

# APLIKASI ASSOCIATION RULES DENGAN MENGGUNAKAN ALGORITMA APRIORI DALAM MENDETEKSI POLA PENYAKIT DBD (Studi Kasus : Pasien DBD Puskesmas Cangkringan Sleman)

Triano Nurhikmat<sup>1)</sup>, Yusnandar<sup>2)</sup>, Ummi Maftuhatul M<sup>3)</sup>, Boki Latupono<sup>4)</sup>, Edy Widodo<sup>5)</sup>,  
<sup>12345</sup>Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Islam Indonesia  
[triano1933@gmail.com](mailto:triano1933@gmail.com), [yusnandarnandar@gmail.com](mailto:yusnandarnandar@gmail.com), [ummimaftuhatul78@gmail.com](mailto:ummimaftuhatul78@gmail.com),  
[boki.latu@gmail.com](mailto:boki.latu@gmail.com), [edy.widodo@uii.ac.id](mailto:edy.widodo@uii.ac.id)

## Abstrak

Penyakit Demam Berdarah Dengue (DBD) merupakan penyakit yang disebabkan oleh virus nyamuk ditularkan melalui nyamuk *Aedes aegypti* dan *Aedes albopictus*. Ancaman penyakit ini berlaku di Indonesia salah satunya di Yogyakarta. Ditjen Pencegahan dan Pengendalian Penyakit Kemenkes RI mencatat pada tahun 2015, Yogyakarta menempati urutan keempat dengan nilai *incident rate* 92,96. Beberapa faktor yang dapat menyebabkan terjadinya gejala DBD adalah demam tinggi, kadar trombosit turun, dan kadar hematokrit naik hingga 20%. Berdasarkan faktor-faktor tersebut dibutuhkan pencarian informasi mengenai pola data penyakit DBD sebagai gambaran umum dalam melihat pola data penyakit tersebut. Oleh karena itu penelitian ini bertujuan untuk melihat pola penyakit DBD berdasarkan faktor-faktor yang mempengaruhinya. Salah satu metode untuk melihat pola data adalah metode Association Rules. Metode Association Rules merupakan suatu prosedur untuk mencari hubungan antar item dalam suatu data set yang ditentukan. Berdasarkan hasil dan pembahasan menghasilkan aturan asosiasi optimum sebanyak 4 kombinasi ( Large 5 itemset ) dengan nilai minimum support 0.41 = 41%, nilai confidence (tingkat kepercayaan) 1 = 100 %, dan lift ratio (tingkat akurasi asosiasi) = 1,619 dari seluruh faktor DBD.

**Kata Kunci:** Association Rule, DBD, Data Mining

## 1. PENDAHULUAN

Penyakit Demam Berdarah *Dengue* (DBD) masih merupakan salah satu masalah kesehatan masyarakat yang utama di Indonesia. Seiring dengan meningkatnya mobilitas dan kepadatan penduduk, jumlah penderita dan luas daerah penyebarannya semakin bertambah. Di Indonesia, demam berdarah pertama kali ditemukan di kota Surabaya pada tahun 1968, sebanyak 58 orang terinfeksi dan 24 orang diantaranya meninggal dunia, dengan Angka Kematian (AK) 41,3 %. Sejak saat itu, penyakit ini menyebar luas keseluruh provinsi di Indonesia (Kementrian Kesehatan RI, 2010). Salah satu provinsi yang terkena dampak dari penyakit ini adalah Yogyakarta. Ditjen Pencegahan dan Pengendalian Penyakit Kemenkes RI mencatat pada tahun 2015, Yogyakarta menempati urutan keempat setelah Bali, Kalimantan Timur, Kalimantan Utara, dan DKI Jakarta dengan nilai *Incidence Rate* (IR) DBD per 100.000 penduduk di Indonesia sebesar 92,96. Dilansir dari salah satu media informasi Yogyakarta yaitu Tribun Jogja, pada tahun 2016 mencatat jumlah penderita demam berdarah di Yogyakarta sudah mencapai 1.705 orang, sehingga pemerintah harus mewaspadaai mewabahnya penyakit ini. DBD merupakan penyakit infeksi yang diakibatkan oleh satu dari 4 virus *dengue* berbeda dan ditularkan melalui nyamuk terutama *Aedes aegypti* dan *Aedes albopictus* yang ditemukan didaerah

tropis dan subtropis diantaranya kepulauan di Indonesia hingga bagian utara Australia (Vyas,2013).

Ancaman penyakit DBD ini berlaku di semua wilayah di Indonesia. Salah satu wilayah yang tidak luput dari bahaya penyakit ini adalah Kota Yogyakarta. Kota Yogyakarta merupakan salah satu kota yang waspada terhadap ancaman penyakit DBD. Puskesmas Cangkringan merupakan salah satu bentuk pelayanan kesehatan di daerah Sleman, Yogyakarta. Hasil rekapitulasi data kesehatan mengenai penyakit DBD mencatat bahwa Yogyakarta merupakan provinsi ke empat dari 5 provinsi yang memiliki nilai *incidence rate* cukup tinggi. Penyakit ini menjadi ancaman bagi masyarakat karena infeksi dari virus tersebut menyebabkan berbagai gejala seperti demam, pusing, nyeri pada bola mata, otot, sendi, dan ruam. Orang yang terinfeksi virus *dengue* juga sering kali mengalami kelelahan jangka panjang. Penyakit ini dapat ditandai dengan turunnya kadar trombosit, naiknya hematokrit mencapai 20%, demam tinggi dan sebagainya.

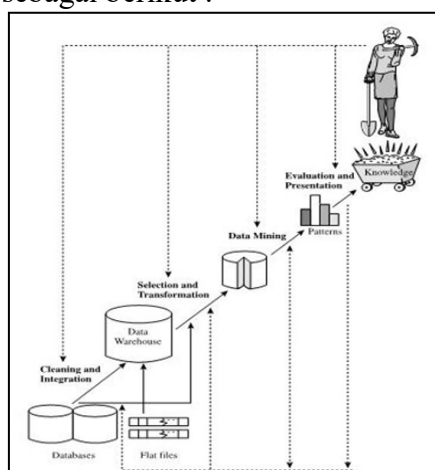
Berdasarkan data diatas, untuk menemukan pola penyakit diperlukan metode terhadap pola data. Identifikasi dan pencarian informasi mengenai pola penyakit DBD diperlukan sebagai gambaran umum dalam melihat pola penyakit DBD berdasarkan faktor pengaruhnya. Salah satu metode yang cocok untuk digunakan menganalisa hubungan pola data antara *itemset* satu dengan yang lainnya adalah metode *Data Mining Association Rule* dengan menggunakan *Algoritma Apriori*, dapat menampilkan informasi berupa *support* dan *confidence* hubungan antara pola penyakit DBD dengan atribut pendukung seperti jenis kelamin, umur, hail pemeriksaan, diagnosis, kadar trombosit dan persentase hematokrit. *Association Rule* (Aturan Asosiatif) berusaha menemukan aturan-aturan tertentu yang mengasosiasikan data satu dengan data yang lainnya (Fadlina, 2014). Beberapa penelitian terdahulu yang berbasis karya ilmiah menggunakan *Data Mining Association Rule* dibidang kesehatan diantaranya [1] (Kuswardani, 2011). : Metode *Association Rule* untuk analisis citra Ct organ pasien kanker ovarium hongon. [2] (Budiono, Fahmi, & Pujiono, 2014) : Penerapan Metode *Association Rule* Menggunakan *Algoritma Apriori* Untuk Mengidentifikasi Pola Penyakit Radang Sendi. [3] (Anggraeni, 2012) : Aplikasi *Data Mining* Analisis Data Transaksi Penjualan obat menggunakan *Algoritma Apriori*. Sehingga Penelitian ini bertujuan untuk mengetahui karakteristik penyakit DBD di Puskesmas Cangkringan Sleman. Serta mengetahui kecenderungan pola hubungan atribut pendukung terhadap penyakit DBD di Puskesmas Cangkringan Sleman.

## 2. METODE PENELITIAN

Penelitian ini dilakukan bertempat di Puskesmas Cangkringan Sleman dari tanggal 27 September - 28 Desember 2017. Sumber data dalam penelitian ini diambil dari hasil test laboratorium yang terkumpul di puskesmas Cangkringan Sleman, Yogyakarta. Metode pengumpulan data yang dilakukan adalah metode pengumpulan data sekunder, yaitu dimana data-data yang menjadi bahan penelitian diperoleh langsung dari tempat penelitian.

Penelitian ini menggunakan satu variabel *decision* dan lima variabel *condition*, beberapa variabel yang dimaksud adalah Jenis Kelamin, Umur, Hasil *Test*, *Diagnosis*, *Kadar Trombosit* (Komponen darah), dan *Persentase Hematokrit* (jumlah sel darah). Tujuan dalam penelitian ini untuk mengetahui

karakteristik penyakit DBD dan menemukan kecenderungan pola hubungan atribut pendukung terhadap penyakit DBD di Puskesmas Cangkringan Sleman. Berdasarkan tujuan tersebut, digunakan metode *data mining* dengan teknik *Algoritma Apriori Association rule*. *Data mining* merupakan istilah yang sering dikatakan sebagai suatu cara untuk menguraikan serta mencari penemuan berupa pengetahuan di dalam suatu *database*. *Data mining* adalah proses pemilihan atau “menambang” pengetahuan dari sekumpulan data dalam jumlah yang banyak (Jiawei & Kamber, 2000). *Data mining* juga sering disebut sebagai kegiatan mengeksplorasi dan menganalisis data dalam jumlah yang besar untuk menemukan *pattern* dan *rule* yang berarti (Berry & Gordon S, 2004). Adapun tahapan *data mining* sebagai berikut :



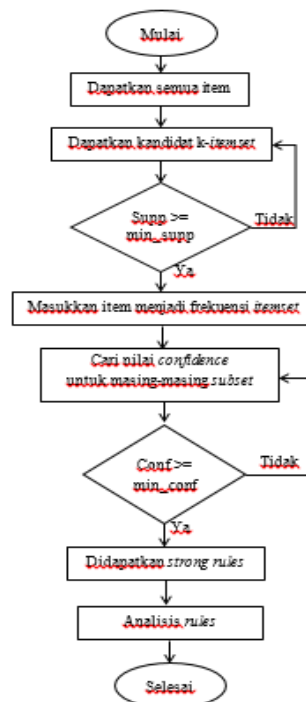
**Gambar 1.** Bidang Ilmu Data Mining

Sumber: (Han, J. dan M. Kamber. 2006)

Berdasarkan gambar diatas berikut uraian tahapan *Data Mining* :

- a. Data Cleaning : Proses pembersihan data dari *dataset*.
- b. Integrasi Data : Mengidentifikasi entitas-entitas pada atribut-atribut *dataset*.
- c. Transformasi Data : Proses mengubah suatu *dataset* kedalam bentuk yang sesuai untuk dianalisis.
- d. Proses Mining : Proses teknik *data mining* yang digunakan untuk menggali informasi pada data bersekala besar.
- e. Evaluasi Pola : Proses identifikasi pola pada hasil *mining*.
- f. Persentasi : Proses menginterpretasi hasil yang didapatkan dari hasil analisis.

Dalam Han dan Kamber, 2006, didefinisikan bahwa “*apriori is an influential algorithm for mining frequent itemsets for Boolean association rules*”. Apriori adalah suatu algoritma yang sudah sangat dikenal dalam melakukan pencarian *frequent itemset* dengan *association rule*. Sesuai dengan namanya, algoritma ini menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya untuk memproses informasi selanjutnya. Algoritma inilah yang biasanya dipakai dalam proses *data mining* untuk *marketbasket analysis*. Algoritma apriori memakai pendekatan *iterative (level-wise search)*, dimana *k-itemset* dipakai untuk menyelidiki  $(k+1)$ -*itemset*. Adapun tahapan dalam algoritma apriori digambarkan dalam diagram alir berikut :



Gambar 2. Diagram Alir Algoritma Apriori

*Association rule mining* adalah suatu prosedur untuk mencari hubungan antar item dalam suatu data set yang ditentukan (Han & M. Kamber, 2006). *Association rule* meliputi dua tahap yaitu mencari kombinasi yang paling sering terjadi dari suatu itemset dan mendefinisikan *condition* dan *result* (untuk *conditional association rule*) (Rindengan, 2012). Menurut (Susanto & Sudiyanto, 2014) Aturan asosiasi yang berbentuk “*if...then...*” atau “*jika...maka...*” merupakan pengetahuan yang dihasilkan dari fungsi aturan asosiasi (Fadlina, 2014). Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu, yaitu :

#### a. Nilai Support

*Support* suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item/itemset* dari keseluruhan transaksi. Menurut (Brin, S. & Silverstain, 1997) *Support* dirumuskan sebagai berikut :

$$Supp(X \Rightarrow Y) = \frac{n(X \cap Y)}{n(D)} \quad (1)$$

#### b. Nilai Confidence

*Confidence* merupakan suatu ukuran yang menunjukkan hubungan antar 2 item secara *conditional* (misal, seberapa sering item B dibeli jika orang membeli item A). Menurut (Zhao, 2013) *Confidence* dirumuskan sebagai berikut :

$$Conf = P(X \cap Y) = \frac{P(X \cap Y)}{P(X)} \quad (2)$$

#### c. Nilai Lift Ratio

*Lift Ratio* merupakan suatu ukuran dalam mengetahui kekuatan suatu aturan asosiasi. Menurut (Brin, S. & Silverstain, 1997) *lift ratio* didefinisikan sebagai :

$$Lift(X \Rightarrow Y) = \frac{P(X \cap Y)}{P(X)P(Y)} \quad (3)$$

### 3. HASIL PENELITIAN DAN PEMBAHASAN

#### a. Karakteristik Pasien DBD

Sebelum menentukan pola asosiasi diantara atribut-atribut penyakit DBD terlebih dahulu akan mendeskripsikan keadaan pasien DBD berdasarkan hasil diagnosis di puskesmas Cangkringan Sleman. Variabel (*decision*) yang digunakan adalah hasil diagnosis pasien yang terdiri dari Demam Berdarah Dengue (DBD), Demam Dengue (DD), dan Sehat. Jika dilihat dari tabel 1 menunjukkan jumlah pasien yang di diagnosis DBD terdapat 21 pasien dari 34 pasien, untuk hasil diagnosis DD sebanyak 10 pasien dan hasil diagnosis sehat terdapat 3 orang dari 34 pasien.

**Tabel 1.** Jumlah dan persentase pasien berdasarkan hasil diagnosis

Diagnosis	Jumlah pasien	Persentase
DBD	21	61.76%
DD	10	29.41%
Sehat	3	8.82%

Dalam mendiagnosis penyakit demam berdarah terdapat beberapa parameter yang dapat dijadikan keputusan salah satunya adalah kadar trombosit dan persentase hematokrit.

**Tabel 2.** Hasil Diagnosa

Diagnosis	Kadar trombosit (ribu/mm <sup>3</sup> )	Persentase Hematokrit (%)
DBD	49 ribu/mm <sup>3</sup>	45.14%
DD	61 ribu/mm <sup>3</sup>	44.29%
Sehat	47 ribu/mm <sup>3</sup>	45.52%

Berdasarkan Tabel 2 diatas menunjukkan hasil rata-rata diagnosa pasien DBD di Puskesmas Cangkringan. Jika dilihat dari diagnosa DBD memiliki rata-rata kadar Trombosit sebesar 49 ribu/mm<sup>3</sup> dan persentase Hematokrit sebesar 45,14 %. Berdasarkan Trombosit normal : 15-40 ribu/mm<sup>3</sup>. Menurut Amrina dkk pada tahun 2014 dalam jurnal kesehatan Andalas menyatakan kesimpulan bahwa hubungan antara Trombosit dan Hematokrit pada pasien DBD berbanding terbalik. Semakin besar nilai suatu variabel maka nilai variabel lainnya akan semakin kecil. Kesimpulan yang didapatkan yaitu semakin naik nilai persentase Hematokrit maka kadar trombosit akan semakin menurun (Rasyada, 2014). Begitu juga dari tabel 2 terlihat keadaan pasien DBD mengalami penurunan kadar Trombosit dan mengalami kenaikan persentase Hematokrit. Hasil diagnosa yang menyebutkan pasien terkena DD memiliki rata-rata kadar Trombosit sebesar 61 ribu/mm<sup>3</sup> dan persentase Hematokrit sebesar 44,29 %. Perbedaan antara DBD dan DD terlihat pada gejala awal. Diagnosa dari DBD dan DD memiliki persamaan, adapun Perbedaan utama DD dan DBD adalah pada DBD ditemukan adanya kebocoran cairan darah pada pembuluh darah, yang menyebabkan darah mengental.

### b. Association Rule

Penentuan pola asosiasi dari faktor-faktor terjadinya DBD dalam penelitian ini digunakan teknik Association Rules. *Association Rule* merupakan salah satu teknik data mining untuk menentukan pola hubungan “jika-maka” antara suatu kombinasi *item Association Rules* akandicari menggunakan algoritma *Apriori*, denganbatasan minimal *support* dan *confidence* yang telah ditentukan oleh peneliti. Dalam penelitian ini dilakukan *trial and error* untuk menemukan aturan asosiasi yang optimal dari penentuan pola penyakit DBD.

#### 1) Analisis 3 Kombinasi *itemset*

Kombinasi tiga *itemset* menunjukkan hubungan dua atribut, misal jika terjadi A diikuti dengan B maka akan terjadi C. pembentukan pola frekuensi tiga *itemset* dibentuk dari *items* sejenis yang memenuhi nilai *support* minimal yaitu dengan cara mengkombinasi semua item kedalam pola tiga kombinasi. Penelitian ini menggunakan beberapa nilai minimum *support* dan *confidence*. Lihat tabel 3 berikut ini :

**Tabel 3.** Kombinasi 3 *Itemset*

Kombinasi	Support	Confidance	Aturan
3	0.1	1	114
	0.2	1	42
	0.3	1	11
	0.4	1	5

Berdasarkan Tabel 3 didapatkan aturan yang paling optimum dengan menggunakan nilai minimum *support*  $0.4 = 40\%$  dan *confidecne*  $1 = 100\%$ . Berikut aturan yang terbentuk.

**Tabel 4.** Aturan Asosiasi 1

No	Aturan	Supp	Conf	Lift
1	{DBD, Hm_naik} => Demam	0.41	1	1.06
2	{Demam, Hm_naik} => DBD	0.41	1	1.62
3	{DBD, Tb_rendah} => Demam	0.44	1	1.06
4	{Laki-laki, Pendarahan} => Demam	0.41	1	1.06
5	{DBD, Laki-laki} => Demam	0.41	1	1.06

Berdasarkan batasan yang telah ditentukan, diketahui bahwa terdapat 5 aturan asosiasi yang terbentuk dengan iterasi sebanyak 3 kali (*large 3-itemset*). Tabel 4 pada baris ke-2 menunjukkan bahwa setiap faktor DBD dengan hasil pemeriksaan demam dan persentase Hematokrit naik maka akan menghasilkan diagnosa DBD dengan nilai *support* (tingkat dominasi *itemset*)  $0.41 = 41\%$ , nilai *confidence* (tingkat kepercayaan)  $1 = 100\%$ , dan *lift ratio* (tingkat akurasi asosiasi)  $= 1,62$  dari seluruh faktor DBD.

#### 2) Analisis 4 Kombinasi *itemset*

Kombinasi tiga *itemset* menunjukkan hubungan dua atribut, misal jika terjadi A di ikuti dengan B dan C maka akan terjadi D. Penelitian ini kembali menggunakan beberapa nilai minimum *support* dan *confidence* dengan kombinasi 4 *itemset*. Lihat tabel 5 berikut ini :



**Tabel 5.** Kombinasi 4 *Itemset*

Kombinasi	Support	Confidance	Aturan
4	0.1	1	15
	0.2	1	11
	0.3	1	2

Berdasarkan Tabel 3 didapatkan aturan yang paling optimum dengan menggunakan nilai minimum  $support_{0.3} = 30\%$  dan  $confidance_{1} = 100\%$ . Berikut aturan yang terbentuk :

**Tabel 6.** Aturan Asosiasi 2

No	Aturan	Support	Confidance	Lift
1	{DBD, Hm_naik, Tb_rendah} => Demam	0.38	1	1.06
2	{Demam, Hm_naik, Tb_rendah} => DBD	0.38	1	1.62

Berdasarkan batasan yang telah ditentukan, diketahui bahwa terdapat 2 aturan asosiasi yang terbentuk dengan iterasi sebanyak 4 kali (*large 4-itemset*). Tabel 6 pada baris ke-2 menunjukkan bahwa setiap faktor DBD dengan hasil pemeriksaan demam, persentase Hematokrit naik dan kadar Trombosit rendah maka akan menghasilkan diagnosa DBD dengan nilai  $support$  (tingkat dominasi *itemset*)  $0.38 = 38\%$ , nilai  $confidence$  (tingkat kepercayaan)  $1 = 100\%$ , dan  $lift\ ratio$  (tingkat akurasi asosiasi) = 1,619 dari seluruh faktor DBD. Sehingga dalam penelitian ini hanya akan dipilih nilai support dan confidence yang paing besar yaitu dengan menggunakan nilai minimum  $support_{0.4}$  dan  $confidence_{1}$ , seluruh faktor DBD.

#### 4. SIMPULAN

Adapun kesimpulan dalam penelitian ini adalah :

- Berdasarkan karakteristik jenis kelamin didapatkan bahwa di Puskesmas Cangkringan Sleman lebih banyak laki-laki dibanding dengan perempuan yang didiagnosa penyakit DBD dengan presentase yakni 68 % untuk pasien laki-laki dan 32 % perempuan dari 34 pasien.
- Hubungan antara kadar trombosit dan hematokrit berbanding terbalik dengan hasil semakin besar nilai suatu variabel maka nilai variabel lainnya akan semakin kecil.
- Analisis dengan 3 iterasi menghasilkan 5 (*Large 5 itemset*) aturan asosiasi dengan nilai support (tingkat dominasi *itemset*)  $0.4 = 40\%$ , nilai  $confidence$  (tingkat kepercayaan)  $1 = 100\%$  menghasilkan setiap faktor DBD dengan hasil pemeriksaan demam dan persentase Hematokrit naik maka akan menghasilkan diagnosa DBD dengan nilai support (tingkat dominasi *itemset*)  $0.411 = 41\%$ , nilai  $confidence$  (tingkat kepercayaan)  $1 = 100\%$ , dan  $lift\ ratio$  (tingkat akurasi asosiasi) = 1,619 dari seluruh faktor DBD.

Adapun Saran dalam penelitian ini adalah :

- Menambahkan pengujian pada pola asosiasi yang terbentuk dengan menggunakan data latih dan data uji sehingga dapat memberikan informasi mengenai tingkat akurasi rule yang dihasilkan.

- b. Menambahkan data sehingga pola asosiasi dapat lebih maksimal dalam menggambarkan pola asosiasi penyakit DBD.

## 5. DAFTAR PUSTAKA

- Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Verkamo, I. (1996). Fast Discovery of Association Rules. *AKDDM, AAAI/MIT Press*, 307-328.
- Anggraeni, H. (2012). Aplikasi Data Mining Analisis Data Transaksi Penjualan obat menggunakan Algoritma Apriori . *Jurnal Masyarakat Iformtik*, 2086-4930.
- Budiono, Fahmi, A., & Pujiono. (2014). Penerapan Metode Association Rule Menggunakan Algoritma Apriori Untuk Mengidentifikasi Pola Penyakit Radang Sendi. *Techno.COM*, 115-124.
- Berry, M., & Gordon S, L. (2004). *Data Minig Techniques : For Marketing, Sales, and Customer Relationship Management.* : John Willey & Sonns, Inc.
- Brin, S., M., & Silverstain. (1997). Beyond Market Baskets : General-Izing Association Rule to Correlations. *Proceedings ACM SIGMOID Conference on Management of Data (SIGMOID1997)*, (hal. 265-276).
- Budiono, Fahmi, A., & Pujiono. (2014). Penerapan Metode Association Rule Menggunakan Algoritma Apriori Untuk Mengidentifikasi Pola Penyakit Radang Sendi. *Techno.COM*, 115-124.
- Fadlina. (2014). Data Mining untuk Analisis Tingkat Kejahatan Jalanan dengan Algoritma Association Rule Metode Apriori. *Jurnal Informasi dan Teknologi Ilmiah (INTI)*, Vol 3, No.1 PP 144-154.
- Han, J., & M. Kamber. (2006). *Data Mining : Concepts and Techniques, Second Edition.* San Fransisco: Morgan Kaufmann Publisher.
- Jiawei, H., & Kamber, M. (2000). *Data Mining: Concepts and Techniques : Chapter 6. Mining Association Rules in Large Databases.* Los Angles: Simon Fraser University.
- Kementrian Kesehatan RI. (2010). *Demam Berdarah Dengue di Indonesia Tahun 1968-2009.* Jakarta: Buletin Jendela Epidemiologi Agustus 2010.
- Kuswardani, d. (2011). Metode Association Rule untuk Analisis Citra CT Organ Pasien Kanker Ovarium Kursor. *Jurnal Science*, vol 6 No. 2.
- Mansjoer. (2001). *Penyakit Autoimun dan Gejalanya.* Surabaya: Fakultas Kedokteran Unair.
- Rasyada, A. (2014). Hubungan Nilai Hematokrit pada Jumlah Trombosit Pada Penderita Dema Berdarah Dengue. *Jurnal Kesehatan Andalas*, 3(3).
- Rindengan, A. (2012). *Perandingan Association Rule Berbentuk Binner dan Fuzzy C-Partition Pada Analisis Market Basket dalam Data Mining.* Manado: Skripsi Universitas Sam Ratulangi.
- Susanto, H., & Sudiyanto. (2014). Data Mining untuk Memprediksi Prestasi Sisa Berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan, dan Prestasi Masa Lalu. *Jurnal Pendidikan Vokasi*, Vol 4 No2.
- Ulmer, D. (2002). *Mining an Online Auctions Data Warehouse.* Pace Univeristy: Proceeding of MASPLAS'02 The Mid-Atlantic Students Workshop on Prograamming Languages and System.



- Waluyo, S., & Putra, B. (2012). *100 Question and*. Jakarta: PT Elex Media Komputindo.
- Zhao, Y. (2013). *R and Data Mining : Examples and Case Studies*. London: Elsevier.