

Implementasi *Bayesian Optimization* untuk Meningkatkan Akurasi Prediksi Penjualan Menggunakan Algoritma *XGBoost* (Studi Kasus: CV XYZ)

Nendi Setiawan¹, Hadi Zakaria^{2*}

Fakultas Ilmu Komputer, Program Studi Teknik Informatika
Universitas Pamulang
dosen00274@unpam.ac.id

Abstrak

CV. XYZ merupakan perusahaan percetakan daring yang menghadapi tantangan dalam memprediksi penjualan akibat fluktuasi permintaan dan keterbatasan metode analisis yang digunakan. Prediksi penjualan yang tidak akurat dapat menyebabkan ketidakseimbangan produksi dan menurunkan efisiensi operasional. Penelitian ini bertujuan untuk meningkatkan akurasi prediksi penjualan dengan mengimplementasikan model Extreme Gradient Boosting (XGBoost) yang dioptimalkan menggunakan Bayesian Optimization. XGBoost dipilih karena kemampuannya dalam menangani data kompleks dan nonlinier, sedangkan Bayesian Optimization digunakan untuk menyempurnakan pemilihan hiperparameter secara efisien. Kebaruan penelitian ini terletak pada integrasi Bayesian Optimization sebagai strategi tuning hiperparameter XGBoost untuk prediksi penjualan pada sektor percetakan digital. Data historis penjualan dianalisis menggunakan bahasa pemrograman Python dengan SQLite sebagai basis penyimpanan data. Hasil eksperimen menunjukkan bahwa kombinasi kedua metode tersebut mampu meningkatkan performa prediksi secara signifikan, dengan nilai RMSE sebesar 8.366,6, MAE sebesar 8.000, dan skor R^2 sebesar 0,9748. Pendekatan ini memberikan solusi praktis bagi CV. XYZ dalam meningkatkan perencanaan produksi serta pengambilan keputusan berbasis data di sektor percetakan digital.

Kata kunci: prediksi penjualan, XGBoost, Bayesian Optimization, machine learning.

Abstract

CV. XYZ is an online printing company that faces challenges in predicting sales due to demand fluctuations and limitations in the analytical methods employed. Inaccurate sales predictions can lead to production imbalances and reduced operational efficiency. This study aims to improve sales prediction accuracy by implementing an Extreme Gradient Boosting (XGBoost) model optimized using Bayesian Optimization. XGBoost is selected for its capability to handle complex and non-linear data, while Bayesian Optimization is employed to efficiently optimize hyperparameter selection. The novelty of this study lies in the integration of Bayesian Optimization as a hyperparameter tuning strategy for the XGBoost model in sales prediction within the digital printing sector. Historical sales data are analyzed using Python, with SQLite utilized as the data storage system. Experimental results demonstrate that the proposed approach significantly improves prediction performance, achieving an RMSE of 8,366.6, an MAE of 8,000, and an R^2 score of 0.9748. This approach provides a practical solution for CV. XYZ in enhancing production planning and data-driven decision-making in the digital printing industry.

Keywords: sales forecasting, XGBoost, Bayesian Optimization, machine learning.

PENDAHULUAN

CV. XYZ merupakan perusahaan percetakan digital yang melayani

pemesanan dan pembayaran secara daring di berbagai wilayah Indonesia. Dalam

operasionalnya, perusahaan menghadapi kesulitan dalam memprediksi volume penjualan secara akurat akibat meningkatnya jumlah transaksi serta fluktuasi permintaan pelanggan. Ketidaktepatan dalam melakukan prediksi penjualan sering kali menyebabkan penumpukan stok, kekurangan bahan baku, serta pemborosan waktu dan biaya produksi, yang berdampak pada menurunnya efisiensi operasional perusahaan. Selain itu, ketiadaan sistem prediksi berbasis data menyebabkan CV. XYZ masih mengandalkan metode perkiraan manual yang rentan terhadap kesalahan dan kurang mampu menangkap pola penjualan yang bersifat kompleks dan dinamis. Kondisi ini menghambat perusahaan dalam menyesuaikan strategi produksi terhadap perubahan tren pasar secara cepat, sehingga pengambilan keputusan menjadi kurang optimal dan berpotensi menimbulkan kerugian finansial. Untuk mengatasi permasalahan tersebut, diperlukan penerapan pendekatan prediksi penjualan berbasis pembelajaran mesin yang mampu mengolah data historis secara efektif. Salah satu metode yang memiliki kinerja baik dalam menangani data kompleks adalah Extreme Gradient Boosting (XGBoost), namun performanya sangat bergantung pada pemilihan hiperparameter yang tepat. Oleh karena itu, penelitian ini menerapkan Bayesian Optimization sebagai strategi optimasi hiperparameter guna meningkatkan akurasi prediksi penjualan. Kebaruan penelitian ini terletak pada integrasi Bayesian Optimization sebagai strategi tuning hiperparameter model XGBoost untuk prediksi penjualan pada sektor percetakan digital, yang sebelumnya belum banyak diterapkan secara spesifik pada konteks

tersebut. Pendekatan yang diusulkan diharapkan mampu memberikan solusi praktis dalam meningkatkan perencanaan produksi, efisiensi operasional, serta pengambilan keputusan berbasis data di CV. XYZ.

Implementasi mengacu pada pelaksanaan suatu rencana, kebijakan, atau keputusan untuk mencapai sasaran tertentu. Tahapan ini melibatkan sejumlah tindakan, termasuk perencanaan, distribusi sumber daya, serta eksekusi, untuk menjamin bahwa ide-ide teoritis dapat diwujudkan dalam bentuk praktik. Proses kompleks ini memainkan peranan penting di berbagai bidang, seperti pendidikan, pemerintahan, dan inisiatif komunitas. Bagian-bagian berikut akan menjelaskan poin-poin utama dari implementasi (Junaini Muhamad et al., 2023).

XGBoost, atau Extreme Gradient Boosting, merupakan suatu algoritma untuk pembelajaran mesin yang dikembangkan sebagai peningkatan dari metode Gradient Boosting. Algoritma ini dibuat untuk efisiensi, fleksibilitas, dan tingkat akurasi yang tinggi dalam pemodelan prediktif. Cara kerja XGBoost melibatkan pembuatan model secara bertahap, dengan setiap model baru berusaha untuk memperbaiki kesalahan yang ada pada model sebelumnya. XGBoost mengedepankan pendekatan berbasis pohon keputusan dan menggunakan teknik regularisasi untuk mencegah terjadinya overfitting, sehingga menjadikannya sangat handal dalam kompetisi di bidang data science (Dharmawan & Sanjaya ER, 2024).

Bayesian Optimization adalah strategi desain sekuensial untuk optimisasi global terhadap black-box function, yang tidak mengasumsikan bentuk fungsi tertentu dan digunakan ketika evaluasi fungsi tersebut

mahal atau memakan waktu terlalu lama. Metode ini membangun model surrogate—biasanya Gaussian Process—untuk memprediksi nilai dan ketidakpastian fungsi. Berdasarkan model ini, fungsi akuisisi digunakan untuk menentukan titik pengamatan selanjutnya dengan menyeimbangkan antara eksplorasi dan eksploitasi (Putra et al., 2023).

Prediksi penjualan adalah perkiraan penjualan masa depan untuk produk di berbagai bidang, penting untuk kesuksesan bisnis. Mereka memungkinkan perencanaan proaktif berdasarkan permintaan yang diantisipasi, meskipun prediksi dapat membawa bias yang melekat dan berbagai tingkat akurasi tergantung pada model yang dipilih (Xu, 2024).

Berdasarkan permasalahan yang dihadapi, ketidakakuratan dalam memprediksi penjualan berpotensi menimbulkan inefisiensi dalam proses produksi dan pengelolaan persediaan. Untuk menjawab tantangan tersebut, penelitian ini mengusulkan pengembangan sistem prediksi penjualan berbasis web dengan menerapkan model Extreme Gradient Boosting (XGBoost). Model ini dioptimalkan menggunakan metode Bayesian Optimization untuk memperoleh konfigurasi parameter terbaik guna meningkatkan akurasi prediksi. Sistem dirancang untuk menganalisis data historis transaksi, termasuk variabel seperti waktu pemesanan, jenis produk, jumlah pesanan, dan pola pembelian pelanggan. Melalui pendekatan ini, perusahaan diharapkan mampu memperkirakan permintaan pasar secara lebih tepat, meminimalkan risiko kelebihan maupun kekurangan stok, serta mendukung efisiensi dalam perencanaan operasional.

Rumusan Masalah

Berdasarkan identifikasi masalah diatas, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana merancang sistem prediksi penjualan yang mampu menghasilkan estimasi secara lebih akurat dan mencerminkan kondisi pasar secara aktual?
2. Bagaimana memanfaatkan volume data transaksi yang terus meningkat untuk mengidentifikasi pola pembelian pelanggan secara menyeluruh?
3. Bagaimana mengimplementasikan mekanisme optimasi model prediktif agar dapat meningkatkan akurasi hasil prediksi penjualan secara signifikan?

Tujuan Penelitian

Berdasarkan batasan masalah yang telah ditetapkan, tujuan penelitian ini adalah:

1. Mengembangkan sistem prediksi penjualan berbasis pembelajaran mesin yang mampu meningkatkan kecepatan dan akurasi prediksi guna mendukung perencanaan produksi dan pengambilan keputusan di CV. XYZ.
2. Menerapkan model Extreme Gradient Boosting (XGBoost) yang dioptimalkan dengan Bayesian Optimization untuk mengolah dan menganalisis data penjualan secara efektif.
3. Menganalisis kontribusi penerapan teknik pembelajaran mesin dalam meningkatkan kinerja prediksi penjualan pada sektor bisnis lokal, khususnya pada perusahaan percetakan digital berbasis daring.

Manfaat Penelitian

1) Bagi Penulis

1. Meningkatkan pemahaman dan pengalaman dalam penerapan algoritma XGBoost serta optimasi hiperparameter menggunakan Bayesian Optimization pada kasus prediksi penjualan.
2. Mengembangkan kompetensi di bidang data science, analisis prediktif, dan sistem informasi bisnis digital.

2) Bagi Universitas

1. Memberikan kontribusi akademik dalam pengembangan penelitian terkait pembelajaran mesin dan sistem prediksi berbasis data.
2. Menjadi referensi bagi penelitian selanjutnya, khususnya pada bidang prediksi penjualan dan UMKM berbasis digital.

3) Bagi CV. XYZ

1. Menyediakan sistem prediksi penjualan berbasis data untuk mendukung perencanaan produksi yang lebih efisien.
2. Membantu perusahaan mengantisipasi fluktuasi permintaan dan meningkatkan efisiensi operasional serta daya saing.

Tinjauan Pustaka

1) (Dankorpo, 2024)

Penelitian ini membandingkan metode peramalan tradisional dengan algoritma XGBoost pada data penjualan ritel tahun 2019–2023. Hasilnya menunjukkan bahwa XGBoost memberikan peningkatan akurasi dengan penurunan galat sebesar 29,23% (MAE) dan 34,54% (RMSE) dibandingkan metode konvensional. Selain itu, model ini memungkinkan pergeseran dari peramalan

bulanan ke harian, sehingga meningkatkan efisiensi pengambilan keputusan dan pengelolaan persediaan.

2) (Khurshid et al., 2025)

Dalam bidang prediksi medis, penelitian ini menggunakan XGBoost yang dioptimalkan dengan Bayesian Optimization untuk menangani ketidakseimbangan data penyakit diabetes. Hasilnya menunjukkan peningkatan akurasi menjadi 97,26%, F1-score 95,72%, dan MCC 81,18%, sedikit lebih baik dibandingkan grid search. Pendekatan ini terbukti efisien dalam tuning parameter serta meningkatkan akurasi prediksi pada data kompleks.

3) (Shi & Xu, 2020)

Penelitian ini menggabungkan XGBoost dan Bayesian Optimization untuk memprediksi keterlambatan kedatangan kereta. Dengan mempertimbangkan 11 variabel independen, hasil menunjukkan bahwa kombinasi metode ini menghasilkan akurasi lebih tinggi dibandingkan Random Forest, Deep Extreme Learning Machine, dan Gradient Boosting Regression Trees. Pendekatan ini efektif dalam menangani permasalahan prediktif dengan karakteristik sistem yang dinamis.

4) (Yan & Zhou, 2023)

Dalam konteks prediksi jumlah penumpang bus, penelitian ini menerapkan XGBoost yang dioptimalkan menggunakan Bayesian Optimization serta metode interpretasi SHAP. Hasilnya menunjukkan model BO-XGBoost unggul dalam akurasi dan stabilitas dibandingkan SVM, KNN, dan Random Forest, sekaligus lebih efisien daripada metode tuning acak.

Penelitian terdahulu menunjukkan bahwa algoritma XGBoost memiliki keunggulan dalam meningkatkan akurasi prediksi pada berbagai studi kasus. Namun,

mayoritas penelitian masih menggunakan pendekatan tuning hiperparameter yang kurang efisien, seperti grid search atau penyesuaian manual, yang berpotensi menghasilkan parameter suboptimal. Meskipun Bayesian Optimization telah mulai digunakan sebagai alternatif optimasi, penerapannya dalam prediksi penjualan pada usaha kecil dan menengah, khususnya di sektor percetakan digital, masih terbatas. Oleh karena itu, penelitian ini mengintegrasikan Bayesian Optimization dengan algoritma XGBoost untuk prediksi penjualan berbasis data historis pada studi kasus CV. XYZ.

LANDASAN TEORI

Implementasi

Implementasi adalah tindakan menerapkan mode yang harus diperhatikan untuk mengimplementasikan kebijakan lembaga pengembangan, memastikan proses mencapai hasil yang diinginkan. (Humaira et al., 2024)

Menurut Sudarsono dalam bukunya “Analisis Kebijakan Publik”, implementasi adalah suatu kegiatan yang berkaitan dengan penyelesaian suatu pekerjaan, melalui penggunaan sarana (tools) untuk memperoleh hasil akhir yang diinginkan. Begitu juga menurut Daniel A. Mazmanian dan Paul Sabatier, konsep implementasi merupakan pemahaman yang terjadi setelah penyusunan rencana yang menjadi fokus implementasi kebijakan rancangan pemerintah.

XGBoost

XGBoost adalah ansambel pohon keputusan yang dapat diskalakan berdasarkan peningkatan gradien, dirancang untuk efisiensi dan kecepatan. Ini meminimalkan fungsi kerugian sambil

mengendalikan kompleksitas pohon melalui variasi fungsi kerugian. Sebaliknya, peningkatan gradien tradisional membangun ekspansi aditif dari fungsi tujuan tanpa tingkat pengoptimalan dan skalabilitas yang sama. XGBoost memungkinkan penyetelan parameter halus dengan algoritma yang efisien secara komputasi, sedangkan peningkatan gradien biasanya membutuhkan waktu komputasi yang lebih lama untuk penyetelan (Yan & Zhou, 2023).

Algoritma

Pengertian algoritma sangat lekat dengan kata logika, yaitu kemampuan seseorang manusia untuk berfikir dengan akal tentang suatu permasalahan menghasilkan sebuah kebenaran, dibuktikan dan dapat diterima akal, logika seringkali dihubungkan dengan kecerdasan, seseorang yang mampu berlogika dengan baik sering orang menyebutnya, sebagai pribadi yang cerdas. Dalam menyelesaikan suatu masalahpun logika mutlak diperlukan (Purnama Indah, 2024).

Bayesian Optimization

Bayesian Optimization adalah suatu metodologi optimisasi sekuensial untuk fungsi objektif ‘kotak hitam’ (black-box) yang mahal untuk dievaluasi. Metode ini menggunakan model surrogate umumnya Gaussian Process sebagai representasi probabilistik terhadap fungsi objektif, yang kemudian dimanfaatkan oleh ‘acquisition function’ untuk memilih titik pengamatan berikutnya. Dengan demikian, eksplorasi dan eksploitasi dapat diatur secara efisien untuk menemukan nilai optimal dengan jumlah evaluasi yang minimal (Garnett Roman, 2023).

- 1) Bayesian Optimization terdiri dari dua komponen utama: Model Probabilistik dari Fungsi Tujuan (Objective Function) Digunakan untuk memprediksi hasil dari parameter-parameter yang belum dicoba. Biasanya menggunakan Gaussian Process (semacam fungsi yang memetakan parameter ke nilai error).
- 2) Fungsi Akuisisi (Acquisition Function) Digunakan untuk menentukan parameter mana yang akan dicoba selanjutnya, berdasarkan prediksi model probabilistik.

Akurasi

Akurasi mengacu pada sejauh mana hasil kerja sesuai dengan data atau informasi yang benar. Hasil yang akurat mengurangi kemungkinan kesalahan. Pentingnya: Akurasi sangat penting dalam pekerjaan yang melibatkan angka, data, atau informasi penting lainnya, seperti laporan keuangan atau analisis data (Rachman Azis et al., 2024).

Prediksi Penjualan

Menurut Tambun Rosmaida dalam (Yunita Irma, 2022). Prediksi penjualan adalah teknis perkiraan permintaan konsumen dengan menggunakan data dan asumsi – asumsi tertentu dalam memproyeksikan penjualann pada periode anggaran yang teal ditentukan. Proyeksi penjualan in merupakan suatu penilaian (judgement) dari manajemen mengenai permintaan konsumen yang potensial dalam kurun waktu tertentu.

Machine Learning

Machine Learning adalah bagian dari cabang artificial intelligence yang didefinisikan sebagai suatu pendekatan yang memungkinkan sistem untuk belajar

dari pengalaman mereka dan beradaptasi dengan data baru, sehingga mereka dapat mengambil tindakan secara otomatis dan membuat keputusan yang tepat. Dengan bantuan Machine Learning, sistem dapat mempelajari data secara mandiri dan menemukan pola-pola yang relevan dalam data tersebut. Ini membuat proses pengambilan keputusan lebih efisien.

Menurut (Setiawan, 2021). Machine learning adalah bidang AI yang menggunakan metode statistik untuk memproses data dan membuat informasi. Sistem belajar dari data yang disediakan dengan mengenali fitur-fitur penting dalam data. Secara umum, dalam Machine learning, data dibagi menjadi dua bagian: data pelatihan dan data uji. Data pelatihan digunakan oleh sistem untuk mempelajari model dan fungsi, yang merupakan langkah penting dalam pemrosesan data. Data uji kemudian digunakan untuk memeriksa kapasitas sistem dengan mengenali model - model ini dan mengklasifikasikannya dalam jenis tertentu sesuai dengan perbandingan fungsi antara pelatihan dan pengujian data.. Pendekatan ini memungkinkan Machine learning untuk memahami dan memproses data secara mandiri.

Demikian juga menurut (Yessy Asri, 2024). bahwa machine learning sebagai teknologi inti dalam kecerdasan buatan yang berfungsi untuk melatih komputer agar mampu mempelajari data dan mengambil keputusan secara otomatis. Teknologi ini dinilai sangat efektif dalam mengelola data kompleks, seperti ulasan pengguna aplikasi, guna memperoleh wawasan yang berguna.

METODE PENELITIAN

Dalam mengimplementasikan sistem prediksi penjualan menggunakan algoritma XGBoost yang dioptimalkan dengan Bayesian Optimization, serta membangun aplikasi berbasis web yang mendukung pengambilan keputusan penjualan, penulis menggunakan beberapa metode penelitian sebagai berikut:

Metode Observasi

Observasi dilakukan secara langsung terhadap proses bisnis dan alur penjualan di CV Printnesia Digital Printing. Tujuan dari metode ini adalah untuk memahami kebutuhan perusahaan dalam melakukan prediksi penjualan, termasuk tantangan yang dihadapi dalam menyusun estimasi permintaan dan pengelolaan stok.

Metode Wawancara

Wawancara dilakukan dengan pihak terkait, termasuk staf operasional dan manajer penjualan di CV Printnesia Digital Printing, salah satunya Ibu Ayu Zulaikha, S.I.Kom. Tujuannya adalah untuk menggali informasi lebih dalam terkait sistem pencatatan penjualan yang digunakan, proses pengambilan keputusan, serta kendala dalam menghadapi fluktuasi permintaan pasar.

Studi Literatur

Studi literatur dilakukan untuk mendalami teori-teori dan pendekatan yang relevan dengan penelitian ini, termasuk algoritma XGBoost, teknik optimasi hiperparameter menggunakan Bayesian Optimization, serta penerapan metode prediksi penjualan pada sektor industri digital printing dan e-commerce. Referensi diperoleh dari jurnal ilmiah, buku teks,

artikel penelitian, serta dokumentasi dari sumber terpercaya.

ANALISIS DAN PERANCANGAN

Analisa Sistem

Analisis Sistem adalah langkah untuk menyelidiki dan menilai sistem yang sedang beroperasi dengan maksud untuk memahami cara kerja sistem tersebut, menemukan masalah yang ada, serta merumuskan kebutuhan untuk sistem baru atau perbaikan dari sistem yang sudah ada. Langkah ini adalah fase awal yang sangat krusial dalam pengembangan sistem informasi karena menjadi landasan dalam menetapkan cara sistem akan dirancang dan dibangun. (Roji Fatkhur Mukhamad, 2025).

Analisa Sistem Saat ini

Proses prediksi penjualan di CV. XYZ masih bersifat manual dan mengandalkan metode statistik sederhana, sehingga kurang mampu menangkap dinamika pasar secara akurat. Meskipun perusahaan memiliki volume data transaksi yang besar, belum ada sistem yang memanfaatkan data tersebut untuk mengenali pola pembelian pelanggan secara menyeluruh. Tidak adanya sistem yang mampu menyesuaikan parameter secara otomatis menyebabkan estimasi penjualan cenderung tidak akurat dan tidak adaptif terhadap perubahan pasar. Hal ini mengindikasikan perlunya implementasi sistem prediktif berbasis machine learning yang terotomatisasi dan dioptimalkan guna meningkatkan akurasi dan kecepatan pengambilan keputusan bisnis.

Analisa Sistem Usulan

Dalam rangka mengatasi keterbatasan sistem manual yang saat ini digunakan oleh CV. XYZ, penulis mengusulkan sebuah

sistem prediksi penjualan berbasis machine learning yang memanfaatkan algoritma Extreme Gradient Boosting (XGBoost). Sistem ini dirancang untuk secara otomatis memproses data historis penjualan dan menghasilkan prediksi yang lebih akurat dan adaptif terhadap perubahan pola pembelian konsumen.

Metode XGBoost dipilih karena kemampuannya yang unggul dalam menangani data berskala besar, performa klasifikasi dan regresi yang tinggi, serta toleransi terhadap ketidakseimbangan data (imbalanced data). Untuk mengoptimalkan kinerja algoritma tersebut, sistem ini juga mengimplementasikan teknik Bayesian Optimization. Teknik ini digunakan dalam proses hyperparameter tuning, yaitu pencarian nilai parameter model yang paling optimal untuk meminimalkan kesalahan prediksi. Berbeda dengan metode pencarian konvensional seperti grid search atau random search, Bayesian Optimization bekerja secara efisien dengan cara memperkirakan kombinasi parameter terbaik berdasarkan hasil eksperimen sebelumnya.

Hasil dari proses prediksi ini akan disajikan dalam bentuk numerik dan visual, sehingga dapat memberikan wawasan yang informatif bagi pihak manajemen dalam menyusun strategi penjualan. Selain itu, sistem ini juga memungkinkan untuk dilakukan pelatihan ulang (retraining) secara berkala jika terdapat data baru, sehingga model tetap relevan dan adaptif terhadap dinamika pasar yang terus berubah. Dengan penerapan sistem usulan ini, diharapkan proses prediksi penjualan yang sebelumnya bersifat subjektif dan tidak terukur dapat digantikan oleh sistem yang lebih objektif, terukur, dan mampu

meningkatkan efektivitas pengambilan keputusan dalam operasional bisnis.

Perancangan Model Konseptual

Perancangan Basis Data

Perancangan basis data dilakukan untuk menggambarkan struktur data yang digunakan dalam sistem secara logis dan terorganisir. Tujuannya adalah memastikan setiap data yang disimpan dapat saling terhubung dan mendukung proses prediksi penjualan secara efektif.

Dalam perancangan ini, ditentukan sejumlah entitas utama seperti Customer, Produk, Transaksi, Hasil Prediksi, dan Admin. Setiap entitas memiliki atribut-atribut yang saling berkaitan, misalnya entitas Transaksi menghubungkan data Customer dengan Produk melalui informasi tanggal, jumlah pesanan, dan total pembelian. Hubungan antarentitas dijelaskan melalui model konseptual berupa Entity Relationship Diagram (ERD) dan kemudian diturunkan menjadi Logical Record Structure (LRS) untuk implementasi pada database relasional. Proses normalisasi juga diterapkan untuk menghindari redundansi data dan menjaga integritas informasi pada sistem.

Perancangan UML.

Perancangan sistem secara konseptual digambarkan menggunakan pendekatan UML, yang menjelaskan bagaimana sistem bekerja dan bagaimana setiap komponen saling berinteraksi.

- 1) Use Case Diagram digunakan untuk menjelaskan interaksi antara aktor dan sistem, di mana aktor utama adalah admin yang dapat melakukan login, mengelola data, melakukan pelatihan model, dan melihat hasil prediksi.

- 2) Activity Diagram menggambarkan alur aktivitas dari proses prediksi penjualan, mulai dari input data transaksi hingga keluarnya hasil prediksi oleh model XGBoost.
- 3) Sequence Diagram menunjukkan urutan interaksi antara komponen sistem, termasuk proses pengiriman data ke model, eksekusi prediksi, dan penyimpanan hasil ke database.
- 4) Terakhir, Class Diagram mendeskripsikan struktur kelas dalam sistem beserta atribut dan metode yang digunakan, seperti kelas Transaksi, ModelPrediksi, dan HasilPrediksi.

Implikasi Perancangan

Perancangan konseptual yang disusun pada penelitian ini memiliki implikasi penting terhadap efektivitas pengembangan sistem prediksi penjualan berbasis algoritma XGBoost yang dioptimalkan dengan Bayesian Optimization. Melalui tahapan analisa sistem, perancangan basis data, dan pemodelan UML, penelitian ini tidak hanya menghasilkan rancangan sistem yang terstruktur tetapi juga mampu memastikan bahwa setiap komponen sistem saling terhubung secara logis.

Implikasi pertama terlihat pada peningkatan akurasi dan konsistensi data, karena rancangan basis data yang disusun secara terormalisasi meminimalkan duplikasi dan kesalahan input, sehingga data historis penjualan dapat digunakan dengan lebih andal untuk pelatihan model. Kedua, melalui perancangan UML, sistem mampu menggambarkan alur proses prediksi secara menyeluruh — mulai dari input transaksi hingga keluarnya hasil prediksi — yang memastikan integrasi yang baik antara antarmuka pengguna, modul prediksi, dan database.

Selain itu, rancangan konseptual ini juga memiliki implikasi terhadap kemudahan implementasi dan pemeliharaan sistem. Dengan adanya representasi model yang jelas, pengembang dapat menerjemahkan rancangan ke dalam bentuk implementasi teknis secara efisien tanpa kehilangan konteks bisnis yang telah dianalisis sebelumnya. Implikasi lainnya adalah peningkatan efisiensi operasional perusahaan, karena sistem yang dirancang mampu mendukung pengambilan keputusan berbasis data, mempercepat proses perencanaan produksi, serta mengurangi risiko overstock maupun understock.

Secara keseluruhan, perancangan konseptual ini menjadi dasar penting bagi keberhasilan tahap implementasi, karena menjembatani antara analisa kebutuhan bisnis dengan penerapan algoritma pembelajaran mesin dalam sistem prediksi penjualan yang adaptif dan akurat.

Perhitungan Manual Evaluasi Model Algoritma XGBoost.

Data Hasil Prediksi

Tabel 1. Hasil Prediksi Perhitungan Berdasarkan Rumus:

Metric	Value	Description
MAE	8,000	Menunjukkan rata-rata kesalahan yang rendah antara nilai penjualan yang diprediksi dan actual.
RMSE	8,366.6	Menunjukkan kinerja prediksi yang stabil dan konsisten.
R²	0.9748	Sebesar 97,48% variasi penjualan dapat dijelaskan oleh model.

a MAE

Hitung selisih absolut:

$$|230,000 - 220,000| = 10,000$$

$$|180,000 - 175,000| = 5,000$$

$$|250,000 - 260,000| = 10,000$$

$$|300,000 - 295,000| = 5,000$$

$|150,000 - 140,000| = 10,000$
 Total: $10,000 + 5,000 + 10,000 + 5,000$
 $+ 10,000 = 40,000$
 $MAE = 40,000 / 5 = 8,000$

TP (Tinggi → Tinggi): 51
 TN (Rendah → Rendah): 1187
 FP (Rendah → Tinggi): 19
 FN (Tinggi → Rendah): 30

b RMSE

Hitung kuadrat error:
 $(-10,000)^2 = 100,000,000$
 $(-5,000)^2 = 25,000,000$
 $(+10,000)^2 = 100,000,000$
 $(-5,000)^2 = 25,000,000$
 $(-10,000)^2 = 100,000,000$
 Total = $350,000,000$
 $MSE = 350,000,000 / 5 = 70,000,000$
 $RMSE = \sqrt{70,000,000} \approx 8,366.6$

3. R2 Score (Koefisien Detetminasi):

Dengan:

$SS_{res} = \sum (y_i - \hat{y}_i)^2 = 350,000,000$ (dari RMSE tadi)

Hitung \bar{y} (rata-rata y_{test}):

$Y = 230000 + 180000 + 250000 + 300000 + 150000 \ / \ 5 = 222.000$

Hitung $SS_{tot} = \sum (y_i - \bar{y})^2$:

$(230,000 - 222,000)^2 = 64,000,000$
 $(180,000 - 222,000)^2 = 1,764,000,000$
 $(250,000 - 222,000)^2 = 784,000,000$
 $(300,000 - 222,000)^2 = 6,084,000,000$
 $(150,000 - 222,000)^2 = 5,184,000,000$

Total $SS_{tot} \approx 13,880,000,000$

$R2 = 1 - 350,000,000 / 13,880,000,000$
 $= 1 - 0.0252 = 0.9748$

1. Confusion Matrix



Gambar 1. Confusion Matrix — Classification of Sales Status (Laku / Tidak Laku)

Perhitungan Matrix Evaluasi:

a) Akurasi (Accuracy)

$Accuracy = (TP + TN) / (TP + TN + FP + FN)$
 $= (51 + 1187) / (51 + 1187 + 19 + 30)$
 $= 1238 / 1287 \approx 0.9619$ atau 96.19%

b) Presisi (Precision)

$Precision = TP / (TP + FP)$
 $= 51 / (51 + 19)$
 $= 51 / 70 \approx 0.7286$ atau 72.86%

c) Recall (Sensitivity)

$Recall = TP / (TP + FN)$
 $= 51 / (51 + 30)$
 $= 51 / 81 \approx 0.6296$ atau 62.96%

d) F1-Score

$F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall)$
 $= 2 \times (0.7286 \times 0.6296) / (0.7286 + 0.6296)$
 $= 2 \times 0.4586 / 1.3582$
 $= 0.9172 / 1.3582 \approx 0.6751$ atau 67.51%

Pembahasan

Hasil evaluasi pada Tabel 1 menunjukkan bahwa model XGBoost yang dioptimalkan menggunakan Bayesian Optimization berhasil mencapai tingkat akurasi yang tinggi. Nilai MAE sebesar 8.000 menunjukkan bahwa rata-rata penyimpangan antara nilai penjualan yang diprediksi dan aktual tergolong kecil. Nilai RMSE sebesar 8.366,6 mengindikasikan bahwa performa prediksi model stabil dan konsisten pada berbagai dataset.

Selain itu, koefisien determinasi (R^2) sebesar 0,9748 menunjukkan bahwa sekitar 97,48% variasi penjualan dapat dijelaskan oleh model. Nilai yang tinggi ini membuktikan bahwa model XGBoost yang telah dioptimalkan mampu mempelajari pola penjualan dengan efektif serta

menghasilkan prediksi yang mendekati kondisi bisnis sebenarnya.

Kesimpulan

Sistem prediksi yang diusulkan, yaitu kombinasi antara XGBoost dan Bayesian Optimization, terbukti berhasil meningkatkan akurasi peramalan penjualan pada CV. XYZ.

Temuan utama dari penelitian ini meliputi:

- a. Performa prediksi yang tinggi ($R^2 = 0,9748$).
- b. Pengurangan kesalahan estimasi manual melalui pemilihan parameter yang teroptimasi.
- c. Pengembangan platform berbasis web yang ramah pengguna dan mendukung pengambilan keputusan berbasis data.

Hasil ini menunjukkan bahwa penggabungan teknik machine learning dengan metode optimisasi memiliki potensi besar dalam meningkatkan kemampuan peramalan di bidang digital printing maupun industri sejenis lainnya.

Kemudian Gambar 1. menggambarkan kinerja klasifikasi model dalam mengidentifikasi status penjualan produk menggunakan confusion matrix. Hasil pengujian menunjukkan terdapat 1187 True Negative (produk yang benar diprediksi tidak terjual) dan 51 True Positive (produk yang benar diprediksi terjual).

Kesalahan klasifikasi yang terjadi relatif sedikit, yaitu 19 False Positive (produk yang tidak terjual namun diprediksi terjual) dan 30 False Negative (produk yang terjual namun diprediksi tidak terjual).

Hasil ini menunjukkan bahwa model mampu mengenali pola penjualan dengan baik, terutama dalam mengidentifikasi produk yang tidak terjual. Tingkat akurasi klasifikasi yang tinggi juga sejalan dengan

nilai R^2 yang kuat pada Tabel 1, sehingga dapat disimpulkan bahwa model XGBoost yang dioptimalkan dengan Bayesian Optimization memiliki kinerja prediktif dan klasifikasi yang andal serta konsisten.

KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan Bayesian Optimization pada algoritma XGBoost mampu meningkatkan kinerja prediksi penjualan dibandingkan konfigurasi parameter standar, sebagaimana ditunjukkan oleh perbaikan nilai MAE, RMSE, dan koefisien determinasi (R^2). Optimasi hiperparameter terbukti berperan penting dalam meningkatkan kemampuan model dalam menangkap pola data historis penjualan secara lebih akurat. Namun, penelitian ini masih terbatas pada satu studi kasus dan penggunaan data historis tertentu, sehingga penelitian selanjutnya disarankan untuk menguji pendekatan ini pada dataset yang lebih beragam dan konteks bisnis yang berbeda guna meningkatkan generalisasi dan validitas hasil.

DAFTAR PUSTAKA

- Dankorpho, P. (2024). *Sales forecasting for retail business using XGBoost algorithm*. *Journal of Computer Science and Technology Studies*, 6(2), 136–141. <https://doi.org/10.32996/JCSTS.2024.6.2.15>
- Garnett, R. (2023). *Bayesian Optimization*. Google Books. https://www.google.co.id/books/editio n/Bayesian_Optimization/8d2iEAAAQ BAJ
- Junaini, M., Utomo, P. D., & Ali, M. (2023). *Implementasi Program Lombok Timur Berkembang (Berantas Rentenir Melalui Kredit Tanpa Bunga) pada UMKM di Kecamatan Selong*.

- Jurnal Pendidikan Ekonomi dan Kewirausahaan (JPEK), 7(2). <https://doi.org/10.29408/JPEK.V7I2.24736>
- Khurshid, M. R., Manzoor, S., Sadiq, T., Hussain, L., Khan, M. S., & Dutta, A. K. (2025). *Unveiling diabetes onset: Optimized XGBoost with Bayesian optimization for enhanced prediction*. PLOS ONE, 20(1), e0310218. <https://doi.org/10.1371/journal.pone.0310218>
- Putra, R. R., Tofik, I. G., & Malyan, A. B. J. (2023). *Buku ajar: Pengantar deep learning dalam pemrosesan citra* (14th ed.).
- Rachman, A., Rosnawati, A., Yolanda, G., Saleh, Z., & Aroujo, L. A. (2024). *Perencanaan sumber daya manusia*. Google Books. https://www.google.co.id/books/editio n/Perencanaan_Sumber_Daya_Manusia/-MA7EQAAQBAJ
- Roji, F. M. (2025). *Buku ajar analisis dan perancangan sistem*. Google Books. https://www.google.co.id/books/editio n/Buku_ajar_Analisis_dan_Perancangan_Siste/5UBcEQAAQBAJ
- Setiawan, W. (2021). *Deep learning using convolutional neural networks: Theory and applications* (In Bahasa Indonesia). Google Books. <https://books.google.co.id/books?id=sE9LEAAAQBAJ>
- Shi, R., & Xu, X. (2020). *A train arrival delay prediction model using XGBoost and Bayesian optimization*. In 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC 2020) (pp. 1–6). <https://doi.org/10.1109/ITSC45102.2020.9294186>
- Xu, M. (2024). *Sales prediction based on Lasso regression*. In Highlights in Science, Engineering and Technology IFMPT (Vol. 2024).
- Yan, M., & Zhou, K. (2023). *Bus passenger flow prediction using BO-optimized XGBoost*. In Proceedings - 2023 2nd International Conference on Big Data, Information and Computer Network (BDICN 2023) (pp. 134–141). <https://doi.org/10.1109/BDICN58493.2023.00035>
- Yessy, A. (2024). *Machine learning & deep learning: Analisis sentimen menggunakan ulasan peng....* Google Books. https://www.google.co.id/books/editio n/MACHINE_LEARNING_DEEP_LEARNING_Analisis/Yu7uEAAAQBAJ