

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Desain Penelitian**

Penelitian ini menggunakan jenis penelitian eksperimen, dengan tahapan penelitian sebagai berikut:

##### **1. Pengumpulan Data**

Pengumpulan data merupakan langkah awal pada suatu penelitian. Data yang digunakan pada penelitian ini adalah data penjualan supermarket.

##### **2. Pengolahan Data Awal**

Pengolahan awal (*Preprocessing*) merupakan tahap untuk mempersiapkan data yang telah diperoleh dari tahap pengumpulan data, yang akan digunakan pada tahap selanjutnya.

##### **3. Desain Eksperimen**

Tahapan ini akan membahas desain eksperimen yang digunakan pada penelitian.

##### **4. Eksperimen dan Pengujian**

Tahapan ini akan membahas tahapan penelitian dan teknik pengujian yang akan digunakan.

##### **5. Evaluasi Penelitian**

Tahapan ini akan membahas hasil evaluasi dari eksperimen yang telah digunakan.

#### **3.2 Pengumpulan Data**

Teknik pengumpulan data dilakukan secara kuantitatif, dengan mengambil *dataset* analisis keranjang pasar dari <http://inf.abdn.ac.uk/~hnguyen/teaching/CS5553/prac05.php>. Dalam *dataset* ini mendefinisikan transaksi penjualan pada sebuah supermarket. Terdapat 304 atribut yang 303 merupakan *item*/barang dan salah satunya merupakan ID transaksi penjualan. Terdiri dari 1361 record transaksi.

Tabel 3.2.1 Contoh Dataset

| Basket ID | Lemons | Standard coffee | Frozen Chicken Wings | 98pct. Fat Free Hamburger | Sugar Cookies | Onions |
|-----------|--------|-----------------|----------------------|---------------------------|---------------|--------|
| C11867    | false  | false           | true                 | true                      | false         | false  |
| C5096     | false  | false           | false                | true                      | false         | true   |
| C4295     | false  | false           | false                | false                     | false         | true   |
| C2837     | true   | false           | false                | false                     | true          | false  |
| C2693     | true   | false           | true                 | false                     | false         | true   |
| C3497     | false  | true            | false                | true                      | false         | false  |

**Keterangan:**

1. **Basket ID** merupakan ID Transaksi penjualan barang.
2. **Lemons, Standard coffee, Frozen Chicken Wings, 98pct. Fat Free Hamburger, Sugar Cookies, dan Onions** merupakan barang-barang yang dijual di Supermarket.
3. **False** menandakan bahwa barang tersebut tidak dibeli oleh pembeli.
4. **True** menandakan bahwa barang tersebut dibeli oleh pembeli.

**3.3 Pengolahan Data Awal**

Data yang sudah dikumpulkan akan diolah dengan beberapa tahap sehingga menjadi sebuah dataset yang akan diimplementasikan terhadap metode-metode data mining. Semua *record* transaksi yang ada pada dataset penjualan digunakan dalam tahap pengolahan data. Tahapan pengolahan data yang dilakukan yaitu:

**3.3.1 Data Preprocessing**

Data *Preprocessing* yaitu dilakukan pembersihan dan persiapan data untuk menghilangkan kosistensi data, data tidak lengkap dan redundant data yang terdapat pada data awal. Data *Preprocessing* juga melakukan pengubahan status barang yang semula *true / false* menjadi 1 / 0, yang akan digunakan untuk proses asosiasi.

**3.3.2 Menghitung Jumlah Barang yang Dibeli / Tidak Dibeli**

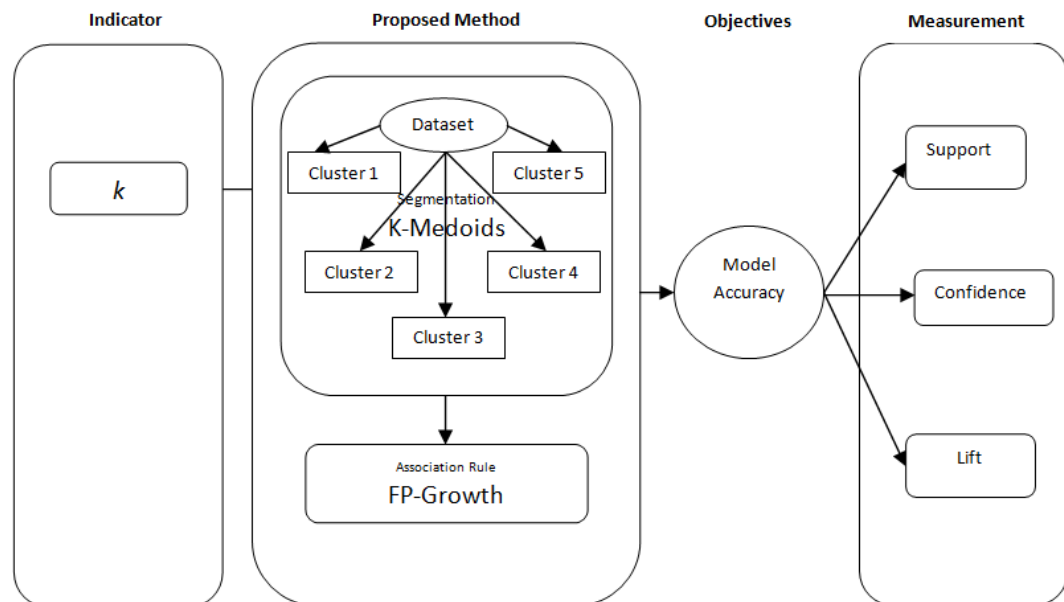
Penghitungan jumlah barang yang dibeli maupun yang tidak dibeli dilakukan dengan menambah 2 atribut di dalam dataset, kedua atribut tersebut digunakan untuk melakukan proses *clustering* data.

| Basket ID | Jumlah Barang yang dibeli | Jumlah Barang yang tak dibeli | Hair Conditioner | Lemons | Standard coffee |
|-----------|---------------------------|-------------------------------|------------------|--------|-----------------|
| C11867    | 1                         | 302                           | 0                | 0      | 0               |
| C5096     | 2                         | 301                           | 0                | 0      | 0               |
| C4295     | 1                         | 302                           | 0                | 0      | 0               |
| C2837     | 1                         | 302                           | 0                | 0      | 0               |
| C2693     | 1                         | 302                           | 0                | 0      | 0               |
| C3497     | 1                         | 302                           | 0                | 0      | 0               |
| C2696     | 1                         | 302                           | 0                | 0      | 0               |
| C1895     | 17                        | 286                           | 0                | 0      | 0               |
| C9524     | 1                         | 302                           | 0                | 0      | 0               |

**Gambar 3.3.1 Menghitung Jumlah Beli / Tidak beli**

### 3.4 Desain Eksperimen

Eksperimen pada penelitian ini adalah penggunaan *K-Medoids* untuk *clustering* data dan *FP-Growth* untuk proses asosiasi.



**Gambar 3.4.1 Desain Eksperimen**

Eksperimen menggunakan algoritma *K-Medoids* untuk *clustering* pada dataset dan menerapkan algoritma *FP-Growth* untuk pendekatan asosiasi pada setiap *cluster*. Sebelum data yang sudah dikumpulkan diolah untuk menemukan pola asosiasi dengan menggunakan *FP-Growth*, terlebih dahulu dilakukan proses *clustering* data dengan menggunakan *K-Medoids*. *Clustering* data dilakukan dengan cara membagi data yang ada menjadi 5 *cluster* ( $k$ ) berdasarkan jumlah barang yang dibeli maupun

yang tidak dibeli. Tujuan dari *clustering* adalah agar data yang di asosiasi menjadi lebih kecil sehingga pola yang dihasilkan dapat lebih akurat.

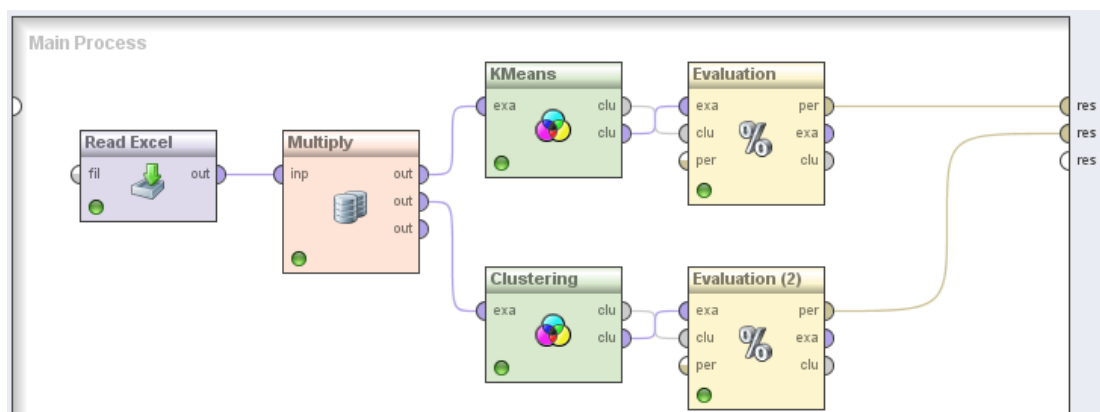
Masing-masing dari *cluster* yang terbentuk akan dilakukan proses asosiasi menggunakan algoritma *FP-Growth* untuk menentukan rekomendasi produk kepada pelanggan. Hasil dari proses asosiasi ini diukur menggunakan *Support*, *Confidence*, *Lift Ratio*. Akurasi dari penerapan algoritma *FP-Growth* yang di *clustering* dahulu menggunakan *K-Medoids* akan dibandingkan dengan akurasi dari penerapan algoritma *FP-Growth* saja.

### 3.5 Eksperimen dan Pengujian dengan *RapidMiner*

Dalam melakukan penelitian ini diperlukan eksperimen dan proses pengujian model yang diusulkan. Proses eksperimen dan pengujian menggunakan bagian dari dataset yang ada. Sebelum eksperimen dilakukan, dataset yang ada harus sudah di *clustering* menjadi beberapa *cluster*. Setiap *cluster* yang dihasilkan dari proses *clustering* akan dilakukan asosiasi dengan *FP-Growth* untuk menghasilkan rekomendasi produk.

#### 3.5.1 Perbandingan *K-Medoids* dan *K-Means*

*Cluster validity index* merupakan fungsi statistik yang digunakan untuk mengevaluasi hasil dari algoritma clustering secara kuantitatif [22]. Secara umum, *cluster validity index* bertujuan untuk menentukan apakah jumlah kelas yang terbentuk optimal atau tidak. Pada penelitian ini menggunakan *Davies-Bouldin index*.



**Gambar 3.5.1 Evaluasi *K-Medoids* dan *K-Means***

*Davies-Bouldin index* mengukur rata-rata kemiripan antar *cluster* dan salah satu yang paling mirip. Semakin rendah nilai *Davies-Bouldin index*, maka cluster tersebut semakin baik [22].

## Davies Bouldin

Davies Bouldin:  $\infty$

Gambar 3.5.2 *Davies-Bouldin index K-Means*

## Davies Bouldin

Davies Bouldin: -1.032

Gambar 3.5.3 *Davies-Bouldin index K-Medoids*

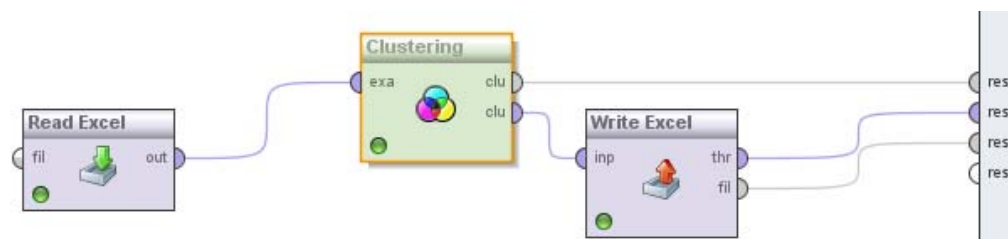
### 3.5.2 Clustering Data Penjualan

Dataset diambil dari data penjualan dalam bentuk excel diimport ke dalam rapid miner.

|                                     |                                     |                                     |                          |                          |                          |                          |                          |
|-------------------------------------|-------------------------------------|-------------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Basket ID                           | Jumlah Bara                         | Jumlah Bara                         | Hair Conditi             | Lemons                   | Standard cc              | Frozen Chic              | 98pct. Fat F             |
| polyno...                           | integer                             | integer                             | integer                  | integer                  | integer                  | integer                  | integer                  |
| attribute                           | attribute                           | attribute                           | attribute                | attribute                | attribute                | attribute                | attribute                |
| C11867                              | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C5096                               | 2                                   | 301                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C4295                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C2837                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C2693                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C3497                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C2696                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C1895                               | 17                                  | 286                                 | 0                        | 0                        | 0                        | 0                        | 1                        |
| C9524                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C9527                               | 1                                   | 302                                 | 0                        | 0                        | 0                        | 0                        | 0                        |
| C8726                               | 2                                   | 301                                 | 0                        | 0                        | 0                        | 0                        | 0                        |

Gambar 3.5.4 *Import Data Penjualan*

Kemudian data yang sudah diimport dilakukan proses *clustering* dengan membagi data tersebut menjadi 5 *cluster* (k). Hasil dari proses *clustering* tersebut akan di ekspor dalam bentuk file excel.



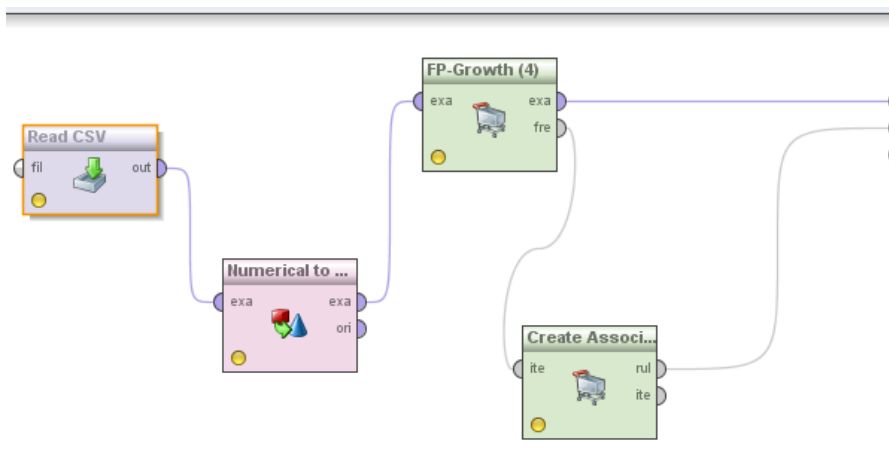
**Gambar 3.5.5 Proses *Clustering* Data**

| Row No. | Basket ID | cluster   | Jumlah Barang yang dibeli | Jumlah Barang yang tak dibeli |
|---------|-----------|-----------|---------------------------|-------------------------------|
| 1       | C11867    | cluster_0 | 1                         | 302                           |
| 2       | C5096     | cluster_0 | 2                         | 301                           |
| 3       | C4295     | cluster_0 | 1                         | 302                           |
| 4       | C2837     | cluster_0 | 1                         | 302                           |
| 5       | C2693     | cluster_0 | 1                         | 302                           |
| 6       | C3497     | cluster_0 | 1                         | 302                           |
| 7       | C2696     | cluster_0 | 1                         | 302                           |
| 8       | C1895     | cluster_3 | 17                        | 286                           |
| 9       | C9524     | cluster_0 | 1                         | 302                           |
| 10      | C9527     | cluster_0 | 1                         | 302                           |
| 11      | C8726     | cluster_0 | 2                         | 301                           |
| 12      | C9383     | cluster_0 | 2                         | 301                           |
| 13      | C8582     | cluster_3 | 18                        | 285                           |
| 14      | C10133    | cluster_3 | 16                        | 287                           |
| 15      | C6983     | cluster_0 | 1                         | 302                           |
| 16      | C3641     | cluster_1 | 3                         | 300                           |
| 17      | C5246     | cluster_0 | 1                         | 302                           |
| 18      | C10136    | cluster_0 | 1                         | 302                           |
| 19      | C7787     | cluster_0 | 1                         | 302                           |

**Gambar 3.5.6 Hasil *Clustering* Data**

### 3.5.3 Asosiasi *Cluster* Data Penjualan

Setiap *cluster* akan dilakukan asosiasi dengan *FP-Growth* untuk menghasilkan rekomendasi produk. Setelah proses *FP-Growth* untuk menghasilkan rekomendasi produk maka dilanjutkan dengan proses *Create Association Rule*.



**Gambar 3.5.7 Proses Asosiasi *FP-Growth***

| Size | Support | Item 1                     | Item 2                     | Item 3           |
|------|---------|----------------------------|----------------------------|------------------|
| 1    | 0.994   | Whole Corn                 |                            |                  |
| 1    | 0.994   | Orange Flavored Fruit Bars |                            |                  |
| 1    | 0.970   | Hair Conditioner           |                            |                  |
| 2    | 0.994   | Whole Corn                 | Orange Flavored Fruit Bars |                  |
| 2    | 0.970   | Whole Corn                 | Hair Conditioner           |                  |
| 2    | 0.970   | Orange Flavored Fruit Bars | Hair Conditioner           |                  |
| 3    | 0.970   | Whole Corn                 | Orange Flavored Fruit Bars | Hair Conditioner |

**Gambar 3.5.8 Frequent Itemset beserta nilai *Support***

```

Association Rules
[Whole Corn] --> [Hair Conditioner] (confidence: 0.976)
[Orange Flavored Fruit Bars] --> [Hair Conditioner] (confidence: 0.976)
[Whole Corn] --> [Orange Flavored Fruit Bars, Hair Conditioner] (confidence: 0.976)
[Orange Flavored Fruit Bars] --> [Whole Corn, Hair Conditioner] (confidence: 0.976)
[Whole Corn, Orange Flavored Fruit Bars] --> [Hair Conditioner] (confidence: 0.976)
[Whole Corn] --> [Orange Flavored Fruit Bars] (confidence: 1.000)
[Orange Flavored Fruit Bars] --> [Whole Corn] (confidence: 1.000)
[Hair Conditioner] --> [Whole Corn] (confidence: 1.000)
[Hair Conditioner] --> [Orange Flavored Fruit Bars] (confidence: 1.000)
[Hair Conditioner] --> [Whole Corn, Orange Flavored Fruit Bars] (confidence: 1.000)
[Whole Corn, Hair Conditioner] --> [Orange Flavored Fruit Bars] (confidence: 1.000)
[Orange Flavored Fruit Bars, Hair Conditioner] --> [Whole Corn] (confidence: 1.000)

```

**Gambar 3.5.9 Rule yang dihasilkan beserta nilai *Confidence***

### 3.6 Evaluasi Penelitian

Hasil akurasi untuk menentukan rekomendasi produk kepada pelanggan yang didapat dari penelitian ini dapat diukur dengan menggunakan *Lift Ratio*. *Lift Ratio* merupakan nilai yang menunjukkan kevalidan proses transaksi dan memberikan informasi apakah benar produk A dibeli bersamaan dengan produk B. Sebuah transaksi dikatakan valid jika mempunyai nilai *Lift Ratio* lebih dari 1, yang berarti

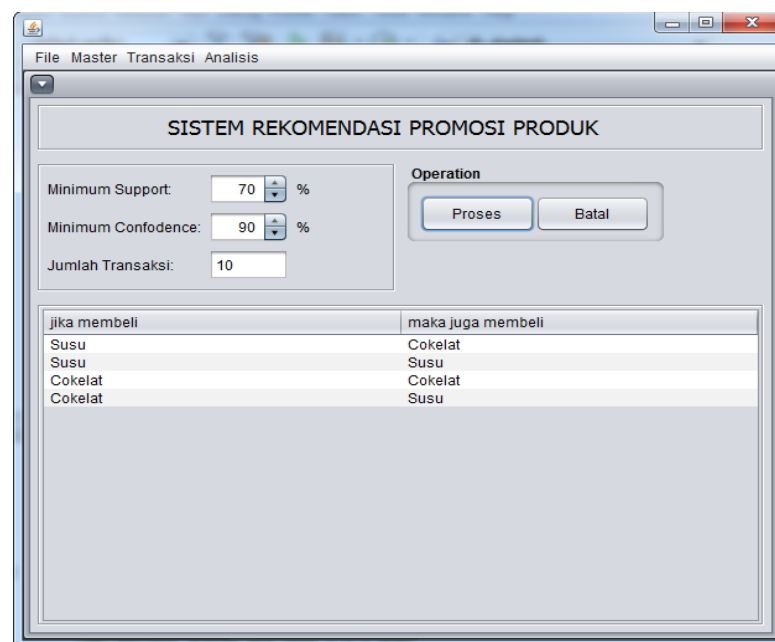
bahwa dalam transaksi tersebut, produk A dan B benar-benar dibeli secara bersamaan [12]. *Lift Ratio* mengukur seberapa penting *rule* yang telah terbentuk berdasarkan nilai *support* dan *confidence*.

| No. | Premises                                     | Conclusion                                   | Support | Confidence | Lift  |
|-----|----------------------------------------------|----------------------------------------------|---------|------------|-------|
| 1   | Whole Corn                                   | Hair Conditioner                             | 0.970   | 0.976      | 1.006 |
| 2   | Orange Flavored Fruit Bars                   | Hair Conditioner                             | 0.970   | 0.976      | 1.006 |
| 3   | Whole Corn                                   | Orange Flavored Fruit Bars, Hair Conditioner | 0.970   | 0.976      | 1.006 |
| 4   | Orange Flavored Fruit Bars                   | Whole Corn, Hair Conditioner                 | 0.970   | 0.976      | 1.006 |
| 5   | Whole Corn, Orange Flavored Fruit Bars       | Hair Conditioner                             | 0.970   | 0.976      | 1.006 |
| 6   | Whole Corn                                   | Orange Flavored Fruit Bars                   | 0.994   | 1          | 1.006 |
| 7   | Orange Flavored Fruit Bars                   | Whole Corn                                   | 0.994   | 1          | 1.006 |
| 8   | Hair Conditioner                             | Whole Corn                                   | 0.970   | 1          | 1.006 |
| 9   | Hair Conditioner                             | Orange Flavored Fruit Bars                   | 0.970   | 1          | 1.006 |
| 10  | Hair Conditioner                             | Whole Corn, Orange Flavored Fruit Bars       | 0.970   | 1          | 1.006 |
| 11  | Whole Corn, Hair Conditioner                 | Orange Flavored Fruit Bars                   | 0.970   | 1          | 1.006 |
| 12  | Orange Flavored Fruit Bars, Hair Conditioner | Whole Corn                                   | 0.970   | 1          | 1.006 |

**Gambar 3.6.1** Evaluasi dengan *Support, Confidence, Lift*

### 3.7 Pembangunan Sistem

Tahapan pembangunan sistem adalah tahapan terakhir dari penelitian ini, dimana tahapan pembangunan sistem dilakukan untuk menerapkan hasil eksperimen kedalam sebuah aplikasi yang nantinya dapat dipergunakan pengguna secara umum untuk menentukan rekomendasi produk. Pada tahapan ini dimulai dengan mendisain *input* dan *output* sistem dimana disain ini akan dipergunakan pada aplikasi rekomendasi produk. Sedangkan untuk pemrograman aplikasi rekomendasi produk ini menggunakan bahasa pemrograman *Java*.



**Gambar 3.7.1** Tampilan Sistem Rekomendasi Promosi Produk