

# PENERAPAN *DATA MINING* UNTUK MENENTUKAN STRATEGI PENJUALAN PADA TOKO BUKU GRAMEDIA MENGGUNAKAN METODE *CLUSTERING*

Tri Utami Putri<sup>1</sup>, M.Izman H,S.T.,M.M.,PhD<sup>2</sup>, Susan Dian PS M.Kom<sup>3</sup>  
Mahasiswa Universitas Bina Darma<sup>1</sup>, Dosen Universitas Bina Darma<sup>2,3</sup>  
Jalan Jenderal Ahmad Yani No.12 Palembang  
Pos-el : putrie\_utami@ymail.com<sup>1</sup>, herdians1816@gmail.com.<sup>2</sup>,  
susandian@mail.binadarma.ac.id<sup>3</sup>

---

*Abstract: In a business world that is always dynamic and competitive, the perpetrators must always think of ways to continue to survive and if possible develop their business scale. Gramedia Bookstore Palembang is a company engaged in the retail business with the main product book and stationery. This of course lead to competition between companies is very large. Data mining is mining or the discovery of new information by looking for certain patterns or rules from the data set. Data mining can analyze large data into information in the form of a pattern that has meaning for decision support. One data mining technique used is clustering. Clustering is in the use of data mining techniques are used to classify the data based on the data similarity so to then be used as a consideration in determining an effective sales strategy.*

*Keywords: business competition, business decisions, data mining, clustering*

---

*Abstrak :* Dalam rangka menghadapi persaingan dan meningkatkan pendapatan perusahaan, pihak terkait dalam perusahaan tersebut dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Toko Buku Gramedia Palembang merupakan perusahaan yang bergerak dibidang bisnis ritel dengan produk utama buku dan alat-alat tulis. Hal tersebut tentu saja menimbulkan persaingan bisnis antar perusahaan yang sangat besar. *Data mining* merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari kumpulan data tersebut. Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Salah satu Teknik *data mining* yang digunakan adalah *clustering*. *Clustering* merupakan teknik didalam pemanfaatan data mining yang digunakan dengan mengelompokan data-data berdasarkan kemiripan data tersebut. Digunakan untuk membantu menemukan sejumlah aturan dari data transaksi penjualan produk buku pada Toko Buku Gramedia Palembang, sehingga untuk selanjutnya dapat digunakan sebagai pertimbangan dalam menentukan strategi penjualan yang efektif.

**Kata Kunci :** Persaingan bisnis, keputusan bisnis, *data mining, clustering*

## 1. PENDAHULUAN

Dalam rangka menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, pimpinan perusahaan maupun manajemen dalam suatu perusahaan tersebut dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan

membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut.

Pihak eksekutif perusahaan mengharapkan adanya teknologi yang mampu menghasilkan suatu informasi yang siap digunakan untuk membantu mereka dalam mengambil keputusan strategis perusahaan. Mereka ingin mengetahui produk apa yang harus ditingkatkan, seberapa besar pencapaian hasil yang diperoleh oleh perusahaan. Untuk memenuhi kebutuhan

pengusaha di atas, banyak cara yang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan data perusahaan (*Data Mining*).

Ketersediaan data yang banyak dan kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis dan dukungan infrastruktur di bidang teknik informatika merupakan cikal-bakal dari lahirnya teknologi *data mining*. *Data mining* merupakan penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Enur Irdiansyah,2010).

Penggunaan teknik *data mining* diharapkan dapat membantu mempercepat proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola informasi yang terkandung didalam data transaksi menjadi sebuah pengetahuan (*knowledge*) yang baru. Lewat pengetahuan yang didapat, perusahaan dapat meningkatkan pendapatannya dan pada akhirnya di masa yang akan datang perusahaan dapat lebih kompetitif. Salah satu metode yang digunakan yaitu teknik *clustering*. Teknik *clustering* digunakan pada *data mining* untuk mengelompokan objek-objek yang memiliki kemiripan dalam kelas atau segmen yang sama, sementara objek-objek yang terletak pada kelas yang berbeda menunjukkan karakteristik yang berbeda juga (Kurniawan,2007).

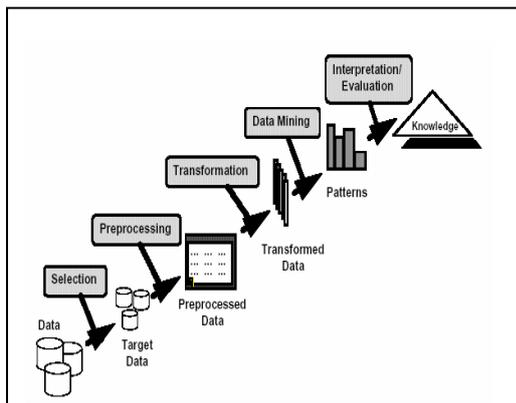
Toko Buku Gramedia merupakan salah satu perusahaan yang bergerak dibidang bisnis ritel dengan produk utama buku dan alat-alat tulis. Dimana perusahaan ini setiap harinya harus

memenuhi kebutuhan konsumen dan dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dianalisis lebih lanjut.

Pada Toko Buku Gramedia, terdapat beberapa permasalahan yang kerap muncul mengenai penjualan buku. Perusahaan sulit mendapatkan informasi-informasi strategis seperti tingkat penjualan buku pelajaran sekolah. Ketersediaan data penjualan yang besar pada Toko Buku Gramedia Palembang tidak digunakan semaksimal mungkin, sehingga data penjualan tersebut tidak dimanfaatkan secara optimal untuk merancang sebuah strategi bisnis dalam meningkatkan penjualan. Melihat kondisi diatas maka penulis tertarik untuk melakukan penerapan *data mining* untuk menentukan strategi penjualan pada Toko Buku Gramedia Palembang.

## 2. METODOLOGI PENELITIAN

Adapun metode untuk menganalisis data dalam penerapan data mining ini adalah dengan menggunakan tahapan *Knowledge Discovery in Database* (KDD) yang terdiri dari beberapa tahapan, yaitu *data selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation*. (Kusrini, 2009:7). Berikut adalah tahapan *Knowledge Discovery in Database* (KDD) yang ditunjukkan pada Gambar 1 .



**Gambar 1. Tahapan *Knowledge Discovery in Database (KDD)***

## 2.1 *Data Selection*

Data yang digunakan dalam penelitian ini berasal dari Toko Buku Gramedia Palembang yaitu data transaksi penjualan buku pelajaran SD,SMP, SMA dan SMK dari tahun 2011, 2012 dan 2013. Dari semua atribut yang ada pada tabel transaksi penjualan tersebut terdapat 6 atribut yang akan digunakan dalam proses *knowledge discovery in databases (kdd)*. Atribut tersebut yaitu:

1. KODE BUKU merupakan atribut yang terdapat pada tabel transaksi penjualan yang menjelaskan tentang informasi mengenai kode buku .
2. KATEGORI merupakan atribut yang terdapat pada tabel transaksi penjualan yang menjelaskan tentang kategori jenis buku tersebut.
3. KELAS merupakan atribut yang terdapat pada tabel transaksi penjualan berdasarkan kategori jenis buku.
4. PENERBIT merupakan atribut yang terdapat pada tabel transaksi penjualan

yang menjelaskan tentang penerbit buku tersebut.

5. TAHUN merupakan atribut yang terdapat pada tabel transaksi penjualan yang menjelaskan tentang tahun penjualan buku tersebut.

6. JUMLAH merupakan atribut yang terdapat pada tabel transaksi penjualan yang menjelaskan jumlah penjualan per tahun.

## 2.2 *Pre-Proseccing*

Pada tahapan *pre-processing* ini akan dilakukan beberapa langkah untuk menghasilkan *dataset* yang bersih sehingga dapat digunakan dalam tahap berikutnya yaitu *mining* dengan tujuan memperoleh pola penjualan yang terjadi terhadap seluruh transaksi penjualan yang telah dilakukan. Berikut langkah-langkah yang dimaksud diatas :

- a. *Integrasi Data*, tahap *integrasi* data adalah tahap penggabungan data dari berbagai sumber. Proses penggabungan dilakukan dengan merelasikan proses integrasi data tersebut dengan *query* seperti yang terlihat pada gambar 4.1 .

### Kode Program 4.1. *Query* Integrasi Data

```
SELECT [2011].[KODEBUKU], [2011].[KATEGORI], [2011].[PENERBIT],
[2011].[JUMLAH 2011], [2012].[JUMLAH 2012], [2013].[JUMLAH 2013]
FROM (2011 INNER JOIN 2012 ON [2012].[KODE BUKU]) INNER JOIN
2013 ON [2012].[KODE BUKU]=[2013].[KODE BUKU];
```

Query diatas menjelaskan bahwa atribut tersebut diambil dari proses penggabungan tiga data *source* yaitu tabel penjualan tahun 2011 , tabel penjualan tahun 2012, dan tabel penjualan tahun 2013. Dimana atribut penghubungnya

adalah Kode Buku . Dari proses tersebut didapatkan sebuah tabel baru yang saya beri nama tabel integrasi seperti di bawah ini.

**Tabel 1 Sebagian Data set Hasil Integrasi**

KODE BUKU	KATEGORI	KELAS	PENERBIT	TAHUN	JUMLAH
BKA001	SD		4 YMARA WIDYA	2011	1788
BKA002	SD		5 ERLANGGA	2011	1870
BKA003	SD		6 ERLANGGA	2011	1885
BKB004	SD		1 ESIS	2011	2152
BKB005	SD		2 ESIS	2011	2107
BKB006	SD		3 ERLANGGA	2011	2142
BKB007	SD		4 ERLANGGA	2011	2433
BKB008	SD		5 ESIS	2011	2213
BKB009	SD		6 ERLANGGA	2011	2354
BKB 010	SD		2 ERLANGGA	2011	2425
BKB 011	SD		2 ESIS	2011	2495
BKB 012	SD		3 ESIS	2011	2475
BKB013	SD		4 GANESHA	2011	2278
BKB014	SD		5 YUDHISTIRA	2011	24444
BKB015	SD		6 YUDHISTIRA	2011	1821
BKB016	SD		2 YUDHISTIRA	2011	1817
BKB017	SD		3 YUDHISTIRA	2011	2117

b. *Data Cleaning* , tahap pembersihan data merupakan tahap awal dari proses KDD. Seluruh atribut pada *dataset* di atas selanjutnya akan diseleksi untuk mendapatkan atribut-atribut yang berisi nilai yang relevan. Tidak *missing value*, dan tidak *redundant*, dikatakan *missing value* jika atribut-atribut dalam *dataset* tidak berisi nilai atau kosong, sementara itu data dikatakan *redundant* jika dalam satu *dataset* yang sama terdapat lebih dari satu *record* yang berisi nilai yang sama. Untuk menampilkan data *missing value* dapat menggunakan *query* seperti yang terlihat pada kode program 3.

```

Kode Program 4.3. Query menampilkan Missing Value

SELECT [2011].[KODEBUKU], [2011].[KATEGORI], [2011].[PENERBIT],
[2011].[JUMLAH 2011], [2012].[JUMLAH 2012], [2013].[JUMLAH 2013]
FROM (2011 INNER JOIN 2012 ON [2011].[KODE BUKU]= [2012].[KODE
BUKU]) INNER JOIN 2013 ON [2012].[KODE BUKU]) INNER JOIN 2013
ON [2012].[KODE BUKU] = [2013].[KODE BUKU] WHERE
(((2011.JUMLAH 2011) IS NULL)) OR (((2012.JUMLAH 2012) IS
NULL))OR (((2013.JUMLAH 2013) IS NULL));

```

Query diatas artinya data yang akan ditampilkan adalah data yang dimana kondisi salah satu dari atribut jumlah 2011, jumlah 2012, dan jumlah 2013 terdapat data yang *NULL* atau kosong. Dari *query* tersebut didapatkan hasil seperti pada tabel 2 dibawah ini.

**Tabel 2 Sebagian Hasil dari Cleaning**

KODE BUKU	KATEGORI	KELAS	PENERBIT	TAHUN	JUMLAH
BKS308	SMK		3 ERLANGGA	2011	
BKS309	SMK		2 PHIBETA	2011	1784
BKA001	SD		4 YMARA WIDYA	2012	1844
BKA002	SD		5 ERLANGGA	2012	
BKA003	SD		6 ERLANGGA	2012	1815
BKB004	SD		1 ESIS	2012	1757
BKB005	SD		2 ESIS	2012	2065
BKB006	SD		3 ERLANGGA	2012	
BKB007	SD		4 ERLANGGA	2012	2237
BKB008	SD		5 ESIS	2012	2400
BKB009	SD		6 ERLANGGA	2012	2299
BKB 010	SD		2 ERLANGGA	2012	2429
BKB 011	SD		2 ESIS	2012	2449
BKB 012	SD		3 ESIS	2012	2482
BKB013	SD		4 GANESHA	2012	2458
BKB014	SD		5 YUDHISTIRA	2012	2429

Pada tabel 2 ditemukan 3 record yang masing-masing pada atribut memiliki data kosong yaitu atribut jumlah 2011, jumlah 2012, dan jumlah 2013 data tersebut harus dibersihkan dengan cara dihapus . Sehingga jumlah *dataset* yang awalnya 311 menjadi 308 record.

### 2.3 Transformasi Data

*Transformasi* data yaitu pada tahap ini data diubah menjadi bentuk yang sesuai untuk di *mining*. Karena dalam penelitian ini hanya dilakukan uji coba secara teoritis. Maka data yang telah di-*cleaning* di *transformasi* menjadi data yang siap di *mining*.

Sebelum melakukan proses *mining* dengan menggunakan aplikasi WEKA maka *dataset* di atas terlebih dahulu di-export kedalam format (.xlsx) kemudian dilakukan perubahan *format file* menjadi (*CSV*) *Comma Delimited* . Setelah selesai dilakukan perubahan *format file* maka data sudah siap untuk ke tahap selanjutnya yaitu tahap *mining*. Berikut *dataset* yang telah di-export ke dalam *format "CSV"*.

**Table 3 Sebagian Dataset dengan format “csv”**

KODE BUKU	KATEGORI	KELAS	PENERBIT	TAHUN	JUMLAH
BKA001	SD	4	YMARA WIDYA	2011	1788
BKA002	SD	5	ERLANGGA	2011	1870
BKA003	SD	6	ERLANGGA	2011	1885
BKB004	SD	1	ESIS	2011	2152
BKB005	SD	2	ESIS	2011	2107
BKB006	SD	3	ERLANGGA	2011	2142
BKB007	SD	4	ERLANGGA	2011	2433
BKB008	SD	5	ESIS	2011	2213
BKB009	SD	6	ERLANGGA	2011	2354
BKB 010	SD	2	ERLANGGA	2011	2425
BKB 011	SD	2	ESIS	2011	2495
BKB 012	SD	3	ESIS	2011	2475
BKB013	SD	4	GANESHA	2011	2278
BKB014	SD	5	YUDHISTIRA	2011	24444
BKB015	SD	6	YUDHISTIRA	2011	1821
RKR016	SD	2	YUDHISTIRA	2011	1817

## 2.4 Data Mining

*Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar (Turban,dkk .2005:3).

Pemilihan teknik dan algoritma yang tepat sangat bergantung pada proses KDD secara keseluruhan. Pada penelitian ini penerapan data mining menggunakan teknik *Clustering* dan algoritma *K-Means*

### 2.4.1 Penerapan *Clustering* dengan Algoritma *K-Means*

Setelah melakukan proses transformasi data kedalam bentuk data yang sesuai untuk penerapan data *mining* dengan teknik *custering* maka tahapan ini dapat dilakukan. *Clustering* merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar dari pada kesamaan tersebut dengan data kelompok lainnya (Santoso,2007).

Dalam tahapan penemuan aturan *clustering* ini, langkah yang dilakukan adalah Penerapan Data Mining Untuk Menentukan Strategi Penjualan Pada Toko Buku Gramedia Palembang Menggunakan Metode Clusterng (Tri Utami Putri )

mengembangkan aturan yang mungkin akan menjadi informasi data penjualan. Data tahapan awal untuk menghasilkan informasi aturan *clustering* dari setiap data transaksi penjualan dihitung dengan *dataset*. *Cluster* secara umum merupakan wujud himpunan bagian dari suatu himpunan data dan metode *clustering* dapat diklasifikasikan berdasarkan himpunan bagian yang dihasilkan : apakah *fuzzy* atau *crisp (hard)*.

### 2.4.2 Algoritma *K-Means*

*K-Means* merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data kedalam kelompok sehingga data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Eko Prasetyo,2012).

Pengelompokan data dengan metode *K-means* ini secara umum dilakukan dengan algoritma sebagai berikut :

1. Tentukan jumlah kelompok.
2. Alokasikan data ke dalam kelompok secara acak.
3. Hitung pusat kelompok (sentroid/rata-rata) dari data yang ada di masing-masing kelompok.
4. Alokasikan masing-masing data ke sentroid / rata-rata terdekat.
5. Kembali ke langkah 3, apabila masih ada data yang berpindah kelompok, atau apabila

perubahan nilai pada fungsi objektif yang digunakan masih diatas nilai ambang yang ditentukan.

Pada contoh ini akan dilakukan pengelompokkan 10 data pada set buku pelajaran SD pengelompokkan dua dimensi. Dimensi datanya dua fitur dengan tujuan memudahkan dalam visualisasi kordinat catersius nantinya. Fitur yang digunakan dalam pengelompokkan adalah x dan y. lalu akan dilakukan pengukuran jarak, jumlah kelompok (K) adalah 2. Ambang batas atau threshold (T) yang digunakan untuk perubahan fungsi objektif adalah 0.1. Kondisi awal data dapat dilihat pada tabel 4 .

**Tabel 4 Kelompok 1**

Data Anggota	Fitur X	Fitur Y
2	4	1
5	2	3
8	3	5
9	2	6
M	Jumlah x	Jumlah y
4	11	15
Rata-rata	2.425	2.444

**Tabel 5 Kelompok 2**

Data Anggota	Fitur X	Fitur Y
3	6	1
4	1	2
6	5	4
7	2	5
10	3	8
M	Jumlah x	Jumlah y
5	17	19
Rata-rata	1.817	1.769

Sentroid yang didapat :

Kelompok 1	Fitur x	Fitur y
1	2.425	2.444
2	1.817	1.769
3	2.425	2.444

Fungsi Objektif

$$J \text{ lama} = 0$$

$$J \text{ baru} = 22,4007$$

Jarak data ke sentroid :

Data Ke-i	Fitur x	Fitur y	Kelompok 1	Kelompok 2	Kelompok 3
1	1	1	0.000		
2	4	1			1788
3	6	1		1870	
4	1	2		1885	
5	2	3			2152
6	5	3		2107	
7	2	5		2142	
8	3	5			2433
9	2	6			2260
10	3	8		2291	
			0.000	10.295	8.633

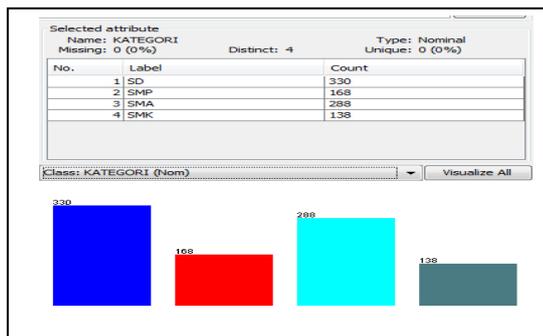
Fungsi Objektif 1.662

### 3. HASIL

#### Proses Data Mining Menggunakan WEKA

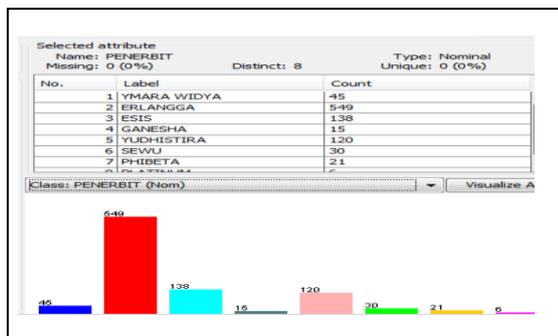
Setelah dijelaskan proses penerapan *data mining* dengan teknik *clustering* secara teoritis pada penjelasan diatas, maka kali ini akan dijelaskan proses *data mining* secara aplikatif dimana proses data mining yang akan dilakukan menggunakan software *data mining WEKA V3.6.10*.

Selanjutnya adalah pemilihan atribut, yaitu atribut kategori terdiri dari 4 atribut yaitu SD,SMP,SMA dan SMK seperti yang ditunjukkan pada Gambar 4 berikut ini.



Gambar 4 Atribut Kategori

Selanjutnya pemilihan atribut penerbit terdiri dari 8 atribut yaitu Ymara Widya , Erlangga, Esis, Ganesh, Yudhistira, Sewu, Phibeta, dan Platinum. Seperti yang ditunjukkan pada gambar 5 .



Gambar 5 Atribut Penerbit

Langkah selanjutnya adalah menentukan metode serta algoritma, namun langkah ini harus disesuaikan dengan kebutuhan dalam pembahasan penelitian. Dalam hal ini teknik yang dipilih adalah teknik *clustering* serta algoritma yang dipilih adalah algoritma *K-means*. Kemudian menentukan parameter jumlah cluster yang ingin kita *clusterisasi* dalam hal ini diubah menjadi 5 cluster. Hasil *clusterer output* seperti ditunjukkan pada gambar 6 .

```
Kode Program 5.13. Hasil Perhitungan Cluster Kategori
Classes to Clusters:
0 1 2 3 4 <-- assigned to cluster
45 56 62 74 90 | SD
138 28 47 40 53 | SMP
288 40 95 72 81 | SMA
138 43 29 47 | SMK

Cluster 0 <-- No class
Cluster 1 <-- SD
Cluster 2 <-- SMP
Cluster 3 <-- SMA
Cluster 4 <-- SMK

Incorrectly clustered instances: 680.0 73.5931 %
```

Kode Program Hasil Cluster Kategori

Dari penjelasan tabel diatas bahwa tabel ini hasil akhir perhitungan *cluster* yang telah diproses dalam data mining. Di dalam tabel mempunyai beberapa atribut yaitu atribut kategori dan penerbit . Berdasarkan *cluster* kategori buku yang paling banyak yaitu buku SD dan SMA yang terdapat pada *cluster* 1 dan *cluster* 3 dengan *Incorrectly clustered instances* : 680.0 73.5931%.

```
Kode Program 5.14. Hasil Perhitungan Cluster Penerbit
Class attribute: PENERBIT
Classes to Clusters:
0 1 2 3 4 <-- assigned to cluster
1 5 8 22 9 | YMAR WIDYA
27 58 147 146 127 | ERLANGGA
27 0 56 43 12 | ESIS
3 0 11 0 1 | GANESHA
21 7 37 27 18 | YUDHISTIRA
0 7 11 8 8 | SEWU
0 2 3 16 0 | PHIBETA
0 0 0 4 2 | PLATINUM

Cluster 0 <-- YUDHISTIRA
Cluster 1 <-- SEWU
Cluster 2 <-- ESIS
Cluster 3 <-- ERLANGGA
Cluster 4 <-- YMAR WIDYA

Incorrectly clustered instances: 685.0 74.1342 %
```

Kode Program Hasil Cluster Penerbit

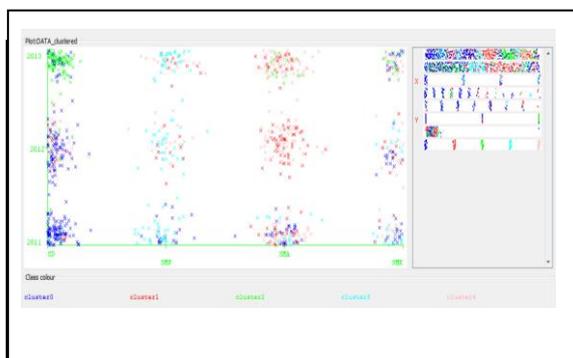
*cluster* berdasarkan penerbit , penerbit buku yang paling banyak yaitu Erlangga dan Eksis

yaitu pada *cluster 2* dan *cluster 3* dengan Incorrectly clustered instances : 685.0 74.1342 %

Dari penjelasan tabel diatas bahwa tabel ini hasil akhir yang telah diproses dalam data mining. Di dalam tabel mempunyai beberapa atribut yaitu atribut kategori dan penerbit . Berdasarkan kategori buku yang paling banyak yaitu buku SD dan SMA yang terdapat pada *cluster 2* dan *cluster 3*. Untuk *cluster* berdasarkan penerbit buku yang paling banyak yaitu Erlangga dan Eksis yaitu pada *cluster 2* dan *cluster 3*.

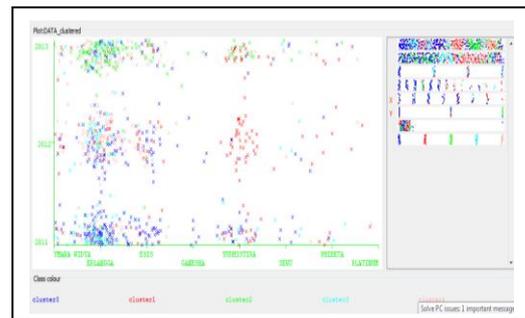
Namun, pada tahapan akhir dari proses Knowledge Discovery in Database (KDD), perlu dilakukan proses evaluation/interpretation. Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

Pada grafik ini merupakan hasil akhir tabel fitted mode yang dihasilkan . Di grafik ini memberikan informasi jumlah tingkat penjualan buku pelajaran di Toko Buku Gramedia Palembang tahun 2011,2012, dan 2013 berdasarkan kategori buku dan penerbit SD,SMP,SMA dan SMK. Seperti yang ditunjukkan pada gambar 8 dibawah ini.



**Gambar 8 Hasil Penjualan Berdasarkan Kategori Buku**

Pada Grafik diatas pada tahun 2011 berdasarkan kategori, buku yang paling banyak terjual yaitu buku SMP dan SMA pada *cluster 1* dan *cluster 2*. Tahun 2012 buku yang paling banyak terjual yaitu buku SMP dan SMK pada *cluster 1* dan *cluster 4* , dan buku yang paling banyak terjual pada tahun 2013 yaitu buku SMA pada *cluster 2*.



X : Penerbit dan Y: Tahun

**Gambar 9 Hasil Penjualan Berdasarkan Penerbit Buku**

Pada Grafik diatas tahun 2011 berdasarkan penerbit, buku yang paling banyak terjual yaitu buku Erlangga dan Esis pada *cluster 0* dan *cluster 3*. Pada tahun 2012 buku yang banyak terjual yaitu buku Erlangga pada *cluster 3* dan buku yang paling banyak terjual tahun 2013 yaitu buku Ymara Widya dan Erlangga pada *cluster 3* dan *4* .

Dari seluruh penjelasan yang telah dipaparkan baik teoritis maupun menggunakan aplikatif *Weka*, seorang manajer dapat mengambil suatu kutupusan yang dilakukan untuk menentukan strategi penjualan yang ada di Toko Buku Gramedia Palembang, karena keputusan yang diambil akan menentukan peningkatan penjualan barang yang diproduksi oleh perusahaan, dari nilai yang sudah di proses dan telah

mendapatkan hasil yang di inginkan maka seorang manajer berperan besar dalam bagian mengatur strategi yang telah di rancang supaya strategi tersebut berjalan dengan baik. Informasi yang akurat dan tepat waktu merupakan informasi yang baik dapat membantu memaksimalkan penjualan perusahaan dan menggunakan sumber daya perusahaan yang langkah dengan efisien, untuk mempersiapkan dan menyesuaikan rencana pemasaran.

#### 4. SIMPULAN

Berdasarkan dari penelitian yang telah dilaksanakan dan sudah diuraikan dalam penerapan *data mining* dari data transaksi penjualan buku tersebut, maka penulis dapat menarik kesimpulan sebagai berikut :

1. Penerapan *data mining* dengan teknik *clustering* dan algoritma *k-means* yang dilakukan menghasilkan sebuah informasi mengenai kategori dan penerbit buku dimana dari informasi tersebut didapat jumlah buku pelajaran yang paling banyak terjual yaitu buku SD berjumlah 7.0577 dan buku SMA berjumlah 5.5713 .
2. Dari hasil diatas maka dapat diketahui penerbit buku yang paling banyak diminati pada Toko Buku Gramedia Palembang yaitu buku pelajaran penerbit Erlangga dan Esis.
3. Semakin kecil batasan nilai data transaksi penjualan yang ditentukan maka semakin sedikit pula pola/aturan yang dapat dihasilkan dan waktu yang diperlukan lebih sedikit.

4. Perhitungan yang dilakukan secara teoritis dan aplikatif menghasilkan nilai data transaksi penjualan yang ditentukan.
5. Waktu yang diperlukan untuk pemrosesan tergantung pada spesifikasi komputer, jumlah transaksi yang diolah dan jumlah item yang terlibat sehingga perumusan yang tepat untuk perhitungan waktu belum dapat ditentukan.

Pada penelitian yang akan datang, ada beberapa hal yang perlu dilakukan agar penelitian ini menjadi lebih baik lagi, yaitu dapat mencoba menggunakan *dataset* penjualan buku yang berbeda seperti buku *fiksi* dan *non fiksi* . Sehingga dapat dihasilkan pengelompokan yang lebih beragam dengan jumlah transaksi yang lebih besar lagi sehingga nilai data transaksi penjualan yang dihasilkan menghasilkan tingkat akurasi yang lebih tinggi dan dapat menggunakan teknik *data mining* dan algoritma lainnya ataupun juga menggabungkan menjadi dua teknik *data mining* , seperti *market basket analysis*, *association rules* sehingga akan menghasilkan hasil yang variatif dan tentunya bernilai informasi yang tinggi.

## DAFTAR RUJUKAN

- Irdiansyah, Enur. (2010). "Penerapan Data Mining Pada Penjualan Produk Minuman Di PT. Pepsi Cola Indobeverages Menggunakan Metode Clustering". Bandung. Universitas Komputer Indonesia.
- Kusrini dan Luthfi. E. Taufiq. (2009). *Algoritma Data Mining*. Yogyakarta: Andi.
- Kurniawan, Mohammad. (2007). "Konsep Penjualan VS Konsep Pemasaran". Semarang.
- Prasetyo, Eko. (2012). *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta : Andi.
- Sugiyono. (2006). *Metode Penelitian Bisnis*. Bandung : CV. Alfabeta.
- Santana, Juandkk. (2010). *Data Mining Meramalkan Bisnis Perusahaan*. Jakarta : PT Elex Media Komputino