

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

**I. Pendahuluan**

**1.1. Latar Belakang**

Seiring dengan pertumbuhan bisnis di era globalisasi dan kemajuan dibidang teknologi informasi yang cepat memberikan pengaruh yang cukup besar baik dalam bidang industri maupun jasa. Hal ini juga membawa suatu perubahan besar dalam tingkat persaingan antara perusahaan, sehingga pelaku-pelaku perusahaan tersebut harus selalu menciptakan berbagai teknik untuk terus *survive*.

Dalam rangka menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, pimpinan perusahaan maupun manajemen dalam suatu perusahaan tersebut di tuntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat di analisis 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-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*. 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.

Toko Buku Gramedia merupakan salah satu perusahaan besar yang bergerak dibidang bisnis ritel dengan produk utama buku dan alat-alat tulis. Dimana perusahaan ini setiap harinya harus memenuhi kebutuhan konsumen dan di tuntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat di analisis lebih lanjut.

Pada Toko Buku Gramedia, terdapat beberapa permasalahan yang kerap muncul mengenai penjualan buku. Perusahaan sulit mendapatkan informasi-informasi strategis seperti tingkat penjualan per periode . Ketersediaan data penjualan yang besar di Toko Buku Gramedia Palembang tidak digunakan semaksimal mungkin, sehingga data penjualan tersebut tidak dimanfaatkan secara optimal dan belum adanya sistem pendukung keputusan dan metode yang dapat

digunakan untuk merancang sebuah strategi bisnis dalam meningkatkan penjualan. Karena sebab itu, maka penulis berkeinginan menuangkan kedalam bentuk skripsi dan memberi judul “**Penerapan Data Mining Untuk Menentukan Strategi Penjualan Pada Toko Buku Gramedia Palembang Dengan Menggunakan Metode Clustering**”.

## **1.2. Perumusan Masalah**

Berdasarkan latar belakang diatas maka permasalahan yang dapat diambil adalah bagaimana cara menggunakan konsep *data mining* dengan menggunakan metode clustering dalam merumuskan strategi penjualan di Toko Buku Gramedia Palembang.

## **1.3. Batasan Masalah**

Adapun ruang lingkup batasan masalah yang dibahas agar lebih terarah dan tidak keluar dari pokok permasalahan adalah cara pemanfaatan data transaksi penjualan buku pelajaran SD,SMP,SMA tahun 2011-2013 dan buku fiksi maupun non fiksi pada Toko Buku Gramedia Palembang.

## **1.4 Tujuan dan Manfaat Penelitian**

### **1.4.1 Tujuan Penelitian**

Tujuan dari penelitian ini adalah membantu memberikan informasi dari data penjualan sebagai solusi dalam meningkatkan penjualan di Toko Buku Gramedia Palembang.

#### **1.4.2 Manfaat Penelitian**

Adapun manfaat penelitian ini adalah sebagai berikut :

- a. Memberikan informasi bagaimana cara pemanfaatan basis data penjualan yang besar, sehingga menjadi *data mining* yang bisa mencari solusi untuk meningkatkan penjualan .
- b. Sebagai bahan masukan untuk mendukung pengambilan keputusan strategi penjualan yang tepat.

## II. TINJAUAN PUSTAKA

### 2.1 Strategi Penjualan

Strategi penjualan merupakan hal yang sangat penting bagi perusahaan dimana strategi penjualan merupakan suatu cara untuk mencapai tujuan dari sebuah perusahaan. Dalam merancang dan mengembangkan produk, baik yang berupa jasa maupun barang, tidak terlepas dari konsep pemasaran yang bertujuan memenuhi kebutuhan yang memuaskan pelanggan. Konsep pemasaran adalah menggambarkan bagaimana sebuah perusahaan harus beroperasi untuk merencanakan , menentukan harga, mempromosikan, dan mendistribusikan barang dan jasa yang bisa memuaskan kebutuhan pembeli dalam mencapai tujuan perusahaan.

Strategi penjualan meyakini bahwa bila konsumen dan pelaku bisnis dibiarkan saja, maka mereka tidak akan membeli dalam jumlah yang memadai produk-produk perusahaan. Perusahaan haruslah melakukan usaha penjualan dan promosi yang memahami segala kebutuhan dan keinginan dari seluruh konsumen dalam ruang lingkup atau batasan sumber daya yang ada. Proses dalam pemenuhan kebutuhan dan keinginan konsumen inilah yang menjadi strategi penjualan yang terdiri dari 4P yaitu :

#### 1. Produk (*product*)

Sebuah produk dirancang untuk memuaskan kebutuhan konsumen.

Strategi produk meliputi sejumlah keputusan tentang kegunaan kualitas,merk dagang kemasan, desain, dan macam model lainnya.

#### 2. Harga (*price*)

Selain harga yang ditetapkan untuk sebuah produk yang di jual kepada konsumen, penentuan harga mencakup beberapa kebijakan manajemen mengenai diskon, harga, dan sebagainya.

### 3. Lokasi (*place*)

Menempatkan produk berarti menyediakan produk pada tempat yang tepat dan waktu yang tepat pula. Strategi penjualan distribusi produk meliputi sejumlah keputusan seperti lokasi dan daerah toko.

### 4. Promosi (*promotion*)

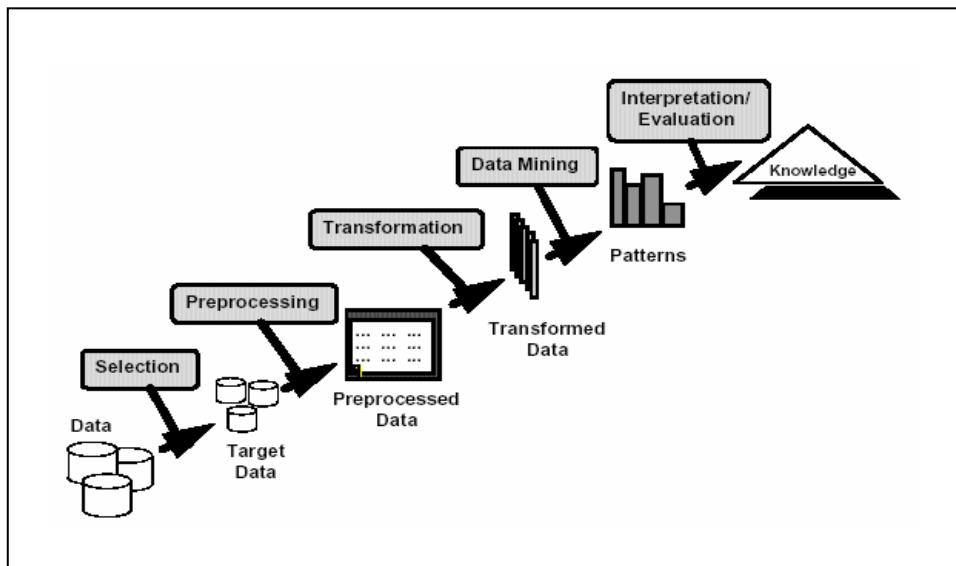
Promosi bermaksud untuk menginformasikan dan membujuk target konsumen dalam hal nilai dari produk yang di jual. Sarana promosi utama adalah melalui iklan, dan penjualan secara besar-besaran.

## **2.2 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).

Istilah data mining dan knowledge discovery in databases (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining*. Proses KDD

(*Knowlegge discovery in databases*) secara garis besar dapat dijelaskan sebagai berikut (Kusrini, 2009:7).



**Gambar 1** Tahapan *Knowledge Discovery in Databases*

### 1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge data discovery* (KDD) dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

### 2. *Preprocessing* atau *Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *knowledge data discovery*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak juga dilakukan proses *enrichment*, yaitu proses memperkaya data

yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi.

### 3. *Transformation*

*Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *knowledge data discovery* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

### 4. *Data mining*

*Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan.

### 5. *Interpretation* atau *evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada pada sebelumnya.

## 2.3 Pengelompokan *Data Mining*

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose ,2005) :

### 1. Deskripsi

Deskripsi adalah menggambarkan pola dan kecenderungan yang terdapat dalam data yang memungkinkan memberikan penjelasan dari suatu pola atau kecenderungan tersebut.

## 2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numeric dari pada kearah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai variabel target sebagai nilai prediksi.

## 3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, akan tetapi dalam prediksi nilai dari hasil akan ada di masa mendatang.

## 4. Klasifikasi

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan memprediksikan kelas untuk data yang tidak diketahui kelasnya.

## 5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.

## 6. Asosiasi

Asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

#### 2.4 *Clustering*

*Clustering* adalah salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok obyek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok obyek lainnya, sehingga obyek yang berada dalam kelompok yang sama relatif lebih homogen dari pada obyek yang berada pada kelompok yang berbeda. Jumlah kelompok yang dapat diidentifikasi tergantung pada banyak dan variasi data obyek. Tujuan dari pengelompokan sekumpulan data obyek kedalam beberapa kelompok yang mempunyai karakteristik tertentu dan dapat dibedakan satu sama lainnya adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan. Model yang diambil diasumsikan bahwa data yang dapat digunakan adalah data yang berupa data interval, frekuensi dan biner.

Dengan menggunakan *clusterisasi*, kita dapat mengidentifikasi daerah yang padat, menemukan pola-pola distribusi secara keseluruhan, dan menemukan keterkaitan yang menarik antara atribut-atribut data. Dalam *data mining*, usaha difokuskan pada metode-metode penemuan untuk cluster pada basis data berukuran besar secara efektif dan efisien. Beberapa kebutuhan *clusterisasi* dalam *data mining* meliputi skalabilitas, kemampuan untuk menangani tipe atribut yang

berbeda, mampu menangani dimensionalitas yang tinggi, menangani data yang mempunyai *noise*, dan dapat diterjemahkan dengan mudah.

Algoritma *clustering* diklasifikasikan menjadi dua algoritma, yaitu *hierarchical clustering*, adalah sebuah metode hierarkis yang menciptakan komposisi hierarkis yang diterapkan pada objek data contohnya *Hierarchical Agglomerative Clustering* (HAC). *Non-hierarchical clustering*, pada umumnya disebut algoritma *partitional clustering*, memberikan sejumlah  $n$  objek dan  $k$  yang merupakan jumlah dari *cluster* yang terbentuk. Algoritma *partitional clustering* mengolah objek ke dalam  $k$ -kelompok berdasarkan kriteria optimasi tertentu, dimana setiap kelompok merupakan representasi sebuah *cluster*. contoh algoritma *partitional clustering* antara lain *K-Means*.

## 2.5 Algoritma *K-Means*

Teknik *clustering* yang akan digunakan adalah algoritma *K-means*. Dalam statistik dan mesin pembelajaran, pengelompokan *K-Means* merupakan metode analisis kelompok yang mengarah pada pemartisian N objek pengamatan ke dalam K kelompok (*cluster*) di mana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat.

*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).

## 2.6 Weka

*Weka* adalah aplikasi data mining open source berbasis Java. Aplikasi ini dikembangkan pertama kali oleh Universitas Waikato di Selandia Baru sebelum menjadi bagian dari Pentaho. Weka terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi / formulasi dari sekumpulan data sampling. Walaupun kekuatan *Weka* terletak pada algoritma yang makin lengkap dan canggih, kesuksesan data mining tetap terletak pada faktor pengetahuan manusia implementornya. Tugas pengumpulan data yang berkualitas tinggi dan pengetahuan pemodelan dan penggunaan algoritma yang tepat diperlukan untuk menjamin keakuratan formulasi yang diharapkan.

*Weka* terdiri dari 4 buah tools yang bisa digunakan yaitu Explorer, Experimenter, Knowledgeflow, dan Simple CLI. Dari tools diatas dapat digunakan untuk menjalankan aplikasi :

1. Explorer digunakan untuk menggali lebih jauh data dengan aplikasi weka.
2. Experimenter digunakan untuk melakukan percobaan dengan pengujian statistic skema belajar.
3. KnowledgeFlow digunakan untuk pengetahuan pendukung.
4. Simple CLI antar muka dengan menggunakan tampilan command-line yang memungkinkan langsung mengeksekusi perintah weka untuk sistem operasi yang tidak menyediakan secara langsung.

## 2.7 Penelitian Sebelumnya

Penelitian sebelumnya digunakan untuk dapat dijadikan bahan pertimbangan dan diharapkan dapat membantu dalam pembuatan sistem yang baru.

Menurut Goldie Gunadi dengan judul penelitian “ **Penerapan Metode Data Mining Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan FP-Growth** ”. *Data mining* merupakan proses analisa data untuk menemukan suatu pola 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 dapat digunakan adalah association data mining atau yang biasa disebut Market Basket Analysis didefinisikan sebagai suatu itemset yang dibeli secara bersamaan oleh pelanggan dalam suatu transaksi. Market basket analysis adalah suatu alat yang ampuh untuk pelaksanaan strategi cross-selling. Metode ini dimulai dengan mencari sejumlah frequent itemset dan dilanjutkan dengan pembentukan aturan-aturan asosiasi (association rules). Algoritma Apriori dan frequent pattern growth (FP-growth) adalah dua algoritma yang sangat populer untuk menemukan sejumlah frequent itemset dari data-data transaksi yang tersimpan dalam basis data. Dalam penelitian ini algoritma Apriori dan frequent pattern growth (FP-growth) digunakan untuk membantu menemukan sejumlah aturan asosiasi dari basis data transaksi penjualan produk buku untuk selanjutnya dapat digunakan sebagai pertimbangan dalam membuat strategi pemasaran dan penjualan yang efektif.

Menurut Bayu Adhi Tama dengan judul penelitian “**Penetapan Strategi Penjualan Menggunakan Association Rules dalam Konteks CRM**” Persaingan usaha telah memaksa perusahaan perdagangan untuk lebih selektif dalam menerapkan strategi pemasarannya. *Customer Relationship Management (CRM)* hadir sebagai sebuah solusi. Bagian CRM seperti *cross/up selling, product affinity analysis*, dan *product bundling* dapat digunakan sebagai sebuah solusi alternatif. Transaksi penjualan dimungkinkan dapat ditingkatkan melalui penjualan *additional products* dari produk utama dimana pelanggan telah berkomitmen untuk membelinya. Paper ini akan melakukan analisis terhadap data transaksi penjualan sebuah perusahaan ritel umum yang bergerak di bidang fotografi yang memiliki variasi produk yang sangat beragam. Keanekaragaman produk ini menghasilkan kemungkinan kombinasi produk yang lebih beragam pula. Hasil penelitian ini menunjukkan bahwa penetapan strategi penjualan dengan menggunakan data pada tengah semester pertama, tren penjualan dari masing-masing kombinasi produk pada semester kedua mengalami peningkatan yang signifikan.

### **BAB III METODELOGI PENELITIAN**

#### **3.1 Waktu dan Tempat**

Penelitian dilakukan di Toko Buku Gramedia Jl.Kolonel Atmo Palembang dari bulan Nopember 2013 sampai dengan bulan Februari 2014 .

#### **3.2 Alat dan Bahan**

Adapun alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut :

##### *1. Hardware*

Kebutuhan perangkat keras (*hardware*) yang digunakan : *Notebook AXIOO D 2500 Processor : Intel® Atom™ cpu D2500@1.86 GHZ, harddisk dengan kapasitas 250 GB .*

##### *2. Software*

Kebutuhan perangkat lunak (*software*) yang digunakan :

- a. Sistem Operasi Windows XP
- b. Microsoft Excel 2007 sebagai pengolah database awal
- c. dan Weka aplikasi yang digunakan untuk menampilkan presentasi knowledge yang dapat disimpulkan dari hasil penerapan model *data mining*.

#### **3.3 Metode Penelitian**

Metode penelitian yang digunakan adalah metode analisis deskriptif dengan pendekatan kuantitatif artinya penelitian yang dilakukan adalah menekankan

analisa pada data-data numeric (angka), yang bertujuan untuk mendapatkan gambaran yang jelas mengenai suatu keadaan berdasarkan data yang diperoleh dengan cara menyajikan, mengumpulkan dan menganalisis data tersebut sehingga menjadi informasi baru yang dapat digunakan untuk menganalisa mengenai masalah yang sedang diteliti.

Menurut Sugiyono (2005:21) “Metode deskriptif adalah metode yang digunakan untuk menggambarkan analisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas”.

### **3.4 Metode Pengumpulan Data**

Dalam penyusunan penelitian ini, untuk mendapatkan data dan informasi yang di butuhkan, maka metode yang digunakan dalam proses pengumpulan data dilakukan sebagai berikut :

#### **1. Studi Pustaka**

Yaitu mempelajari data dengan cara mengumpulkan dan mengkaji penjualan buku selama periode analisis.

#### **2. Wawancara**

Merupakan teknik pengumpulan data dengan cara mengadakan tanya jawab secara langsung kepada supervisor penjualan pada Toko Buku Gramedia Palembang.

### 3.5 Metode Analisis Data

Metode analisis data yang digunakan adalah *clustering*. Adapun untuk menganalisis data dalam penerapan *data mining* ini menggunakan tahapan *Knowledge Discovery in Databases (KDD)* yang terdiri dari beberapa tahapan, (kusrini,2009:7) yaitu :

#### 1. *Data selection*

Pada tahap ini penulis mengumpulkan data-data yang terdapat pada Toko Buku Gramedia Palembang. Data yang diambil berupa data laporan penjualan tahun 2011-2013 yang terdiri dari tabel transaksi penjualan berisi tentang informasi data penjualan yang nantinya akan digunakan dalam penerapan data mining.

#### 2. *Pre-processing*

Sebelum proses *data mining* dapat dilaksanakan perlu dilakukan proses *cleaning*. *Dataset* transaksi penjualan akan diseleksi dari data yang mengandung *missing value* dan data dengan nilai yang *redundant*. Karena seleksi data relevan yang berarti atribut-atributnya ditentukan oleh output yang ingin dihasilkan.

#### 3. *Transformasi data*

Data yang telah di *cleaning* di *transformasi* menjadi data yang siap diminingkan. Karena software yang digunakan untuk data mining yaitu WEKA maka dataset di atas terlebih dahulu di-export kedalam format (.xlsx) kemudian dari format (.xlsx) akan dirubah lagi menjadi format “csv” (Comma Delimited). Karena Software yang digunakan hanya dapat membaca file data

dengan format “csv” (Comma Delimited). Maka data siap untuk dilakukan proses selanjutnya.

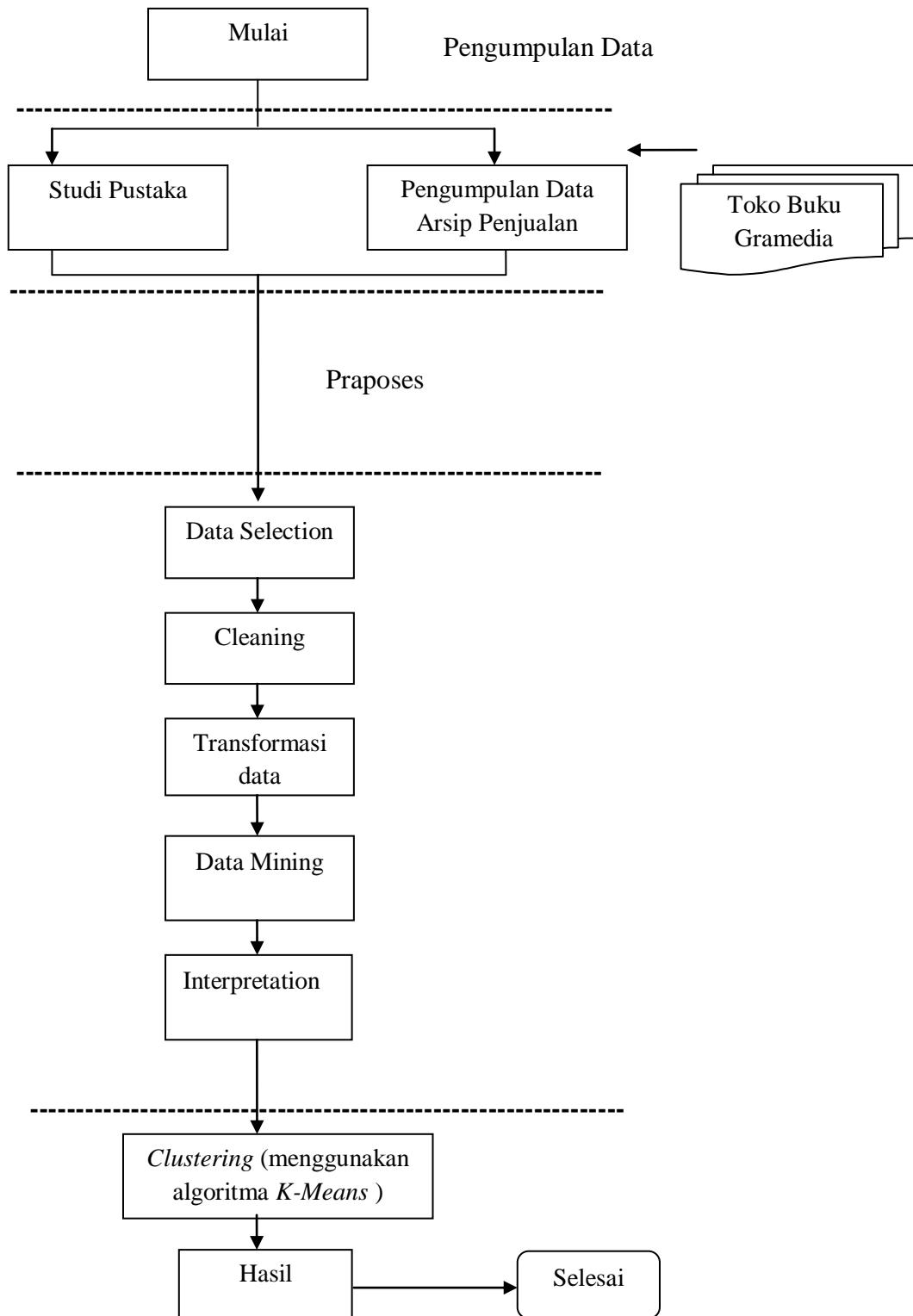
#### 4. *Data Mining*

Merupakan tahapan untuk menemukan pola atau informasi dalam sekumpulan data dengan menggunakan teknik dan algoritma tertentu. Pada penelitian ini penerapan data mining menggunakan teknik *Clustering* dan algoritma *K-Means* untuk menemukan informasi mengenai data transaksi penjualan. Setelah dijelaskan proses penerapan *data mining* dengan teknik *clustering* maka akan dijelaskan proses data mining secara aplikatif dengan menggunakan software *Weka* v3.6.10 yang merupakan data yang telah di transformasi kedalam format *Microsoft Excel 2007 (.CSV)*.

#### 5. *Interpretation / Evalution*

Merupakan pola informasi yang dihasilkan dari proses *data mining*. Informasi yang dihasilkan pada software *Weka* akan menampilkan, dan memberikan informasi tingkat banyaknya penjualan barang yang banyak terjual , dari sinilah seorang manajer dapat mengambil suatu keputusan yang dilakukan untuk menentukan strategi penjualan karena keputusan yang diambil akan menentukan peningkatan penjualan barang yang diproduksi oleh perusahaan.

### 3.6 Tahapan Penelitian



Gambar 4 Tahapan Penelitian

Adapun tahapan penelitian yang akan dilakukan penulis dalam proses penelitian ini adalah sebagai berikut :

1. Studi Pustaka

Dengan mempelajari literatur yang berkaitan dengan konsep data mining menggunakan metode *clustering* dan algoritma *K-Means*. Sumber literatur berupa data penjualan buku selama periode analisis.

2. Pengumpulan Data Arsip

Untuk mengetahui informasi yang dibutuhkan, penulis melakukan pengumpulan data arsip penjualan tahun 2011-2013 .

3. Praposes data meliputi :

a. *Data Selection*

Untuk memilih himpunan data (*dataset*) yang akan digunakan pada penulisan ini yaitu berupa data transaksi penjualan yang berisi tentang informasi penjualan.

b. *Cleaning*

Untuk membersihkan data yaitu, melengkapi data, menghapus data duplikat, menghilangkan *noise*.

c. *Transformasi data*

Untuk memformat data agar bisa dicluster atau bisa dengan menambahkan atribut baru.

d. *Data mining*

Menerapkan teknik *clustering* dan algoritma *K-Means* untuk menemukan informasi mengenai data transaksi penjualan. Dan akan dijelaskan proses *data mining* secara aplikatif dengan software *Weka*.

#### 4. *Clustering menggunakan algoritma Fuzzy C-means*

Tahapan proses dimana data yang sudah dipraproses di *cluster* dengan menggunakan algoritma FCM didasarkan pada minimasi fungsi objektif yang diformulasikan dalam persamaan :

$$J_m = \sum_{i=1}^N \sum_{j=1}^c (U_{ij})^m \|Z_i - C_j\|^2$$

Selanjutnya algoritma FCM disusun dengan langkah sebagai berikut :

Langkah 1 : Tentukan himpunan data Z yaitu (data awal). Tentukan jumlah *cluster* yang diharapkan  $1 < c < N$ , nilai pembobot  $m > 1$ , toleransi penghentian  $\epsilon > 0$ .

Langkah 2 : Inisialisasi matriks partisi secara acak,  $U(0) \in M_{fc}$ .  
Ulangi untuk  $i=3,4,5$

Langkah 3 : Hitung *cluster center (means)*.

$$V_i^{(i)} = \frac{\sum_{k=1}^N (U_{ik}^{(i-1)})^m Z_k}{\sum_{k=1}^N (U_{ik}^{(i-1)})^m}, 1 \leq i \leq c$$

Keterangan :

$U_{ij}$  : merupakan level keanggotaan dari  $Z_i$  dalam *Cluster j*<sup>2</sup>

$Z_i$  : merupakan nilai data ke- $I$  dari d-dimensi data

$C_j$  : merupakan nilai ke- $j$  dari d-dimensi *cluster center*

$M$  : merupakan sembarang bilangan *real* lebih besar dari 1

$N$  : merupakan suatu himpunan dari  $N$  yang dinotasikan dengan  $Z$ (data awal)

Contoh :

Untuk proses awal maka algoritma *k-means* yang ada akan dilakukan sebagai berikut :

Diberikan data sebagai data awal

$$Z = ( 244 \ 247 \ 267 \ 267 \ 215 \ 372 \ 381 \ 369 \ 326 \ 247 \ 216 \ 286 ) \\ 233 \ 254 \ 246 \ 253 \ 277 \ 375 \ 317 \ 375 \ 326 \ 277 \ 233 \ 266$$

Akan dilaksanakan *clustering* dengan FCM.

Langkah 1:

- Jumlah cluster yang diharapkan  $c = 3$
- Nilai pembobot  $m = 3$
- Toleransi penghentian  $\epsilon = 1.10^{-6}$
- Maksimum iterasi = 100

Langkah 2 : Inisialisasi matriks partisi awal secara acak.

$$U = \begin{pmatrix} 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 \\ 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 \\ 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 & 0,5 \end{pmatrix}$$

Langkah 3 : Hitung *cluster center* (*means*)

$$v_1 = \frac{0,5 \times 244 + 0,5 \times 247 + 0,5 \times 267 + 0,5 \times 267 + 0,5 \times 215 + 0,5 \times 372 + 0,5 \times 381 + 0,5 \times 369 + 0,5 \times 326 + 0,5 \times 247 + 0,5 \times 216 + 0,5 \times 286}{3}, 1 \leq i \leq c$$

$$v_2 = \frac{0,5 \times 233 + 0,5 \times 254 + 0,5 \times 246 + 0,5 \times 253 + 0,5 \times 277 + 0,5 \times 375 + 0,5 \times 317 + 0,5 \times 375 + 0,5 \times 326 + 0,5 \times 277 + 0,5 \times 233 + 0,5 \times 266}{3}, 1 \leq i \leq c$$

$$V = \begin{pmatrix} 40,67 & 41,16 & 44,5 & 44,5 & 35,83 & 63,5 & 61,5 & 54,34 & 41,16 & 36,00 & 47,66 \\ 38,84 & 42,34 & 41,00 & 42,16 & 46,17 & 62,5 & 52,83 & 54,34 & 38,84 & 44,34 \end{pmatrix}$$

## Rancangan Database

### Tabel Pengarang

Field	type	ukuran	Keterangan
Kd_pengarang	Varchar	20	PK
Nm_pengarang	Varchar	25	
Alamat_pengarang	Varchar	30	
Email_pengarang	Varchar	10	

### Tabel kategori buku

Field	Type	ukuran	Keterangan
Kd_kategori_buku	Varchar	5	PK
Nm_kategori_buku	Varchar	25	

### Tabel Buku

Field	type	ukuran	Keterangan
Kd_buku	Varchar	20	PK
Judul_buku	Varchar	25	
Stock	Varchar	30	
Harga	Varchar	10	

### Tabel penerbit

Field	type	ukuran	Keterangan
Kd_penerbit	Varchar	20	PK
Nm_penerbit	Varchar	25	
Email_penerbit	Varchar	30	

### Tabel distributor supplier

Field	type	ukuran	Keterangan
Kd_supplier	Varchar	20	PK
Nm_supplier	Varchar	25	
Email_supplier	Varchar	30	

## Daftar Pustaka

- Azwar Saifuddin. (2005). Metode Penelitian. Yogyakarta : Pustaka Pelajar.
- Gunadi,Goldie.(2012)."Penerapan Metode *Data Mining* Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan FP-Growth ".Universitas Budi Luhur.
- Irdiansyah,Enur.(2009)." Penerapan Data Mining Pada Penjualan Produk Minuman Di PT. Pepsi Cola Indobeverages Menggunakan Mertode Clustering". Bandung. Univeritas Komputer Indonesia.
- Kusrini dan Luthfi. E. Taufiq.(2009).Algoritma *Data Mining*. Yogyakarta: Andi. Kanzunnudin,Mohammad.(2009)."Konsep Penjualan VS Konsep Pemasaran". Semarang.
- Nango,Dwi Noviati.(2012)."Penerapan Algoritma *K-Means* Untuk *Clustering* Data Anggaran Pendapatan Belanja Daerah Di Kabupaten XYZ".Universitas Negeri Gorontalo.
- Prasetyo,Eko.(2012).*Data Mining* Konsep dan Aplikasi Menggunakan Matlab. Yogyakarta: Andi.
- Sugiyono. (2006), Metode Penelitian Bisnis. Bandung : CV.Alfabeta.
- Tama,Bayu Adhi.(2010)." Penetapan Strategi Penjualan Menggunakan *Association Rules* dalam Konteks CRM".Universitas Sriwijaya.

