

Peramalan Stok Gula Merah dengan ARIMA dan Prediksi Biaya Distribusi menggunakan XGBoost

^{1*}Ellok Sintha Maydiana, ²Patmi kasih, ³Muh Aris Saputra

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹_elloksinthamaydiana@gmail.com, ²patmikasih@unpkdr.ac.id, ³aris.saputra@unpkdr.ac.id

Penulis Korespondens : Ellok Sintha Maydiana

Abstrak— Ketidakpastian persediaan stok barang dan fluktuasi biaya logistik distribusi memicu inefisiensi operasional pada sektor perdagangan gula merah. Penelitian ini bertujuan mengembangkan sistem manajemen penjualan terintegrasi menggunakan arsitektur hibrida dengan kerangka kerja Laravel dan React.js, serta layanan kecerdasan buatan berbasis Python. Metode peramalan stok dieksekusi melalui algoritma *Autoregressive Integrated Moving Average* (ARIMA), sedangkan estimasi biaya pengiriman logistik menggunakan algoritma Extreme Gradient Boosting (XGBoost). Hasil pengujian komputasi peramalan murni ke depan (out-of-sample) dari model ARIMA(1,1,1) merekam tingkat kesalahan dinamis dengan *Mean Absolute Percentage Error* (MAPE) di rentang 32,29% hingga 61,52%, menjustifikasi kemampuan model dalam mendeteksi tren dasar persediaan. Sementara itu, model regresi XGBoost meraih akurasi Koefisien Determinasi (R²) memukau sebesar 95,5% dengan Mean Absolute Error (MAE) senilai Rp12.053,21. Kesimpulannya, integrasi sistem teknologi mutakhir ini terbukti andal menciptakan tata kelola anggaran berbasis data (data-driven).

Kata Kunci— ARIMA, biaya distribusi, gula merah, manajemen penjualan, XGBoost

Abstract— *Inventory stock uncertainty and fluctuating distribution logistic costs trigger operational inefficiencies in the brown sugar trading sector. This study aims to develop an integrated sales management system utilizing a hybrid architecture with Laravel and React.js frameworks, alongside a Python-based artificial intelligence service. The stock forecasting method employs the Autoregressive Integrated Moving Average (ARIMA) algorithm, while the delivery cost estimation utilizes Extreme Gradient Boosting (XGBoost). The out-of-sample computational testing of the ARIMA(1,1,1) model recorded a dynamic error rate with a Mean Absolute Percentage Error (MAPE) ranging from 32.29% to 61.52%, justifying the model's ability to detect baseline inventory trends. Meanwhile, the XGBoost regression model achieved an impressive Coefficient of Determination (R²) of 95.5% with a Mean Absolute Error (MAE) of Rp12,053.21. In conclusion, the integration of this technology proves reliable in creating a data-driven budget management.*

Keywords— ARIMA, brown sugar, distribution cost, sales management, XGBoost

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Perkembangan teknologi informasi saat ini telah menjadi instrumen vital dalam mereduksi ketidakefisienan pada rantai pasok industri agrobisnis. Distributor Gula *Express* merupakan entitas niaga yang memfokuskan usahanya pada perdagangan gula merah. Pengelolaan data yang selama ini manual memicu keputusan pengadaan stok yang tidak tepat [1]. Selain itu, penetapan anggaran logistik pengiriman yang sering kali subjektif memicu inefisiensi kas perusahaan [2].

Dibutuhkan sistem manajemen penjualan yang dilengkapi dengan kecerdasan komputasi [3]. Studi oleh Fadillah dkk [4] menyimpulkan bahwa algoritma statistik deret waktu ARIMA sangat efektif memetakan fluktuasi pasokan pangan tanpa rekayasa fitur rumit. Riset terkait algoritma XGBoost juga membuktikan keunggulannya dalam mengkalkulasi hubungan non-linear multi-variabel tabular [5]. Penelitian ini merancang sebuah Sistem Manajemen Penjualan Gula Merah hibrida yang memadukan ARIMA untuk peramalan stok persediaan dan XGBoost untuk estimasi anggaran logistik..

II. METODE

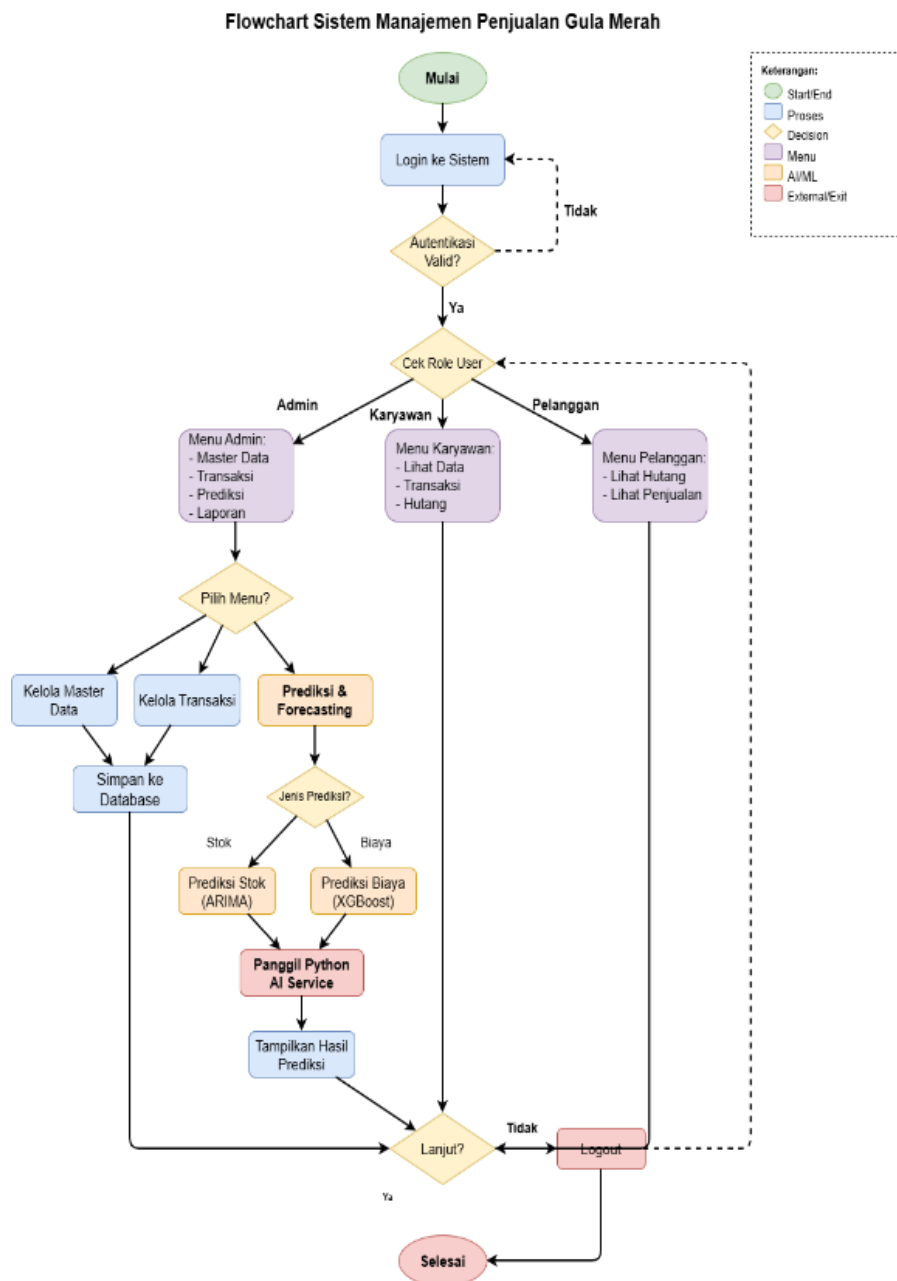
Desain penelitian ini menggunakan pendekatan sekuensial air terjun (*Waterfall*). Kerangka sistem backend menggunakan bahasa PHP melalui framework Laravel, sedangkan antarmuka interaktif frontend direalisasikan dengan kerangka kerja React.js dan Inertia.js berbasis TypeScript [6]. Modul kecerdasan buatan dipisah (isolated) ke dalam layanan mandiri Python (*Python AI Service*) melalui arsitektur REST API.

Subyek dan bahan yang diteliti bertumpu pada dataset operasional transaksi riil Distributor Gula *Express*. Pengambilan data dilakukan terhadap catatan historis dari bulan Juli 2024 hingga Mei 2026 yang diekstraksi ke dalam basis data relasional [3]. Teknik pengambilan sampel data dilakukan menggunakan metode purposive sampling berdasarkan kriteria kelengkapan atribut transaksi. Untuk model peramalan ARIMA, pengambilan data difokuskan pada ekstraksi riwayat faktur penjualan harian, yang kemudian diagregasi (dijumlahkan) secara periodik per jenis varietas gula merah [7]. Proses ini mentransformasi data relasional menjadi format data deret waktu (time-series) yang berurutan secara kronologis [4]. Sementara itu, pengambilan data untuk model XGBoost dilakukan dengan menarik log riwayat surat jalan pengiriman logistik. Data mentah ini kemudian dibersihkan dari missing values dan dibentuk menjadi format tabular matriks fitur [5].

Pada tahap pengujian model cerdas, dataset dibagi secara proporsional dengan rasio 80:20, di mana 80% data bertindak sebagai data latih (training data) dan 20% sisanya didedikasikan sebagai data uji peramalan murni ke depan (out-of-sample). Variabel yang diukur pada modul peramalan persediaan meliputi riwayat kuantitas jumlah permintaan harian per jenis varietas gula merah. Model peramalan menggunakan algoritma ARIMA dengan konfigurasi orde (1,1,1). Nilai *differencing* tingkat pertama ($d=1$) digunakan untuk mengurangi pengaruh tren pada data historis sehingga pola data menjadi lebih stabil untuk dimodelkan. Implementasi model dilakukan menggunakan library Statsmodels pada Python yang secara otomatis menerapkan proses *differencing* sesuai parameter yang ditentukan [8]. Data yang teridentifikasi non-stasioner disesuaikan melalui proses diferensiasi tingkat pertama ($d=1$), sehingga model dieksekusi menggunakan ordo *Autoregressive Integrated Moving Average* atau ARIMA(1,1,1).

Di sisi lain, model *regresi ensemble Extreme Gradient Boosting (XGBoost)* dikonfigurasi untuk mengelola beban komputasi prediksi anggaran biaya distribusi [9]. Desain eksperimen untuk modul regresi ini melibatkan variabel independen (prediktor) berupa jarak rute pengiriman dalam kilometer, berat kapasitas muatan dalam kilogram, serta klasifikasi tipe kendaraan armada ekspedisi (truk sedang dan pickup) yang dikonversi menjadi format matriks melalui teknik Label

Encoding [10]. Target variabel (dependen) dari analisis ini adalah nominal estimasi ongkos distribusi riil dalam satuan mata uang Rupiah.



Gambar 1. Flowchart System

Gambar 1 menyajikan alur prosedur eksekusi sistem secara menyeluruh. Proses dimulai dengan pengambilan data operasional riil yang tersimpan dalam basis data, mencakup data barang, pelanggan, dan riwayat distribusi. Setelah data ditarik, sistem akan mengevaluasi fungsionalitas utama yang diminta pengguna. Jika permintaan berupa peramalan stok persediaan barang, sistem akan memanggil subsistem berbasis Python untuk mengeksekusi algoritma deret waktu ARIMA

(1,1,1) dengan differencing tingkat pertama ($d=1$) untuk mengurangi tren pada data historis sebelum proses peramalan dilakukan. [8] dan differencing tingkat pertama ($d=1$). Sebaliknya, jika permintaan berupa estimasi biaya distribusi logistik, sistem akan mengaktifkan prosedur komputasi model ensemble XGBoost [9].

Prosedur ini melibatkan rekayasa fitur variabel input seperti jarak km, berat kargo kg, serta klasifikasi kategorikal kendaraan armada logistik melalui teknik Label Encoding. Kedua prosedur analisis tersebut secara akumulatif akan melahirkan visualisasi dasbor analitik berbasis data-driven yang siap dijadikan sebagai landasan objektif perencanaan arus kas pendistribusian masa depan.

III. HASIL DAN PEMBAHASAN

A. Fungsionalitas Sistem Utama

Sistem informasi manajemen menyajikan dasbor spesifik dengan hierarki pengguna melalui pengamanan lapis ganda (*Two-Factor Authentication/2FA*). Berdasarkan uji prosedural Black-Box Testing, modul master data, kartu pemutakhiran stok, hingga entri faktur penjualan berhasil beroperasi penuh tanpa mendeteksi keberadaan galat teknis.

B. Evaluasi Model Peramalan ARIMA

Pengujian prediksi out-of-sample model ARIMA(1,1,1) diukur menggunakan dataset dari 1 November 2025 hingga 1 April 2026. Target peramalan adalah 4 minggu ke depan secara riil. Hasil komputasi menunjukkan bahwa nilai *Mean Absolute Percentage Error (MAPE)* berfluktuasi secara dinamis antara 32,29% (Gula Ceplik) hingga tertinggi 61,52% (Gula Tebu). Kenaikan indeks rasio error ini terbilang rasional dan wajar untuk studi peramalan murni (out-of-sample), di mana kejutan permintaan pasar (*market shock*) sesungguhnya di masa depan tidak dapat dilacak secara matematis melalui data latih masa lampau [7]. Kendati memproduksi MAPE di rentang puluhan persen, skema peramalan ini tetap terbukti sukses memberikan sumbu panduan (*baseline*) objektif bagi manajer pergudangan dalam menangkal krisis kekosongan barang.

Gambar 2. Form Parameter Prediksi ARIMA

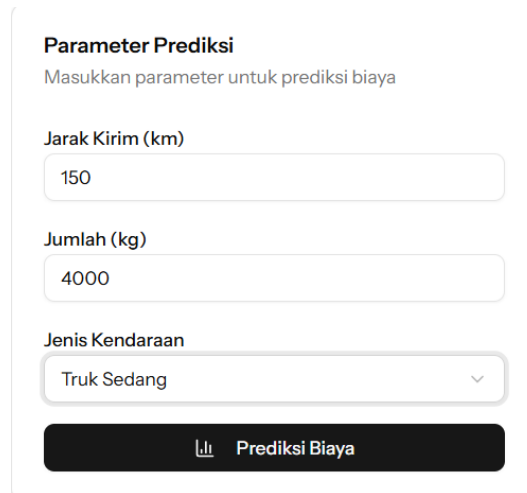
Gambar 2 menunjukkan antarmuka modul peramalan stok yang digunakan pengguna untuk menentukan parameter prediksi. Pengguna dapat memilih jenis barang, menentukan rentang data historis yang digunakan sebagai data latih, serta memasukkan jumlah minggu yang akan diprediksi. Pada pengujian penelitian ini, data historis digunakan pada rentang 1 November 2025 hingga 1 April 2026 dengan horizon prediksi selama 4 minggu ke depan. Parameter tersebut kemudian dikirim ke layanan Python AI Service untuk diproses menggunakan model ARIMA(1,1,1) sehingga menghasilkan estimasi kebutuhan stok pada periode mendatang.

Tabel 1 Matriks Evaluasi Kinerja Model ARIMA
(Periode Uji Nov 2025 - Apr 2026)

Kategori Gula	MAE (Kg)	RMSE (Kg)	MAPE (%)
Gula Tebu	3.345,26	4.119,62	61,52%
Gula Kelapa	492,89	503,06	32,86%
Gula Ceplik	732,10	1.052,61	32,29%
Gula Tanggung	2.358,63	2.617,62	48,57%

C. Evaluasi Model Prediksi Biaya XGBoost

Pengujian presisi anggaran operasional dilakukan oleh fungsi regresi pohon keputusan beruntun XGBoost. Validasi ketepatan regresi sukses menyentuh angka Koefisien Determinasi (R^2) fantastis sebesar 95,5%. Dalam simulasi distribusi sejauh 150 km, kapasitas kargo 4.000 kg dengan kendaraan truk sedang, sistem mendeteksi simpangan eror MAE hanya sebesar Rp12.053,21 [10]. Deviasi dalam takaran belasan ribu rupiah pada operasional skala niaga membuktikan kapabilitas algoritma ini.



Parameter Prediksi
Masukkan parameter untuk prediksi biaya

Jarak Kirim (km)
150

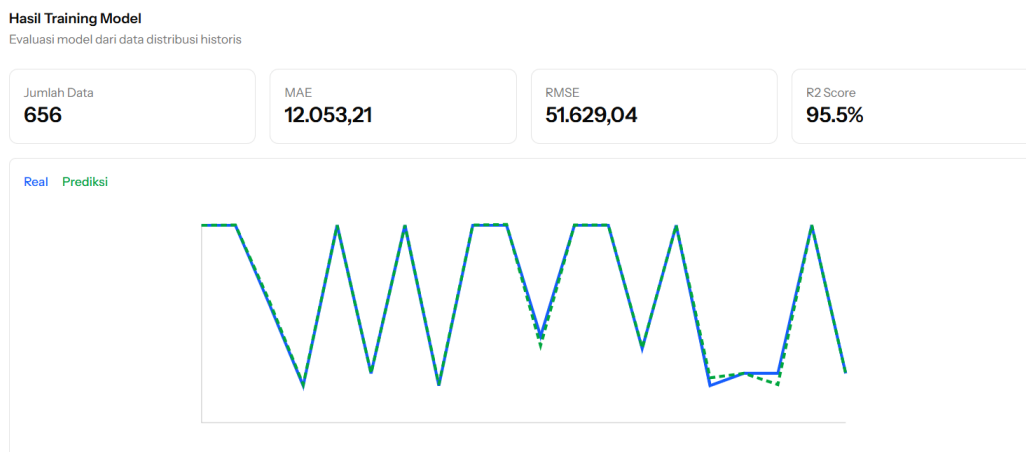
Jumlah (kg)
4000

Jenis Kendaraan
Truk Sedang

[📄](#) **Prediksi Biaya**

Gambar 3. Form Prediksi Biaya Distribusi

Gambar 3 memperlihatkan form parameter prediksi biaya distribusi berbasis algoritma XGBoost. Variabel independen yang digunakan meliputi jarak pengiriman (km), berat muatan (kg), dan jenis kendaraan. Ketiga variabel tersebut dipilih karena memiliki pengaruh langsung terhadap besaran biaya distribusi. Pada penelitian ini dilakukan simulasi menggunakan parameter jarak 150 km, muatan 4.000 kg, dan kendaraan truk sedang. Data input kemudian diproses oleh model XGBoost untuk menghasilkan estimasi biaya distribusi dengan tingkat akurasi yang diukur menggunakan metrik MAE, RMSE, dan koefisien determinasi (R^2).



Gambar 4 Grafik Komparasi Biaya Distribusi Aktual dan Prediksi XGBoost

Gambar 3 menunjukkan perbandingan antara biaya distribusi aktual dan hasil prediksi menggunakan algoritma XGBoost. Grafik memperlihatkan bahwa pola prediksi mengikuti tren data aktual dengan cukup baik sehingga menunjukkan kemampuan model dalam mempelajari hubungan antara variabel jarak pengiriman, berat muatan, dan jenis armada. Hasil evaluasi memperoleh nilai koefisien determinasi (R^2) sebesar 95,5%, yang menunjukkan bahwa sebagian besar variasi biaya distribusi dapat dijelaskan oleh model. Selain itu, nilai MAE sebesar Rp12.053,21 dan RMSE sebesar Rp51.629,04 mengindikasikan tingkat kesalahan prediksi yang

relatif rendah sehingga model layak digunakan sebagai alat bantu estimasi biaya distribusi pada Distributor Gula Express.

IV. KESIMPULAN

Kesimpulan pertama dari penelitian ini menunjukkan bahwa integrasi algoritma *Autoregressive Integrated Moving Average* (ARIMA) terbukti mampu mendigitalisasi pengambilan keputusan pada pengadaan stok persediaan komoditas gula merah. Berdasarkan evaluasi komputasi peramalan murni ke depan (out-of-sample), model ARIMA(1,1,1) merekam tingkat galat *Mean Absolute Percentage Error* (MAPE) yang berfluktuasi secara dinamis di rentang 32,29% hingga 61,52%. Meskipun rentang galat ini merepresentasikan tingginya volatilitas kejutan pasar (*market shock*) di lapangan yang tidak linear, skema peramalan ini sukses menciptakan parameter panduan dasar (*baseline*) objektif yang jauh lebih andal dibandingkan insting tebakan manual. Pencapaian ini memastikan pihak manajerial memiliki landasan antisipatif untuk mereduksi risiko krisis kekosongan barang secara sistematis. Kesimpulan kedua berkaitan dengan implementasi model *regresi ensemble Extreme Gradient Boosting* (XGBoost) yang dihela untuk mengatasi inefisiensi dan subjektivitas pada penganggaran armada distribusi logistik harian.

Algoritma ini memvalidasi keandalannya dengan meraih tingkat kepresisian struktural Koefisien Determinasi (R^2) yang sangat superior sebesar 95,5%. Melalui konversi variabel matriks jarak rute dan kapasitas muatan, model membukukan margin penyimpangan rata-rata (MAE) yang sangat minim, yakni hanya sebesar Rp12.053,21. Tingginya ketajaman komputasi ini menegaskan bahwa XGBoost sangat efektif dalam mentransformasi perencanaan anggaran ekspedisi konvensional menjadi sebuah ekosistem keputusan yang presisi, transparan, dan sepenuhnya berbasis data (data-driven).

Adapun upaya peningkatan dan perbaikan pada bidang studi yang sama untuk penelitian masa depan diharapkan mampu mentransformasi pilar peramalan univariat yang digunakan saat ini menjadi kerangka prediktif multivariat, misalnya penerapan algoritma SARIMAX maupun Long Short-Term Memory (LSTM). Integrasi algoritma multivariat akan membuka peluang bagi model untuk menyerap variabel intervensi pasar eksternal seperti fluktuasi indeks komoditas nasional, efek musim panen, dan anomali momen perayaan keagamaan. Di samping itu, otomatisasi mekanisme penyesuaian parameter (auto-retraining pipeline) serta integrasi pencatatan seluler pintar (mobile native) yang disematkan langsung pada supir armada logistik sangat disarankan guna menghasilkan data analitik waktu nyata yang jauh lebih murni dan utuh.

DAFTAR PUSTAKA

- [1] J. Vicky, P. Putra, U. Mahdiah, and A. Sanjaya, "Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) 929 Penggunaan Metode Trend Moment Untuk Proses Peramalan Jumlah Stok Penjualan Snack," Online, 2023.
- [2] G. Aprilia Putri Agita, R. Aswi Ramadhani, and A. Bagus Setiawan, "Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) 2025," vol. 9, pp. 2549–7952, 2025.

- [3] C. Coronel and S. Morris, *Database Systems: Design, Implementation, & Management*. Boston: Cengage Learning, 2020.
- [4] B. D. Fadillah and N. Hendrastuty, "Prediksi Stok Barang di Toko Eko Helm Menggunakan Metode Time series Analysis," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 2, pp. 278–291, Apr. 2025, doi: 10.30591/jpit.v10i2.8584.
- [5] Muhammad Rizky Nurhambali, M. R. Angraini, and Y. Fitrianto, "Simulation Study to Identify Factors Affecting the Performance of LSTM and XGBoost for Anomaly Detection on Labeled Time Series Data," *JUITA: Jurnal Informatika*, vol. 13, no. 2, pp. 219–228, 2025, doi: 10.30595/juita.v13i2.26604.
- [6] A. Yaqin and G. Ramadhani, "Penilaian Kredit Menggunakan Algoritma XGBoost dan Logistic Regression," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 8, no. 1, 2023.
- [7] W. Nurlela *et al.*, "Analisis Metode Moving Average, Exponential Smoothing, dan Arima dalam Peramalan Permintaan untuk Pengendalian Stok Floor Rear (Studi Kasus : PT. SAI)," *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, vol. 4, no. 3, pp. 1066–1075, 2025.
- [8] G. Box, G. M. , Jenkins, G. C. , Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, New Jersey, USA, 2023.
- [9] K. Dyansyah, A. B. Setiawan, and P. Kasih, "Penerapan Estimasi Pose dengan Model MoveNet," 2025.
- [10] A. A. Saputra, B. N. Sari, C. Rozikin, U. Singaperbangsa, and K. Abstrak, "Penerapan Algoritma Extreme Gradient Boosting (Xgboost) Untuk Analisis Risiko Kredit," *Jurnal Ilmiah Wahana Pendidikan*, vol. 10, no. 7, pp. 27–36, 2024, doi: 10.5281/zenodo.10960080.