J-CEKI : Jurnal Cendekia Ilmiah
Vol.4, No.2, Februari 2025

Penerapan Data Mining Untuk Memprediksi Penjualan Produk Terlaris UD Timbul Jaya Menggunakan Metode K-Nearest Neighbor

Muhammad Bakriansyah[1], Ainun Subair[2], Wahyudi Setiawan[2]
[1,2,3] Universitas Trunojoyo Madura
E-mail: bakriansyah44@gmail.com

Article History:
Received: 26 November 2024
Revised: 06 Januari 2025
Accepted: 11 Januari 2025

Keywords: *Data Mining, K Nearest Neighbor, Product*

Abstract: *Dalam era digital serta dengan adanya kemajuan teknologi yang berkembang di berbagai sektor. Volume pengolahan data yang dihasilkan semakin besar dan kompleks dalam berbagai bidang dan organisasi. Pengolahan serta analisis data yang efisien menjadi sangat penting untuk memperoleh suatu informasi berharga dari suatu data yang telah ada. UD TIMBUL JAYA adalah salah satu dari berbagai toko yang bergerak dalam penyediaan bahan bangunan yang terletak di Kecamatan Geger.Toko ini menyedikan berbagai jenis bahan yang terkait dengan aktivitas pembangunan seperti semen , paku , asbes , langit-langit , bata ringan serta bahan lainnya. Menggunakan metode K-Nearest Neighbor dalam penerapan data mining untuk prediksi penjualan, metode ini dapat digunakan untuk mengklasifikasikan data penjualan berdasarkan jumlah penjualan produk yang paling laris.*

PENDAHULUAN

Dalam era digital serta dengan adanya kemajuan teknologi yang berkembang di berbagai sektor. Volume pengolahan data yang dihasilkan semakin besar dan kompleks dalam berbagai bidang dan organisasi. Pengolahan serta analisis data yang efisien menjadi sangat penting untuk memperoleh suatu informasi berharga dari suatu data yang telah ada. Sering kali alat serta teknik analisis data secara tradisional tidak dapat kita gunakan untuk mengolah suatu informasi dengan ukuran data yang banyak dan besar, karenanya diperlukan adanya kontribusi teknologi. Proses untuk memprediksi secara sistematis dan terperinci menjadi sesuatu hal yang paling mungkin terjadi di masa kini dan masa yang akan datang. Proses ini berdasarkan informasi masa lalu dan masa kini yang telah dimiliki, sehingga perbedaan antar sesuatu yang terjadi serta hasil ramalan dapat diminimalkan. Namun prediksi hanya melainkan berusaha mencari jawaban yang semirip mungkin dengan suatu akan terjadi.

UD TIMBUL JAYA adalah salah satu dari berbagai toko yang bergerak dalam penyediaan bahan bangunan yang terletak di Kecamatan Geger. Toko ini menyediakan berbagai jenis bahan yang terkait dengan aktivitas pembangunan seperti semen , paku , asbes , langit-langit , bata ringan serta bahan lainnya. Toko ini setiap hari diharuskan menyediakan dan memenuhi kebutuhan, serta dituntut untuk dapat mengambil keputusan yang baik dan tepat dalam

menentukan strategi penjualan dan penyediaan stok. Dengan menggunakan strategi dalam memanfaat kan Data transaksi penjualan yang tersimpan memungkinkan manajemen untuk memahami kebiasaan belanja pelanggan serta perilaku mereka terhadap produk yang sering dibeli. Dari data tersebut, akan diperoleh wawasan atau informasi mengenai produk-produk yang paling diminati oleh konsumen. Dengan adanya pengendalian produk tentunya akan meminimalisir adanya produk yang menumpuk serta mengurangi persentase kerugian dan juga dapat mengelola modal dengan lebih efektif. Oleh karena itu, mengelola data transaksi sangatlah penting untuk membantu manajemen dalam merencanakan persediaan stok. Untuk mengelola data transaksi, dibutuhkan penerapan *data mining* dan Algoritma Regresi Linier dengan memanfaatkan data historis guna mengetahui produk yang paling laris di Toko UD TIMBUL JAYA. Algoritma Regresi Linier adalah salah satu metode yang digunakan untuk memodelkan hubungan antara satu atau lebih variabel dengan variabel lainnya. Teknik ini juga dapat digunakan untuk memprediksi dampak dari hubungan sebab akibat. Berdasarkan beberapa penelitian sebelumnya yang telah menggunakan metode K-Nearest Neighbor dalam penerapan *data mining* untuk prediksi penjualan, metode ini dapat digunakan untuk mengklasifikasikan data penjualan berdasarkan jumlah penjualan produk yang paling laris. Berdasarkan hasil analisis tersebut, penulis melakukan penelitian dengan judul "Penerapan *Data Mining* untuk Prediksi Penjualan Produk Terlaris di UD TIMBUL JAYA Menggunakan Metode K-Nearest Neighbor" guna mempermudah dan membantu pihak UD TIMBUL JAYA dalam menentukan produk apa saja yang perlu disediakan dalam jumlah besar.

METODE PENELITIAN

Jenis Penelitian

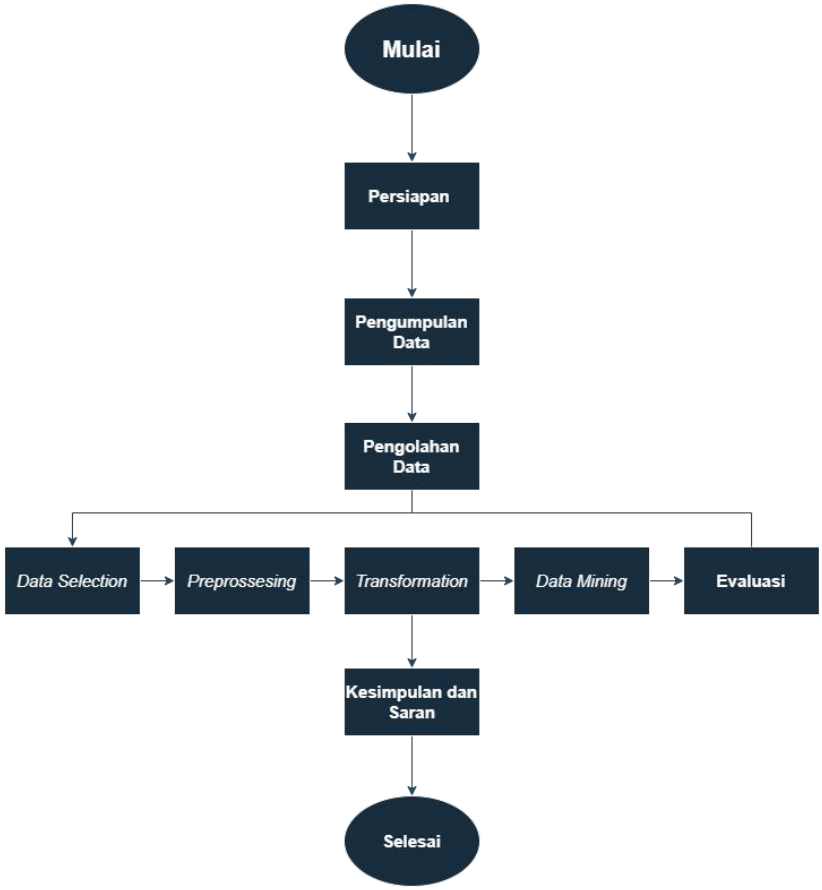
Penelitian ini menggunakan metode kuantitatif dengan pendekatan rumusan masalah deskriptif dan asosiatif, karena variabel-variabel yang diteliti memiliki hubungan serta bertujuan untuk memberikan gambaran yang terstruktur dan faktual mengenai fakta serta hubungan antarvariabel tersebut. Model penelitian kuantitatif ini melibatkan dua variabel, yaitu variabel persediaan produk sebagai variabel bebas yang dilambangkan dengan (X) dan variabel penjualan sebagai variabel terikat yang dilambangkan dengan (Y).

Objek Penelitian

Objek penelitian ini merupakan UD TIMBUL JAYA yang bergerak dalam bidang penyedia bahan bangunan,dimana data yang akan di gunakan adalah data penjualan yang telah diperoleh dari beberapa narasumber seperti pemilik UD TIMBUL JAYA tersebut serta beberapa karyawan UD TIMBUL JAYA melalui metode observasi dan wawancara.

Tahapan Penelitian

Untuk tahapan penelitian ini perlu disusun tahapan kerja yang jelas dan terstruktur guna mempermudah pemecahan masalah.Kerangka kerja ini mewakili langkah yang telah tersusun secara sistematis,Seperti yang telah termuat dalam gambar 1.



Gambar 1. Tahapan Penelitian

1. Identifikasi Masalah.
2. Pengumpulan data.
3. Pengolahan Data
4. Kesimpulan dan Saran

Analisis Data

Analis data ialah suatu metode untuk mengolah data menjadi suatu bentuk informasi untuk membantu memahami suatu karakteristik data dan memecahkan suatu masalah, terutama yang berkaitan dengan suatu penelitian. Pada tabel 1 adalah data penjualan awal yang diterima dari narasumber.

Tabel 1. Data penjualan UD TIMBUL JAYA

NO	Nama Produk	Stok	Terjual
1	Semen Gresik	2000	1900
2	Semen Merdeka	2050	1950
3	Semen Conch	1950	1850
4	Perekat Bata Ringan	1300	1140
5	Semen Putih	1200	945
6	Asbes Gelombang 3 meter	1425	1400
7	Asbes Gelombang 2,4 meter	1400	1350
8	Asbes Gelombang 2,1 meter	1335	1075

9	Asbes Gelombang 1,8 meter	1100	925
10	Plafon GRC	1175	1010
11	Plafon Pink Board	1025	950
12	Plafon 1*1	1050	900
13	Plafon 0.5*1	1150	1000
14	Besi 6	2340	2100
15	Besi 8 ulir	2200	2000
16	Besi 8 polos	2500	2300
17	Besi 10 ulir	2450	2400
18	Besi 10 polos	2325	2150
19	Besi 12 ulir	1990	1760
20	Besi 12 polos	1895	1700
21	Triplek Merah	1670	1650
22	Triplek Putih	1850	1760
23	Seng gelombang 1,8 meter	1475	1375
24	Seng gelombang 2,1 meter	1550	1460
25	Seng gelombang 2,4 meter	1600	1420

Data Mining

Menurut Firdaus, *Data mining* adalah pencarian pengetahuan dalam basis data pada proses identifikasi pola-pola yang valid, yang berpotensi bermanfaat dan dapat dipahami dengan mudah. *Data mining* merupakan proses yang bertujuan untuk menemukan pengetahuan (*knowladge discovery*) yang ditambang dari sekumpulan data yang memiliki volume yang sangat besar.

K-Nearest Neighbor

Algoritma *K-Nearest Neighbor (KNN)* adalah metode yang digunakan untuk mengklasifikasikan objek baru berdasarkan nilai (K) dari tetangga terdekatnya. KNN termasuk dalam algoritma supervised learning, di mana hasil dari *query instance* baru diklasifikasikan berdasarkan kategori mayoritas pada K tetangga terdekat. Tujuan algoritma K-NN adalah untuk mengklasifikasikan objek baru berdasarkan atribut dan data pelatihan. Dengan menerapkan algoritma *K-Nearest Neighbor*, UD TIMBUL JAYA akan lebih mudah dalam mengelola stok dan mengetahui produk yang paling diminati oleh konsumen. Terdapat beberapa metode dalam Algoritma *K-Nearest Neighbor* untuk menentukan nilai tetangga terdekat, yaitu: *Cosine Distance, Euclidean Distance, Manhattan Distance, Hamming Distance, dan Correlation Distance*. Dalam penelitian ini, penulis hanya menggunakan *Euclidean Distance*. Berikut adalah rumus untuk menghitung *Euclidean Distance*.

$$\sqrt{\sum_{i=1}^{K}(Xi - Yi)^2 + (Xi - Yi)^2 + \cdots + \cdots = \ldots\ldots\ldots\ldots\ldots}$$

Keterangan:

Nilai Xi : Suatu nilai yang terdapat pada data training

Nilai Yi : Suatu nilai yang terdapat pada data testing .

Nilai K: Dimensi suatu atribut.

Langkah-langkah menggunakan Algoritma *K Nearest Neighbor(KNN)* :

1. Menentukan parameter K (jumlah tetangga paling dekat) dianjurkan ganjil seperti 1,3,5….
2. Menghitung kuadrat jarak Eucliden objek
3. Mengurutkan hasil kuadrat jarak *Eucliden* dari yang tertinggi ke yang terendah.
4. Kumpulkan kategori Y (klasifikasi nearest neighbor berdasarkan nilai k).
5. Menggunakan kategori nearest neighbor yang paling mayoritas maka dapat diprediksi objek yang baru.

HASIL DAN PEMBAHASAN

Implementasi *K-Nearest Neighbor* untuk prediksi penjualan UD TIMBUL JAYA.Berdasarkan ketentuan dalam penentuan produk terlaris maka didapatkan hasil transformasi data pada tabel 2.

Tabel 2. Data setelah di transformasi

NO	Nama Produk	Stok	Terjual	Keterangan
1	Semen Gresik	2550	2500	Laris
2	Semen Merdeka	2450	2400	Laris
3	Semen Conch	2300	2210	Laris
4	Perekat Bata Ringan	1300	1200	Tidak laris
5	Semen Putih	1200	845	Tidak laris
6	Asbes Gelombang 3 meter	1550	1425	Tidak laris
7	Asbes Gelombang 2,4 meter	1465	1305	Tidak laris
8	Asbes Gelombang 2,1 meter	1355	1075	Tidak laris
9	Asbes Gelombang 1,8 meter	1100	1015	Laris
10	Plafon GRC	1150	1000	Tidak laris
11	Plafon Pink Board	1000	900	Tidak laris
12	Plafon 1*1	1025	995	Laris
13	Plafon 0.5*1	1050	1000	Laris
14	Besi 6	2700	2650	Laris
15	Besi 8 ulir	2650	2635	Laris

16	Besi 8 polos	2565	2535	Laris
17	Besi 10 ulir	2550	2225	Tidak laris
18	Besi 10 polos	2345	2150	Tidak laris
19	Besi 12 ulir	1990	1760	Tidak laris
20	Besi 12 polos	1895	1850	Laris
21	Triplek Merah	800	680	Tidak laris
22	Triplek Putih	850	760	Laris
23	Seng gelombang 1,8 meter	650	655	Laris
24	Seng gelombang 2,1 meter	670	635	Laris
25	Seng gelombang 2,4 meter	700	675	Laris

Langkah selanjutnya yaitu melakukan normalisasi data dengan menggunakan normalisasi min max untuk menemukan hasil yang lebih akurat,seperti pada tabel 3.

Rumus normalisasi min max:

$$x' = {x - min(x)}/{max\ (x) - min(x)}$$

Keterangan:

x’ : nilai skala baru

x : nilai data asli sebelum dilakukan normalisasi

min : nilai terendah dari data

max : nilai tertinggi dari data

Tabel 3. Perhitungan minmax data

No	MinMax Kuantitas Produk	MinMax Kuantitas Terjual	No	MinMax Kuantitas Produk
1	0,661016949	0,666666667	1	0,661016949
2	0,694915254	0,7	2	0,694915254
3	0,627118644	0,633333333	3	0,627118644
4	0,186440678	0,16	4	0,186440678
5	0,118644068	0,03	5	0,118644068

6	0,271186441	0,333333333	6	0,271186441
7	0,254237288	0,3	7	0,254237288
8	0,210169492	0,116666667	8	0,210169492
9	0,050847458	0,016666667	9	0,050847458
10	0,101694915	0,073333333	10	0,101694915
11	0	0,033333333	11	0
12	0,016949153	0	12	0,016949153
13	0,084745763	0,066666667	13	0,084745763
14	0,891525424	0,8	14	0,891525424
15	0,796610169	0,733333333	15	0,796610169
16	1	0,933333333	16	1
17	0,966101695	1	17	0,966101695
18	0,881355932	0,833333333	18	0,881355932
19	0,654237288	0,573333333	19	0,654237288
20	0,589830508	0,533333333	20	0,589830508
21	0,437288136	0,5	21	0,437288136
22	0,559322034	0,573333333	22	0,559322034
23	0,305084746	0,316666667	23	0,305084746
24	0,355932203	0,373333333	24	0,355932203
25	0,389830508	0,346666667	25	0,389830508
No	MinMax Kuantitas Produk	MinMax Kuantitas Terjual	No	MinMax Kuantitas Produk

Proses perhitungan *K Nearest Neighbor* yaitu :

1. Menentukan parameter K = 7.
2. Hitung jarak *Euclidean Distance* seperti tabel 4.

Tabel 4. Perhitungan Jarak Euclidean

NO	Nama Produk	Keterangan	Hasil
1	Seng gelombang 2,1 meter	Laris	1,048411002
2	Seng gelombang 1,8 meter	Laris	1,085570249
3	Seng gelombang 2,4 meter	Laris	1,011022593
4	Triplek Merah	Tidak laris	0,436251856
5	Triplek Putih	Laris	0,1990778
6	Plafon Pink Board	Tidak laris	0,634374475
7	Semen Putih	Tidak laris	0,600160107
8	Plafon 1*1	Laris	0,395466101
9	Plafon 0.5*1	Laris	0,121686836
10	Asbes Gelombang 1,8 meter	Laris	0,281479564
11	Plafon GRC	Tidak laris	0,169967317
12	Asbes Gelombang 2,1 meter	Tidak laris	0,068787486
13	Perekat Bata Ringan	Tidak laris	0,263446515
14	Asbes Gelombang 2,4 meter	Tidak laris	1,261099971
15	Asbes Gelombang 3 meter	Tidak laris	1,167680029
16	Besi 12 polos	Laris	1,388844444
17	Besi 12 ulir	Tidak laris	1,388851339
18	Semen Conch	Laris	1,267153175
19	Besi 10 polos	Tidak laris	0,998456468
20	Semen Merdeka	Laris	0,936370075
21	Besi 10 ulir	Tidak laris	0,828719777
22	Semen Gresik	Laris	0,939004806
23	Besi 8 polos	Laris	0,636630917

24	Besi 8 ulir	Laris	0,703972032
25	Besi 6	Laris	0,70298652

Mengurutkan hasil seperti tabel seperti pada tabel 5.

Tabel 5.Pengurutan Data

NO	Nama Produk	Keterangan	Hasil
12	Asbes Gelombang 2,1 meter	Tidak laris	0,068787486
9	Plafon 0.5*1	Laris	0,121686836
11	Plafon GRC	Tidak laris	0,169967317
5	Triplek Putih	Laris	0,1990778
13	Perekat Bata Ringan	Tidak laris	0,263446515
10	Asbes Gelombang 1,8 meter	Laris	0,281479564
8	Plafon 1*1	Laris	0,395466101
4	Triplek Merah	Tidak laris	0,436251856
7	Semen Putih	Tidak laris	0,600160107
6	Plafon Pink Board	Tidak laris	0,634374475
23	Besi 8 polos	Laris	0,636630917
25	Besi 6	Laris	0,70298652
24	Besi 8 ulir	Laris	0,703972032
21	Besi 10 ulir	Tidak laris	0,828719777
20	Semen Merdeka	Laris	0,936370075
22	Semen Gresik	Laris	0,939004806
19	Besi 10 polos	Tidak laris	0,998456468
3	Seng gelombang 2,4 meter	Laris	1,011022593
1	Seng gelombang 2,1 meter	Laris	1,048411002

2	Seng gelombang 1,8 meter	Laris	1,085570249
15	Asbes Gelombang 3 meter	Tidak laris	1,167680029
14	Asbes Gelombang 2,4 meter	Tidak laris	1,261099971
18	Semen Conch	Laris	1,267153175
16	Besi 12 polos	Laris	1,388844444
17	Besi 12 ulir	Tidak laris	1,388851339

Maka dari tabel dapat di simpulkan bahwa produk semen kuda merah dengan dapat dikategorikan "Laris",Karena sesuai dengan nilai k=7 terdapat hasil "Laris" sebanyak 4 dan "Tidak laris" sebanyak 3.

KESIMPULAN

Berdasarkan hasil analisis, penelitian ini menyimpulkan bahwa menggunakan teknik *data mining* serta algoritma *K-Nearest Neighbor* menghasilkan prediksi dengan tingkat akurasi yang cukup tinggi. Oleh karena itu, penerapan metode *K-Nearest Neighbor* dalam sebuah sistem aplikasi dapat membantu UD TIMBUL JAYA dalam mengatasi masalah terkait pengelolaan stok dan kebutuhan lainnya.

DAFTAR REFERENSI

H. A. Dwi Fasnuari, H. Yuana, and M. T. Chulkamdi, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Penyakit Diabetes Melitus," Antivirus J. Ilm. Tek. Inform., vol. 16, no. 2, pp. 133–142, 2022, doi: 10.35457/antivirus.v16i2.2445.

S. Mujilahwati and L. D. Windasari, "Implementasi Metode k-Nearest Neighbor (k-NN) untuk Memprediksi Penjualan Buah di Indonesia berbasis Website," Semin. Nas. Teknol. Sains, vol. 3, no. 1, 2024, doi: 10.29407/stains.v3i1.4077.

E. V Rahcmadani, S. F. Pane, and N. H. Harani, Algoritma C4.5 dan K-Nearest Neighbors (KNN) untuk Memetakan Matakuliah dan Keterlambatan Kelulusan Mahasiswa. Kreatif. [Online]. Available: https://books.google.co.id/books?id=BGv9DwAAQBAJ

M. Mustika et al., DATA MINING DAN APLIKASINYA. CV WIDINA MEDIA UTAMA, 2021.

Y. R. Amalia, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Elektronik Terlaris Menggunakan Metode K-Nearest Neighbor (Studi Kasus : PT.Bintang Multi Sarana Palembang)," Skripsi, pp. i–90, 2018, [Online]. Available: http://eprints.radenfatah.ac.id/id/eprint/3302%0A

S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," Build. Informatics, Technol. Sci., vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.

S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," Build. Informatics, Technol. Sci., vol. 3, no. 4, pp. 639–648, 2022.

R. Iriane, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Data Mining Untuk

Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode K-Nearest Neighbor," Media Online, vol. 3, no. 5, pp. 509–515, 2023, [Online]. Available: https://djournals.com/klik

Hasmawati, Nangi, Jumadil, and M. Muchtar, "Aplikasi Prediksi Penjualan Barang Menggunakan Metode K-Nearest Neighbor (Knn) (Studi Kasus Tumaka Mart)," semanTIK, vol. 3, no. 2, pp. 151–160, 2017, [Online]. Available: http://ojs.uho.ac.id/index.php/semantik/article/view/3658/2775

R. H. Saragih et al., "Penerapan Data Mining Untuk Memprediksi Penjualan Di Erigo Store Dengan K-Nearest Neighbor," J. Ris. Sist. Inf. Dan Tek. Inform., vol. 8, no. 2, pp. 649–661, 2023, [Online]. Available: https://tunasbangsa.ac.id/ejurnal/index.php/jurasik

R. Bahtiar, "Implementasi Data Mining Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode K-Nearest Neighbor," J. Inform. Multi, vol. 1, no. 3, pp. 203–214, 2023.

L. Suryadi, N. E. Pratiwi, F. Ardhy, P. Riswanto, and others, "Penerapan Data Mining Prediksi Penjualan Mebel Terlaris Menggunakan Metode K-Nearest Neighbor (K-Nn)(Studi Kasus: Toko Zerita Meubel)," JUSIM (Jurnal Sist. Inf. Musirawas), vol. 7, no. 2, pp. 174–184, 2022.

P. W. Rahayu et al., Buku Ajar Data Mining. PT. Sonpedia Publishing Indonesia, 2024.

F. Malik Namus Akbar, "Metode KNN (K-Nearest Neighbor) untuk Menentukan Kualitas Air," J. Tekno Kompak, vol. 18, no. 1, pp. 28–40, 2024.

U. Nijunnihayah, S. S. Hilabi, F. Nurapriani, and E. Novalia, "Implementasi Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Alat Kesehatan pada Media Alkes," MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 4, no. 2, pp. 695–701, 2024, doi: 10.57152/malcom.v4i2.1326.

I. Arimi, R. Purwaningsih, and Z. Rosyada, "Metode K-Nearest Neighbor Untuk Memprediksi Penjualan Produk Pada Umkm Pengolahan Ikan Maju Jaya," Kampus Undip Tembalang Semarang, vol. 3, no. 1, pp. 1–13, 2022.

M. Faisal, W. S. Utami, and S. Parmica, "Implementasi Algoritma K-Nearest Neighbor (KNN) Dalam Memprediksi Indeks Kemiskinan," J. Sensi, vol. 9, no. 1, pp. 11–23, 2023, doi: 10.33050/sensi.v9i1.2616.

I. Yolanda and 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, pp. 9–15, 2021.