

# ANALISIS FAKTOR – FAKTOR YANG MEMPENGARUHI KESEMBUHAN DAN KEMATIAN PADA PASIEN STROKE DENGAN MENGGUNAKAN ASSOCIATION RULE MINING

Erina Siska Dewi, Wiwik Anggraeni S.Si, M.Kom, dan Renny P Kusumawardani ST.,MT.

Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

*e-mail:* wiwik.jsi@gmail.com

**Abstrak**— Penyakit stroke mengalami peningkatan jumlah penderita dan resiko kematian yang meningkat setiap tahun, oleh karena itu diperlukan penanganan medis yang tepat pada penderita stroke. Pada penelitian ini akan dianalisis faktor – faktor yang mempengaruhi kesembuhan dan kematian pasien stroke selama 14 hari perawatan dan 6 bulan pemantauan. Untuk mengetahui faktor mana yang benar – benar memiliki pengaruh pada kesembuhan dan kematian pasien stroke, akan dilakukan analisis dengan menggunakan association rule mining dan algoritma apriori.

Keluaran yang dihasilkan dari penggalian didalam data training berupa rule yang memenuhi batas ambang minimum support dan minimum confidence. Nilai minimum support ditentukan berdasarkan proporsi data, sedangkan nilai minimum confidence ditentukan minimal 0.6. Kualitas rule yang dihasilkan diukur dengan menggunakan perhitungan support, confidence, lift dan coverage. Setelah penggalian rule dari data training selesai, rule – rule tersebut akan diuji ke data testing dan diukur akurasi. Hasil penelitian yang diperoleh berupa rule yang menggambarkan pola, pola tersebut dapat digali untuk menemukan faktor – faktor yang mempengaruhi diagnosis penyakit stroke dan hasil akhir dari kondisi pasien (meninggal/cacat/sembuh/belum sembuh). Rata – rata akurasi yang didapatkan sebesar 49,3% untuk data diagnosis dan 96% untuk data kelangsungan hidup pasien.

**Kata Kunci**— Association rule mining, algoritma apriori, stroke.

faktor mengenai kondisi dasar pasien sebelum perawatan rumah sakit (umur, jenis kelamin, kondisi fisik dan gejala) serta tindakan medis yang dilakukan oleh rumah sakit terhadap pasien pada awal perawatan sampai pasien sembuh atau meninggal, masa pemantauan perawatan selama 14 hari dan follow up pasien setelah keluar dari rumah sakit selama 6 bulan, data diperoleh dari The International Stroke Trial.

Metode yang digunakan dalam penelitian ini adalah association rule mining, dimana metode ini dapat menghasilkan output berupa pola yang dapat menggambarkan hubungan atau keterkaitan antar banyak faktor. Pada penelitian tugas akhir ini akan dipilih algoritma apriori karena dibandingkan dengan algoritma lain pada penelitian sebelumnya, algoritma apriori dapat melakukan pencarian dalam waktu yang lebih singkat mengingat data yang akan digunakan cukup besar (Nahar, Imam, Tickle, & Chen, 2013). Hasil dari tugas akhir ini diharapkan dapat membantu penanganan pasien stroke dengan lebih tepat sehingga dapat mengurangi kematian pasien rumah sakit akibat penyakit stroke. Tujuan dari pengerjaan tugas akhir ini adalah mengetahui faktor – faktor yang mempengaruhi kesembuhan dan kematian pasien stroke sehingga dapat digunakan untuk membantu penanganan pasien stroke agar lebih baik, dalam arti mengurangi peluang kematian dan meningkatkan peluang kesembuhan dengan mempertimbangkan faktor – faktor yang berpengaruh terhadap keduanya .

## I. PENDAHULUAN

Penyakit stroke merupakan salah satu penyakit yang memiliki resiko kematian yang sangat tinggi, menurut data WHO penyakit ini merupakan penyakit penyebab kematian dengan jumlah terbesar ketiga bila dibandingkan dengan penyakit lain. Jumlah penderita stroke diseluruh dunia mencapai angka 15 juta orang, dari jumlah ini sebanyak 5 juta orang meninggal dan mengakibatkan kecacatan tubuh permanen bagi 5 juta orang lainnya (Mackay & Mensah). Di Indonesia jumlah penderita penyakit stroke meningkat dari tahun ke tahun, menurut riset yang dilakukan oleh Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan RI jumlah penderita meningkat dari 8,3 per 1000 penduduk (2007) menjadi 12,1 per 1000 penduduk (2013) (Badan Penelitian dan Pengembangan Kesehatan, 2013).

Pada tugas akhir ini peneliti akan mencoba menganalisis faktor – faktor yang dapat mempengaruhi kesembuhan dan kematian pasien stroke pada saat pasien stroke dalam masa perawatan rumah sakit. Dalam tugas akhir pasien stroke yang menjadi objek penelitian sedang dalam masa perawatan rumah sakit. Faktor – faktor yang akan dianalisis berupa faktor –

## II. TINJAUAN PUSTAKA

### A. Penyakit Stroke

Penyakit stroke disebabkan oleh terganggunya aliran darah atau asupan darah ke otak, gangguan tersebut disebabkan oleh pembuluh darah yang pecah atau tersumbat. (WHO). Berikut merupakan jenis – jenis penyakit stroke :

#### 1. Stroke Ringan (TIA)

Stroke ringan atau transient ischemic attack (TIA) terjadi ketika asupan oksigen ke otak terhalang sebentar, lalu kembali normal jika asupan terhalang lebih dari 24 jam, maka akan dikategorikan sebagai stroke biasa.

#### 2. Stroke Iskemik

Stroke digolongkan menjadi stroke iskemik jika pembuluh darah yang memasok darah ke otak tersumbat, stroke jenis ini yang umumnya diderita oleh pasien (hampir 90% stroke yang ada adalah

iskemik). Penyebab stroke iskemik adalah penumpukan lemak yang melapisi dinding pembuluh darah sampai lapisan tersebut menjadi plak, plak mengakibatkan darah sulit mengalir dan dapat menyebabkan trombus atau bekuan darah.

**3. Stroke Hemoragik**

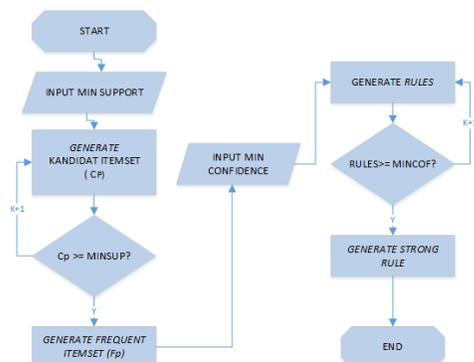
Stroke hemoragik penyebabnya adalah pembuluh darah yang bocor atau pecah di dalam atau di sekitar otak sehingga asupan darah ke bagian otak yang dituju menjadi terhenti

**Subtipe penyakit stroke :**

- a. *Total Anterior Circulation Stroke (TACS)*
- b. *Partial Anterior Circulation Stroke (PACS)*
- c. *Lacunar Stroke (LACS)*
- d. *Posterior Circulation (POCS)*

**B. Association rule mining**

Association rule mining merupakan salah satu dari prosedur data mining yang banyak digunakan. Berfungsi untuk mencari pola yang sering muncul diantara banyak transaksi atau untuk menemukan rule atau aturan antara suatu kombinasi item. Rule berbentuk  $X \rightarrow Y$ , dimana X merupakan antesenden (*left hand side*) dan y adalah konsekwen (*right hand side*). X dan Y merupakan disjoint itemset ( $X \cap Y = \emptyset$ ). Rule tersebut menggambarkan bahwa Y akan terjadi ketika X terjadi, contohnya seseorang yang membeli buku akan membeli pensil dan juga bolpoin ( $\{buku\} \rightarrow \{pensil, bolpoin\}$ ) (Ishibuchi, Kuwajima, & Nojima, 2007).



**Gambar 1 Alur kerja association rule mining**

Secara garis besar metode pengerjaan penelitian ini terdapat pada gambar 2.

**1. Support**

Adalah presentase dari transaksi yang mengandung itemset (X dan Y).

$$Support(X \rightarrow Y) = \frac{\text{jumlah transaksi yang mengandung X dan Y}}{\text{jumlah transaksi keseluruhan}} \quad (1)$$

**2. Confidence**

Merupakan rasio antara jumlah transaksi yang berisi X dan Y dan jumlah transaksi yang berisi X. Minimum confidence menggambarkan batas ambang kuatnya hubungan antar item dalam aturan asosiatif.

$$Confidence(X \rightarrow Y) = \frac{\text{jumlah transaksi yang mengandung X dan Y}}{\text{jumlah transaksi yang mengandung X}} \quad (2)$$

**3. Lift**

Lift mengukur tingkat kepentingan dari suatu rule.

$$lift(X \rightarrow Y) = lift(Y \rightarrow X) = \frac{conf(X \rightarrow Y) / supp(Y)}{conf(Y \rightarrow X) / supp(X)} = \frac{P(X \text{ and } Y) / (P(X)P(Y))}{P(X \text{ and } Y) / (P(X)P(Y))} \quad (3)$$

**4. Coverage**

Mengukur seberapa besar cakupan implementasi rule pada keseluruhan data. Dapat dihitung dari support lhs (*left hand side*).

**Tahapan Association Rule Mining**

**a. Frequent itemset generation dengan menggunakan algoritma apriori.**

Sebelum menjalankan algoritma ini, parameter minsup atau persentase kemunculan suatu rules harus ditentukan terlebih dahulu. Selain itu, minimum confidence atau persentase kemunculan rules terhadap kemunculan antesedent-nya juga harus ditentukan terlebih dahulu. Algoritmanya adalah sebagai berikut :

1. misal k =1
2. bentuk frequent item sets yang terdiri dari k-item
3. ulangi hingga tidak ada lagi frequent item sets yang baru
4. bentuk kandidat item sets dengan panjang k+1 dari frequent item sets dengan panjang k
5. buang kandidat item sets yang berisi subset dengan panjang k yang tidak frequent
6. hitung support dari setiap kandidat dengan scanning basisdata
7. eliminasi kandidat yang infrequent (Minimum Apriori, 2013)

**b. Rule generation**

Setelah mendapatkan frequent itemset hasil dari tahapan sebelumnya, tahap selanjutnya adalah mendapatkan rule yang memenuhi confidence ( lebih besar atau sama dengan batas ambang yang ditentukan). Tidak perlu dilakukan pehitungan support karena rules yang akan diproses dalam tahap ini telah memenuhi minsup yang telah ditentukan. (Tan, Steinbach & Kumar, 2005).

**III. PENYIAPAN DATA**

Data yang akan diolah berasal dari database International Stroke Trial, data akan dibagi menjadi dua bagian, yang pertama adalah data diagnosis dengan variabel hasil diagnosis pasien. Bagian yang kedua adalah data *survival*, berisi informasi perawatan pasien dengan variabel hasil OCCODE (berisi status terakhir pasien/meninggal/sembuh/cacat)

Tahapan pertama yang harus dilakukan sebelum proses association rule mining adalah penyiapan data. Penyiapan data adalah proses pembersihan data dan penyesuaian data terhadap format yang dibutuhkan untuk diproses lebih lanjut. Pada tugas akhir ini penyiapan data terdiri dari beberapa tahapan, yaitu pemilihan atribut yang relevan, konversi atribut bernilai kontinyu menjadi kategorikal dan pembagian data training dan testing dengan proporsi 70% dan 30%.

IV. ASSOCIATION RULE MINING

A. Penggalian Rule Dari Data Training

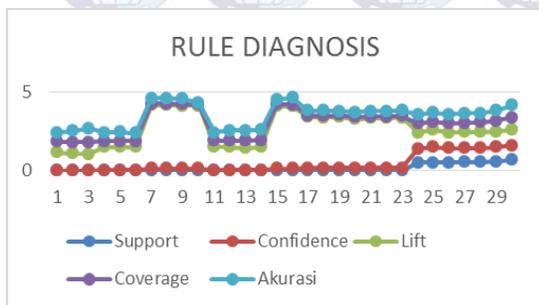
Untuk melakukan tahap ini digunakan aplikasi RStudio, penggalian rule menggunakan metode *association rule mining* dengan algoritma apriori. Karena distribusi data yang berbeda – beda, maka minimum support dan confidence setiap kelas atribut hasil juga berbeda – beda. Hal ini bertujuan agar algoritma ini dapat mengeluarkan pola dari setiap kelas yang terdapat didalam atribut hasil. Yang dimaksud atribut hasil adalah *consequent* atau *right hand side*. Tabel dibawah ini berisi proporsi data dan minimum support & confidence dari setiap kelas. Setelah rule dari data training dihasilkan, akan dilakukan pemangkasan rule (pruning) berdasarkan nilai lift. Rule yang dipangkas merupakan super rule yang memiliki nilai lift tidak lebih besar dari pada rule dibawahnya.

Tabel 1 Proporsi dan minimum support&confidence

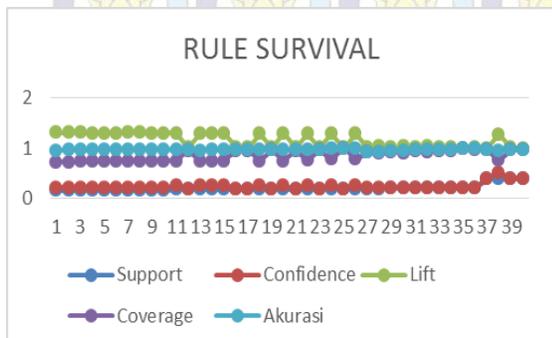
TRAINING	PROPORSI DATA		Minsup	Mincof
	Jumlah	%		
Ischaemic	12181	89.540	0.448	0.537
Haemorrhagic	283	2.080	0.010	0.012
Indeterminate	427	3.139	0.016	0.019
Not a stroke	697	5.123	0.026	0.031
unknown	16	0.118	0.001	0.001
Total	13604	100.000	0.500	0.600
Dead	2068	21.718	0.109	0.130
Recover	1615	16.961	0.085	0.102
Not recover	1887	19.817	0.099	0.119
Dependent	3870	40.643	0.203	0.244
missing status	82	0.861	0.004	0.005
Total	9522	100.000	0.500	0.600

B. Pengujian Rule pada Data Testing

Pengujian rule hasil pruning pada data testing bertujuan untuk mengetahui akurasi rule.



Gambar 2 Kualitas rule diagnosis



Gambar 3 Kualitas rule survival

Pada gambar 3 dan 4 dibawah ini, menunjukkan semakin besar nilai minimum *support* dan minimum *confidence* maka jumlah rule yang dihasilkan semakin sedikit. Nilai akurasi pada data survival lebih stabil dari pada data diagnosis, karena distribusi data survival lebih merata. Nilai support tidak begitu berpengaruh pada fluktuasi nilai kualitas rule, sebaliknya nilai confidence sangat berpengaruh terhadap nilai kualitas rule. Semakin besar nilai confidence maka nilai lift dan coverage juga semakin besar. Namun hal ini tidak berlaku apabila nilai support naik secara signifikan, nilai lift dan coverage akan menurun karna cakupan data semakin kecil meskipun nilai confidencenya besar.

C. Pembobotan faktor – faktor yang berpengaruh terhadap data diagnosis dan data survival

Setiap item didalam rule yang dihasilkan dipertimbangkan sebagai faktor yang mempengaruhi konsekwen atau *right hand side*, besar pengaruh item / faktor tersebut akan diketahui berdasarkan nilai bobot. Bobot didapatkan dari perbandingan kemunculan faktor tersebut terhadap jumlah rule yang dihasilkan. Faktor dianggap memiliki pengaruh besar apabila memiliki nilai 50% atau jika tidak ada yang memiliki bobot 50% akan dipilih item / faktor yang memiliki bobot diatas rata – rata nilai bobot pada kelas faktor yang bersangkutan.

V. HASIL

A. Data Diagnosis

Tabel 2 sampai dengan 6 menunjukkan rule atau pola yang terbentuk pada masing – masing kelas pada atribut diagnosis.

Tabel 2 Rule Ischaemic

No	RULE	Akurasi
1	{RATRIAL=N,RCT=Y} => {DIAGNOSIS=Ischaemic}	58.175%
2	{RDEF1=Y,RDEF2=Y,RDEF7=N} => {DIAGNOSIS=Ischaemic}	56.891%
3	{RCONSC=Y,RDEF2=Y,RDEF7=N}=> {DIAGNOSIS=Ischaemic}	57.658%
4	{RCONSC=F,RDEF6=N} => {DIAGNOSIS=Ischaemic}	60.974%
5	{RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Ischaemic}	60.073%
6	{RCONSC=F} => {DIAGNOSIS=Ischaemic}	78.072%
7	{RDEF6=N} => {DIAGNOSIS=Ischaemic}	66.763%

Tabel 3 Rule Haemorrhagic

No	RULE	Akurasi
1	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
2	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF1=Y,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
3	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
4	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	9.884%
5	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
6	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	44.186%

**Tabel 4 Rule Indeterminate**

No	RULE	Akurasi
1	{RCT=N,RVISINF=N,RDEF1=Y,RDEF2=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	32.203%
2	{RCT=N,RDEF1=Y,RDEF2=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	32.203%
3	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Indeterminate}	25.424%
4	{SEX=F,RCT=N,RDEF1=Y} => {DIAGNOSIS=Indeterminate}	35.254%
5	{RCT=N,RVISINF=N,RDEF1=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%
6	{RCT=N,RDEF1=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%
7	{RCT=N,RVISINF=N,RDEF2=Y,RDEF3=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%

**Tabel 5 Rule Not a stroke**

No	RULE	Akurasi
1	{RATRIAL=N,RVISINF=N,RDEF6=N} => {DIAGNOSIS=Not a stroke}	52.555%
2	{RVISINF=N,RDEF4=N} => {DIAGNOSIS=Not a stroke}	54.015%
3	{RSLEEP=N,RATRIAL=N,RVISINF=N} => {DIAGNOSIS=Not a stroke}	56.934%
4	{RATRIAL=N,RVISINF=N,RDEF5=N} => {DIAGNOSIS=Not a stroke}	51.825%
5	{RVISINF=N,RDEF6=N} => {DIAGNOSIS=Not a stroke}	64.964%
6	{RCONSC=F,RATRIAL=N,RVISINF=N} => {DIAGNOSIS=Not a stroke}	62.774%
7	{RVISINF=N,RDEF5=N} => {DIAGNOSIS=Not a stroke}	64.234%

**Tabel 6 Rule unknown**

No	RULE	Akurasi
1	{RVISINF=N} => {DIAGNOSIS=unknown}	60.000%
2	{RSLEEP=N} => {DIAGNOSIS=unknown}	70.000%
3	{RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=unknown}	90.000%

**B. Data Survival**

Tabel 7 sampai dengan 10 berisi rule yang terbentuk pada masing – masing kelas pada atribut OCCODE.

**Tabel 7 Rule Dead**

No	RULE	Akurasi
1	{DAP=N,DOAC=N,DHAEMD=N} => {OCCODE=DEAD}	93.222%
2	{DAP=N,DOAC=N,DTHROMB=N} => {OCCODE=DEAD}	94.333%
3	{DAP=N,DOAC=N} => {OCCODE=DEAD}	96.111%
4	{DOAC=N,DTHROMB=N} => {OCCODE=DEAD}	96.667%
5	{DOAC=N} => {OCCODE=DEAD}	98.556%
6	{DAP=N,DHAEMD=N,DTHROMB=N} => {OCCODE=DEAD}	92.778%
7	{DAP=N,DHAEMD=N} => {OCCODE=DEAD}	94.667%
8	{DAP=N,DTHROMB=N} => {OCCODE=DEAD}	95.667%
9	{DAP=N} => {OCCODE=DEAD}	97.556%
10	{DTHROMB=N} => {OCCODE=DEAD}	98.000%

**Tabel 8 Rule recover**

No	RULE	Akurasi
1	{DSCH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.83%
2	{DSCH=N,DTHROMB=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.9%
3	{DMAJNCH=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	96.69%
4	{DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	96.9%
5	{DSCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	97.12%
6	{DMAJNCH=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=RECOVER}	97.12%
7	{DTHROMB=N,DMAJNCH=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	95.3%

No	RULE	Akurasi
8	{DTHROMB=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	95.63%
9	{DTHROMB=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=RECOVER}	95.827%
10	{DSCH=N,DTHROMB=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.691%

**Tabel 9 Rule not recover**

No	Rule	Akurasi
1	{DMAJNCH=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=NRECOVER}	95.734%
2	{DCAUSE=unknown,CMPLASP=Y} => {OCCODE=NRECOVER}	96.110%
3	{DSCH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.487%
4	{DIVH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.236%
5	{DSCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.738%
6	{DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	99.624%
7	{DIVH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.487%
8	{DOAC=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	95.232%
9	{DCAUSE=unknown} => {OCCODE=NRECOVER}	100%
10	{DMAJNCH=N,CMPLASP=Y} => {OCCODE=NRECOVER}	95.734%
11	{CMPLASP=Y} => {OCCODE=NRECOVER}	96.110%
12	{DSCH=N,DMAJNCH=N} => {OCCODE=NRECOVER}	96.487%
13	{DIVH=N,DMAJNCH=N} => {OCCODE=NRECOVER}	96.236%
14	{DMAJNCH=N} => {OCCODE=NRECOVER}	99.624%
15	{DSCH=N} => {OCCODE=NRECOVER}	96.738%
16	{DIVH=N} => {OCCODE=NRECOVER}	96.487%

**Tabel 10 Rule dependent**

No	RULE	Akurasi
1	{DIVH=N,DTHROMB=N} => {OCCODE=DEPENDENT}	96.216%
2	{DHAEMD=N,DMAJNCH=N} => {OCCODE=DEPENDENT}	96.396%
3	{DTHROMB=N,DMAJNCH=N} => {OCCODE=DEPENDENT}	97.057%
4	{DTHROMB=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=DEPENDENT}	94.955%

**C. Hasil pembobotan**

Tabel 10 dan 11 berisi hasil pembobotan faktor – faktor pada data diagnosis dan data survival. Faktor yang diambil dari masing – masing kelas adalah yang memiliki bobot diatas 50% atau diatas rata – rata bobot kelas bersangkutan.

**Tabel 11 Tabel bobot faktor pada data diagnosis**

Kelas (Jenis Stroke)	Faktor	Bobot%
Ischaemic	Tidak adanya kelainan detak jantung pada pasien (RATRIAL=N)	14%
Ischaemic	Pasien telah melakukan proses CT-scan (RCT=Y).	14%
Ischaemic	Pasien mengalami gejala defisit pada wajah (RDEF1=Y)	29%
Ischaemic	Pasien mengalami defisit pada kaki (RDEF3=Y)	29%
Ischaemic	Pasien tidak memiliki tanda pendarahan pada otak (RDEF7=N)	29%
Ischaemic	Pasien dalam keadaan sadar (RCONSC=Y)	14%
Ischaemic	Pasien dalam keadaan sangat sadar / fully alert (RCONSC=F)	29%
Ischaemic	Pasien tidak mengalami kelainan visuospasial (RDEF6=N)	29%
Haemorrhagic	Pasien tidak memiliki gejala yang tampak pada saat tidur (RSLEEP=N)	100%
Haemorrhagic	Pasien tidak memiliki kelainan detak jantung (RATRIAL=N)	100%
Haemorrhagic	Pasien tidak melakukan proses CT-scan	100%

Kelas (Jenis Stroke)	Faktor	Bobot%
	(RCT=N)	
Haemorrhagic	Pasien tidak memiliki infark yang terlihat (RVISINF=N)	50%
Haemorrhagic	Pasien memiliki gejala defisit pada wajah (RDEF1=Y)	67%
Haemorrhagic	Pasien memiliki gejala defisit pada tangan/lengan (RDEF2=Y)	67%
Haemorrhagic	Pasien memiliki gejala defisit pada kaki (RDEF3=Y)	100%
Indeterminate	Pasien tidak melakukan CT-scan (RCT=N)	100%
Indeterminate	Tidak terlihat adanya infark (RVISINF=N)	57%
Indeterminate	Terdapat gejala defisit pada wajah (RDEF1=Y)	86%
Indeterminate	Terdapat gejala defisit pada tangan/lengan (RDEF2=Y)	43%
Indeterminate	Terdapat kelainan pemahaman bahasa / dysphasia (RDEF4=Y)	71%
Indeterminate	Tidak ada gejala yang Nampak pada saat pasien tidur (RSLEEP=N)	14%
Indeterminate	Tidak adanya kelainan detak jantung (RATRIAL=N)	14%
Indeterminate	Terdapat gejala defisit pada kaki (RDEF3=Y)	43%
Indeterminate	Jenis kelamin pasien adalah perempuan (SEX=F)	14%
Not a stroke	Pasien tidak memiliki kelainan detak jantung (RATRIAL=N)	57%
Not a stroke	Pasien tidak memiliki infark (RVISINF=N)	100%
Not a stroke	Pasien tidak mengalami gejala kelainan visuospasial (RDEF6=N)	29%
Not a stroke	Pasien tidak mengalami gejala dysphasia (RDEF4=N)	14%
Not a stroke	Pasien tidak memiliki gejala yang tampak pada saat tidur (RSLEEP=N)	14%
Not a stroke	Pasien tidak mengalami gejala hemianopia (RDEF5=N)	29%
Not a stroke	Pasien dalam keadaan sangat sadar / fully alert (RCONS=F)	14%
Unknown	Pasien tidak memiliki infark (RVISINF=N)	33%
Unknown	Pasien tidak memiliki gejala yang tampak pada saat tidur (RSLEEP=N)	33%
Unknown	Pasien mengalami gejala defisit pada tangan/lengan (RDEF2=Y)	33%
Unknown	Pasien mengalami gejala defisit pada kaki (RDEF3=Y)	33%

Tabel 12 Tabel bobot faktor pada data survival

Kelas (Kondisi akhir pasien)	Faktor	Bobot %
Dead	Tidak ada konsumsi obat antiplatelet percobaan (DAP=N)	70%
Dead	Tidak ada konsumsi obat antikoagulan lain (DOAC=N)	50%
Dead	Tidak adanya penambahan volume plasma darah (DHAEMD=N)	30%
Dead	Tidak ada pemecahan gumpalan sumbatan darah / trombolisis (DTHROMB=N)	50%
Recover	Pasien tidak mengkonsumsi heparin yang diberikan dibawah kulit / underskin (DSCH=N)	40%
Recover	Pasien tidak mengalami pendarahan otak yang besar (DMAJNCH=N)	50%
Recover	Penyebab kematian tidak diketahui karena pasien tidak meninggal (DCAUSE=unknown)	100%
Recover	Pasien tidak mengalami pemecahan sumbatan darah / trombolisis (DTHROMB=N)	50%
Recover	Pasien mengkonsumsi heparin sesuai aturan (komplrit) (CMLPLASP=Y)	40%

Kelas (Kondisi akhir pasien)	Faktor	Bobot %
Recover	Pasien mengkonsumsi aspirin sesuai dengan aturan (komplrit) (CMLPLASP=Y)	20%
Dependent	Pasien tidak mengkonsumsi heparin percobaan yang disuntikkan kedalam pembuluh darah / intravenous (DIVH=N)	25%
Dependent	Pasien tidak mengalami pemecahan sumbatan darah / trombolisis (DTHROMB=N)	75%
Dependent	Pasien tidak mengalami peningkatan volume plasma darah / haemodilution (DHAEMD=N)	25%
Dependent	Pasien tidak mengalami pendarahan otak yang besar (DMAJNCH=N)	75%
Dependent	Sebab kematian pasien tidak diketahui karena pasien tidak meninggal (DCAUSE=unknown)	25%
Not Recover	Tidak adanya pendarahan otak yang besar (DMAJNCH=N)	50%
Not Recover	Sebab kematian pasien tidak diketahui karena pasien tidak meninggal (DCAUSE=unknown)	56%
Not Recover	Pasien telah mengkonsumsi aspirin sesuai dengan aturan (komplrit) (CMLPLASP=Y)	25%
Not Recover	Pasien tidak mengkonsumsi heparin yang diberikan dibawah kulit / underskin (DSCH=N)	25%
Not Recover	Tidak mengkonsumsi heparin percobaan yang disuntikkan kedalam pembuluh darah / intravenous (DIVH=N)	25%
Not Recover	Pasien tidak mengkonsumsi obat antikoagulan lain (DOAC=N)	6%

## VI. KESIMPULAN

Kesimpulan yang didapatkan dalam penelitian ini akan dibagi menjadi dua bagian, yang pertama adalah kesimpulan dari proses pengerjaan dan yang kedua adalah kesimpulan mengenai hasil .

### A. Proses association rule mining

1. Distribusi kelas dalam atribut dapat mempengaruhi kualitas data. Data yang memiliki distribusi kelas yang merata memiliki kualitas lebih baik. Dibuktikan dengan rata – rata nilai akurasi data survival lebih baik daripada data diagnosis.
2. Penentuan minimum support dan confidence mempengaruhi banyaknya rule yang dapat dihasilkan,. Minimum support dan minimum confidence ditentukan masing – masing sebesar 50% dan 60% dari jumlah transaksi masing – masing kelas.
3. Penggalan rule pada data diagnosis perlu dilakukan secara terpisah karena atribut yang ada di dalam data tersebut hanya memiliki korelasi kuat terhadap atribut hasil “Diagnosis” dan tidak berkorelasi dengan atribut hasil “OCCODE” pada data survival. Selain itu atribut hasil “Diagnosis” merupakan input dari atribut hasil “OCCODE”.
4. Faktor – faktor yang berpengaruh pada data diagnosis dengan atribut hasil “Diagnosis” adalah sebagai berikut:

#### a. Stroke Ischaemic

- Pasien mengalami gejala defisit pada wajah (RDEF1=Y)
- Pasien mengalami defisit pada kaki (RDEF3=Y)
- Pasien tidak memiliki tanda pendarahan pada otak (RDEF7=N)
- Pasien dalam keadaan sangat sadar / fully alert (RCONS=F)

- Pasien tidak mengalami kelainan visuospatial ( RDEF6=N)
- b. Stroke Haemorrhagic**
- Pasien tidak memiliki gejala yang tampak pada saat tidur (RSLEEP=N)
  - Pasien tidak memiliki kelainan detak jantung (RATRIAL=N)
  - Pasien tidak melakukan proses CT-scan (RCT=N)
  - Pasien tidak memiliki infark yang terlihat (RVISINF=N)
  - Pasien memiliki gejala defisit pada wajah (RDEF1=Y)
  - Pasien memiliki gejala defisit pada tangan/lengan (RDEF2=Y)
  - Pasien memiliki gejala defisit pada kaki (RDEF3=Y)
- c. Indeterminate**
- Pasien tidak melakukan CT-scan (RCT=N)
  - Tidak terlihat adanya infark (RVISINF=N)
  - Terdapat gejala defisit pada wajah (RDEF1=Y)
  - Terdapat kelainan pemahanan bahasa / dysphasia (RDEF4=Y)
- d. Not a stroke (bukan stroke)**
- Pasien terdiagnosis tidak menderita stroke apabila pasien tidak memiliki infark (RVISINF=N)
- e. unknown**
- Pasien tidak memiliki infark (RVISINF=N)
  - Pasien tidak memiliki gejala yang tampak pada saat tidur (RSLEEP=N)
  - Pasien mengalami gejala defisit pada tangan/lengan (RDEF2=Y).
  - Pasien mengalami gejala defisit pada kaki (RDEF3=Y)
5. Faktor – faktor yang berpengaruh pada data survival dengan atribut hasil “OCCODE” adalah sebagai berikut:
- a. Dead (meninggal)**
- Tidak ada konsumsi obat antiplatelet percobaan (DAP=N)
  - Tidak ada konsumsi obat antikoagulan lain (DOAC=N)
  - Tidak ada pemecahan gumpalan sumbatan darah / trombolisis (DTHROMB=N)
- b. Recover (sembuh)**
- Pasien tidak mengalami pendarahan otak yang besar (DMAJNCH=N).
  - Penyebab kematian tidak diketahui karena pasien tidak meninggal (DCAUSE=unknown).
  - Pasien tidak mengalami pemecahan sumbatan darah / trombolisis (DTHROMB=N)
  - Pasien mengkonsumsi heparin sesuai aturan (kompli) (CMPLHEP=Y)
  - Pasien mengkonsumsi aspirin sesuai dengan aturan (kompli) (CMPLASP=Y).
- c. Dependent (sembuh dan cacat)**
- Pasien tidak mengalami pemecahan sumbatan darah / trombolisis (DTHROMB=N)
  - Pasien tidak mengalami peningkatan volume plasma darah / haemodilution (DHAEMD=N)
  - Pasien tidak mengalami pendarahan otak yang besar (DMAJNCH=N)
  - Not recover (belum sembuh)
  - Tidak adanya pendarahan otak yang besar (DMAJNCH=N)

- Sebab kematian pasien tidak diketahui karena pasien tidak meninggal (DCAUSE=unknown)
6. Akurasi rule pada data diagnosis memiliki rata – rata sebesar 49.30% dan data survival sebesar 96.44%. Hal ini disebabkan oleh distribusi kelas yang ada di dalam data survival yang lebih merata.

### B. Saran

1. Untuk hasil yang lebih baik, penelitian sebaiknya dilakukan dengan menggunakan data dengan proporsi kelas yang merata.
2. Algoritma apriori memang efektif untuk menggali rule dari data secara menyeluruh, namun membutuhkan sumber daya yang besar. Jadi diharapkan untuk menggunakan sumber daya yang mencukupi, seperti laptop yang memiliki spesifikasi baik.

### DAFTAR PUSTAKA

- Minimum Apriori*. (2013, Mei 16). Retrieved Oktober 17, 2014, from Nurfadillah100: <http://nurfadillah100.blogspot.com/2013/05/minimum-apriori.html>
- Badan Penelitian dan Pengembangan Kesehatan. (2013). *Riset Kesehatan Dasar Tahun 2013*. Jakarta: Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan RI.
- Budi. (2010, April 17). *Data Mining: Market Basket Analysis for Marketing Strategy*. Retrieved 10 30, 2014, from Statistika Komputasi: <http://statistikakomputasi.wordpress.com/2010/04/17/data-mining-market-basket-analysis-for-marketing-strategy/>
- Dewi, d. (n.d.). *Jenis - jenis stroke*. Retrieved Oktober 17, 2014, from Poliklinik IPDN: <http://poliklinik.ipdn.ac.id/home/artikel-kesehatan/jenis---jenis-stroke>
- Ishibuchi, H., Kuwajima, I., & Nojima, Y. (2007). Prescreening of candidate rules using association rule mining and pareto optimality in genetic rule selection. *Knowledge-based intelligent information and engineering systems*, 509-516.
- Nahar, J., Imam, T., Tickle, K. S., & Chen, Y.-p. P. (2013). Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, 1086 - 1093.
- Oktoaria, R., Maharani, W., & Firdaus, Y. (2010). Content Based Recommender System Menggunakan Algoritma Apriori. *Konferensi Nasional Sistem dan Informatika*.
- Rambe, S. A. (2006, Desember 20). *STROKE: SEKILAS TENTANG DEFINISI, PENYEBAB, EFEK, DAN FAKTOR RESIKO*. Retrieved Oktober 16, 2014, from Departemen Neurologi FK-USU: [http://repository.usu.ac.id/bitstream/123456789/18925/1/ikm-des2006-10%20\(3\).pdf](http://repository.usu.ac.id/bitstream/123456789/18925/1/ikm-des2006-10%20(3).pdf)
- Sandercock, P. A., Niewada, M., & Czlonkowska, A. (2011). The International Stroke Trial database. *trialsjournal*.
- Tan, P. N., Kumar, V., & Srivastava, J. (2004). Selecting the Right Objective Measure for Association Analysis. *Information System*, 293-313.
- Tan, P. N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. USA: Adison-Wesley.
- Wandi, N., Hendrawan, R. A., & Mukhlason, A. (2012). Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalan Association Rule Menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur). *Jurnal Teknik ITS*.