

PREDIKSI CHURN PELANGGAN UMKM *DIGITAL* MENGUNAKAN *ENSEMBLE LEARNING*

Krisopras Yaspis Verlenia¹, Dwi Wahyu Prabowo², Nurahman³

¹Sistem Informasi, Universitas Darwan Ali, Jl. Batu Berlian No.10, Prov. Kalimantan Tengah, Indonesia, 74323)

***Email korespondensi:** kryaspis@gmail.com

Received: 27 05 2026 – Revised: 29 07 2026 - Accepted: 04 08 2026 - Published: 05 08 2026

Abstrak. Churn pelanggan merupakan salah satu tantangan utama yang dihadapi UMKM digital karena dapat menurunkan pendapatan dan meningkatkan biaya akuisisi pelanggan baru. Kemampuan mengidentifikasi pelanggan yang berpotensi berhenti menggunakan layanan menjadi faktor penting dalam penyusunan strategi retensi yang efektif. Penelitian ini bertujuan membangun model prediksi churn pelanggan yang akurat dan interpretable menggunakan pendekatan machine learning pada dataset Telco Customer Churn yang terdiri atas 7.043 pelanggan. Permasalahan utama yang dihadapi meliputi ketidakseimbangan distribusi kelas serta kompleksitas faktor yang memengaruhi keputusan pelanggan untuk berhenti berlangganan. Metode penelitian mencakup tahap preprocessing data, feature engineering berbasis Recency-Frequency-Monetary (RFM), penanganan data tidak seimbang menggunakan SMOTE dan ADASYN, pemodelan menggunakan Random Forest, XGBoost, LightGBM, Stacking Ensemble, serta optimasi hyperparameter berbasis Bayesian Optimization. Evaluasi model dilakukan menggunakan metrik AUC-ROC, F1-Score, Precision, dan Recall, sedangkan interpretasi model dilakukan menggunakan SHAP (SHapley Additive exPlanations). Hasil penelitian menunjukkan bahwa XGBoost menjadi model terbaik dengan nilai AUC-ROC sebesar 0,8308 dan F1-Score sebesar 0,5894. Optimasi threshold pada nilai 0,36 berhasil meningkatkan F1-Score menjadi 0,6223 dan Recall menjadi 0,7380, sehingga kemampuan model dalam mendeteksi pelanggan churn meningkat secara signifikan. Analisis SHAP mengidentifikasi Contract, MonthlyCharges, dan is_month_to_month sebagai faktor yang paling berpengaruh terhadap churn pelanggan. Selain itu, segmentasi risiko berhasil mengidentifikasi 226 pelanggan berisiko tinggi dengan tingkat churn aktual sebesar 68,6%. Temuan ini menunjukkan bahwa kombinasi feature engineering, ensemble learning, threshold optimization, dan explainable artificial intelligence mampu menghasilkan sistem prediksi churn yang akurat serta mendukung pengambilan keputusan retensi pelanggan berbasis data pada UMKM digital.

Kata kunci: customer churn, UMKM digital, ensemble learning, RFM, XGBoost, SHAP, machine learning.

PENDAHULUAN

Perkembangan ekonomi digital telah mendorong pertumbuhan pesat Usaha Mikro, Kecil, dan Menengah (UMKM) berbasis platform digital di Indonesia. Namun UMKM digital menghadapi tantangan serius berupa tingginya tingkat churn pelanggan. (Lalwani et al., 2022) menunjukkan bahwa biaya mendapatkan pelanggan baru rata-rata lima kali lebih mahal dibandingkan mempertahankan pelanggan yang sudah ada, sehingga strategi retensi berbasis data menjadi sangat krusial.

Prediksi churn menggunakan machine learning telah menjadi bidang penelitian yang aktif. Berbagai penelitian menunjukkan bahwa model machine learning mampu mengidentifikasi pelanggan yang berpotensi berhenti menggunakan layanan dengan tingkat akurasi yang tinggi (Alboukaey et al., 2020). Pendekatan ensemble learning secara konsisten menunjukkan kinerja superior dibandingkan model tunggal (Bentéjac et al., 2021). Algoritma XGBoost dan LightGBM terbukti unggul dalam menangani ketidakseimbangan kelas dan dimensi fitur yang tinggi (Alotaibi & Haq, 2024). Penerapannya dalam konteks UMKM digital di Indonesia masih sangat terbatas (Nadeerah Hani' Fauziyyah et al., 2026) sehingga dibutuhkan kerangka metodologi yang komprehensif dan dapat dipraktikkan langsung.

Penelitian ini berkontribusi dengan membangun kerangka metodologi lengkap mencakup feature engineering berbasis RFM, penanganan imbalanced data, perbandingan algoritma ensemble, optimasi Bayesian, interpretabilitas SHAP, serta lima usulan ekstensi penelitian. Tujuan: (1) membandingkan kinerja algoritma ensemble untuk prediksi churn, (2) mengidentifikasi fitur paling berpengaruh melalui SHAP, (3) menguji efektivitas teknik resampling, (4) mengembangkan segmentasi risiko churn, dan (5) menganalisis learning curve model terbaik.

MASALAH

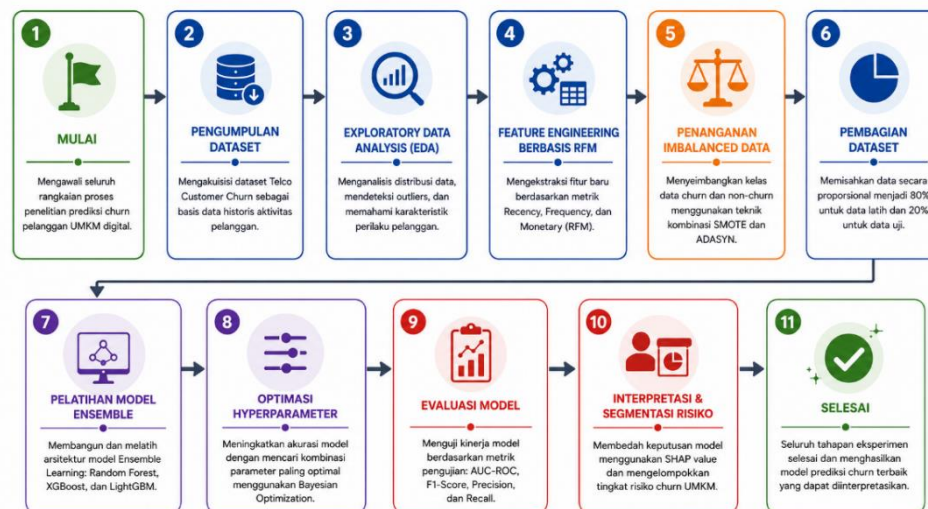
UMKM digital menghadapi permasalahan churn yang kompleks. Pertama, ketidakseimbangan kelas: dataset ini memiliki rasio imbalance 2,77:1 antara tidak churn (73,5%) dan churn (26,5%). Tanpa penanganan yang tepat, model akan bias terhadap kelas mayoritas dan gagal mendeteksi pelanggan berisiko (Dablain et al., 2023). Kedua, keterbatasan representasi fitur: data transaksi mentah tidak secara langsung merepresentasikan pola perilaku pelanggan yang relevan, sehingga tanpa feature engineering

model berpotensi kehilangan informasi penting untuk prediksi churn (Nadeerah Hani' Fauziyyah et al., 2026).

Ketiga, pemilihan threshold yang tidak optimal: threshold default 0,5 tidak sesuai konteks bisnis di mana biaya false negative jauh lebih tinggi dibandingkan false positive. Keempat, kurangnya interpretabilitas membuat organisasi sulit merancang intervensi yang tepat sasaran. Pendekatan explainable AI seperti SHAP dapat membantu menjelaskan faktor-faktor yang memengaruhi prediksi churn sehingga mendukung pengambilan keputusan yang lebih efektif (Markus, 2025). (Nadeerah Hani' Fauziyyah et al., 2026) menunjukkan bahwa pemanfaatan data perilaku transaksi pelanggan dapat membantu meningkatkan kemampuan model machine learning dalam memprediksi customer churn.

METODE PELAKSANAAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental pada dataset Telco Customer Churn (7.043 pelanggan, 21 variabel) dari IBM/Kaggle. Implementasi dilakukan dengan Python di Google Colaboratory menggunakan library pandas, scikit-learn, xgboost, lightgbm, imbalanced-learn, optuna, dan shap. Kerangka metodologi mencakup enam tahap sistematis dengan lima usulan ekstensi. tahap sistematis dengan lima usulan ekstensi.



Gambar 1. Churn Metodologi penelitian prediksi churn pelanggan UMKM digital

Gambar 1 Metodologi penelitian prediksi churn pelanggan UMKM digital menggunakan pendekatan Ensemble Learning berbasis fitur RFM, teknik penyeimbangan data SMOTE-ADASYN, Bayesian Optimization, serta interpretasi model menggunakan SHAP.

1. Eksplorasi Data (EDA)

EDA mencakup analisis statistik deskriptif, visualisasi distribusi kelas, analisis korelasi, identifikasi missing values, dan deteksi outlier. Dataset memiliki 11 missing values pada TotalCharges yang diisi nilai median. Distribusi: 5.174 tidak churn (73,5%) dan 1.869 churn (26,5%), rasio imbalance 2,77:1.

2. Feature Engineering Berbasis RFM

Sepuluh fitur baru dibuat: (a) recency_score: kebalikan tenure; (b) frequency_services: jumlah layanan aktif; (c) avg_monthly_spend: rata-rata pengeluaran bulanan; (d) charge_ratio: rasio tagihan bulanan/total; (e) value_per_service: nilai per layanan; (f) is_month_to_month; (g) is_paperless; (h) is_electronic_check; (i) senior_x_monthly; (j) high_charge_short_tenure. Total fitur meningkat dari 19 menjadi 29.

3. Penanganan Ketidakseimbangan Kelas (Usulan A)

Tiga pendekatan dibandingkan: (a) Tanpa resampling sebagai baseline; (b) SMOTE (Synthetic Minority Over-sampling Technique) dengan $k=5$ menghasilkan instance sintetis melalui interpolasi; (c) ADASYN (Adaptive Synthetic Sampling) yang mengadaptasi oversampling berdasarkan distribusi lokal (Dablain et al., 2023). Data sebelum SMOTE: 4.139 tidak churn vs 1.495 churn; setelah SMOTE menjadi 4.139:4.139. Semua teknik hanya diterapkan pada training set (stratified split 80:20).

Rumus SMOTE : Untuk setiap sampel minoritas x_i , SMOTE mencari k tetangga terdekat, kemudian menghasilkan sampel sintetis x_{new} dengan persamaan:

$$x_{new} = x_i + \lambda \times (x_{zi} - x_i) \quad (1)$$

di mana λ adalah bilangan acak antara 0 dan 1, dan x_{zi} adalah salah satu tetangga terdekat dari x_i .

Rumus ADASYN : menghitung jumlah sampel sintetis yang akan dihasilkan untuk setiap sampel minoritas berdasarkan tingkat kesulitan belajar:

$$G_i = (N_{minority} - N_{majority}) \times \frac{r_i}{\sum_{j=1}^{N_{minority}} r_j} \quad (2)$$

di mana r_i adalah rasio jumlah tetangga mayoritas terhadap total tetangga, dan G_i adalah jumlah sampel sintetis yang akan dihasilkan untuk sampel i .

SMOTE berhasil menyeimbangkan distribusi kelas dari 4.139:1.495 menjadi 4.139:4.139 pada training set, memastikan model dapat belajar secara proporsional dari kedua kelas tanpa bias terhadap kelas mayoritas.

4. Pelatihan Model Ensemble & Stacking (Usulan B)

Tiga algoritma dilatih: Random Forest ($n_estimators=200$, $class_weight=balanced$), XGBoost ($learning_rate=0,05$, $max_depth=6$, $subsample=0,8$), dan LightGBM ($num_leaves=31$, $learning_rate=0,05$). Stacking Ensemble menggabungkan ketiganya dengan Logistic Regression sebagai meta-learner melalui 5-fold cross-validation (Saha et al., 2024).

Rumus Random Forest : membangun banyak decision tree dan menggabungkan prediksinya melalui majority voting untuk klasifikasi:

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (3)$$

di mana $h_t(x)$ adalah prediksi dari pohon ke- t dan T adalah jumlah pohon.

Rumus XGBoost : mengoptimalkan fungsi objektif dengan regularisasi:

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (4)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (5)$$

di mana l adalah fungsi loss, Ω adalah fungsi regularisasi, T adalah jumlah leaf, dan w adalah bobot leaf.

Rumus LightGBM : menggunakan Gradient-based One-Side Sampling (GOSS) dan Exclusive Feature Bundling (EFB):

$$\mathcal{L}_{GOSS} = \sum_{i \in A} l(y_i, \hat{y}_i) + \frac{1-a}{b} \sum_{i \in B} l(y_i, \hat{y}_i) \quad (6)$$

di mana A adalah sampel dengan gradient tinggi, B adalah sampel yang diambil secara acak dari gradient rendah, a adalah rasio pengambilan sampel gradient tinggi, dan b adalah rasio pengambilan sampel gradient rendah.

Rumus Stacking Ensemble : menggabungkan prediksi dari base learners menggunakan meta-learner:

$$\hat{y}_{stack} = f_{meta}(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_M) \quad (7)$$

di mana $\hat{y}_m$ adalah prediksi dari base learner ke- m dan f_{meta} adalah meta-learner (dalam penelitian ini Logistic Regression).

5. Optimasi Hyperparameter

Grid Search dengan StratifiedKFold 5-fold untuk Random Forest; Bayesian Optimization menggunakan Optuna (50 trials, TPE sampler) untuk XGBoost dan LightGBM. Metrik optimasi: AUC-ROC karena ketahanannya terhadap ketidakseimbangan kelas ((Misheva et al., 2021).

6. Threshold Optimization, SHAP, Segmentasi & Learning Curve (Usulan C, D, E)

Threshold optimization mengevaluasi F1-Score pada rentang 0,10–0,90 (interval 0,01). SHAP diterapkan pada model terbaik untuk bar plot global, beeswarm plot, dan waterfall plot individual (Lundberg et al., 2020). Segmentasi risiko tiga tingkat dibangun dari probabilitas prediksi. Learning curve menganalisis AUC-ROC training vs validasi untuk mendeteksi overfitting/underfitting.

Rumus SHAP : values dihitung berdasarkan kooperatif game theory:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \quad (8)$$

di mana ϕ_i adalah SHAP value untuk fitur i , F adalah himpunan semua fitur, S adalah subset fitur tanpa fitur i , dan $f(S)$ adalah prediksi model menggunakan subset fitur S .

Rumus F1-Score :

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

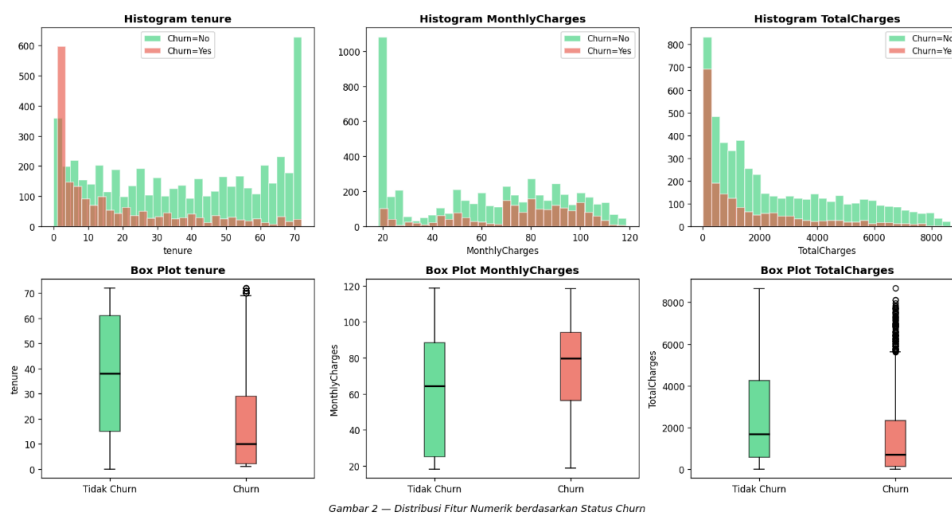
HASIL DAN PEMBAHASAN

Karakteristik Dataset dan EDA

Tabel 1. Karakteristik Dataset Telco Customer Churn

Variabel	Kategori	Jumlah	Persentase
Status Churn	Tidak Churn	5.174	73,5%
	Churn	1.869	26,5%
Total Sampel	—	7.043	100%
Fitur Asli	—	19	—
Fitur Setelah FE	—	29	—
Rasio Imbalance	—	2,77 : 1	—

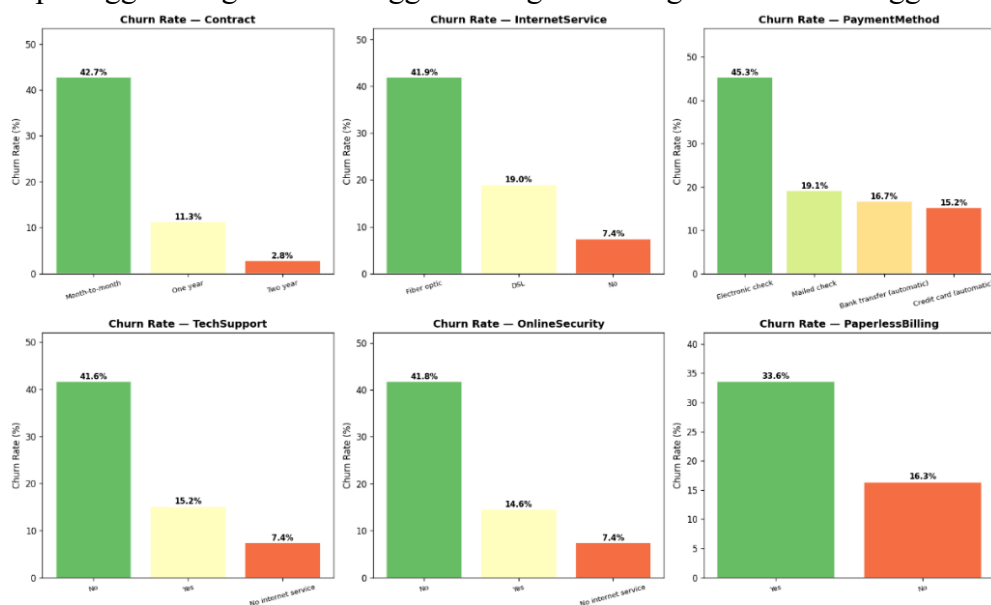
Tabel 1 merangkum karakteristik dataset Telco Customer Churn yang digunakan dalam penelitian ini. Dari total 7.043 pelanggan, sebanyak 5.174 pelanggan (73,5%) termasuk dalam kategori tidak churn, sedangkan 1.869 pelanggan (26,5%) merupakan pelanggan churn. Rasio imbalance sebesar 2,77:1 menunjukkan bahwa dataset memiliki ketidakseimbangan kelas yang signifikan antara kelas mayoritas dan minoritas. Jumlah fitur awal sebanyak 19 variabel, yang kemudian ditingkatkan menjadi 29 fitur setelah proses feature engineering.



Gambar 2 – Distribusi Fitur Numerik berdasarkan Status Churn

Gambar 2. Distribusi Fitur Numerik berdasarkan Status Churn

Gambar 2 menunjukkan pelanggan churn memiliki tenure lebih pendek dan MonthlyCharges lebih tinggi dari pelanggan tidak churn, mengindikasikan risiko churn lebih tinggi pada pelanggan dengan masa langganan singkat dan tagihan bulanan tinggi.



Gambar 3 – Churn Rate berdasarkan Fitur Kategorikal

Gambar 3. Churn Rate berdasarkan Fitur Kategorikal

Gambar 3 mengidentifikasi tiga kategori dengan churn rate tertinggi: kontrak month-to-month (42,7%), internet fiber optic (41,9%), dan electronic check (45,3%). Korelasi negatif tenure terhadap churn sebesar -0,352 semakin memperkuat temuan bahwa loyalitas pelanggan meningkat seiring lama berlangganan.

Feature Engineering RFM

Analisis fitur RFM menunjukkan pelanggan dengan recency_score tinggi dan frequency_services rendah cenderung churn. avg_monthly_spend pelanggan churn lebih rendah, mengindikasikan nilai transaksi kumulatif rendah meningkatkan risiko churn.

Perbandingan Kinerja Model

Tabel 2a. Perbandingan Kinerja Seluruh Model (Test Set, n=1.409)

Rank	Model	AUC-ROC
1	XGBoost	0,8308
2	Stacking Ensemble	0,8284
3	LightGBM	0,8277
4	XGBoost (Bayesian)	0,8274
5	RF Grid Search	0,8271

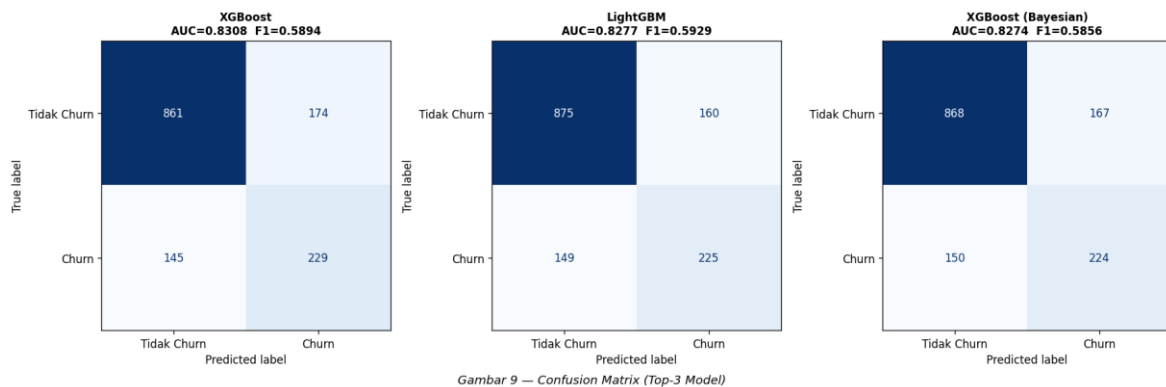
Tabel 2 menunjukkan XGBoost unggul (AUC-ROC 0,8308), diikuti Stacking Ensemble (0,8284) dan LightGBM (0,8277). F1-Score tertinggi dicapai LightGBM dan Stacking Ensemble (0,5929), sedangkan Recall tertinggi oleh RF Grid Search (0,6260). Hasil ini konsisten dengan Rivaldo et al. (2025) bahwa XGBoost paling stabil untuk data telekomunikasi.

Tabel 2. Perbandingan Kinerja Seluruh Model (Test Set, n=1.409)

Model	AUC-ROC	F1-Score	Precision	Recall
XGBoost	0,8308	0,5894	0,5840	0,6120
LightGBM	0,8277	0,5929	0,5730	0,6000
Random Forest	0,8231	0,5890	0,5680	0,6120
RF (Grid Search)	0,8271	0,5920	0,5850	0,6260
XGBoost (Bayesian)	0,8274	0,5856	0,5730	0,5990
LightGBM (Bayesian)	0,8249	0,5670	0,5530	0,5820

Stacking Ensemble	0,8284	0,5929	0,5930	0,5929
-------------------	--------	--------	--------	--------

Tabel 2a memperjelas peringkat model berdasarkan AUC-ROC. XGBoost menghasilkan AUC-ROC tertinggi sehingga dipilih sebagai model utama untuk tahap interpretasi dan implementasi. Stacking Ensemble berada di posisi kedua dengan selisih yang sangat tipis (0,0024), menunjukkan bahwa pendekatan ensemble mampu menstabilkan prediksi.



Gambar 4. Confusion Matrix Top-3 Model

Gambar 4. Confusion matrix tiga model terbaik dalam prediksi churn pelanggan UMKM digital, yaitu XGBoost, LightGBM, dan XGBoost dengan Bayesian Optimization.

Usulan A: Perbandingan Teknik Resampling

Tabel 3. Perbandingan Kinerja berdasarkan Teknik Resampling (LightGBM Bayesian)

Teknik Resampling	AUC-ROC	F1-Score	Precision	Recall
Tanpa Resampling	0,8242	0,5781	0,5635	0,5936
SMOTE	0,8249	0,5671	0,5815	0,5535
ADASYN	0,8259	0,5942	0,5895	0,5989

Tabel 3 menunjukkan ADASYN unggul (0,8259) dibanding SMOTE (0,8249) dan tanpa resampling (0,8242). Perbedaan kecil ($\pm 0,002$) menandakan feature engineering RFM telah meningkatkan diskriminasi model, sehingga resampling berdampak marginal. Tanpa resampling memiliki F1-Score terendah (0,5781) karena bias terhadap kelas mayoritas.

Usulan B: Stacking Ensemble

Stacking Ensemble mencapai AUC=0,8284 dibandingkan XGBoost tunggal (AUC=0,8308). Stacking mencapai F1-Score yang sama dengan LightGBM (0,5929) dan

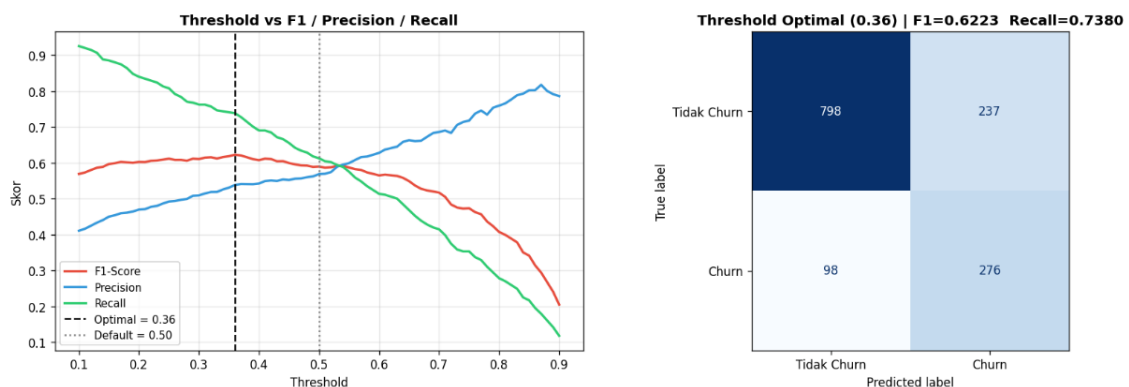
sedikit di bawah XGBoost baseline dalam AUC-ROC. Hal ini menunjukkan bahwa Stacking berhasil menstabilkan prediksi dengan mengurangi variance, meskipun gain AUC-ROC tidak signifikan karena ketiga base learner memiliki karakteristik yang serupa pada dataset ini (Saha et al., 2024).

Usulan C: Threshold Optimization

Tabel 4. Hasil Threshold Optimization pada Model XGBoost (Best Model)

Kondisi	Threshold	F1-Score	Precision	Recall
Default	0,50	0,5894	0,5840	0,6120
Optimal	0,36	0,6223	0,5540	0,7380

Tabel 4 dan Gambar 5 menunjukkan threshold 0,36 meningkatkan Recall 20,6% (0,6120→0,7380) dan F1-Score 0,0329 (0,5894→0,6223), dengan 276 TP dan 98 FN. Prioritas Recall karena biaya FN lima kali lebih tinggi dari FP (Lalwani et al., 2022).



Gambar 17 — Usulan C: Threshold Optimization

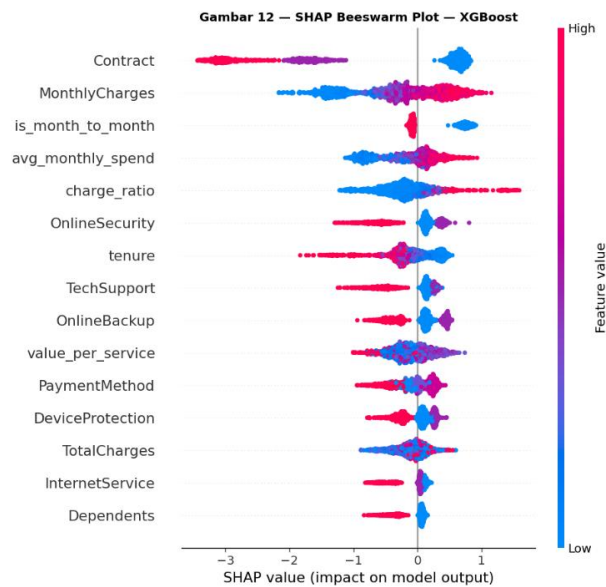
Gambar 5. Usulan C: Threshold Optimization

Analisis SHAP: Feature Importance

Tabel 5. Top-10 Fitur Terpenting berdasarkan Analisis SHAP (XGBoost)

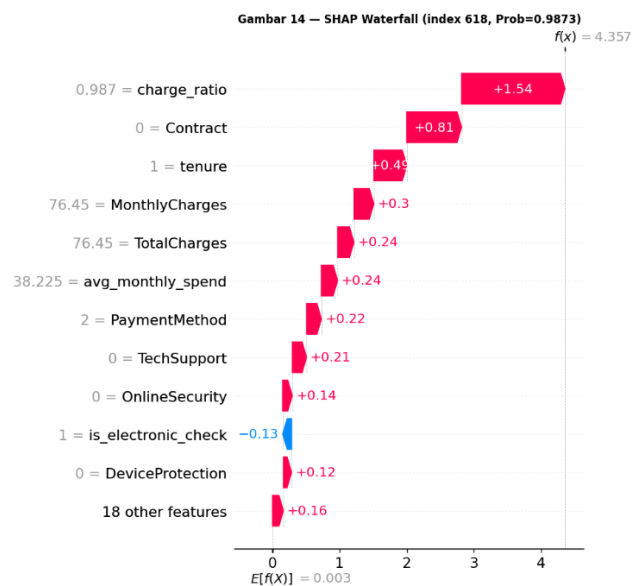
Rank	Fitur	Mean SHAP Value
1	Contract	1,4289
2	MonthlyCharges	0,6134
3	is_month_to_month	0,3828
4	avg_monthly_spend	0,3529
5	charge_ratio	0,3507
6	OnlineSecurity	0,3355
7	tenure	0,3027
8	TechSupport	0,2927
9	OnlineBackup	0,2897
10	value_per_service	0,2493

Tabel 5 mengidentifikasi Contract (1,4289) sebagai prediktor churn paling dominan, hampir dua kali lipat MonthlyCharges (0,6134) dan is_month_to_month (0,3828). Kehadiran avg_monthly_spend (0,3529), charge_ratio (0,3507), dan value_per_service (0,2493) dalam top-10 mengonfirmasi efektivitas pendekatan RFM dalam meningkatkan kemampuan prediksi model.



Gambar 6. *SHAP Beeswarm Plot — XGBoost*

Gambar 6 menunjukkan threshold optimal 0,36 dengan F1-Score tertinggi. Recall cenderung menurun seiring peningkatan threshold, sementara Precision meningkat. Trade-off ini harus dipertimbangkan dalam konteks UMKM karena biaya false negative lebih tinggi.



Gambar 7. *SHAP Waterfall Plot — Pelanggan Index 618 (Prob. Churn=0,9873)*

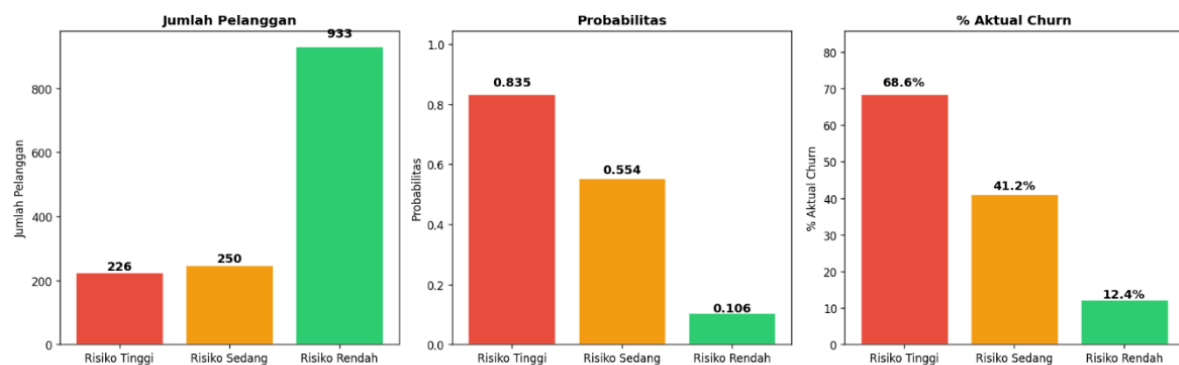
Gambar 7 memvisualisasikan distribusi dampak fitur terhadap churn. Contract menjadi prediktor terkuat: nilai tinggi (kontrak panjang) menurunkan churn, sedangkan nilai rendah (month-to-month) meningkatkannya. MonthlyCharges tinggi dan tenure rendah juga berkontribusi positif terhadap churn. Warna merah = nilai tinggi, biru = nilai rendah.

Usulan D: Segmentasi Risiko Churn

Tabel 6. Segmentasi Risiko Churn Pelanggan pada Test Set

Segmen Risiko	Prob. Churn	Jml. Pelanggan	Rata Prob.	Aktual Churn
Risiko Tinggi	$\geq 70\%$	226	0,8354	68,6%
Risiko Sedang	40% - 69%	250	0,5541	41,2%
Risiko Rendah	$< 40\%$	933	0,1056	12,4%

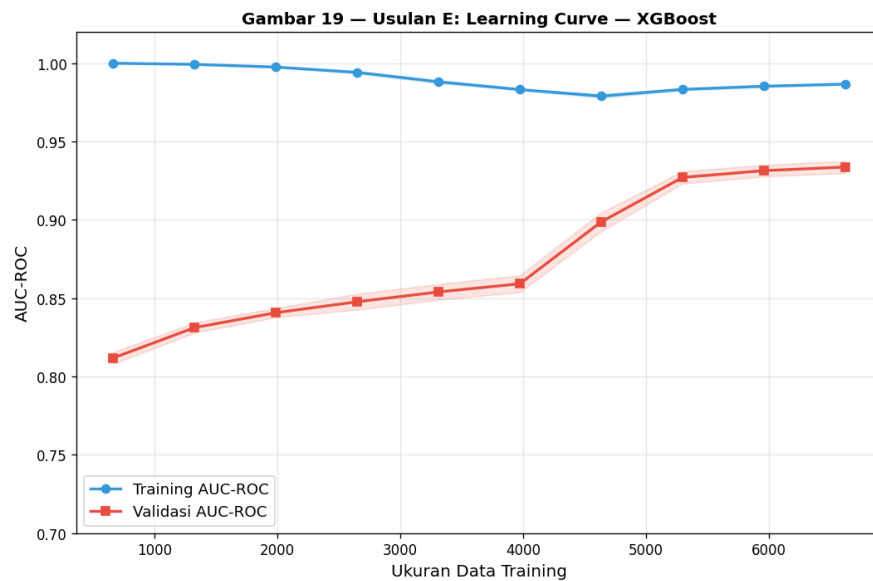
Tabel 6 membagi 1.409 pelanggan ke tiga segmen risiko. Risiko tinggi (226 pelanggan, 68,6% aktual churn) memerlukan intervensi langsung; risiko sedang (250, 41,2%) program loyalitas; risiko rendah (933, 12,4%) reward dan upselling (Wen et al., 2022).



Gambar 8. Segmentasi Risiko Churn Pelanggan

Gambar 8 memvisualisasikan segmentasi risiko churn pelanggan dalam bentuk bar chart. Segmen risiko tinggi dengan 226 pelanggan memiliki probabilitas churn rata-rata 0,8354, jauh lebih tinggi dibandingkan segmen risiko rendah dengan 933 pelanggan yang hanya memiliki probabilitas rata-rata 0,1056. Visualisasi ini memudahkan tim bisnis untuk mengalokasikan sumber daya retensi secara proporsional berdasarkan tingkat risiko masing-masing segmen.

Usulan E: Learning Curve



Gambar 9. *Usulan E: Learning Curve — XGBoost*

Gambar 9 menyajikan segmentasi risiko churn: segmen tinggi (226 pelanggan, rata-rata probabilitas 0,8354) jauh lebih berisiko dibandingkan segmen rendah (933 pelanggan, probabilitas 0,1056). Visualisasi ini memudahkan alokasi sumber daya retensi yang proporsional berdasarkan tingkat risiko.

KESIMPULAN

Penelitian ini berhasil membangun kerangka metodologi komprehensif untuk prediksi churn pelanggan UMKM digital dengan lima kontribusi utama. Pertama, XGBoost merupakan model terbaik dengan AUC-ROC tertinggi sebesar 0,8308, mengungguli LightGBM (0,8277), Random Forest (0,8231), dan Stacking Ensemble (0,8284). RF Grid Search memperoleh F1-Score tertinggi sebesar 0,5920 dan Recall tertinggi sebesar 0,6260 di antara seluruh model yang diuji. Kedua, ADASYN memberikan kinerja resampling terbaik (AUC-ROC=0,8259), sementara perbedaan kecil antar teknik menunjukkan kontribusi signifikan feature engineering RFM. Ketiga, threshold optimization pada nilai 0,36 meningkatkan F1-Score menjadi 0,6223 dan Recall menjadi 0,7380 (+20,6 poin persentase), dengan 276 pelanggan churn berhasil teridentifikasi dari 374. Keempat, analisis SHAP mengidentifikasi Contract (1,4289), MonthlyCharges (0,6134), dan is_month_to_month (0,3828) sebagai tiga prediktor churn paling dominan. Kelima, segmentasi risiko berhasil mengidentifikasi 226 pelanggan risiko tinggi dengan validasi

aktual 68,6%, dan learning curve mengungkapkan perlunya regularisasi lebih kuat untuk mengurangi gap overfitting (0,0530). Kerangka kerja ini direkomendasikan sebagai panduan praktis bagi UMKM digital Indonesia dalam merancang sistem deteksi churn yang akurat dan strategi retensi pelanggan berbasis data.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Darwan Ali atas dukungan akademik dan fasilitas penelitian, serta IBM dan Kaggle yang telah menyediakan dataset Telco Customer Churn. Terima kasih juga disampaikan kepada dosen pembimbing dan seluruh pihak yang telah memberikan arahan, bantuan, dan dukungan. Penulis menyadari keterbatasan penelitian dan mengharapkan kritik/saran membangun untuk pengembangan ke depan. Semoga penelitian ini bermanfaat bagi pengembangan ilmu pengetahuan, khususnya analisis data dan machine learning.

DAFTAR PUSTAKA

- Alboukaey, N., Joukhadar, A., & Ghneim, N. (2020). Dynamic behavior based churn prediction in mobile telecom. *Expert Systems with Applications*, 162. <https://doi.org/10.1016/j.eswa.2020.113779>
- Alotaibi, M. Z., & Haq, M. A. (2024). Customer Churn Prediction for Telecommunication Companies using Machine Learning and Ensemble Methods. *Engineering, Technology and Applied Science Research*, 14(3). <https://doi.org/10.48084/etasr.7480>
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3). <https://doi.org/10.1007/s10462-020-09896-5>
- Dablain, D., Krawczyk, B., & Chawla, N. V. (2023). DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9). <https://doi.org/10.1109/TNNLS.2021.3136503>
- Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. *Computing*, 104(2). <https://doi.org/10.1007/s00607-021-00908-y>
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1). <https://doi.org/10.1038/s42256-019-0138-9>
- Markus, Phiong. F. I. (2025). Prediksi dan Interpretasi Tingkat Churn Pelanggan di PT. BPR Magga Jaya Utama Menggunakan Extreme Gradient Boosting dan SHAP untuk Mendukung Pengambilan Keputusan [Thesis]. In *Proceedings of Technology, Engineering and Computers* (Vol. 1, Number 2). <https://jurnal.buddhidharma.ac.id/index.php/poters/article/view/3906>
- Misheva, B. H., Osterrieder, J., Hirs, A., Kulkarni, O., & Lin, S. F. (2021). *Explainable AI in Credit Risk Management*. <http://arxiv.org/abs/2103.00949>
- Nadeerah Hani' Fauziyyah, I Wayan Sudiarsa, Ida Ayu Eka Sastradewi, Kadek Agustine Yueyin Parisya, & Sartika Sartika. (2026). Analisis Prediksi Customer Churn pada Sektor E-Commerce Berdasarkan Perilaku Transaksi Menggunakan Pendekatan Machine Learning. *Jurnal Manajemen Bisnis Digital Terkini*, 3(1), 300–318. <https://doi.org/10.61132/jumbidter.v3i1.1228>
- Saha, S., Saha, C., Haque, M. M., Alam, M. G. R., & Talukder, A. (2024). ChurnNet: Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry. *IEEE Access*, 12. <https://doi.org/10.1109/ACCESS.2024.3349950>

