

DETEKSI RISIKO PENDERITA BUTA WARNA MENURUN BERBASIS POHON KELUARGA DENGAN ALGORITMA GENETIKA

Deny Salvana Ervan¹, Edy Mulyanto²

^{1,2} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
Jl. Nakula I No. 5-11 Semarang

Email : 111201005204@mhs.dinus.ac.id¹, edy.mulyanto@dsn.dinus.ac.id²

Abstrak

Penyakit menurun merupakan salah satu masalah yang cukup serius dalam suatu silsilah keluarga. Hal ini dikarenakan risiko sifat gen tersebut bisa diturunkan pada generasi selanjutnya dalam satu silsilah tersebut. Salah satu dari penyakit menurun adalah buta warna sebagian yaitu tidak bisa membedakan warna-warna tertentu dengan benar. Salah satu yang jadi masalah adalah ketidak tahuan cara untuk mengetahui bahwa buta warna yang diderita merupakan buta warna menurun atau tidak. Perlu dilakukan perhitungan dengan rumus pewarisan sifat untuk mengetahui riwayat risiko buta warna yang bisa diturunkan. Berdasarkan kasus permasalahan yang terjadi, maka dirancang sistem aplikasi yang terotomatisasi sehingga mampu memprediksi risiko penderita buta warna menurun dalam satu silsilah keluarga. Diimplementasikan ke perangkat lunak yang berbasis pohon keluarga dan memanfaatkan algoritma genetika akan mampu memprediksi secara optimal anggota keluarga yang berisiko menurunkan sifat buta warna. Sehingga nantinya akan membantu untuk mengetahui lebih dini risiko pewarisan sifat tersebut oleh semua anggota keluarga dan riwayat sifat buta warna yang dimiliki tiap keluarga dalam satu silsilah. Uji coba yang dilakukan menggunakan variabel prediktor diantaranya jumlah anak berdasarkan jenis kelaminnya, dan sifat gen yang dimiliki orang tua atau induk. Kedua gen induk pada tahap awal akan dilakukan perhitungan persilangan gen dengan perkalian matriks, kemudian dari hasil risiko akan dilakukan prediksi menggunakan algoritma genetika. Nilai fitness dengan tingkat kesalahan yang kecil menunjukkan bahwa uji coba ini dapat diimplementasikan pada obyek penelitian.

Kata kunci : Pohon Keluarga, Penyakit Turunan, Buta Warna, Algoritma Genetika, Prediksi.

Abstract

Inherited disease is a serious problem in a family tree. This is because the nature of the risk gene that can be passed down to the next generation in a lineage. One of inherited diseases are partial color blind that can not distinguish certain colors correctly. One problem is the unawareness of a way to know that the color blindness suffered an inherited color blind or not. Needs to be calculated with the inheritance formula to know the history of color blind risk that can be lowered. Based on case problems occur, then the application of automated systems designed so as to predict the risk of decreased color-blind people in the family tree. Implemented to software-based family tree and utilizing the genetic algorithm will be able to predict the optimal family members who are at risk of lowering properties of color blindness. So that will help to determine early risk of inherited disease by all members of the family and the history of color blindness properties owned by each family in a genealogy. The test is conducted using predictor variables including the number of children based on gender, and the nature of the gene owned by the parents. Both parents genes at an early stage will be calculated by multiplying the matrix of gene crosses, then the results will be predictive risk using genetic algorithm. Fitness value with a small error rate indicates that this trial can be implemented on the object of research.

Keywords: family tree, inherited disease, color blind, genetic algorithm, prediction.

1. PENDAHULUAN

Berbagai macam penyakit yang timbul dan diderita oleh manusia semakin banyak macamnya, seiring dengan perubahan kondisi alam. Salah satunya adalah penyakit keturunan yang diturunkan melalui gen dalam suatu silsilah keluarga. Gen merupakan pembawa sifat yang menentukan dan mengatur perkembangan, metabolisme tubuh, dan menyampaikan informasi genetik dari suatu generasi ke generasi berikutnya [1]. Untuk menentukan kadar risiko terhadap anggota keluarga yang terlibat dalam silsilah keluarga, dapat menggunakan ilmu pewarisan sifat genetika. Faktor genetik dari ayah atau ibu adalah pengaruh utama yang menjadi faktor penentu risiko terjangkitnya penyakit oleh keturunan atau generasi setelahnya. Masalah yang sering terjadi di masyarakat adalah ketidaktahuan apakah penyakit yang diderita termasuk dalam penyakit yang diturunkan ataukah penyakit yang diderita karena gaya hidup manusia itu sendiri. Untuk mengetahui suatu penyakit yang diderita oleh seseorang termasuk dalam penyakit menurun atau tidak, perlu adanya data pada tiap anggota keluarga yang bersangkutan dan merunut melalui pohon silsilah keluarga.

Dengan Pesatnya perkembangan teknologi, maka manusia membutuhkan sistem yang terotomatisasi untuk mendapatkan data yang ingin diperoleh secara cepat. Maka penulis mencoba menyelesaikan masalah tersebut dengan membuat aplikasi ini. Selain menggunakan rumus pewarisan sifat [2], utamanya adalah rumus untuk menentukan risiko penderita buta warna menurun yang didapat dari gen induk, pada penelitian ini penulis menggunakan algoritma genetika untuk mendapat hasil yang optimal. Dengan menerapkan

algoritma genetika dan mengimplementasikannya pada silsilah keluarga yang berbentuk tree, akan mempermudah untuk mengetahui hubungan antar anggota keluarga dan risiko penyakit yang dideritanya. Algoritma Genetika adalah salah satu pendekatan untuk menentukan global optimum yang didasari oleh Teori Evolusi. Secara garis besar langkah dalam prosedur ini dimulai dengan menetapkan suatu set solusi potensial dan melakukan perubahan dengan beberapa iterasi dengan algoritma genetika untuk mencapai solusi terbaik [3]. Tujuan dari penelitian ini adalah membuat aplikasi untuk memprediksi jumlah anak yang berisiko menderita buta warna menurun yang diturunkan dari gen induk dalam satu silsilah keluarga, dengan menggunakan algoritma genetika.

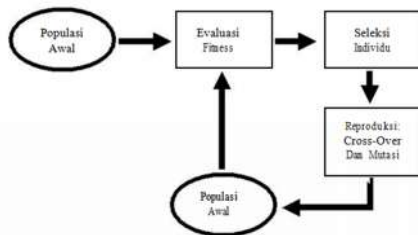
2. METODE

2.1 Algoritma Genetika

Algoritma genetika (*Genetic Algorithm*) adalah salah satu cabang dari kecerdasan buatan. Penemu algoritma genetika adalah John Holland dan dikembangkan oleh muridnya David Goldberg. Algoritma genetika adalah algoritma yang memanfaatkan proses seleksi alamiah yang dikenal dengan proses evolusi. Proses seleksi alamiah ini melibatkan perubahan gen yang terjadi pada individu melalui proses perkembangbiakan. Dalam algoritma genetika ini, proses perkembangbiakan ini menjadi proses dasar yang menjadi perhatian utama, dengan dasar berfikir “Bagaimana mendapat keturunan yang lebih baik”. Algoritma ini bekerja dengan sebuah populasi yang terdiri dari individu-individu, yang masing-masing individu mempresentasikan sebuah solusi yang mungkin bagi persoalan yang ada. Dalam kaitan ini,

individu dilambangkan dengan sebuah nilai *fitness* yang akan digunakan untuk mencari solusi terbaik dari persoalan yang ada [4]. nilai *fitness* adalah nilai yang menyatakan baik tidaknya suatu solusi (individu).

Siklus dari algoritma genetika pertama kali dikenalkan oleh David Goldberg, dimana gambar siklus tersebut dapat dilihat pada gambar 1 [4].



Gambar 1. Siklus Algoritma Genetika

Keterangan [5] :

1. Populasi awal

Proses ini merupakan proses yang digunakan untuk membangkitkan populasi awal secara random sehingga didapatkan solusi awal.

2. Evaluasi fitness

Proses ini merupakan proses untuk mengevaluasi setiap populasi dengan menghitung nilai fitness kromosom dan mengevaluasinya sampai memenuhi kriteria berhenti.

3. Seleksi

Proses seleksi merupakan proses untuk menentukan individu-individu mana saja yang akan dipilih untuk dilakukan crossover.

4. Crossover

Proses crossover ini merupakan proses untuk menambah keanekaragaman string dalam satu populasi.

5. Mutasi

Mutasi merupakan proses mengubah nilai dari satu atau beberapa gen dalam suatu kromosom.

6. Kriteria berhenti

Kriteria berhenti merupakan kriteria yang digunakan untuk menghentikan proses algoritma genetika.

7. Hasil

Hasil merupakan solusi optimum yang didapat algoritma genetika.

2.2 Regresi Berganda

Persamaan regresi yang digunakan untuk membuat taksiran mengenai variabel dependen disebut persamaan regresi estimasi, yaitu suatu formula matematis yang menunjukkan suatu hubungan keterkaitan antara satu atau beberapa variabel yang nilainya sudah diketahui dengan satu variabel lain yang nilainya belum diketahui [5].

Analisis regresi berganda adalah teknik statistik yang dapat digunakan untuk menganalisis hubungan linier satu variabel respon dan beberapa variabel prediktor [6].

3. HASIL DAN PEMBAHASAN

Pada penelitian dan uji coba ini penerapan algoritma genetika terbagi menjadi tahapan sebagai berikut :

1. Menentukan faktor-faktor penunjang nilai fitness.

Dalam proses penyeleksian dari individu, terdapat beberapa faktor yang mempengaruhi seberapa baik individu tersebut. Faktor tersebut adalah jumlah anak berdasarkan jenis kelaminnya, dan sifat gen yang dibawa oleh induknya. Faktor-faktor tersebut nantinya akan diberikan nilai skala atau bobot, sesuai dengan kriteria yang ditentukan.

2. Penentuan nilai faktor

Penskalaan atau pembobotan dari

faktor tergantung dari berapa jumlah pembagian dari tiap faktor. Semakin besar nilainya maka kemungkinan berisiko menurunkan sifat buta warna akan semakin besar.

Tabel 1 : Faktor Sifat Gen

No.	Sifat Gen	Nilai Faktor
1	Wanita normal	3
2	Wanita carrier	4
3	Wanita penderita	1
4	Pria normal	2
5	Pria penderita	5

3. Penentuan nilai fitness

Nilai fitness merupakan suatu ukuran baik tidaknya suatu solusi yang dinyatakan sebagai satu individu, atau dengan kata lain nilai fitness menyatakan nilai dari fungsi tujuan [5].

Dalam permasalahan prediksi risiko buta warna turunan pada penelitian ini, maka nilai fitness bergantung pada faktor jumlah anak berdasar jenis kelaminnya, dan sifat gen yang dimiliki induknya. f adalah nilai fitness. b adalah taksiran parameter regresi, sedangkan x adalah variabel prediktor. x_1 adalah jumlah anak pria, x_2 jumlah anak wanita, x_3 sifat gen ibu, x_4 sifat gen ayah. x_3 dan x_4 ditentukan berdasarkan pembobotan nilai faktor yang dibuat sebelumnya.

$$f1 = -1,267 + 0,348.P + 0,193.L + 0,285.I - 0,095.A$$

$$f2 = 1,267 + 0,652.P - 0,193.L - 0,285.I - 0,095.A$$

$$f3 = 0$$

$$f4 = -3,662 - 0,013.P + 0,0547.L + 0,459.I + 0,684.A$$

$$f5 = 3,662 + 0,013.P + 0,453.L - 0,459.I - 0,684.A$$

Tabel 2 : Hasil pengujian terhadap $f1$

No.	Jumlah Anak		Induk		Risiko		Selisih	
	P	L	I	A	Y_i	Pr	Angka	Dalam %
1	1	2	4	2	0	0	0	0%
2	0	4	4	2	0	0	0	0%
3	2	1	3	5	0	0	0	0%
4	0	5	2	4	0	0	0	0%
5	4	4	4	2	0	1	1	100%
6	6	4	4	2	4	2	2	50%
7	1	3	4	2	1	0	1	100%
8	1	3	4	2	1	0	1	100%
9	0	3	4	2	0	0	0	0%
10	1	2	4	2	1	0	1	100%
11	2	1	4	2	0	0	0	0%
12	2	2	4	2	1	0	1	100%
13	2	1	4	2	0	0	0	0%
14	2	2	4	2	1	0	1	100%
15	1	1	4	2	1	0	1	100%
16	0	2	4	2	0	0	0	0%
17	0	1	4	2	0	0	0	0%
18	0	4	3	5	0	0	0	0%
19	2	2	3	5	0	0	0	0%
20	0	4	1	2	0	0	0	0%
21	3	3	1	2	0	0	0	0%
Rata-rata kesalahan								35,7%

Tabel 3: Hasil pengujian terhadap $f1$

No.	Jumlah Anak		Induk		Risiko		Selisih	
	P	L	I	A	Y ₂	Pr	Angka	Dalam %
1	1	2	4	2	1	0	1	100%
2	0	4	4	2	0	0	0	0%
3	2	1	3	5	2	1	1	50%
4	0	5	2	4	0	0	0	0%
5	4	4	4	2	4	2	2	50%
6	6	4	4	2	2	3	1	50%
7	1	3	4	2	0	0	0	0%
8	1	3	4	2	0	0	0	0%
9	0	3	4	2	0	0	0	0%
10	1	2	4	2	0	0	0	0%
11	2	1	4	2	2	1	1	50%
12	2	2	4	2	1	1	0	0%
13	2	1	4	2	2	1	1	50%
14	2	2	4	2	1	1	0	0%
15	1	1	4	2	0	0	0	0%
16	0	2	4	2	0	0	0	0%
17	0	1	4	2	0	0	0	0%
18	0	4	3	5	0	0	0	0%
19	2	2	3	5	2	1	1	50%
20	0	4	1	2	0	0	0	0%
21	3	3	1	2	3	2	1	33.3%
Rata-rata kesalahan								20,6%

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

1. Berdasarkan penelitian dan pengujian terhadap dataset uji, prediksi risiko penderita buta warna berbasis pohon keluarga menggunakan algoritma genetika mempunyai tingkat kesalahan 35,64%.
2. Dengan mengaplikasikan algoritma genetika untuk proses prediksi anak yang berisiko menderita buta warna menjadi lebih optimal.
3. Perhitungan nilai fitness yang kurang tepat menjadi faktor utama yang mempengaruhi tingkat akurasi prediksi dan optimasi dengan algoritma genetika.

4.2 Saran

1. Penelitian dapat dilanjutkan dengan dataset yang lebih detil dari sisi

variabel terikat dan bebas nya, nilai fitness yang dihasilkan dapat lebih efektif dan menghasilkan nilai yang optimal dengan toleransi kesalahan yang sedikit.

2. Untuk mendapatkan hasil perhitungan nilai fitness dari tiap variabel dependen yang terbentuk akan lebih akurat dan menghasilkan nilai hasil yang detil jika menggunakan tools untuk menghitung nilai regresinya.
3. Untuk mendapatkan hasil yang lebih baik hendaknya dilakukan pembobotan atau penskalaan faktor-faktor yang ada dapat diteliti lebih lanjut .

DAFTAR PUSTAKA

- [1] Furqonita, Deswaty, Biomed M., Seri IPA Biologi SMP Kelas IX,

- Quadra, 2006.
- [2] Sukis Wariyono, Yani Muharomah. Mari belajar ilmu alam sekitar 3: Panduan Belajar IPA terpadu untuk kelas IX SMP/MTs. Jakarta : Pusat Perbukuan, Departemen Pendidikan Nasional, 2008.
- [3] Akhmad Y. Oni Soesanto, “Algoritma Genetik Pada Penyelesaian Akar Persamaan Sebuah Fungsi,” *Jurnal Matematika Murni dan Terapan*, vol.6, no. 2, pp. 47 – 56, Desember 2012.
- [4] “Algoritma Genetika,”. [Online], Available: <http://lecturer.eepis-its.edu/>. [Accesed 11 Februari 2014].
- [5] H. Hazaki. J. L. Buliali. A. Yuniarti, “Aplikasi Untuk Prediksi Jumlah Mahasiswa Pengambil Mata Kuliah Dengan Menggunakan Algoritma Genetika,” Tugas Akhir Dengan Studi Kasus Di Jurusan Teknik Informatika ITS, 2009.
- [6] Anderson. Hair. Tatham. Black, “Multivariate Data Analysis,” Prentice-Hall International, 1998.