

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.

2.1 Adverse Event

Adverse event merupakan tanggapan berbahaya atau yang tidak diinginkan untuk obat yang terjadi pada dosis biasanya digunakan untuk profilaksis, diagnosis, atau terapi penyakit atau untuk modifikasi fungsi psikologis (WHO, 1972 dikutip oleh Mariyono & Suryana, 2008). Menurut FDA (*Food And Drug Administration*) Amerikat Serikat (2009), *adverse event* adalah efek yang tidak diinginkan yang berhubungan dengan penggunaan obat yang timbul sebagai bagian dari aksi farmakologis dari obat yang kejadiannya mungkin tidak diharapkan. *Adverse event* adalah dampak yang membahayakan atau tidak menyenangkan yang disebabkan oleh dosis obat yang digunakan sebagai terapi yang mengharuskan untuk mengurangi dosis atau menyetop pemberian dan meramalkan adanya bahaya pada pemberian selanjutnya (Lawrence, 1998 dikutip oleh Mariyono & Suryana, 2008).

“*Adverse event means any unfordward medical accurrence associated with the use of a drug in humans, whether or not considered drug related*” (*Adverse event* merupakan suatu kejadian yang tidak diinginkan yang terkait dengan penggunaan obat pada manusia, apakah itu termasuk narkoba ataupun tidak) (Goldfarb, 2012). Menurut Francisco (2014) pada jurnal *Clinical and Translational Science Institute*, *adverse event* merupakan suatu kejadian yang merugikan yang timbul dari setiap penggunaan obat yang tidak diketahui kemunculannya. Menurut *World Health Organization*, “*Adverse event medical occurrence temporally associated with the use of a medicinal product, but not necessarily causally related*” (kejadian medis temporal terkait dengan penggunaan produk obat, tetapi belum tentu saling berhubungan

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.

2.2 FDA (*Food and Drug Administration*)

FDA (Food and Drug Administration) atau Badan Pengawas Obat dan Makanan Amerika Serikat adalah salah satu lembaga kesehatan Amerika yang bertanggung jawab melindungi dan mempromosikan kesehatan masyarakat melalui peraturan dan pengawasan keamanan pangan, produk tembakau, suplemen diet, obat-obatan, vaksin, produk biofarmasi, transfusi darah, peralatan medis, produk kedokteran hewan dan kosmetik di Amerika Serikat. *FDA* menetapkan standar yang memberi wewenang kepada perorangan untuk meresepkan obat atau peralatan medis. Selain itu, *FDA* bertugas mengatur berbagai produk yang ditawarkan produsen ke konsumen untuk menjamin keamanannya sehingga apa yang dijanjikan produsen dalam produknya dapat dinikmati dengan benar oleh konsumen.

2.3 *Knowledge Discovery in Database (KDD)*

KDD merupakan suatu area yang mengintegrasikan berbagai metode yang meliputi statistik, basis data, kecerdasan buatan (*Artificial Intelligence*), *machine learning*, pengenalan pola (*Pattern Recognition*), pemodelan yang menangani ketidakpastian, visualisasi data, optimasi, Sistem Informasi Manajemen (SIM), dan sistem berbasis pengetahuan (*knowledge based-system*). Sebagai bagian dari proses yang ada di dalam *KDD*, maka *data mining* didahului dengan proses pemilihan data, pembersihan data, *pre-processing*, dan transformasi data (Larose, 2005). Proses *KDD* secara garis besar dapat dijelaskan sebagai berikut (Fayyad, Piatetsky-shapiro, & Smyth, 1996):

1. *Data Selection*

Menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan akan dilakukan. Data dari hasil seleksi digunakan untuk proses *data mining*, disimpan pada suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/ Cleaning*

Pemrosesan, pendahuluan, dan pembersihan data merupakan operasi dasar seperti penghapusan *noise* dilakukan. Tahap *cleaning* ini mencakup seperti membuang data duplikasi, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan dalam menulis (*tipografi*). Pada tahap ini juga dilakukan proses *enrichment*, yaitu merupakan suatu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan, seperti data atau informasi eksternal.

3. Transformation

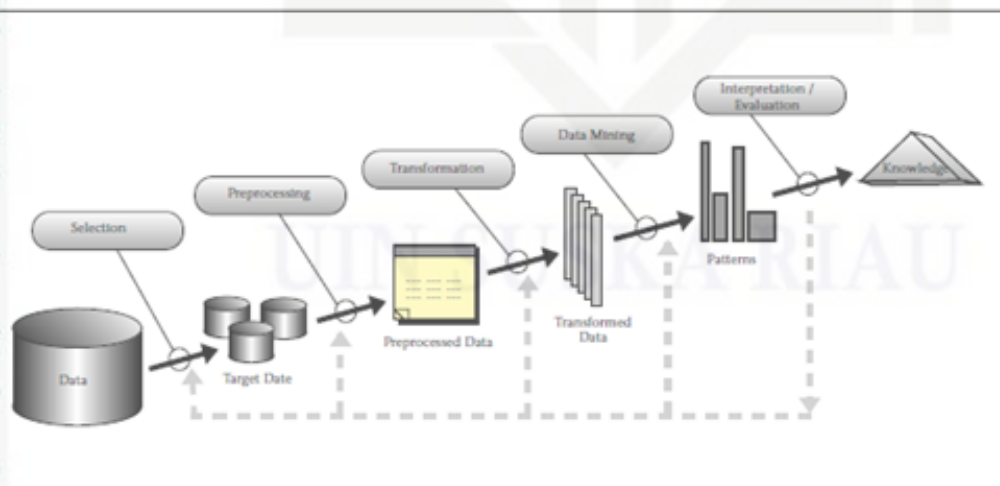
Mentransformasi atau menggabungkan atau mengubah data ke dalam bentuk yang sesuai untuk penggalian lewat operasi *summary* atau *aggregation*. Tahap ini dalam *KDD* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. Data Mining

Pemilihan tujuan dari proses *KDD* misalnya klasifikasi, regresi, clustering dan lain sebagainya. Juga merupakan proses mencari pola atau informasi yang menarik dalam data yang terpilih dengan menggunakan teknik atau metode tertentu.

5. Interpretation/Evaluation

Penerjemahan pola-pola yang dihasilkan dari data mining. Informasi yang dihasilkan haruslah dalam bentuk yang mudah dipahami oleh banyak orang.



Gambar 2.1 Tahap-tahap KDD

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.

2.4 Data Mining

Kebutuhan dunia bisnis yang ingin mendapatkan nilai tambah dari data yang telah terkumpul, mendorong penerapan teknik pengolahan data dari berbagai bidang pengetahuan seperti statistika dan kecerdasan buatan. Penerapan teknik tersebut memberikan tantangan baru yang akhirnya memunculkan metode baru yang disebut *data mining* (Kusrini & Luthfi, 2010).

Data Mining merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*. *Data Mining* juga merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dar berbagai *database* (Kusrini & Luthfi, 2010). Menurut Goela (2012) *data mining* merupakan proses menganalisis suatu data dengan berbagai perspektif dan meringkasnya dalam bentuk informasi yang berguna, sehingga informasi tersebut bisa digunakan untuk meningkatkan pendapatan, memotong biaya atau keduanya.

Konsep *data mining* muncul dikarenakan timbulnya data *explosion* akibat dari penumpukan data oleh sistem pengolahan basis data terpadu di suatu organisasi. Proses *data mining* menggunakan berbagai perangkat analisis data untuk menemukan pola dan hubungan dalam data yang mungkin dapat digunakan untuk membuat prediksi yang valid (Zhao, 2003). *Data Mining* juga merupakan suatu istilah yang digunakan untuk menemukan suatu pengetahuan yang tersembunyi didalam *database*.

2.4.1 Teknik Data Mining

Dengan definisi data mining yang sangat luas, ada banyak jenis teknik analisa yang dapat digolongkan dalam data mining. Beberapa diantaranya yaitu:

1. Analisis Asosiasi

Analisis asosiasi dikenal juga sebagai salah satu metode data mining yang menjadi dasar dari berbagai metode data mining lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien.

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

2. Klasifikasi dan Prediksi

Merupakan suatu proses untuk menemukan sebuah model (fungsi) yang menjelaskan atau mencirikan konsep atau kelas data untuk suatu kepentingan tertentu, yang bisa menggunakan pemodelan untuk memprediksi kelas objek yang labelnya tidak diketahui. Model yang didapat mungkin bisa diwakili dalam berbagai format, seperti aturan klasifikasi IF-THEN, pohon keputusan, formula matematika, atau jaringan saraf tiruan. Klasifikasi bisa digunakan untuk memprediksi label kelas data objek pada data.

3. Analisis *Clustering*

Clustering termasuk metode yang sudah cukup dikenal dan banyak dipakai dalam data mining. Sampai sekarang para ilmuwan dalam bidang data mining masih melakukan berbagai usaha untuk melakukan perbaikan model *clustering* karena metode yang dikembangkan sekarang ini masih bersifat heuristik. Usaha untuk menghitung jumlah *cluster* yang optimal dan pengklasteran yang paling baik masih terus dilakukan. Dengan demikian menggunakan metode yang ada saat ini tidak bias menjamin hasil merupakan yang paling maksimal. Namun, hasil yang biasanya dicapai sudah cukup bagus dari segi praktis.

Berbeda dengan klasifikasi dan prediksi yang menganalisa pelabelan objek data, *clustering* menganalisis objek data tanpa mengkonsultasikan label kelas yang dikenal. Secara umum label kelas bukan didapat dalam pengolahan data sederhana, karena mereka tidak tau bagaimana memulainya. *Clustering* bisa digunakan untuk meng-generate label. Objek yang di *cluster* (dikelompokan) berdasarkan pada prinsip memaksimalkan persamaan dalam kelas dan meminimalkan kesamaan antar kelas.

4. Analisis *Outlier*

Didalam sebuah *database* mungkin berisi objek data yang tidak sesuai dengan kebiasaan umumnya yang disebut *Outlier*. Analisa *outlier* membantu dalam pendekteksian kesalahan dan nilai-nilai abnormal.

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.

2.4.2 Pengelompokan *Data Mining*

Dalam *data mining*, terdapat beberapa teknik yang dimiliki berdasarkan tugas yang bisa dilakukan, yaitu:

1. Deskripsi

Para peneliti/ analis biasanya mencoba menemukan cara untuk mendeskripsikan pola yang tersembunyi dalam data.

2. Estimasi

Estimasi mirip dengan klasifikasi, kecuali variabel tujuan yang lebih ke arah numerik daripada kategori. Misalnya, akan dilakukan estimasi tekanan *systolic* dari pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan dan level sodium darah.

3. Prediksi

Prediksi memiliki kemiripan dengan estimasi dan klasifikasi. Hanya saja, prediksi hasilnya menunjukkan sesuatu yang belum terjadi (mungkin terjadi dimasa depan). Misalnya, ingin diketahui prediksi harga beras tiga bulan yang akan datang.

4. Klasifikasi

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep kelas atau data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Proses klasifikasi biasanya dibagi menjadi dua fase yaitu *learning* dan *test*. Pada fase *learning* sebagian data yang telah diketahui kelas datanya diumpangkan untuk membentuk model prediksi. Kemudian pada fase *test*-nya model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut.

Dalam klasifikasi variabel, tujuan bersifat kategorik. Misalnya kita akan mengklasifikasikan pendapatan dalam tiga kelas, yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah.

5. *Clustering*

Clustering melakukan pengelompokan data tanpa berdasarkan kelas data tertentu. *Clustering* lebih ke arah pengelompokan record, pengamatan, atau kasus dalam kelas yang memiliki kemiripan. Sebuah *cluster* adalah kumpulan record yang

memiliki kemiripan satu dengan yang lain dan memiliki ketidakmiripan dengan *record-record* dalam *cluster* yang lain. Misalnya untuk tujuan audit akuntansi akan dilakukan segmentasi perilaku financial dalam kategori data mencurigakan. Prinsip dari *clustering* adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas.

6. Asosiasi

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu. Pendekatan asosiasi tersebut menekankan sebuah kelas masalah dicirikan dengan analisis keranjang pasar (*market basket analysis*) (Kusrini, 2010).

2.5 Association Rule

Association Rule atau sering disebut *Market Basket Analysis* adalah salah satu teknik data mining dalam pencarian hubungan antar item pada suatu set data (Larose, 2005). *Association Rules* merupakan salah satu teknik hasil riset terbaik dari data mining (Agrawal, Imieliński, & Swami, 1993). Tujuan dari *Association rule* adalah untuk mengekstrak korelasi yang menarik, pola yang sering muncul, asosiasi atau struktur kasual antar item set dalam database transaksi.

Didalam *data mining* asosiasi bertugas untuk menemukan atribut yang muncul dalam satu waktu. Analisis asosiasi merupakan teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. Ada dua hal yang penting dalam aturan asosiasi (Zhao, 2003) :

1. Support

Yaitu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *item set* dari keseluruhan transaksi. Rumus untuk menentukan support (Zhao, 2003) :

$$S(a) = \frac{\sum(Ta)}{\sum T} \times 100\% \dots\dots\dots (Persamaan 2.1)$$

Keterangan :

- S = Support
- $\sum(Ta)$ = Jumlah transaksi mengandung antecedent
- $\sum(T)$ = Jumlah transaksi

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

2. Confidence

Merupakan ukuran yang menunjukkan hubungan antar dua *item* secara *conditional* (berdasarkan suatu kondisi tertentu).

Metodologi dasar analisis asosiasi memiliki dua tahap, yaitu melakukan analisa pola frekuensi tinggi (*frequent pattern*) dan proses pembentukan aturan asosiasi.

$$C = \frac{\sum(Ta + Tc)}{\sum Ta} \times 100\% \dots\dots\dots (Persamaan 2.2)$$

Keterangan :

C = Confidence

$\sum (Ta + Tc)$ = Jumlah transaksi mengandung *antecedent* dan *consequent*

$\sum (Ta)$ = Jumlah transaksi mengandung *antecedent*

2.5.1 Algoritma Apriori

Algoritma apriori merupakan algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan frequent itemsets untuk aturan asosiasi *boolean*. Algoritma apriori merupakan salah satu algoritma yang melakukan pencarian *frequent itemset* dengan melakukan teknik association rule. Penting tidaknya suatu aturan asosiasi dapat diketahui dengan dua parameter, yaitu *support* dan *confidence*. *Support* (nilai penunjang) merupakan persentase kombinasi item tersebut dalam database, sedangkan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiasi.

Proses utama yang dilakukan dalam algoritma apriori untuk menemukan *frequent itemset* yaitu:

1. Join (Penggabungan)

Pada proses ini dilakukan kombinasi antara *item* dengan *item* lainnya hingga tidak dapat dikombinasikan lagi.

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.

Prune (Pemangkasan)

Proses ini merupakan hasil dari *item* yang telah dikombinasikan sebelumnya, kemudian dilakukan pemangkasan dengan menggunakan minimum *support* yang telah ditentukan oleh *user*.

Langkah pertama pada algoritma apriori yaitu, *support* dari setiap item dihitung dengan men-*scan* database. Setelah *support* dari setiap item didapat, item yang memiliki *support* lebih besar dari minimum *support* dipilih sebagai pola frekuensi tinggi dengan panjang 1 atau sering disingkat 1-itemset. Singkatan k-itemset berarti satu set yang terdiri dari k item.

Iterasi kedua menghasilkan 2-itemset yang tiap set-nya memiliki dua item. Pertama dibuat kandidat 2-itemset dari kombinasi semua 1-itemset. Lalu untuk tiap kandidat 2-itemset ini dihitung *support*-nya dengan men-*scan* database. *Support* artinya jumlah transaksi dalam database yang mengandung kedua item dalam kandidat 2-itemset. Setelah *support* dari semua kandidat 2-itemset didapatkan, kandidat 2-itemset yang memenuhi syarat minimum *support* dapat ditetapkan sebagai 2-itemset yang juga merupakan pola frekuensi tinggi.

Prinsip dari algoritma apriori ini adalah bila *itemset* digolongkan sebagai *frequent itemset*, yang memiliki *support* lebih dari yang ditetapkan sebelumnya, maka semua subsetnya juga termasuk golongan *frequent itemset*, dan sebaliknya. Cara algoritma ini bekerja adalah algoritma akan menghasilkan kandidat baru dari k-itemset dari *frequent itemset* pada langkah sebelumnya dan menghitung nilai *support* k-itemset tersebut. *Itemset* yang memiliki nilai *support* di bawah dari *minsup* akan dihapus. Algoritma berhenti ketika tidak ada lagi *frequent itemset* baru yang dihasilkan.

2.6 WEKA

Weka adalah aplikasi *Open Source* untuk pengolahan *data mining* yang berbasis *Java* yang dikembangkan pertama kali di Universitas Waikato Selandia Baru dan memiliki sekumpulan algoritma standar untuk *data mining* untuk melakukan *Preprocessing*, *Classification*, *Clustering*, *Regretion*, *Association Rule*,

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.

dan Visualisasi. Weka dapat dijalankan berbasis GUI atau secara langsung menggunakan *Command Line*.

2.7 Penelitian Terkait

Berikut beberapa penelitian terkait mengenai penelitian ini :

1. “*Identification of Suicide-Related Events Through Network Analysis of Adverse Event Reports*” oleh Alwis Nazir, Takashi Ichinomiya, Nobuteru Miyamura, Yasuaki Sekiya, Yasutomi Kinoshita. Penelitian ini merupakan penelitian terhadap identifikasi kasus-kasus adverse event yang berdampak kepada keinginan untuk melakukan bunuh diri. Penelitian ini menggunakan laporan *FAERS* dari tahun 1997 hingga 2012. Penelitian ini bertujuan untuk menganalisa gejala-gejala yang dapat menimbulkan resiko tertinggi bunuh diri menggunakan metode *Network Analysis*.
2. “*Implementasi Algoritma-Asosiasi sebagai bagian dari pengembangan Data Mining Algorithms Collection*” oleh Yova Ruldeviyani dan Muhammad Fahrian. Penelitian ini membahas tentang penggunaan dan perbandingan algoritma-algoritma yang terdapat dalam metode *Association Rule*.
3. “*Pencarian Hubungan Adverse Event Pada Obat Penenang Dengan Apriori*” oleh Meilinda Heriza. Penelitian ini membahas tentang penggunaan apriori untuk menentukan *Adverse Event* yang terjadi kepada orang-orang yang mengkonsumsi obat anti depresi yang menyebabkan mereka memiliki keinginan untuk melakukan bunuh diri.
4. “*Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan*” oleh Kennedi Tampubolon, Hoga Saragih, Bobby Reza. Penelitian ini membahas tentang bagaimana mengimplementasikan algoritma *apriori* untuk menciptakan lingkungan bisnis yang inteligen, dan membuktikan bahwa *Data Mining* dapat digunakan dalam segala macam bidang termasuk bisnis.

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.

5. “*Medical Data Mining Based on Association Rule*” oleh Ruijuan Hu. Penelitian ini dilakukan untuk membuktikan apakah *Data Mining* dapat digunakan untuk penelitian di bidang kesehatan dengan cara mencari hubungan antara kemungkinan terjadinya kanker payudara dan atribut lainnya.

