

BAB IV

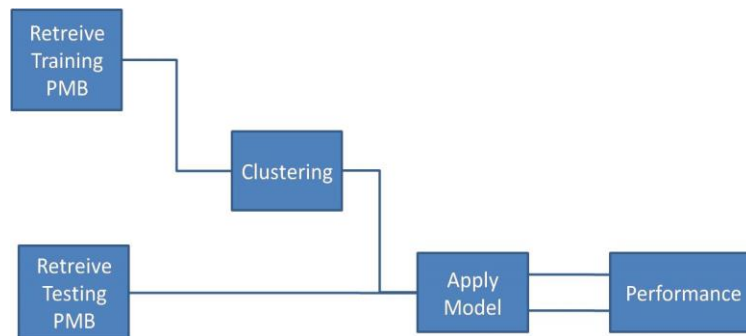
PERANCANGAN SISTEM DAN IMPLEMENTASI

4.1 Perancangan Sistem

Pada sub bab ini akan diuraikan tentang desain model rapidminer

4.1.1 Desain Model

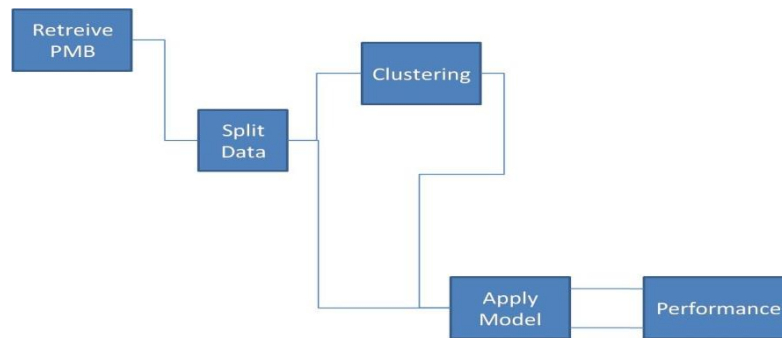
A. Data Normal



Gambar 4.1 Gambar Model Normal

Dengan penggambaran perancangan model clustering k – means. Dengan jumlah data yang di input dari 336 data sampel dibagi menjadi data *training* 90% (302 data) dan data *testing* 10% (34 data), menggunakan metode perhitungan clustering k – means, menggunakan *apply model* untuk menguji data, dan menggunakan *performance* untuk mengukur kinerja sebuah model data tersebut.

B. Split Data



Gambar 4.2 Gambar Desain Model Split Data

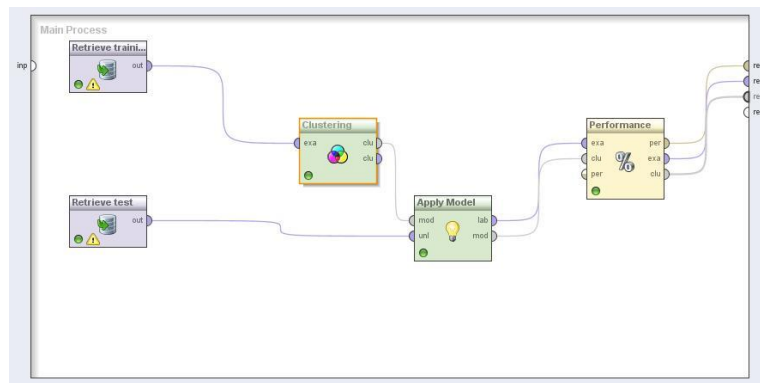
Dengan penggambaran perancangan model clustering k – means. Dengan jumlah data yang di input 336 data, menggunakan *split* data untuk membagi data awal menjadi data *training* dan data *testing*, menggunakan metode perhitungan clustering k – means, menggunakan *apply model* untuk menguji data, menggunakan *performance* untuk mengukur kinerja sebuah model dan data tersebut.

4.2 Implementasi

Pada sub bab ini akan diuraikan tentang *modeling* dan *evaluation* dalam rapidminer.

4.2.1 Model Rapidminer

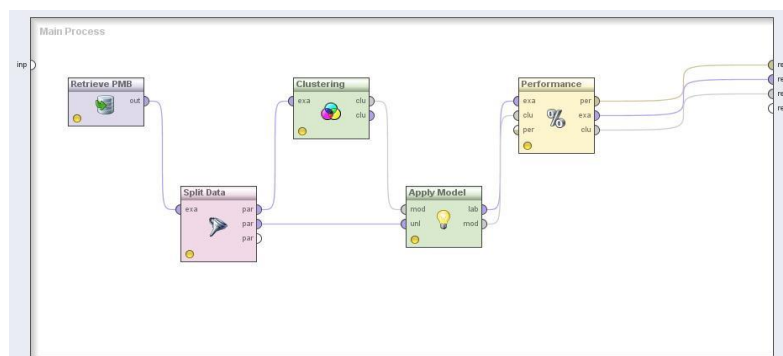
A. Data Normal



Gambar 4.3 Gambar Rapidminer Data Normal

Ini tampilan gambar model clustering k – means. Dengan jumlah data yang di input data *training* dan data *testing*, dan inialisasi jumlah cluster sebanyak 3 (K=3).

B. Split Data



Gambar 4 4 Model Rapidminer Split Data

Ini tampilan gambar model clustering k – means. Dengan jumlah data yang diinput 336 data, data tersebut di split menjadi 90% data *training* dan 10% data *testing* dengan inisialisasi jumlah cluster sebanyak 3 (K=3).

4.2.2 Evaluation

- A. Data Normal
 - a) Cluster Model

```
Cluster Model  
Cluster 0: 116 items  
Cluster 1: 44 items  
Cluster 2: 142 items  
Total number of items: 302
```

Gambar 4.5 Gambar Cluster Model Data Normal

Ini adalah hasil yang didapatkan dari perhitungan clustering k – means (K=3) yaitu pada cluster pertama (0) berjumlah 116 data (*items*), cluster kedua (1) berjumlah 44 data (*items*), cluster ketiga (2) berjumlah 142 data (*items*) dan total semua data berjumlah 302 data.

b) PerformanceVector

```
PerformanceVector  
  
PerformanceVector:  
Avg. within centroid distance: -26.123  
Avg. within centroid distance_cluster_0: -4.799  
Avg. within centroid distance_cluster_1: -68.625  
Avg. within centroid distance_cluster_2: -1.391  
Davies Bouldin: -0.560
```

Gambar 4.6 Gambar PerformanceVector Data Normal

Ini adalah hasil yang didapatkan dari performance sebuah model yang sudah dibuat dapat dilihat dari *davies bouldin* yaitu dengan nilai - 0. 560.

B. Split Data

a) Cluster Model

```
Cluster Model  
  
Cluster 0: 143 items  
Cluster 1: 50 items  
Cluster 2: 109 items  
Total number of items: 302
```

Gambar 4.7 Gambar Cluster Model Split Data

Ini adalah hasil yang didapatkan dari perhitungan clustering k – means (K=3) yaitu pada cluster pertama (0) berjumlah 143 data (*items*), cluster kedua (1) berjumlah 50 data (*items*), cluster ketiga (2) berjumlah 109 data (*items*) dan total semua data berjumlah 302 data.

b) PerformanceVector

PerformanceVector

```
PerformanceVector:  
Avg. within centroid distance: -9.404  
Avg. within centroid distance_cluster_0: -2.676  
Avg. within centroid distance_cluster_1: -29.260  
Avg. within centroid distance_cluster_2: -6.749  
Davies Bouldin: -0.503
```

Gambar 4.8 Gambar PerformanceVector Split Data

Ini adalah hasil yang didapatkan dari performance sebuah model yang sudah dibuat dapat dilihat dari *davies bouldin* yaitu dengan nilai - 0.503.