

# Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode *K-NEAREST NEIGHBOR* (STUDI KASUS : RAKUNI BAKERY, PASTRY AND CAKE)

Sandy Aldy Pradana<sup>1\*</sup>, Mochamad Adhari Adiguna<sup>1</sup>

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

Email: <sup>1\*</sup>[sandialdy99@gmail.com](mailto:sandialdy99@gmail.com), <sup>2</sup>[dosen01864@unpam.ac.id](mailto:dosen01864@unpam.ac.id)

(\* : coressponding author)

**Abstrak**– Pada setiap toko, penjualan adalah hal yang sangat penting agar toko bisa tetap beroperasi dan mendapatkan penghasilan untuk diputar kembali untuk membeli bahan atau barang yang dibutuhkan untuk dijual. Setiap toko akan bersaing melalui mutu dan kualitas produk agar penjualan bisa tetap naik sesuai yang diharapkan. Pemilik toko biasanya akan menggunakan prediksi pada penjualan untuk mengetahui berapa banyak produk yang akan terjual pada bulan yang akan datang. Data mining adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat. Berdasarkan tugasnya, data mining dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, clustering dan asosiasi. Proses dalam tahap data mining terdiri dari tiga langkah Utama, yaitu data Preparation Pada langkah ini, data dipilih, dibersihkan, dan dilakukan preprocessed mengikuti pedoman dan knowledge dari ahli domain yang menangkap dan mengintegrasikan data internal dan eksternal ke dalam tinjauan organisasi secara menyeluruh. Pada pola informasi yang sudah dihasilkan setelah proses data mining harus ditampilkan dalam bentuk yang dapat dengan mudah di pahami oleh pihak yang berkepentingan. Tahap ini akan memeriksa apakah pola dari informasi yang telah di temukan sesuai atau bertentangan dengan fakta atau hipotesis yang ada sebelumnya. Pada tahap ini akan di dapatkan pola informasi penjualan produk dari proses data mining dengan metode KNN. Untuk melakukan prediksi penjualan produk terlaris peneliti menggunakan perhitungan data mining, yang menggunakan teknik klasifikasi dan algoritma K-Nearest Neighbor, dari 10 produk yang terjual didapatkan hasil prediksi pada RapidMiner penjualan yang terlaris dengan pengujian confusion matrix yaitu mantau burger dan pao casio ayam dengan nilai akurasi 85.00% dan pengujian cross validation yaitu mantau burger dan boneka cewe pao dengan nilai akurasi 87.00%.

**Kata Kunci:** Data Mining, Metode *K-Nearest Neighbor*, Prediksi Produk Terlaris

**Abstract**– In every shop, sales are very important so that the store can continue to operate and get income to be rolled back to buy materials or goods needed to be sold. Each store will compete through product quality and quality so that sales can continue to rise as expected. Store owners will usually use predictions on sales to find out how many products will be sold in the coming month. Data mining is a process that uses statistics, mathematics, artificial intelligence and machine learning to extract and identify useful information. Based on activity, data mining is collected into description, estimation, prediction, classification, clustering and association. The process in the data mining stage consists of three main steps, namely data preparation. In this step, data is selected, cleaned, and preprocessed following the guidelines and knowledge from domain experts who capture and integrate internal and external data into an overall organizational review. the information that has been generated after the data mining process must be displayed in a form that can be easily understood by the parties concerned. This stage will check whether the pattern of information that has been found is in accordance with or contrary to the facts or hypotheses that existed before. At this stage, product sales information patterns will be obtained from the data mining process with the KNN method. To predict sales of best-selling products, researchers use data mining calculations, which use classification techniques and the K-Nearest Neighbor algorithm, from the 10 products sold, the prediction results are obtained from RapidMiner Sales, which sell by Confusion Matrix testing, namely monitor burgers and chicken casio pao with accuracy values 85.00% and cross validation testing, namely monitoring burgers and girl pao dolls with an accuracy value of 87.00%.

**Keywords:** Data Mining, Method *K-Nearest Neighbor*, Best Selling Products Prediction

## 1. PENDAHULUAN

Saat ini industri makanan dan minuman menjadi salah satu sektor yang memiliki pengaruh besar terhadap perkembangan ekonomi nasional. Pada tahun 2018 Kementerian Perindustrian

mencatat industri makanan dan minuman bertumbuh 7,91 persen (Kementerian Perindustrian Republik Indonesia, 2019). Salah satu sektor usaha dalam industri makanan yang diminati adalah industri makanan kue. Adanya perkembangan zaman yang cukup pesat sehingga banyak bermunculan usaha-usaha dalam industri makanan kue yang menjual berbagai macam kue. Pengusaha kue memikirkan banyak inovasi baru untuk terus mengembangkan bisnisnya (Budiman et al 2020).

Toko Rakuni Bakery adalah sebuah toko kue dan pastry yang sudah memiliki beberapa cabang didaerah Jakarta dan Tangerang. Toko ini menyediakan berbagai macam varian Cake dan Pastry mulai dari Aneka Roti, Kue Kering(Cookies), Kue Basah, Bolu dan Aneka Dessert..

Dalam penjualan kebutuhan layanan informasi sangatlah penting untuk memprediksi produk terlaris produk. Dalam memprediksi produk terlaris menganalisa secara manual bisa diketahui membutuhkan waktu yang cukup lama. Sedangkan, melalui aplikasi yang menggunakan data mining dapat mempermudah dalam proses prediksi produk terlaris yang memiliki akurasi lebih baik.

Data Mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Salah satu teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model. Kebutuhan untuk prediksi juga dapat memanfaatkan teknik ini. Dalam data mining, pengelompokan data juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui pola universal data-data yang ada. (Puspita, 2020).

## 2. METODOLOGI PENELITIAN

### 2.1 Metode Pengumpulan Data

Berikut merupakan kerangka metodologi penelitian yang akan dijalankan:

#### 1. Studi Pustaka

Studi Pustaka adalah suatu metode pengumpulan data dengan cara mencari, membaca dan mempelajari berbagai macam referensi yang berhubungan dengan penelitian yang sedang dikaji. Seperti buku, jurnal, e-book, skripsi ataupun tesis.

#### 2. Observasi

Observasi adalah suatu metode pengumpulan data dengan cara menguji atau melihat langsung ke lokasi penelitian untuk mengumpulkan informasi yang ada. Observasi yang dilakukan pada penelitian ini dengan mendatangi langsung Toko Rakuni Bakery.

### 2.2 Metode *K-Nearest Neighbor*

*K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek yang berdasarkan dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut. KNN merupakan algoritma *supervised learning* dimana hasil dari query instance yang baru diklasifikasi berdasarkan mayoritas dari kategori pada algoritma KNN. Dimana kelas yang paling banyak muncul yang nantinya akan menjadi kelas hasil dari klasifikasi. Kedekatan didefinisikan dalam jarak metrik, seperti jarak *Euclidean*.

Berisi penjelasan tentang tahapan penelitian yang menggambarkan urutan logis untuk mendapatkan hasil penelitian sesuai dengan harapan dan gambaran sistem. Jika ada gambar dan tabel, itu harus disajikan dengan nama tabel dan gambar yang disertai dengan nomor urut.

Tahapan – tahapan untuk menghitung algoritma K-NN adalah sebagai berikut :

1. Menentukan nilai K (nilai K dipilih secara manual).
2. Menghitung jarak antar data training dan data testing dengan menggunakan rumus Euclidean Distance.
3. Kemudian mengurutkan objek-objek (data yang sudah dihitung jaraknya) tersebut ke dalam kelompok yang berdasarkan jarak terkecil.
4. Menetapkan kelas, dimana kelas yang dipilih adalah kelas dengan jumlah nilai K terbanyak pada data testing.

### 2.3 *K-Nearest Neighbor Dengan Euclidean Distance*

Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode *Euclidean Distance*. *Euclidean Distance* merupakan salah satu metode perhitungan jarak yang digunakan untuk mengukur jarak dari 2 buah titik dalam *Euclidean space* (Nishom, 2019). Berikut ini adalah rumus *Euclidean Distance*:

$$d_e = \sqrt{\sum_{k=1}^m (f d_{i,k} - k_j)^2}$$

Keterangan :

- de = Jarak euclidean
- fdi = Data training
- k = Data testing
- m = Jumlah data penelitian

Nantinya hasil yang didapat dari perhitungan akan digunakan untuk mengetahui jarak antara data training dan testing. Selanjutnya akan dilakukan pengurutan berdasarkan data paling kecil ke data paling besar dengan melihat nilai K sebagai batas dengan jarak tetangga terdekat. Lalu untuk menentukan hasil prediksi akan diambil berdasarkan jumlah mayoritas yang ada pada batas data tetangga terdekat.

## 3. ANALISA DAN PEMBAHASAN

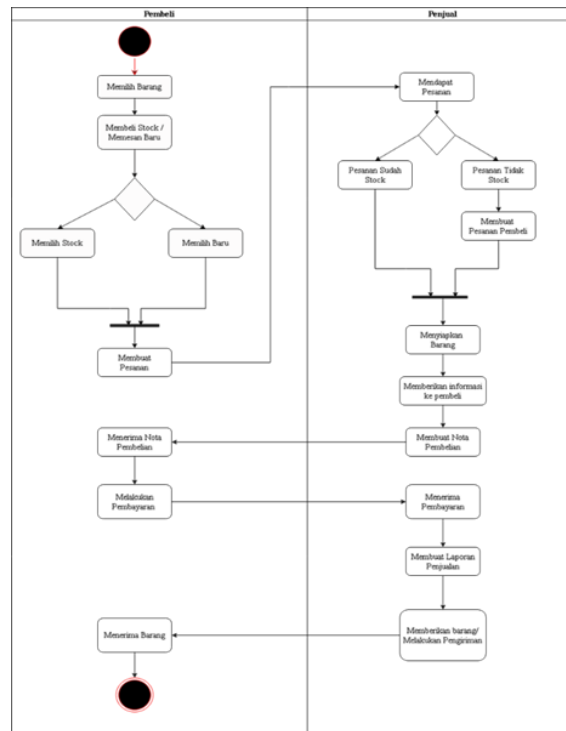
### 3.1 Analisa Sistem

Analisa sistem adalah tahapan penelitian yang berfokus terhadap sistem yang sedang berjalan dan memiliki tujuan untuk mengetahui segala permasalahan yang terjadi serta memudahkan dalam menjalankan tahap berikutnya yaitu tahap perancangan penelitian. Dalam tahapan ini akan digambarkan sistem yang sedang berjalan, penjelasan masalah yang terjadi pada sistem dan memberikan solusi untuk menyelesaikan masalah tersebut dengan mengganti sistem yang berjalan dengan sistem yang diusulkan. Gunalan huruf kecil dan abjed untuk penomoran list.

#### 3.1.1 Analisa Sistem Berjalan

Analisa sistem yang berjalan saat ini bermaksud untuk mengetahui tentang sistem penjualan yang ada pada Rakuni Bakery yang beralamat di Ruko Mendrisio 1 Blok D, Tangerang Selatan, Banten. Tidak adanya sistem prediksi penjualan yang tentunya sangat menyulitkan pihak toko untuk mengetahui apa yang akan terjadi pada penjualan di kemudian hari. Analisa sistem berjalan pada Toko Rakuni Bakery adalah sebagai berikut :

1. Konsumen datang ke toko untuk membeli produk
2. Konsumen memilih produk, terdapat dua kategori dalam memilih produk yang sudah stock (ready) dan tidak stock (pre-order)
3. Penjual mengetahui produk yang diinginkan pembeli
4. Penjual menyiapkan produk
5. Penjual memberikan informasi kepada pembeli bahwa barangnya sudah siap
6. Penjual membuat struk pembelian
7. Pembeli menerima struk pembelian
8. Pembeli melakukan pembayaran
9. Penjual menerima pembayaran
10. Penjual melakukan pengiriman produk
11. Pembeli menerima produk yang dipesan



**Gambar 1.** Activity Diagram Berjalannya Toko

### 3.1.2 Analisa Sistem Usulan

Analisa sitem usulan adalah bentuk pemecahan masalah yang terjadi pada sistem yang lama dengan melakukan pengembangan sistem baru. Pada tahapan ini akan dilakukan proses pengolahan data dengan mengikuti tahapan Knowledge Discovery in Database (KDD) untuk melakukan prediksi penjualan. Berikut ini tahapan-tahapan penelitiannya:

1. Pengumpulan Data
2. Seleksi Data.
3. Preprocessing
4. Transformation
5. Data Mining.
6. Interpretation/Evaluasi
7. Validasi dan Pengujian
8. Analisis Hasil Pengujian

### 3.2 Data Selection

Data selection atau seleksi data merupakan tahapan menyeleksi sekumpulan data operasional pengujian yang dilakukan untuk menggali informasi dalam proses KDD dimulai. Hasil dari data seleksi ini akan digunakan dalam proses data mining, disimpan di suatu berkas dan terpisah dari data operasional pengujian. Data atribut ini meliputi data penjualan yang didalamnya berisikan file Nama Produk, Jumlah, dan Bulan. Semua data ini akan di seleksi dan yang akan digunakan hanya 4 file saja yang digunakan pada proses Knowledge Discovery in Databas (KDD). File yang akan digunakan yaitu:

1. Nama barang merupakan file atribut didalam tabel penjualan yang berisikan informasi dari nama barang yang terjual pada Rakuni Bakery.
2. Kuantitas merupakan file atribut didalam tabel penjualan yang berisikan informasi jumlah barang yang terjual pada Rakuni Bakery.
3. Bulan merupakan file atribut pada label penjualan yang berisikan informasi tentang bulan transaksi penjualan pada Toko Rakuni Bakery.

### 3.3 Preprocessing

Tahapan selanjutnya pada KDD yaitu tahap preprocessing yang merupakan tahap yang harus dilakukan sebelum melakukan data mining. Tahapan ini meliputi proses integrasi data untuk menggabungkan data dari database yang berbeda-beda, selanjutnya agar menghasilkan data yang bersih untuk melakukan tahapan mining maka data harus terlebih dahulu melakukan proses cleaning

Bulan Januari					Bulan Februari				
Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4	Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4
Boneka Cewe Pao	2	3	3	2	Boneka Cewe Pao	4	5	4	6
Boneka Cowo Pao	4	7	8	4	Boneka Cowo Pao	5	6	2	3
Mantau Burger	3	5	3	3	Mantau Burger	4	6	3	5
Mantau Gandum	6	4	5	7	Mantau Gandum	3	4	6	6
Mantau Spesial	3	3	2	3	Mantau Spesial	6	7	7	10
Pao Ayam Kecap	6	3	3	2	Pao Ayam Kecap	3	7	4	3
Pao Casio Ayam	2	1	2	3	Pao Casio Ayam	4	7	8	5
Pao Classic Ayam	1	3	0	5	Pao Classic Ayam	4	3	2	3
Talas Pao	3	2	2	2	Talas Pao	1	4	3	1
Tausa Pao	1	2	1	1	Tausa Pao	3	2	2	2

Bulan Maret					Bulan April				
Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4	Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4
Boneka Cewe Pao	4	3	6	8	Boneka Cewe Pao	3	6	5	6
Boneka Cowo Pao	7	5	6	5	Boneka Cowo Pao	2	1	4	3
Mantau Burger	5	7	6	6	Mantau Burger	5	6	4	8
Mantau Gandum	6	4	5	6	Mantau Gandum	3	2	4	2
Mantau Spesial	5	7	4	3	Mantau Spesial	2	5	7	4
Pao Ayam Kecap	4	5	5	4	Pao Ayam Kecap	6	4	7	4
Pao Casio Ayam	9	7	6	10	Pao Casio Ayam	5	8	3	6
Pao Classic Ayam	3	2	4	3	Pao Classic Ayam	3	4	2	4
Talas Pao	1	0	1	2	Talas Pao	3	1	2	0
Tausa Pao	0	1	2	0	Tausa Pao	2	0	2	1

Bulan Mei					Bulan Juni				
Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4	Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4
Boneka Cewe Pao	3	7	5	3	Boneka Cewe Pao	4	4	3	2
Boneka Cowo Pao	1	4	3	4	Boneka Cowo Pao	3	4	2	3
Mantau Burger	8	5	6	6	Mantau Burger	2	1	3	4
Mantau Gandum	3	4	2	5	Mantau Gandum	2	1	2	4
Mantau Spesial	6	8	4	6	Mantau Spesial	3	2	2	3
Pao Ayam Kecap	4	2	3	4	Pao Ayam Kecap	3	3	2	2
Pao Casio Ayam	4	3	5	4	Pao Casio Ayam	5	6	4	4
Pao Classic Ayam	3	5	5	2	Pao Classic Ayam	2	2	3	3
Talas Pao	4	3	2	4	Talas Pao	4	3	2	3
Tausa Pao	1	3	2	1	Tausa Pao	1	3	1	2

Bulan Juli					Bulan Agustus				
Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4	Nama Produk	Minggu 1	Minggu 2	Minggu 3	Minggu 4
Boneka Cewe Pao	3	4	3	3	Boneka Cewe Pao	6	4	7	7
Boneka Cowo Pao	3	2	3	3	Boneka Cowo Pao	5	5	3	3
Mantau Burger	2	2	4	2	Mantau Burger	5	7	6	4
Mantau Gandum	3	2	3	1	Mantau Gandum	5	6	4	4
Mantau Spesial	3	3	2	2	Mantau Spesial	6	5	4	6
Pao Ayam Kecap	2	4	1	3	Pao Ayam Kecap	3	3	2	4
Pao Casio Ayam	4	3	6	6	Pao Casio Ayam	7	9	7	8
Pao Classic Ayam	2	3	3	2	Pao Classic Ayam	1	1	4	2
Talas Pao	3	4	2	3	Talas Pao	5	4	4	4
Tausa Pao	1	3	0	3	Tausa Pao	3	5	2	2

### 3.4 Transformation

Di dalam tahapan ini seluruh data operasional menghasilkan data pengelompokan atribut yang akan digunakan. Pada proses data mining, yaitu atribut dari data bulan dan klasifikasi dari kriteria data yang akan menjadi bahan utama untuk melakukan proses data mining.

No	Nama Produk	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6	Bulan 7	Bulan 8	Bulan 9	Bulan 10	Bulan 11	Bulan 12	Jumlah
1	Boneka Cewe Pao	2	4	4	3	3	4	3	6	3	3	5	4	44
2	Boneka Cowo Pao	4	5	7	2	1	3	3	5	2	1	4	0	37
3	Mantau Burger	3	4	5	5	8	2	2	5	7	5	7	4	57
4	Mantau Gandum	6	3	6	3	3	2	3	5	4	8	8	7	58
5	Mantau Spesial	3	6	5	2	6	3	3	6	7	7	10	5	63
6	Pao Ayam Kecip	6	3	4	6	4	3	2	3	4	3	5	3	46
7	Pao Casio Ayam	2	4	9	5	4	5	4	7	5	4	5	5	59
8	Pao Classic Ayam	1	4	3	3	3	2	2	1	3	1	10	3	36
9	Talies Pao	3	1	1	3	4	4	3	5	5	4	7	4	44
10	Tausa Pao	1	3	0	2	1	1	1	3	2	2	3	2	21
11	Boneka Cewe Pao	3	5	3	6	7	4	4	4	5	2	7	3	53
12	Boneka Cowo Pao	7	6	5	1	4	4	2	5	0	0	3	0	37
13	Mantau Burger	5	6	7	6	5	1	2	7	6	9	5	3	62
14	Mantau Gandum	4	4	4	2	4	1	2	6	3	5	7	6	48
15	Mantau Spesial	3	7	7	5	8	2	3	5	5	5	14	8	72
16	Pao Ayam Kecip	3	7	5	4	2	3	4	3	7	3	7	3	51
17	Pao Casio Ayam	1	7	7	8	3	6	3	9	7	5	4	7	67
18	Pao Classic Ayam	3	3	2	4	5	2	3	1	3	3	15	4	48
19	Talies Pao	2	4	0	1	3	3	4	4	3	3	4	3	34
20	Tausa Pao	2	2	1	0	3	3	3	5	1	4	2	1	27
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
100	Boneka Cewe Pao	2	6	8	6	3	2	3	7	3	3	6	4	53

Dari data diatas kita akan transformasi dengan merubah jumlah menjadi 3 kategori yaitu Sangat Laris, Cukup Laris, Kurang Laris. Berikut adalah prosesnya:

- Menghitung rentang dengan (jumlah maksimal – jumlah minimal)  
 $72 - 21 = 51$
- Lalu hitung rentang dibagi jumlah kategori  
Rentang :  $3 = 51 : 3 = 17$
- Menentukan kategori 1 dengan menjumlahkan, jumlah minimal :  
Rentang  $21 + 17 = 38$  (Kurang Laris)
- Menentukan kategori 2 dengan menjumlahkan, kategori 1 + rentang  
 $38 + 17 = 55$  (Cukup Laris)
- Menentukan kategori 3 dengan menjumlahkan, Kategori 2 + rentang  
 $55 + 17 = 72$  (Sangat Laris)
- Maka didapat tabel kategori sebagai berikut :

Jumlah	Kategori
$\leq 38$	Kurang Laris
$>38 \leq 55$	Cukup Laris
$>55 \leq 72$	Sangat Laris

No	Nama Produk	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6	Bulan 7	Bulan 8	Bulan 9	Bulan 10	Bulan 11	Bulan 12	Jumlah	Keterangan
1	Boneka Cewe Pao	2	4	4	3	3	4	3	6	3	3	5	4	44	Cukup Laris
2	Boneka Cowo Pao	4	5	7	2	1	3	3	5	2	1	4	0	37	Kurang Laris
3	Mantau Burger	3	4	5	5	8	2	2	5	7	5	7	4	57	Sangat Laris
4	Mantau Gandum	6	3	6	3	3	2	3	5	4	8	8	7	58	Sangat Laris
5	Mantau Spesial	3	6	5	2	6	3	3	6	7	7	10	5	63	Sangat Laris
6	Pao Ayam Kecip	6	3	4	6	4	3	2	3	4	3	5	3	46	Cukup Laris
7	Pao Casio Ayam	2	4	9	5	4	5	4	7	5	4	5	5	59	Sangat Laris
8	Pao Classic Ayam	1	4	3	3	3	2	2	1	3	1	10	3	36	Kurang Laris
9	Talies Pao	3	1	1	3	4	4	3	5	5	4	7	4	44	Cukup Laris
10	Tausa Pao	1	3	0	2	1	1	1	3	2	2	3	2	21	Kurang Laris
11	Boneka Cewe Pao	3	5	3	6	7	4	4	4	5	2	7	3	53	Cukup Laris
12	Boneka Cowo Pao	7	6	5	1	4	4	2	5	0	0	3	0	37	Kurang Laris
13	Mantau Burger	5	6	7	6	5	1	2	7	6	9	5	3	62	Sangat Laris
14	Mantau Gandum	4	4	4	2	4	1	2	6	3	5	7	6	48	Cukup Laris
15	Mantau Spesial	3	7	7	5	8	2	3	5	5	5	14	8	72	Sangat Laris
16	Pao Ayam Kecip	3	7	5	4	2	3	4	3	7	3	7	3	51	Cukup Laris
17	Pao Casio Ayam	1	7	7	8	3	6	3	9	7	5	4	7	67	Sangat Laris
18	Pao Classic Ayam	3	3	2	4	5	2	3	1	3	3	15	4	48	Cukup Laris
19	Talies Pao	2	4	0	1	3	3	4	4	3	3	4	3	34	Kurang Laris
20	Tausa Pao	2	2	1	0	3	3	3	5	1	4	2	1	27	Kurang Laris
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
100	Boneka Cewe Pao	2	6	8	6	3	2	3	7	3	3	6	4	53	Cukup Laris

Data yang sudah siap selanjutnya akan diklasifikasikan menjadi dua data yaitu data training dan data testing dengan komposisi 80 : 20. Pembagian data training dan data testing akan dilakukan secara random dengan menggunakan Sampling Split Data yang terdapat pada aplikasi RapidMiner.

### 3.5 Penerapan *K-Nearest Neighbor*

Berikut adalah perhitungan jarak antara data training dan data testing :

**Perhitungan jarak pada data testing D1 =**

$$a1 = \sqrt{\begin{matrix} (2-3)^2 + (4-5)^2 + (4-6)^2 + (3-4)^2 + \\ (3-4)^2 + (4-3)^2 + (3-3)^2 + \\ (6-5)^2 + (3-2)^2 + (3-3)^2 + (5-4)^2 + (4-1)^2 \end{matrix}} = 4,582$$

$$a2 = \sqrt{\begin{matrix} (4-3)^2 + (5-5)^2 + (7-6)^2 + (2-4)^2 + \\ (1-4)^2 + (3-3)^2 + (3-3)^2 + \\ (5-5)^2 + (2-2)^2 + (1-3)^2 + (4-4)^2 + (0-1)^2 \end{matrix}} = 4,242$$

$$a3 = \sqrt{\begin{matrix} (3-3)^2 + (4-5)^2 + (5-6)^2 + (5-4)^2 + \\ (8-4)^2 + (2-3)^2 + (2-3)^2 + \\ (5-5)^2 + (7-2)^2 + (5-3)^2 + (7-4)^2 + (4-1)^2 \end{matrix}} = 8,246$$

$$a4 = \sqrt{\begin{matrix} (6-3)^2 + (3-5)^2 + (6-6)^2 + (3-4)^2 + \\ (3-4)^2 + (2-3)^2 + (3-3)^2 + \\ (5-5)^2 + (4-2)^2 + (8-3)^2 + (8-4)^2 + (7-1)^2 \end{matrix}} = 9,848$$

$$a5 = \sqrt{\begin{matrix} (3-3)^2 + (6-5)^2 + (5-6)^2 + (2-4)^2 + \\ (6-4)^2 + (3-3)^2 + (3-3)^2 + \\ (6-5)^2 + (7-2)^2 + (7-3)^2 + (10-4)^2 + (5-1)^2 \end{matrix}} = 10,198$$

$$a6 = \sqrt{\begin{matrix} (6-3)^2 + (3-5)^2 + (4-6)^2 + (6-4)^2 + \\ (4-4)^2 + (3-3)^2 + (2-3)^2 + \\ (3-5)^2 + (4-2)^2 + (3-3)^2 + (5-4)^2 + (3-1)^2 \end{matrix}} = 5,916$$

$$a7 = \sqrt{\begin{matrix} (2-3)^2 + (4-5)^2 + (9-6)^2 + (5-4)^2 + \\ (4-4)^2 + (5-3)^2 + (4-3)^2 + \\ (7-5)^2 + (5-2)^2 + (4-3)^2 + (5-4)^2 + (5-1)^2 \end{matrix}} = 6,928$$

$$a8 = \sqrt{\begin{matrix} (1-3)^2 + (4-5)^2 + (3-6)^2 + (3-4)^2 + \\ (3-4)^2 + (2-3)^2 + (2-3)^2 + \\ (1-5)^2 + (3-2)^2 + (1-3)^2 + (10-4)^2 + (3-1)^2 \end{matrix}} = 8,888$$

$$a9 = \sqrt{\begin{matrix} (3-3)^2 + (1-5)^2 + (1-6)^2 + (3-4)^2 + \\ (4-4)^2 + (4-3)^2 + (3-3)^2 + \\ (5-5)^2 + (5-2)^2 + (4-3)^2 + (7-4)^2 + (4-1)^2 \end{matrix}} = 8,426$$

$$a10 = \sqrt{\begin{matrix} (1-3)^2 + (3-5)^2 + (0-6)^2 + (2-4)^2 + \\ (1-4)^2 + (1-3)^2 + (1-3)^2 + \\ (3-5)^2 + (2-2)^2 + (2-3)^2 + (3-4)^2 + (2-1)^2 \end{matrix}} = 8,485$$

$$a11 = \sqrt{\frac{(3-3)^2 + (5-5)^2 + (3-6)^2 + (6-4)^2 + (7-4)^2 + (4-3)^2 + (4-3)^2}{(4-5)^2 + (5-2)^2 + (2-3)^2 + (7-4)^2 + (3-1)^2}} = 6,928$$

$$a12 = \sqrt{\frac{(7-3)^2 + (6-5)^2 + (5-6)^2 + (1-4)^2 + (4-4)^2 + (4-3)^2 + (2-3)^2}{(5-5)^2 + (0-2)^2 + (0-3)^2 + (3-4)^2 + (0-1)^2}} = 6,633$$

$$a13 = \sqrt{\frac{(5-3)^2 + (6-5)^2 + (7-6)^2 + (6-4)^2 + (5-4)^2 + (1-3)^2 + (2-3)^2}{(7-5)^2 + (6-2)^2 + (9-3)^2 + (5-4)^2 + (3-1)^2}} = 6,480$$

$$a14 = \sqrt{\frac{(4-3)^2 + (4-5)^2 + (4-6)^2 + (2-4)^2 + (4-4)^2 + (1-3)^2 + (2-3)^2}{(6-5)^2 + (3-2)^2 + (5-3)^2 + (7-4)^2 + (6-1)^2}} = 7,416$$

$$a15 = \sqrt{\frac{(3-3)^2 + (7-5)^2 + (7-6)^2 + (5-4)^2 + (8-4)^2 + (2-3)^2 + (3-3)^2}{(5-5)^2 + (5-2)^2 + (5-3)^2 + (14-4)^2 + (8-1)^2}} = 3,872$$

**Perhitungan jarak pada data testing D2=**

$$b1 = \sqrt{\frac{(2-7)^2 + (4-6)^2 + (4-6)^2 + (3-4)^2 + (3-5)^2 + (4-2)^2 + (3-3)^2}{(6-4)^2 + (3-5)^2 + (3-8)^2 + (5-8)^2 + (4-7)^2}} = 9,643$$

$$b2 = \sqrt{\frac{(4-7)^2 + (5-6)^2 + (7-6)^2 + (2-4)^2 + (1-5)^2 + (3-2)^2 + (3-3)^2}{(5-4)^2 + (2-5)^2 + (1-8)^2 + (4-8)^2 + (0-7)^2}} = 12,206$$

$$b3 = \sqrt{\frac{(3-7)^2 + (4-6)^2 + (5-6)^2 + (5-4)^2 + (8-5)^2 + (2-2)^2 + (2-3)^2}{(5-4)^2 + (7-5)^2 + (5-8)^2 + (7-8)^2 + (4-7)^2}} = 7,483$$

$$b4 = \sqrt{\frac{(6-7)^2 + (3-6)^2 + (6-6)^2 + (3-4)^2 + (3-5)^2 + (2-2)^2 + (3-3)^2}{(5-4)^2 + (4-5)^2 + (8-8)^2 + (8-8)^2 + (7-7)^2}} = 4,123$$

$$b5 = \sqrt{\frac{(3-7)^2 + (6-6)^2 + (5-6)^2 + (2-4)^2 + (6-5)^2 + (3-2)^2 + (3-3)^2}{(6-4)^2 + (7-5)^2 + (7-8)^2 + (10-8)^2 + (5-7)^2}} = 6,324$$



$$b_6 = \sqrt{\begin{matrix} (6-7)^2 + (3-6)^2 + (4-6)^2 + (6-4)^2 + \\ (4-5)^2 + (3-2)^2 + (2-3)^2 + \\ (3-4)^2 + (4-5)^2 + (3-8)^2 + (5-8)^2 + (3-7)^2 \end{matrix}} = 8,544$$

$$b_7 = \sqrt{\begin{matrix} (2-7)^2 + (4-6)^2 + (9-6)^2 + (5-4)^2 + \\ (4-5)^2 + (5-2)^2 + (4-3)^2 + \\ (7-4)^2 + (5-5)^2 + (4-8)^2 + (5-8)^2 + (5-7)^2 \end{matrix}} = 9,380$$

$$b_8 = \sqrt{\begin{matrix} (1-7)^2 + (4-6)^2 + (3-6)^2 + (3-4)^2 + \\ (3-5)^2 + (2-2)^2 + (2-3)^2 + \\ (1-4)^2 + (3-5)^2 + (1-8)^2 + (10-8)^2 + (3-7)^2 \end{matrix}} = 11,704$$

$$b_9 = \sqrt{\begin{matrix} (3-7)^2 + (1-6)^2 + (1-6)^2 + (3-4)^2 + \\ (4-5)^2 + (4-2)^2 + (3-3)^2 + \\ (5-4)^2 + (5-5)^2 + (4-8)^2 + (7-8)^2 + (4-7)^2 \end{matrix}} = 9,949$$

$$b_{10} = \sqrt{\begin{matrix} (1-7)^2 + (3-6)^2 + (0-6)^2 + (2-4)^2 + \\ (1-5)^2 + (1-2)^2 + (1-3)^2 + \\ (3-4)^2 + (2-5)^2 + (2-8)^2 + (3-8)^2 + (2-7)^2 \end{matrix}} = 14,212$$

$$b_{11} = \sqrt{\begin{matrix} (3-7)^2 + (5-6)^2 + (3-6)^2 + (6-4)^2 + \\ (7-5)^2 + (4-2)^2 + (4-3)^2 + \\ (4-4)^2 + (5-5)^2 + (2-8)^2 + (7-8)^2 + (3-7)^2 \end{matrix}} = 9,591$$

$$b_{12} = \sqrt{\begin{matrix} (7-7)^2 + (6-6)^2 + (5-6)^2 + (1-4)^2 + \\ (4-5)^2 + (4-2)^2 + (2-3)^2 + \\ (5-4)^2 + (0-5)^2 + (0-8)^2 + (3-8)^2 + (0-7)^2 \end{matrix}} = 13,416$$

$$b_{13} = \sqrt{\begin{matrix} (5-7)^2 + (6-6)^2 + (7-6)^2 + (6-4)^2 + \\ (5-5)^2 + (1-2)^2 + (2-3)^2 + \\ (7-4)^2 + (6-5)^2 + (9-8)^2 + (5-8)^2 + (3-7)^2 \end{matrix}} = 6,855$$

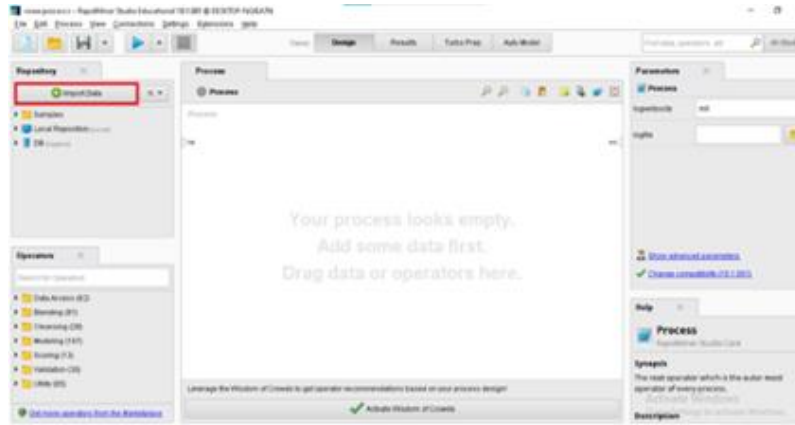
$$b_{14} = \sqrt{\begin{matrix} (4-7)^2 + (4-6)^2 + (4-6)^2 + (2-4)^2 + \\ (4-5)^2 + (1-2)^2 + (2-3)^2 + \\ (6-4)^2 + (3-5)^2 + (5-8)^2 + (7-8)^2 + (6-7)^2 \end{matrix}} = 6,557$$

$$b_{15} = \sqrt{\begin{matrix} (3-7)^2 + (7-6)^2 + (7-6)^2 + (5-4)^2 + \\ (8-5)^2 + (2-2)^2 + (3-3)^2 + \\ (5-4)^2 + (5-5)^2 + (5-8)^2 + (14-8)^2 + (8-7)^2 \end{matrix}} = 8,660$$

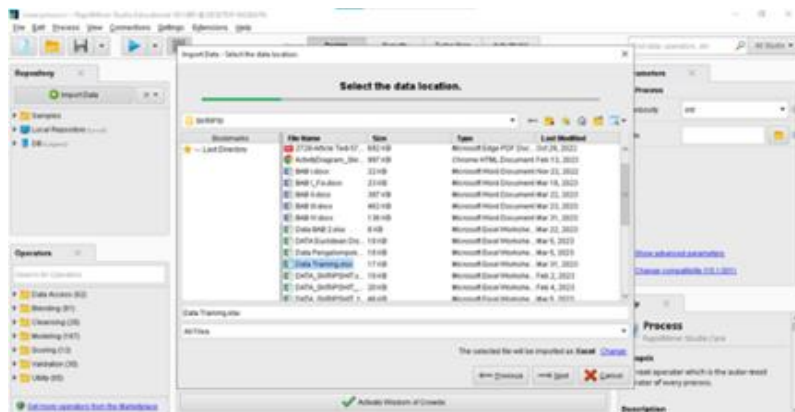
## 4. IMPLEMENTASI

### 4.1 Implementasi Pada Rapid Miner

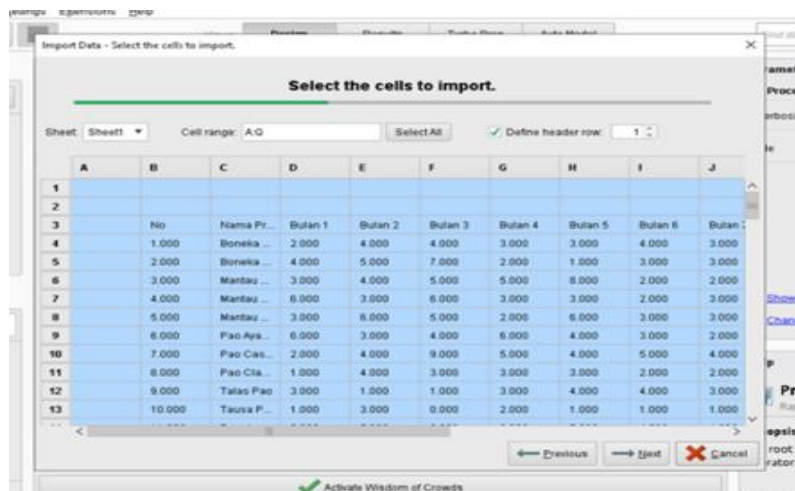
Metode yang digunakan pada pengumpulan data dalam program aplikasi ini adalah sebagai berikut:



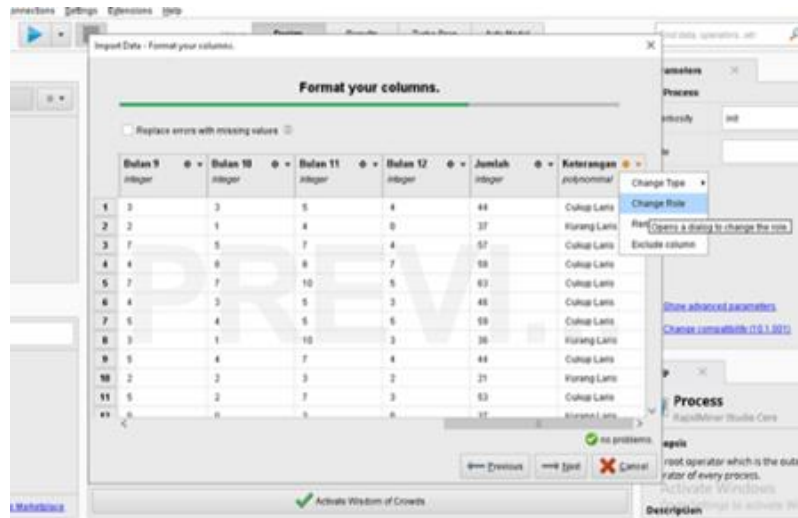
Gambar 1. Import Data



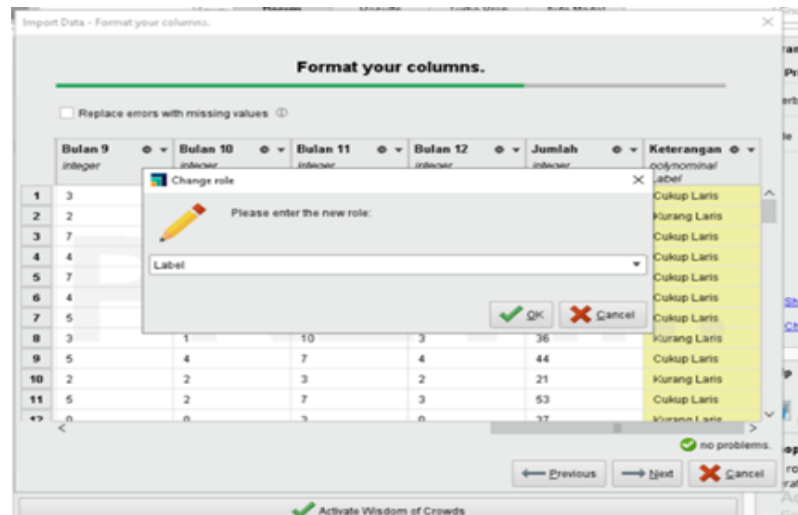
Gambar 2. Select Data



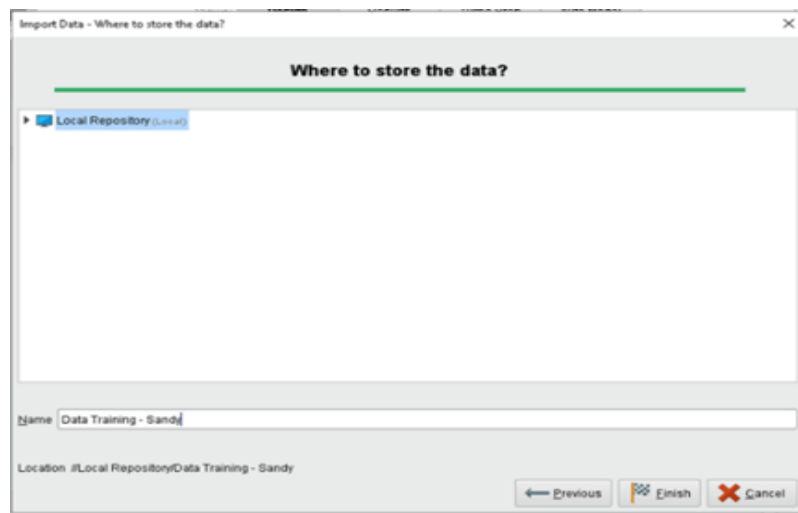
Gambar 3. Menunjukkan Bahwa Tampilan Table Terdapat 100 Data Yang Telah Diimport Ke Dalam Rapidminer



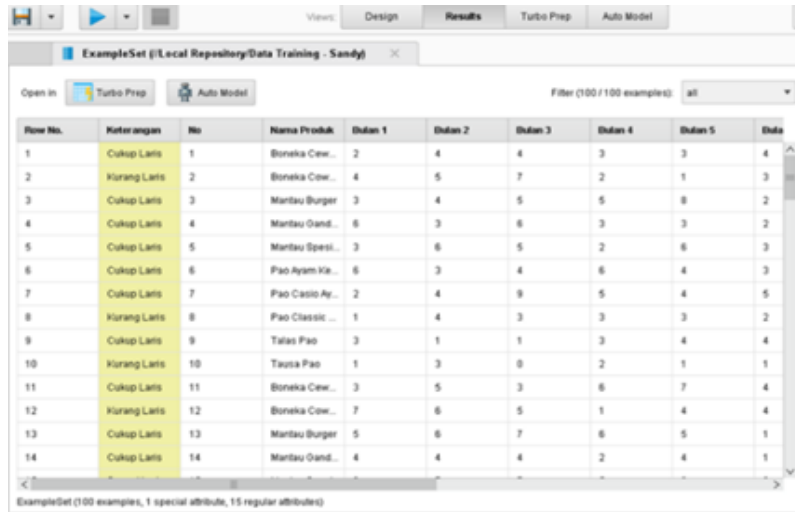
**Gambar 4.** Tampilan *Change Role* Pada *Rapidminer*



**Gambar 5.** Tampilan *Label* Pada *Rapidminer*

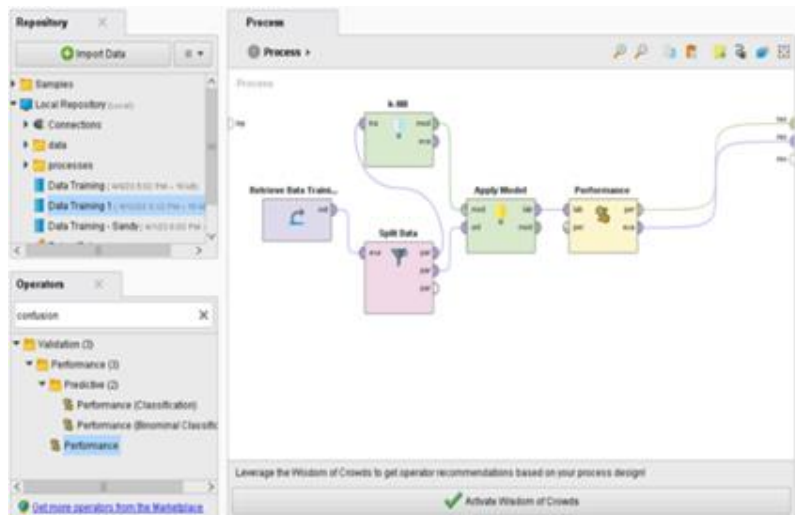


**Gambar 6.** Tampilan Akhir *Import Data*

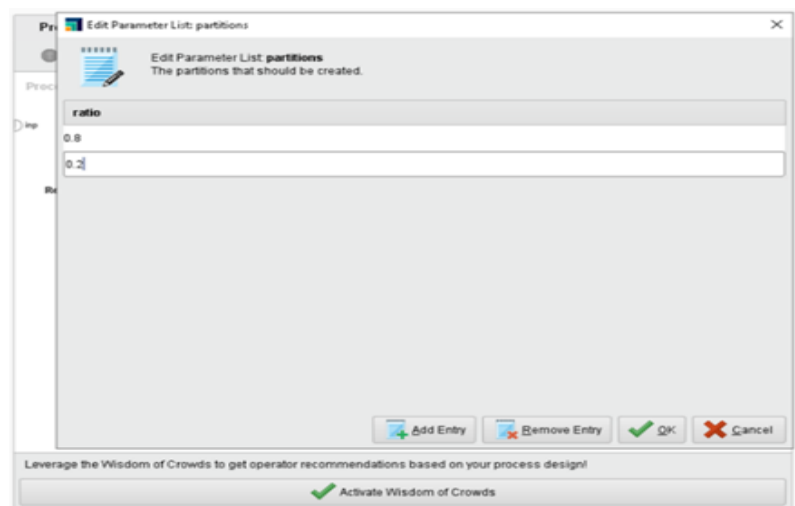


Row No.	Keterangan	No	Nama Produk	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Data
1	Cukup Laris	1	Boneka Cew...	2	4	4	3	3	4
2	Kurang Laris	2	Boneka Cew...	4	5	7	2	1	3
3	Cukup Laris	3	Mantau Burger	3	4	5	5	8	2
4	Cukup Laris	4	Mantau Gand...	6	3	6	3	3	2
5	Cukup Laris	5	Mantau Spesi...	3	6	5	2	6	3
6	Cukup Laris	6	Pao Ayam Ka...	6	3	4	6	4	3
7	Cukup Laris	7	Pao Classic Ar...	2	4	9	5	4	5
8	Kurang Laris	8	Pao Classic ...	1	4	3	3	3	2
9	Cukup Laris	9	Talao Pao	3	1	1	3	4	4
10	Kurang Laris	10	Taurus Pao	1	3	0	2	1	1
11	Cukup Laris	11	Boneka Cew...	3	5	3	6	7	4
12	Kurang Laris	12	Boneka Cew...	7	6	5	1	4	4
13	Cukup Laris	13	Mantau Burger	5	6	7	6	5	1
14	Cukup Laris	14	Mantau Gand...	4	4	4	2	4	1

**Gambar 7.** Tampilan Hasil *Import Data*

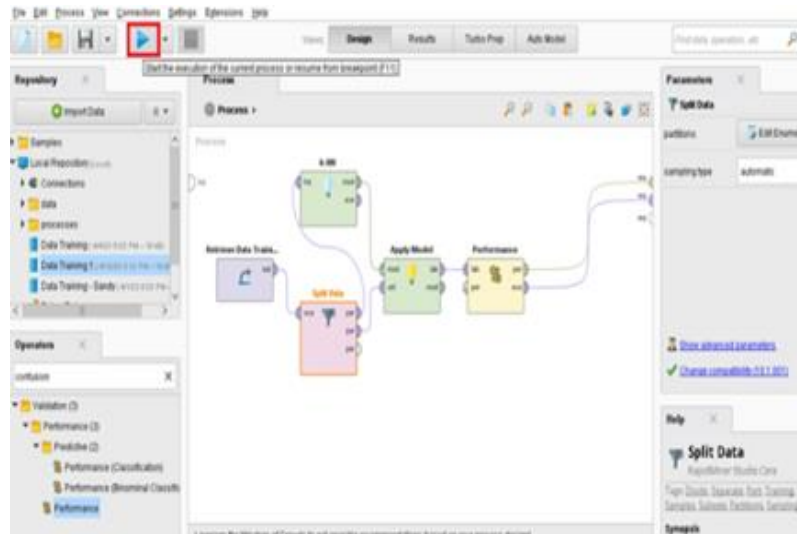


**Gambar 8.** Urutan Operator

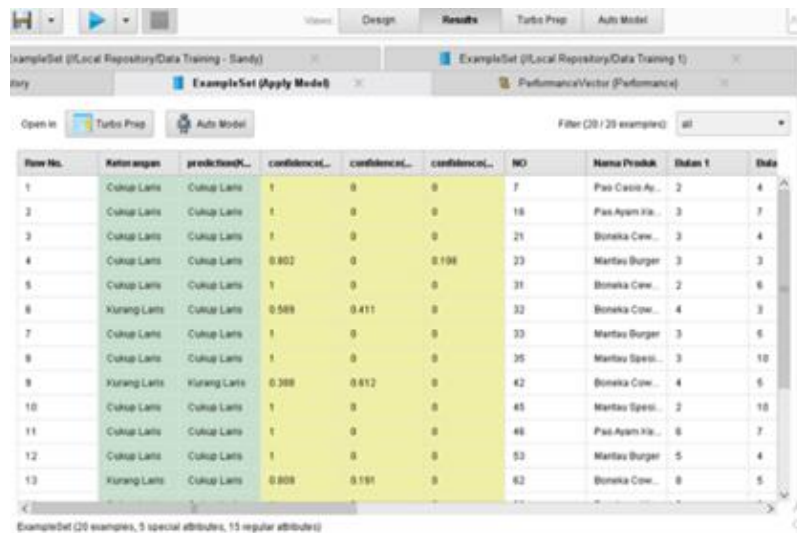


The dialog box is titled 'Edit Parameter List: partitions'. It contains a text input field with the value '0.2'. Below the input field are buttons for '+ Add Entry', '- Remove Entry', 'OK', and 'Cancel'. At the bottom, there is a message: 'Leverage the Wisdom of Crowds to get operator recommendations based on your process design!' and a checkbox for 'Activate Wisdom of Crowds' which is checked.

**Gambar 9.** Tahapan *Add Entry*



**Gambar 10.** Tahapan *Run*



Row No.	Keterangan	prediction	confidence	confidence	confidence	NO	Nama Produk	Ekuitas	Ekuitas
1	Cukup Laris	Cukup Laris	1	0	0	7	Pao Cacio Ap...	2	4
2	Cukup Laris	Cukup Laris	1	0	0	18	Pao Ayam Ka...	3	7
3	Cukup Laris	Cukup Laris	1	0	0	21	Boneka Cow...	3	4
4	Cukup Laris	Cukup Laris	0.802	0	0.198	23	Martabak Bur...	3	3
5	Cukup Laris	Cukup Laris	1	0	0	31	Boneka Cow...	2	6
6	Kurang Laris	Cukup Laris	0.589	0.411	0	32	Boneka Cow...	4	3
7	Cukup Laris	Cukup Laris	1	0	0	33	Martabak Bur...	3	5
8	Cukup Laris	Cukup Laris	1	0	0	35	Martabak Spe...	3	10
9	Kurang Laris	Kurang Laris	0.388	0.612	0	42	Boneka Cow...	4	5
10	Cukup Laris	Cukup Laris	1	0	0	45	Martabak Spe...	2	10
11	Cukup Laris	Cukup Laris	1	0	0	46	Pao Ayam Ka...	5	7
12	Cukup Laris	Cukup Laris	1	0	0	53	Martabak Bur...	5	4
13	Kurang Laris	Cukup Laris	0.808	0.191	0	62	Boneka Cow...	8	5

**Gambar 11.** Hasil Prediksi Data

Table View  Plot View

accuracy: 85.00%

	true Cukup Laris	true Kurang Laris	true Sangat Laris	class precision
pred. Cukup Laris	15	3	0	83.33%
pred. Kurang Laris	0	2	0	100.00%
pred. Sangat Laris	0	0	0	0.00%
class recall	100.00%	40.00%	0.00%	

**Gambar 12.** Hasil Prediksi Data Confusion Matrix

## 5. KESIMPULAN

Untuk melakukan prediksi penjualan produk terlaris peneliti menggunakan perhitungan data mining, yang menggunakan teknik klasifikasi dan algoritma K-Nearest Neighbor, dari 10 produk yang terjual didapatkan hasil prediksi pada RapidMiner penjualan yang terlaris dengan pengujian confusion matrix yaitu mantau burger dan pao casio ayam dengan nilai akurasi 85.00% dan pengujian cross validation yaitu mantau burger dan boneka cewe pao dengan nilai akurasi 87.00%. Dengan menerapkan data mining k-nearest neighbor pada toko, pemilik dapat memperoleh informasi potensi produk terlaris secara efisien dan memiliki acuan untuk memproduksi produk cake dan dessert yang paling banyak diminati konsumen.

## REFERENCES

- Alfani P R, A. W., Rozi, F., Sukmana, F., Teknik Informatika, J., AKAKOM Yogyakarta Jl Raya Janti No, S., Teknologi Informasi, J., & Bhinneka PGRI Jl Mayor Sujadi Timur No, U. (n.d.). *PREDIKSI PENJUALAN PRODUK UNILEVER MENGGUNAKAN METODE K-NEAREST NEIGHBOR*.
- Baharuddin, M. M., Azis, H., & Hasanuddin, T. (2019). ANALISIS PERFORMA METODE K-NEAREST NEIGHBOR UNTUK IDENTIFIKASI JENIS KACA. *ILKOM Jurnal Ilmiah*, 11(3), 269–274. <https://doi.org/10.33096/ilkom.v11i3.489.269-274>
- Budiman1, D. F., Yuwono, E. C., & Yusuf, V. (n.d.). *PERANCANGAN DESAIN KEMASAN PRODUK TOKO KUE KERING “DEWI” SEMARANG*.
- Ginanjari Mabrur, A., & Lubis, R. (2012). PENERAPAN DATA MINING UNTUK MEMREDIKSI KRITERIA NASABAH KREDIT. In *Jurnal Komputer dan Informatika (KOMPUTA) 53 Edisi. I*.
- Handoko, D., Tambunan, H. S., & Hardinata, J. T. (2021). Analisis Penjualan Produk Paket Kuota Internet Dengan Metode K-Nearest Neighbor. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 6(1), 111. <https://doi.org/10.30645/jurasik.v6i1.275>
- Harun, R., Chandra Pelangi, K., & Lasena, Y. (2020). PENERAPAN DATA MINING UNTUK MENENTUKAN POTENSI HUJAN HARIAN DENGAN MENGGUNAKAN ALGORITMA K NEAREST NEIGHBOR (KNN). In *Jurnal Manajemen informatika & Sistem Informasi* (Vol. 3). Retrieved from Online website: <http://e-journal.stmiklombok.ac.id/index.php/misi>
- Junaedi, I., Nuswantari, N., & Yasin, V. (n.d.). *UNTUK DATA MINING ANALISIS TINGKAT RISIKO KEMATIAN NEONATUM PADA BAYI*. Retrieved from <http://journal.stmikjayakarta.ac.id/index.php/jisicomTelp.+62-21-3905050>,
- Karsito, & Monika Sari, W. (2018). Prediksi Potensi Penjualan Produk Delifrance Dengan Metode Naive Bayes Di Pt. Pangan Lestari. *Jurnal Teknologi Pelita Bangsa*, 9(1), 67–78. Retrieved from <https://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/465>
- Meliala, D. M., & Hasugian, P. (2020). Perbandingan Algoritma K-Nearest Neighbor Dengan Decision Tree Dalam Memprediksi Penjualan Makanan Hewan Peliharaan Di Petshop Dore Vet Clinic. *Respati*, 15(3), 35. <https://doi.org/10.35842/jtir.v15i3.369>
- Nasution, T. (2020). Implementasi Algoritma K-Nearest Neighbor Untuk Penentuan Kelulusan Mahasiswa Tepat Waktu. *Jurnal Perangkat Lunak*, 2(1), 1–14. <https://doi.org/10.32520/jupel.v2i1.944>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Nofitri, R., & Irawati, N. (2019). ANALISIS DATA HASIL KEUNTUNGAN MENGGUNAKAN SOFTWARE RAPIDMINER. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 5(2), 199–204. <https://doi.org/10.33330/jurtek.v5i2.365>
- Puspita Hidayanti, W. (2020). Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada “Lombok Vape On.” *Jurnal Informatika Dan Teknologi*, 3(2).



- Riyanto, U. (2018). *ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE DALAM MENGLASIFIKASIKAN JUMLAH PEMBACA ARTIKEL ONLINE*.
- Tempola, F., Muhammad, M., & Khairan, A. (2018). *PERBANDINGAN KLASIFIKASI ANTARA KNN DAN NAIVE BAYES PADA PENENTUAN STATUS GUNUNG BERAPI DENGAN K-FOLD CROSS VALIDATION COMPARISON OF CLASSIFICATION BETWEEN KNN AND NAIVE BAYES AT THE DETERMINATION OF THE VOLCANIC STATUS WITH K-FOLD CROSS VALIDATION*. 5(5), 577–584. <https://doi.org/10.25126/jtiik20185983>
- Tumanggor, A., & Hasugian, S. (2021). Penerapan Data Mining Untuk Memprediksi Tingkat Kemampuan Anak Dalam Mengikuti Mata Pelajaran Dengan Metode C4.5 Pada SDN 105351 Bakaran Batu. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 4(1).