

Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Berbasis Web Pada Toko Sembiring

Muhammad Yusuf Virasdi^{1*}, Ari Syaripudin¹

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspiptek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: ^{1*}yusufvirasdi2410@gmail.com, ²dosen00671@unpam.ac.id

(* : coressponding author)

Abstrak– Toko Sembiring didirikan pada tahun 2000 oleh Pak Pio Sembiring. Toko ini menjual berbagai barang kebutuhan seperti Sembako, makanan ringan, minuman ringan, keperluan dapur seperti bumbu masakan, kemudian juga ada sabun mandi, sabun cuci baju serta menjual gas dan air mineral dalam galon. Namun seiring dengan mulai banyaknya pesaing, angka penjualan di Toko Sembiring menurun, karena sistem yang sedang berjalan pada Toko Sembiring ini masih bersifat manual yaitu belum adanya system yang dapat mengetahui produk mana saja yang paling diminati oleh konsumen serta belum tersedianya metode yang baik dalam menentukan produk yang paling diminati. Dengan banyaknya barang yang terjual, Toko Sembiring mengalami kesulitan untuk menentukan produk apa yang sudah sesuai dengan target penjualan atau tidak dalam 1 tahun terakhir. Sebuah aplikasi akan dirancang agar bisa melakukan suatu prediksi penjualan dengan menggunakan Metode K-Nearest Neighbor. Tingkat akurasi sistem prediksi yang menggunakan metode K-Nearest Neighbor ini adalah 93,3%. Sehingga diharapkan aplikasi ini bisa memudahkan Toko Sembiring dalam memprediksi penjualan kedepannya agar tercipta sistem kerja yang lebih cepat dan efisien.

Kata Kunci: *K-Nearest Neighbor*, Prediksi, Penjualan

Abstract– Toko Sembiring was founded in 2000 by Mr. Pio Sembiring. This shop sells various necessities such as groceries, snacks, soft drinks, kitchen needs such as cooking spices, then there is also bath soap, laundry soap and sells gas and mineral water in gallons. However, along with the increasing number of competitors, sales figures at the Sembiring Store have decreased, because the system currently running at the Sembiring Store is still manual, that is, there is no system that can find out which products are most in demand by consumers and there is no good method for determining products. the most interested. With so many items sold, Toko Sembiring has difficulty determining what products have met the sales target or not in the last 1 year. An application will be designed to be able to predict sales using the K-Nearest Neighbor Method. The accuracy rate of the prediction system using the K-Nearest Neighbor method is 93.3%. So it is hoped that this application can facilitate the Sembiring Store in predicting future sales in order to create a faster and more efficient work system.

Keywords: *K-Nearest Neighbor*, Prediction, Sales

1. PENDAHULUAN

Prediksi atau peramalan (forecasting) merupakan sebuah kegiatan meramalkan penjualan dimasa mendatang berarti menentukan perkiraan besarnya volume penjualan, bahkan menentukan potensi penjualan dan luas pasar yang dikuasai dimasa mendatang. Selain itu juga dapat membantu perusahaan dalam melakukan perencanaan penyediaan stok. Karena prediksi ini memberikan output terbaik bagi perusahaan sehingga dapat meminimalisir kesalahan dalam perencanaan.

Pada setiap perusahaan, penjualan merupakan hal yang sangat penting agar perusahaan tetap bisa beroperasi dan mendapatkan penghasilan semaksimal mungkin. Setiap perusahaan akan bersaing melalui mutu dan kualitas produk agar penjualan bisa naik sesuai harapan. Saat ini prediksi atau peramalan pada penjualan produk sangat dibutuhkan. Karena dengan adanya prediksi ini, perusahaan akan tahu produk apa yang akan laku lebih banyak dimasa mendatang.

Toko Sembiring didirikan pada tahun 2000 oleh Pak Pio Sembiring. Toko ini menjual berbagai barang kebutuhan seperti Sembako, makanan ringan, minuman ringan, keperluan dapur seperti bumbu masakan, kemudian juga ada sabun mandi, sabun cuci baju serta menjual gas dan air mineral dalam galon. Namun seiring dengan mulai banyaknya pesaing, angka penjualan di Toko Sembiring menurun, karena sistem yang sedang berjalan pada Toko Sembiring ini masih bersifat manual yaitu belum adanya system yang dapat mengetahui produk mana saja yang paling diminati oleh konsumen serta belum tersedianya metode yang baik dalam menentukan produk yang paling

diminati. Dengan banyaknya barang yang terjual, Toko Sembiring mengalami kesulitan untuk menentukan produk apa yang sudah sesuai dengan target penjualan atau tidak dalam 1 tahun terakhir.

Untuk mengatasi permasalahan tersebut, dibutuhkan suatu system yang dapat memprediksi penjualan barang yang terdapat di Toko Sembiring. Dalam hal ini data mining sangat berpengaruh dalam proses menggali sebuah data dari suatu kumpulan data yang berupa pengetahuan yang selama ini tidak diketahui secara manual. Berdasarkan hasil analisis diatas, maka penulis akan melakukan penelitian dengan judul “Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Berbasis Web Pada Toko Sembiring”. Dengan metode ini diharapkan dapat mempermudah Toko Sembiring dalam menentukan produk apa saja yang harus lebih diperbanyak untuk menaikkan jumlah penjualan pada toko.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan gabungan dari berbagai bidang ilmu, antara lain basis data, information retrieval, statistika, algoritma dan machine learning. Bidang ini telah berkembang sejak lama namun makin terasa pentingnya sekarang ini dimana muncul keperluan untuk mendapatkan informasi yang lebih dari data transaksi maupun fakta yang terkumpul selama bertahun-tahun. Data mining adalah cara menemukan informasi tersembunyi dalam sebuah basis data dan merupakan bagian dari proses Knowledge Discovery in Databases (KDD) untuk menemukan informasi dan pola yang berguna dalam data (Jiawei & Kamber, 2006).

Kegiatan data mining biasanya dilakukan pada sebuah data warehouse yang menampung data dalam jumlah besar dari suatu organisasi. Proses data mining mencari informasi baru, berharga dan berguna di dalam sekumpulan data bervolume besar dengan melibatkan manusia dan computer serta bersifat iterative baik melalui proses otomatis ataupun manual (Jiawei & Kamber, 2006). Secara umum, data mining terbagi dalam 2 sifat:

- a. Predictive : menghasilkan model berdasarkan sekumpulan data yang dapat digunakan untuk memperkirakan nilai data yang lain. Ada beberapa metode yang termasuk Predictive Data Mining adalah:
 1. Klasifikasi : Pembagian data ke dalam beberapa kelompok yang telah ditentukan sebelumnya.
 2. Regresi : memetakan data ke suatu prediction variable.
 3. Time Series Analysis : pengamatan perubahan nilai atribut dari waktu ke waktu.
- b. Descriptive : mengidentifikasi pola atau hubungan dalam data untuk menghasilkan informasi baru. Metode yang termasuk dalam Descriptive Data Mining adalah :
 1. Clustering : identifikasi kategori untuk mendeskripsikan data.
 2. Association Rules : identifikasi hubungan antara data yang satu dengan lainnya.
 3. Summarization : pemetaan data ke dalam subset dengan deskripsi sederhana.
 4. Sequence Discovery : identifikasi pola sekuensial dalam data.

2.2 Prediksi

Forecasting adalah suatu usaha untuk meramalkan keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Dalam kehidupan sosial segala sesuatu itu serba tidak pasti, sukar untuk diperkirakan secara tepat. Dalam hal ini perlu diadakan Forecast. Forecasting yang dibuat selalu diupayakan agar dapat meminimalisir kesalahan meramal (forecast error) yang biasanya diukur dengan mean squared error, mean absolute error, dan sebagainya (Bruce, Richard, O’Connell, 1993) Data mining merupakan gabungan dari berbagai bidang ilmu, antara lain basis data, information retrieval, statistika, algoritma dan machine learning. Bidang ini telah berkembang sejak lama namun makin terasa pentingnya sekarang ini dimana muncul keperluan untuk mendapatkan informasi yang lebih dari data transaksi maupun fakta yang terkumpul selama bertahun-tahun. Data mining adalah cara menemukan informasi tersembunyi dalam sebuah basis data dan merupakan bagian dari proses Knowledge Discovery in Databases (KDD) untuk menemukan informasi dan pola yang berguna dalam data (Jiawei & Kamber, 2006).

2.3 K-Nearest Neighbor

K-Nearest Neighbor (K-NN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari *label class* pada K-NN. Tujuan dari algoritma K-NN adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. Algoritma K-NN bekerja berdasarkan jarak terpendek dari *query instance* ke *training* data untuk menentukan K-NN nya. Salah satu cara untuk menghitung jarak dekat atau jauh tetangga menggunakan metode *Euclidean distance* (Yisheng & Shuming, 2009).

Euclidian Distance berfungsi menguji ukuran yang bisa digunakan sebagai interpretasi kedekatan jarak antara dua obyek. Jarak *Euclidean* dapat didefinisikan sebagai jarak antara dua titik pada data latih dan pada data uji dengan persamaan sebagai berikut :

$$d_i = \sqrt{\sum_{i=1}^n (x_{2i} - x_{1i}) \dots \dots \dots (1)}$$

Keterangan:

d_i : jarak *Euclidean*
 x_i : Data *Training*
 x_2 : Data *Testing*
 i : Variabel data
 n : Dimensi atribut / data

Keunggulan dari metode KNN ini adalah *relative* tidak terpengaruh dari error dari data dan juga dapat digunakan dengan kumpulan data dengan jumlah besar. Namun kekurangan metode ini adalah proses pelaksanaannya lambat. Untuk meningkatkan akurasi dari metode yang digunakan dalam aplikasi yang akan dibangun kita menggunakan metode tambahan yang merupakan metode optimasi yang meminimalkan fungsi kesalahan untuk menentukan hasil akhir, *Simple Unweighted voting*. Adapun langkah kerja untuk perhitungan dengan menggunakan metode tersebut yaitu:

- Tentukan K (Jumlah tetangga terdekat).
- Hitung Jarak antara data yang diuji dengan data training.
- Urutkan berdasarkan jarak terdekat dan tentukan apakah termasuk dalam K (jumlah tetangga terdekat).
- Hitung Rata-rata dari data yang termasuk dalam K.

3. ANALISA DAN PEMBAHASAN

Penelitian ini merupakan penelitian lapangan (*field research*) adalah sebuah peneliti yang dilakukan dalam situasi alamiah tetapi didahului oleh semacam *intervensi* (campur tangan) dari pihak peneliti. Dimaksudkan *Intervensi* disini adalah agar fenomena yang dikehendaki peneliti dapat segera tampak dan diamati (Sugiyono, 2017). Dengan begitu akan terjadi semacam kendali atau kontrol parsial terhadap situasi pada Toko Sembiring. Metode yang akan digunakan dalam penelitian ini adalah metode penelitian deskriptif dengan pendekatan kualitatif. Metode penelitian kualitatif adalah metode penelitian yang berlandaskan pada filsafat postpositivisme, digunakan untuk meneliti pada kondisi obyek yang alamiah, (sebagai lawannya adalah eksperimen) dimana peneliti adalah sebagai instrument kunci, teknik pengumpulan data dilakukan secara triangulasi (gabungan), analisis data bersifat induktif/kualitatif, dan hasil penelitian kualitatif lebih menekankan makna dari generalisasi (Sugiyono, 2017).

3.1 Metode K-NN

Dataset yang telah ditentukan dan disimpan di dalam *database* bisa digunakan sebagai data *training* yaitu data yang sudah diketahui status kelasnya seperti terlihat pada table dibawah berisi 120 dataset, terdapat atribut-atribut yang digunakan untuk melakukan perhtiungan diantaranya Kuantitas produk, Kuantitas terjual dan juga Harga.

Tabel 1. Dataset

No	Merk	Kunatitas Produk	Kuantitas Terjual	Harga	Kategori
1	Gas LPG 3Kg	3000	2700	21000	Laris
2	Gas LPG 12Kg	3500	3100	220000	Laris
3	Air Mineral Galon Aqua 19L	4500	3800	19000	Laris
4	Air Mineral Galon Vit 19L	4500	3750	19000	Laris
5	Minyak Goreng Bimoli 1L	6500	5700	19500	Laris
6	Minyak Goreng Kunci Mas 2L	5000	2500	43000	Tidak Laris
7	Minyak goreng Sania 1L	6500	5950	19000	Laris
8	Minyak goreng Fortuna 1L	5500	5000	19000	Laris
9	Minyak goreng Tropical 1L	6000	5000	21500	Laris
10	Beras RojoLele 5Kg	3500	3000	62000	Laris
11	Beras Sentra Ramos 5Kg	3500	2500	60000	Laris
12	Beras Gurih 5Kg	3500	3000	64000	Laris
13	Beras BMW Pandan Wangi 5Kg	3500	3000	75000	Laris
14	Beras Sania 5Kg	3000	2750	62000	Laris
15	Beras Kembang 5Kg	3500	2800	60000	Tidak Laris
16	Beras Merah 5Kg	2000	1600	75000	Laris
17	Fanta 1L botol	3500	3000	11000	Laris
18	Coca-Cola 1L botol	3500	3100	11000	Laris
19	Sprite 1L botol	3500	3150	11000	Laris
20	Fanta 390ML botol	3000	2000	7500	Laris
21	Coca-Cola 390ML	3000	2500	7500	Laris
22	Sprite 390ML	3000	2150	7500	Laris
23	Fanta 250ML	2000	1750	5000	Laris
24	Coca-Cola 250ML	2000	1800	5000	Laris
25	Sprite 250ML	2000	1700	5000	Laris
26	Air Mineral Aqua 1,5L	2500	2000	5000	Laris
27	Air Mineral Vit 1,5L	2500	2000	5000	Laris
28	Air Mineral LeMinerale 1,5L	2500	2150	5000	Laris
29	Air Mineral Aqua 600ML	3500	3100	3000	Laris
30	Air Mineral Vit 600ML	3500	2000	3000	Tidak Laris
31	Air Mineral LeMinerale 600ML	3500	3000	3000	Laris
32	Air Mineral Cleo 600ML	2500	1500	4000	Tidak Laris
33	Air Mineral 330ML	4000	3000	2500	Laris
34	Air Mineral VIt 330ML	4500	4100	2500	Laris
35	Air Mineral Leminerale 330ML	4000	3050	2500	Laris
....
120	Tebs Sparkling kaleng	1500	900	4000	Laris

Pada sistem ini terdapat beberapa tahapan dalam menyelesaikan masalah prediksi penjualan diantaranya:

a. Menentukan Parameter K (Jumlah Tetangga Terdekat)

Menentukan parameter k biasanya menggunakan angka dengan bilangan yang ganjil untuk menghindari kemunculan nilai jarak yang sama dalam proses klasifikasi. Oleh karena itu nilai K yang digunakan adalah 15.

b. Menentukan Kuadrat Jarak Euclidean Terhadap Data Training Yang Diberikan.

Selanjutnya yang akan dilakukan adalah melakukan input data *testing* produk baru dengan memasukan atribut-atribut kriteria diantaranya adalah atribut Kuantitas produk, Kuantitas terjual dan juga Harga. Contohnya di *input* data produk baru dengan data seperti berikut:

Nama Produk : Hydro Coco 1L
 Kuantitas Produk : 2500
 Kuantitas Terjual : 1800
 Harga : 25000
 Status : Belum diklasifikasi
 Jumlah Tetangga : 15

Untuk memperoleh jarak yang memiliki ukuran yang kecil dari perhitungan maka kita akan menggunakan normalisasi data. Misalnya menggunakan data *testing* Produk diatas.

Rumusnya adalah sebagai berikut:

$$Normalisasi = \frac{data_x - data_{min}}{data_{max} - data_{min}}$$

Penjelasan :

x : Data
 min : Data Minimum
 max : Data maximum

Pada normalisasi data Kuantitas produk, dimana dari *dataset* yang ada pada table 1 diketahui Kuantitas produk minimal adalah 1000 dan maksimalnya adalah 7000.

$$Normalisasi = \frac{(2500-1000)}{(7000-1000)} = 0.25$$

Pada normalisasi data Kuantitas terjual, dimana dari table 1 diketahui Kuantitas terjual minimal adalah 500 dan maksimalnya adalah 6800.

$$Normalisasi = \frac{(1800-500)}{(6800-500)} = 0.206$$

Pada normalisasi data Harga, diketahui dari table 1 adalah Harga minimal 1000 dan maksimalnya adalah 220000.

$$Normalisasi = \frac{(25000-1000)}{(220000-1000)} = 0.109$$

Dengan melakukan cara perhitungan yang sama pada seluruh data maka didapatkan hasil seperti pada tabel dibawah.

Tabel 2. Normalisasi Data

No	Merk	Kuantitas Produk	Kuantitas Terjual	Harga	Kategori
1	Gas LPG 3Kg	0.333	0.349	0.091	Laris
2	Gas LPG 12Kg	0.417	0.413	1	Laris

3	Air Mineral Galon Aqua 19L	0.583	0.524	0.082	Laris
4	Air Mineral Galon Vit 19L	0.583	0.516	0.082	Laris
5	Minyak Goreng Bimoli 1L	0.917	0.825	0.084	Laris
6	Minyak Goreng Kunci Mas 2L	0.667	0.317	0.192	Tidak Laris
7	Minyak goreng Sania 1L	0.917	0.865	0.082	Laris
8	Minyak goreng Fortuna 1L	0.75	0.714	0.082	Laris
9	Minyak goreng Tropical 1L	0.833	0.714	0.094	Laris
10	Beras RojoLele 5Kg	0.417	0.397	0.279	Laris
11	Beras Sentra Ramos 5Kg	0.417	0.317	0.269	Laris
12	Beras Gurih 5Kg	0.417	0.397	0.288	Laris
13	Beras BMW Pandan Wangi 5Kg	0.417	0.397	0.338	Laris
14	Beras Sania 5Kg	0.333	0.357	0.279	Laris
15	Beras Kembang 5Kg	0.417	0.365	0.269	Tidak Laris
16	Beras Merah 5Kg	0.167	0.175	0.338	Laris
17	Fanta 1L botol	0.417	0.397	0.046	Laris
18	Coca-Cola 1L botol	0.417	0.413	0.046	Laris
19	Sprite 1L botol	0.417	0.421	0.046	Laris
20	Fanta 390ML botol	0.333	0.238	0.03	Laris
21	Coca-Cola 390ML	0.333	0.317	0.03	Laris
22	Sprite 390ML	0.333	0.262	0.03	Laris
23	Fanta 250ML	0.167	0.198	0.018	Laris
24	Coca-Cola 250ML	0.167	0.206	0.018	Laris
25	Sprite 250ML	0.167	0.19	0.018	Laris
26	Air Mineral Aqua 1,5L	0.25	0.238	0.018	Laris
27	Air Mineral Vit 1,5L	0.25	0.238	0.018	Laris
28	Air Mineral LeMinerale 1,5L	0.25	0.262	0.018	Laris
29	Air Mineral Aqua 600ML	0.417	0.413	0.009	Laris
30	Air Mineral Vit 600ML	0.417	0.238	0.009	Tidak Laris
31	Air Mineral LeMinerale 600ML	0.417	0.397	0.009	Laris
32	Air Mineral Cleo 600ML	0.25	0.159	0.014	Tidak Laris
33	Air Mineral 330ML	0.5	0.397	0.007	Laris
34	Air Mineral VIt 330ML	0.583	0.571	0.007	Laris
35	Air Mineral Leminerale 330ML	0.5	0.405	0.007	Laris
....
120	Tebs Sparkling kaleng	0.083	0.063	0.014	Laris

Setelah melakukan *normalisasi* data, kemudian langkah selanjutnya menggunakan rumus *Euclidean distance* untuk mencari jarak data *Hydro Coco* 1L terhadap data *training*.

Rumus *Euclidean distance*:

$$d_i = \sqrt{\sum_{i=1}^n (x_{2i} - x_{1i}) \dots \dots \dots (1)}$$

Metode yang digunakan pada pengumpulan data dalam program aplikasi ini adalah sebagai berikut:

Keterangan:

d_i : jarak *Euclidean*

X_i : Data *Training*

x_2 : Data *Testing*

i : Variabel data

n : Dimensi atribut / data

Sehingga jika dilakukan perhitungan menggunakan rumus diatas menghasilkan perhitungan sebagai berikut.

$$d_i : \sqrt{(0.333 - 0.25)^2 + (0.349 - 0.206)^2 + (0.091 - 0.109)^2}$$

$$d_i : \sqrt{(0.083)^2 + (0.143)^2 + (-0.018)^2}$$

$$d_i : \sqrt{(0.006) + (0.020) + (0.000)}$$

$$d_i : \sqrt{(0.026)}$$

c. Hasil diurutkan secara *ascending*

Selanjutnya data diurutkan secara *ascending* dari data yang memiliki jarak terkecil sampai dengan data jarak terbesar dari data *testing*. Kemudian diambil data teratas sesuai dengan jumlah nilai yang sudah ditentukan diawal yaitu 15 data teratas yang memiliki nilai jarak paling kecil dengan data *testing*. Maka hasilnya adalah sebagai berikut.

Tabel 3. Mengambil 15 data teratas

No.	Merk	Kategori	Jarak
74	Sabun cuci baju Rinso	Laris	0.033
56	Susu Ultra Milk 1L	Laris	0.05
72	Sabun cuci baju DAIA	Laris	0.051
98	Sasa Saus Sambal Botol 335ML	Laris	0.056
109	Tepung terigu Segitiga Biru 1Kg	Laris	0.056
57	Susu Diamond Milk 1L	Laris	0.061
82	Pasta gigi Ciptadent 190g	Laris	0.075
77	Pasta gigi Pepsodent 190g	Laris	0.078
78	Pasta gigi Pepsodent Herbal	Laris	0.082
81	Pasta gigi Ciptadent 120g	Laris	0.086
37	Larutan Penyegar Cap Kaki Tiga botol	Laris	0.086
64	Pembersih Toilet Wipol	Laris	0.087
83	Pasta gigi Kodomo	Laris	0.088
47	HydroCoco 330ML	Laris	0.091
50	Adem Sari Cingku Botol	Laris	0.095

Setelah didapatkan 15 data teratas berdasarkan jarak terkecil, maka hasil yang akan didapatkan adalah:

Laris : 15

Tidak Laris : 0

Dari banyaknya jumlah yang tertera diatas, diambil dari data mayoritas, maka didapatkan kesimpulan bahwa Penjualan dari produk *HydroCoco 1L* adalah Laris. Sehingga penjualan produk tersebut bisa dilanjutkan oleh Toko Sembiring.

4. IMPLEMENTASI

4.1 Pengujian Akurasi Sistem

Pengujian akurasi sistem disini yaitu pengujian data *Training* yang ada dengan data *Training* sendiri untuk menguji keakuratan sistem. Karena K-Nearest Neighbor menggunakan nilai data *Training* bervariasi sebagai acuan dalam menentukan keputusan. Dalam pengujian akurasi ini, nilai tetangga yang digunakan adalah 15 (limabelas).

Tabel 4. Pengujian Akurasi

No	Merk	<i>Training</i>	<i>k-NN</i>	Hasil
1	Gas LPG 3Kg	Laris	Laris	Sesuai
2	Gas LPG 12Kg	Laris	Laris	Sesuai
3	Air Mineral Galon Aqua 19L	Laris	Laris	Sesuai
4	Air Mineral Galon Vit 19L	Laris	Laris	Sesuai
5	Minyak Goreng Bimoli 1L	Laris	Laris	Sesuai
6	Minyak Goreng Kunci Mas 2L	Tidak Laris	Laris	Tidak Sesuai
7	Minyak goreng Sania 1L	Laris	Laris	Sesuai
8	Minyak goreng Fortuna 1L	Laris	Laris	Sesuai
9	Minyak goreng Tropical 1L	Laris	Laris	Sesuai
10	Beras RojoLele 5Kg	Laris	Laris	Sesuai
11	Beras Sentra Ramos 5Kg	Laris	Laris	Sesuai
12	Beras Gurih 5Kg	Laris	Laris	Sesuai
13	Beras BMW Pandan Wangi 5Kg	Laris	Laris	Sesuai
14	Beras Sania 5Kg	Laris	Laris	Sesuai
15	Beras Kembang 5Kg	Tidak Laris	Laris	Tidak Sesuai
16	Beras Merah 5Kg	Laris	Laris	Sesuai
17	Fanta 1L botol	Laris	Laris	Sesuai
18	Coca-Cola 1L botol	Laris	Laris	Sesuai
19	Sprite 1L botol	Laris	Laris	Sesuai
20	Fanta 390ML botol	Laris	Laris	Sesuai
21	Coca-Cola 390ML	Laris	Laris	Sesuai
22	Sprite 390ML	Laris	Laris	Sesuai
23	Fanta 250ML	Laris	Laris	Sesuai
24	Coca-Cola 250ML	Laris	Laris	Sesuai
25	Sprite 250ML	Laris	Laris	Sesuai
26	Air Mineral Aqua 1,5L	Laris	Laris	Sesuai
27	Air Mineral Vit 1,5L	Laris	Laris	Sesuai
28	Air Mineral LeMinerale 1,5L	Laris	Laris	Sesuai
29	Air Mineral Aqua 600ML	Laris	Laris	Sesuai
30	Air Mineral Vit 600ML	Tidak Laris	Laris	Tidak Sesuai
31	Air Mineral LeMinerale 600ML	Laris	Laris	Sesuai
32	Air Mineral Cleo 600ML	Tidak Laris	Laris	Tidak Sesuai
33	Air Mineral 330ML	Laris	Laris	Sesuai
34	Air Mineral Vit 330ML	Laris	Laris	Sesuai

35	Air Mineral Leminerale 330ML	Laris	Laris	Sesuai
....
120	Tebs Sparkling kaleng	Laris	Laris	Sesuai

Berdasarkan hasil pengujian akurasi sistem untuk data *Training* diatas menyatakan bahwa dari 120 data *training*, terdapat 8 data *training* yang hasil statusnya tidak sesuai dengan pengujian sistem.

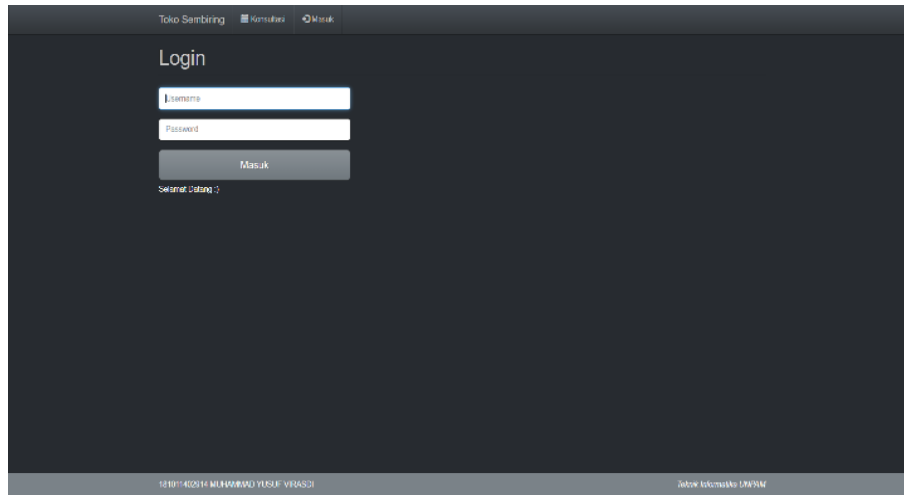
Presentase Akurasi

$$= \frac{\text{Jumlah Data Training Sesuai}}{\text{Jumlah Data Training}} \times 100\%$$

$$= \frac{112}{120} \times 100\% = 93,3\%$$

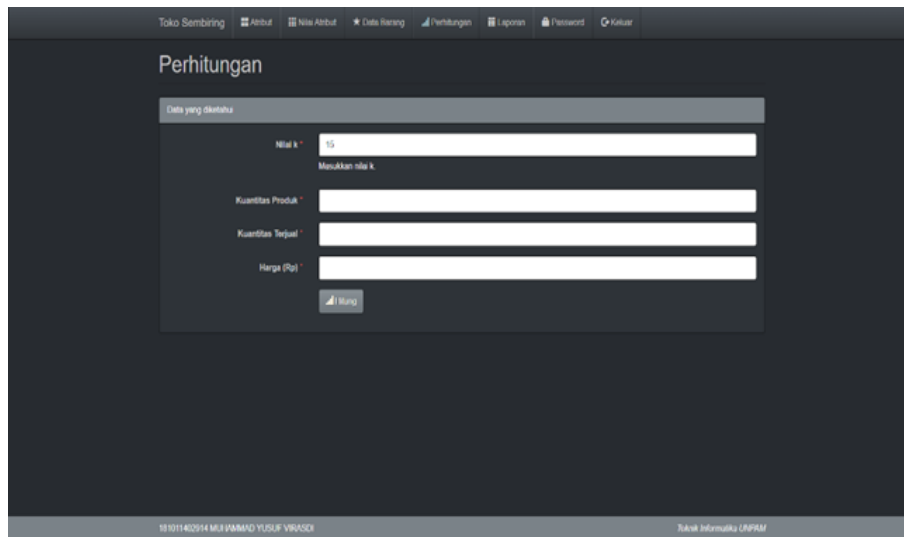
4.2 Implementasi Aplikasi

a. Tampilan Halaman *Login*



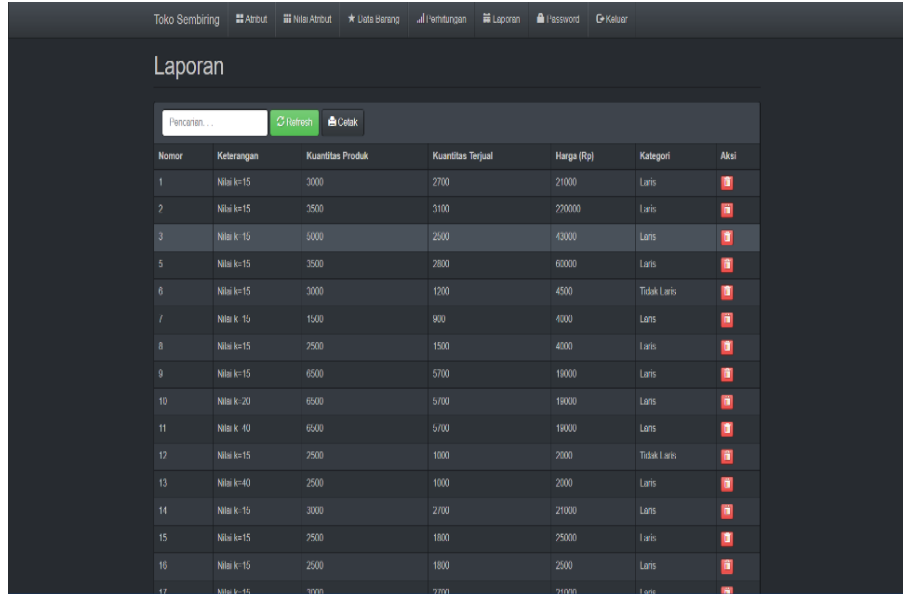
Gambar 1. Tampilan Halaman *Login*

b. Tampilan Halaman *Form Perhitungan*



Gambar 2. Tampilan Halaman *Form Perhitungan*

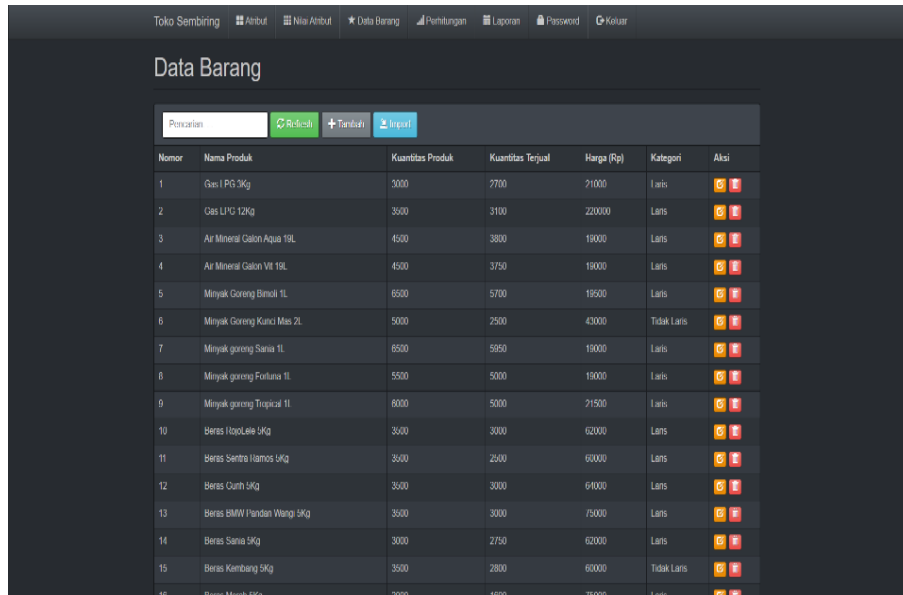
c. Tampilan Halaman Hasil Perhitungan



Nomor	Keterangan	Kuantitas Produk	Kuantitas Terjual	Harga (Rp)	Kategori	Aksi
1	Nilai k=15	3000	2700	21000	Laris	[C] [E]
2	Nilai k=15	3500	3100	22000	Laris	[C] [E]
3	Nilai k=15	5000	2500	43000	Laris	[C] [E]
5	Nilai k=15	3500	2800	60000	Laris	[C] [E]
6	Nilai k=15	3000	1200	4500	Tidak Laris	[C] [E]
7	Nilai k=15	1500	900	4000	Laris	[C] [E]
8	Nilai k=15	2500	1500	4000	Laris	[C] [E]
9	Nilai k=15	6500	5700	18000	Laris	[C] [E]
10	Nilai k=20	6500	5700	18000	Laris	[C] [E]
11	Nilai k=40	6500	5700	18000	Laris	[C] [E]
12	Nilai k=15	2500	1000	2000	Tidak Laris	[C] [E]
13	Nilai k=40	2500	1000	2000	Laris	[C] [E]
14	Nilai k=15	3000	2700	21000	Laris	[C] [E]
15	Nilai k=15	2500	1800	25000	Laris	[C] [E]
16	Nilai k=15	2500	1800	2500	Laris	[C] [E]
17	Nilai k=40	3000	2700	21000	Laris	[C] [E]

Gambar 3. Tampilan Halaman Hasil Perhitungan

d. Tampilan Halaman Data Training



Nomor	Nama Produk	Kuantitas Produk	Kuantitas Terjual	Harga (Rp)	Kategori	Aksi
1	Gas LPG 3Kg	3000	2700	21000	Laris	[C] [E]
2	Gas LPG 12Kg	3500	3100	22000	Laris	[C] [E]
3	Air Mineral Galon Aqua 19L	4500	3800	18000	Laris	[C] [E]
4	Air Mineral Galon Vit 19L	4500	3750	18000	Laris	[C] [E]
5	Minyak Goreng Bendi 1L	6500	5700	19500	Laris	[C] [E]
6	Minyak Goreng Kanci Mas 2L	5000	2500	43000	Tidak Laris	[C] [E]
7	Minyak goreng Sasia 1L	6500	5950	19000	Laris	[C] [E]
8	Minyak goreng Fortuna 1L	5500	5000	19000	Laris	[C] [E]
9	Minyak goreng Tropical 1L	6000	5000	21500	Laris	[C] [E]
10	Beras InapLela 5Kg	3500	3000	62000	Laris	[C] [E]
11	Beras Santra Hama 5Kg	3500	2500	60000	Laris	[C] [E]
12	Beras Putih 5Kg	3500	3000	64000	Laris	[C] [E]
13	Beras BMW Handan Wang 5Kg	3500	3000	75000	Laris	[C] [E]
14	Beras Sasia 5Kg	3000	2750	62000	Laris	[C] [E]
15	Beras Kambang 5Kg	3500	2800	60000	Tidak Laris	[C] [E]
16	Beras Merah 5Kg	3000	2500	75000	Laris	[C] [E]

Gambar 4. Tampilan Halaman Data Training

5. KESIMPULAN

Dari hasil pengujian fungsional program, fitur-fitur pengelolaan data (simpan, edit, hapus, tambah, report) pada sistem menunjukkan hasil berjalan tanpa error. Hasil pengujian akurasi metode *k-nn* yang diterapkan dalam prediksi penjualan di Toko Sembiring dengan menggunakan 120 data *Training* menghasilkan nilai akurasi sebesar 93,3%. Kemudian Pada penelitian tentang penerapan metode *k-nearest neighbor (k-NN)* untuk memprediksi penjualan di Toko Sembiring dapat dikembangkan dengan metode lain yaitu dengan metode *Naïve Bayes*, *Decision Tree* dan sebagainya.

REFERENCES

- Yolanda, I., & Fahmi, H. (2020). Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT. Nippon Indosari Corpindo Tbk Menggunakan Metode K-Nearest Neighbor. *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, 3(1.1), 9-15
- Muttaqin, A. G., Auliasari, K., & Wahyuni, F. S. (2020). Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Berbasis Web Pada Pt. Wika Industri Energy. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 4(2), 1-6.
- WPR, A. A., Rozi, F., & Sukmana, F. (2021). Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 6(1), 155-160.
- Jiawei, H., Kamber, M., & Pei, I. (2006). *Data mining: concepts and techniques*. Morgan kaufmann.
- Bowerman, B. L., & O'Connell, R. T. (1993). *Forecasting and Time Series*. Wadsworth. Inc, United States.
- Lv, Y., Tang, S., & Zhao, H. (2009, April). Real-time highway traffic accident prediction based on the k-nearest neighbor method. In *2009 international conference on measuring technology and mechatronics automation (Vol. 3, pp. 547-550)*. IEEE.
- Annastasya, E. F. (2015). *Analisis pengungkapan CSR melalui website dan annual report pada Industri Perbankan Indonesia tahun 2011-2013 (Doctoral dissertation, Universitas Negeri Malang)*.
- Aswadi, M. (2015). *Database Dasar With Xampp*. Surabaya: CV. Garuda Mas Sejahtera dari <http://elexmedia.id>.
- Minarni, M., & Susanti, S. (2014). Sistem Informasi Inventory Obat Pada Rumah Sakit Umum Daerah (RSUD) Padang. *Jurnal Momentum ISSN: 1693-752X*, 16(1).
- Suhaidi, M. (2016). *Konsep Dasar Pemrograman Web dengan PHP dan MySQL*. CV. Budi Utama.
- Supono, P. Virdiandry. 2016. *Pemrograman Web dengan menggunakan PHP dan Framework Codeigniter*. Yogyakarta: Deepublish.
- HERDIANTO, E. (2013). *PERANCANGAN SISTEM INFORMASI PENJUALAN MEUBEL PADA PT. ANDARA FURNITURE MENGGUNAKAN VISUAL BASIC 6.0 (Doctoral dissertation, Universitas Muhammadiyah Ponorogo)*.
- Abdillah, A. A., Adinugraha, M. T., & Bianca, I. (2019). Rancang Bangun Aplikasi Manajemen Gudang Suku Cadang Alat Berat PNJ. *Politeknologi Vol. 18 No. 3*, 307-313.
- Jacobson, I., & Booch, J. R. G. (2021). *The unified modeling language reference manual*.
- PWA, A. A. R., Maulana, M. H., Andini, C. D., & Nadziroh, F. (2018). Sistem Peminjaman Ruang Online (SPRO) dengan Metode UML (Unified Modeling Language). *Jurnal Teknologi dan Terapan Bisnis*, 1(1), 1-8.
- Wibisono, W., & Baskoro, F. (2002). Pengujian perangkat lunak dengan menggunakan model behaviour UML. *Jurnal Ilmiah Teknologi Informasi*, 1(1), 43-50.
- Rosa, A.S., Shalahuddin, M. (2018). *Rekayasa Perangkat Lunak (terstruktur dan berorientasi objek)*. Bandung: Informatika Bandung.
- Mohammed, M. A., Muhammed, D. A., & Abdullah, J. M. (2015). *Practical Approaches of Transforming ER Diagram into Tables. International Journal of Multidisciplinary and Scientific Emerging Research*, 4.
- Hidayat, T., & Muttaqin, M. (2020). Pengujian sistem informasi pendaftaran dan pembayaran wisuda online menggunakan black box testing dengan metode equivalence partitioning dan boundary value analysis.
- Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.