

**PENERAPAN METODE *K-NEAREST NEIGHBOR* UNTUK MEMPREDIKSI
PENJUALAN PRODUK REMPAH-REMPAHTERLARIS
PADA TOKO HJ FATIMAH**

Ella Octa Teres¹, Rizaldi^{1*}, Santoso¹

¹Sistem Informasi, STMIK Royal Kisaran

**Email: rizaldipiliang@gmail.com*

Abstract: Hj Fatimah Store is a retail business that sells a variety of spice products. The store is not computerized yet, so employees still record reports manually. Hj Fatimah Store faces challenges in predicting the sales of its best-selling spice products, leading to stockouts, difficulties in inventory planning, and the lack of a computerized system to forecast product sales. In this context, data mining plays a crucial role in uncovering knowledge from data sets that is not easily discovered through manual analysis. The system aims to predict the best-selling spice products at Hj Fatimah Store by applying the K-Nearest Neighbor (KNN) method. This method is expected to assist in stock planning and address stockout issues. K-Nearest Neighbor is a classification technique that groups data based on the proximity or distance between data points. The purpose of applying the K-Nearest Neighbor algorithm is to classify new objects using attributes and training samples. The data to be used in this system includes sales at Hj Fatimah Store from June 2022 to May 2023. The results of the implementation and testing demonstrate that the designed system for Hj Fatimah Store aims to predict the best-selling spice products, facilitating inventory planning and helping the store manage data more effectively and produce accurate reports..

Keywords: K-Nearest Neighbor; HJ Fatimah Shop; Rempah-Rempah ; Laris

Abstrak: Toko Hj Fatimah merupakan usaha yang bergerak dibidang retail, yang menjual berbagai macam produk rempah-rempah. Toko Hj Fatimah ini belum terkomputerisasi sehingga pegawai masih saja pencatatan laporan secara manual. Toko Hj Fatimah mengalami kesulitan dalam memprediksi penjualan produk rempah-rempah terlaris sehingga mengalami kehabisan stok, sulit dalam perencanaan persediaan stok dan belum adanya sistem komputerisasi yang dapat memprediksi penjualan produk. Dalam hal ini, data mining sangat berperan penting dalam mengungkap pengetahuan dari kumpulan data yang tidak mudah diketahui melalui analisis manual. Sistem ini bertujuan untuk memprediksi penjualan produk rempah-rempah terlaris di Toko Hj Fatimah dengan menerapkan metode K-Nearest Neighbor (KNN). Metode ini diharapkan dapat membantu perencanaan stok dan mengatasi masalah kehabisan stok. K-Nearest Neighbor adalah teknik klasifikasi yang mengelompokkan data berdasarkan kedekatan atau jarak antara titik data. Tujuan penerapan algoritma K-Nearest Neighbor adalah untuk mengklasifikasikan objek baru menggunakan atribut dan sampel pelatihan. Data yang akan digunakan dalam sistem ini mencakup penjualan di Toko Hj Fatimah dari Juni 2022 hingga Mei 2023. Dari hasil implementasi dan pengujian hasil dapat dijelaskan bahwa sistem yang dirancang pada toko HJ Fatimah merupakan suatu rancangan untuk memprediksi penjualan produk rempah-rempah terlaris pada toko Hj Fatimah yang dapat memudahkan pihak toko dalam perencanaan kebutuhan stok produk rempah-rempah. Dengan rancangan ini diharapkan mampu mempermudah pihak toko HJ Fatimah bisa mengelola data dengan baik serta mendapatkan hasil laporan.

Kata Kunci: K-Nearest Neighbor; Toko HJ Fatimah; Rempah-Rempah; Laris

PENDAHULUAN

Dalam era perkembangan teknologi dan informasi yang semakin maju, kebutuhan akan informasi cepat sangat penting bagi lembaga dan bisnis, baik yang dijalankan secara pribadi maupun oleh organisasi. Misalnya, dalam transaksi jual beli, diperlukan keefektifan dalam memperoleh data dan informasi baru untuk masa depan [1]. Kebutuhan informasi bisnis dalam penjualan dapat diperoleh melalui prediksi atau peramalan berdasarkan data penjualan sebelumnya. Prediksi atau peramalan penjualan (forecasting) adalah metode perhitungan yang digunakan untuk memperkirakan kondisi masa depan dengan menguji data dari masa lalu [2].

Rempah-rempah adalah tanaman herba yang digunakan sebagai bumbu masakan untuk memberikan aroma dan meningkatkan selera makan. Bagian tanaman yang termasuk rempah-rempah meliputi batang, daun, kulit kayu, umbi, rimpang (rhizome), akar, biji, bunga, atau bagian lainnya. Produk rempah-rempah dijual dalam jumlah besar dan memiliki beragam permintaan dari konsumen [3]. Toko Hj Fatimah mengalami kesulitan dalam memprediksi penjualan produk rempah-rempah terlaris sehingga mengalami kehabisan stok, sulit dalam perencanaan persediaan stok dan belum adanya sistem komputerisasi yang dapat memprediksi penjualan produk. Dalam konteks ini, data mining memiliki peran penting dalam mengungkap informasi tersembunyi dari kumpulan data yang tidak dapat ditemukan secara manual [4].

Data mining adalah proses yang menggunakan satu atau lebih teknik pembelajaran mesin (*machine learning*) untuk menganalisis dan mengungkap pengetahuan tersembunyi dalam database [5]. Oleh karena itu, data mining sangat penting untuk memperoleh informasi mengenai produk rempah-rempah yang paling laris dan diminati oleh konsumen. Akan tetapi untuk menentukan produk rempah-rempah mana saja yang terlaris dibutuhkan sebuah metode algoritma K-Nearest Neighbor untuk mempermudah pada toko Hj Fatimah. Data mining adalah proses menggunakan pembelajaran mesin, pengenalan pola, basis data, statistik, dan teknik visualisasi untuk memecahkan masalah dalam mengekstraksi informasi dari repositori basis data yang besar [6].

Penelitian sebelumnya yang menggunakan metode *K-Nearest Neighbor*, berjudul "Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT. Nippon Indosari Corpindo Tbk," menunjukkan bahwa metode ini dapat digunakan untuk mengklasifikasikan data penjualan roti berdasarkan produk yang paling laris. Metode *K-Nearest Neighbor* tersebut membantu pihak toko dalam mengelompokkan data penjualan [7].

METODE

K-Nearest Neighbor (KNN) adalah metode supervised yang mengklasifikasikan instance baru berdasarkan mayoritas label kelas dari tetangga terdekatnya. Tujuan utama dari algoritma ini adalah mengklasifikasikan objek baru menggunakan atribut dan data pelatihan. KNN bekerja dengan mengukur jarak terpendek antara instance query dan data pelatihan untuk menentukan tetangga terdekatnya [8]. Dalam penelitian ini, algoritma *K-Nearest Neighbor* menggunakan perhitungan jarak Euclidean, karena metode ini paling umum digunakan untuk data numerik. Berikut adalah rumus perhitungan jarak Euclidean [9].

$$\sqrt{\sum_{i=1}^K (X_i - Y_i)^2 + (X_i - Y_i)^2 + \dots + \dots = \dots \dots \dots}$$

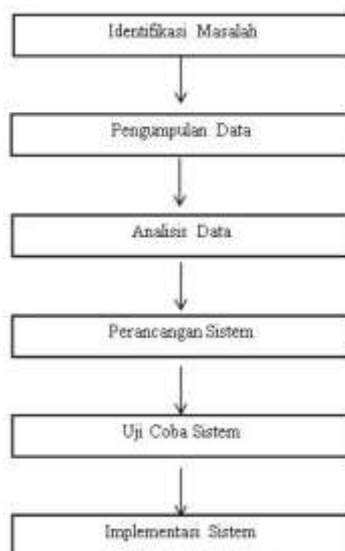
Keterangan Rumus:

Nilai X_i : nilai yang ada pada data *training* Nilai Y_i : nilai yang ada pada data *testing*

Nilai K merupakan dimensi atribut [10].

Metode Penelitian

Metode penelitian yang diterapkan dalam studi ini adalah metode penelitian kuantitatif. Metode kuantitatif memungkinkan peneliti untuk menemukan pengetahuan melalui data yang dinyatakan dalam bentuk angka. Angka-angka ini kemudian digunakan untuk menganalisis informasi. Teknik pengumpulan data kuantitatif dilakukan dengan cara survei langsung ke toko Hj Fatimah, sehingga mengetahui data yang diinginkan seperti data penjualan produk rempah-rempah.



Gambar 1. Kerangka Kerja Penelitian

Teknik Pengumpulan Data

Dalam Penelitian ini penulis mengumpulkan data dengan berbagai teknik sebagai berikut:

1. Wawancara adalah cara pengumpulan data yang dilakukan melalui percakapan antara peneliti dengan pemilik toko Ibu Alfridayati.
2. Observasi adalah teknik pengumpulan data yang dilakukan secara langsung datang ke lokasi yang diteliti, yaitu pada toko Hj Fatimah.
3. Studi Pustaka adalah pengumpulan data dengan cara membaca buku dan literatur lainnya yang dapat dijadikan acuan dalam penelitian skripsi.

HASIL DAN PEMBAHASAN

Implementasi Sistem

Implementasi adalah proses penerapan atau pelaksanaan sebuah aplikasi. Pada implementasi sistem ini, akan ditampilkan rancangan antarmuka pengguna. Rancangan antarmuka ini khususnya mencakup antarmuka admin. Antarmuka admin terdiri dari beberapa menu pilihan, yaitu menu Login, Home, Data Training, Data Testing, dan Hasil KNN.

Sebelum sistem ini diimplementasikan pada server internet, sistem ini akan dijalankan terlebih dahulu pada server komputer lokal. Pastikan Apache web server lokal dan database MySQL telah diaktifkan sebelum menjalankan sistem. Setelah keduanya aktif, buka aplikasi web browser, seperti Mozilla Firefox atau browser lainnya. Ketik alamat berikut pada address bar web browser: <http://localhost/KNN/>.

Data Training

Data training adalah data yang sudah di kelompokkan atau data yang akan dimasukkan dalam perhitungan *K-Nearest Neighbor* (KNN). Adapun data training tersebut adalah sebagai berikut:

Tabel 1. Data Testing

No	Nama	Data Periode	Data Periode	Data Periode	Data Periode	Klasifikasi
		Jun 22 - Agt 22	Sept 22 - Okt 22	Des 22 - Feb 23	Mar 23 - Mei 23	
1	Akar Usar	52	53	77	73	Laris
2	Kecibling	163	119	119	113	Tidak Laris
3	Kumis Kucing	152	125	108	109	Tidak Laris
4	Sambiroto	62	47	106	124	Laris
5	Melati Cina	73	55	54	63	Tidak Laris

Data Testing

Data testing adalah data data yang akan diprediksi dan dimasukkan dalam perhitungan *K-Nearest Neighbor*. Data testing tersebut yang akan diprediksi ialah data pada bulan juni 2023. Adapun data testing tersebut adalah sebagai berikut

Tabel 2. Data Testing

No	Nama	Jun-23		TOTAL
		Awal	Jual	
1	Akar Usar	18	12	30
2	Kecibling	23	19	42
3	Kumis Kucing	21	16	37
4	Sambiroto	28	21	49
5	Melati Cina	16	6	22

Perhitungan K-Nearest Neighbor

Dalam perhitungan *K-Nearest Neighbor* (KNN) ini dimulai dengan mencari perhitungan jarak antara data baru dan masing-masing data lainnya. Adapun cara

menghitung jarak maka menggunakan metode Euclidean Distance adalah sebagai berikut:

Data 1

$$\text{dis} = \sqrt{(52 - 30)^2 + (53 - 30)^2 + (77 - 30)^2 + (73 - 30)^2} = 71,21095421$$

Data 2

$$\text{dis} = \sqrt{(163 - 42)^2 + (119 - 42)^2 + (119 - 42)^2 + (113 - 42)^2} = 177,595045$$

Data 3

$$\text{dis} = \sqrt{(152 - 37)^2 + (125 - 37)^2 + (108 - 37)^2 + (109 - 37)^2} = 176,6182324$$

Data 4

$$\text{dis} = \sqrt{(62 - 49)^2 + (47 - 49)^2 + (106 - 49)^2 + (124 - 49)^2} = 95,115719$$

Data 5

$$\text{dis} = \sqrt{(73 - 22)^2 + (55 - 22)^2 + (54 - 22)^2 + (63 - 22)^2} = 79,96874389$$

Dari perhitungan Euclidean distance diatas, jika kita rangkum dari jarak terdekat, hasilnya sebagai berikut:

Tabel 3. Data Jarak Terdekat

No	Nama	Euclidean
1	Akar Usar	71,21095421
2	Kecibling	177,595045
3	Kumis Kucing	176,6182324
4	Sambiroto	95,115719
5	Melati Cina	79,96874389

Hasil Perhitungan K-Nearest Neighbor Pada Nilai K

Dari hasil data jarak terdekat dan nilai K=3 maka data baru tersebut dapat dilihat pada tabel dibawah ini:

Tabel 4. Hasil Data Jarak Terdekat dan Nilai K=3

No	Nama	Euclidean	Klasifikasi
1	Akar Usar	71,21095421	Tidak Laris
2	Kecibling	177,595045	Laris
3	Kumis Kucing	176,6182324	Laris
4	Sambiroto	95,115719	Laris
5	Melati Cina	79,96874389	Tidak Laris

Halaman *Login Admin*

Halaman login admin digunakan untuk mengautentikasi hak admin dalam mengakses sistem. Admin dengan kredensial yang valid, seperti nama pengguna: admin dan kata sandi: admin, dapat masuk. Berikut adalah tampilan halaman login..

Gambar 2. Tampilan *Login Admin*

Halaman *Data Training*

Halaman data *training* merupakan data untuk input data *training* serta mulai pelatihan dengan menginput nilai K. Berikut tampilan halaman data *training*.

ID	Nama	Nilai Jun 2020-Apr 2022	Nilai Sept 2020-Okt 2022	Nilai Nov 2020-Jan 2022	Nilai Feb 2020-Mai 2022	Klasifikasi	Aksi Data
1	Adi Lili	82	83	77	73	Laba	Q 100 W 100
2	Kucing	905	119	119	113	Tidak Laba	Q 100 W 100
3	Kucing Kucing	550	125	100	100	Tidak Laba	Q 100 W 100
4	Santoso	82	47	536	124	Laba	Q 100 W 100
5	Mahar Cini	73	85	54	83	Tidak Laba	Q 100 W 100

Gambar 3. Tampilan Halaman *Data Training*

Halaman *Data Testing*

Halaman data *testing* merupakan halaman data yang akan di prediksi. Berikut tampilan halaman data *testing*.

ID	Nama	Nilai Jan 2023-Ags 2023	Nilai Sept 2023-Oct 2023	Nilai Nov 2023-Jan 2024	Nilai Feb 2023-Mar 2024	Beli Data
1	Kayu Cendur	30	30	30	30	<button>30.000</button> <button>30.000</button>
2	Kayu Cendur	42	42	42	42	<button>42.000</button> <button>42.000</button>
3	Kayu Cendur	37	37	37	37	<button>37.000</button> <button>37.000</button>
4	Kayu Cendur	40	40	40	40	<button>40.000</button> <button>40.000</button>
5	Kayu Cendur	22	22	22	22	<button>22.000</button> <button>22.000</button>
6	Kayu Cendur	35	35	35	35	<button>35.000</button> <button>35.000</button>
7	Kayu Cendur	18	18	18	18	<button>18.000</button> <button>18.000</button>

Gambar 4. Tampilan Halaman Data *Testing*

Halaman klasifikasi KNN ini merupakan halaman hasil dari perhitungan data training dan data testing. Berikut tampilan halaman klasifikasi KNN:

ID	Nama	Nilai Jan 2023-Ags 2023	Nilai Sept 2023-Oct 2023	Nilai Nov 2023-Jan 2024	Nilai Feb 2023-Mar 2024	Klasifikasi
1	Kayu Cendur	30	30	30	30	Tidak Laris
2	Kayu Cendur	42	42	42	42	Laris
3	Kayu Cendur	37	37	37	37	Laris
4	Kayu Cendur	40	40	40	40	Laris
5	Kayu Cendur	22	22	22	22	Tidak Laris
6	Kayu Cendur	35	35	35	35	Laris
7	Kayu Cendur	18	18	18	18	Tidak Laris
8	Kayu Cendur	30	30	30	30	Laris

Gambar 5. Tampilan Klasifikasi *KNN*

Berdasarkan hasil implementasi dan pengujian, dapat dijelaskan bahwa sistem yang dirancang di <http://localhost/KNN/> mempermudah admin dalam mengelola data training dan testing. Sistem ini memungkinkan input data training dan testing serta menampilkan hasil prediksi menggunakan metode KNN.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan di Toko HJ Fatimah dengan judul “Penerapan Metode K-Nearest Neighbor Untuk Memprediksi Penjualan Rempah-Rempah Terlaris Pada Toko HJ Fatimah”. Terciptanya sebuah aplikasi berbasis website dengan perhitungan K-Nearest Neighbor yang dapat membantu memprediksi penjualan produk rempah-rempah dan mengurangi kemungkinan kesalahan dalam menetapkan jumlah stok yang harus disediakan oleh pemilik Toko HJ Fatimah agar permintaan konsumen dapat terpenuhi dengan baik. Dari hasil menerapkan metode K-Nearest

Neighbor untuk prediksi rempah- rempah pada Toko HJ Fatimah yang telah dilakukan, mendapatkan jumlah hasil prediksi klasifikasi KNN diantaranya produk yang laris sebanyak 13 dan produk tidak laris sebanyak 32. Hasil prediksi ini dapat membantu pemilik toko dalam pengambilan keputusan terkait pengelolaan stok. Dengan mengetahui produk mana yang laris dan mana yang tidak, pemilik toko dapat menyesuaikan jumlah stok yang harus disediakan, sehingga dapat mengurangi pemborosan dan memastikan ketersediaan produk yang diminati oleh konsumen.

DAFTAR PUSTAKA

- [1] D. W. T. Putra and R. Andriani, “Unified Modelling Language (UML) dalam Perancangan Sistem Informasi Permohonan Pembayaran Restitusi SPPD,” *J. TeknoIf*, vol. 7, no. 1, p. 32, 2019, doi: 10.21063/jtif.2019.v7.1.32-39.
- [2] S. Widaningsih, “Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm,” *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [3] 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.
- [4] M. H. Iku, I. Surya Kumala, and Y. A. Mustofa, “Metode K-Nearest Neighbor untuk Memprediksi Harga Eceran Beras di Pasar Tradisional Gorontalo,” *J. Nas. cosPhi*, vol. 3, no. 2, pp. 2597–9329, 2019.
- [5] Y. Anggraini, D. Pasha, D. Damayanti, and A. Setiawan, “Sistem Informasi Penjualan Sepeda Berbasis Web Menggunakan Framework Codeigniter,” *J. Teknol. dan Sist. Inf.*, vol. 1, no. 2, pp. 64–70, 2020, doi: 10.33365/jtsi.v1i2.236.
- [6] 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.
- [7] Kaharuddin, Kusri, and E. T. Luthfi, “Klasifikasi Jenis Rempah-Rempah Berdasarkan Fitur Warna Rgb Dan Tekstur Menggunakan Algoritma K-Nearest Neighbor,” *J. Inf. Interaktif*, vol. 4, no. 1, pp. 17–22, 2019.
- [8] D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [9] 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.
- [10] A. Alfani W.P.R., F. Rozi, and F. Sukmana, “Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jupi.v6i1.1910.