ABSTRAK

PT Alis Jaya Ciptatama menghadapi kendala operasional berupa deviasi yang signifikan antara durasi pengeringan papan kayu teoritis (14 hari) dengan kondisi aktual di lapangan yang bervariasi antara 9 hingga 43 hari. Kadar air (*moisture content*) merupakan kriteria fundamental yang menentukan stabilitas dimensi dan kualitas kayu, sehingga ketidakpastian waktu pengeringan ini menyebabkan hambatan arus produksi (*bottleneck*) serta kesulitan dalam perencanaan logistik pengiriman barang. Penelitian ini bertujuan untuk mengembangkan model *machine learning* yang mampu memprediksi durasi waktu pengeringan secara akurat berdasarkan data historis operasional perusahaan periode Januari 2024 hingga Maret 2026.

Metodologi penelitian mencakup tahap *data preprocessing* melalui agregasi data pada level *batch*, pembersihan data, dan penanganan *outlier*. Selain itu, dilakukan *feature engineering* untuk menghasilkan fitur interaksi cuaca dan beban fisik muatan. Ukuran *dataset* awal hanya terdiri dari 166 *batch* dilakukan augmentasi untuk memperkaya variasi sampel hingga mencapai 311 data latih untuk meningkatkan kemampuan generalisasi model. Lima algoritma regresi dibandingkan dalam penelitian ini, yaitu *Multiple Linear Regression*, *Random Forest, Extra Trees, Gradient Boosting,* dan *XGBoost*.

Hasil penelitian menunjukkan bahwa algoritma *XGBoost* hasil *hyperparameter tuning* memberikan performa terbaik dan ditetapkan sebagai model final. Pengujian pada skenario dengan penyertaan fitur cuaca (Skenario 1) terbukti lebih unggul dengan nilai $R2$ sebesar 0,8401 dan MAE sebesar 1,7335 hari pada data uji independen. Performa ini jauh melampaui model *baseline Multiple Linear Regression* yang hanya memperoleh $R2$ sebesar 0,0164, mengonfirmasi adanya hubungan non-linear yang kompleks antara variabel prediktor dengan durasi pengeringan. Analisis *feature importance* mengungkapkan bahwa fitur *bulan_in* memberikan kontribusi paling dominan sebesar 29,73%, diikuti oleh fitur interaksi *hujan_x_tebal* (7,81%) dan *target mc_akhir_ket_max* (7,68%). Model ini diimplementasikan ke dalam aplikasi berbasis web menggunakan *framework* Streamlit untuk mendukung pengambilan keputusan penjadwalan produksi dengan 82,4% prediksi berada dalam rentang toleransi kesalahan ±3 hari.

Kata Kunci: *Machine learning*, *Kiln dry*ing, *XGBoost*, Prediksi Durasi, Industri Pengolahan Kayu.

ABSTRACT

PT Alis Jaya Ciptatama faces operational obstacles in the form of significant deviations between the theoretical wood board drying duration (14 days) and actual field conditions, which vary between 9 to 43 days. Moisture content is a fundamental criterion that determines the dimensional stability and quality of wood, so this drying time uncertainty causes production flow obstacles (bottlenecks) and difficulties in logistics planning for shipments. This research aims to develop a machine learning model capable of accurately predicting drying duration based on the company's historical operational data for the period of January 2024 to March 2026.

The research methodology includes the data preprocessing stage through batch-level data aggregation, data cleaning, and outlier handling; feature engineering to produce weather interaction and physical load features; as well as data augmentation using Gaussian noise, KNN interpolation, and targeted oversampling techniques to enrich the variety of training samples. Five regression algorithms are compared in this research, namely Multiple Linear Regression, Random Forest, Extra Trees, Gradient Boosting, and XGBoost.

The research results show that the XGBoost algorithm from hyperparameter tuning results provides the best performance and is established as the final model. Testing on a scenario with the inclusion of weather features (Scenario 1) is proven to be superior with an R2 value of 0.8401 and an MAE of 1.7335 days on independent test data. This performance far exceeds the Multiple Linear Regression baseline model which only obtained an R2 of 0.0164, confirming a complex non-linear relationship between the predictor variables and drying duration. Feature importance analysis reveals that the bulan_in feature provides the most dominant contribution at 29.73%, followed by the hujan_x_tebal interaction feature (7.81%) and the mc_akhir_ket_max target (7.68%). This model is implemented into a web-based application using the Streamlit framework to support production scheduling decision-making with 82.4% of predictions within the error tolerance range of ±3 *days.*

Keywords*:* *Machine learning, Kiln drying, XGBoost, Duration Prediction, Wood Processing Industry*.