

ABSTRAK

Pajak kendaraan bermotor merupakan sumber utama Pendapatan Asli Daerah (PAD). Namun, tingkat kepatuhan pembayaran pajak kendaraan bermotor di Kabupaten Aceh Utara masih suboptimal, terutama terkait dengan keterlambatan pembayaran. Pendekatan berbasis data diperlukan untuk memprediksi dan memahami pola kepatuhan wajib pajak secara lebih akurat. Studi ini bertujuan untuk membandingkan kinerja metode Regresi Logistik dan Random Forest dalam memprediksi kepatuhan pembayaran pajak kendaraan bermotor, serta mengidentifikasi faktor-faktor yang mempengaruhi perilaku kepatuhan wajib pajak di Samsat Aceh Utara (Sat). Studi ini menggunakan data sekunder berupa transaksi pembayaran pajak kendaraan bermotor di Samsat Aceh Utara periode 2022–2024, dengan total 100.000 observasi. Variabel respons adalah status kepatuhan pembayaran pajak (patuh dan tidak patuh), sedangkan variabel prediktor meliputi umur kendaraan, jenis kepemilikan, jenis kendaraan, dan merek kendaraan. Data dibagi menjadi 70% data pelatihan dan 30% data pengujian. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, recall, dan Area Under Curve (AUC). Hasil analisis menunjukkan bahwa Random Forest memiliki kinerja prediksi yang lebih baik daripada Regresi Logistik, dengan akurasi dan nilai AUC yang lebih tinggi. Usia kendaraan dan jenis kepemilikan merupakan variabel yang paling berpengaruh dalam memprediksi kepatuhan pembayaran pajak, sedangkan merek kendaraan memiliki pengaruh yang relatif lebih kecil. Regresi Logistik memberikan interpretasi yang jelas tentang hubungan antar variabel, tetapi memiliki kemampuan diskriminasi yang lebih rendah daripada Random Forest. Random Forest terbukti lebih efektif sebagai model prediksi untuk kepatuhan pembayaran pajak kendaraan bermotor di Samsat Aceh Utara. Penerapan model prediksi berbasis pembelajaran mesin berpotensi mendukung pembuatan kebijakan yang lebih tepat sasaran dalam upaya meningkatkan kepatuhan pembayaran pajak kendaraan bermotor, terutama dalam mengurangi keterlambatan pembayaran.

Kata kunci: Pajak, Pajak Kendaraan Bermotor, Regresi Logistik, Random Forest,

Klasifikasi

ABSTRACT

Motor vehicle tax is a major source of Regional Original Income (PAD). However, the level of motor vehicle tax payment compliance in North Aceh Regency is still suboptimal, particularly related to late payments. A data-driven approach is needed to predict and understand taxpayer compliance patterns more accurately. This study aims to compare the performance of the Logistic Regression and Random Forest methods in predicting motor vehicle tax payment compliance, as well as to identify factors that influence taxpayer compliance behavior at the North Aceh Samsat (Sat). This study uses secondary data in the form of motor vehicle tax payment transactions at the North Aceh Samsat for the 2022–2024 period, totaling 100,000 observations. The response variable is the tax payment compliance status (compliant and non-compliant), while the predictor variables include vehicle age, type of ownership, vehicle type, and vehicle brand. The data is divided into 70% training data and 30% testing data. The performance evaluation model is conducted using accuracy, precision, recall, and Area Under Curve (AUC) metrics. The analysis results show that Random Forest has better predictive performance than Logistic Regression, with higher accuracy and AUC values. Vehicle age and type of ownership are the most influential variables in predicting tax payment compliance, while vehicle brand has a relatively smaller influence. Logistic Regression provides a clear interpretation of the variable relationship, but has lower discrimination ability than Random Forest. Random Forest has proven to be more effective as a prediction model for motor vehicle tax payment compliance at the North Aceh Samsat. The application of machine learning-based predictive models has the potential to support more targeted policy making in an effort to improve motor vehicle tax payment compliance, especially in reducing late payments.

Keywords: *Tax, Motor Vehicle Tax, Logistic Regression, Random Forest;*

Classificatio