



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1 Youtube

Youtube adalah salah satu layanan Video-on-Demand (VoD) paling populer, yang memungkinkan pengguna untuk *streaming* konten video melalui web browser. Didirikan oleh Chad Hurley, Steven Chen, dan Jawed Karim, YouTube diluncurkan pada Juni 2005. Situs web ini menyediakan antarmuka yang sangat sederhana dan terintegrasi di mana pengguna dapat mengunggah, menerbitkan, dan melihat video *streaming* tanpa pengetahuan teknis tingkat tinggi. YouTube sebanding dengan start-up video lainnya pada saat itu hingga Google mengakuisisi situs tersebut dengan \$ 1,65 miliar pada Oktober 2006 (Burgess 2009). Situs ini semakin populer, dan sejak 2008 secara konsisten masuk dalam sepuluh besar situs yang paling banyak dikunjungi secara global (Morreale 2014). Sejak dibeli oleh Google, YouTube telah berevolusi dari situs tempat video amatir dan bebas iklan diposkan ke tujuan daring yang sekarang digunakan oleh video profesional dan komersial. Tetapi ada sisi lain sebagaimana Morreale (2014) menulis, “Tagline-nya *'Broadcast Yourself'* mengundang pengguna biasa untuk mengambil bagian aktif dalam menciptakan bahan yang mereka konsumsi. Pada saat yang sama, yang kurang jelas adalah bahwa YouTube adalah bisnis yang tujuannya untuk menghasilkan laba (Morreale 2014).

2.2 SocialBlade.com

Social Blade adalah situs web statistik yang memungkinkan Anda melacak statistik dan mengukur pertumbuhan di berbagai platform media sosial termasuk YouTube, Twitch, dan Instagram. Ketika Social Blade dibuat, pertama kali hanya membuat peringkat berdasarkan jumlah *subscribers* dan *video views*, tetapi kedua hal ini bukan indikator yang akurat tentang apa yang sebenarnya orang-orang lakukan di YouTube. Bisa saja sebuah channel Youtube Seseorang dapat memiliki *subscribers* palsu dalam jumlah yang cukup banyak, akan tetapi jumlah *video views* jauh lebih sedikit. Sistem pemeringkatan SB bertujuan untuk mengukur pengaruh

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

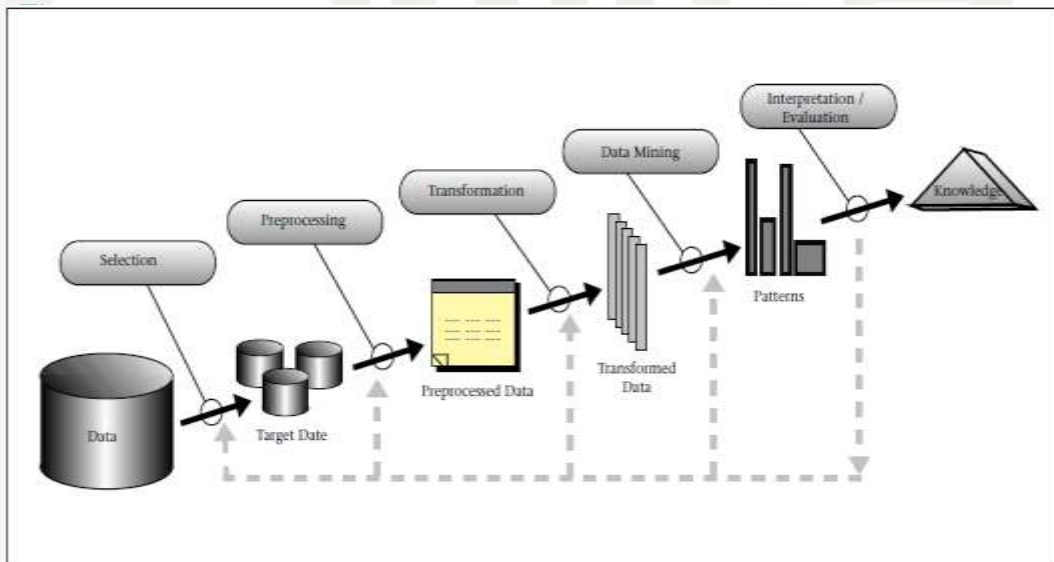
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

channel berdasarkan berbagai metrik termasuk jumlah rata-rata *video views*, *subscribers*, dan *video uploads* (socialblade.com 2019).

2.3 Tahapan Knowledge Discovery in Database (KDD)

Knowledge discovery in database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data untuk menemukan keteraturan, pola atau hubungan dalam *dataset* yang besar. *Data mining* merupakan salah satu tahapan dari KDD. Terdapat beberapa tahapan dalam KDD yakni diantaranya : seleksi data, *pre-proccesing/cleaning*, transformasi data, *data mining*, evaluasi pola dan presentasi pengetahuan.



Gambar 2.1 Tahapan-tahapan KDD (Fayyad, Piatetsky-Shapiro dan Smyth 1996)

2.3.1 Data Selection

Data yang ada direduksi dan diambil data yang relevan dari *database* untuk dianalisis. Beberapa cara seleksi data antara lain adalah :

- A. *Dimensionality reduction*, adalah mengurangi jumlah variabel acak atau atribut yang menjadi pertimbangan.
- B. *Numerosity reduction*, menggunakan parametric atau non-parametrik model untuk mendapatkan representasi yang lebih kecil dari data asli.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

C. *Data compression*, menerapkan transformasi untuk mengurangi representasi dari data asli.

2.3.2 Data Preprocessing

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten. membersihkan data dapat dilakukan dengan mengisi data yang memiliki *missing value*, memperbaiki kesalahan data, mengidentifikasi dan menghilangkan *outliers*, serta memeriksa data yang tidak konsisten. Data yang memiliki *noise* dapat menyebabkan kekeliruan dalam proses *mining*, sehingga mempengaruhi performansi hasil yang akan didapat. Dengan pembersihan data ini dapat dilihat data mana yang terdapat kesalahan sehingga data tersebut dapat ditangani sebelum menjadi inputan pada proses *mining* nantinya. Beberapa cara mengatasi *missing value*:

1. Mengabaikan tupel: yaitu mengabaikan tupel dari kelas yang hilang. Metode ini sangat tidak efektif terutama ketika persentase kelas yang hilang bervariasi.
2. Mengisi nilai yang hilang secara manual: cara ini akan memakan waktu. Terutama pada data yang besar dan *missing value* yang banyak.
3. Menggunakan konstanta yang umum: mengganti semua nilai atribut yang hilang dengan konstanta yang sama seperti label. Meskipun metode ini sederhana, tetapi agak sulit diterapkan.
4. Menggunakan ukuran *tendency* tengah atribut untuk mengisi *missing value* : contoh ukuran *tendency* yang biasa digunakan adalah *mean* dan median. Untuk data yang normal, *mean* dapat digunakan. Sedangkan median dapat digunakan untuk data yang bias.
5. Menggunakan kemungkinan nilai untuk mengisi *missing value* : dapat ditentukan dengan regresi

Menurut (Larose 2006), jika jumlah data yang digunakan banyak atau database besar maka proses ini bisa diabaikan dan jika jumlah data yang digunakan kecil atau terbatas, maka data yang hilang bisa digantikan dengan nilai rata-rata dari atribut yang berkaitan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.3.3 Data Transformation

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Beberapa cara yang dapat dilakukan dalam transformasi data adalah :

1. *Smoothing*, menghilangkan *noise* dari data.
2. *Attribute construction*, menambahkan atau membuat atribut baru untuk membantu proses *mining*.
3. *Aggregation*, ringkasan atau operasi agregasi diterapkan pada data.
4. *Normalization*, data atribut dibuat dalam skala tertentu sehingga menjadi kisaran data yang lebih kecil sehingga sebaran datanya tidak terlalu jauh.
5. *Discretization*, nilai-nilai baku atribut numerik diubah menjadi data dengan interval label. Misalnya usia yang terdiri dari anak-anak, remaja dan dewasa.
6. *Concept hierarchy generation for nominal data*, banyak hirarki untuk atribut nominal yang implisit dalam skema *database* dan dapat otomatis didefinisikan.

2.3.4 Data Mining

Data Mining merupakan proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Terdapat berbagai macam teknik dalam *data mining*, dimana pemilihannya bergantung pada tujuan dan proses pencarian pengetahuannya.

2.3.5 Evaluasi Pola

Tahapan evaluasi pola untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses *data mining*, mencoba metode *data mining*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang diluar dugaan mungkin bermanfaat.

2.3.5 Presentasi Pengetahuan

Presentasi pengetahuan merupakan visualisasi penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses *data mining* ini adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami *data mining*. Karenanya presentasi hasil *data mining* dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses *data mining*. Dalam presentasi, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining*.

2.4 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowlegde*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery in Database* (KDD) adalah penerapan metode saintifik pada *data mining*. Dalam konteks ini data mining merupakan satu langkah dari proses KDD (Hermawati F. A., 2013).

2.4.1 Karakteristik Data Mining

Berdasarkan definisi-definisi yang telah disampaikan, hal penting terkait data mining adalah :

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2.4.2 Pengelompokan *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

1. Deskripsi

Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa datang.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori yang berguna untuk menggolongkan data kedalam kelompok-kelompok yang ada. Misalnya penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

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 lainnya dan memiliki ketidakmiripan dengan record-record dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (Kusrini & Luthfi, 2009).

2.5 Association Rule

Association rule mining adalah suatu prosedur untuk menemukan aturan asosiatif antara suatu kombinasi item dalam suatu dataset yang ditentukan. Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu. Umumnya ada dua ukuran, yaitu *support* dan *confidence*. Kedua ukuran ini nantinya berguna dalam menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh user. Batasan tersebut umumnya terdiri dari *minimum support* dan *minimum confidence*. Metodologi dasar analisis asosiasi terbagi menjadi dua tahap (Rizky 2014):

2.5.1 Analisa Pola Frekuensi Tinggi

Pada tahap ini untuk mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah item diperoleh dengan rumus berikut :

$$Support(A) =$$

$$\frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \dots\dots\dots(2.1)$$

Pada rumus ini dijelaskan bahwa nilai *support* (A) diperoleh dengan cara mencari jumlah transaksi mengandung di bagi dengan total transaksi. Sementara itu, nilai *support* dari 2 item diperoleh dari rumus berikut :

$$Support(A, B) = P(A \cap B)$$

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$Support(A, B) =$

$$\frac{\Sigma \text{Transaksi Mengandung A dan B}}{\Sigma \text{Total}} \dots\dots\dots(2.2)$$

Pada rumus ini dijelaskan bahwa *support* (A,B) diperoleh dengan cara transaksi yang mengandung item A dan B di bagi dengan jumlah seluruh transaksi.

Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah di cari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan diperoleh dari rumus berikut :

$Confidence = P(B|A) =$

$$\frac{\Sigma \text{Jumlah Transaksi Mengandung A dan B}}{\Sigma \text{Transaksi mengandung A}} \dots\dots\dots(2.3)$$

Pada rumus diatas dijelaskan bahwa *confidence* (A, B) diperoleh dengan cara data transaksi yang mengandung A dan B dibagi dengan transaksi mengandung A (Rizka Nurul Arifin, 2014).

2.6 Algoritma ECLAT

Algoritma ECLAT merupakan salah satu algoritma yang diusulkan M Zaki dalam metode *Association Rules* yang memiliki waktu pencarian *frequent itemset* yang lebih cepat dibanding Apriori dengan pendekatan teknik BDF (*Breath Deep First*) yang hanya melakukan scan terhadap database sebanyak 3 (tiga) kali dengan mentransformasikan tabel sehingga dapat mengurangi proses yang terjadi di bagian *input* maupun *output*. Selain itu, algoritma ini juga direpresentasikan secara vertikal dalam *database* (untuk setiap *item*, dibuat *list/array* transaksi) sehingga mengurangi pemakaian memori yang membuat algoritma ini bekerja lebih cepat (Borgelt 2010).

Equivalence Class Transformation (ECLAT) merupakan algoritma yang sangat sederhana untuk menemukan itemset yang paling sering muncul, pada dasarnya algoritma ECLAT melakukan pencarian secara depth-first search pada database dengan tata letak vertikal, jika database berbentuk horizontal maka harus dikonservikan ke bentuk vertikal terlebih dahulu (Norsyanah 2016).

Contoh penerapan algoritma ECLAT adalah sebagai berikut :

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Format data awal transaksi berbentuk horizontal, seperti pada Tabel 2.1 di bawah ini:

Tabel 2.1 Transaksi Dalam Format Horizontal

| TID | Item |
|-----|------------------------------|
| 1 | Jagung, Gandum, Telur |
| 2 | Gandum, Tepung |
| 3 | Gandum, Beras |
| 4 | Jagung, Gandum, Tepung |
| 5 | Jagung, Beras |
| 6 | Gandum, Beras |
| 7 | Jagung, Beras |
| 8 | Jagung, Gandum, Beras, Telur |
| 9 | Jagung, Gandum, Beras |

Untuk pembentukan *itemset* menggunakan ECLAT, transaksi di atas diolah dahulu ke dalam bentuk vertikal. Berikut tampilan data dalam bentuk vertikal dimana pada kasus ini *minimum support* = 2, seperti pada Tabel 2.2 di bawah ini :

Tabel 2.2 Transaksi dalam format vertical

| Item | TID |
|--------|---------------|
| Jagung | 1,4,5,7,8,9 |
| Gandum | 1,2,3,4,6,8,9 |
| Beras | 3,5,6,7,8,9 |
| Tepung | 2,4 |
| Telur | 1,8 |

Dikarenakan seluruh *itemset* pada gambar di atas memenuhi *minimum support*, maka seluruh *itemset* di atas digunakan untuk pencarian penyilangan 2-*itemset* selanjutnya. Seperti pada Tabel 2.3.

Tabel 2.3 Hasil Penyilangan 2-itemset

| 2-Itemset | TID |
|----------------|-----------|
| Jagung, gandum | 1,4,8,9 |
| Jagung, beras | 5,7,8,9 |
| Jagung, tepung | 4 |
| Jagung, telur | 1,8 |
| Gandum, beras | 3,6,7,8,9 |
| Gandum, tepung | 2,4 |
| Gandum, telur | 1,8 |
| Beras, tepung | 0 |
| Beras, telur | 8 |
| Tepung, telur | 0 |

Dikarenakan pada hasil di atas terdapat beberapa *itemset* yang tidak memenuhi syarat *minimum support*, maka *itemset* tersebut dihapuskan. Berikut tabel *frequent 2-itemset* dari penyilangan *2-itemset* sebelumnya. Seperti pada Tabel 2.4 di bawah ini:

Tabel 2.4 Frequent 2-itemset

| 2-Itemset | TID |
|------------------|------------|
| Jagung, gandum | 1,4,8,9 |
| Jagung, beras | 5,7,8,9 |
| Jagung, telur | 1,8 |
| Gandum, beras | 3,6,7,8,9 |
| Gandum, tepung | 2,4 |
| Gandum, telur | 1,8 |

Selanjutnya yaitu langkah penyilangan *3-itemset*. Hasil penyilangan dapat dilihat pada Tabel 2.5 di bawah ini :

Tabel 2.5 Hasil Penyilangan 3-itemset

| 3-Itemset | TID |
|------------------------|------------|
| Jagung, gandum, beras | 8,9 |
| Jagung, gandum, tepung | 4 |
| Jagung, gandum, telur | 1,8 |

Dari hasil di atas, diketahui terdapat *itemset* yang tidak memenuhi syarat *minimum support* dan akan dihilangkan. *Frequent 3-itemset* dapat dilihat pd Tabel 2.6.

Tabel 2.6 Frequent 3-itemset

| 3-Itemset | TID |
|-----------------------|------------|
| Jagung, gandum, beras | 8,9 |
| Jagung, gandum, telur | 1,8 |

Penerapan algoritma ECLAT dalam Pseudocode dapat dilihat pada gambar di bawah ini:

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

Input : database D , minimum support , a set of atom of a sublattice S
Output : Frequent itemsets F
Procedure Eclat ( S )
For all atoms  $A_i \in S$ 
     $T_i = \theta$ ;
    For all atoms  $A_j \in S$  , with  $j > i$  do
         $R = A_i \cup A_j$ ;
         $L(R) = L(A_i) \cap L(A_j)$ 
        If  $support(R) \geq minsup$  then
             $T_i = T_i \cup \{R\}$ ;
             $F_{i,R} = F_{i,R} \cup \{R\}$ ;
        end
    end
end
    
```

Gambar 2.2 Pseudocode Algoritma ECLAT (Xu 2011)

2.7 Penelitian Terkait

Berikut adalah beberapa penelitian sebelumnya yang menjadi acuan dan dasar penelitian ini :

1. Joseph Eric Samodra, 2015

Judul penelitian Joseph Eric Samodra yaitu, "Implementasi Algoritma Eclat Untuk Frequent Pattern Mining Pada Penjualan Barang" support antara 5% sampai 20%. Minimal support 15% dan 20% tidak melaporkan peraturan penjualan.

2. Budiana, 2015

Judul implementasi *data mining* pada penjualan produk di PT. Focus Gaya Graha Menggunakan Metode *Association Rule* yaitu ECLAT. Penelitian ini bertujuan untuk mengetahui pola pembelian konsumen agar dapat mempermudah dalam menentukan produk apa yang akan diproduksi. Menggunakan data penjualan dari bulan Januari-Februari 2015 dengan nilai *minimum support* 3 dan nilai *minimum confidence* 50%. Aturan asosiasi sistem yang dibangun adalah MB.0602 – Beige → MB.0603 – Beige *confidence* 50%, LPP.32 – Walnut → LPP.22 – Italian Walnut *confidence* 60%, MB.0603 – Beige → MB.0602 – Beige *confidence* 60%, LHS.0701 – Beige → LHS.0703 – Walnut *confidence* 66,67%, LHS.0703 → LHS.0701 – Beige *confidence* 66,67% dan LPP.22 – Italian Walnut → LPP.32 – Walnut *confidence* 100%. Kesimpulan yang didapat yaitu nilai *minimum support* dan *minimum confidence* mempengaruhi terhadap banyaknya *rule* yang terbentuk.


Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Maka dari itu untuk nilai *minimum support* lebih baik bernilai kecil, sedangkan untuk nilai *minimum confidence* lebih baik bernilai besar karena jika seperti itu akan menghasilkan *rule* yang lebih bervariasi dengan nilai kepastian yang tinggi.

3. Enny Norsyanah, 2016

Judul Penerapan Algoritma Eclat Dalam Menentukan Metode Kontrasepsi Yang Dipilih. Tujuan dari penelitian ini adalah mencari metode kontrasepsi yang dipilih berdasarkan umur, lama menikah, pekerjaan dan pendidikan. Hasil dari penelitian ini yaitu pengguna KB terbanyak adalah pengguna yang berumur antara 26-35 tahun, dengan usia pernikahan 1-10 tahun, bekerja sebagai Ibu Rumah Tangga (IRT) dan berpendidikan SMA/ sederajat, dengan metode KB yang paling banyak adalah jenis suntik dengan nilai *support* tertinggi sebesar 8,50% dengan nilai *confidence* sebesar 55,26%.

2 Adita dkk, 2015

Judul penelitian Rizky Adita dkk adalah, " *Identification of Best Algorithm in Association Rule Mining Based on Performane* ". Tujuan dari penelitian ini adalah Membandingkan peforma dari tiga algoritma asosiasi, yaitu: FP-Growth, Apriori dan ECLAT. Yang menunjukkan bahwa ECLAT merupakan yang terbaik karena memiliki waktu eksekusi yang lebih cepat dengan nilai *support* dan *confidence* yang lebih baik.