

## ARTIKEL

# Perbandingan Metode *Naive Bayes* Dan *ID3* Untuk Klasifikasi Jamur Beracun Dan Tidak Beracun



Oleh:

**MUHAMMAD LUTFI ARISANDI**

**NPM : 14.1.03.02.0029**

Dibimbing oleh :

1. Ahmad Bagus Setiawan, S.T, M. M, M. Kom
2. Ardi Sanjaya, M.Kom.

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS NUSANTARA PGRI KEDIRI**

**TAHUN 2019**



**SURAT PERNYATAAN  
ARTIKEL SKRIPSI TAHUN 2019**

Yang bertanda tangan di bawah ini:

Nama Lengkap : Muhammad Lutfi Arisandi  
NPM : 14.1.03.02.0029  
Telepon/HP : 083846204233  
Alamat Surel (Email) : Doi.seejati@gmail.com  
Judul Artikel : Perbandingan Metode *Naive Bayes* Dan *ID3* Untuk  
Klasifikasi Jamur Beracun Dan Tidak Beracun  
Fakultas – Program Studi : Fakultas Teknik – Teknik Informatika  
Nama Perguruan Tinggi : UNIVERSITAS NUSANTARA PGRI KEDIRI  
Alamat Perguruan Tinggi : Jln. Kh. Ahmad Dahlan No.76, Mojoroto, Kediri

Dengan ini menyatakan bahwa :

- artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
- artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri 12 Februari 2019
Pembimbing I  Ahmad Bagus Setiawan, S.T, M. M, M. Kom NIDN. 0703018704	Pembimbing II  Ardi Sanjaya, M.Kom NIDN. 0706118101	Penulis,  Muhammad Lutfi Arisandi NPM. 14.1.03.02.0029

Muhammad Lutfi Arisandi | 14.1.03.02.0029  
Fakultas Teknik – Teknik Informatika

simki.unpkediri.ac.id  
|| 1 ||

## Perbandingan Metode *Naive Bayes* Dan *ID3* Untuk Klasifikasi Jamur Beracun Dan Tidak Beracun

Muhammad Lutfi Arisandi

14.1.03.02.0029

Fakultas Teknik – Teknik Informatika

Email : [Doi.seejati@gmail.com](mailto:Doi.seejati@gmail.com)

Ahmad Bagus Setiawan, S.T, M. M, M. Kom<sup>1</sup>, Ardi Sanjaya, M.Kom<sup>2</sup>  
UNIVERSITAS NUSANTARA PGRI KEDIRI

### ABSTRAK

Indonesia berada di daerah tropis, iklim tersebut cukup mendukung untuk pertumbuhan jamur. Jamur adalah organisme bersel tunggal atau banyak dengan tidak memiliki klorofil. Dalam famili *Agaricus* dan *Lepiota* yang secara liar hidup di alam terbuka dengan bentuk yang beraneka ragam, Bagi beberapa masyarakat Indonesia masih belum mengetahui ciri-ciri jamur beracun dan tidak beracun karena belum adanya informasi yang terperinci mengenai ciri-ciri jamur. Maka dibutuhkan sistem pengklasifikasian jamur beracun dan tidak beracun untuk membedakannya yang dapat diakses dengan mudah oleh masyarakat dengan jaringan internet.

Dalam kategori klasifikasi memiliki banyak pilihan algoritma, beberapa algoritma yang sering digunakan adalah algoritma *naive bayes* dan *ID3*. Pada penelitian ini akan dilakukan perbandingan dari kedua metode *Naive Bayes* dan *ID3*. Metode *Naive Bayes* mengklasifikasikan data jamur dengan menggunakan teknik probabilitas sedangkan metode *ID3* mengklasifikasikan jamur dengan sebuah pohon keputusan yang didapatkan dari pelatihan data *training*.

Pada pengujian sistem ini, rata-rata tingkat keakuratan yang dihasilkan dari kedua metode adalah 79,52 tingkat keakuratan dari metode *ID3* dan 63,338 tingkat keakuratan dari metode *Naive Bayes*. Jadi dapat disimpulkan bahwa metode *ID3* lebih unggul dari pada metode *Naive Bayes*.

**KATA KUNCI** : Klasifikasi, Jamur, *Naive Bayes*.

### I. LATAR BELAKANG

Indonesia berada di daerah tropis, iklim tersebut cukup mendukung untuk pertumbuhan jamur. Jamur adalah organisme eukariotik yang bersel tunggal atau banyak dengan tidak memiliki klorofil. Salah satunya famili *Agaricus* dan *Lepiota* yang secara liar hidup di alam terbuka dengan bentuk yang beraneka ragam,

warna yang bermacam-macam serta sifat yang belum banyak diketahui.

Pada umumnya jamur terbagi dua yaitu jamur beracun dan tidak beracun. Bagi beberapa masyarakat Indonesia masih belum mengetahui ciri-ciri jamur beracun dan tidak beracun karena belum adanya informasi yang terperinci mengenai ciri-ciri jamur dan bentuk, Maka dibutuhkan pengklasifikasian jamur beracun dan tidak beracun untuk

membedakannya. Dalam sistem klasifikasi ada beberapa metode yang sangat baik dalam klasifikasi, yaitu : *Naive Bayes*, *KNN*, *C45*, dan *ID3*. Dari beberapa metode tersebut dibutuhkan sebuah sistem yang membandingkan tingkat keakurasian dari kedua metode yang akan digunakan untuk mengklasifikasikan jamur beracun dan tidak beracun

Dalam penelitian yang dilakukan oleh Dewi Sartika (2017), menjelaskan tentang Perbandingan Algoritma Klasifikasi *Naive Bayes*, *Nearest Neighbour*, dan *Decision Tree* pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. Hasil perbandingan menunjukkan bahwa metode decision tree memiliki tingkat akurasi tertinggi dibandingkan algoritma naive bayes dan nearest neighbour yaitu mencapai 75.6%.

Penelitian Mila Listiana (2015), perbandingan algoritma decision tree (c4.5) dan naive bayes pada data mining untuk identifikasi tumbuh kembang anak balita (studi kasus puskesmas kartasura). Hasil akhir dari penelitian ini adalah metode *naive bayes* lebih baik digunakan dari pada metode *decision tree* dengan nilai total 250,67% untuk *decition tree* dan 258.03% untuk *naive bayes*.

Berdasarkan kedua penelitian tersebut, dapat dilihat bahwa kedua algoritma, Naive Bayes dan ID3 mempunyai kinerja yang baik dalam pengklasifikasiannya. Namun, belum diketahui algoritma mana diantara keduanya yang lebih unggul kinerjanya. Oleh karena itu algoritma ini perlu dibandingkan. Pada penelitian ini membandingkan kinerja dari algoritma Naive Bayes dan ID3 dalam melakukan klasifikasi jamur beracun dan tidak beracun. Hasil akhir dari penelitian ini dapat mengetahui algoritma mana yang memiliki akurasi yang paling tinggi.

## II. METODE

### A. Metode Iterative Dichotomiser 3

Algoritma pada metode ini menggunakan konsep dari *entropy* informasi. Algoritma ini melakukan pencarian secara menyeluruh pada semua kemungkinan pohon keputusan. Secara ringkas, langkah kerja algoritma ID3 :

1. Hitung nilai *Entropy* dari setiap atribut dengan menggunakan rumus

$$Entropy(S) = (-P_+ \log_2 P_+) + (-P_- \log_2 P_-)..... (1)$$

Keterangan pada Rumus diatas terdiri dari :

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

**P+** = jumlah yang bersolusi positif (mendukung) pada data sample dibagi dengan total kasus.

**P-** = jumlah yang bersolusi negatif (tidak mendukung) pada data sample dibagi dengan total kasus.

2. Hitung nilai information gain untuk setiap atribut dengan rumus

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \dots\dots(2)$$

Keterangan pada Rumus diatas terdiri dari :

**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.

**|S<sub>v</sub>|** = jumlah sample untuk nilai V.

**|S|** = jumlah seluruh sample data.

**Entropy(S<sub>v</sub>)** = *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 (Suyanto ,2017).

## B. Metode Naive Bayes

Metode ini menggunakan teorema bayes yang ditemukan oleh Thomas bayes di abad ke-18 (Suyanto ,2017). *Naive bayes* adalah perhitungan statistik dengan menghitung probabilitas kemiripan kasus lama yang ada dibasis kasus dengan kasus baru. *Naive bayes* memiliki tingkat akurasi yang tinggi dan kecepatan yang baik ketika diterapkan pada *database* yang besar. Persamaan dari *naive bayes* sebagai berikut :

$$P(C_i | X) = \frac{P(X|C_i)P(C_i)}{P(X)} \dots\dots(3)$$

Keterangan :

**X** : Kriteria suatu kasus berdasarkan masukan

**C<sub>i</sub>** : Kelas solusi pola ke-i, dimana i adalah jumlah label kelas

$P(C_i | X)$  : Probabilitas kemunculan label kelas  $C_i$  dengan kriteria masukan X

$P(X | C_i)$  : Probabilitas kriteria masukan X dengan label kelas  $C_i$

$P(C_i)$  : Probabilitas label kelas  $C_i$

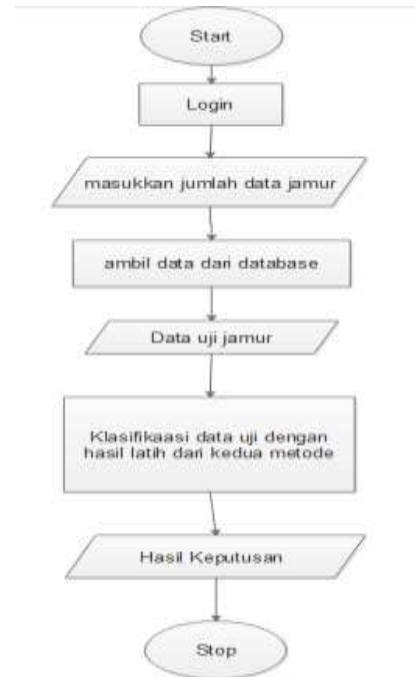
### III. HASIL DAN KESIMPULAN

#### A. Implementasi sistem

Adanya Evaluasi sistem sangatlah penting untuk menjadi koreksi dimana apa yang kita rancang tidak melenceng dari apa yang telah kita design. Sehingga kita dapat mendapatkan tolak ukur yang baik dari apa yang kita design atau rancang dapat terealisasi dengan baik. Implementasi yang baik adalah implementasi yang tidak kurang dari design yang telah direncanakan. Berikut adalah bagian-bagian yang telah terealisasi dalam bentuk aplikasi.

##### 1. Flowchart Sistem

Berikut merupakan perancangan *Flowchart* dari aplikasi data mining untuk klasifikasi jamur beracun dan tidak beracun. Dapat dilihat pada gambar 3.1.



Gambar 3.1 Flowchart Sistem  
Klasifikasi Jamur

Keterangan :

- Start: aplikasi dijalankan
- Proses : login
- Mengisi : masukkan jumlah data jamur yang ingin diuji
- Proses: ambil data jamur dari database
- Keluaran : menampilkan Data Jamur yang akan diuji
- Proses: Klasifikasi menggunakan hasil latih dari kedua metode
- Hasil: Keputusan Klasifikasi
- Stop : sistem selesai

Hasil klasifikasi jamur menampilkan tingkat akurasi dari kedua metode dan memberikan informasi metode mana yang lebih baik.

## B. Tampilan Program

Aplikasi perbandingan metode *naive bayes* dan *ID3* untuk klasifikasi jamur beracun dan tidak beracun ini dibuat sangat sederhana supaya dapat dipahami oleh pengguna. Tampilan beranda dapat dilihat pada gambar 3.2.



Gambar 3.2 Tampilan Beranda

Pada tampilan beranda terdapat menu-menu sebagai berikut :

### 1. Menu Beranda

Menu ini berfungsi untuk kembali ke halaman awal dari sistem ini.

### 2. Menu Uji Data

Menu Uji data merupakan yang akan menguji data dengan hasil dari perhitungan dari metode *Naive Bayes* dan *ID3* yang hasil dari pengujian tersebut menampilkan tingkat keakurasian dari kedua metode.

### 3. Menu Bantuan

Modul bantuan merupakan sebuah modul yang berisi

tentang informasi mengenai tata cara pengguna dalam menggunakan sistem tersebut.

### 4. Menu Tentang

Pada saat pengguna ingin mengetahui informasi dari aplikasi ini, maka pengguna bisa mengakses modul tentang.

### 5. Menu Log Out

Pada saat pengguna harus melakukan login supaya dapat masuk ke dalam sistem.

Pada menu Uji Data pengguna akan memasukkan jumlah data yang akan di uji dan akan dilakukan pengujian dengan hasil yang didapat dari proses *training* dari kedua metode tersebut. Tampilan halaman Uji Coba dapat dilihat pada gambar 3.3



Gambar 3.1 Halaman Uji Data

Pada Halaman Uji terdapat 2 *Button*, yaitu :

### 1. *Button* Tampilkan

*Button* ini berfungsi untuk menampilkan data yang akan



diuji. data yang ditampilkan sudah disimpan di dalam *database*.

2. *Button* Uji Data

*Button* ini berfungsi untuk menguji data yang di masukkan tadi dengan menggunakan hasil *training* dari kedua metode dan hasilnya akan menampilkan tingkat akurasi dari kedua metode beserta hasil klasifikasi.

**C. Skenario Uji Coba**

Dari uji coba yang telah dilakukan oleh peneliti dengan data *training* 250 dan data *testing* 20, data *training* 170 dan data *testing* 70, data *training* 170 dan data *testing* 100. Hasil uji coba dengan metode *naive bayes* dan *ID3* dapat dilihat pada tabel berikut :

Tabel 3.1 hasil uji coba *training* 250 dan *testing* 20

no	metode	Data uji coba	benar	salah	akurasi
1	<i>Naive Bayes</i>	20	16	4	$\frac{16}{20} \times 100\% = 80\%$
2	<i>ID3</i>	20	14	6	$\frac{14}{20} \times 100\% = 70\%$

Tabel 3.2 hasil uji coba *training* 170 dan *testing* 70

no	metode	Data uji coba	benar	salah	akurasi
1	<i>Naive Bayes</i>	70	40	30	$\frac{40}{70} \times 100\% = 57,14\%$

2	<i>ID3</i>	70	62	8	$\frac{62}{70} \times 100\% = 88,57\%$
---	------------	----	----	---	--

Tabel 3.3 hasil uji coba *training* 170 dan *testing* 100

no	metode	Data uji coba	benar	salah	akurasi
1	<i>Naive Bayes</i>	100	62	38	$\frac{62}{100} \times 100\% = 62\%$
2	<i>ID3</i>	100	80	20	$\frac{80}{100} \times 100\% = 80\%$

Dari tiga skenario yang telah dilakukan diatas dapat disimpulkan bahwa data *training* 250, 170 dan 170 dengan data *testing* 20, 70 dan 100 hanya dua uji coba saja yang mendapatkan tingkat keakurasian pada *ID3* yang lebih unggul dapat dilihat pada tabel 5.4

Tabel 5.4 rata-rata akurasi

no	Uji coba	Presentase Keberhasilan (%)	
		<i>ID3</i>	<i>Naive Bayes</i>
1	Skenario 1	70	80
2	Skenario 2	88,57	57,14
3	Skenario 3	80	62
Jumlah		238,57	199,14
Rata- rata presentase keberhasilan		$\frac{238,57}{3} = 79,52$	$\frac{199,14}{3} = 63,38$

Pada skenario pertama metode *naive bayes* lebih unggul dikarenakan perhitungan probabilitas kondisionalnya pada data *training* tidak terlalu banyak bernilai nol karena jumlah data



training mempengaruhi hasil perhitungan probabilitas. Pada skenario 2 dan 3 *naive bayes* kurang unggul dikarenakan jumlah data training sedikit dan ada beberapa dari perhitungan probabilitas kondisionalnya bernilai nol maka akan mempengaruhi probabilitas prediksinya. Dari tabel 5.17 dapat disimpulkan bahwa metode *ID3* lebih unggul dengan nilai rata-ratanya adalah 79,52, sedangkan untuk metode *naive bayes* nilai rata-rata dibawah metode *ID3* dengan nilai 63,38.

#### IV. PENUTUP

##### A. Simpilan

Berdasarkan uraian pembahasan dari penelitian yang telah dilakukan maka dapat di simpulkan sebagai berikut:

1. Metode *Naive Bayes* dan *ID3* dapat digunakan untuk klasifikasi jamur beracun dan tidak beracun.
2. Dari pengujian sistem ini didapatkan rata-rata presentase akurasi dari metode *ID3* adalah 79,52 dan metode *Naive Bayes* adalah 63,38.

##### B. Tujuan

Dari hasil penelitian dapat ditarik kesimpulan bahwa tujuan yang ingin dicapai oleh peneliti telah berhasil, tujuan yang terjcapai yaitu :

1. Dengan adanya sistem ini dapat menjadi referensi bagi masyarakat dalam mengenali jamur beracun dan tidak beracun.
2. Perbandingan dari hasil klasifikasi jamur beracun dan tidak beracun dengan metode *Naive Bayes* dan *ID3* telah berhasil, dan metode *ID3* lebih baik untuk melakukan klasifikasi dari pada metode *Naive Bayes*.

#### V. DAFTAR PUSTAKA

- Dan,Thanh Thi Bi., Sihwi ,Sari Widya & Anggrainingsih,Rini. *Implementasi Iterative Dichotomiser 3 Pada Data Kelulusan Mahasiswa S1 Di Universitas Sebelas Maret*. URL :<https://jurnal.uns.ac.id/itsmart/article/view/1770> Diunduh 10 November 2017.
- Kusrini, Luthfi Taufiq Emha, (2009), *Algoritma Data Mining*,Penerbit Andi,Yogyakarta.
- Listiana, Mila. *Perbandingan Algoritma Decision Tree (C4.5) dan Naive Bayes Pada Data Mining Untuk Identifikasi Tumbuh Kembang Anak Balita (Studi Kasus Puskesmas Kartasura)*. URL :



- <http://eprints.ums.ac.id/36124/2/NASKAH%20PUBLIKASI.pdf>  
Diunduh 6 Febuari 2018.
- Pradana, Galih Adi Surya, (2009), Identifikasi Jamur Beracun Pada Jenis Jamur Famili Agaricus Dan Lepiota Berdasarkan Klasifikasi. Skripsi. Dipublikasikan. Kediri: Teknik Informatika Universitas Nusantara PGRI.
- Raharjo, Budi. 2015. Mudah Belajar PHP. Bandung : Informatika Bandung.
- Sari, Ita Purnama dan Saputra, Erik Hadi. 2014. Sistem Informasi Raport Berbasis Web di SMP N 4 Temanggung. URL : <https://ojs.amikom.ac.id/index.php/dasi/article/view/190>.  
Diunduh 10 Desember 2017.
- Sartika, Dewi dan Sensuse, Dana Indra. Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. URL : [http://www.mdp.ac.id/jatisi/vol-3-no-2/Jatisi%20Vol%203%20No%201%20Maret%202017\\_new.pdf](http://www.mdp.ac.id/jatisi/vol-3-no-2/Jatisi%20Vol%203%20No%201%20Maret%202017_new.pdf).  
Diunduh 6 Febuari 2018.
- Suyanto. 2017. Data Mining Untuk Klasifikasi dan Klasterisasi Data. Bandung : Informatika Bandung. UCI Machine Learning Respository. UCI machine learning respository : mushroom data set. URL: <https://archive.ics.uci.edu/ml/datasets/mushroom>. Diakses pada 10 Oktober 2017
- Wibowo, Agung. Purwarupa Sistem Pakar Indentifikasi Jamur Layak Konsumsi Berbasis Web. URL : <https://jurnal.unimed.ac.id/2012/index.php/cess/article/view/6539>.  
Diunduh 26 September 2017.