

# Prediksi Keberhasilan Menindaklanjuti Pelanggan pada Dealer Mobil dengan Komparasi Algoritma *Random Forest* dan *XGBoost*

DOI: <http://dx.doi.org/10.35889/progresif.v21i2.2776>

Creative Commons License 4.0 (CC BY – NC) 

**Helma Nopijani Heidy<sup>1\*</sup>, Dimas Eko Putro<sup>2</sup>, Muhammad Fadli<sup>3</sup>,  
 Erliyan Redy Susanto<sup>4</sup>**

<sup>1,2,3,4</sup>Magister Ilmu Komputer, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia

<sup>3</sup>Ekonomi dan Bisnis, Politeknik Negeri Lampung, Bandar Lampung, Indonesia

\*e-mail *Corresponding Author*: [helmanopijaniheidy@teknokrat.ac.id](mailto:helmanopijaniheidy@teknokrat.ac.id)

## Abstract

*The automotive industry is facing intense competition in boosting vehicle sales, where the follow-up process with prospective customers plays a crucial role in sales conversion. This study develops a predictive model for the success of follow-ups at car dealerships by comparing two machine learning algorithms: Random forest and XGBoost. A dataset of Honda car dealership customers from 2023 was processed through a preprocessing stage, including handling data imbalance and encoding categorical data. The models were evaluated using accuracy, precision, recall, and F1-score metrics. The results show that XGBoost outperforms with an accuracy of 91.67%, compared to Random forest's 88.89%. Both models demonstrate balanced performance across positive and negative classes, indicating a significant improvement over previous approaches. This study recommends expanding the dataset and developing a prediction-based decision support system to enhance the marketing effectiveness of car dealerships.*

**Keywords:** *Machine learning; Random forest; XGBoost*

## Abstrak

Industri otomotif menghadapi persaingan ketat dalam meningkatkan penjualan kendaraan, di mana proses tindak lanjut (*Follow-up*) kepada calon pelanggan menjadi faktor krusial dalam konversi penjualan. Penelitian ini mengembangkan model prediksi keberhasilan *Follow-up* pada dealer mobil dengan membandingkan dua algoritma *machine learning*, yaitu *Random forest* dan *XGBoost*. Dataset pelanggan dealer mobil Honda tahun 2023 diproses melalui tahap *preprocessing*, termasuk penanganan ketidakseimbangan data menggunakan encoding data kategorikal. Model dievaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan *XGBoost* unggul dengan akurasi 91,67%, lebih baik dibanding *Random forest* dengan akurasi 88,89%. Kedua model menunjukkan performa yang seimbang pada kelas positif dan negatif, menandai peningkatan signifikan dari pendekatan sebelumnya. Penelitian merekomendasikan perluasan dataset dan pengembangan sistem pendukung keputusan berbasis prediksi untuk meningkatkan efektivitas pemasaran dealer mobil.

**Kata kunci:** *Machine learning; Random forest; XGBoost*

## 1. Pendahuluan

Industri otomotif merupakan salah satu sektor yang sangat kompetitif, di mana efektivitas strategi pemasaran menjadi faktor penentu keberhasilan penjualan. Salah satu aspek penting dari strategi ini adalah proses tindak lanjut (*follow-up*) terhadap calon pelanggan, yang bertujuan untuk mengubah prospek menjadi konsumen aktual. Mengingat keputusan pembelian mobil sering kali melibatkan pertimbangan yang kompleks dan waktu yang panjang, pendekatan *follow-up* yang tepat menjadi sangat krusial dalam meningkatkan konversi penjualan [1].

Pada praktiknya, banyak dealer mobil masih menerapkan proses *follow-up* secara manual dan berdasarkan intuisi individu, tanpa menggunakan pendekatan berbasis data. Hal ini

menyebabkan tidak efisiennya alokasi sumber daya tenaga pemasaran, serta sulitnya mengukur faktor-faktor yang memengaruhi keberhasilan tindak lanjut. Dari observasi data internal, tercatat bahwa tingkat keberhasilan *follow-up* pelanggan pada beberapa tim penjualan masih berada di bawah 60%, menandakan perlunya upaya sistematis untuk meningkatkan akurasi dalam menentukan prospek yang potensial.

Seiring berkembangnya teknologi kecerdasan buatan (AI), khususnya *machine learning*, pendekatan prediktif berbasis data menjadi solusi yang potensial dalam mengoptimalkan strategi *follow-up*. Algoritma seperti *Random forest* dan *XGBoost* telah terbukti mampu melakukan klasifikasi dan prediksi yang andal pada berbagai domain, termasuk perilaku konsumen [2]. Keduanya dapat mengolah data pelanggan dalam jumlah besar, mengenali pola-pola tersembunyi, serta memberikan rekomendasi prospek yang layak ditindaklanjuti. Model prediktif ini memungkinkan dealer untuk fokus pada pelanggan dengan kemungkinan konversi tinggi, sehingga efisiensi kerja meningkat dan hasil penjualan dapat ditingkatkan secara signifikan.

Beberapa studi sebelumnya telah membandingkan performa algoritma *Random forest* dan *XGBoost* dalam konteks customer retention dan purchase behavior. *XGBoost* unggul dalam metrik akurasi, *precision*, *recall*, *F1-score*, dan ROC AUC pada dataset multisektoral, terutama setelah menerapkan teknik penyeimbangan data seperti SMOTE [3]. Selanjutnya, studi lain menunjukkan bahwa *XGBoost* mencapai akurasi hingga 94.5% dalam prediksi perilaku pembelian, melebihi model lain seperti *Logistic Regression* dan SVM [4].

Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Random forest* dan *XGBoost* dalam membangun model prediksi keberhasilan *follow-up* pelanggan di sebuah dealer mobil. Evaluasi model dilakukan berdasarkan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam bentuk solusi berbasis data untuk membantu tim pemasaran dalam menentukan strategi tindak lanjut yang lebih efektif, serta memberikan sumbangan akademik dalam bidang penerapan *machine learning* di industri otomotif.

## 2. Tinjauan Pustaka

Penelitian terkait prediksi keberhasilan pekerjaan telah banyak dilakukan dengan pendekatan *machine learning* pada berbagai bidang. Prediksi churn pelanggan menggunakan algoritma *Random forest* dan *XGBoost* pada data pelanggan layanan telekomunikasi [5]. Mereka menggunakan parameter histori pembelian, frekuensi kontak, dan lama berlangganan. Hasil penelitian menunjukkan bahwa *XGBoost* memiliki keunggulan dalam akurasi dan kestabilan prediksi dibandingkan algoritma lainnya.

*Machine learning* diterapkan untuk mengidentifikasi pelanggan potensial pada dealer mobil dengan algoritma *Decision Tree* dan *Random forest*. Penelitian ini memproses variabel seperti usia kendaraan, jenis kelamin pelanggan, status *follow-up* sebelumnya, dan lokasi domisili. Hasilnya menunjukkan bahwa *Random forest* mampu memberikan hasil klasifikasi lebih baik dibanding metode dasar lainnya [1].

Algoritma *XGBoost* dan Naive Bayes dibandingkan dalam prediksi respon pelanggan terhadap promosi *e-commerce*. Mereka mengolah parameter seperti waktu kunjungan, frekuensi pembelian, serta segmentasi usia. Dalam penelitian ini, *XGBoost* unggul dari sisi presisi dan *recall* dalam memprediksi keberhasilan promosi [6].

Beragam studi sebelumnya telah membuktikan efektivitas algoritma *Random forest* dan *XGBoost* dalam prediksi sales dan perilaku konsumen. Secara langsung membandingkan SVM, *Random forest*, dan *XGBoost* dalam prediksi penjualan dan menemukan performa kuat dari kedua algoritma pohon keputusan [7]. Begitu pula dalam konteks retail besar, hasil *XGBoost* lebih unggul dibanding metode klasik dalam memprediksi volume penjualan Walmart [8]. Lebih lanjut, penelitian tentang churn pelanggan menunjukkan bahwa penggunaan teknik *oversampling* (SMOTE, ADASYN) dapat meningkatkan akurasi pada dataset tidak seimbang dengan model *Random forest* maupun *XGBoost* [9]. Dalam domain perilaku pembelian konsumen, *XGBoost* menunjukkan keunggulan signifikan dibanding KNN, SVM, dan *Random forest* dari sisi stabilitas dan metrik F1, *recall*, *precision* [10]. Studi lokal di Indonesia juga telah menunjukkan *capability Random forest* untuk prediksi harga mobil bekas, memperkuat relevansi aplikasi ML di industri *otomotif domestic* [11].

Penggunaan *XGBoost* dalam sistem pendukung keputusan untuk prediksi konversi penawaran pinjaman bank. Variabel yang digunakan meliputi histori transaksi, status pekerjaan, dan wilayah geografis. Mereka menyimpulkan bahwa *XGBoost* sangat cocok digunakan pada data yang kompleks dan tidak seimbang, karena fitur regularisasi yang dimilikinya [12].

Pembelajaran mesin untuk memprediksi tingkat keberhasilan pasien dalam terapi pernapasan menggunakan dataset medis. Parameter yang digunakan antara lain usia, jenis kelamin, riwayat medis, serta frekuensi terapi. Dalam penelitiannya, *XGBoost* memberikan

performa terbaik dibandingkan *Logistic Regression* dan *Random forest* dalam hal akurasi dan *F1-score* [2].

Berdasarkan studi-studi tersebut, dapat disimpulkan bahwa pendekatan prediktif berbasis *machine learning*, khususnya *Random forest* dan *XGBoost*, telah terbukti efektif dalam memprediksi keberhasilan suatu tindakan berdasarkan data historis pelanggan. Namun, sebagian besar penelitian sebelumnya fokus pada domain selain otomotif dan tidak secara spesifik mengkaji keberhasilan tindak lanjut (*follow-up*) pelanggan sebagai objek prediksi. Penelitian ini menawarkan novelty dengan menerapkan kedua algoritma tersebut dalam konteks dealer mobil, memproses data real pelanggan yang mencakup status *follow-up*, data kontak, lokasi, hingga data kendaraan. Selain itu, penelitian ini mengevaluasi kinerja model secara menyeluruh menggunakan empat metrik evaluasi utama (akurasi, *precision*, *recall*, dan *F1-score*) untuk menghasilkan model yang akurat dan praktis sebagai dukungan pengambilan keputusan bagi tim pemasaran dealer.

### 3. Metodologi

#### 3.1 Jenis dan Sumber Data

Penelitian ini memanfaatkan dataset pelanggan dari dealer mobil Honda yang dikumpulkan melalui sistem *Customer Relationship Management (CRM)*. Dataset tersebut mencakup informasi demografis pelanggan, data kendaraan, serta catatan hasil *Follow-up* selama tahun 2023. Penggunaan data serupa telah terbukti efektif dalam penelitian terkait prediksi perilaku konsumen untuk meningkatkan akurasi model.

#### 3.2 Pengolahan Data (*Preprocessing*)

*Random forest* dan *XGBoost* dalam memprediksi kredit bermasalah dan menekankan pentingnya teknik *preprocessing* untuk meningkatkan akurasi prediksi [13].

Tahapan *preprocessing* meliputi beberapa langkah penting:

- 1) Pembersihan data: Proses menghilangkan data yang duplikat serta menangani nilai yang hilang (*missing values*) dengan teknik imputasi median untuk data numerik dan modus untuk data kategorikal.
- 2) Transformasi data kategorikal: Variabel seperti model kendaraan dan alamat diubah menjadi format numerik menggunakan teknik *one-hot encoding*.

#### 3.3 Variabel dan Fitur

Variabel fitur yang digunakan mencakup nomor polisi kendaraan, model kendaraan, tahun pembuatan, alamat, nomor telepon pelanggan, serta status hasil *Follow-up* sebelumnya. Target klasifikasi berupa dua kelas, yaitu *Follow-up* berhasil (1) dan tidak berhasil (0), sesuai dengan skema klasifikasi biner.

#### 3.4 Pembagian Dataset

Dengan terbatasnya data uji yaitu 36 data uji dalam 1501 baris dan 20 kolom maka dataset dibagi menjadi data pelatihan (80%) dan data pengujian (20%) menggunakan *stratified sampling* untuk menjaga proporsi kelas target.

#### 3.5 Implementasi Model

##### 1) *Random forest*

Model *Random forest* dikembangkan menggunakan library *Scikit-learn*, dengan proses optimasi *hyperparameter* seperti jumlah pohon (*n\_estimators*), kedalaman maksimal pohon (*max\_depth*), dan jumlah fitur yang dipilih saat pemisahan (*max\_features*). *Random forest* bekerja dengan membangun banyak pohon keputusan dan menggabungkan prediksi dari setiap pohon untuk menghasilkan prediksi akhir yang lebih stabil dan akurat.

*Random forest* memiliki keunggulan dalam mengurangi overfitting dibanding *Decision Tree* tunggal karena sifat *ensemble*-nya. Dalam konteks pemasaran, algoritma ini terbukti lebih akurat dalam memprediksi tingkat kepuasan pelanggan.

##### 2) *XGBoost*

Model *XGBoost* diimplementasikan menggunakan library *XGBoost* dengan pendekatan *boosting* dan penyetelan parameter seperti *learning rate*, *max\_depth*, *subsample*, dan *n\_estimators*. *XGBoost* menggunakan teknik *gradient boosting*, di mana model dibangun secara bertahap untuk meminimalkan fungsi kerugian (*loss function*).

Secara matematis, prediksi model *XGBoost* pada iterasi ke-*t* adalah:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i), \quad f_t \in \mathcal{F} \tag{1}$$

dengan:

- $y_i^{(t)}$  adalah prediksi pada literasi ke- $t$ ,
- $f_t(x_i)$  adalah pohon regresi baru,
- $F$  adalah ruang fungsi pohon keputusan.

Fungsi objektif yang diminimalkan dalam *XGBoost* adalah:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{t=1}^T \Omega(f_t) \quad (2)$$

Dalam berbagai studi, kombinasi *Random forest* dan *XGBoost* telah menunjukkan peningkatan performa yang signifikan dalam prediksi klasifikasi *customer churn* ([14]) al ini menunjukkan potensi model untuk diadaptasi pada prediksi keberhasilan *follow-up* pelanggan.

### 3.6 Evaluasi Model

Model dievaluasi menggunakan metrik: akurasi, *precision*, *recall*, dan *F1-score*. Metrik tersebut digunakan untuk menilai keseimbangan kemampuan model dalam mengenali kelas positif dan negatif secara akurat.

### 3.7 Analisis Perbandingan

Performa *Random forest*, *XGBoost*, dan *Logistic Regression* dibuat perbandingan dan dinyatakan bahwa *XGBoost* memiliki performa terbaik dari segi akurasi dan ketahanan terhadap overfitting [2]. Perbandingan performa antara kedua model dilakukan dengan menganalisis nilai metrik evaluasi pada data testing.

## 4. Hasil dan Pembahasan

### 4.1 Sampel Data Penelitian:

Dataset yang digunakan dalam penelitian ini merupakan data hasil tindak lanjut (*follow-up*) pelanggan dari sebuah dealer mobil. Beberapa fitur penting yang digunakan antara lain: jenis kelamin, status *follow-up* sebelumnya, lokasi, sumber data, jenis kendaraan, serta label target yang menunjukkan keberhasilan atau kegagalan *follow-up*. Dataset awal terdiri dari 180 entri, kemudian dibagi menjadi data pelatihan (80%) dan data pengujian (20%) menggunakan teknik *stratified sampling* untuk menjaga distribusi kelas target.

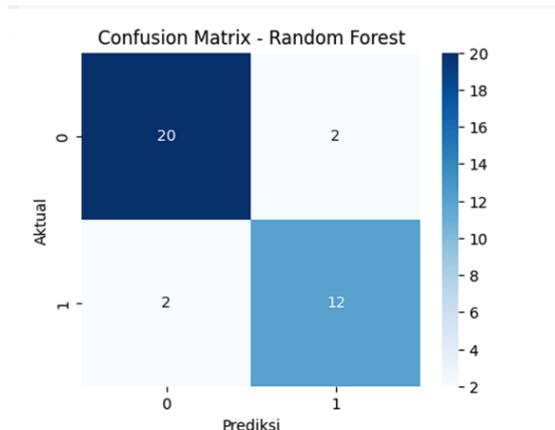
NO	Nama Pelanggan	Model	Tahun	HP	TANGGAL	PIC	HASIL FU
1	RETA DWI LESTARI	BRIO	2018	81369325555	25-01-2025	SCO	BELUM ADA MINAT GANTI KENDARAAN ,MASIH PAKE MOBIL YG LAMA
2	RIYANTO.	CIVIC	2018	82299179211	25-01-2025	SCO	MASIH PAKE MOBIL YANG LAMA MASIH NYAMAN DIGUNAKAN
3	BUDI PURNOMO	CITY	2018	81369470104	25-01-2026	SCO	BELUM RESPON
4	DAMARA SEPLYAN SIRYA	HR-V	2018	81377922333	25-01-2027	SCO	BLM BISA DIGANGGU
5	APRIJIA. S,SOS	CIVIC	2018	8127201978	25-01-2028	SCO	LAGI DIJALAN
6	ACHMAD RIZAL, SH	HR-V	2018	8127936630	25-01-2029	SCO	BLM ADA MINAT GANTI MOBIL
7	DWI SAFITRI	BRIO	2018	8972380183	25-01-2030	SCO	SUDAH BELI MOBIL BARU AVANZA
8	LUKMAN LUDIN	BRIO	2018	81379964553	25-01-2031	SCO	NO TIDAK AKTIF
9	ALIFAH HIDAYAH	HR-V	2021	82252415563	25-01-2032	SCO	LANJUT INFORMASI WA
10	DWI SAFITRI	JAZZ	2019	82185541517	25-01-2033	SCO	BLM BISA DIGANGGU
11	SAPTO SUDARSONO	BRIO	2018	81379357458	25-01-2034	SCO	BELUM ADA UANG
12	MARYULI,SE	HR-V	2016	8127468110	25-01-2035	SCO	MASIH ADA TANGGUNGAN ANAK KULIAH
13	ISTANA LAMPUNG RAYA, PT	BRIO	2018	82223071993	25-01-2036	SCO	BLM ADA MINAT GANTI MOBIL
14	M.IRWAN NASITION	BRIO	2018	82183948148	25-01-2037	SCO	BELUM RESPON
15	ARI DESMITA KUNDO, SH	CR-V	2018	81272229444	28-01-2025	SCO	LAGI DIJALAN
16	MONI	BRIO	2018	87899532099	28-01-2025	SCO	NO TIDAK AKTIF

Gambar 1 Dataset awal

### 4.2 Evaluasi Performa Algoritma dan Analisis Perbandingan

#### 1) Hasil Evaluasi Model *Random forest*

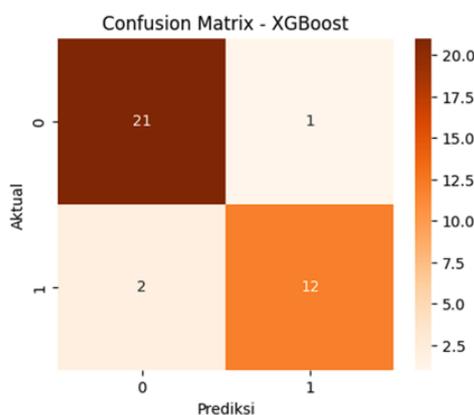
Model *Random forest* yang dikembangkan menunjukkan performa yang memuaskan pada data pengujian dengan *precision* 0,91 untuk kelas negatif dan 0,86 untuk kelas positif. *Recall* untuk kelas negatif dan positif masing-masing sebesar 0,91 dan 0,86, sedangkan *F1-score* mencapai 0,91 untuk kelas negatif dan 0,86 untuk kelas positif. Secara keseluruhan, model ini memperoleh akurasi sebesar 88,89%.



Gambar 2 Confusion Matrix-Random forest

2) Hasil Evaluasi Model XGBoost

Pemodelan XGBoost pada data kampanye pemasaran menghasilkan prediksi yang lebih akurat dibandingkan model klasifikasi lainnya, terutama pada data tidak seimbang [12]. Model XGBoost memberikan hasil evaluasi yang lebih baik dengan precision sebesar 0,91 (kelas negatif) dan 0,92 (kelas positif), recall 0,95 dan 0,86, serta F1-score 0,93 dan 0,89 untuk kelas negatif dan positif secara berurutan. Akurasi keseluruhan mencapai 91,67%.



Gambar 3 Confusion Matrix-XGBoost

3) Perbandingan Performa Kedua Model

Random Forest				
	precision	recall	f1-score	support
0.0	0.91	0.91	0.91	22
1.0	0.86	0.86	0.86	14
accuracy			0.89	36
macro avg	0.88	0.88	0.88	36
weighted avg	0.89	0.89	0.89	36
Akurasi: 0.8888888888888888				
XGBoost				
	precision	recall	f1-score	support
0.0	0.91	0.95	0.93	22
1.0	0.92	0.86	0.89	14
accuracy			0.92	36
macro avg	0.92	0.91	0.91	36
weighted avg	0.92	0.92	0.92	36
Akurasi: 0.9166666666666666				

Gambar 4 Perbandingan Performa 2 Model

Perbandingan metrik evaluasi menunjukkan bahwa *XGBoost* secara konsisten memberikan hasil yang lebih baik dibandingkan *Random forest*, khususnya pada akurasi dan *recall* kelas negatif. Hal ini menunjukkan bahwa *XGBoost* lebih mampu mengurangi kesalahan dalam mengklasifikasikan calon pelanggan yang tidak berhasil di-*Follow-up*, sehingga meningkatkan efisiensi sumber daya pemasaran. *XGBoost* menunjukkan performa klasifikasi kredit yang lebih baik dibanding *Random forest*, terutama dalam hal presisi pada data tidak seimbang [14]. *XGBoost* memberikan hasil akurasi dan *F1-score* yang lebih tinggi dibandingkan *LightGBM* dalam memprediksi churn pelanggan, membuktikan efektivitasnya untuk klasifikasi biner [15]. Namun, perbedaan performa kedua model tidak terlalu signifikan pada *recall* kelas positif (keduanya 0,86), yang menandakan bahwa masih terdapat ruang untuk peningkatan dalam mendeteksi keberhasilan *Follow-up* secara tepat.



**Gambar 5** Grafik Perbandingan Akurasi

### 4.3 Pembahasan

Klasifikasi performa UMKM menunjukkan bahwa kombinasi *preprocessing* yang baik dengan model *XGBoost* menghasilkan peningkatan akurasi yang signifikan, mendukung strategi serupa pada prediksi *Follow-up* pelanggan [16]. Berdasarkan hasil ini, beberapa rekomendasi pengembangan dapat diajukan seperti: Perluasan Dataset, dengan memperbesar dan memperkaya data untuk membantu model belajar pola dan mengurangi risiko *overfitting*. Peningkatan *Preprocessing* juga direkomendasikan dengan menggunakan teknik lanjutan seperti ADASYN dan penghilangan *noise* data untuk meningkatkan kualitas input model.

Temuan dari penelitian ini menunjukkan bahwa kedua model mampu memprediksi keberhasilan *follow-up* pelanggan dengan akurasi tinggi. Hal ini menjawab permasalahan yang diangkat dalam pendahuluan mengenai ketidakefisienan strategi *follow-up* manual, karena kedua model *machine learning* ini dapat digunakan untuk memprioritaskan pelanggan potensial secara otomatis dan berbasis data. Keunggulan *XGBoost* yang terlihat dalam hasil evaluasi memperkuat posisinya sebagai algoritma unggulan untuk klasifikasi kompleks.

Penelitian ini juga memberikan kontribusi terhadap literatur sebelumnya yang membahas penerapan *machine learning* pada klasifikasi perilaku pelanggan. *Random forest* telah digunakan secara luas dan terbukti andal, sementara *XGBoost* semakin populer karena efisiensinya [5]. Penelitian ini mengintegrasikan temuan tersebut dalam konteks industri otomotif, khususnya pada tahap *follow-up* pelanggan, yang sebelumnya belum banyak dikaji secara mendalam.

### 5. Simpulan

Penelitian ini telah berhasil mengembangkan dan membandingkan dua model *machine learning*, yaitu *Random forest* dan *XGBoost*, dalam memprediksi keberhasilan tindak lanjut (*follow-up*) pelanggan pada dealer mobil. Berdasarkan evaluasi yang dilakukan menggunakan metrik akurasi, precision, *recall*, dan *F1-score*, model *XGBoost* menunjukkan performa yang lebih unggul dengan akurasi mencapai 91,67%. Nilai precision, *recall*, dan *F1-score* pada model ini juga lebih tinggi untuk kedua kelas target dibandingkan dengan model *Random forest* yang memperoleh akurasi sebesar 88,89%. Hal ini menunjukkan bahwa *XGBoost* memiliki kemampuan yang lebih baik dalam menangani kompleksitas data dan ketidakseimbangan kelas.

Walaupun demikian, kedua model menunjukkan performa yang relatif seimbang dalam mendeteksi keberhasilan maupun kegagalan *follow-up* pelanggan. Hasil ini menunjukkan adanya peningkatan akurasi klasifikasi berkat penerapan teknik *preprocessing* data seperti encoding pada variabel kategorikal. Namun, terdapat beberapa faktor yang membatasi performa optimal model, antara lain keterbatasan ukuran dataset yang relatif kecil dan pola perilaku pelanggan yang

kompleks. Kondisi ini mengindikasikan bahwa diperlukan pendekatan pembelajaran yang lebih canggih atau pengumpulan data yang lebih representatif untuk meningkatkan akurasi dan generalisasi model prediktif pada masa mendatang.

### Referensi

- [1] A. G. Wicaksono and I. Nurhidayat, "Penerapan Algoritma Machine Learning untuk Prediksi Pelanggan Potensial pada Dealer Mobil," *J. Inform. Univ. Pamulang*, vol. 8, no. 1, pp. 12–20, 2023, [Online]. Available: <https://ojs.unpam.ac.id/index.php/informatika/article/view/11726>
- [2] B. S. C. Putra, I. Tahyudin, B. A. Kusuma, and K. N. Isnaini, "Efektivitas Algoritma Random Forest, XGBoost, dan Logistic Regression dalam Prediksi Penyakit Paru-paru," *Techno.Com*, vol. 23, no. 4, pp. 210–218, 2024, [Online]. Available: <https://publikasi.dinus.ac.id/index.php/technoc/article/view/11705>
- [3] A. Sharma, N. Patel, and R. Gupta, "Enhancing Predictive Customer Retention Using Machine Learning Algorithms : A Comparative Study of Random Forest , XGBoost , and Neural Networks Authors :," pp. 1–23.
- [4] "Emre Deniz ID \*,1 and Semanur Çökekoglu Bülbül ~.pdf."
- [5] A. Sudrajat and B. Subekti, "Analisis Perbandingan Algoritma Random Forest dan XGBoost untuk Prediksi Churn Pelanggan," *J. Ris. Inform.*, vol. 2, no. 3, pp. 184–191, 2020, [Online]. Available: <https://jurnal.polibatam.ac.id/index.php/juri/article/view/1757>
- [6] N. M. Sari and A. Wibowo, "Perbandingan Metode Random Forest dan Support Vector Machine untuk Prediksi Keberhasilan Penjualan," *J. Ilm. Teknol. Inf. Terap.*, vol. 6, no. 1, pp. 20–27, 2020, [Online]. Available: <https://ejournal.undip.ac.id/index.php/jtit/article/download/31328/19016>
- [7] Y. Jin, "Sales Prediction Based on Machine Learning Approach : Support Vector Machine , Xgboost and Random Forest," vol. 88, pp. 104–109, 2024.
- [8] T. Yang, "Sales Prediction of Walmart Sales Based on OLS , Random Forest , and XGBoost Models," vol. 49, pp. 244–249, 2023.
- [9] M. Imani and A. Beikmohammadi, "Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE , ADASYN , and GNUS Under Varying Imbalance Levels," vol. 13, no. 2, p. 8, 2025.
- [10] W. Wang *et al.*, "A User Purchase Behavior Prediction Method Based on XGBoost," 2023.
- [11] B. Kriswantara and R. Sadikin, "Used Car Price Prediction with Random Forest Regressor Model," vol. 6, no. 1, pp. 40–49, 2022, doi: 10.52362/jisicom.v6i1.752.
- [12] D. Lestari and B. Haryanto, "Evaluasi Model Machine Learning untuk Prediksi Respon Konsumen pada Kampanye Marketing," *ROY Inform.*, vol. 7, no. 1, pp. 58–65, 2023.
- [13] R. Yusuf and M. A. Pratama, "Implementasi Random Forest dan XGBoost dalam Prediksi Kredit Macet," *J. Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 40–48, 2022.
- [14] J. M. A. S. Dachy and P. Sitompul, "Analisis Perbandingan Algoritma XGBoost dan Algoritma Random Forest Ensemble Learning pada Klasifikasi Keputusan Kredit," *J. Ris. Rumpun Mat. dan Ilmu Pengetah. Alam*, vol. 2, no. 2, pp. 87–103, 2023, doi: 10.55606/jurrimipa.v2i2.1470.
- [15] A. P. Sari and E. Prasetyo, "Prediksi Customer Churn Menggunakan Algoritma XGBoost dan LightGBM," *J. Ilmu Komput. dan Sist. Inf.*, vol. 9, no. 2, pp. 124–132, 2021.
- [16] M. Erkamim, S. Suswadi, M. Z. Subarkah, and E. Widarti, "Komparasi Algoritme Random Forest dan XGBoosting dalam Klasifikasi Performa UMKM," *J. Sist. Inf. Bisnis*, vol. 13, no. 2, pp. 127–134, 2023, [Online]. Available: <https://ejournal.undip.ac.id/index.php/jsinbis/article/view/52428>