

OPTIMASI BASIS PENGETAHUAN MENGGUNAKAN ALGORITMA *FP-GROWTH* UNTUK MEMBANGUN STRUKTUR *BAYESIAN NETWORK* (Studi Kasus : Penyakit Mata di Rumah Sakit Mata Pekanbaru)

Yelfi Vitriani, MMSI¹, Westi Oktaviana²

^{1,2} Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau
Jl. HR. Soebrantas No. 155 Simpang Baru, Panam, Pekanbaru, 28293
Email: yelfivitriani@yahoo.com, westi.oktaviana@uin-suska.ac.id

ABSTRAK

Salah satu kelemahan *Bayesian Network* adalah sulit dalam mengambil kesepakatan beberapa pakar, pakar akan sulit menentukan nilai probabilitas, dan pakar akan memerlukan waktu yang lama hanya untuk membangun struktur *Bayesian Network*. Untuk mengatasi kelemahan *Bayesian Network* maka diperlukanlah ilmu kecerdasan buatan lain yaitu data *mining* dengan teknik *Association Rule* menggunakan algoritma *FP-Growth*. Penelitian ini mengambil kasus pada penyakit mata dengan tujuan untuk membangun struktur *Bayesian Network* dan menghasilkan nilai probabilitas untuk mendapatkan mana gejala yang paling berpengaruh dalam penyakit mata tersebut. Pengujian dilakukan dengan *tools* data *mining* WEKA 3.7.10 didapat 24 *rules* yang memenuhi ketentuan dan hasil pengujian kualitatif sebesar 99% benar, sehingga mendapatkan nilai probabilitas untuk penyakit *presbiopia*, jenis kelamin yang paling berpengaruh adalah perempuan, dibuktikan dengan nilai probabilitas sebesar 60%. Untuk umur yang paling berpengaruh adalah dewasa tengah dari 31-59 tahun sebesar 65%, dan gejala yang paling berpengaruh adalah kabur dekat sebesar 98%. Sedangkan untuk penyakit *conjungtivitis*, jenis kelamin yang paling berpengaruh laki-laki sebesar 53%. Untuk umur yang paling berpengaruh adalah dewasa tengah dari 31-59 tahun sebesar 43%, dan gejala yang paling berpengaruh adalah mata lengket sebesar 100%. Berdasarkan hasil pengujian tersebut dapat disimpulkan bahwa teknik *Association Rule* berhasil dalam mengatasi kelemahan *Bayesian Network* berdasarkan fakta dan data.

Kata Kunci: *Association Rule, Bayesian Network, Data Mining, FP-Growth, Penyakit Mata*

ABSTRACT

One of the weaknesses of the Bayesian Network is that it is difficult to get agreement from some experts, for an expert it will be difficult to determine the probability value, and an expert will take a long time just to build the Bayesian Network structure. To overcome the weakness of the Bayesian Network then required another artificial intelligence science that is data mining with Association Rule technique using FP-Growth algorithm. This research takes the case on eye disease with the aim of building a Bayesian Network structure and generating probability values to get where the most influential symptoms in eye disease. The method test is carried out by using data mining tools WEKA 3.7.10, with the results obtained by 24 rules that meet the provisions and qualitative test results of 99% correct and get the probability value for presbyopia disease, with the greatest influence on women, evidenced by probability value of 60%. For the most influential age was the mature middle of 31-59 years at 65%, and the most influential symptom was a near blur of 98%. As for conjunctivitis disease with the largest influence on men by 53%. For the most influential age is the middle adult from 31-59 years by 43%, and the most influential symptoms are the sticky eye of 100%. Based on the results of these tests can be concluded that the Association Rule technique succeeded in overcoming Bayesian Network weaknesses based on facts and data.

Keywords: *Association Rule, Bayesian Network, Data Mining, Eye Disease, FP-Growth*

Corresponding Author

Yelfi Vitriani

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarfi Kasim Riau,

Email: yelfivitriani@yahoo.com

Pendahuluan

Knowledge Base atau basis pengetahuan adalah bagian terpenting dari kecerdasan buatan seperti sistem pakar, mesin learning, data mining, sistem keputusan dan komponen kecerdasan buatan lainnya (Ma, zhang, Dong, dkk, 2008). Salah satu bagian terpenting dalam *knowledge base* atau basis pengetahuan adalah Akuisisi pengetahuan (Ma, zhang, Dong, dkk, 2008). Akuisisi pengetahuan sudah berkembang, dulu secara tradisional akuisisi pengetahuan dilakukan dengan cara manual atau mewawancarai pakar atau ahli, sehingga di dapatkan data dengan menuliskan hasil wawancara dan menggambarkannya. Salah satu cara menggambarkan akuisisi pengetahuan dengan algoritma *Bayesian Network* yaitu dengan struktur *Bayesian Network*. *Bayesian Network* atau *Belief Network* adalah sebuah graf *acyclic* berarah sebagai mesin inferensi yang merepresentasikan sebuah informasi peluang dari hubungan antara beberapa variabel (Ghosh J. K., 1999).

Metode *Bayesian Network* ini saat populer di salah satu ilmu kecerdasan buatan yaitu pada sistem pakar. Sebelumnya telah dilakukan penelitian tentang sistem pakar untuk mendiagnosa penyakit mata menggunakan metode *Naive Bayes* (Rahmad, Yanti, dan Nazri, 2014). Pada penelitian tersebut dapat mengetahui klasifikasi penyakit mata dan informasi tentang pengobatan terbaik. Selain itu metode *Bayesian Network* juga pernah diterapkan untuk mendiagnosa penyakit mata pada sistem pakar (Kurniawan, 2011) Metode *Bayesian Network* memiliki banyak kelebihan diantaranya *Bayesian Network* sangat bagus untuk kasus yang tidak pasti dan sangat mudah diterapkan pada bidang apapun (Marcot, 2001), *Bayesian Network* ini memiliki keakuratan yang baik (Uusitalo, 2007).

Salah satu bagian terpenting dalam *knowledge base* atau basis pengetahuan adalah Akuisisi pengetahuan (Ma, zhang, Dong, dkk, 2008). Akuisisi pengetahuan sudah berkembang, dulu secara tradisional akuisisi pengetahuan dilakukan dengan cara manual atau mewawancarai pakar atau ahli, sehingga di dapatkan data dengan menuliskan hasil wawancara dan menggambarkannya. Salah satu cara menggambarkan akuisisi pengetahuan dengan algoritma *Bayesian Network* yaitu dengan struktur *Bayesian Network*.

Walaupun memiliki kelebihan dan banyak digunakan oleh para peneliti, *Bayesian Network* juga memiliki kelemahan yaitu sulit untuk mencapai kesepakatan beberapa pakar dalam membangun struktur *Bayesian Network*, pakar juga akan memerlukan waktu yang lama hanya untuk membangun struktur *Bayesian Network*, Pakar akan sulit menentukan *conditional*

probability berdasarkan pendapat pakar. Oleh karena itu untuk mengatasi kelemahan *Bayesian Network* dan dalam menengahi kasus akuisisi pengetahuan yang selama ini tradisional yaitu manual atau mewawancarai pakar, maka diperlukanlah ilmu kecerdasan buatan lain yaitu data mining dengan teknik *Association Rule* untuk membangun struktur *Bayesian Network* berdasarkan fakta dan data

Ilmu pengetahuan dari pakar dapat di ganti dengan menggunakan pengetahuan dari data set yang akan digunakan. Jika dalam sistem pakar nilai *probability* ditentukan berdasarkan kesepakatan para pakar maka dengan menggunakan data mining nilai *probability* didapatkan berdasarkan fakta dan data. Pada penelitian ini menggunakan salah satu algoritma pada *Association Rules* data mining yaitu algoritma *FP-growth*. Menurut David Samuel, bahwa Algoritma *FP-Growth* dapat langsung mengekstrak *item* yang sering muncul (*frequent itemset*) dari pohon *prefix* atau yang disebut dengan *FP tree* yang telah terbentuk sebelumnya. Sehingga dinilai cukup efektif dan efisien dalam menentukan *frequent itemset*. Algoritma ini dipilih dikarenakan dapat mengoptimalkan penentuan pencarian *frequent itemset* yang lebih ringkas (Samuel, 2008).

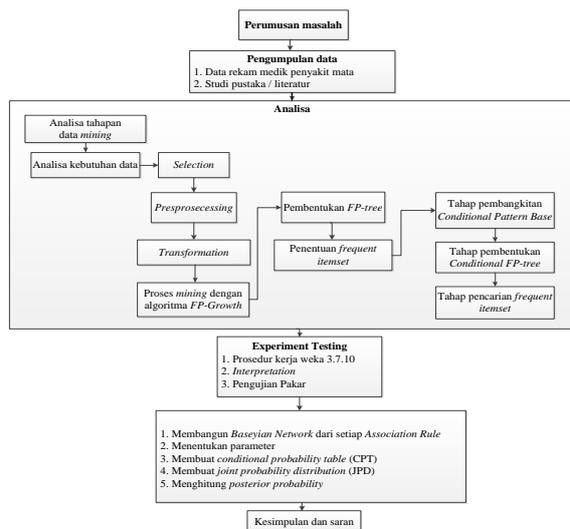
Sebelumnya telah dilakukan penelitian tentang implementasi algoritma *FP-growth* (Pinheiro, 2013) Tujuan penelitian tersebut untuk menemukan pola dari data kanker hati. Sehingga menggunakan algoritma *Frequent Pattern Growth (FP-Growth)* untuk menemukan pola tersebut. Pada penelitian ini mengambil kasus pada data rekam medik penyakit mata, dari salah satu rumah sakit mata Pekanbaru. Dalam bidang medis penyakit mata merupakan salah satu penyakit yang sering terjadi pada manusia dengan berbagai kesulitan dalam menentukan mana gejala yang paling berpengaruh dalam penyakit mata tersebut. Kesulitan itu dapat dibantu dengan merepresentasikan gejala penyakit mata kedalam sebuah model grafis dalam *Baseyan Network* (Kurniawan, 2011) dan menghasilkan nilai probabilitas pada setiap gejala penyakit mata berdasarkan fakta dan data. Sehingga akan dilakukan Optimasi basis pengetahuan menggunakan algoritma *FP-Growth* untuk membangun struktur *Bayesian Network* Penyakit mata.

Proses dari penelitian ini nantinya akan dilakukan dengan mengumpulkan data rekam medik penyakit mata dari salah satu rumah sakit mata Pekanbaru secara manual. Data Penyakit Mata selanjutnya akan dilakukan dengan menganalisa tahapan data mining, dimana data tersebut akan melalui tahapan; *Selection*, *Preprocessing*, *Transformation*, dan Proses

mining dengan Algoritma *FP-Growth*. Tahapan Proses mining dengan Algoritma *FP-Growth* ada dua yaitu Pembentukan *FP-tree* dan Pembentukan *frequent itemset*. Tahapan selanjutnya adalah Pembangkitan *Conditional Pattern Base*, Pembentukan *Conditional FP-tree* dan Pencarian *frequent itemset*. Setelah tahapan mining selesai, akan dilakukan Pengujian ke dalam tools Weka 3.7.10. Hasil pengujian Weka yang memenuhi ketentuan diuji kembali dengan Pengujian Pakar dan nantinya akan dilanjutkan dengan membangun Struktur Bayesian Network dan mencari nilai probabilitas pada setiap gejala penyakit mata.

Metodologi Penelitian

Metodologi penelitian merupakan pedoman dalam melakukan penelitian. Di dalam metodologi penelitian terdapat langkah-langkah penelitian yang sudah ditentukan sebelumnya agar penelitian tidak menyimpang dari tujuan. Tahapan penelitian dapat dilihat pada gambar berikut ini.



Gambar Metodologi Penelitian

A. Analisa

Setelah mendapatkan literatur yang berhubungan dengan permasalahan yang akan diteliti pada penelitian ini, maka tahap selanjutnya yaitu menganalisa permasalahan.

1. Perumusan masalah merupakan tahapan awal dari metodologi penelitian. Adapun perumusan masalah pada penelitian ini adalah “Bagaimana menemukan pola hubungan probabilitistik gejala penyakit terhadap diagnosa penyakit mata menggunakan teknik *Association Rule*

untuk membangun struktur *Bayesian Network* menggunakan algoritma *FP-Growth*”.

2. Pengumpulan data bertujuan untuk memperoleh data-data dan informasi yang berhubungan dengan penelitian yang akan dilakukan. Pada penelitian ini, akan dilakukan pengumpulan data sebagai berikut :
 - a. Data Rekam medik penyakit mata
 Pada penelitian ini merupakan data rekam medik penyakit mata dari Rumah Sakit Mata kota Pekanbaru. Data yang digunakan pada tahun 2016 sebanyak 1599 data.
 - b. Studi Pustaka/literatur merupakan tahapan metodologi penelitian lanjutan dari tahapan sebelumnya. Tahapan ini dilakukan dengan tujuan untuk mendapatkan teori-teori dari berbagai sumber, misalnya buku, jurnal atau *paper*, *media online* dan penelitian-penelitian sebelumnya yang berkaitan dengan penelitian yang akan dilaksanakan ini.
3. Analisa

Tahapan analisa merupakan tahapan pemahaman terhadap suatu permasalahan sebelum mengambil tindakan dan keputusan.

- a. Analisa Kebutuhan Data
 Pada tahapan ini dilakukan pengumpulan terhadap data yang akan digunakan pada penelitian.
- b. Analisa Tahapan Data Mining
 - a. *Selection*
 - b. *Preprocessing*
 - c. *Transformation*
 - d. Proses *mining* dengan Algoritma *FP-growth*

B. Experiment Testing

- a. Prosedur Weka 3.7.0

Setelah hasil analisa didapatkan dalam bentuk *rules* yang menyatakan hubungan antar atribut, selanjutnya data diuji dengan menggunakan WEKA 3.7.0. Pengujian dilakukan dengan memasukkan *minimum support* dan *minimum confidence* yang berbeda-beda dengan algoritma yang sama dan jumlah data yang sama dengan tujuan untuk mengetahui

pengaruhnya terhadap *rules* yang terbentuk serta mencari nilai *minimum support* dan *minimum confidence* yang tepat untuk menghasilkan *rules* yang memenuhi ketentuan penelitian.

b. *Interpretation*

Dalam tahapan ini hasil dari proses *mining* berupa pola-pola atau aturan (*rules*) yang diperoleh akan ditampilkan dalam bentuk yang mudah dimengerti.

c. Pengujian Pakar

Pada tahapan ini akan diuji kembali hasil dari pengujian *association rule mining* dengan algoritma *FP-Growth* terhadap pakar untuk membangun Struktur *Bayesian Network*.

C. Bayesian Network

Bayesian Network merupakan metode yang digunakan untuk menarik kesimpulan besar kemungkinan gejala yang berpengaruh terhadap penyakit mata. Terdapat beberapa langkah untuk menerapkan *Bayesian Network*. Langkah-langkah tersebut diantaranya:

- a. Membangun struktur *Bayesian Network* penyakit mata
- b. Menentukan parameter
- c. Membuat *conditional probability table* (CPT)
- d. Membuat *joint probability distribution* (JPD)
- e. Menghitung *posterior probability*

D. Kesimpulan dan Saran

Pada tahapan ini berisikan kesimpulan dari hasil penelitian yang nantinya akan mendapatkan dalam mencari pola hubungan probabilistik antara gejala penyakit mata terhadap penyakit mata untuk mendapatkan mana gejala yang paling berpengaruh dalam penyakit mata tersebut, membangun stuktur *Bayesian Network* dan mendapatkan nilai probabilitas.

ANALISA

Berikut adalah beberapa tahapan analisis yang telah dilakukan:

A. Analisa Kebutuhan Data

Pada tahap ini dilakukan analisa terhadap kebutuhan data. Jumlah data yang dikumpulkan sebanyak 1.599 data dengan 6 atribut.

Tabel 4.1 Atribut Rekam Medik

<i>Atribut</i>	<i>Keterangan</i>
No.Rekam Medik	Nomor Pasien
Tanggal	Tanggal melakukan <i>checkup</i>
Jenis Kelamin	Jenis Kelamin pasien
Umur	Umur
Gejala atau Keluhan	Keluhan Pasien terhadap penyakit yang dideritanya
Penyakit	Penyakit Pasien

B. Analisa Tahapan Data Mining

a. *Selection*

Pada tahapan ini dilakukan pemilihan data yang akan digunakan pada penelitian. Pemilihan atribut yang akan digunakan pada tahapan ini hanya 4 yaitu Umur, Jenis kelamin, Gejala, dan penyakit dari 6 atribut yang tersedia pada data *rekam medik* penyakit mata.

b. *preprocessing*

Setelah data diseleksi proses selanjutnya yaitu *pre-processing* dengan menghilangkan *noise* atau membersihkan data yang *missing value*, *inconsistent data*, *outlier*, dan lain-lain.

c. *Transformasi*

Proses selanjutnya adalah transformasi data. Data ditransformasi dengan menginisialisasi atribut dan mentransformasi data ke format yang dapat dibaca oleh *tools* WEKA 3.7.0.

Tabel 4.2 Atribut Dan Kode Atribut

No	Kolum	Kode atribut	Atribut
1	Umur	U1	Bayi
		U2	Anak awal
		U3	Anak tengah
		U4	Remaja awal
		U5	Remaja akhir
		U6	Dewasa awal
		U7	Dewasa Tengah
		U8	Dewasa akhir

No	Kolum	Kode atribut	Atribut
2	JK	P	Perempuan
		L	Laki-laki
3	Gejala	G1	Kabur jauh
		G2	Kabur dekat
		G3	Silau
		G4	Berasap
		G5	Berlemak
		G6	Pusing
		G7	Mata mengganjal
		G8	Berair
		G9	Mata gatal
		G10	Mata lengket
		G11	Mata merah
		G12	Bengkak
		G13	Mata sakit
		G14	Mata mudah lelah
		G15	Perih
		G16	Berpasir
		G17	Kering
		G18	Kabur
		G19	Trauma
		P1	Astigmatisme Miopia
		P2	Conjungtivitis
		P3	Dry Eye
4	Penyakit	P4	KSI
		P5	KSM
		P6	Miopia
		P7	Presbiopia
		P8	Pterigium
		P9	Trauma

- d. Proses Mining dengan Algoritma FP-Growth membahas bagaimana cara menerapkan proses Mining dengan Algoritma FP-Growth untuk mendapatkan *frequent itemset* yang memenuhi *minimum support* dan *minimum confidence*. Untuk penentuan *minimum support* dan *minimum confidence* tidak boleh terlalu tinggi karena hal tersebut akan mengakibatkan *rule* tidak akan muncul, begitu juga sebaliknya tidak boleh terlalu rendah karena akan mengakibatkan *rule* yang didapat akan terlalu banyak.

PENGUJIAN

Tahapan ini merupakan tahapan pengujian proses mining menggunakan Weka. Sebelum melakukan pengujian, data penyakit mata ditransformasi ke dalam format *.arff agar data dapat dibaca oleh WEKA 3.7.10. Dengan

ketentuan jika *items* termasuk dalam satu kasus maka menjadi t (TRUE), sedangkan yang tidak termasuk dalam kasus menjadi ? (FALSE). Transformasi data dalam format *.arff dapat dilihat pada tabel 5.1 dibawah ini.

Tabel 5.1 Transformasi data dalam format *.arff

U1	U2	..	P	L	G1	G2	..	P1	..	P
?	?	?	?	t	t	?	?	t	?	?
?	?	?	?	t	t	?	?	t	?	?
?	?	?	t	?	?	?	?	t	?	?
?	?	?	t	?	?	?	?	t	?	?
?	?	?	?	t	t	?	?	t	?	?
?	?	?	?	t	t	?	?	t	?	?
?	?	t	?	t	t	?	?	t	?	?
?	?	?	?	t	t	?	?	t	?	?

A. Hasil Pengujian

Pengaruh *minimum support* dan *minimum confidence* terhadap jumlah *rules* yang dihasilkan adalah berbanding terbalik. Hal ini dapat dilihat dari semakin tingginya nilai *minimum support* (minsup) dan *minimum confidence* (mincon) maka semakin berkurangnya *rules* yang dihasilkan. Jumlah *rules* dapat dilihat pada Tabel 5.2 di bawah ini.

Tabel 5.2 Pengaruh minsup dan mincon

Minsup	3%	10%	20	..	90	100
Minconf	%	%	%	%	%	%
10%	13173	1208	54	0	0	0
20%	11962	1206	54	0	0	0
30%	10430	1167	54	0	0	0
40%	8298	1126	54	0	0	0
50%	6459	1038	51	0	0	0
60%	4815	815	50	0	0	0
70%	3895	732	44	0	0	0
80%	2870	459	38	0	0	0
90%	2208	359	24	0	0	0
100%	0	0	0	0	0	0

B. Interpretasi

Hasil pengujian *association rule mining* dengan algoritma FP-Growth didapat 2208 *rules* yang memenuhi *minimum support* dan *minimum confidence* dengan nilai *lift ratio* ≥ 1 sehingga *rules* yang dihasilkan layak untuk digunakan. Kemudian dari *rules* yang didapat, hanya 24 *rules* yang memenuhi *antecedent* dan *consequent* pada penelitian ini yaitu dengan *antecedent* adalah jenis kelamin, umur, dan gejala sedangkan *consequent* adalah Penyakit mata. Dalam hal

ini terdapat 9 jenis Penyakit Mata, namun yang merupakan *frequent itemset* dan termasuk kedalam *rules* yang memenuhi ketentuan hanya 2 jenis penyakit mata yakni P7 dan P2 yang berarti Penyakit *Presbiopia* dan *Conjunctivitis*. Pengujian dilakukan dengan minimum support 3% dan minimum confidence 90%. *Rules* yang memenuhi ketentuan *antecedent* dan *consequent* dapat dilihat pada Tabel 5.3 di bawah ini.

Tabel 5.3 Rules yang memenuhi ketentuan Jika Maka

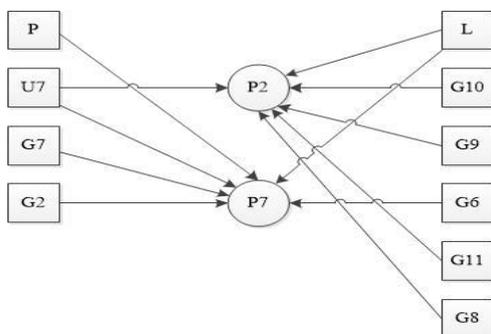
<i>Antecedent</i>	<i>C</i>	<i>S</i>	<i>Confid</i>	<i>Lif R</i>
P∧U7∧G7	P7	4.53	1.0	7.27
L∧U7∧G9	P2	3.54	1.0	5.4
L∧U7∧G10	P2	3.13	1.0	5.4
L∧U7∧G2	P7	4.28	1.0	7.27
L∧U7∧G7	P7	3.29	1.0	7.25
P∧U7∧G6∧G2	P7	4.94	1.0	7.27
P∧U7∧G6∧G7	P7	4.36	1.0	7.27
P∧U7∧G2∧G7	P7	4.36	1.0	7.27
L∧U7∧G6∧G2	P7	3.45	1.0	7.27
L∧U7∧G6∧G7	P7	3.21	1.0	7.25
L∧U7∧G11∧G9	P2	3.13	1.0	5.4

C. Pengujian Pakar

Pada tahapan ini akan diuji kembali hasil dari pengujian *association rule mining* dengan algoritma *FP-Growth* dalam hal ini terdapat 24 *rule* dapat dilihat pada Tabel 5.2 di atas yang akan diuji oleh pakar mata dan hasil pengujian sebesar 99% benar.

D. Membangun Struktur Bayesian Network penyakit mata

Struktur *Bayesian Network* akan merepresentasikan antara jenis penyakit mata dengan umur, jenis kelamin, dan gejala penyakit mata berdasarkan *Association Rules* yang terbentuk. Struktur *Bayesian Network* dapat dilihat pada Gambar 5.4 di bawah ini.



Gambar 5.4 Struktur Bayesian Network Penyakit mata

E. Menentukan Parameter

Nilai *prior probability* atau nilai kepercayaan dari gejala penyakit mata merupakan nilai yang muncul untuk menjelaskan besar kepercayaan dari setiap gejala pada penyakit mata berdasarkan data. Untuk setiap gejala yang direpresentasikan pada struktur *Bayesian Network* mempunyai estimasi parameter yang didapat dari data yang telah ada dapat dilihat pada Tabel 5.5 di bawah ini.

Tabel 5.1 Prior Probability

Kode Gejala	Nama	Nilai
P	Perempuan	0.54
L	Laki-laki	0.45
U7	Dewasa tengah	0.42
G2	Kabur dekat	0.12
G6	Pusing	0.30
	Mata	0.10
G7	mengganjal	0.10
G8	Berair	0.20
G9	Mata gatal	0.18
G10	Mata lengket	0.14
G11	Mata merah	0.25

F. Conditional probability Table (CPT)

Conditional probability adalah probabilitas suatu *event* B terjadi apabila *event* A sudah terjadi. Sebuah Tabel yang berisi probabilitas dari setiap kemungkinan nilai dari A dan B disebut dengan *conditional probability table* (CPT). Contoh *Conditional probability table* dari penyakit mata terhadap gejala penyakit mata dapat dilihat pada Tabel 5.6 di bawah ini.

Tabel 5.2 P untuk P7

P	P7	
	present	Absent
Present	0.58	0.46
Absent	0.42	0.54

Pada Tabel 5.5 diatas Nilai *Conditional probability* P untuk P7 dimana P (perempuan) adalah P7 (*Presbiopia*) present = 0.58 dan dimana P (perempuan) tetapi tidak P7 (*Presbiopia*) absent = 0.46.

G. Membuat Joint Probability Distribution (JPD)

Langkah keempat adalah menentukan nilai *joint probability distribution* (JPD). Probabilitas kemunculan bersama untuk semua kombinasi

kemungkinan nilai-nilai yang terdapat pada variabel A dan B disebut *joint probability distribution* (JPD). Sama seperti CPT, *joint probability distribution* dari suatu variabel A dan B adalah sebuah Tabel yang berisi probabilitas untuk setiap nilai A dan B yang dapat terjadi. Untuk mendapatkan *joint probability distribution* yaitu dengan cara menghitung hasil kali antara *conditional probability* dengan *prior probability*.

Misalkan akan dihitung *joint probability distribution* P. *Prior probability* P present adalah 0.58, sedangkan *absent* 0.42. Dapat dilihat pada Tabel 5.3 Sehingga dapat diperoleh *joint probability distribution* (JPD) dari P yaitu:

P	P7	
	Present	Absent
Present	$0.54 \times 0.58 = 0.3132$	$0.46 \times 0.46 = 0.2116$
Absent	$0.54 \times 0.42 = 0.2268$	$0.46 \times 0.54 = 0.2484$

H. Penghitungan Posterior Probability

Langkah kelima adalah menghitung nilai *posterior probability*. Untuk mendapatkan nilai *posterior probability*, dapat dihitung dari hasil JPD yang telah diperoleh, kemudian nilai inilah yang digunakan untuk menghitung probabilitas kemunculan suatu gejala. Berikut ini diberikan contoh cara menghitung *posterior probability* P, dilihat dari Tabel 5.15. Berdasarkan JPD tersebut, dapat dihitung *posterior probability* dari P :

$$P(A|B) = \frac{P(B \cap A)}{P(B)}$$

$$P(A|B) = \frac{P(A)P(B|A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$

$$= \frac{0.3132}{0.3132 + 0.2116} = 0.59679878$$

Jadi, jika seseorang berjenis kelamin perempuan, maka dia berkemungkinan mengalami penyakit Presbiopia sebesar 0.59679878 atau 60%

Sehingga didapatkan nilai *posterior probability* untuk umur, jenis kelamin, dan gejala penyakit mata 12 atribut dapat dilihat pada Tabel 5.7 di bawah ini.

Tabel 5.7 nilai posterior probability

Kode	Nama	Kode P	Nama P	Nilai	P
P	Perempuan	P7	Presbiopia	0.59679878	60%
L	Laki-laki	P2	Conjungtivitis	0.53191489	53%
L	Laki-laki	P7	Presbiopia	0.46240602	46%
U7	Dewasa tengah	P7	Presbiopia	0.64855356	65%
U7	Dewasa tengah	P2	Conjungtivitis	0.42687747	43%
G2	Kabur dekat	P7	Presbiopia	0.98933075	98%
G6	Pusing	P7	Presbiopia	0.64343164	64%
G7	Mata tidak nyaman	P7	Presbiopia	0.95174263	95%
G8	Berair	P2	Conjungtivitis	0.92307692	92%
G9	Mata gatal	P2	Conjungtivitis	0.98596371	98%
G10	Mata lengket	P2	Conjungtivitis	1.0	100%
G11	Mata merah	P2	Conjungtivitis	0.72727273	72%

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dan lembar pengujian yang telah dilakukan bersama pakar terdapat beberapa kesimpulan, diantaranya adalah sebagai berikut:

1. Menemukan hubungan *probability*
 - a. Jenis kelamin, terdapat dua jenis kelamin yang berpengaruh untuk penyakit *presbiopia* yaitu perempuan dan laki-laki, tetapi jenis kelamin yang paling berpengaruh adalah perempuan, dibuktikan dengan nilai probabilitas sebesar 60% sedangkan untuk penyakit *conjungtivitis* jenis kelamin yang paling berpengaruh adalah laki-laki dibuktikan dengan nilai probabilitas sebesar 53%.
 - b. Umur, pada penelitian ini terdapat delapan bagian umur, tetapi hanya satu umur yang didapat dan paling berpengaruh untuk penyakit *presbiopia* yaitu Dewasa tengah dari

- umur 31-59, dibuktikan dengan nilai probabilitas sebesar 65% dan untuk penyakit *conjungtivitis* sebesar 43%.
- c. Gejala, terdapat tiga gejala yang berpengaruh untuk penyakit *presbiopia* yaitu Kabur dekat, pusing, dan mata tidak nyaman, tetapi diantara tiga gejala tersebut yang paling berpengaruh adalah kabur dekat dibuktikan dengan nilai probabilitas sebesar 98%, sedangkan untuk penyakit *conjungtivitis* terdapat empat gejala yang berpengaruh yaitu berair, mata gatal, mata lengket, dan mata merah, tetapi gejala yang paling berpengaruh adalah mata lengket dibuktikan dengan nilai probabilitas sebesar 100%.
 2. Berhasil dalam membangun struktur *Bayesian Network* berdasarkan fakta dan data, dengan hasil pengujian rules diuji kembali dengan pakar dan hasil pengujian sebesar 99% benar.
 3. Berhasil dalam mengatasi kelemahan *Bayesian Network* yang sulit dalam menentukan nilai probabilitas pada setiap gejala penyakit berdasarkan fakta dan data.

Daftar Pustaka

- [1.] Ma, D. L., zhang, w. j., Dong, B., Yang, P., dan Lu, H. X. (2008). Establishing Knowledge Base of Expert System With Association Rules. *International*.
- [2.] Ghosh, J. K. (1999). *Probabilistic Bayesian Network Model Building of Heart Disease*. Columbia.
- [3.] Rahmad, K., Yanti, N., dan Nazri, M. Z. (2014). "Expert system for self-diagnosing of eye diseases using Naive Bayes". *Advanced informatics: Concept, Theroey and application (ICAICTA), International Conference of IEEE*.
- [4.] Kurniawan, R. (2011). *Sistem pakar untuk mendiagnosa penyakit mata menggunakan metode Bayesian Network*. Pekanbaru: Universitas Sultan Syarif Kasim Riau.
- [5.] Marcot, B. G. (2001). Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from environmental impact statement. *Forest ecology and management* . 153(1) 29-42.
- [6.] Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological modelling*,. 203(3), 312-318.
- [7.] Hart, B., Pollino, C. (2008). *Increased use of Bayesian Network models will improve ecological risk assessments*.
- [8.] Kragt, M. E. (2009). *A beginners guide to Bayesian Network modelling for integrated catchment management. Landscape Logic*.
- [9.] Samuel, D. (2008). *Penerapan Struktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung: 1.
- [10.] Pinheiro, F. K. (2013). Extracting association rules from liver cancer data using the FP-growth algorithm. In *Computational Advances in Bio and Medical Sciences (ICCABS). IEEE 3rd International Conference*.
- [11.] Fayyad, U., Piatetsky-Shapiro, G., dan Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence*.
- [12.] Han, J. (2006). *Data Mining: Concepts and Techniques*. 2nd edition. Elsevier Inc.
- [13.] Tian, D. G. (2013). A Bayesian Association Rule Mining Algorithm. In *2013 IEEE International Conference on Systems, Man, and Cybernetics (pp. 3258-3264). IEEE*.
- [14.] Heckerman, D. (1998). *A tutorial on learning with Bayesian Networks. In Learning in graphical model*. Springer Netherlands.