

IDENTIFIKASI JAMUR MENGUNAKAN METODE K- NEAREST NEIGHBOR DENGAN EKSTRAKSI CIRI MORFOLOGI

by Ahmad Rofiqul Muslikh

Submission date: 14-Mar-2019 09:34AM (UTC+0700)

Submission ID: 1092993177

File name: 44-140-1-PB.pdf (572.59K)

Word count: 1754

Character count: 10574



IDENTIFIKASI JAMUR MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN EKSTRAKSI CIRI MORFOLOGI

Anis Zubair¹⁾, Ahmad Rofiqul Muslikh²⁾

^{1,2)} Fakultas Teknologi Informasi, Universitas Merdeka Malang

Email: anis.zubair@unmer.ac.id

Email: rofickachmad@unmer.ac.id

Abstrak

Jamur merupakan tumbuhan yang banyak di konsumsi oleh masyarakat umum, tetapi tidak semua jamur bisa dikonsumsi langsung, dikarenakan jenis – jenis jamur yang layak konsumsi dan masih terlalu sulit untuk membedakan, maka ada beberapa cara untuk mengidentifikasi jamur yaitu dengan cara morfologi. Dengan cara ekstraksi morfologi jamur beracun dan tidak beracun diidentifikasi dengan lebih akurat. Proses ekstraksi morfologi disini menggunakan pendekatan klasifikasi KNN (K-Nearest Neighbor). Dengan menggunakan pendekatan KNN tersebut dan Dengan 10 kali percobaan dengan menggunakan berbagai macam nilai k, akhirnya dengan menggunakan nilai k=60, akurasi yang didapatkan mencapai 99%.

Kata kunci : identifikasi jamur, KNN, klasifikasi jamur, identifikasi morfologi, ekstraksi morfologi

Abstract

Fungus is a plant widely consumed by the general public, but not all mushrooms can be consumed directly, due to the types of mushrooms that are feasible to consume and still too difficult to distinguish, then there are several ways to identify the fungus is by morphology. By way of morphological extraction of toxic and non-toxic fungi are identified more accurately. The morphological extraction process here uses the KNN (K-Nearest Neighbor) classification approach. Using these KNN approaches and With 10 experiments using various k values, finally using k = 60, the accuracy reached 99%.

Keyword : mushroom identification, mushrooms clasication, KNN, morphology identification, morphology extraction

PENDAHULUAN

Jamur termasuk kingdom fungi, oleh karena itu jamur tidak memiliki daun dan akar yang sejati, serta tidak mempunyai klorofil sehingga tidak dapat melakukan fotosintesis seperti tumbuhan pada umumnya. Jamur digolongkan atau diklasifikasikan tersendiri karena tidak dapat

digolongkan dalam tumbuhan atau hewan. Bentuk jamur ada yang dapat dilihat secara langsung atau bentuknya makroskopis dan ada yang harus diamati menggunakan mikroskop atau bentuknya mikroskopis. Pada umumnya jamur mempunyai sel banyak (multiseluler) misalnya jamur merang dan jamur tempe, tetapi ada juga



yang bersel tunggal (uniseluler) seperti ragi atau yeast (*Saccharomyces*). Jamur yang multiseluler tersusun atas benang-benang yang disebut dengan hifa. Apabila dilihat dengan mikroskop hifa memiliki bentuk bersekat-sekat (berseptata) dan ada yang tidak bersekat (Gudang Biologi, 2015).

Untuk mengenali jamur mana yang dapat dikonsumsi dan beracun, ada beberapa cara yang dapat digunakan. Salah satu aspek yang dapat menjadi tolok ukur dalam identifikasi suatu jamur ialah ciri morfologi. Ciri morfologi yang dimaksud adalah bentuk payung, warna, tekstur payung, dan ciri lain yang terlihat. Penelitian ini menggunakan ekstraksi ciri morfologi dengan K-Nearest Neighbor sebagai classifier. Ekstraksi ciri morfologi ini digunakan untuk membantu identifikasi jamur, sehingga nantinya akan dapat diketahui termasuk kelas apakah suatu jamur tersebut.

KAJIAN LITERATUR

Ekstraksi ciri morfologi pada penelitian ini menggunakan klasifikasi K-Nearest Neighbor. Tahap-tahap penggunaan klasifikasi K-Nearest Neighbor disajikan di bawah ini.



Gambar 1 Flowchat k-NN

1. Menentukan Nilai k

Tahap pertama adalah menentukan nilai k. Penentuan nilai k dapat ditentukan secara bebas.

Penelitian ini menggunakan berbagai macam nilai k yang berbeda. Hasilnya digunakan sebagai perbandingan antara nilai k yang satu dengan yang lainnya.



2. Menghitung Jarak

Tahap kedua yaitu menghitung jarak.

Rumus jarak yang digunakan adalah *Euclidean Distance*.

Rumus untuk menghitung *Euclidean Distance* dinyatakan sebagai berikut.

$$D(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

³ *D* adalah jarak antara data latih (*x*) dan data uji (*y*) yang akan diklasifikasikan.

3. Menentukan Nilai Nearest Neighbor Value

Tahap ketiga adalah menentukan nilai tetangga terdekat dari tiap-tiap jarak.

Hasil yang didapat dari penghitungan ini berupa kategori kelas yang akan dijadikan acuan ketika pengklasifikasian berdasarkan nilai tetangga terdekat.

4. Membuat Klasifikasi Berdasarkan Nearest Neighbor Mayoritas

Tahap terakhir adalah membuat klasifikasi data uji dapat dilakukan berdasarkan kelas mayoritas. Misalnya jika mayoritas nilai tetangga terdekat bernilai kelas “jamur beracun” maka kelas data uji yang belum diketahui termasuk dalam kelas “jamur beracun”.

METODE PENELITIAN

Tahapan Penelitian

1. Pengumpulan Data

Data yang digunakan adalah data jamur yang didapat dari situs archive.ics.uci.edu (UCI Machine Learning Repository).

Data ini didonorkan oleh Jeff Schlimmer tahun 1987. Pada penelitian ini terdapat 2 kelas, yaitu jamur pangan dan jamur beracun. Untuk jumlah tiap-tiap kelas, terdiri atas 4208 data yang termasuk dalam kategori jamur pangan dan 3916 data yang termasuk dalam kategori jamur beracun, sehingga jumlah keseluruhan data yang digunakan sebanyak 8124 data. Setiap inisial dari tiap-tiap atribut dan kelas merupakan representasi dari jenis atribut yang bersangkutan. Ekstraksi ciri yang digunakan dalam data latih adalah ciri morfologi. Pada penelitian ini yang digunakan adalah 21 atribut. Atribut Tipe Kerudung tidak digunakan karena hanya memiliki satu nilai.

Atribut No	Bentuk-Topi	Permukaan-Topi	Warna-Topi	Memar	Bau	Lampiran Lamela	Jarak Lamela	Ukuran Lamela	Warna Lamela	Bentuk Tangkai	Akor Tangkai
1	x	s	n	t	p	f	c	n	k	e	e
2	x	s	y	t	a	f	c	b	k	e	c
3	b	s	w	t	l	f	c	b	n	e	c
4	x	y	w	t	p	f	c	n	n	e	e
5	x	s	g	f	n	f	w	b	k	l	e
6	x	y	y	t	a	f	c	b	n	e	c
7	b	s	w	t	a	f	c	b	g	e	c
8	b	y	w	t	l	f	c	b	n	e	c
9	x	y	w	t	p	f	c	n	p	e	e
10	b	s	y	t	a	f	c	b	g	e	c

Permukaan Tangkai Diatas	Permukaan Tangkai Dibawah	Warna Tangkai Diatas Cincin	Warna Tangkai Dibawah Cincin	Warna Kerudung	Jumlah Cincin	Tipe Cincin	Warna Spora	Populasi	Habitat	Tipe Kerudung	Kelas
s	s	w	w	w	o	p	k	s	u	p	(Ctrl)
s	s	w	w	w	o	p	n	n	g	p	u
s	s	w	w	w	o	p	n	n	m	p	e
s	s	w	w	w	o	p	k	s	u	p	p
s	s	w	w	w	o	e	n	a	g	p	o
s	s	w	w	w	o	p	k	n	g	p	e
s	s	w	w	w	o	p	k	n	m	p	e
s	s	w	w	w	o	p	n	s	m	p	e
s	s	w	w	w	o	p	k	v	g	p	p
s	s	w	w	w	o	p	k	s	m	p	e

Gambar 2 Data Latih



2. Pengolahan Data

Setelah data latih didapatkan, data tersebut diolah terlebih dahulu sebelum pada akhirnya digunakan dalam menentukan kelas pada model klasifikasi. Inisial masing-masing atribut telah dijelaskan sebelumnya, selanjutnya hal yang perlu dilakukan ialah mengubah inisial-inisial tersebut menjadi angka. Angka dimulai dari 1 dan seterusnya sebanyak n sesuai dengan jenis-jenis yang terdapat dalam atribut yang bersangkutan.

3. Penentuan Nilai k

Selanjutnya adalah menentukan nilai k . Nilai k bersifat bebas, yang artinya dapat ditentukan dengan nilai berapa saja dengan batas maksimal jumlah data latih.

4. Pembuatan Klasifikasi Menggunakan XLSTAT

Langkah selanjutnya yaitu proses pengklasifikasian menggunakan metode K-Nearest Neighbor. Pada penelitian ini menggunakan third party Add-ins dalam Ms. Excel yang bernama XLSTAT. Add-ins ini dapat secara cepat membantu menentukan prediksi atau identifikasi suatu obyek berdasarkan data latih yang telah tersedia.

5. Penghitungan Tingkat Akurasi

Untuk menghitung akurasi, peneliti menggunakan rumus Confusion Table.

Rancangan Penelitian

Penelitian ini menggunakan 5 Sheet pada Microsoft Excel. Tiap-tiap sheet diberi nama. Namanya adalah Data Latih, Pengolahan, Akurasi Nilai k , Hasil, dan Grafik Akurasi.

1. Sheet Data Latih

Data mentah berada di sel A1 sampai X8126.

2. Sheet Pengolahan

Keterangan jenis atribut dan kelas berada di sel A3 sampai D175. Data latih yang telah diolah berada di sel F1 sampai AD8126.

3. Sheet Akurasi Nilai k

Kelas data latih (aktual) berada di sel A1 sampai C8125. Untuk prediksi dengan berbagai nilai k beserta confusion table masing-masing, diletakkan sebagaimana berikut.

- $k=7000$, F1 sampai K8125.
- $k=8000$, N1 sampai S8125.
- $k=6000$, V1 sampai AA8125.
- $k=5000$, AD1 sampai AI8125.
- $k=4000$, AL1 sampai AQ8125.
- $k=3000$, AT1 sampai AY8125.
- $k=800$, BB1 sampai BG8125.
- $k=500$, BJ1 sampai BO8125.
- $k=100$, BR1 sampai BW8125.
- $k=60$, BZ1 sampai CE8125.



4. Sheet Hasil

Hasil klasifikasi berada di sel A1 sampai B11, dengan 10 nilai k yang dijadikan percobaan.

5. Sheet Grafik Akurasi

Sheet ini berisi visualisasi berupa grafik koordinat x,y antara nilai k dengan akurasi.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Penelitian

Dengan 10 kali percobaan dengan menggunakan berbagai macam nilai k, akhirnya dengan menggunakan nilai k=60, akurasi yang didapatkan mencapai 99%.

No	Kelas (Aktual)	Prediksi (k=60)	Confusion Table	
1	1	Beracun	Beracun	
2	2	Dapat Dimakan	Dapat Dimakan	32
3	2	Dapat Dimakan	Beracun	3884
4	1	Beracun	Dapat Dimakan	4158
5	2	Dapat Dimakan	Beracun	
6	2	Dapat Dimakan	Dapat Dimakan	
7	2	Dapat Dimakan	Dapat Dimakan	
8	2	Dapat Dimakan	Dapat Dimakan	
9	1	Beracun	Beracun	
10	2	Dapat Dimakan	Dapat Dimakan	
11	2	Dapat Dimakan	Dapat Dimakan	
12	2	Dapat Dimakan	Dapat Dimakan	
13	2	Dapat Dimakan	Dapat Dimakan	
14	1	Beracun	Beracun	
15	2	Dapat Dimakan	Dapat Dimakan	
16	2	Dapat Dimakan	Dapat Dimakan	
17	2	Dapat Dimakan	Dapat Dimakan	
18	1	Beracun	Beracun	
19	1	Beracun	Beracun	
20	1	Beracun	Beracun	
21	2	Dapat Dimakan	Dapat Dimakan	
22	1	Beracun	Beracun	
23	2	Dapat Dimakan	Dapat Dimakan	

Akurasi	99%
---------	-----

Gambar 2 Akurasi

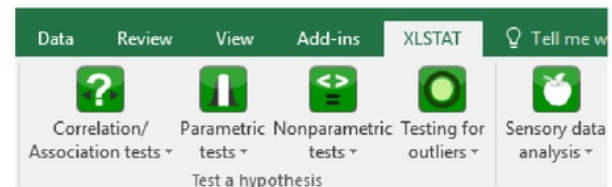
Pembahasan

1. Menentukan Nilai k

Dalam menentukan nilai k, peneliti menggunakan 10 nilai k antara lain : 7000, 8000, 6000, 5000, 4000, 3000, 800, 500, 100, dan 60.

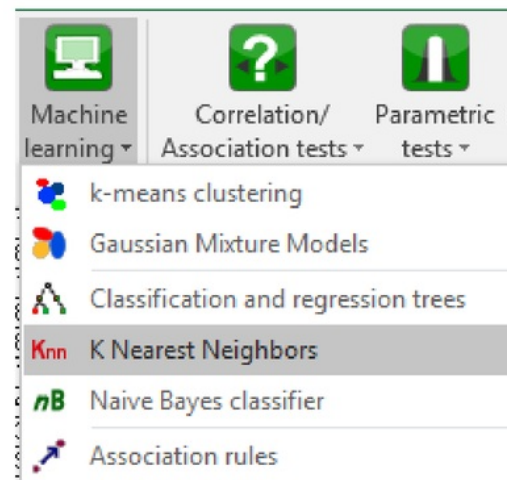
2. Membuat Klasifikasi Menggunakan XLSTAT

Untuk memulai menggunakan XLSTAT, klik tab “XLSTAT” pada menu ribbon Microsoft Excel.



Gambar 3 Menu Ribbon XLSTAT

Lalu pilih bagian “Machine Learning” dan klik pada “K Nearest Neighbor”.

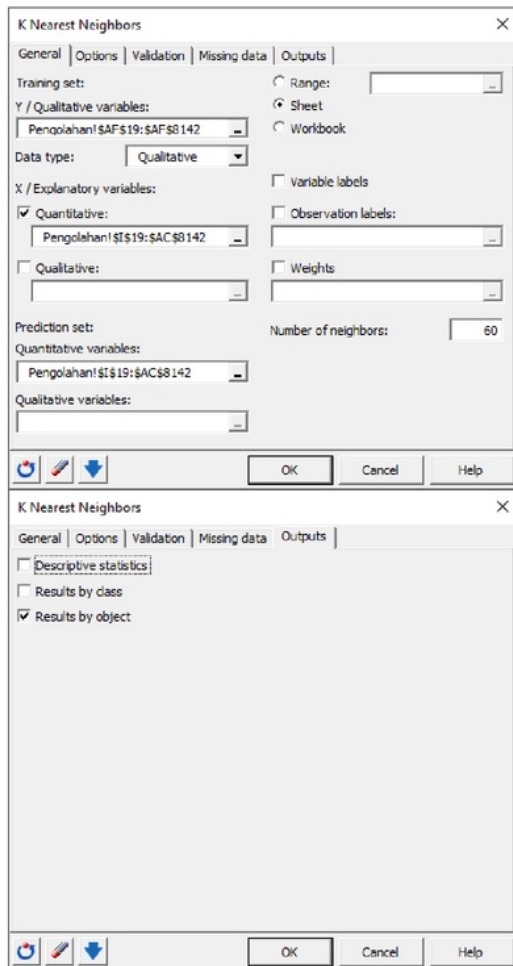


Gambar 4 Menu Machine Learning

Selanjutnya, pada Tab “General”, untuk “Y/Quantitative variables:” diisikan range label kelas data latih yang berupa data kuantitatif mulai dari data ke-1 sampai data ke-8124. Untuk “X/Explanatory variables:”, diisikan range antara sel



atribut pertama hingga sel atribut terakhir. Untuk “Number of neighbors” diisikan nilai k yang akan diujikan. Untuk Tab “Output”, yang dicentang hanyalah “Result by objects”, yang nantinya akan menghasilkan hasil tiap-tiap kelas dari data latih.



Gambar 5 Konfigurasi K-Nearest Neighbor

Hasil dari XLSTAT akan membuat sheet baru dengan isi prediksi label kelas

berdasarkan nilai k yang diinputkan sebelumnya.

Observation	Class
PredObs1	Beracun
PredObs2	Dapat Dimakan
PredObs3	Dapat Dimakan
PredObs4	Beracun
PredObs5	Dapat Dimakan
PredObs6	Dapat Dimakan
PredObs7	Dapat Dimakan
PredObs8	Dapat Dimakan
PredObs9	Beracun
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Beracun
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Dapat Dimakan
PredObs1	Beracun

Gambar 5 Prediksi Label Kelas

3. Menghitung Tingkat Akurasi

Untuk kasus nilai k=60, confusion tabel dijelaskan sebagaimana tabel berikut

Tabel 1 Confusion Table

		Prediksi	
		Beracun	Dapat Dimakan
Aktual	Beracun	3884	32
	Dapat Dimakan	50	4158

Akurasi dihitung berdasarkan rumus di bawah ini.

$$\text{Akurasi} = (\text{Jumlah beracun} + \text{jumlah dapat dimakan}) / \text{Jumlah total jamur}$$

$$\text{Akurasi} = (3884 + 4158) / 8124$$

$$\text{Akurasi} = 99\%$$

Jadi, dapat disimpulkan bahwa tingkat akurasi dari k=60 mencapai 99%.

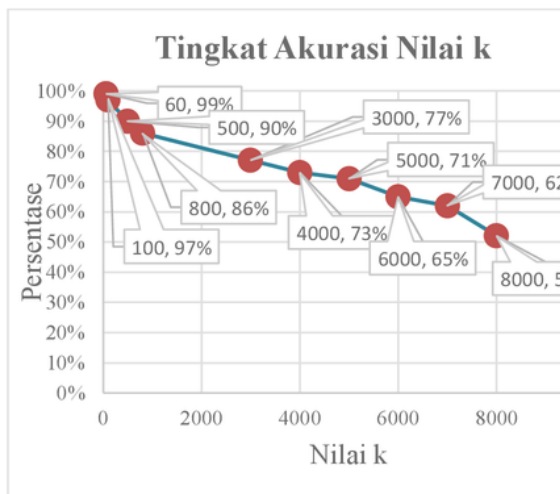


Dengan cara yang sama, tabel berikut menjelaskan tingkat akurasi dari bermacam-macam nilai k yang dapat digunakan sebagai perbandingan.

Tabel 2 Nilai k dan Akurasi

Nilai k	Akurasi
60	99%
100	97%
500	90%
800	86%
3000	77%
4000	73%
5000	71%
6000	65%
7000	62%
8000	52%

Adapun grafik antara nilai k dengan akurasi masing-masing k sebagai berikut.



Gambar 6 Akurasi Nilai k

4

KESIMPULAN DAN SARAN

Kesimpulan

Dari penelitian ini, dapat disimpulkan beberapa hal dalam identifikasi jamur menggunakan metode K-Nearest Neighbor dengan ekstraksi ciri morfologi, yaitu :

1. Metode klasifikasi K-Nearest Neighbor sangat baik dalam pengklasifikasian jamur.
2. Ekstraksi ciri morfologi yang menghasilkan akurasi tertinggi berada pada nilai k=60 yaitu mencapai 99%.
3. Dalam range antara k=8000 hingga k=60, tingkat akurasi mengalami kenaikan yang berarti semakin tinggi nilai k, maka akurasi semakin menurun.

Saran

Saran untuk penelitian selanjutnya adalah

1. Dengan data yang sama dapat dicoba dengan nilai k yang lebih beragam agar akurasi dapat melebihi 99%.
2. Obyek jamur dapat diperluas dengan menambahkan jenis jamur yang tidak hanya makroskopik, namun juga mikroskopik yang mana akan berpengaruh pada penambahan atribut.

REFERENSI

1. Gudang Biologi. 2015. Ciri-Ciri Morfologi dan Fisiologi Jamur (Fungi)
<http://www.gudangbiologi.com/201>



Seminar Nasional Sistem Informasi 2017, 14 September 2017
Fakultas Teknologi Informasi – UNMER Malang

5/12/ciri-ciri-morfologi-dan-

fisiologi-jamur-fungi.html

UC Irvine Machine Learning Repository.

1987. Mushroom Data Set

<https://archive.ics.uci.edu/ml/dataset/mushroom>

ts/mushroom

IDENTIFIKASI JAMUR MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN EKSTRAKSI CIRI MORFOLOGI

ORIGINALITY REPORT

11%

SIMILARITY INDEX

10%

INTERNET SOURCES

1%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1	lokaltuban.blogspot.com Internet Source	7%
2	www.webinfobits.com Internet Source	1%
3	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	1%
4	media.neliti.com Internet Source	1%
5	www.unwahas.ac.id Internet Source	1%
6	lingkaran-pengetahuan.blogspot.com Internet Source	<1%
7	nugashare.blogspot.co.id Internet Source	<1%
8	anandageo.blogspot.com Internet Source	<1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off