PROJECT REPORT

A COMPARATIVE ANALYSIS OF GAUSSIAN NAIVE BAYES, MULTINOMIAL NAIVE BAYES, AND K-NEAREST NEIGHBOR PERFORMANCE TO ANALYZE RAIN PREDICTION IN THE SEMARANG CITY AREA

SOEN ANITA SANJAYA
21.K1.0012

Faculty of Computer Science
Soegijapranata Catholic University
2024

ABSTRACT

Semarang, known for its heavy rainfall, faces a significant impact on the daily lives of its residents. Given the erratic rainfall patterns and frequent changes, accurate predictions have become crucial for disaster response and preparedness. While various algorithmic methods have been used, implementing Gaussian Naive Bayes, Multinomial Naive Bayes, and K-Nearest Neighbor (KNN) remains less than optimal, often being perceived as having lower accuracy than other algorithms. However, several studies indicate that Naive Bayes methods are widely employed and have even achieved up to 84.26% accuracy. This research aims to explore and compare the performance of these three algorithms in predicting rainfall in Semarang. By utilizing historical weather data from 2013 to 2024 obtained from dataonline.bmkg.go.id, this study evaluates the performance of each algorithm. Through data processing and analysis with Gaussian Naive Bayes, Multinomial Naive Bayes, and KNN, the expected outcome is to identify which algorithm provides the most precise and reliable rainfall predictions based on precision, accuracy, recall, and F1 score metrics.

Keywords: Rainfall Prediction, Gaussian Naive Bayes, Multinomial Naive Bayes, K-Nearest Neighbor, Semarang, Data Mining, Weather Forecasting