

Sistem Prediksi Produksi Minuman Menggunakan Metode K-Nearest Neighbor Berbasis Web

Faisal¹, Syarifuddin², Jufri³, Sriwahyuningsih⁴

^{1,2}Universitas Dipa: Jl.Perintis Kemerdekaan KM.09, 0411-587194/0411-588283

³Program Studi Teknik Informatika, Universitas Dipa, Makassar

e-mail: ¹fhaizalh66@gmail.com, ²syarifaaab@gmail.com, ³jufri.idp@undipa.ac.id, ⁴sriwahyuningsihpiu.idp@gmail.com

Abstrak

Produksi minuman ini sangat mudah ditemukan di berbagai tempat seperti pusat perbelanjaan, sekitar sekolah, pasar, rumah sakit, pinggiran jalan, taman kota, kampus dan tempat umum lainnya. Pengusaha akan menggunakan prediksi atau peramalan pada produksi minuman untuk mengetahui minuman diwaktu apa banyak terjual pada masa yang akan datang. Namun terdapat satu kasus dimana prediksi ini justru terbalik, dalam artian apabila waktu minuman diprediksi akan naik penjualannya, justru malah terjadi penurunan pada minuman tersebut diperiode yang telah di prediksi ini terjadi karena metode yang di gunakan kurang akurat dalam memprediksi penjualan dimasa yang akan datang ini juga berdampak pada sulitnya penyediaan bahan yang tepat. Pada penelitian ini dibuatkan program prediksi penjualan menggunakan metode K-Nearest Neighbor berbasis website dengan menggunakan 500 data. Hasil prediksi yang diperoleh adalah jenis minuman yang masuk kategori tinggi masuk dalam prediksi penjualan periode bulan berikutnya.

Kata Kunci: *Prediksi, Minuman, K-Nearest Neighbor, Website.*

Abstract

The production of this drink is very easy to find in various places such as shopping centers, around schools, markets, hospitals, roadsides, city parks, campuses and other public places. Entrepreneurs will use predictions or forecasting on beverage production to find out at what time drinks will be sold in the future. However, there is one case where this prediction is actually reversed, in the sense that when the drink is predicted to increase in sales, a decrease in the drink in the predicted period occurs because the method used is less accurate in predicting future sales as well. impact on the difficulty of providing the right materials. In this study a sales prediction program was made using the K-Nearest Neighbor method based on a website using 500 data. The prediction results obtained are types of drinks that are in the high category included in the sales prediction for the next month's period.

Keyword: *Prediction, Drink, K-Nearest Neighbor, Website.*

1. PENDAHULUAN

Produksi minuman adalah usaha yang bergerak dibidang penjualan minuman dengan berbagai macam varian rasa seperti greentea, tahitea, red velvet, dan varian lainnya. Produksi minuman ini sangat mudah ditemukan di berbagai tempat seperti pusat perbelanjaan, sekitar sekolah, pasar, rumah sakit, pinggiran jalan, taman kota, kampus dan tempat umum lainnya.

Pada umumnya usaha akan menggunakan prediksi atau peramalan pada produksi minuman untuk mengetahui minuman diwaktu apa banyak terjual pada masa yang akan datang. Namun terdapat satu kasus dimana prediksi ini justru terbalik, dalam artian apabila waktu

minuman diprediksi akan naik penjualannya, justru malah terjadi penurunan pada minuman tersebut diperiode yang telah di prediksi ini terjadi karena metode yang di gunakan kurang akurat dalam memprediksi penjualan dimasa yang akan datang ini juga berdampak pada sulitnya penyediaan bahan yang tepat.

Berdasarkan uraian tersebut, dilakukan penelitian untuk membuat program prediksi penjualan berbasis web yang dapat digunakan untuk memprediksi hasil penjualan minuman menggunakan metode K-Nearest Neighbor berbasis website dengan tujuan untuk mengetahui sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi dari masa sebelumnya dan masa sekarang.

Hasil yang akan diharapkan dari penelitian ini, pengusaha kecil dapat mengetahui atau memprediksi waktu minuman banyak terjual berdasarkan waktu atau musim dan juga dapat mengklasifikasikan produk minuman mana yang memiliki potensial akan laris terjual dan yang tidak, sehingga pengusaha kecil dapat mengurangi produksi minuman atau memperbanyak minuman berdasarkan hasil prediksi sistem.

1.1 Pokok Masalah

Pokok permasalahan dalam penelitian ini yaitu bagaimana penerapan metode KNN untuk memprediksi waktu dan musim banyaknya minuman terjual.

1.2 Tujuan Penelitian

Penelitian dilakukan dengan tujuan untuk mengetahui hasil prediksi banyaknya minuman terjual berdasarkan waktu atau musim sehingga pengusaha kecil dapat mengatur jumlah produksi minuman.

2. METODE PENELITIAN

2.1 Landasan Teori

Adapun landasan teori yang digunakan sebagai berikut

2.2.1 Data Mining

Data mining dapat di artikan sebagai teknik menemukan suatu pola yang tidak diketahui dengan cara menggali informasi terpendam dari database yang sangat besar [1]. Tugas data mining berdasarkan fungsionalitasnya dapat dikelompokkan 6 kelompok yaitu [2]:

1. Klasifikasi (*classification*)
2. Klasterisasi (*clustering*)
3. Regresi (*regression*)
4. Deteksi anomali (*anomaly detection*)
5. Pembelajaran aturan asosiasi (*association rule learning*)
6. Perangkuman (*summarization*)

2.2.2 Teknik Klasifikasi

Klasifikasi adalah suatu proses untuk menyatakan suatu objek ke salah satu kategori yang sudah didefinisikan sebelumnya. Proses pembelajaran fungsi target (model klasifikasi) yang memetakan setiap sekumpulan atribut x (input) ke salah satu kelas y yang didefinisikan sebelumnya. Inputnya adalah sekumpulan record (training set) dari setiap record tersebut terdiri atas sekumpulan atribut, salah satu atribut adalah kelas. Mencari model untuk atribut kelas sebagai fungsi dari nilai-nilai untuk atribut yang lain (*teknik.klasifikasi.https://ocw.upj.ac.id/files/Handout-TIF311-DM-4.PDF*)[3].

2.2.3 K-Nearest Neighbor

K-Nearest Neighbour (KNN) merupakan kelompok instance based learning. Algoritma KNN juga termasuk salah satu teknik dalam *lazy learning*. KNN bekerja dengan cara mencari k-objek yang terdapat di dalam data training terdekat dengan objek yang ada pada data testing (data baru) [4].

Untuk merumuskan bagaimana jarak diantara titik pada data training (x) dan titik pada data testing (y), maka akan menggunakan rumus *Euclidean*.

$$D(x, y) = \sqrt{\sum_{k=1}^n (X_k - Y_k)^2} \dots \dots \dots (1)$$

Keterangan :

D merupakan jarak titik data training x dengan titik data testing y yang akan diklasifikasikan, dimana x adalah x_1, x_2, \dots, x_i dan y adalah y_1, y_2, \dots, y_i serta merepresentasikan nilai atribut dan n adalah dimensi atribut.

2.2.4 Website

Website merupakan kumpulan halaman digital yang berisi informasi berupa teks, animasi, gambar, suara dan video atau gabungan dari semuanya yang terkoneksi oleh internet, sehingga dapat dilihat oleh siapapun yang terkoneksi jaringan internet (Ani Oktarini Sari, 2019) [5].

2.2.5 Pengumpulan Data

Metode yang digunakan untuk pengumpulan data yaitu observasi, Observasi merupakan pengamatan dan pencatatan secara sistematis mengenai data yang diteliti. Observasi ini menjadi salah satu Teknik pengumpulan data apabila sesuai dengan tujuan penelitian, yang direncanakan dan dicatat secara sistematis serta dikontrol keadaan (rehabilitas) dan kesahihannya (validasinya) (Sugiyono, Metode Penelitian Kuantitatif Kualitatif dan R&D (Bandung : Alfabeta, 2012). [6]

2.2.6 Jenis Penelitian

Adapun jenis penelitian yang digunakan adalah penelitian kuantitatif, Penelitian kuantitatif adalah penelitian empiris di mana data-datanya dalam bentuk sesuatu yang dapat dihitung. Penelitian kuantitatif memperhatikan pengumpulan dan analisis data dalam bentuk numerik. Menurut Sugiyono (2018;13)[7]

2.2.7 Penelitian terkait

Berikut terdapat beberapa penelitian terkait yang dapat diuraikan sebagai berikut :

Kurniawan, 2017 : Implementasi Metode *K-Nearest Neighbor (KNN)* Untuk Mengelompokkan Status Ekonomi Masyarakat Desa Kedak Kecamatan Semen : Algoritma *K-Nearest Neighbor* dapat digunakan untuk mengelompokkan status masyarakat miskin, sedang, dan kaya berdasarkan kriteria yang diberikan oleh kepala desa kepada peneliti : menggunakan metode *K-Nearest Neighbor (KNN)* dengan memperhitungkan segala kriteria yang dipilih, setelah perhitungan semuanya selesai maka proses terakhir yaitu mencari nilai terdekat antara data lama (*data training*) dengan data baru (*data testing*) dan mencari nilai kemunculan data terbanyak yang dijadikan acuan sebagai hasil dari algoritma metode.

Asahar Johar, 2016 : Implementasi Metode *K-Nearest Neighbor (KNN)* Dan *Simple Additive Weighting (Saw)* Dalam Pengambilan Keputusan Seleksi Penerimaan Anggota Paskibraka: Hasil dari aplikasi ini yaitu berupa rekomendasi nama peserta yang lolos dan tidak lolos seleksi berdasarkan hasil perbandingan nilai masing-masing peserta: Metode *simple Additive Weighting* digunakan untuk melakukan perbandingan. Aplikasi ini dibangun dengan menggunakan bahasa pemrograman *PHP*. Metode pengembangan sistem yang digunakan untuk membangun aplikasi ini adalah *model waterfall* dan *Unified Modelling Language (UML)* sebagai perancangan sistem.

Lestari, 2017: Perbandingan Klasifikasi Beasiswa Toyota Astra Menggunakan *K-Nearest Neighbor Classifier* Dan *Naïve Bayes* Sebagai Penentu Metode Klasifikasi: Hasil bahwa pengimplementasian algoritma *k-nearest neighbor classifier* untuk klasifikasi penerimaan beasiswa Toyota Astra bisa dilakukan secara akurat dengan nilai akurasi 90.26%: mengimplementasikan sistem pendukung keputusan penerimaan beasiswa Toyota Astra. Data pendaftar dan penerima beasiswa didapatkan dari Badan Administrasi Akademik dan Kemahasiswaan (BAAK) ITS berupa data beasiswa Toyota Asrtra dari tahun 2015 sampai 2016 dalam lingkup institut.

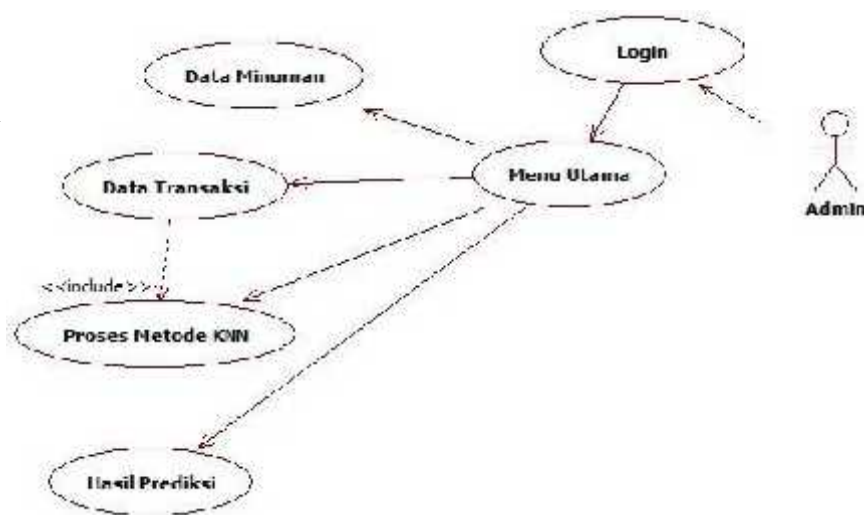
Mochammad Faris, 2021: Penentuan Penerima Bantuan Sosial Bagi Siswa Yang Terkena Dampak Covid-19 Menggunakan Metode *K-Nearest Neighbor (KNN)*: sistem ini dapat dijalankan dengan baik pada 2 browser yaitu *google chrome* dan *mozilla firefox*. pengujian perhitungan algoritma *K-Nearest Neighbor* dengan nilai parameter $K = 5$ menggunakan *confusion matrix* di dapat nilai akurasi sebesar 98%, nilai presisi sebesar 100%, dan nilai *recall* sebesar 96%: menerapkan metode *K-Nearest Neighbor* akan memudahkan pihak sekolah dalam mengklasifikasi data siswa yang baru, apakah termasuk dalam kategori terdampak covid-19 atau tidak, sehingga sekolah dapat memberikan bantuan atau tidak secara cepat dan akurat.

3. HASIL DAN PEMBAHASAN

3.1 Use Case Diagram

Use Case Diagram adalah diagram yang dibuat pada awal penelitian untuk memberikan penjelasan umum mengenai sistem dengan dunia luar serta fitur apa saja yang dimiliki oleh sistem.

adalah
Diagram
penelitian



Berikut
Use Case
pada
ini:

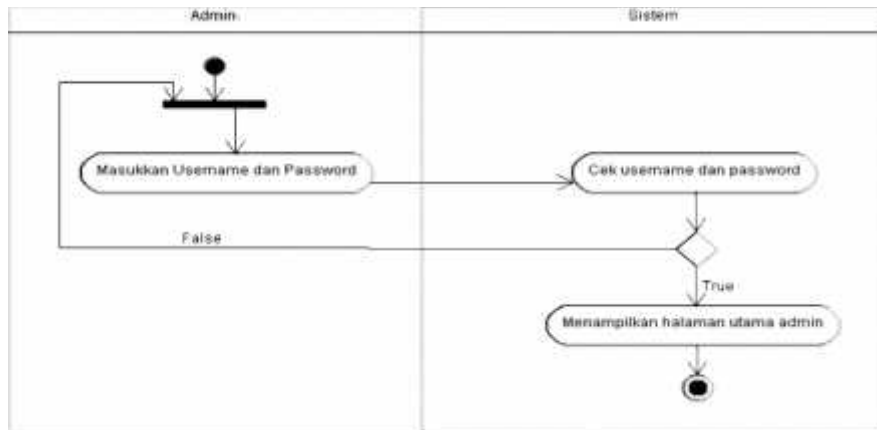
Gambar 1 Usecase Diagram

Pada gambar 1 menggambarkan bagaimana cara admin berinteraksi dengan sistem yang akan dibuat. Pada sistem server ini, terdapat 2 aktor yaitu aktor sebagai admin dapat melakukan mengolah data pelanggan, data minuman memasukkan data latih atau transaksi penjualan,

melakukan proses perhitungan berdasarkan data transaksi penjualan dan melihat hasil perhitungan menggunakan algoritma KNN, sedangkan aktor pelanggan dapat melihat data pelanggan dan melihat hasil klasifikasi data transaksi penjualan dengan melihat produksi yang banyak diminati dan tidak diminati pelanggan.

3.2 Activity Diagram

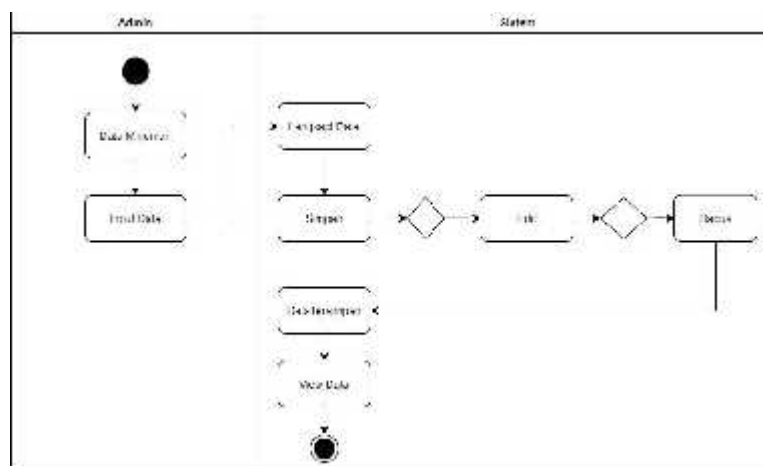
1. Activity Diagram Login



Gambar 2 Activity Diagram Login

Pada gambar 2 merupakan aktifitas untuk manajemen sistem. Aktifitas ini dimulai dari admin masuk dalam web browser, lalu akan tampil form login dimana admin dapat memasukkan username dan password, jika data username dan password salah atau tidak benar maka akan tampil pesan kesalahan, sebaliknya jika username dan password benar maka sistem akan menampilkan halaman utama.

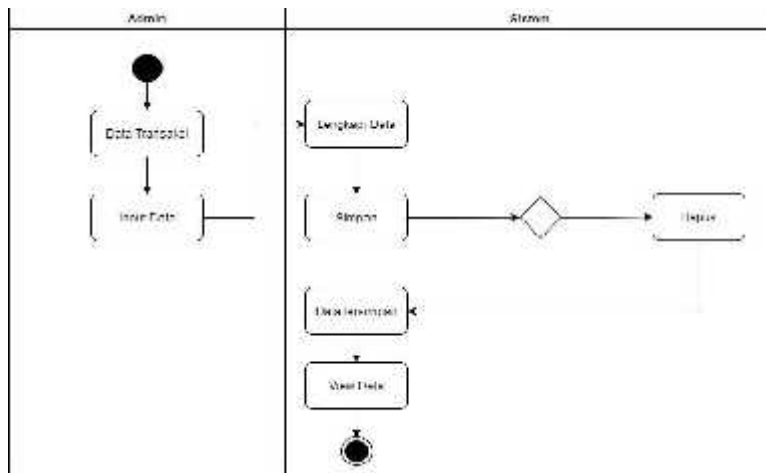
2. Activity Diagram Data Minuman



Gambar 3 Activity Diagram Data Minuman

Pada gambar 3 dijelaskan bahwa pengguna memilih menu data minuman, lalu sistem akan menampilkan form data minuman. Pengguna dapat menambahkan data minuman, sehingga sistem akan menyimpan data pada database dan bila ditemukan data yang ada, sistem akan

menampilkan pesan error dan sistem akan meminta pengguna untuk mulai memasukkan data lengkap.



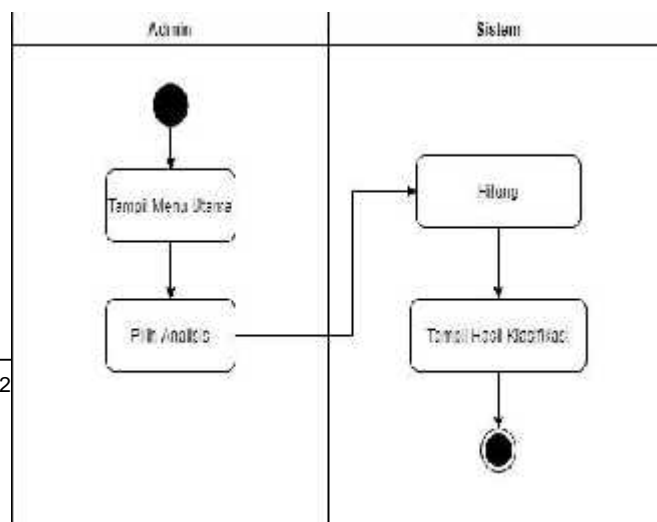
3. Activity Diagram Data

Transaksi/Data latihan

Gambar 4 Activity Diagram Data Transaksi

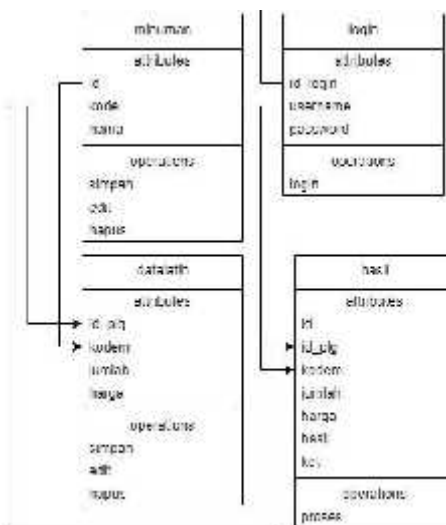
Pada gambar 4 dijelaskan bahwa pengguna memilih menu data transaksi, lalu sistem akan menampilkan form daftar data transaksi. Pengguna dapat menambahkan data transaksi, sehingga sistem akan menyimpan data pada database dan bila ditemukan data yang ada, sistem akan menampilkan pesan error dan sistem akan meminta pengguna untuk mulai memasukkan data lengkap

4. Activity Diagram Data Proses Perhitungan



Gambar 5 Activity Diagram Data Proses Perhitungan

Pada gambar 5 memilih menu data menampilkan form dapat melakukan sistem akan database.



dijelaskan bahwa pengguna proses, lalu sistem akan daftar data nilai. Pengguna proses data nilai, sehingga menyimpan data pada

3.3 Class Diagram

Gambar 6 Class Diagram

Pada gambar 4.2 admin melakukan login terlebih dahulu dengan input password dan username, jika password atau username salah maka sistem akan menampilkan pesan salah kombinasi username dan password. Jika username dan password benar maka akan otomatis masuk pada dashboard dan admin dapat mengelola data alternatif, mengelola data kriteria, mengelola aturan fuzzy, mengelola nilai bobot mengelola hasil maka sistem akan otomatis menampilkan hasil perhitungan.

3.4 Perancangan Database

Berikut perancangan database yang dikumpulkan dalam bentuk penyajian sebagai berikut :

1. Tabel User

Tabel 1 Tabel User

No	Nama Item Data	Type	Lebar	Keterangan
----	----------------	------	-------	------------

1	id_login	int	3	Primary key*
2	username	Varchar	200	-
3	password	Varchar	200	-

2. Tabel Minuman

Tabel 2 Tabel Minuman

No	Nama Item Data	Type	Lebar	Keterangan
1	id	Int	11	Primary Key
2	kode	Varchar	12	-
3	nama	Varchar	50	-

3. Tabel Transaksi/Data Latih

Tabel 3 Tabel Transaksi/ Data latih

No	Nama Item Data	Type	Lebar	Keterangan
1.	id	int	11	Primary Key
2.	id_plg	varchar	15	-
3	kodem	varchar	20	
4	jumlah	varchar	50	
5	harga	varchar	12	
6.	tgl	date		

4. Tabel Hasil

Tabel 4 Tabel Hasil

No	Nama Item Data	Type	Lebar	Keterangan
1.	id	int	11	Primary Key
2.	id_plg	varchar	15	-
3	kodem	varchar	11	
4	jumlah	varchar	200	
5	harga	varchar	50	
6	hasil	varchar	200	
7	ket	text		

3.5 Penerapan KNN

Data yang digunakan sebanyak 10 data. dari 500 data dibagi menjadi data training dan data testing.

1. Menentukan jumlah data training dan data testing dengan k-fold.
2. Menentukan nilai k- (tetangga terdekat). $K=3$.
3. Menghitung jarak setiap data testing satu persatu ke data training menggunakan rumus Euclidean Distance.

Data Testing Pertama

$\sqrt{(4-1)^2 + (60000-15000)^2}$ <i>Jarak = √(</i>	$= 45000$
$\sqrt{(4-3)^2 + (60000-45000)^2}$ <i>Jarak = √(</i>	$= 30000$
$\sqrt{(4-1)^2 + (60000-15000)^2}$ <i>Jarak = √(</i>	$= 45000$
$\sqrt{(4-6)^2 + (60000-90000)^2}$ <i>Jarak = √(</i>	$= 30000$
$\sqrt{(4-2)^2 + (60000-30000)^2}$ <i>Jarak = √(</i>	$= 30000$
$\sqrt{(4-1)^2 + (60000-15000)^2}$ <i>Jarak = √(</i>	$= 45000$
$\sqrt{(4-2)^2 + (60000-30000)^2}$ <i>Jarak = √(</i>	$= 30000$

Data Testing Kedua

$\sqrt{(3-1)^2 + (45000-15000)^2}$ <i>Jarak = √(</i>	$= 30000$
$\sqrt{(3-3)^2 + (45000-45000)^2}$ <i>Jarak = √(</i>	$= 15000$
$\sqrt{(3-1)^2 + (45000-15000)^2}$ <i>Jarak = √(</i>	$= 30000$
$\sqrt{(3-6)^2 + (45000-90000)^2}$ <i>Jarak = √(</i>	$= 45000$
$\sqrt{(3-2)^2 + (45000-30000)^2}$ <i>Jarak = √(</i>	$= 15000$
$\sqrt{(3-1)^2 + (45000-15000)^2}$ <i>Jarak = √(</i>	$= 30000$

$\text{Jarak} = \sqrt{(3-2)^2 + (45000-30000)^2} = 15000$

Data Testing Ketiga

$\sqrt{(5-1)^2 + (75000-15000)^2}$ <i>Jarak = √(</i>	$= 60000$
$\sqrt{(5-3)^2 + (75000-45000)^2}$ <i>Jarak = √(</i>	$= 45000$

$$\begin{array}{l} \text{Jarak} = \sqrt{(5-1)^2 + (75000-15000)^2} = 60000 \\ \text{Jarak} = \sqrt{(5-6)^2 + (75000-90000)^2} = 15000 \\ \text{Jarak} = \sqrt{(5-2)^2 + (75000-30000)^2} = 45000 \\ \text{Jarak} = \sqrt{(5-1)^2 + (75000-15000)^2} = 60000 \\ \text{Jarak} = \sqrt{(5-2)^2 + (75000-30000)^2} = 45000 \end{array}$$

4. Selanjutnya mengurutkan jarak yang telah dihitung secara ascending dari jarak terkecil ke terbesar.

Tabel 5 Hasil Perhitungan Jarak Data Testing dan Training

Jarak	Jarak	Jarak
45000	30000	60000
30000	15000	45000
45000	30000	60000
30000	45000	15000
30000	15000	45000
45000	30000	60000
30000	15000	45000

Tabel 6 Hasil Prediksi

No	Nama Minuman	Hasil
1	green tea	rendah
2	taro	rendah
3	taro	rendah
4	coklat	rendah
5	banana coffe	rendah
6	milo	tinggi
7	lemon tea	tinggi
8	kopi susu	rendah

Berdasarkan dari 500 data training dan satu kali uji perhitungan manual maka disimpulkan bahwa jenis minuman MILO dan LEMON TEA menghasilkan hasil prediksi TINGGI.

3.6 Hasil Pengujian

Untuk menguji apakah aplikasi berjalan sesuai dengan yang diharapkan, maka peneliti melakukan pengujian *black box* yang merupakan salah satu metode pengujian untuk mengecek fungsionalitas khususnya input dan output dari aplikasi tersebut. Berikut tabel rekapitulasi hasil pengujian pada sistem:

Tabel 2 Rekapitulasi Hasil Pengujian

No.	Modular	Kesimpulan
1	Tombol <i>Sign in</i>	Berhasil
2	Tombol <i>Create</i>	Berhasil
3	Tombol Proses	Berhasil
4	Tombol Hasil	Berhasil

3.7 Hasil Penelitian

Setelah melakukan proses pengumpulan data, perancangan system, dan pengujian pada system menggunakan metode black box. Peneliti telah berhasil merancang sistem sesuai dengan permasalahan yang ada pada penelitian ini. Berdasarkan dari 500 data training dan satu kali uji perhitungan manual maka disimpulkan bahwa jenis minuman MILO dan LEMON TEA menghasilkan hasil prediksi TINGGI.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut:

1. Informasi yang didapat dari aplikasi ini berdasarkan data training dari 500 data dengan satu kali uji sistem maka dapat disimpulkan bahwa hasil yang didapat yaitu jenis minuman MILO dan LEMON TEA menghasilkan nilai kategori TINGGI untuk prediksi periode bulan berikutnya berdasarkan masa periode setahun terakhir.
2. Penelitian ini berhasil mengimplementasikan Metode K-Nearest Neighbor menggunakan bahasa pemrograman PHP dan database MySQL dengan klasifikasi pengelompokan data transaksi penjualan berdasarkan kategori prediksi Tinggi, Sedang dan Rendah.
3. Dengan adanya sistem ini, maka dapat membantu pengusaha kecil dalam memprediksi waktu minuman banyak terjual berdasarkan waktu atau musim serta dapat mengklasifikasikan produk minuman yang memiliki potensi akan laris terjual.
4. Uji coba dengan menggunakan metode pengujian *Black Box* pada sistem yang dibuat dapat dikatakan berhasil karena fungsional sistem berjalan dengan baik.

5. SARAN

Saran yang dapat diberikan untuk pengembangan sistem adalah sebagai berikut:

1. Diharapkan hasil klasifikasi dengan menggunakan metode K-Nearest Neighbor dapat dibandingkan dengan menggunakan metode data mining lainnya
2. Metode pengklasifikasian dengan metode K-NN lebih ditingkatkan tingkat akurasi pengklasifikasiannya, salah satunya menggunakan komparasi pengklasifikasian menggunakan tools Matlab atau Python.

DAFTAR PUSTAKA

- [1] F. Sulianta and D. Juju, Data Mining Meramalkan Bisnis Perusahaan, Jakarta: Elex Media

Komputindo, 2010

- [2] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan PT. Gramedia," *Telematika*, pp. 118-132, 2012.
- [4] P. P. Widodo, R. T. Handayanto and H. , Penerapan Data Mining Dengan Matlab, Bandung: REKAYASA SAINS, 2013.
- [5] L. J. Siagian, Otomatisasi Pengujian Perangkat Lunak (Software Tes Automation), Yogyakarta: Deepublish, 2018.
- [7] Metode Penelitian Kuantitatif Kualitatif dan R&D (Sugiyono, 2018;13)
- [8] Susanto, C., Salman, N., & Sunarya, P. A. Data Management Application Research and Community Service Development Center (P4M) Web Based on STMIK Dipanegara Makassar.