

## Diagnosa Gejala yang Muncul Bersamaan pada Penderita Tuberculosis Menggunakan Algoritma Apriori dengan Substitusi Metode Bayesians pada Nilai Confidence

Delfany Arcadia Valeska<sup>1\*)</sup>, Fajri Rakhmat Umbara<sup>2)</sup>, Puspita Nurul Sabrina<sup>3)</sup>

<sup>1)2)3)</sup> Informatika, Fakultas Sains dan Informatika, Universitas Jenderal Achmad Yani

Correspondence author : [delfanyav18@if.unjani.ac.id](mailto:delfanyav18@if.unjani.ac.id), Cimahi, Indonesia

DOI: <https://doi.org/10.37012/jtik.v8i1.1105>

### Abstrak

*Tuberculosis* merupakan salah satu penyakit infeksi yang menyerang saluran pernapasan manusia yang berasal dari bakteri *mycobacterium*. Berdasarkan data WHO pada tahun 2012, *Tuberculosis* adalah salah satu penyakit menular yang menyebabkan masalah kesehatan terbesar kedua di dunia. Agar permasalahan tersebut dapat diatasi, diperlukan metode yang dapat mendiagnosa gejala yang muncul bersamaan pada penyakit *Tuberculosis*. Pada penelitian ini, digunakan metode *Association Rules* dengan algoritma Apriori yang dimodifikasi dengan metode Bayesians untuk mendapatkan aturan asosiasi yang memenuhi nilai *minimum support* dan nilai *minimum confidence* dengan beberapa jumlah *itemset* dalam keseluruhan gejala *tuberculosis* untuk mencapai nilai persentase *support* dan nilai persentase *confidence* terbaik dalam mendiagnosa gejala yang muncul bersamaan pada penderita *tuberculosis*. Modifikasi algoritma apriori dilakukan dengan mensubstitusi metode bayesians pada formula *confidence* di apriori standar. Tujuan penggunaan algoritma Apriori pada permasalahan ini adalah untuk mengetahui gejala apa saja yang muncul bersamaan pada penderita *tuberculosis*. Adapun hasil dari penelitian ini adalah rule pertama pada apriori non modifikasi untuk nilai minimum support 50% menghasilkan nilai confidence sebesar 100% dan nilai confidence pada apriori yang dimodifikasi sebesar 39,6%. Sedangkan rule kedua pada apriori non modifikasi untuk nilai minimum support 50% menghasilkan nilai confidence sebesar 75% dan nilai confidence pada apriori yang dimodifikasi sebesar 90%.

**Kata kunci:** Tuberculosis, Algoritma Apriori, Metode Bayesians, Nilai Confidence

### Abstract

*Tuberculosis* is an infectious disease that attacks the human respiratory tract, which originates from the *mycobacterium*. Based on WHO data in 2012, tuberculosis is an infectious disease that causes the second largest health problem in the world. So that these problems can be overcome, we need a method that can diagnose symptoms that appear together in Tuberculosis disease. In this study, the Association Rules method was used with the Apriori algorithm modified with the Bayesian method to obtain association rules that meet the minimum support value and minimum confidence value with several number of itemsets in all tuberculosis symptoms to achieve the best support percentage value and confidence percentage value in diagnosing symptoms coexisting in patients with tuberculosis. Modification of the a priori algorithm is carried out by substituting the Bayesian method in the standard a priori confidence formula. The purpose of using the Apriori algorithm in this problem is to find out what symptoms appear together in tuberculosis sufferers. The results of this study are that the first rule on the non-modified a priori for a minimum support value of 50% produces a confidence value of 100% and the confidence value for the modified a priori is 39.6%. Whereas the second rule on the non-modified a priori for a minimum support value of 50% produces a confidence value of 75% and the confidence value for the modified a priori is 90%.

**Keywords:** Tuberculosis, Apriori Algorithm, Bayesian Method, Confidence Value

## PENDAHULUAN

Tuberculosis (TBC) merupakan salah satu penyakit menular, yang mana terjadi infeksi pada saluran pernafasan yang berasal dari kuman *mycobacterium* (Kurniasari & Kurniasari, 2013). Tuberculosis ini merupakan penyakit menular dan terdapat berbagai kasus yang bersifat mematikan. Tuberculosis dapat disebarkan melalui udara, penyebaran berlangsung saat penderita Tuberculosis tersebut batuk atau bersin (Aini et al., 2017). Ketika mengalami gejala batuk disertai dengan demam, terkadang masyarakat saat ini masih menganggap itu bukan penyakit yang serius dan tidak melakukan pemeriksaan lebih lanjut untuk memastikan tentang gejala yang dirasakan.

Sebagai salah satu jenis penyakit berbahaya, diperlukan adanya kesadaran dan pemahaman masyarakat mengenai penyakit tuberculosis ini dengan baik. Di Indonesia sendiri, TBC termasuk dalam satu dari 10 penyebab kematian. Berdasarkan data WHO pada tahun 2019, jumlah kasus TBC di Indonesia sebanyak 843.000 orang. Menurut data TBC Indonesia tahun 2020, jumlah kasus TBC meningkat menjadi 845.000 dan jumlah kematian lebih dari 98.000 orang (ADM PLK, 2021).

Jumlah pasien TBC di Indonesia mengalami peningkatan. Hal tersebut disebabkan karena masih banyaknya masyarakat yang tidak menerapkan pola hidup sehat, dan juga kurangnya informasi mengenai gejala yang muncul pada penyakit TBC (Santosa et al., 2018). Oleh karena itu untuk mengetahui pola data gejala yang muncul bersamaan pada penderita tuberculosis, diperlukan suatu cara untuk menganalisa pola gejala tuberculosis yaitu dengan memanfaatkan *data mining*.

Penelitian yang telah dilakukan untuk diagnosa terhadap penyakit tuberculosis yaitu dengan menerapkan Algoritma *Decision Tree C4.5* (Santosa et al., 2018). C4.5 merupakan salah satu algoritma yang melakukan klasifikasi pada data mining. Algoritma ini bekerja dengan membentuk sebuah pohon keputusan (*decision tree*), hasil yang ditampilkan berupa rule. Penelitian ini menghasilkan sebuah sistem yang dapat membantu masyarakat dalam memperoleh informasi yang tepat serta cepat untuk melakukan diagnosa terhadap penyakit Tuberculosis. Penelitian ini telah menghasilkan sistem pakar untuk melakukan diagnosa terhadap penyakit Tuberculosis dan telah dilakukan pengujian kepada data pasien sebanyak 100 data, yang mana 50 merupakan data training dan 50 merupakan data testing. Dengan penggunaan *Confusion Matrix*, maka didapatkan nilai akurasi sebesar 90%.

Penelitian sebelumnya yang berhubungan dengan modifikasi Algoritma Apriori adalah Algoritma Apriori yang dimodifikasi dengan menggunakan *Top Down Approach* (Shah, 2017). Modifikasi ini dilakukan dengan menghilangkan kumpulan item yang jarang dari sekumpulan kandidat. Dengan *top down approach*, berhasil ditemukan *frequent itemsets* tanpa melalui beberapa iterasi, sehingga meminimalisir waktu dan ruang.

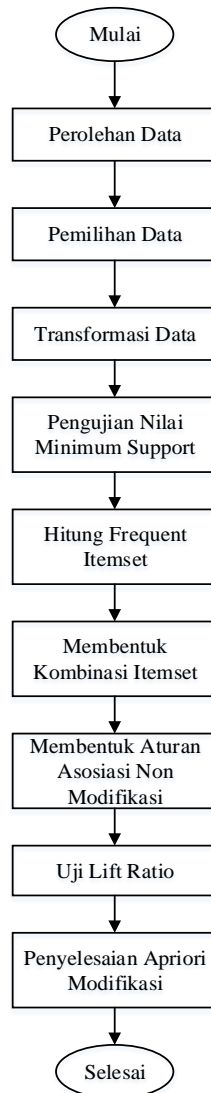
Penelitian lain untuk modifikasi Algoritma Apriori adalah penggunaan algoritma baru yang disebut *Reduced Apriori Algorithm with Tag* (RAAT) (Yu et al., 2008). Algoritma tersebut mengurangi satu operasi pemangkasan yang berlebihan dari C2 dan juga mengoptimalkan operasi subset melalui tag transaksi untuk mempercepat perhitungan support.

Banyak penelitian sebelumnya yang sudah menggunakan berbagai metode dalam melakukan diagnosa terhadap penyakit tuberculosis. Dan untuk kesempatan ini penulis mencoba menerapkan Metode Association Rule menggunakan Algoritma Apriori yang dimodifikasi dengan metode Bayesian (Ginting et al., 2018), dimana modifikasi dilakukan pada penentuan nilai *confidence* dengan mensubstitusi metode peluang bayesian terhadap aturan yang dihasilkan untuk mendiagnosa gejala yang muncul bersamaan pada penderita tuberculosis. Modifikasi ini dilakukan karena nilai *confidence* berperan penting dalam Algoritma Apriori, karena semakin tinggi nilai *confidence* pada aturan yang terbentuk, semakin tinggi pula tingkat kepercayaan dan keterkaitan antar item (Fauzy et al., 2016). Metode *Association Rule* merupakan teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item (Putra, 2018). Sedangkan Algoritma A Priori merupakan suatu algoritma untuk mencari pola hubungan antar satu atau lebih itemset (Purba & Buulolo, 2020). Dengan diketahui pola gejala yang muncul bersamaan pada penderita tuberculosis, maka penyakit tersebut dapat ditangani dengan baik serta dapat mencegah atau mengurangi penyakit tersebut (Kumar et al., 2012). Adapun tujuan penelitian ini adalah mengetahui bagaimana peran apriori yang dimodifikasi dengan Bayesian dibandingkan tanpa Bayesian pada nilai *confidence* terhadap gejala tuberculosis yang muncul bersamaan.

## METODE

Pada metode penelitian ini terdapat beberapa tahapan yang akan dilakukan untuk menemukan hasil yang diinginkan, tahapan pertama adalah perolehan data, tahapan kedua

yaitu pemilihan data yang kemudian dilanjut dengan proses transformasi data, tahapan selanjutnya yakni menghitung *frequent itemset* dan membentuk kombinasi *itemset*, lalu pembentukan aturan asosiasi atau pola kombinasi dari suatu item menggunakan apriori standard dan apriori yang dimodifikasi dengan metode bayesian.



**Gambar 1.** Tahap Penelitian

1. Perolehan Data

Pada tahap ini, peneliti memperoleh data dari *Repository Kaggle Machine Learning* dalam bentuk excel.

2. Pemilihan Data

Pemilihan data meliputi proses memilih atribut apa saja yang akan digunakan pada proses algoritma apriori.

### 3. Transformasi Data

Pada tahapan ini, data yang telah diperoleh dan dilakukan pemilihan atribut kemudian diubah kedalam format data standar untuk proses algoritma apriori.

### 4. Pengujian Nilai Minimum Support

Pengujian ini dilakukan dengan melakukan analisa pengaruh minimum support terhadap jumlah *frequent itemset* yang dibangkitkan algoritma apriori.

### 5. Hitung *Frequent Itemset*

Proses perhitungan *frequent itemset* ditentukan dengan menggunakan nilai *support* pada setiap item yang ada dengan nilai minimum support yang sudah ditentukan.

### 6. Membentuk Kombinasi Itemset

Pembentukan kombinasi *itemset* ditentukan dengan nilai support yang telah dilakukan perhitungan sebelumnya dengan nilai min *support* yang sudah ditentukan.

### 7. Membentuk Aturan Asosiasi Non Modifikasi

Pada tahapan ini hasil *frequent itemset* dari pola kombinasi *itemset* yang memenuhi *minimum support* dari tiap pola kombinasi dijadikan sebagai *Association Rule mining*.

### 8. Uji *Lift Ratio*

Uji *lift ratio* ini digunakan untuk menguji kekuatan rule yang sudah terbentuk.

### 9. Penyelesaian Apriori Modifikasi

Apriori modifikasi ini dilakukan dengan mensubstitusi atau menggantikan formula dari perhitungan *confidence apriori* nonmodifikasi dengan formula peluang bayesian.

## HASIL DAN PEMBAHASAN

Data gejala tuberculosis yang digunakan untuk penelitian ini diperoleh dari *Repository Kaggle Machine Learning*. Dataset ini diperoleh dalam bentuk excel dengan total 2000 record data dan 16 atribut yaitu *no*, *id*, *name*, *gender*, *date*, *time* dan 13 *symptoms* pada penderita tuberculosis. Gejala tuberculosis yang digunakan pada penelitian ini menggunakan bahasa inggris. Dan dibawah ini merupakan dataset gejala tuberculosis, namun disini peneliti hanya menampilkan beberapa data.

**Tabel 1. Dataset Gejala Tuberculosis**

no	id	name	gender	date	time	fever for two weeks	coughing blood	sputum mixed with blood	night sweats	chest pain	back pain in certain parts	shortness of breath
1	4842212462	Raye	Female	6/20/2020	10:47 AM	1	0	0	0	1	0	0
2	6947717450	Brittney	Female	12/31/2020	11:51 AM	1	1	0	1	0	1	1
3	6627053772	Sioux	Male	6/12/2020	6:55 PM	1	1	0	1	0	1	0
4	2885316438	Haley	Male	12/14/2020	10:42 AM	0	1	1	1	0	0	1
5	1090197098	Zelda	Male	9/16/2020	5:48 PM	1	1	0	1	0	1	0
6	3248727829	Karlen	Female	1/17/2021	10:05 AM	0	0	0	1	0	1	1
7	4007497230	Guillema	Female	1/15/2020	11:52 AM	0	0	0	1	0	1	1
8	190127767	Seymour	Female	3/19/2020	3:55 PM	0	1	1	0	1	1	1
9	6034623235	Cordellie	Male	9/1/2020	12:28 PM	0	1	0	1	0	1	1
10	2727691440	Duke	Female	4/4/2020	12:50 PM	1	1	1	0	0	0	0
11	5964617905	Virgie	Female	9/19/2020	12:23 PM	1	0	0	1	0	0	1
12	8328172445	Baxter	Male	9/6/2020	10:57 AM	1	0	0	0	1	0	1
13	7175681267	Eugenio	Female	1/6/2021	8:15 PM	0	0	0	1	0	0	0
14	2097994059	Justinian	Female	11/14/2020	5:24 PM	1	1	0	1	0	1	0
15	3621698019	Salvidor	Female	4/5/2020	10:14 AM	0	0	0	0	1	0	1
16	1834750105	Kaia	Female	2/4/2020	6:43 PM	1	0	0	0	1	0	0
17	6796566092	Tuesday	Male	6/23/2020	2:44 PM	1	0	0	0	1	0	1
18	7489396670	Emmey	Female	5/21/2020	6:56 PM	0	1	1	1	0	1	0
19	2812277351	Elmore	Female	7/28/2020	1:15 PM	1	0	0	1	1	1	0
20	3345283948	Joye	Male	5/17/2020	3:45 PM	1	0	0	1	0	0	0
21	4488609252	Hettie	Female	1/15/2020	2:23 PM	0	1	1	0	1	1	1
22	6643747937	Corrina	Female	2/13/2020	5:37 PM	1	1	0	1	0	1	0
23	8244650435	Falkner	Female	10/23/2020	7:24 PM	0	1	0	1	1	0	0
24	5025335442	Baldwin	Male	12/4/2020	7:09 AM	0	1	0	1	0	1	1
25	2010516257	Cob	Male	3/2/2020	10:44 AM	1	0	1	0	1	0	1

Tahap pemilihan data ini meliputi proses memilih atribut apa saja yang akan digunakan pada proses algoritma apriori. Untuk penelitian ini menggunakan atribut *no*, *name*, *gender*, dan 13 *symptoms* pada penderita tuberculosis.

**Tabel 2. Data Sampel**

no	name	gender	fever for two weeks	coughing blood	sputum mixed with blood	night sweats	chest pain	back pain in certain parts	shortness of breath	weight loss	body feels tired	lumps that appear around the armpits and neck	cough and phlegm continuously for two weeks to four weeks	swollen lymph nodes	loss of appetite
1	Raye	Female	1	0	0	0	1	0	0	1	1	0	1	1	0
2	Brittney	Female	1	1	0	1	0	1	1	0	1	0	0	0	1
3	Sioux	Male	1	0	0	0	0	1	0	0	0	0	0	1	0
4	Haley	Male	0	0	1	0	0	0	1	1	0	1	1	1	0
5	Zelda	Male	1	0	0	0	0	1	0	1	0	1	1	1	1

Pada dataset sampel diatas, angka 1 mengartikan bahwa pasien memiliki gejala TBC tersebut dan angka 0 mengartikan sebaliknya yakni pasien tidak memiliki gejala tersebut.

Pada tahapan transformasi data, data ditransformasikan kedalam format data standar untuk proses algoritma apriori. Pada tabel dibawah ini diambil beberapa data gejala penderita tuberculosis yang digunakan sebagai contoh dalam menjalankan apriori non-modifikasi seperti berikut :

**Tabel 3.** Data Sampel 5 Pasien

Pasien	Gejala
1	Fever for two weeks, chest pain, weight loss, body feels tired, cough and pleghm continuously for two weeks to four weeks, swollen lymph nodes
2	Fever for two weeks, coughing blood, night sweats, back pain in certain parts, shortness of breath, body feels tired, loss of appetite
3	Fever for two weeks, back pain in certain parts, swollen lymph nodes
4	Sputum mixed with blood, shortness of breath, weight loss, cough and pleghm continuously for two weeks to four weeks, swollen lymph nodes
5	Fever for two weeks, back pain in certain parts, weight loss, lumps that appear around the armpits and neck, cough and pleghm continuously for two weeks to four weeks, swollen lymph nodes, loss of appetite

Untuk memudahkan pengerjaannya, penulis mempersingkat gejala diatas menjadi inisial seperti yang dapat dilihat pada tabel berikut.

**Tabel 4.** Penamaan Inisial Gejala

Gejala	Inisial
<i>fever for two weeks</i>	F
<i>coughing blood</i>	CP
<i>sputum mixed with blood</i>	S
<i>night sweats</i>	NS
<i>chest pain</i>	CP
<i>back pain in certain parts</i>	BP
<i>shortness of breath</i>	SOB
<i>weight loss</i>	WL
<i>body feels tired</i>	BFT
<i>lumps that appear around the armpits and neck</i>	L
<i>cough and pleghm continuously for two weeks to four weeks</i>	C
<i>swollen lymph nodes</i>	SLN
<i>loss of appetite</i>	LOA

Kemudian berdasarkan Tabel 4 diatas, data diubah menjadi seperti format seperti berikut :

**Tabel 5.** Transformasi Data

Pasien	Gejala
1	{F, CP, WL, BFT, C, SLN}
2	{F, CB, NS, BP, SOB, BFT, LOA}
3	{F, BP, SLN}
4	{SMB, SOB, WL, C, SLN}
5	{F, BP, WL, L, C, SLN, LOA}

Pengujian nilai *minimum support* dilakukan dengan melakukan analisa pengaruh *minimum support* terhadap jumlah *frequent itemset* yang dibangkitkan algoritma apriori. Pada penelitian ini, nilai minimum support yang diuji mulai dari 10%, 30%, 50%, dan 70%. Pengujian ini dilakukan agar mendapatkan nilai *minimum support* terbaik untuk perhitungan sebuah *itemset*, karena semakin tinggi nilai minimum support yang digunakan maka akan semakin sedikit jumlah *frequent itemset* yang dibangkitkan (Fauzy et al., 2016).

Berikut ini hasil dari pengujian nilai minimum support berdasarkan 5 sampel data :

**Tabel 6.** Pengujian Nilai Minimum Support

Min. Support	Jumlah Frequent Itemset	Jumlah Rule
10%	13	90
30%	10	24
50%	7	24
70%	4	1

Proses perhitungan *frequent itemset* ditentukan dengan menetapkan nilai *support* pada setiap item yang ada dengan nilai min. support = 50%. Dapat dihitung menggunakan persamaan berikut :

$$support (A) = \frac{\sum \text{Penderita yang mengandung gejala } (A)}{\sum \text{Penderita}} * 100\%$$

Maka akan mendapatkan nilai *support* pada setiap *item* yang ada.

Dari data pada tabel 6 diatas, dibentuk pola kombinasi *itemsets* seperti pada tabel berikut:

**Tabel 7.** Tabel K1 itemset

Gejala	Frequent Item	Support
F	4	80%
SLN	4	80%
CB	4	80%
NS	4	80%
WL	3	60%
C	3	60%
BP	3	60%

Tabel 6 diatas menunjukkan bahwa ada 7 item yang bisa dilanjutkan ke dalam proses berikutnya yakni 2-itemset karena memiliki nilai frekuensi kemunculannya di atas 50% dan terdapat juga beberapa item yang tidak memenuhi support. Untuk data yang tidak dipakai untuk 2-itemset tidak dimunculkan pada tabel di atas dikarenakan Item yang mempunyai *frequent item* dibawah 3 atau nilai support dibawah 50% akan di hapus.



Pembentukan kombinasi 2 *itemset* ditentukan dengan nilai support yang telah dilakukan perhitungan sebelumnya dengan nilai min *support* = 50%. Dengan nilai minimum *support* yang telah ditentukan yaitu 50%. Calon kombinasi 2 *itemset* yang tidak memenuhi syarat minimum *support* akan dihilangkan, maka terbentuk 2 *itemset*.

**Tabel 8.** Tabel K2 itemset

Gejala	Frequent Item	Support
CB, NS	4	80%
F, SLN	3	60%
F, CB	3	60%
F, NS	3	60%
F, BP	3	60%
SLN, CB	3	60%
SLN, NS	3	60%
SLN, WL	3	60%
SLN, C	3	60%
CB, BP	3	60%
NS, BP	3	60%
WL, C	3	60%

Hasil *frequent itemset* dari pola kombinasi 2 *itemset* yang memenuhi support minimal dari tiap pola kombinasi diatas dijadikan sebagai Asosiasi Rule mining seperti pada tabel 9 dibawah ini. Kemudian dihitung nilai *confidence* nya menggunakan rumus berikut :

$$confidence = \frac{\sum \text{Penderita yang mengandung gejala A dan B}}{\sum \text{Penderita mengandung gejala A}} * 100\%$$

**Tabel 9.** Asosiasi Rule Mining yang dihasilkan dari pola kombinasi 2 item

Gejala	Support	Support Itemset	Confidence
CB, NS	4	4	100%
NS, CB	4	4	100%
F, SLN	4	3	75%
SLN, F	4	3	75%
F, CB	4	3	75%
CB, F	4	3	75%
F, NS	4	3	75%
NS, F	4	3	75%
F, BP	4	3	75%
BP, F	3	3	100%

SLN, CB	4	3	75%
CB, SLN	4	3	75%
SLN, NS	4	3	75%
NS, SLN	4	3	75%
SLN, WL	4	3	75%
WL, SLN	3	3	100%
SLN, C	4	3	75%
C, SLN	3	3	100%
CB, BP	4	3	75%
BP, CB	3	3	100%
NS, BP	4	3	75%
BP, NS	3	3	100%
WL, C	3	3	100%
C, WL	3	3	100%

Untuk rule yang sudah terbentuk, kemudian dilakukan sebuah evaluasi untuk menguji kekuatannya dengan cara uji lift ratio. Nilai lift ratio sebuah rule didapatkan melalui perbandingan confidence dari rule tersebut dengan benchmark confidence. Perhitungan Lift Ratio dapat dilihat pada rumus berikut (Fauzy et al., 2016) :

$$Lift\ Ratio = \frac{Confidence}{Benchmark\ Confidence}$$

Untuk melakukan perhitungan benchmark confidence dapat menggunakan rumus berikut :

$$Benchmark\ Confidence = \frac{\sum\ Penderita\ dengan\ gejala\ dalam\ consequent}{\sum\ Penderita\ dalam\ basis\ data}$$

Lift Ratio yang lebih besar dari 1 menunjukkan adanya manfaat dari rule tersebut. Semakin tinggi nilai lift ratio, maka semakin tinggi pula kekuatan asosiasinya.

Dengan data sampel yang sama dengan proses apriori non-modifikasi di atas, apriori yang di modifikasi fokus pada penentuan nilai *confidence* yakni nilai kepercayaan atau kepastian dari hubungan itemset-itemset tersebut. Apriori modifikasi ini dilakukan dengan mensubstitusi atau menggantikan formula dari perhitungan *confidence apriori non modifikasi* dengan formula peluang bayesian.

Bayesian digunakan untuk menghitung peluang atau probabilitas suatu peristiwa berdasarkan hubungannya dengan peristiwa lain (Studi Sistem Informasi & Triguna Dharma, 2017). Rumus probabilitas bersyarat  $P(F_i|E)$  untuk sembarang kejadian E dalam algoritma Bayes dapat dituliskan dengan persamaan berikut :

$$P(F_i|E) = \frac{P(F_i) * P(E|F_i)}{P(F_1) * P(E|F_1) + P(F_2) * P(E|F_2) + \dots + P(F_n) * P(E|F_n)}$$

Dimana :

$P(F_i|E)$  : Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis  $F_i$  terjadi jika diberikan bukti (*evidence*) E terjadi

$P(E|F_i)$  : Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis  $F_i$

$P(F_i)$  : Probabilitas awal (priori) hipotesis  $F_i$  terjadi tanpa memandang bukti apapun

$P(E)$  : Probabilitas awal (priori) bukti E terjadi tanpa memandang hipotesis / bukti yang lain.

Kemudian aturan asosiasi yang dihasilkan oleh apriori non modifikasi dijadikan input untuk menghitung nilai *confidence* pada apriori modifikasi dengan bayesian menjadi teknik perhitungan nilai *confidencenya*. Sebagai contoh diambil dari sampel data 1 dan 2 atau *rule* pertama dan kedua pada K-2 Itemset.

**Rule 1 : {CB, NS} →** Ketika pasien mengalami gejala CB (*Coughing Blood*), maka pasien juga mengalami gejala NS (*Night Sweats*).

**Tabel 10.** Tabel Kondisi Peluang {CB, NS}

		Kondisi 1		Jumlah
		Dengan Gejala CB	Tanpa Gejala CB	
Kondisi 2	Dengan Gejala NS	1	1	3
	Tanpa Gejala NS	2	1	2
	Jumlah	3	2	5

$$P(A) = \{CB, NS\}$$

$B_1 = CB; B_2 = NS$ , maka

$$P(B_1) = \frac{n(B_1)}{n(S)} = \frac{n(CB)}{n(S)} = \frac{3}{5} = 0,6$$

$$P(B_2) = \frac{n(B_2)}{n(S)}$$

$$= n(NS) / n(S) = 3/5 = 0,6$$

$$P(A|B1) = n(A) / n(B1)$$

$$= n (CB, NS) / n(CB) = 1/3 = 0,33$$

$$P(A|B2) = n(A)/n(B2)$$

$$= n (CB, NS) / n(NS) = 1/3 = 0,33$$

$$P(A) = P(B1) * P(A|B1) + P(B2) * P(A|B2)$$

$$= 0,6 * 0,33 + 0,6 * 0,33$$

$$= 0,396 \rightarrow 39,6\%$$

**Rule 2: {F, SLN}** → Ketika pasien mengalami gejala F (*Fever for two weeks*), maka pasien juga mengalami gejala SLN (*Swollen Lymph Nodes*).

**Tabel 11.** Tabel Kondisi Peluang {F, SLN}

		Kondisi 1		Jumlah
		Dengan Gejala F	Tanpa Gejala F	
Kondisi 2	Dengan Gejala SLN	3	1	4
	Tanpa Gejala SLN	1	1	2
	Jumlah	4	2	6

$$P(A) = \{F, SLN\}$$

B1 = F; B2 = SLN, maka

$$P(B1) = n(B1) / n(S)$$

$$= n(F) / n(S) = 4/6 = 0,6$$

$$P(B2) = n(B2) / n(S)$$

$$= n(SLN) / n(S) = 4/6 = 0,6$$

$$P(A|B1) = n(A) / n(B1)$$

$$= n (F, SLN) / n(F) = 3/4 = 0,75$$

$$P(A|B2) = n(A)/n(B2)$$

$$= n (F, SLN) / n(SLN) = 3/4 = 0,75$$

$$P(A) = P(B1) * P(A|B1) + P(B2) * P(A|B2)$$

$$= 0,6 * 0,75 + 0,6 * 0,75$$

$$= 0,9 \rightarrow 90\%$$

Dari perhitungan apriori standar dan modifikasi dapat di lihat hasil nilai confidence yang didapatkan, dan hasilnya dapat dilihat sebagai berikut:

**Tabel 12.** Tabel Hasil Sampel Data 1 dan 2 dari K-2 Itemset

Gejala	Confidence	
	Non Modifikasi	Modifikasi
CB, NS	100%	39,6%
F, SLN	75%	90%

Penurunan dan kenaikan nilai yang terjadi terhadap nilai *confidence* dipengaruhi oleh frekuensi relasi item pada kondisi “**dengan gejala**” dan “**tanpa gejala**” pada item-item tersebut.

Untuk pencarian nilai *confidence apriori* non modifikasi memfokuskan terhadap relasi gejala misal A, B yang dibagi dalam jumlah gejala A dengan kondisi **dengan gejala** tanpa menghiraukan kondisi **tanpa gejala**. Namun pada apriori modifikasi yang menggunakan formula bayes memperhitungkan frekuensi dari relasi gejala-gejala dengan menghiraukan kondisi **dengan gejala** dan **tanpa gejala**.

## KESIMPULAN DAN REKOMENDASI

Dalam penelitian ini, menghasilkan sebuah rule atau pola sebuah gejala yang muncul bersamaan pada penderita tuberculosis dengan nilai *minimum support* dan *minimum confidence* yang telah didefinisikan. Adapun hasil dari penelitian ini adalah rule pertama pada apriori non modifikasi untuk nilai minimum support 50% menghasilkan nilai *confidence* sebesar 100% dan nilai *confidence* pada apriori yang dimodifikasi sebesar 39,6%. Sedangkan rule kedua pada apriori non modifikasi untuk nilai minimum support yang sama menghasilkan nilai *confidence* sebesar 75% dan nilai *confidence* pada apriori yang dimodifikasi sebesar 90%.

## REFERENSI

- ADM PLK. (2021). *Waspadai Tbc Di Kala Pandemi - Pusat Layanan Kesehatan*.  
<http://plk.unair.ac.id/waspadai-tbc-di-kala-pandemi/>
- Aini, N., Ramadiani, R., & Hatta, H. R. (2017). Sistem Pakar Pendiagnosa Penyakit Tuberkulosis. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 12(1), 56.  
<https://doi.org/10.30872/jim.v12i1.224>
- Fauzy, M., Saleh W, K. R., & Asror, I. (2016). Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung. *Jurnal Ilmiah Teknologi Infomasi Terapan*, 2(3).  
<https://doi.org/10.33197/jitter.vol2.iss3.2016.111>
- Ginting, D. S., Ilmu, F., Dan, K., Informasi, T., & Utara, U. S. (2018). *Modifikasi algoritma apriori dengan substitusi metode bayesian pada nilai confidence terhadap aturan asosiasi*.
- Kumar, P., Dangwal, D., & Puri, N. (2012). Diagnosis of Tuberculosis using Association Rule Method. *Journal of Information and Operation Management*, 3(1), 133–135.
- Kurniasari, A., & Kurniasari, A. (2013). *Sistem Pakar Untuk Mendiagnosa Awal Penyakit Tuberkulosis ( Tb ) Menggunakan Metode Dempster- Oleh : Sistem Pakar Untuk Mendiagnosa Awal Penyakit Tuberkulosis ( Tb ) Menggunakan Metode Dempster-Shaffer*.
- Purba, C. V., & Buulolo, E. (2020). Implementasi Algoritma Apriori Untuk Menentukan Pola Data Penyakit Pada Anak Usia Dini (Studi Kasus: RS. Estomihi). *JURIKOM (Jurnal Riset Komputer)*, 7(2), 308. <https://doi.org/10.30865/jurikom.v7i2.2113>
- Putra, Y. (2018). Assosiaton rule(Algortima Apriori). *Assosiation Algoritma Apriori*, 1–9.
- Santosa, I., Rosiyah, H., & Rahmanita, E. (2018). Implementasi Algoritma Decision Tree C .45 Untuk Diagnosa Penyakit Tubercolusis ( Tb ). *Jurnal Ilmiah NERO*, 3(3), 169–176.
- Shah, A. (2017). Association rule mining with modified apriori algorithm using top down approach. *Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, ICATccT 2016*, 747–752.  
<https://doi.org/10.1109/ICATCCT.2016.7912099>

- Studi Sistem Informasi, P., & Triguna Dharma, S. (2017). Sistem Pakar Untuk Mendiagnosa Penyakit Anemia Dengan Menggunakan Metode Teorema Bayes \* Trinanda Syahputra #1 , Muhammad Dahria #2 , Prilla Desila Putri #3. *Saintikom*, 16(3), 284–294.
- Yu, W., Wang, X., Wang, F., Wang, E., & Chen, B. (2008). The research of improved Apriori algorithm for mining association rules. *International Conference on Communication Technology Proceedings, ICCT*, 513–516. <https://doi.org/10.1109/ICCT.2008.4716098>