Implementasi Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Kebutuhan Stok Produk Pada Lunart Corp Clothing

Afiana Septi Laili[1], Haris Yuana[2], Abdi Pandu Kusuma[3]
Teknik Informatika, Fakultas Teknik dan Informatika, Universitas Islam Balitar
e-mail: afianasepti21@gmail.com[1], harisyuana2010@gmail.com[2], pans.uib3blitar@gmail.com[3]

Kata Kunci:

Data Mining
K-Nearest Neighbor
Prediksi Stok
Model Klasifikasi

ABSTRAK

Lunart Corp merupakan usaha di bidang produksi dan sablon kaos. Dalam pengelolaan stok, perusahaan sering kesulitan menentukan kondisi stok karena banyaknya variasi produk berdasarkan jenis kain, warna, dan ukuran. Penelitian ini bertujuan mengklasifikasikan kebutuhan stok produk kaos di *Lunart Corp* menggunakan metode *K-Nearest Neighbor* (KNN). Data yang digunakan sebanyak 119 data penjualan tahun 2024–2025 dengan atribut bulan, tahun, jenis kain, warna, dan ukuran. Penelitian ini merupakan penelitian terapan dengan alur pengumpulan data, pengolahan data, penerapan metode KNN, dan pengujian hasil menggunakan *confusion matrix*. Hasil pengujian menunjukkan bahwa metode KNN mampu mengklasifikasikan stok produk dengan akurasi sebesar 86,11%, *precision* 88,46%, *recall* 92%, dan *f1-score* 90,20%. Nilai tersebut menunjukkan bahwa model cukup baik dalam mengenali pola data dan mengelompokkan kondisi stok produk. Dengan demikian, metode KNN dapat digunakan sebagai pendukung pengelolaan persediaan stok kaos di *Lunart Corp*.

Keyword:

Data Mining
K-Nearest Neighbor
Stock Prediction
Classification Model

ABSTRACT

Lunart Corp. is a company engaged in the production and screen printing of t-shirts. In managing product inventory, the company often faces difficulties in determining stock levels due to the wide variety of products based on fabric type, color, and size. This study aims to classify t-shirt stock requirements at Lunart Corp. using the K-Nearest Neighbor (KNN) method. The data used were 119 sales records from 2024–2025 with attributes of month, year, fabric type, color, and size. This applied research followed a flow that included data collection, data processing, application of the KNN method, and testing of the results using a confusion matrix. The test results showed that the KNN method was able to classify product stock with an accuracy of 86.11%, precision of 88.46%, recall of 92%, and f1-score of 90.20%. These values indicate that the model has a fairly good ability to recognize data patterns and group product stock conditions. Therefore, the KNN method can be used to support t-shirt inventory management at Lunart Corp.

Copyright © SENASIF 2026
ISSN: 2598-0076

I. PENDAHULUAN

Perkembangan pencatatan data transaksi yang semakin cepat menghasilkan data dalam jumlah besar yang dapat dimanfaatkan sebagai pendukung pengambilan keputusan bisnis [1]. *Lunart Corp* merupakan usaha di bidang produksi dan sablon kaos yang melayani pencetakan desain pada media pakaian sesuai permintaan pelanggan. Produk yang dikelola memiliki variasi jenis kain, khususnya Combed 30s dan Combed 24s, serta beragam warna dan ukuran, sehingga diperlukan pengelolaan data penjualan dan stok yang terstruktur agar permintaan pelanggan dapat dipenuhi tepat waktu.

Namun, pengelolaan persediaan stok di *Lunart Corp* masih dilakukan secara konvensional [2]. Penentuan *restock* umumnya baru dilakukan setelah terjadi kekosongan stok, tanpa didasarkan pada analisis data penjualan sebelumnya. Pencatatan penjualan juga masih berupa transaksi harian yang memuat nama, alamat, jenis kain, ukuran, dan jumlah pembelian, sehingga data hanya berfungsi sebagai arsip dan belum dimanfaatkan untuk mengetahui pola kebutuhan stok berdasarkan jenis kain dan ukuran yang paling diminati. Dampaknya, terjadi kekosongan pada produk tertentu sementara produk lain justru menumpuk. Kondisi ini membuat perputaran modal tidak maksimal dan menimbulkan ketidakpastian gambaran laba penjualan.

Data mining diharapkan dapat membantu memberikan solusi karena mampu menggali informasi bernilai dari data historis untuk mendukung pengambilan keputusan [1]. Salah satu tekniknya adalah klasifikasi, yaitu metode *supervised learning* yang memetakan data ke dalam kelas-kelas yang telah ditentukan berdasarkan hubungan antara atribut masukan dan atribut sasaran [3]–[5]. Metode *K-Nearest Neighbor* (KNN) merupakan salah satu algoritma klasifikasi yang bekerja berdasarkan konsep *learning by analogy*, yaitu menentukan kelas suatu objek dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut [6].

Penerapan KNN pada data penjualan telah dilakukan oleh beberapa peneliti. Alfani dkk. [7] memprediksi penjualan produk Unilever dengan akurasi terbaik 86,66%, sedangkan Handoko dkk. [8] menerapkan KNN pada penjualan paket kuota internet dengan akurasi 71,43%. KNN juga digunakan pada bidang lain, misalnya prediksi kualitas udara yang memperoleh akurasi 96% pada K = 7 [9], prediksi kelulusan mahasiswa dengan akurasi 96,18% yang lebih tinggi daripada *Naive Bayes* [10], analisis sentimen dengan akurasi 84,93% [11], serta prediksi arus lalu lintas dengan akurasi rata-rata 91,1% [12], serta prediksi kasus Covid-19 dengan akurasi 96,98% [13]. Hasil-hasil tersebut menunjukkan bahwa KNN memiliki kinerja yang kompetitif pada berbagai jenis data.

Pada permasalahan klasifikasi, kinerja model umumnya dievaluasi menggunakan *confusion matrix*, yaitu tabel yang membandingkan label aktual dengan label hasil klasifikasi melalui empat komponen: *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [14]. Dari komponen tersebut dapat dihitung akurasi, *precision*, *recall*, dan *f1-score*. Evaluasi ini

Copyright © SENASIF 2026
ISSN: 2598-0076

lebih sesuai untuk kasus klasifikasi karena keluaran model berupa kategori, bukan nilai numerik kontinu.

Berdasarkan uraian tersebut, penelitian ini menerapkan metode KNN untuk mengklasifikasikan produk ke dalam kategori Perlu *Restock* dan Aman di *Lunart Corp* dengan memanfaatkan variabel jenis kain, ukuran, bulan, dan riwayat jumlah penjualan. Kebaruan penelitian ini terletak pada penggunaan fitur temporal hasil rekayasa fitur, penerapannya pada industri sablon pakaian yang permintaannya sangat bervariasi, serta pengintegrasian model ke dalam sistem berbasis web. Tujuan penelitian ini adalah (1) menjelaskan penerapan metode KNN dalam mengklasifikasikan kebutuhan stok penjualan di *Lunart Corp* dan (2) mengukur kinerja algoritma tersebut menggunakan *confusion matrix* berdasarkan nilai akurasi, *precision*, dan *recall*. Analisis dibatasi pada data historis penjualan perusahaan, tanpa membahas faktor eksternal seperti tren pasar, promosi, atau perilaku konsumen.

II. METODE

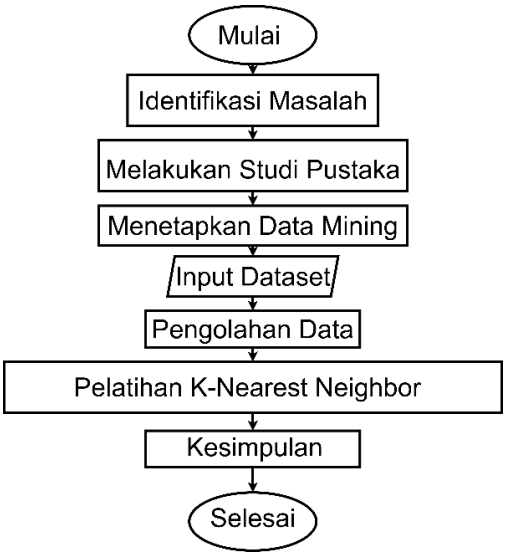
Penelitian ini merupakan penelitian terapan yang berfokus pada penyelesaian masalah nyata, yaitu klasifikasi kebutuhan stok produk berdasarkan data penjualan *Lunart Corp* yang berlokasi di Sumbergempol, Kabupaten Blitar. Penelitian dilaksanakan pada bulan November 2025 sampai Juli 2026. Lokasi dipilih karena perusahaan menghadapi masalah nyata dalam pengelolaan stok, yaitu pencatatan stok yang tidak diperbarui secara berkala, proses *restock* yang baru dilakukan ketika stok habis, dan kesulitan mengetahui jumlah stok yang tersedia.

Data dikumpulkan melalui tiga teknik. Wawancara dilakukan dengan kepala produksi untuk mengetahui kondisi pengelolaan stok, kendala pencatatan penjualan, dan proses pengambilan keputusan *restock*. Observasi dilakukan dengan mengamati langsung proses transaksi, pencatatan penjualan, dan pengelolaan persediaan. Studi literatur dilakukan dengan menelaah buku, jurnal, dan sumber tertulis lain yang relevan. Data utama berupa data sekunder riwayat penjualan tahun 2024 sampai 2025 sebanyak 119 data dengan total 5.051 unit kaos terjual, yang memuat bulan, tahun, jenis kain, warna, ukuran, dan jumlah terjual (*Qty*).

Tahapan penelitian ditunjukkan pada Gambar 1, yang meliputi identifikasi masalah, studi pustaka, penetapan metode *data mining*, input dataset, pengolahan data, pelatihan KNN, dan penarikan kesimpulan. Pengolahan data mengikuti prinsip *Knowledge Discovery in Database* (KDD) yang terdiri atas seleksi, pembersihan, dan transformasi data sebelum dianalisis [15]. Seleksi dilakukan dengan menghapus atribut yang tidak berpengaruh langsung, seperti nama pembeli. Setiap data kemudian diberi label berdasarkan jumlah terjual, yaitu kurang dari 50 unit sebagai Perlu *Restock* (label 1) dan 50 unit atau lebih sebagai Aman (label 0).

Copyright © SENASIF 2026

ISSN: 2598-0076



Gambar 1. Tahap-tahap Penelitian

Atribut kategorikal diubah menjadi numerik: jenis kain dengan *One-Hot Encoding* dan ukuran (S sampai 3L) dengan *Ordinal Encoding* agar urutan ukuran tetap terjaga. Rekayasa fitur dilakukan per kelompok SKU (jenis dan ukuran) untuk membentuk fitur *qty_lag1* (jumlah terjual pada periode sebelumnya), *qty_trend* (selisih jumlah saat ini dan sebelumnya), serta *qty_trend_sign* (arah tren naik, tetap, atau turun). Karena KNN sangat sensitif terhadap perbedaan skala, fitur numerik dinormalisasi menggunakan *Min-Max Scaling* yang mengubah nilai ke rentang 0 sampai 1 [16], sebagaimana pada Persamaan (1).

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

dengan X' adalah nilai hasil normalisasi, X adalah nilai asli, serta X_{min} dan X_{max} adalah nilai minimum dan maksimum atribut. Data kemudian dibagi menjadi data latih dan data uji dengan proporsi 70:30 menggunakan *stratified split*. Ketimpangan kelas pada data latih diatasi dengan teknik *Synthetic Minority Over-sampling Technique* (SMOTE). Proses penskalaan dan penyeimbangan hanya dilakukan pada data latih agar tidak terjadi kebocoran informasi dari data uji.

Nilai K ditentukan menggunakan *Elbow Method*, yaitu menguji beberapa nilai K ganjil dan mengamati perubahan tingkat kesalahan (*error rate*) untuk menemukan titik siku. Nilai K ganjil dipilih untuk menghindari hasil pemungutan suara yang seri. Jarak antardata dihitung dengan *Euclidean Distance* seperti pada Persamaan (2), dikombinasikan dengan pembobotan berdasarkan jarak sehingga tetangga yang lebih dekat memiliki pengaruh lebih besar [6].

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{2}$$

Dengan d(x,y) adalah jarak, x_i dan y_i adalah nilai variabel ke-i pada data latih dan data uji, serta n adalah dimensi data. Kinerja model diuji menggunakan *confusion matrix* dengan metrik akurasi, *precision*, *recall*, dan *f1-score* yang dirumuskan pada Persamaan (3) sampai (6) [14].

Copyright © SENASIF 2026
ISSN: 2598-0076

$$\text{Akurasi} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Seluruh proses dikerjakan dengan bahasa pemrograman Python menggunakan pustaka *pandas*, *scikit-learn*, dan *Matplotlib*, sedangkan antarmuka sistem dibangun menggunakan kerangka kerja *Streamlit*.

III. HASIL DAN PEMBAHASAN

Deskripsi Data dan Pelabelan

Dataset yang digunakan berasal dari riwayat penjualan *Lunart Corp* tahun 2024–2025 sebanyak 119 data dalam format CSV. Hasil pemeriksaan menunjukkan bahwa seluruh data lengkap (tanpa *missing value*) sehingga tidak diperlukan imputasi maupun penghapusan baris. Variabel yang digunakan dijelaskan pada Tabel 1, sedangkan contoh data beserta label hasil pelabelan ditampilkan pada Tabel 2.

Tabel 1. Variabel Dataset Penjualan Lunart Corp

No	Variabel	Keterangan
1	Bulan	Bulan terjadinya transaksi
2	Tahun	Tahun terjadinya transaksi
3	Jenis	Jenis kain produk (Combed 24s dan 30s)
4	Warna	Warna produk
5	Ukuran	Ukuran produk (S sampai 3L)
6	Qty Terjual	Jumlah produk yang terjual

Tabel 2. Contoh Dataset dan Hasil Pelabelan

No	Bulan	Tahun	Jenis	Warna	Ukuran	Qty Terjual	Label
1	1	2025	30s	Hitam	S	15	Perlu restock
2	1	2025	30s	Hitam	M	5	Perlu restock
3	1	2025	24s	Hitam	L	156	Aman
4	1	2025	24s	Putih	M	34	Perlu restock
…	…	…	…	…	…	…	…
117	11	2025	30s	Hitam	L	147	Aman
118	12	2025	30s	Putih	XL	16	Perlu restock
119	12	2025	30s	Hitam	3L	40	Perlu restock

Pelabelan berdasarkan ambang batas 50 unit menghasilkan kelas Perlu *Restock* sebagai kelas mayoritas, sedangkan kelas Aman hanya berjumlah 26 data pada data latih. Kondisi ini menunjukkan

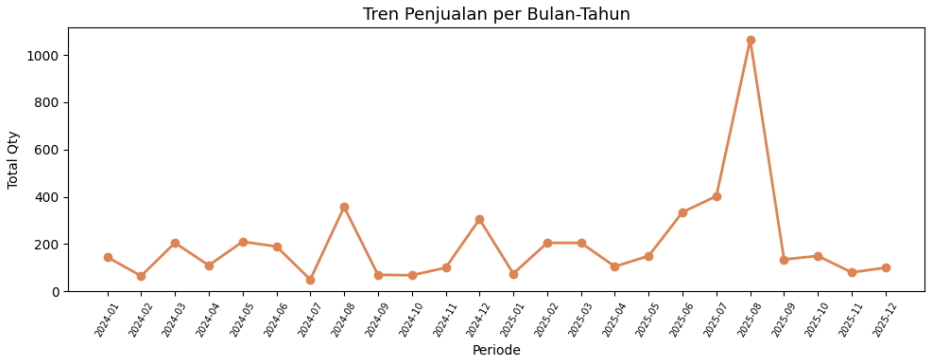
Copyright © SENASIF 2026
ISSN: 2598-0076

adanya ketimpangan kelas yang berpotensi membuat model bias terhadap kelas mayoritas bila tidak ditangani.

Eksplorasi Data

Distribusi jumlah terjual memiliki nilai *skewness* 1,78 yang menunjukkan distribusi condong ke kanan. Artinya, sebagian besar transaksi berada pada volume kecil hingga sedang, tetapi terdapat beberapa lonjakan penjualan yang sangat tinggi. Deteksi pencilan dengan metode *Interquartile Range* (IQR) menghasilkan batas atas 112,75 unit dan menemukan 6 data pencilan dengan nilai tertinggi 200 unit. Data tersebut tidak dihapus karena merepresentasikan lonjakan permintaan yang nyata dan penting bagi model.

Analisis tren waktu pada Gambar 2 menunjukkan bahwa penjualan tertinggi terjadi pada bulan Agustus 2025 dengan 1.065 unit, yang mengindikasikan adanya faktor musiman atau promosi pada bulan tersebut. Dari sisi produk, kaos berwarna merah memiliki rata-rata permintaan tertinggi, yaitu 58,24 unit per transaksi, dan ukuran XL serta L mendominasi dengan rata-rata masing-masing 57,38 dan 47,32 unit. Analisis korelasi menunjukkan bahwa ukuran berkorelasi negatif sebesar −0,17 dan bulan sebesar −0,20 terhadap status *restock*; semakin besar ukuran, kecenderungan masuk kategori Perlu *Restock* semakin rendah karena volume penjualannya tinggi.



Gambar 2. Tren Penjualan per Bulan

Pra-Pemrosesan dan Rekayasa Fitur

Pada tahap pra-pemrosesan, kolom yang tidak relevan dihapus, kemudian dibentuk fitur *qty_lag1*, *qty_trend*, dan *qty_trend_sign*. Distribusi data menunjukkan bahwa 75,6% transaksi mengalami kenaikan penjualan dibandingkan periode sebelumnya. Analisis korelasi pasca-rekayasa fitur memperlihatkan bahwa *qty_lag1* dan *qty_trend_sign* memiliki keterkaitan yang berarti dengan variabel target, sehingga penambahan dimensi tren dan riwayat penjualan memperkuat informasi yang tersedia bagi model.

Data dibagi menjadi 83 data latih dan 36 data uji. Setelah fitur numerik diskalakan dengan *MinMaxScaler*, SMOTE dengan parameter k_neighbors = 1 diterapkan pada data latih. Jumlah data kelas Aman meningkat dari 26 menjadi 57 melalui penambahan sampel sintetis sehingga kedua kelas seimbang dengan masing-masing 57 data. Hal ini penting agar model tidak condong pada kelas mayoritas.

Copyright © SENASIF 2026
ISSN: 2598-0076

Pelatihan dan Pengujian Model

Penentuan nilai K dengan *Elbow Method* dilakukan pada beberapa nilai K ganjil, dan titik siku yang paling jelas diperoleh pada K = 7. Setelah titik tersebut, penambahan nilai K tidak lagi menurunkan tingkat kesalahan secara berarti. Model dilatih menggunakan data hasil SMOTE dengan K = 7, jarak *Euclidean*, dan pembobotan berdasarkan jarak, lalu diuji pada 36 data uji yang belum pernah dilihat model.

Hasil pengujian ditampilkan pada Tabel 3. Model mencapai akurasi 86,11%, yang berarti sebagian besar data uji diklasifikasikan dengan benar. Nilai *precision* sebesar 88,46% menunjukkan bahwa sebagian besar produk yang diprediksi Perlu *Restock* memang benar membutuhkan *restock*, sedangkan *recall* sebesar 92% menunjukkan bahwa model mampu menemukan hampir seluruh produk yang sebenarnya perlu di-*restock*. *F1-score* sebesar 90,20% menandakan keseimbangan yang baik antara ketepatan dan kelengkapan. Nilai *recall* yang tinggi penting bagi perusahaan karena kegagalan mendeteksi kebutuhan *restock* dapat berujung pada kekosongan stok yang menghambat penjualan.

Tabel 3. Hasil Evaluasi Model KNN (K = 7) pada 36 Data Uji

Metrik	Nilai (%)
Akurasi	86,11
Precision	88,46
Recall	92,00
F1-score	90,20

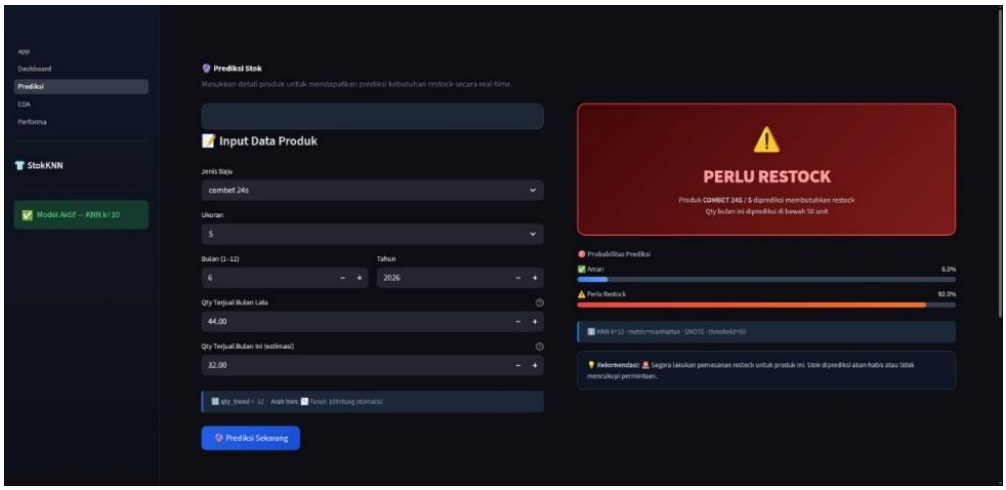
Implementasi Sistem Berbasis Web

Model yang telah divalidasi diintegrasikan ke dalam sistem berbasis web bernama StokKNN *Lunart Corp*. Seluruh komponen *pipeline*, yaitu model KNN, *scaler*, *encoder*, dan parameter rekayasa fitur, disimpan dalam berkas berekstensi .pkl menggunakan pustaka *joblib* untuk menjamin konsistensi transformasi data antara fase pelatihan dan fase prediksi. Antarmuka sistem terdiri atas empat halaman, yaitu Dashboard, Prediksi Stok, EDA dan Analisis, serta Performa Model.

Prediksi status stok pada data baru dilakukan melalui halaman Prediksi Stok dengan mengisi jenis bahan, ukuran, bulan, dan jumlah stok saat ini, seperti pada Gambar 3. Masukan tersebut diproses otomatis oleh *pipeline*, kemudian hasilnya ditampilkan dalam kartu berwarna merah untuk status Perlu *Restock* dan hijau untuk status Aman. Dengan demikian, pemilik usaha dapat mengambil keputusan *restock* secara langsung tanpa bergantung pada kondisi stok yang terlihat secara fisik.

Copyright © SENASIF 2026

ISSN: 2598-0076



Gambar 3. Tampilan Halaman Prediksi Stok

Pembahasan

Hasil penelitian menunjukkan bahwa KNN mampu mengklasifikasikan kebutuhan stok di *Lunart Corp* dengan kinerja yang baik, sejalan dengan penelitian sebelumnya yang memanfaatkan *data mining* untuk mendukung keputusan pada data penjualan [7]. Akurasi 86,11% setara dengan hasil Alfani dkk. [7] sebesar 86,66% dan lebih tinggi daripada Handoko dkk. [8] sebesar 71,43%, meskipun lebih rendah daripada penelitian pada data lain yang mencapai 96% [9], [10]. Perbedaan ini wajar karena karakteristik dan jumlah data setiap penelitian berbeda.

Pada tahap pengolahan data, tantangan utama adalah atribut kategorikal yang tidak dapat langsung dihitung jaraknya, perbedaan rentang nilai antaratribut, serta ketimpangan kelas. Ketiganya diatasi dengan *encoding*, normalisasi *Min-Max*, dan SMOTE. Normalisasi dan pengujian beberapa nilai K juga diterapkan pada penelitian KNN lain untuk menjaga kontribusi setiap atribut tetap seimbang [16], [17]. Pada tahap pemodelan, penentuan nilai K menjadi penting karena K yang terlalu kecil membuat model sensitif terhadap *noise*, sedangkan K yang terlalu besar mengaburkan batas antarkelas [6]. *Elbow Method* membantu menentukan K secara objektif berdasarkan data, dan hasil tersebut sejalan dengan temuan bahwa pemilihan K yang tepat meningkatkan kinerja klasifikasi [14].

Penelitian ini memiliki keterbatasan. Jumlah data hanya 119 dengan kelas Aman yang sedikit pada kondisi awal, sehingga keseimbangan kelas masih bergantung pada data sintetis hasil SMOTE. Penelitian ini juga hanya menggunakan satu algoritma tanpa pembanding dan terbatas pada satu studi kasus, sehingga generalisasi ke jenis usaha lain perlu diuji lebih lanjut. Meskipun demikian, penerapan model pada sistem web membuat proses penentuan *restock* menjadi lebih terstruktur dibandingkan pencatatan manual sebelumnya.

IV. SIMPULAN

Penelitian ini menunjukkan bahwa metode *K-Nearest Neighbor* dapat diterapkan untuk mengklasifikasikan kebutuhan stok produk pada *Lunart Corp* melalui tahapan yang sistematis, mulai

Copyright © SENASIF 2026
ISSN: 2598-0076

dari penyeleksian dan pelabelan data, pembersihan, penyandian atribut, penyamaan skala fitur, penyeimbangan kelas, hingga pelatihan model dengan nilai K yang ditentukan menggunakan metode siku. Hasil pengujian dengan *confusion matrix* menunjukkan bahwa model memiliki kinerja yang baik dan sangat andal dalam mengenali produk yang perlu segera ditambah persediaannya, sehingga risiko kekosongan stok dapat ditekan. Model yang dihasilkan juga telah diintegrasikan ke dalam sistem berbasis web sehingga dapat membantu pemilik usaha mengambil keputusan penambahan stok secara lebih mudah dan terencana dibandingkan cara konvensional. Penelitian selanjutnya disarankan menambah jumlah data dan variabel, menggunakan bentuk pengujian yang lain, serta membandingkan metode ini dengan algoritma klasifikasi lain atau teknik gabungan beberapa model agar hasil yang diperoleh lebih baik.

DAFTAR RUJUKAN

[1] N. Riskya and S. Yuliana, "Penerapan Data Mining Untuk Prediksi Perilaku Pelanggan Menggunakan Multiple Linear Regression," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3, 2023, doi: 10.23960/jitet.v11i3.3194.

[2] A. F. Jose, S. Sucipto, M. N. Muzaki, and M. N. F. Hidayat, "Sistem Informasi Manajemen Stok Mengimplementasikan Metode Min-Max Pada Toko Grosir Amanah," *JSITIK J. Sist. Inf. dan Teknol. Inf. Komput.*, vol. 3, no. 2, pp. 112–127, 2025, doi: 10.53624/jsitik.v3i2.563.

[3] A. Nata and S. Suparmadi, "Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar," *J. Sci. Soc. Res.*, vol. 5, no. 3, p. 697, 2022, doi: 10.54314/jssr.v5i3.1041.

[4] S. Hidayat and R. Idmayanti, "Klasifikasi Data Penduduk Berdasarkan Tingkat Ekonomi di Nagari Batahan Tengah Menggunakan Metode Naive Bayes," *J. Bitwise*, vol. 1, no. 2, pp. 66–73, 2025.

[5] R. S. Nurhalizah, R. Ardianto, and P. Purwono, "Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review," *J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, pp. 61–72, 2024, doi: 10.54082/jiki.168.

[6] S. R. Raysyah, V. Arinal, and D. I. Mulyana, "Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode KNN Dan PCA," *JSii (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 88–95, 2021, doi: 10.30656/jsii.v8i2.3638.

[7] A. Alfani W.P.R., F. Rozi, and F. Sukmana, "Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jipi.v6i1.1910.

[8] D. Handoko, H. S. Tambunan, and J. T. Hardinata, "Analisis Penjualan Produk Paket Kuota Internet Dengan Metode K-Nearest Neighbor," *JURASIK (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 6, no. 1, p. 111, 2021, doi: 10.30645/jurasik.v6i1.275.

Copyright © SENASIF 2026
ISSN: 2598-0076

[9] A. Amalia, A. Zaidia, and I. N. Isnainiyah, "Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 4, no. 2, pp. 496–507, 2022.

[10] K. Kartarina, N. K. Sriwinarti, and N. L. P. Juniarti, "Analisis Metode K-Nearest Neighbors (K-NN) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 3, no. 2, pp. 106–112, 2021, doi: 10.35746/jtim.v3i2.159.

[11] J. Supriyanto, D. Alita, and A. R. Isnain, "Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 74–80, 2023, doi: 10.33365/jatika.v4i1.2468.

[12] Y. Liu, H. Yu, and H. Fang, "Application of KNN prediction model in urban traffic flow prediction," in *Proc. 2021 5th Asian Conf. Artif. Intell. Technol. (ACAIT)*, 2021, pp. 389–392, doi: 10.1109/ACAIT53529.2021.9731348.

[13] M. V. Overbeek, A. G. Y. Stephanno, E. E. Surbakti, and F. A. T. Tobing, "Covid-19 Prediction in Indonesia Capital Buffer with K Nearest Neighbor Regression," in *Proc. Int. Conf. Informatics Comput. Sci. (ICICoS)*, 2021, pp. 238–243, doi: 10.1109/ICICoS53627.2021.9651749.

[14] M. Iqbal Mubarok, P. Purwantoro, and C. Carudin, "Penerapan Algoritma K-Nearest Neighbor (KNN) Dalam Klasifikasi Penilaian Jawaban Ujian Esai," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 5, pp. 3446–3452, 2024, doi: 10.36040/jati.v7i5.7676.

[15] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 9, no. 2, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.

[16] M. Ula, R. Zulhusna, R. Putra Fhonna, and A. Pratama, "Penerapan Model Klasifikasi K-Nearest Neighbor Dalam Pencarian Kesesuaian Pekerjaan," *Metik J.*, vol. 6, no. 1, pp. 18–23, 2022, doi: 10.47002/metik.v6i1.343.

[17] R. Rahmadini, E. E. Lorencis Lubis, A. Priansyah, Y. R.W.N, and T. Meutia, "Penerapan Data Mining Untuk Memprediksi Harga Bahan Pangan Di Indonesia Menggunakan Algoritma K-Nearest Neighbor," *J. Mhs. Akunt. Samudra*, vol. 4, no. 4, pp. 223–235, 2023, doi: 10.33059/jmas.v4i4.7074.

Copyright © SENASIF 2026
ISSN: 2598-0076