

PENERAPAN METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI PENJUALAN SEPEDA MOTOR TERLARIS

Rismala¹, Irfan Ali², Ade Rizki Rinaldi³

¹ Program Studi Komputerisasi Akuntansi D3, STMIK IKMI Cirebon
^{2,3} Program Studi S1 Rekayasa Perangkat Lunak, STMIK IKMI Cirebon
 Jl. Perjuangan No 10B, Majasem, Kec. Kesambi, Cirebon, Indonesia
 rismala71002@gmail.com

ABSTRAK

Penggunaan sepeda motor di Indonesia setiap tahunnya selalu meningkat. Seiring dengan meningkatnya jumlah kebutuhan akan alat transportasi, banyak produsen sepeda motor yang mengeluarkan produk dengan bermacam-macam merek dan desain untuk memenuhi permintaan konsumen. Suatu strategi untuk meningkatkan penjualan produk yang akan dijual sangat dibutuhkan pengembang karena tingkat persaingan di dunia bisnis, terutama di industri penjualan, salah satunya dengan memanfaatkan data penjualan setiap harinya. Guna mengetahui produk mana yang paling banyak di jual (laris) dan produk mana yang kurang dalam penjualannya (tidak laris). Penelitian ini melakukan proses data mining pada data penjualan dari PT. Sumber Rejeki Jabar dari bulan Januari – Desember tahun 2022. Hasil nilai akurasi terhadap data penjualan sepeda motor dari bulan Januari - Desember tahun 2022 di PT. Sumber Rejeki Jabar dengan nilai K sebesar 5 yaitu sebesar 96,15%.

Kata kunci: Sepeda Motor, Data Mining, K-Nearest Neighbor

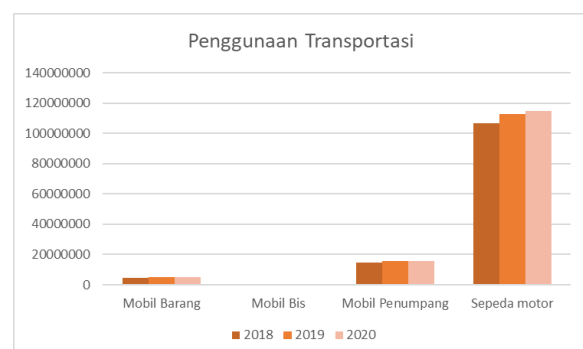
1. PENDAHULUAN

Semakin banyak penduduk dalam suatu daerah, semakin besar kebutuhan akan transportasi untuk memenuhi kebutuhan mobilitas.[1] kebutuhan transportasi tidak hanya meliputi transportasi umum seperti bus, kereta api, dan taksi, tetapi juga kendaraan pribadi seperti mobil dan sepeda motor. Terutama penggunaan sepeda motor karena dengan ukurannya yang kecil, cepat, dan harganya tidak terlalu mahal dibandingkan dengan transportasi lainnya. Dari kasus tersebut menjadi salah satu faktor pendukung dalam penjualan sepeda motor di seluruh dealer di Indonesia, salah satunya di PT. X yang merupakan dealer resmi sepeda motor Honda. Persaingan dalam dunia bisnis menuntut para pengembang untuk menemukan suatu strategi yang dapat meningkatkan akan produk yang dijual, pemanfaatan data penjualan setiap harinya dapat menjadi strategi yang efektif untuk meningkatkan penjualan dan pemasaran produk.[2] Dengan mengumpulkan dan menganalisis data penjualan setiap harinya, para pengembang dapat memahami perilaku dan preferensi konsumen, tren pasar, dan kinerja produk.

Dari permasalahan tersebut dapat dilihat dari data penjualan mengenai jenis unit sepeda motor apa yang paling banyak terjual atau tidak. Dengan data tersebut bisa menjadi acuan perusahaan untuk mengetahui minat *customer* dalam penjualan sepeda motor. Maka penelitian ini dapat digunakan algoritma *K-Nearest Neighbor* untuk membantu dalam memprediksi penjualan pada PT. X untuk menentukan banyaknya sepeda motor yang terjual (laris) atau tidak (tidak laris). Data mining adalah suatu proses untuk mengeksplorasi dan menganalisis data untuk mendapatkan pola-pola tersembunyi atau informasi penting yang dapat digunakan untuk membuat

keputusan atau mengambil tindakan yang lebih baik.[3] data mining melibatkan penggunaan Teknik-teknik analisis data yang canggih untuk mengekstraksi informasi dari data yang sangat besar, beragam, dan kompleks.

Berdasarkan penelitian terdahulu yang dilakukan oleh Sri Puspita Dewi dkk [4] dengan judul “Penerapan Data Mining Untuk prediksi Penjualan Produk terlaris Menggunakan metode *K-Nearest Neighbor*”. Dengan hasil prediksi yang menunjukkan bahwa prediksi menggunakan metode *K-Nearest Neighbor* berhasil ditetapkan, didapatkan dari 25 data testing memperoleh hasil dengan nilai $k = 7$ memiliki keputusan yaitu terlaris = 4 dan tidak laris = 3. Dari penelitian tersebut bisa menjadi salah satu acuan dalam tugas akhir ini untuk menentukan penjualan sepeda motor masuk kedalam kategori laris ataupun tidak laris. Selain itu dilihat dari penelitian tersebut menunjukkan bahwa pemilihan algoritma K-NN ini menjadi salah satu cara yang paling mudah untuk digunakan dalam tugas akhir ini mengenai penentuan label laris dan tidak laris pada penjualan sepeda motor di PT. Sumber Rejeki Jabar Cirebon.



Gambar 1. Data Penggunaan Transportasi

Menurut data dari Badan Pusat Statistik jumlah penggunaan transportasi pada setiap tahunnya selalu bertambah, terutama penggunaan kendaraan bermotor. Dari data tersebut menjadi salah satu acuan pada data penjualan motor di PT. X, bahwasannya penjualan pada perusahaan pun mengalami peningkatan pada setiap bulannya dari semua jenis tipe sepeda motor yang ada. Maka dari itu perusahaan memerlukan prediksi pada penjualan sepeda motor sesuai dengan kategori laris ataupun tidak laris, guna mengetahui tipe motor seperti apa yang paling banyak diminati masyarakat saat ini. Dengan metode yang digunakan untuk prediksi ialah metode *K-Nearest Neighbor* (K-NN). Dikarenakan K-NN memiliki beberapa kelebihan yaitu Tangguh terhadap *training data* yang *noise*, tidak memerlukan *pre-processing* data yang rumit, mudah diimplementasikan dan dipahami dan efektif apabila data latihnya besar.[5] Dengan metode tersebut dapat membantu dalam penentuan total akurasi pada label laris dan tidak laris nya sepeda motor pada PT. X. Sehingga pihak perusahaan dapat mengetahui nilai akurasi pada penjualannya.

2. TINJAUAN PUSTAKA

2.1. Data

Sebagai informasi, data harus diolah lebih lanjut menggunakan model.[6] Data dapat didefinisikan sebagai kumpulan item yang belum diproses atau belum diolah. Data memiliki berbagai jenis, seperti : Data numerik, Data kategorikal dan Data Ordinal, Data teks, Data gambar, Data audio dan Data video.

Semua jenis data dapat diolah dan dianalisis menggunakan Teknik-teknik pembelajaran mesin atau *data science* untuk menghasilkan informasi yang berguna dan berarti. Dengan demikian, diperlukan metode pemrosesan, seperti perhitungan dan pengukuran data yang tersedia, untuk menafsirkannya.[7]

2.2. Data Mining

Data mining adalah proses menganalisis data untuk mengidentifikasi korelasi yang berbeda dan menarik kesimpulan tentang apa yang tidak diketahui sehingga pemilik data dapat memahami dan berguna bagi pemilik data.[3]

Data mining adalah metode untuk menganalisis pengetahuan secara mandiri dengan menggunakan satu atau lebih algoritma pembelajaran komputer.[8]

2.3. Prediksi

Prediksi adalah proses memperkirakan sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi yang tersedia saat ini, untuk mengurangi kesalahan (ketidaksesuaian antara apa yang sebenarnya terjadi dengan hasil yang diprediksi).[9] Dalam konteks *data science*, prediksi sering dilakukan dengan menggunakan algoritma yang dilatih pada data historis atau data yang dikumpulkan sebelumnya.

2.4. K-Nearest Neighbor

Salah satu metode yang menggunakan algoritma terpandu adalah *K-Nearest Neighbor* (K-NN). Dimana hasil kueri *instance* baru dikategorikan sesuai dengan sebagian besar kategori K-NN. Pendekatan K-NN mengklasifikasikan informasi berdasarkan seberapa dekat atau jauh suatu informasi dari informasi lain.[10]

Metode K-NN cukup sederhana namun sangat akurat. Tujuan algoritma ini adalah untuk mengklasifikasikan objek baru menggunakan atribut dan data pelatihan. Jalur pendek antara sampel uji dan sampel pelatihan berfungsi sebagai dasar untuk metode K-NN langsung yang menghitung K-NN. Mayoritas K-NN kemudian digunakan sebagai prediksi dari sampel uji setelah dikumpulkan.[11]

2.5. Rapid Miner

Salah satu program untuk memproses penambahan data disebut *RapidMiner*. Analisis teks adalah fokus dari pekerjaan yang dilakukan oleh *RapidMiner text mining*, yang melibatkan penggalian pola dari kumpulan data besar dan menggabungkannya dengan Teknik statistik, kecerdasan buatan, dan basis data.[12] Tujuan dari analisis teks ini adalah untuk menghasilkan informasi yang akurat dan bermakna dari teks yang sangat besar atau kompleks.

RapidMiner adalah perangkat lunak open source yang kuat untuk knowledge discovery dan data mining yang menawarkan berbagai fitur yang berguna.[13] Beberapa fitur utama RapidMiner meliputi:

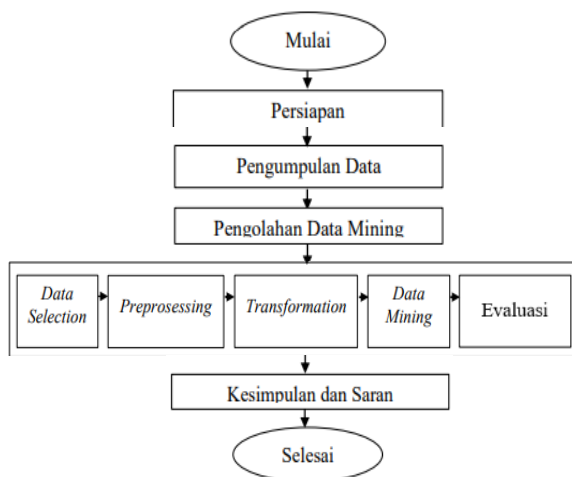
1. Berlisensi gratis (open source): RapidMiner tersedia secara gratis untuk penggunaan non-komersial dan komersial, dan tersedia dalam bentuk kode sumber yang dapat diakses oleh siapa saja.
2. Multiplatform karena deprogram dalam bahasa java: RapidMiner ditulis dalam bahasa pemrograman Java, yang memungkinkannya untuk berjalan pada berbagai platform, termasuk Windows, Mac OS X, dan Linux.
3. Internal data yang berbasis XML, yang memfasilitasi pertukaran data eksperimen: Data dalam RapidMiner disimpan dalam format XML yang fleksibel, sehingga memudahkan pertukaran data antara pengguna dan memungkinkan integrasi dengan alat lain.
4. Berisi scripting language untuk mengotomatisasi eksperimen: RapidMiner menyertakan bahasa scripting sendiri, yang memungkinkan pengguna untuk mengotomatisasi eksperimen dan memperluas fungsionalitas.
5. Memiliki GUI (Graphical User Interface), command line mode (batch mode), dan java API yang dapat dipanggil dari program lain: RapidMiner memiliki antarmuka pengguna grafis yang intuitif dan mudah digunakan, tetapi juga menyediakan mode baris perintah dan API Java

untuk memungkinkan integrasi yang lebih kuat dengan alat lain.

6. Dapat diperluas dengan menyertakan ekstensi dan plugin: RapidMiner dapat diperluas dengan mudah melalui penyertaan ekstensi dan plugin, memungkinkan pengguna untuk menyesuaikan fungsionalitas sesuai dengan kebutuhan mereka.
7. Alat plotting untuk model dan tampil data multidimensi: RapidMiner menyertakan alat plotting yang kuat untuk memvisualisasikan model dan data multidimensi, yang membantu pengguna untuk memahami dan menganalisis data dengan lebih baik.

3. METODE PENELITIAN

Tahapan dalam proses penyelesaian permasalahan bertujuan agar penelitian tersebut menghasilkan data yang benar dan tepat. Tahapan penelitian dapat dilihat pada Gambar 2 di bawah ini.



Gambar 2. Tahapan Penelitian

Tahapan penelitian ini dituangkan dalam diagram alir ini. Mencakup tahapan yang akan diambil dalam penelitian secara keseluruhan. Tahapan yang akan dilakukan yaitu:

3.1. Persiapan

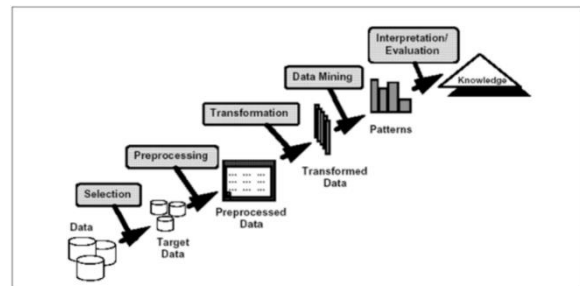
Tahap ini adalah tahapan yang menjadi subjek (populasi) PT. X, objek yang diambil yaitu sepeda motor.

3.2. pengumpulan data

Merupakan proses memperoleh Informasi atau fakta yang relevan dengan tujuan tertentu. Data yang diperoleh merupakan data primer yaitu data yang dikumpulkan dan diolah sendiri dari PT. X. Data yang digunakan adalah data penjualan dari Bulan Januari – Desember Tahun 2022. Selain itu observasi dan wawancara pun digunakan untuk memperoleh data.

3.3. Pengolahan data mining

Tahapan *Knowledge Discovery in Database* (KDD) digunakan untuk mengolah data yang terkumpul.



Gambar 3. Proses Knowledge Discovery in Database (KDD)

3.4. Data Selection

data operasional yang akan digunakan perlu adanya pemilihan sebelum tahapan penggalian informasi dalam KDD dimulai tahapan ini ada pada tahapan *data selection*. Data yang digunakan adalah data harian penjualan dari Bulan Januari – Desember Tahun 2022 pada PT. X.

Tabel 1. Data penjualan Harian

No	Tgl Pembelian	Customer	Beat Sporty Cbs	Beat Cbs iss Deluxe	Vario 125 Cbs	Total
1	03-01-2022	Suhardiman		1		1
2	03-01-2022	Asmat			1	1
3	03-01-2022	Sanuri	1			1

3.5. Pre-processing Data / Cleaning

Tahapan dalam pengolahan data yang bertujuan untuk membersihkan, menormalkan, dan mengubah format data agar mudah dipahami dan digunakan untuk analisis.

3.6. Transformation

Tahapan *Transformation* dimana data yang dipilih diubah dengan melakukan ringkasan (agregasi). Sehingga data tersebut dapat digunakan untuk data mining.

3.7. Data Mining

Merupakan tahapan dalam proses KDD yang bertujuan untuk mencari pola atau informasi yang berguna dari data yang telah dipilih dan diolah. Metode yang digunakan dalam penelitian ini adalah metode *K-Nearest Neighbor* (K-NN). Untuk mengukur performa dari algoritma K-NN menggunakan empat metode evaluasi yaitu:

- a. *Accuracy*, merupakan rasio prediksi secara keseluruhan. Dinyatakan sebagai rasio benar (*true positif* dan *true negative*) dengan total yang dievaluasi, dapat dihitung menggunakan persamaan

$$Accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \times 100\% \quad (1)$$

- b. *Precision*, merupakan rasio prediksi yang mengukur seberapa banyak prediksi positif yang benar dari seluruh hasil yang diprediksi positif oleh model. Selain itu didefinisikan sebagai rasio antara *true positive* dan jumlah seluruh hasil yang diprediksi positif oleh model. Persamaannya sebagai berikut (2).

$$Precision = \frac{TP+TN}{(TP+FP+FN+TN)} \times 100\% \quad (2)$$

- c. *Recall*, merupakan rasio prediksi yang mengukur seberapa banyak data yang bernilai positif dapat terdeteksi oleh model. Dapat didefinisikan juga sebagai rasio antara *true positive* dan jumlah seluruh data yang sebenarnya bernilai positif. Persamaannya yaitu (3).

$$Recall = \frac{TP+TN}{(TP+FP+FN+TN)} \times 100\% \quad (3)$$

- d. *Specificity*, merupakan kebenaran memprediksi perbandingan dengan keseluruhan data negatif menggunakan persamaan (4).

$$Specificity = \frac{TN}{(TP+FP+FN+TN)} \times 100\% \quad (4)$$

3.8. Evaluasi

Merupakan proses untuk mengevaluasi kualitas model yang telah dibuat berdasarkan data yang telah diproses oleh algoritma data mining. Pada tahapan ini didapatkan pola penjualan sepeda motor dari proses data mining dengan metode *K-Nearest Neighbor*, pola atau informasi yang dihasilkan dari proses *data mining* adalah berupa *rules* yang didapat dari perhitungan *K-Nearest Neighbor*.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Penggunaan data harian penjualan di rekap menjadi data bulanan, karena data yang akan dipakai dalam proses *data mining* adalah data di setiap bulannya. Data penjualan pun dikelompokkan menjadi dua kategori yaitu Laris dan Tidak Laris, dengan ketentuan penjualan yang lebih dari sama dengan 3 itu masuk kedalam kategori laris.

Tabel 2 Tabel Data Selection

Tipe	Bulan	QTY	Keterangan
Beat Sporty Cbs	Januari	5	Laris
Beat Cbs iss Deluxe	Januari	18	Laris
Vario 125 Cbs	Januari	10	Laris
Adv 160 Cbs	Januari	1	Tidak Laris

4.2. Pre-processing Data / Cleaning

Dalam tahapan ini tidak ada penggabungan atau pembersihan data karena data tidak memiliki *missing value* atau bernilai 0 (no).

Name	Type	Missing	Statistics
KETERANGAN	Nominal	0	Laris Tidak Laris
TIPE	Nominal	0	VARIO 150 CBS (12) ADV 160 ABS (12) ADV 160 ABS (12) ADV 160 ABS2 (12) [27 more]
BULAN	Nominal	0	SEPTEMBER (29) AGUSTUS (29) AGUSTUS (29) APRIL (29) [10 more]
QTY	Integer	0	45 45 45 45 45

Gambar 4. Preprocessing and Cleaning Data

4.3. K-Nearest Neighbor

Pengujian metode menggunakan *Confusion Matrix* setelah dilakukan perhitungan *Accuracy*, *Precision*, *Recall*, dan *Specificity*. Dapat disimpulkan hasil yang didapatkan dapat dilihat pada Tabel 3.

Tabel 3. Pengujian Metode Menggunakan Confusion Matrix

Accuracy	Precision	Recall	Specificity
96,15%	94,20%	100%	89,74%

Pengukuran *Performance* :

a. Accuracy

$$Akurasi = \frac{TP+TN}{TP+FP+FN+TN} \times 100\%$$

$$Akurasi = \frac{65+35}{65+4+0+35} \times 100\%$$

$$Akurasi = 96,15\%$$

b. Precision

$$Precision = \frac{TP}{TP+FP} \times 100\%$$

$$Precision = \frac{65}{65+4} \times 100\%$$

$$Precision = 94,20\%$$

c. Recall

$$Recall = \frac{TP}{TP+FN} \times 100\%$$

$$Recall = \frac{65}{65+0} \times 100\%$$

$$Recall = 100\%$$

d. Specificity

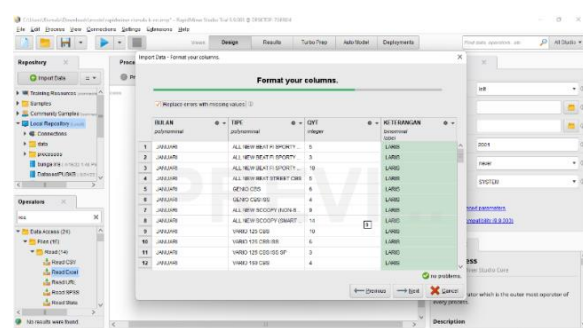
$$Specificity = \frac{TN}{TN+FP} \times 100\%$$

$$Specificity = \frac{35}{35+4} \times 100\%$$

$$Specificity = 89,74$$

4.4. Penerapan Algoritma K-NN dengan Rapidminer

Pada tahapan ini diawal dengan pemilihan data lalu penentuan label kelas, dapat dilihat pada Gambar 3

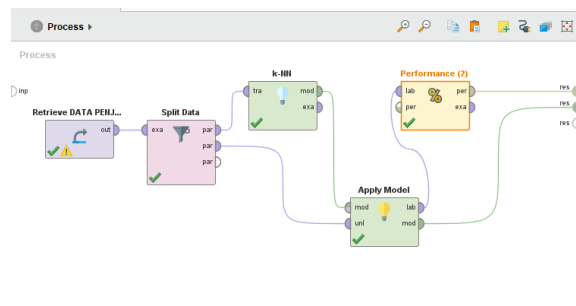


Gambar 5. Penentuan Label Kelas

Model yang digunakan sesuai dengan proses algoritma K-NN yang terdiri dari operator retrieve data set, operator split data sebagai tahapan transformation

untuk mengatur data training sebesar 70% dan data testing sebesar 30% dari total seluruh data set, operator K-NN sebagai bentuk proses algoritma K-NN itu sendiri, untuk menggabungkan operator K-NN dan split data menjadi satu kesatuan model dan operator performance untuk mengukur akurasi kinerja dari model K-NN yang telah terbentuk digunakan operator apply model.

Setelah model K-NN diterapkan pada data testing, Langkah terakhir adalah mengukur performa model menggunakan operator *performance*. Operator ini digunakan untuk menghitung nilai akurasi kinerja dari model K-NN yang telah terbentuk dengan membandingkan hasil prediksi model dengan nilai sebenarnya pada data *testing*.



Gambar 6. Model K-NN

Gambar 7 Pembagian Data Training dan Data Testing

Dapat dilihat hasil uji coba menggunakan algoritma K-NN dengan performa prediksi dikatakan laris 65 *true* laris dan 4 *true* tidak laris dengan *class precision* nya 94,20%, sedangkan prediksi dari tidak laris yaitu *true* laris 0 dan 35 *true* tidak laris dengan *class precision* nya 100%. Sedangkan total *class recall* nya *true* laris yaitu sebesar 100%, sedangkan *true* tidak larisnya sebesar 89,74% dan nilai *accuracy* nya sebesar 96,15%. Dan dengan hasil 244 data sebagai data *Training* dengan dimensi kelasnya laris dan tidak laris.

	pred. LARIS	pred. TIDAK LARIS	class recall
accuracy: 96.15%	65	0	100.00%
	4	35	89.74%
			94.20%
			100.00%

Gambar 8. Hasil Nilai Accuracy

KNNClassification

Weighted 5-Nearest Neighbour model for classification.
The model contains 244 examples with 3 dimensions of the following classes:
LARIS
TIDAK LARIS

Gambar 9. Klasifikasi K-NN

5. KESIMPULAN DAN SARAN

Berdasarkan uji coba yang dilakukan menunjukkan bahwa algoritma K-NN dapat digunakan dalam klasifikasi data penjualan sepeda motor pada PT. X. Berdasarkan pengujian didapatkan hasil akurasi sebesar 96,15% yang artinya data set dapat digunakan pada tahapan selanjutnya sebagai data yang valid untuk digunakan.

Dengan tingkat akurasi yang cukup tinggi, model K-NN dapat menjadi salah satu solusi untuk memperediksi penjualan sepeda motor berdasarkan data penjualan yang telah ada sebelumnya. Dengan demikian, PT. X dapat menggunakan model ini untuk membantu dalam pengambilan keputusan dan perencanaan strategi pemasaran yang lebih efektif.

Meskipun algoritma K-NN telah berhasil menghasilkan akurasi yang cukup tinggi dalam klasifikasi data penjualan sepeda motor pada PT X, namun agar memperoleh hasil yang lebih baik dari penelitian ini sebaiknya dilakukan penelitian lebih lanjut dengan menerapkan metode lainnya sehingga hasil yang didapatkan bisa lebih variatif dan dapat digunakan sebagai bahan perbandingan pada hasil yang didapatkan pada hasil ini.

DAFTAR PUSTAKA

- [1] R. A. Robin, N. Fadhly, dan A. Achmad, "Analisis Jalur Trans Kutaraja di Kota Banda Aceh di Tinjau Dari Pengembangan Wilayah," *J. Penelit. Transp. Darat*, vol. 23, no. 1, hal. 88–105, 2021, doi: 10.25104/jptd.v23i1.1765.
- [2] C. Anisa1 dan Andri2, "PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK PREDIKSI PENJUALAN OBAT PADA APOTEK KIMIA FARMA ATMO PALEMBANG," *J. Ilmu Komput. Dan ...*, vol. 4, no. 2, hal. 15–23, 2021, [Daring]. Tersedia pada: <http://ejournal.sisfokomtek.org/index.php/jikom/article/view/155>
- [3] Yuli Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5," *J. Edik Inform.*, vol. 2, no. 2, hal. 213–219, 2019.
- [4] S. P. Dewi, N. Nurwati, dan E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, hal. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [5] R. N. Sukmana, Abdurrahman, dan Y. Wicaksono, "Implementasi K-Nearest Neighbor Untuk Menentukan Prediksi Penjualan (Studi Kasus : PT Maksipus Utama Indonesia)," *J. Teknol. Inf. dan Komun. Vol. 8 No. 2, Desember*

- 2020, vol. 8, no. 2, hal. 31–38, 2020.
- [6] Nawassyarif, M. Julkarnain, dan K. Rizki Ananda, “Sistem Informasi Pengolahan Data Ternak Unit Pelaksana Teknis Produksi Dan Kesehatan Hewan Berbasis Web,” *J. Inform. Teknol. dan Sains*, vol. 2, no. 1, hal. 32–39, 2020, doi: 10.51401/jinteks.v2i1.556.
- [7] S. Arlis, “Sistem Keamanan Aktivitas Komputer Anak Berbasis Opensource,” *J. Teknol. Inf. Pendidik.*, vol. 9, no. 3, hal. 1–13, 2020.
- [8] J. Halim, “Penerapan Data Mining Untuk Mengukur Tingkat Kepuasan Siswa Terhadap Pelayanan Di Bimbingan Belajar Al-Misbah Dengan Menggunakan Metode K-Means,” *J. SAINTIKOM*, vol. 16, no. 1, hal. 1–6, 2019, [Daring]. Tersedia pada: <https://prpm.trigunadharma.ac.id/public/fileJurnal/hplH11Jufri.pdf>
- [9] Mukhlisin, M. Imrona, dan D. T. Murdiansyah, “Prediksi Harga Beras Premium dengan Metode Algoritma K-Nearest Neighbor,” *e-Proceeding Eng.*, vol. 7, no. 1, hal. 2714–2724, 2019.
- [10] S. H. Putra dan B. T. Putra, “Klasifikasi Harga Cell Phone menggunakan Metode K-Nearest Neighbor (KNN),” *Pros. Annu. Res. Semin.*, vol. 4, no. 1, hal. 242–245, 2018.
- [11] I. Yolanda dan H. Fahmi, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT . Nippon Indosari Corpindo Tbk Menggunakan Metode K-Nearest Neighbor,” vol. 3, no. 3, hal. 9–15, 2021.
- [12] J. Suntoro, *DATA MINING Algoritma dan Implementasi Menggunakan Bahasa Pemrograman PHP*, vol. 9, no. 9. 2019.
- [13] S. Mulyati, S. M. Husein, dan R. Ramdhan, “Rancang Bangun Aplikasi Data Mining Prediksi Kelulusan Ujian Nasional Menggunakan Algoritma (Knn) K-Nearest Neighbor Dengan Metode Euclidean Distance Pada Smpn 2 Pagedangan,” *JIKA (Jurnal Inform.)*, vol. 4, no. 1, hal. 65, 2020, doi: 10.31000/jika.v4i1.2288.