

# Klasifikasi Pasien Penderita Diabetes Melitus Tipe Dua Menggunakan Metode Analisis Diskriminan *Hybrid* Algoritma Genetika

Ramadhana Dio Gradiana dan Irhamah  
Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Institut Teknologi Sepuluh Nopember (ITS)  
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia  
*e-mail*: irhamahn@statistika.its.ac.id

**Abstrak**—Kebiasaan dan gaya hidup masyarakat Indonesia yang tidak sehat menyebabkan 8,4 juta jiwa menderita diabetes melitus sehingga menempatkan Indonesia pada peringkat keempat dunia dalam jumlah penderita diabetes melitus. Dari seluruh pasien diabetes melitus tersebut, 90% diantaranya menderita diabetes melitus tipe dua. Salah satu cara penanganannya adalah dengan pemberian obat. Namun, kondisi gula darah setiap pasien berbeda dalam merespon obat yang diberikan. Ada kondisi di mana gula darah pasien telah teregulasi dan ada pula kondisi di mana gula darah belum mampu teregulasi oleh obat yang diberikan. Atas kedua kondisi tersebut, dapat dilakukan klasifikasi yang dapat mengelompokkan pasien-pasien ke dalam kondisi gula darah teregulasi dan belum teregulasi salah satunya dengan analisis diskriminan. Namun, seiring berkembangnya ilmu pengetahuan, metode diskriminan dapat di *hybrid* dengan metode lain salah satunya algoritma genetika. Oleh karena itu dalam penelitian ini dilakukan perbandingan ketepatan klasifikasi dari analisis diskriminan dan *hybrid* analisis diskriminan-algoritma genetika. Hasilnya, ketepatan klasifikasi dari metode *hybrid* analisis diskriminan-algoritma genetika lebih tinggi dibandingkan dengan metode analisis diskriminan.

**Kata Kunci**—Algoritma genetika, analisis diskriminan, diabetes melitus, ketepatan klasifikasi.

## I. PENDAHULUAN

DIABETES melitus atau biasa dikenal dengan penyakit *Diabetes* adalah penyakit yang terjadi karena kadar gula dalam darah meningkat akibat gangguan sistem metabolisme dalam tubuh, dimana organ pankreas tidak mampu memproduksi hormon insulin sesuai kebutuhan tubuh. Pada tahun 2000, jumlah penderita diabetes melitus di Indonesia mencapai 8,4 juta jiwa. Artinya sekitar 3,5% penduduk Indonesia menderita penyakit ini. Jumlah ini tentunya akan terus bertambah dan diprediksikan pada tahun 2030 jumlah penderita diabetes di Indonesia akan mencapai angka 21,3 juta penderita [1]. Organisasi kesehatan dunia *World Health Organization* (WHO) mengklasifikasikan diabetes melitus ke dalam tiga tipe yaitu tipe satu, tipe dua dan tipe lain. Namun, 90% diantara penderita diabetes melitus terklasifikasikan ke dalam tipe dua [2].

Penanganan yang dapat dilakukan terhadap pasien diabetes melitus tipe dua adalah dengan memberikan obat dan melakukan terapi pada pasien. Namun, terkadang gula darah pasien belum mampu teregulasi oleh obat yang diberikan sehingga pihak medis harus mengganti dosis atau mengganti obat lain.

Penanganan pada kondisi ini akan lebih mudah jika pihak medis melakukan klasifikasi terhadap pasien diabetes melitus tipe dua yang gula darahnya telah teregulasi oleh obat yang diberikan dan belum teregulasi oleh obat yang telah diberikan.

Pengklasifikasian pasien diabetes melitus pada kedua kondisi gula darah pasca pemberian obat dapat dilakukan dengan menggunakan analisis diskriminan dan *hybrid* analisis diskriminan algoritma genetika. Pengklasifikasian diharapkan akan berguna sebagai informasi tambahan dalam bidang kesehatan khususnya dalam pengklasifikasian pasien diabetes melitus. Salah satu penelitian yang menggunakan algoritma genetika adalah penelitian dengan judul Implementasi Algoritma Genetika Pada Struktur *Backpropagation Neural Network* untuk Klasifikasi Kanker Payudara. Penelitian tersebut menyimpulkan bahwa metode *Neural Network* yang dioptimalkan parameternya menggunakan algoritma genetika menghasilkan nilai rata-rata akurasi yang tinggi [3]. Oleh karena itu, untuk memperoleh akurasi klasifikasi yang lebih tinggi, peneliti menggunakan analisis diskriminan *hybrid* algoritma genetika untuk mengklasifikasikan pasien diabetes melitus tipe dua.

## II. TINJAUAN PUSTAKA

### A. Analisis Diskriminan

Analisis diskriminan adalah suatu teknik analisa statistik untuk mengklasifikasikan objek ke dalam suatu kelompok tertentu berdasarkan variabel bebasnya [4]. Analisis diskriminan adalah suatu metode yang dapat menghasilkan pemisahan yang terbaik antara berbagai macam populasi [5]. Dalam analisis diskriminan asumsi yang harus dipenuhi antara lain.

1. Variabel yang diamati menyebar secara normal multivariat (*multivariate normality*)
2. Semua kelompok populasi mempunyai matrik varian kovarian yang sama
3. Terdapat perbedaan rata-rata antar kelompok.

Setelah semua asumsi dipenuhi, kita dapat melanjutkan analisis diskriminan. Didalam analisis diskriminan terdapat fungsi diskriminan yang biasa disebut dengan *canonical root*. Nilai prediksi dari fungsi diskriminan disebut dengan skor diskriminan [6]. Persamaan umum dari fungsi diskriminan untuk setiap objek adalah sebagai berikut.

$$Z_{jk} = a + w_1x_{1k} + w_2x_{2k} + \dots + w_nx_{nk} \quad (1)$$

dimana

- $Z_{jk}$  = skor diskriminan dari fungsi ke- $j$  objek ke- $k$
- $a$  = konstanta
- $w_i$  = koefisien diskriminan untuk variabel independen ke  $i$
- $x_{ik}$  = variabel independen ke- $i$  objek ke- $k$

Setelah model dibentuk, akurasi klasifikasi dapat dihitung. Jika menggunakan skor diskriminan dan *optimal cutting score*, prosedur pengklasifikasian adalah sebagai berikut: sampel ke- $k$  masuk ke dalam grup A jika  $Z_k < Z_{ct}$ , atau sampel ke- $k$  masuk ke dalam grup B jika  $Z_k > Z_{ct}$ . Nilai  $Z_n$  merupakan skor diskriminan untuk setiap individu ke- $k$ . Sedangkan  $Z_{ct}$  merupakan *critical cutting score value* yang didapatkan dari persamaan berikut [6].

$$Z_{CT} = \frac{Z_A + Z_B}{2} \quad (2)$$

dimana

- $Z_A$  = centroid untuk grup A
- $Z_B$  = centroid untuk grup B

Hasil dari prosedur klasifikasi akan ditampilkan dalam bentuk matriks sebagai berikut.

Tabel 1.  
Matriks klasifikasi

Aktual	Prediksi	
	Negatif	Positif
Negatif	TN	FP
Positif	FN	TP

Persentase ketepatan pengelompokan dapat dihitung dari matriks klasifikasi yang tepat menempatkan hasil prediksi terhadap data aktual dari setiap kelompok. Setelah diperoleh klasifikasi untuk masing-masing observasi, nilai *classification accuracy* (%) dapat dihitung. Ketepatan klasifikasi merupakan suatu ukuran yang menentukan apakah klasifikasi yang telah dilakukan akurat atau tidak.

### B. Analisis Diskriminan Hybrid Algoritma Genetika

Algoritma Genetika dikembangkan oleh John Holland pada tahun 1975. Algoritma genetika adalah metode untuk memecahkan masalah optimasi yang didasarkan pada seleksi alam dan proses yang mendorong evolusi biologis [7]. Pada setiap iterasi, algoritma genetika memilih individu secara acak dari populasi saat ini untuk menjadi ‘orang tua’ atau ‘induk’ dan menggunakan mereka untuk menghasilkan ‘anak’ untuk generasi berikutnya. Algoritma genetika digunakan untuk permasalahan pencarian dengan melakukan minimisasi biaya dan probabilitas yang tinggi untuk mendapatkan solusi global optimum [8].

Di dalam metode algoritma genetika, terdapat 7 komponen yang membangun metode ini. Ketujuh komponen-komponen tersebut adalah sebagai berikut [9].

1. Skema Pengkodean
2. Nilai *Fitness*
3. Seleksi Orang Tua
4. Pindah silang (*Crossover*)
5. Mutasi
6. Elitisme
7. Penggantian Populasi

### C. Diabetes melitus

Diabetes melitus yang juga dikenal di negara Indonesia sebagai penyakit kencing manis adalah kelainan yang disebabkan metabolik oleh banyak faktor dengan gangguan metabolisme karbohidrat, lemak serta protein. *World Health Organization* (WHO) telah mengklasifikasikan jenis penyakit diabetes melitus berdasarkan perawatan dan simtoma. Berdasarkan tipe-tipe diabetes melitus yang ada, jumlah penderita diabetes melitus tipe dua merupakan yang terbanyak dibandingkan dengan tipe-tipe lainnya. WHO melaporkan bahwa 90% penderita diabetes melitus merupakan penderita diabetes melitus tipe dua [10]. Menurut dr. Ipung Paruhito S.Pd., salah satu dokter spesialis penyakit dalam di Surabaya, penanganan pasien diabetes melitus dilakukan dengan pemberian obat dan terapi medis. Namun, pemberian obat kadang tidak langsung cocok dengan pasien karena kondisi tubuh pasien berbeda-beda. Dosis obat dan jenis obat dikatakan telah cocok dengan pasien apabila gula darah telah mampu teregulasi oleh obat yang diberikan. Sebaliknya, apabila gula darah belum mampu teregulasi oleh obat yang diberikan, maka langkah yang harus dilakukan adalah mengganti dosis obat atau mengganti jenis obat yang digunakan untuk penyembuhan. Dua kondisi ini yang dapat di klasifikasikan oleh pihak medis sehingga pihak medis dapat melakukan langkah selanjutnya untuk penanganan pasien. Kondisi dimana gula darah sudah teregulasi oleh obat yang diberikan atau belum dapat diketahui dari gula darah puasa dan gula darah 2 jam *post prandial* [2].

## III. METODOLOGI PENELITIAN

### A. Sumber data

Data yang digunakan dalam penelitian ini yaitu data sekunder yaitu data rekam medik pasien diabetes melitus yang menjalani rawat jalan. Data pasien diambil di klinik penyakit dalam dr. Ipung Paruhito Sp.PD. Data pasien yang menjadi sampel penelitian merupakan data seluruh pasien diabetes melitus tipe dua pada tahun 2012 sampai 2013 yaitu sebanyak 130 sampel dengan rincian 76 data pasien dengan kondisi gula darah telah teregulasi dan 54 data pasien dengan kondisi gula darah belum teregulasi.

### B. Variabel Penelitian

Variabel-variabel penelitian yang digunakan dalam penelitian ini antara lain sebagai berikut.

Variabel dependen:

Kondisi gula darah setelah pemberian obat:

- 1 : Gula darah telah teregulasi
- 2 : Gula darah belum teregulasi

Variabel prediktor:

1. Gula darah 2 jam *post prandial* (mg/dl)
2. Tekanan darah *sistolic* (mm Hg).
3. Tekanan darah *diastolic* (mm Hg)
4. Gula darah puasa (mg/dl)
5. Umur (tahun)

C. Langkah Analisis

Langkah analisis pada penelitian ini antara lain.

1. Melakukan analisis statistik deskriptif
2. Melakukan pengolahan data dengan analisis diskriminan.
  - a. Melakukan uji asumsi analisis diskriminan
  - b. Menentukan fungsi diskriminan
  - c. Membentuk matriks klasifikasi diskriminan
  - d. Menghitung ketepatan klasifikasi
3. Melakukan pengolahan data dengan metode analisis diskriminan *hybrid* algoritma genetika
  - a. Merepresentasikan kromosom awal sebanyak 100 kromosom dimana salah satu individu disisipi dengan parameter yang telah didapatkan dari fungsi diskriminan.
  - b. Menentukan fungsi obyektif yaitu meminimumkan nilai *fitness* yakni kesalahan klasifikasi.
  - c. Melakukan proses seleksi sebanyak N kromosom dari sejumlah P induk yang berasal dari populasi dengan seleksi *roulette wheel*.
  - d. Melakukan proses pindah silang jika nilai bilangan random *r* antara [0,1] yang dibangkitkan kurang dari probabilitas pindah silang (*Ps*).
  - e. Melakukan proses elitisme dimana dua kromosom dengan nilai *fitness* terbaik akan bertahan ke generasi selanjutnya.
  - f. Melakukan pergantian populasi lama dengan generasi baru dengan cara memilih kromosom terbaik berdasarkan nilai *fitness*-nya.
  - g. Melakukan pengecekan terhadap solusi yang telah didapatkan.
4. Membandingkan hasil analisis antara metode analisis diskriminan dan *hybrid* analisis diskriminan algoritma genetika berdasarkan nilai persentase ketepatan klasifikasinya.
5. Menarik kesimpulan dan membuat saran untuk penelitian selanjutnya.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik pasien diabetes melitus tipe dua

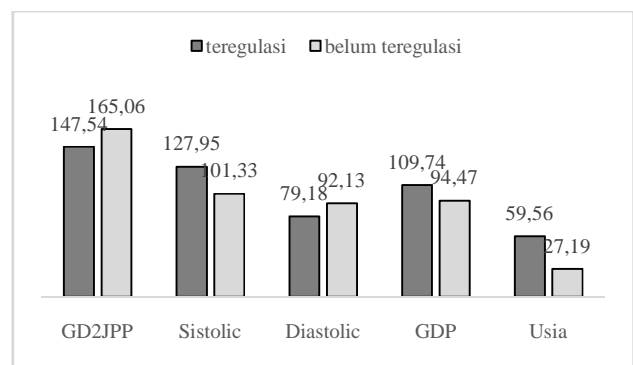
Karakteristik dari pasien diabetes melitus tipe dua yang akandisajikan berupa statistik deskriptif yang terdiri dari rata-rata, simpangan baku, nilai tertinggi dan nilai terendah untuk masing-masing variabel. Statistik deskriptif untuk seluruh pasien diabetes melitus tipe dua yang menjadi sampel disajikan pada tabel berikut.

Tabel 2. Statistik deskriptif untuk seluruh pasien diabetes melitus tipe dua

Variabel Prediktor	Rata-rata	Simpangan baku	Max	Min
Gula darah 2jpp	219,10	113,50	609	72
Sistolic	132,16	19,74	190	90
Diastolic	81,635	9,86	120	50
Gula darah puasa	159,06	80,89	445	38
Umur	56,59	10,53	90	36

Berdasarkan Tabel 2 terlihat bahwa nilai rata-rata gula darah 2 jam *post prandial* sebesar 219,10 mg/dl. Nilai tersebut lebih tinggi dari batas minimumdiagnostik diabetes melitus yang di-tentukan oleh WHO yaitu sebesar 200 mg/dl. Hal serupa terlihat pada rata-rata gula darah puasa yakni sebesar 159,06 mg/dl. Nilai ini juga lebih tinggi dari batas minimum diagnostik yang ditentukan WHO yaitu sebesar 140 mg/dl.

Gambar 1 menunjukkan perbedaan rata-rata untuk setiap variabel pada masing-masing kelompok. Kondisi gula darah te-lah teregulasi tercapai apabila nilai rata-rata antara gula darah 2 jam *post prandial* dan gula darah puasa kurang dari 140 mg/dl. Untuk kondisi gula darah pasien diabetes melitus tipe dua yang belum teregulasi memiliki rata-rata antara gula darah 2 jam *post prandial* dan gula darah puasa lebih besar dari 140 mg/dl. Ber-dasarkan nilai rata-rata pasien diabetes melitus tipe dua dengan kondisi gula darah telah teregulasi pada Gambar 1, diperoleh rata-rata antara gula darah 2 jam *post prandial* dan gula darah puasa sebesar 128,64 mg/dl. Untuk rata-rata antara gula darah 2 jam *post prandial* dan gula darah puasa pada pasien diabetes melitus tipe dua dengan kondisi gula darah belum teregulasi sebesar 274,85 mg/dl.



Gambar 1. Diagram batang rata-rata setiap variabel prediktor untuk masing-masing kelompok

B. Analisis diskriminanda pasien diabetes melitus tipe dua

Sebelum melakukan analisis diskriminan dan melakukan klasifikasi, data yang ada dibagi menjadi dua yaitu data *training* dan data *testing* dengan proporsi 80:20 untuk masing-masing kelompok. Kelompok pertama sebanyak 76 data dibagi menjadi 61 data *training* dan 15 data *testing*. Sedangkan kelompok kedua sebanyak 54 data dibagi menjadi 43 data *training* dan 11 data *testing*.

Pengujian asumsi yang pertama adalah uji distribusi normal multivariat. Berdasarkan pengujian asumsi distribusi normal multivariat didapatkan hasil nilai  $\chi^2_{(5,0,5)}$  sebesar 4,351. Dengan nilai tersebut didapatkan proporsi nilai jarak mahalalanobis  $d_j^2$  yang kurang dari nilai  $\chi^2_{(5,0,5)}$  adalah sebanyak 59,6%. Karena proporsinya disekitar 50%, maka dapat disimpulkan bahwa data yang digunakan memenuhi asumsi distribusi multivariat normal.

Uji asumsi selanjutnya adalah uji kesamaan matriks varian-kovarian. Berdasarkan hasil penghitungan diperoleh nilai sta-tistik uji *Box's M* sebesar 36,510. Apabila nilai ini dibandingkan dengan  $\chi^2_{15;0,05}$  yaitu sebesar 24,996 maka

kesimpulannya  $H_0$  ditolak artinya matriks varian-kovarian dari dua kelompok berbeda. Kondisi ini menyebabkan asumsi kesamaan matriks varian-kovarian tidak terpenuhi. Namun, Marks dan Dunn pada tahun 1974 berpendapat bahwa apabila jumlah sampel yang diambil cukup besar, maka analisis diskriminan tetap dapat dilakukan walaupun matriks varian-kovarian dari kedua kelompok berbeda [11]. Sehingga analisis diskriminan tetap dilanjutkan.

Uji asumsi selanjutnya adalah menguji apakah terdapat perbedaan rata-rata setiap variabel pada masing-masing kelompok. Berdasarkan hasil perhitungan menunjukkan keseluruhan variabel prediktor yang menjadi variabel penelitian signifikan terhadap model. Hal tersebut ditunjukkan dari *pvalue* semua variabel prediktor yang kurang dari nilai  $\alpha=5\%$ . Namun, berdasarkan nilai *Wilks' Lambda*, variabel yang memiliki perbedaan rata-rata yang tinggi pada masing-masing kelompoknya adalah variabel gula darah 2 jam *post prandial*, gula darah puasa, dan usia. Hal tersebut dikarenakan nilai statistik uji *Wilks' Lambda* yang jauh dari nilai 1. Namun, meskipun tidak semua variabel memiliki perbedaan rata-rata yang tinggi antar kelompok, asumsi adanya perbedaan rata-rata pada setiap variabel antar kelompok terpenuhi.

Langkah selanjutnya adalah membentuk fungsi diskriminan dengan metode *stepwise*. Metode *stepwise* yang digunakan adalah metode *backward elimination*. Metode pemilihan variabel akan berhenti ketika tidak ada lagi variabel prediktor yang memiliki  $F_{hitung}$  lebih dari  $F_{tabel} = 3,84$ .

Tabel 3. Metode *backward elimination* dalam pemilihan variabel prediktor

Step	Variabel	$F_{hitung}$
0	gd2jpp	135,270
	sistolik	7,118
	diastolic	9,925
	gdp	116,186
	usia	13,128
	1	sistolik
diastolic		5,194
gdp		5,226
usia		11,704
2	sistolik	3,037
	diastolic	5,893
3	gdp	2,805
	sistolik	0,001
	gdp	2,910

Pada *step nol (step awal)*, seluruh variabel prediktor dimasukkan kedalam model. Variabel prediktor yang terpilih merupakan variabel prediktor yang mengandung nilai  $F_{hitung}$  tertinggi yakni variabel gula darah 2 jam *post prandial* dengan nilai  $F_{hitung}$  sebesar 135,270. Kemudian, pada *step pertama*, variabel prediktor yang mengandung nilai  $F_{hitung}$  tertinggi yaitu variabel usia dengan nilai  $F_{hitung}$  sebesar 11,704. Pada *step kedua* variabel prediktor yang mengandung nilai  $F_{hitung}$  tertinggi yaitu variabel *diastolic* dengan nilai  $F_{hitung}$  sebesar 5,893. Pada *step ketiga* sudah tidak ada lagi variabel prediktor yang mengandung nilai  $F_{hitung}$  lebih dari 3,84 sehingga proses

pemilihan variabel prediktor dengan metode *backward elimination* berhenti. Hasil pemilihan variabel-variabel dan koefisien fungsi diskriminan yang diperoleh dari metode *stepwise* adalah sebagai berikut.

$$Z = -2,912 + 0,013(Gd2jpp) + 0,031(Diastolic) - 0,042(Umur)$$

Nilai-nilai dari fungsi diskriminan merupakan nilai kontribusi setiap variabel terhadap klasifikasi. Kontribusi terbesar terdapat pada variabel usia dengan koefisien sebesar 0,042. Namun nilai kontribusi variabel usia bernilai negatif, artinya setiap kenaikan satu-satuan pada usia akan memperkecil peluang suatu observasi masuk kedalam kelompok 2 atau memperbesar peluang masuk kelompok 1. Pada variabel tekanan darah *diastolic*, nilai koefisiennya sebesar 0,031 dan variabel gula darah 2 jam *post prandial* sebesar 0,013. Keduanya bertanda positif artinya setiap kenaikan satu-satuan terhadap kedua variabel, akan memperbesar peluang suatu observasi untuk masuk kedalam kelompok 2 dan memperkecil peluang untuk masuk kelompok 1. Pengklasifikasian dilakukan dengan membandingkan  $Z_{score}$  dengan *critical cutting score* yaitu sebesar 0,23 artinya suatu observasi akan diklasifikasikan pada kelompok satu apabila  $Z_{score}$  observasi tersebut kurang dari 0,23. Apabila  $Z_{score}$  observasi lebih dari 0,23 maka observasi tersebut masuk ke dalam kelompok kedua. Setelah dilakukan klasifikasi terhadap data *testing* didapatkan hasil sebagai berikut.

Tabel 4. Matriks klasifikasi analisis diskriminan data asli

		Prediksi		Total
		Kelompok 1	Kelompok 2	
Aktual	Kelompok 1	13	2	15
	Kelompok 2	2	9	11

Tabel 4 merupakan tabel hasil klasifikasi dengan menggunakan fungsi yang diperoleh dari analisis diskriminan. Ketepatan klasifikasi dari 26 data *testing* adalah sebesar 84,6% sedangkan kesalahan klasifikasinya sebesar 15,4%. Berdasarkan nilai ketepatan dan kesalahan klasifikasi ini, dapat dikatakan klasifikasi yang telah dilakukan telah baik.

C. Analisis diskriminan hybrid algoritma genetika pasien dia-betes melitus tipe dua

Langkah awal yang dilakukan dalam algoritma genetika adalah menginisialisasi kromosom-kromosom sebanyak 100 buah. Setiap kromosom terdiri dari empat gen. Salah satu kromosom disisipi oleh parameter yang didapat dari analisis diskriminan dan direpresentasikan ke salah satu kromosom dengan ilustrasi berikut.

Kromosom	-2,912	0,013	0,031	-0,042
----------	--------	-------	-------	--------

Gambar 2. Contoh salah satu kromosom yaitu parameter fungsi diskriminan

Kromosom pada Gambar 2 akan menjalani proses seleksi *roulette*, *crossover single point*, dan elitisme dengan kromosom lainnya. Proses-proses tersebut dilakukan untuk mendapatkan parameter fungsi analisis diskriminan yang meminimumkan kesalahan klasifikasi. Namun, sebelum

melalui proses-proses ter-sebut, nilai *fitness* setiap populasi dihitung. Nilai *fitness* dari penelitian ini adalah kesalahan klasifikasi. Berikut adalah nilai-nilai *fitness* untuk setiap populasi.

Tabel 5.  
Nilai *Fitness* untuk masing-masing kromosom untuk data asli

Populasi ke-	Kromosom				Nilai <i>Fitness</i>
1	0.2166	0.5691	-0.026	-2.262	0,076
2	0.2166	0.5691	-0.026	-2.295	0,115
⋮	⋮	⋮	⋮	⋮	⋮
100	0.2965	0.4275	-0.123	-1.841	0,192

Nilai *fitness* terdapat pada setiap kromosom dengan nilai yang berbeda-beda. Nilai *fitness* ini merupakan acuan dalam melakukan proses seleksi, *crossover*, dan elitismesehingga didapatkan nilai *fitness* terbaik. Metode pemilihan kromosom yang pertama adalah seleksi *roulette wheel* yang merupakan salah satu cara dalam penentuan induk kromosom yang akan digunakan pada proses pindah silang. Pemilihan induk dengan seleksi *roulette wheel* melibatkan nilai *fitness* relatif dan *fitness* kumulatif. Berdasarkan penghitungan didapatkan nilai *fitness* relatif dan *fitness* kumulatif sebagai berikut.

Tabel 6.  
Contoh nilai *fitness* relatif dan kumulatif

Populasi ke-	<i>Fitness</i>	<i>Fitness</i> Relatif (rk)	<i>Fitness</i> Kumulatif (kk)	Bilangan Random
1	0,0769	0,002731464	0,002731464	0,499084
2	0,4231	0,01502838	0,017759844	0,009868
⋮	⋮	⋮	⋮	⋮
100	0,5769	0,020491308	1	0,051255

Setelah diperoleh nilai *fitness* relatif dan kumulatif, bilangan random dibangkitkan antara 0 dan 1 sebanyak ukuran populasi yaitu 100. Bilangan random ini digunakan sebagai pembanding untuk setiap individu apakah individu tersebut layak menjadi induk atau tidak. Berdasarkan contoh pada Tabel 6, bilangan random pertama lebih besar dari nilai *fitness* kumulatif pertama ( $k_1$ ) sehingga kromosom pertama tidak bisa menjadi induk. Pada individu kedua terdapat bilangan random lebih kecil dari *fitness* kumulatif kedua ( $k_2$ ) sehingga populasi kedua terpilih untuk bertahan menjadi induk di generasi selanjutnya. Selain menggunakan *roulette wheel* dalam memperoleh solusi terbaik juga bisa dilakukan dengan metode *crossover*. Metode *crossover* yang digunakan pada penelitian ini adalah *single point crossover*. Salah satu contoh pindah silang pada penelitian ini adalah sebagai berikut.

Induk 1	-0,388	0,569	-0,026	-2,262
Induk 2	0,216	0,883	-0,729	-2,262
Anak 1	-0,388	0,569	-0,729	-2,262
Anak 2	0,216	0,883	-0,026	-2,262

Gambar 3. Contoh kromosom yang melakukan *crossover*

Bilangan random dibangkitkan dan dibandingkan dengan probabilitas *crossover*  $P_s=0,8$ . Ternyata, nilai bilangan random lebih kecil dari  $P_s$  saat terletak pada titik kedua, maka kromo-som induk akan menghasilkan anak yang berpindah silang pada titik kedua. Pada Gambar 3 juga menampilkan anak hasil pin-dah silang dari kedua induk. Metode terakhir yang digunakan untuk memperoleh nilai *fitness* terbaik adalah elitisme. Dengan menggunakan elitisme, parameter yang menghasilkan nilai *fitness* terbaik akan dipertahankan untuk generasi selanjutnya. Jumlah populasi terbaik yang dipertahankan untuk generasi se-lanjutnya adalah dua populasi. Pada contoh yang ditampilkan pada tabel 7, populasi yang bertahan ke generasi selanjutnya adalah populasi yang menghasilkan nilai *fitness* terkecil. Sehingga akan diperoleh hasil sebagai berikut.

Tabel 7.  
Pemilihan kromosom yang mengalami elitisme pada generasi pertama

Populasi ke-	Kromosom				<i>Fitness</i>
1	0,472	0,3166	-0,038	-1,232	0,0769
2	0,589	0,316	-0,038	-1,232	0,0769
⋮	⋮	⋮	⋮	⋮	⋮
100	0,690	0,316	-0,038	0,182	0,5769

Pada generasi kedua juga akan melakukan proses elitisme hingga generasi-generasi selanjutnya sampai kondisi nilai *fitness* telah konvergen. Setelah menjalani proses pemilihan *fitness* terbaik, diperoleh hasil nilai *fitness* terbaik untuk metode algoritma genetika sebesar 0,0769 atau ketepatan klasifikasi sebesar  $1-0,0769 = 0,9231$  atau 92,31%. Parameter yang diperoleh dari hasil algoritma genetika antara lain  $a=-2,6393$ ;  $w_1=0,2739$ ;  $w_3=-0,0169$ ; dan  $w_5=-1,0334$ .

D. Perbandingan hasil analisis diskriminan dan analisis diskriminan hybrid algoritma genetika

Setelah melakukan analisis dengan menggunakan kedua metode langkah selanjutnya adalah melakukan perbandingan. Perbandingan ketepatan klasifikasi antara kedua metode adalah sebagai berikut

Tabel 8.  
Perbandingan kesalahan klasifikasi

Metode	Parameter	Ketepatan Klasifikasi
Analisis diskriminan	$a = -2,912$ $w_1 = 0,013$ $w_3 = 0,031$ $w_5 = -0,042$	84,6%
Hybrid analisis diskriminan dan algoritma genetika	$a = -2,6393$ $w_1 = 0,2739$ $w_3 = -0,0169$ $w_5 = -1,0334$	92,31%

Berdasarkan Tabel 8 untuk metode analisis diskriminan, nilai ketepatan klasifikasi diperoleh sebesar 84,6% sedangkan, ketepatan klasifikasi dari metode algoritma genetika adalah 92,31%. Dengan hasil ini, dapat disimpulkan bahwa metode

*hybrid* analisis diskriminan-algoritma genetika merupakan metode yang lebih baik dalam kasus klasifikasi pasien diabetes melitus tipe dua karena menghasilkan ketepatan klasifikasi yang lebih tinggi dibandingkan metode analisis diskriminan.

#### KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan, pengklasifikasian dengan menggunakan metode analisis diskriminan menghasilkan nilai ketepatan klasifikasi sebesar 84,6%. Pengklasifikasian dengan menggunakan metode *hybrid* analisis diskriminan dan algoritma genetika menghasilkan ketepatan klasifikasi sebesar 92,31%. Perbandingan ketepatan klasifikasi antara kedua metode menyimpulkan bahwa metode *hybrid* diskriminan-algoritma genetika menghasilkan ketepatan klasifikasi yang lebih tinggi dibandingkan analisis diskriminan.

Saran yang dapat diberikan untuk penelitian selanjutnya adalah menambah jumlah variabel penelitian yang berhubungan dengan kondisi gula darah pada pasien diabetes melitus tipe dua. Variabel-variabel tersebut antara lain variabel berat badan, tinggi badan, pola hidup, kebiasaan merokok, dan lain-lain.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Institut Teknologi Sepuluh Nopember Surabaya khususnya kepada Jurusan Statistika yang telah memfasilitasi hal-hal yang berhubungan dengan penelitian ini. Penulis juga mengucapkan terima kasih kepada dr. Ipung Paruhito Sp.PD. beserta para perawat yang bersedia membantu dan memberikan data pasien diabetes melitus tipe dua di klinik penyakit dalam yang bertempat di Jl. Ngagel Jaya Tengah 115 Surabaya.

#### DAFTAR PUSTAKA

- [1] Liputan 6. 2011. *Diabetes Melitus Indonesia Duduki Peringkat Ke-4-Dunia*. Diakses pada 30 April 2013, dari <http://health.liputan6.com/>
- [2] Tjokronegoro, A & Utama, H. 1996. *Ilmu Penyakit Dalam*. Jakarta: Gaya Baru.
- [3] Zamani. 2013. *Implementasi Algoritma Genetika Pada Struktur Backpropagation Neural Network Untuk Klasifikasi Kanker Payudara*. Surabaya: ITS.
- [4] Dillon, William R. & Matthew Goldstein. 1984. *Multivariate Analysis Methods and Application*. USA: John Wiley & Sons. Inc.
- [5] Johnson, N. & Wichern, D. 2002. *Applied Multivariate Statistical Analysis, 6th Edition*. New Jersey: Prentice Hall, Englewood Cliffs.
- [6] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. 2010. *Multivariate Data Analysis (7th Edition ed.)*. New Jersey: Pearson Prentice Hall.
- [7] Wikipedia, 2013. Diakses 24 Oktober 2013, dari Algoritma Genetik: [http://id.wikipedia.org/wiki/Algoritma\\_genetik](http://id.wikipedia.org/wiki/Algoritma_genetik)
- [8] Engelbrecht, A.P. 2002. *Computational Intelligence*. South Africa: John Wiley & Sons. Inc.
- [9] Suyanto. 2005. *Algoritma Genetika dalam Matlab*. Yogyakarta : Andi Offset.
- [10] Wikipedia, 2013. Diakses 23 Oktober 2013, dari Diabetes Melitus: [http://id.wikipedia.org/wiki/Diabetes\\_melitus](http://id.wikipedia.org/wiki/Diabetes_melitus)
- [11] Marks, S. and Dunn, O.J. 1974. *Discriminant Functions When Covariance Matrices Are Unequal*. USA : J Amer Stat Assoc.