
IMPLEMENTASI METODE *RANDOM FOREST* DALAM PREDIKSI PENYEBAB TUNGGAKAN PEMBAYARAN PAJAK KENDARAAN BERMOTOR DI SAMSAT MEDAN UTARA

Muhammad Ridwan Hasibuan^{1*}, Sayuti Rahman², Calvin Chiuloto³

1,2,3) Prodi Teknik Informatika, Fakultas Teknik dan Komputer, Universitas Harapan Medan, Indonesia

*Corresponding Email: ridwanhsb@gmail.com

Abstrak

Penelitian ini bertujuan untuk menerapkan metode Random Forest dalam melakukan prediksi penyebab tunggakan pembayaran pajak kendaraan bermotor di Samsat Medan Utara. Metode Random Forest digunakan untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap tunggakan pembayaran pajak kendaraan bermotor, sehingga dapat membantu dalam pengambilan keputusan yang lebih efektif dalam menangani tunggakan tersebut. Data yang digunakan dalam penelitian ini merupakan data historis tentang pembayaran pajak kendaraan bermotor dan faktor-faktor terkait dari Samsat Medan Utara. Proses analisis dimulai dengan tahap pra-pemrosesan data, termasuk pemilihan fitur yang relevan dan penanganan missing data. Selanjutnya, model Random Forest dikembangkan dan dilatih menggunakan data yang ada. Hasil prediksi kemudian dievaluasi menggunakan metrik evaluasi yang tepat. Dengan menerapkan metode Random Forest, penelitian ini dapat memberikan informasi yang berharga tentang faktor-faktor yang berkontribusi terhadap tunggakan pembayaran pajak kendaraan bermotor di Samsat Medan Utara. Hasil prediksi yang akurat dapat membantu pihak berwenang untuk mengidentifikasi faktor-faktor risiko yang berpotensi menyebabkan tunggakan pembayaran pajak dan mengambil tindakan yang sesuai untuk mengurangi jumlah tunggakan tersebut. Penelitian ini diharapkan dapat memberikan kontribusi positif dalam pengelolaan pembayaran pajak kendaraan bermotor di Samsat Medan Utara dan dapat menjadi acuan dalam pengambilan keputusan yang lebih efektif dalam menangani masalah tunggakan pajak kendaraan bermotor.

Kata Kunci: Prediksi, Tunggakan, Pajak Kendaraan, Rando Forest.

Abstract

This study aims to apply the Random Forest method to predict the causes of motor vehicle tax payment delays at Samsat Medan Utara. The Random Forest method is used to identify factors contributing to motor vehicle tax payment delays, thus assisting in more effective decision-making in addressing these delays. The data used in this study consists of historical data on motor vehicle tax payments and related factors from Samsat Medan Utara. The analysis process begins with data pre-processing, including the selection of relevant features and handling missing data. Subsequently, the Random Forest model is developed and trained using the available data. The prediction results are then evaluated using appropriate evaluation metrics. By applying the Random Forest method, this study provides valuable information about the factors contributing to motor vehicle tax payment delays at Samsat Medan Utara. Accurate prediction results can help authorities identify risk factors that may lead to tax payment delays and take appropriate actions to reduce the number of delays. This research is expected to make a positive contribution to the management of motor vehicle tax payments at Samsat Medan Utara and serve as a reference for more effective decision-making in addressing motor vehicle tax payment issues.

Keywords: Prediction, Arrears, Vehicle Tax, Random Forest.



PENDAHULUAN

Pajak merupakan salah satu sumber penerimaan daerah yang utama, sehingga perlu adanya target yang jelas untuk mencapainya (Sururama et al., 2023). Pajak juga memiliki kontribusi wajib yang harus dipenuhi oleh individu atau badan usaha sesuai dengan ketentuan undang-undang, bersifat memaksa, dan digunakan untuk keperluan negara demi kemakmuran rakyat secara luas (UU No. 28 Tahun 2007 Tentang KUP). Salah satu jenis pajak daerah yang memberikan kontribusi besar terhadap Pendapatan Asli Daerah (PAD) adalah Pajak Kendaraan Bermotor (PKB) (Damayanti et al., 2023). Dalam hal ini, Unit Pelayanan Teknis Daerah Pengelolaan Pendapatan Daerah (UPTD-PPD) Samsat Medan Utara berperan sebagai lembaga pelayanan pajak kendaraan bermotor, meliputi pembayaran pajak, balik nama kendaraan bermotor, hingga penagihan pajak.

Sistem administrasi pelayanan pajak di Samsat menggunakan konsep *one roof system* (Satuan Administrasi Manunggal Satu Atap), yang bertujuan mempercepat dan mempermudah layanan masyarakat (Waruwu, 2022). Meskipun demikian, masih banyak kendaraan bermotor yang menunggak pajak, sehingga berdampak pada penurunan penerimaan daerah (Ulya, 2022). Untuk mengatasi permasalahan ini, diperlukan analisis mendalam terhadap data yang tersedia. Teknik *data mining* menjadi salah satu solusi untuk menganalisis data dalam jumlah besar guna mendapatkan informasi yang relevan. Penelitian ini menggunakan algoritma *random forest*, salah satu metode dalam *decision tree*, untuk memprediksi tunggakan pembayaran pajak kendaraan bermotor. Algoritma ini dapat memproses data yang kompleks dan menghasilkan prediksi yang akurat.

Penelitian terdahulu menunjukkan potensi penerapan algoritma prediksi untuk mengatasi permasalahan pajak. (Mustakim et al., n.d.) menggunakan algoritma *Support Vector Regression* untuk memprediksi jumlah tunggakan pajak kendaraan bermotor, menghasilkan akurasi hingga 96% dengan nilai error yang sangat kecil. Hasil penelitian tersebut menunjukkan variasi tunggakan pajak antar kecamatan. Di sisi lain, (Tenggono et al., 2021) menggunakan metode *Naïve Bayes* untuk memprediksi kepatuhan wajib pajak dengan akurasi sebesar 53,48%, menunjukkan bahwa tingkat kesadaran masyarakat terhadap pajak bumi dan bangunan masih rendah.

Berdasarkan hasil penelitian sebelumnya, faktor-faktor seperti kurangnya pemahaman wajib pajak terhadap peraturan, rendahnya kesadaran, kendaraan yang sudah tidak layak pakai, serta kendala keuangan menjadi penyebab utama tunggakan pajak (Karlina & Ethika, 2020). Oleh karena itu, penelitian ini mengusulkan penerapan metode random forest untuk memprediksi penyebab tunggakan pajak kendaraan bermotor di Samsat Medan Utara.

KAJIAN TEORI

Pajak Kendaraan Bermotor (PKB)

Pajak Kendaraan Bermotor (PKB) merupakan salah satu jenis pajak daerah yang dikenakan atas kepemilikan dan/atau penguasaan kendaraan bermotor, baik kendaraan roda dua maupun roda empat, yang beroperasi di jalan. PKB termasuk dalam kategori pajak objektif, di mana penetapan besarnya didasarkan pada objek pajak, yaitu kendaraan bermotor, tanpa memperhatikan kondisi subjektif pemiliknya (Bintary, 2020).

Dasar hukum pengenaan PKB diatur dalam Undang-Undang No. 28 Tahun 2009 tentang Pajak Daerah dan Retribusi Daerah. Adapun ketentuan pokok PKB meliputi (Faharudin & Wulandari, 2024):

1. Subjek Pajak: Orang pribadi atau badan yang memiliki atau menguasai kendaraan bermotor.
2. Objek Pajak: Semua kendaraan bermotor, baik yang beroda dua, tiga, empat, atau lebih, termasuk alat berat dan alat besar yang menggunakan jalan umum.
3. Tarif Pajak: Berkisar antara 1% hingga 2% dari nilai jual kendaraan bermotor (NJKB), tergantung pada kebijakan pemerintah daerah.
4. Pemanfaatan Pajak: Hasil penerimaan PKB digunakan untuk pembiayaan pembangunan daerah, termasuk peningkatan infrastruktur transportasi.

Random Forest

Metode *Random Forest* adalah salah satu algoritma *ensemble learning* yang digunakan untuk tugas klasifikasi dan regresi (Nugroho & Harini, 2024). Algoritma ini bekerja dengan membangun sejumlah besar *decision trees* selama proses pelatihan dan menggabungkan hasilnya untuk menghasilkan prediksi yang lebih akurat dan stabil (Baskoro et al., 2021). *Random Forest* pertama kali diperkenalkan oleh Leo Breiman pada tahun 2001

sebagai pengembangan dari metode *bagging* dan *decision trees* (Harahap et al., n.d.).

Karakteristik utama dari metode *random forest* adalah (Yohannes et al., 2019):

1. *Ensemble Learning: Random Forest* menggabungkan banyak *decision trees* untuk meningkatkan performa prediksi dibandingkan dengan menggunakan satu pohon keputusan.
2. Pengacakan (*Randomization*): Algoritma ini memperkenalkan elemen pengacakan dalam:
 - Pemilihan subset data selama proses pelatihan (*bagging*).
 - Pemilihan subset fitur untuk setiap node pada pohon keputusan.
3. *Out-of-Bag (OOB) Error: Random Forest* memiliki kemampuan untuk mengestimasi akurasi model menggunakan data yang tidak digunakan selama pelatihan (*out-of-bag samples*).
4. *Robust terhadap Overfitting*: Karena menggunakan banyak pohon, algoritma ini lebih tahan terhadap overfitting dibandingkan dengan *decision tree* tunggal.

Dimana langkah kerja dari metode *random forest* yaitu (Khalim et al., 2023):

1. Pembuatan Subset Data (*Bootstrap Sampling*)
Dari dataset asli, *Random Forest* membuat beberapa subset data dengan pengambilan sampel secara acak dengan pengembalian (*with replacement*). Data yang tidak terpilih disebut sebagai *out-of-bag samples*.
2. Pembuatan Pohon Keputusan
Untuk setiap subset data, algoritma membangun pohon keputusan. Pada setiap node, hanya subset fitur yang dipilih secara acak yang dipertimbangkan untuk pemisahan.
3. Penggabungan Hasil
 - Untuk tugas klasifikasi, hasil akhir ditentukan berdasarkan mayoritas suara (*majority voting*).
 - Untuk tugas regresi, rata-rata nilai prediksi dari semua pohon digunakan sebagai hasil akhir.

Untuk kinerja algoritma ini biasanya dievaluasi dengan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score* untuk klasifikasi, atau *mean squared error (MSE)* dan *R*-

squared untuk regresi. Dengan keunggulan utamanya pada kestabilan prediksi dan akurasi yang tinggi, *Random Forest* sering digunakan sebagai algoritma pilihan dalam berbagai proyek berbasis data (Putri & Arianto, 2024).

Penelitian Terdahulu

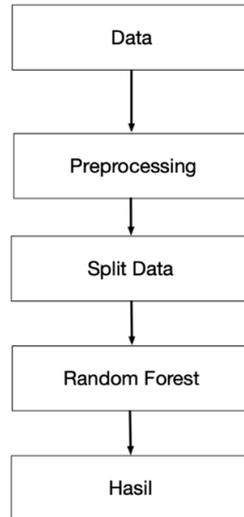
Berikut adalah penelitian terdahulu yang menggunakan metode *random forest* disajikan pada tabel berikut:

Tabel 1. Penelitian Terdahulu

No	Peneliti	Penelitian Yang Dibahas
1.	(Speiser et al., 2019)	Penelitian " <i>A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling</i> " membahas dan membandingkan berbagai metode seleksi variabel yang digunakan dalam algoritma <i>Random Forest</i> untuk meningkatkan performa prediksi klasifikasi. Fokus penelitian ini adalah mengidentifikasi pendekatan terbaik dalam memilih fitur yang relevan untuk model prediksi klasifikasi, dengan mempertimbangkan efisiensi dan akurasi.
2.	(Wang et al., 2020)	Penelitian " <i>An Improved Random Forest-Based Rule Extraction Method for Breast Cancer Diagnosis</i> " mengusulkan metode baru untuk meningkatkan kemampuan algoritma <i>Random Forest</i> dalam mendiagnosis kanker payudara. Fokus utama penelitian ini adalah mengembangkan teknik ekstraksi aturan (<i>rule extraction</i>) dari model <i>Random Forest</i> , sehingga hasil prediksi lebih interpretatif dan transparan, terutama untuk mendukung pengambilan keputusan medis.
3.	(Wu & Chang, 2024)	Penelitian " <i>Ransomware Detection on Linux Using Machine Learning with Random Forest Algorithm</i> " berfokus pada pengembangan sistem deteksi ransomware berbasis <i>machine learning</i> yang dirancang khusus untuk lingkungan Linux. Ransomware merupakan ancaman keamanan siber yang mengenkripsi data pengguna dan meminta tebusan untuk pemulihan. Dengan meningkatnya serangan ransomware di sistem Linux, diperlukan solusi yang efektif untuk mendeteksinya.

METODE PENELITIAN

Pada penelitian ini, penulis menggunakan multilayer preceptron, untuk membangun jaringan *Feedforward Neural Network* dengan algoritma *multilayer preceptron* memerlukan beberapa langkah. Adapun langkah penelitian dapat dilihat pada gambar 1. berikut.



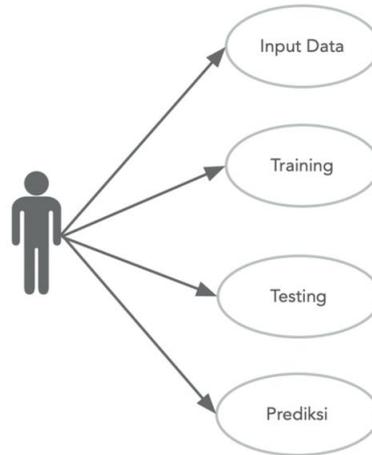
Gambar 1. Langkah Penelitian

Data yang digunakan pada penelitian ini adalah data tunggakan pembayaran pajak yang diperoleh dari samsat medan utara dengan jumlah 1000. Sebelum dilakukan proses training, setiap data akan dinormalisasikan terlebih dahulu dengan interval antara 0 dan 1. Intuisi algoritma *random forrest* dapat dibagi menjadi dua tahap. Pada tahap pertama, kami secara acak memilih fitur “k” dari total m fitur dan membangun *random forrest*. Pada tahapapan yang dilakukan adalah sebagai berikut:

- Pilih secara acak k fitur dari total m fitur di mana $k < m$.
- Di antara k fitur, hitung simpul d menggunakan titik pisah terbaik.
- Pisahkan node menjadi node anak menggunakan split terbaik.
- Ulangi 1 sampai 3 langkah sampai l jumlah node telah tercapai.

Untuk proses bisnis system menggunakan model perancangan sistem *Unified Modelling Language* (UML). Dimana tools yang digunakan dalam UML adalah *usecase diagrams*, *activity diagrams*, dan *squence diagrams*.

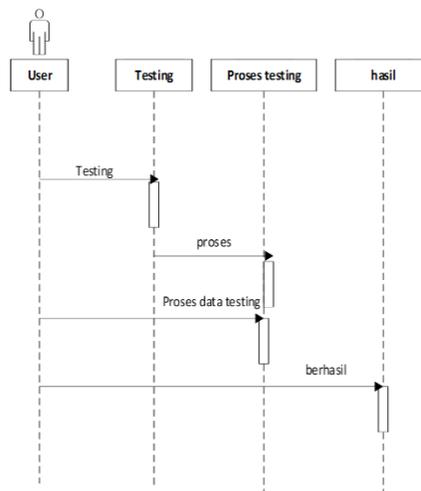
Use case diagram menggambarkan fungsionalitas yang di harapkan dari sebuah pembuatan model dalam melakukan prediksi penyebab tunggakan pembayaran pajak kendaraan bermotor yang mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan proses yang akan di buat, berikut ini perancangan *use case* sistem.



Gambar 2. *Usecase Diagrams System*

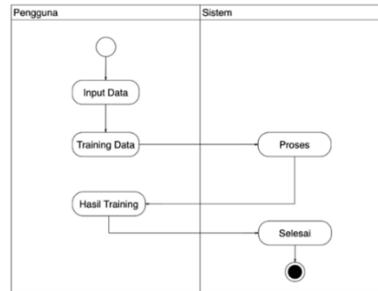
Berdasarkan gambar 2. dijelaskan bahwa dalam user terdapat aktivitas seperti input data, tranning, testing, prediksi yang berfungsi sebagai proses dalam melakukan prediksi penyebab tunggakan pembayaran pajak kendaraan bermotor menggunakan metode *Random Forrest*.

Berikut ini adalah *Sequence Diagram testing* yang menggambarkan prosedur informasi training datainput yang dilakukan dapat dilihat pada gambar 3. berikut.



Gambar 3. *Sequence Diagrams Testing*

Untuk proses *activity diagrams training* dapat lihat pada gambar berikut.



Gambar 4. Proses *Activity Diagrams Training*

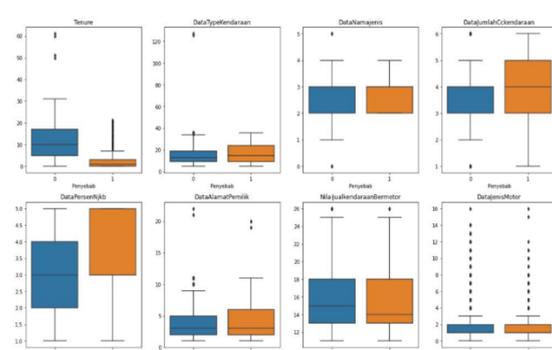
Berdasarkan gambar 4. terdapat aktivitas input data dan training yang mana proses nya akan dilakukan oleh pengguna yang harus memasukan data terlebih dahulu kemudian melakukan proses training sebanyak 70% dari data keseluruhan.

HASIL DAN PEMBAHASAN

Hasil

Penelitian ini bertujuan untuk memprediksi penyebab tunggakan pembayaran pajak kendaraan bermotor menggunakan metode *Random Forest*. Data yang digunakan terdiri dari 6000 data historis tunggakan pajak pada tahun 2021. Dataset ini telah melalui proses *preprocessing*, seperti pembagian data menjadi *training* dan *testing*, normalisasi data, serta penghapusan *outlier*. Implementasi dilakukan menggunakan *Python* dengan dukungan *library* seperti *Pandas* melalui platform *Jupyter Notebook* yang terdapat dalam *Anaconda*.

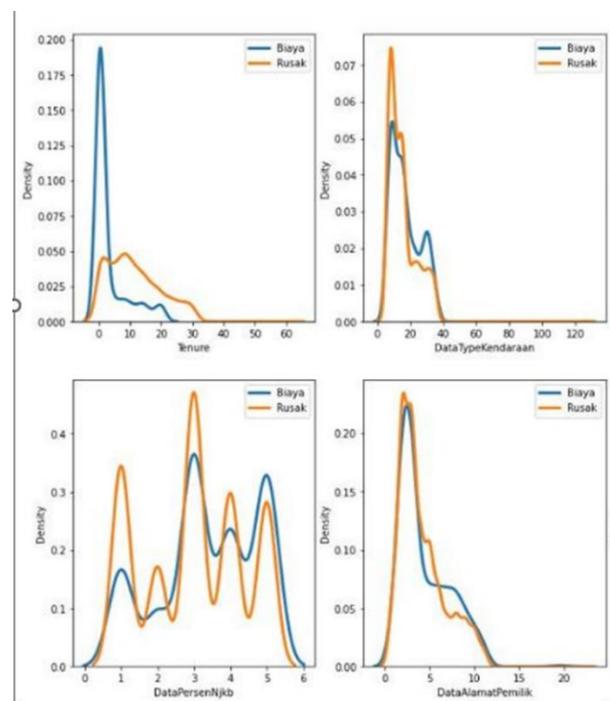
Selanjutnya dilakukan visualisasi data menggunakan Visualisasi OHLC untuk mempermudah menarik kesimpulan dengna melihat data tren atau pola yang ada. Berikut visualisasi OHLC disajikan pada gambar 5. berikut.



Gambar 5. Visualisasi Data Menggunakan OHLC

Berdasarkan pada gambar 5. merupakan visualisasi pada data data yang terdapat pada dataset seperti data type kendaraan, nama jenis, jumlah cc kendaraan dan lainnya, data tersebut menunjukkan jumlah data pada masing masing variabel yang digambarkan dalam bentuk visualisasi.

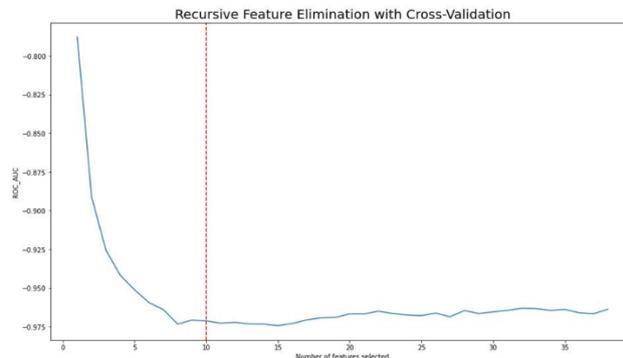
Representasi grafis dari penyebab tunggakan akan menggunakan *density numeric* yang akan menampilkan visualisasi data dari tunggakan pembayaran pajak kendaraan bermotor yang akan ditampilkan pada gambar 6. berikut ini.



Gambar 6. Visualisasi *Density Numeric*

Berdasarkan pada gambar 6. merupakan visualisasi antara variabel dengan label penyebab tunggakan pajak pada kendaraan bermotor. Pada gambar 6. masing masing variabel memiliki pola dalam menggambarkan penyebab tunggakan pajak kendaraan bermotor. Penggunaan *density plot* untuk melihat distribusi data yang terdapat pada data tunggakan pajak kendaraan bermotor.

Pada visualisasi *cross validation* menampilkan diagram yang mana terdapat informasi hasil dari algoritma *random forest* dalam melakukan prediksi penyebab tunggakan pembayaran pajak kendaraan bermotor. Berikut ini hasil *recursive elimination* dengan *cross validasi* yang akan ditampilkan pada gambar 7. berikut.



Gambar 7. Visualisasi Cross Validation

Setelah melalui beberapa tahapan, maka selanjutnya dilakukan perhitungan menggunakan *random forrest*. Pada bagian ini akan dihitung proses dari evaluasi data testing dengan menggunakan nilai precision, recall, F1 score, yang menghasilkan akurasi prediksi sebesar 98%.

```
<class 'sklearn.ensemble._forest.RandomForestClassifier'>
CV score : 0.9681448957734855
      precision    recall  f1-score   support

     0       0.98     0.99     0.99     941
     1       0.97     0.88     0.92     185

 accuracy          0.98          1126
 macro avg         0.97     0.94     0.95     1126
 weighted avg      0.98     0.98     0.97     1126
```

Gambar 8. Hasil Dari *Random Forrest*

Berdasarkan hasil dari evaluasi testing yang menghasilkan akurasi 98% dapat disimpulkan bahwa model dalam melakukan prediksi sudah baik namun perlu dilakukan validasi model dengan menggunakan data baru.

Tahap akhir dilakukan pengujian model, dari hasil pengujian disajikan pada tabel 2. berikut.

Tabel 2. Hasil Pengujian Model

TYPE_KB	Tanggal	NJKB	PKB	Label	Prediksi
NF125 TD	01/05/2025	10.300.000	1,750	rusak	Rusak
NF100 L (SUPRA FIT)	01/05/2024	6.600.000	1,750	biaya	Rusak
NF11T11CO1 M/T (110 CC)	01/05/2022	8.500.000	1,750	rusak	Biaya
RX-K 135CC	01/05/2022	9.800.000	1,750	rusak	Rusak

NF125 TR	01/05/2026	11.500.000	1,750	rusak	Rusak
H1B02N42L0 AT	01/05/2026	10.600.000	1,750	biaya	Rusak
P 150 X	01/05/2026	1.800.000	1,750	biaya	Biaya
AFX12U21C08 M/T (125 CC)	01/05/2024	12.900.000	1,750	biaya	Biaya
NF125 TD M/T	01/05/2022	12.400.000	1,750	biaya	Biaya
X1HO2N35M1 A/T	01/05/2024	14.200.000	1,750	biaya	Biaya
28D (AL115S/MIO)	01/05/2025	8.000.000	1,750	biaya	Biaya
A5C02R37M2 M/T	01/05/2026	20.600.000	1,750	biaya	Biaya
H1B02N41LO AT	01/05/2026	10.100.000	1,750	biaya	Biaya

Berdasarkan tabel 2, ada 10 data yang dilakukan pengujian yang menghasilkan 7 data berhasil terprediksi benar dan 3 data terprediksi salah.

Pembahasan

Berdasarkan hasil evaluasi dan pengujian, metode *Random Forest* menunjukkan performa yang baik dalam memprediksi penyebab tunggakan pajak kendaraan bermotor. Akurasi sebesar 98% menunjukkan potensi penggunaan model ini dalam mendukung pengambilan keputusan di Samsat Medan Utara. Visualisasi data dan analisis korelasi memberikan wawasan mengenai pola dan distribusi variabel yang relevan. Namun, hasil pengujian model dengan data baru menunjukkan adanya kelemahan dalam memprediksi kategori tertentu. Kesalahan prediksi ini dapat disebabkan oleh:

1. Ketidakseimbangan data pada kategori tertentu.
2. Keterbatasan jumlah data *testing* yang digunakan.
3. Variabel prediktor yang mungkin kurang mencakup semua faktor penyebab tunggakan.

Untuk meningkatkan performa model, beberapa langkah dapat dilakukan, seperti penambahan variabel baru, optimasi parameter *Random Forest*, atau penerapan metode validasi yang lebih komprehensif. Dengan demikian, penelitian ini memberikan dasar yang



kuat untuk pengembangan lebih lanjut dalam memprediksi dan mengelola tunggakan pembayaran pajak kendaraan bermotor.

SIMPULAN

Berdasarkan penelitian yang dilakukan, metode *Random Forest* diterapkan untuk memprediksi penyebab tunggakan pembayaran pajak kendaraan bermotor di Samsat Medan Utara, dan terbukti efektif dalam mengidentifikasi faktor-faktor yang berkontribusi terhadap tunggakan tersebut. Dengan memanfaatkan data historis tentang pembayaran pajak kendaraan bermotor beserta faktor-faktor terkait, penelitian ini berhasil mengembangkan model prediksi *Random Forest* yang menunjukkan tingkat akurasi yang memadai. Hasil prediksi ini memberikan informasi berharga bagi pihak berwenang dalam menangani masalah tunggakan pajak. Selain itu, melalui penerapan metode ini, faktor-faktor risiko yang berpotensi menyebabkan tunggakan dapat diidentifikasi secara lebih efektif, sehingga memungkinkan pihak berwenang untuk mengambil langkah strategis guna mengurangi jumlah tunggakan pajak kendaraan bermotor di Samsat Medan Utara.

DAFTAR PUSTAKA

- Baskoro, B. B., Susanto, I., & Khomsah, S. (2021). Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR). *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, 3(2), 21–29.
- Bintary, A. A. (2020). Analisis Kepatuhan Wajib Pajak Kendaraan Bermotor dalam upaya Meningkatkan Penerimaan Pajak Daerah pada Kantor Bersama Samsat Jakarta Timur Tahun 2015-2018. *Jurnal Pajak Vokasi (JUPASI)*, 1(2), 86–101.
- Damayanti, A. Y., Afifah, A. N., & Sunaningsih, S. N. (2023). Analisis Kontribusi Pemungutan Pajak Kendaraan Bermotor (PKB) terhadap Peningkatan Pendapatan Asli Daerah (PAD) Di Kota Magelang Tahun 2018-2021. *Jurnal Maneksi (Management Ekonomi Dan Akuntansi)*, 12(2), 241–245.
- Faharudin, F., & Wulandari, S. (2024). TINJAUAN YURIDIS PEMUNGUTAN PAJAK HOTEL DAN HIBURAN BERDASARKAN UNDANG-UNDANG NOMOR 28



- TAHUN 2009 TENTANG PAJAK DAERAH DAN RETRIBUSI DAERAH. *Jurnal Ilmu Hukum Kanturuna Wolio*, 32–40.
- Harahap, B., Umbara, R. F., & Danang Triantoro, M. T. (n.d.). *Prediksi Indeks Harga Saham Menggunakan Metode Gabungan Support Vector Regression dan Random Forest Prediction Of Stock Market Index Using Hybrid of Support Vector Regression and Random Forest*.
- Karlina, U. W., & Ethika, M. H. (2020). Pengaruh pengetahuan wajib pajak, kesadaran wajib pajak, dan sanksi perpajakan terhadap kepatuhan wajib pajak kendaraan bermotor. *Jurnal Kajian Akuntansi Dan Auditing*, 15(2), 143–154.
- Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 498–504.
- Mustakim, M., Pratama, Y. R., & others. (n.d.). Prediksi Jumlah Tunggakan Pajak Jatuh Tempo Menggunakan Algoritma Support Vector Regression. *Seminar Nasional Teknologi Informasi Komunikasi Dan Industri*, 115–122.
- Nugroho, A., & Harini, D. (2024). Teknik Random Forest untuk Meningkatkan Akurasi Data Tidak Seimbang. *JSITIK: Jurnal Sistem Informasi Dan Teknologi Informasi Komputer*, 2(2), 128–140.
- Putri, F., & Arianto, D. B. (2024). PERBANDINGAN PERFORMA RANDOM FOREST DAN GRADIENT BOOSTING DALAM PREDIKSI PADA DATASET CUSTOMER SHOPPING TRENDS. *Kohesi: Jurnal Sains Dan Teknologi*, 5(11), 1–10.
- Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134, 93–101.
- Sururama, R., Weku, J., & Syahbana, A. (2023). Pengelolaan Pajak Restoran Dalam Meningkatkan Pendapatan Asli Daerah (PAD) di Kota Medan. *JEKP (Jurnal Ekonomi Dan Keuangan Publik)*, 10(1), 1–20.
- Tenggono, A., Adryansyah, M., & others. (2021). Prediksi Potensi Kepatuhan Wajib Pajak PBB-P2 Menggunakan Metode Naïve Bayes di Kecamatan Seberang Ulu I Kota Palembang. *Teknomatika*, 11(01), 51–58.



- Ulya, H. (2022). *Efektivitas Program Pemutihan Pajak Kendaraan Bermotor dalam Meningkatkan Kepatuhan Wajib Pajak Membayar Pajak dalam Perspektif Masalah Mursalah (Studi Kasus Pada Kantor Bersama SAMSAT Natal)*. Universitas Islam Negeri Sumatera Utara Medan.
- Wang, S., Wang, Y., Wang, D., Yin, Y., Wang, Y., & Jin, Y. (2020). An improved random forest-based rule extraction method for breast cancer diagnosis. *Applied Soft Computing*, 86, 105941.
- Waruwu, H. (2022). *Kualitas Pelayanan Pajak Kendaraan Bermotor di Kantor Sistem Administrasi Manunggal Satu Atap (Samsat) Medan Selatan Kota Medan*. Universitas Medan Area.
- Wu, Y., & Chang, Y. (2024). Ransomware detection on linux using machine learning with random forest algorithm. *Authorea Preprints*.
- Yohannes, Y., Devella, S., & Pandrean, A. H. (2019). Penerapan Speeded-Up Robust Feature pada Random Forest Untuk Klasifikasi Motif Songket Palembang. *Jurnal Teknik Informatika Dan Sistem Informasi*, 5(3).