

IMPLEMENTASI METODE *ROUGH SET* DAN *CERTAINTY FACTOR* UNTUK DETEKSI DINI PENYAKIT MENULAR SEKSUAL (STUDI KASUS: PUSKESMAS DINOYO KOTA MALANG)

Ratri Arijaya¹, Imam Cholissodin², Arief Andy Soebroto³

^{1,2,3}Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Jl. Veteran No.8 Malang, Informatika, Gedung A PTIHK – UB
Email : ratriarijaya@gmail.com¹, imamcs@ub.ac.id², ariefas@ub.ac.id³

ABSTRAK

Penyebaran jumlah penderita pengidap penyakit menular seksual di Indonesia maupun dunia dari tahun ke tahun terus meningkat seiring dengan perkembangan pola hidup manusia di era modern saat ini. Penelitian dari WHO (*World Health Organization*) mencatat jumlah penderita penyakit menular seksual di wilayah bagian Asia Tenggara sangat besar dengan perbandingan berupa jumlah kasus baru dari empat infeksi menular seksual pada tahun 2008 dan setiap negara memiliki jumlah kasus baru yang berbeda-beda. Keadaan ini jelas meningkat disebabkan sejumlah faktor termasuk meningkatnya jumlah pemuda aktif secara seksual, perubahan sosial, dan perubahan pola dalam praktik pencegahan dan pengobatan di setiap negara. Di Indonesia sendiri penyebaran penyakit menular seksual sangat luas dan tidak terkendali disebabkan karena kurangnya edukasi terhadap masyarakat dan prasarana untuk memberikan layanan bagi pasien penyakit menular seksual. Penyakit menular seksual memiliki beberapa gejala yang mudah dirasakan dan dilihat oleh penderita namun karena kurangnya edukasi masyarakat dan beberapa faktor alasan yang lain sehingga masyarakat tidak memeriksakan gejala yang dialami lebih lanjut di klinik terdekat. Oleh sebab itu maka dibutuhkan sistem yang dapat mempercepat dan memudahkan pendeteksian dini penyakit menular seksual. Sistem aplikasi untuk deteksi dini penyakit menular seksual digunakan metode 2 metode kombinasi *roughset* dan *certainty-factor*. *Roughset* merupakan suatu metode analisis data yang dikembangkan untuk melakukan pencarian *decision rule* dalam sistem informasi. *Certainty factor* merupakan metode perhitungan nilai keyakinan suatu masukan gejala yang dapat memberikan nilai keyakinan sistem dari hasil deteksi dini penyakit menular seksual oleh sistem. Hasil pengujian akurasi pada sistem aplikasi deteksi dini penyakit menular seksual dengan cara random sampling data uji sebesar 97,81166667%.

Kata Kunci: Metode *roughset*, *Certainty factor*, penyakit menular seksual.

ABSTRACT

*Spreading the number of people with sexually transmitted diseases in Indonesia and the world from year to year continues to increase along with the development of the current lifestyle in modern era. Research from the WHO (World Health Organization) noted the number of people with sexually transmitted diseases in the region of Southeast Asia is very large by comparison of the number of new cases of four sexually transmitted infections in 2008, and every country has a different number of new cases. This situation clearly increased due to several factors including the increasing number of sexually active youth, social change, and changes in the prevention and treatment practices in each country. In Indonesia, the spread of sexually transmitted diseases is very broad and unmanageable due to lack of education of the people and infrastructure to provide services for patients with sexually transmitted diseases. Sexually transmitted diseases have few symptoms easily felt and seen by the patient but due to lack of public education and some other factors are reasons that people do not check their symptoms at a nearby clinic. Therefore, the system needs to accelerate and facilitate the early detection of sexually transmitted diseases. Application systems for early detection of sexually transmitted diseases used 2 methods combination method *roughset* and *certainty-factor*. *Roughset* is a data analysis method was developed to search decision rule in information systems. *Certainty factor* is the method of calculating the value of the assurance an input symptoms that can give the assurance system from the early detection of sexually transmitted disease by the system. The test results on the accuracy of the application system of early detection of sexually transmitted diseases by means of random sampling test data of 97.81166667%.*

Keyword: *roughset method, Certainty factor, sexually transmitted diseases.*

1. PENDAHULUAN

Mikroorganisme penyakit yang jumlahnya lebih dari 30 macam jenis dapat ditularkan melalui hubungan seksual dengan manifestasi klinis bervariasi menurut jenis kelamin dan umur. Dengan perkembangan di bidang sosial, demografik, serta meningkatnya migrasi penduduk kemungkinan tertular Infeksi menular seksual (IMS) akan meningkat pesat. Infeksi menular

seksual (IMS) ditularkan melalui hubungan seksual antara manusia satu dengan manusia yang lain. Penularannya juga dapat terjadi dari ibu kepada janin dalam kandungan atau saat kelahiran. Infeksi menular seksual dapat ditemui melalui produk darah/transfer jaringan yang telah tercemar dan juga dapat ditularkan melalui alat kesehatan (Kesehatan, 2011).

Sarana yang belum memadai untuk pendeteksi dini terhadap penyakit menular seksual yang ada dan kurangnya edukasi tentang penyakit menular seksual menyebabkan masyarakat kurang peduli terhadap penularan penyakit IMS. Penyakit yang diderita belum tentu disebabkan oleh kesalahan diri sendiri namun juga bisa disebabkan oleh pasangan. Kegiatan seksual dengan pasangan resmi adalah kegiatan yang wajar untuk melangsungkan garis keturunan, akhirnya masyarakat cenderung kurang begitu peduli terhadap kesehatan alat reproduksi mereka karena mereka saling percaya satu sama lain (Kesehatan, 2011).

Pemberian layanan yang dilakukan oleh dokter terhadap penderita penyakit menular seksual yang telah dilakukan banyak mengalami kendala yang sering dihadapi. Dokter sering mendapatkan data masukan gejala yang diberikan pasien tidak lengkap karena pasien tidak tahu dengan pasti gejala penyakit yang diderita. Dokter kesulitan untuk mengambil keputusan yang tepat dalam mendiagnosa penyakit yang diderita pasien karena kurangnya keterangan gejala yang diberikan oleh penderita penyakit menular seksual. Dokter kesulitan melihat statistik jumlah penderita penyakit menular seksual yang tersebar di area operasional puskesmas dinoyo karena belum adanya sistem untuk pencatatan data penderita penyakit menular seksual, Karena dengan melihat statistik data penderita yang dibuat maka pihak puskesmas dan dokter dapat mengambil kesimpulan yang diambil apakah program yang dijalankan oleh puskesmas dan dokter dalam mengurangi jumlah penderita berhasil atau tidak. Dari permasalahan di atas, maka perlu dibuat suatu sistem cerdas yang mampu untuk membantu dokter dalam mengambil keputusan dari beberapa kemungkinan hasil deteksi infeksi penyakit menular seksual dan perkembangan jumlah penderita.

Penelitian sebelumnya yang terkait dengan diagnosa menggunakan sistem pakar dilakukan oleh Sianturi tentang penggunaan faktor kepastian (*certainty factor*). Sistem pakar mengkombinasikan kaidah-kaidah penarikan kesimpulan (*inference rule*) dengan basis pengetahuan tertentu yang diberikan oleh satu pakar atau lebih dalam bidang tertentu. Penghitungan prosentase keyakinan diawali dengan pemecahan sebuah kaidah yang memiliki premis majemuk, menjadi kaidah-kaidah yang memiliki premis tunggal. Aturan baru dihitung nilai *certainty factor*-nya, sehingga diperoleh nilai *certainty factor* untuk masing-masing aturan, kemudian nilai *certainty factor* tersebut dikombinasikan satu sama lain. Hasil pengujian dengan perhitungan *CF* sistem yang telah dikategorikan tersebut memiliki nilai kepastian 91,5% dan sistem hanya menghitung nilai kepastian dari salah satu kategori (*rule based*) (Sianturi, 2014).

Penelitian lain dilakukan oleh (Listiana, 2011), penerapan menggunakan metode *rough set* untuk deteksi dini penyakit sapi. Kasus tersebut disimpulkan

bahwa *rough set* tidak membutuhkan data awal atau data tambahan seperti probabilitas pada teori probabilitas atau tingkat keanggotaan pada *fuzzy set*. Metode *Rough set* terdapat kumpulan objek yang disebut sebagai *information system (IS)*. *IS (Information system)* merupakan objek-objek diklasifikasikan ke dalam area-area tertentu yang disebut dengan *lower approximation*, *upper approximation*, *boundary region*, dan *outside region*. Pengelompokan area tersebut dapat dilakukan perhitungan dependensi antar atribut, reduksi atribut dan *rule generation* sehingga diperoleh *rule* dari *data set* yang digunakan. Hasil dari penelitian penerapan metode *rough set* dan *qualitative measure* dapat membantu klasifikasi menjadi lebih jelas dengan adanya bobot pada *rule* yang dihasilkan. Penerapan *positive region* untuk proses *reduct* dinilai kurang efisien untuk penyakit dengan atribut banyak (Listiana, 2011).

Penelitian yang dilakukan oleh Ren menunjukkan bahwa metode yang didasarkan pada *rough set* dan *certainty factor* digunakan untuk memecahkan dua masalah utama dalam pencarian informasi. Pertama adalah aturan ekstraksi dari sistem informasi yang lengkap menggunakan teori *rough set*. Kedua menghitung hasil kepercayaan yang diperoleh dari bukti penalaran dengan *certainty factor*. Hasil yang didapatkan adalah menyelesaikan *data set* yang tidak lengkap dikonversikan dengan *mean completer algorithm*. Mengekstrak aturan menggunakan metode *rough set* kemudian menghitung keyakinan dari setiap aturan dengan menghasilkan keputusan yang diambil dari *rule* dan bobot alternatif (Ren, 2010).

Kasus dan kendala dokter yang ada di puskesmas dinoyo kota Malang dan beberapa penelitian, maka di dalam skripsi ini penulis mengambil judul "Implementasi Metode *Rough Set* dan *Certainty Factor* untuk Deteksi Dini Penyakit Menular Seksual (Studi Kasus: Puskesmas Dinoyo Kota Malang)" untuk memudahkan dokter dalam mendiagnosa penyakit menular seksual secara maksimal, efisien dan efektif.

2. PERMASALAHAN

Dari penjelasan pendahuluan, penelitian ini merumuskan bagaimana mengimplementasikan metode *roughset* dan *certainty factor* untuk deteksi dini penyakit menular seksual dan mengetahui tingkat akurasi hasil pakar dengan sistem dari hasil perhitungan *roughset* dan *certainty factor*.

3. TINJAUAN PUSTAKA

3.1 Studi terkait

Penelitian yang dilakukan oleh Ren menunjukkan bahwa metode yang didasarkan pada *rough set* dan *certainty factor* digunakan untuk memecahkan dua masalah utama dalam pencarian informasi. Pertama adalah aturan ekstraksi dari sistem informasi yang lengkap menggunakan teori *rough set*. Kedua menghitung hasil kepercayaan yang diperoleh dari bukti penalaran dengan *certainty factor*. Hasil yang

didapatkan adalah menyelesaikan *data set* yang tidak lengkap dikonversikan dengan *mean completer algorithm*. Mengekstrak aturan menggunakan metode *rough set* kemudian menghitung keyakinan dari setiap aturan dengan menghasilkan keputusan yang diambil dari *rule* dan bobot alternatif (Ren, 2010)].

3.2 Pengumpulan Data

Lokasi pengumpulan data penelitian ini berada di puskesmas dinoyo kota malang menggunakan data hasil pemeriksaan pasien sebanyak 109 data dan berdasarkan batasan masalah yang sudah ditentukan.

3.3 Dasar Teori

Penyakit menular seksual

PMS atau *Sexually Transmitted Disease* adalah suatu gangguan penyakit yang ditularkan dari satu orang ke orang lain melalui kontak hubungan seksual. Penyakit menular seksual lebih berisiko bila melakukan hubungan seksual dengan berganti-ganti pasangan. PMS yang sering terjadi adalah *Gonorrhoe*, *Sifilis*, *Harpes*.

Penyakit menular seksual menyebabkan infeksi saluran reproduksi penderita harus cepat ditangani sebelum menimbulkan komplikasi yang dapat membahayakan nyawa penderita. Bila tidak diobati secara tepat, infeksi dapat menjalar dan menyebabkan penderita sakit berkepanjangan, kemandulan dan kematian (Kesehatan, 2011).

Gejala PMS yang mungkin timbul antara

lain:

- Keluarnya sekret atau nanah dari penis, vagina atau anus.
- Nyeri atau rasa panas waktu kencing.
- Benjolan, bintil atau luka pada penis, vagina, anus atau mulut.
- Pembengkakan dipangkal paha.
- Pendarahan setelah berhubungan kelamin.
- Nyeri pada perut bawah (wanita).
- Nyeri pada buah pelir

Jenis jenis penyakit menular seksual antara lain

- gonore, sifilis, herpes genitalis, klamidia, trikomoniasis vaginalis, kutil kelamin, chankroid, limfoganuloma venereum, granuloma inguinale, servicitis, kandidiasis, vaginosis bakterial, moluskum kontagiosum, proctitis, konjuktivitis neonatorum, radang panggul.

Metode Roughset

Algoritma *Rough Set*, diusulkan pada tahun 1982 oleh Zdzislaw Pawlak. Metodologinya berkaitan dengan klasifikasi dan analisis data yang tidak tepat, tidak pasti atau tidak lengkap pada sistem informasi. Salah satu pendekatan non-statistik pertama dalam analisis data. Konsep dasar metode *rough set* adalah pendekatan *lower* dan *upper* dari satu set. Teori ini telah ditemukan di banyak bidang, seperti sistem pendukung keputusan, sistem pakar, rekayasa dan lain-lain. Setiap set pada *rough set* memiliki batas, yaitu benda yang tidak dapat

diklasifikasikan dengan pasti pada sebuah *data set*. Analisis data pada tabel informasi *rough set* disebut tabel keputusan, kolom yang diberi label disebut atribut, atribut dari tabel keputusan dibagi menjadi dua kelompok yang saling berhubungan yaitu atribut kondisi dan atribut keputusan (Rissino, 2009).

Langkah-langkah pembentukan rule menggunakan metode roughset secara umum:

Langkah 0 Menentukan data dari sistem informasi yang akan dijadikan rule sebagai basis pengetahuan.

Langkah 1 *Indiscernibility relation* merupakan hubungan antara dua objek atau lebih, di mana semua nilai-nilai yang identik saling berkaitan dengan subset/bagian sebuah atribut.

Langkah 2 *Set approximation* merupakan hubungan antar objek data pada tabel informasi yang memiliki batasan-batasan kriteria pada suatu objek data tertentu untuk mempermudah pengelompokan anggota data.

Langkah 3 Menentukan data konsisten dan tidak konsisten dalam *Dataset information table* mungkin memiliki data yang berisi data konflik dalam suatu kelas dan saling bertentangan hasil atribut keputusan. (Cohagan, 2010).

Langkah 4 *Reduct* dan *core* adalah dua konsep yang paling penting teori *rough set*. *Reduct* adalah penyeleksian atribut minimal (*interesting atribut*) dari sekumpulan atribut kondisi. *Core* merupakan *intersection/irisan* dari *reduct* yang telah terbentuk dari tabel informasi.

Langkah 5 *Rule induction* adalah salah satu teknik yang paling penting dalam *machine learning*. Dalam tabel informasi ada keteraturan pola data yang sering dinyatakan dalam bentuk *rule*.

Langkah 6 Dalam algoritma *rough set* banyak digunakan metode untuk membangun *rule induction* dari atribut keputusan pada tabel informasi. Metode yang bisa digunakan harus sesuai dengan tabel informasi yang ada dengan melihat jumlah atribut pada tabel informasi metode *rule induction* antara lain *Boolean reasoning*, GDT-RS, LEM1, LEM 2, CN2, AQ, PRIM, MODLEM dan lainnya (Ning, 2000).

Metode Certainty factor

Faktor kepastian (certainty factor) diperkenalkan oleh Shortliffe buchanan dalam pembuatan MYCIN. *Certainty Factor (CF)* merupakan nilai parameter MYCIN untuk menunjukkan besarnya kepercayaan pada sistem pakar. *Certainty Factor* menggunakan suatu nilai untuk mengasumsikan derajat keyakinan seorang pakar terhadap suatu data. Nilai maksimum dari *certainty factor* adalah +1.0 (pasti benar) dan minimum -1.0 (pasti palsu). Dalam *certainty factor* basis pengetahuan terdiri dari seperangkat aturan yang memiliki sintaks sebagai berikut (Munandar, et al., 2012):

IF $\langle evidence \rangle$
THEN $\langle hypothesis \rangle \{cf\}$

Dimana cf merupakan keyakinan *hypothesis H* dengan nilai dari *evidence E*, Kemudian *certainty factor* memperkenalkan konsep keyakinan dan ketidakyakinan yang kemudian diformulasikan dalam rumusan dasar sebagai berikut (Munandar, et al., 2012):

1. *Certainty Factor* untuk kaidah dengan premis tunggal

(single premis rules) :

$$CF(H, E) = CF(E) * CF(rule)$$

$CF(rule)$ ditentukan oleh para ahli, sedangkan $CF(E)$ ditentukan oleh pengguna atau user.

2. *Certainty Factor* untuk kaidah dengan premis majemuk

(multiple premis rules)

Untuk disjungsi:

$$IF e1 AND e2... AND en THEN h(CF_{rule}):$$

$$CF(A AND B) = Minimum(CF(a), CF(b)) * CF_{rule}$$

Untuk konjungsi:

$$IF e1 OR e2... OR en THEN h(CF_{rule})$$

$$CF(A OR B) = Maximum(CF(a), CF(b)) * CF_{rule}$$

3. *Certainty Factor* untuk kaidah dengan kesimpulan yang serupa

(similarly concluded rules) :

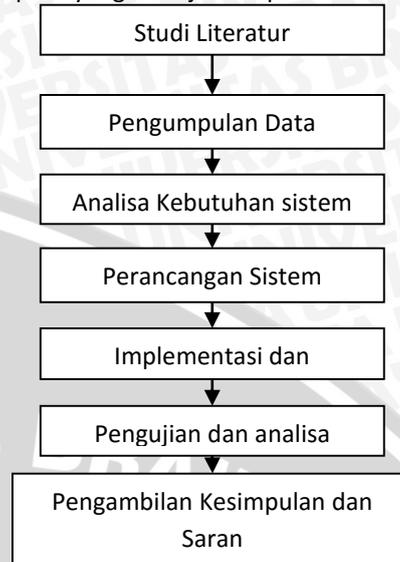
$$CF_{Combine}(CF1, CF2) = \begin{cases} CF_1 + CF_2(1 - CF_1) \rightarrow \text{if } CF_1 > 0 \text{ and } CF_2 > 0 \\ \frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)} \rightarrow \text{if } CF_1 < 0 \text{ or } CF_2 < 0 \\ CF_1 + CF_2(1 + CF_1) \rightarrow \text{if } CF_1 < 0 \text{ and } CF_2 < 0 \end{cases}$$

$CF[H, E] = M[H, E] - MD[H, E]$, Dimana $CF[H, E]$ merupakan faktor kepastian dan $M[H, E]$ merupakan ukuran kepercayaan hipotesis nilai H , jika E (antara 0 dan 1), $MD[H, E]$ merupakan ukuran ketidakpercayaan hipotesis H , jika diberikan nilai e (antara 0 dan 1).

4. METODOLOGI

Tahapan penelitian menjelaskan langkah-langkah yang akan dilakukan dalam penyusunan skripsi, yaitu perancangan, implementasi, pengujian dan analisis dari aplikasi perangkat lunak yang akan dibuat dan penerapan metode *rougth set* dan *certainty factor*

didalam sistem pakar. Secara garis besar, langkah-langkah penelitian yang dilakukan untuk membuat sistem seperti yang ditunjukkan pada Gambar.



Gambar 1. Diagram Alur

5. PERANCANGAN

siklus perancangan aplikasi pendeteksi dini penyakit menuar seksua menggunakan algoritma metode *rough set* dan *certainty factor* yang dikombinasikan untuk menyelesaikan permasalahan. Isi dari perancangan yaitu proses perhitungan penelitian, siklus pembentukan *rule* menggunakan *ROSE 2*, perhitungan nilai *certainty factor* masing masing *rule* dan hasil akhir berupa hasil kombinasi perhitungan. Gambar diagram alir perancangan sistem.

- $\|C\|$ = Kasus yang memiliki kondisi atribut sesuai dengan beberapa aturan misalkan aturan/rule $A \rightarrow B$, kondisi atribut adalah A .
- a_j = Nilai atribut ke-j yang sesuai dengan beberapa aturan.
- V_{aj} = Banyaknya nilai atribut yang mungkin ada pada nilai atribut ke-j.
- $t1$ = Banyaknya jumlah data kasus yang sesuai dengan rule $A \rightarrow B$.
- $t2$ = Banyaknya jumlah data kasus yang sesuai dengan kondisi atribut A ditambahkan dengan kondisi atribut A yang masih bernilai *missing value* pada data lain dalam satu tabel
- n = banyaknya jumlah atribut yang terpakai dalam suatu *rule* tertentu.

Nilai certainty factor setiap rule

Dengan persamaan diatas maka terbentuk nilai certainty factor setiap rule dengan melihat data dari tabel informasi.

rule	Rule	Nilai cf
1	(G3 = y) & (G10 = y) => (Penyakit = Chancroid)	1
2	(G12 = y) & (G18 = y) => (Penyakit = Granulomainguinale)	1
3	(G6 = t) & (G9 = y) & (G11 = t) & (G21 = t) => (Penyakit = HerpesGenitalis)	1
4	(Resiko = 7) & (G18 = y) & (G24 = y) => (Penyakit = HerpesGenitalis)	1
5	(G2 = t) & (G3 = t) & (G4 = t) & (G11 = t) & (G12 = t) & (G23 = t) & (G28 = t) & (G29 = t) => (Penyakit = Kandidiasis)	0,85714 2857
6	(JenisKelamin = w) & (G2 = y) & (G22 = y) => (Penyakit = Kandidiasis)	1
7	(G1 = t) & (G15 = y) & (G20 = t) & (G23 = t) => (Penyakit = Klamidia)	1
8	(JenisKelamin = w) & (G4 = y) => (Penyakit = Klamidia)	1
9	(G14 = y) & (G16 = y) => (Penyakit = Klamidia)	1
10	(G28 = y) => (Penyakit = KonjungtivitisNeonatorum)	1
11	(G2 = t) & (G11 = y) & (G13 = t) => (Penyakit = KutilKelamin)	1
12	(G3 = y) & (G4 = y) => (Penyakit = Limfogranulamavenerium)	1

13	(G1 = y) & (G12 = y) => (Penyakit = MoluskuMkontagiosum)	1
14	(Resiko = 3) & (G4 = y) => (Penyakit = PembengkakanSkrotum)	1
15	(G13 = y) & (G16 = y) & (G19 = t) => (Penyakit = Proctitis)	1
16	(G29 = y) => (Penyakit = RadangPanggul)	1
17	(JenisKelamin = w) & (G2 = y) & (G3 = t) & (G9 = t) & (G15 = t) & (G16 = t) & (G23 = t) => (Penyakit = Servicitis)	1
18	(G18 = y) & (G26 = y) => (Penyakit = Sifilis)	1
19	(G8 = y) & (G9 = y) & (G20 = t) & (G22 = t) => (Penyakit = Sifilis)	1
20	(G2 = y) & (G9 = t) & (G26 = t) & (G27 = y) => (Penyakit = Trikomoniasisvaginalis)	1
21	(G11 = t) & (G15 = y) & (G17 = t) & (G23 = y) => (Penyakit = Uretritis)	1
22	(JenisKelamin = w) & (G4 = t) & (G7 = t) & (G11 = t) & (G12 = t) & (G18 = t) & (G19 = t) & (G26 = t) & (G28 = t) => (Penyakit = Uretritis)	1
23	(G2 = y) & (G8 = t) & (G9 = t) & (G13 = t) & (G15 = t) => (Penyakit = Uretritis)	1
24	(G7 = t) & (G10 = t) & (G15 = t) & (G25 = y) => (Penyakit = Vaginosisbakterial)	1
25	(Resiko = 3) & (G11 = y) => (Penyakit = HerpesGenitalis)	0,5
26	(G1 = y) & (G3 = t) & (G14 = t) & (G22 = y) & (G25 = t) => (Penyakit = HerpesGenitalis)	0,2
27	(G2 = t) & (G9 = y) & (G13 = y) & (G15 = t) => (Penyakit = Kandidiasis)	0,33333 3333
28	(Resiko = 3) & (G11 = y) => (Penyakit = Uretritis)	0,5
29	(G1 = y) & (G3 = t) & (G14 = t) & (G22 = y) & (G25 = t) => (Penyakit = KutilKelamin)	0,6
30	(G1 = y) & (G3 = t) & (G14 = t) & (G22 = y) & (G25 = t) => (Penyakit = Uretritis)	0,2
31	(G2 = t) & (G9 = y) & (G13 = y) & (G15 = t) => (Penyakit = Klamidia)	0,33333 3333
32	(G2 = t) & (G9 = y) & (G13 = y) & (G15 = t) => (Penyakit = Servicitis)	0,33333 3333

Perhitungan Nilai akhir



Inputan data pasien penyakit menular seksual dengan nilai cf masing-masing gejala

Data pasein	Jawaban	Nilai Cf
Nama	Pasien 1	
Jenis Kelamin	Laki-laki	1
Resiko	R3	1
G1	Tidak	0.95
G2	Tidak	0.95
G3	Tidak	0.95
G4	Tidak	0.95
G5	Tidak	0.95
G6	Ya	0.95
G7	Ya	0.95
G8	--	
G9	--	
G10	Ya	0.95
G11	Tidak	0.95
G12	Tidak	0.95
G13	-	--
G14	-	--
G15	-	-
G16	-	-
G17	-	-
G18	-	-
G19	-	-
G20	--	--
G21	-	-
G22	Tidak	0.95
G23	Tidak	0.95
G24	--	-
G25	--	-
G26	Ya	0.95
G27	Tidak	0.95
G28	Tidak	0.95
G29	Tidak	0.95
Penyakit	Kandidiasis	0.95

Dihitung nilai masing-masing rule dengan data yang dimasukan oleh user jika kondisi atribut masukan user sesuai dengan kondisi atribut setiap rule maka dihitung nilai cf nya:

- $(G2 = t) \& (G3 = t) \& (G4 = t) \& (G11 = t) \& (G12 = t) \& (G23 = t) \& (G28 = t) \& (G29 = t) \Rightarrow$ (Penyakit = Kandidiasis)

Gejala pada pasien 1 mempunyai kesamaan dan memiliki semua nilai attribut kondisi sesuai rule 5 sehingga dapat dihitung nilai cfnya:

Menggunakan cara disjungsi karena persamaan aturan yang terbentuk yaitu dengan persamaan *if and then*

$$IF e1 AND e2 \dots AND en THEN h(CF_{rule}).$$

$$CF(A AND B) = Minimum(CF(a), CF(b)) * CF_{rule}$$

Diambil nilai cf minimum dari hasil inputan yang dimasukan oleh penderita.

$$(G2 = t) \& (G3 = t) \& (G4 = t) \& (G11 = t) \& (G12 = t)$$

$$\& (G23 = t) \& (G28 = t) \& (G29 = t) \rightarrow \text{(Penyakit = Kandidiasis)}$$

$$cf(G2 = t \& G3 = t \& G4 = t \& G11 = t \& G12 = t$$

$$\& G23 = t \& G28 = t \& G29 = t \rightarrow \text{Penyakit = Kandidiasis})$$

$$= Minimum \left(\begin{matrix} cf(G2), cf(G3), cf(G4), cf(G11), cf(G12), \\ cf(G23), cf(G28), cf(G29) \end{matrix} \right) * cf_{ruleKandidiasis}$$

$$= Minimum(0.95, 0.95, 0.95, 0.95, 0.95, 0.95, 0.95, 0.95) * 0.8571428$$

$$= 0.95 * 0.8571428$$

$$= 0.8075 * 100\%$$

$$= 80,75$$

Maka hasil dari perhitungan certainty factor pada rule 5 mendapatkan nilai keyakinan sebesar 80,75 %

7. Pengujian Sistem

Hasil akurasi dari pengujian sistem aplikasi deteksi dini penyakit menular seksual sebagai berikut:

Pengujian Akurasi

No	Hasil pakar	Hasil sistem	Analisa/pencocokan hasil
1	Pembengkakan Skrotum	Vaginosis bakterial	Salah
2	Klamidia	Klamidia	Benar
3	Klamidia	Klamidia	Benar
4	Sifilis	Sifilis	Benar
5	Vaginosis bakterial	Vaginosis bakterial	Benar
6	Moluskum Kontagiosum	Moluskum Kontagiosum	Benar
7	Moluskum Kontagiosum	Moluskum Kontagiosum	Benar
8	Vaginosis bakterial	Vaginosis bakterial	Benar
9	Limfogranulama venereum	Limfogranulama venereum	Benar
10	Herpes Genitalis	Herpes Genitalis	Benar
11	Proctitis	Proctitis	Benar

1	Kutil	Kutil	Benar
2	Kelamin	kelamin	
1	Proctitis	Proctitis	Benar
3			
1	Uretritis/gonore	Uretritis/gonore	Benar
4			
1	Trikomoniasis vaginalis	Trikomoniasis vaginalis	Benar
5			
1	Konjungtivitis Neonatorum	Konjungtivitis Neonatorum	Benar
6			
1	Chancroid	Chancroid	Benar
7			
1	Granuloma inguinale	Granuloma inguinale	Benar
8			
1	Vaginosis bakterial	Vaginosis bakterial	Benar
9			
2	Klamidia	Klamidia	Benar
0			
2	Servicitis	Servicitis	Benar
1			
2	Servicitis	Servicitis	Benar
2			
2	Kandidiasis	Kandidiasis	Benar
3			
2	Proctitis	Proctitis	Benar
4			
2	Herpes Genitalis	Herpes Genitalis	Benar
5			
2	Trikomoniasis vaginalis	Trikomoniasis vaginalis	Benar
6			
2	Radang Panggul	Radang Panggul	Benar
7			
2	Klamidia	Klamidia	Benar
8			
2	Uretritis/gonore	Uretritis/gonore	Benar
9			
3	Sifilis	Sifilis	Benar
0			
3	Granuloma inguinale	Granuloma inguinale	Benar
1			
3	Konjungtivitis Neonatorum	Konjungtivitis Neonatorum	Benar
2			

Analisa Hasil

Hasil akurasi pengujian diatas mendapatkan hasil akurasi sistem sebesar 96,87% dihitung dengan membagi hasil diagnosa aplikasi yang benar

dengan hasil seluruh data uji sebanyak 32 dari pakar.

- Pada data no 1 pada Tabel 6.2.1 hasil dari pakar adalah pembengkakan skrotum namun yang dihasilkan sistem yaitu vagionosis bakterial dengan hasil berbeda tersebut dapat disimpulkan bahwa nilai *certainty factor* dari hasil keputusan yang dihasilkan dalam perhitungan antara pembengkakan skrotum dan *vagionosis bakterial* nilai *certainty factor* vagionosis bakterial lebih besar.

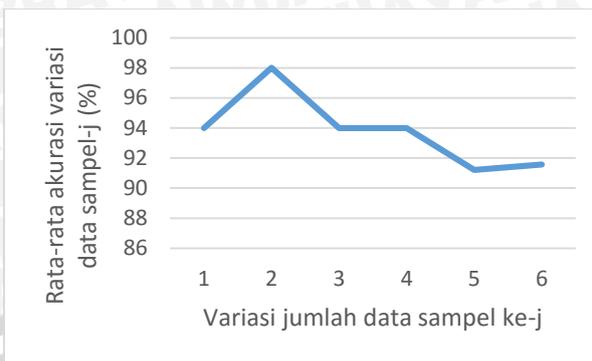
Kesimpulan dari pengujian ini adalah bahwa data pasien penyakit menular seksual yang didiagnosa oleh dokter sebelumnya dan kemudian hasil diagnosa oleh sistem aplikasi mempunyai nilai keakuratan sebesar 96,87% jadi pembuatan sistem aplikasi ini dinyatakan berhasil.

Pengujian akurasi sistem dengan *sampling bootstrap*

	Variasi jumlah data sampel ke-j						Rata-rata akurasi percobaan sampel ke-i
	1	2	3	4	5	6	
Pengambilan data sampel ke-i	5 data	10 data	15 data	20 data	25 data	32 data	
	1	100	100	93,3	95	100	87,5
	2	80	90	100	95	88	87,5
	3	100	90	93,3	95	88	78,12
	4	100	100	93,3	90	100	90,62
5	100	100	100	90	80	75	93,33



6	60	100	10	0	100	84	100	90,6
			86					6666
			,6					667
7	100	100	6	95	96	100	100	96,2
			,3					7666
			3	90	92	87	96,	6666
8	100	100	3	90	92	87	87	95,3
			,6					6666
			6	100	88	75	93,	94,7
9	100	100	6	100	88	75	75	35
			,3					94,4
			3	90	96	5	87,	7166
10	100	100	3	90	96	5	5	667
Rata-rata akurasi variasi jumlah data sampel-j	94	98	97	94	91,2	91,561		



Gambar 2 Rata-rata akurasi variasi jumlah data sampel-j

Analisa Hasil

Analisa hasil pengujian akurasi sistem dengan *sampling bootstrap* memiliki nilai rata-rata akurasi diatas >90% dan mendekati 96,87 % dari hasil pengujian sebelumnya penjelasan sebagai berikut:

Hasil akurasi percobaan sampel ke-j berdasarkan informasi dari Gambar 1 dan 2 diatas dapat disimpulkan bahwa akurasi hasil pengujian yang dilakukan secara *sampling bootstrap* dan bukan *sampling bootstrap* bahwa nilai akurasi sistem yang didapatkan tidak jauh berbeda dengan hasil pengujian yang tidak dilakukan dengan *sampling bootstrap* sehingga dengan contoh kasus diatas.

Bahwa metode *sampling bootstrap* dapat diimplementasikan ke dalam penelitian yang memiliki jumlah uji yang lebih besar untuk mencari nilai akurasi hasil sistem

Pengujian akurasi sistem dengan variasi nilai missing value dan variasi jumlah data sampel dengan *sampling bootstrap*

Hasil pengujian variasi *missing value* pada atribut yang diacak secara random dan diuji akurasinya dengan data yang telah diambil kemudian dicocokkan dengan pakar seperti pada proses pengujian sebelumnya bisa dilihat dibawah ini

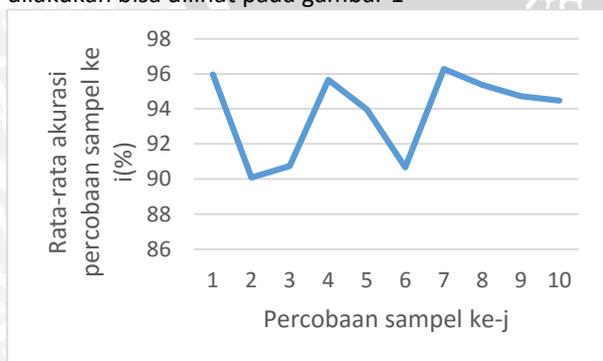
Pengujian Variasi *missing value* dan data

Variasi	Variasi jumlah data sampel ke-j						Rata-rata akurasi variasi nilai missing value ke i
	1	2	3	4	5	6	
missing value atribut data uji	5 data	10 data	15 data	20 data	25 data	32 data	

Hasil akurasi setiap percobaan dihasilkan dari setiap jumlah data yang diambil dihitung dengan aplikasi dan dicocokkan dengan hasil pakar :

$$Akurasi = \frac{HasilAplikasi}{HasilSistem} \times 100\%$$

Hasil rata-rata akurasi yang ditampilkan dalam bentuk grafik pengujian *sampling bootstrap* yang telah dilakukan bisa dilihat pada gambar 1

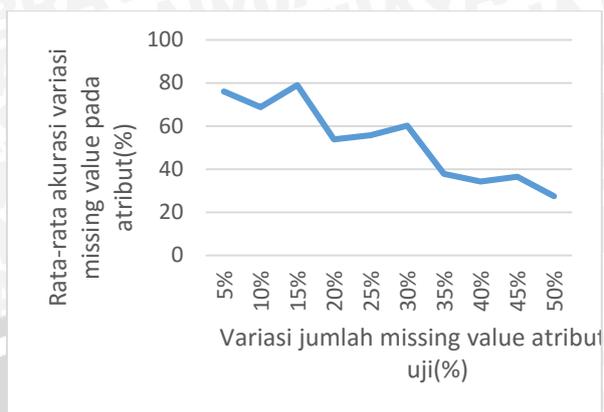


Gambar 1 Rata-rata akurasi percobaan sampel ke-i

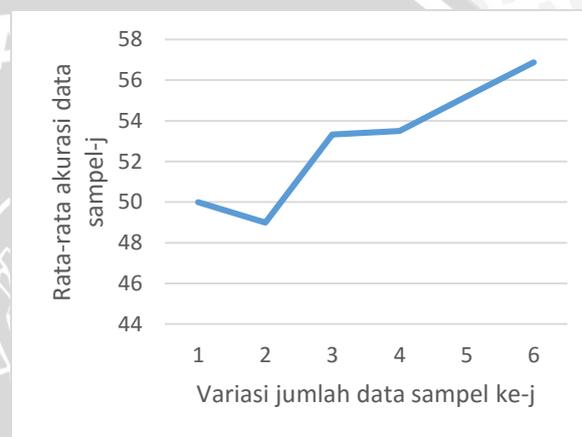


5%	60	70	93,3	75	80	78,1	76,075
10%	80	60	66,6	60	72	84,3	68,833
15%	100	80	66,6	60	80	87,5	79,026
20%	40	40	46,6	55	76	65,6	53,88
25%	40	70	33,3	75	60	56,2	55,763
30%	80	50	80	65	36	50	60,166
35%	20	40	46,6	35	48	37,5	37,86
40%	40	10	40	40	44	31,2	34,208
45%	20	50	46,6	45	20	37,5	36,526
50%	20	20	13,3	35	36	40,6	27,491
Rata-rata akurasi data sampel ke-j	50	49	53,3	53,5	55,2	56,8	

Hasil nilai rata-rata akurasi berdasarkan peningkatan persentase *missing value* pada atribut yang diacak secara random dan pengambilan data uji yang dilakukan secara sampling. Hasil ditunjukkan pada gambar



Gambar 0 Rata-rata akurasi variasi nilai *missing value*



Gambar 4 Hasil akurasi pengujian *missing value sampling data*

Analisa Hasil

Berdasarkan hasil pengujian variasi *missing value* dan variasi sampling data uji berdasarkan Gambar grafik 3 dan 4 dapat disimpulkan bahwa nilai akurasi sistem cenderung menurun saat nilai *missing value* pada atribut data uji diberikan persentase semakin besar karena data uji yang dimasukan tidak terakomodasi terhadap *rule* yang ada didalam sistem karena kurangnya informasi gejala penyakit yang sengaja dihilangkan, sehingga sistem kesulitan mendeteksi penyakit dengan informasi atribut data yang sedikit. Dan untuk nilai akurasi variasi jumlah data sampel yaitu 5, 10, 15, 20, 25, 32 data nilai akurasinya cenderung sama satu sama lain yaitu antara 45%-60%

Pengujian *missing value* pada atribut data uji terhadap akurasi hasil sistem

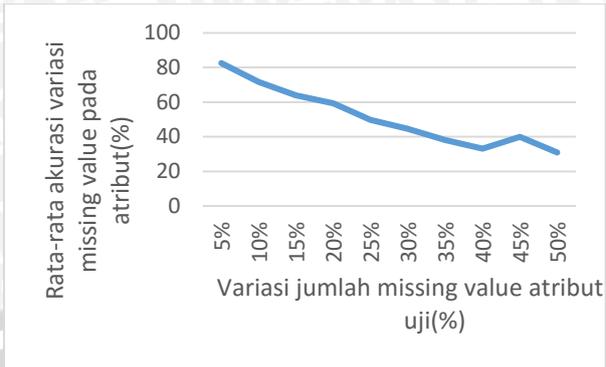
Perbedaan hasil pengujian ini dengan hasil pengujian sebelumnya yaitu terletak pada jumlah data yang dipakai. Jumlah data yang diambil untuk pengujian selalu berjumlah 32 dilakukan sebanyak 6 kali dan

dengan variasi persentase pemberian *missing value* pada atribut data uji.

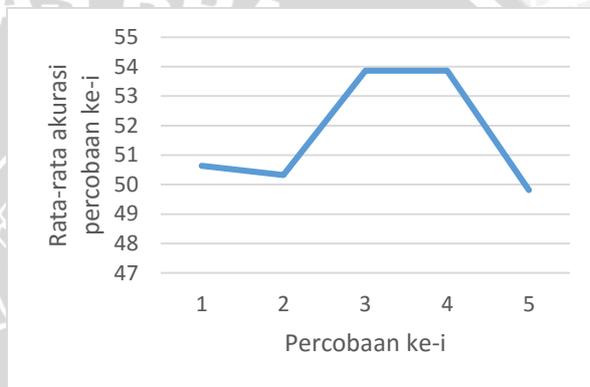
Variasi *missing value*

Variasi nilai <i>missing value</i> pada atribut data uji	Percobaan ke-i					Rata-rata akurasi variasi nilai <i>missing value</i>
	1	2	3	4	5	
5%	32 da ta	32 da ta	32 da ta	32 da ta	32 da ta	82,576
10%	70 ,9 6	74 ,1 9	80 ,6 4	74 ,1 9	58 ,0 6	71,608
15%	61 ,2 9	61 ,2 9	64 ,5 1	64 ,5 1	67 ,7 4	63,868
20%	54 ,8 3	58 ,0 6	61 ,2 9	64 ,5 1	58 ,0 6	59,35
25%	54 ,8 3	48 ,3 8	51 ,6 1	51 ,6 1	41 ,9 3	49,672
30%	45 ,1 6	51 ,6 1	41 ,9 3	45 ,1 6	38 ,7	44,512
35%	45 ,1 6	25 ,8	35 ,4 8	41 ,9 3	41 ,9 3	38,06
40%	25 ,8	29 ,0 3	38 ,7	38 ,7	29 ,0 3	33,0575
45%	38 ,7	45 ,1 6	38 ,7	38 ,7	38 ,7	39,992
50%	29 ,0 3	22 ,5 8	41 ,9 3	32 ,2 5	29 ,0 3	30,964
Rata-rata akurasi percobaan	50 ,6 4	50 ,3 19	53 ,8 66	53 ,8 65	15 ,55 6	49,8

Hasil rata-rata akurasi pengujian variasi *missing value* pada atribut data uji ditunjukkan pada grafik



Gambar 5 Rata-rata akurasi variasi *missing value* pada atribut



Gambar 6 Rata-rata akurasi percobaan ke-i Gambar akurasi Variasi *missing value*

Analisa hasil

Berdasarkan hasil pengujian yang telah dilakukan dengan melihat grafik Gambar 6.4.1 dan Gambar 6.4.2 Hasil akurasi pengujian variasi *missing value* mempunyai nilai akurasi sistem cenderung turun saat nilai *missing value* semakin besar karena data yang diuji yang memiliki nilai *missing value* pada atribut tidak terakomodasi oleh *rule* yang ada pada sistem aplikasi. Dengan membandingkan rata-rata akurasi pada Gambar 6.4.2 dengan pengujian sebelumnya yaitu pada Gambar 6.3.2 bahwa hasil rata-rata akurasi sistem bernilai antara 45-60% meskipun cara pengambilan data yang diuji berbeda sehingga dapat disimpulkan lagi bahwa pengujian sampling dapat dilakukan saat data yang akan diuji lebih besar

Terdapat nilai *Miss* pada atribut data uji yang sengaja dihilangkan/diberikan nilai *MISS* untuk menghitung berapa akurasi sistem terhadap hasil



pakar. Pemberian nilai *Miss* dihitung dengan rumus:

$$Miss = \frac{\text{persentaseMissValue}}{100} \times (\text{jumlahDataUji} * \text{JumlahAtributKondisiDataUji})$$

8. PENUTUP

1. Kesimpulan

Berdasarkan hasil perancangan dan pengujian yang dilakukan pada aplikasi deteksi dini penyakit menular seksual menggunakan metode *Rough set-Certainty factor*, maka didapatkan kesimpulan sebagai berikut :

1. Proses identifikasi penyakit menular seksual dilakukan dengan memasukan fakta gejala yang muncul pada tubuh pasien. Setiap gejala yang dipilih memiliki bobot *certainty factor* yang digunakan dalam perhitungan metode *certainty factor*. Nilai *certainty factor* setiap penyakit menular seksual dihitung dengan mengkombinasikan dengan *rule* yang dihasilkan menggunakan metode *Rough set rule induction LEM 2* dan proses pembentukan *rulanya* menggunakan tool ROSE 2. Hasil diagnosa yang tertinggi akan dijadikan kuputusan hasil akhir.
2. Pendekatan metode *rough set* dalam analisis data pasien penyakit menular seksual puskesmas Dinoyo kota Malang dapat menghasilkan *decision rule* dari data. Dalam pembentukan *rule* menggunakan LEM2 pada aplikasi ROSE 2 sangat efektif untuk membentuk basis pengetahuan penyakit menular seksual yang akan digunakan dalam perhitungan mencari nilai *certainty* penyakit. Pendekatan metode *rough set* dalam penelitian ini digunakan untuk menentukan klasifikasi atau aturan penyakit menular seksual dari data informasi pasien dari puskesmas Dinoyo kota Malang, Kemudian nilai keyakinan hasil sistem untuk memutuskan deteksi dini penyakit menular seksual dihitung dengan nilai *certainty factor* tiap gejala yang dialami pasien dengan nilai *certainty factor* dari setiap *rule/aturan* yang terbentuk dengan dengan perhitungan probabilitas atribut pada data latih yang

digunakan. Basis pengetahuan tidak hanya didapatkan dari pakar/ahli maupun observasi melainkan bisa diketahui dari *rule* yang terbentuk menggunakan pendekatan metode *rough set* dengan sistem informasi yang ada dengan berbagai macam *rule induction* yang ada dalam *rough set*.

3. Dengan pemberian nilai *missing value* pada atribut data uji yang digunakan dapat membuat nilai hasil akurasi dari aplikasi menurun dalam penentuan deteksi dini penyakit menular seksual karena kurangnya informasi data gejala yang dapat digunakan dan tidak terakomodasi oleh *rule* gejala yang ada. Aplikasi deteksi dini penyakit menular seksual metode *rough set-certainty factor* telah mampu memenuhi seluruh kebutuhan fungsional. Hal ini berdasarkan hasil pengujian blackbox dengan akurasi 100% yang membuktikan bahwa seluruh fungsi telah sesuai dengan yang diharapkan. Penerapan metode *sampling bootstrap* untuk pengujian memberikan kesimpulan bahwa hasil akurasi tidak berbeda jauh dengan pengujian tanpa menggunakan metode *sampling bootstrap* dengan pemberian nilai *missing value* pada atribut yang sama. Pengujian variasi *missing value* terendah yaitu 5% dengan pengambilan *sampling data* pada data uji memiliki akurasi terbaik bernilai 93,33%, sedangkan untuk pengujian bukan *sampling* dengan menggunakan nilai *missing value* terendah sebesar 5% memiliki nilai akurasi 87,09%

4.1. Saran

Aplikasi deteksi dini penyakit menular seksual yang dibangun ini masih mempunyai beberapa kekurangan yang dapat dilengkapi sehingga aplikasi memiliki performa yang lebih baik. Saran yang dapat diberikan untuk pengembangan aplikasi antara lain :

1. Aplikasi dapat dikembangkan lebih lanjut dengan mengkombinasikan metode *Rough set-certainty factor* dengan metode *rule induction* selain LEM 2 dalam pembentukan *rule* masing-masing

keputusan, Karena dalam penerapan metode LEM 2 dalam penelitian ini terdapat data pasien yang tidak terdeteksi/terdiagnosa penyakit menular seksual karena masih kurangnya *rule* yang dijadikan basis pengetahuan disebabkan data latih yang dipakai untuk pembentukan *rule* masih sedikit.

2. Perlu adanya penelitian lebih lanjut untuk mencari *rule induction* yang lebih baik dan cepat dalam pembentukan *rule* sebagai basis pengetahuan untuk data informasi yang memiliki atribut dan jumlah data banyak. Sehingga *rule* yang terbentuk nanti dapat mengakomodasi gejala yang dimasukkan oleh pasien.
3. Perlu dilakukan percobaan pembentukan *rule* dengan berbagai macam jumlah variasi data untuk pembentukan *rule* sebagai basis pengetahuan

DAFTAR PUSTAKA

- Alfaris, S., 2014. Sistem pakar untuk mendiagnosa penyakit polip nasi (polip hidung) menggunakan metode certainty factor. *Sistem pakar untuk mendiagnosa penyakit polip nasi (polip hidung) menggunakan metode certainty factor*, VII(1), pp. 147-151.
- Belton, D., 1998. *Boolean Algebra*. [Online] Available at: <http://www.ee.surrey.ac.uk/Projects/Labview/boolalgebra/#booleantheorems> [Diakses 28 11 2015].
- Buaton, R., 2013. Sistem pakar. *Sistem Pakar*, 1(1), pp. 108-107.
- Busse, J. W. G., 2014. Chapter 1 Rule induction. Dalam: U. o. Kansas, penyunt. *Rule induction*. Kansas: University of Kansas, pp. 1-5.
- Cohagan, C. e. a., 2010. Mining Inconsistent Data with the Bagged MLEM2 Rule Induction Algorithm. *2010 IEEE International Conference on Granular Computing*, 1(Inkonsisten), pp. 115-120.
- Hong, T. P., 2002. Learning rule from incomplete training examples by rough set. *Learning rule from incomplete training examples by rough set*, (-), pp. 286-287.
- Sianturi, E., 2014. sistem pakar diagnosa gejala kecanduan game online dengan menggunakan metode certainty factor.. *sistem pakar diagnosa gejala kecanduan game online dengan menggunakan metode certainty factor.*, VII(3), pp. 1-5.
- spesialis.info, 2016. *www.spesialis.info*. [Online] Available at: [http://www.spesialis.info/?penyebab-oftalmia-neonatorum-\(konjungtivitis-neonatorum\),914](http://www.spesialis.info/?penyebab-oftalmia-neonatorum-(konjungtivitis-neonatorum),914) [Diakses 30 Januari 2016].
- Sungkono, J., 2013. RESAMPLING BOOTSTRAP PADA R. *RESAMPLING BOOTSTRAP PADA R*, Magistra No. 84(-), pp. 47-54.
- Yanti, 2016. *www.sridianti.com*. [Online] Available at: <http://www.sridianti.com/gejala-bakterial-vaginosis-dan-pengobatannya.html> [Diakses 17 Januari 2016].