

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pajak kendaraan bermotor merupakan salah satu komponen utama dalam Pendapatan Asli Daerah (PAD) yang berperan penting dalam pembangunan infrastruktur, layanan publik, dan penyelenggaraan pemerintahan daerah. Namun, di Kabupaten Aceh Utara, tingkat kepatuhan wajib pajak kendaraan bermotor masih jauh dari optimal (Nisa, 2024). Hal ini menimbulkan kekhawatiran, mengingat ketidakpatuhan wajib pajak dapat menghambat pencapaian target PAD yang telah ditetapkan, serta berimplikasi terhadap kinerja keuangan dan pembangunan daerah secara keseluruhan.

Pajak kendaraan bermotor menjadi sumber utama Pendapatan Asli Daerah (PAD) yang mendukung kemandirian fiskal pemerintah daerah. Namun, dalam pelaksanaannya, masih terdapat kesenjangan penerimaan (tax gap), yaitu perbedaan antara potensi pajak yang seharusnya diterima dengan jumlah yang benar-benar dibayarkan oleh wajib pajak (Wasana et al., 2021). Otoritas pajak daerah melalui Samsat berupaya meminimalkan kesenjangan tersebut melalui edukasi dan pemanfaatan teknologi informasi (Maulia & Lestari, n.d.).

Pemerintah daerah dan Samsat Aceh Utara telah melakukan berbagai upaya edukasi dan penegakan peraturan, pendekatan yang digunakan selama ini cenderung bersifat reaktif dan kurang berbasis data. Tidak adanya sistem prediktif yang mampu mengidentifikasi potensi ketidakpatuhan secara dini menjadi salah satu kendala utama dalam perumusan strategi kebijakan yang tepat sasaran dengan memanfaatkan perkembangan teknologi informasi yang menghadirkan metode-metode analisis data yang canggih, seperti Regresi Logistik dan *Random Forest*, yang terbukti efektif dalam berbagai studi untuk klasifikasi data biner, termasuk prediksi perilaku kepatuhan.

Regresi Logistik merupakan salah satu metode statistik yang sering digunakan untuk memprediksi variabel biner, seperti kepatuhan atau ketidakpatuhan wajib

pajak (Wasana et al., 2021). Metode ini memiliki kelebihan dalam interpretasi

hasil dan identifikasi pengaruh variabel bebas terhadap variabel terikat. Klasifikasi dengan Regresi Logistik dilakukan melalui suatu model matematika yang menghubungkan antara variabel respon dengan variabel bebas (Yuniarsyih R.A et al., 2025).

Random Forest menawarkan keunggulan dalam menangani data yang kompleks dan menghasilkan prediksi dengan akurasi, precision, recall, dan AUC tinggi. *Random Forest* mampu mengolah data dengan jumlah variabel yang besar dan mendeteksi interaksi antara variabel-variabel tersebut. *Random Forest* mengklasifikasikan objek dengan mengombinasikan sejumlah pohon keputusan (Sugianto & Triloka, n.d.)

Dengan membandingkan performa metode Regresi Logistik dan *Random Forest* dalam memprediksi kepatuhan pembayaran pajak kendaraan bermotor bermotor di Samsat Aceh Utara, diharapkan dapat diperoleh pemahaman yang lebih baik mengenai metode yang paling efektif untuk mengolah data kepatuhan pajak di wilayah tersebut.

Ada beberapa penelitian terdahulu yang dilakukan oleh Taly Purwapada tahun 2019 dimana penelitian ini menunjukkan bahwa penerapan skema under sampling, oversampling dan combine sampling pada model Regresi Logistik memberikan efek meningkatnya rata-rata nilai sensitivity dan turunnya rata-rata nilai akurasi, precision, recall, dan AUC dan specificity. Sedangkan pada model *Random Forest*, efek tersebut hanya terlihat dari hasil skema under sampling saja. Proses feature selection dapat menurunkan varian nilai akurasi, precision, recall, dan AUC, specificity, sensitivity dan AUC pada model Regresi Logistik dan *Random Forest* hanya pada skema tertentu. Model terbaik secara keseluruhan adalah model model Regresi Logistik dengan skema combine sampling dan tanpa proses feature selection dengan rata-rata nilai (Purwa, 2019).

Penelitian yang dilakukan oleh Ayu Sulistiara, Najmah Istikaanah, Niken Larasati pada tahun 2023 dimana membandingkan kedua metode tersebut menggunakan data cuaca di wilayah Jawa Tengah yang dibagi dalam tiga proporsi data latih yang berbeda, yaitu 60%, 70% dan 80%, dan evaluasi modelnya

menggunakan nilai area under curve (AUC). Rata-rata AUC dari metode Regresi Logistik dan *Random Forest* berturut-turut adalah 0,6923 dan 0,7419. Berdasarkan hasil analisis kedua metode tersebut, nilai AUC tertinggi didapatkan dari hasil klasifikasi menggunakan metode *Random Forest* (Larasati, 2023).

Penelitian yang dilakukan oleh Muhamad Syukron, Rukun Santoso, dan Tatik Widiyarih pada tahun 2020 dimana menunjukkan bahwa *Random Forest* dan XGboost memiliki akurasi, precision, recall, dan AUC sekitar 74% tetapi nilai recall kurang dari 2%. *Random Forest* SMOTE dan SMOTE XGboost memiliki nilai akurasi, precision, recall, dan AUC & recall lebih dari 75%. *Random Forest* SMOTE memiliki akurasi, precision, recall, dan AUC lebih tinggi untuk memprediksi kelas fibrosis sementara SMOTE XGboost lebih baik dalam memprediksi kelas sirosis. Variabel yang lebih berpengaruh dalam menentukan stadium hepatitis adalah variabel hasil pemeriksaan laboratorium (Syukron et al., 2020).

Rendahnya tingkat kepatuhan pembayaran pajak kendaraan bermotor di Samsat Aceh Utara, yang berdampak pada optimalisasi penerimaan Pendapatan Asli Daerah (PAD) (Wasana et al., 2021). Meski berbagai upaya telah dilakukan untuk meningkatkan kepatuhan wajib pajak, belum ada pendekatan prediktif berbasis data yang secara komprehensif mampu mengidentifikasi pola dan faktor yang memengaruhi perilaku wajib pajak. Namun, hingga saat ini belum ada penelitian komprehensif di wilayah Aceh Utara yang secara langsung membandingkan performa kedua metode ini dalam konteks prediksi kepatuhan pajak kendaraan bermotor. Samsat dapat menerapkan sistem pendeteksi dini berbasis data yang membantu dalam pengambilan keputusan yang lebih strategis dan efisien.

Berdasarkan latar belakang diatas, penelitian ini penting untuk dilakukan karena rendahnya tingkat kepatuhan wajib pajak kendaraan bermotor di Kabupaten Aceh Utara masih menjadi permasalahan yang belum terpecahkan secara efektif. Pendekatan konvensional yang selama ini digunakan, seperti sosialisasi dan penegakan hukum, belum sepenuhnya mampu meningkatkan kepatuhan secara signifikan. Pemanfaatan metode prediksi berbasis data untuk mendukung pengambilan keputusan yang lebih tepat sasaran dapat memberikan solusi

alternatif yang cerdas dan berbasis bukti dalam memetakan pola kepatuhan wajib pajak. Oleh karena itu, peneliti membuat penelitian yang berjudul **“Perbandingan Metode Regresi Logistik dan *Random Forest* Dalam Memprediksi Kepatuhan Pembayaran Pajak Kendaraan di Samsat Aceh Utara”**

1.1. Rumusan Masalah

Berdasarkan penjelasan dari latar belakang diatas, maka rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana perbandingan kinerja metode Regresi Logistik dan *Random Forest* dalam memprediksi kepatuhan pembayaran pajak kendaraan bermotor di Samsat Aceh Utara?
2. Faktor-faktor apa saja yang paling berpengaruh terhadap kepatuhan wajib pajak kendaraan bermotor di wilayah Samsat Aceh Utara?

1.2. Tujuan Penelitian

Berdasarkan penjelasan dari rumusan masalah diatas, maka tujuan pada penelitian ini adalah sebagai berikut :

1. Menganalisis dan membandingkan kinerja metode Regresi Logistik dan *Random Forest* dalam memprediksi kepatuhan pembayaran pajak kendaraan bermotor di Samsat Aceh Utara.
2. Mengidentifikasi faktor-faktor signifikan yang memengaruhi kepatuhan wajib pajak kendaraan bermotor berdasarkan data historis dari Samsat Aceh Utara.

1.3. Manfaat Penelitian

Berdasarkan penjelasan dari tujuan diatas, maka manfaat pada penelitian ini adalah sebagai berikut :

1. Mampu memberikan informasi dan rekomendasi berbasis data kepada pihak Samsat Aceh Utara mengenai metode Regresi Logistik dan *Random Forest* dalam analisis data kepatuhan pembayaran pajak.
2. Menyediakan model prediksi yang akurat untuk membantu pengambilan keputusan dalam perumusan kebijakan strategis peningkatan kepatuhan pajak kendaraan

bermotor.

1.4. Ruang Lingkup dan Batasan Penelitian

Berdasarkan penjelasan dari manfaat diatas, maka ruang lingkup dan batasan penelitian ini adalah sebagai berikut :

1. Objek penelitian adalah wajib pajak kendaraan bermotor di Samsat Aceh Utara yang telah melakukan pembayaran. Ketidakpatuhan didefinisikan sebagai pembayaran yang dilakukan setelah masa berlaku pajak berakhir, sehingga penelitian ini merefleksikan ketepatan waktu pembayaran, bukan ketidakpatuhan total.
2. Penelitian membandingkan dua metode analisis, yaitu Regresi Logistik dan *Random Forest*.
3. Analisis difokuskan pada kinerja model prediksi yang meliputi akurasi, precision, recall, dan AUC, serta interpretasi faktor-faktor yang memengaruhi kepatuhan pajak dan perbandingan hasil kedua metode.
4. Variabel independen yang digunakan berasal dari arsip Samsat Aceh Utara, meliputi jenis kendaraan, jenis kepemilikan, umur kendaraan, jenis bahan bakar, dan merek kendaraan. Variabel dependen ditentukan berdasarkan status kepatuhan pembayaran pajak, yang diperoleh dari perbandingan antara tanggal pembayaran dan masa berlaku akhir pajak.
5. Data yang digunakan berupa transaksi pembayaran pajak kendaraan bermotor di Samsat Aceh Utara periode 2022–2024 dan hanya mencakup wajib pajak yang telah melakukan pembayaran, sehingga hasil prediksi lebih mencerminkan ketepatan waktu pembayaran pajak