

Pemerintah Indonesia berkomitmen mengurangi emisi karbon dengan meratifikasi Perjanjian Paris melalui Undang-Undang Nomor 16 Tahun 2016. Pemerintah berupaya dengan menargetkan penurunan emisi Gas Rumah Kaca (GRK) sebesar 31,89% pada 2030 untuk mencapai *net zero emission* pada 2060 atau lebih cepat. Untuk mencapai target ini, peralihan menuju penggunaan energi terbarukan menjadi langkah krusial.

Jumlah energi surya lebih dari 3.200 GW yang dimiliki Indonesia menjadi potensi untuk dikelola melalui Pembangkit Listrik Tenaga Surya (PLTS). Pengelolaan PLTS juga membutuhkan dukungan perencanaan operasional yang lebih efisien serta manajemen risiko yang lebih baik untuk mendapatkan keefektifan pemanfaatan energi. Salah satu pendekatan yang mendukung adalah dengan pengembangan model prediksi produksi energi pada PLTS yang akurat dan dapat diandalkan. Penelitian ini bertujuan untuk mengembangkan model prediksi produksi energi pada PLTS dengan menggunakan *machine learning* pada dataset yang berlokasi di Indonesia. Dua model algoritma XGBoost dan LightGBM yang memiliki performa baik dalam tugas tersebut berdasarkan penelitian terdahulu dibandingkan. Pengoptimalan parameter pada model dilakukan dengan menggunakan *Tree-structured Parzen Estimator* (TPE) untuk meningkatkan akurasi prediksi serta membandingkan pada kondisi yang lebih optimal.

Hasil penelitian didapatkan XGBoost memberikan performa yang sedikit lebih unggul, dimana didapatkan hasil R^2 0,99472, RMSE 0,76864, dan MAE 0,39478 dibandingkan dengan LightGBM R^2 0,99466, RMSE 0,77294, dan MAE 0,41702. Namun, LightGBM memberikan stabilitas hasil prediksi yang lebih baik dengan nilai standar deviasi 0,65079 dibandingkan dengan XGBoost 0,65951 dan LightGBM juga memberikan risiko *overfitting* yang lebih kecil dibandingkan dengan XGBoost. Kedua model tersebut dikembangkan dengan *hyperparameter* optimal yang didapatkan untuk model XGBoost adalah *no of estimators* 912, *learning rate* 0,08, *max depth* 4, *subsample* 0,7, *colsample bytree* 0,97, *gamma* 0,22, dan *min child weight* 10,9. Sedangkan, untuk LightGBM didapatkan *no of estimators* 892, *learning rate* 0,14, *max depth* 4, *num leaves* 6, dan *min data in leaf* 29. Model yang dikembangkan dengan *hyperparameter optimization* dibandingkan dengan menggunakan *hyperparameter default* menghasilkan akurasi prediksi dan stabilitas prediksi yang meningkat.

Kata kunci : Model Prediksi, Produksi Energi PLTS, XGBoost, LightGBM, *Tree-structured Parzen Estimators*

ABSTRACT

The Indonesian government is committed to reducing carbon emissions by ratifying the Paris Agreement through Undang-Undang Nomor 16 Tahun 2016. The government aims to reduce greenhouse gas emissions by 31.89% by 2030 to achieve net zero emissions by 2060 or sooner. Achieving this target requires a crucial shift towards the use of renewable energy.

Indonesia's solar energy potential exceeds 3,200 GW, which represents a significant opportunity to be managed through solar power plants. Effective management of solar power plant also requires support in efficient operational planning and better risk management. One approach that supports this is the development of accurate and reliable energy production prediction models for solar power plant. This research aims to develop energy production prediction models for solar power plant using machine learning on datasets located in Indonesia. Two algorithms, XGBoost and LightGBM, known for their strong performance in such tasks, are compared. Hyperparameter optimization for the models is performed using Tree-structured Parzen Estimator (TPE) to enhance prediction accuracy and compare performance under optimal conditions.

The research results XGBoost slightly outperforms with an R^2 of 0.99472, RMSE of 0.76864, and MAE of 0.39478 compared to LightGBM's with R^2 of 0.99466, RMSE of 0.77294, and MAE of 0.41702. However, LightGBM exhibited better prediction stability with a lower standard deviation of 0.65079 compared to XGBoost's 0.65951. Additionally, LightGBM showed a lower risk of overfitting compared to XGBoost. Both models are developed with optimal hyperparameters, where for XGBoost, the optimal settings are: number of estimators 912, learning rate 0.08, max depth 4, subsample 0.7, colsample bytree 0.97, gamma 0.22, and min child weight 10.9. For LightGBM, the optimal settings are: number of estimators 892, learning rate 0.14, max depth 4, num leaves 6, and min data in leaf 29. The models developed with hyperparameter optimization compared to using default hyperparameters resulted in improved prediction accuracy and prediction stability.

Keywords: Prediction Model, Solar Power Plants Energy Production, XGBoost, LightGBM, Tree-structured Parzen Estimators