

PREDIKSI TINGKAT CHURN PELANGGAN DI PT. BPR MAGGA JAYA UTAMA MENGGUNAKAN EXTREME GRADIENT BOOSTING DAN SHAP UNTUK MENDUKUNG PENGAMBILAN KEPUTUSAN

Phiong Markus¹, Indah Fenriana^{2*}

^{1,2} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

*Corresponding Author, email: indah.fenriana@ubd.ac.id

ABSTRAK

Customer churn adalah penurunan persentase pelanggan dalam suatu industri yang terjadi dalam jangka waktu tertentu yang dapat mempengaruhi pertumbuhan suatu bisnis. Dalam industri perbankan terus selalu memberikan inovasi terbaru baik secara pelayanan, fasilitas, dan kepuasan. Mempertahankan loyalitas nasabah merupakan tantangan utama di tengah persaingan yang semakin ketat. Salah satu permasalahan yang kerap terjadi adalah *customer churn*, yaitu berhentinya nasabah dalam menggunakan layanan bank. Penelitian ini bertujuan untuk membangun sistem prediksi churn menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) dan melakukan interpretasi prediksi dengan metode *Shapley Additive Explanations* (SHAP) guna mendukung pengambilan keputusan strategis di PT. BPR Magga Jaya Utama. Data yang digunakan terdiri dari 54.213 sampel nasabah dengan 13 atribut utama, diolah menggunakan pendekatan machine learning yang mencakup proses pra-pemrosesan data, balancing dengan SMOTE, pelatihan model, dan evaluasi kinerja. Hasil penelitian menunjukkan bahwa XGBoost mampu memprediksi churn secara akurat dengan metrik evaluasi yang tinggi, sedangkan metode SHAP berhasil menginterpretasikan kontribusi masing-masing fitur terhadap prediksi churn secara transparan. Sistem prediksi ini diintegrasikan dalam sebuah *platform* berbasis *web* untuk memudahkan analisis dan pengambilan keputusan oleh pihak manajemen. Dengan implementasi sistem ini, PT. BPR Magga Jaya Utama diharapkan dapat meningkatkan retensi nasabah dan mengurangi tingkat churn secara signifikan melalui intervensi yang lebih tepat sasaran.

Kata kunci: *Churn*, Xgboost, SHAP, Perbankan, *Machine Learning*, Retensi Pelanggan.

I. PENDAHULUAN

Industri perbankan memiliki peran penting dalam perekonomian, terutama dalam menghimpun dana masyarakat dan menyalurkannya kembali dalam bentuk kredit untuk mendukung kegiatan ekonomi. tantangan utama yang dihadapi bank adalah bagaimana mempertahankan nasabah agar tetap menggunakan layanan mereka. Keberhasilan suatu bisnis adalah mampu mendapatkan nasabah baru dan mempertahankan nasabah lama untuk menjadi loyalitas terhadap bank (Singh et al., 2024). Hal ini membuka peluang pemasaran baru seiring meningkatnya interaksi pelanggan, sementara strategi pemasaran tetap berfokus pada mempertahankan

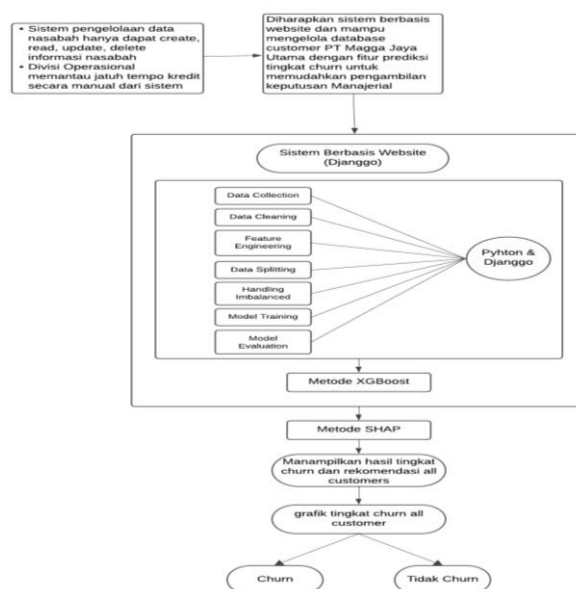
pelanggan (Soni & Nelson, 2023). Pelanggan yang berhenti menggunakan layanan bank, atau dikenal sebagai churn, merupakan isu krusial karena mencerminkan kepuasan pelanggan, kualitas layanan, dan daya saing perusahaan di pasar (Miryam Clementine & Arum, 2022). Dapat terjadi karena berbagai faktor seperti ketidakpuasan terhadap layanan, penawaran kompetitor yang lebih menarik, hingga perubahan kebutuhan finansial nasabah. Teknologi digital juga mempercepat proses perpindahan nasabah antar layanan, terutama jika sistem bank tidak mampu menyediakan fitur lengkap dan proses transaksi yang cepat serta *real-time* (Husein & Harahap, 2021).

Salah satu teknologi yang dapat digunakan adalah metode *Extreme Gradient Boosting* (XGBoost), yang merupakan algoritma berbasis *decision tree* dengan pendekatan *ensemble learning*. Metode ini terbukti mampu menghasilkan model prediksi yang andal dalam tugas klasifikasi dan regresi (Liang, 2021). Dengan memanfaatkan data historis nasabah seperti riwayat transaksi, interaksi, dan pola perilaku, model ini dapat memprediksi dengan lebih akurat siapa saja yang berpotensi berhenti menggunakan layanan (Faran & Triayudi, 2024). Agar hasil prediksi lebih dapat dipahami oleh pihak manajemen dan pengambil keputusan, maka digunakan pendekatan interpretasi model bernama *Shapley Additive Explanations* (SHAP). Teknik ini menghitung kontribusi setiap fitur dalam model terhadap hasil prediksi berdasarkan nilai *Shapley* dari teori permainan (Syamkalla et al., 2024). SHAP memberikan transparansi dan penjelasan yang akurat terhadap pengaruh masing-masing variabel input terhadap keputusan model, baik untuk prediksi individual maupun keseluruhan (Baruah & Sarma, 2024).

Sistem prediksi churn membantu bank mengantisipasi potensi churn lebih cepat dengan menganalisis data historis pelanggan, seperti riwayat transaksi, interaksi, dan pola perilaku (Nalatissifa & Pardede, 2021). Berbasis website yang menyajikan informasi dalam bentuk teks, gambar, audio, atau animasi, baik secara statis maupun dinamis, dan saling terhubung membentuk sistem terintegrasi (Riko Rivanthio, 2020). Proses ini memanfaatkan kecerdasan buatan (AI) yang memungkinkan sistem mempelajari data, mengenali pola, dan membuat keputusan secara cerdas (Eriana et al., 2023).

II. METODOLOGI

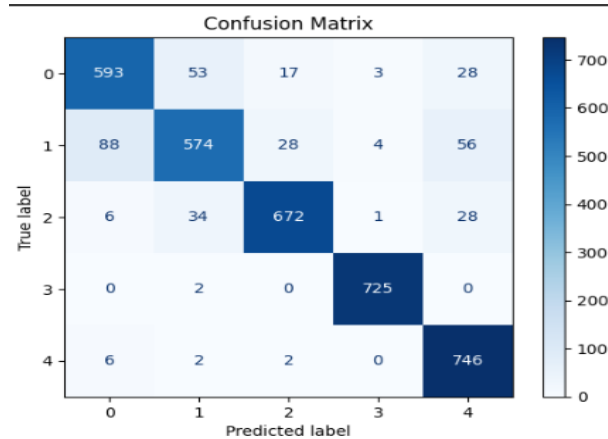
Penelitian ini mengintegrasikan pendekatan kuantitatif berbasis machine learning ke dalam sistem berbasis web menggunakan Django dan Python, dengan menganalisis data yang dimiliki untuk mengidentifikasi pola dan menarik kesimpulan yang akurat dari aktivitas yang diamati sebelumnya (Amirulhaq Iskandar et al., 2023), seperti tergambar pada diagram. Sistem ini bertujuan mempermudah pengelolaan database nasabah PT. BPR Magga Jaya Utama sekaligus memprediksi tingkat churn pelanggan menggunakan algoritma Extreme Gradient Boosting (XGBoost). Seluruh proses mulai dari pengumpulan data, pembersihan, rekayasa fitur, pembagian data, penanganan data tidak seimbang, pelatihan, hingga evaluasi model dibangun di atas framework Django yang menerapkan pola arsitektur Model, View, dan Template (MVT) (Saputra & Fathoni Aji, 2022). Model dikembangkan dengan XGBClassifier menggunakan parameter seperti `random_state = 42`, `eval_metric = 'mlogloss'`, dan regularisasi `reg_alpha = 0.5`, `reg_lambda = 1.0` untuk menghindari overfitting dan menjaga akurasi. Evaluasi model menggunakan metrik akurasi, presisi, recall, dan F1-score. Setelah itu, interpretasi model dilakukan dengan metode SHAP guna menjelaskan kontribusi masing-masing fitur, di mana SHAP (Shapley Additive Explanations) merupakan metode yang menghitung pengaruh setiap fitur terhadap output model machine learning (Khairunnisa, 2023), seperti Plafond dan Saldo Tabungan, terhadap prediksi churn.



Gambar 1. Prosedur Penelitian

III. HASIL DAN PEMBAHASAN

Model prediksi churn pelanggan pada PT. BPR Magga Jaya Utama dikembangkan menggunakan algoritma *Extreme Gradient Boosting* (XGBoost). Model ini dievaluasi melalui confusion matrix guna menilai akurasi klasifikasi dan kemampuan diskriminatif terhadap setiap kelas kolektibilitas.



Gambar 2. Confusion Matrix

Berdasarkan hasil pengujian terhadap 3.668 data uji, confusion matrix menunjukkan bahwa model mampu mengklasifikasikan mayoritas data dengan tepat. Sebagai contoh, untuk kelas 4, sebanyak 746 *instance* berhasil diklasifikasikan dengan benar, sementara hanya terdapat sedikit kesalahan klasifikasi ke kelas lain. Hasil serupa juga tercermin pada kelas 0, 1, 2, dan 3 dengan prediksi yang akurat, serta performa memadai pada kelas.

Selanjutnya, performa model dikaji lebih lanjut melalui metrik evaluasi utama yang meliputi *precision*, *recall*, dan *F1-score*. Seperti ditunjukkan pada Tabel 1 model menunjukkan *precision* tertinggi sebesar 0.99 pada kelas 3 dan terendah 0.76 pada kelas 0. Sementara itu, nilai *recall* berkisar antara 0.68 hingga 0.99, dengan *F1-score* tertinggi 0.99 dan terendah 0.75. Secara keseluruhan, akurasi model mencapai 88%, menunjukkan bahwa XGBoost mampu mengklasifikasi churn dengan akurasi dan stabilitas yang baik.

Tabel 1. Hasil Evaluasi Model XGBoost

Performance	XGBoost		
	Precision	Recall	F1-Score
0.0	0.76	0.87	0.81
1.0	0.84	0.68	0.75
2.0	0.93	0.88	0.91

Performance	XGBoost		
	Precision	Recall	F1-Score
3.0	0.99	0.99	0.99
4.0	0.88	0.98	0.93
Accuracy	0.88%		

Secara umum, Setiap pohon dalam model XGBoost berupaya memperbaiki kesalahan sebelumnya untuk menyelesaikan masalah regresi, klasifikasi, dan perangkaian (Alotaibi & Haq, 2024). Hasil evaluasi membuktikan bahwa XGBoost tidak hanya efektif dalam klasifikasi multikelas kolektibilitas, tetapi juga mampu meningkatkan akurasi secara signifikan.

IV. SIMPULAN

Prediksi *churn* terhadap nasabah PT Magga Jaya Utama memungkinkan identifikasi terhadap faktor-faktor yang berkontribusi terhadap keputusan nasabah untuk menghentikan layanan perbankan, Pemahaman terhadap faktor ini penting untuk membantu bank mengembangkan strategi retensi pelanggan yang lebih efektif. Prediksi ini mampu meningkatkan pengambilan langkah proaktif oleh divisi terkait untuk mempertahankan nasabah menggunakan layanan pada PT Magga Jaya Utama seperti menyediakan penawaran khusus, meningkatkan layanan pelanggan, dan membangun relasi dengan nasabah melalui komunikasi.

DAFTAR PUSTAKA

- 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), 14572–14578. <https://doi.org/10.48084/etasr.7480>
- Amirulhaq Iskandar, M., Latifa, U., Singaperbangsa Karawang, U., & Ronggo Waluyo, J. H. (2023). WEBSITE PREDIKSI CUSTOMER CHURN UNTUK MEMPERTAHANKAN PELANGGAN PADA PERUSAHAAN TELEKOMUNIKASI. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 2).

- Baruah, P., & Sarma, B. (2024). *Predictive Analysis on Churn Customers in Business Industries using Supervised Machine Learning Algorithms and Smote*. <https://www.researchgate.net/publication/384652783>
- Eriana, E. S., Kom, S., Kom, M., & Zein, D. A. (2023). *ARTIFICIAL INTELLIGENCE (AI) PENERBIT CV. EUREKA MEDIA AKSARA*.
- Faran, J., & Triayudi, A. (2024). ANALYSIS OF THE EFFECTIVENESS OF POLYNOMIAL FIT SMOTE MESH ON IMBALANCE DATASET FOR BANK CUSTOMER CHURN PREDICTION WITH XGBOOST AND BAYESIAN OPTIMIZATION. *Jurnal Teknik Informatika (JUTIF)*, 5(3), 661–667. <https://doi.org/10.52436/1.jutif.2024.5.3.1284>
- Husein, A. M., & Harahap, M. (2021). Pendekatan Data Science untuk Menemukan Churn Pelanggan pada Sector Perbankan dengan Machine Learning. *Data Sciences Indonesia (DSI)*, 1(1), 8–13. <https://doi.org/10.47709/dsi.v1i1.1169>
- Khairunnisa, A. (2023). PERBANDINGAN MODEL RANDOM FOREST DAN XGBOOST UNTUK PREDIKSI KEJAHATAN KESUSILAAN DI PROVINSI JAWA BARAT. *JIKO (Jurnal Informatika Dan Komputer)*, 7(2), 202. <https://doi.org/10.26798/jiko.v7i2.799>
- Liang, S. (2021). Comparative Analysis of SVM, XGBoost and Neural Network on Hate Speech Classification. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 896–903. <https://doi.org/10.29207/resti.v5i5.3506>
- Miryam Clementine, & Arum. (2022). Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Naïve Bayes dan ID3. *Jurnal Processor*, 17(1), 9–18. <https://doi.org/10.33998/processor.2022.17.1.1170>
- Nalatissifa, H., & Pardede, H. F. (2021). Customer Decision Prediction Using Deep Neural Network on Telco Customer Churn Data. *Jurnal Elektronika Dan Telekomunikasi*, 21(2), 122. <https://doi.org/10.14203/jet.v21.122-127>
- Riko Rivanthio, T. (2020). *PERANCANGAN PENGAJUAN SIDANG LAPORAN PRAKTEK KERJA LAPANGAN MAHASISWA BERBASIS WEBSITE PADA SEKOLAH TINGGI ANALIS BAKTI ASIH BANDUNG* (Vol. 7, Issue 1).
- Saputra, D., & Fathoni Aji, R. (2018). ANALISIS PERBANDINGAN PERFORMA WEB SERVICE REST MENGGUNAKAN FRAMEWORK LARAVEL, DJANGO DAN RUBY ON RAILS UNTUK AKSES DATA

DENGAN APLIKASI MOBILE (Studi Kasus: Portal E-Kampus STT Indonesia Tanjungpinang). *Bangkit Indonesia*, 2.

Singh, P. P., Anik, F. I., Senapati, R., Sinha, A., Sakib, N., & Hossain, E. (2024).

Investigating customer churn in banking: A machine learning approach and visualization app for data science and management. *Data Science and Management*, 7(1), 7–16. <https://doi.org/10.1016/j.dsm.2023.09.002>

Soni, P. K., & Nelson, L. (2023). PCP: Profit-Driven Churn Prediction using Machine Learning Techniques in Banking Sector. *International Journal of Performability Engineering*, 19(5), 303–311. <https://doi.org/10.23940/ijpe.23.05.p2.303311>

Syankalla, M. T., Khomsah, S., & Nur, Y. S. R. (2024). Implementasi Algoritma Catboost Dan Shapley Additive Explanations (SHAP) Dalam Memprediksi Popularitas Game Indie Pada Platform Steam. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(4), 777–786. <https://doi.org/10.25126/jtiik.1148503>