

## DATA MINING UNTUK MENENTUKAN PARAMETER KUNCI SYARAT TUMBUH OPTIMAL BEBERAPA TANAMAN INDUSTRI DENGAN MENGGUNAKAN METODE C4.5

*Data Mining to Determine Key Parameters Requirements for Optimal Growth of Some Industrial Plants by Using the C45 Method*

**Aneke Rintiasti, dan Raflin Fauzi**

Baristand Industri Surabaya, Jl. Jagir Wonokromo 360 Surabaya,  
e-mail: [anekerintiasti@gmail.com](mailto:anekerintiasti@gmail.com)

### **Abstract**

*The plantation industry cannot be separated from land management. Rule of expert knowledge information is obtained through input parameters of land conditions to determine which industrial plants are most suitable to be planted with the Data Mining process. Industrial plantations studied included clove land, robusta coffee land, arabica coffee fields, sugar cane fields, cocoa fields and tea fields. Every industrial plant is used 15 (fifteen) parameters for optimal growth requirements. The values of the growth conditions are data that will be processed in Data Mining by mapping it into the decision tree using the WEKA application using the C45 method. In this decision tree a simpler rule will be obtained, so that it can be used to compile questions and provide answers like an expert. The decision tree was successfully obtained for 6 (six) industrial plants with Kappa Statistic worth 1 which means it can be trusted. The questions generated by the decision tree have been simplified from 15 (fifteen) parameters to only 5 (five) key parameters. Key parameters according to the purpose of the decision tree simplify the rule, because experts do not need to ask all parameters to get conclusions.*

**Keywords:** *Data Mining, Weka, Optimal Growing Requirements, Industrial Plant*

### **Abstrak**

*Industri perkebunan tidak dapat dipisahkan dengan pengolahan lahan. Rule informasi pengetahuan pakar didapatkan melalui inputan parameter kondisi lahan untuk menentukan tanaman industri yang paling cocok ditanam dengan proses Data Mining. Lahan tanaman Industri yang diteliti terdiri dari lahan cengkeh, lahan kopi robusta, lahan kopi arabika, lahan tebu, lahan kakao dan lahan teh. Setiap tanaman industri digunakan 15 (lima belas) parameter syarat tumbuh optimal. Nilai-nilai syarat tumbuh tersebut merupakan data yang akan diolah dalam Data Mining dengan dipetakan kedalam pohon keputusan menggunakan aplikasi WEKA dengan metode C45. Pada pohon keputusan ini akan didapatkan rule yang lebih sederhana, sehingga dapat digunakan untuk menyusun pertanyaan dan memberikan jawaban layaknya seorang pakar. Pohon keputusan berhasil didapatkan untuk 6 (enam) tanaman industri dengan Kappa Statistic bernilai 1 yang berarti dapat dipercaya. Pertanyaan yang dihasilkan oleh pohon keputusan telah disederhanakan dari 15 (lima belas) parameter menjadi hanya 5 (lima) parameter kunci. Parameter kunci sesuai dengan tujuan pohon keputusan menyederhanakan rule, karena pakar tidak perlu menanyakan semua parameter untuk mendapatkan kesimpulan.*

**Kata Kunci :** *Data Mining, Weka, Syarat Tumbuh Optimal, Tanaman Industri*

### **PENDAHULUAN**

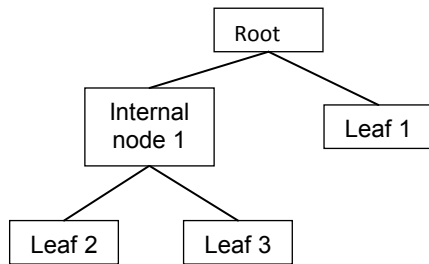
Nilai ekspor non migas tahun 2015 mengalami penurunan sebesar 9,77 persen dibandingkan tahun 2014 yang mencapai USD 145,96 Miliar. (Lakip Kemenperin, 2015). Nilai ekspor produk non migas salah satunya berasal dari industri perkebunan tentunya tidak dapat dipisahkan dengan pengolahan lahan perkebunan. Pengetahuan para pelaksana perkebunan dalam menentukan karakteristik lahan sesuai jenis

tanaman industri yang akan ditanamnya, tentunya akan meminimalisasi berbagai permasalahan. Tujuan penelitian ini adalah mendapatkan parameter kunci dan nilainya syarat tumbuh optimal tanaman industri.

Salah satu metode data mining yang umum digunakan adalah pohon keputusan (Konsep Pohon Keputusan, 2014). Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan *rule*. Konsep dari

pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan.

Pohon keputusan merupakan himpunan aturan *IF, THEN*. Setiap *path* dalam *tree* dihubungkan dengan sebuah aturan, dimana premis terdiri atas sekumpulan *node-node* yang ditemui, dan kesimpulan dari aturan terdiri atas kelas yang terhubung dengan *leaf* dari *path*.



**Gambar 1.** Konsep Dasar Pohon Keputusan

Bagian awal dari pohon keputusan ini adalah titik akar (*root*), sedangkan setiap cabang dari pohon keputusan merupakan pembagian berdasarkan hasil uji, dan titik akhir (*leaf*) merupakan pembagian kelas yang dihasilkan.

Algoritma Pohon Keputusan C4.5 atau *Classification Version 4.5* adalah pengembangan dari algoritma ID3. Oleh karena pengembangan tersebut, algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan algoritma ID3. Pada algoritma ID3 Hitung *Entropy* dan *Information gain* dari setiap atribut dengan menggunakan rumus:

$$Entropy(S) = -P+ \log_2 P+ - P- \log_2 P-$$

Dimana:

S = ruang (data) *sample* yang digunakan untuk *training*.

P+ = jumlah yang bersolusi positif (mendukung) pada data *sample* untuk kriteria tertentu.

P- = jumlah yang bersolusi negatif (tidak mendukung) pada data *sample* untuk kriteria tertentu.

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{nilai}(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

Dimana:

S= ruang (data) *sample* yang digunakan untuk *training*.

A= atribut.

V= suatu nilai yang mungkin untuk atribut A.

Nilai(A) = himpunan yang mungkin untuk atribut A.

|Sv| = jumlah *sample* untuk nilai V.

|S| = jumlah seluruh *sample* data.

*Entropy*(Sv) = *entropy* untuk *sample-sample* yang memiliki nilai V.

Tujuan dari pengukuran nilai *information gain* adalah untuk memilih atribut yang akan dijadikan cabang pada pembentukan pohon keputusan. Pilih atribut yang memiliki nilai *information gain* terbesar.

Algoritma C4.5 yang merupakan suksesor dari ID3 menggunakan *gain ratio* untuk memperbaiki *information gain*, dengan rumus *gain ratio*:

$$Gain Ratio(S, A) = \frac{Gain(S, A)}{Split Info(S, A)}$$

Dimana:

S = ruang (data) *sample* yang digunakan untuk *training*.

A = atribut.

Gain(S,A) = *information gain* pada atribut A

SplitInfo(S,A) = *split information* pada atribut A

Atribut dengan nilai *Gain Ratio* tertinggi dipilih sebagai *atribut test* untuk simpul. Dengan *gain* adalah *information gain*. Pendekatan ini menerapkan normalisasi pada *information gain* dengan menggunakan apa yang disebut sebagai *split information*. *SplitInfo* menyatakan *entropy* atau informasi potensial dengan rumus:

$$Split Info(S, A) = - \sum_{i=1}^i \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

Dimana:

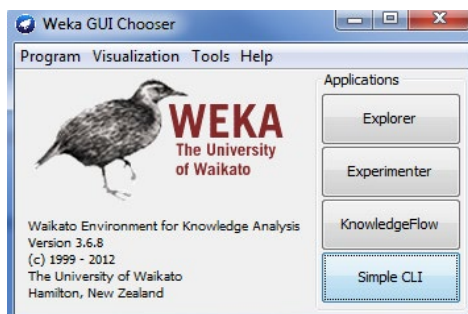
S = ruang (data) *sample* yang digunakan untuk *training*.

A= atribut.

Si= jumlah *sample* untuk atribut

WEKA (*Waikato Environment for Knowledge Analysis*) adalah *tools open source* (berlisensi GPL) yang berisi kumpulan

algoritma *machine learning* dan data *preprocessing* (Yulyanto, 2014). WEKA dapat digunakan mulai dari tahap *preprocessing*, proses, evaluasi sampai dengan visualisasi. Berbagai algoritma dapat dibandingkan untuk memilih yang terbaik untuk masalah yang akan dipecahkan. Basis weka adalah Java sehingga bersifat *multiplatform* dan dapat digunakan sebagai *library* dalam aplikasi Java. Weka menyediakan tiga *interface* untuk mengolah dataset.



Gambar 3. Aplikasi WEKA

#### a. Explorer

Digunakan untuk mencari algoritma yang paling cocok untuk data. Semua data *load* ke memori sehingga dapat cepat diproses, tapi hanya dapat digunakan untuk data dengan jumlah terbatas.

#### b. Eksperimenter

Digunakan untuk mencari parameter yang cocok. Mirip dengan *explorer* tetapi prosesnya dapat diotomatisasi. Eksperimen ukuran besar (*multi machine* dengan RMI) dapat dilakukan dengan *interface* ini.

#### c. Knowledgeflow

Digunakan untuk memproses data *stream*. Mendukung *incremental learning*. *Simple CLI (Command Line Interface)* menggunakan *command line* untuk *interface*.

## METODOLOGI

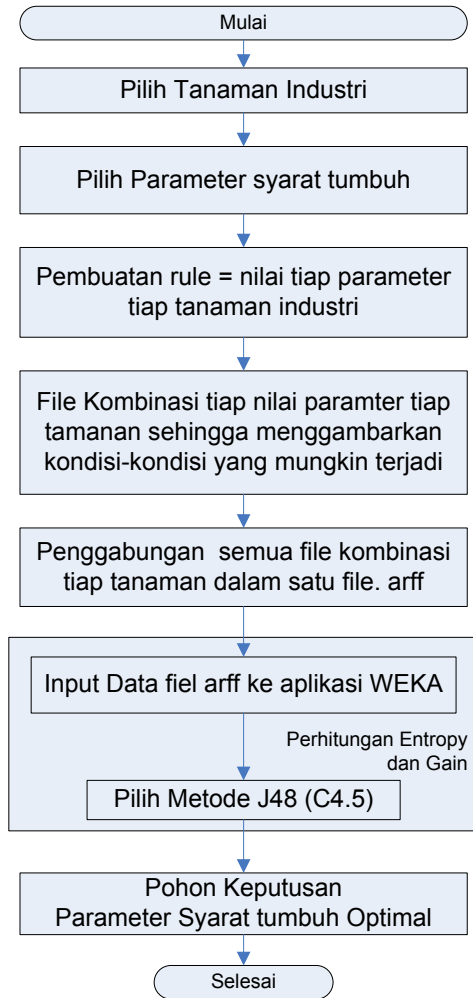
Rismayanti (2016) menggunakan algoritma C4.5 untuk membantu dalam menentukan mahasiswa yang layak dan tidak layak menerima beasiswa dari STTH Medan. Dalam melakukan analisa dan perancangan untuk menentukan penerima beasiswa, maka ada 4 atribut prediktor yang digunakan yaitu IPK, Semester, Penghasilan Orang Tua dan Jumlah Tanggungan Orang Tua, sedangkan untuk atribut prediksi ada dua yaitu Diterima dan Ditolak. Hasil implementasi dilakukan dengan menggunakan *software Data Mining Rapid Miner*.

Haryanto (2017) mengidentifikasi calon pegawai suatu perusahaan. Algoritma C4.5 dapat digunakan untuk melakukan prediksi dan klasifikasi terhadap calon pegawai yang berpotensi untuk masuk ke dalam perusahaan dengan cara membuat pohon keputusan berdasarkan data data yang sudah ada dan melakukan prediksi terhadap calon pegawai baru yang ingin masuk ke perusahaan. Berdasarkan metode pengukuran akurasi *ten-fold cross validation* telah didapatkan hasil pengukuran tingkat keberhasilan prediksi calon pegawai baru sebesar 71%.

Elisa (2017) menggunakan metode C4.5 untuk menganalisa faktor-faktor kecelakaan kerja konstruksi Proyek PT. Arupadhatu Adisesanti sebagai berikut: 1. Pekerja dan Cara Kerja, 2. Lingkungan Tempat Kerja, 3. Alat Pelindung Diri.

Metode penelitian ini dapat dilihat pada Gambar 4. Data karakteristik syarat tumbuh berasal dari Buku Literatur.

Data karakteristik tersebut diolah dalam WEKA untuk mengetahui parameter kunci syarat tumbuh optimal.



**Gambar 4.** Metode Penelitian

**HASIL DAN PEMBAHASAN**

Tanaman Industri yang digunakan dalam Sistem Pakar ini ada 6 (enam) tanaman industri seperti yang tertuang dalam tabel 1.

**Tabel 1. Tanaman Industri**

Kode Tanaman	Nama Tanaman	Varietas
1	Cengkeh	-
2	Teh	
3	Kopi	Robusta
4	Kopi	Arabika
5	Tebu	
6	Cokelat	

Dari setiap tanaman industri akan digunakan 15 (lima belas) parameter syarat tumbuh optimal sebagai target pencocokan dengan karakteristik lahan. Parameter-parameter tersebut dapat dilihat pada tabel 2.

**Tabel 2. Syarat Tumbuh Optimal**

Kode	Nama Parameter	satuan
1	Curah Hujan	mm
2	Jenis Tanah	
3	pH	
4	Ketinggian	dpl
5	Kecepatan Angin	
6	Kelembapan Udara	
7	Intensitas Penyerapan	
8	Suhu Udara	Celcius
9	Bulan kering	
10	Air tanah/drainase	
11	Kedalaman efektif	
12	Bahan organik	
13	Unsur Hara	
14	Aerasi	
15	Kemiringan	derajat

**Nilai Syarat Tumbuh**

Berikut ini adalah nilai parameter syarat tumbuh optimal yang harus dipenuhi untuk setiap tanaman industri. Berdasarkan (Rukmana, 2016) didapatkan syarat tumbuh optimal untuk tanaman industri Cengkeh.

Tabel 3. Tabel *Rule* Tanaman Cengkeh

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	1	1	1500-2500
2	1	2	Latosol
3	1	2	Andosol
4	1	3	05-Jul
5	1	4	200-600
6	1	5	tidak ada angin kencang
7	1	6	60-70%
8	1	7	60%
9	1	8	24-28 celcius
10	1	9	<2
11	1	10	>3m
12	1	11	Tidak dipersyaratkan
13	1	12	Tidak dipersyaratkan
14	1	13	Tidak dipersyaratkan
15	1	14	Tidak dipersyaratkan
16	1	15	Tidak dipersyaratkan

Tabel 4. *Rule* Tanaman Kopi Arabika

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	4	1	2000-3000
2	4	2	Latosol
3	4	2	Andosol
4	4	2	Posolik Merah Kuning
5	4	3	6.5
6	4	4	800-1500
7	4	5	Tidak dipersyaratkan
8	4	6	Tidak dipersyaratkan
9	4	7	Tidak dipersyaratkan
11	4	8	17-21
12	4	9	3

**Tabel 4. Rule Tanaman Kopi Arabika (Continue)**

13	4	10	Tidak dipersyaratkan
14	4	11	100 cm
15	4	12	3%
16	4	13	Kalium
17	4	14	baik
18	4	15	Tidak dipersyaratkan

**Tabel 5. Rule Tanaman Kopi Robusta**

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	3	1	2000-3000
2	3	2	Latosol
3	3	2	Andosol
4	3	2	Posolik Merah Kuning
5	3	3	5.5
6	3	4	400-800
7	3	5	Tidak dipersyaratkan
8	3	6	Tidak dipersyaratkan
9	3	7	Tidak dipersyaratkan
10	3	8	21-24
11	3	9	3
12	3	10	Tidak dipersyaratkan
13	3	11	100 cm
14	3	12	3%
15	3	13	Kalium
16	3	14	baik
17	3	15	Tidak dipersyaratkan

Berdasarkan Rukmana (2015) didapatkan syarat tumbuh tanaman teh.

**Tabel 6. Rule Tanaman Teh**

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	2	1	1500 - 2000
2	2	2	Latosol
3	2	2	Andosol

Tabel 6. *Rule* Tanaman Teh (Continue)

4	2	2	Posolik Merah Kuning
5	2	3	4,5 - 5,6
6	2	4	800-1200
7	2	5	Tidak dipersyaratkan
8	2	6	> 70 %
9	2	7	> 6 jam/hari
10	2	8	20-22
11	2	9	< 2
12	2	10	Tidak dipersyaratkan
13	2	11	< 35 %
14	2	12	Tidak dipersyaratkan
15	2	13	Tidak dipersyaratkan
16	2	14	Tidak dipersyaratkan
17	2	15	Tidak dipersyaratkan

Tabel 7. *Rule* Tanaman Kakao

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	6	1	1200-2500 mm
2	6	2	Regosol
3	6	3	6,0 -7,5
4	6	3	5,6 -7,2
5	6	4	≤ 800 m
6	6	5	tidak ada angin kencang
7	6	6	60-70%
8	6	7	20%
9	6	8	30-32 celcius
10	6	8	18-21 celcius
11	6	10	>3m

Berdasarkan Rukmana (2013) didapatkan syarat tumbuh optimal tanaman tebu.

Tabel 8. *Rule* Tanaman Tebu

Kode Rule	Nama Tanaman	Nama Parameter	Nilai Parameter
1	6	1	1000 - 1300
2	6	2	Regosol

Tabel 8. Rule Tanaman Tebu (C)

3	6	2	Alluvial
4	6	2	Grumosol
5	6	2	Latosol
6	6	2	Posolik Merah Kuning
7	6	3	6 - 7,5
8	6	4	0 - 500
9	6	5	< 10 km /jam
10	6	6	40 % - 60 %
11	6	7	12 -14 jam
12	6	8	24 - 25
13	6	8	> 34
14	6	9	> 3
15	6	10	Tidak dipersyaratkan
16	6	11	< 8 %
17	6	12	Tidak dipersyaratkan
18	6	13	Tidak dipersyaratkan
19	6	14	Tidak dipersyaratkan
20	6	15	Tidak dipersyaratkan

**Pembuatan Pohon Keputusan**

Untuk mendapatkan pohon keputusan dengan metode C45 atau *Tree J48* disini digunakan aplikasi data *mining* WEKA. Langkah pertama adalah membuat file yang berisi informasi relasi, atribut dan data. Data yang digunakan adalah data kombinasi dari semua kemungkinan parameter Gambar 5.

```

cengkeh,1500,latosol,5,200,dibawah_10_
km_per_jam,60,60,24,2,>3,?,?,??,??
cengkeh,1500,andosol,5,200,dibawah_10_km_
per_jam,60,60,24,2,>3,?,?,??,??
cengkeh,1600,latosol,5,200,dibawah_10_km_per_
jam,60,60,24,2,>3,?,?,??,??
cengkeh,1600,andosol,5,200,dibawah_10_
km_per_jam,60,60,24,2,>3,?,?,??,??
cengkeh,1700,latosol,5,200,dibawah_10_
km_per_jam,60,60,24,2,>3,?,?,??,??
cengkeh,1700,andosol,5,200,dibawah_10_km_
per_jam,60,60,24,2,>3,?,?,??,??
.... dst
    
```

Gambar 5. Persiapan Pengolahan Data



```
kopiarabika,2000,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2100,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2200,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2300,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2400,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2500,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
... dst
```

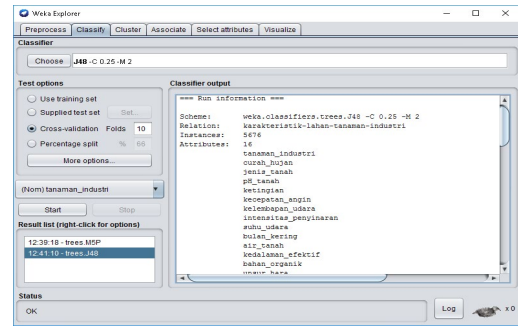
```
kopirobusta,2000,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
kopirobusta,2100,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
kopirobusta,2200,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
kopirobusta,2300,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
kopirobusta,2400,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
kopirobusta,2500,latosol,5.5,800,?,?,?,21,3,?,100,7%,K,baik,?
..... dst
```

**Gambar 6.** File Arf sebagai masukan file pada WEKA

Data-data dalam Gambar 6 akan digabungkan dalam file berformat .arf untuk menjadi masukan aplikasi WEKA dan akan dihitung *entropy* dan *gain* sehingga didapatkan suatu pohon keputusan. Gambar 6 adalah contoh file.arf. Gambar 7 adalah hasil input Data pada *software* WEKA.

```
@relation karakteristik-lahan-tanaman-industri
@attribute tanaman_industri {cengkeh, kopi_arabika, kopi_robusta, kakao, tebu, coklat}
@attribute curah_hujan numeric
@attribute jenis_tanah {aluvial, regosol, grumosol, latosol, andosol, posolik_merah_kuning}
@attribute pH_tanah numeric
@attribute ketinggian numeric
@attribute kecepatan_angin {dibawah_10_km_per_jam, diatas_10_km_per_jam}
@attribute kelembapan_udara numeric
@attribute intensitas_penyinaran numeric
@attribute suhu_udara numeric
@attribute bulan_kering numeric
@attribute air_tanah numeric
@attribute kedalaman_efektif numeric
@attribute bahan_organik numeric
@attribute unsur_hara {N, P, K}
@attribute aerasi numeric {baik, tidak_baik}
@attribute kemiringan numeric
@data
```

```
kopiarabika,2000,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2100,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2200,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2300,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2400,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
kopiarabika,2500,latosol,6.5,800,?,?,?,17,3,?,100,7%,K,baik,?
...dst
```

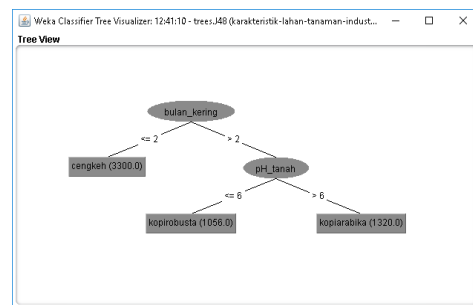


**Gambar 7.** Input data dalam aplikasi

Pohon keputusan dibuat setelah menghitung *entropy* dan *gain* tertinggi. Nilai *gain* tertinggi dijadikan sebagai *node* 1. Kemudian dihitung *entropy* dan *gain* dari tanaman yang tersisa. Nilai *gain* tertinggi menjadi cabang dari *node* 1. Begitu seterusnya sampai semua tanaman industri terpetakan dalam *tree*.

Dalam penelitian ini diketahui bahwa nilai *entropy* dan *gain* berubah tergantung berapa banyak tanaman industri yang dianalisa. Hal ini terlihat pada beberapa percobaan.

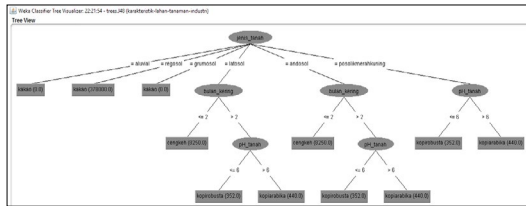
Pohon keputusan ketika memiliki data untuk 3 tanaman industri yaitu Cengkeh, Kopi Robusta, Kopi Arabika, Gambar 8, *rule* hanya menggunakan 2 parameter kunci yaitu bulan kering dan pH tanah. Hal ini berarti seorang pakar hanya cukup menanyakan nilai nilai bulan kering kemudian pH Tanah tersebut untuk menentukan tanaman terbaik apa yang dapat ditanam.



**Gambar 8.** Pohon Keputusan untuk 3 jenis tanaman

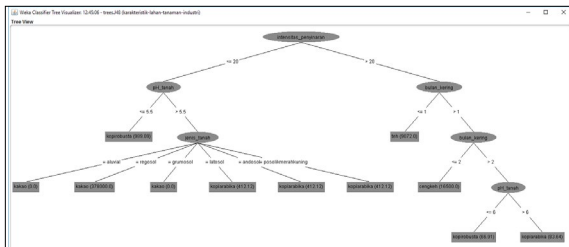
Pohon keputusan mengalami perubahan ketika memiliki data untuk 4 tanaman industri (Cengkeh, Kopi Robusta, Kopi Arabika dan Kakao) Gambar 9. Parameter

kunci jika data 4 tanaman industri diolah adalah jenis tanah, bulan kering, pH Tanah.



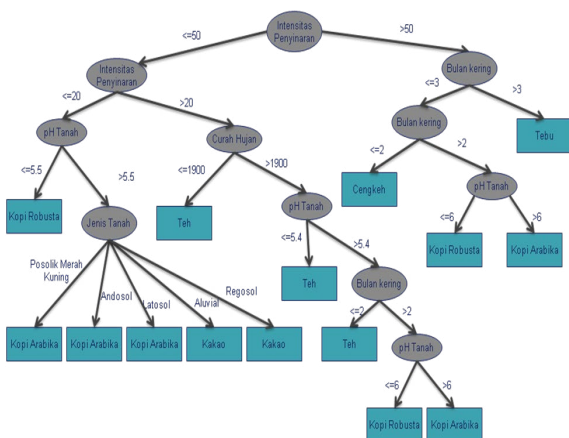
**Gambar 9.** Pohon keputusan untuk 4 jenis tanaman

Pohon keputusan untuk lima tanaman industri (Cengkeh, Kopi Robusta, Kopi Arabika, Kakao dan Teh) pada Gambar 10. Parameter Kunci menjadi intensitas penyinaran, pH tanah, bulan kering dan jenis tanah.



**Gambar 10.** Pohon keputusan untuk 5 jenis tanaman

Pohon keputusan untuk 6 tanaman industri (Cengkeh, Kopi Robusta, Kopi Arabika, Kakao, Teh dan Tebu) pada Gambar 11. Parameter Kunci menjadi intensitas penyinaran, pH tanah, bulan kering, jenis tanah dan curah hujan. Parameter kunci yang muncul untuk 6 tanaman industri adalah atribut dengan nilai *entropy* dan *gain* tertinggi.



**Gambar 11.** Pohon keputusan untuk 6 jenis tanaman

Pohon keputusan Gambar 11, dapat dilihat bahwa yang dievaluasi pertama kali adalah parameter Intensitas penyinaran, jika intensitas penyinaran  $\leq 50$ , maka dilakukan evaluasi lagi apakah  $\leq 20$  atau  $>20$ . Jika Intensitas Penyinaran  $\leq 20$ , maka dievaluasi pH Tanah, jika  $\leq 5.5$  maka tanaman yang cocok adalah Kopi Robusta. Jika pH tanah  $> 5.5$ , maka harus dievaluasi Jenis Tanahnya, jika jenis tanah Posolik Merah Kuning, maka tanaman paling cocok adalah Kopi Arabika. Begitu seterusnya.

```

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances    987106
100 %
Incorrectly Classified Instances    0
0 %
Kappa statistic                    1
Mean absolute error                 0
Root mean squared error             0
Relative absolute error             0 %
Root relative squared error         0 %
Total Number of Instances          987106
    
```

**Gambar 12.** Kappa Statistic bernilai 1

Untuk memperkuat keyakinan kebenaran maka dihitung nilai *kappa statistic*. Pada Gambar 12 adalah nilai Kappa yang dihasilkan terhadap *tree* untuk 6 tanaman industri adalah bernilai 1 yang berarti meyakinkan.

Dengan telah ditemukannya parameter kunci syarat tumbuh, petani atau pengambil kebijakan di dunia pertanian dan perkebunan, dapat memilih jenis tanaman industri yang dibudidayakan sehingga dapat menghasilkan produk yang optimal.

Parameter kunci ini juga dapat dirujuk sebagai acuan dalam pembuatan Sistem Pakar (Merlina, 2012).

## KESIMPULAN

1. Pohon keputusan berhasil didapatkan untuk 6 (enam) tanaman industri dengan *kappa statistic* bernilai 1 yang berarti dapat dipercaya.
2. Pertanyaan yang dihasilkan oleh pohon keputusan telah menyederhanakan dari 15 (lima belas) parameter menjadi 5

(lima) parameter kunci yaitu: intensitas penyinaran, pH tanah, jenis tanah, curah hujan, bulan kering. Parameter kunci ini muncul sesuai dengan tujuan pohon keputusan menyederhanakan *rule*, karena pakar tidak perlu menanyakan semua parameter untuk mendapatkan kesimpulan.

#### DAFTAR PUSTAKA

- Erlin, Elisa, "Analisa dan Penerapan Algoritma C45 dalam Data Mining untuk mengidentifikasi factor-faktor penyebab kecelakaan kerja konstruksi PT. Arupadhatu Adisesanti", 2017, Sistem Informasi Universitas Putera Batam, JOIN, ISSN 2527-9165.
- Ferdiyan Fandy, Hansung Seng, "Penerapan Algoritma C4.5 untuk memprediksi penerimaan calon pegawai barudi PT. WISE", 2017, Jatsi, Vol 3 no 2 Maret 2017.
- Kementerian Perindustrian, 2015, Laporan Kinerja Kementerian Perindustrian.
- Konsep Pohon Keputusan ID3 dan C4.5, <http://pohonkeputusan.com/konsep-pohon-keputusan-id3-dan-c45/>, 2014, diakses pada 28 Oktober 2016.
- Merlina, Nita M.Kom, "Perancangan Sistem Pakar, Ghalia Indonesia", 2012, Bogor.
- Rismayanti, "Implementasi Algoritma C45 untuk menentukan penerima beasiswa di STT Harapan Medan", 2016, Jurnal Media Infotama Vol. 12 No. 2, September 2016.
- Rukmana, H. Rahmat dkk, "Untung Selangit dari Agribisnis The", 2015, Lily Publisher, Yogyakarta
- Rukmana, H. Rahmat dkk, "Untung Selangit dari Agribisnis Cengkeh", 2016, Lily Publisher, Yogyakarta
- Rukmana, H. Rahmat, "Untung Selangit dari Agribisnis Tebu", 2015, Lily Publisher, Yogyakarta
- Yulyanto, Mochamad Tri, "Aplikasi Pohon dalam Sistem Pakar/ Expert System Klasifikasi Hewan dengan Bahasa Prolog", 2014, Program Studi Teknik Informatika, Institut Teknologi Bandung
- Wibisono, Yudi, "Modul Praktikum WEKA, Ilmu Komputer Universitas Pendidikan Indonesia", 2013, (cs.upi.edu).