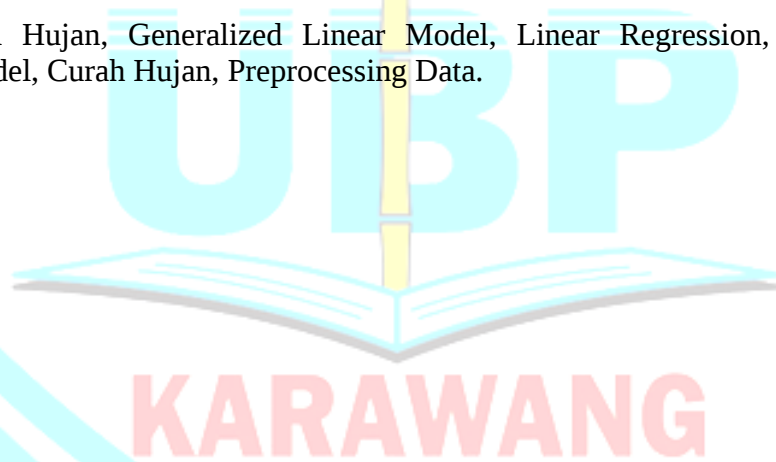


ABSTRAK

Prediksi curah hujan sangat penting bagi berbagai aktivitas yang dipengaruhi kondisi cuaca, khususnya di negara beriklim tropis mengalami kondisi ini secara signifikan. Prediksi curah hujan yang akurat sangat penting untuk mendukung berbagai aspek perencanaan kota, termasuk pengelolaan sumber daya air dan mitigasi risiko bencana banjir. Penelitian ini membandingkan dua algoritma *machine learning*, *Generalized Linear Model* (GLM) dan *Linear Regression*, dalam memprediksi curah hujan berdasarkan fitur cuaca seperti suhu, kelembaban, tekanan, angin, tutupan awan, dan data historis. Selanjutnya diproses melalui *encoding* yang dimana akan mengubah nilai kategorikal menjadi nilai numerik, normalisasi yang melibatkan penyesuaian ulang nilai nilai dalam dataset, dan penanganan *class imbalance* untuk melakukan duplikasi *sample* pada kelas minoritas. Setelah dibagi menjadi data latih dan uji, kedua algoritma diterapkan dan dievaluasi menggunakan akurasi, RMSE, dan MAE. Hasilnya, GLM memiliki akurasi sebesar 90.17% lalu untuk RMSE sebesar 0.3949 dan MAE 0.3836, sedangkan *Linear Regression* lebih baik dalam nilai MAE sebesar 0.2656 dan RMSE 0.3218 untuk akurasi sebesar 89.26%. Dengan pendekatan analisis yang tepat, pola tersebut dapat dimanfaatkan untuk mendukung keputusan dan perencanaan secara lebih terarah..

Kata Kunci: Prediksi Hujan, Generalized Linear Model, Linear Regression, Machine Learning, Evaluasi Model, Curah Hujan, Preprocessing Data.



ABSTRACT

Rainfall prediction is very important for various activities that are influenced by weather conditions, especially in tropical countries that experience this condition significantly. Accurate rainfall prediction is very important to support various aspects of city planning, including water resource management and flood disaster risk mitigation. This study compares two machine learning algorithms, Generalized Linear Model (GLM) and Linear Regression, in predicting rainfall based on weather features such as temperature, humidity, pressure, wind, cloud cover, and historical data. Furthermore, it is processed through encoding which will change categorical values into numeric values, normalization which involves readjusting the values in the dataset, and handling class imbalance to duplicate samples in the minority class. After being divided into training and test data, both algorithms are applied and evaluated using accuracy, RMSE, and MAE. As a result, GLM has an accuracy of 90.17% then for RMSE of 0.3949 and MAE of 0.3836, while Linear Regression is better in MAE values of 0.2656 and RMSE of 0.3218 for an accuracy of 89.26%. With the right analysis approach, the pattern can be used to support decisions and planning in a more targeted manner.

Keyword: Rain Prediction, Generalized Linear Model, Linear Regression, Machine Learning, Model Evaluation, Rainfall, Data Preprocessing.

