

PENENTUAN PENGGUNAAN LULUR DAN MASKER ORGANIK SESUAI DENGAN DIAGNOSA JENIS KULIT WAJAH MENGGUNAKAN METODE *DECISION TREE* ALGORITMA C4.5

Prameswari Reksa Agami¹⁾, Intan Yuniar Purbasari²⁾, Basuki Rahmat³⁾
E-mail : ¹⁾ reksaprameswari@gmail.com , ²⁾ intanyuniar.if@upnjatim.ac.id ,
³⁾ basukirahmat.if@upnjatim.ac.id

¹²³ Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur

Abstrak

Perawatan kulit atau *skincare* merupakan salah satu upaya untuk mendukung kesehatan dan kebersihan kulit, memelihara, merawat dan mempertahankan kondisi kulit, dengan melakukan perawatan kulit maka kebersihan dan kesehatan kulit dapat terjaga. Saat ini masih banyak wanita yang belum mengenali jenis kulitnya sendiri sehingga memiliki resiko kesalahan dalam pemilihan perawatan wajah yang sesuai untuk jenis kulit wajah. Maka dari itu diperlukan adanya sistem pendukung keputusan untuk mengenali jenis kulit yang dimiliki sehingga peluang kesalahan dalam memilih produk perawatan kulit bisa diminimalisir. Metode yang diterapkan pada penelitian ini adalah *decision tree* dengan menggunakan algoritma C4.5. Metode tersebut diterapkan untuk memperoleh keputusan akhir yang didapatkan dari pohon keputusan. yang digunakan sebagai acuan untuk mengetahui dan menentukan jenis kulit wajah yang dimiliki. Dari penelitian yang dilakukan dapat diketahui bahwa terdapat 5 output jenis kulit wajah dan 8 jenis produk perawatan yang mana hasil pada setiap orang akan berbeda sesuai dengan ciri-ciri yang dimiliki. Sistem ini memiliki nilai akurasi 100% dan evaluasi nilai *presisi* dan *recall* pada algoritma C4.5 yang diterapkan bernilai 1.

Kata kunci: *Skincare*, *Decision Tree*, Algoritma C4.5.

1. PENDAHULUAN

Kecantikan merupakan salah satu faktor yang cukup penting bagi kaum wanita dalam menunjang penampilannya. Salah satu tuntutan dari kemajuan zaman yang semakin modern saat ini adalah penampilan fisik yang baik dan menarik, dengan penampilan fisik yang baik maka dapat digunakan sebagai modal untuk bersosialisasi dengan masyarakat umum dan akan mendapatkan simpati yang tinggi dari masyarakat. Ketidakpuasan terhadap penampilan fisik yang dimiliki telah mendorong kaum wanita untuk melakukan perawatan-perawatan fisik demi mendapatkan penampilan fisik yang dianggap menarik [1].

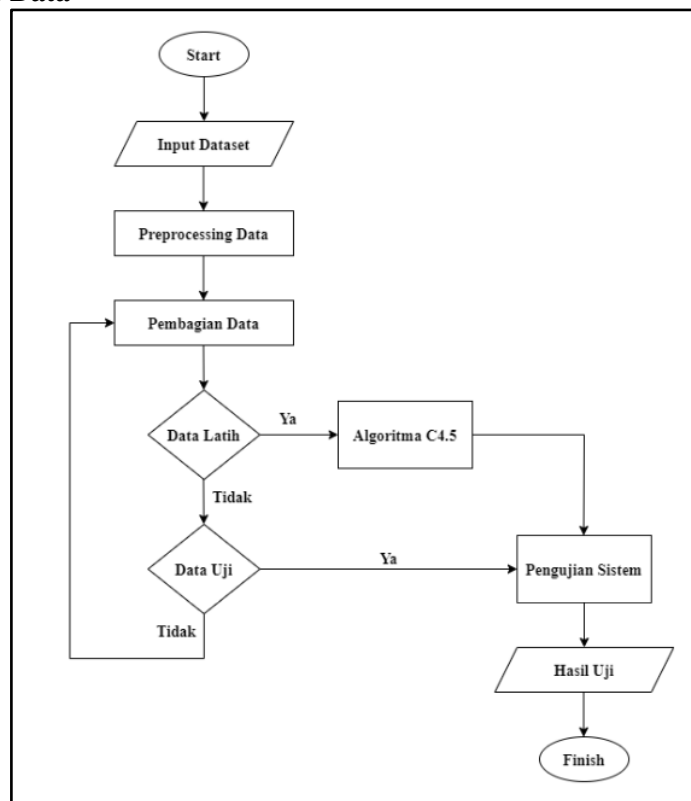
Roro Mendut *Skincare* adalah salah satu perawatan kulit yang memadukan ramuan rempah-rempah tradisional Jawa yang alami dalam komposisi dan pengolahan produk-produknya [2]. Produk dari Roro Mendut *Skincare* yang banyak diminati oleh pelanggan adalah produk lulur wajah dan masker wajah . Saat ini masih banyak pelanggan dari Roro Mendut *Skincare* yang belum mengenali jenis kulit wajah mereka sendiri dikarenakan Roro Mendut *Skincare* belum memiliki fasilitas konsultasi secara langsung dengan dokter melainkan hanya tanya jawab mengenai produk-produk saja melalui admin. Roro Mendut *Skincare* melakukan penjualan produk-produk melalui *marketplace* dan memasarkan melalui distributor serta *reseller*, sehingga pelanggan kesulitan menentukan produk perawatan wajah yang sesuai dengan jenis kulit wajah mereka. Konsultasi *online* dengan

admin dianggap kurang tepat dan tidak efisien dikarenakan pelanggan harus menunggu balasan sesuai dengan antrian *chat* dari admin Roro Mendut *Skincare* dan seringkali terjadi kesalahpahaman antara pelanggan dan admin, hal tersebut juga menjadi salah satu faktor kesalahan pelanggan dalam memilih produk perawatan kulit yang sesuai dengan jenis kulit mereka.

Berdasarkan permasalahan diatas maka dalam penelitian ini penulis akan membuat sebuah sistem dengan judul “Penentuan Penggunaan Lulur Dan Masker Organik Sesuai Dengan Diagnosa Jenis Kulit Wajah Menggunakan Metode *Decision Tree* Algoritma C4.5”. *Output* dari sistem berupa diagnosa jenis kulit wajah termasuk jenis kulit apa dan rekomendasi lulur dan masker wajah yang cocok untuk jenis kulit tersebut. Jenis kulit wajah terbagi menjadi 5 jenis kulit yaitu kulit normal, kulit berminyak, kulit kering, kulit sensitif, dan kulit kombinasi. Dan untuk varian lulur Roro Mendut ada varian lulur spirulina, lulur brotowali, lulur rempah hitam, dan lulur kulit manggis, sedangkan untuk varian masker Roro Mendut ada varian masker spirulina, masker brotowali, masker rempah hitam, dan masker kulit manggis.

2. METODOLOGI

2.1 Analisis Data



Gambar 1. Tahapan Penelitian

Proses atau tahapan pada penelitian yang terdapat di gambar 1 memerlukan *input* atau masukan berupa *dataset* dan akan menghasilkan *output* atau keluaran berupa hasil uji sistem. Hasil uji sistem tersebut dapat menghasilkan hasil diagnosis jenis kulit wajah untuk menentukan masker wajah organik yang sesuai dengan jenis kulit wajah. Dataset yang merupakan masukan pada penelitian tersebut akan diproses terlebih dahulu sebelum dijadikan model pada pohon keputusan dengan menggunakan algoritma C4.5. Sebelum menghasilkan *output* atau keluaran berupa hasil uji, maka akan dilakukan beberapa uji sistem.

Proses selanjutnya adalah tahapan *preprocessing* data yang merupakan tahapan penerapan algoritma C4.5. *Preprocessing* data adalah tahapan yang digunakan untuk mengubah data mentah menjadi data dengan format yang sesuai untuk penerapan teknik *data mining* yang dalam penelitian ini adalah algoritma C4.5. Setelah melalui tahapan *preprocessing* data, maka data akan dibagi terlebih dahulu menjadi data latih dan data uji, kemudian data yang menjadi data latih akan diolah berdasarkan algoritma C4.5.

2.2 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau pengelompokkan pada *dataset*. Dasar dari algoritma C4.5 adalah pembentukan pohon keputusan (*decision tree*). Cabang-cabang dari pohon keputusan merupakan pertanyaan klasifikasi sedangkan untuk daun-daunnya merupakan kelas-kelas atau kelompoknya. Tujuan dari algoritma C4.5 untuk melakukan klasifikasi sehingga hasil dari pengolahan *dataset* berupa pengelompokkan data kedalam kelas-kelas tertentu [3]. Algoritma C4.5 membangun pohon keputusan dari serangkaian data pelatihan yang mirip dengan Algoritma ID3, dengan menggunakan konsep entropi informasi. C4.5 juga dikenal sebagai klasifikasi statistik [4].

Dalam memilih sebuah atribut menjadi akar, dilakukan perhitungan nilai dari atribut yang ada. Nilai *gain* yang paling tinggi dijadikan *root* pada pohon keputusan. Untuk menghitung nilai *gain* digunakan rumus berikut [5] :

$$Gain(A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy \quad (2.1)$$

Keterangan :

- S = Himpunan kasus
- A = Atribut
- n = Jumlah sampel
- |S_i| = Jumlah kasus pada partisi ke-i
- |S| = Jumlah kasus dan S

Sedangkan untuk perhitungan nilai *entropy* dapat dilakukan dengan rumus :

$$Entropy(S_1, S_2, \dots, S_n) = \sum_{i=1}^n P_i * \log_2 (P_i) \quad (2.2)$$

Keterangan :

- S = Himpunan Kasus
- n = Jumlah Sampel
- P_i = Proporsi Kelas

2.3 Confusion Matrix

Confusion matrix merupakan matriks yang digunakan untuk mengetahui dan mengukur performa dari proses klasifikasi dengan menggunakan *machine learning*, yang biasanya memiliki dua *output* atau lebih [6]. Pada penelitian ini penulis menggunakan *confusion matrix multiclass* karena memiliki *output* dari program sebanyak 5 *output*.

Confusion matrix multiclass merupakan *confusion matrix* yang digunakan untuk permasalahan klasifikasi dengan *output* lebih dari dari dua. *Output* yang diuji menggunakan matriks ini merupakan *output* pilihan ganda, yang bukan merupakan output positif dan negatif saja.

Tabel 1. Confusion Matrix Multiclass

		<i>Predicted Number</i>			
		<i>Class 1</i>	<i>Class 2</i>	...	<i>Class n</i>
<i>Actual Number</i>	<i>Class 1</i>	X ₁₁	X ₁₂	...	X _{1n}
	<i>Class 2</i>	X ₂₁	X ₂₂	...	X _{2n}

	<i>Class n</i>	X _{n1}	X _{n2}	...	X _{nn}

Tabel 1 menyajikan jumlah data prediksi dan data sebenarnya dari algoritma pada sebuah sistem yang dengan simbol x dari masing-masing class i. Nilai TP, nilai TN, nilai FP, dan nilai FN dapat dihitung dengan menggunakan persamaan 2.7, 2.8, 2.9, dan 2.10

dibawah ini. Kecuali pada persamaan 2.7, persamaan-persamaan tersebut memiliki syarat $j \neq I$ dan $k \neq I$ [7].

$$TTP_{all} = \sum_{j=1}^n X_{jj} \quad (2.7)$$

$$TFP_i = \sum_{j=1}^n X_{ji} \quad (2.8)$$

$$TFN_i = \sum_{j=1}^n X_{ij} \quad (2.9)$$

$$TTN_i = \sum_{j=1}^n \sum_{k=1}^n X_{jk} \quad (2.10)$$

Dari nilai TTP, nilai TFP, nilai TFN dan nilai TTN yang telah didapatkan dari masing-masing *class* I, maka dapat dilakukan perhitungan untuk nilai *presisi*, *recall* dan *specificity* pada masing-masing *class* I, serta akurasi keseluruhan dapat dihitung dengan persamaan 2.11, 2.12, 2.13, 2.14.

$$\text{Akurasi keseluruhan} = \frac{TTP_{all}}{\text{Jumlah seluruh data}} \times 100\% \quad (2.11)$$

$$P_i = \frac{TTP_{all}}{TTP_{all} + TFP_i} \quad (2.12)$$

$$R_i = \frac{TTP_{all}}{TTP_{all} + TFN_i} \quad (2.13)$$

$$S_i = \frac{TTN_i}{TTN_i + TFP_i} \quad (2.14)$$

3. HASIL DAN PEMBAHASAN

Berikut ini adalah penjelasan mengenai penerapan metode *decision tree* Algoritma C4.5 pada sistem. Penjelasan seperti pada gambar 1 tahapan penelitian.

3.1 Input Dataset

Dataset yang digunakan pada penelitian ini adalah menggunakan data dari pengguna Ella *Skincare* sebagai data uji dan pengguna Roro Mendut *Skincare* sebagai data latih. Data latih berjumlah 100 data dan data latih berjumlah 50 data.

No	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	Hasil
1	T	T	T	T	T	T	T	T	Y	T	T	Y	Y	T	T	Y	T	T	Y	T	T	T	T	Y	Kombinasi
2	T	T	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	T	Y	T	T	T	T	Y	Kombinasi
3	T	Y	Y	Y	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	Normal
4	T	T	T	T	T	T	T	Y	Y	T	T	T	T	T	T	T	Y	T	T	Y	T	T	T	T	Berminyak
5	T	T	T	T	T	T	T	Y	T	T	T	T	Y	T	T	Y	T	T	Y	T	T	T	T	Y	Kombinasi
6	T	Y	Y	Y	T	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	Normal
7	T	T	T	T	T	T	Y	Y	T	T	T	T	T	T	T	T	T	T	Y	T	T	Y	Y	T	Berminyak
8	T	T	T	T	T	T	T	T	Y	T	T	Y	Y	T	T	T	T	T	Y	T	T	T	T	Y	Kombinasi
9	T	T	T	T	T	T	T	Y	T	T	Y	T	T	T	Y	T	T	Y	T	Y	T	T	T	Y	Kombinasi
10	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	Y	T	T	Y	T	Y	T	Berminyak
11	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	Sensitif
12	Y	Y	Y	Y	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	T	T	T	T	Y	Normal
13	T	T	Y	Y	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	Normal
14	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	Sensitif
15	T	Y	Y	Y	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	Y	T	T	T	T	Normal
16	T	T	T	T	T	T	T	Y	T	T	T	Y	T	T	Y	T	T	Y	T	Y	T	T	T	Y	Kombinasi
17	T	T	T	T	T	Y	T	Y	T	T	T	T	T	T	T	T	Y	T	Y	T	Y	T	Y	T	Berminyak
18	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	Y	T	T	Y	T	Y	T	Y	Berminyak
19	T	T	T	T	Y	Y	T	T	T	T	T	T	T	T	Y	Y	T	T	T	Y	T	Y	T	T	Kering
20	T	T	T	T	T	Y	Y	T	T	T	T	T	T	T	T	T	Y	T	Y	T	T	Y	T	Y	Berminyak
21	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	Y	T	T	Y	T	Y	T	Y	Berminyak
22	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	Y	T	T	Y	T	Y	T	Y	Berminyak
...
95	Y	T	T	T	Y	Y	T	T	T	T	T	T	T	T	Y	Y	T	T	T	Y	Y	T	T	T	Kering
96	Y	Y	Y	Y	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	Normal
97	Y	Y	Y	Y	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	Normal
98	T	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	T	T	Y	T	T	T	T	Y	Kombinasi
99	T	T	T	T	T	T	T	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	T	Y	T	T	Sensitif
100	Y	T	T	T	Y	Y	T	T	T	T	T	T	T	T	Y	Y	T	T	T	Y	T	T	T	T	Kering

Gambar 2. Data Uji

Dari data latih pada gambar 4.21 maka akan dikategorikan menjadi tiga kategori yaitu gejala, efek, dan tampilan, untuk memudahkan pada proses perhitungan algoritma C4.5.

Penentuan Penggunaan Lulur Dan Masker Organik Sesuai Dengan Diagnosa Jenis Kulit Wajah Menggunakan Metode *Decision Tree* Algoritma C4.5

No	Gejala	Efek	Tampilan	Hasil
1	G12, G13, G09	G19, 16	G24	Kombinasi
2	G12, G13	G19	G24	Kombinasi
3	G02, G03, G04	G14, G15	G20	Normal
4	G08, G09	G18	G21	Berminyak
5	G13, G09	G19, G16	G24	Kombinasi
6	G02, G03, G04	G14	-	Normal
7	G07, G08	-	G21, G23	Berminyak
8	G12, G13, G09	G19	G24	Kombinasi
9	G12, G09	G19, G16	G24	Kombinasi
10	G07	G18	G21, G23	Berminyak
11	G10, G11	G16	-	Sensitif
12	G01, G02, G03, G04	G15	-	Normal
13	G03, G04	G14, G15	G20	Normal
14	G10, G11	G16	-	Sensitif
15	G02, G03, G04	G14	G20	Normal
16	G13, G09	G19, G16	G24	Kombinasi
17	G07, G08	G18	G21, G23	Berminyak
18	G07	G18	G21, G23	Berminyak
19	G05, G06	G16, G17	G21	Kering
20	G07, G08	G18	G21, G23	Berminyak
...
93	G10, G11	G16	G22	Sensitif
94	G01, G02, G03, G04	G14	G20	Normal
95	G01, G05, G06	G16, G17	G21, G22	Kering
96	G01, G02, G03, G04	G14, G15	G20	Normal
97	G01, G02, G03, G04	G14, G15	G20	Normal
98	G12, G13	G19	G24	Kombinasi
99	G10, G11	G16	G22	Sensitif
100	G01, G05, G06	G16, G17	G21	Kering

Gambar 3. Data Latih Berdasarkan Kategori

3.2 Perhitungan *Entropy* dan *Gain*

Perhitungan *entropy* dilakukan untuk data latih yang telah tersedia untuk mengetahui nilai informasi pada masing-masing atribut dan masing-masing kategorinya. Perhitungan juga dilakukan untuk data latih yang tersedia untuk digunakan sebagai pemilihan atribut yang akan digunakan sebagai akar pada pohon keputusan.

	Jumlah	Normal	Kering	Berminyak	Sensitif	Kombinasi	Entropy	Gain
Total	100	24	5	30	18	23	2.164295856	
Gejala								1.708094626
G01	22	19	3	0	0	0	0.574635698	
G02	23	23	0	0	0	0	0	
G03	23	23	0	0	0	0	0	
G04	19	19	0	0	0	0	0	
G05	5	0	5	0	0	0	0	
G06	5	0	5	0	0	0	0	
G07	24	0	0	24	0	0	0	
G08	25	0	0	25	0	0	0	
G09	33	0	0	17	0	16	0.999337504	
G10	18	0	0	0	18	0	0	
G11	18	0	0	0	18	0	0	
G12	21	0	0	0	0	21	0	
G13	21	0	0	0	0	21	0	
Efek								1.621048489
G14	21	21	0	0	0	0	0	
G15	14	14	0	0	0	0	0	
G16	38	0	5	0	16	17	1.429598335	
G17	5	0	5	0	0	0	0	
G18	29	0	0	29	0	0	0	
G19	23	0	0	0	0	23	0	
Tampilan								1.873351198
G20	19	19	0	0	0	0	0	
G21	32	0	4	28	0	0	0.543564443	
G22	18	0	3	0	15	0	0.650022422	
G23	25	0	0	25	0	0	0	
G24	22	0	0	0	0	22	0	

Gambar 4. Hasil Perhitungan Data Latih Node 1

Pada gambar 4 diketahui bahwa perhitungan *entropy* total data latih menghasilkan nilai *entropy* sebesar 2,164295856 dikarenakan jumlah kelas pada data berjumlah lima kelas. Hasil perhitungan *gain* pada data latih menghasilkan *gain* terbesar pada kategori tampilan yaitu nilai *gain* sebesar 1,873351198, sehingga kategori tampilan akan dijadikan sebagai *node* awal pada pohon keputusan.

Setelah ditemukan nilai *gain* terbesar untuk menentukan *node* awal maka dilakukan langkah yang sama untuk menghasilkan *node* selanjutnya. Perhitungan ulang pada *node* selanjutnya dilakukan dengan menghitung data kategori tampilan dengan atribut G21 dan G22 dikarenakan atribut yang lain memiliki nilai *entropy* sebesar 0 sehingga telah didapatkan hasil jenis kulit. Perhitungan *node* selanjutnya pada atribut G21 dan G22 disajikan pada gambar 5 dan gambar 6.

	Jumlah	Normal	Kering	Beminyak	Sensitif	Kombinasi	Entropy	Gain
Total	32	0	4	28	0	0	0.543564443	
Gejala								0.543564443
G01	3	0	3	0	0	0	0	
G02	0	0	0	0	0	0	0	
G03	0	0	0	0	0	0	0	
G04	0	0	0	0	0	0	0	
G05	4	0	4	0	0	0	0	
G06	4	0	4	0	0	0	0	
G07	22	0	0	22	0	0	0	
G08	23	0	0	23	0	0	0	
G09	15	0	0	15	0	0	0	
G10	0	0	0	0	0	0	0	
G11	0	0	0	0	0	0	0	
G12	0	0	0	0	0	0	0	
G13	0	0	0	0	0	0	0	
Efek								0.543564443
G14	0	0	0	0	0	0	0	
G15	0	0	0	0	0	0	0	
G16	4	0	4	0	0	0	0	
G17	4	0	4	0	0	0	0	
G18	27	0	0	27	0	0	0	
G19	0	0	0	0	0	0	0	

Gambar 5. Hasil Perhitungan Data Latih *Node* G21

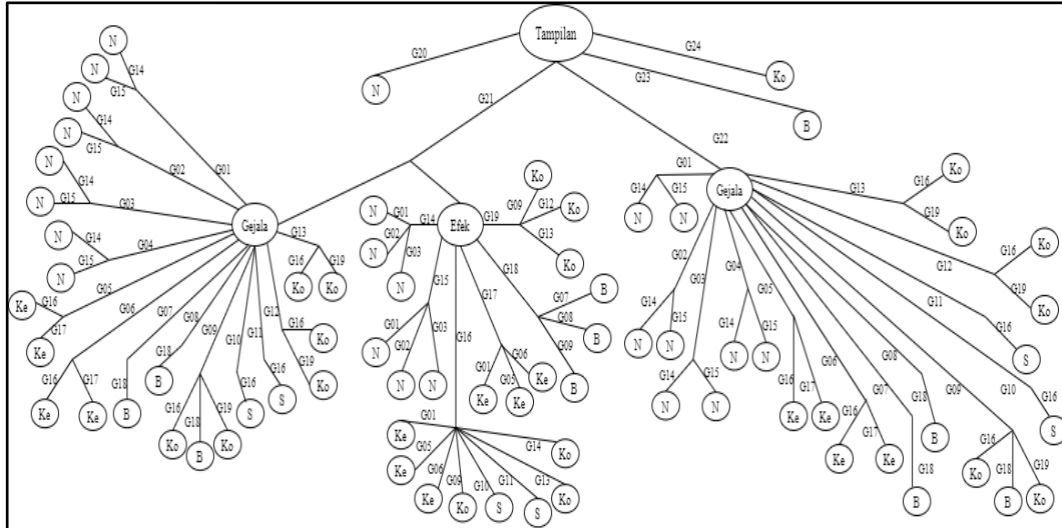
Pada perhitungan *node* 2 atribut G21 diketahui nilai *gain* dari kategori gejala dan efek memiliki nilai yang sama yaitu sebesar 0,543564443. Dengan kasus seperti ini maka solusi yang akan dilakukan adalah dengan membangun *node* dari dua pohon keputusan alternatif dengan masing-masing atribut dari kategori gejala dan efek.

	Jumlah	Normal	Kering	Beminyak	Sensitif	Kombinasi	Entropy	Gain
Total	18	0	3	0	15	0	0.650022422	
Gejala								0.650022422
G01	2	0	2	0	0	0	0	
G02	0	0	0	0	0	0	0	
G03	0	0	0	0	0	0	0	
G04	0	0	0	0	0	0	0	
G05	3	0	3	0	0	0	0	
G06	3	0	3	0	0	0	0	
G07	0	0	0	0	0	0	0	
G08	0	0	0	0	0	0	0	
G09	0	0	0	0	0	0	0	
G10	15	0	0	0	15	0	0	
G11	15	0	0	0	15	0	0	
G12	0	0	0	0	0	0	0	
G13	0	0	0	0	0	0	0	
Efek								0.031167079
G14	0	0	0	0	0	0	0	
G15	0	0	0	0	0	0	0	
G16	16	0	3	0	13	0	0.69621226	
G17	3	0	3	0	0	0	0	
G18	0	0	0	0	0	0	0	
G19	0	0	0	0	0	0	0	

Gambar 6. Hasil Perhitungan Data Latih *Node* G22

Pada perhitungan *node* 3 atribut G22 menghasilkan nilai *gain* terbesar pada kategori gejala yaitu sebesar 0,650022422. Sehingga kategori gejala dijadikan akar untuk perhitungan pohon keputusan *node* selanjutnya. Perhitungan *node* selanjutnya adalah dengan melakukan perhitungan nilai *entropy* dan *gain* pada setiap atribut pada kategori gejala. Perhitungan *entropy* dan *gain* dilakukan sampai dengan menemukan nilai keputusan akhir untuk membangun pohon keputusan.

3.3 Pohon Keputusan



Gambar 7. Pohon Keputusan

Pada tahapan ini adalah proses mengubah dari nilai *gain* tertinggi menjadi sebuah pohon keputusan. Hasil dari perhitungan data latih terdapat empat *node* yang dijadikan sebagai akar acuan pohon keputusan. Hasil dari perhitungan *gain* pada sistem didapatkan kategori tampilan yang memiliki nilai *gain* terbesar, maka dijadikan *node* awal sebagai pembentukan pohon keputusan.

Kemudian pada kategori tampilan memiliki atribut yang belum memiliki nilai pasti untuk keputusan akhir yaitu G21 dan G22. Kemudian dihitung kembali *entropy* dan *gain* untuk atribut G21 dan G22. Untuk atribut G21 menghasilkan nilai *gain* yang sama antara kategori gejala dan kategori efek, sehingga kategori gejala dan efek dijadikan *node* akar selanjutnya untuk pohon keputusan alternatif dikarenakan memiliki nilai *gain* yang sama. Kemudian pada atribut G22 menghasilkan nilai *gain* tertinggi pada kategori gejala sehingga kategori gejala dijadikan *node* akar selanjutnya untuk mendapatkan pohon keputusan dengan nilai pasti mendapatkan keputusan akhir.

3.4 Uji Akurasi

Proses selanjutnya adalah melakukan uji coba pohon keputusan dan aturan akhir terhadap data uji yang tersedia dari proses pembagian data pada langkah awal penelitian. Data uji terdiri dari 50 data yang memuat masing-masing atribut dan hasil. Uji akurasi dihitung menggunakan persamaan 2.11.

$$\begin{aligned} \text{Akurasi keseluruhan} &= \frac{\text{TTP}_{\text{all}}}{\text{Jumlah seluruh data}} \times 100\% \\ &= \frac{50}{50} \times 100\% = 100\% \end{aligned}$$

3.5 Evaluasi Algoritma

Pada tahap ini, hasil dari pohon keputusan dan data uji akan diukur melalui perhitungan *presisi* dan *recall* dengan menggunakan *confusion matrix*. *Confusion matrix* yang akan digunakan dalam pengujian dan pengukuran kinerja pada penelitian ini adalah

confusion matrix multiclass, yaitu *confusion matrix* yang memuat hasil klasifikasi dengan lebih dari dua kelas. Alasan penggunaan *confusion matrix multiclass* pada pengukuran kinerja pada penelitian ini adalah karena *output* kelas yang dihasilkan oleh *dataset* dan sistem berjumlah 5 kelas.

Tabel 2. Confusion Matrix Hasil Uji Coba

		Predicted Number				
		Normal	Kering	Berminyak	Sensitif	Kombinasi
Actual Number	Normal	9	0	0	0	0
	Kering	0	4	0	0	0
	Berminyak	0	0	18	0	0
	Sensitif	0	0	0	2	0
	Kombinasi	0	0	0	0	17

Berdasarkan hasil uji coba perbandingan data uji dan hasil *iterasi* aturan pohon keputusan di subbab sebelumnya, tabel 2 merepresentasikan hasil dari angka tiap kelas yang ada pada data uji dengan angka tiap kelas yang menjadi hasil prediksi dengan aturan pohon keputusan yang sudah dibuat. Berdasarkan data dalam *confusion matrix* diatas, maka dapat dihitung nilai *true positive* (TP), *true negative* (TN), *false positive* (FP), dan juga *false negative* (FN). Setelah mendapatkan seluruh angka yang dibutuhkan, dapat dihitung nilai *presisi* dan *recall* yang akan menjadi tolak ukur hasil penelitian menggunakan algoritma C4.5. Tahapan pertama adalah mencari nilai TP, TN, FP dan FN sesuai dengan persamaan 2.7 sampai 2.10 pada *confusion matrix* yang telah dibuat.

$$\begin{aligned}
 TTP_{all} &= \sum_{j=1}^n X_{jj} \\
 TTP_{all} &= X_{nn} + X_{kk} + X_{bb} + X_{ss} + X_{koko} \\
 TTP_{all} &= 9 + 4 + 18 + 2 + 17 = 50
 \end{aligned}$$

Tabel 3. Nilai FP, FN, dan TN

Nilai	Normal	Kering	Berminyak	Sensitif	Kombinasi
FP	0	0	0	0	0
FN	0	0	0	0	0
TN	41	46	32	48	33

Setelah dilakukan perhitungan nilai FP, FN, dan TN dari semua kelas, tahap selanjutnya adalah melakukan perhitungan untuk nilai *presisi* dan *recall*. Untuk mengetahui nilai dari variabel tersebut, sama seperti perhitungan sebelumnya, perhitungan akan dilakukan untuk tiap-tiap kelas sesuai dengan persamaan 2.12 dan 2.13.

Tabel 4. Hasil Presisi dan Recall

Nilai	Normal	Kering	Berminyak	Sensitif	Kombinasi
Presisi	1	1	1	1	1
Recall	1	1	1	1	1

Dengan hasil pengukuran kinerja dengan serangkaian perhitungan yang telah dilakukan, dapat disimpulkan bahwa implementasi algoritma C4.5 untuk sistem diagnosa jenis kulit wajah membuahkan hasil yang baik dengan akurasi 1 atau 100%. Tentunya hasil ini dipengaruhi oleh beberapa faktor, yakni pengolahan data, pembagian data latih dan data uji, serta keragaman data yang terdapat pada *dataset* yang disediakan untuk penelitian algoritma C4.5.

4. KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dan saran yang bisa diambil dari penelitian yang telah dilakukan. Kesimpulan diambil secara garis besar dari hasil penelitian,

sedangkan saran diberikan untuk pengembangan dan penyempurnaan penelitian yang lebih baik lagi.

4.1 Kesimpulan

Berdasarkan hasil analisis dan pengujian terhadap sistem diagnosa jenis kulit wajah dengan menggunakan algoritma C4.5 ini dapat diambil beberapa kesimpulan, diantaranya :

1. Sistem diagnosa jenis kulit wajah untuk merekomendasikan produk perawatan yang sesuai dengan jenis kulit wajah dapat dibangun dengan algoritma C.45
2. Dari hasil pengukuran kinerja algoritma, diperoleh hasil untuk akurasi, *presisi*, dan *recall* dengan nilai sama yaitu 1 atau 100%.
3. Pada studi algoritma C4.5 untuk kelas berjumlah lebih dari 2, jumlah data, rasio pembagian data dan keragaman data adalah faktor yang berpengaruh dalam pengujian algoritma dan pengukuran kinerja pada algoritma C4.5 yang telah dirancang.

4.2 Saran

Dari serangkain proses dan tahapan penelitian yang telah dilakukan, diberikan saran untuk membuat penelitian ini menjadi lebih baik lagi di masa depan, diantaranya :

1. *Dataset* yang digunakan dalam penelitian ini dapat dikembangkan menjadi *dataset* yang lebih baik lagi, baik dari segi jumlah data latih maupun data uji, rasio pembagian data latih dan data uji maupun keragaman pada *dataset* yang digunakan. Jumlah atribut yang digunakan bisa dikembangkan lagi sehingga lebih detail ketika dijadikan sebagai *dataset*.
2. Sistem diagnosa jenis kulit untuk merekomendasikan produk perawatan wajah yang sesuai dengan menggunakan algoritma C4.5 ini dapat dikembangkan lebih baik lagi dalam bentuk *platform* lainnya seperti *mobile*.

5. DAFTAR RUJUKAN

- [1] L. E. Sari and W. Hadikurniawati, "Sistem Pendukung Keputusan Penentuan Skincare Untuk Kulit Wajah Menggunakan Metode Decision Tree (Studi Kasus : Kosmetik Wardah)," pp. 978–979, 2020.
- [2] "Roro Mendut Skincare." [Online]. Available: <http://roromendutskincare.com/>. [Accessed: 20-Feb-2021].
- [3] E. Elisa, "Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti," *J. Online Inform.*, vol. 2, no. 1, p. 36, 2017.
- [4] M. A. Nishara Banu and B. Gomathy, "Disease forecasting system using data mining methods," *Proc. - 2014 Int. Conf. Intell. Comput. Appl. ICICA 2014*, pp. 130–133, 2014.
- [5] B. Sugara, D. Widyatmoko, B. S. Prakoso, and D. M. Saputro, "Penerapan Algoritma C4.5 untuk Deteksi Dini Autisme Pada Anak," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2018, no. Sentika, pp. 87–96, 2018.
- [6] D. Iskandar and Y. K. Suprpto, "Perbandingan Akurasi Klasifikasi Tingkat," *J. Ilm. NERO*, vol. 2, no. 1, pp. 37–43, 2015.
- [7] A. Maulana and A. A. Fajrin, "Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor," *Klik - Kumpul. J. Ilmu Komput.*, vol. 5, no. 1, p. 27, 2018.