

## Prediksi Penjualan Obat Dan Alat Kesehatan Terlaris Menggunakan Algoritma *K-Nearest Neighbor*

Abdul Azis<sup>a</sup>, Ahmad Turmudi Zy<sup>b</sup>, Aswan S Sunge<sup>c</sup>

<sup>abc</sup> Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa  
<sup>bbc</sup> email: aziss305@gmail.com<sup>a</sup>, turmudi@pelitabangsa.ac.id<sup>b</sup>, aswan.sunge@pelitabangsa.ac.id<sup>c</sup>,

Submitted: 25-10-2023, Reviewed: 07-11-2023, Accepted 25-12-2023  
<https://doi.org/10.47233/jteksis.v6i1.1078>

### Abstract

*As-Shofwan Hospital faces problems in the management of drug and medical device supplies, with excess and lack of stock on certain products. Overstock has the potential to lead to wastage of resources, while understock can hinder the patient care process. This research aims to improve the efficiency of inventory management and the availability of health services at the hospital. By identifying and understanding the demand pattern for drugs and medical devices, to avoid the risk of excess or lack of stock, and increase responsiveness to patient needs at As-Shofwan Hospital. In this context, machine learning algorithms, especially the *K-Nearest Neighbors (KNN)* algorithm, can be used to analyze predictions of drugs or medical devices that are most in demand or often prescribed by doctors. The results showed that the *K-Nearest Neighbors (KNN)* algorithm provides accurate predictions with accuracy, recall, and precision values reaching 1.0 in the classification of drugs and medical devices, so that it can be relied on in supporting hospitals in more efficient drug stock management.*

**Keywords:** Analysis, Medicine, Medical Devices, *K-Nearest Neighbor*

### Abstrak

Rumah Sakit As-Shofwan menghadapi permasalahan dalam manajemen persediaan obat dan alat kesehatan, dengan kelebihan dan kekurangan stok pada produk tertentu. Kelebihan stok berpotensi menimbulkan pemborosan sumber daya, sementara kekurangan stok dapat menghambat proses perawatan pasien. Penelitian ini bertujuan untuk meningkatkan efisiensi manajemen persediaan dan ketersediaan pelayanan kesehatan di rumah sakit tersebut. Dengan mengidentifikasi dan memahami pola permintaan obat dan alat kesehatan, untuk menghindari risiko kelebihan atau kekurangan stok, serta meningkatkan responsivitas terhadap kebutuhan pasien di Rumah Sakit As-Shofwan. Dalam konteks inilah algoritma *machine learning*, khususnya algoritma *K-Nearest Neighbors (KNN)* dapat digunakan untuk melakukan Analisa prediksi obat atau alat kesehatan yang paling laris atau sering diresepkan oleh dokter. Hasil penelitian menunjukkan bahwa algoritma *K-Nearest Neighbors (KNN)*, memberikan prediksi yang akurat dengan nilai *accuracy*, *recall*, dan *precision* mencapai 1.0 dalam klasifikasi obat dan alat kesehatan, sehingga dapat diandalkan dalam mendukung rumah sakit dalam manajemen stok obat yang lebih efisien.

**Keywords:** Analisis, Obat, Alat Kesehatan, *K-Nearest Neighbor*

*This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license*



### PENDAHULUAN

Rumah sakit, sebagai pusat pelayanan kesehatan, memegang tanggung jawab besar dalam menyediakan obat dan alat kesehatan yang dibutuhkan untuk perawatan pasien. Dinamika dalam pengadaan, distribusi, dan penggunaan obat dan alat kesehatan menjadi elemen utama dalam menjaga kelancaran proses perawatan dan memastikan pasien mendapatkan layanan kesehatan yang optimal. Oleh karena itu, pemahaman mendalam terkait tren permintaan, distribusi, dan prediksi obat serta alat kesehatan menjadi esensial untuk mengoptimalkan manajemen sumber daya, pengambilan keputusan, dan ketersediaan pelayanan kesehatan yang berkualitas.

Rumah Sakit As-Shofwan, sebagai lembaga kesehatan yang berkomitmen untuk memberikan layanan terbaik, menghadapi tantangan kompleks

dalam manajemen ketersediaan obat dan alat kesehatan. Permasalahan yang muncul mencakup dua sisi yang kontras, yaitu kecenderungan terdapatnya persediaan obat dan alat kesehatan yang sering berlebih dan kekurangan pada produk tertentu. Kondisi berlebihan dapat berpotensi mengakibatkan pemborosan sumber daya, termasuk obat yang memiliki masa kedaluwarsa dan alat kesehatan yang kurang dimanfaatkan. Di sisi lain, kekurangan pada obat tertentu dapat menghambat proses perawatan pasien dan menghadirkan risiko serius terhadap ketersediaan pelayanan kesehatan. Oleh karena itu, penanganan yang efisien terhadap manajemen persediaan obat dan alat kesehatan menjadi krusial untuk memastikan ketersediaan yang tepat, efisien, dan berkualitas dalam upaya memberikan perawatan yang optimal bagi pasien.

Untuk mengatasi permasalahan manajemen persediaan obat dan alat kesehatan yang sering berlebih atau kekurangan pada Rumah Sakit As-Shofwan, salah satu solusi yang dapat diimplementasikan adalah menggunakan pendekatan algoritma K-Nearest Neighbors (KNN). Algoritma KNN dapat digunakan untuk menganalisis pola permintaan berdasarkan data historis penjualan, memberikan prediksi yang akurat terkait produk-produk yang paling diminati oleh dokter spesialis dan paling dibutuhkan oleh pasien.

Algoritma KNN dipilih berdasarkan hasil penelitian sebelumnya yang dilakukan oleh A. Pratiwi dalam konteks analisis penjualan terlaris. Penelitian tersebut membuktikan keefektifan algoritma KNN dalam optimasi produksi, pengaturan stok, dan alokasi sumber daya di bidang bisnis. Meskipun konteksnya berbeda, kerangka kerja analisis yang diperkenalkan oleh A. Pratiwi menjadi dasar bagi penerapan algoritma KNN dalam penelitian ini, dengan fokus khusus pada dinamika permintaan dan distribusi obat serta alat kesehatan di rumah sakit [1].

Pendekatan menggunakan KNN akan melibatkan pengumpulan data historis penjualan obat dan alat kesehatan, termasuk jumlah dan jenisnya. Dengan menggunakan algoritma KNN, model dapat mempertimbangkan karakteristik pasien, diagnosis dokter, serta tren musiman atau perubahan musiman lainnya. Analisis ini akan membantu Rumah Sakit As-Shofwan dalam menentukan kebijakan stok yang lebih tepat, mengidentifikasi produk-produk yang sering diminati, serta memprediksi kebutuhan persediaan obat dan alat kesehatan secara lebih akurat.

Selain itu, pendekatan ini dapat diperkuat dengan penerapan teknik Machine Learning lainnya untuk meningkatkan akurasi prediksi. Integrasi sistem yang mendukung monitoring real-time dari kondisi stok dan permintaan dapat memberikan manfaat tambahan dalam menjaga keseimbangan persediaan. Dengan adanya solusi ini, Rumah Sakit As-Shofwan diharapkan dapat meningkatkan efisiensi manajemen persediaan, mengurangi risiko kelebihan stok dan kekurangan stok, serta memberikan pelayanan kesehatan yang lebih optimal dan responsif terhadap kebutuhan pasien.

Analisis adalah tindakan memisahkan atau menguraikan suatu entitas menjadi elemen-elemen terkecil. Dari konsep tersebut, dapat disimpulkan bahwa analisis adalah proses berpikir yang bertujuan untuk mengurangi atau memecahkan permasalahan dari elemen ke elemen terkecil [6].

Obat adalah substansi yang direncanakan untuk digunakan dalam mendiagnosis, mencegah, mengurangi, menghilangkan, menyembuhkan penyakit atau tanda penyakit, cedera, atau gangguan fisik dan mental pada manusia atau

hewan, serta memperbaiki tubuh atau bagian tubuh. Besarnya efektivitas obat tergantung pada *biosis* dan kepekaan organ tubuh. Setiap orang berbeda kepekaan dan kebutuhan *biosis* obat nya [7].

Alat Kesehatan (Alkes) adalah peralatan, apparatus, dan implan yang tidak mengandung obat yang digunakan untuk mencegah, mendiagnosis, menyembuhkan, dan meringankan penyakit, mengobati orang sakit, memulihkan kesehatan seseorang, dan menciptakan struktur atau meningkatkan fungsi kekebalan tubuh. Ini juga mencakup *reagen in vitro* dan kalibrator, perangkat lunak, zat atau bahan yang digunakan sendiri atau dalam kombinasi untuk mencegah pembuahan, desinfeksi peralatan medis, dan pengujian *in vitro* terhadap spesimen tubuh manusia, serta obat-obatan yang tidak mencapai efeknya dalam tubuh manusia melalui *mekanisme farmakologis, imunologis*, atau metabolik [8].

*K-Nearest Neighbor* (KNN) adalah metode yang sering digunakan dalam klasifikasi data. Teknik ini digunakan untuk mengelompokkan objek berdasarkan data pembelajaran dengan objek tetangga yang memiliki jarak terdekat atau perbedaan nilai yang kecil [22].

## METODE PENELITIAN

### 2.1. Metode Pengumpulan Data

Sumber pengumpulan data yang dilakukan penulis untuk mendapatkan data yang valid adalah sebagai berikut:

#### A. Data Sekunder

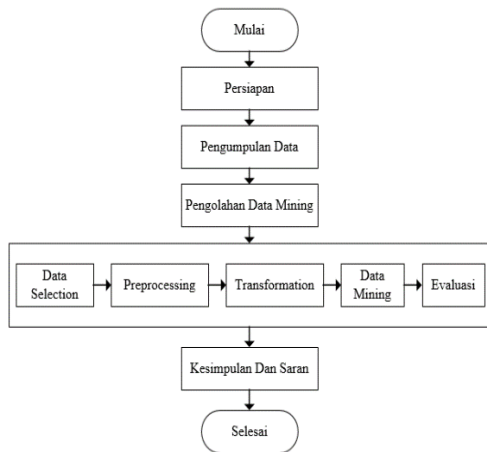
Data Sekunder adalah bentuk data historis yang berisi informasi mengenai variabel-variabel yang telah terkumpul dan diarsipkan sebelumnya oleh entitas lain. Data sekunder berasal dari dalam perusahaan itu sendiri dalam penelitian ini, dimana peneliti menggunakan data sekunder internal yang mencakup data yang dikumpulkan oleh perusahaan secara individu, seperti laporan hasil penjualan untuk tujuan akuntansi, laporan aktivitas pemasaran, dan pengetahuan mengenai pelanggan

#### B. Studi Kepustakaan

Studi kepustakaan adalah teknik pengumpulan data dari suatu penelitian terdahulu untuk memperoleh informasi tentang penelitian sejenis atau yang ada kaitannya dengan penelitian tersebut. Pengumpulan data yang dilakukan pada penelitian ini secara langsung dari sumber-sumber lain

### 2.2 Tahapan Penelitian

Tahapan penelitian adalah menggambarkan proses penelitian yang akan ditempuh sekaligus menggambarkan penelitian secara keseluruhan. Tahapan penelitian yang dilakukan peneliti dapat dilihat pada *flowchart* berikut:



Gambar 1. Tahapan Penelitian

Tahapan penelitian ini dituangkan dalam diagram alir ini. Mencakup tahapan yang akan diambil dalam penelitian secara keseluruhan. Tahapan yang akan dilakukan yaitu:

- A. Persiapan, tahap ini adalah tahapan yang menjadi subjek analisis di Rumah Sakit As-Shofwan, objek yang diambil yaitu sepeda obat dan alat kesehatan.
- B. Pengumpulan data merupakan proses memperoleh informasi atau fakta yang relevan dengan tujuan tertentu. Data yang diperoleh merupakan data primer yaitu data yang dikumpulkan dan diolah sendiri dari Rumah Sakit As-Shofwan Data yang digunakan adalah data penjualan di Bulan Agustus Tahun 2023.
- C. Pengolahan data mining, tahapan ini adalah proses pengelolaan data dimana tahapan yang dilakukan adalah *knowledge discovery in database* (KDD) digunakan untuk mengolah data yang terkumpul.
  1. *Data Selection*, data operasional yang akan digunakan perlu adanya pemilihan sebelum tahapan penggalian informasi dalam KDD dimulai tahapan ini ada pada tahapan data selection.
  2. *Pre-processing Data / Cleaning*, Tahapan dalam pengolahan data yang bertujuan untuk membersihkan, menormalkan, dan mengubah format data agar mudah dipahami dan digunakan untuk analisis.
  3. *Transformation*, tahapan adalah dimana data yang dipilih diubah dengan melakukan *agregasi*, sehingga data tersebut dapat digunakan untuk pengujian data mining.
  4. *Data Mining*, merupakan tahapan dalam proses KDD yang bertujuan untuk mencari pola atau informasi yang berguna dari data yang telah dipilih dan diolah. Metode yang digunakan dalam penelitian ini adalah metode *K-Nearest Neighbor* (KNN).

5. Evaluasi merupakan proses untuk mengevaluasi kualitas model yang telah dibuat berdasarkan data yang telah diproses oleh algoritma data mining.

- D. Kesimpulan Dan Saran, pada tahap ini akan ditarik sebuah kesimpulan dari hasil uji data mining pada obat dan alat Kesehatan sehingga dibuat sebuah kesimpulan barang terlaris, serta saran dari peneliti untuk penelitian selanjutnya.

### 2.3 Pemodelan

Pada penelitian ini, proses pemodelan menggunakan data *mining* dan algoritma *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan objek berdasarkan data pembelajaran yang memiliki jarak terdekat dengan objek yang akan diklasifikasikan. Untuk pemilihan atribut sendiri terdiri dari  $n$  *Neighbors* (biasa disebut  $k$ ). Ketika melakukan pengujian, parameter  $k$  dipilih berdasarkan nilai  $k$  yang paling optimal yang telah diidentifikasi selama tahap pelatihan. Nilai  $k$  yang paling baik ditemukan melalui eksperimen, yang melibatkan perhitungan kuadrat jarak *Euclidean* antara setiap objek dan data sampel yang disediakan.

$$ED = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Berdasarkan rumus diatas dimana matriks ED adalah jarak skala dari kedua vektor  $a$  (data latih) dan  $b$  (data uji) dari Matriks dengan ukuran dimensi. Mengurutkan hasil objek-objek tersebut kedalam kelompok yang mempunyai jarak euclidean terkecil. Mengumpulkan kategori  $y$  (*Klasifikasi Nearest Neighbor* berdasarkan nilai  $k$ ) dengan menggunakan kategori *neighbor* yang paling mayoritas, maka dapat diprediksikan kategori objek tersebut.

## HASIL DAN PEMBAHASAN

Setelah tujuan dan rencana ditetapkan langkah selanjutnya melakukan pengumpulan data awal, deskripsi data, dan eksplorasi. Penelitian ini menggunakan data transaksi penjualan obat dan alat Kesehatan periode bulan Agustus 2023. *Dataset* dengan 383 sampel, memiliki 6 atribut prediktor dan 1 atribut kelas yang digunakan sebagai target. *Dataset* yang digunakan pada penelitian ini dapat dilihat pada gambar

	NO. TRS	NO.FAKTUR	TANGGAL	DISTRIBUTOR	OBAT/ALKES	QTY
0	DO0001230A	FP.23070966	1/8/2023	PT. SAKYO	ANTASIDA TAB	500
1	DO0001230A	FP.23070967	1/8/2023	PT. SAKYO	OBH SYRUP	20
2	DO0001231A	FP.23070970	1/8/2023	PT. SAKYO	AMBROXOL 30MG TABLET	1000
3	DO0001231A	FP.23070971	1/8/2023	PT. SAKYO	RANITIDIN TAB	1000
4	DO0001232A	SI-AUM/VII/2023/1281	1/8/2023	PT. AIRLANGGA	LANCET KLUNING	400
...	...	...	...	...	...	...
378	DO0001391A	INV23293	30-08-2023	PT. TRINITI	CONCENTRATE CLEANER	1
379	DO0001391A	INV23294	30-08-2023	PT. TRINITI	DILUENT	1
380	DO0001391A	INV23295	30-08-2023	PT. TRINITI	REAGEN LYSE 0.5 L	1
381	DO0001392A	000929/FKT/IM/VIII/2023	30-08-2023	PT. INTAN MEDIKA INDONESIA	NATRIUM DIKLOFENAK 25MG	300
382	DO0001392A	000929/FKT/IM/VIII/2024	30-08-2023	PT. INTAN MEDIKA INDONESIA	RAMPRIL 5 MG TAB	200

383 rows x 6 columns

Gambar 2. *Dataset* Penelitian

Pada langkah ini, peneliti melakukan identifikasi dan konstruksi jawaban berdasarkan data yang telah dikumpulkan. Data *preparation* mencakup semua tindakan yang diperlukan untuk membangun *Dataset* yang akan digunakan dalam proses pemodelan dengan algoritma *K-Nearest Neighbor*. Proses persiapan data dalam penelitian ini mencakup pemilihan data, pembersihan data, integrasi data, dan normalisasi data.

1. Normalisasi Data

Di termin ini data yg dipilih merupakan data yang *berafiliasi* menggunakan jumlah penjualan antara lain qty. pada proses ini meliputi *cleaning* dan *selection* data.

2. Kategori Data

Pada tahap ini data ditentukan menjadi 3 kategori yaitu katagori sangat laris, kategori cukup laris, dan kategori kurang laris, penilain kategori dapat dilihat pada tabel 1.

Tabel 1. Nilai Kategori

Jumlah	Kategori
<= 500	Kurang Laris
>500 <= 2000	Cukup Laris
>2000 <=9000	Sangat Laris

3. Pada tahap *preprocessing* data, diperlukan transformasi data menggunakan metode normalisasi, di mana nilai-nilai variabel disesuaikan ke dalam rentang 0-1. Normalisasi angka pada setiap variabel ini merupakan langkah yang penting sebelum menghitung jarak oleh algoritma *K-Nearest Neighbor*. Tujuannya adalah untuk menghindari dominasi parameter tertentu dalam perhitungan jarak antar data, sambil tetap mempertahankan hasil yang akurat..

3.1 Proses Pengujian Data

Dalam proses pengujian, penulis memanfaatkan aplikasi *Jupyter Notebook* untuk melakukan evaluasi data. Sejalan dengan itu, dalam penelitian ini, proses pengkajian data mengikuti langkah-langkah dalam *Knowledge Discovery In Database* (KDD) dengan tujuan menghasilkan informasi sesuai dengan urutan yang telah ditentukan, berikut adalah langkah-langkah dalam proses ini.

1. Data Seleksi, Pada tahap ini dilakukan pemilihan data awal yang akan digunakan untuk analisis prediksi. Pemilihan data awal sangat penting untuk memastikan kesesuaian relevansi data dengan tujuan penelitian. Dalam penelitian ini menggunakan data yang dipilih adalah data yang berhubungan dengan jumlah penjualan antara lain qty. Berikut adalah proses *selection* data yang dilakukan penulis.

```
# Membuat kamus untuk menyimpan jumlah QTY untuk setiap kode barang (DATA SELEKSI)
kode_qty_map = {}

# Menghitung total QTY untuk setiap kode barang dari semua kolom
kode_columns = [col for col in data.columns if col.startswith('OBAT/ALKES')]
qty_columns = [col for col in data.columns if col.startswith('QTY')]

for i in range(len(kode_columns)):
    for kode_barang, qty in zip(data[kode_columns[i]], data[qty_columns[i]]):
        if kode_barang in kode_qty_map:
            kode_qty_map[kode_barang] += qty
        else:
            kode_qty_map[kode_barang] = qty

# Menampilkan hasil rangkuman dalam bentuk tabel
table_data = []
for kode_barang, qty in kode_qty_map.items():
    table_data.append([kode_barang, qty])

headers = ['OBAT/ALKES', 'Jumlah (QTY)']

# Mengubah data menjadi DataFrame
df_summary = pd.DataFrame(table_data, columns=headers)

# Menampilkan DataFrame
df_summary
```

Gambar 3. Proses Seleksi *Dataset*

menghitung dan menyajikan ringkasan jumlah (QTY) untuk setiap kode barang pada *Dataset*, langkah-langkahnya dimulai dengan membuat kamus kosong (*kode\_qty\_map*) untuk menyimpan informasi jumlah (QTY) setiap nama obat atau alkes.

	OBAT/ALKES	Jumlah (QTY)
0	ANTASIDA TAB	2500
1	OBH SYRUP	20
2	AMBROXOL 30MG TABLET	1800
3	RANITIDIN TAB	1000
4	LANCET KUNING	800
...	...	...
260	PREABOR TABLET	60
261	ACETYLCYSTEINE TAB (PT NULAB)	500
262	CONCENTRATE CLEANER	1
263	NATRIUM DIKLOFENAK 25MG	300
264	RAMIPRIL 5 MG TAB	200

265 rows x 2 columns

Gambar 4. Hasil Proses Seleksi

2. Data Transformasi, proses ini dilakukan untuk mengubah data menjadi bentuk lebih sesuai atau lebih mudah diolah dalam analisis atau pemodelan.

```
# Menggunakan LabelEncoder untuk mengubah kolom 'OBAT/ALKES' menjadi data numerik (DATA TRANSFORMASI)
label_encoder = LabelEncoder()
summary['OBAT/ALKES'] = label_encoder.fit_transform(summary['OBAT/ALKES'])

# Menggunakan .loc[row_indexer, col_indexer] untuk menetapkan nilai pada DataFrame
summary.loc[:, 'TARGET'] = summary['Jumlah (QTY)'].apply(lambda x: 'No' if x < 1000 else 'Yes')

# Tambahkan kolom 'NORMALIZED' yang berisi hasil normalisasi ke DataFrame 'summary'
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
normalized_data = scaler.fit_transform(summary[['Jumlah (QTY)']])
summary.loc[:, 'NORMALIZED'] = normalized_data
```

Gambar 5. Proses Transformasi Data

Transformasi data atribut “jumlah” dan target yang bedasarkan kurang dari 1000 menjadi ‘No’ dan menjadi ‘Yes’ apabila jumlah lebih dari 1000.

	OBAT/ALKES	Jumlah (QTY)	TARGET	NORMALIZED
0	15	2500	Yes	2.489333
1	153	20	No	-0.430044
2	11	1800	Yes	1.665315
3	188	1000	Yes	0.723581
4	124	800	No	0.488147
...	...	...	...	...
257	213	40	No	-0.406500
258	177	60	No	-0.382957
259	7	500	No	0.134997
260	145	300	No	-0.100437
261	186	200	No	-0.218153

262 rows x 4 columns

Gambar 6. Hasil Transformasi

3. Pembagian Data, membagi data menjadi dua bagian yaitu data latih (*Training Data*) dan data uji (*Test Data*). Tujuan dari pembagian data adalah untuk melatih model uji prediksi produk terlaris.

```
#PEMBAGIAN DATA
imputer = SimpleImputer(strategy='mean')
X = summary.drop(columns=['TARGET'])
X = imputer.fit_transform(X)
y = summary['TARGET']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

X_train_df = pd.DataFrame(X_train)

print("Data Latih (X_train):")
print(tabulate(X_train_df, tablefmt="grid"))

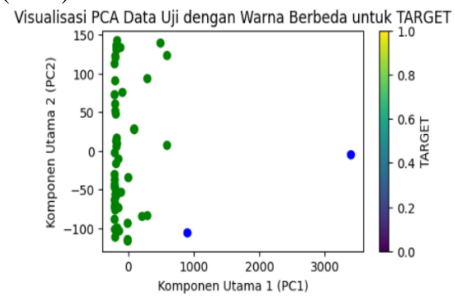
y_train_table = pd.DataFrame(y_train, columns=['TARGET'])
print("Label Target Data Latih (y_train):")
print(tabulate(y_train_table))
```

Gambar 7. Proses Pembagian Data

proses dalam pembagian data yang menghasilkan data training dan data testing yang ditampilkan pada gambar dilakukan pembagian data sebesar 80% data latih dan data sebesar 20% data uji. Dimana mendapatkan hasil jumlah data X\_train 209 data, X\_test 53 data, y\_train 209 data dan y\_test 53 data.

4. Hasil Dan Evaluasi, *Principal Component Analysis* (PCA) yang bertujuan untuk memvisualisasikan data hasil prediksi dalam bentuk yang lebih mudah dipahami dan divisualisasikan dalam ruang dua dimensi. *Principal Component Analysis* (PCA) digunakan untuk mengurangi dimensi data asli yang mungkin memiliki banyak fitur atau atribut menjadi dua komponen utama yang paling mengandung variasi dalam data. Dan dapat dijelaskan gambar diatas untuk *Principal Component 1* adalah jumlah (Qty) sedangkan *Principal Component 2* yaitu mengacu pada rentang nilai pada sumbu y yang digunakan untuk memvisualisasikan data setelah dilakukan reduksi dimensi

dengan *Principal Component Analysis* (PCA).



Gambar 8. Visualisasi Prediksi Dengan PCA

### 3.2 Perhitungan *K-Nearest Neighbor*

Berdasarkan dari tahapan data mining untuk algoritma *K-Nearest Neighbor*, adapun langkah-langkahnya sebagai berikut:

1. Penentuan nilai K, penentuan nilai K yang digunakan tidak memiliki aturan yang baku, namun pada penelitian ini nilai k yang digunakan adalah 3.

```
#PEMODELAN ALGORITMA KNN
knn = KNeighborsClassifier(n_neighbors=3)
knn.fit(X_train, y_train)

KNeighborsClassifier
KNeighborsClassifier(n_neighbors=3)
```

Gambar 9. Proses Pemodelan KNN

Setelah model KNN dilatih, langkah terakhir adalah menggunakan model tersebut untuk membuat prediksi pada data uji yang diberikan, yang ditempatkan dalam X\_test. Dalam konteks penelitian atau skripsi, ini berarti Anda akan menggunakan model KNN yang telah dilatih untuk melakukan klasifikasi atau prediksi pada data baru yang belum pernah dilihat sebelumnya. Hasil prediksi ini disimpan dalam variabel y\_pred.

2. Menentukan Barang Kategori Sangat laris dengan nilai kode barang lebih dari 2000.

```
top_10_barang = df.summary(-df.summary
['OBAT/ALKES'].isin(excluded_items)).sort_values(by='Jumlah (QTY)', ascending=False).head(10)

# Tambahkan kolom 'Category' berdasarkan Jumlah (QTY)
top_10_barang['Category'] = 'Kurang Laris'
top_10_barang.loc[top_10_barang['Jumlah (QTY)'] >= 2000, 'Category'] = 'Sangat Laris'
top_10_barang.loc[(top_10_barang['Jumlah (QTY)'] >= 900) &
(top_10_barang['Jumlah (QTY)'] < 2000), 'Category'] = 'Cukup Laris'

# Pilih kolom yang diperlukan untuk tabel hasil
result_table = top_10_barang[['OBAT/ALKES', 'Category', 'Jumlah (QTY)']]

# Tabel untuk kategori Cukup Laris
cukup_laris_data = df.summary[(df.summary['Jumlah (QTY)'] >= 900) &
(df.summary['Jumlah (QTY)'] < 2000)].sort_values(by='Jumlah (QTY)', ascending=False)
cukup_laris_data['Category'] = 'Cukup Laris'
cukup_laris_data[['OBAT/ALKES', 'Jumlah (QTY)', 'Category']]

# Hitung jumlah barang yang cukup Laris
jumlah_barang_cukup_laris = cukup_laris_data.shape[0]

# Set opsi untuk menampilkan tabel tanpa penomoran
pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
pd.set_option('display.width', None)

# Tampilkan tabel untuk kategori Cukup Laris tanpa penomoran
print("Tabel OBAT/ALKES Cukup Laris")
print(cukup_laris_data.to_string(index=False))

# Tampilkan jumlah barang yang cukup Laris
jumlah_barang_cukup_laris = cukup_laris_data.shape[0]
```

Gambar 10. Proses Menentukan Kategori Sangat Laris

- Menentukan barang dengan kategori cukup laris dengan nilai kode barang lebih dari 500-1999.

```
top_10_barang = df_summary[~df_summary
['OBAT/ALKES'].isin(excluded_items)].sort_values(by='Jumlah (QTY)', ascending=False).head(10)

# Tambahkan kolom 'Category' berdasarkan Jumlah (QTY)
top_10_barang['Category'] = 'Kurang Laris'
top_10_barang.loc[top_10_barang['Jumlah (QTY)'] >= 2000, 'Category'] = 'Sangat Laris'
top_10_barang.loc[(top_10_barang['Jumlah (QTY)'] >= 500 &
(top_10_barang['Jumlah (QTY)'] < 2000), 'Category'] = 'Cukup Laris'

# Pilih kolom yang diperlukan untuk tabel Hasil
result_table = top_10_barang[['OBAT/ALKES', 'Category', 'Jumlah (QTY)']]

# Tabel untuk kategori Cukup Laris
cukup_laris_data = df_summary[(df_summary['Jumlah (QTY)'] >= 500 &
(df_summary['Jumlah (QTY)'] < 2000)].sort_values(by='Jumlah (QTY)', ascending=False)
cukup_laris_data['Category'] = 'Cukup Laris'
cukup_laris_data = cukup_laris_data[['OBAT/ALKES', 'Jumlah (QTY)', 'Category']]

# Hitung jumlah barang yang cukup Laris
jumlah_barang_cukup_laris = cukup_laris_data.shape[0]

# Set opsi untuk menampilkan tabel tanpa pemotongan
pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
pd.set_option('display.width', None)

# Tampilkan tabel untuk kategori Cukup Laris tanpa pemotongan
print("Tabel OBAT/ALKES Cukup Laris")
print(cukup_laris_data.to_string(index=False))

# Tampilkan jumlah barang yang cukup Laris
print("Jumlah OBAT/ALKES Cukup Laris:", jumlah_barang_cukup_laris)
```

Gambar 11. Proses Menentukan Kategori Cukup Laris

- Menentukan barang dengan kategori kurang laris dengan nilai kode barang kurang dari 500.

```
['OBAT/ALKES'].isin(excluded_items)].sort_values(by='Jumlah (QTY)', ascending=False).head(10)

# Tambahkan kolom 'Category' berdasarkan Jumlah (QTY)
top_10_barang['Category'] = 'Kurang Laris'
top_10_barang.loc[top_10_barang['Jumlah (QTY)'] >= 2000, 'Category'] = 'Sangat Laris'
top_10_barang.loc[(top_10_barang['Jumlah (QTY)'] >= 500 &
(top_10_barang['Jumlah (QTY)'] < 2000), 'Category'] = 'Cukup Laris'

# Pilih kolom yang diperlukan untuk tabel Hasil
result_table = top_10_barang[['OBAT/ALKES', 'Category', 'Jumlah (QTY)']]

# Tabel untuk kategori Sangat Laris
sangat_laris_data = df_summary[(df_summary['Jumlah (QTY)'] >= 2000)].sort_values(by='Jumlah (QTY)', ascending=False)
sangat_laris_data['Category'] = 'Sangat Laris'
sangat_laris_data = sangat_laris_data[['OBAT/ALKES', 'Category', 'Jumlah (QTY)']]

# Tabel untuk kategori Cukup Laris
cukup_laris_data = df_summary[(df_summary['Jumlah (QTY)'] >= 500) &
(df_summary['Jumlah (QTY)'] < 2000)].sort_values(by='Jumlah (QTY)', ascending=False)
cukup_laris_data['Category'] = 'Cukup Laris'
cukup_laris_data = cukup_laris_data[['OBAT/ALKES', 'Category', 'Jumlah (QTY)']]

# Tabel untuk kategori Tidak Laris
tidak_laris_data = df_summary[(df_summary['Jumlah (QTY)'] < 500)].sort_values(by='Jumlah (QTY)', ascending=False)
tidak_laris_data['Category'] = 'Kurang Laris'
tidak_laris_data = tidak_laris_data[['OBAT/ALKES', 'Jumlah (QTY)', 'Category']]

# Set opsi untuk menampilkan tabel tanpa pemotongan
pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
pd.set_option('display.width', None)

# Tampilkan tabel untuk kategori Tidak Laris tanpa pemotongan
print("Tabel OBAT/ALKES Tidak Laris")
print(tidak_laris_data.to_string(index=False))
```

Gambar 12. Proses Menentukan Kategori Kurang Laris

### 3.3 Perhitungan Matrik Evaluasi

*Confusion Matrix* merupakan hasil dari klasifikasi yang berupa tabel berdasarkan jumlah data data uji yang benar dan salah. Berikut adalah hasil dari penggunaan *tools Jupyter Notebook* dan hitungan manual

```
#PENGEHITUNGAN Matrik EVALUASI
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='weighted')
recall = recall_score(y_test, y_pred, average='weighted')
f1 = f1_score(y_test, y_pred, average='weighted')
conf_matrix = confusion_matrix(y_test, y_pred)

# Mengubah nilai akurasi, presisi, dan recall menjadi persentase
accuracy_percent = accuracy * 100
precision_percent = precision * 100
recall_percent = recall * 100

# Menampilkan hasil evaluasi
print("Accuracy: {:.2f}%".format(accuracy_percent))
print("Precision: {:.2f}%".format(precision_percent))
print("Recall: {:.2f}%".format(recall_percent))
print("Confusion Matrix:")
print(conf_matrix)

Accuracy: 100.00%
Precision: 100.00%
Recall: 100.00%
Confusion Matrix:
[[ 51  0]
 [ 0  2]]
```

Gambar 13. Proses Perhitungan Matrix Evaluasi

Proses Metrik Evaluasi menggambarkan hasil dari perhitungan menggunakan *tools Jupyter*

*Notebook* dengan menyatakan bahwa *accuracy*, *precision*, dan *recall* menampilkan hasil 1.0 apabila diubah menjadi persen menjadi 100 %.

Dilihat dari hasil *Confusion Matrix* pada pengujian menggunakan *jupyter notebook* dapat ditunjukkan seperti table.

Table 2. Confusion Matrix

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi i	TRUE	TP	FN
	FALSE	FN	TN
		51	0
		0	2

Nilai *Accuracy* dihitung dengan cara membagi yang akan terjadi penjumlahan data sah yang bernilai (*true positive*) serta data akbar yang bernilai negatif (*true negative*) dibagi menggunakan hasil penjumlahan data benar yang bernilai (*true positive*) dan benar yang bernilai negatif (*true negative*) data galat yang bernilai positif (*false positive*) serta data galat yang bernilai negatif (*false negative*).

### 3.3 Analisa Hasil

Berdasarkan hasil yang telah dilakukan bahwa variabel atau atribut yang digunakan dalam penelitian ini (jumlah) berpengaruh signifikan terhadap penelitian ini terbukti dengan menggunakan algoritma *K-Nearest Neighbor* mampu menghasilkan prediksi yang baik dengan nilai *Accuracy* 1.0, *Recall* 1.0 dan *Precision* 1.0 hal ini dikarenakan kedekatan jarak antara nilai yang satu dengan nilai yang lain dalam suatu variabel atau atribut. Proses pengujian ini dilakukan untuk mengetahui obat dan alat kesehatan terlaris berdasarkan data historis penjualan dalam satu tahun.

Sehingga didapatkan hasil prediksi barang dengan tiga kategori yaitu barang Sangat Laris, Cukup Laris, Kurang Laris.

- Obat Dan Alat Kesehatan Sangat Laris

Table 3. Kategori Sangat Laris

Obat Dan Alkes	Jumlah	Kategori
Klip Obat Besar 15 X 10 cm	9000	Sangat Laris
Klip Obat Uk. 13x8,7	4000	Sangat Laris
Alkohol Swab	4000	Sangat Laris
Kertas Puyer	3600	Sangat Laris
Omeprazole 20mg Nulab	3100	Sangat Laris
Curcuma Sanbe Tablet	2900	Sangat Laris
Ranitidine Inj	2600	Sangat Laris
Antasida Tab	2500	Sangat Laris

Cefixime 200mg Tab Nulab	2500	Sangat Laris
Betahistine 6mg Tab	2160	Sangat Laris
Ondan 4mg Inj Interbat	2100	Sangat Laris
Handsocon Latex Uk.M	2000	Sangat Laris
Handsocon Latex Uk.L	2000	Sangat Laris

## 2. Obat Dan Alat Kesehatan Cukup Laris

Table 4. Kategori Cukup Laris

Obat Dan Alkes	Jumlah	Kategori
Ceftriaxone Inj	1900	Cukup Laris
Ambroxol 30 mg Tablet	1800	Cukup Laris
Amlodipin 5 Mg Tab Nulab	1620	Cukup Laris
...	...	...
Simvastatin 20mg	500	Cukup Laris
Microtube	500	Cukup Laris
Acetylcysteine Tab (Pt Nulab)	500	Cukup Laris

## 3. Obat Dan Alat Kesehatan Kurang Laris

Table 5. Kategori Kurang Laris

Obat Dan Alkes	Jumlah	Kategori
Otsu Nacl 100ml	480	Kurang Laris
L-Bio Sachet	450	Kurang Laris
Nozepav 5 mg Supp	10	Kurang Laris
...	...	...
Diazepam 5 Mg Suppo	7	Kurang Laris
Kassa Roll X-Ray	4	Kurang Laris
Film Fujifilm Uk.35x35	4	Kurang Laris
Ett Non Cuff No 4	3	Kurang Laris
Ett Non Cuff No 5,5	3	Kurang Laris
Film Fujifilm Uk.24x30	2	Kurang Laris
Oil Imersi 50ml	1	Kurang Laris
Concentrate Cleaner	1	Kurang Laris
Reagen Triglicerida	1	Kurang Laris

## SIMPULAN

Berdasarkan hasil analisa dan pengujian terhadap data transaksi penjualan produk yang diambil satu bulan terakhir (Periode bulan Agustus 2023), untuk menentukan prediksi obat dan alat kesehatan terlaris menggunakan metode *K-Nearest Neighbor* maka dapat disimpulkan metode *K-Nearest Neighbor* dapat digunakan untuk prediksi obat dan alat kesehatan terlaris dengan menggunakan *variable* Obat/Alkes dan

QTY Berdasarkan hasil perhitungan data *mining* menggunakan teknik klasifikasi dan *K-Nearest Neighbor*. Didapatkan hasil prediksi penjualan Obat/Alkes terlaris di Rumah Sakit As-Shofwan adalah Klip Obat Besar 15 X 10 Cm, Klip Obat Uk. 13X87, Alkohol Swab, Kertas Puyer, Omeprazole 20Mg Nulab, Curcuma Sanbe Tablet, Ranitidine Inj, Antasida Tab, Cefixime 200Mg Tab Nulab, Betahistine 6Mg Tab, Ondan 4Mg Inj Interbat, Handsocon Latex Uk.M, Handsocon Latex Uk.L dan *tools Jupyter Notebook* dapat digunakan untuk melakukan prediksi obat dan alat kesehatan terlaris dengan penerapan pengujian menggunakan metode *K-Nearest Neighbor*. Berdasarkan hasil dari pengolahan *Dataset*, nilai akurasi, *recall* dan *precision* terhadap klasifikasi penjualan obat dan alat kesehatan terlaris menggunakan *K-Nearest Neighbor* sebesar *accuracy: 1.0, recall: 1.0* dan *precision: 1.0*.

Pada penelitian selanjutnya penulis memberikan saran untuk membandingkan metode algoritma lain dan menggunakan data penjualan yang melibatkan rentang waktu lebih dari satu bulan, bertujuan untuk meningkatkan akurasi prediksi dengan memperoleh wawasan yang lebih mendalam tentang tren dan pola popularitas obat dan alat kesehatan.

## DAFTAR PUSTAKA

- [1] A. Pratiwi, "( PREDICTIVE ANALYSIS OF BEST-SELLING PLASTIC MILLS USING THE K-NEAREST NEIGHBOR Disusun oleh :," 2023.
- [2] Z. Zhai, H. Jiang, L. Lu, and Y. Liu, "Adaptive truncation coding for computed tomography images," *Proc. 2014 Int. Symp. Inf. Technol. ISIT 2014*, pp. 115–118, 2015, doi: 10.1201/b18776-23.
- [3] r a Manullang and f a Sianturi, "Penerapan Algoritma K-Nearest Neighbor Untuk Memprediksi Kelulusan Mahasiswa," *J. Ilmu Komput. Dan ...*, vol. 4, no. 2, pp. 15–23, 2021, [Online]. Available: <http://ejournal.sisfokomtek.org/index.php/jikom/article/view/155>
- [4] R. Rismala, I. Ali, and A. Rizki Rinaldi, "Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Sepeda Motor Terlaris," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 585–590, 2023, doi: 10.36040/jati.v7i1.6419.
- [5] A. A. Putri, "Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor ( Studi Kasus : PT . Central Brastagi Utama )," vol. 1, no. 6, pp. 354–361, 2021.
- [6] Y. Septiani, E. Aribbe, and R. Diansyah, "ANALISIS KUALITAS LAYANAN SISTEM INFORMASI AKADEMIK UNIVERSITAS ABDURRAB TERHADAP KEPUASAN PENGGUNA MENGGUNAKAN METODE SEVQUAL (Studi Kasus : Mahasiswa Universitas Abdurrab Pekanbaru)," *J. Teknol. Dan Open Source*, vol. 3, no. 1, pp. 131–143, 2020, doi: 10.36378/jtos.v3i1.560.

- [7] A. Bisri and M. Man, "Machine Learning Algorithms Based on Sampling Techniques for Raisin Grains Classification," *Int. J. Informatics Vis.*, vol. 7, no. 1, pp. 7–14, 2023, doi: 10.30630/joiv.7.1.970.
- [8] Y. Dani and M. A. Ginting, "Classification of Predicting Customer Ad Clicks Using Logistic Regression and k-Nearest Neighbors," *Int. J. Informatics Vis.*, vol. 7, no. 1, pp. 98–104, 2023, doi: 10.30630/joiv.7.1.1017.
- [9] A. Lailiyah, "Tingkat Pengetahuan Masyarakat Tentang Penggunaan Dan Penyimpanan Obat Bebas Dan Bebas Terbatas Di Apotek Sambeng Farma," *Skripsi. Univ. Muhammadiyah Gresik*, pp. 3–16, 2019.
- [10] Y. Sainika, S. Astiti, D. J. A. Kusuma, and A. W. Muhammad, "Analysis of E-learning readiness level of public and private universities in central Java, Indonesia," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 7, no. 1, pp. 15–30, 2021, doi: 10.26594/register.v7i1.2042.
- [11] H. S. Amalia, U. Athiyah, and A. W. Muhammad, "The Application of Modified K-Nearest Neighbor Algorithm for Classification of Groundwater Quality Based on Image Processing and pH, TDS, and Temperature Sensors," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 9, no. 1, pp. 42–54, 2023, doi: 10.26594/register.v9i1.2827.
- [12] Monita, "Mengenal Klasifikasi, Jenis dan Regulasi Terkait Pengadaan Alat Kesehatan Dalam Rangka Melakukan Pengawasan dan Mencegah Adanya Fraud/Kecurangan di Masa Pandemi Covid 19," pp. 1–15, 2020.
- [13] M. Rizq Daffa Jodi, "Fakultas Komputer Algoritma dan Struktur data," *Fak. Kompiter*, vol. 1, pp. 1–10, 2020.
- [14] V. No, O. Hal, S. Mendrofa, and J. I. Sihotang, "Analisis Tingkat Kematangan Learning Management System SMAN 1 Parongpong Menggunakan Framework COBIT 5 Domain DSS03 &," vol. 5, no. 4, pp. 492–502, 2023.
- [15] M. Afdhal, V. Ariandi, and R. Rita, "Memprediksi Penjualan Pada Toko Hanifah Metode C.45," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 4, no. 2, pp. 248–255, 2022, doi: 10.47233/jteksis.v4i1.460.
- [16] O. S. Udang, M. Tabaru, E. A. M. Sampetoding, and E. S. Manapa, "Pengolahan Data Siswa SMA Negeri 1 Sambuara Kabupaten Kepulauan Talaud Pada Aplikasi DAPODIK," *J. Dyn. Sainst*, vol. 6, no. 1, pp. 7–11, 2021, doi: 10.47178/dynamicsaint.v6i1.1193.
- [17] M. Safii and A. Trydillah, "Implementasi Data Mining Dalam Menentukan Pola Pembelian Obat Dengan Metode Algoritma Apriori," *METHOMIKA J. Manaj.* ..., vol. 3, no. 1, pp. 66–71, 2019, [Online]. Available: <https://ejournal.methodist.ac.id/index.php/methomika/article/view/178%0Ahttps://ejournal.methodist.ac.id/index.php/methomika/article/download/178/151>
- [18] R. Handayani, R. F. Runtuwene, and S. A. P. Sambul, "Pengaruh Startegi Pemasaran Terhadap Peningkatan Penjualan Produk Ikan Kaleng Isabella pada PT.Sinar Purefoods Internasional Bitung," *J. Adm. Bisnis*, vol. 6, no. 2, pp. 34–40, 2018.
- [19] A. A. Khaleel, A. N. Kareem, and L. H. Mahdi, "Predictive analytics on COVID-19 data using Hive based on Hadoop cluster," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 31, no. 2, pp. 945–956, 2023, doi: 10.11591/ijeecs.v31.i2.pp945-956.
- [20] S. Syofian and A. Nugraha, "Prediksi Sistem Stok Barang Toko Elektronik Abc Dengan Algoritma Apriori Dan Metode Moving Average," *J. Sains dan Teknol.*, vol. 11, no. 1, pp. 27–32, 2021.
- [21] V. No, O. Hal, Y. Bahtera, E. Uliyanti, and P. Br, "Analisis Perbandingan Metode Marker dan Markerless Angka 0-9 3D Pada Teknologi Augmented Reality," vol. 5, no. 4, pp. 454–459, 2023.
- [22] 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.
- [23] A. T. Zy, A. S. Sunge, R. Riani, and E. Widodo, "Analisis Sentimen Terhadap Masyarakat Indonesia Di Masa PPKM Menggunakan Algoritma Naïve Bayes," *J. SIGMA*, vol. 13, no. 2, pp. 87–94, 2022.
- [24] L. Ardiani, H. Sujaini, and T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 2, p. 183, 2020, doi: 10.26418/justin.v8i2.36776.
- [25] Karsito and S. Susanti, "Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah Dengan Algoritma Naïve Bayes Di Perumahan Azzura Residencia," *J. Teknol. Pelita Bangsa*, vol. 9, pp. 43–48, 2019.
- [26] Sunge, Aswan Supriyadi, and Ahmad Turmudi Zy. "ANALISIS PREDIKSI PENJUALAN DENGAN METODE REGRESI LINEAR DI PT. EAGLE INDUSTRY INDONESIA." *Jurnal Informatika Teknologi dan Sains (Jinteks)* 5.3 (2023): 398-403.