

Meningkatkan Efisiensi Prediksi Risiko Diabetes Melitus dengan Metode *Fuzzy Decision Tree*

DOI: <http://dx.doi.org/10.35889/progresif.v21i2.2680>

Creative Commons License 4.0 (CC BY –NC)



Deddy Kurniawan^{1*}, Wahyu Yanuarta², Tina Tri Wulansari³, Niken Ayu Dwi Febriani⁴

Sistem Informasi, Universitas Mulia, Samarinda, Indonesia

*e-mail *Corresponding Author*: deddy.kurniawan@universitasmulia.ac.id

Abstract

The main problem in Fuzzy Logic (FL)-based prediction models is that the number of rules increases as the data dimension increases, reducing the efficiency of the system in interpretation and prediction. This study aims to unravel the complexity and improve the accuracy value of DM predictions using the Fuzzy Decision Tree (FDT) method based on Iterative Dichotomiser 3 (ID3). The research data was obtained from the National Institute of Diabetes and Digestive and Kidney Diseases with the parameters of glucose, BMI, HDL, and systolic blood pressure. The process includes data fuzzification, the formation of a decision tree with ID3, and the application of two thresholds, namely FCT and LDT. The results showed that the FDT model succeeded in reducing the number of rules by 25%, from 81 rules to 60 rules. The application of ID3-based FDT succeeded in increasing the accuracy value of DM predictions by 80%. The conclusion of the study states that the FDT model is able to unravel the complexity of the prediction model by using a simpler number of rules and can maintain and increase the accuracy value of the DM prediction model.

Keywords: Diabetes Melitus; Fuzzy Decision Tree; Fuzzy Logic; Fuzzy Sugeno; Risk Prediction

Abstrak

Permasalahan utama dalam model prediksi berbasis *Fuzzy Logic* (FL) adalah meningkatnya jumlah aturan seiring bertambahnya dimensi data, mengurangi efisiensi sistem dalam interpretasi dan prediksi. Penelitian ini bertujuan mengurai kompleksitas dan meningkatkan nilai akurasi prediksi DM menggunakan metode *Fuzzy Decision Tree* (FDT) berbasis *Iterative Dichotomiser 3* (ID3). Data penelitian diperoleh dari *National Institute of Diabetes and Digestive and Kidney Diseases* dengan parameter glukosa, BMI, HDL, dan tekanan darah sistolik. Proses meliputi fuzzifikasi data, pembentukan pohon keputusan dengan ID3, serta penerapan dua *threshold*, yaitu FCT dan LDT. Hasil penelitian menunjukkan bahwa model FDT berhasil mengurangi jumlah aturan sebesar 25%, dari 81 aturan menjadi 60 aturan. Penerapan FDT berbasis ID3 berhasil meningkatkan nilai akurasi prediksi DM sebesar 80%. Simpulan penelitian menyatakan bahwa model FDT mampu untuk mengurai kompleksitas model prediksi dengan menggunakan menghasilkan jumlah aturan yang lebih sederhana dan dapat menjaga serta meningkatkan nilai akurasi model prediksi DM.

Kata kunci: Diabetes Melitus; Pohon Keputusan Fuzzy; Logika Fuzzy; Fuzzy Sugeno; Prediksi Risiko

1. Pendahuluan

Diabetes Melitus (DM) telah menjadi krisis kesehatan global yang memerlukan respons mendesak. Laporan *International Diabetes Federation* (2024) menyebutkan bahwa 589 juta orang dewasa (20-79 tahun) hidup dengan DM pada tahun 2024, dengan proyeksi peningkatan menjadi 853 juta pada 2050 [1]. Pada tahun yang sama, DM menyebabkan 3,4 juta kematian, menegaskan urgensi sistem prediksi risiko yang akurat dan efisien untuk deteksi dini. Deteksi terlambat tidak hanya meningkatkan biaya perawatan hingga 30%, tetapi juga memperburuk prognosis pasien [2]. Sistem pendukung keputusan berbasis *Fuzzy Logic* (FL) menjadi harapan

untuk mengatasi masalah ini, namun kompleksitas aturan dan keterbatasan interpretasi klinis masih menjadi tantangan utama.

Sistem prediksi risiko DM berbasis FL secara umum menghasilkan jumlah aturan berdasarkan kombinasi parameter dan level membership function yang terdapat pada dataset [3]. Proses pembentukan *rule fuzzy* didasarkan pada kombinasi logika (seperti *AND/OR*) dari level *membership function* setiap parameter, di mana peningkatan jumlah parameter dan level *membership function* menyebabkan pertumbuhan jumlah aturan secara eksponensial [4], [5]. Ketimpangan antara kebutuhan sistem yang sederhana dan akurat dengan kompleksitas model yang ada menjadi akar masalah yang mendorong perlunya solusi inovatif.

Untuk mengatasi gap tersebut, penelitian ini mengusulkan metode *Fuzzy Decision Tree* (FDT) berbasis algoritma *Iterative Dichotomiser 3* (ID3). FDT menggabungkan fleksibilitas FL dalam menangani ketidakpastian data medis dengan struktur hierarkis pohon keputusan yang efisien. Algoritma ID3 dipilih karena kemampuannya meminimalkan entropi dan membentuk aturan optimal melalui kriteria *information gain* [6]. Studi [7] membuktikan bahwa integrasi FL dan pohon keputusan berhasil memperlancar proses pembuatan aturan tanpa intervensi dalam memprediksi jumlah produksi genteng, sementara [8] menunjukkan keunggulan FDT dalam meningkatkan akurasi klasifikasi data medis. Dalam penelitian ini, dua ambang batas *Fuzziness Control Threshold* (FCT) dan *Leaf Decision Threshold* (LDT) diterapkan untuk mengontrol kompleksitas aturan dan memastikan konsistensi interpretasi. Kombinasi ini tidak hanya mengurangi jumlah aturan, tetapi juga mempertahankan akurasi prediksi, bahkan pada dataset terbatas.

Tujuan penelitian ini adalah mengatasi ketimpangan antara kebutuhan sistem yang sederhana dan akurat dengan mengoptimalkan model prediksi berbasis FL melalui FDT berbasis algoritma ID3, sehingga mengurangi kompleksitas aturan tanpa mengorbankan akurasi model prediksi.

2. Tinjauan Pustaka

FL telah terbukti sangat efektif dalam menangani ketidakpastian yang melekat dalam diagnosis medis, memungkinkan profesional kesehatan untuk membuat keputusan yang lebih tepat [9]. Pendekatan ini memungkinkan integrasi pengetahuan ahli dan data pasien, menghasilkan peningkatan akurasi diagnostik dan hasil pasien. Kemampuan FL untuk menggabungkan beberapa variabel dan menangani data yang tidak tepat memungkinkan analisis yang lebih bernuansa dibandingkan dengan metode konvensional, yang mungkin bergantung pada ambang batas kaku dan klasifikasi biner [10].

Hasil kontribusi model prediksi DM menggunakan konvensional FL telah banyak dilakukan, seperti penelitian yang bertujuan mengurai representasi kaku dari penyakit diabetes (*Diabetes/1* dan *No diabetes/0*) dari Kumpulan data publik Kaggle *National Institute of Diabetes and Digestive and Kidney Diseases* (NIDDK) menggunakan logika fuzzy Sugeno dengan parameter glukosa, BMI, HDL, dan tekanan darah sistolik. Model ini menggunakan kombinasi fungsi keanggotaan segitiga (*trimf*) dan trapesium (*trapmf*) untuk menghasilkan tiga tingkat risiko (NR, R, VR) [3]. Dalam upaya kontribusi lain dilakukan pengujian terhadap 768 titik data dengan melibatkan parameter kadar insulin dan glukosa dan label output dari data bernilai biner. Model FL Tsukamoto yang digunakan berhasil membuat prediksi untuk 65% data yang digunakan [11]. Pada penelitian lain yang dilakukan [12] menghasilkan model prediksi berbasis FL untuk tujuan klasifikasi DM yang diujicobakan pada kumpulan data PIMA dengan parameter *Number of pregnancies*, *Plasma glucose concentration*, *Diastolic blood pressure*, *Triceps skinfold thickness*, *Serum insulin*, *BMI*, *Age*, *Diabetes Pedigree*. Model prediksi FL berhasil melakukan klasifikasi dengan menghasilkan nilai akurasi 96,47%. Penelitian lainnya dari [10] menggunakan

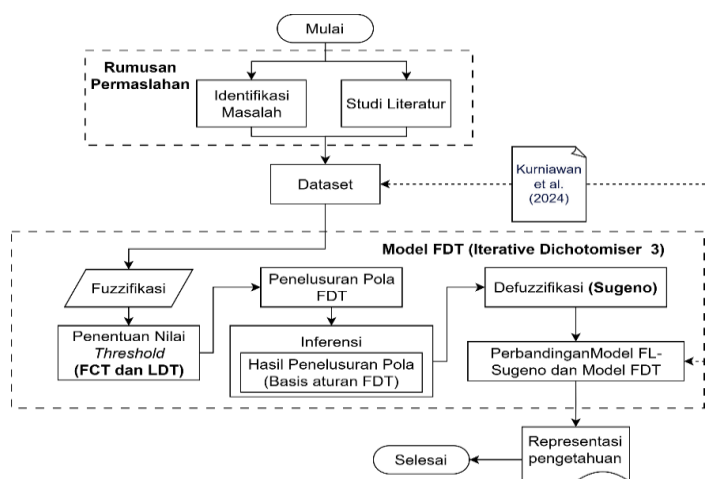
Berfokus pada upaya optimalisasi kinerja FL dalam menghasilkan rule terbaik dengan efisien adalah tantangan utama yang akan di selesaikan dalam penelitian ini. Pada penelitian diluar prediksi DM yang berkaitan dengan bagaimana penggunaan FDT digunakan juga telah banyak dilakukan, penelitian [7] berhasil menghasilkan rule terbaik dari FL di telusuri dengan FDT berbasis pada algoritma C 4.5. Model prediksi di tetapkan untuk tujuan melakukan prediksi jumlah produksi genteng. Penerapan FDT dengan kombinasi AHP digunakan sebagai alat melakukan evaluasi terhadap kualitas pengajaran guru terhadap Bahasa Inggris [13]. Hasil menunjukkan model kombinasi FDT lebih unggul dari algoritma lainnya. Pada penelitian lain FDT juga telah digunakan untuk membantu menterjemahkan hasil analisis suara pernapasan yang di hasilkan dari algoritma CNN dengan nilai akurasi mencapai 90% [14]. Penggunaan FDT untuk tujuan

klasifikasi kepribadian juga telah berhasil dilakukan [15]. FDT berbasis C 4.5 diterapkan dengan menggunakan pruning atau pemangkasan, sehingga ada kemungkinan aturan yang dihasilkan memiliki akurasi yang tinggi.

Penelitian ini mengusulkan pendekatan konferhensif baru dalam lingkup medis yang fokus kepada pengujian terkait penerapan FDT berbasis ID3 untuk prediksi DM masih sangat minim. Hanya [19] yang telah mencoba melakukan hal serupa, tetapi dengan *dataset* dan parameter berbeda dalam pengujian yang dilakukan (tidak spesifik pada glukosa, BMI, HDL, sistolik). Kebaruan model prediksi DM diperkenalkan pada penelitian ini sebagai inovasi kritis dalam prediksi risiko DM melalui implementasi *Fuzzy Decision Tree* (FDT) berbasis algoritma ID3 yang diperkuat dengan mekanisme *dual threshold* (FCT dan LDT). Berbeda dengan model FL konvensional [3] yang menghasilkan 81 aturan melalui kombinasi eksponensial parameter, pendekatan kami mengurai kompleksitas *rule explosion* secara sistematis dengan seleksi hierarkis atribut berbasis information gain (ID3) untuk membentuk struktur pohon optimal, pemangkasan dinamis menggunakan FCT untuk membatasi percabangan redundan dan stabilisasi keputusan akhir via LDT untuk menjamin konsistensi *leaf-node*.

3. Metodologi

Tahapan penelitian yang dilakukan dalam penelitian dengan pendekatan DSM ditunjukkan pada Gambar 1. Perumusan soal mengacu pada topik utama penelitian untuk menangani beban komputasi FL dalam pembentukan aturan *fuzzy* dengan penelitian ini menggunakan pendekatan *Fuzzy Decision Tree* (FDT) berdasarkan algoritma Iterative Dichotomiser 3 (ID3) untuk mengoptimalkan aturan prediksi risiko *Diabetes Melitus* (DM).



Gambar 1. *Framework* Penelitian

Tahap penelitian diawali dengan pendataan, dimana dataset diperoleh dari *National Institute of Diabetes and Digestive and Kidney Diseases* (NIDDK). Setelah data dikumpulkan, tahap pra-pemrosesan data dilakukan untuk memastikan kualitas data yang akan digunakan. Tahap ini termasuk membersihkan data untuk menghilangkan nilai yang hilang atau tidak valid. Selanjutnya, pembentukan set *fuzzy* dilakukan, di mana setiap parameter diubah menjadi domain *fuzzy* menggunakan fungsi keanggotaan trapesium (*trapmf*) dan segitiga (*trimf*).

Pada tahap fuzzifikasi, nilai numerik dari setiap parameter diubah menjadi nilai *fuzzy* berdasarkan tingkat keanggotaan yang telah ditentukan. Misalnya, kadar glukosa 87 mg/dL dapat dikategorikan dalam kisaran "Normal" dengan tingkat keanggotaan tertentu. Setelah itu, proses pembentukan FDT dilakukan dengan menggunakan algoritma ID3 (*Iterative Dichotomiser 3*). Pada tahap ini, entropi dan perolehan informasi dihitung untuk menentukan atribut terbaik dalam pemisahan cabang pohon keputusan. Selain itu, dua ambang batas utama diterapkan dalam pembentukan aturan, yaitu *Fuzziness Control Threshold* (FCT) untuk mengontrol kompleksitas aturan dan *Leaf Decision Threshold* (LDT) untuk menentukan kapan pemisahan cabang harus dihentikan.

Setelah pohon keputusan terbentuk, pembentukan aturan *fuzzy* dilakukan dengan menerjemahkan setiap cabang pohon menjadi aturan *IF-THEN* yang digunakan dalam proses

inferensi fuzzy. Misalnya, aturan yang dihasilkan bisa menjadi "IF Glukosa Tinggi AND BMI Lemak THEN Risiko Diabetes Sangat Berisiko (VR)". Tahap selanjutnya adalah defuzzifikasi, yaitu proses mengubah hasil *fuzzy* kembali ke nilai numerik (nilai tajam) menggunakan metode *Center of Gravity* (CoG) mengacu pada persamaan 1.

$$z = \frac{\sum_{i=1}^n \alpha_i \times index_i}{\sum_{i=1}^n \alpha_i} \quad (1)$$

Nilai yang dihasilkan dalam perhitungan persamaan 3 menjadi nilai *Decision Value* (DV) yang dikategorikan menjadi 3 tingkat risiko yang meliputi, Tidak Risiko (NR), Risiko (R) dan Sangat Risiko (VR). Ketiga tingkat risiko DM tersebut kemudian dinilai dengan keberadaan nilai indeks dan nilai *Up Threshold* (UT) dan *Bottom Threshold* (BT) untuk menetapkan batas yang tegas pada setiap tingkat risiko. Detail DV model ditunjukkan pada Tabel 1.

Tabel 1. Pendefinisian nilai DV

DV	Indeks	BT	UT
Tidak Berisiko (NR)	0,300	0,200	0,400
Berisiko (R)	0,500	0,400	0,600
Sangat Berisiko (VR)	0,700	0,600	0,800

Nilai *crisp* yang dihasilkan kemudian digunakan sebagai prediksi akhir tingkat risiko DM. Pada tahap terakhir, model dievaluasi dengan membandingkan hasil prediksi antara model FL-Sugeno dan FDT. Evaluasi ini meliputi pengukuran efisiensi dalam jumlah aturan, akurasi model, dan sensitivitas dalam mendeteksi kategori berisiko.

3.1. Dataset Penelitian

Studi ini menggunakan 50 data sampel yang saat pada pengujian model FL-Sugeno [3]. Data pengujian berisikan parameter Glukosa, BMI, HDL, Sistolik, label diabetes dalam dataset NIDDK secara aktual (AL) bernilai Diabetes dan Tidak diabetes, dan dilakukan penambahan hasil prediksi FL-Sugeno yang memiliki 3 tingkat label risiko diabetes yang meliputi Tidak Risiko (NR), Risiko (R) dan Sangat Risiko (VR). Statistik berdasarkan nilai CD adalah 27 data pasien berlabel Diabetes dan 23 berlabel Tidak diabetes, sedangkan berdasarkan tingkat risiko (FL-Sugeno) terdapat sebanyak 5 kasus NR, 21 kasus R dan 25 kasus VR. Rincian data sampel ditunjukkan pada Tabel 2.

Tabel 2. Detail Nilai Sampel Dataset Penelitian

No	Glukosa	BMI	HDL	Sistolik	AL	Tingkat Risiko
1	87	19	63	120	Tidak diabetes	R
2	251	33,4	48	178	Diabetes	VR
3	299	42,5	43	136	Diabetes	VR
...	...	...	...	...	...	...
38	77	22,5	49	118	Tidak diabetes	R
39	161	29,6	87	160	Diabetes	R
40	111	43	50	126	Diabetes	VR

3.2. Model Pohon Keputusan Fuzzy berbasis ID3

Persiapan awal dilakukan dengan menentukan domain dari set *fuzzy* detail nilai berdasarkan nilai kelompok rentang setiap parameter (Tabel 3).

Tabel 3. Perincian Nilai Data Parameter

Parameter	Rentang Nilai	Set Data
Glukosa	< 70 Rendah; 70 - 99 Normal; > 126 Tinggi	52 - 385
BMI	< 18,5 Kurus; 18,5 – 24,9 Ideal; > 25 Gemuk	16 - 51,4
HDL	< 40 Rendah; 40 - 59 Normal; > 60 Tinggi	14 - 120
Sistolik	< 90 Rendah; 90 - 119 Normal; > 119 Tinggi	98 - 250

Kami menetapkan tiga tingkat Variabel Linguistik (LV) dan jenis kurva trapesium (Trapmf) dan segitiga (Trimf) yang digunakan untuk menggambarkan persimpangan pada masing-masing fungsi keanggotaan. Rincian pembentukan domain set *fuzzy* ditunjukkan pada Tabel 4.

Dalam pendekatan FDT pada tahap fuzzifikasi, terdapat 2 nilai ambang batas yang didefinisikan untuk membantu meningkatkan akurasi, dimana ketika setiap kasus yang dicari dibebaskan untuk dapat terus masuk ke setiap *leaf-node* hingga dimasukkan ke dalam setiap kelas DT, akan ada peluang besar untuk mengurangi nilai akurasi model navigasi pembelajaran dalam batas pembentukan DT [16]. Nilai ambang batas FCT memiliki fungsi membatasi perluasan DT dengan menggunakan nilai ambang batas (θ_r), maka keberadaan nilai ambang batas LDT berperan dalam memastikan proporsi data pada setiap kelas akhir data yang harus dipenuhi berdasarkan nilai ambang batas (θ_n).

ID3 adalah salah satu algoritma dasar yang bertipe DT. ID3 melakukan pengukuran ketidakpastian atau ketidakaturan data pada kumpulan data selama proses penelusuran dan pembentukan DT menggunakan *nilai entropi* [17]. Hasil perhitungan nilai entropi dievaluasi dengan menerapkan pemilihan parameter yang paling signifikan melalui pembuatan subset dataset. Pembentukan cabang pohon dilakukan berdasarkan nilai evaluasi perolehan informasi. Perolehan informasi berperan dalam menggambarkan perubahan ketidakpastian kumpulan data selama proses pembentukan subset baru berdasarkan parameter yang telah ditentukan sebelumnya [18], [19]. Rumus untuk menghitung *entropi* dan *perolehan informasi* ditunjukkan pada persamaan 2 dan 3. Hasil pembentukan DT berdasarkan penelusuran ID3 kemudian dijadikan dasar *aturan* dalam proses inferensi dengan diterjemahkan ke dalam bentuk *aturan IF-THEN* yang akan digunakan dalam *mesin inferensi fuzzy*.

Tabel 4. Domain Fuzzy

Parameters	LV	Kurva	Domain
Glukosa	Rendah	Trapmf	[0 – 84,5]
	Normal	Trimf	[70 – 99]
	Tinggi	Trapmf	[84,5 – 400]
BMI	Thin	Trapmf	[0 – 21,75]
	Ideal	Trimf	[19 – 25]
	Fat	Trapmf	[21,75 – 61]
HDL	Rendah	Trapmf	[0 – 49,5]
	Normal	Trimf	[40 – 60]
	Tinggi	Trapmf	[49,5 – 130]
Sistolik	Rendah	Trapmf	[0 – 104,5]
	Normal	Trimf	[90 – 119]
	Tinggi	Trapmf	[104,5 – 300]

$$Entropy(S) = \sum_{x \in i} -p(i) \log_2 p(i) \quad (2)$$

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n * \frac{|S_i|}{S} * Entropy(S_i) \quad (3)$$

4. Hasil dan Pembahasan

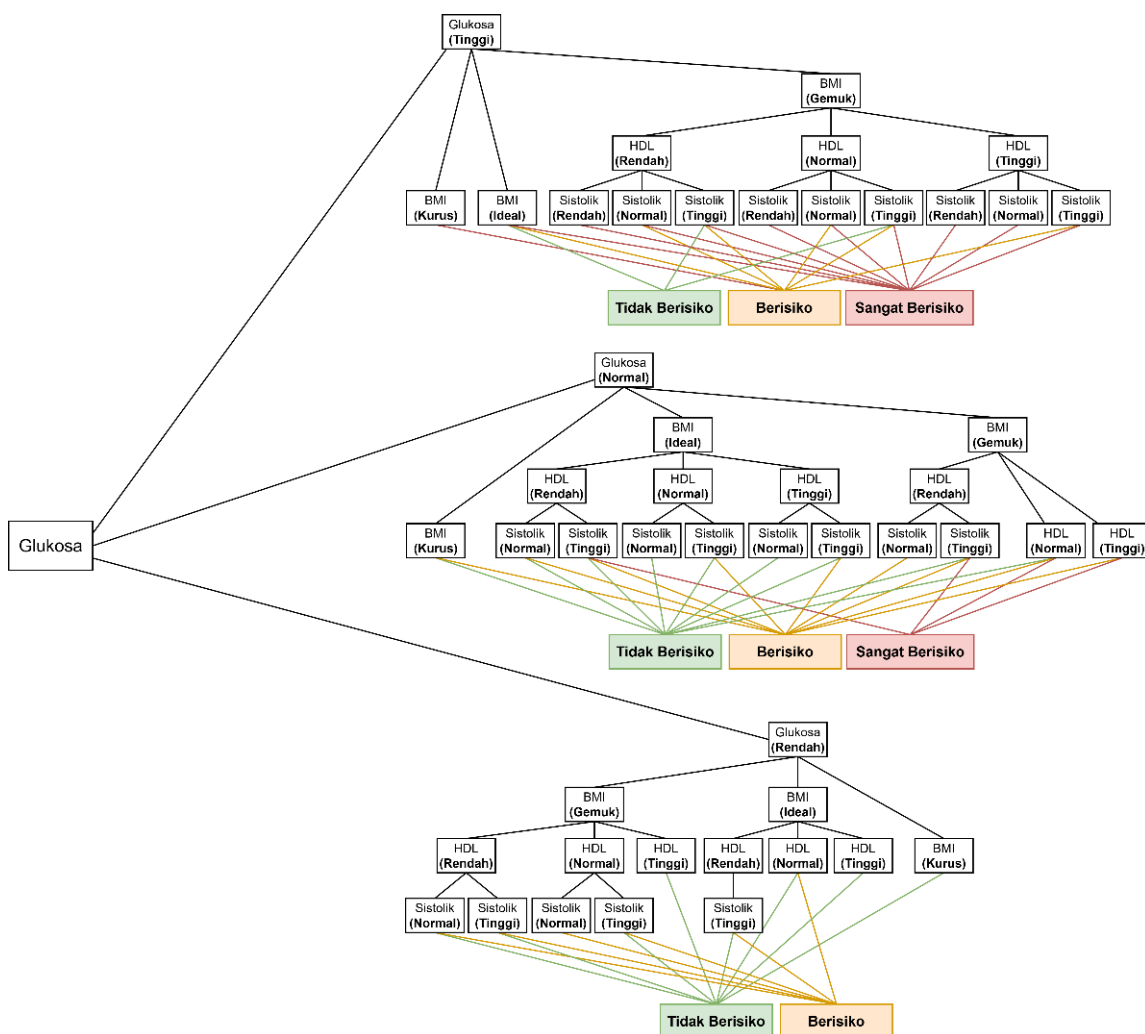
Tujuan utama penelitian ini difokuskan pada bagaimana mencapai efisiensi dalam pembentukan *pola aturan fuzzy* menggunakan metode Decision Tree atau dikenal dengan *konsep* FDT. Model prediksi berbasis FDT menggunakan algoritma ID3 dibangun dengan menerapkan 2 nilai ambang batas FCT (θ_r) dan LDT (θ_n) masing-masing 90% dan 30%.

Secara rinci, aturan *pola fuzzy* dari kedua model dibandingkan dengan melihat hasil implementasi yang dilakukan. Terjadi perubahan pada setiap tingkat prediksi risiko menderita penyakit DM, pada tingkat NR pada model FL-Sugeno memiliki 33 pola aturan, sedangkan pada model FDT menjadi lebih baik menjadi 18 pola aturan. Pada tingkat R, berubah dari 30 pola aturan menjadi 24 pola aturan, dan di tingkat VR, sedikit berubah dari 18 pola aturan menjadi 16 pola aturan. Perubahan penurunan cukup tinggi terlihat pada level VR dan R dengan nilai perubahan rata-rata 35%. Pada tingkat VR, kedua model hanya mengalami perubahan pola dari 18 menjadi 16, yang menunjukkan bahwa kedua model sama-sama konsisten selama proses pembentukan aturan. Perincian nilai perbandingan yang menunjukkan bagaimana perubahan

dari kedua model ditunjukkan pada Tabel 5 dan visualisasi rantai rule FDT ditampilkan pada Gambar 2.

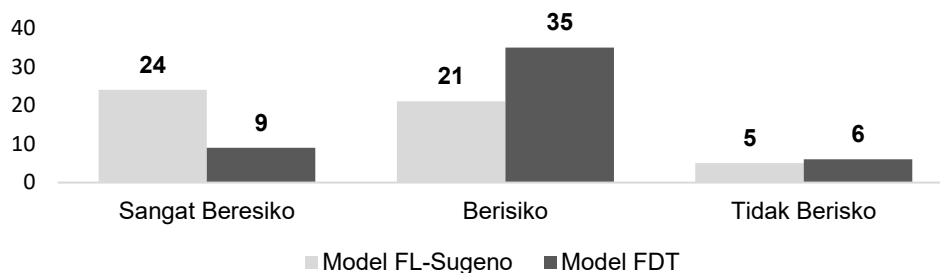
Tabel 5. Perbandingan Perubahan Pola Aturan Model FL-Sugeno Dan FDT

Tingkat Risiko DM	Model FL-Sugeno	Model FDT
Tidak Berisiko (NR)	33	18
Berisiko (R)	30	24
Sangat Berisiko (VR)	18	16



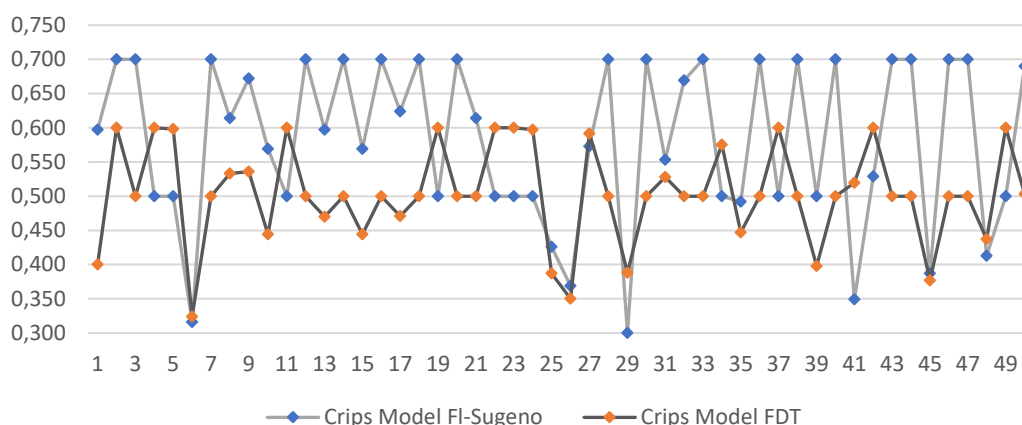
Gambar 2. Hasil Pohon Keputusan FDT

Kumpulan aturan yang dihasilkan FDT berbasis ID3 dapat memangkas pembentukan aturan sebesar 25% dari jumlah *aturan awal* (model FL-Sugeno) sebanyak 81 menjadi 60 aturan. Pada hasil evaluasi awal disajikan pada Gambar 3, perubahan hasil prediksi pada tingkat risiko DM dan nilai *fuzzy crisp* kedua model di analisis lebih jauh untuk melihat sensitifitas kedua model dalam menentukan tingkat risiko DM. Model FL-Sugeno lebih sensitif dalam menghasilkan prediksi pada tingkat Tidak berisiko/NR dengan perbandingan 24 : 9 untuk FL-Sugeno dan FDT. Sedangkan pada hasil prediksi Berisiko/R model FDT menunjukkan sensitifitas yang lebih tinggi dengan perbandingan 21 : 35 untuk model FL-Sugeno dan FDT. Dan pada tingkat risiko Sangat berisiko/VR kedua model relative memiliki sensitifitas yang berimbang dengan hasil prediksi 5:6 untuk model FL-Sugeno dan FDT.



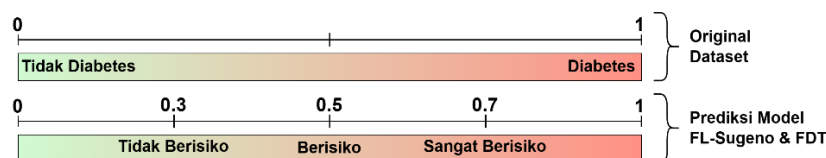
Gambar 3. Hasil Prediksi Tingkat Risiko DM Model FL-Sugeno dan FDT

Pada evaluasi lanjutan dilakukan untuk melihat secara detail nilai perubahan yang terjadi pada nilai *crisp fuzzy* dari kedua model dalam melakukan prediksi risiko DM. Detail evaluasi nilai *crisp fuzzy* yang dihasilkan dari kedua model disajikan pada Gambar 4. Penurunan nilai *crisp fuzzy* sebanyak 0.200 secara dominan terjadi pada kasus dengan indeks risiko 0.7 (VR) menjadi indeks risiko 0.5 (R). Sedangkan peningkatan nilai *crisp fuzzy* sebanyak 0.008 hingga 0.100 secara dominan terjadi pada kasus dengan indeks risiko 0.5 (R) menjadi indeks risiko 0.6 yang menyebabkan peningkatan tingkat risiko menjadi VR. Pada tingkat risiko dengan indeks 0.300 (NR) terlihat mengalami sedikit sekali perubahan naik ataupun turun pada nilai *crisp fuzzy* yang dihasilkan sehingga relative tidak mengalami perubahan tingkat risiko pada kedua model prediksi.



Gambar 4. Detail Perubahan Nilai *Crisp Fuzzy* Model FL-Sugeno dan FDT

Pada tahap akhir dilakukan analisis mendalam dengan tujuan untuk melihat bagaimana hasil prediksi tingkat risiko kedua model tersebut diterjemahkan kedalam notasi kasus sebenarnya pada AL. Hasil prediksi FL-Sugeno dan FDT dikonversi berdasarkan letak nilai indeks (0.3, 0.5, 0.7) tingkat risiko terhadap nilai indeks label pada AL dengan kondisi Tidak diabetes diwakili indeks 0 dan Diabetes diwakili indeks 1. Skema konversi hasil prediksi disajikan pada Gambar 5.



Gambar 5. Skema Konversi Hasil Prediksi Model FL-Sugeno dan FDT

Nilai prediksi (*crisp fuzzy*) kedua model dengan nilai 0 hingga < 0.500 akan dikelompokkan kedalam Tidak diabetes dan hasil nilai *crisp fuzzy* dengan rentang nilai ≥ 0.500 hingga 1 dikelompokkan kedalam Diabetes. Model FL-Sugeno berhasil memprediksi sesuai dengan AL sebanyak 35 kasus DM atau memiliki akurasi sebesar 70% [3]. Sedangkan pada model prediksi FDT yang ditawarkan sebagai hasil akhir penelitian ini memiliki hasil prediksi yang lebih baik dengan berhasil memprediksi sesuai nilai AL sebanyak 40 kasus DM atau memiliki

akurasi sebesar 80%. Hingga proses akhir perbandingan kedua model, proses prediksi kasus DM dengan model FDT berbasis ID3 tidak hanya mampu mengurai kompleksitas basis inferensi *fuzzy* dengan memangkat 25% aturan akan tetapi juga mampu meningkatkan nilai akurasi model prediksi DM.

5. Simpulan

Penelitian ini mengimplementasikan metode *Fuzzy Decision Tree* (FDT) berbasis algoritma Iterative Dichotomiser 3 (ID3) untuk mengurai kompleksitas model dalam melakukan prediksi DM. FDT dirancang untuk mengatasi tantangan dalam FL yang secara umum melakukan pembentukan aturan *fuzzy* dengan didasarkan pada kombinasi logika (seperti AND/OR) dari level *membership function* setiap parameter, di mana peningkatan jumlah parameter dan level *membership function* menyebabkan pertumbuhan jumlah aturan secara eksponensial. Hasil penelitian menunjukkan bahwa model FDT berhasil mengurangi jumlah aturan sebesar 25%, dari 81 aturan pada model FL-Sugeno menjadi 60 aturan pada model FDT dan meningkatkan nilai akurasi prediksi model menjadi sebesar 80% berdasarkan label aktual pada dataset NIDDK yang menjadi sumber data yang digunakan.

Dengan demikian, model FDT menunjukkan keunggulan dalam kompleksitas aturan yang menjadi dasar prediksi tingkat risiko DM. Pengurangan aturan yang signifikan ini menjadikan FDT model yang lebih efisien dan lebih cocok untuk aplikasi dalam sistem prediksi DM yang membutuhkan kesederhanaan dan interpretasi yang lebih baik, terutama di dunia medis yang membutuhkan kecepatan dan akurasi dalam pengambilan keputusan.

Ke depannya, penelitian ini dapat dikembangkan lebih lanjut dengan menggunakan dataset yang lebih besar atau memperkenalkan peningkatan ambang batas yang lebih dinamis untuk meningkatkan akurasi dan meningkatkan kapasitas model untuk menangani data yang lebih kompleks.

Daftar Referensi

- [1] D. J Magliano and E. J Boyko, "IDF Diabetes Atlas 11th Edition," 2025. Accessed: May 14, 2025. [Online]. Available: <https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/>
- [2] E. Dritsas and M. Trigka, "Data-Driven Machine-Learning Methods for Diabetes Risk Prediction," *Sensors*, vol. 22, no. 14, pp. 1–18, Jul. 2022, doi: 10.3390/s22145304.
- [3] D. Kurniawan, T. T. Wulansari, M. R. Ibrahim, and R. M. Fajar, "Model Berbasis Logika Fuzzy untuk Mengukur Risiko Menderita Diabetes Melitus," *Progresif: Jurnal Ilmiah Komputer*, vol. 20, no. 1, pp. 337–347, Feb. 2024, doi: 10.35889/progresif.v20i1.1587.
- [4] D. Kurniawan and D. N. Utama, "Decision Support Model using FIM Sugeno for Assessing the Academic Performance," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 605–611, Jan. 2021, doi: 10.25046/aj060165.
- [5] D. N. Utama and D. Kurniawan, "Fuzzy based Decision Support Model for Deciding the Students' Academic Performance," *International Journal of Emerging Technology and Advanced Engineering*, vol. 11, no. 10, pp. 118–130, Oct. 2021, doi: 10.46338/ijetae1021_15.
- [6] B. Zhu, J. Wang, and X. Zhang, "Fuzzy Decision Tree Based on Fuzzy Rough Sets and Z-Number Rules," *Axioms*, vol. 13, no. 12, pp. 1–23, Nov. 2024, doi: 10.3390/axioms13120836.
- [7] T. Tundo and F. Mahardika, "Fuzzy Inference System Tsukamoto–Decision Tree C 4.5 in Predicting the Amount of Roof Tile Production in Kebumen," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 7, no. 2, pp. 533–544, Apr. 2023, doi: 10.31764/jtam.v7i2.13034.
- [8] Z. Liu, A. Liu, G. Zhang, and J. Lu, "An Empirical Study of Fuzzy Decision Tree for Gradient Boosting Ensemble," Springer, 2022, pp. 716–727. doi: 10.1007/978-3-030-97546-3_58.
- [9] T. Mahboob Alam *et al.*, "Disease Diagnosis System Using IoT Empowered with Fuzzy Inference System," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 5305–5319, 2022, doi: 10.32604/cmc.2022.020344.
- [10] S. A. Abdulkareem, H. Y. Radhi, Y. A. Fadil, and H. Mahdi, "Soft computing techniques for early diabetes prediction," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 2, pp. 1167–1176, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp1167-1176.
- [11] S. Oktova Pramudya *et al.*, "Sistem Pendeteksi Diabetes Menggunakan Algoritma Tsukamoto Pada Bahasa Pemrograman Python," *Seminar Nasional Informatika Bela*

- Negara (SANTIKA)*, vol. 3, no. 1, pp. 2747–0563, 2023, Accessed: May 23, 2025. [Online]. Available: <https://santika.upnjatim.ac.id/submissions/index.php/santika/article/view/191>
- [12] K. M. Aamir, L. Sarfraz, M. Ramzan, M. Bilal, J. Shafi, and M. Attique, "A fuzzy Rule-Based System for Classification of Diabetes," *Sensors*, vol. 21, no. 23, pp. 1–17, Dec. 2021, doi: 10.3390/s21238095.
 - [13] H. Zhu, "English Teaching Quality Evaluation Based on Analytic Hierarchy Process and Fuzzy Decision Tree Algorithm," *Comput Intell Neurosci*, vol. 2022, pp. 1–10, Jul. 2022, doi: 10.1155/2022/5398085.
 - [14] J. Li *et al.*, "Explainable CNN With Fuzzy Tree Regularization for Respiratory Sound Analysis," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 6, pp. 1516–1528, Jun. 2022, doi: 10.1109/TFUZZ.2022.3144448.
 - [15] S. Aisah, F. R. Umbara, and H. Ashaury, "Klasifikasi Kerpibadian Berdasarkan Big Five Personality Menggunakan Metode Fuzzy Decision Tree Dengan Algoritma C4.5," *Jurnal Teknologi Informatika dan Komputer*, vol. 8, no. 1, pp. 333–349, Mar. 2022, doi: 10.37012/jtik.v8i1.1110.
 - [16] H. B. Santoso, "Fuzzy Decision Tree to Predict Student Success in Their Studies," *International Journal of Quantitative Research and Modeling*, vol. 1, no. 3, pp. 135–144, 2020.
 - [17] P. G. Ahire, S. Sanket Kolhe, K. Kirange, H. Karale, and A. Bhole Assistant Professor, "Implementation of Improved ID3 Algorithm to Obtain more Optimal Decision Tree," *International Journal of Engineering Research*, vol. 11, no. 02, pp. 44–47, 2015, [Online]. Available: www.ijerd.com
 - [18] A. S. Fitrani, M. A. Rosid, Y. Findawati, Y. Rahmawati, and A. K. Anam, "Implementation of ID3 Algorithm Classification Using Web-based Weka," *J Phys Conf Ser*, vol. 1381, no. 1, pp. 1–9, Nov. 2019, doi: 10.1088/1742-6596/1381/1/012036.
 - [19] F. Romansyah, I. S. Sitanggang, and S. Nurdianti, "Fuzzy Decision Tree dengan Algoritme ID3 pada Data Diabetes," *Internetworking Indonesia Jurnal*, vol. 1, no. 2, pp. 45–52, 2009, [Online]. Available: <https://www.researchgate.net/publication/298455671>