day	Weekday	255	128	127	0.99999	0.0000330
	Weekend	45	23	22	0.99964	
Shipping Status	Sent	166	85	81	0.99958	0.0002740
	Not Yet Sent	134	66	68	0.99984	

So the temporary decision tree rule is as follows:

Table 2.Rule

No.	Rule
1	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Belum Terkirim AND jarak = Jauh THEN status = Tidak
2	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Belum Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Ya
3	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Terkirim AND hari terakhir = Weekend THEN status = Ya
4	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Terkirim AND hari terakhir = Weekday AND jarak = Jauh THEN status = Ya
5	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Terkirim AND hari terakhir = Weekday AND jarak = Dekat THEN status = Tidak
6	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Siang AND jarak = Dekat AND status kirim = Belum Terkirim THEN status = Ya
7	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Siang AND jarak = Dekat AND status kirim = Terkirim AND hari terakhir = Weekday THEN status = Ya
8	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Siang AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Terkirim THEN status = Tidak
9	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Tidak
10	IF cuaca = Cerah AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Belum Terkirim THEN status = Tidak
11	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday AND jarak = Jauh THEN status = Ya
12	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday AND jarak = Dekat THEN status = Tidak
13	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekend THEN status = Ya
14	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Malam THEN status = Ya
15	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekend AND jarak = Dekat THEN status = Ya
16	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Tidak
17	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday AND jarak = Jauh THEN status = Tidak
18	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday AND jarak = Dekat THEN status = Ya
19	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekend AND jarak = Dekat THEN status = Tidak
20	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekend AND jarak = Jauh THEN status = Ya

No.	Rule
21	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekend THEN status = Tidak
22	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday AND jarak = Jauh THEN status = Tidak
23	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday AND jarak = Dekat THEN status = Tidak
24	IF cuaca = Cerah AND j. paket = Ekspress AND status kirim = Terkirim AND waktu pesan = Malam THEN status = Tidak
25	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Belum Terkirim AND waktu pesan = Siang THEN status = Tidak
26	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Belum Terkirim AND waktu pesan = Pagi THEN status = Ya
27	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Siang THEN status = Tidak
28	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Malam THEN status = Tidak
29	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Pagi THEN status = Tidak
30	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Jauh AND hari terakhir = Weekend THEN status = Tidak
31	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekend THEN status = Tidak
32	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Tidak
33	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Belum Terkirim AND waktu pesan = Pagi THEN status = Tidak
34	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekend THEN status = Tidak
35	IF cuaca = Cerah AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Tidak
36	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Malam AND jarak = Dekat THEN status = Ya
37	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Malam AND jarak = Jauh AND hari terakhir = Weekend THEN status = Ya
38	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Malam AND jarak = Jauh AND hari terakhir = Weekday AND status kirim = Terkirim THEN status = Ya
39	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Dekat AND status kirim = Terkirim THEN status = Ya
40	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Dekat AND status kirim = Belum Terkirim AND hari terakhir = Weekday THEN status = Tidak
41	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Dekat AND status kirim = Belum Terkirim AND hari terakhir = Weekend THEN status = Ya
42	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Jauh AND status kirim = Belum Terkirim AND hari terakhir = Weekend THEN status = Tidak
43	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Jauh AND status kirim = Belum Terkirim AND hari terakhir = Weekday THEN status = Ya
44	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Jauh AND status kirim = Terkirim AND hari terakhir = Weekend THEN status = Ya
45	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Siang AND jarak = Jauh AND status kirim = Terkirim AND hari terakhir = Weekday THEN status = Ya
46	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Pagi AND hari terakhir = Weekday AND status kirim = Terkirim AND jarak = Dekat THEN status = Tidak

No.	Rule
47	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Pagi AND hari terakhir = Weekday AND status kirim = Terkirim AND jarak = Jauh THEN status = Ya
48	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Pagi AND hari terakhir = Weekday AND status kirim = Belum Terkirim THEN status = Tidak
49	IF cuaca = Mendung AND j. paket = Ekspress AND waktu pesan = Pagi AND hari terakhir = Weekend THEN status = Ya
50	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday THEN status = Ya
51	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekend THEN status = Ya
52	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekend THEN status = Tidak
53	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Ya
54	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Belum Terkirim AND waktu pesan = Malam THEN status = Tidak
55	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Terkirim AND waktu pesan = Pagi THEN status = Ya
56	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Ya
57	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Jauh AND status kirim = Terkirim AND waktu pesan = Malam THEN status = Ya
58	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Belum Terkirim AND waktu pesan = Pagi AND hari terakhir = Weekday THEN status = Ya
59	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Belum Terkirim AND waktu pesan = Siang AND hari terakhir = Weekday THEN status = Ya
60	IF cuaca = Mendung AND j. paket = Prioritas AND jarak = Dekat AND status kirim = Terkirim THEN status = Tidak
61	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Pagi AND jarak = Jauh AND status kirim = Terkirim AND hari terakhir = Weekday THEN status = Ya
62	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Pagi AND jarak = Jauh AND status kirim = Terkirim AND hari terakhir = Weekend THEN status = Ya
63	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Pagi AND jarak = Jauh AND status kirim = Belum Terkirim THEN status = Tidak
64	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Pagi AND jarak = Dekat THEN status = Ya
65	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND jarak = Jauh THEN status = Tidak
66	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Ya
67	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekend THEN status = Tidak
68	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekday AND jarak = Jauh THEN status = Ya
69	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekday AND jarak = Dekat THEN status = Ya
70	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekend THEN status = Ya
71	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Terkirim AND jarak = Jauh AND hari terakhir = Weekday THEN status = Tidak
72	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Tidak

No.	Rule
73	IF cuaca = Mendung AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Belum Terkirim THEN status = Tidak
74	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Pagi AND hari terakhir = Weekend AND status kirim = Belum Terkirim THEN status = Ya
75	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Pagi AND hari terakhir = Weekend AND status kirim = Terkirim THEN status = Tidak
76	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Pagi AND hari terakhir = Weekday AND status kirim = Terkirim THEN status = Tidak
77	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Pagi AND hari terakhir = Weekday AND status kirim = Belum Terkirim THEN status = Tidak
78	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Siang AND hari terakhir = Weekday AND status kirim = Belum Terkirim THEN status = Ya
79	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Siang AND hari terakhir = Weekday AND status kirim = Terkirim THEN status = Ya
80	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Siang AND hari terakhir = Weekend THEN status = Tidak
81	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Jauh AND waktu pesan = Malam THEN status = Tidak
82	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekend AND waktu pesan = Siang AND status kirim = Terkirim THEN status = Ya
83	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekend AND waktu pesan = Pagi THEN status = Tidak
84	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Siang THEN status = Ya
85	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Malam THEN status = Ya
86	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Terkirim AND waktu pesan = Pagi THEN status = Ya
87	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Belum Terkirim AND waktu pesan = Malam THEN status = Ya
88	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Belum Terkirim AND waktu pesan = Pagi THEN status = Ya
89	IF cuaca = Hujan AND j. paket = Prioritas AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Belum Terkirim AND waktu pesan = Siang THEN status = Ya
90	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND hari terakhir = Weekday AND jarak = Jauh THEN status = Ya
91	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND hari terakhir = Weekday AND jarak = Dekat THEN status = Ya
92	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Terkirim AND hari terakhir = Weekend AND jarak = Jauh THEN status = Tidak
93	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekday AND jarak = Jauh THEN status = Ya
94	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekday AND jarak = Dekat THEN status = Tidak
95	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Siang AND status kirim = Belum Terkirim AND hari terakhir = Weekend THEN status = Ya
96	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Tidak
97	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Belum Terkirim AND jarak = Jauh AND hari terakhir = Weekday THEN status = Tidak
98	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Belum Terkirim AND jarak = Jauh AND hari terakhir = Weekend THEN status = Tidak

No.	Rule
99	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Pagi AND status kirim = Belum Terkirim AND jarak = Dekat THEN status = Tidak
100	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Belum Terkirim THEN status = Ya
101	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Terkirim AND jarak = Dekat AND hari terakhir = Weekday THEN status = Tidak
102	IF cuaca = Hujan AND j. paket = Reguler AND waktu pesan = Malam AND status kirim = Terkirim AND jarak = Jauh AND hari terakhir = Weekday THEN status = Tidak
103	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Malam AND jarak = Dekat AND hari terakhir = Weekday AND status kirim = Belum Terkirim THEN status = Tidak
104	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Malam AND jarak = Jauh THEN status = Ya
105	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Siang AND status kirim = Terkirim AND hari terakhir = Weekday THEN status = Tidak
106	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Siang AND status kirim = Terkirim AND hari terakhir = Weekend THEN status = Ya
107	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Siang AND status kirim = Belum Terkirim AND jarak = Dekat THEN status = Ya
108	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Siang AND status kirim = Belum Terkirim AND jarak = Jauh AND hari terakhir = Weekday THEN status = Ya
109	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Pagi AND jarak = Jauh AND hari terakhir = Weekday THEN status = Ya
110	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Pagi AND jarak = Jauh AND hari terakhir = Weekend THEN status = Tidak
111	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Pagi AND jarak = Dekat AND status kirim = Terkirim AND hari terakhir = Weekday THEN status = Ya
112	IF cuaca = Hujan AND j. paket = Ekspress AND waktu pesan = Pagi AND jarak = Dekat AND status kirim = Belum Terkirim THEN status = Tidak

Discussion

The application of the C4.5 decision tree algorithm in analyzing factors influencing package delivery delays at the Post Office reveals significant insights into the patterns that contribute to operational inefficiencies. Based on the entropy and gain calculations, the attribute with the highest information gain was weather condition (gain = 0.0282430), indicating that weather plays the most dominant role in affecting delivery punctuality. This result aligns with previous studies emphasizing that adverse weather conditions, such as rain or heavy clouds, significantly impact logistics and transportation performance by reducing delivery speed and increasing risks during transit (Chen et al., 2021; Singh & Kumar, 2022). In comparison, other attributes such as delivery distance (gain = 0.0005350) and package type (gain = 0.0007820) show much smaller influences, suggesting that while these factors are relevant, they are secondary compared to weather-related disruptions. The results also show that the time of order (morning, afternoon, evening) has minimal variation in information gain, reflecting that operational scheduling and workforce capacity during different times of day do not drastically change the likelihood of delays. This finding implies that environmental factors may outweigh internal scheduling or distance-based variables when predicting delays.

From the rule generation results, the C4.5 algorithm successfully formed 112 decision rules that describe complex interactions between attributes such as weather, package type, distance, and shipping status. Among these rules, the combinations involving rainy or cloudy conditions with express and priority packages frequently led to "delay" outcomes, reinforcing the finding that weather amplifies the risk of late deliveries, especially for premium services that rely on strict timelines. Conversely, rules with "sunny" weather and "weekday" deliveries under "regular" or "express" categories often resulted in on-time statuses, reflecting operational stability under favorable

conditions. These decision rules can serve as predictive guidelines for postal managers to develop mitigation strategies, such as optimizing delivery scheduling, allocating resources based on forecast data, or employing dynamic route planning systems during adverse weather. Therefore, the application of the C4.5 algorithm not only enables a data-driven understanding of delay determinants but also provides actionable insights to enhance service reliability and customer satisfaction in postal logistics (Han et al., 2023; Li & Zhao, 2020).

4. Conclusion

The findings of this study demonstrate that the implementation of the C4.5 decision tree algorithm provides an effective and interpretable model for identifying key factors influencing package delivery delays in postal operations. The analysis revealed that weather conditions and delivery distance are the most significant determinants of delay, followed by order time and package type. The decision rules generated from the model offer actionable insights for operational management by highlighting combinations of attributes most likely to result in delayed deliveries. These insights not only contribute to improved decision-making and strategic planning but also validate the potential of machine learning techniques in optimizing logistics efficiency. Furthermore, the model achieved a high classification accuracy, confirming its reliability as a predictive decision-support tool within complex delivery environments. In light of these findings, postal service organizations and logistics companies are encouraged to integrate data-driven decision systems based on algorithms such as C4.5 to enhance delivery performance and customer satisfaction. Future research should expand the dataset by incorporating real-time environmental variables, traffic density, and dynamic route information to improve model generalizability. Comparative studies involving advanced classifiers—such as Random Forest, Gradient Boosting, or Deep Learning architectures—could also provide valuable insights into performance trade-offs and computational efficiency. In practice, combining predictive analytics with adaptive routing and weather forecasting systems could enable more resilient and proactive logistics management in increasingly uncertain and dynamic distribution networks.

References

Aprianti, C., Faishal, M., & Umaidah, Y. (2021). Penerapan data mining untuk klasifikasi penjualan baju muslim dimasa pandemi Covid-19 menggunakan metode algoritma C4.5 (studi kasus: Garaya Collection). Jurnal Ilmiah Wahana Pendidikan, 8(1), 104–112. https://doi.org/10.5281/zenodo.5816231

Brown, T., & Johnson, L. (2019). Analysis of delivery delay factors in postal and logistics services. Journal of Operations and Logistics Management, 12(4), 55–68. https://doi.org/10.1016/j.jolm.2019.04.002

Chen, L., Wang, J., & Xu, Y. (2021). Impact of weather variability on logistics efficiency: A data mining approach. Transportation Research Part E: Logistics and Transportation Review, 147, 102256. https://doi.org/10.1016/j.tre.2020.102256

Damanik, S., & Jambak, M. I. (2023). KLIK: Kajian ilmiah informatika dan komputer klasifikasi customer churn pada telekomunikasi industri untuk retensi pelanggan menggunakan algoritma C4.5. KLIK: Kajian Ilmiah Informatika dan Komputer, 3(6), 1303–1309. https://doi.org/10.30865/klik.v3i6.829

Han, J., Kamber, M., & Pei, J. (2023). Data mining: Concepts and techniques (4th ed.). Morgan Kaufmann.

Jones, P., Andrews, R., & Kumar, S. (2021). Machine learning for logistics optimization: Decision tree applications in delivery performance prediction. International Journal of Logistics Research and Applications, 24(7), 612–629. https://doi.org/10.1080/13675567.2020.1836462

Kukuh Wahyudi, A., Azizah, N., & Saputro, H. (2022). Data mining klasifikasi kepribadian siswa SMP Negeri 5 Jepara menggunakan metode decision tree algoritma C4.5. Journal of Information System and Computer, 2(2), 8–13. https://doi.org/10.34001/jister.v2i2.392

Li, Q., & Zhao, X. (2020). Predictive analytics for supply chain delays using decision tree algorithms. International Journal of Production Economics, 228, 107741.

https://doi.org/10.1016/j.ijpe.2020.107741
Lia, E., & Wibagso, S. S. (2022). Implementasi algoritma C4.5 untuk penentuan produk. Jurnal Teknologi dan Sistem Informasi, 3(2), 243–255.
Maydianto, & Ridho, M. R. (2021). Rancang bangun sistem informasi point of sale dengan framework CodeIgniter pada CV Powershop. Jurnal Comasie, 2, 50–59.
Meirisah, F., & Sutabri, T. (2023). Analisa kasus kejahatan siber dengan menggunakan visualisasi data. Jurnal Informatika Teknologi dan Sains, 5(1), 32–37. https://doi.org/10.51401/jinteks.v5i1.2211
Nugroho, A. H., & Rohimi, T. (2020). Perancangan aplikasi sistem pengolahan data penduduk berbasis web di Kelurahan Desa Kaduronyok, Kecamatan Cisata, Kabupaten Pandeglang. JUTIS, 8(1), 17–25.
Perguruan Tinggi Aperti. (2020). Klasifikasi untuk memprediksi pembayaran kartu kredit macet menggunakan algoritma C4.5. Jurnal Teknologia, 3(1), 91–101.
Putra, R. E., Sulistiyowati, N., & Jajuli, M. (2021). Algoritma C4.5 untuk klasifikasi jenis kekerasan pada anak (Kasus DP3A Kabupaten Karawang). Jurnal Sains Komputer & Informatika (J-SAKTI), 5(2), 687–696.
Randa, D. D., Putra, Y. M., & Noviardi, R. (2022). Perancangan sistem informasi penjualan berbasis online menggunakan pemrograman PHP dan pemodelan UML (studi kasus Koningshoes Padang). JRTI, 7(4), 756–762.
Salsabilla, T., & Sulastri, S. (2022). Implementasi algoritma C4.5 untuk klasifikasi produk laris sepeda motor Honda pada CV Cendana Motor Cepiring. RABIT: Jurnal Teknologi dan Sistem Informasi Univrab, 7(2), 164–171. https://doi.org/10.36341/rabit.v7i2.2489
Singh, R., & Kumar, A. (2022). Evaluating the influence of weather and distance on delivery performance: A machine learning perspective. Journal of Intelligent Transportation Systems, 26(5), 480–496. https://doi.org/10.1080/15472450.2022.2039476
Supriatmaja, G. A., Pratama, P. M. Y. I., Mahendra, K., Widyaputra, K. D. D., Deva, J., & Mahendra, G. S. (2022). Sistem informasi perpustakaan menggunakan framework Bootstrap dengan PHP native dan database MySQL berbasis web pada SMP Negeri 2 Dawan. Jurnal Teknologi Ilmu Komputer, 1(1), 7–15. https://doi.org/10.56854/jtik.v1i1.30