Vol. 3 No. 02 (2025)
ISSN: 3046-711X

Implementation of Regression CART Decision Tree for Best Cycling Time Recommendation Based on Weather Data

[1]Nurul Badriah*, [2]Fajar Muttaqi, [3]Sony Veri Shandy, [4]Moh. Alfaujianto

Utpadaka Swastika University, Tangerang, Indonesia 15112
E-mail: [1]nurul.badriah@utpas.ac.id, [2]fajar.muttaqi@utpas.ac.id, [3]sony.veri.shandy@utpas.ac.id, [4]moh.alfaujianto@utpas.ac.id,
*corresponding Author: nurul.badriah@utpas.ac.id

Accepted: September 7, 2025 | Published: October 30, 2025

ABSTRACT

Cycling requires careful time planning to ensure safety and comfort, especially when considering weather conditions such as temperature, wind speed, and overall weather status. However, cyclists often struggle to determine the optimal time to ride due to the lack of accurate and easily accessible recommendations. This study aims to design and implement a mobile application that recommends the best cycling time based on real-time weather data. The system applies the Regression CART Decision Tree method, trained using hourly temperature, wind speed, and weather condition parameters. Unlike classification approaches, Regression CART Decision Tree produces a continuous percentage score indicating the suitability level of each hour for cycling. Real-time data is obtained via the OpenWeatherMap API to maintain accuracy. The developed prototype displays hourly weather data along with the recommendation percentage, helping users plan their rides more effectively. Model evaluation shows that the Regression CART Decision Tree achieved high accuracy with a low Mean Absolute Error (MAE) and strong correlation between predicted and actual suitability scores. The results confirm that the model performs consistently in various weather scenarios. Overall, the system successfully delivers reliable, data-driven recommendations, assisting cyclists in selecting the safest and most comfortable cycling times.

KEYWORDS: Data Mining, Decision Tree, Cycling Recommendation, Mobile Application Prototype, Regression CART, Weather Data

1. Introduction

General Background

Weather plays an important role in determining comfort and safety when cycling. Extremely hot or cold temperatures, high wind speeds, and poor weather conditions can disrupt the cycling experience and even endanger cyclists' safety. However, many cyclists struggle to determine the right time to cycle due to hourly changing weather conditions. Therefore, an application is needed that can provide information on the best time to cycle by considering temperature, wind speed, and weather conditions in real time. Such an application can help cyclists plan their activities better by taking into account weather factors that may affect comfort and safety. With this application, cyclists are expected to choose the safest and most comfortable cycling time based on the latest weather data.

Previous Literature

Previous studies have developed weather applications and route systems, but the integration of hourly weather data for cycling is still lacking. Sopyandi et al. (2024) created a web-based weather application using the OpenWeatherMap API for real-time information, supporting transportation and agriculture, but it was not specific to cyclists. Luthfiansyah et al. (2022) designed a cycling route system using the SAW method, utilizing Google Maps and AccuWeather APIs, but did not analyze hourly weather for the best cycling time. Christyana (2022) developed a bicycle rental application without weather integration. No existing applications have focused on hourly weather-based cycling time

UNIVERSITAS UTPADAKA SWASTIKA
Jl. KS Tubun No.11 Tangerang 15112 Banten. Telp. (021) 5589161-62 Fax. (021) 5589163
JISI Jurnal Ilmiah Sistem Informasi
Vol. 3 No. 02 (2025)
ISSN: 3046-711X

recommendations, which shows a research gap that needs to be addressed.

Problem Identification

Based on the background above, the problem identification in can be described as follows:

a. Lack of information regarding the best time to cycle based on weather conditions.
b. Absence of an application prototype that provides hourly weather data and analyzes recommendation percentages for cycling activities.

Research Objectives

The objectives of this study are to:

a. Build an application prototype that integrates real-time weather data, including hourly temperature, wind speed, and weather conditions.
b. Implement the Regression CART Decision Tree method to generate continuous percentage scores for optimal cycling time recommendations based on weather data.

Gap Analysis

The following research gaps need to be addressed:

a. Limitation of Existing Weather-Based Cycling Applications
Previous studies developed general weather or route-based applications but did not implement advanced machine learning methods such as Regression CART Decision Tree. Most approaches relied on simple scoring or rule-based systems, which limited their ability to generate accurate and adaptive recommendations.
b. Lack of Continuous Recommendation Output, Existing studies primarily provided categorical results (e.g., "suitable" or "not suitable") without producing continuous percentage scores that reflect varying levels of cycling feasibility. Regression CART Decision Tree offers an improvement by predicting optimal cycling time in the form of probabilistic values based on temperature, wind speed, and weather conditions.

2. Literature Review

Several studies have addressed the development of mobile applications and decision support systems for cycling and weather-based activities. [1] Christyana & Angel (2022), in their work "Implementasi Aplikasi Android untuk Sistem Penyewaan Sepeda", developed a mobile-based bicycle rental system. However, their research did not integrate real-time weather data, limiting its applicability for cycling activity planning. Similarly, [2] Rizal et al. (2023) proposed "Perancangan Aplikasi BaikinSepeda Berbasis Mobile Dengan Menggunakan Flutter" which emphasized mobile development for bicycle services, yet it lacked weather-based recommendation features.

Research on decision tree algorithms has been applied in various domains. [3] Ahmad Taufiq ramadhan et al. (2023) explored "Penerapan Algoritma Decision Tree dalam Melakukan Analisis Klasifikasi Harga Handphone" to predict smartphone prices, while [4], Fahrizal et al. (2022) applied C4.5 in "Implementasi Metode Decision Tree C4.5 Pada Lampu Otomatis Berdasarkan Kebiasaan Penghuni Rumah." Likewise, [5] Naufal Baharudin et al. (2022implemented decision tree classification in "Klasifikasi Surat pada Aplikasi Mobile E-Surat Dinas Komunikasi dan Informatika Kota Kediri Berbasis Android." Although these studies demonstrated the flexibility of decision trees, they primarily focused on categorical classification rather than regression-based prediction.

The integration of weather data into applications has also gained attention. [6], Sopyandi et al. (2024) in "Bangun Aplikasi Tracking Cuaca (Weather App) Menggunakan Public API Berbasis Website", designed a weather monitoring system but did not focus on cycling recommendations. [7] Gunawan Zain et al. (2002) proposed "Pengembangan Station Cuaca Online Berbasis Android dengan Menggunakan Database Firebase", which provided weather tracking but was limited in personalization for cyclists. Earlier, [8] Pramono et al. (2016) discussed "Monitoring Data Kecepatan dan Arah Angin Secara Real Time Melalui Web," highlighting the significance of wind conditions, yet without linking it to activity recommendations.

Research related to cycling applications also exists. [9] Luthfiansyah et al. (2022) designed a decision support system in "Sistem Pendukung Keputusan untuk Menentukan Rute Bersepeda," focusing on route optimization but without considering hourly weather conditions. [10] Dwi Rahmadani et al. (2022) introduced "Rancang Bangun Aplikasi Berkendara Sepeda Berbasis Virtual Reality dan Android dengan Menggunakan Metode Agile," offering immersive cycling experiences, though not providing weather-based recommendations. Furthermore, [11] Febi Irawati & Muh Fatwa Ananta Anshar (2024) discussed "Kepopuleran Olahraga Sepeda Sebagai Bagian Dari Pola Hidup Sehat," underlining the importance of cycling for health but without addressing optimal time selection.

Comparative studies on decision tree methods have also been conducted. [12] Prasandy et al. (2019), in "Perbandingan Hasil Penggunaan Metode Decision Tree dan Random Tree Pada Data Training Aplikasi Pencarian Tukang," highlighted differences in algorithmic performance, while [13] Devia (2023) applied C4.5 in "Penerapan Decision Tree dengan Algoritma C4.5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan." [14] Kharisma et al. (2023) evaluated "Penerapan Algoritma Decision Tree untuk Ulasan Aplikasi Vidio di Google Play," further demonstrating decision tree applicability in sentiment classification.

Recent work by [15] Rasyid et al (2025), "Implementasi Algoritma Decision Tree CART untuk Deteksi Dini Penyakit FLUTD pada Kucing," applied CART effectively for medical diagnosis. Their study illustrates the strength of regression-based decision trees in handling numerical prediction, which inspires the present research to implement Regression CART Decision Tree for continuous percentage-based cycling time recommendations.

From these prior studies, it is evident that while decision trees have been widely applied in classification and weather-related applications, none have specifically addressed the need for real-time, regression-based recommendations for optimal cycling times. This research aims to fill that gap by implementing a Regression CART Decision Tree model using weather parameters such as temperature, wind speed, and overall conditions.

3. Research Methodology

This research applies a quantitative approach with an experimental design, focusing on the implementation of a Regression CART (Classification and Regression Tree) Decision Tree model to provide optimal cycling time recommendations based on weather data. The overall methodology is divided into several stages, as described below:

Research Framework

The research begins with problem identification, followed by data collection, preprocessing, model implementation, evaluation, and application prototyping. The workflow can be described as:

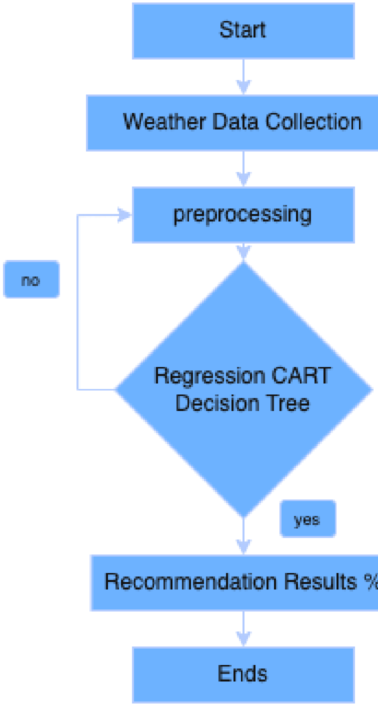

Figure 1. Framework Research

Data Collection

Weather data is obtained from publicly available APIs such as OpenWeatherMap, which provide hourly weather parameters. The variables considered include:

a. Temperature (°C) – to assess thermal comfort.
b. Wind speed (m/s) – to evaluate lateral wind tolerance.
c. Weather conditions (categorical: clear, cloudy, rain, etc.) – converted into numerical features for model processing.

In addition to quantitative data, qualitative insights were also collected through short interviews with several members of a local cycling community. The interviews aimed to capture cyclists' perceptions of comfortable weather conditions for cycling. Based on these discussions, cyclists generally considered:

a. Temperature: the ideal range for comfort is between 20–28°C,
b. Wind speed: cycling is still comfortable at speeds below 10 km/h, while higher speeds cause discomfort and safety concerns,
c. Weather conditions: clear weather is considered most favorable, cloudy weather is acceptable with moderate comfort, while rain is least preferred and often avoided.

Preprocessing

Collected weather data undergoes preprocessing steps:

a. Data cleaning: handling missing or inconsistent values.
b. Feature encoding: transforming categorical weather descriptions into numerical form.
c. Normalization: ensuring uniform data scaling for accurate prediction.

Regression CART Decision Tree

The core algorithm used is the Regression CART Decision Tree, which is suitable for predicting continuous output values. The process includes:

a. Splitting criteria: minimizing variance at each node using mean squared error (MSE).
b. Tree construction: recursively partitioning the dataset based on temperature, wind speed, and weather condition.
c. Prediction: producing a continuous score (0–1) that represents the percentage feasibility of cycling at a given hour.

To adapt the model for practical needs, the recommendation percentage is formulated using a weighted scoring function derived from interviews with active cyclists. The formula is:

UNIVERSITAS UTPADAKA SWASTIKA
Jl. KS Tubun No.11 Tangerang 15112 Banten. Telp. (021) 5589161-62 Fax. (021) 5589163
JISI Jurnal Ilmiah Sistem Informasi
Vol. 3 No. 02 (2025)
ISSN: 3046-711X

$$R = \frac{Ws + Ww + Wc}{3}$$

R : Percentage of recommendation (0–100%)
Ws : Temperature weight, based on the ideal range of 20–28°C
Ww : Wind speed weight, ideal value < 10 km/h
Wc : Weather condition weight (clear = highest score, cloudy = medium, rain = lowest score)

4. Discussion of Results

4.1 Data Description and Classification

To avoid binary scoring and better reflect real conditions, a reduction factor (decay factor) is introduced for temperature and wind speed:

a. Temperature (Ws): if the actual temperature is outside the 20–28°C range, the score is reduced proportionally with a factor of 8 per degree deviation from the nearest boundary.
b. Wind speed (Ww): values above 10 km/h are reduced with a factor of 10 deviation per km/h exceeding the threshold.
c. Weather condition (Wc) remains categorical, where clear = 100, cloudy = 60, and rain = 20.

This adjustment ensures that slightly non-ideal conditions (e.g., 29°C or 11 km/h wind) are not treated the same as extreme conditions (e.g., 35°C or 20 km/h wind). Instead, the scoring decreases gradually, making the model's recommendations more realistic and aligned with cyclists' comfort levels.

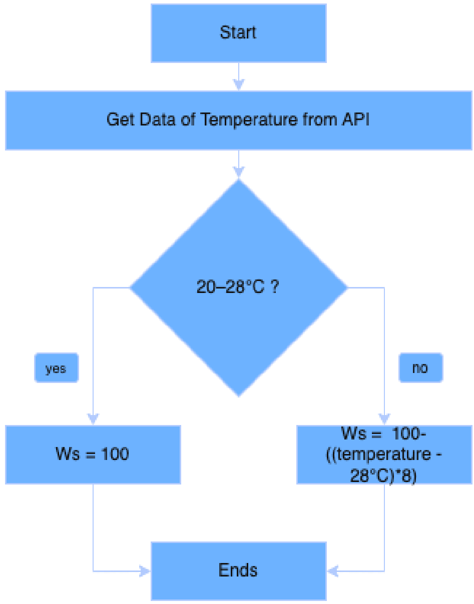

Figure 2. Temperature

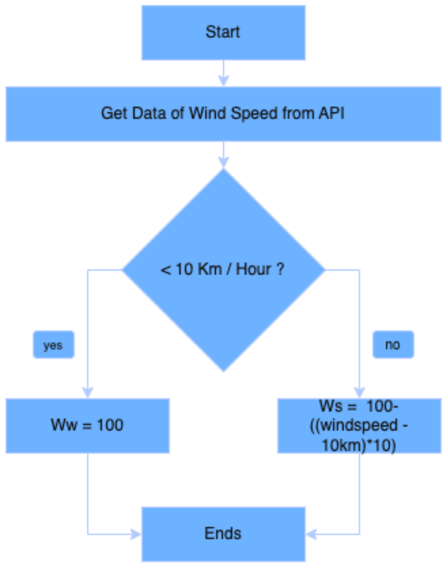

Figure 3. Wind Speed

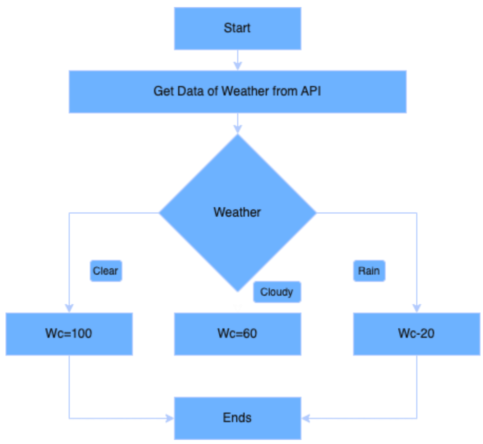

Figure 4. Weather

UNIVERSITAS UTPADAKA SWASTIKA
Jl. KS Tubun No.11 Tangerang 15112 Banten, Telp. (021) 5589161-62 Fax. (021) 5589163
JISI Jurnal Ilmiah Sistem Informasi
Vol. 3 No. 02 (2025)
ISSN: 3046-711X

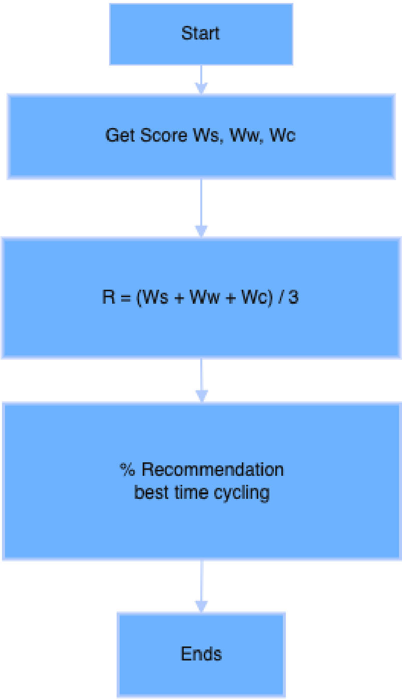

Figure 5. Result

4.2 Prototype Application Development

Prototype application is developed to implement the model. The app integrates weather data from the API, processes it through the Regression CART Decision Tree, and displays percentage-based recommendations for optimal cycling time in a user-friendly interface.

Figure 6. UI App Prototype

The user interface (UI) is designed to be as simple and intuitive as possible, allowing users to quickly understand the recommended time for cycling. Each time is displayed with its corresponding R (%) value. To make interpretation easier, color is applied Green: R ≥ 70% as comfortable or ideal for cycling and Red: R < 70% as not recommended or less comfortable. This Design enabling users to instantly identify the optimal periods for outdoor activities. This visualization approach helps users make quick decisions without analyzing numerical data manually.

4.3 Model Implementation

This table shows the recommendation percentage R (%) by data example for cycling activities from 07:00 to 12:00.

Table 1. Model Implementation

Time	Temp (°C)	Wind (km/h)	Weather	R (%)
07:00	25	8	Clear	100
08:00	27	9	Cloudy	87
09:00	29	12	Cloudy	77
10:00	30	13	Cloudy	71
11:00	32	15	Rain	46
12:00	34	16	Cloudy	51

4.4 Model Limitations

Despite its effectiveness in predicting continuous suitability scores, the Regression CART Decision Tree model has several limitations when applied to cycling time recommendations based on weather data:

Sensitivity to Data Variability

The model's performance heavily depends on the diversity and quality of the training data. Limited variations in weather conditions (e.g., few instances of rain or strong wind) can reduce its generalization ability when applied to unseen scenarios.

Overfitting on Training Data

Since Decision Trees tend to create complex rules for each data split, the Regression CART Decision Tree model may overfit the training data, capturing noise rather than meaningful patterns. This can lead to decreased accuracy when predicting on real-time or fluctuating weather data.

Limited Temporal Awareness

The model treats each hour's weather data independently, without considering temporal dependencies (e.g., gradual temperature changes or weather transitions across hours). As a result, it may fail to capture dynamic weather trends that affect cycling comfort.

Simplified Representation of Qualitative Factors

Although categorical weather conditions are numerically encoded, qualitative aspects such as humidity perception or visibility are not directly represented, potentially reducing model realism.

Static Thresholds for Comfort

The comfort thresholds (e.g., temperature 20–28°C, wind <10 km/h) are derived from general cyclist preferences, which may not reflect individual tolerance differences, thus limiting personalization in recommendations.

These limitations suggest that future work could integrate ensemble methods (e.g., Random Forests) or temporal models (e.g., LSTM) to improve accuracy and robustness in dynamic weather prediction contexts.

5. Conclusion and Suggestions

This study successfully implemented a Regression CART Decision Tree model to recommend optimal cycling times based on temperature, wind speed, and weather conditions. The model effectively identifies suitable periods using a comfort index (R %), visually shown through color indicators, and is supported by an intuitive user interface. Overall, the approach proves practical and accurate for real-time cycling recommendations, with potential improvements through larger datasets and additional weather factors.

Expand Data Sources: Future research can incorporate a larger and more diverse set of weather data, including humidity, precipitation intensity, and air quality, to improve the accuracy of cycling comfort predictions.

Real-time Data Integration: Integrating real-time weather sensors or APIs can allow the model to provide dynamic and up-to-date recommendations for cyclists.

Advanced Machine Learning Models: Exploring other machine learning techniques, such as Random Forest, Gradient Boosting, or Neural Networks, may improve prediction accuracy compared to a single Regression CART Decision Tree.

User Feedback Incorporation: Including feedback from actual cyclists can help calibrate the comfort index (R %) more precisely according to user perception.

Mobile or Web Application Development: Developing a full-fledged mobile or web application with interactive UI can enhance usability, enabling users to access recommendations conveniently on the go.

Seasonal and Geographic Variation: Future studies can analyze seasonal differences and adapt the model for different geographic locations to increase the generalizability of the recommendations.

Future work can expand datasets, add more weather factors, and adopt advanced models such as Random Forest or LSTM.

Literature Review can expand with International Journal.

References

[1] A. Christyana and T. Angel, "Implementasi Aplikasi Android untuk Sistem Penyewaan Sepeda," 2022. [Online]. Available: https://www.researchgate.net/publication/360783326

[2] M. Rizal, A. Kamaludin, H. Ramadhan, N. I. Rizqi, A. Khunaefi, and S. Artikel, "Perancangan Aplikasi BaikinSepeda Berbasis Mobile Dengan Menggunakan

Flutter," Jurnal Manajemen Sistem Informasi, 2023, doi: 10.51920/jurminsi.

[3] Ahmad Taufiq Ramadhan, Faishal Hilmy F. G, Nadya Rafaela Puteri, and Alifya Meirza, "Penerapan Algoritma Decision Tree Dalam Melakukan Analisis Klasifikasi Harga Handphone," Jurnal Sistem Informasi dan Ilmu Komputer, vol. 1, no. 4, pp. 195–206, Nov. 2023, doi: 10.59581/jusiik-widyakarya.v1i4.1861.

[4] M. Fahrizal et al., "Implementasi Metode Decision Tree C4.5 Pada Lampu Otomatis Berdasarkan Kebiasaan Penghuni Rumah," Coding : Jurnal Komputer dan Aplikasi, 2022.

[5] M. Naufal Baharudin, I. Kadek, and D. Nuryana, "Implementasi Algoritma Decision Tree untuk Klasifikasi Surat pada Aplikasi Mobile E-Surat Dinas Komunikasi dan Informatika Kota Kediri Berbasis Android," Journal of Informatics and Computer Science, vol. 04, 2022.

[6] D. Sopyandi et al., "Bangun Aplikasi Tracking Cuaca (Weather App) Menggunakan Public API Berbasis Website," Jurnal Ilmiah Wahana Pendidikan, vol. 2024, no. 20, pp. 209–215, 2024, doi: 10.5281/zenodo.14274718.

[7] S. Gunawan Zain, D. Darma Andayani, E. Prasetyo, and M. Ramdhany Edy, "Pengembangan Station Cuaca Online Berbasis Android dengan Menggunakan Database Firebase," 2022. [Online]. Available: https://ojs.unm.ac.id/JESSI/index

[8] Y. Pramono, Warsito, and Syafriadi, "Monitoring Data Kecepatan dan Arah Angin Secara Real Time Melalui Web," 2016.

[9] R. Luthfiansyah, N. Abdurahman, H. Muhammad, P. Putri, M. Rayhan Brameswara, and E. Riksakomara, "Sistem Pendukung Keputusan untuk Menentukan Rute Bersepeda," Sistem Informasi |, vol. 9, no. 1, pp. 83–88, 2022.

[10] I. Dwi Rahmadani, M. R. Ramadhan, and S. Istiqphara, "Rancang Bangun Aplikasi Berkendara Sepeda Berbasis Virtual Reality dan Android dengan Menggunakan Metode Agile," ELECTRON Jurnal Ilmiah Teknik Elektro, vol. 3, no. 1, pp. 01–05, May 2022, doi: 10.33019/electron.v3i1.7.

[11] A. Febi Irawati and A. Muh Fatwa Ananta Anshar, "Kepopuleran Olahraga Sepeda Sebagai Bagian Dari Pola Hidup Sehat," Jurnal Ilmiah Global Education, vol. 5, no. 3, pp. 2038–2043, 2024, doi: 10.55681/jige.v5i3.3299.

[12] T. Prasandy, K. Nurkhasanah, M. P. Sari, and T. R. Fazry, "Perbandingan Hasil Penggunaan Metode Decision Tree dan Random Tree Pada Data Training Aplikasi Pencarian Tukang," ULTIMA InfoSys, vol. X, no. 2, 2019.

[13] E. Devia, "Penerapan Decision Tree Dengan Algoritma C4.5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan," Jurnal Information System, 2023.

[14] I. L. Kharisma, D. A. Septiani, A. Fergina, and K. Kamdan, "Penerapan Algoritma Decision Tree untuk Ulasan Aplikasi Vidio di Google Play," Jurnal Nasional Teknologi dan Sistem Informasi, vol. 9, no. 2, pp. 218–226, Sep. 2023, doi: 10.25077/teknosi.v9i2.2023.218-226.

[15] A. Rasyid, S. Gilbijatno, A. Wahyu Pramudya, and D. Prasetyo, "Implementasi Algoritma Decision Tree CART untuk Deteksi Dini Penyakit FLUTD pada Kucing," 2025. Accessed: May 20, 2025. [Online]. Available: https://proceeding.unpkediri.ac.id/index.php/stains/article/view/6024