

**SISTEM KONTROL BARANG
DENGAN PENERAPAN *DATA MINING*
(STUDI KASUS: APOTEK UNIVERSITAS AHMAD DAHLAN)**

SKRIPSI

**Disusun Untuk Memenuhi Sebagian Persyaratan
Mencapai Derajat Sarjana**



Disusun Oleh:

**IRFAN AFANDI
1600018085**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN
YOGYAKARTA
2021**

HALAMAN PERSETUJUAN

SKRIPSI

SISTEM KONTROL BARANG
DENGAN PENERAPAN DATA MINING
(STUDI KASUS: APOTEK UNIVERSITAS AHMAD DAHLAN)

Dipersembahkan dan disusun oleh:

Irfan Afandi
160018085

Program Studi Teknik Informatika
Fakultas Tenolog Industri
Universitas Ahmad Dahlan
Yogyakarta

Telah disetujui oleh:
Pembimbing



Sri Winiarti, S.T., M.Cs.

HALAMAN PENGESAHAN

SKRIPSI SISTEM KONTROL BARANG DENGAN PENERAPAN DATA MINING (STUDI KASUS: APOTEK UNIVERSITAS AHMAD DAHLAN)

yang dipersembahkan dan disusun oleh:

Irfan Afandi
1600018085

telah dipertahankan di depan Dewan Penguji pada tanggal 22 Januari 2021
dan dinyatakan telah memenuhi syarat

Susunan Dewan Penguji :

Ketua : Sri Winiarti, S.T., M.Cs.
Penguji 1 : Anna Hendri Soleliza Jones, S.Kom., M.Cs.
Penguji 2 : Dewi Soyusiawaty, S.T., M.T.

acc 05/02/2021

Yogyakarta, 17 Februari 2021
Dekan
Fakultas Teknologi Industri
Universitas Ahmad Dahlan



Sunardi, S.T., M.T., Ph.D.
NIY. 60010313

PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Irfan Afandi
NIM : 1600018085
Email : irfan1600018085@webmail.uad.ac.id
Fakultas : Teknologi Industri
Program Studi : Teknik Informatika
Judul TA/Skripsi : Sistem kontrol Barang Dengan Penerapan *Data Mining* (Studi Kasus: Apotek Universitas Ahmad Dahlan)

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Ahmad Dahlan maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Yogyakarta, 22 Januari 2021



Irfan Afandi

PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : Irfan Afandi

NIM : 1600018085

Email :irfan1600018085@webmail.uad.ac.id

Fakultas : Teknologi Industri

Program Studi : Teknik Informatika

Judul TA/Skripsi: Sistem kontrol Barang Dengan Penerapan *Data Mining* (Studi Kasus: Apotek Universitas Ahmad Dahlan)

Dengan ini saya menyerahkan hak sepenuhnya kepada Perpustakaan Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tugas akhir elektronik sebagai berikut:

☒ Saya mengizinkan karya tersebut diunggah ke dalam aplikasi Repository Perpustakaan Universitas Ahmad Dahlan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Yogyakarta, 22 Januari 2021



Irfan Afandi

Mengetahui,

Dosen Pembimbing



Sri Winiarti, S.T., M.Cs.

HALAMAN PERSEMBAHAN



Dengan Rahmat Allah Yang Maha Pengasih lagi Maha Penyayang atas nikmat kekuatan, kemudahan, kelancaran yang telah Engkau berikan. Shalawat serta salam senantiasa saya haturkan kepada junjungan kita, Nabi Muhammad ﷺ atas teladannya. Dengan ini saya persembahkan karya ini untuk:

1. Kedua orang tua saya Bapak Mat Harlim dan Ibu Siti Muyasaroh yang senantiasa memberikan do'a, nasehat, dukungan, semangat, inspirasi dan segalanya. Terimakasih telah mendidik saya dengan sangat baik, memberikan saya kepercayaan, serta pengorbanan yang mengiringi langkah saya dari sejak saya kecil, terimakasih untuk segalanya.
2. Kakak kandungku Evi Rahmawati yang selalu memberi do'a, dukungan, semangat serta nasehat kepada saya.
3. Saudaraku Syarifudin, Debi, Andrie, Asrori Terimakasih selama ini sudah menyemangati dalam menempuh perkuliahan sampai selesai.
4. Sahabat seperjuangan Arfiansyah, Luthfiantoro, Hermawan, Ian, Oei, Alvin, Febrian, Habibila, Akmal, Aprizal, Ardiansyah, Gema, Erizal terimakasih untuk solidaritas yang luar biasa dari selama awal hingga akhir perkuliahan.
5. Arfiansyah terimakasih telah banyak membagikan ilmunya, memberikan banyak solusi dan motivasi dalam menyelesaikan skripsi ini.
6. Sahabat-sahabatku Dicky, Veri, Alif, Rahayu, Rika sahabat kecilku hingga sama-sama kita berjuang menempuh pendidikan di SMK. Terimakasih atas dukungan, candaan, dan kebersamaannya yang memberikan warna tersendiri.
7. Teknik Informatika Angkatan 2016. Khususnya kelas B yang telah memberikan banyak pengajaran selama kuliah.

MOTTO



Artinya:

Dan Katakanla “Karena sesungguhnya sesudah kesulitan itu ada kemudahan(5) Sesungguhnya sesudah kesulitan itu ada kemudahan (6)”
(Q.S At-Insyirah: 5-6)

“Tiap kali kita mencari dan mengejar sesuatu yang tinggi dan mulia, pasti kita akan menemui banyak kesulitan dan rintangan. Jangan kita berpikir bahwa mencapai derajat yang tinggi dan mulia adalah pekerjaan yang mudah dan tidak memerlukan usaha keras.”

Al-Ghazali, Ihya Ulumuddin

“All’s well that carries on well”

Amit Abraham

KATA PENGANTAR

Assalamu'alaikum wr.wb

Dengan mengucap Alhamdulillahirobbil'alamin, segala puji syukur kehadiran Allah Yang Maha Esa atas limpahan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul **"SISTEM KONTROL BARANG DENGAN PENERAPAN DATA MINING (STUDI KASUS: APOTEK UNIVERSITAS AHMAD DAHLAN)"**. Skripsi ini disusun guna memenuhi persyaratan menyelesaikan derajat Sarjana di Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Ahmad Dahlan Yogyakarta.

Dalam penyusunan skripsi ini tentunya penulis menyadari bahwa banyak pihak yang telah memberikan bantuannya. Oleh karena itu, penulis mengucapkan terimakasih kepada:

1. Bapak Dr. Muchlas, M.T. selaku Rektor Universitas Ahmad Dahlan.
2. Bapak Sunardi, S.T., M.T., Ph.D. selaku Dekan Fakultas Teknologi Industri Universitas Ahmad Dahlan.
3. Ibu Nur Rochmah Dyah Pujiastuti, S.T., M.Kom. selaku Kepala Program Studi Teknik Informatika.
4. Ibu Ika Arfiani, S.T., M.Cs. selaku dosen pembimbing akademik yang telah memberikan arahan selama masa perkuliahan.
5. Ibu Sri Winiarti, S.T., M.Cs. selaku dosen pembimbing skripsi atas bimbingan dan arahan yang telah diberikan sehingga peneliti dapat menyelesaikan penelitian ini.
6. Ibu Anna Hendri Soleliza Jones, S.Kom., M.Cs. selaku dosen penguji I yang telah menyetujui, menerima dan memberikan saran serta kritik pada skripsi ini.
7. Ibu Dewi Soyusiawaty, S.T., M.T. selaku dosen penguji II yang telah menyetujui, menerima dan memberikan saran serta kritik pada skripsi ini.
8. Segenap dosen Teknik Informatika Universitas Ahmad Dahlan, yang telah membagikan ilmunya sehingga skripsi ini dapat selesai.
9. Semua pihak yang tidak bisa disebutkan satu persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, kritik serta saran yang membangun selalu penulis harapkan demi penelitian yang lebih baik kedepannya.

Wassalamu'alaikum wr.wb.

Yogyakarta, 22 Januari 2021



Penulis

DAFTAR ISI

HALAMAN JUDUL

.....

i

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN TIDAK PLAGIAT	iv
PERNYATAAN PERSETUJUAN AKSES	iv
HALAMAN PERSEMBAHAN	vi
MOTTO	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
DAFTAR LISTING	xv
ABSTRAK.....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Identifikasi Masalah.....	4
1.3. Batasan Masalah	4
1.4. Rumusan Masalah	5
1.5. Tujuan Penelitian	5
1.6. Manfaat Penelitian	6
BAB II LANDASAN TEORI	7
2.1 Kajian Penelitian Terdahulu.....	7
2.2 Landasan Teori.....	14
2.2.1 Pengertian Data Mining.....	14

	2.2.2 Fungsi Data Mining	14
	2.2.3 Proses Data Mining	15
	2.2.4 Tahap - Tahap Data Mining	18
	2.2.5 Klasifikasi	20
	2.2.6 Algoritma K-Nearest Neighbor	22
	2.2.7 Confusion Matrix	27
	2.2.8 Kontrol Barang	29
BAB III	METODE PENELITIAN	32
	3.1. Objek Penelitian	32
	3.2. Metode Pengumpulan Data	33
	3.3. Alat dan Bahan penelitian	34
	3.4. Tahapan Penelitian	35
	3.5. Tahapan Alur Sistem	36
	3.6. Processing	38
	3.7. Implementasi	39
	3.8. K- Nearest Neighbour (KNN)	39
	3.9. Pengujian	40
BAB IV	HASIL DAN PEMBAHASAN	41
	4.1 Hasil Pengumpulan Data	41
	4.2 Analisis Kebutuhan Sistem	41
	4.3 Processing	42
	4.4 Implementasi	50
	4.5 K- Nearest Neighbour	61
	4.6 Pengujian	64
BAB V	KESIMPULAN DAN SARAN	66
	5.1 Kesimpulan	66
	5.2 Saran	66
	DAFTAR PUSTAKA	67

DAFTAR GAMBAR

Gambar 2.1	Tahapan <i>Data Mining</i>	18
Gambar 3.1	Tahapan Penelitian.....	35
Gambar 3.2	Tahapan alur sistem.....	37
Gambar 3.3	flowchart <i>processing</i>	39
Gambar 4.1	tampilan Hasil <i>upload</i> data	51
Gambar 4.2	Tampilan Data <i>Cleaning</i>	52
Gambar 4.3	Tampilan seebelum <i>Selection</i> Data	53
Gambar 4.4	Tampilan sesudah <i>Selection</i> Data	53
Gambar 4.5	sebelum <i>Transformasi</i> Data	54
Gambar 4.6	sesudah <i>Transformasi</i> Data	55
Gambar 4.7 Hasil	56
Gambar 4.8	Hasil Data <i>Training</i>	57
Gambar 4.9	Hasil Data <i>Testing</i>	59
Gambar 4.	10 <i>Save</i> data	60

Gambar 4.11	Data	
	Training	
	62	
Gambar 4.12	Data	
	Testing	
	62	
Gambar 4.13		
Sorting	
	
	63	
Gambar 4.14	Tetangga terdekat dan label	
	mayoritas	
	63	
Gambar 4.15	Hasil	
	KNN	
	64	

DAFTAR TABEL

Tabel 2.1	Jurnal Penelitian Terdahulu.....	9
Tabel 2.2	data <i>training</i>	23
Tabel 2.3	perhitungan jarak dengan euclidean distance.....	24
Tabel 2.4	pengurutan jarak terdekat data baru dengan data training.....	25
Tabel 2.5	penentuan kategori yang termasuk K=3	26
Tabel 2.6	penentuan kategori yang termasuk K=3	26
Tabel 2.7	hasil klasifikasi berdasarkan kategori mayoritas	27
Tabel 2.8	Confusion Matrix.....	28
Tabel 2.9	Data obat.....	30
Tabel 3.1	Hasil Keterangan Data.....	32
Tabel 4.1	Data awal obat.....	42
Tabel 4.2	Data <i>Cleaning</i>	43
Tabel 4.3	<i>Selection</i> data stok barang.....	44

Tabel 4.4	Hasil <i>Transform</i>	45
Tabel 4.5	Perhitungan Data <i>Testing</i> Pertama.....	46
Tabel 4.6	Ranking data	49
Tabel 4.7	Hasil Data Testing Terbaik	49
Tabel 4.8	Hasil Data Testing terbaik dengan Keterangan.....	49
Tabel 4.9	Hasil Klasifikasi Data yang dicari	50
Tabel 4.10	Hasil pengujian akurasi.....	65

DAFTAR LISTING

Listing 4.1 <i>upload</i> data	51
Listing 4.2 tampilan <i>cleaning</i>	53
Listing 4.3 tampilan <i>selection</i>	54
Listing 4.4 tampilan <i>Transformasi</i> data	55
Listing 4.5 tampilan Hasil <i>KNN</i>	57
Listing 4.6 tampilan data <i>training</i>	58

SISTEM KONTROL BARANG DENGAN PENERAPAN DATA MINING (STUDI KASUS: APOTEK UNIVERSITAS AHMAD DAHLAN)

**Irfan Afandi
160018085**

ABSTRAK

Kontrol barang merupakan salah satu aktivitas yang selalu menjadi perhatian penting pada perusahaan karena transaksi penjualan yang dilakukan sangat berhubungan dengan stok ketersediaan barang. Ketersediaan stok pada perusahaan kesehatan di Apotek Universitas Ahmad Dahlan (UAD) pernah terjadi kehabisan stok yang mengakibatkan kehilangan konsumen dan omset karena pengontrolan stok barang masih dilakukan secara langsung pada gudang apotek dan hasil disimpan file *excel*. Penelitian ini membuat aplikasi sistem kontrol barang menggunakan konsep *data mining* menggunakan metode *K-Nearest Neighbour* (KNN) untuk menghasilkan rekomendasi sebagai acuan pada penyetoran pada tahun berikutnya.

Penelitian ini menggunakan metode KNN dengan dilakukan tahapan- tahapan *data mining* berupa *Data Cleaning, integration, Selection, Transformation, Knowledge Discovery (Data mining)*, *Pattern Evaluation*, dan *Knowledge Presentation* untuk menghasilkan klasifikasi berupa *order*, *Tidak order*, dan *order* bulan depan pada barang pada apotek dengan data tahun 2016. Pengujian sistem dengan perhitungan *euclidean distance* dan pengujian keakurasian data memakai *confusion matrix*.

Pengujian sistem ini pada kontrol barang menggunakan data tahun 2016 sebanyak 700 dari 1000 data yang telah dilakukan tahapan *data mining* dengan *data testing* 210 mendapatkan akurasi 95% dan *error rate* 5% sehingga disimpulkan metode KNN mendapatkan hasil yang bagus dalam penelitian sistem kontrol barang pada Apotek UAD.

Kata Kunci: *Data Mining*, Sistem Kontrol Barang, Stok, *K-Nearest Neighbour*

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Kesehatan merupakan hak asasi manusia dan sekaligus merupakan investasi untuk keberhasilan pembangunan Bangsa Indonesia. Oleh karena itu perlu dilakukan pembangunan kesehatan secara menyeluruh dan berkesinambungan, dengan tujuan guna meningkatkan kesadaran, kemauan dan kemampuan hidup sehat bagi setiap orang agar terwujud derajat kesehatan masyarakat yang setinggi-tingginya, salah satu caranya adalah dengan meningkatkan akses dan mutu pelayanan kesehatan.

Salah satu unsur penting dalam memelihara dan meningkatkan mutu pelayanan kesehatan adalah ketersediaan obat yang memadai, dimana diantara berbagai alternatif yang ada, intervensi dengan obat merupakan intervensi yang paling banyak digunakan dalam penyelenggaraan upaya kesehatan dan biaya yang digunakan untuk obat merupakan bagian yang cukup besar dari seluruh biaya kesehatan.

Perencanaan kebutuhan obat merupakan salah satu aspek penting dan menentukan dalam pengelolaan obat khususnya pada Apotek Universitas Ahmad Dahlan. Di Apotek Universitas Ahmad Dahlan pada pendataan stok barang masih dengan melakukan pengecekan secara langsung pada gudang dalam merekomendasikan suatu ordernya barang atau tidak dengan

menyimpan data pada excel. Oleh karena itu pernah terjadi kehabisan stok barang yang akan berdampak kehilangan konsumen dan omset pada apotek. Oleh karena itu perlu di cari sebagai patokan berupa data transaksi dari tahun 2016 sebagai acuan persediaan barang, karena perencanaan kebutuhan obat akan mempengaruhi pengadaan, pendistribusian dan pemakaian obat di unit pelayanan kesehatan. Dimana dengan perencanaan kebutuhan obat yang tepat, berdampak membuat pengadaan menjadi efektif dan efisien sehingga tersedia obat dengan jenis dan jumlah yang cukup sesuai dengan kebutuhan pelayanan kesehatan dengan mutu yang terjamin serta dapat diperoleh pada saat yang diperlukan.

Data Mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan (Daniel, 2005). Data Mining memiliki tahap proses dari Seleksi data, Pre-Processing/cleaning Data, Transformasi hingga melakukan evaluasi pada dataset kelayakan untuk pengujian Data Mining, dengan menggunakan Data mining dapat membantu pemrosesan data dalam skala besar, dan dapat mewujudkan Hasil Penelitian yang akan dicari dalam sistem kontrol barang.

Penelitian terkait penerapan data mining dalam penyelesaian kasus kontrol persediaan barang sudah banyak dilakukan. (Simbolon, 2019), Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori oleh Priska. sistem pengecekan persediaan barang pada algoritma apriori dengan

mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam *database*. *Support* (nilai penunjang), adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi.

Penelitian yang dilakukan oleh (Kafil et al., 2019), Penerapan Metode *K-Nearest Neighbors* untuk prediksi penjualan berbasis *web* pada *boutiq dealove Bondowoso*. Menggunakan metode *K-NN* untuk prediksi penjualan pada *Butik Dealove* dengan masa periode setiap bulan dari Bulan Januari 2016 – September 2018. Menggunakan sistem peramalan berbasis *Web PHP Native* agar lebih tepat dan akurat dalam melakukan perhitungan

Penelitian yang dilakukan oleh (Hasmawati et al., 2019), Menggunakan Algoritma *K-Nearest Neighbor* digunakan karena memiliki akurasi yang tinggi dengan rasio kesalahan kecil, dan memprediksi atau meramalkan penjualan barang di masa mendatang berdasarkan data yang telah direkam sebelumnya.

Penelitian yang dilakukan oleh (Winiarti et al., 2018), Menggunakan Algoritma *K-means* digunakan karena untuk mengoptimalkan data transaksi pasien balita gizi buruk. Penderita gizi buruk dikelompokkan menurut nilai gizi balita menggunakan metode data mining dengan algoritma *k-means clustering*.

Berdasarkan penelitian yang sudah dipaparkan sebelumnya, dengan Metode *Algoritma Apriori, K-Means, Naive bayes, K-Nearest Neighbor* yang menjadi pembedaan pada penelitian ini ialah akan menggunakan

menggunakan Metode *K-Nearest Neighbor* (KNN) dengan memaparkan hasil sistem kontrol barang dengan penerapan data mining dengan mengambil data laporan transaksi tahun 2016 dibelakang dan mencari produk yang paling laku terjual yang akan direkomendasikan barang (obat) untuk melakukan order dan sebagai acuan untuk tahun berikutnya agar mencukupi dalam penyediaan stok barang di Apotek Universitas Ahmad Dahlan.

1.2. Identifikasi Masalah

Berdasarkan uraian dari latar belakang masalah, maka didapatkan identifikasi masalah yaitu :

1. Apotek UNIVERSITAS AHMAD DAHLAN (UAD) sistem dalam pendataan stok barang masih dengan melakukan pengecekan secara langsung pada gudang dalam merekomendasikan suatu *order*nya barang atau tidak dengan menyimpan data pada excel.
2. Pada APOTEK UNIVERSITAS AHMAD DAHLAN (UAD) pernah terjadi kehabisan stok barang, berdampak kehilangan konsumen dan omset pada APOTEK UAD.

1.3. Batasan Masalah

Batasan masalah yang diangkat didapat berdasarkan identifikasi masalah yang dibatasi pada permasalahan:

1. Pada Penelitian ini, Metode data mining yang digunakan adalah *k-nearest neighbor* (K-NN).

2. Data yang digunakan dalam penelitian ini adalah data *kontrol* barang pada tahun 2016 yang di ambil dari pihak APOTEK Universitas Ahmad Dahlan (UAD).
3. Hasil pengelompokkan pada sistem kontrol barang dengan data obat di tahun 2016 yang akan dituju ada 3 yaitu *order*, tidak *order* dan *order* bulan depan

1.4. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, terdapat beberapa permasalahan yang dapat dirumuskan sebagai berikut :

1. Bagaimana membuat aplikasi sistem kontrol barang pada suatu apotek dengan menggunakan konsep data mining.
2. Bagaimana mengetahui penerapan algoritma KNN dapat berfungsi untuk mengklasifikasi order barang di apotek.

1.5. Tujuan Penelitian

Melalui penelitian ini, maka tujuan yang ingin dicapai oleh penulis adalah :

1. membuat aplikasi sistem kontrol barang pada suatu apotek dengan menggunakan konsep data mining.
2. mengetahui penerapan algoritma KNN dapat berfungsi untuk mengklasifikasi order barang di apotek.

1.6. Manfaat Penelitian

Manfaat yang ingin diperoleh dari penelitian ini adalah dapat membantu pihak Apotek UAD dalam *kontrol* barang yang dapat menghasilkan rekomendasi *order* barang pada catatan transaksi sehingga pengguna sistem dapat memiliki acuan untuk tahun berikutnya dalam penyetoran persediaan barang.

BAB II

LANDASAN TEORI

2.1 Kajian Penelitian Terdahulu

Pada bagian ini akan dijelaskan kajian terdahulu sebagai acuan dalam penelitian. Selain kajian terdahulu, akan dijelaskan tentang teori yang mendukung penelitian ini.

Penelitian yang dilakukan oleh (Simbolon, 2019), Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori oleh Priska. sistem pengecekan persediaan barang pada algoritma apriori dengan mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam *database*. *Support* (nilai penunjang), adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi.

Penelitian yang dilakukan oleh (Nurdiawan et al., 2018), Metode *Naive Bayes* Dalam Memprediksi Stok Barang. Sistem Prediksi Stok barang pada Metode *Naive Bayes* dengan cara melakukan pengolahan data penjualan dari produk Aspira untuk Wilayah Sumatera Barat. Data yang dikumpulkan terhitung dari bulan januari 2011 sampai dengan bulan desember 2015 sebagai bahan dalam pengujian metode yang digunakan dalam memprediksi stok barang.

Penelitian yang dilakukan oleh (Kafil et al., 2019), Penerapan Metode *K-Nearest Neighbors* untuk prediksi penjualan berbasis *web* pada *boutiq dealove Bondowoso*. Menggunakan metode K-NN untuk prediksi penjualan pada Butik Dealove dengan masa periode setiap bulan dari Bulan Januari 2016 – September 2018. Menggunakan sistem peramalan berbasis Web PHP Native agar lebih tepat dan akurat dalam melakukan perhitungan

Penelitian yang dilakukan oleh (Winiarti et al., 2018), Menggunakan Algoritma K-means digunakan karena untuk mengoptimalkan data transaksi pasien balita gizi buruk. Penderita gizi buruk dikelompokkan menurut nilai gizi balita menggunakan metode data mining dengan algoritma k-means clustering.

Berdasarkan Penelitian yang sudah ada, maka dilakukan pengembangan penelitian selanjutnya yang disajikan pada tabel 2.1.

Tabel 2.1 Jurnal Penelitian Terdahulu

NAMA PENULIS	Priska Hartinah Simbolon (2019)	Nurdiawan (2018)	Mohammad Kafil(2019)	Sri Winiarti, Heman Yuliansyah, Aprial Andu Purnama (2018)	Irfan Afandi (2019)
NAMA JURNAL	Data Mining Pada Sistem Persediaan Barang Menggunakan <i>Algoritma Apriori</i>	Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode <i>Naive Bayes Classifier</i> untuk Optimasi Strategi Pemasaran	penerapan Metode <i>K-Nearest Neighbors</i> untuk prediksi penjualan berbasis <i>web</i> pada boutiq dealove Bondowoso	Identifikasi Status Gizi Balita Menggunakan Pendekatan Data Mining	
VOLUME, NOMOR	Vol. 6, No.4	-	Vol. 3 No. 2	Vol.9, No.1	
BULAN, TAHUN	Agustus, 2019	April, 2018	septemberi, 2019	2018	
PENERBIT	Riset Komputer (JURIKOM)Studi Teknik Informatika, STMIK Budi Darma	IlJurnal Teknologi Informasi dan Komunikasi	JATI (Jurnal Mahasiswa Teknik Informatika)	Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia	
MASALAH PENELITIAN	Persediaan barang yang tidak dilakukan secara optimal akan menimbulkan kekosongan salah satu barang yang tersedia.	Kurang optimalnya perusahaan dalam mengolah data penjualan. Belum adanya analisa data yang dapat membantu dalam mengoptimasi pemasaran pada cv tersebut. Penerapan metode yang	sulitnya mendata banyak barang yang keluar setiap hari, perhitungan stok barang yang tidak akurat, serta proses penjumlahan harga yang sering terjadi	Pemerintah Indonesia melalui Puskesmas (Puskesmas) telah melakukan pendataan status gizi balita dengan menggunakan aplikasi berbasis Excel. Namun, hasil pengelompokan data status gizi tidak dapat ditampilkan secara otomatis. Data itu	Belum Memiliki sistem dari pihan Apotek dalam melakukan sistem kontrol barang dan melihat tren barang yang sering dibeli pada sistem, karena hal itu pernah terjadi kehabisan stok barang, yang akan berdampakkehilangan konsumen

		dapat menganalisa data	kesalahan, peramalan penjualan yang kurang akurat karena dikerjakan secara manual karena belum adanya system pendukung untuk mempermudah peramalan penjualan	tersedia di PUSKESMAS masih belum dapat menentukan status gizi balita, sesuai standar yang ditetapkan pemerintah Indonesia. Ketika ada permintaan data terkait status gizi masyarakat, maka proses pemetaan dilakukan secara manual. Proses ini menjadi tidak maksimal karena membutuhkan proses yang lama dan dapat terjadi duplikasi data jika ribuan data yang ada diolah secara manual.	dan omset pada apotek.
TUJUAN PENELITIAN	Tujuan dari penelitian ini adalah mengetahui prosedur pendataan persediaan barang pada toko Srikandi Cash Credit Electronic dan Furniture saat ini, menerapkan algoritma apriori pada persediaan barang di	Mengetahui seberapa besar dampak penerapan metode Naive Bayes Classifier dalam upaya optimasi strategi pemasaran. Membuat pengetahuan baru dari hasil analisa menggunakan metode Naive Bayes Classifier. Menggunakan	untuk prediksi penjualan pada Butik Dealove dengan masa periode setiap bulan dari Bulan Januari 2016 – September 2018. Menggunakan sistem peramalan berbasis Web PHP Native agar lebih tepat dan akurat dalam melakukan perhitungan	Penelitian ini bertujuan untuk mengoptimalkan data transaksi pasien balita gizi buruk. Penderita gizi buruk dikelompokkan menurut nilai gizi balita menggunakan metode data mining dengan algoritma k-means clustering.	membuat aplikasi sistem kontrol barang pada suatu apotek dengan menggunakan konsep data mining dan mengetahui penerapan algoritma KNN dapat berfungsi untuk mengklasifikasi order barang di apotek.

	Srikandi Cash Credit Electronic dan Furniture, dan menguji penerapan algoritma apriori pada persediaan barang menggunakan software tanagra.	hasil analisa untuk optimasi strategi pemasaran			
METODE/TEORI YANG DIGUNAKAN	Menggunakan Algoritma apriori, tahapan Analisa pola frekuensi tinggi Pada tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database, tahapan Pembentukan aturan asosiasi Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum	penerapan metode Naive Bayes Classifier pada data penjualan dan transaksi-transaksi penjualan untuk menciptakan pengetahuan baru yang kemudian digunakan untuk optimasi strategi pemasaran produk agar bisa tercapai dengan maksimal	metode K-NN untuk prediksi penjualan pada Butik Dealove dengan masa periode setiap bulan dari Bulan Januari 2016 – September 2018. Menggunakan sistem peramalan berbasis Web PHP Native agar lebih tepat dan akurat dalam melakukan perhitungan. Salah satu metode yang dapat digunakan adalah Forecasting. Forecasting yaitu kegiatan memperkirakan atau memprediksikan apa yang akan	Teknik pengelompokan menggunakan k-means clustering. Algoritma pengelompokan k-means adalah algoritme paling sederhana dan paling umum yang digunakan untuk mengelompokkan objek berdasarkan atribut / fitur ke dalam jumlah k cluster, di mana k adalah bilangan bulat positif dan ditentukan oleh pengguna. Pengelompokan dilakukan dengan meminimalkan jumlah kuadrat jarak antara data dan centroid cluster yang sesuai.	Menggunakan Metode K-Nearest Neighbor Menemukan Parameter K (jumlah tetangga paling dekat). Menghitung kuadrat jarak eucliden objek terhadap data training yang diberikan. Mengurutkan hasil No 2 secara ascending (berurutan dari nilai tinggi ke rendah). Mengumpulkan Kategori Y(Klasifikasi Nearest Neighbor berdasarkan nilai K). Dengan menggunakan kategori Nearest Neighbor yang paling mayoritas maka dapat

	untuk confidence dengan menghitung confidence aturan asosiatif A ke B, dan pengambilan data dengan aturan asosiatif (Association rule) untuk menentukan hubungan asosiatif suatu kombinasi item.		terjadi pada masa yang akan datang dengan kurun waktu tertentu		diprediksikan kategori objek.
HASIL PENELITIAN	didapatkan rule persediaan barang yang paling tinggi adalah M. Rias kerang dan Kt.kartini dengan support 50% dan confidence 85%. Kemudian itemsets selanjutnya yakni Sb.wisdom holy black dan M. Rias kerang dengan support 41% dan	Hasil akurasi model naive bayes menunjukkan tingkat akurasinya 97.22%, artinya model klasifikasi kelulusan menggunakan naïve bayes terbukti baik hal ini dilihat dari tingkat akurasinya yang mencapai 97.22% akan tetapi hal ini perlu di tinjau ulang dari sudut pandang kompleksitas dan jumlah datasetnya.	Hasil pengujian keakuratan metode menggunakan menggunakan 12 data traning dan 12 data testing diperoleh hasil nilai akurasi 83,3% dan nilai errorsebesar 16,7 %	Penelitian ini membangun perangkat lunak yang dapat digunakan untuk mengidentifikasi status gizi balita dengan menggunakan teknik data mining, dengan algoritma pengelompokan k-means. Tes dilakukan oleh melakukan validasi silang dan memberikan 90% validasi bahwa Sistem dapat menentukan status gizi balita dengan memproduksi 5 klaster yaitu, gizi baik, gizi sedang,	hasil penelitian yang telah dilakukan dengan <i>data testing</i> 210 mendapatkan hasil akurasi 95% dan <i>error rate</i> 5%, maka dapat disimpulkan penelitian dengan metode <i>K-Nearest Neighbour (KNN)</i> mendapatkan hasil yang bagus dalam penelitian sistem kontrol barang pada Apotek Universitas Ahmad Dahlan.

	<p>confidence 83%, Sb.wisdom holy black dan kt.kartini dengan support 41% dan confidence 83%, dan Kt.kartini, Sb.wisdom holy black dan M. Rias kerang dengan support 33% dan confidence 80%.</p> <p>5</p>			malnutrisi, lebih banyak gizi dan obesitas.	
--	---	--	--	---	--

2.2 Landasan Teori

2.2.1 Pengertian Data Mining

Data Mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan (Daniel, 2005).

Dan menurut (Vercellis, 2009), Data mining adalah aktivitas yang menggambarkan sebuah proses analisis yang terjadi secara iteratif pada database yang besar, dengan tujuan mengekstrak informasi dan knowledge yang akurat dan berpotensi berguna untuk knowledge workers yang berhubungan dengan pengambilan keputusan dan pemecahan masalah.

(Begg, 2010), Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting.

Data Mining (Witten, 2011) didefinisikan sebagai proses penemuan pola dalam data.

2.2.2 Fungsi Data Mining

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Menurut (Larose, 2005) data mining mempunyai empat fungsi dasar yaitu:

1) Fungsi Prediksi (prediction)

Proses untuk menemukan pola dari data dengan menggunakan beberapa variabel untuk memprediksi variabel lain yang tidak diketahui jenis atau nilainya.

2) Fungsi Deskripsi (description).

Proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data.

3) Fungsi Klasifikasi (classification).

Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan class atau konsep dari suatu data.

Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.

4) Fungsi Asosiasi (association).

Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data.

2.2.3 Proses Data Mining

Proses yang umumnya dilakukan oleh data mining antara lain: deskripsi, prediksi, estimasi, *klasifikasi*, *clustering* dan asosiasi. Secara rinci proses data mining dijelaskan sebagai berikut (Larose, 2005):

a. Deskripsi

Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi

aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada *domain* aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (*knowledge*) pada sistem. Tugas deskriptif merupakan tugas data mining yang sering dibutuhkan pada teknik *postprocessing* untuk melakukan *validasi* dan menjelaskan hasil dari proses data *mining*. *Postprocessing* merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.

b. Prediksi

Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga saham dalam tiga bulan yang akan datang.

c. Estimasi

Estimasi hampir sama dengan prediksi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan

darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi.

d. *Klasifikasi*

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. *Klasifikasi* melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

e. *Clustering*

Clustering merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah *kluster* adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam *kluster* lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu *cluster* dan semakin besar perbedaan tiap cluster maka kualitas analisis *cluster* semakin baik.

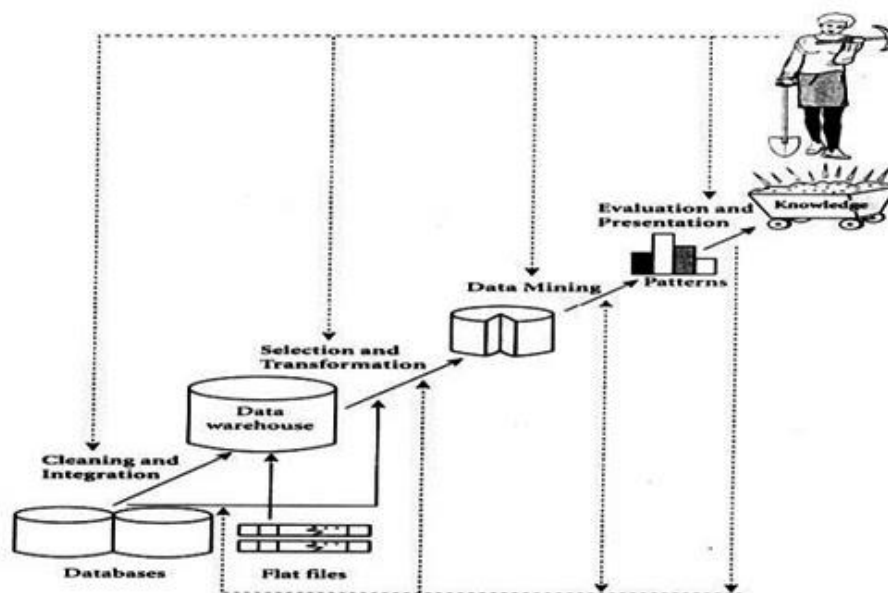
f. *Asosiasi*

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut

analisis keranjang belanja (market basket analysis). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

2.2.4 Tahap - Tahap Data Mining

(Han & Kamber, 2012) menjelaskan tahapan *data mining* melalui skema seperti pada Gambar 2.1.



Gambar 2.1 Tahapan *Data Mining*

Menurut (Han & Kamber, 2012) *knowledge Discovery in Database* (KDD) terdiri atas 7 tahap yaitu :

a. *Data Cleaning*

Pada tahap ini, melakukan kegiatan membuang data yang tidak konsisten dan bersifat *noise* dari data yang terdapat diberbagai basis data yang mungkin berbeda format maupun *platform*. Pada proses cleaning mencakup antara lain membuang duplikasi data atau data

ganda, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data.

b. *Data integration*

Menyatukan sumber data yang terbesar atau terpecah sehingga dikumpulkan menjadi data *warehouse*.

c. *Data Selection*

Data yang ada dalam *warehouse* kemudian direduksi untuk mendapatkan hasil yang akurat, yaitu data yang relevan dengan tugas analisa dikembalikan kedalam *database*.

d. *Data Transformation*

Data berubah atau bersatu menjadi bentuk yang paling tepat untuk *mining* dengan ringkasan performa atau operasi. Metode *transformasi* pada *data mining* adalah:

- 1) *Centering* mengurangi setiap data dengan rata-rata dari setiap atribut yang ada.
- 2) *Normalization* membagi setiap data yang di-*center-ing* dengan standar deviasi dari atribut bersangkutan.
- 3) *Scaling* mengubah data sehingga berada dalam skala tertentu.

e. *Knowledge Discovery (Data Mining)*

Tahapan ini merupakan proses esensial dimana metode intelegent (algoritma data mining) digunakan untuk mengekstrak pola data. Data yang telah ditransformasi, kemudian ditambang dengan

berbagai teknik. Proses *data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan fungsi-fungsi tertentu. Fungsi atau algoritma dalam *data mining* sangat bervariasi, dimana pemilihannya bergantung pada tujuan dan proses pencarian pengetahuan secara menyeluruh.

f. *Pattern Evaluation*

Tahap ini untuk mengidentifikasi pola yang benar-benar menarik (*interestingness measures*) yang mewakili pengetahuan berdasarkan atas beberapa tindakan yang menarik. Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

g. *Knowledge Presentation*

Tahap ini merupakan gambaran teknik *visualisasi* dan pengetahuan yang digunakan untuk memberikan pengetahuan yang telah ditambang kepada *user*. Pada tahap ini, dipresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna atau pihak yang berkepentingan.

2.2.5 Klasifikasi

Klasifikasi adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang

bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (*Han & Kamber, 2012*).

Beberapa metode yang dapat digunakan untuk menyelesaikan suatu permasalahan klasifikasi dengan metode *Backpropagation neural network*, *support vector classification (SVC)*, *extreme learning machine (ELM)*, *K- Nearest Neighbour (KNN)*, *Naïve Bayes*.

Pengukuran terhadap kinerja suatu sistem klasifikasi merupakan hal yang penting. Kinerja sistem klasifikasi menggambarkan seberapa baik sistem dalam mengklasifikasikan data. Confusion matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. (E. Prasetyo, 2012)

Klasifikasi data terdiri dari 2 langkah proses. Pertama adalah *learning* (fase training), dimana algoritma klasifikasi dibuat untuk menganalisa data training lalu direpresentasikan dalam bentuk *rule* klasifikasi. Proses kedua adalah klasifikasi, dimana data tes digunakan untuk memperkirakan akurasi dari rule klasifikasi (*Han & Kamber, 2012*). Hasil klasifikasi pada data obat ditahun 2016 dipenelitian ini berupa order, tidak order, direktur. Proses klasifikasi didasarkan pada empat komponen (*Han & Kamber, 2012*).

a. Kelas

Variabel dependen yang berupa kategorikal yang merepresentasikan 'label' yang terdapat pada objek.

b. *Predictor*

Variabel independen yang direpresentasikan oleh karakteristik (atribut) data.

c. *Training dataset*

Satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.

d. *Testing dataset*

Berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi dievaluasi.

2.2.6 Algoritma K-Nearest Neighbor

k-nearest neighbor (KNN) termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. KNN dilakukan dengan mencari kelompok k objek dalam data *training* yang paling dekat (mirip) dengan objek pada data baru atau data *testing*.

Algoritma KNN memiliki beberapa tahapan yaitu :

- a. Menemukan Parameter K (jumlah tetangga paling dekat).
- b. Menghitung kuadrat jarak *eucliden* objek terhadap data *training* yang diberikan.

- c. Mengurutkan hasil No 2 secara *ascending* (berurutan dari nilai tinggi ke rendah).
- d. Mengumpulkan Kategori Y (Klasifikasi Nearest Neighbor berdasarkan nilai K).
- e. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksikan kategori objek.

Teknik dalam mencari data training pada penelitian ini menggunakan Rumus *Euclidean Distance*.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \dots\dots\dots(2.1)$$

Keterangan :

d : jarak kedekatan

x : data *training*

y : data *testing*

A. Contoh kasus *K-Neares Neighbor*

Diberikan data *Training* dua atribut Tidak *Order* dan *Order* untuk mengklasiikasikan sebuah data apakah tergolong Tidak *Order* atau *Order*, mencari data baru yang akan diklasifikasikan, yaitu X = 15 dan Y = 6. Jadi termasuk klasifikasi apa data baru ini ? Tidak *Order* atau *Order* ?, berikut pada Tabel 2.2 data *training*.

Tabel 2.2 data *training*

NO	Nama Obat	Minimal Stok (X)	Sisa Stok (Y)	Klasifikasi
1	Parasol SPF 33 Cr 20	3	7	Tidak Order
2	Paratusin syr 60	4	8	Tidak Order

3	Paratusin tab	9	10	Tidak Order
4	Peditox	25	6	Order
5	Persidal 2 mg	37	7	Order
6	Pharmaton Formula	47	3	Order
7	Pharmaton Vit Kapl	25	7	Order
8	Pharolit Sach	33	5	Order
9	Phenytoin 100	46	5	Order
10	Pi Kang Shuang Biru	15	6	?

a. Langkah penyelesaian pertama

Pertama, Menentukan parameter K. Misalnya Untuk Parameter tetangga terdekat **K = 3**.

b. Langkah penyelesaian kedua

Kedua, Menghitung hitung jarak antara data baru dengan semua data training, menggunakan *Euclidean Distance*. Hasil hitung seperti pada Tabel 2.3 :

Tabel 2.3 perhitungan jarak dengan euclidean distance

X	Y	Euclidean Distance (15,6)
3	7	$\sqrt{(3-15)^2 + (7-6)^2} = \sqrt{(12)^2 + (1)^2} = \sqrt{145} = 12.04$
4	8	$\sqrt{(4-15)^2 + (8-7)^2} = \sqrt{(11)^2 + (2)^2} = \sqrt{122} = 11.04$
9	10	$\sqrt{(9-15)^2 + (10-6)^2} = \sqrt{(6)^2 + (4)^2} = \sqrt{52} = 7.21$
25	6	$\sqrt{(25-15)^2 + (6-6)^2} = \sqrt{(10)^2 + (0)^2} = \sqrt{100} = 10$
37	7	$\sqrt{(37-15)^2 + (7-6)^2} = \sqrt{(22)^2 + (1)^2} = \sqrt{485} = 22.02$
47	3	$\sqrt{(47-15)^2 + (3-6)^2} = \sqrt{(32)^2 + (3)^2} = \sqrt{1033} = 32.14$
25	7	$\sqrt{(6-15)^2 + (7-6)^2} = \sqrt{(10)^2 + (1)^2} = \sqrt{101} = 10.04$
33	5	$\sqrt{(33-15)^2 + (5-6)^2} = \sqrt{(18)^2 + (1)^2} = \sqrt{324} = 18$
46	5	$\sqrt{(46-15)^2 + (5-6)^2} = \sqrt{(31)^2 + (1)^2} = \sqrt{961} = 31$

c. Langkah penyelesaian ketiga

Ketiga, urutkan jarak dari data baru dengan data training dan menentukan tetangga terdekat berdasarkan jarak minimum K . seperti pada Tabel 2.4.

Tabel 2.4 pengurutan jarak terdekat data baru dengan data training

X	Y	Euclidean Distance (15,6)	Urutan Jarak	Apakah Termasuk 3k-NN?
3	7	12.04	5	Tidak($K>3$)
4	8	11.04	4	Tidak($K>3$)
9	10	7.21	1	Ya ($K<3$)
25	6	10	2	Ya ($K<3$)
37	7	22.02	7	Tidak($K>3$)
47	3	32.14	9	Tidak($K>3$)
25	7	10.04	3	Ya ($K=3$)
33	5	18	6	Tidak($K>3$)
46	5	31	8	Tidak($K>3$)

Dari kolom 3 dan 4 (urutan jarak) hasil mengurutkan dari yang terdekat ke terjauh antara jarak data baru dengan data training. Pada kolom 7 (Apakah termasuk 3 K -NN?) maksudnya adalah K -NN menjadi 3 K -NN , karena nilai K ditentukan sama dengan 3

d. Langkah penyelesaian ke-empat

Ke-empat, tentukan kategori dari tetangga terdekat. perhatikan baris 3, 4, dan 7 pada gambar sebelumnya (diatas). Kategori Ya diambil jika nilai $K \leq 3$. Jadi baris 3, 4, dan 7 termasuk kategori Ya dan sisanya Tidak.

Tabel 2.5 penentuan kategori yang termasuk K=3

X	Y	Euclidean Distance (15,6)	Urutan Jarak	Apakah Termasuk 3k-NN?	Kategori Ya untuk KNN
3	7	12.04	5	Tidak($K > 3$)	-
4	8	11.04	4	Tidak($K > 3$)	-
9	10	7.21	1	Ya ($K < 3$)	Tidak Order
25	6	10	2	Ya ($K < 3$)	Order
37	7	22.02	7	Tidak($K > 3$)	-
47	3	32.14	9	Tidak($K > 3$)	-
25	7	10.04	3	Ya ($K = 3$)	Order
33	5	18	6	Tidak($K > 3$)	-
46	5	31	8	Tidak($K > 3$)	-

Kategori ya untuk *K-NN* pada kolom 10, mencakup baris 3,4, dan 5.

Hasil yang diberikan kategori berdasarkan tabel awal. baris 3 memiliki kategori Tidak *Order*, dan 4,7 memiliki kategori *Order*. seperti pada Tabel 2.6.

Tabel 2.6 penentuan kategori yang termasuk K=3

X	Y	Euclidean Distance (15,6)	Urutan Jarak	Apakah Termasuk 3k-NN?	Kategori Ya untuk KNN
3	7	12.04	5	Tidak($K > 3$)	-
4	8	11.04	4	Tidak($K > 3$)	-
9	10	7.21	1	Ya ($K < 3$)	Tidak Order
25	6	10	2	Ya ($K < 3$)	Order
37	7	22.02	7	Tidak($K > 3$)	-
47	3	32.14	9	Tidak($K > 3$)	-
25	7	10.04	3	Ya ($K = 3$)	Order
33	5	18	6	Tidak($K > 3$)	-
46	5	31	8	Tidak($K > 3$)	-

e. Langkah penyelesaian ke-lima

Ke-lima, gunakan kategori mayoritas yang sederhana dari tetangga yang terdekat tersebut sebagai nilai prediksi data yang baru. seperti pada Tabel 2.7.

Tabel 2.7 hasil klasifikasi berdasarkan kategori mayoritas

NO	Nama Obat	Minimal Stok	Sisa Stok	Klasifikasi
1	Parasol SPF 33 Cr 20	3	7	Tidak Order
2	Paratusin syr 60	4	8	Tidak Order
3	Paratusin tab	9	10	Tidak Order
4	Peditox	25	6	Order
5	Persidal 2 mg	37	7	Order
6	Pharmaton Formula	47	3	Order
7	Pharmaton Vit Kapl	25	7	Order
8	Pharolit Sach	33	5	Order
9	Phenytoin 100	46	5	Order
10	Pi Kang Shuang Biru	15	6	Order

Data yang dimiliki pada baris 3, 4 dan 7 mempunyai 2 kategori *Order* dan 1 kategori *Tidak Order*. Dari jumlah mayoritas (***Order*** > ***Tidak Order***) tersebut, Dapat disimpulkan bahwa data baru (**X=15 dan Y=6**) termasuk dalam kategori ***Order***.

2.2.7 Confusion Matrix

Sebelum tahap model *klasifikasi*, perlu dilakukan pengujian keakruasian data untuk mengukur performa model *klasifikasi* yang akan dihasilkan. Dalam pengukuran Model *klasifikasi* menggunakan *Confision Matrix*.

Confusion Matriks memiliki dimensi 2x2, yang terdiri dari baris *horizontal* yang merupakan kelas prediksi (*predicted classes*) dan kolom *vertical* adalah kelas sebenarnya (*actual classes*). kelas yang digunakan pada matriks dasar yaitu kelas *Yes* dan *No*. *Confusion Matriks* dapat dilihat pada Tabel 3.1 (D. M. P. M. L. T. and Teachniques, 2011).

Tabel 2.8 Confusion Matrix

	<i>Actual Class</i>		
		Yes	No
Predicted Class	Yes	TP (<i>true positive</i>)	FP (<i>false positive</i>)
	No	FN (<i>false negative</i>)	TN (<i>true negative</i>)

Keterangan :

TP : kelas prediksi Y namun kelas sebenarnya Y.

FP : kelas prediksi Y namun kelas sebenarnya N.

FN : kelas prediksi N namun kelas sebenarnya Y.

TN : kelas prediksi N namun kelas sebenarnya N.

Dari *confusion matrix* dasar tersebut, dihitung akurasi untuk mengetahui performa dari model klasifikasi. (D. M. P. M. L. T. and Teachniques, 2011)

1) Akurasi

Akurasi merupakan tahapan pengklasifikasi dalam memprediksi kelas dari masing-masing *istance* dalam kelas.

$$\frac{(TP + TN)}{(TP + FP + TN + FN)} * 100\% \quad (2.2)$$

2.2.8 Kontrol Barang

Persediaan (kontrol) adalah kemampuan suatu perusahaan dalam mengatur dan mengelola setiap kebutuhan barang baik barang mentah, barang setengah jadi dan barang jadi agar selalu tersedia baik dalam kondisi pasar yang stabil maupun berfluktuasi.(Irwandi, 2015). Menurut (Schroeder, 2000), persediaan adalah stock bahan yang digunakan untuk memudahkan produksi atau untuk memuaskan permintaan pelanggan. Konsep persediaan merupakan suatu aktiva yang meliputi barang-barang milik perusahaan dengan maksud untuk dijual dalam suatu periode usaha tertentu, atau persediaan barang-barang yang masih dalam pengerjaan/proses produksi, ataupun persediaan bahan baku yang menunggu penggunaannya dalam suatu proses produksi (Rangkuti, 2004). Barang adalah segala jenis sesuatu yang dapat digunakan sebagai alat pemuas kebutuhan manusi.(Irwandi, 2015).

A. Aturan Kontrol Obat

Apoteker bertanggung jawab terhadap aturan kontrol obat yang berlaku pada Apotek. Aturan Kontrol Obat yang baik akan membantu apoteker untuk dapat mengontrol kebutuhan supply dan demand karena persediaan berperan sebagai penyangga dalam supply dan demand. Dalam Aturan Pengontrolan suatu obat memiliki Aturan Kriteria dari Transaksi, Stok Minimal dan Total Stok Untuk mengetahui

suatu Ordernya barang. Berdasarkan hal tersebut, menurut (Yunarto & Santika, 2005), Ada 3 Aturan Kriteria Kontrol Obat menurut fungsinya:

1. Rata-Rata Pemakaian (*Transaksi*)

Rata-Rata Pemakaian ialah Jumlah obat yang dipakai dalam waktu tertentu.

2. Stok Minimal

Jumlah acuan dalam proses pengontrolan obat, yang akan dilakukan order pada Apotek.

3. Total Stok

Jumlah Keseluruhan obat yang telah di order, yang akan dibutuhkan pihak konsumen.

Untuk Data Aturan kontrol barang Dapat dilihat pada Tabel 2.9

Tabel 2.9 Data obat

NO	Nama Obat	<i>Transaksi</i>	Stok Minimal	Total Stok
1	Parasol SPF 33 Cr 20	2.00	3	9
2	Paratusin syr 60	6.00	4	14
3	Peditox	64.00	25	70
4	Persidal 2 mg	493.00	37	500
5	Pharmaton Formula	336.00	47	339
6	Pharmaton Vit Kapl	88.00	25	95
7	Pharolit Sach	205.00	33	210
8	Phenytoin 100	175.00	46	180
9	Pi Kang Shuang Biru	59.00	15	65
10	Paratusin tab	540.00	9	550

Pada Tabel 2.9 adalah Data berupa Kriteria yang harus dipenuhi dalam Aturan Kontrol Obat, dengan adanya kriteria hasil transaksi, Stok Minimal dan Total Stok

maka pengelolaan suatu stok barang berupa obat pada apotek akan mengetahui hasil berupa order, order bulan depan dan tidak order dari patokan kriteria yang ada.

BAB III

METODE PENELITIAN

3.1. Objek Penelitian

Objek dalam penelitian ini adalah APOTEK UNIVERSITAS AHMAD DAHLAN (UAD). Untuk penentuan rekomendasi *order*, *order* bulan depan, Tidak *Order* pada barang dengan algoritma *K- Nearest Neighbor* (K-NN). pada penentuan rekomendasi dinyatakan *order*, *order* bulan depan, dan Tidak *Order* dapat dilihat Tabel 3.1 berikut.

Tabel 3.1 Hasil Keterangan Data

Klasifikasi	Keterangan
Order	ketika sisa stok lebih sedikit dari minimal stok
Order Bulan Depan	ketika jumlah data minimal stok lebih banyak dari pada data sisa stok dengan perbandingan 10
Tidak Order	untuk Tidak order ketika jumlah data sisa stok lebih banyak dari data minimal stok

Pada pihak apotek untuk Kriteria pada pemberian minimal stok selalu berubah seiringnya tahun yang akan datang dengan patokan transaksi obat apakah lebih meningkat, apabila barang yang sering laku terjual maka minimal stok lebih besar dibandingkan dari minimal stok barang lainnya dengan kategori dikit dalam melakukan penyetokan barang.

3.2. Metode Pengumpulan Data

Dalam penelitian ini yang dibutuhkan informasi transaksi barang pada APOTEK UAD untuk dapat mengetahui rekomendasi order pada penelitian ini, ada beberapa metode pengumpulan data yang digunakan dalam penelitian ini yaitu sebagai berikut :

1. Studi Pustaka

Studi pustaka dilakukan untuk mengetahui kegiatan penelitian yang sudah pernah dilakukan sebelumnya. Metode ini dilakukan dengan membaca buku, paper, dan jurnal terkait algoritma *K-Nearest Neighbor* (KNN).

2. Wawancara

Wawancara merupakan menyusun daftar pertanyaan kemudian melakukan tanya jawab beberapa poin penting pengelolaan data apotek. Wawancara dilakukan agar mendapatkan data yang lebih akurat dan spesifik. Data Yang diperoleh dalam wawancara dengan Tetie Herlina selaku dari pihak Apotek UAD, mendapatkan berupa permasalahan pada sistem kontrol barang pada apotek UAD yang diteliti.

3. Observasi

Observasi merupakan sebuah teknik pengamatan secara langsung datang dilokasi penelitian disaat pengambilan data barang pada APOTEK UAD. Data yang didapat dari hasil observasi pada apotek UAD berupa data obat pada tahun 2016.

3.3. Alat dan Bahan penelitian

Dalam melakukan penelitian ini dibutuhkan beberapa alat untuk membantu jalannya penelitian. Spesifikasi yang digunakan dalam penelitian sebagai berikut:

1. Kebutuhan Perangkat Keras (*Hardware*)

Perangkat Keras yang digunakan dalam penelitian ini adalah :

a. *Processor : Intel core i5-8250U up to 3.4 GHZ*

b. *Ram : 8 GB*

c. *Hard Disk : 1 TB*

2. Kebutuhan Perangkat Lunak (*Software*)

Perangkat Lunak yang digunakan dalam penelitian ini adalah :

a. *Microsoft Excel*

b. *Anaconda 3*

c. *Visual studio code*

d. *Browser : Google chrome*

e. *Library python berupa:*

1) *Pandas*

2) *Os*

3) *numpy*

4) *Glob*

5) *String*

6) *sklearn.model_selection*

- 7) *train_test_split*
- 8) *sklearn.neighbors*
- 9) *KneighborsClassifier*
- 10) *Flask*
- 11) *App*
- 12) *app.module.Engine*
- 13) *dataset,knn,knntest,tests,upload*
- 14) *request,jsonify,render_template,url_for,redirect*
- 15) *werkzeug.utils*
- 16) *secure_filename*

3.4. Tahapan Penelitian

Pada tahapan ini memiliki beberapa tahapan-tahapan dalam melakukan proses penelitian untuk mencapai tujuan.

Pada tahapan ini menjelaskan tahapan-tahapan dapat dilihat pada Gambar 3.1 berikut.



Gambar 3.1 Tahapan Penelitian

1. Pengumpulan Data

Dalam Pengumpulan data telah dilakukan dengan melalui wawancara dengan pihak Apotek Universitas Ahmad Dahlan (UAD) dengan

mendapatkan data penjualan obat dan stok obat pada tahun 2016, berekstensi file excel yang akan dilakukan pada penelitian ini.

2. Kajian Pustaka

Pada penelitian sistem kontrol barang di apotek uad, hasil tahap akhir penelitian akan di bandingkan dengan kajian penelitian terdahulu yang telah melakukan penelitian dengan unsur yang sama dengan metode yang berbeda.

3. Analisis Kebutuhan Sistem

Pada analisis kebutuhan sistem mendeskripsikan berupa awal proses pada inputan sistem,kebutuhan proses, dan output proses seperti apa yang akan dilakukan.

4. Pembuatan sistem

Pada pembuatan sistem mendeskripsikan tahapan-tahapan awal sampai akhir dari sistem yang akan dirancang pada penelitian.

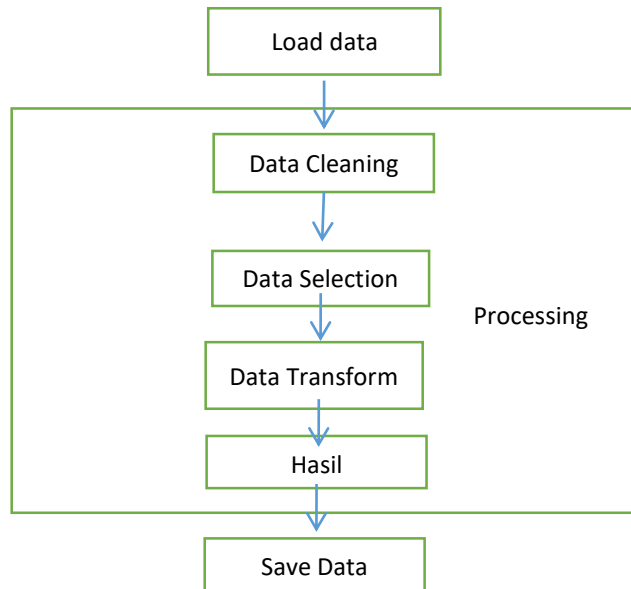
5. Penyusunan Laporan

Akhir dari proses setelah melakukan penelitian, dengan adanya laporan sebagai media penyimpanan berbentuk file berisi informasi yang telah diteliti dengan disimpan pada laporan yang telah dilakukan penelitian.

3.5. Tahapan Alur Sistem

Pada tahapan ini menjelaskan tahapan-tahapan pengklasifikasian data obat untuk memudahkan dalam penyusunan pembahasan didalam BAB IV.

Pada tahapan ini menjelaskan tahapan-tahapan pengklasifikasian data obat dilihat pada Gambar 3.1 berikut.



Gambar 3.2 Tahapan alur sistem

1. Data Cleaning

Melakukan pembersihan data obat, yaitu menghilangkan data yang sama atau data yang tidak konsisten dan juga memperbaiki kesalahan yang ada, misalnya pada data obat terdapat beberapa yang tidak konsisten dengan bernilai kosong yang tidak dapat diproses.

2. Seleksi Data

Pada data obat tidak semuanya dipakai, oleh karena itu hanya mengambil data yang penting dan sesuai untuk analisis untuk dilanjutkan pada tahap proses selanjutnya.

Data yang diambil pada tahap seleksi adalah *Transaksi*, Stok Minimal, Total Stok.

3. Data *Transformation*

Pada tahap *Transformation* merubah beberapa data dalam kategori tertentu sesuai untuk proses data mining. Pada proses transformasi data, variabel total stok dan *Transaksi* menjadi variabel Sisa stok. Hasil dari *transfromasi* data ialah *order*, *order* bulan depan, dan Tidak *order*.

4. Hasil

Pada Tahapan Hasil menampilkan hasil data yang telah diklasifikasi pada data obat dengan pembagian sebagai berikut:

1) Data training

Hasil pemetakan dari total dataset yang telah diolah maka dapat diambil 70% dari total data sebagai data training.

2) Data testing

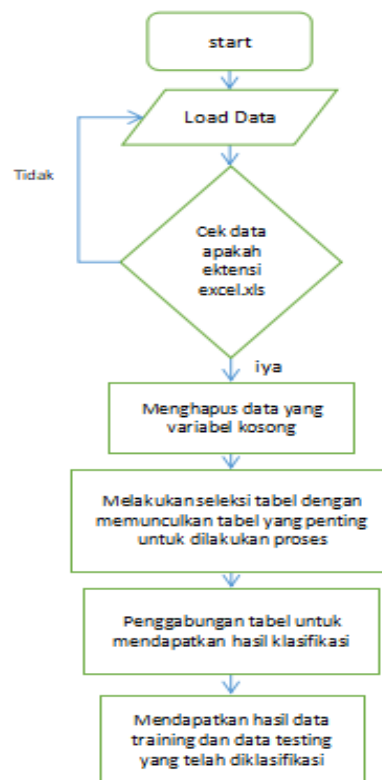
Hasil pemetakan dari seluruh dataset dengan mengambil 30% dan sisa data 70% sebagai data *training*. Data *testing* ialah data yang akan diuji dengan pada sistem dengan mencari hasil akurasi nya .

5. Save Data

Proses penyimpanan data dari hasil proses sistem yang telah dilakukan mendapatkan data *training* dan data *testing* yang akan disimpan pada file excel.

3.6. Processing

Tahapan processing ialah deskripsi suatu langkah-langkah dari alur sistem, Alur processing dijelaskan pada Gambar 3.3.

Gambar 3.3 flowchart *processing*

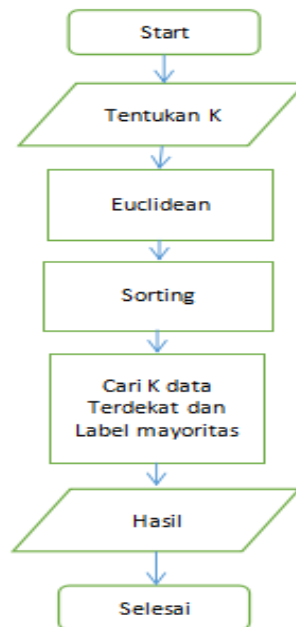
3.7. Implementasi

Implementasi adalah penerapan perancangan kedalam program yang akan dibuat. Program *data mining* ini diimplementasikan dengan bahasa pemrograman *PHP* dan *Framework flash* Pada bagian *controller* bahasa pemrograman yang digunakan ialah *python* yang berenteksi *py* dan untuk tampilan interface yaitu menggunakan bahasa *HTML*.

3.8. K- Nearest Neighbour (KNN)

Proses melakukan klasifikasi terhadap *object* data *training* dan data *testing* dengan beberapa tahapan dengan menggunakan rumus *euclidean*

distance untuk mencari tetangga K jarak tetangga terdekat. Tahapan KNN dijelaskan pada Gambar 3.4.



Gambar 3.4 flowchart KNN

3.9. Pengujian

Tahap pengujian dilakukan untuk meminimalisir kesalahan (error) dan memastikan bahwa keluaran yang dihasilkan harus sesuai yaitu dengan *Euclidean Distance* dan *confusion matrix* yang digunakan untuk mengetahui nilai akurasi dari data *testing* melalui tahapan data mining pada pemrosesan klasifikasi dengan algoritma *K-NN*.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Pengumpulan Data

Tahapan sebelum pembuatan perangkat lunak adalah mengkaji situasi dan mempelajari aplikasi yang akan dibangun. Berdasarkan hasil wawancara yang telah dilakukan dengan pihak Apotek Universitas Ahmad Dahlan memperoleh data transaksi barang berupa penjualan obat dan stok obat pada tahun 2016. Dari hasil wawancara tersebut diperoleh data 1035 data obat.

4.2 Analisis Kebutuhan Sistem

Analisis kebutuhan sistem bertujuan untuk mengetahui kebutuhan pengguna aplikasi yang akan dibangun. Di dalam analisis kebutuhan, program yang diharapkan dapat memecahkan masalah terhadap sistem kontrol barang pada Apotek UAD, pemaparan kemampuan analisis kebutuhan tersebut sebagai berikut.

1. kebutuhan input

Program mampu memasukkan data transaksi obat ditahun 2016 yang merupakan file excel yang akan diupload pada sistem.

2. Kebutuhan proses

- a. Program mampu melakukan proses pembersihan data, dimana data obat yang kosong dan tidak konsisten akan dibersihkan.

- b. Program mampu melakukan proses seleksi data, dimana data obat yang diseleksi merupakan data yang akan dilakukan proses data mining.
 - c. Program mampu melakukan proses transformasi data, dimana atribut data diubah menjadi kategori tertentu untuk mempermudah dalam melakukan proses mining.
3. kebutuhan output
- a. Program mampu menghasilkan pengklasifikasian data obat dari proses data mining yang terdiri hubungan antara variabel yang digunakan.
 - b. Program mampu menghasilkan perhitungan dari akurasi data Testing pada data obat yang telah dilakukan proses mining.

4.3 Processing

Setelah mendapatkan gambaran sistem yang akan dibuat, selanjutnya dilakukan pengembangan. Adapun pengembangan sistem sesuai dengan langkah - langkah data mining sesuai contoh kasus berikut.

1. Load Data

Tahap Dimana data stok barang yang berbentuk File *Excel* diinsertkan pada *button load* data kemudian di *upload*. Berikut data stok barang sebanyak 10 data dari 1035 data yang akan diolah. Diantara 9 data training 1 data testing.

Tabel 4.1 Data awal obat

NO	Nama Obat	Transaksi	Harga	Diskon	Pajak	Jumlah	Stok Minimal	Total Stok
1	Parasol SPF 33 Cr 20	2	104.384	0	10.438	114.822	3	9
2	Paratusin syr 60	6	124.703	0	12.470	137.173	4	14
3	ANTALGIN 500 STRIP	540	118.646	0	11.852	150.499		
4	Peditox	64	334.161	0	33.416	367.577	25	70
5	Persidal 2 mg	493	4.316.138	0	428.942	4.832.680	37	500
6	Pharmaton Formula	336	1.271.597	0	126.824	1.398.421	47	339
7	Pharmaton Vit Kapl	88	334.299	0.00	33.429.	367.729	25	95
8	Pharolit Sach	205	201.527	0.00	20.079	222.407	33	210
9	Phenytoin 100	175	123.750	0.00	12.045	210.495	46	180
10	Pi Kang Shuang Biru	59	695.066	0.00	69.506	764.573	15	65
11	Paratusin tab	540	452.256	0.00	45.129	504.785	9	550
12	Antangin cair	715	1.460.128	0	146.012	1.612.841		

2. Pembersihan data (*Cleaning*)

Setelah data stok barang sudah di *upload*, maka langkah selanjutnya melakukan tahap *cleaning* yaitu tahap membersihkan data stok barang dari data kosong dan yang tidak konsisten pada variabel yang dibutuhkan. Pada tabel 4.1 data nomor 3 dan 12 dihapus karena data yang dimiliki kosong atau tidak diisi. Berikut hasil tabel data *cleaning*:

Tabel 4.2 Data *Cleaning*

NO	Nama Obat	Transaksi	Harga	Diskon	ajak	Jumlah	Stok Minimal	Total Stok
1	Parasol SPF 33 Cr 20	2	104.384	0	10.438	114.822	3	9
2	Paratusin syr 60	6	124.703	0	12.470	137.173	4	14
3	Peditox	64	334.161	0	33.416	367.577	25	70
4	Persidal 2 mg	493	4.316.138	0	428.942	4.832.680	37	500
5	Pharmaton Formula	336	1.271.597	0	126.824	1.398.421	47	339
6	Pharmaton Vit Kapl	88	334.299	0	33.429	367.729	25	95
7	Pharolit Sach	205	201.527	0	20.079	222.407	33	210
8	Phenytoin 100	175.00	123.750	0.00	12,045.00	210.495	46	180
9	Pi Kang Shuang Biru	59.00	695,066.54	0.00	69,506.60	764,573.14	15	65
10	Paratusin tab	540.00	452,256.00	0.00	45,129.50	504,785.50	9	550

3. Seleksi data (*Selection*)

Tahap *selection* adalah tahap memilih variabel dalam data stok barang yang akan digunakan agar sesuai dalam pengolahan data, variabel yang tidak digunakan akan dihilangkan. Variabel data obat antara lain : Nama obat, *Transaksi*, Harga, Diskon, Pajak, Jumlah, Stok minimal, Total Stok. Kemudian diseleksi sehingga hanya ada variabel : *Transaksi*, Stok minimal, Total Stok. Berikut hasil data selection:

Tabel 4.3 *Selection* data stok barang

NO	Nama Obat	<i>Transaksi</i>	Stok Minimal	Total Stok
1	Parasol SPF 33 Cr 20	2.00	3	9
2	Paratusin syr 60	6.00	4	14
3	Peditox	64.00	25	70
4	Persidal 2 mg	493.00	37	500
5	Pharmaton Formula	336.00	47	339
6	Pharmaton Vit Kapl	88.00	25	95
7	Pharolit Sach	205.00	33	210
8	Phenytoin 100	175.00	46	180
9	Pi Kang Shuang Biru	59.00	15	65
10	Paratusin tab	540.00	9	550

4. *Transformasi data*

Transformasi data adalah tahap merubah beberapa data dalam kategori tertentu sesuai untuk proses data *mining*. Pada proses *transformasi data*, variabel total stok dan *Transaksi* menjadi variabel Sisa stok. Hasil *transformasi data* bisa dilihat pada tabel 4.4.

Pengkategorian keterangan *Klasifikasi* mencari order, Tidak Order dan diretur yaitu:

- Dinyatakan Order apabila jumlah data minimal stok lebih banyak dari pada data sisa stok.
- Dinyatakan Tidak Order apabila jumlah data sisa stok lebih banyak dari data minimal stok.

Tabel 4.4 Hasil *Transform*

NO	Nama Obat	Stok Minimal	Sisa Stok	Keterangan
1	Parasol SPF 33 Cr 20	3	7	Tidak Order
2	Paratusin syr 60	4	8	Tidak Order
3	Peditox	25	6	Order
4	Persidal 2 mg	37	7	Order
5	Pharmaton Formula	47	3	Order

6	Pharmaton Vit Kapl	25	7	Order
7	Pharolit Sach	33	5	Order
8	Phenytoin 100	46	5	Order
9	Pi Kang Shuang Biru	15	6	?
10	Paratusin tab	9	10	Tidak Order

Penyelesaian Teknik data mining dengan menggunakan metode K-

Nearest Neighbour (KNN) :

- Menentukan Parameter K = 3
- Menghitung kuadrat jarak *euclid* (*euclidean distance*) masing-masing obyek terhadap data sampel.

Tabel 4.5 Perhitungan Data *Testing* Pertama

No	Nama Obat	Distance
1	d1,d9	12.04
2	d2,d9	11.04
3	d3,d9	10
4	d4,d9	22.02
5	d5,d9	32.14
6	d6,d9	10.04
7	d7,d9	18
8	d8,d9	31
9	d10,d9	7.21

Baris Total Pada kolom *Distance* pada tabel 4.5 dihitung dengan perhitungan Rumus *Euclidean Distance* sebagai berikut.

$$\begin{aligned}
 1) \text{ d1,d9} &= \sqrt{(3-15)^2 + (7-6)^2} \\
 &= \sqrt{(12)^2 + (1)^2} \\
 &= \sqrt{145}
 \end{aligned}$$

$$= 12.04$$

$$2) \quad d_{2,d9} = \sqrt{(4-15)^2 + (8-7)^2}$$

$$= \sqrt{(11)^2 + (2)^2}$$

$$= \sqrt{122}$$

$$= 11.04$$

$$3) \quad d_{3,d9} = \sqrt{(25-15)^2 + (6-6)^2}$$

$$= \sqrt{(10)^2 + (0)^2}$$

$$= \sqrt{100}$$

$$= 10$$

$$4) \quad d_{4,d9} = \sqrt{(37-15)^2 + (7-6)^2}$$

$$= \sqrt{(22)^2 + (1)^2}$$

$$= \sqrt{485}$$

$$= 22.02$$

$$5) \quad d_{5,d9} = \sqrt{(47-15)^2 + (3+6)^2}$$

$$= \sqrt{(32)^2 + (3)^2}$$

$$= \sqrt{1033}$$

$$= 32.14$$

$$6) \quad d_{6,d9} = \sqrt{(6-15)^2 + (7-6)^2}$$

$$= \sqrt{(10)^2 + (1)^2}$$

$$= \sqrt{101}$$

$$= 10.04$$

$$7) \ d_{7,d9} = \sqrt{(33-15)^2 + (5-6)^2}$$

$$= \sqrt{(18)^2 + (1)^2}$$

$$= \sqrt{324}$$

$$= 18$$

$$8) \ d_{8,d9} = \sqrt{(46-15)^2 + (5-6)^2}$$

$$= \sqrt{(31)^2 + (1)^2}$$

$$= \sqrt{961}$$

$$= 31$$

$$9) \ d_{10,d9} = \sqrt{(9-15)^2 + (10-6)^2}$$

$$= \sqrt{(6)^2 + (4)^2}$$

$$= \sqrt{52}$$

$$= 7.21$$

- c. Mengurutkan objek-objek kedalam kelompok yang mempunyai jarak euclidean terkecil. Untuk melakukan pengurutan membuat urutan dari data terkecil ke terbesar.

Tabel 4.6 Ranking data

No	Nama Obat	Distance	Ranking
1	d1,d9	12.04	5
2	d2,d9	11.04	4
3	d3,d9	10	2
4	d4,d9	22.02	7
5	d5,d9	32.14	9
6	d6,d9	10.04	3
7	d7,d9	18	6
8	d8,d9	31	8
9	d10,d9	7.21	1

- d. Mengumpulkan kategori klasifikasi, pada tahap ini mengambil data sesuai ranking terkecil berdasarkan $K=3$. Hasilnya sebagai berikut.

Tabel 4.7 Hasil Data Testing Terbaik

No	Nama Obat	Distance	Ranking
9	d10,d9	7.21	1
3	d3,d9	10	2
6	d6,d9	10.04	3

- e. Dengan menggunakan Kategori mayoritas, maka dapat hasil klasifikasi dari dari ketiga data yang terbaik. Dari ketiga data tersebut harus lihat hasil kolom keterangan yang dominan muncul pada ketiga data dari hasil perangkingan.

Tabel 4.8 Hasil Data Testing terbaik dengan Keterangan

No	Nama Obat	Distance	Ranking	Keterangan
9	d10,d9	7.21	1	Tidak Order
3	d3,d9	10	2	Order

6	d6,d9	10.04	3	Order
---	-------	-------	---	-------

- f. Data terbaik dari ketiga data d10,d3,d6 adalah Order, karena kategori Order mayoritas data yang paling dominam muncul dari ketiga data terbaik. Maka d9 mendapatkan Hasil kategori Order dengan acuan ketiga data terbaik.

Tabel 4.9 Hasil Klasifikasi Data yang dicari

No	Nama Obat	Stok minimal	Sisa Stok	Keterangan
9	Pi Kang Shuang Biru	15	6	Order

4.4 Implementasi

Tahapan ini merupakan tahapan dimana perancangan dan pembuatan untuk membangun sebuah sistem kontrol barang dengan menggunakan bahasa pemograman *PHP* dan *Framework flash* Pada bagian *controller* bahasa pemograman yang digunakan ialah *python* yang berenteksi *py* dan untuk tampilan *interface* yaitu menggunakan bahasa *HTML*. Tampilan-tampilan yang telah dibuat sebagai berikut.

1. Tampilan *load* Data

Upload Datasets Process Save Data									
Dataset Data Cleaning Data Selection Data Transform KNN Predic									
Jumlah Dataset : 1000									
	No.	Nama Obat	Transaksi	Harga	Diskon	Pajak	Jumlah	Stok Minimal	Total Stok
0	1	Abate 10 g	425.0	924343.96	0.00	92434.40	1017878.36	15.0	430.0
1	2	Acid Salisil	206.6	17596.99	0.00	1759.71	62256.70	10.0	220.0
2	3	Aclonac 25 (P)	50.0	48702.50	0.00	4870.25	56572.75	18.0	510.0
3	4	Aclonac 50 (P)	1039.0	898226.00	0.00	88976.48	1040402.48	16.0	1050.0
4	5	Acnol Lotion 10 ml	1.0	9296.00	0.00	929.60	10225.60	2.0	7.0
5	6	Actazolam 1 mg	100.0	338433.00	0.00	33843.30	396076.30	13.0	110.0
6	7	Actifed hijau 60ml	29.0	1047282.56	0.00	104598.43	1151880.99	15.0	40.0
7	8	Actifed Kuning 60ml	5.0	173958.80	0.00	17266.06	191224.86	3.0	7.0

Gambar 4.1 tampilan Hasil *upload* data

Berdasarkan Gambar 4.1 yang merupakan hasil tampilan dari proses *upload* data yang telah dilakukan oleh user, dimana akan ditampilkan atribut didalam file *excel* yaitu No, Nama obat, *Transaksi*, Harga, Diskon, Pajak, Jumlah, Stok Minimal dan Total Stok.

Berikut *listing* programnya.

Listing 4.1 *upload* data

1	def index():
2	if request.method == 'POST':
3	file = request.form['upload-file']
4	data = pd.read_excel(file)
5	jumlah = len(data)
6	dataset = pd.DataFrame(data)
7	Return render_template('index.html', ses = [dataset.to_html()], countDataset = jumlah)
8	return render_template('index.html', ses = '')

Berikut ini penjelasan dari listing 5.1 :

- 1) Pada baris 1 - 2 merupakan perintah panggilan dengan inisiasi *POST* dalam melakukan penguploadan file
- 2) Pada baris 3 - 4 merupakan perintah untuk upload file dengan data harus file *excel*.
- 3) Baris 5- 7 merupakan perintah menampilkan jumlah seluruh dataset yang diupload pada tampilan *index.html*

2. Tampilan Hasil *Cleaning* Data

Data Cleaning Data Seletion Data Transform KNN Precic									
Jumlah DataCleaning : 700									
	No.	Nama Obat	Transaksi	Harga	Diskon	Pajak	Jumlah	Stok Minimal	Total Stok
0	1	Abate 10 g	425.0	924343.96	0.00	92434.40	1017878.36	15.0	430.0
1	2	Acid Salisil	206.6	17596.99	0.00	1759.71	62256.70	10.0	220.0
2	3	Aclonac 25 (P)	50.0	48702.50	0.00	4870.25	56572.75	18.0	510.0
3	4	Aclonac 50 (P)	1039.0	898226.00	0.00	88976.48	1040402.48	16.0	1050.0
4	5	Acnol Lotion 10 ml	1.0	9296.00	0.00	929.60	10225.60	2.0	7.0
5	6	Actazolam 1 mg	100.0	338433.00	0.00	33843.30	396076.30	13.0	110.0
6	7	Actifed hijau 60ml	29.0	1047282.56	0.00	104598.43	1151880.99	15.0	40.0
7	8	Actifed Kuning 60ml	5.0	173958.80	0.00	17266.06	191224.86	3.0	7.0

Gambar 4.2 Tampilan Data *Cleaning*

Berdasarkan Gambar 4.2 yang merupakan hasil tampilan dari proses data *cleaning* yang telah dilakukan oleh user, dimana akan ditampilkan menghapus atribut data yang kosong dan tidak relevan. Dari hasil data *cleaning* total data tersisa 700 dari 1000 data.

Berikut *listing* programnya.

Listing 4.2 tampilan *cleaning*

1	def dataset():
2	df_dropna = df.dropna(axis=0)

Berikut ini penjelasan dari listing 5.2 :

Pada baris 1 dan 2 ialah perintah menghapus kolom data yang bernilai kosong.

3. Tampilan Hasil *secelection* Data

Tampilan Hasil sebelum *Selection* Data dan sesudah *Selection* Data

Datasets Process Save Data									
Data Cleaning Data Seletion Data Transform KNN Precic									
Jumlah DataCleaning : 700									
	No.	Nama Obat	Transaksi	Harga	Diskon	Pajak	Jumlah	Stok Minimal	Total Stok
0	1	Abate 10 g	425.0	924343.96	0.00	92434.40	1017878.36	15.0	430.0
1	2	Acid Salisil	206.6	17596.99	0.00	1759.71	62256.70	10.0	220.0
2	3	Aclonac 25 (P)	50.0	48702.50	0.00	4870.25	56572.75	18.0	510.0
3	4	Aclonac 50 (P)	1039.0	898226.00	0.00	88976.48	1040402.48	16.0	1050.0
4	5	Acnol Lotion 10 ml	1.0	9296.00	0.00	929.60	10225.60	2.0	7.0
5	6	Actazolam 1 mg	100.0	338433.00	0.00	33843.30	396076.30	13.0	110.0
6	7	Actifed hijau 60ml	29.0	1047282.56	0.00	104598.43	1151880.99	15.0	40.0
7	8	Actifed Kuning 60ml	5.0	173958.80	0.00	17266.06	191224.86	3.0	7.0

Gambar 4.3 Tampilan sebelum *Selection* Data

Datasets Process Save Data									
Data Cleaning Data Seletion Data Transform Hasil KNN									
Jumlah Data Selection : 700									
	Transaksi	Stok Minimal	Total Stok						
0	425.0	15.0	430.0						
1	206.6	10.0	220.0						
2	50.0	18.0	510.0						
3	1039.0	16.0	1050.0						
4	1.0	2.0	7.0						
5	100.0	13.0	110.0						
6	29.0	15.0	40.0						
7	5.0	3.0	7.0						
8	22.0	11.0	50.0						

Gambar 4.4 Tampilan sesudah *Selection* Data

Berdasarkan Gambar 4.4 yang merupakan hasil tampilan *selection* data setelah tahapan *cleaning* data. pada dari proses *selection* data yang telah dilakukan oleh user, dimana akan menampilkan atribut yang diperlukan yaitu *Transaksi*, *Stok Minimal* dan *Total Stok*.

Berikut *listing* programnya.

Listing 4.3 tampilan *selection*

1	def dataset() :
2	selection = df_dropna.drop(['No.', 'Harga', 'Diskon', 'Pajak', 'Jumlah'], axis=1)
3	a = df_dropna.drop(['No.', 'Harga', 'Diskon', 'Pajak', 'Jumlah'], axis=1)

Berikut ini penjelasan dari listing 5.3 :

Pada baris 1-3 perintah melakukan *selection* dengan menghilangkan variabel data No, Nama Obat, Harga, Diskon, Pajak dan Jumlah pada sistem.

4. Tampilan Hasil *Transformasi* Data

a. Tampilan Hasil sebelum *Transformasi* Data dan sesudah *Transformasi* Data

Datasets

Process

Save Data

Data Cleaning

Data Selection

Data Transform

Hasil KNN

Jumlah Data Selection : 700

	Transaksi	Stok Minimal	Total Stok
0	425.0	15.0	430.0
1	206.6	10.0	220.0
2	50.0	18.0	510.0
3	1039.0	16.0	1050.0
4	1.0	2.0	7.0
5	100.0	13.0	110.0
6	29.0	15.0	40.0
7	5.0	3.0	7.0
8	22.0	11.0	50.0

Gambar 4.5 sebelum *Transformasi* Data

Datasets Process Save Data			
Data Cleaning Data Seletion Data Transform Hasil KNN			
Jumlah Data Transform : 700			
	Stok Minimal	sisa	Keterangan
0	15.0	5.0	Order
1	10.0	13.4	Order Bulan Depan
2	18.0	460.0	Tidak Order
3	16.0	11.0	Order
4	2.0	6.0	Order Bulan Depan
5	13.0	10.0	Order
6	15.0	11.0	Order
7	3.0	2.0	Order
8	11.0	28.0	Tidak Order
9	15.0	13.0	Order

Gambar 4.6 sesudah *Transformasi* Data

Berdasarkan Gambar 4.6 yang merupakan hasil tampilan *Transformasi* data setelah tahapan *selection* data, dimana akan menampilkan penggabungan atribut Total Stok dengan *Transaksi*. Hasil dari *Transformasi* data mendapatkan atribut sisa stok dan mendapatkan hasil *klasifikasi* Order, Tidak order, Diretur pada tabel keterangan.

Berikut *listing* programnya.

Listing 4.4 tampilan *Transformasi* data

1	<code>transform['sisa']= transform['Total Stok']-transform['Quantity']</code>
2	<code>transform.loc[transform['sisa'] <= transform['Stok Minimal'],'Keterangan'] = 'Order'</code>
3	<code>transform.loc[(transform['sisa'] <= transform['Stok Minimal']+10) & (transform['sisa'] > transform['Stok Minimal']),'Keterangan'] = 'Order Bulan Depan'</code>
4	<code>transform.loc[transform['sisa'] > transform['Stok Minimal']+10,'Keterangan'] = 'Tidak Order'</code>

Berikut ini penjelasan dari listing 5.4 :

- 1) Pada baris 1 melakukan perintah *transformasi* dengan penggabungan tabel total stok dengan tabel *Transaksi* dengan kondisi (-) yang akan ditampilkan pada hasil tabel sisa.
- 2) Pada baris 2 ialah kondisi ketika sisa stok lebih kecil atau sama dengan stok minimal maka hasilnya Order.
- 3) Pada baris 3 ialah kondisi ketika sisa stok lebih kecil atau sama dengan stok minimal dengan patokan perbandingan 10 maka hasilnya Order bulan depan.
- 4) Pada baris 4 ialah kondisi ketika sisa stok lebih besar dan dengan stok minimal maka hasilnya Tidak Order.

b. Tampilan Hasil

Datasets Process Save Data				
Process Dataset				
Data Cleaning Data Seletion Data Transform Hasil KNN				
Data Training Data Testing				
Jumlah Data Training : 490				
	Nama Obat	Stok Minimal	sisa	Keterangan
0	Abate 10 g	15	5	Order
2	Aclonac 25 (P)	18	460	Tidak Order
3	Aclonac 50 (P)	16	11	Order
4	Acnol Lotion 10 ml	2	6	Order Bulan Depan
5	Actazolam 1 mg	13	10	Order
8	Actifed merah 60 ml	11	28	Tidak Order
9	Acyclovir 200	15	13	Order

Gambar 4.7 Hasil

Berdasarkan Gambar 4.7 yang merupakan hasil tampilan menu hasil *knn*, dimana akan menampilkan berupa bagian Hasil Data *training*, Data *testing* dan akurasi. Berikut *listing* programnya.

Listing 4.5 tampilan Hasil *KNN*

1	<code>train = dataset()[4].sort_index(ascending=True)</code> <code>jumlahTraining = len(train)</code>
2	<code>jumlahTesting = len(test)</code>
3	<code>return render_template('kp.html',</code>

Berikut ini penjelasan dari *listing* 5.5 :

- 1) Perintah baris 1 - 2 pemanggilan *library python* untuk *sorting* pada urutan *index* perdata yang akan ditampilkan pada jumlah *training* dan fungsi *testing*.
- 2) Pada baris ke 3 perintah pemanggilan *library* yang dipakai akan di tampilkan di *kp.html*.

c. Tampilan Data *Training*

Datasets Process Save Data				
Process Dataset				
Data Cleaning Data Seletion Data Transform Hasil KNN				
Data Training Data Testing				
Jumlah Data Training : 490				
	Nama Obat	Stok Minimal	sis	Keterangan
0	Abate 10 g	15	5	Order
2	Aclonac 25 (P)	18	460	Tidak Order
3	Aclonac 50 (P)	16	11	Order
4	Acnol Lotion 10 ml	2	6	Order Bulan Depan
5	Actazolam 1 mg	13	10	Order
8	Actifed merah 60 ml	11	28	Tidak Order
9	Acyclovir 200	15	13	Order

Gambar 4.8 Hasil Data *Training*

Berdasarkan Gambar 4.8 menampilkan Data *training* berjumlah 497 data dari 70% pembagian dengan data *testing*. Tampilan data testing berisi atribut stok minimal, sisa stok, dan keterangan.

Berikut *listing* programnya.

Listing 4.6 tampilan data *training*

1	<code>def knn(x,y):</code>
2	<code>x_train, x_test, y_train, y_test = train_test_split(x, y,</code> <code>test size=0.30, random state=5)</code>
3	<code>knn.fit(x_train,y_train.values.ravel())</code>

Berikut ini penjelasan dari listing 5.6 :

- 1) Pada baris 1 ialah membuat inisiasi variabel x dan y sebagai pemanggilan data *training* dan *testing* pada library program.
- 2) Pada baris 2 dan 43 pemanggilan perintah library dengan inisiasi x dan y dengan perintah *train_test_split* (pemisahan data testing 30% pada data *training* yang otomatis mendapatkan data 70% pada program) dan pengambilan data testing diambil secara acak.

d. Tampilan Data *Testing*

Upload Datasets Process Save Data				
Proccess Dataset				
Dataset Data Cleaning Data Seletion Data Transform Hasil				
Data Training Data Testing				
Jumlah Data Testing : 210				
ID	Nama Obat	Stok Minimal	sis	Keterangan
27	Alcohol 70% 300 cc	18	48	Tidak Order
941	Ob Herbat Junior 60	3	6	Order Bulan Depan
222	Canesten 10 gr	5	3	Order
452	Esperson cr 5	4	24	Tidak Order
841	Minyak Gondopuro Ika	2	3	Order Bulan Depan
818	Mensana Femmy	10	1	Order
541	Glucophage 500	40	0	Order
271	Cimetidin	10	0	Order

Gambar 4.9 Hasil Data *Testing*

Berdasarkan Gambar 4.8 menampilkan Data testing berjumlah 214 data dari 30% pembagian dengan data *training*. Tampilan data *training* berisi atribut stok minimal, sisa stok, dan keterangan.

Berikut *listing* programnya.

Listing 4.7 tampilan data testing

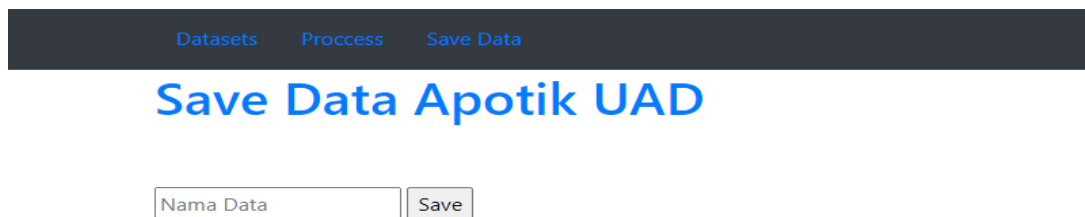
1	def knn(x,y):
2	x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.30, random state=5)
3	knn.fit(x_train,y_train.values.ravel())

Berikut ini penjelasan dari listing 5.7 :

- 1) Pada baris 1 ialah membuat inisiasi variabel x dan y sebagai pemanggilan data *training* dan *testing* pada library program.

- 2) Pada baris 3 dan 4 pemanggilan perintah library dengan inisiasi x dan y dengan perintah *train_test_split* (pemisahan data testing 30% pada data *training* yang otomatis mendapatkan data 70% pada program) dan pengambilan data *testing* diambil secara acak.

e. *Save data*



Gambar 4.10 *Save data*

Proses penyimpanan data dari hasil proses sistem yang telah dilakukan mendapatkan data *training* dan data *testing* yang akan disimpan pada file *excel*.

Berikut *listing* programnya.

Listing 4.8 tampilan data testing

1	<code>xlsx = '.xlsx'</code>
2	<code>workbook = xlsxwriter.Workbook(n)</code>
3	<code>train.to_excel(writer, sheet_name='Training')</code>
4	<code>test.to_excel(writer, sheet_name='Testing')</code>

Berikut ini penjelasan dari listing 5.8 :

- 1) Pada baris 1 dan 2 ialah membuat ekstensi *format* data yang akan disimpan dan pemanggilan library *xlsxwriter* .
- 2) Pada baris 3 dan 4 ialah proses pemanggilan *library train* dan *test* yang akan di simpan pada file *excel* nantinya dengan *sheet* (halaman) yang berbeda sesuai keterangan data.

4.5 K- Nearest Neighbour

a. Menentukan Nilai K

Pada penelitian ini Nilai K = 3 yaitu order, order bulan depan, Tidak Order

b. Menggunakan Rumus *Euclidean Distance*

Pada Tahapan ini dalam perhitungan dengan menggunakan Rumus *euclidean Distance* yaitu dengan perbandingan satu persatu dari data testing dengan data training maka akan mendapatkan hasil yang akan dilanjutkan pada tahapan selanjutnya.

1) Data Training

Data *Training* ialah Data yang akan dilatih pada proses KNN. Untuk Data *Training* bisa dilihat pada Gambar 4.11.

ID	Stok Minimal	sis	Keterangan
0	3	7	Order Bulan Depan
4	47	64	Tidak Order
6	33	55	Tidak Order
8	15	6	Order
9	9	10	Order Bulan Depan
10	15	6	Order
11	9	10	Order Bulan Depan
14	3	8	Order Bulan Depan
15	3	18	Tidak Order
16	15	2	Order
17	15	2	Order
18	80	2	Order
19	40	80	Tidak Order
20	10	20	Order Bulan Depan
23	20	20	Order
25	2	2	Order
26	3	1	Order
27	95	31	Order
28	50	3	Order
29	4	4	Order
30	3	10	Order Bulan Depan
31	80	2	Order
32	200	1	Order

Gambar 4.11 Data Training

2) Data Testing

Data testing ialah Data yang akan diuji pada proses KNN. Untuk Data Testing bisa dilihat pada Gambar 4.12

ID	Stok Minimal	sis	Predict	K3
21	30	70	Tidak Order	Tidak Order
7	46	75	Tidak Order	Tidak Order
5	25	7	Order	Order
2	25	26	Order	Order Bulan Depan
3	37	7	Order	Order
22	40	27	Order	Order
13	2	5	Order Bulan Depan	Order Bulan Depan
24	2	3	Order Bulan Depan	Order Bulan Depan
12	50	85	Tidak Order	Order
1	4	3	Order Bulan Depan	Order

Gambar 4.12 Data Testing

c. Sorting

Proses melakukan sorting dengan jumlah $K = 3$, dari hasil perhitungan dengan menggunakan Rumus Euclidean Distance. Hasil sorting bisa dilihat pada Gambar 4.13.

d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
19	0	17	20	19	4	4	0	29	26
6	6	25	11	31	23	9	16	16	28
4	4	9	23	16	16	11	15	15	30
Tidak Order	Order Bulan Depan	Order	Order Bulan	Tidak Order	Tidak Order	Tidak Order	Order Bulan	Order	Order
Tidak Order	Tidak Order	Order	Order Bulan Depan	Order	Order	Order Bulan Depan	Order	Order	Order
Tidak Order	Tidak Order	Order Bulan Depan	Order	Order	Order	Order Bulan	Tidak Order	Tidak Order	Order Bulan

Gambar 4.13 Sorting

d. Mencari K tetangga terdekat dan label mayoritas

Pada proses mencari K tetangga terdekat dengan hasil $K = 3$ pada hasil Sorting. Dilihat pada gambar 4.14.

d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
19	0	17	20	19	4	4	0	29	26
6	6	25	11	31	23	9	16	16	28
4	4	9	23	16	16	11	15	15	30
Tidak Order	Order Bulan Depan	Order	Order Bulan	Tidak Order	Tidak Order	Tidak Order	Order Bulan	Order	Order
Tidak Order	Tidak Order	Order	Order Bulan Depan	Order	Order	Order Bulan Depan	Order	Order	Order
Tidak Order	Tidak Order	Order Bulan Depan	Order	Order	Order	Order Bulan	Tidak Order	Tidak Order	Order Bulan

Gambar 4.14 Tetangga terdekat dan label mayoritas

e. Hasil

Pada hasil $k=3$ telah didapatkan dengan membandingkan dengan tetangga terdekat dan label mayoritas yang sering muncul, maka hasilnya dapat dilihat pada gambar 4.15.

ID	Stok Minimal	sis	Predict	K3
21	30	70	Tidak Order	Tidak Order
7	46	75	Tidak Order	Tidak Order
5	25	7	Order	Order
2	25	26	Order	Order Bulan Depan
3	37	7	Order	Order
22	40	27	Order	Order
13	2	5	Order Bulan Depan	Order Bulan Depan
24	2	3	Order Bulan Depan	Order Bulan Depan
12	50	85	Tidak Order	Order
1	4	3	Order Bulan Depan	Order

Gambar 4.15 Hasil KNN

4.6 Pengujian

Pengujian akurasi akan dilakukan untuk mengetahui sebaik apa metode *k-nearest neighbour* dalam hal uji akurasi. Uji akurasi dilakukan menggunakan *confusion matrix*.

Tabel 4.10 Hasil pengujian akurasi

Nilai	Pengujian Data training =490 Data testing =210
Accuracy	TP = 120 FP =0 TN =80 FN =10
Rumus	$\frac{TP+ TN}{N(\text{Total Data})} \times 100\%$
Hasil	$\frac{120 +80}{210} \times 100\%$ = 95%
Error rate	5%
keterangan	TP = Hasil Order Berelasi dengan Hasil Order TN =Hasil Order Bulan Depan Berelasi dengan Order Bulan Depan & Hasil Tidak Order Berelasi dengan Tidak Order FN = Hasil Tidak Order Berelasi dengan Order FP = Hasil Order Berelasi dengan Order bulan depan, Hasil Order Berelasi dengan Tidak Order, Order Bulan depan berelasi dengan Tidak Order

Pada hasil penelitian yang telah dilakukan dengan data *testing* 210 mendapatkan hasil akurasi 95% dan *error rate* 5%, maka dapat disimpulkan penelitian dengan metode *K-Nearest Neighbour (KNN)* mendapatkan hasil yang bagus dalam penelitian sistem kontrol barang pada Apotek Universitas Ahmad Dahlan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dalam penelitian ini, penerapan data mining telah berhasil dalam melakukan proses untuk *klasifikasi* data obat pada sistem kontrol barang pada apotek universitas ahmad dahlan. Berdasarkan perhitungan dengan *k-nearest neighbour*, maka dapat disimpulkan bahwa.

- 1) Dapat diterapkannya metode *K-Nearest Neighbour* untuk *klasifikasi* data obat.
- 2) Terujinya metode *K-Neasrest Neighbour* dalam hal keakurasian. Hasil keakurasian sistem pengujian dengan data *training* 490 dan data *testing* 210, didapatkan akurasi sebesar 95%.

5.2 Saran

Untuk memperbaiki atau menyempurnakan penelitian ini dapat diberikan saran. Penelitian selanjutnya dapat menggunakan variabel lain yang belum ada dalam penelitian ini.

1. Penelitian selanjutnya dapat menggunakan variabel lain yang belum ada dalam penelitian ini.
2. Penggunaan data yang terbaru untuk melakukan penelitian selanjutnya.

DAFTAR PUSTAKA

- Daniel, T. (2005). *An Introduction to Data Mining*.
- Hasmawati, M. M. K., Studi, K. N. N., & Tumaka, K. (2019). *Aplikasi prediksi penjualan barang menggunakan metode k- nearest neighbor (knn) (studi kasus tumaka mart)*. July.
- Irwandi. (2015). Maulan Irwadi, S.E., M.Si., Ak. CA. *Jurnal Akuntansi Politeknik Sekayu (ACSY)*, 12(3), 210. PENERAPAN REORDER POINT UNTUK PERSEDIAAN BAHAN BAKU PRODUKSI ALAT PABRIK KELAPA SAWIT PADA PT. SWAKARYA ADHI USAHA KABUPATEN BANYUASIN%0AMaulan
- Kafil, M., & Industri, F. T. (2019). *PENERAPAN METODE K-NEAREST NEIGHBORS*. 3(2), 59–66.
- Nurdiawan, U., & Pemasaran, S. (2018). *Jurnal Teknologi Informasi dan Komunikasi STMIK Subng, April 2018 ISSN: 2252-4517*. April, 84–95.
- Simbolon, P. H. (2019). *Implementasi Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori (Studi Kasus : Srikandi Cash Credit Elektronik dan Furniture)*. 6(4), 401–406.
- Han & Kamber, (2012) (testimony of Data Mining Concepts and Teachniques).
- Teachniques, D. M. P. M. L. T. and. (2011). *Witten*.
- Vercellis, Carlo. (2009). *Business intelligence : Datamining and optimization for decision making*. Chichester: John Wiley& Sons.
- Connolly, Thomas and Begg, Carolyn, 2010, *Database Systems: A Practical Approach to Design, Implementation, and Management*, Fifth Edition. Pearson Education, Boston.
- Winiarti, S., Yuliansyah, H., Purnama, A. A., & Dahlan, U. A. (2018). *Identification of Toddlers ' Nutritional Status using Data Mining Approach*. 9(1), 164–169.
- Larose, Daniel T. 2005. *Discovering Knowledge in Data : An Introduction to Data Mining*. John Willey & Sons, Inc.
- Yunarto, Holy Iacun & Santika, Martinus Getty. (2005), "Business Concepts Implementation Series In Inventory Management", Jakarta: Penerbit PT Elex

Media Komputindo.

E.Prasetyo, Data Mining: Konsep dan Aplikasi menggunakan Matlab, 1 ed.
Yogyakarta: Andi Offset, 2012.

Schroeder, Roger G. 2000. Operations Management: Contemporary Concepts and Cases. International Edition. Mc Graw Hill company. New York

Rangkuti, Freddy. 2004. Manajemen Persediaan Aplikasi di Bidang Bisnis. Jakarta :
PT. Raja Grafindo Persada.