



Penerapan Metode Algoritma Extreme Gradient Boosting Dalam Memprediksi Penjualan Thrifting Pada Toko Cchase

Muhammad Ichwan Gifari¹, Yoshida Sary²

^{1,2} Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

¹ichwangifari07@gmail.com, ²yoshidasary@umsu.ac.id

ARTICLE INFORMATION

Received: June 30, 2025
 Revised: August 24, 2025
 Available online: September 12, 2025

KEYWORDS

Thrifting, Sales Prediction, Machine Learning, XGBoost

CORRESPONDENCE

Phone: +62 877-7744-0149
 E-mail: ichwangifari07@gmail.com

ABSTRACT

Thrifting nowadays is very easy to access since thrift stores have been rapidly emerging in urban areas. However, one of the common problems faced by thrift shops is related to stock management and sales prediction. For example, at Cchase store, stock estimation is often inaccurate due to the wide variety of thrift items and the uncertain supply, which leads to the risk of unsold inventory. To overcome this issue, this study applies machine learning technology using the Extreme Gradient Boosting (XGBoost) algorithm. This method was chosen because it has been proven to provide accurate prediction results on data with complex patterns. The data used consist of one year of sales records, which were processed, split into training and testing sets, and then evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) metrics. The results show that the model achieved an RMSE of 0.8105 and an MAE of 0.6643, indicating that the model performs well in predicting sales. Furthermore, the prediction results for the upcoming month reveal the top three product categories with the highest sales, namely crewnecks, hoodies, and t-shirts. These findings are expected to help thrift business owners manage stock more efficiently and develop more effective sales strategies.

1. PENDAHULUAN

Thrifting merupakan praktik membeli dan menjual kembali pakaian bekas yang semakin populer karena menawarkan keunikan, keterjangkauan, dan nilai estetika [1]. Tren ini didorong oleh meningkatnya minat masyarakat, khususnya kalangan remaja, terhadap gaya busana yang unik, terbatas, serta memiliki nilai vintage. Berbeda dengan produk fashion baru, pakaian bekas yang ditawarkan thrift shop sering kali berasal dari merek luar negeri dengan jumlah terbatas, sehingga memberi nilai eksklusif bagi konsumen [2]. Perkembangan ini menjadikan thrifting sebagai salah satu sektor bisnis ritel yang tumbuh pesat di berbagai kota, termasuk Medan, di mana toko Cchase menjadi salah satu contoh pelaku usaha thrift shop yang berkembang [3]. Industri thrifting mengalami pertumbuhan yang sangat pesat. Namun di awal tahun 2020 kita di kejudkan dengan merebaknya virus covid 19 yang membuat pemerintah harus membatasi kegiatan diberbagai sector khususnya di indonesia [4]. Itu merupakan tantangan yang dihadapi pengusahan di bidang thrif.

Selain itu pemerintah melalui PERMENDAG No. 40 Tahun 2022 menetapkan bahwa impor pakaian bekas termasuk dalam kategori barang yang dilarang masuk ke Indonesia [5]. Dengan aturan ini, pelaku usaha thrift hanya diperbolehkan menjual barang preloved dari koleksi pribadi atau melalui jalur legal tertentu.

Hal ini menimbulkan tantangan tersendiri bagi para pelaku usaha karena sebagian besar pasokan thrift shop masih bergantung pada barang impor [6]. Selain aspek regulasi, masalah lain yang dihadapi thrift shop adalah pengelolaan stok dan prediksi permintaan pasar. Tidak seperti produk retail baru yang diproduksi massal, barang thrift memiliki variasi yang tinggi dan ketersediaan yang tidak menentu. Kesalahan dalam menentukan jumlah stok dapat menyebabkan penumpukan barang yang sulit terjual sehingga berdampak pada perputaran modal dan keuntungan usaha [7]. Bagi thrift shop skala kecil hingga menengah, kondisi ini dapat menghambat keberlangsungan bisnis jika tidak diatasi dengan strategi yang tepat.

Seiring dengan berkembangnya teknologi digital, penerapan machine learning dalam pengelolaan bisnis menjadi solusi yang relevan. Salah satu metode yang relevan adalah Extreme Gradient Boosting (XGBoost), sebuah algoritma yang mampu memproses data dalam jumlah besar, mengenali pola non-linier yang kompleks, serta menghasilkan prediksi yang akurat [8] [9]. Dengan penerapan XGBoost, pelaku usaha thrift dapat memproyeksikan permintaan pasar, menentukan jenis barang yang paling diminati, serta merancang strategi stok yang lebih efisien. Hal ini diharapkan tidak hanya membantu mencegah risiko penumpukan produk, tetapi juga meningkatkan daya saing thrift shop dalam menghadapi perubahan tren pasar yang dinamis [9].

2. METODE

Metode penelitian yang digunakan adalah pendekatan prediktif dengan menggunakan model algoritma XGBoost untuk memprediksi penjualan produk thrifting pada toko Cchase. Algoritma ini dipilih karena memiliki kemampuan yang unggul dalam menangani dengan kompleksitas tinggi. Metode ini dipilih karena kemampuannya dalam mencegah overfitting, menangani data yang tidak terstruktur dan memiliki missing value juga menghasilkan prediksi yang cepat dan akurat pada data besar.

1. Pengumpulan Dataset

Pengumpulan data dilakukan melalui dua pendekatan utama, yaitu wawancara semi terstruktur dan dokumentasi. wawancara dilakukan dengan narasumber yang berupa pemilik dari toko Cchase untuk memahami lebih dalam mengenai pola operasional penjualan dan faktor yang memengaruhi naik turunnya penjualan. Data yang dikumpulkan seperti tanggal, hari, minggu penjualan, bulan penjualan, merk, kategori produk, dll.

2. Data Processing

Setelah data terkumpul, dilakukan, data preprocessing yaitu proses pembersihan dan transformasi data mentah agar dapat digunakan dalam pemodelan. Pada tahap ini, data dicek untuk memastikan tidak ada nilai kosong, duplikasi, atau kesalahan format. Selain itu, seluruh variabel kategorikal seperti kategori, merk, hari, metode pembayaran, ukuran, dan kondisi produk diubah menjadi bentuk numerik menggunakan LabelEncoder. Transformasi ini penting karena algoritma machine learning hanya dapat bekerja dengan data numerik, sehingga semua informasi kategorikal perlu dikonversi terlebih dahulu.

3. Data Preparation

Langkah berikutnya data preparation, pada tahap ini ditentukan variabel yang digunakan sebagai fitur (X) dan variabel target (y). Fitur merupakan sekumpulan variabel input yang memengaruhi hasil prediksi, seperti minggu ke, kategori produk, harga, harga akhir, diskon, dan metode pembayaran. Sedangkan target adalah total penjualan yang ingin diprediksi. Data kemudian dibagi menjadi dua bagian, yaitu data latih (training) dan data uji (testing).

4. Train Model

Pada tahap berikutnya train model, yaitu melatih algoritma menggunakan data latih. Model yang digunakan adalah XGBoost, salah satu algoritma berbasis boosting yang efektif dalam menangani data tabular dan dikenal mampu memberikan prediksi dengan akurasi tinggi. Selama proses pelatihan, model belajar mengenali pola hubungan antara variabel input dengan

target sehingga mampu memprediksi jumlah penjualan berdasarkan fitur yang diberikan.

5. Prediksi

Dilakukan tahap prediksi pada data uji, yaitu data dari minggu terakhir. Pada tahap ini, model menghasilkan estimasi penjualan yang kemudian dibandingkan dengan data aktual. Perbandingan ini menjadi dasar untuk melakukan evaluasi model. Evaluasi dilakukan dengan menghitung nilai RMSE (Root Mean Squared Error) dan MAE (Mean Absolute Error). RMSE memberikan gambaran rata-rata kesalahan prediksi dengan penekanan lebih besar pada error yang tinggi, sedangkan MAE menunjukkan rata-rata kesalahan absolut yang lebih mudah dipahami karena langsung dalam satuan penjualan.

Tabel 1. Kriteria RMSE dan MAE

Nilai RMSE dan MAE	Kategori Kemampuan Prediksi
< 10%	Sangat Baik
10% – 20%	Baik
20% – 50%	Layak
> 50%	Buruk

6. Evaluasi Model

Hasil evaluasi menunjukkan model bekerja dengan cukup baik, maka tahap selanjutnya adalah prediksi bulan depan. Untuk melakukannya, data minggu terakhir disalin lalu nilai minggu ke ditambah empat, sehingga dapat mewakili prediksi empat minggu ke depan. Model kemudian digunakan untuk menghasilkan prediksi penjualan di periode tersebut. Hasil prediksi selanjutnya diolah dengan cara mengelompokkan berdasarkan kategori produk, lalu dihitung total penjualan per kategori.

3. HASIL DAN PEMBAHASAN

1. Hasil Evaluasi Model

```

HASIL EVALUASI
RMSE: 0.8105
MAE: 0.6643

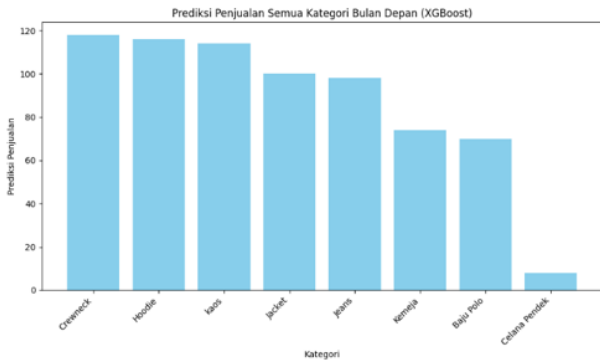
Prediksi Penjualan Semua Kategori Bulan Depan
Kategori      Prediksi Penjualan
1. Crewneck   118
2. Hoodie     116
3. kaos       114
4. Jacket     100
5. Jeans      98
6. Kemeja     74
7. Baju Polo  70
8. Celana Pendek 8
    
```

Gambar 1. Hasil Evaluasi Model

Dari hasil di atas terlihat bahwa model memiliki performa evaluasi yang cukup baik, dengan nilai RMSE sebesar 0.8105 dan MAE sebesar 0.6643. Nilai tersebut menunjukkan bahwa tingkat kesalahan prediksi relatif kecil sehingga model dapat diandalkan untuk memproyeksikan penjualan. Selanjutnya, dilakukan prediksi penjualan untuk periode 1 bulan ke depan (4 minggu) dengan hasil agregasi ke setiap kategori produk. Dengan demikian, kategori Crewneck, Hoodie, dan Kaos menempati posisi teratas, sementara kategori Celana Pendek diproyeksikan memiliki penjualan paling rendah di bulan mendatang. Setelah

mendapatkan hasil dari evaluasi model maka dibuatlah visualisasi hasil prediksi penjualan dengan library Matplotlib. Visualisasi pertama menggunakan diagram batang (bar chart) untuk menunjukkan prediksi penjualan per kategori produk pada bulan depan. Visualisasi kedua menampilkan perbandingan antara data aktual dan prediksi penjualan dalam bentuk kumulatif menggunakan grafik garis (line chart).

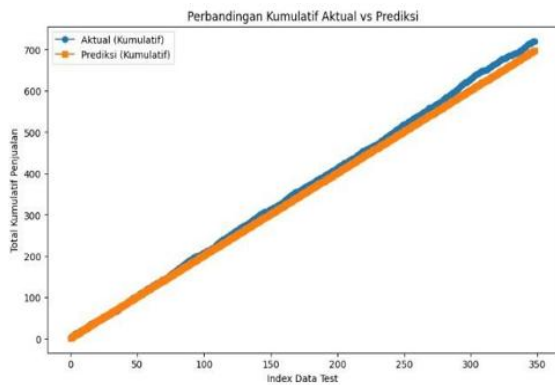
2. Hasil Bar Chart Prediksi Penjualan



Gambar 2. Bar Chart Prediksi Penjualan

Hasil visualisasi prediksi penjualan untuk semua kategori produk pada bulan depan menggunakan algoritma XGBoost. Visualisasi ditampilkan dalam bentuk diagram batang (bar chart) dengan warna biru muda (skyblue) untuk memberikan tampilan yang lebih jelas. Sumbu X merepresentasikan kategori produk seperti Crewneck, Hoodie, Kaos, Jacket, Jeans, Kemeja, Baju Polo, dan Celana Pendek, sedangkan sumbu Y menunjukkan jumlah prediksi penjualan pada masing-masing kategori. Dari grafik ini terlihat bahwa kategori dengan penjualan tertinggi adalah Crewneck, Hoodie, dan Kaos yang diprediksi terjual lebih dari 110 item, sedangkan kategori dengan prediksi penjualan terendah adalah Celana Pendek dengan jumlah kurang dari 10 item.

3. Hasil Nilai Aktual Dan Prediksi



Gambar 3. Line Chart Nilai Aktual dan Prediksi

Dari gambar di atas terlihat visualisasi perbandingan kumulatif nilai aktual dan prediksi penjualan. Garis biru merepresentasikan data aktual kumulatif, sedangkan garis oranye menunjukkan hasil prediksi kumulatif dari model XGBoost. Pola kedua garis terlihat sangat berdekatan dan hampir sejajar dari awal hingga akhir, yang menandakan bahwa model mampu mengikuti tren penjualan dengan baik. Perbedaan yang muncul antara garis aktual dan prediksi relatif kecil, sehingga dapat disimpulkan

bahwa model memiliki performa prediksi yang cukup akurat dalam menggambarkan pola kumulatif penjualan.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model machine learning dengan algoritma XGBoost berhasil dibangun dan diterapkan secara efektif untuk melakukan prediksi penjualan pada toko thrifting lokal (Cchase) dengan memanfaatkan data historis yang tersedia. Hasil evaluasi kinerja model menunjukkan nilai RMSE sebesar 0.8105 dan MAE sebesar 0.6643, yang menandakan bahwa tingkat kesalahan prediksi relatif kecil sehingga model dapat dikategorikan cukup andal dan layak digunakan dalam mendukung pengambilan keputusan bisnis. Prediksi penjualan untuk periode satu bulan ke depan (empat minggu) mengidentifikasi tiga kategori utama yang diperkirakan memberikan kontribusi terbesar terhadap total penjualan, yaitu Crewneck dengan 118 item, Hoodie dengan 116 item, dan Kaos dengan 114 item, yang menunjukkan bahwa tren permintaan konsumen cenderung berfokus pada kategori pakaian kasual dengan daya tarik tinggi di pasar thrifting. Selain itu, temuan penelitian ini memperlihatkan bahwa penerapan algoritma XGBoost tidak hanya relevan bagi perusahaan atau industri berskala besar, tetapi juga dapat diadaptasi secara efektif pada usaha kecil dengan karakteristik data yang unik, seperti variasi stok terbatas, pasokan tidak konsisten, serta dinamika permintaan yang cepat berubah. Dengan demikian, penelitian ini memberikan kontribusi empiris dan praktis bahwa penggunaan algoritma XGBoost dapat menjadi solusi inovatif bagi pelaku usaha thrifting untuk meminimalisasi risiko penumpukan barang, mengoptimalkan strategi pengelolaan stok, serta meningkatkan efektivitas perencanaan penjualan di masa mendatang.

DAFTAR PUSTAKA

- [1] Jury, C. W. A., & Fimelia, Q. (2024). Analisis Pengambilan Keputusan Kementerian Perdagangan Republik Indonesia Dalam Larangan Impor Pakaian Bekas (Thrifting) Menurut Teori Rasional. *Jurnal Ilmu Hukum dan Tata Negara*, 2(2), 200-217. <https://doi.org/10.55606/birokrasi.v2i2.1185>
- [2] Utarsih, H. (2024). Analisis Kualitas Produk dan Harga Terhadap Keputusan Pembelian pada Thrifting Pasar Cimol Gedebage. *Journal of Innovation in Management, Accounting and Business*, 3(2), 170-177. <https://doi.org/10.56916/jimab.v3i2.909>
- [3] Adji, N. L., & Claretta, D. (2022). Fenomena Thrift Shop Dikalangan Remaja: Studi Fenomenologi tentang Thrift Shop di Kalangan Remaja Surabaya. *Journal Of Communication And Islamic Broadcasting*, 2(4), 36-44. <https://doi.org/10.47467/dawatuna.v2i4.2201>
- [4] Sary, Y. S., Martiano, M., Hutagalung, F. S., & Al-iksani, F. A. I. (2022). Pelatihan dan Pendampingan Pemanfaatan Elearning Berbasis LMS Upaya Peningkatan Kualitas Pembelajaran SMA Muhammadiyah 1 Dimasa Pandemi. *ABDI SABHA (Jurnal Pengabdian kepada Masyarakat)*, 3(2), 293-300. <https://doi.org/10.53695/jas.v3i2.712>

- [5] Santika, D. N. A. W., & Bagiastra, I. N. (2021). Legalitas Thrift Shop dan Preloved di Indonesia. *Jurnal Kertha Desa*, 9(6),24-33.
<https://ojs.unud.ac.id/index.php/kerthadesa/article/view/70205>
- [6] Kirana, D. H., & Febriola, B. (2025). Strategi Penanggulangan Dampak Larangan Penjualan Pakaian Bekas Impor di Indonesia Terhadap Pemasukan Pedagang. *Indonesian Journal of Business and Management*, 5(1), 58-68. <https://doi.org/10.53363/buss.v5i1.337>
- [7] Widiana, S. A., Firdaus, I., Tenda, E., & Ketaren, E. (2023). Sistem Informasi Prediksi Penjualan Produk Thrift di Toko Manado Menggunakan Algoritma K-Nearest Neighbor (K-NN). *Jurnal TIMES*, 12(2), 52-57.
<https://doi.org/10.51351/jtm.12.2.2023708>
- [8] Riza, F. (2022). Analisis dan Prediksi Data Penjualan Menggunakan Machine Learning dengan Pendekatan Ilmu Data. *Data Sciences Indonesia*, 1(2), 62-68.
<https://doi.org/10.47709/dsi.v1i2.1308>
- [9] Asselman, A., Khaldi, M., & Aammou, S. (2023). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interactive Learning Environments*, 31(6), 3360-3379.
- [10] Pratama, A. F., Suryani, D., Izzudin, F. S., & Adzikirani, A. (2024). Analisis Prediktif Menggunakan Machine Learning untuk Menanggulangi Masalah Reject Produk pada Proses Produksi PT. XYZ. *Journal of Innovation Research and Knowledge*,4(5),2971-2984.
<https://bajangjournal.com/index.php/JIRK/article/view/8725>