

## ABSTRAK

Perkembangan teknologi informasi menuntut setiap perusahaan untuk mampu memanfaatkan data penjualan sebagai dasar pengambilan keputusan bisnis. UD Timbul Jaya, sebagai usaha dagang bahan bangunan di Kabupaten Bangkalan, menghadapi fluktuasi penjualan yang belum dapat diprediksi secara akurat, sehingga sering terjadi kelebihan atau kekurangan stok barang. Untuk mengatasi hal tersebut, penelitian ini menerapkan metode Regresi Linier Sederhana sebagai salah satu teknik data mining prediktif untuk memodelkan hubungan antara waktu dan jumlah penjualan produk. Data yang digunakan merupakan data penjualan lima produk utama, yaitu Semen, Besi, Pasir, Batu Bata, dan Cat Tembok selama periode Januari 2022 hingga Mei 2025. Proses penelitian meliputi tahap pengumpulan data, pembersihan data, transformasi, pembangunan model, serta evaluasi model menggunakan metrik  $R^2$ , MSE, RMSE, dan MAE. Hasil analisis menunjukkan bahwa model regresi linier memberikan tingkat akurasi terbaik pada produk Besi ( $R^2 = 0,62$ , RMSE = 11,65) dan Cat Tembok ( $R^2 = 0,59$ , RMSE = 12,07), sedangkan produk Batu Bata dan Pasir memiliki akurasi lebih rendah karena fluktuasi musiman. Secara keseluruhan, model ini mampu menjelaskan 47% variasi data penjualan dan dapat digunakan sebagai alat prediksi awal untuk mendukung perencanaan stok dan strategi penjualan berbasis data di UD Timbul Jaya.

*Kata Kunci: Data Mining, Prediksi Penjualan, Regresi Linier, Evaluasi Model, UD Timbul Jaya*

## ABSTRACT

*The The development of information technology requires every company to be able to utilize sales data as a basis for business decision-making. UD Timbul Jaya, a building materials trading company located in Bangkalan Regency, faces fluctuations in sales that cannot be accurately predicted, often resulting in either excess or shortage of stock. To address this issue, this study applies the Simple Linear Regression method as one of the predictive data mining techniques to model the relationship between time and sales volume. The data used in this research consists of sales records for five main products, namely cement, iron, sand, bricks, and wall paint, covering the period from January 2022 to May 2025. The research process includes data collection, cleaning, transformation, model development, and model evaluation using metrics such as the coefficient of determination, mean squared error, root mean squared error, and mean absolute error. The analysis results show that the linear regression model provides the best accuracy for iron and wall paint products, while brick and sand products have lower accuracy due to seasonal fluctuations. Overall, the model can explain a significant portion of the sales variation and can be used as an initial predictive tool to support stock planning and data-driven sales strategies at UD Timbul Jaya*

*Keywords: Data Mining, Sales Prediction, Linear Regression, Data Analysis, UD Timbul Jaya*