

Penerapan Metode *K-Nearest Neighbor* untuk Prediksi Penjualan Sparepart Mobil pada Gudang Marketplace

Indah Nur Qoriah¹, Leny Tritanto Ningrum²

^{1,2}Program Studi Sistem Informasi, Fakultas Informatika dan Komputer, Universitas Binaniaga Indonesia
Email: indahnq1107@gmail.com
Email: lenytrinie@unbin.ac.id

ABSTRACT

This study focuses on predicting car spare parts sales in a marketplace warehouse by employing the K-Nearest Neighbor (K-NN) algorithm. K-NN is utilized as an effective data mining technique for classification in order to construct an accurate predictive model. The dataset used in this research consists of key attributes, including spare part name, sales volume, quantity, price, condition, and brand. A web-based prototype was developed using PHP, enabling accessibility across devices and providing itemset predictions through Euclidean distance analysis. To assess model performance, evaluation was carried out using a confusion matrix, which produced an accuracy of 72%, precision of 86%, and recall of 100%. Furthermore, a feasibility test of the system achieved a perfect score of 100%, while user testing resulted in 83.03%, both of which indicate that the system is highly feasible and user-friendly. The research findings confirm that the K-NN method can deliver effective and reliable predictions of car spare parts sales. Beyond its theoretical contribution in expanding insights into K-NN applications, this study also offers practical value by supporting inventory management efficiency and assisting company management in making strategic business decisions.

Keywords: K-NN, Data Mining, Confusion Matrix, Classification, Prediction, Spare Parts, Sales.

ABSTRAK

Penelitian ini berfokus pada prediksi penjualan sparepart mobil di gudang marketplace dengan memanfaatkan algoritma K-Nearest Neighbor (K-NN). Algoritma K-NN digunakan sebagai teknik data mining yang efektif dalam klasifikasi data untuk membangun model prediksi yang akurat. Dataset penelitian mencakup atribut utama, yaitu nama sparepart, jumlah penjualan, quantity, harga, kondisi, dan brand. Sistem yang dikembangkan berupa prototipe berbasis web menggunakan bahasa pemrograman PHP, sehingga mudah diakses melalui berbagai perangkat, dan menghasilkan prediksi itemset berdasarkan analisis jarak Euclidean. Evaluasi model dilakukan dengan metode confusion matrix yang menghasilkan akurasi sebesar 72%, precision 86%, dan recall 100%. Selain itu, uji kelayakan sistem memperoleh nilai 100%, sedangkan uji pengguna mencapai 83,03%, yang menandakan bahwa aplikasi yang dibangun sangat layak dan mudah digunakan. Hasil penelitian ini membuktikan bahwa metode K-NN mampu memberikan prediksi penjualan sparepart mobil secara efektif dan andal. Secara teoritis, penelitian ini berkontribusi dalam memperluas pemahaman penerapan metode K-NN, sedangkan secara praktis, penelitian ini memberikan manfaat nyata dalam meningkatkan efisiensi manajemen persediaan serta mendukung pengambilan keputusan strategis bagi perusahaan.

Kata Kunci: K-NN, Data Mining, Confusion matrix, Klasifikasi, Prediksi, Spare part, Penjualan.

A. PENDAHULUAN

1. Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah mendorong transformasi signifikan pada sektor *e-commerce*, termasuk dalam industri penjualan *spare part* mobil melalui *marketplace*. Kusumaningsih (2021) menyatakan bahwa *marketplace* adalah *platform* yang berfungsi sebagai perantara antara penjual dan pembeli untuk memfasilitasi transaksi produk secara daring. Keunggulan *marketplace* terletak pada kemudahan akses, variasi produk, serta fleksibilitas transaksi. Namun, tantangan yang dihadapi pengelola gudang adalah bagaimana memprediksi penjualan secara akurat agar ketersediaan stok tetap terjaga, biaya penyimpanan terkendali, dan risiko kelebihan maupun kekurangan stok dapat diminimalisasi. Menurut data Kominfo (2024), jumlah pengguna internet di Indonesia pada Januari 2024 mencapai 221.563.479 jiwa atau 79,5% dari total populasi 278,6 juta

orang. Fakta ini memperkuat posisi *marketplace* sebagai *platform* utama yang sangat potensial bagi pelaku usaha, khususnya Usaha Mikro, Kecil, dan Menengah (UMKM), dalam meningkatkan daya saing serta pendapatan di tengah persaingan pasar yang semakin ketat.

Kontribusi UMKM terhadap perekonomian nasional juga sangat besar. Berdasarkan Undang-Undang Nomor 20 Tahun 2008 (RI, 2023), UMKM merupakan usaha produktif yang dimiliki perorangan maupun badan usaha dengan memenuhi kriteria tertentu. Data Kementerian Koperasi dan UKM (2019) menunjukkan bahwa Indonesia memiliki 65,4 juta UMKM yang mampu menyerap lebih dari 123,3 ribu tenaga kerja, sehingga sektor ini menjadi tulang punggung ekonomi. Namun, pandemi COVID-19 memberikan dampak signifikan pada pola konsumsi masyarakat. Roy Nicholas Mandey (2021) menjelaskan bahwa Indeks Penjualan Riil pada Maret 2021 berada di angka -17,1%, dengan rata-rata 4–5 toko ritel yang tutup setiap hari sepanjang tahun 2020. Kondisi ini mempercepat peralihan aktivitas belanja masyarakat ke *platform* daring, sekaligus memperkuat urgensi penelitian terkait strategi penjualan di *marketplace*.

Selain faktor eksternal tersebut, kebutuhan *spare part* mobil semakin meningkat seiring dengan bertambahnya jumlah kendaraan bermotor. Data Korlantas Polri (2024) mencatat bahwa jumlah kendaraan di Indonesia mencapai 160.652.675 unit, terdiri dari 19.906.353 mobil pribadi, 269.476 bus, 6.120.307 kendaraan angkut barang, dan 134.181.607 sepeda motor. Kondisi ini menunjukkan bahwa *spare part* merupakan komoditas dengan potensi pasar yang besar. Wikipedia (2024) mendefinisikan *spare part* sebagai komponen cadangan yang digunakan untuk perawatan maupun perbaikan kendaraan, sekaligus menjadi bagian vital dalam rantai suplai. Dengan permintaan yang terus meningkat, manajemen persediaan *spare part* menjadi tantangan strategis yang membutuhkan metode prediksi penjualan yang akurat dan dapat diandalkan.

Berbagai penelitian terdahulu telah mengkaji metode prediksi penjualan dengan pendekatan data mining. Sahara Abdy, dkk. (2022) menyatakan bahwa algoritma K-Nearest Neighbor (K-NN) dapat digunakan untuk memprediksi penjualan *spare part* mobil dengan memanfaatkan data historis menggunakan perhitungan jarak Euclidean. Sementara itu, Whendasgoro dan Joseph (2022) menambahkan bahwa penerapan normalisasi data dengan teknik Z-Score Normalization mampu meningkatkan kinerja algoritma K-NN secara signifikan. Walaupun penelitian-penelitian tersebut menunjukkan hasil yang positif, sebagian besar masih berfokus pada analisis data secara umum dan belum menitikberatkan pada implementasi model prediksi penjualan *spare part* mobil di gudang *marketplace*. Dengan demikian, masih terdapat kesenjangan penelitian terkait bagaimana metode K-NN dapat dioptimalkan dalam konteks manajemen persediaan yang spesifik di *marketplace*.

Dibandingkan dengan metode lain, algoritma K-NN memiliki sejumlah keunggulan yang menjadikannya pilihan utama dalam penelitian ini. Metode *decision tree* misalnya, sering kali menghadapi masalah *overfitting* pada dataset yang kompleks. Sementara itu, metode *naïve bayes* mengasumsikan independensi antar variabel yang dalam praktiknya sulit terpenuhi pada data penjualan yang penuh dengan keterkaitan antar atribut. K-NN, di sisi lain, bersifat non-parametrik dan sederhana namun efektif, karena melakukan klasifikasi berdasarkan kedekatan jarak antar data. Hal ini membuat K-NN mampu beradaptasi dengan baik terhadap variasi data penjualan *spare part* yang dinamis. Seperti dijelaskan oleh Han, dkk. (2012), K-NN merupakan salah satu algoritma *lazy learning* yang bekerja optimal pada dataset berukuran sedang dengan atribut numerik yang terukur. Keunggulan inilah yang menjadi alasan utama pemilihan metode K-NN dalam penelitian ini.

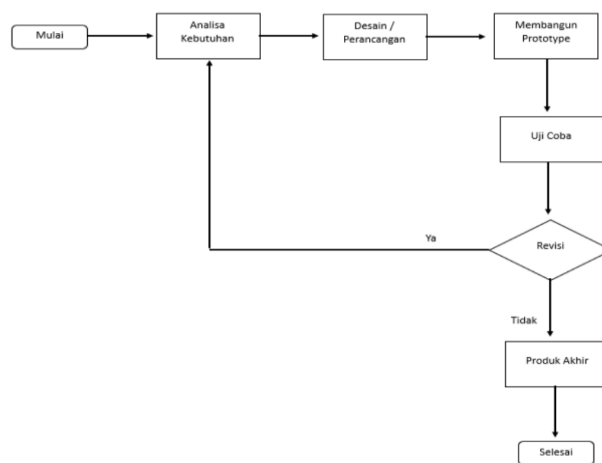
Oleh karena itu, penelitian ini tidak hanya relevan secara praktis dalam membantu pengelola gudang *marketplace* memprediksi penjualan dan mengelola persediaan, tetapi juga penting secara akademis sebagai pengembangan pengetahuan di bidang *data mining* dan sistem prediksi. Dari sisi praktis, hasil penelitian dapat digunakan sebagai acuan bagi perusahaan dalam mengambil keputusan strategis terkait persediaan, pengendalian biaya, dan perencanaan distribusi. Dari sisi teoretis, penelitian ini memperluas pemahaman mengenai penerapan algoritma K-NN pada kasus nyata yang spesifik, yaitu penjualan *spare part* mobil, yang hingga saat ini masih jarang menjadi fokus kajian. Dengan adanya kombinasi kontribusi praktis dan teoretis, penelitian ini diharapkan memberikan nilai tambah baik bagi dunia akademik maupun industri.

Berdasarkan latar belakang tersebut, penelitian ini dilakukan dengan tujuan untuk menerapkan metode K-NN dalam memprediksi penjualan *spare part* mobil pada gudang *marketplace* dengan mempertimbangkan atribut historis penjualan, jumlah data, quantity, dan harga. Penelitian ini juga bertujuan untuk mengukur tingkat akurasi serta efektivitas model yang dikembangkan melalui evaluasi berbasis confusion matrix. Selain itu, penelitian ini diarahkan untuk mengembangkan prototipe aplikasi berbasis web yang dapat diakses secara luas guna mendukung proses pengambilan keputusan strategis dalam manajemen persediaan. Dengan demikian, penelitian ini diharapkan tidak hanya memberikan kontribusi teoritis dalam pengembangan metode data mining, tetapi juga manfaat praktis bagi pengelolaan persediaan *spare part* di *marketplace*.

B. METODE

1. Tahapan Penelitian

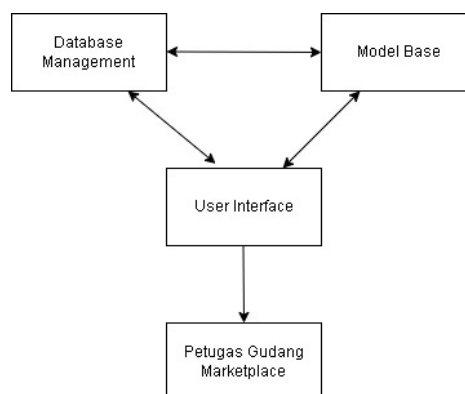
Prosedur pengembangan pada penelitian dan pengembangan ini mengikuti suatu program terstruktur yang digambarkan pada skema Gambar 1. Prosedur ini dimulai dengan Analisis Kebutuhan, yaitu tahapan menganalisis data yang diperlukan untuk prediksi penjualan sparepart mobil, membuat gambaran awal aplikasi, menentukan tujuan pengembangan, serta menetapkan metode K-NN sebagai algoritma inti berdasarkan jurnal yang relevan. Setelah kebutuhan teridentifikasi, dilanjutkan dengan fase Desain Produk, di mana perancangan aplikasi dilakukan untuk memastikan kesesuaian dengan tujuan dan kebutuhan pengguna. Fase berikutnya adalah membangun Prototipe, yaitu perwujudan rancangan menjadi model prototype aplikasi yang dapat diuji coba. Proses Uji Coba kemudian dilakukan kepada ahli sistem, ahli materi, dan pengguna akhir untuk mengetahui keberhasilan dan kesesuaian aplikasi dengan kebutuhan. Tahap Revisi dilakukan berdasarkan saran dan evaluasi dari proses uji coba. Jika prototype masih memiliki kekurangan, proses akan kembali ke tahap Desain Produk. Jika aplikasi telah sesuai dan lolos evaluasi, maka akan menjadi Produk Akhir yang sudah layak untuk digunakan dalam membantu prediksi penjualan sparepart.



Gambar 1 Tahapan Penelitian

2. Metode Penelitian

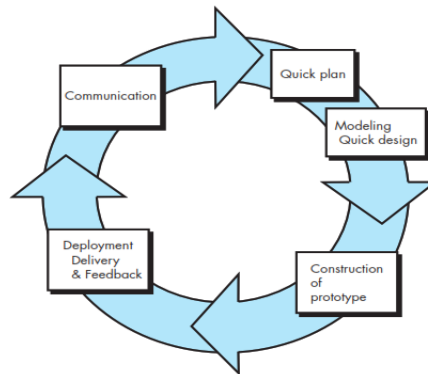
Model teoritis utama yang diterapkan adalah Algoritma K-Nearest Neighbor (K-NN). K-NN merupakan metode supervised learning non-parametrik yang digunakan untuk klasifikasi, di mana prediksi dilakukan berdasarkan kedekatan data baru terhadap K data training terdekat. Dalam konteks ini, K-NN berfungsi mengklasifikasikan sparepart mobil menjadi kategori Laris atau Tidak Laris berdasarkan atribut penjualan historis. Model konseptual yang diterapkan adalah kerangka Sistem Pendukung Keputusan (SPK). SPK digunakan sebagai landasan sistem untuk membantu pihak manajemen gudang dalam pengambilan keputusan terkait stok sparepart berdasarkan hasil prediksi. Model konseptual ini secara umum terdiri dari tiga komponen utama: Database Management, Model Base, dan User Interface, yang secara visual digambarkan sebagai berikut



Gambar 2 Model Konseptual
Sumber: (Sari, 2018)

- Komponen pertama adalah Database Management, yang memiliki fungsi krusial dalam mengorganisasi data yang relevan dengan masalah, yaitu data hasil tes diagnostik peserta didik baru.
- Selanjutnya, Model Base bertugas merepresentasikan permasalahan penentuan pemetaan kompetensi ke dalam format kuantitatif, yang secara esensial memungkinkan penganalisisan solusi alternatif secara komprehensif.
- Komponen terakhir, User Interface, merupakan antarmuka yang bertindak sebagai penghubung dan gabungan antara Database Management dan Model Base, sehingga mampu menampilkan visualisasi sistem yang mudah dimengerti dan dapat dimanfaatkan secara langsung oleh pihak sekolah.

Model procedural yang digunakan dalam penelitian ini yaitu model prototyping.



Gambar 3 Tahapan tahapan Prototyping
 Sumber: (Butsianto & Arifin, 2020)

Model ini sangat efisien karena memfasilitasi interaksi intensif antara pengembang dan pemakai selama proses pembuatan, memungkinkan pengembang memodelkan perangkat lunak dengan mudah dan pengguna mendapatkan gambaran awal serta melakukan pengujian sederhana. Tahapan metode prototyping yang diterapkan dalam penelitian ini melibatkan empat langkah utama. Pertama, Communication, yang dimulai dengan komunikasi untuk menentukan tujuan keseluruhan sistem prediksi, mengidentifikasi persyaratan yang sudah diketahui, dan menguraikan area yang memerlukan definisi lebih lanjut. Kedua, Quick Plan, di mana iterasi prototype direncanakan dengan cepat dan diikuti oleh pemodelan dalam bentuk "desain cepat". Ketiga, Modelling Quick Design & Construction of Prototype, di mana fokus desain cepat adalah pada aspek perangkat lunak yang akan dilihat oleh pengguna akhir, yang kemudian mengarah langsung pada pembangunan prototype. Keempat adalah Deployment Delivery & Feedback, di mana prototype digunakan dan dievaluasi oleh para pemangku kepentingan. Umpan balik yang diberikan digunakan untuk menyempurnakan persyaratan lebih lanjut. Proses iterasi ini berlanjut hingga prototype memenuhi kebutuhan pengguna dan pengembang memiliki pemahaman yang utuh mengenai produk akhir yang harus dihasilkan.

3. Uji coba produk

Instrumen pengumpulan data dalam penelitian ini mencakup dua jenis, yakni uji ahli dan uji pengguna. Pengujian oleh ahli dilakukan menggunakan metode Blackbox Testing, sedangkan pengujian pengguna memanfaatkan Post-Study System Usability Questionnaire (PSSUQ). Rumus yang digunakan adalah sebagai berikut;

$$\text{Persentase Kelayakan}(\%) = \frac{\text{Skor yang diobservasi}}{\text{Skor yang diharapkan}} \times 100\%$$

Perhitungan hasil uji mengacu pada rumus persentase kelayakan. Mengacu pada Arikunto (2009, hlm. 44), tingkat kelayakan dibagi ke dalam lima kategori dengan rentang persentase 0% hingga 100%, di mana 100% menunjukkan nilai tertinggi dan 0% nilai terendah.

Tabel 1 Rentang Kategori Kelayakan

Persentase Pencapaian	Interpretasi
< 21%	Sangat Tidak Layak
21% - 40%	Tidak Layak
41% - 60%	Cukup Layak
61% - 80%	Layak

Persentase Pencapaian	Interpretasi
81% - 100%	Sangat Layak

Sumber : (Arikunto, 2009)

4. Uji Hasil

Performa algoritma K-Nearest Neighbor (K-NN) dievaluasi menggunakan Confusion Matrix, yaitu tabel yang menggambarkan perbandingan antara hasil prediksi model dan nilai aktual. Confusion Matrix menyajikan empat kategori utama: True Positive (TP) dan True Negative (TN), yang menunjukkan prediksi benar masing-masing untuk kelas positif dan negatif, serta False Positive (FP) dan False Negative (FN), yang menandai kesalahan prediksi. Dengan memetakan seluruh keluaran klasifikasi ke dalam empat kategori ini, tingkat akurasi dan kesalahan dapat diukur secara menyeluruh. Analisis kuantitatif dilakukan melalui empat metrik evaluasi utama berikut:

- a. Accuracy, mengukur proporsi total prediksi yang benar terhadap keseluruhan data:

$$accuracy = \frac{TP}{\text{jumlah total dataset}}$$

- b. Precision, menunjukkan tingkat ketepatan prediksi positif:

$$Precision = \frac{TP}{TP + FP}$$

- c. Recall, menilai kemampuan model menemukan seluruh data positif yang sebenarnya:

$$Recall = \frac{TP}{TP + FN}$$

- d. F1-Score, memberikan rata-rata harmonis antara precision dan recall:

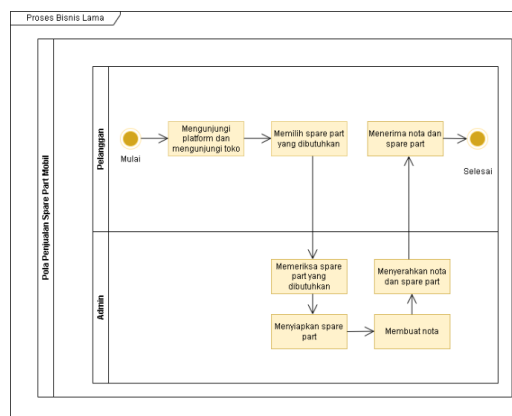
$$f1 - Score = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

Keempat metrik tersebut memberikan gambaran menyeluruh mengenai kekuatan dan kelemahan model K-NN, sehingga memudahkan proses evaluasi kinerja dan perbandingan dengan metode klasifikasi lainnya.

C. HASIL DAN PEMBAHASAN

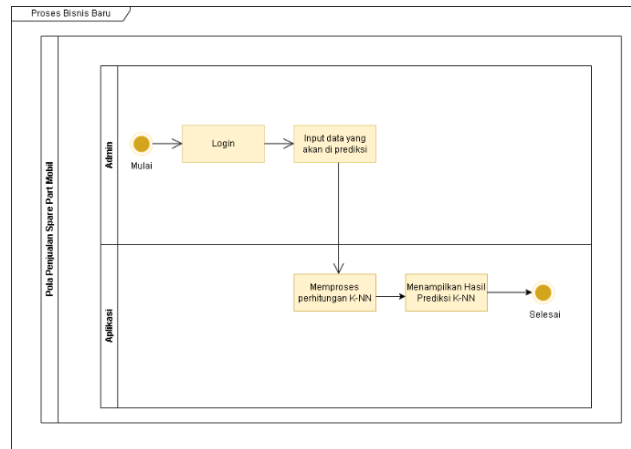
1. Hasil Analisis Proses

Analisis kebutuhan dilakukan untuk merancang sistem prediksi penjualan spare part mobil menggunakan algoritma K-Nearest Neighbor (K-NN). Proses bisnis lama pada gudang masih memiliki kelemahan, yaitu belum adanya sistem prediksi yang efektif sehingga sering terjadi overload pada penyimpanan. Hal ini ditunjukkan pada proses bisnis lama (Gambar 4)



Gambar 4 Proses Bisnis Lama

Sebagai solusi, dirancang proses bisnis baru (Gambar 5) dengan tahapan: pengguna melakukan login, memilih data penjualan yang akan dihitung, menentukan quantity, harga, kondisi, serta merek (brand), kemudian data diproses oleh sistem untuk menghasilkan prediksi penjualan spare part mobil. Hasil prediksi ini membantu pihak gudang dalam mengidentifikasi spare part yang berpotensi laris maupun tidak laris, sehingga pengelolaan stok menjadi lebih optimal.

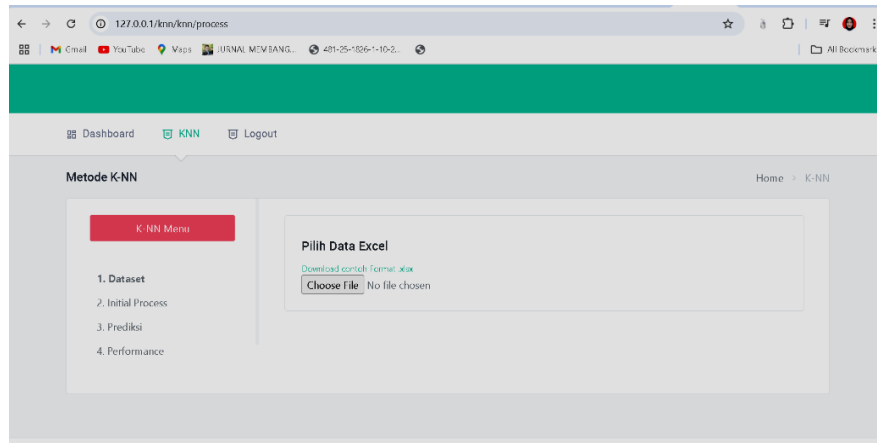


Gambar 5 Proses Bisnis Baru

Penelitian ini menerapkan metode analisis data historis penjualan suku cadang mobil guna membangun model prediksi penjualan yang mendukung pengelolaan stok gudang marketplace. Data yang digunakan meliputi jenis suku cadang, jumlah penjualan per periode, faktor musiman, promosi, dan variabel lain yang berpotensi memengaruhi permintaan. Seluruh data dikumpulkan dari catatan transaksi penjualan pada periode tertentu dan diolah menggunakan pendekatan machine learning berbasis algoritma prediksi, sehingga pola dan tren permintaan dapat diidentifikasi secara lebih akurat dibandingkan metode konvensional. Proses analisis mencakup pembersihan data, eksplorasi pola penjualan, pemodelan prediksi, dan evaluasi kinerja model. Hasil analisis menunjukkan bahwa pemanfaatan data historis mampu memproyeksikan jumlah stok ideal yang perlu disiapkan untuk setiap jenis suku cadang. Prediksi yang dihasilkan membantu gudang marketplace meminimalkan risiko kekurangan stok yang dapat menyebabkan keterlambatan pengiriman, sekaligus menekan potensi kelebihan stok yang meningkatkan biaya penyimpanan dan risiko kerusakan barang. Dengan demikian, metode prediksi penjualan yang diusulkan terbukti mendukung pengambilan keputusan manajemen gudang secara lebih efisien dan efektif.

Tahap awal pengembangan aplikasi prediksi penjualan spare part mobil pada gudang marketplace dilakukan melalui analisis kebutuhan untuk mengidentifikasi permasalahan dan menentukan data yang diperlukan. Data yang dihimpun mencakup nama spare part, histori penjualan tahun 2023, jumlah (quantity), dan harga, yang diperoleh dari operator gudang dengan persetujuan kepala gudang. Hasil pengumpulan menunjukkan banyak spare part tersimpan di gudang tanpa prediksi penjualan yang jelas, sehingga diperlukan sistem yang mampu memberikan proyeksi agar pengelolaan stok lebih terarah. Proses analisis kebutuhan juga dilengkapi dengan wawancara yang dilaksanakan pada 1 Juli 2024 bersama operator gudang, Lily Nur Khikmah. Dari wawancara diketahui bahwa alur kerja saat ini dimulai dari pencatatan permintaan spare part oleh bengkel atau pelanggan, dilanjutkan pengecekan ketersediaan stok, kemudian proses penjualan dan pembuatan invoice. Selanjutnya, operator gudang merekap data penjualan dan memperbarui stok dengan menambah atau mengurangi jumlah sesuai proyeksi permintaan. Apabila stok tidak tersedia, operator akan mengajukan permintaan restock kepada bagian supply chain dengan jumlah tertentu. Temuan ini menjadi dasar perancangan sistem prediksi penjualan berbasis algoritma K-Nearest Neighbor (K-NN) yang diharapkan mampu meningkatkan efisiensi pengelolaan stok serta meminimalkan risiko kelebihan persediaan.

Tampilan ini merepresentasikan keseluruhan fitur utama yang terdapat dalam aplikasi. Setelah berhasil masuk, pengguna diarahkan ke halaman utama yang menyediakan fasilitas import data untuk mengunggah data penjualan yang akan diprediksi. Selanjutnya, aplikasi menampilkan pratinjau data prediksi, yaitu data yang telah diunggah agar dapat diverifikasi sebelum diproses lebih lanjut. Pengguna juga dapat menambahkan data baru melalui form tambah data prediksi dengan mengisi kolom atribut yang diperlukan. Hasil prediksi dapat dilihat pada halaman pratinjau hasil prediksi, di mana data baru yang dimasukkan akan otomatis ditambahkan ke dataset. Terakhir, aplikasi menyediakan halaman pratinjau hasil performa akurasi yang menampilkan tingkat ketepatan model dalam memproses data penjualan. Tampilan halaman utama dapat dilihat pada gambar 7.



Gambar 7 Tampilan Aplikasi

D. PEMBAHASAN

1. Perhitungan Metode K-NN

a. Tahapan Pemahaman Bisnis (Business Understanding)

Permasalahan yang dihadapi adalah ketidakakuratan prediksi penjualan spare part mobil yang berpotensi menimbulkan overstock di gudang. Untuk mengatasinya, penelitian ini mengembangkan prototipe aplikasi dengan algoritma K-NN guna menentukan pola persediaan dan menghasilkan itemset spare part yang paling sering dibeli pelanggan.

b. Tahapan Pemahaman Data (Data Understanding)

Penelitian ini menggunakan 500 data penjualan spare part mobil periode Januari–Desember 2023, dengan 100 data sebagai data uji. Tahap pengolahan mencakup data preparation dan transformasi untuk mendukung proses data mining.

c. Tahap Persiapan Data (Data Preparation)

Tahap persiapan data dilakukan untuk menyiapkan dataset penjualan spare part mobil yang akan dianalisis menggunakan algoritma K-Nearest Neighbor (K-NN). Data awal diperoleh dari catatan transaksi gudang marketplace sepanjang Januari–Desember 2023. Atribut asli mencakup nama spare part, jumlah penjualan per bulan, quantity, harga, kondisi, brand, dan kategori penjualan. Langkah pertama adalah menghapus kolom penjualan bulanan Januari–Desember 2023 agar dataset lebih ringkas. Selanjutnya hanya atribut yang relevan dengan proses pemodelan yang dipertahankan. Contoh hasil pembersihan dan seleksi atribut dapat dilihat pada Tabel 2, yang menampilkan 100 entri spare part dengan berbagai merek seperti Corolla, C-HR, dan Vellfire, lengkap dengan data jumlah penjualan tahunan.

Tabel 2 Data Persiapan

No	Spare Part	Jumlah penju an 1th	Qty	Harga	Kondisi	Brand	Kategori
1	Subwoofer	118	300	Rp 1.700.000	bekas	corolla	T
2	Tweeter	89	300	Rp 1.700.000	bekas	corolla	T
3	Side Airbag	62	300	Rp 1.700.000	bekas	corolla	T
..	..	..	..	..	..	..	..
100	Multifunctio n monitor	257	450	Rp 5.000.000	baru	vellfire	L

Untuk mendukung pemrosesan algoritma, data kategorikal diubah menjadi nilai numerik. Setelah transformasi, dataset uji siap dipakai dalam proses pemodelan K-NN. Hasil transformasi dapat dilihat pada Tabel 3:

Tabel 3 Hasil Transformasi Data

No	Spare Part	Jumlah penjualan 1th	Qty	Harga	Kondisi	Brand	Kategori
1	Subwoofer	118	300	Rp	0	16	T

No	Spare Part	Jumlah penjualan 1th	Qty	Harga	Kondisi	Brand	Kategori
				1.700.000			
2	Tweeter	89	300	Rp 1.700.000	0	16	T
3	Side Airbag	62	300	Rp 1.700.000	0	16	T
..	..	..	..	..	..	..	..
100	Multifunction monitor	257	450	Rp 5.000.000	1	15	L

d. Modelling Phase

Menggunakan algoritma K-Nearest Neighbor (K-NN) untuk memproyeksikan penjualan spare part mobil pada periode mendatang. Variabel yang dianalisis meliputi nama spare part, jumlah penjualan tahunan, quantity, harga, kondisi, brand, dan kategori. Total dataset yang digunakan sebanyak 500 entri. Untuk keperluan pemodelan K-NN, data dibagi menjadi data latih dan data uji. Hasil evaluasi menunjukkan bahwa komposisi 80% data latih dan 20% data uji memberikan tingkat akurasi terbaik dibandingkan komposisi lainnya. Dari total 500 data, 400 data digunakan sebagai data latih dan 100 data sebagai data uji, dengan 20 data representatif dipilih untuk pengujian akhir algoritma K-NN. Contoh data uji yang digunakan dapat dilihat pada Tabel 4.

Tabel 4 Data Uji

No	Spare Part	Jumlah penjualan 1th	Qty	Harga	Kondisi	Brand	Kategori
1	Subwoofer	118	300	0,19	0	16	T
2	Tweeter	89	300	0,19	0	16	T
3	Side Airbag	62	300	0,19	0	16	T
..	..	..	..	..	..		
100	Multifunction monitor	257	450	0,56	1	15	L

e. Menghitung jarak Euclidean

Setiap data uji dihitung jaraknya terhadap data latih menggunakan rumus Euclidean Distance untuk mengetahui kedekatan antar data. Rumus *Euclidean Distance* adalah sebagai berikut :

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

Keterangan:

d : Euclidean Distance

x : Representasi dari satu titik atau data yang terdiri dari jumlah atribut atau fitur

y : Representasi dari titik atau data kedua yang dibandingkan dengan x

Beberapa perhitungan tabel 4 sebagai berikut:

Data ke-1:

$$d = \sqrt{(385 - 329)^2 + (450 - 350)^2 + (0,78 - 0,28)^2 + (1 - 0)^2 + (1 - 9)^2}$$

$$d = 114,90$$

Data ke-2

$$d = \sqrt{(385 - 348)^2 + (450 - 350)^2 + (0,78 - 0,28)^2 + (1 - 0)^2 + (1 - 9)^2}$$

$$d = 106,93$$

Data ke-3

$$d = \sqrt{(385 - 290)^2 + (450 - 350)^2 + (0,78 - 0,28)^2 + (1 - 0)^2 + (1 - 9)^2}$$

$$d = 138,17$$

Nilai K = 2 dipilih sehingga klasifikasi dilakukan berdasarkan dua tetangga terdekat. Kategori tiap spare part ditentukan melalui mayoritas label dari dua tetangga terdekat. Dari 100 data uji, sekitar 30% mengalami perubahan kategori setelah klasifikasi. Contoh perubahan yang menonjol ditunjukkan pada Tabel 5. Beberapa item yang semula berlabel Tidak Laris diprediksi ulang menjadi Laris, seperti:

Tabel 5 Perbandingan Data Asli dengan Data setelah perhitungan

No	Spare Part	Kategori Lama	Kategori Baru
1	Subwoofer	T	L
2	Tweeter	T	L
3	Side Airbag	T	L
..	..	..	..
100	Multifunction monitor	L	L

f. Evaluasi Model K-NN Menggunakan Confusion Matrix

Tahap evaluasi dilakukan untuk mengukur kinerja model klasifikasi K-NN dengan Confusion Matrix. Pengujian dilakukan pada 100 data uji, dengan hasil perbandingan kategori asli dan prediksi ditampilkan pada Tabel 6.

Tabel 6 Confusion Matrix

	True L	True T	Class Precission
Pred L	59 (TP)	21 (FN)	86%
Pred T	7 (FP)	13 (TN)	86%
Class Recall	100%	100%	

Berdasarkan tabel di atas, diperoleh metrik evaluasi sebagai berikut:

a.Accuracy

$$akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

$$akurasi = \frac{59 + 13}{29 + 13 + 7 + 21}$$

$$akurasi = 0.72 (72\%)$$

b.Precision

$$Precision = \frac{TP}{TP + FP}$$

$$Precios = \frac{59}{59 + 7}$$

$$Precision = 0.89 (89\%)$$

c.Recall

$$Recall = \frac{TP}{TP + FN}$$

$$Recall = \frac{59}{59 + 21}$$

$$Recall = 0.74 (74\%)$$

d.F1 Score

$$F1 - Score1 = 2 \times \frac{Precision \times Recall}{Precision + Recal}$$

$$F1 - Score1 = 2 \times \frac{0.89 \times 0.74}{0.89 + 0.74}$$

$$F1 - Score = 0.81(81\%)$$

Hasil ini menunjukkan bahwa algoritma K-NN mampu memberikan ketepatan prediksi sebesar 72 %, dengan precision 89 %, recall 74 %, dan F1-score 81 %. Nilai-nilai tersebut menandakan bahwa model cukup andal dalam mengklasifikasikan spare part ke kategori Laris maupun Tidak Laris, meskipun masih terdapat beberapa kesalahan klasifikasi terutama pada data yang sebenarnya Laris namun diprediksi Tidak Laris.

E. KESIMPULAN

Penelitian ini telah berhasil mengembangkan prototipe aplikasi berbasis algoritma K-Nearest Neighbor (K-NN) untuk memprediksi penjualan dan menentukan prioritas stok spare part mobil di gudang marketplace. Hasil penelitian menunjukkan bahwa penerapan algoritma K-NN terbukti mampu memprediksi penjualan dengan akurasi yang sangat baik melalui klasifikasi item yang dievaluasi menggunakan confusion matrix. Prototipe aplikasi yang dikembangkan memudahkan identifikasi pola penjualan dan mendukung pengelolaan persediaan dengan lebih efektif. Evaluasi kinerja model menunjukkan bahwa metode K-NN efektif digunakan untuk proses prediksi penjualan, sehingga dapat menjadi solusi yang tepat untuk meningkatkan efisiensi dan efektivitas pengelolaan persediaan di marketplace. Selain itu, hasil uji pengguna dengan kuesioner PSSUQ menunjukkan bahwa aplikasi ini layak digunakan dengan skor kepuasan keseluruhan 83,04%. Secara spesifik, aplikasi ini dinilai memiliki kegunaan sistem yang baik dengan skor 80,95%, kualitas informasi yang sangat baik dengan skor 90,48%, dan kualitas antarmuka yang baik dengan skor 75%. Dengan demikian, penelitian ini dapat disimpulkan bahwa algoritma K-NN dapat digunakan sebagai metode yang efektif untuk memprediksi penjualan dan menentukan prioritas stok spare part mobil di gudang marketplace. Pengembangan prototipe aplikasi berbasis K-NN ini dapat menjadi kontribusi yang signifikan bagi pengelolaan persediaan di marketplace dan dapat digunakan sebagai acuan untuk penelitian lanjutan.

F. DAFTAR PUSTAKA

- [1] Abdy, S., Gultom, E. R. B., Ramadhany, S., & Afifudin, A. (2022). Prediksi Penjualan Sparepart Mobil Terlaris Menggunakan Metode K-Nearest Neighbor. *JURIKOM (Jurnal Riset Komputer)*, 9(6), 2003-2009.
- [2] Anonim. (2022). *BAB II: Tinjauan Pustaka dan Dasar Teori*. Yogyakarta: Universitas Teknologi Digital Indonesia.
- [3] Ardiansyah, A., Mahendra, G. S., Rahayu, P. W., et al. (2023). *Buku Ajar Sistem Pendukung Keputusan*. PT Sonpedia Publishing Indonesia.
- [4] Arhami, M., & Nasir, M. (2021). *Data Mining: Algoritma dan Implementasi*. Penerbit Andi.
- [5] Elisa, E., Tukino, T., & Handoko, K. (2022). Penerapan Forecasting Methods Untuk Penjualan Produk Umkm Dengan Algoritma K-Nearest Neighbor. *Jurnal Tekinkom (Teknik Informasi dan Komputer)*, 5(2), 455-463.
- [6] Hasibuan, A. M., & Sriani, S. (2024). Klasifikasi Kelayakan Peminjaman Nasabah Menggunakan Algoritma K-Nearest Neighbor (K-NN): Classification of Customer Loan Eligibility Using The K-Nearest Neighbor (K-NN) Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(4), 1525-1532.
- [7] Iku, M. H., Mustofa, Y. A., & Kumala, I. S. (2019). Metode k-nearest neighbor untuk memprediksi harga eceran beras di pasar tradisional gorontalo. *Jurnal Cosphi*, 3(2).
- [8] Johan, A. (2020). *Data Mining: Algoritma dan Implementasi*. Deepublish.
- [9] Khairunnisa, Nurhadi, N., Jatmiko, A. R., et al. (2023). *Buku Ajar Logika & Algoritma*. PT Sonpedia Publishing Indonesia.
- [10] Maesaroh, S., Iskandar, D., Sari, M. M., et al. (2023). *Rekayasa Perangkat Lunak*. Sada Kurnia Pustaka.
- [11] Mesran, Windarto, A. P., Wanto, A., et al. (2020). *Sistem Pendukung Keputusan & Data Mining: Metode dan Penerapannya Dalam Berbagai Kasus*. Green Press.
- [12] Nijunnihayah, U., Hilabi, S. S., Nurapriani, F., & Novalia, E. (2024). Implementasi Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Alat Kesehatan pada Media Alkes: Implementation of the K-Nearest Neighbor Algorithm to Predict Sales of Medical Devices in Medical Devices. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 695-701.

- [13] Prayoga, I. F. (2023). Sistem Pendukung Penerimaan Beasiswa Sekolah Menggunakan Metode K-Nearest Neighbor (KNN) di SMA PGRI. *Jurnal Teknologi Pintar*, 3(10).
- [14] Putri, D. P. (2021). *Sistem Prediksi Cuaca Menggunakan Metode K-Nearest Neighbor*. Skripsi, Universitas Batam.
- [15] Rafsyanzani S, M. F. (2021). *Implementasi Metode K-Nearest Neighbor Dalam Peramalan Penjualan Mobil Bekas Di Kota Batam* (Doctoral dissertation, Prodi Teknik Informatika).
- [16] Ramadhani, F. D., & Ardiansyah, M. (2022). *Sistem Prediksi Penjualan dengan Metode Single Exponential Smoothing dan Prototype*.
- [17] Raymana Aprilian, Roni Habibi, M Yusril Helmi Setyawan (2020). *Algoritma KNN dalam Memprediksi Cuaca untuk Menentukan Tanaman yang Cocok*. Jakarta: Kreatif.
- [18] Sabda, M. A., & Suhardi, S. (2023). Implementasi Data Mining Dalam Memprediksi Penjualan Parfum Terlaris Menggunakan Metode K-Nearest Neighbor. *Jurnal Sistem Komputer dan Informatika (JSON)*, 5(2), 415-422.
- [19] Saputro, B. (2020). *Manajemen Penelitian Pengembangan (Research & Development) Bagi Penyusun Skripsi, Tesis, dan Disertasi*. Aswaja Presindo.
- [20] Sembiring, F. (2022). *Buku Ajar Dasar Pemrograman (Python)*. Nusa Putra Press.
- [21] Setyobudhi, C. T. (2023). *Buku Ajar Pemrograman Struktur Data*. Mega Press Nusantara.
- [22] Siagian, S. B. (2022). *Implementasi Gain Ratio Pada Metode KNN Dalam Memprediksi Penjualan Sparepart Electronic Di Service Center Panasonic Lhokseumawe* (Doctoral dissertation, Universitas Islam Negeri Sumatera Utara).
- [23] Sutrisno, D., & Farizky, M. (2021). *Implementasi Metode K-Nearest Neighbor (KNN) untuk Prediksi Cuaca Berbasis Android*. *Prosiding Seminar Nasional Energi dan Teknologi Informasi (SENESTIK)*, Institut Teknologi Adhi Tama Surabaya.
- [24] Virza, V. P., Pranot, G. T., & Putra, F. E. (2023). Klasifikasi Kebutuhan Sparepart Dengan Algoritma K-Nearest Neighbor Untuk Meningkatkan Penjualan Sparepart. *Bulletin of Information Technology (BIT)*, 4(3), 287-293.
- [25] Wahana Komputer. (2019). *Pemrograman Web dengan Menggunakan PHP dan MySQL: Edisi Revisi*. Andi Offset.
https://www.google.co.id/books/edition/Pemrograman_Web_dengan_Menggunakan_PHP_d/KuKTDwAAQB_AJ?hl=id
- [26] Whendasmoro, R. G., & Joseph, J. (2022). Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 872-876.
- [27] Pemerintah Kota. *Perwali No. 13 Tahun 2024 tentang Pedoman Penerimaan Peserta Didik Baru*. Pemerintah Kota, 2024.
- [28] Pramudya, Aditya, et al. "Pengelompokan Hasil Belajar Siswa SDN Tunas Jaya dengan Algoritma K-Means." *JATI (Jurnal Mahasiswa Teknik Informatika)*, Dec. 2023.
- [29] Pressman, Roger, and Bruce Maxim. *Software Engineering: A Practitioner's Approach*. McGraw-Hill, 2019.
- [30] Riani, Adella Putri, et al. "Penerapan K-Means Clustering dalam Pengelompokan Hasil Belajar Peserta Didik dengan Metode Elbow." *Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD*, vol. 4, no. 2, 2023, pp. 164-172.
- [31] Riadi, and Muchlisin. www.kajianpustaka.com/2022/02/sistem-pendukung-keputusan-spk.html. diakses pada tanggal 18 Aug. 2024.
- [32] Rosa, A. S., and M. Shalahuddin. *Rekayasa Perangkat Lunak: Terstruktur dan Berorientasi Objek*. Informatika Bandung, 2013.
- [33] Safiq, Rosad. *Bahasa Pemrograman Java Tingkat Dasar*. Nas Media Pustaka, 2022.
- [34] Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta, 2013.
- [35] Wali, Muhammad, et al. *Pengantar 15 Bahasa Pemrograman Terbaik di Masa Depan (Referensi & Coding)*. Sonpedia Publishing, 2023.
- [36] Wu, Junjie. *Advances in K-Means Clustering: A Data Mining Thinking*. Springer-Verlag, 2012.