

Model Berbasis Logika *Fuzzy* untuk Mengukur Risiko Menderita Diabetes Melitus

Deddy Kurniawan^{1*}, Tina Tri Wulansari², Muhammad Rivani Ibrahim³, Rasyid Maulana Fajar⁴

^{1,2,4}Sistem Informasi, Universitas Mulia, Samarinda, Indonesia

³Prodi Manajemen Informatika, Universitas Mulia, Samarinda, Indonesia

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

Abstract

Diabetes Mellitus (DM) is a global health issue. DM is a non-communicable disease that spreads quickly. In general, type 2 diabetes mellitus (DMt2) is the most common type of diabetes suffered by people caused by irregular lifestyles. From these health issues, early identification of the risk of individuals having the opportunity to suffer from DMt2 is needed as an early warning of DMt2. A prediction model was proposed in the results of this study using the fuzzy logic (FL) Sugeno technique as the primary basis for predicting DMt2 risk numbers. The prediction model uses four parameters (glucose, BMI, HDL, and systolic) considered relevant to DMt2 cases sourced from Kaggle's public dataset. A combination of triangular (trimf) and trapezium (trapmf) curves is used for the three linguistic levels of each parameter. The final model interprets each predicted outcome into three risk levels, including no risk (NR), risky (R), and very risky (VR). The results of the FL-Sugeno model verification and validation test are based on the application of all parameters used in the model with a prediction accuracy rate of 100%.

Keywords: *Fuzzy Logic; Sugeno; Risk Suffering; Diabetes Mellitus*

Abstrak

Diabetes Melitus (DM) adalah masalah kesehatan global. DM adalah penyakit tidak menular yang menyebar dengan cepat. Secara umum, diabetes melitus tipe 2 (DMt2) adalah jenis diabetes yang paling umum diderita oleh orang-orang yang disebabkan oleh gaya hidup yang tidak teratur. Dari masalah kesehatan tersebut, identifikasi dini risiko individu berkesempatan menderita DMt2 diperlukan sebagai peringatan dini DMt2. Model prediksi diusulkan dalam hasil penelitian ini dengan menggunakan teknik *fuzzy logic* (FL) Sugeno sebagai dasar utama untuk memprediksi angka risiko DMt2. Model prediksi menggunakan empat parameter (glukosa, BMI, HDL, dan sistolik) yang dianggap relevan dengan kasus DMt2 yang bersumber dari dataset publik Kaggle. Kombinasi kurva segitiga (trimf) dan trapesium (trapmf) digunakan untuk tiga tingkat linguistik dari setiap parameter. Model yang diusulkan menafsirkan setiap hasil yang diprediksi menjadi tiga tingkat risiko, termasuk tidak ada risiko (NR), berisiko (R), dan sangat berisiko (VR). Hasil pengujian verifikasi dan validasi model FL-Sugeno berdasarkan penerapan seluruh parameter yang digunakan pada model dengan tingkat akurasi prediksi sebesar 100%.

Kata kunci: *Logika Fuzzy; Sugeno; Risiko menderita; Diabetes Melitus*

1. Pendahuluan

Menurut *World Health Organization* (WHO), Diabetes Melitus (DM) yang dikelompokkan menjadi DM tipe 1 (DMt1) dan DM tipe 2 (DMt2) dan diabetes gestasional merupakan Penyakit Tidak Menular (PTM) yang saat ini tersebar luas secara global [1], [2]. DM terjadi layaknya silent killer, dimana setiap individu sebagai penderitanya tidak menyadari gejala penyakit ini [3]. Berdasarkan hasil penelitian komprehensif yang telah dilakukan terhadap DMt2 [4] dan [5] menunjukkan bahwa gaya hidup yang tidak teratur dan kadar berlebihan dalam mengonsumsi makanan yang mengandung gula menjadi penyebab sebagian besar orang terkena penyakit DMt2. DMt2 adalah gangguan metabolisme kronis dengan beberapa etiologi yang ditandai dengan peningkatan kadar gula (glukosa) dan gangguan metabolisme karbohidrat, lipid, dan protein karena insufisiensi fungsi insulin [3].

Secara khusus, peningkatan jumlah kasus yang berisiko menderita DMt2 terjadi pada individu dengan *Index Massa Tubuh* (BMI) yang lebih tinggi [6], [7]. Banyak faktor yang dapat mempengaruhi kemungkinan seseorang menderita DMt2, dibagi menjadi faktor-faktor yang dapat diubah dan yang tidak dapat diubah [8]. Dengan kondisi ini, penilaian berupa representasi risiko menderita penyakit DMt2 menjadi sangat diperlukan. Risiko setiap individu tidak dikonfirmasi menderita DMt2 dapat menjadi peringatan dini dengan mempertimbangkan semua faktor yang berkaitan erat dengan kasus DMt2.

Sejauh ini, banyak model prediksi yang dibangun untuk melakukan prediksi DMt2 masih memberikan *output* yang kaku dan tidak memihak dengan merepresentasikan hasil prediksi mereka ke dalam kelompok label akhir prediktif dengan notasi 1 dan 0 untuk prediksi menderita DMt2 dan tidak menderita DMt2. Untuk prediksi yang lebih spesifik yang mencari risiko menderita DMt2, 1 atau 0 dapat menghasilkan nilai bias [9]. Melihat kesenjangan, teori logika *fuzzy* (FL) adalah pendekatan yang lebih mendalam terhadap nilai-nilai logis (1 atau 0) menggunakan tingkat kebenaran [10]. FL memiliki konsep melacak nilai-nilai *output* seperti otak manusia yang memikirkan berbagai faktor yang mempengaruhi keputusan tertentu. Konsep tersebut diterjemahkan menggunakan *membership function* (MF) yang terdiri dari beberapa *Linguistic Value* (LV) untuk menggambarkan nilai *input* [11], [12] atau contoh, atas pertimbangan nilai kadar gula dalam tubuh manusia, yang dapat disusun dengan nilai LV rendah, normal dan tinggi. Teori FL telah banyak digunakan untuk memecahkan masalah di berbagai bidang, misalnya [13] Membangun model berbasis *fuzzy* untuk manajemen layanan teknologi informasi, [14], [15] menghasilkan model sistem pendukung keputusan pembelian untuk jumlah barang di perusahaan farmasi berdasarkan aturan *fuzzy*, [15] membangun model pendukung keputusan untuk penilaian kota hijau Indonesia berdasarkan aturan *fuzzy*.

Penelitian ini mengusulkan model prediksi yang dapat mengukur risiko menderita DMt2 berbasis FL dengan teknik Sugeno (FL-Sugeno) *orde-nol* untuk penentuan nilai keputusan (DV) yang berupa menggunakan nilai konstan untuk setiap aturan fuzzy yang terbentuk [16]. Penerapan FL-Sugeno mewakili hasil nilai risiko menjadi tiga tingkat risiko: tidak berisiko (NR), berisiko (R), dan sangat berisiko (VR). Dari hasil penilaian akhir (DV) yang dihasilkan akan lebih dapat membantu individu dalam melihat cerminan kondisi secara lebih detail dalam bentuk tingkat risiko.

2. Tinjauan Pustaka

Di saat perkembangan teknologi dan pengetahuan terus meningkat pesat, penerapan *Machine Learning* (ML), *Artificial Intelligence* (AI), *data mining*, dan *expert system* yang berkolaborasi dengan lingkungan *Big Data* telah merambah ke semua bidang ilmu, seperti kesehatan, bisnis, pendidikan dan sebagainya.

Mengurai penelitian yang berkontribusi terhadap masalah kesehatan ini di mana penelitian telah selesai [3] menghasilkan model prediksi dengan menerapkan algoritma *Support Vector Machine* (SVM) yang memanfaatkan *dataset* publik berlabel dalam membangun model dan menghasilkan *output* prediksi yang dikelompokkan menjadi dua kelompok, yaitu terpapar DMt2 yang diwakili oleh angka 1 dan tidak terpapar DMt2 yang diwakili oleh angka 0.

Penelitian lain [17] mengusulkan model prediksi dengan menerapkan algoritma *Artificial Neural Network* (ANN) dengan *dataset* sumber yang sama untuk [3] dalam membangun model yang mengelompokkan label akhir prediksi dengan notasi 1 dan 0 untuk prediksi yang menderita DMt2 dan tidak menderita DMt2.

Hasil penelitian lebih lanjut [18] mengusulkan model dengan pendekatan klasifikasi dalam masalah ini. Model tersebut dibangun menggunakan model *hybrid* dengan menggabungkan algoritma *Firefly* dengan algoritma *K-Nearest Neighbors* (KNN), yang mempertimbangkan kedekatan setiap data dalam proses klasifikasi. Model ini dibangun dengan *dataset* berlabel publik dengan prediksi akhir memiliki diabetes mewakili angka 1 dan tidak memiliki diabetes mewakili angka 0.

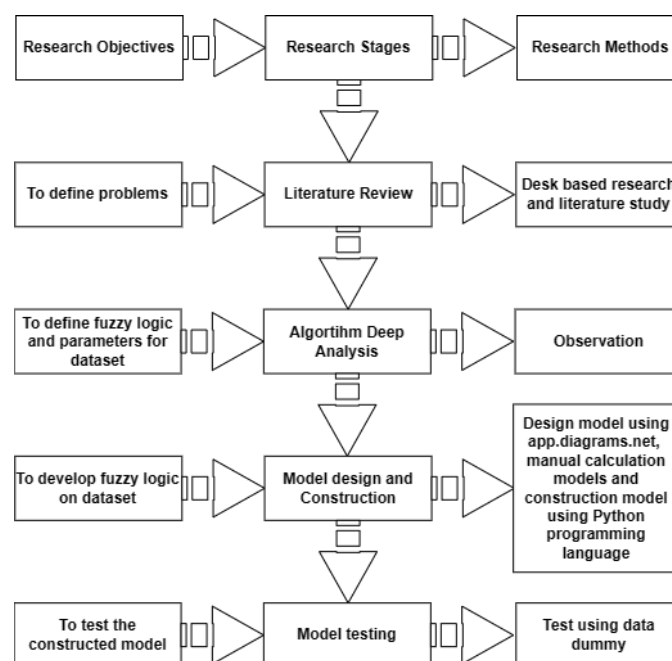
Penelitian [19] mengusulkan model prediksi menggunakan algoritma *regresi linier* dengan tema penelitian yang menyelaraskan fokus dengan penelitian ini terhadap prediksi risiko DM. Tetapi hasil akhir dari prediksi model masih memberikan nilai pasti 0 dan 1.

Akhirnya, berdasarkan hasil penelitian [20], mengusulkan model dalam bentuk klasifikasi menggunakan algoritma berbasis pohon keputusan tipe C4.5 dijalankan terhadap *dataset* publik. Hasil prediksi dengan model yang dibangun berjalan sangat baik dengan menghasilkan hasil akhir dengan label 0 dan 1.

Berdasarkan hasil penelitian pada isu kesehatan ini terdapat sebuah bias yang berasal dari hasil penilaian akhir terhadap identifikasi penyakit DMt2 terhadap individu dengan menggunakan konstanta mutlak bernilai 0 (tidak menderita) dan 1 (menderita). Pada penelitian ini berusaha menghilangkan bias pada penilaian akhir saat ini dengan memberikan kejelasan yang lebih detail dengan memanfaatkan sebaran nilai yang di petakan kedalam tiga (3) tingkat risiko menderita DMt2 yakni, tidak berisiko (NR), berisiko (R), dan sangat berisiko (VR). Setiap tingkat risiko diberikan nilai index dan *Bottom Trashould* (BT) dan *Up Trashould* (UT). Nilai *index* akan berperan dalam menggambarkan sejauh mana suatu tingkat risiko individu di dalam rentang yang ditentukan oleh *Bottom Trashould* dan *Up Trashould*.

3. Metodologi

Ada empat tahap untuk melakukan penelitian ini, dan tahap penelitian mengadopsi tahap penelitian berbasis FL [21]. Tahap-tahap yang digambarkan digambarkan pada Gambar 1. Dalam penelitian ini, FL akan diterapkan pada kumpulan dataset publik yang terkait dengan masalah DM untuk menghitung tingkat risiko individu dengan pemulung yang menderita penyakit DMt2.



Gambar 1. Blok Diagram Metodologi Penelitian

Pada tahap pertama, tinjauan literatur menunjukkan hal utama dari masalah. Banyak penelitian yang telah dilakukan sebelumnya mengenai prediksi kejadian kasus DMt2 masih belum memberikan model yang dapat menghasilkan angka risiko tertentu namun cenderung menghasilkan hasil akhir berupa pengelompokan label mutlak berupa 1 atau 0 yang mewakili menderita DMt2 dan tidak menderita DMt2. Model-model ini masih mempertimbangkan hasil perhitungan bias yang dihasilkan oleh model prediksi saat ini. Oleh karena itu, FL akan digunakan dalam menghitung seberapa besar atau kecil seseorang berisiko untuk kemungkinan menderita DMt2. FL merupakan pertimbangan penting untuk menilai angka risiko sehingga dapat diukur secara lebih spesifik, tepat, dan tidak memihak.

Metode FL, yang meliputi nilai linguistik fuzzy (LV), nilai domain, fungsi keanggotaan (MF), dan basis aturan fuzzy (FR) dengan menerapkan FL-Sugeno *orde-nol* berdasarkan empat parameter yang telah diputuskan untuk digunakan lebih lanjut dalam penelitian berdasarkan analisis mendalam dari *dataset*, semuanya akan diterapkan pada fase kedua. Pada tahap ketiga, bahasa pemrograman Python akan digunakan untuk mengembangkan model berdasarkan hasil definisi yang dilakukan pada tahap sebelumnya.

Pada penelitian ini akan dijalankan terhadap *dataset* publik populer yang bersumber dari situs penyedia data Kaggle dengan *dataset* "*Predict diabetes base on diagnostic measures*" dengan total 250 rekaman data kasus DMt2 yang di dominasi kejadian kasus "No Diabetes".

Skema pengujian model pada akhir penelitian dilakukan dengan melakukan pengambilan sampel yang seimbang (*balanced dataset*) dari setiap label kasus “No Diabetes” dan “Diabetes” dengan total data sampel sebanyak 50 kasus DMt2. Detail Keseluruhan parameter yang tersedia pada *dataset* disajikan pada Tabel 1.

Tabel 1. Keseluruhan Parameter *Dataset*

| No | Parameter | Deskripsi | Nilai |
|----|------------------|--|-------------------------------|
| 1 | cholesterol | Kolesterol total dalam darah | 78-443 |
| 2 | glucose | Glukosa | 48-384 |
| 3 | hdl_chol | <i>High-density</i> lipoprotein yang merupakan kolesterol baik | 12-120 |
| 4 | chol_hdl_ratio | Perbandingan antar kolesterol di dalam darah | 2-193 |
| 5 | age | Umur | 19-92 |
| 6 | gender | Jenis kelamin | Male, female |
| 7 | height | Tinggi badan | 52-76 |
| 8 | weight | Berat badan | 99-325 |
| 9 | bmi | <i>Index</i> masa tubuh atau proporsional tubuh | 16-558 |
| 10 | systolic_bp | Tekanan darah sistolik | 90-250 |
| 11 | diastolic_bp | Tekanan darah diastolik | 48-124 |
| 12 | waist | Lingkar perut | 25-56 |
| 13 | hip | Lingkar panggul | 30-64 |
| 14 | waist_hip_ratio | Perbandingan lingkar perut dan lingkar panggul | 1-114 |
| 15 | Diabetes (Label) | Suatu penyakit metabolik yang diakibatkan oleh meningkatnya kadar glukosa atau gula darah. | No diabetes (0), diabetes (1) |

Keseluruhan parameter dari *dataset* terlebih dahulu dilakukan uji korelasi terhadap parameter label sebagai tujuan akhir. Disamping uji korelasi dilakukan pula penghapusan beberapa parameter dari keseluruhan parameter dari Tabel 1, misal pada parameter tinggi badan (*height*) dan berat badan (*weight*) dapat diwakili dengan parameter yang merepresentasikan kedua parameter tersebut. Hasil akhir dari uji korelasi parameter disajikan pada (Tabel 2) yang akan menjadi parameter utama dalam membangun model pada penelitian ini.

Tabel 2. Nilai Korelasi Parameter

| Parameter | Nilai Korelasi |
|-----------|----------------|
| HDL | 0.6605 |
| BMI | 0.2062 |
| Sistolik | 0.0201 |
| Glukosa | -0.0915 |

HDL adalah parameter yang paling berpengaruh, dengan nilai 0,6605. Secara umum, pada kasus DMt2, peningkatan angka HDL dipengaruhi oleh parameter BMI yang lebih signifikan dimana nilai korelasi BMI adalah 0,2062. Parameter glukosa adalah parameter kritis yang selalu menentukan kejadian DMt2 dari catatan kasus dalam *dataset* yang digambarkan dengan nilai negatif (-0,0915). Hasil identifikasi nilai korelasi dihasilkan menggunakan alat Rapidminer. Rincian nilai masing-masing parameter *dataset* disajikan dalam Tabel 3, bersama dengan pembentukan rentang nilai sebagai domain *dataset* yang akan digunakan selanjutnya dalam pembentukan domain himpunan *fuzzy*.

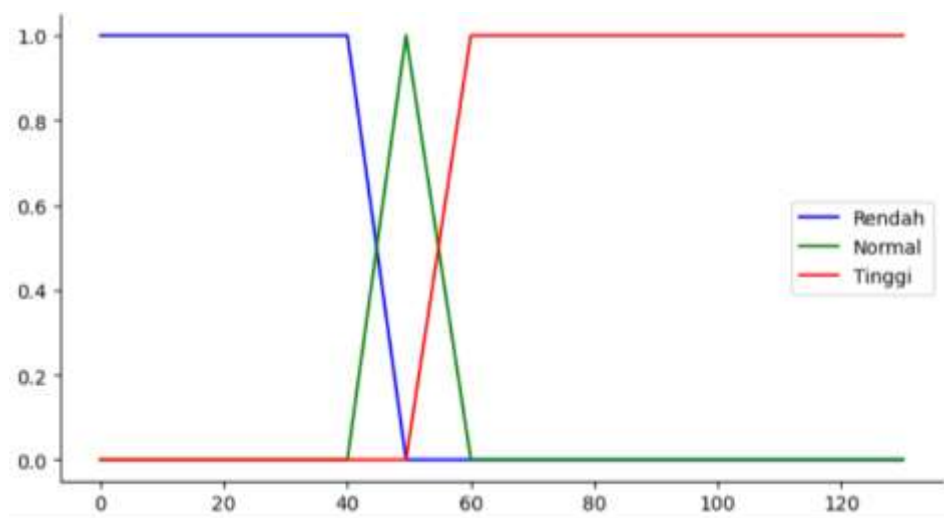
Tabel 3. Rincian Nilai dan Nilai Kelompok Rentan

| Parameter | Kelompok Rentan | Himpunan Data Nilai |
|-----------|--|---------------------|
| HDL | < 40 Rendah; 40-59 Normal; > 60 Tinggi | 14-120 |
| BMI | < 18.5 Thin; 18.5-24.9 Ideal; > 25 Gemuk | 16-51.4 |
| Sistolik | < 90 Rendah; 90-119 Normal; > 119 Tinggi | 98-250 |
| Glukosa | < 70 Rendah; 70-99 Normal; > 126 Tinggi | 52-385 |

Berdasarkan data dalam Tabel 3, selanjutnya ditentukan domain himpunan *fuzzy* (Tabel 4). Setiap parameter diturunkan menjadi tiga tingkat LV, misalnya, kurus, ideal, dan gemuk untuk parameter BMI. Setiap LV menentukan penentuan nilai domain untuk setiap LV yang ada, LV kurus [0-21,75], LV ideal [19-25], dan LV gemuk [21,75-61]. Penentuan nilai domain dilakukan dengan memberikan *overlap* pada pembentukan kurva dengan menggunakan kombinasi jenis kurva trapesium (*trapmf*) dan segitiga (*trimf*). Rincian pembentukan set *fuzzy* disajikan pada Tabel 3. Setiap parameter berdasarkan Tabel 4 direpresentasikan ke dalam grafik MF seperti pada Gambar 2 (salah satu contoh MF untuk parameter HDL).

Tabel 4. Domain Himpunan *Fuzzy*

| Parameters | LV | Kurva | Domain |
|------------|--------|--------|-------------|
| HDL | Rendah | trapmf | [0-49.5] |
| | Normal | trimf | [40-60] |
| | Tinggi | trapmf | [49.5-130] |
| BMI | Kurus | trapmf | [0-21.75] |
| | Ideal | trimf | [19-25] |
| | Gemuk | trapmf | [21.75-61] |
| Sistolik | Rendah | trapmf | [0-104,5] |
| | Normal | trimf | [90-119] |
| | Tinggi | trapmf | [104.5-300] |
| Glukosa | Rendah | trapmf | [0-84.5] |
| | Normal | trimf | [70-99] |
| | Tinggi | trapmf | [84.5-400] |



Gambar 2. MF Parameter HDL

Empat parameter utama semuanya memiliki MF dengan LV dan nilai batas yang sama: 'Rendah,' 'Normal' dan 'Tinggi' (Gambar. 2). Dalam parameter pertama HDL, trapesium MF (0, 0, 40, 49,5) digunakan untuk mewakili LV 'rendah', segitiga MF (40, 49,5, 60) untuk LV 'normal' dan trapesium MF (49,5, 60, 130, 130) untuk LV 'tinggi'. Ketiga LV yang kemudian diterapkan dalam representasi kurva MF pada empat parameter diberikan batas nilai yang ditetapkan berdasarkan analisis mendalam tentang kondisi statistik *dataset* dan studi literatur.

Tiga parameter yang tersisa termasuk BMI dengan MF (0, 0, 18,5, 21,75) dengan LV 'kurus', MF (18,5, 21,75, 25) untuk LV 'ideal', dan MF (21,75, 25, 61, 61) dengan LV 'gemuk'. Pada parameter sistolik memiliki MF (0, 0, 90, 104,5) untuk LV 'rendah', kemudian MF (90, 104,5, 119) untuk LV 'normal', dan MF (104,5, 119, 300, 300) untuk LV 'tinggi'. Terakhir, parameter glukosa memiliki MF (0, 0, 70, 84,5) untuk LV 'rendah', MF (70, 84,5, 99) untuk LV 'normal', dan MF (84,5, 99, 400, 400) untuk LV 'tinggi'.

Dalam keseluruhan proses, perhitungan *fuzzy* akan mencakup semua inferensi di mana dasar aturan dibuat berdasarkan seluruh LV dari empat parameter yang akan digunakan [9]. Semua basis aturan (Tabel 5) dibentuk berdasarkan hasil studi pustaka kejadian kasus DMt2 yang melibatkan empat parameter yang digunakan dalam penelitian ini. Basis aturan *fuzzy* akan menghasilkan nilai keputusan akhir (DV) untuk tingkat risiko menderita DMt2.

Tabel 5. FR untuk Menilai Resiko DMt2

| No | Glukosa | BMI | HDL | Sistolik | Risiko |
|-----|---------|-------|--------|----------|----------------|
| 1 | Rendah | Kurus | Rendah | Rendah | Tidak Berisiko |
| 2 | Rendah | Kurus | Rendah | Normal | Tidak Berisiko |
| 3 | Rendah | Kurus | Rendah | Tinggi | Tidak Berisiko |
| ... | ... | ... | ... | ... | ... |
| 35 | Normal | Kurus | Tinggi | Normal | Berisiko |
| 36 | Normal | Kurus | Tinggi | Tinggi | Berisiko |
| 37 | Normal | Ideal | Rendah | Rendah | Tidak Berisiko |
| 38 | Normal | Ideal | Rendah | Normal | Tidak Berisiko |
| 39 | Normal | Ideal | Rendah | Tinggi | Tidak Berisiko |
| ... | ... | ... | ... | ... | ... |
| 79 | Tinggi | Gemuk | Tinggi | Rendah | Berisiko |
| 80 | Tinggi | Gemuk | Tinggi | Normal | Berisiko |
| 81 | Tinggi | Gemuk | Tinggi | Tinggi | Berisiko |

Sebagai *decision value* (DV) diambil dengan menghitung nilai *Center of Gravity* (CoG) menggunakan formula 1 yang secara spesifik nilai DV dari model terhadap kondisi individu dikategorikan kedalam tingkat risiko kemungkinan menderita DMt2 ke dalam tingkat: tidak berisiko (NR), berisiko (R), dan sangat berisiko (VR), dengan nilai *index* masing-masing 0,300, 0,500, dan 0,700. Definisi DV diberikan nilai *Up Trashould* (UT) dan *Bottom Trashould* (BT) untuk menetapkan batasan tegas pada setiap tingkat risiko. Detail DV model ditunjukkan pada Tabel 4.

$$DV = \frac{\sum_{i=1}^n \alpha_i * Index_i}{\sum_{i=1}^n \alpha_i} \quad (1)$$

Keterangan:

DV = *decision value*

α_i = Nilai alpa predikat dari basis FR

Index = Nilai pada tingkat risiko DMt2

Tabel 6. *Index* DV Model

| DV | <i>Index</i> | <i>Bottom Trashould</i> (BT) | <i>Up Trashould</i> (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 | 0.700 |

Pada akhir penelitian dilakukan uji verifikasi dan validasi terhadap model prediksi FL-Sugeno. Uji verifikasi berfungsi untuk memeriksa derajat kebenaran model yang telah berhasil dibangun berdasarkan konsep teori yang digunakan untuk menyelesaikan suatu permasalahan [22]. Sedangkan pada proses validasi model dilakukan dengan melakukan perbandingan hasil prediksi data model terhadap data aktual yang ada di lapangan yang dihasilkan secara manual [22]. Hal ini dilakukan untuk mengetahui sejauh mana model dapat menghasilkan nilai akurat dan valid dalam konteks yang sesuai.

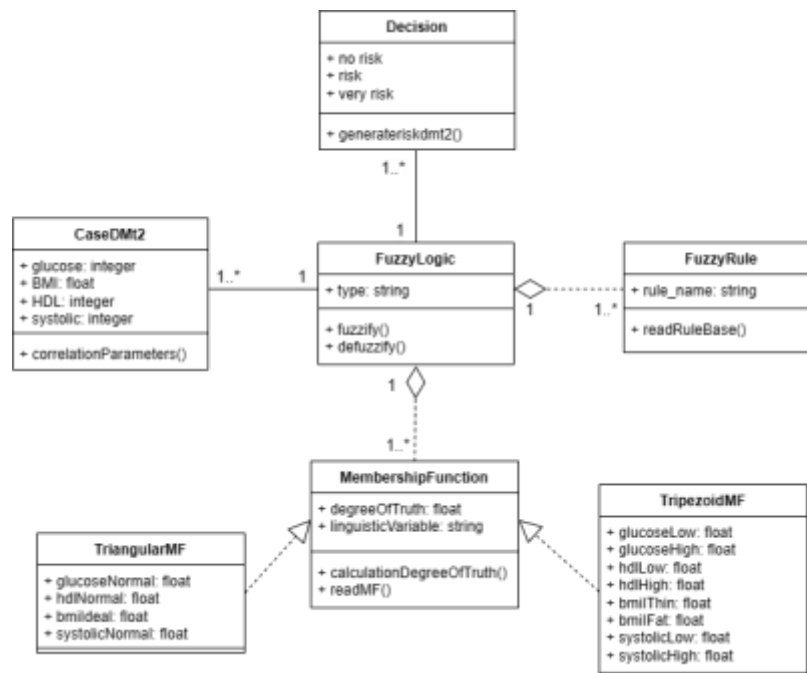
4. Hasil dan Pembahasan

Berdasarkan langkah-langkah metodologis pada Gambar 1, pemilihan *dataset* ditentukan terhadap *dataset* publik yang bersumber dari situs *dataset* besar Kaggle dengan *dataset* "*Predict diabetes base on diagnostic measures*". Detail himpunan data yang akan digunakan ditunjukkan pada Tabel 7.

Tabel 7. Sampel Dataset

| No | Glukosa | BMI | HDL | Sistolik | Label |
|-----|---------|------|-----|----------|----------------|
| 1 | 87 | 19 | 63 | 120 | Tidak Diabetes |
| 2 | 251 | 33,4 | 48 | 178 | Diabetes |
| ... | ... | ... | ... | ... | ... |
| 49 | 161 | 29,6 | 87 | 160 | Diabetes |
| 50 | 111 | 43 | 50 | 126 | Diabetes |

Model prediksi dalam penelitian ini diilustrasikan dengan menggunakan diagram kelas yang ditunjukkan pada Gambar 3. Ada beberapa langkah yang diperlukan untuk membangun model prediksi ini, termasuk *CaseDMt2*, *Decision*, *FuzzyLogic*, *FuzzyRule*, *MembershipFunction*, *TriangularMF*, dan *TrapezoidMF*. Kelas *CaseDMt2* menjelaskan parameter terhadap kasus DMt2 terhadap individu.

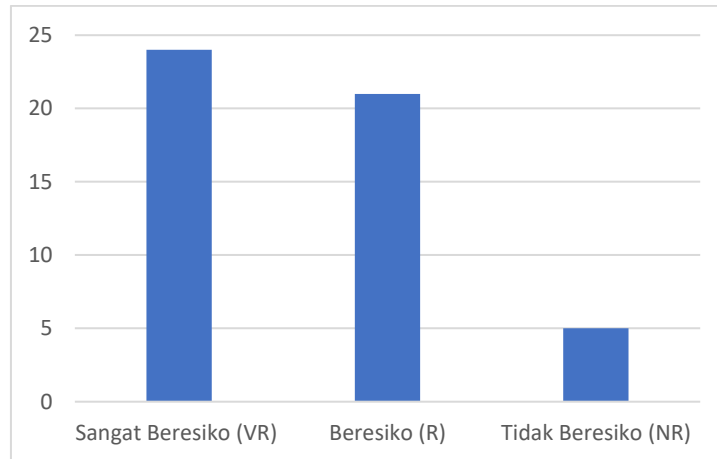


Gambar 3. Kelas Diagram Model FL-Sugeno DMt2

Output model dari *dataset* yang diuji diwakili oleh DV, yang memiliki tiga tingkat risiko berdasarkan Tabel 6, di mana 50 dummy data kasus yang diuji ke model menghasilkan *output* yang lebih spesifik yang dapat menggambarkan tingkat risiko penderitaan individu dari DMt2 dengan hasil *output* prediksi model terperinci pada Tabel 8 dan detail sebaran data pada Gambar 4.

Tabel 8. Perbandingan Hasil Prediksi Model

| Glucose | BMI | HDL | Systolic | DV | Tingkat Risiko | Label Dataset |
|---------|------|-----|----------|-------|----------------|---------------|
| 87 | 19 | 63 | 120 | 0,503 | R | No diabetes |
| 251 | 33,4 | 48 | 178 | 0,700 | VR | Diabetes |
| 299 | 42,5 | 43 | 136 | 0,700 | VR | Diabetes |
| ... | ... | ... | ... | ... | ... | ... |
| 77 | 22,5 | 49 | 118 | 0,413 | R | No diabetes |
| 161 | 29,6 | 87 | 160 | 0,500 | R | Diabetes |
| 111 | 43 | 50 | 126 | 0,690 | VR | Diabetes |



Gambar 4. Distribusi Hasil Prediksi Model

Keseluruhan prediksi kasus didominasi pada tingkat menderita VR sebanyak 24 kasus, tingkat menderita R dan NT berturut-turut sebesar 21 dan 5. Lebih lanjut dalam salah satu contoh kasus, nilai input dalam parameter glukosa, BMI, HDL, dan sistolik adalah 251, 33,4, 48 dan 178 dengan model prediksi yang dibangun menghasilkan output DV 0,700 pada *index* VR (sangat berisiko). Dalam kasus lain, hasil output cenderung mendekati ambang batas UT atau BT yang ditentukan untuk setiap *index* (Tabel 6). Dengan data parameter berturut-turut sebesar 77, 25,8, 50, dan 122, menghasilkan nilai DV sebesar 0,597, mendekati nilai UP pada *index* R (berisiko), yaitu 0,600, dan BT pada *index* VR (sangat berisiko) sebesar 0,600. Hasil *output* menggambarkan bahwa individu cenderung berada pada tingkat VR (sangat berisiko), dengan nilai yang dihasilkan pada DV jauh dari nilai *index* R (berisiko), yaitu sebesar 0,500. Hasil prediksi model secara keseluruhan diilustrasikan pada Gambar.4, dengan sebagian besar hasil prediksi data berada pada tingkat risiko (R) dan sangat berisiko (VR) setelah proses penyaringan tingkat risiko.



Gambar 5. Keseluruhan Hasil Prediksi Model

Dalam tujuan akhir penelitian yang bertujuan menghilangkan bias dari hasil penelitian terdahulu yang merepresentasikan hasil akhir berupa pengelompokan label mutlak berupa 1 atau 0 yang mewakili menderita DMt2 dan tidak menderita DMt2 pada dataset penelitian. Hasil prediksi model yang disulkan dapat menjadi solusi baru dalam melakukan prediksi DMt2 secara spesifik

dengan memberikan informasi terkait tingkat menderita individu yang digambarkan kedalam tiga tingkat risiko (NR, R dan VR). Informasi ini terhadap kondisi individu terhadap tingkat risiko menderita DMt2 menjadi sangat berharga bagi individu tersebut maupun tenaga medis terkait dalam melakukan monitoring dan diagnosa lanjutan terhadap kondisi aktual individu sebagai pengetahuan penting dalam pengambilan keputusan akhir yang akan diberikan. Detail hasil perbandingan hasil prediksi Model DMt2 baru terhadap model prediksi terdahulu disajikan pada Tabel 9.

Tabel 9. Perbandingan hasil Prediksi Model dengan Label *Dataset*

| DV | Tingkat Menderita | Simbol Label | Label <i>Dataset</i> |
|-------|-------------------|--------------|----------------------|
| 0,503 | Beresiko | 0 | No diabetes |
| 0,700 | Sangat Beresiko | 1 | Diabetes |
| 0,700 | Sangat Beresiko | 1 | Diabetes |
| ... | ... | ... | ... |
| 0,413 | Beresiko | 0 | No diabetes |
| 0,500 | Beresiko | 1 | Diabetes |
| 0,690 | Sangat Beresiko | 1 | Diabetes |

Hasil uji verifikasi dilakukan dengan pertama membandingkan seluruh kebutuhan *input* model (parameter) berdasarkan referensi parameter yang bersumber kepada detail gambaran dari *dataset* publik " *Predict diabetes base on diagnostic measures*" pada Tabel 1. selanjutnya melihat seluruh tahapan yang ada pada teori *fuzzy (fuzzy-inferensi-defuzzy)* dan formula yang dibutuhkan dalam membangun model serta hasil nilai prediksi yang dihasilkan oleh model terhadap perhitungan aktual secara konvensional. Hasil uji verifikasi model memperoleh nilai 1 yang bermakna bahwa seluruh komponen yang di ujikan (parameter, prosedur, formula dan hasil) telah sesuai dengan seluruh referensi yang digunakan sebagai dasar membangun model FL-Sugeno untuk melakukan prediksi angka resiko menderita DMt2. detail uji verifikasi disajikan pada Tabel 10.

Tabel 10. Hasil Uji Verifikasi Model FL-Sugeno

| Model | Komponen | Model | Referensi | Verifikasi |
|-----------|-------------|-------------|-------------|------------|
| FL-Sugeno | ∑ Parameter | 4 | 4 | 1.00 |
| | Prosedur | √ | √ | |
| | Formula | √ | √ | |
| | Hasil | 0,300-0,700 | 0,300-0,700 | |

Pengujian kedua dilakukan dengan memvalidasi penggunaan nilai *input* pada setiap parameter dan *output* (DV) pada model FL-Sugeno terhadap referensi utama yang digunakan (Tabel 3). Setiap nilai prediksi di ujikan seluruh data *dummy* yang telah ditentukan dengan jumlah data sebanyak 50 data kasus DMt2. Hasil validasi model berdasarkan seluruh komponen uji meliputi nilai *input* dan *output* model di sajikan pada Tabel 11.

Tabel 11. Hasil Uji Validasi Model FL-Sugeno

| Model | Parameter | Model | Referensi | Nilai | | Hasil Prediksi | | | |
|------------|-----------|---------|-----------|-------|----|----------------|----|----|----|
| | | | | T | F | TP | FP | TN | FN |
| FL-Sugeno | BMI | 14-120 | 14-120 | 1 | 1 | | | | |
| | HDL | 16-51.4 | 16-51.4 | 1 | 1 | | | | |
| | Glukosa | 98-250 | 98-250 | 1 | 1 | | | | |
| | Sistolik | 52-385 | 52-385 | 1 | 1 | | | | |
| Total Data | | | | 50 | 50 | | | | |

Hasil validasi model FL-Sugeno secara keseluruhan di rangkum pada Tabel 12. Dimana terlihat nilai sensitifitas sebesar 1.00 dimana ini bermakna bahwa model FL-Sugeno memiliki seluruh nilai parameter sebesar 100% benar berdasarkan referensi yang digunakan. Sedangkan

Nilai Prediksi Positif (NPP) mendapatkan nilai 1.00 dimana ini menunjukkan bahwa model berhasil melakukan prediksi dengan benar dan akurat jika dibandingkan dengan hasil perhitungan nilai konvensional yang menerapkan seluruh parameter yang dijadikan acuan berdasarkan referensi yang ada.

Tabel 12. Rangkuman Hasil Validasi Model FL-Sugeno

| | | Nilai Prediksi | | |
|----------|-------|----------------|------------|---------------------|
| | | True | False | |
| Validasi | True | TP = 50 | FN = 0 | Sensitifitas = 1.00 |
| | False | FP = 0 | TN = 0 | |
| | | | NPP = 1.00 | |

5. Simpulan

Penelitian ini menggunakan empat parameter untuk memprediksi risiko menderita DMt2: glukosa, BMI, HDL, dan sistolik. Tiga tingkat LV untuk setiap parameter dibentuk berdasarkan FL pemikiran dasar utama yang digunakan dalam penelitian ini. Model tersebut dibangun dengan FL-Sugeno dalam memprediksi tingkat risiko yang kemungkinan menderita DMt2, yang dikelompokkan menjadi tiga tingkat risiko yaitu, tidak berisiko (NR), berisiko (R), dan sangat berisiko (VR). Model berikutnya yang dihasilkan dari penelitian menunjukkan kinerja positif dalam memprediksi risiko individu memiliki kesempatan untuk menderita DMt2, hasil kinerja di buktikan dengan rangkuman kinerja model FL-Sugeno pada Tabel 12 yang mewakili kemampuan model dalam melakukan prediksi angka risiko menderita DMt2 berdasarkan penerapan seluruh parameter yang digunakan pada model dengan tingkat akurasi prediksi sebesar 100% di bandingkan dengan hasil perhitungan aktual dengan menggunakan metode konvensional.

Selanjutnya, berdasarkan hasil pengujian model FL-Sugeno dari sisi verifikasi dan validasi model memberikan hasil prediksi yang diharapkan dapat membantu tim medis menentukan apakah individu berpotensi mengembangkan kasus DMt2 [23]. Tantangan selanjutnya yang dapat dilakukan untuk mengoptimalkan hasil prediksi model dapat dilakukan dengan membangun model yang dilengkapi dengan banyak parameter yang lebih relevan sebagai prediktor terbaik untuk kasus DMt2. Integrasi teknologi kesehatan adalah peluang yang sangat baik di masa depan untuk membangun integrasi data dari perangkat dengan model prediksi berbasis *fuzzy* ini.

Daftar Referensi

- [1] U. I. Lestari, A. Y. Nadhiroh, and C. Novia, "Penerapan Metode K-Nearest Neighbor Untuk Sistem Pendukung Keputusan Identifikasi Penyakit Diabetes Melitus," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 4, pp. 2071–2082, Dec. 2021, doi: 10.35957/jatisi.v8i4.1235.
- [2] J. Ginting, R. Ginting, and H. Hartono, "Deteksi Dan Prediksi Penyakit Diabetes Melitus Tipe 2 Menggunakan Machine Learning (Scooping Review)," *Jurnal Keperawatan Priority*, vol. 5, no. 2, pp. 93–105, Aug. 2022, doi: 10.34012/jukep.v5i2.2671.
- [3] S. W. Hovi, A. I. Hadiana, and F. R. Umbara, "Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)," *Informatics and Digital Expert (INDEX)*, vol. 4, no. 1, pp. 40–45, Jul. 2022, doi: 10.36423/index.v4i1.895.
- [4] P. Piko, N. A. Werissa, S. Fiatal, J. Sandor, and R. Adany, "Impact of Genetic Factors on the Age of Onset for Type 2 Diabetes Mellitus in Addition to the Conventional Risk Factors," *J Pers Med*, vol. 11, no. 6, pp. 1–17, Dec. 2020, doi: 10.3390/jpm11010006.
- [5] A. Izzo, E. Massimino, G. Riccardi, and G. Della Pepa, "A Narrative Review on Sarcopenia in Type 2 Diabetes Mellitus: Prevalence and Associated Factors," *Nutrients*, vol. 13, no. 183, pp. 1–18, Jan. 2021, doi: 10.3390/nu13010183.
- [6] M. A. B. Khan, M. J. Hashim, J. K. King, R. D. Govender, H. Mustafa, and J. Al Kaabi, "Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends," *J Epidemiol Glob Health*, vol. 10, no. 1, p. 107, 2019, doi: 10.2991/jegh.k.191028.001.
- [7] R. Ohno *et al.*, "Association of Body Mass Index and Its Change With Incident Diabetes Mellitus," *J Clin Endocrinol Metab*, Jun. 2023, doi: 10.1210/clinem/dgad374.

- [8] K. R. Widiyarsi, I. M. K. Wijaya, and P. A. Suputra, "Diabetes Melitus Tipe 2: Faktor Risiko, Diagnosis, Dan Tatalaksana," *Ganesha Medicine*, vol. 1, no. 2, p. 114, Dec. 2021, doi: 10.23887/gm.v1i2.40006.
- [9] G. I. Perdana and D. N. Utama, "The Multistage Fuzzy Logic-Based Decision Support Model to Determine the Salary of New Employees," *Journal of Computer Science*, vol. 19, no. 7, pp. 812–824, 2023, doi: 10.3844/jcssp.2023.812.824.
- [10] L. A. Zadeh, "Fuzzy logic," *Computer (Long Beach Calif)*, vol. 21, no. 4, pp. 83–93, Apr. 1988, doi: 10.1109/2.53.
- [11] 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.
- [12] 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.
- [13] R. Yandri, Suharjo, D. N. Utama, and A. Zahra, "Evaluation Model for the Implementation of Information Technology Service Management using Fuzzy ITIL," *Procedia Comput Sci*, vol. 157, pp. 290–297, 2019, doi: 10.1016/j.procs.2019.08.169.
- [14] G. K. Ramadhan and D. N. Utama, "Fuzzy Tsukamoto based Decision Support Model for Purchase Decision in Pharmacy Company," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 4, pp. 3868–3874, Nov. 2019, doi: 10.35940/ijrte.D8243.118419.
- [15] D. N. Utama and N. Wibowo, "A Simple Fuzzy Decision Support Model for Determining the Indonesian Green Cities," in *IOP Conference Series: Earth and Environmental Science*, Apr. 2023, p. 012058. doi: 10.1088/1755-1315/1169/1/012058.
- [16] R. Furqan Ramadhan and K. Eliyen, "Implementasi Metode Fuzzy Sugeno Pada Pemodelan Pertumbuhan Tanaman Kedelai Dengan Variasi Dosis Pupuk," *Jurnal Mnemonic*, vol. 6, no. 1, pp. 28–34, May 2023, doi: 10.36040/mnemonic.v6i1.5831.
- [17] C. Maisyarah, E. Haryatmi, R. Y. Fajriatifah, and Y. H. Puspita, "Prediksi Penyakit Diabetes menggunakan Algoritma Artificial Neural Network," *Jurnal Data Science & Informatika*, vol. 2, no. 1, pp. 46–52, 2022.
- [18] D. Abdianto Nggego, R. R. Taufik Bau, and N. Patawaran, "FA-KNN: Hybrid Algoritma Untuk Klasifikasi Penyakit Diabetes Melitus," *MJRICT: Musamus Journal Of Research Information and Communication Technology*, vol. 5, no. 2, pp. 71–80, 2023, [Online]. Available: <https://ejournal.unmus.ac.id/index.php/mjrict>
- [19] Q. R. Cahyani, M. J. Finandi, J. Rianti, D. L. Arianti, A. D. Pratama, and Putra, "Prediksi Risiko Penyakit Diabetes Menggunakan Algoritma Regresi Logistik Diabetes Risk Prediction Using Logistic Regression Algorithm," *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, vol. 1, no. 2, pp. 2828–9099, 2022, doi: 10.55123/jomlai.v1i2.598.
- [20] R. Putri Fadhilah *et al.*, "Klasifikasi Penyakit Diabetes Mellitus Berdasarkan Faktor-Faktor Penyebab Diabetes Menggunakan Algoritma C4.5," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 4, pp. 1265–1270, 2022, [Online]. Available: www.kaggle.com
- [21] G. I. Perdana, M. D. Devanda, and D. N. Utama, "Fuzzy Based Butterfly Life Cycle Algorithm for Measuring Company's Growth Performance," *Journal of Advances in Information Technology*, vol. 14, no. 1, pp. 1–6, 2023, doi: 10.12720/jait.14.1.1-6.
- [22] D. N. Utama, L. I. Lazuardi, H. A. Qadrya, B. M. Caroline, T. Renanda, and A. P. Sari, "Worth eat: An intelligent application for restaurant recommendation based on customer preference (Case study: Five types of restaurant in Tangerang Selatan region, Indonesia)," in *2017 5th International Conference on Information and Communication Technology (ICoICT)*, IEEE, May 2017, pp. 1–4. doi: 10.1109/ICoICT.2017.8074654.
- [23] S. Kodama *et al.*, "Predictive ability of current machine learning algorithms for type 2 diabetes mellitus: A meta-analysis," *J Diabetes Investig*, vol. 13, no. 5, pp. 900–908, May 2022, doi: 10.1111/jdi.13736.