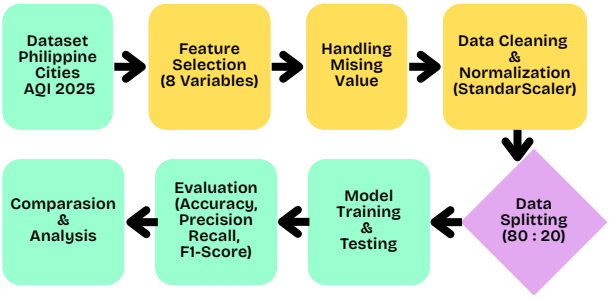
COMPARATIVE PERFORMANCE OF C4.5 AND CART FOR AIR QUALITY CLASSIFICATION IN PHILIPPINE CITIES

SUBIAN SAIDI, RIDHO SHOLEHURROHMAN, NUR HAMZAH, WARTARIUS, WIRA ADIGUNA SUTAWA, MUHAQIQIN

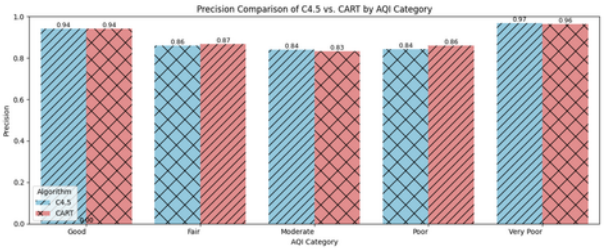
1. INTRODUCTION

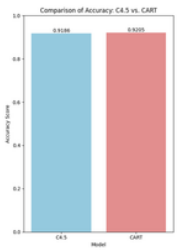
- Air pollution is a global environmental problem that has a significant impact on human health and sustainable development.
- In the Philippines, air quality has become a major concern due to high pollutant concentrations resulting from transportation activities, industrial activities, and biomass burning.
- This study compares the performance of the C4.5 and CART algorithms in classifying air quality levels in Philippine cities using the Philippine Cities Air Quality Index 2025 dataset.
- The objective is to determine the better-performing model and provide recommendations for an interpretable machine-learning-based air quality monitoring system.

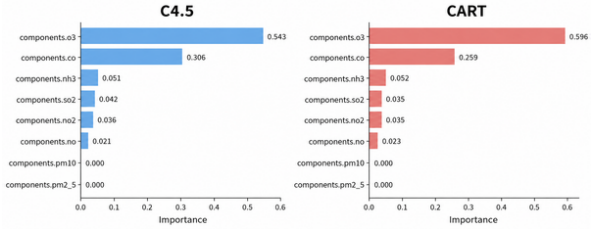
2. METHOD

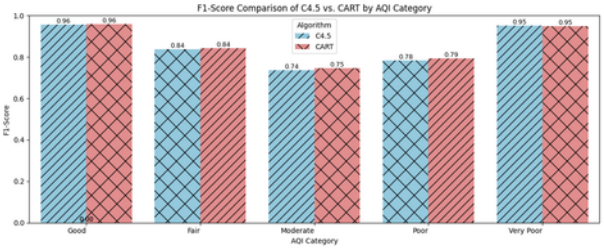


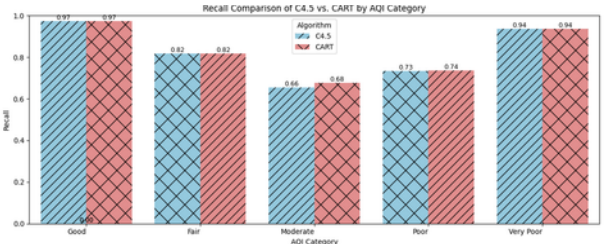
3. RESULT











Model	Metric (macro average)			
	Accuracy	Precision	Recall	F1-Score
C4.5	0.918585	0.890707	0.82315	0.853648
CART	**0.920455**	**0.893282**	**0.828179**	**0.857931**

- CART demonstrates slightly higher performance than C4.5 across all evaluation metrics.
- O_3 (ozone) is the most important predictor in both models, followed by CO (carbon monoxide).
- Both models show lower performance for the Poor class, indicating the need for improved detection of this category.
- Decision tree models are interpretable, making them suitable for supporting air quality warning systems in the Philippines.

4. CONCLUSION

- C4.5 and CART are capable of classifying air quality levels in Philippine cities with high performance.
- CART provides a small numerical advantage, achieving 92.05% accuracy and an 85.79% F1-score, compared with C4.5.
- O_3 and CO are important predictors of air quality classification.
- These results can serve as a basis for developing an interpretable machine-learning-based air quality monitoring and early warning system

MATHEMATICS DEPARTMENT, FACULTY OF MATHEMATICS AND NATURAL SCIENCES, UNIVERSITAS LAMPUNG