

BAB II
TINJAUAN PUSTAKA

A. Penelitian Terdahulu

| Peneliti(Tahun) | Objek Penelitian | Metode | Hasil |
|-------------------|--|-------------------|---|
| (Sari,2013) | Data transaksi merek pakaian | Algoritma apriori | Mengetahui merek pakaian yang paling diminati |
| (Fajri,2016) | Data hasil <i>Interview</i> | Algoritma apriori | Menghasilkan aturan (<i>rules</i>) yang merupakan kumpulan <i>frequent itemset</i> dengan nilai <i>confidence</i> yang tinggi yang digunakan untuk memilih program studi yang diinginkan. |
| (Failasuf a,2014) | Data transaksi penjualan pada Pamella <i>Supermarket</i> | Algoritma apriori | Mengetahui pola pembelian konsumen pada masing-masing cabang Pamella <i>Supermarket</i> berdasarkan lokasi |
| (Dodi,2021) | Data transaksi penjualan pada Toko | Algoritma apriori | Mengetahui pola pembelian konsumen pada bulan Januari dan |

| | | | |
|--|----------------------------|--|---------------|
| | Bangunan Wahyu Agung | | Februari 2021 |
|--|----------------------------|--|---------------|

Selanjutnya, penelitian yang diusulkan melalui penulis menggunakan seperangkat aturan apriori. Pada tampilan ini, perkembangan data transaksi sebagai obyek. Statistik yang digunakan adalah dari Januari hingga Februari 2021. Penelitian ini diselesaikan dalam beberapa tahapan yang meliputi seleksi record, pembersihan/preprocessing, transformasi, fact mining, & interpretasi/evaluasi biaya. dalam studi ini, Anda akan membuat pembelanjaan pelindung untuk gaya berdasarkan sepenuhnya pada statistik transaksi yang diproses (Kamber, 2006).

B. Data Mining

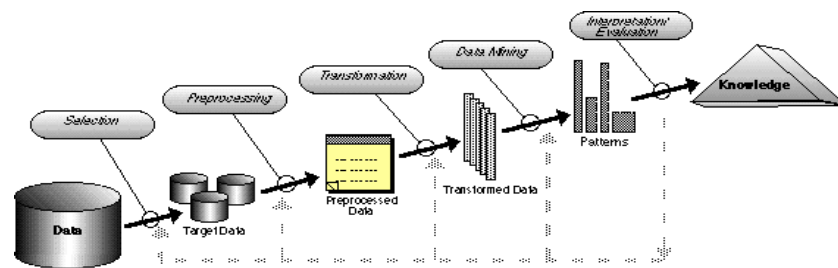
Penambangan informasi adalah penentuan atau "penambangan" banyak informasi (Kamber, 2006). Data mining, juga dikenal sebagai pengenalan pola, adalah metode untuk memproses data, yang menentukan pola tersembunyi dari data yang diproses. Data yang diolah menggunakan teknik data mining ini kemudian menghasilkan wawasan baru dari data lama, dan hasil pengolahan data tersebut dapat digunakan untuk menentukan keputusan di masa mendatang. (Ong, 2013).

Data mining juga dapat diartikan sebagai rangkaian kegiatan untuk menemukan pola-pola menarik dalam jumlah data yang besar, kemudian menyimpan data tersebut dalam database, data warehouse atau penyimpanan data. Ada beberapa ilmu yang mendukung teknik data mining, antara lain analisis data, pemrosesan sinyal, dan jaringan saraf pengenalan pola.(Prilianti & Wijaya, 2014).

Penambangan informasi penting untuk ukuran KDD (*Knowledge Discovery in Database*), yang terdiri dari beberapa fase, seperti penentuan informasi, prapemrosesan, transformasi, penambangan informasi, dan penilaian.. (Last, 2001).

Data mining merupakan kegiatan yang melibatkan pengumpulan & penggunaan data historis buat menemukan hukum, pola, atau interaksi pada perpaduan data yang besar. Salah satu tugas menurut data mining merupakan clustering. Tujuan primer menurut clustering merupakan buat mengelompokkan banyak data/objek ke pada grup-grup sebagai akibatnya setiap grup berisi data sebanyak-banyaknya.(Santosa, 2007).

Metode analisis data menggunakan adalah clustering. Untuk analisis data pada aplikasi data mining, mengadopsi langkah database knowledge discovery (KDD), yang terdiri dari beberapa tahap, (Bidang et al., n.d.)yaitu :



Gambar 2.1 Stages of knowledge discovery in databases

1. Pemilihan data

Sebelum dimulainya tahap ekstraksi data di Knowledge Data Discovery (KDD), diperlukan pemilihan data (seleksi) dari kumpulan data operasional. Simpan data terpilih yang digunakan dalam proses penambangan data dalam sebuah file, terpisah dari database operasional.

2. Preprocessing atau pembersihan

Sebelum proses data mining, perlu dilakukan pembersihan data yang menjadi fokus penemuan data pengetahuan. Proses pembersihan tersebut meliputi namun tidak terbatas pada menghilangkan duplikasi data, pengecekan data yang tidak konsisten, dan koreksi kesalahan pada data, seperti kesalahan pencetakan. Proses pengayaan juga dilakukan yaitu proses pengayaan data atau data yang ada dengan data

yang lain. Data yang berkaitan dengan KDD dan yang diperlukan, seperti data atau data.

3. Konversi

Encoding adalah proses mengubah data yang dipilih untuk membuat data cocok untuk proses data mining. Proses encoding dalam penemuan data pengetahuan adalah proses kreatif, yang sangat tergantung pada jenis atau mode data yang akan dicari dalam database.

4. Data Mining

Data mining adalah proses menggunakan teknik atau metode tertentu untuk menemukan pola atau data yang menarik dalam data yang dipilih. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Memilih metode atau algoritma yang tepat sebenarnya tergantung pada tujuan KDD dan proses secara keseluruhan.

5. Interpretasi atau Evaluasi

Pola data yang dihasilkan dari proses data mining harus ditampilkan dengan cara yang mudah dipahami oleh pemangku kepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretasi. Tahap ini meliputi verifikasi apakah pola data yang ditemukan bertentangan dengan fakta atau asumsi sebelumnya.

C. Pengelompokan Data Mining

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

1. Deskripsi

Deskripsi tersebut meliputi mendeskripsikan pola dan tren yang terdapat pada data, serta dapat menjelaskan pola atau tren tersebut.

2. Estimasi

Estimasi secara praktis setara dengan kapabilitas, namun yang penting adalah penilaian variabel objektif lebih sebagai angka daripada

penilaian. Model ini dibuat dengan memanfaatkan catatan total yang menawarkan manfaat dari variabel objektif sebagai nilai sekarang

3. Prediksi

Prediksi namun hampir mirip dengan kelas dan estimasi, biaya yang diantisipasi akan ada di masa depan

4. Klarifikasi

Klasifikasi adalah cara untuk menemukan versi atau fitur yang menggambarkan dan membedakan kelas catatan atau ide. Tujuannya adalah untuk mengharapkan kategori informasi yang kelasnya tidak diketahui.

5. Pengklusteran

Pengelompokan adalah catatan, pengamatan, atau minat dan dokumen kategori objek yang sebanding. Cluster adalah sekelompok statistik yang mungkin sama seperti setiap statistik yang berbeda namun unik dari statistik di cluster yang berbeda.

6. Asosiasi

Asosiasi dalam penambangan fakta terdiri dari menemukan atribut yang tampak bersama, dalam bisnis global, yang lebih sering disebut analisis keranjang belanja.

D. Definisi matematis jarak Euclidean

Jarak Euclidean adalah metrik yang umum digunakan untuk menghitung kesamaan dua vektor. Semakin besar jarak antara dua vektor, semakin kecil derajat kemiripan atau kemiripannya. Sebaliknya, semakin kecil jarak antara dua vektor, semakin besar derajat kemiripan atau kemiripannya. Desain pengelompokan data menggunakan algoritma k-means berbasis peta panas (Tulus, 2014).

E. Clustering

Clustering merupakan salah satu teknik yang berkaitan dengan pengelompokan objek berdasarkan karakteristik atau kesamaan dalam data mining. Salah satu klaster memiliki derajat kesamaan (homogenitas) yang tinggi dalam karakteristik dan terdapat perbedaan yang besar antar klaster (heterogenitas). Cluster juga termasuk dalam data mining, yang merupakan pembelajaran tanpa pengawasan. (Harrington, 2012).

Pengelompokan akan mengelompokkan data menjadi beberapa kelompok (cluster) berdasarkan kesamaan karakteristik masing-masing data pada kelompok yang ada. (Prasetyo, 2012).

Pengelompokan mengacu pada pengelompokan catatan, pengamatan, atau contoh ke dalam item serupa. Cluster adalah sekumpulan record yang bisa sama seperti satu sama lain tetapi satu-satunya dari record di perusahaan yang berbeda. Perbedaan antara pengelompokan dan peringkat adalah bahwa tidak ada variabel tujuan untuk pengelompokan. Tantangan pengelompokan tidak mencoba mengetik, memperkirakan, atau mengharapkan nilai variabel target. Dalam penilaian, algoritma pengelompokan mencoba untuk melembagakan semua statistik ke dalam subkelompok atau bisnis yang sangat homogen, di mana kesamaan statistik di dalam klaster dimaksimalkan, dan kesamaan data di luar klaster diminimalkan. (Larose, 2005).

F. Algoritma Apriori

Algoritma dasar diusulkan oleh Agrawal dan Sirkait pada tahun 1994 untuk menentukan himpunan aturan asosiasi Boolean yang sering (Sensue, 2012). Algoritme ini mengontrol pertumbuhan kandidat kumpulan elemen frequent dari hasil kumpulan elemen frequent dengan dukungan berbasis pemangkasan untuk menghapus kumpulan spektrum yang tidak menarik. (Wandi, 2012).(Mara & Satyahadewi, 2013).

Metode apriori adalah metode untuk menemukan pola hubungan antara satu atau lebih elemen dalam suatu kumpulan data. Algoritma apriori banyak digunakan dalam data transaksi atau biasa disebut keranjang belanja. Misalnya supermarket memiliki keranjang belanja. Melalui algoritma apriori pemilik supermarket dapat mengetahui pola pembelian konsumen. Jika konsumen membeli barang A dan B, mereka memiliki Ada kemungkinan 50% bahwa komoditas C akan dibeli. Mengingat data transaksi saat ini, model ini sangat penting. (Frank, 2005).

Masalah yang dipelajari dalam artikel ini adalah bagaimana menggunakan metode apriori dalam pengklasifikasian karakteristik suatu objek, Secara umum, dalam pembentukan gaya asosiasi dengan menggunakan algoritma apriori, terdapat dua tingkatan: pertama, mencari frequent itemset (kumpulan objek yang memenuhi nilai panduan minimum). kedua, membentuk *affiliation sample* dari *common itemset* yang telah diperoleh penggunaan *confidence price*. berikutnya adalah bagian dari atribut dari seperangkat aturan apriori:

1. Nilai Support (nilai penunjang):
persentase kombinasi item dalam database.
2. Nilai confidence (nilai kepastian): kuatnya hubungan antar-item dalam aturan asosiasi.
3. Itemset : himpunan item-item yang ada dalam transaksi.
4. k-itemset : itemset yang berisi k item
5. Candidate k-itemset (C_k) : calon kitemset dari data transaksi.
6. Frequent k-itemset (L_k) : itemset yang memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan. (Agusta et al., 2007).

G. Aplikasi Website Framework CodeIgniter

Aplikasi berbasis website yang digunakan dalam penelitian ini menggunakan framework CodeIgniter. Tujuan dari penggunaan *Aplikasi berbasis website* ini adalah membandingkan hasil dengan perhitungan secara teoritis dengan hasil yang didapatkan dengan proses di *Aplikasi berbasis website* ini. Alat penelitian *Aplikasi berbasis website*, seperti tampak pada Gambar 2.2 adalah *Aplikasi berbasis website open source* berbasis framework CodeIgniter yang terkoneksi dengan database mysql.



Gambar 2.2 Aplikasi berbasis website menggunakan framework CI

CodeIgniter adalah jaringan aplikasi bersih open source yang digunakan untuk membangun paket Hypertext Preprocessor dinamis. CodeIgniter adalah personal home page framework dengan versi MVC (version, View, Controller) untuk membangun situs web dinamis menggunakan php yang dapat mempercepat pengembang untuk membuat program internet. selain ringan dan cepat, CodeIgniter juga memiliki seluruh dokumentasi yang luar biasa yang diamati melalui contoh implementasi kode.

Seluruh dokumentasi ini adalah salah satu alasan kuat mengapa banyak orang memilih CodeIgniter sebagai framework favorit mereka. karena kelebihan yang dimiliki CodeIgniter, penulis home page pribadi Rasmus

Lerdorf memuji CodeIgniter di frOSCon (Agustus 2008) dengan mengatakan bahwa dia menyukai CodeIgniter karena “it is faster, lighter and the least like a framework.”

CodeIgniter pertamakali dikembangkan pada tahun 2006 oleh Rick Ellis. Dengan logo api yang menyala, CodeIgniter dengan cepat “membakar” semangat para web developer untuk mengembangkan web dinamis dengan cepat dan mudah menggunakan framework PHP yang satu ini.

Keunggulan Codeigniter

1. Size yang kecil

Seluruh kerangka Codeigniter paling efektif 2MB. Ukuran yang kecil ini memudahkan kita untuk mempelajari Codeigniter, menjalankannya, dan menggantinya.

2. Sangat cepat

Sebagai pelanggan, kita umumnya cenderung menyukai paket yang mungkin cepat, baik loading maupun saat pemakaiannya. Codeigniter memiliki kekuatan respons pemuatan kurang dari 50 milidetik.

3. Fleksibel

Codeigniter memiliki komponen dan kemampuan yang berdiri sendiri tanpa terlalu bergantung satu sama lain, membuat Codeigniter sangat bersih untuk diupgrade dan akurat jika terjadi masalah..

4. Arsitektur yang bersifat MVC (Model-View-Controller)

Saat mengembangkan dan menjalankan program berbasis web, tata letak arsitektur MVC adalah tren perusahaan untuk sebuah kerangka kerja. MVC sangat berguna untuk memisahkan fakta, logika perusahaan komersial, dan presentasi.

5. Dapat dimodifikasi

Pada awalnya Codeigniter sudah memiliki library dan help dari awal kami gunakan. Jika apa yang Anda cari tidak selalu ada, Anda dapat dengan mudah mengunggah perpustakaan dan pembantu pribadi Anda.

6. Mudah dipelajari dan dikuasai

Bagi Anda yang sudah terbiasa dengan PHP, Codeigniter adalah kerangka kerja yang sangat mudah untuk dipelajari. Dalam waktu singkat, pelanggan dapat membuat dan mengembangkan program penggunaan Codeigniter.

H. Pengertian flowchart (bagan alir)

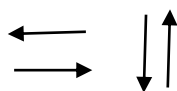
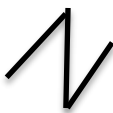
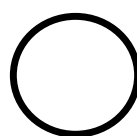
Flowchart atau bagan waft adalah suatu pendekatan analitis yang digunakan untuk menjelaskan semua aspek dari sebuah mesin arsip secara pasti, tepat, dan logis. Flowchart menggunakan seperangkat simbol yang modis untuk menggambarkan prosedur pemrosesan transaksi yang digunakan oleh perusahaan, serta untuk menggambarkan arus fakta dalam perangkat..(Krismiaji, n.d.).

Secara populer, diagram alur atau diagram alur dapat diartikan sebagai rantai langkah/prosedur yang dicapai dalam sistem untuk menyelesaikan masalah. Bagan apung atau bagan melayang memiliki simbol-simbol berikut dan fungsinya masing-masing:

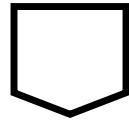
1. “Flow direction symbol” (simbol penghubung atau alur)

Sesuai dengan namanya, gambar ini memiliki fungsi menghubungkan simbol-simbol yang terdapat di dalam waft chart. Gambar ini juga disebut garis penghubung.

Tabel 2.1 menghubungkan simbol atau alur

| Symbol | Nama symbol | Fungsi |
|---|--------------------|--|
|  | Arus/flow | Tunjukkan alur prosesnya. |
|  | Communication link | menggambarkan konversi atau konversi data dari satu lokasi ke lokasi lain. |
|  | On page connector | Menggambarkan hubungan suatu proses |

dengan proses lain yang masih pada lembar atau halaman yang sama.








Off page connector



Menggambarkan sambungan satu proses dengan proses lainnya yang berbeda lembar atau halaman.

2. "Processing symbol" (simbol proses)

Simbol ini menggambarkan proses dalam diagram alir atau operasi dalam proses.

Tabel 2.2 symbol proses





| Symbol | Nama symbol | Fungsi |
|---|--------------------|--|
|  | Proses | Menggambarkan proses atau operasi yang ada pada flowchart. |
|  | Processing manualy | Menjelaskan operasi atau proses yang dilakukan secara manual atau tidak dapat dilakukan oleh komputer. |
|  | Decision | Untuk menjelaskan kondisi tertentu dengan dua pilihan ya / tidak. |
|  | Predefined press | Menyatakan penyediaan tempat penyimpanan suatu pengolahan untuk member nilai awal. |
|  | Terminal | Menunjukkan awal dan akhir program atau proses. |

| | | |
|---|------------------|---|
|  | Off-line storage | Menjelaskan awal dan akhir suatu program atau proses. |
|  | Manual input | Menjelaskan proses input yang dijalankan secara manual. |

3. "Input-output symbol"

Simbol ini menunjukkan jenis peralatan sebagai input atau output proses dan media.

Tabel 2.3 input-output symbol

| Symbol | Nama symbol | Fungsi |
|---|--------------------|--|
|  | Input-output | Menjelaskan proses input atau output, terlepas dari jenis alat yang digunakan. |
|  | Punched card | Menunjukkan input dari kartu atau output yang direkam pada media kartu.. |
|  | Magnetic-tape unit | Menggambarkan masukan dari pita magnetik atau keluaran yang disimpan pada pita magnetik. |
|  | Disk storage | Menunjukkan input atau output dari sumber disk yang disimpan pada disk optik. |



Document

Deskripsi proses
pencetakan laporan
pada printer



Display

Deskripsi perangkat
keluaran yang Anda
gunakan adalah layar.
