

PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN PRODUK ELEKTRONIK TERLARIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR

Sriani¹, Aidil Halim Lubis², Ridho Rizky Nasution³

Universitas Islam Negeri Sumatera Utara, Medan

e-mail: ¹sriani@uinsu.ac.id, ²aidilhalimlubis@uinsu.ac.id, ³ridhorisky07@gmail.com

Abstract: Sales of electronic products are an integral part of a market that continues to grow rapidly. In this context, the application of Data Mining techniques becomes important to uncover hidden patterns and insights in sales data, with the ultimate goal of predicting best-selling electronic products. This research focuses on the application of the K-Nearest Neighbor (KNN) method as a tool to achieve accurate predictions. The KNN method is an algorithm based on the concept that entities tend to have similar performance if they are close together in feature space. In the context of electronic product sales, KNN is used to identify patterns from large amounts of sales data, which includes variables such as price and quantity sold. By analyzing these patterns, the KNN algorithm can predict which electronic products are likely to become best-selling products in the future. This research involves important steps such as data pre-processing, selection of K parameters in KNN, model validation, and accuracy measurements. The results of this research are accuracy in making predictions with a value of 92%. The total number of products sold is 491 products. Overall, this research shows that the application of Data Mining using the KNN method has the potential to increase understanding of electronic product sales trends and provide valuable information for predicting best-selling products in the future. Prediction accuracy can be improved by considering additional variables and combining other analysis methods.

Keywords: Prediction, Goods, KNN, Best Sellers

Abstrak: Penjualan produk elektronik merupakan bagian integral dari pasar yang terus berkembang dengan cepat. Dalam konteks ini, penerapan teknik Data Mining menjadi penting untuk mengungkap pola dan wawasan yang tersembunyi dalam data penjualan, dengan tujuan utama memprediksi produk elektronik terlaris. Penelitian ini fokus pada penerapan metode K-Nearest Neighbor (KNN) sebagai alat untuk mencapai prediksi yang akurat. Metode KNN adalah algoritma yang berdasarkan pada konsep bahwa entitas cenderung memiliki kinerja serupa jika mereka berdekatan dalam ruang fitur. Dalam konteks penjualan produk elektronik, KNN digunakan untuk mengidentifikasi pola dari sejumlah besar data penjualan, yang mencakup variabel-variabel seperti harga dan jumlah terjual. Dengan menganalisis pola ini, algoritma KNN dapat memprediksi produk elektronik yang kemungkinan besar menjadi produk terlaris di masa depan. Penelitian ini melibatkan langkah-langkah penting seperti pra-pemrosesan data, pemilihan parameter K dalam KNN, validasi model, dan pengukuran akurasi. Hasil penelitian ini berupa akurasi dalam melakukan prediksi dengan nilai 92%. Dari total keseluruhan jumlah produk yang terjual adalah 491 produk. Secara keseluruhan penelitian ini menunjukkan bahwa penerapan Data Mining dengan metode KNN memiliki potensi untuk meningkatkan pemahaman tentang tren penjualan produk elektronik dan memberikan informasi berharga untuk prediksi produk terlaris di masa mendatang. Keakuratan prediksi dapat ditingkatkan dengan mempertimbangkan variabel tambahan dan menggabungkan metode analisis lainnya.

Kata kunci: Promosi, Brosur, Multimedia Development Life Cycle, Augmented Reality

PENDAHULUAN

Penjualan merupakan faktor terpenting bagi sebuah perusahaan. Karena dengan adanya penjualan, maka suatu perusahaan akan mendapat keuntungan yang lebih supaya bisa melanjutkan usaha tersebut. Namun persaingan bisnis di era perdagangan bebas seperti sekarang ini sangat ketat, setiap perusahaan dituntut untuk mempersiapkan dirinya secara profesional dan fleksibel sehingga perusahaan tidak hanya mampu bertahan, tetapi juga mampu tumbuh dan berkembang (Jannah et al., 2023). Untuk itu perusahaan harus memiliki strategi yang tepat agar bisa memenuhi kebutuhan pasar. Maka dari itu persaingan terletak pada bagaimana sebuah perusahaan dapat menghadirkan produk yang lebih murah, lebih baik namun dengan stok barang yang selalu terpenuhi (Hardiyanto, 2020).

Okey Elektronik merupakan toko yang bergerak di bidang penjualan barang elektronik. Okey Elektronik menjual berbagai produk elektronik seperti televisi, kipas angin, mesin cuci dan produk-produk elektronik lainnya. Dilihat dari peminatnya konsumen terhadap barang elektronik ini, maka dibutuhkan proses prediksi barang elektronik yang banyak dibeli konsumen agar tidak terjadi kekurangan stok barang atau kelebihan stok barang yang mengakibatkan kerugian pada pihak toko tersebut. Masalah yang sering dialami pihak toko adalah bagaimana cara pihak toko dapat menyediakan stok barang yang jumlahnya dapat di perkirakan dengan cara melihat data-data penjualan sebelumnya.

Sebagai contoh masalah yang dihadapi pihak toko adalah penjualan suatu produk yang meningkat penjualannya setiap tahun seperti mesin cuci yang pada 3 tahun terakhir penjualannya terus meningkat dari tahun 2020 yang hanya terjual 20 unit, 2021 terjual 18 unit dan 2022 terjual hanya 28 unit, disini pihak toko tidak dapat mengambil keputusan yang tepat dalam

menyediakan stok barang mesin cuci dikarenakan pihak toko masih meragukan berapa jumlah yang tepat untuk persediaan stok mesin cuci untuk tahun depan. Masalah lain yang sering dialami pihak toko adalah kelebihan stok barang yang mengakibatkan kerugian pihak toko sebagai contoh barang elektronik yang berukuran kecil seperti setrika dimana penjualan produk ini pada 3 tahun terakhir tidak menentu untuk penjualan produknya tersebut, pada tahun 2020 penjualan setrika ini mencapai 25 unit, tahun 2021 mencapai 29 unit dan tahun 2022 mencapai 16 unit disini pihak toko tidak memprediksi berapa jumlah barang yang harus di stok untuk tahun berikutnya sehingga pihak toko berpikir hanya karena barangnya kecil dan harganya terjangkau pihak toko menyetok barang tersebut sehingga menumpuk yang mengakibatkan kerugian pada pihak toko karena barang yang di stok terlalu berlebih sehingga membuat produk tersebut ada yang rusak dan penjualan produk tersebut terjual dengan waktu yang sangat lama.

Prediksi ini bertujuan agar memudahkan pihak toko dalam menyetok sebuah barang dan nanti hasilnya dari prediksi ini akan di sampaikan pihak toko barang mana saja yang paling banyak diminati oleh konsumen, karena sering terjadinya kasus kehabisan stok barang yang di cari konsumen, dan membuat konsumen terkadang tidak puas akan layanan yang ada sebagaimana toko penjualan yang kurang adanya rekapan barang yang harus di sediakan di toko (Aripin, 2021).

Prediksi ini dimaksudkan untuk memberikan indikasi jumlah dana yang dibutuhkan untuk bulan berikutnya sehubungan dengan data dari masa lalu atau periode waktu lain yang telah ditinjau sehingga dapat digunakan sebagai acuan untuk memperkirakan jumlah dana yang dibutuhkan untuk bulan berikutnya (R. D. Putri, 2022). Karena kekuatan esensialnya untuk mengkorelasikan kelas dengan properti yang sudah ada, prediksi juga memberikan peluang untuk

kualifikasi tambahan yang melampaui perkiraan deret waktu saja. Jenis prediksi yang bisa berupa kuantitatif (juga dikenal sebagai berbentuk angka) atau kualitatif (tidak berbentuk angka).

Selama ini belum ada sistem yang mengatur prediksi atau peramalan untuk penjualan produk elektronik, sehingga sering terjadi penumpukan barang, barang rusak atau bahkan kekurangan stok barang yang mengakibatkan kerugian bagi perusahaan (A. A. Putri, 2021).

METODE

Teknik pengumpulan data yaitu dengan melakukan observasi, wawancara dan studi pustaka:

1. Observasi
Yaitu dengan pengumpulan data dengan pengamatan secara langsung pada dalam proses prediksi penjualan produk elektronik terlaris di kota pinang menggunakan metode K-Nearest Neighbor dengan segala aspek yang berhubungan langsung dengan penelitian (Suhardi et al., 2023).
2. Wawancara
Yaitu dengan melakukan kegiatan wawancara untuk mencari informasi dengan menggunakan tanya jawab dengan pihak yang berwenang mengenai penjualan produk elektronik terlaris (Irwanto, 2021). Data yang digunakan untuk dilakukan metode K-Nearest Neighbor.
3. Studi Pustaka
Melakukan studi pustaka mengenai teori-teori dan konsep yang berhubungan dengan penelitian, seperti teori mengenai keamanan data dengan menggunakan metode K-Nearest Neighbor. Referensi yang digunakan penelitian adalah buku, jurnal ilmiah online dan situs (Fadli, 2021).

K-Nearest Neighbor (KNN)

K-Nearest Neighbor (K-NN) adalah algoritma yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) suatu data dengan data yang lain (Wijaya & Ridwan, 2019) (Dunggio & Bode, 2019). Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel latih. Algoritma K-Nearest Neighbor merupakan algoritma pada data mining yang termasuk dalam klasifikasi data mining (Dinanti & Purwadi, 2023). Proses pada algoritma K-Nearest Neighbor dimulai dengan penentuan nilai K terlebih dahulu. Setelah penentuan nilai K maka selanjutnya proses yang dilakukan menghitung kedekatan nilai antar objek (Furqan et al., 2022). Pada K-Nearest Neighbor objek yang terpilih untuk memenuhi nilai K dihitung berdasarkan dengan euclidean distance. Euclidean distance merupakan proses perhitungan jarak terdekat antara objek baru dengan objek lama yang tersimpan pada data (Mahalisa & Arminarahmah, 2022). Adapun rumus yang digunakan untuk menyelesaikan algoritma K-Nearest Neighbor yaitu:

Contoh perhitungan K-NN (Sakariana et al., 2020):

1. Menentukan parameter K sebagai banyaknya jumlah tetangga terdekat dengan objek baru.
2. Menghitung jarak antar objek/data baru terhadap semua objek/data yang telah di training.
3. Urutkan hasil perhitungan tersebut.
4. Tentukan tetangga terdekat berdasarkan jarak minimum ke K.
5. Tentukan kategori dari tetangga terdekat dengan objek /data.
6. Gunakan kategori mayoritas sebagai klasifikasi objek /data baru.

HASIL DAN PEMBAHASAN

Data Selection

Data yang digunakan dalam penelitian ini yaitu data penjualan produk elektronik berdasarkan penjualan 3 tahun

terakhir dari tahun 2020, 2021 dan 2022 yang berasal Toko Okey Elektronik. Adapun atribut yang digunakan dalam penentuan prediksi penjualan elektronik terlaris adalah atribut Nama Barang, Harga dan Terjual.

Preprocessing

Tahap selanjutnya yaitu tahap *preprocessing* yang merupakan tahapan yang harus dilakukan sebelum melakukan data mining (Nanda, 2021). Tahapan yang akan dilakukan agar menghasilkan data yang bersih ketika melakukan tahapan mining, sebelum kita ke tahap mining terlebih dahulu melakukan proses *cleaning*.

Tabel 1. Preprocssing Data

Nama Barang	Harga	Terjual
Ketel 7 Liter	140000	25
TV Polytron	1000000	2
Mesin Cuci Polytron 700	1350000	2
Microphone Tda	230000	3
Blender Maspion 3 Tabung	475000	5
.....		
Dispenser Sijempol	130000	2

Transformation

Dari data transformasi dengan merubah menjadi 3 kategori yaitu laris, cukup laris dan sangat laris.

Tabel 2. Kategori

Jumlah Terjual	Kategori
2-3	Cukup Laris
4-7	Laris
8++	Terlaris

Tabel 3. Tabel Penjualan Setelah Transformasi

Nama Barang	Harga	Terjual	Label
Ketel 7 Liter	140000	25	Terlaris
TV Polytron	1000000	2	Cukup Laris
Mesin Cuci Polytron	1350000	2	Cukup Laris

700			
Microphone Tda	230000	3	Cukup Laris
Blender Maspion 3 Tabung	475000	5	Laris
.....			
Dispenser Sijempol	130000	2	Cukup Laris

Data yang sudah siap selanjutnya akan diklasifikasikan menjadi dua data yaitu data *training* dan data *testing*, pembagian data *training* dan data *testing* dibagi menjadi 80% untuk data *training* dan 20% data *testing*.

Data Mining

Tahapan ini merupakan tahapan untuk mencari kecocokan dari pola dan informasi menarik dalam data yang terpilih dengan menggunakan metode tertentu berdasarkan proses KDD ini secara keseluruhan. Karena jumlah dataset yang cukup banyak dan tentunya akan mempengaruhi proses perhitungan manual menjadi panjang, maka penulis hanya memakai 10 data training dan 3 data testing yang akan dijadikan sebagai proses perhitungan manual. Berikut adalah tabel datanya:

Tabel 4. Data Training

No	Nama Barang	Harga	Terjual	Label
1	Ketel 7 Liter	140000	25	Terlaris
2	Tv Pollytron	1000000	2	Cukup Laris
3	Mesin Cuci Polytron 700	1350000	2	Cukup Laris
4	Microphone Tda	230000	3	Cukup Laris
5	Blender Maspion 3	475000	5	Laris

	Tabung			
6	Ceret Lirik	75000	3	Cu kup Lar is
7	Setrika Maspion	85000	2	Cu kup Lar is
8	Stand Fan Homey	75000	2	Cu kup Lar is
9	Parabola Optus	180000	4	Lar is
1 0	Kipas Maspion 167	430000	3	Cu kup Lar is

Penerapan Algoritma KNN

Dalam konteks penelitian ini, algoritma K-NN digunakan untuk menentukan produk elektronik laris, cukup laris dan terlaris. Algoritma ini bekerja dengan mencari data terdekat dari suatu titik yang akan diklasifikasikan dan menentukan label atau kelas data tersebut berdasarkan mayoritas label data terdekat. Pada penelitian ini terdapat 2 kriteria seperti nilai harga dan jumlah terjual. Berikut ini adalah keterangan 2 kriteria yang sudah di berikan skor sesuai kebutuhan:

Tabel 5. Kriteria Harga

No	Kriteria Harga	Skor
1	0 – 100.000	1
2	=>100.000 – 200.000	2
3	=>300.000 – 500.000	3
4	=>600.000	4

Tabel 6. Jumlah Terjual

No	Kriteria Jumlah Terjual	Skor
1	=>2-3	2
2	=>4– 7	3
3	=>8	4

Berikut ini dataset dari penelitian ini yang akan dibagi menjadi data pelatihan dan data pengujian.

Tabel 7. Data Pelatihan

No	Nama Barang	Harga	Terjual	Label
1	Ketel 7 Liter	2	4	Terlaris
2	Tv Pollytron	4	2	Cukup Laris
3	Mesin Cuci Polytron 700	4	2	Cukup Laris
4	Microphone Tda	3	2	Cukup Laris
5	Blender Maspion 3 Tabung	3	3	laris
6	Ceret Lirik	1	2	Cukup Laris
7	Setrika Maspion	1	2	Cukup Laris
8	Stand Fan Homey	1	2	Cukup Laris
9	Parabola Optus	2	3	Laris
1 0	Kipas Maspion 167	3	2	Cukup Laris

Tabel 8. Data Pengujian

No	Nama Barang	Harga	Terjual	Label
1	Wall fan aoyama	2	3	?
2	Megicom Cosmos	2	2	?
3	Setrika Philips	3	4	?

Hitung jarak *euclidean* antara data baru dan setiap data pada *dataset*.

Pada penelitian ini akan menentukan produk laris, cukup laris atau sangat laris dengan data *testing* sebagai berikut:

a) Harga = 2

b) Terjual = 3

Berikut ini perhitungan dari jarak *euclidean* antara data baru pada setiap *dataset* dengan rumus:

$$dis = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \dots}$$

$$d = ((x_1 - x_2)^2 + (y_1 - y_2)^2)$$

perhitungan terhadap produk 1

$$d = ((2-2)^2 + (3-4)^2)$$

$$d = 1$$

perhitungan terhadap produk 2

$$d = ((2-4)^2 + (3-2)^2)$$

$$d = 5$$

perhitungan terhadap produk 3

$$d = ((2-4)^2 + (3-2)^2)$$

$$d = 5$$

Urutkan data berdasarkan jarak terdekat dari yang terkecil hingga terbesar:

Tabel 9. Hasil Perhitungan Jarak

Euclidean Distance

No	Nama Barang	Nilai	Label
1	Ketel 7 liter	1	Terlaris
2	Tv Pollytron	5	Cukup Laris
3	Mesin Cuci Polytron 700	5	Cukup Laris
4	Microphone Tda	2	Cukup Laris
5	Blender Maspion 3 Tabung	1	Laris
6	Ceret Lirik	2	Cukup Laris
7	Setrika Maspion	2	Cukup Laris
8	Stand Fan Homey	2	Cukup Laris
9	Parabola Optus	0	Laris
10	Kipas Maspion 167	2	Cukup Laris

Ambil k (dalam kasus ini 3) data terdekat dan hitung frekuensi dalam produk laris dan cukup laris yang terdapat pada tabel berikut ini:

Tabel 10. Hasil Dari Data Terdekat

Nama Barang	Nilai
Ketel 7 Liter	1
Blender Maspion 3 Tabung	1
Parabola Optus	0

Tabel 11. Hasil Data Testing

Nama Barang	Hasil Testing
Wall Fan Aoyama	Laris

Untuk produk yang berada pada urutan 32 tidak dapat ditentukan label dikarenakan perlu dilakukan perhitungan ulang terhadap dari keseluruhan data terhadap data *testing*.

a) Harga = 2

b) Terjual = 2

perhitungan terhadap produk 1

$$d = ((2-2)^2 + (2-4)^2)$$

$$d = 4$$

perhitungan terhadap produk 2

$$d = ((2-4)^2 + (2-2)^2)$$

$$d = 4$$

perhitungan terhadap produk 3

$$d = ((2-4)^2 + (2-2)^2)$$

$$d = 4$$

Urutkan data berdasarkan jarak terdekat dari yang terkecil hingga terbesar:

Tabel 12. Hasil Perhitungan Jarak

Euclidean Distance

No	Nama Barang	Nilai	Label
1	Ketel 7 Liter	4	Terlaris
2	Tv Pollytron	4	Cukup Laris
3	Mesin Cuci Polytron 700	4	Cukup Laris
4	Microphone Tda	1	Cukup Laris
5	Blender Maspion 3 Tabung	2	Laris

6	Ceret Listrik	1	Cukup Laris
7	Setrika Maspion	1	Cukup Laris
8	Stand Fan Homey	1	Cukup Laris
9	Parabola Optus	1	Laris
10	Kipas Maspion 167	1	Cukup Laris

Ambil k (dalam kasus ini 3) data terdekat dan hitung frekuensi dalam produk laris dan cukup laris yang terdapat pada tabel berikut ini.

Tabel 13. Hasil Dari Data Terdekat

Nama Barang	Nilai
Ceret Listrik	1
Setrika Maspion	1
Stand Fan Homey	1

Dari hasil data dengan nilai K dihasilkan bahwa dari label yang terdapat pada produk megicom cosmos masuk kedalam label laris dikarenakan hasil dari nilai k terdapat produk ceret listrik, setrika maspion dan stand fan homey dengan label cukup laris. Berikut ini hasil dari data *testing* yang terdapat pada tabel berikut ini:

Tabel 14. Hasil Data *Testing*

Nama Barang	Hasil <i>Testing</i>
Megicom Cosmos	Cukup laris

Untuk produk yang berada pada urutan ke 3 tidak dapat ditentukan label dikarenakan perlu dilakukan perhitungan ulang terhadap dari keseluruhan data terhadap data *testing*.

a) Harga = 3

b) Terjual = 4

perhitungan terhadap produk 1

$$d = ((3-2)^2 + (4-4)^2)$$

$$d = 1$$

perhitungan terhadap produk 2

$$d = ((3-4)^2 + (4-2)^2)$$

$$d = 5$$

perhitungan terhadap produk 3

$$d = ((3-4)^2 + (4-2)^2)$$

$$d = 5$$

Urutkan data berdasarkan jarak terdekat dari yang terkecil hingga terbesar:

Tabel 15. Hasil Perhitungan Jarak Euclidean Distance

No	Nama Barang	Nilai	Label
1	Ketel 7 Liter	1	Terlaris
2	Tv Pollytron	5	Cukup Laris
3	Mesin Cuci Polytron 700	5	Cukup Laris
4	Microphone Tda	4	Cukup Laris
5	Blender Maspion 3 Tabung	1	Laris
6	Ceret Listrik	8	Cukup Laris
7	Setrika Maspion	8	Cukup Laris
8	Stand Fan Homey	8	Cukup Laris
9	Parabola Optus	2	Laris
10	Kipas Maspion 167	4	Cukup Laris

Ambil k (dalam kasus ini 3) data terdekat dan hitung frekuensi dalam produk laris dan cukup laris yang terdapat pada tabel berikut ini:

Tabel 16. Hasil Dari Data Terdekat

Nama Barang	Nilai
Ketel 7 Liter	1
Blender Maspion 3 Tabung	1
Parabola Optus	2

Dari hasil data dengan nilai K dihasilkan bahwa dari label yang terdapat pada produk setrika philips masuk kedalam label laris dikarenakan hasil dari nilai k terdapat produk ketel 7 liter dengan label terlaris blender dan parabola termasuk dengan label laris. Berikut ini hasil dari data *testing* yang terdapat pada tabel berikut ini:

Tabel 17. Hasil Data *Testing*

Nama Barang	Hasil <i>Testing</i>
Setrika Philips	Laris

Berikut ini hasil dari data *testing* yang terdapat pada tabel berikut ini:

Tabel 18. Hasil Data *Testing*

N o	Nama produk	Harg a	Terjua l	Prediks i
1	Wall fan aoyama	2	3	Laris
2	Megico m	2	2	Cukup Laris
3	Serika philips	3	4	Laris

Akurasi

Dalam proses perhitungan data dengan menggunakan algoritma K-nn didapatkan hasil prediksi laris, cukup laris dan sangat laris pada produk elektronik, pada penelitian ini menggunakan 3 data uji, berikut ini hasil data pengujian tanpa menggunakan label.

Tabel 19. Hasil Prediksi

No	Nama Produk	Label	Hasil prediksi
1	Wall fan aoyama	Laris	Laris
2	Megicom cosmos	Cukup laris	Cukup Laris
3	Setrika philips	Terlaris	Laris

Dari tabel 20 kemudian dihitung akurasi dari jumlah data pengujian dengan rumus:

$$akurasi = \frac{\text{jumlah data terprediksi benar}}{\text{jumlah data uji}} \times 100\%$$

$$akurasi = \frac{2}{3} \times 100\%$$

Akurasi = 66,6%

Kemudian disimpulkan bahwa akurasi pengujian 66,6% dikarenakan dari jumlah 3 data testing yang berhasil diprediksi dengan benar berjumlah 2 data.

Implementasi

Tampilan Halaman Login Administrator

Halaman menu utama dari sistem prediksi produk terlaris dengan algoritma KNN.

Gambar 1. Halaman Login Administrator

Tampilan Menu Data Barang

Halaman menu data barang berfungsi melakukan *input* data barang sesuai kebutuhan pada sistem prediksi produk terlaris.

Gambar 2. Halaman Menu Data Barang

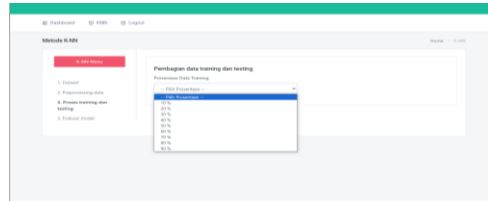
Tampilan Proses *Processing*

Processing data merupakan proses menganalisis, dan menyajikan data agar menjadi dataset yang dapat digunakan. Proses ini akan menentukan atribut yang digunakan dan label target yang digunakan.

Gambar 3. Halaman Proses Preprocessing

Tampilan Pembagian Data Training Dan Testing

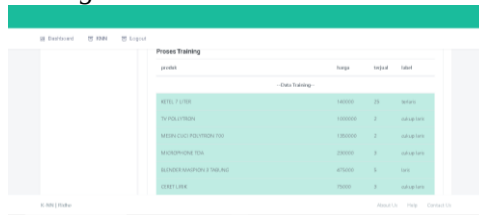
Menu pembagian data *training* dan *testing* bertujuan untuk melatih model dalam melakukan prediksi dan data *testing* digunakan untuk mengukur kinerja model.



Gambar 4. Halaman Pembagian Data Training dan Testing

Tampilan Proses Training

Tampilan proses *training* akan menampilkan *training* data dari keseluruhan data berdasarkan jumlah data *training* yaitu dengan presentase 80% dari jumlah data. Langkah ini melibatkan memberikan data kepada model agar model dapat mempelajari pola atau fitur-fitur dalam data tersebut. Model mencoba untuk menyesuaikan dengan data *training*.



Gambar 5. Tampilan Proses Training

Tampilan Proses Testing

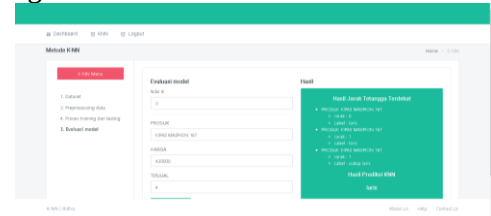
Proses data *testing* adalah subset dari data yang tidak digunakan selama proses pelatihan model, tetapi disimpan sebagai data yang belum pernah dilihat oleh model. Data ini digunakan untuk menguji sejauh mana model yang dilatih mampu menggeneralisasi pada data baru. Tujuan dari pengujian ini adalah untuk mengukur kinerja model pada data yang belum pernah dilihat sebelumnya dan mengevaluasi sejauh mana model dapat melakukan prediksi yang akurat.



Gambar 6. Tampilan Proses Testing

Hasil Evaluasi Dengan Algoritma KNN

Pada hasil evaluasi akan menampilkan hasil evaluasi dalam melakukan prediksi dengan menggunakan algoritma KNN.



Gambar 7. Tampilan Hasil Prediksi

SIMPULAN

Berdasarkan rumusan masalah penelitian dan hasil dari penerapan algoritma KNN dalam melakukan prediksi produk elektronik terlaris, maka diperoleh kesimpulan sebagai berikut:

1. Sistem prediksi produk elektronik terlaris dalam melakukan prediksi dengan menggunakan algoritma KNN dapat membantu perusahaan untuk meningkatkan penjualan dan menambah stok.
2. Akurasi yang dihasilkan pada sistem ini dalam melakukan prediksi produk elektronik terlaris menghasilkan akurasi 85,8% dengan jumlah pembagian data 80% untuk training dan 20% untuk testing.
3. Sistem ini dapat memberikan pertimbangan dan gambaran penyediaan stok berdasarkan jumlah terjual.

DAFTAR PUSTAKA

- Aripin, Z. (2021). E-Business strategi, model, dan penerapannya. Deepublish.
- Dinanti, A., & Purwadi, J. (2023). Analisis Performa Algoritma K-Nearest Neighbor dan Reduksi Dimensi Menggunakan Principal Component Analysis. *Jambura Journal of Mathematics*, 5(1), 155–165.
- Dunggio, F. Y. R., & Bode, A. (2019).

- Penerapan Metode Local Binary Pattern Untuk Pengenalan Jenis Daun Tanaman Obat Tradisional Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Cosphi*, 3(1).
- Fadli, M. R. (2021). Memahami desain metode penelitian kualitatif. *Humanika*, 21(1), 33–54. <https://doi.org/10.21831/hum.v21i1.38075>
- Furqan, M., Sriani, S., & Sari, S. M. (2022). Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia. *Techno.Com*, 21(1), 52–61. <https://doi.org/10.33633/tc.v21i1.5446>
- Hardiyanto, F. (2020). Analisis marketing Syariah dalam menghadapi covid 19 (studi kasus ARPI hijab Kuningan). ., 1(1), 23–32.
- Irwanto, I. (2021). Perancangan Sistem Informasi Sekolah Kejuruan dengan Menggunakan Metode Waterfall (Studi Kasus SMK PGRI 1 Kota Serang-Banten). *Lectura : Jurnal Pendidikan*, 12(1), 86–107. <https://doi.org/10.31849/lectura.v12i1.6093>
- Jannah, M., Sartika, D., & Purnamasari, E. (2023). Prediksi Penjualan Produk Pada PT Bintang Sriwijaya Palembang Menggunakan K-Nearest Neighbour: Prediksi Calon Mahasiswa Penerima KIP Pada Universitas Indo Global Mandiri menggunakan Algoritma Decision Tree. *Jurnal Software Engineering and Computational Intelligence*, 1(2), 80–89.
- Mahalisa, G., & Arminarahmah, N. (2022). Diabetes Classification Analysis Using the Euclidean Distance Method Based on the K-Nearest Neighbors Algorithm. *J. Teknol. Komput. Dan Sist. Inf*, 5(3), 178–182.
- Nanda, I. D. (2021). Penerapan Local Binary Pattern Dan K-Nearest Neighbor Mendeteksi Penyakit Pada Daun Mangga. Universitas Islam Negeri Sumatera Utara.
- Putri, A. A. (2021). Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus: PT. Central Brastagi Utama). *RESOLUSI: Rekayasa Teknik Informatika Dan Informasi*, 1(6), 354–361.
- Putri, R. D. (2022). Prediksi Penjualan Produk Elektronik Yang Terlaris Pada CV. Istana Komputer Palembang Menggunakan Algoritma Regresi Linear Sederhana. *Jurnal Mantik*, 6(2), 2254–2263.
- Sakariana, M. I. D., Indriati, & Dewi, C. (2020). Analisis Sentimen Pemindahan Ibu Kota Indonesia Dengan Pembobotan Term BM25 Dan Klasifikasi Neighbor Weighted K-Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(3), 748–755.
- Suhardi, Lubis, A. H., Aprilia, A., & Ningrum, I. A. (2023). Penerapan Metode Simple Multi Attribute Rating Technique pada Pemilihan Cafe Terfavorit. *Sistem Pendukung Keputusan Dengan Aplikasi*, 2(1), 1–11. <https://doi.org/10.55537/spk.v2i1.114>
- Wijaya, N., & Ridwan, A. (2019). Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors Dengan Ekstraksi Fitur HSV dan LBP. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 8(1), 74–78.