



**PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI TINGKAT KEGANASAN
HAMA PADA TANAMAN PADI
(Studi Kasus : Dinas Pertanian Kabupaten Kerinci)**

Sularno¹⁾, Putri Anggraini²⁾

¹Sistem Informasi, Universitas Dharma Andalas, Jl. Sawahan No.103, Padang
email: soelarno@unidha.ac.id

²Sistem Informasi, Universitas Dharma Andalas, Jl. Sawahan No.103, Padang
email: putri_anggraini@unidha.ac.id

Submitted: 02-11-2017, Reviewed: 29-11- 2017, Accepted 29-11-2017

<http://doi.org/10.22216/jsi.v3i2.2779>

Abstract

Rice crop is a food plant susceptible to pests. The introduction of pests that attack is a very important first step to support the success in control efforts. Pests of rice plants can be an obstacle for farmers to be able to increase production. because the pest can damage the rice plants to make crop failures. With the holding of classification on rice pests using C4.5 algorithm is expected farmers can immediately know the type of rice pest and the level of ferocity. So it is expected to be able to handle the air properly, in order to avoid damage and crop failure.

Keywords: *Pests, Rice Crops, Algorithms C4.5, Classification*

Abstrak

Tanaman padi merupakan tanaman pangan yang rentan terserang hama. Pengenalan terhadap jenis hama yang menyerang merupakan langkah awal yang sangat penting untuk menunjang keberhasilan dalam usaha pengendaliannya. Hama tanaman padi tersebut dapat menjadi kendala bagi petani untuk bisa meningkatkan produksi. karena hama tersebut dapat merusak tanaman padi hingga membuat gagal panen. Dengan diadakannya klasifikasi pada hama padi menggunakan algoritma C4.5 diharapkan para petani dapat segera mengetahui jenis hama padi dan tingkat keganasannya. Sehingga diharapkan mampu menangani hawa secara tepat, agar tidak terjadi kerusakan dan gagal panen.

Keywords: *Hama, Tanaman Padi, Algoritma C4.5, Klasifikasi*

PENDAHULUAN

Data mining adalah kegiatan meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan (Liliana Swastina, 2013). Untuk menerapkan konsep *data mining* banyak cara yang dapat digunakan diantaranya adalah menggunakan konsep beberapa algoritma yang ada pada *data mining*. Algoritma C4.5 adalah salah satu metode untuk

membuat *decision tree* berdasarkan *training* data yang telah disediakan (Sarah Faradillah, 2013). Metode C4.5 adalah metode klasifikasi data dengan teknik pohon keputusan yang terkenal dan disukai karena memiliki kelebihan-kelebihan (Rina Dewi Indah sari, 2014). Pemanfaatan atau penerapan *data mining* tersebut menjadi salah satu elemen yang terpenting untuk dicermati perkembangannya. Tujuan dari semua itu hanyalah untuk memudahkan pemakai dalam melaksanakan



pekerjaan dan mengkomidasi setiap bentuk kebutuhannya.

Berdasarkan hal diatas, maka penulis sangat tertarik untuk menentukan tingkat keganasan hama pada tanaman padi menggunakan bahasa sehari-hari yang nantinya dalam pengambilan keputusan akan lebih mudah dipahami guna menjadikan sebagai pembelajaran untuk kedepannya dalam menghasilkan padi yang bermutu. Penulis mencoba untuk memberikan informasi dalam menentukan tingkat keganasan hama pada tanaman padi di Dinas Pertanian Kabupaten Kerinci. Selama ini informasi yang diberikan dalam tingkat keganasan hama pada tanaman padi hanya bisa dilihat berdasarkan penurunan produksi hasil panen padi dan gagal panen, yang menunjukkan kualitas padi tidak bermutu.

Knowledge Discovery in Databases (KDD)

Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Benni R Siburian, 2014). *Knowledge Discovery in Database* (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *Knowlegde Discovery in Database* melibatkan hasil proses *data mining* (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami (Tampubolon, Saragih, Reza, 2013). Fayyad (1996) dan Dunham (2003) menggunakan istilah *Knowledge Discovery in Databases* (KDD) untuk menunjuk pada keseluruhan proses pencarian pengetahuan dalam kumpulan data jumlah besar. Mereka menggolongkan *data mining* sebagai salah satu langkah dalam proses KDD dikarenakan oleh penerapannya terhadap suatu algoritma spesifik dalam mencari pola-pola (model) dalam data (Azhari dan Anshori, 2009).

Algoritma C4.5

Algoritma C4.5 adalah salah satu metode untuk membuat *decision tree* berdasarkan *training* data yang telah disediakan. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi data *kontinu*, dan *pruning* (Alimancon Sijabat, 2015). Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan yang merupakan pengembangan dari algoritma ID3, algoritma tersebut digunakan untuk membentuk pohon keputusan. Pohon keputusan dianggap sebagai salah satu pendekatan yang paling populer, dalam klasifikasi pohon keputusan terdiri dari sebuah node yang membentuk akar, node akar tidak memiliki inputan (Ari Puspita dan Mochamad Wahyudi, 2015). Tiga prinsip kerja algoritma C4.5 pada tahap belajar dari data adalah :

1. Pembuatan pohon keputusan.
Obyektif dari algoritma pohon keputusan adalah mengkonstruksi struktur data pohon (dinamakan pohon keputusan) yang dapat digunakan untuk memprediksi kelas dari sebuah kasus atau record baru yang belum memiliki kelas. C4.5 mengkonstruksi pohon keputusan dengan strategi *divide* dan *conquer*. Pada awalnya, hanya dibuat node akar dengan menerapkan algoritma *divide* dan *conquer*. Algoritma ini memilih pemecahan kasus-kasus yang terbaik dengan menghitung dan membandingkan *gain ratio*, kemudian pada node-node yang terbentuk di level berikutnya, algoritma *divide* dan *conquer* akan diterapkan lagi. Demikian seterusnya sampai terbentuk daun-daun. Algoritma C4.5 dapat menghasilkan pohon keputusan, dengan simbol kotak menyatakan simpul dan elips menyatakan daun.
2. Pemangkasan pohon keputusan dan evaluasi (opsional).
Karena pohon yang dikonstruksi dapat berukuran besar dan tidak mudah



“dibaca”, C4.5 dapat menyederhanakan pohon dengan melakukan pemangkasan berdasarkan nilai tingkat kepercayaan (*confidence level*). Selain untuk pengurangan ukuran pohon, pemangkasan juga bertujuan untuk mengurangi tingkat kesalahan prediksi pada kasus (*record*) baru.

3. Pembuatan aturan-aturan dari pohon keputusan (opsional).

Aturan-aturan dalam bentuk *if-then* diturunkan dari pohon keputusan dengan melakukan penelusuran dari akar sampai ke daun. Setiap *node* dan syarat pencabangnya akan diberikan di *if*, sedangkan nilai pada daun akan menjadi ditulis di *then*. Setelah semua aturan dibuat, maka aturan akan disederhanakan (digabung atau diperumum).

Secara garis besar langkah-langkah yang dilakukan oleh algoritma C4.5 dalam membentuk pohon keputusan adalah sebagai berikut (Rina Dewi Indah sari dan Yuwono Sindunata, 2014):

- Pada saat awal pembentukan pohon akan di mulai dengan membuat suatu node yang melambangkan *training sample*.
- Jika sampel-sampel tersebut memiliki kelas yang sama, maka node tersebut dijadikan *leaf node* dengan label kelas tersebut.
- Jika sampel-sampel tersebut tidak memiliki kelas yang sama maka algoritma ini akan mencari gain ratio tertinggi dari atribut yang tersedia, sebagai cara untuk memilih atribut yang paling berpengaruh paada *training sample* yang tersedia. Nantinya atribut ini akan dijadikan atribut “penguji” atau “keputusan” pada node tersebut. Hal yang perlu diperhatikan adalah ketika atribut tersebut bernilai kontinu, maka atribut tersebut harus di diskritkan terlebih dahulu.

- Cabang untuk setiap node akan dibentuk berdasarkan nilai-nilai yang diketahui dari atribut pengujian.
- Algoritma ini akan terus melakukan proses yang sama *rekursif* untuk membentuk suatu pohon keputusan untuk setiap sample di setiap bagiannya.
- Proses *rekursif* ini akan berhenti, ketika salah satu dari kondisi di bawah telah terpenuhi Kondisi-kondisinya adalah:
 - Semua sample yang diberikan pada node adalah berasal dari satu kelas yang sama.
 - Tidak ada atribut lainnya yang dapat digunakan untuk mempartisi sample lebih lanjut.
 - Tidak ada sample yang memenuhi *test-attribute*. Dalam hal ini, sebuah daun dibuat dan dilabeli dengan kelas yang memiliki sample terbanyak (*majority voting*).

Untuk memilih atribut sebagai akar, didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *Gain* digunakan rumus seperti tertera dalam persamaan berikut (Sarah Faradillah, 2013) :

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

Keterangan :

- S : Himpunan Kasus
- A : Atribut
- n : Jumlah partisi atribut A
- |S_i| : Jumlah kasus pada partisi ke i
- |S| : Jumlah kasus dalam S

Sementara itu perhitungan nilai *entropy* dapat dilihat pada persamaan berikut:



Entropy (S)

$$= \sum_{i=1}^n \frac{|Si|}{|S|} - pi$$

* $\log_2 pi$

..... (2)

Di mana :

- S : himpunan kasus
- A : fitur
- n : jumlah partisi S
- pi : proporsi dari Si terhadap S

METODE PENELITIAN

Untuk mencapai keakuratan dan ketelitian data serta informasi dalam penelitian ini, maka pengumpulan data di lakukan dengan 3 cara :

1. Penelitian lapangan (*Field Research*)

Penelitian ini dilakukan untuk mendapatkan data secara langsung dari Dinas Pertanian Kabupaten Kerinci dengan melakukan Observasi dan wawancara langsung ke instansi serta pada pihak – pihak yang terkait sehingga data yang dimuat dalam laporan ini nyata kebenarannya.

2. Penelitian Perpustakaan (*Library Research*)

Penelitian ini dilakukan dengan membaca buku – buku literature, yang di jadikan sebagai bahan.

3. Penelitian Laboratorium (*Laboratory Research*)

Penelitian yang didukung perangkat komputer sebagai alat bantu dalam percobaan penyelesaian masalah. Data yang akan digunakan dalam penyusunan sistem informasi berbasis komputer harus diolah dengan program aplikasi yang sesuai dengan kebutuhan pemakai, oleh sebab itu penerapan sistem informasi berbasis komputer dapat dilakukan dengan menggunakan komputer yang dibutuhkan.

4. Merancang Program aplikasi untuk membantu mengolah data yang di dapat pada saat melakukan penelitian.

5. Adapun spesifikasi *hardware* dan *software* yang di gunakan sebagai berikut :

1. *Hardware* (perangkat keras)

Merupakan komponen secara fisik yang dapat dilihat dan disentuh.

Spesifikasi *Hardware* yang dibutuhkan adalah :

1. Processor intel core i3
2. Flash Disk 4 GB
3. Printer cannon PIXMA iP 1980
4. Memory 2 GB DDR3 dan 320 GB HDD
5. *Mouse* : *Optic Lexron*
6. Monitor
7. Laptop

2. *Software* (perangkat lunak)

1. Sistem Operasi Windows 7
2. Microsoft *Office* 2007 dan *microsoft visio*
3. Weka
4. Serta *Software* pendukung lainnya.

HASIL DAN PEMBAHASAN

Pemilihan Variabel

Data hama tanaman padi pada Dinas Pertanian Kabupaten Kerinci tahun ajaran 2015 memiliki *format* seperti berikut :

1. No
2. Nama Hama
3. Ciri - ciri
4. Keganasan
5. Keterangan

Dari data-data tersebut, yang diambil sebagai variabel keputusannya adalah nilai keganasan ganas, dan tidak ganas. Sedangkan yang diambil sebagai variabel penentu dalam pembentukan pohon keputusan adalah jumlah populasi, cara berkembang biak, perjalanan dan tingkat kerusakan.

Melakukan Pra-Proses



No	Nama Hama	Jumlah Populasi per (induk)	Berkembang Biak	Masa Berkembang Biak (hari)	Tingkat Kerusakan (%)	Keganasan
1	Tikus	10	Melahirkan	25	35	Ganas
2	Walang sangit	200	bertelur	21	30	Ganas
3	Kepik hijau	250	bertelur	25	15	Tidak Ganas
4	Wereng Coklat	300	bertelur	28	30	Ganas
5	Burung pipit	5	bertelur	15	20	Tidak Ganas
6	Ulat grayak	230	bertelur	30	20	Tidak Ganas
7	Hama putih palsu	150	bertelur	30	20	Tidak Ganas
8	belalang	100	bertelur	15	20	Tidak Ganas
9	Penggerek batang putih	260	bertelur	28	30	Ganas
10	Penggerek batang kuning	150	bertelur	31	30	Ganas
11	Babi hutan	10	melaahirkan	120	15	Tidak Ganas
12	Penggerek batang Padi Bergaris	150	bertelur	35	30	Ganas
13	Penggerek batang Merah Jambu	100	bertelur	30	30	Ganas
14	Keong emas	80	bertelur	26	15	Tidak Ganas
15	Hama putih	150	bertelur	25	20	Tidak Ganas
16	ganjur	200	bertelur	20	30	Ganas
17	Orong-orong	50	bertelur	15	15	Tidak Ganas
18	Kepinding tanah	70	bertelur	18	20	Tidak Ganas
19	Lalat bibit	80	bertelur	20	15	Tidak Ganas
20	uret	60	bertelur	18	20	Tidak Ganas
21	bekicot	70	bertelur	20	20	Tidak Ganas
22	Tungau merah	50	bertelur	10	20	Tidak Ganas
23	Wereng hijau	250	bertelur	15	30	Ganas
24	Wereng Punggung Putih	200	bertelur	15	30	Ganas
25	Kepik hitam	150	bertelur	20	20	Tidak Ganas
26	Ulat tanduk hijau	50	bertelur	25	20	Tidak Ganas
27	kepiting	30	bertelur	25	15	Tidak Ganas

28	Kutu akar padi	150	bertelur	20	20	Tidak Ganas
29	Ulat kantung	250	bertelur	18	15	Tidak Ganas
30	Ulat tanah	80	bertelur	28	15	Tidak Ganas

Tabel 1 Format Data Hama Tanaman padi Setelah Pemilihan Variabel

Setelah data dalam format di atas, kemudian dilakukan pra-proses dengan mengelompokkan data-data tersebut ke dalam kelas (K) dengan Sturges (Riduwan, 2003) dengan menggunakan Rumus :

$$\text{Jumlah Kelas (K)} = 1 + 3.3 \log n$$

Dimana : n = jumlah data

$$K = 1 + 3.3 \log(30)$$

$$K = 1 + 3.3 * 1.4771$$

$$K = 1 + 4.874500$$

$$K = 5.874500 \implies 6$$

Berarti dengan jumlah data 30, maka akan mendapatkan 6 kelas, data tersebut akan dikelompokkan berdasarkan atribut sebagai berikut.

Jumlah Populasi	Klasifikasi
0 – 50	1
51 – 100	2
101 – 150	3
151 – 200	4
201 – 250	5
>251	6

Tabel 2 Klasifikasi Jumlah Populasi

Berkembang Biak	Klasifikasi
Melahirkan	A
Bertelur	B

Tabel 3 Klasifikasi Perkembang Biakaan

Masa Berkembang Biak	Klasifikasi
0 - 22	1
23 - 44	2
45 - 66	3
67 - 88	4
89 - 110	5
>111	6



Tabel 4 Masa Berkembang Biak

Tingkat Kerusakan	Klasifikasi
0,00 – 4	1
5 – 9	2
10 – 14	3
15 – 19	4
20 – 24	5
>25	6

Tabel 5 Tingkat Kerusakan

Format data akhir setelah dilakukan pra-proses tampak seperti tabel 6.

N o.	Nama Hama	Jumlah Populasi per(induk)	Berkembang Biak	Masa Berke mbang Biak (hari)	Tingkat Kerusakan (%)	Keganasan
1	Tikus	1	A	2	35	Ganas
2	Walangsangit	4	B	1	30	Ganas
3	Kepik hijau	5	B	2	15	Tidak Ganas
4	Wering Coklat	6	B	2	30	Ganas
5	Burung pipit	1	B	1	20	Tidak Ganas
6	Ulat grayak	5	B	2	20	Tidak Ganas
7	Hama putih palsu	3	B	2	20	Tidak Ganas
8	belalang	2	B	1	20	Tidak Ganas
9	Penggerak batang putih	6	B	2	30	Ganas
10	Penggerak batang kuning	3	B	2	30	Ganas
11	Babi hutan	1	A	6	15	Tidak Ganas
12	Penggerak batang Padi	3	B	2	30	Ganas

	Bergaris					
13	Penggerak batang Merah Jambu	2	B	2	30	Ganas
14	Keong emas	1	B	2	15	Tidak Ganas
15	Hama putih	3	B	2	20	Tidak Ganas
16	ganjur	4	B	1	30	Ganas
17	Orong-orong	1	B	1	15	Tidak Ganas
18	Kepinding tanah	2	B	1	20	Tidak Ganas
19	Lalat bibit	2	B	1	15	Tidak Ganas
20	uret	2	B	1	20	Tidak Ganas
21	bekicot	2	B	1	20	Tidak Ganas
22	Tungau merah	1	B	1	20	Tidak Ganas
23	Wering hijau	5	B	1	30	Ganas
24	Wering Punggung Putih	4	B	1	30	Ganas
25	Kepik hitam	3	B	1	20	Tidak Ganas
26	Ulat tanduk hijau	1	B	2	20	Tidak Ganas
27	kepiting	1	B	2	15	Tidak Ganas
28	Kutu akar padi	3	B	1	20	Tidak Ganas
29	Ulat kantung	5	B	1	15	Tidak Ganas
30	Ulat tanah	1	B	2	15	Tidak Ganas

Tabel 6 Format Data Akhir

Format data akhir pada tabel 5 di atas didapat berdasarkan dari atribut yang sudah dikelompokkan atau diklasifikasi, misalkan data



pada tabel 4.1 jumlah populasi hama adalah “10”, setelah diklasifikasi menjadi “1”, Berkembang Biak adalah “Melahirkan”, maka setelah klasifikasi menjadi “A”, Masa berkembang biak adalah “25” berubah menjadi “2, tingkat kerusakan adalah “35” berubah menjadi “6” dan seterusnya.

Pohon Keputusan

Dari format data akhir keganasan hama tanaman padi maka akan dilakukan klasifikasi data algoritma C4.5 dengan membuat pohon keputusan. Seperti yang telah dijelaskan sebelumnya, algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Dalam kasus yang tertera pada tabel 4.5 di atas, akan dibuat pohon keputusan untuk menentukan klasifikasi keganasan hama tanaman padi (ganas dan tidak ganas) dengan melihat jumlah populasi, berkembang biak, masa berkembang biak dan tingkat kerusakan.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus (2.1), sedangkan untuk menghitung nilai *entropy* dapat dilihat pada rumus (2.2).

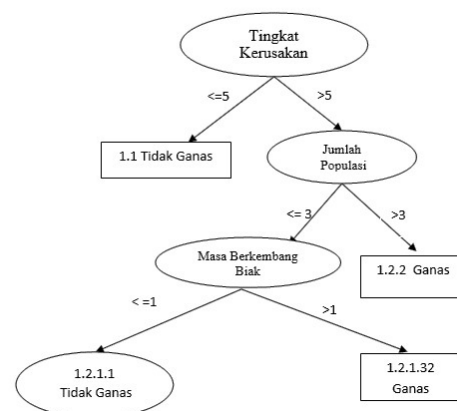
Dengan menggunakan dua persamaan di atas maka akan didapatkan *entropy* dan *gain* yang digunakan sebagai akar dalam membuat pohon keputusan.

- Menghitung jumlah kasus, jumlah kasus untuk keputusan “Ganas”, jumlah kasus untuk keputusan “Tidak Ganas”, dan kasus yang dibagi berdasarkan atribut jumlah populasi, berkembang biak, masa berkembang biak dan tingkat kerusakan. Setelah itu, lakukan perhitungan *gain* untuk setiap atribut.

Node		Jumlah Kasus (S)	Ganas (S1)	Tidak Ganas (S2)	Entropy	Gain
1	Jumlah Populasi	10	1	9		
2		2	3	1	0,91834	
3		3	3	2	0,91834	
	Masa Berkembang Biak					0,31150375
1		3	0	3	0	
2		5	3	2	0,9709508	
3		0	0	0		
4		0	0	0		
5		0	0	0		
6		0	0	0		
	Berkembang Biak					-0,03
B		8	3	5		

Tabel 7 Hasil Perhitungan Node terakhir

Dari hasil tabel 8 dapat diketahui bahwa atribut *gain* tertinggi adalah Masa Berkembang Biak, yaitu sebesar 0,31150375, berarti Masa Berkembang Biak dapat menjadi node akar selanjutnya. Ada dua nilai atribut dari Masa Berkembang Biak yaitu “1” dan “2”. Dari nilai atribut tersebut, nilai “2” sudah Ganas, sehingga tidak perlu dilakukan perhitungan. Pohon keputusan yang terbentuk pada saat ini adalah seperti terlihat pada gambar 5.



Gambar 1 Pohon Keputusan Hasil Perhitungan Node terakhir

Berdasarkan pohon keputusan terakhir yang terbentuk pada gambar 4.5 di atas, maka



aturan atau *rule* yang terbentuk adalah sebagai berikut :

1. Jika Tingkat Kerusakan ≤ 5 , maka Tingkat Keganasan = Tidak Ganas
2. Jika Tingkat Kerusakan > 5 dan Jumlah Populasi > 3 , Maka Tingkat Keganasan = Ganas
3. Jika Tingkat Kerusakan > 5 dan Jumlah Populasi ≤ 3 dan Masa Berkembang Biak > 1 , Maka Tingkat Keganasan = Ganas
4. Jika Tingkat Kerusakan > 5 dan Jumlah Populasi ≤ 3 dan Masa Berkembang Biak ≤ 1 , Maka Tingkat Keganasan = Tidak Ganas

Berdasarkan dari *rule/knowledge* yang dihasilkan terdapat beberapa *rule* yang cukup sesuai dengan kejadian yang terjadi didalam menentukan Tingkat Keganasan hama tanaman padi, dimana rata-rata hama tanaman padi yang memiliki tingkat kerusakan tinggi dan jumlah populasinya yang tinggi dapat dikategorikan hama tanaman padi yang ganas.

Implementasi

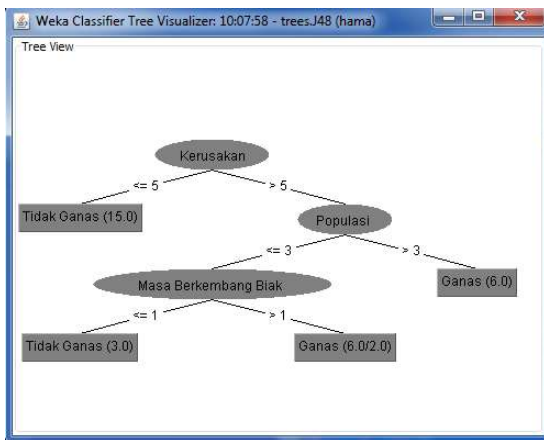
Pengujian terhadap hasil analisa, sangat penting untuk dilakukan untuk memastikan apakah hasil analisa tersebut sesuai. Hasil pengolahan data yang dikerjakan secara manual pada BAB IV dapat diuji kebenarannya menggunakan salah satu *software datamining WEKA*.

No.	Nama Hama	Jumlah Populasi per(indak)	Berkembang Biak	Masa Berkembang Biak (hari)	Tingkat Kerusakan (%)	Keganasan
1	Tikus	10	Malahirkan	25	35	Ganas
2	Walang sangir	200	bersebar	21	30	Ganas
3	Kepik hijau	250	bersebar	25	15	Tidak Ganas
4	Wereng coklat	300	bersebar	28	30	Ganas
5	Burung pipit	5	bersebar	15	20	Tidak Ganas
6	Ulat grayak	230	bersebar	30	20	Tidak Ganas
7	Hama putih palsu	150	bersebar	30	20	Tidak Ganas
8	belalang	100	bersebar	15	20	Tidak Ganas
9	Penggerak batang putih	200	bersebar	28	30	Ganas
10	Penggerak batang kuning	150	bersebar	31	30	Ganas
11	Babi hutan	10	malahirkan	120	15	Tidak Ganas
12	Penggerak batang Padi Bergaris	150	bersebar	35	30	Ganas
13	Penggerak batang Merah Jambu	100	bersebar	30	30	Ganas
14	Keceng emas	80	bersebar	26	15	Tidak Ganas
15	Hama putih	150	bersebar	25	20	Tidak Ganas
16	ganjur	200	bersebar	20	30	Ganas
17	Orong-orong	50	bersebar	15	15	Tidak Ganas
18	Kepinding tanah	70	bersebar	18	20	Tidak Ganas
19	Lalat bibit	80	bersebar	20	15	Tidak Ganas
20	uret	60	bersebar	18	20	Tidak Ganas
21	belicot	70	bersebar	20	20	Tidak Ganas
22	Tungau merah	50	bersebar	10	20	Tidak Ganas
23	Wereng hijau	200	bersebar	15	30	Ganas
24	Wereng Punggung Putih	200	bersebar	15	30	Ganas
25	Kepik hitam	150	bersebar	20	20	Tidak Ganas
26	Ulat tanduk hijau	50	bersebar	25	20	Tidak Ganas
27	kepiting	30	bersebar	25	15	Tidak Ganas
28	Kuru akar padi	150	bersebar	20	20	Tidak Ganas
29	Ulat kampong	250	bersebar	18	15	Tidak Ganas
30	Ulat tanah	80	bersebar	28	15	Tidak Ganas

Gambar 2 Rekap Data Hama Tanaman Padi Dalam File Microsoft Excel 2007 (hama.xls)

No.	Nama Hama	Jumlah Populasi per(indak)	Berkembang Biak	Masa Berkembang Biak (hari)	Tingkat Kerusakan (%)	Keganasan
1	Tikus	1	A	2	6	Ganas
2	Walang sangir	4	B	1	6	Ganas
3	kepik hijau	5	B	2	4	Tidak Ganas
4	Wereng coklat	6	B	2	6	Ganas
5	Burung pipit	1	B	1	6	Tidak Ganas
6	Ulat grayak	5	B	2	5	Tidak Ganas
7	Hama putih palsu	3	B	2	6	Tidak Ganas
8	belalang	2	B	1	5	Tidak Ganas
9	Penggerak batang putih	6	B	2	6	Ganas
10	Penggerak batang kuning	3	B	2	6	Ganas
11	Babi hutan	1	A	6	4	Tidak Ganas
12	Penggerak batang Padi Bergaris	3	B	2	6	Ganas
13	Penggerak batang Merah Jambu	2	B	2	6	Ganas
14	Keceng emas	1	B	2	4	Tidak Ganas
15	Hama putih	3	B	2	5	Tidak Ganas
16	ganjur	4	B	1	6	Ganas
17	Orong-orong	1	B	1	4	Tidak Ganas
18	Kepinding tanah	2	B	1	6	Tidak Ganas
19	Lalat bibit	2	B	1	4	Tidak Ganas
20	uret	2	B	1	6	Tidak Ganas
21	belicot	2	B	1	5	Tidak Ganas
22	Tungau merah	1	B	1	5	Tidak Ganas
23	Wereng hijau	5	B	1	6	Ganas
24	Wereng Punggung Putih	4	B	1	6	Ganas
25	kepik hitam	3	B	1	5	Tidak Ganas
26	Ulat tanduk hijau	1	B	2	6	Tidak Ganas
27	kepiting	1	B	2	4	Tidak Ganas
28	Kuru akar padi	3	B	1	5	Tidak Ganas
29	Ulat kampong	5	B	1	4	Tidak Ganas
30	Ulat tanah	1	B	2	4	Tidak Ganas

Gambar 3 File Data Hama Yang Telah Ditransformasi



Gambar 4 Hasil Visualisasi Tree

Dimana *rule* yang dihasilkan pada gambar 5.10 adalah :

- Jika Kerusakan ≤ 5 , maka Tingkat Keganasan = Tidak Ganas
- Jika Kerusakan > 5 dan Populasi > 3 , Maka Tingkat Keganasan = Ganas
- Jika Kerusakan > 5 dan Populasi ≤ 3 dan Masa Berkembang Biak > 1 , Maka Tingkat Keganasan = Ganas
- Jika Kerusakan > 5 dan Populasi ≤ 3 dan Masa Berkembang Biak ≤ 1 , Maka Tingkat Keganasan = Tidak Ganas

Setelah dilakukan pengujian pada software *WEKA* 3.6 untuk memprediksi tingkat keganasan hama tanaman padi seperti yang sudah dijelaskan di atas ternyata hasil *rule* yang dihasilkan pada *WEKA* pada gambar 5.10 hampir sama dengan *rule* yang dihasilkan pada proses manual penemuan *knowledge* menggunakan pohon keputusan.

SIMPULAN

Dari uraian pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

1. Metode pohon keputusan yang diproses dengan Weka lebih efektif dan fleksibel jika digunakan pada proses pengklasifikasian keganasan hama tanaman padi.

2. Pemilihan variabel (atribut kondisi dan atribut keputusan) yang akan digunakan dalam memprediksi juga sangat mempengaruhi rule atau knowledge yang dihasilkan.
3. Sistem yang dibangun dapat membantu dalam pengklasifikasian tingkat keganasan hama tanaman padi yang umumnya masih dilakukan secara manual.
4. Algoritma C4.5 dengan metode pohon keputusan dapat memberikan informasi eksekutif dan sistem digunakan untuk menggambarkan proses yang terkait dengan pengklasifikasian keganasan hama tanaman padi.
5. Sistem pengklasifikasian tingkat keganasan hama tanaman padi menggunakan algoritma C4.5 dapat digunakan dalam pengambilan keputusan untuk mencari alternatif yang baik.

DAFTAR PUSTAKA

- Liliana Swastina, (2013), Penerapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa, Jurnal GEMA AKTUALITA, Vol. 2 No. 1, Juni 2013.
- Sarah Faradillah, (2013, Implementasi Data Mining Untuk Pengenalan Karakteristik Transaksi Customer Dengan Menggunakan Algoritma C4.5, Pelita Informatika Budi Darma, Volume : V, Nomor: 3, Desember 2013
- Rina Dewi Indah sari, (2014), Penerapan Data Mining untuk Analisa Pola Perilaku Nasabah Dalam Pengkreditan Menggunakan Metode C.45 Studi Kasus Pada Ksu Insan Kamil Demak, Jurnal Ilmiah Teknologi dan Informasi ASIA Vol. 8 No 2, Agustus 2014



Benni R Siburian, (2013), aplikasi data mining untuk menampilkan tingkat kelulusan mahasiswa dengan algoritma apriori, Pelita Informatika Budi Darma., Vol. VII, No. 2, Agustus 2014.

Kennedi Tampubolon, Hoga Saragih, & Bobby Reza, (2013), Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan, Majalah Ilmiah Informasi dan Teknologi Ilmiah (INTI), Vol. I, No. 1, Oktober 2013.

Azhari & Anshori, (2009), pendekatan aturan asosiasi untuk analisis pergerakan saham, Seminar Nasional Informatika 2009 (semnasIF 2009).

Alimancon Sijabat, (2015), Penerapan Data Mining Untuk Pengolahan Data Siswa Dengan Menggunakan Metode Decision Tree (Studi Kasus : Yayasan Perguruan Kristen Andreas), Majalah Ilmiah Informasi dan Teknologi Ilmiah (INTI) Volume : V, Nomor : 3 , Pebruari 2015

Ari Puspita dan Mochamad Wahyudi, (2015), Konferensi Nasional Ilmu Pengetahuan dan Teknologi (KNIT) 2015 8 Agustus 2015, Bekasi, Indonesia