

PERBANDINGAN KINERJA ALGORITMA *DEMPSTER SHAFER* DAN *FUZZY-NAIVE BAYES* DALAM KLASIFIKASI PENYAKIT DEMAM BERDARAH DAN TIFUS

¹⁾Angelina Sagita Sastrawan, ²⁾I Gede Aris Gunadi, ³⁾I N. Sukajaya

^{1,2,3}Program Studi Ilmu Komputer, Program Pascasarjana
Universitas Pendidikan Ganesha
Singaraja, Indonesia

e-mail: angelina.sagita@gmail.com, igedearisgunadi@undiksha.ac.id,

nyoman.sukajaya@undiksha.ac.id

ABSTRAK

Menurut data tahun 2017 dan 2018, penyakit Demam Berdarah dan Tifus selalu masuk sebagai 10 (sepuluh) besar penyakit dengan jumlah rawat inap terbanyak di Rumah Sakit Umum Daerah Kabupaten Buleleng. Ada baiknya ketika gejala sudah mulai muncul, langsung diberikan penanganan yang tepat sehingga penyakit dapat diobati sejak dini dan jumlah pasien rawat inap menjadi berkurang. Hal ini menjadi dasar perlunya pendeteksian penyakit dilihat dari gejala awal yang dialami. Dalam penelitian ini deteksi penyakit dibatasi pada Demam Berdarah dan Tifus. Metode yang digunakan adalah *Dempster Shafer* atau sistem pakar dengan representasi pengetahuan dari pakar yaitu seorang Dokter Spesialis yang bekerja di Rumah Sakit Umum Daerah Kabupaten Buleleng. Metode ini sekaligus akan dibandingkan dengan metode *Machine Learning* yaitu kombinasi dari *fuzzy logic* dan *naive bayes* yang sesuai dengan karakteristik data gejala yang ada. Model klasifikasi menggunakan metode *k-fold cross validation*, sedangkan pengujian akurasi menggunakan *confusion matriks*. Untuk menguji akurasi hasil klasifikasi digunakan matriks konfusi. Dengan $k = 4$ pada metode *k-fold cross validation*, dan perhitungan akurasi, presisi, dan recall pada *confusion matriks* didapat bahwa nilai akurasi menggunakan metode *Dempster Shafer* lebih besar daripada metode *Fuzzy Naive Bayes*.

Kata kunci: deteksi penyakit, metode dempster shafer, metode fuzzy-naive bayes

ABSTRACT

According to data from 2017 and 2018, Dengue Fever and Typhus Fever are always included as 10 (ten) major diseases with the highest number of hospitalizations at the Buleleng Regency General Hospital. It's good when the symptoms have begun to appear, immediately given appropriate treatment so that the disease can be treated early and the number of inpatients becomes reduced. This is the basis for the need for disease detection seen from the initial symptoms experienced. In this study disease detection was limited to Dengue Fever and Typhus. The method used is *Dempster Shafer* or expert system with knowledge representation from experts, namely a Specialist Doctor who works at the Buleleng Regency General Hospital. This method will also be compared with the *Machine Learning* method, which is a combination of *fuzzy logic* and *naïve bayes* that is in accordance with the characteristics of existing symptom data. The classification model uses the *k-fold cross validation* method, while the accuracy test uses *confusion matrix*. To test the accuracy of the classification results, a *confusion matrix* is used. With $k = 4$ in the *k-fold cross validation*, and the calculation of accuracy, precision, and recall in the *confusion matrix*, it is found that the accuracy value using the *Dempster Shafer* method is greater than the *Fuzzy Naive Bayes* method.

Keywords : detection disease, dempster-shafer method, fuzzy-naive bayes method

I. PENDAHULUAN

Kesehatan yang baik merupakan hal utama yang diinginkan oleh setiap orang, karena dengan tubuh yang sehat, segala aktifitas dapat dilakukan dengan baik. Namun demikian, tetap ada kemungkinan seseorang terserang penyakit yang diakibatkan oleh bakteri, virus, dan cara penyebaran lainnya, bahkan sampai harus memeriksakan diri ke fasilitas kesehatan terdekat. Sesuai dengan data dari Buku Profil Rumah Sakit Umum Tahun 2017, jumlah kunjungan pasien pada Pelayanan Gawat Darurat adalah 27.194 orang atau rata-rata berjumlah 74 orang perhari, dan jumlah kunjungan pasien pada Poliklinik adalah 13.859 orang atau jika dirata-ratakan berjumlah 380 orang perhari. Akan sangat bermanfaat jika di awal terserang penyakit, telah dapat diketahui kemungkinan besar penyakit yang diderita, sehingga penanganan selanjutnya lebih mudah dilaksanakan dan diobati. Terutama untuk beberapa penyakit yang memiliki kemiripan gejala namun penanganan kedua penyakit ini jauh berbeda, beberapa diantaranya adalah Demam Berdarah (DB) dan Tifus. Hal ini menjadi dasar bahwa pendeteksian harus lebih baik dilaksanakan agar mendapatkan penanganan awal yang tepat dan dapat mengurangi jumlah pasien yang rawat inap karena sudah dapat ditangani sejak dini

Salah satu metode yang dapat diterapkan dalam proses deteksi awal penyakit DB dan Tifus adalah metode Dempster Shafer. Metode Dempster Shafer dapat memberikan tingkat keyakinan tentang jenis penyakit sesuai dengan gejala yang dirasakan oleh user [7]. Metode *Dempster Shafer* memanfaatkan nilai pakar dalam perhitungannya sehingga sangat bergantung dari pakar yang memberi nilai keyakinan. Dalam beberapa kasus, pakar tidak dapat memberikan nilai keyakinan sehingga perhitungan pada metode Dempster Shafer tidak dapat dilaksanakan [6]. Oleh karena itu, proses pendeteksian juga diterapkan dengan memanfaatkan analisa dari pola data yang telah ada sebelumnya dengan metode Naive Bayes yang dikombinasikan dengan fuzzy logic.

Dari latar belakang diatas, dalam penelitian ini akan dilakukan perbandingan sistem mana yang lebih akurat dalam membantu mendeteksi penyakit Demam Berdarah dan Tifus, antara metode sistem pakar menggunakan metode *Dempster Shafer* dan metode machine learning dengan *Naive Bayes* yang dikombinasikan dengan *fuzzy logic*. Penelitian ini juga bertujuan untuk dapat digunakan oleh masyarakat awam atau paramedis dalam mendeteksi awal penyakit demam berdarah dan tifus, dan mengetahui nilai akurasi terbaik dari penerapan kedua metode diatas untuk mendeteksi diagnosis penyakit demam berdarah dan tifus.

II. TINJAUAN PUSTAKA

A. Demam Berdarah dan Tifus

Demam Berdarah Dengue (DBD) adalah penyakit infeksi yang disebabkan oleh virus Dengue, sedangkan demam Tifoid (Tifus) adalah demam yang disebabkan oleh kuman *Salmonella typhi* atau *Salmonella para typhi* dengan metode penularan ke manusia melalui makanan dan atau minuman yang tercemar dengan feses manusia. Demam tifoid banyak ditemukan di masyarakat perkotaan maupun di perdesaan dan erat kaitannya dengan kualitas hygiene pribadi dan sanitasi lingkungan yang kurang baik [4]. Gejala pasien yang terkena penyakit tifus adalah demam, sakit kepala (pusing), nyeri otot, pegal-pegal, insomnia, anoreksia, mual/muntah, diare, nyeri abdomen, BAB berdarah, dan lidah putih, sedangkan gejala pasien dengan penyakit DBD adalah demam, nyeri kepala, nyeri retroorbital, mialgia/atralgia, ruam, gusi berdarah, mimisan, nyeri perut, mual/muntah, hematemesis, dan dapat juga melena [5].

B. Metode Dempster Shafer

Metode Dempster Shafer adalah teori matematika untuk pembuktian berdasarkan *belief functions* (fungsi kepercayaan) yang mengkombinasikan satu kemungkinan dengan fakta yang ada [3]. Teori *Dempster Shafer* ditulis dalam suatu interval [Belief,Plausibility]. Belief (Bel) adalah ukuran kekuatan evidence dalam mendukung kepercayaan, sedangkan plausibility (Pls) adalah ukuran

ketidakpercayaan terhadap evidence/ gejala. Jika yakin akan X, maka dapat dikatakan bahwa $Bel(X)=1$ dan $Pls(X)=0$.

Dalam teori *Dempster Shafer* dikenal juga dengan adanya frame of discernment (FOD) yang dinotasikan dengan θ dan mass function yang dinotasikan dengan (m) seperti pada persamaan (1) berikut :

$$m3(Z) = \frac{\sum_{X \cap Y = Z} m1(X)m2(Y)}{1 - \sum_{X \cap Y = \emptyset} m1(X)m2(Y)} \quad (1)$$

Dimana :

$m3(Z)$ = mass function dari evidence (Z)

$m1(X)$ = mass function dari evidence (X), yang diperoleh dari nilai keyakinan suatu evidence dikalikan dengan nilai disbelief dari evidence.

$m2(Y)$ = mass function dari evidence (Y), yang diperoleh dari nilai keyakinan suatu evidence dikalikan dengan nilai disbelief dari evidence.

$\sum X \cap Y = Z$ $m1(X). m2(Y)$ = merupakan nilai kekuatan dari evidence Z yang diperoleh dari kombinasi sekumpulan evidence X dan Y yang memiliki irisan.

$\sum X \cap Y = \emptyset$ $m1(X). m2(Y)$ = merupakan nilai kekuatan dari kombinasi sekumpulan X dan Y evidence yang tidak memiliki irisan

C. Metode Fuzzy Naive Bayes

Metode *Fuzzy-Naive Bayes* mengkombinasikan Logika *Fuzzy* dan *Naive Bayes* dalam perhitungan di dalamnya. Logika *fuzzy* bekerja dengan menggunakan derajat keanggotaan dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang telah ditentukan [8]. Metode *Naive Bayes* merupakan pengklasifikasian dengan memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya, dengan perhitungan sesuai persamaan 2 berikut ini :

$$P(X|Ci) = P(x1|Ci) \dots x P(xn|Ci) x P(Ci) \quad (2)$$

Dimana :

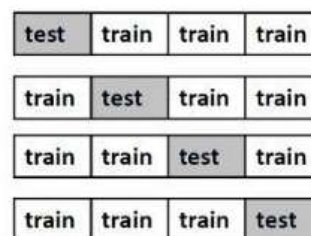
X = gejala yang muncul

$x1, x2, \dots, xn$ = gejala ke-1 s/d ke-n

Ci = class ke-i

D. Metode Pengukuran Klasifikasi

Dalam penelitian ini, pengukuran klasifikasi menggunakan metode *k-fold cross validation* dan *confusion matriks*. Metode *cross validation* merupakan metode yang digunakan untuk mengevaluasi kinerja model maupun algoritma [1]. Dataset akan dibagi menjadi data training dan data testing, dimana proses training dan testing dilakukan sebanyak k kali secara berulang-ulang. Pemodelan secara berulang diyakini lebih kuat terhadap spesifikasi salah model yang diakibatkan oleh suatu kebetulan pola data. Dengan menggunakan $k=4$, metode *4-fold cross validation* akan membagi dataset secara acak ke dalam 4 subset yaitu D1, D2, D3, dan D4 yang diilustrasikan seperti Gambar 1 berikut ini :



Gambar 1. Ilustrasi *4-fold cross validation*

Sebagai evaluasi akurasi model klasifikasi untuk memperkirakan objek yang benar atau salah, maka digunakan metode *confusion matriks* yaitu sebuah matriks dari prediksi yang akan dibandingkan dengan kelas sebenarnya. Ilustrasi *confusion matrik* dapat digambarkan seperti Tabel 1 di bawah ini :

Tabel 1. Ilustrasi *Confusion Matriks*

Kelas	Terklasifikasi Positif	Terklasifikasi Negatif
Positif	TP (<i>True Positif</i>)	FN (<i>False Negatif</i>)
Negatif	FP (<i>False Positif</i>)	TN (<i>True Negatif</i>)

Berdasarkan nilai True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP) diperoleh nilai akurasi, *presisi* dan *recall*. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar, yaitu perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai *presisi* menggambarkan jumlah data kategori positif yang diklasifikasikan secara benar dibagi dengan total data yang diklasifikasi positif. *Recall*

menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar oleh sistem.

Ketiga penilaian kinerja klasifikasi ini didapat dengan rumus pada persamaan sebagai berikut :

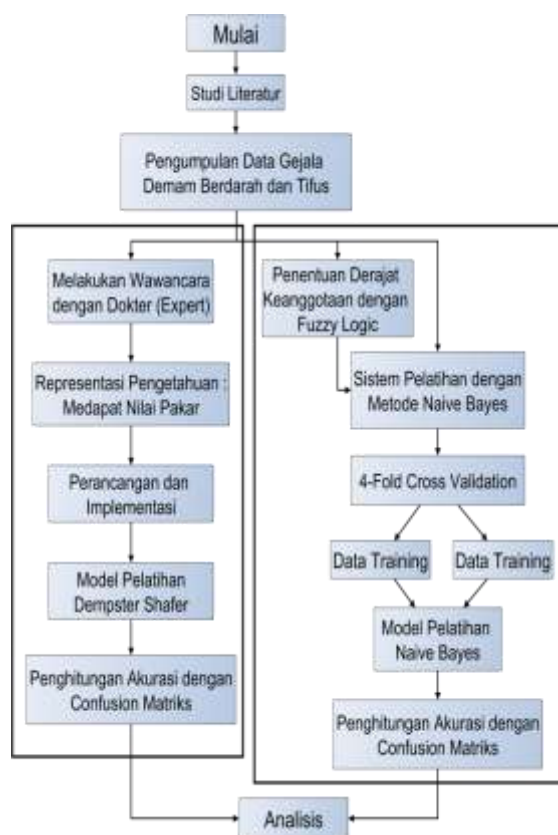
$$\text{Akurasi} = \frac{TP+TN}{N} \times 100\% \quad (3)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (5)$$

III. METODE PENELITIAN

Penelitian ini dilakukan dengan alur sebagai berikut :



Gambar 2. Diagram Alir Penelitian

3.1 Pengumpulan Data

Metode pengumpulan data pada penelitian ini adalah dengan melaksanakan pencatatan data pasien Demam Berdarah dan Tifus dengan jumlah 100 data dari aplikasi Sistem Informasi Manajemen Rumah Sakit (SIM RS). SIM RS adalah sistem komputerisasi yang dimiliki oleh RSUD Kabupaten Buleleng untuk memproses dan mengintegrasikan seluruh alur proses bisnis layanan kesehatan dalam bentuk jaringan koordinasi, pelaporan dan prosedur administrasi untuk memperoleh informasi

secara cepat, tepat dan akurat. Dari aplikasi SIM RS didapat rekapitulasi data pasien dengan jenis penyakit yang diderita dari rentang waktu Agustus 2018 sampai dengan April 2019. Dari rekapitulasi ini, kombinasi nomor rekam medis dan jenis penyakit khusus Demam Berdarah dan Tifus digunakan untuk mempermudah pencarian dokumen rekam medis lengkap dengan data gejala yang dialami oleh masing-masing pasien. Jumlah komposisi data adalah masing-masing 50 data pasien dengan gejala Demam Berdarah dan Tifus.

Adapun data gejala yang muncul pada ditunjukkan pada Tabel 2 di bawah ini :

Tabel 2. Data Gejala

Ko de	Nama Gejala	Kode	Nama Gejala
E1	Lama Panas	E11	Lidah Putih
E2	Jenis Panas	E12	Nyeri Perut
E3	Mual	E13	Nyeri Sendi
E4	Muntah	E14	Nyeri Kepala
E5	Mimisan	E15	Nyeri Ulu Hati
E6	Gusi Berdarah	E16	Batuk
E7	Lemas	E17	Pilek
E8	Diare	E18	Sesak
E9	Sembelit	E19	Pusing
E10	Bintik Merah	E20	Nyeri Seluruh Badan

3.2 Metode Dempster Shafer dan Fuzzy-Naive Bayes

Dari 100 data gejala pasien yang telah dikumpulkan, akan diolah dengan dua metode yaitu metode Dempster Shafer dan metode Fuzzy-Naive Bayes. Proses awal pada metode Dempster Shafer adalah wawancara dengan Dokter Spesialis yang bekerja di RSUD Kabupaten Buleleng sebagai pakar. Setelah terjadi proses akuisisi pengetahuan mengenai gejala yang digunakan dalam perhitungan Dempster Shafer, pakar memilih 12 (dua belas) gejala penyakit yang menunjukkan adanya penyakit Demam Berdarah ataupun Tifus seperti ditunjukkan pada Tabel 3 berikut ini :

Tabel 3. Nilai Keyakinan Pakar

Ko de	Gejala	DB	Tif us	Nilai Keya kinan
G1	Lama Panas			
	1-5 hari	Ya	-	0,8
	>6 hari	-	Ya	0,8

G2	Jenis Panas			
	Naik Turun	-	Ya	0,8
	Tinggi	Ya	-	0,8
G3	Mimisan	Ya	-	0,6
G4	Gusi Berdarah	Ya	-	0,6
G5	Bintik Merah	Ya	Ya	0,45
G6	Nyeri Sendi	Ya	Ya	0,45
G7	Sakit Kepala	Ya	Ya	0,3
G8	Mual	Ya	Ya	0,5
G9	Muntah	Ya	Ya	0,5
G10	Diare	-	Ya	0,6
G11	Sembelit	-	Ya	0,6
G12	Lidah Putih	-	Ya	0,75

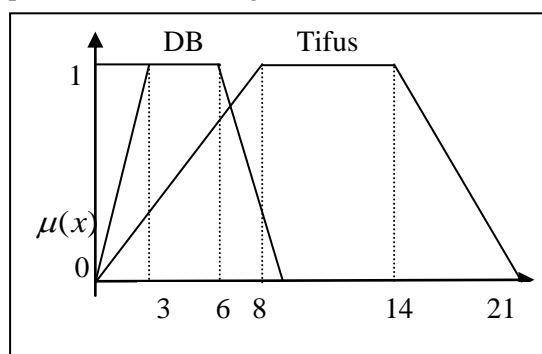
Nilai keyakinan pada tabel diatas digunakan sebagai dasar perhitungan dengan metode Dempster Shafer, sehingga didapat nilai kemungkinan pasien terkena penyakit Demam Berdarah atau Tifus dari gejala awal yang dialami.

Berbeda dengan metode *Dempster Shafer*, metode *Fuzzy-Naive Bayes* memanfaatkan data sebelumnya untuk digunakan sebagai data training dalam klasifikasi sebuah data.

Perhitungan fuzzy akan mencari nilai keanggotaan pada Gejala Lama Panas dengan metode Fuzzy Logic. Menurut Kriteria WHO, salah satu gejala suatu penyakit terdiagnosis DB adalah adanya gejala demam atau riwayat demam akut antara 2-7 hari. Demam akut atau demam tinggi dapat diartikan sebagai demam yang memiliki batasan nilai atau derajat suhu tubuh diatas 39,5°C [2].

Gejala dari penyakit Tifus antara lain demam diawal kebanyakan samar-samar. Pada minggu ke-2 intensitas demam makin tinggi, kadang terus-menerus (kontinu), bila pasien membaik maka pada minggu ke 3 suhu badan berangsur turun dan dapat normal kembali pada akhir minggu ke 3.

Data gejala diatas dapat direpresentasikan melalui bentuk kurva dimana adalah derajat keanggotaan yang ditampilkan pada Gambar 3 sebagai berikut :



Gambar 3. Kurva Gejala Lama Panas

Sesuai gambar kurva diatas, fungsi keanggotaan DB akan menjadi :

$$\mu(x) = \begin{cases} 0; & x \geq 9 \\ (x-0)/(3-0); & 0 \leq x \leq 3 \\ 1; & 3 \leq x \leq 6 \\ (9-x)/(9-6); & 6 \leq x \leq 9 \end{cases}$$

Fungsi keanggotaan Tifus setelah disesuaikan dengan gambar kurva diatas akan menjadi :

$$\mu(x) = \begin{cases} 0; & x \geq 21 \\ (x-0)/(8-0); & 0 \leq x \leq 8 \\ 1; & 8 \leq x \leq 14 \\ (21-x)/(21-14); & 14 \leq x \leq 21 \end{cases}$$

Nilai gejala lama panas yang telah diproses menggunakan metode *fuzzy* akan dikombinasikan bersama gejala lain yang muncul menggunakan metode *Naive Bayes* dengan menghitung probabilitas kemunculan tiap gejala sesuai dengan Tabel 3 pada data training untuk setiap kelas.

3.3. Metode Pengujian

Model yang didapatkan dari penerapan metode *Dempster Shafer* dan *Fuzzy-Naive Bayes* ini akan diuji menggunakan *k-fold cross validation*. Dengan menentukan k=4 maka langkah-langkah pengujian data dengan *4-fold cross validation* dan *confusion matrik* pada dataset berjumlah 100 data yang digunakan akan dibagi rata menjadi 4 subset data, yaitu D1 adalah data ke-1 s/d 25, D2 adalah data ke-26 s/d 50, D3 adalah data ke-51 s/d 75, dan D4 adalah data ke-76 s/d 100. Dalam proses penghitungannya, salah satu subset data dipilih untuk digunakan sebagai data testing dan subset data lainnya sebagai data training. Pembagian masing-masing subset diatur seperti tabel dibawah ini :

Tabel 4. Pembagian Dataset

No	Nilai k	Data Training	Data Test
1.	k=1	D2, D3, D4	D1
2.	k=2	D1, D3, D4	D2
3.	k=3	D1, D2, D4	D3
4.	k=4	D1, D2, D3	D4

Tingkat akurasi dihitung pada setiap iterasi (iterasi-1, iterasi-2, iterasi-3, iterasi-4), kemudian dihitung rata-rata tingkat akurasi dari seluruh iterasi untuk mendapatkan tingkat akurasi data keseluruhan. Setelah data terkumpul dan dilakukan iterasi dengan metode *4-fold cross validation*, maka dilanjutkan dengan pengukuran klasifikasi menggunakan *confusion matrik*. Hasil keluaran klasifikasi pada penelitian ini termasuk *multi class* yaitu Demam Berdarah (DB) sebagai *class* pertama dan Tifus sebagai *class* kedua. Analisa data dilakukan dengan membandingkan nilai akurasi, *presisi*, dan *recall* dari penerapan metode *Dempster Shafer* dan metode *Fuzzy-Naive Bayes* untuk mendeteksi penyakit demam berdarah dan tifus.

IV. HASIL DAN PEMBAHASAN

C. Hasil Klasifikasi dengan Metode *Dempster Shafer*

Hasil klasifikasi iterasi ke-1 yaitu data testing D1 atau data ke-1 sampai dengan data ke-25 menggunakan nilai pakar adalah sebagai berikut :

Tabel 5. Klasifikasi D1 *Dempster Shafer*

ID Pasien	Prediksi Sistem	ID Pasien	Prediksi Sistem
ID001	DB	ID014	DB
ID002	DB	ID015	Tifus
ID003	DB	ID016	Tifus
ID004	DB	ID017	DB/Tifus
ID005	DB	ID018	DB
ID006	Tifus	ID019	Tifus
ID007	Tifus	ID020	Tifus
ID008	Tifus	ID021	Tifus
ID009	Tifus	ID022	Tifus
ID010	Tifus	ID023	Tifus
ID011	Tifus	ID024	DB
ID012	Tifus	ID025	DB
ID013	Tifus		

Berdasarkan hasil pada Tabel 5, dapat dilihat bahwa sistem dapat memprediksi 24 data dari 25 data yang diuji. Hasil prediksi didapat dengan memilih penyakit yang memiliki nilai keyakinan yang lebih tinggi antara DB dan tifus. Data lainnya berjumlah 1 (satu) data belum dapat diprediksi secara pasti karena setelah gejala yang dialami oleh pasien diproses menggunakan perhitungan dengan metode *Dempster Shafer*, nilai keyakinan kedua penyakit ini bernilai sama.

Hasil klasifikasi menggunakan metode *Dempster Shafer* akan dibandingkan dengan

nilai aktual dari data penyakit pasien untuk mendapatkan nilai TP, FP, FN, dan TN sesuai dengan Tabel 1. Ilustrasi *Confusion Matriks*. Berdasarkan hasil prediksi sistem pada iterasi ke-1, maka didapat nilai *confusion matriks* seperti pada Tabel 6 berikut ini :

Tabel 6. *Confusion Matriks* DS Iterasi ke-1

	Terklasifikasi Positif	Terklasifikasi Negatif
Kelas Positif	TP = 9	FN = 2
Kelas Negatif	FP = 0	TN = 13

Dengan menggunakan perhitungan yang sama menggunakan metode *Dempster Shafer* (DS), didapat nilai *confusion matriks* hasil prediksi pada iterasi k=2, k=3, dan k=4 ditunjukkan seperti pada Tabel 7 berikut ini :

Tabel 7. *Confusion Matriks* Metode DS

	k=2	k=3	k=4
Nilai TP	12	7	10
Nilai FP	2	2	2
Nilai FN	0	3	1
Nilai TN	8	10	8

Dari nilai *confusion matriks* pada setiap iterasi di tabel 7, dapat dilihat bahwa ketepatan hasil prediksi yang ditunjukkan oleh nilai TP dan TN paling tinggi ada pada iterasi k=2 sejumlah 20 data, sedangkan pada iterasi k=3 dan k=4 masing-masing sejumlah 17 dan 18 data.

D. Hasil Klasifikasi dengan Metode *Fuzzy-Naive Bayes*

Pengolahan data dengan metode *fuzzy-naive bayes* menggunakan data gejala yang sama dengan metode *Dempster Shafer*. Nilai akurasi pada hasil klasifikasi kedua metode ini akan dibandingkan sehingga didapatkan metode yang terbaik dalam mendeteksi penyakit demam berdarah dan tifus. Pada iterasi ke-1 yaitu pada k = 1, data training yang digunakan adalah data ke-26 sampai dengan data ke-100, sedangkan data testing yang digunakan adalah data ke-1 sampai dengan data ke-25 atau disebut data D1. Hasil klasifikasi dapat ditunjukkan seperti pada Tabel 8 berikut ini :

Tabel 8. Klasifikasi D1 *Fuzzy-Naive Bayes*

ID Pasien	Prediksi Sistem	ID Pasien	Prediksi Sistem
ID001	DB	ID014	DB
ID002	DB	ID015	Tifus
ID003	DB	ID016	DB

ID004	Tifus	ID017	DB
ID005	DB	ID018	Tifus
ID006	Tifus	ID019	Tifus
ID007	DB	ID020	Tifus
ID008	Tifus	ID021	DB/Tifus
ID009	Tifus	ID022	Tifus
ID010	DB	ID023	Tifus
ID011	Tifus	ID024	DB
ID012	DB/Tifus	ID025	DB
ID013	Tifus		

Klasifikasi menggunakan metode *fuzzy-naive bayes* dengan menggunakan data training sejumlah 75 data dan data testing sejumlah 25 data seperti pada tabel 8 di atas, dapat memprediksi 23 dari 25 data yang diuji. Data lainnya sejumlah 2 data belum dapat diprediksi karena kedua data ini memiliki nilai perhitungan yang sama ketika diproses menggunakan metode *fuzzy-naive bayes*.

Hasil klasifikasi pada Tabel 8 akan dibandingkan dengan data penyakit aktualnya untuk mendapatkan nilai TP, FP, FN, dan TN. Nilai *confusion matriks* sesuai nilai tersebut dapat ditunjukkan pada Tabel 9 di bawah ini :

Tabel 9. *Confusion Matriks* FNB Iterasi ke-1

	Terklasifikasi Positif	Terklasifikasi Negatif
Kelas Positif	TP = 9	FN = 3
Kelas Negatif	FP = 1	TN = 10

Nilai *confusion matriks* hasil prediksi pada iterasi k=2, k=3, dan k=4 ditunjukkan seperti pada Tabel 10 berikut ini

Tabel 11. Perbandingan nilai akurasi, presisi dan recall

N o.	Evaluasi	Data 1-25 (D1)		Data 26-50 (D2)		Data 51-75 (D3)		Data 76-100 (D4)	
		DS	FNB	DS	FNB	DS	FNB	DS	FNB
1.	Akurasi	88%	76%	80%	72%	68%	72%	72%	52%
2.	Presisi	100%	90%	85,71%	88,89%	77,78%	100%	83,33%	57,14%
3.	Recall	81,82%	75%	100%	61,54%	70%	41,67%	90,91%	40%

Pada perhitungan dengan menggunakan data testing D1, dapat dilihat bahwa metode *Dempster Shafer* memiliki nilai akurasi, presisi, dan recall yang lebih besar dibandingkan menggunakan metode *Fuzzy-Naive Bayes*. Hal ini dipengaruhi oleh nilai ketepatan pendeteksian yang ditunjukkan oleh nilai *True Positif* (TP) dan nilai *True Negatif* (TN) pada *confusion matriks*. Dari jumlah data testing sebanyak 25 data, nilai TP dan TN pada metode *Dempster Shafer* yang

Tabel 10. *Confusion Matriks* Metode *Fuzzy-Naive Bayes*

	k=2	k=3	k=4
Nilai TP	8	5	4
Nilai FP	1	0	3
Nilai FN	5	7	6
Nilai TN	10	13	9

Dari rangkuman nilai *confusion matriks* pada setiap iterasi di tabel 9 dan tabel 10, dapat dilihat bahwa ketepatan hasil prediksi yang ditunjukkan oleh nilai TP dan TN paling tinggi ada pada iterasi k=1 sejumlah 19 data, sedangkan pada iterasi k=2, k=3, dan k=4 masing-masing sejumlah 18, 18, dan 13 data. Nilai dari setiap iterasi akan dicari nilai rata-ratanya sehingga didapatkan nilai akurasi, *presisi* dan *recall* dari metode *Fuzzy-Naive Bayes*.

E. Perhitungan Akurasi

Sesuai dengan tabel 6 dan 7 pada metode *Dempster Shafer* dan tabel 9 dan 10 pada metode *Fuzzy Naive Bayes*, didapat nilai TP, FP, FN, dan FN pada data D1, D2, D3, dan D4. Dari nilai tersebut, sesuai dengan persamaan (2), (3), dan (4) akan dicari nilai akurasi, *presisi* dan *recall* pada setiap dataset untuk menentukan metode mana yang lebih tepat dalam mengklasifikasi pasien dengan penyakit demam berdarah atau tifus. Data perbandingan nilai akurasi, *presisi* dan *recall* berdasarkan nilai *confusion matriks* pada kedua metode tersebut dirangkum dalam bentuk tabel dan ditunjukkan seperti pada Tabel 11 berikut ini :

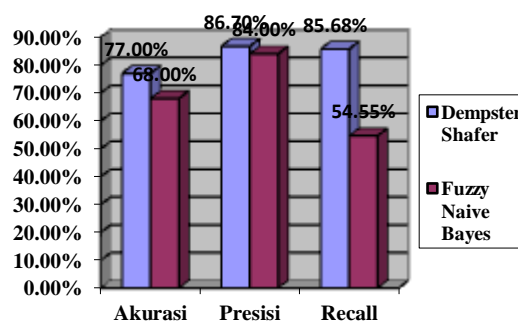
didapatkan adalah sejumlah 9 dan 13 data, sedangkan nilai TP dan TN pada metode *Fuzzy-Naive Bayes* adalah 9 dan 10 data, yang apabila dijumlahkan bernilai 22 dan 19 data. Dari proses perhitungan metode *Fuzzy-Naive Bayes* juga diketahui bahwa ada beberapa gejala yang tidak muncul pada data training di data ke-26 sampai dengan data ke-100, sehingga probabilitas gejala menjadi bernilai 0 saat proses klasifikasi data testing di data ke-1 sampai dengan data ke-25.

Pada data D2 dengan komposisi jumlah pasien DB dan Tifus sebanyak 37 dan 38 data, metode *Dempster Shafer* memiliki nilai presisi yang lebih rendah dibandingkan dengan metode *Fuzzy-Naive Bayes* karena perbandingan jumlah data positif yang terklasifikasi benar terhadap jumlah data positif pada metode *Dempster Shafer* hanya 85,71% sedangkan pada metode FNB bernilai 88,89%. Jika diperhatikan pada tabel 7 dan tabel 10, sistem dapat memprediksi nilai positif untuk metode *Dempster Shafer* adalah sebanyak 14 data dengan termasuk di dalamnya 2 data salah klasifikasi, dibandingkan metode *Fuzzy-Naive Bayes* yaitu 9 data dengan 1 data salah terklasifikasi. Hal ini yang mengakibatkan nilai presisi atau pengukuran tingkat kebenaran sistem pada *Fuzzy-Naive Bayes* lebih besar dibandingkan metode *Dempster Shafer*. Untuk nilai akurasi dan *recall*, metode *Dempster Shafer* memiliki nilai yang lebih besar dibandingkan metode *Fuzzy-Naive Bayes*.

Untuk data testing D3, ketepatan prediksi pada metode *Fuzzy-Naive Bayes* lebih banyak dengan jumlah TP dan TN adalah 5 dan 13 dibandingkan metode *Dempster Shafer* dengan TP dan TN berjumlah 7 dan 10. Hal ini mengakibatkan metode *Fuzzy-Naive Bayes* memiliki nilai akurasi dan *presisi* yang lebih unggul dibandingkan metode *Dempster Shafer* yaitu dengan nilai 72% dan 100%. Untuk nilai *recall*, metode *Dempster Shafer* memiliki nilai lebih besar yaitu 70% dibandingkan metode *Fuzzy-Naive Bayes* dengan nilai 41,67%. Ini menunjukkan bahwa ketepatan metode *Dempster Shafer* dalam memprediksi data positif lebih besar dibandingkan metode *Fuzzy-Naive Bayes*.

Pada data D4 dengan komposisi jumlah pasien DB dan Tifus sebanyak 37 dan 38 data, metode *Dempster Shafer* memiliki nilai akurasi sebesar 72% yang lebih unggul dibandingkan metode *Fuzzy-Naive Bayes* dengan nilai 52%. Begitu pula untuk presisi dan *recall*, metode *Dempster Shafer* memiliki nilai yang lebih tinggi dibandingkan metode *Fuzzy-Naive Bayes*.

Apabila nilai rata-rata akurasi, presisi dan *recall* dari data keseluruhan digambarkan dalam bentuk grafik, maka dapat ditunjukkan pada Gambar di bawah ini :



Gambar 2. Grafik Data Akurasi

Nilai akurasi mengukur nilai ketepatan seluruh hasil prediksi pada sistem dengan nilai diagnosis sebenarnya pada keseluruhan data. Dengan nilai akurasi sebesar 77%, menunjukkan bahwa sistem prediksi dengan metode *Dempster Shafer* dapat mengklasifikasikan data gejala awal pasien dengan tepat sebesar 77 data dari 100 data yang ada. Sedangkan sistem prediksi yang menerapkan metode *Fuzzy-Naive Bayes* memiliki akurasi lebih rendah dibandingkan metode *Dempster Shafer* yaitu sebesar 68%.

Nilai *presisi* menggambarkan kehandalan sistem dalam memberikan hasil prediksi yang benar. Pada metode *Dempster Shafer*, jumlah data yang diprediksi benar dibandingkan dengan jumlah data yang diprediksi oleh sistem memiliki nilai 86,70%. Hal ini menunjukkan bahwa dari 100 hasil prediksi sistem dengan metode *Dempster Shafer*, sebesar 86,70% prediksi benar terdiagnosis penyakit Demam Berdarah atau Tifus. Untuk metode *Fuzzy-Naive Bayes*, memiliki perbedaan nilai *presisi* yang tidak terlalu signifikan yaitu 84%.

Nilai *recall*, menunjukkan seberapa relevan hasil prediksi yang diberikan oleh sistem dibandingkan dengan data positif secara aktual. Perbedaan nilai *recall* pada metode *Dempster Shafer* dan *Fuzzy-Naive Bayes* adalah masing-masing bernilai 85,68% dan 54,55%. Perbedaan nilai yang cukup besar ini menandakan bahwa untuk memprediksi data positif, sistem dengan metode *Dempster Shafer* lebih handal dibandingkan dengan sistem dengan metode *Fuzzy-Naive Bayes*.

V. PENUTUP

Dari hasil perhitungan dan perbandingan nilai akurasi dari metode *Dempster Shafer* dan

Fuzzy Naive Bayes, didapat kesimpulan dalam penelitian ini adalah sebagai berikut :

1. Akurasi dari metode *Dempster Shafer* sangat tergantung dari nilai pakar yang menentukan bobot dari setiap gejala yang dialami oleh pasien. Proses akuisisi pengetahuan pakar terhadap data gejala pada rekam medik yang ditunjukkan menjadi tabel nilai pakar menjadi faktor penentu dalam perhitungan menggunakan metode ini.
2. Metode alternatif lain dengan memanfaatkan pola data gejala pada pasien yaitu metode *Fuzzy-Naive Bayes* dipengaruhi oleh komposisi data training yang digunakan sebagai model pembelajaran dalam klasifikasi pada data testing. Penggunaan metode pengujian dengan *k-fold cross validation* tepat digunakan untuk menghindari faktor kebetulan dari pengambilan data training yang acak, dan nilai akurasi metode ini menjadi lebih akurat.
3. Dari perbandingan nilai akurasi pada kedua metode diatas, metode *Dempster Shafer* lebih tepat digunakan dalam penelitian ini karena kemampuan dalam mendeteksi apakah gejala tersebut dapat didiagnosa sebagai Demam Berdarah atau Tifus lebih akurat yaitu sebesar 77%.

Saran yang muncul setelah dilakukan penelitian ini adalah perlu dilakukannya *preprocessing* untuk menentukan fitur gejala yang berpengaruh dalam penentuan penyakit Demam Berdarah dan Tifus. Penelitian selanjutnya juga diharapkan dapat memperluas ruang lingkup deteksi penyakit dengan mengakomodir data pasien dengan kategori non diagnosis yaitu terkena penyakit keduanya, bukan Demam Berdarah, atau bukan Tifus untuk memberikan ruang kemungkinan bagi gejala awal yang mirip dialami namun tidak terdiagnosis sebagai Demam Berdarah atau Tifus atau malah terkena kedua penyakit tersebut.

DAFTAR PUSTAKA

- [1] Bahri, S., Marisa Midyanti, D., & Hidayati, R. (2018). Perbandingan Algoritma Naive Bayes dan C4.5 Untuk Klasifikasi Penyakit Anak. *Seminar Nasional Aplikasi Teknologi Informasi (SNATi)*, 11–2018.
- [2] Kurniati, H. S. (2016). Gambaran Pengetahuan Ibu dan Metode Penanganan Demam Pada Balitita Di Wilayah Puskesmas Pisangan Kota Tangerang Selatan. *UINJKT Respository*. Retrieved from http://repository.uinjkt.ac.id/dspace/bitstream/123456789/33032/1/Hizah_Septi_Kurniati-FKIK.pdf.
- [3] Mau, S. (2014) ‘Sistem Pendukung Keputusan Seleksi Beasiswa Menggunakan Teorema Bayes dan Dempster Shafer’, *Pekommas*, pp. 23-32
- [4] Menteri Kesehatan RI. (2006). “*Documentslide.Com_Kmk-No-364-Ttg-Pedoman-Pengendalian-Demam-Tifoidpdf.Pdf*” (pp. 20–35). pp. 20–35
- [5] Menteri Kesehatan Republik Indonesia. (2014). Peraturan Menteri Tentang Panduan Praktik Klinis. *Lincolin Arsyad*, 3(2), 1–46. <https://doi.org/http://dx.doi.org/110.21043/equilibrium.v3i2.1268>
- [6] Nababan, M., Regasari, R., Putri, M. (2017) ‘Pembangkitan Nilai Belief Pada Dempster Shafer Dengan Particle Swarm Optimization (PSO) Untuk Penentuan Pasal Kasus Penganiayaan’, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK)*, pp. 1036-1044
- [7] Suhendro, Nina Sevani (2018) ‘Deteksi Penyakit Demam Berdarah Dan Tipus Application of Dempster-Shafer To Detect’, *Jurnal Teknik dan Ilmu Komputer*, pp. 459–471
- [8] Zuhri (2015) ‘Penentuan Prioritas Perguruan Tinggi Negeri di Indonesia Menggunakan Metode Analisa *Fuzzy Topsis*’, *Jurnal Ilman*, pp. 50-60.