

ABSTRAK

WULAN NOVITASARI. Prediksi Jumlah Kasus ISPA di Tanjungpinang Menggunakan Algoritma XGBoost dengan Optimasi *Hyperparameter*. Dibimbing oleh NOLA RITHA dan FORTIA MAGFIRA.

Infeksi Saluran Pernapasan Akut (ISPA) merupakan salah satu masalah kesehatan utama di Kota Tanjungpinang dengan jumlah kasus yang tinggi dan dipengaruhi oleh kondisi iklim yang tidak stabil. Hubungan antara faktor iklim dan jumlah kasus ISPA bersifat kompleks sehingga sulit dimodelkan menggunakan metode statistik konvensional. Penelitian ini bertujuan menganalisis performa algoritma XGBoost dengan optimasi *hyperparameter* menggunakan *Grid Search* dan *Random Search* dalam memprediksi jumlah kasus ISPA harian berdasarkan faktor iklim, serta menentukan metode optimasi dengan performa terbaik. Data yang digunakan berupa data harian periode Januari 2022–April 2026 sebanyak 1.580 data, dengan variabel suhu, kelembapan, curah hujan, kecepatan angin, dan fitur lag jumlah kasus ISPA. Evaluasi model dilakukan menggunakan RMSE dan MAE. Hasil penelitian menunjukkan bahwa optimasi *hyperparameter* mampu meningkatkan performa model XGBoost *Default*. Model dengan optimasi *Random Search* menghasilkan performa terbaik dengan RMSE 40,090 dan MAE 28,639, lebih baik dibandingkan *Grid Search* (RMSE 40,096; MAE 28,815) maupun *Default* (RMSE 42,576; MAE 30,052). Analisis SHAP menunjukkan bahwa fitur lag jumlah kasus ISPA tujuh hari sebelumnya (*ispalag_7*) merupakan variabel yang memberikan kontribusi terbesar terhadap hasil prediksi. Model ini berpotensi mendukung pengambilan keputusan dalam mitigasi dan penanganan ISPA di Kota Tanjungpinang.

Kata kunci: *Grid Search*, ISPA, *Random Search*, SHAP, XGBoost.

ABSTRACT

WULAN NOVITASARI. Prediction of ISPA Cases in Tanjungpinang Using the XGBoost Algorithm with Hyperparameter Optimization. Supervised by NOLA RITHA and FORTIA MAGFIRA.

Acute Respiratory Infection (ARI) is one of the major health problems in Tanjungpinang City, with a high number of cases influenced by unstable climate conditions. The relationship between climate factors and the number of ARI cases is complex, making it difficult to model using conventional statistical methods. This study aims to analyze the performance of the XGBoost algorithm with hyperparameter optimization using Grid Search and Random Search in predicting the daily ARI cases based on climate factors, and to determine the optimization method with the best performance. The dataset consisted of 1,580 daily records collected from January 2022 to April 2026, including temperature, humidity, rainfall, wind speed, and lag features of ARI cases. Model performance was evaluated using RMSE and MAE. The results showed that hyperparameter optimization improved the performance of the Default XGBoost model. The model optimized with Random Search achieved the best performance with an RMSE of 40.090 and MAE of 28.639, outperforming Grid Search (RMSE = 40.096; MAE = 28.815) and the Default model (RMSE = 42.576; MAE = 30.052). SHAP analysis indicated that the seven-day lag feature of ARI cases (*ispalag_7*) contributed the most to the prediction results. The proposed model has the potential to support decision-making for ARI mitigation and response efforts in Tanjungpinang City.

Keywords: Grid Search, ARI, Random Search, SHAP, XGBoost.