

Scientific Article

## IDENTIFIKASI OTOMATIS LIMA JENIS RESAK (*Vatica* spp.) BERDASARKAN BEBERAPA KARAKTER MORFOLOGI DAUN DAN ALGORITMA PEMBELAJARAN MESIN

*Automatic identification of five Vatica species (Vatica spp.) based on some leaf morphological characters and machine learning algorithms*

Muhammad Farhan Kurnia<sup>1</sup>, Siti Sumiati Solihat<sup>1</sup>, Gut Windarsih<sup>1</sup>, Didi Usmadi<sup>2\*</sup>

<sup>1</sup>Program Studi Biologi - Fakultas Sains - Universitas Islam Negeri Sultan Maulana Hasanuddin Banten

Jl. Syech Nawawi Al Bantani, Kp. Andamu'i, Kel. Sukawana, Kec. Curug, Kota Serang, Banten 42171

<sup>2</sup>Pusat Riset Ekologi dan Etnobiologi - Badan Riset dan Inovasi Nasional (BRIN)

Kawasan Sains dan Teknologi Soekarno, Jl. Raya Jakarta-Bogor Km. 46, Cibinong, Bogor, Jawa Barat 16911

### Informasi Artikel

Diterima/Received : 12 Desember 2022

Disetujui/Accepted : 14 April 2023

Diterbitkan/Published : 30 April 2023

\*Koresponden E-mail : didi020@brin.go.id

DOI: <https://doi.org/10.55981/bkr.2023.740>

### Cara mengutip

Kurnia MF, Solihat SS, Windarsih G, Usmadi D. 2023. Identifikasi otomatis lima jenis resak (*Vatica* spp.) berdasarkan beberapa karakter morfologi daun dan algoritma pembelajaran mesin. Buletin Kebun Raya 26(1): 26–37.

DOI: <https://doi.org/10.55981/bkr.2023.740>

### Kontributor

#### Kontributor Utama/Main author:

Muhammad Farhan Kurnia

Siti Sumiati Solihat

Gut Windarsih

Didi Usmadi

#### Kontributor Anggota/Author member:

-

**Keywords:** automatic identification, Dipterocarpaceae, leaf character, machine learning, *Vatica*

**Kata Kunci:** Dipterocarpaceae, identifikasi otomatis, karakter daun, pembelajaran mesin, *Vatica*

### Abstract

Resak (*Vatica* spp.) is a genus of the timber family, Dipterocarpaceae, with several species categorized as threatened. One of the most important components of conservation efforts is accurate species identification. This study aimed to identify the morphological traits of resak leaves, the similarity between species, and the effectiveness of five machine learning methods in automatically recognizing resak species. The measured morphological characters included the color, size, shape, and texture of the leaves of five species of *Vatica* growing at Bogor Botanic Garden. The differences in the mean value of each morphological character were analyzed using analysis of variance and the Tukey test. Morphological diversity and similarity were analyzed using Principal Component Analysis and Cluster Analysis. Automatic identification was conducted using five machine learning algorithms, namely BayesNet, K-Nearest Neighbor, Artificial Neural Network, Random Forest, and Support Vector Machine. The results showed that leaf morphological characters (color, size, shape, and texture) in five species of resak have significant differences. All morphological characters significantly affect the differences in resak leaf characteristics. At an 80% similarity level, the five species were grouped into three clusters, namely cluster I (*V. granulata*, *V. pauciflora*, *V. venulosa*), cluster II (*V. bantamensis*) and cluster III (*V. rassak*). The best machine learning algorithm for identifying resak species based on leaf morphological characters is K-Nearest Neighbor with an overall accuracy value of 0.92, a Kappa coefficient of 0.90, an average precision of 0.93, and an average recall of 0.92.

### Abstrak

Resak (*Vatica* spp.) merupakan salah satu marga yang termasuk dalam suku tumbuhan berkayu, Dipterocarpaceae, dengan beberapa jenis di antaranya termasuk jenis terancam. Kemampuan identifikasi jenis dengan benar merupakan salah satu aspek yang penting dalam upaya konservasi. Penelitian ini bertujuan untuk mengetahui karakter morfologi daun resak, kemiripan antar jenis, dan performa dari lima algoritma pembelajaran mesin dalam mengidentifikasi otomatis jenis resak. Karakter morfologi yang diukur adalah warna, ukuran, bentuk, dan tekstur daun dari lima jenis *Vatica* spp. koleksi Kebun Raya Bogor. Perbedaan nilai rata-rata setiap karakter morfologi dianalisis menggunakan analisis sidik ragam dan uji Tukey. Keragaman dan kemiripan morfologi dianalisis menggunakan analisis komponen utama dan analisis kluster. Identifikasi otomatis dilakukan menggunakan lima algoritma pembelajaran mesin, yaitu BayesNet, *K-Nearest Neighbor*, *Artificial Neural Network*, *Random Forest*, dan *Support Vector Machine*. Hasil analisis menunjukkan bahwa karakter morfologi daun (warna, ukuran, bentuk, dan tekstur) pada kelima jenis resak mempunyai perbedaan yang signifikan. Semua karakter morfologi secara signifikan mempengaruhi perbedaan penciri daun resak. Pada tingkat kemiripan 80%, kelima jenis resak dikelompokkan menjadi tiga kluster, yaitu kluster I (*V. granulata*, *V. pauciflora*, dan *V. venulosa*), kluster II (*V. bantamensis*), dan kluster III (*V. rassak*). Algoritma pembelajaran mesin yang terbaik dalam melakukan identifikasi otomatis jenis resak menggunakan karakter morfologi daun adalah *K-Nearest Neighbor* dengan nilai *overall accuracy* 0,92, koefisien Kappa 0,90, rata-rata *precision* 0,93, dan rata-rata *recall* 0,92.

## PENDAHULUAN

Dipterocarpaceae merupakan salah satu suku tumbuhan yang mendominasi hutan hujan tropis di Indonesia, terutama hutan tropis dataran rendah di Pulau Sumatra dan Kalimantan (Achmalidi 2001). Dipterocarpaceae dapat ditemukan pada daerah dataran rendah sampai ketinggian kurang dari 1.500 m dpl. (Rosinta et al. 2019). Suku ini terdiri atas 16 marga dengan jumlah anggota mencapai 539 jenis (Govaerts et al. 2021). Di Indonesia ditemukan sembilan marga dari suku Dipterocarpaceae, yaitu *Anisoptera*, *Cotylelobium*, *Dipterocarpus*, *Dryobalanops*, *Hopea*, *Parashorea*, *Shorea*, *Upuna*, dan *Vatica* (Ibrizah et al. 2018; Govaerts et al. 2021). Anggota suku ini yang tumbuh di Indonesia dikenal memiliki nilai ekonomi dan fungsi ekologi yang tinggi (Fajri 2008; Ngatiman & Saridan 2012; Petrus et al. 2021). Dipterocarpaceae merupakan penghasil kayu utama untuk ekspor dan pemenuhan kebutuhan dalam negeri (Saridan 2012), sedangkan produk non kayu berupa balsam, damar, lemak, buah, minyak atsiri, dan kapur barus (Widiyono 2021). Dominansi di hutan dataran rendah mengakibatkan Dipterocarpaceae mempunyai peran yang sangat besar dalam menjaga kestabilan ekosistem hutan tersebut (Wardani et al. 2017).

Jenis-jenis dari marga *Vatica* umumnya dikenal dengan nama resak. Resak dapat dimanfaatkan kayunya sebagai bahan pembuatan kapal, bahan bangunan, dan furnitur (Lemmens et al. 1994). Resak umumnya memiliki karakter kayu yang keras dan cukup awet dengan tingkat keawetan kelas III-V dan kelas kuat II (Widiati et al. 2023). Konversi hutan menjadi penggunaan lahan lain (pertanian, perumahan, dan pertambangan), perubahan iklim, kebakaran hutan, tumbuhan invasif, dan eksploitasi resak yang tinggi mengakibatkan keberadaan jenis-jenis tersebut di alam semakin berkurang dan berpotensi mengalami kepunahan. Beberapa jenis *Vatica* sudah masuk dalam daftar merah IUCN, diantaranya *Vatica bantamensis* (Hassk.) B.&H. ex Miq (CR = *Critically Endangered*), *V. venulosa* Blume (CR), dan *V. pauciflora* (Korth.) Blume (Vu = *Vulnerable*) (Asthon 1998; Pooma & Newman 2017; Robiansyah 2018). Oleh karena itu diperlukan upaya-upaya konservasi terhadap jenis-jenis tersebut.

Identifikasi jenis tumbuhan menjadi salah satu landasan penting dalam upaya konservasi jenis terancam kepunahan (Schmeller et al. 2008; Jalonen et al. 2014). Pengamatan ciri morfologi menjadi upaya pengenalan jenis tumbuhan yang umum dilakukan di lapangan (McLellan 2000). Analisis karakterisasi morfologi dalam identifikasi jenis dapat dilakukan menggunakan organ vegetatif dan generatif (Nasution 2016). Organ daun menjadi bagian yang umum digunakan dalam identifikasi morfologi, karena secara umum daun dimiliki oleh setiap jenis

tumbuhan, memiliki ketersediaan yang cukup banyak, dan terdapat ciri spesifik pada setiap jenis (Khalid et al. 2008). Identifikasi suatu jenis tumbuhan umumnya menggunakan spesimen herbarium (Esa et al. 2016). Keterbatasan akses terhadap spesimen herbarium, jumlah ahli botani yang terbatas, dan keterampilan serta pengetahuan ahli botani yang kurang merupakan masalah yang dihadapi dalam identifikasi jenis tumbuhan (Kho et al. 2017; Pujiati & Rochmawati 2022).

Seiring perkembangan teknologi, identifikasi suatu jenis tumbuhan dapat menggunakan program algoritma komputer yang dinilai lebih cepat dan memiliki akurasi yang tinggi (Yusniar & Kustiyo 2014; Ni'mah et al. 2018; Azadnia & Kheiralipour 2021). Algoritma pembelajaran mesin merupakan pendekatan pemodelan terkomputerisasi yang dapat mempelajari pola dari data, sehingga dapat melakukan tugas tertentu tanpa harus diprogram ulang berkali-kali (Singh et al. 2016). Pemanfaatan algoritma pembelajaran mesin telah banyak diterapkan untuk mengidentifikasi suku, jenis, dan varietas tumbuhan. Hussein et al. (2020) menggunakan *Random Forest* dan *Linear Discriminant Analysis* untuk mengidentifikasi jenis-jenis dari suku Dipterocarpaceae, Annonaceae, dan Euphorbiaceae. Algoritma *K-Nearest Neighbor* (KNN) digunakan untuk identifikasi *Shorea* spp. (Yusniar & Kustiyo 2014), tumbuhan obat (Ni'mah et al. 2018), dan varietas bugenvil (Furqan et al. 2020). Algoritma *Artificial Neural Networks* telah digunakan untuk identifikasi tumbuhan obat (Azadnia & Kheiralipour 2021), tumbuhan buah tropika (Agmalaro et al. 2013), *Ficus* spp. (Kho et al. 2017), dan anggrek (Nasution 2016). Algoritma *Support Vector Machine* telah digunakan untuk identifikasi *Ficus* spp. (Kho et al. 2017), dan tumbuhan obat (Begue et al. 2017).

Karakterisasi daun *Vatica* spp. telah dilakukan oleh beberapa peneliti. Srinual & Thammathaworn (2008) melakukan karakterisasi anatomi daun sembilan jenis *Vatica* yang tumbuh di Thailand berupa stomata, kelenjar resin, sel sklereid, serat sel-sel, dan jenis kristal dalam daun. Meekiong et al. (2012) meneliti karakter morfologi mikro dari daun *Vatica badiifolia* P.S.Ashton, *V. coriacea* P.S.Ashton, *V. globosa* P.S.Ashton, *V. pedicellata* Brandis, dan *V. sarawakensis* F.Heim dengan mengukur ukuran stomata, kelenjar resin, dan trikoma. Theera et al. (2012) menggunakan spesimen herbarium untuk melakukan karakterisasi morfologi daun *V. diospyroides* Symington berupa panjang dan lebar daun, serta panjang tangkai daun.

Penelitian ini difokuskan pada karakter morfologi daun berupa ukuran, bentuk, warna, dan tekstur daun dari lima jenis resak yang tumbuh di Kebun Raya Bogor. Identifikasi otomatis resak yang tumbuh di Indonesia dengan menggunakan berbagai algoritma pembelajaran mesin hingga saat ini belum pernah dilakukan. Penelitian

ini bertujuan untuk mengetahui karakter morfologi daun dari lima jenis resak dan kemiripan antar jenisnya, serta performa dari lima macam algoritma pembelajaran mesin dalam pengidentifikasian jenis resak.

## BAHAN DAN METODE

### Lokasi dan waktu penelitian

Penelitian ini dilakukan di Kebun Raya Bogor yang terletak di Kelurahan Paledang, Kecamatan Bogor Tengah, Kota Bogor, Provinsi Jawa Barat pada bulan Oktober–November 2022. Kota Bogor mempunyai ketinggian tempat 190–330 m dpl, beriklim tropis dengan suhu udara rata-rata 24,60–26,80 °C, kelembapan udara rata-rata 80–86%, dan curah hujan tahunan 4.212 mm (BPS 2022).

### Alat dan bahan

Bahan yang digunakan dalam kegiatan ini berupa daun dari lima jenis resak koleksi Kebun Raya Bogor, yaitu *Vatica bantamensis* (Hassk.) Benth. & Hook.f. ex Miq., *V. granulata* Slooten, *V. pauciflora* (Korth.) Blume, *V. rassak* (Korth.) Blume, dan *V. venulosa* Blume (Tabel 1). Sampel dari setiap jenis *Vatica* terdiri atas dua individu, sehingga total koleksi tanaman yang digunakan dalam penelitian ini sebanyak 10 tanaman. Alat yang digunakan berupa kamera *smartphone* Samsung Galaxy A22, penggaris, dan *photobox*.

**Tabel 1.** Tanaman koleksi Kebun Raya Bogor yang diambil sampel daun

No.	Nama jenis	Vak	Nomor koleksi	Umur (tahun)
1	<i>Vatica bantamensis</i>	I.K	63a	32
2	<i>Vatica bantamensis</i>	I.K	63b	32
3	<i>Vatica granulata</i>	VII.B	150b	7
4	<i>Vatica granulata</i>	VI.C	312c	33
5	<i>Vatica pauciflora</i>	XXV.A	182a	62
6	<i>Vatica pauciflora</i>	XXV.A	182b	62
7	<i>Vatica rassak</i>	XIX	34a	7
8	<i>Vatica rassak</i>	XIX	34b	7
9	<i>Vatica venulosa</i>	XXV.A	235a	27
10	<i>Vatica venulosa</i>	XXV.A	235b	27

### Tahapan pelaksanaan penelitian

Pengambilan sampel daun menggunakan metode *purposive sampling* (Arini et al. 2020). Daun yang dipilih sebagai sampel adalah daun yang dewasa, segar (baru dipetik), bentuk utuh (tidak ada bagian daun yang hilang, berlubang, atau robek), dan terbebas dari hama serta penyakit. Selain itu, sampel daun yang diambil mewakili daun yang terkena cahaya matahari langsung (kanopi bagian atas) dan terlindungi (di bawah kanopi). Sampel daun dikumpulkan sebanyak 25 helai daun dari setiap

individu tanaman, sehingga total sampel daun yang digunakan sebanyak 250 sampel.

Ruang untuk pengambilan citra sampel daun menggunakan *photobox* yang dilengkapi lampu LED berwarna putih. *Photobox* membantu penyinaran lebih merata dan konstan serta mengurangi adanya bayangan objek, sehingga kondisi ruang pengambilan citra menjadi seragam pada semua sampel daun. Sampel daun tanpa tangkai daun (*petiole*) disusun dengan rapi pada *photobox* dengan latar belakang kertas berwarna putih dan diletakkan penggaris pada salah satu sisi. Pengambilan citra sampel daun dilakukan dari lubang bagian atas *photobox* dengan jarak 32 cm dari atas sampel daun. Sampel daun difoto menggunakan kamera dengan resolusi 3000×4000 piksel dan direkam dalam format ekstensi *Joint Photographic Group* (\*.jpg).

Parameter yang diukur dari setiap citra sampel daun resak berupa morfologi daun yang mencakup warna, ukuran, bentuk, dan tekstur daun. Jumlah karakter morfologi daun yang diukur sebanyak 15 karakter. Karakter warna mencakup warna merah, hijau, dan biru (*Red Green Blue/RGB*). Karakter ukuran mencakup panjang, lebar, luas, dan keliling. Panjang daun diukur dari pangkal sampai ke ujung daun, sedangkan lebar daun diukur pada titik terlebar daun. Karakter bentuk mencakup *circularity* (sirkularitas), *aspect ratio* (rasio panjang terhadap lebar), *round* (kebulatan), dan *solidity* (soliditas). Karakter tekstur daun yang digunakan mencakup *Gray Level Co-Occurrence Matrix* (GLCM). GLCM merupakan suatu matriks yang merepresentasikan hubungan ketetanggaan antar-piksel dalam gambar pada berbagai arah orientasi sudut dan jarak (Haralick et al. 1973). Tekstur GLCM yang digunakan dalam penelitian ini berupa kontras, korelasi, *Inverse Defference Moment* (IDM), dan entropi. Perhitungan nilai tekstur daun menggunakan nilai *window size* sebesar 1 piksel dan arah perhitungan berupa 0°. Pengukuran karakter morfologi daun dilakukan menggunakan perangkat lunak ImageJ versi 1.35t (Schneider et al. 2012).

### Analisis data

Nilai rata-rata setiap karakter morfologi daun pada masing-masing jenis resak dianalisis dengan menggunakan analisis sidik ragam (*Analysis of Variance/ANOVA*). Apabila hasil analisis sidik ragam menunjukkan perbedaan rata-rata yang signifikan, maka dilakukan uji lanjut menggunakan uji Tukey berpasangan (McHugh 2011). Keragaman morfologi pada daun resak dianalisis menggunakan analisis komponen utama (*Principal Component Analysis/PCA*). Analisis komponen utama dilakukan untuk mengetahui karakter yang memiliki nilai kontribusi tinggi terhadap variasi data dan karakter yang menjadi penciri suatu jenis (Setiawati 2013; Rosdayanti et al. 2019).

Kemiripan antar jenis resak berdasarkan karakter morfologi daun ditentukan menggunakan analisis kluster (*Cluster Analysis/CA*). Analisis kluster dilakukan untuk mengelompokkan berbagai jenis menjadi beberapa kelompok berdasarkan kesamaan karakter yang dimiliki (Mattjik & Sumertajaya 2011). Pembuatan dendogram dilakukan menggunakan algoritma UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*), di mana penghitungan indeks kesamaan (*similarity index*) dilakukan menggunakan indeks Bray-Curtis (Hammer 2022). Algoritma UPGMA dan indeks Bray-Curtis telah umum dan populer digunakan untuk dalam pembuatan dendogram dan analisis indeks kesamaan pada data yang bersifat kontinu (Hammer 2022). Analisis sidik ragam, analisis komponen utama, dan analisis kluster dilakukan menggunakan perangkat lunak PAST (PAleontological STatistics) versi 4.11 (Hammer et al. 2001).

Seleksi karakter morfologi yang digunakan untuk identifikasi otomatis dilakukan menggunakan metode *Correlation-based Feature Selection* (CFS) dengan metode pencarian berupa Best First (Hall 2000; Bouckaert et al. 2016). Metode seleksi CFS menghasilkan hasil akurasi yang tinggi dalam mengidentifikasi sel leukemia dan kanker payudara (Kumar et al. 2017). Metode seleksi tersebut memilih kumpulan variabel (*feature subsets*) yang mempunyai korelasi antar variabel yang rendah, namun mempunyai korelasi yang tinggi antara variabel dan target klasifikasi (Hall 1998). Kumpulan variabel yang terpilih adalah kumpulan variabel yang menghasilkan nilai *merit* tertinggi (Hall 2000). Karakter morfologi hasil seleksi selanjutnya digunakan dalam penyusunan dan evaluasi model. Data karakter morfologi daun resak dibagi secara acak menjadi data latih dan data uji. Kombinasi persentase data latih dan data latih diujicobakan, sehingga diperoleh nilai kombinasi persentase data yang menghasilkan performa model yang tertinggi. Dari hasil uji coba model diperoleh kombinasi terbaik dari data morfologi tersebut yaitu 80% (200 data latih) digunakan untuk membangun model dan 20% (50 data uji) digunakan untuk memvalidasi model yang dihasilkan (*model validation*).

Identifikasi otomatis jenis resak berdasarkan karakter morfologi daun dilakukan dengan menggunakan lima algoritma pembelajaran mesin, yaitu BayesNet, *K-Nearest Neighbor* (KNN), *Artificial Neural Network* (ANN), *Random Forest* (RF), dan *Support Vector Machine* (SVM). Nilai semua parameter pada masing-masing algoritma pembelajaran mesin ditentukan berdasarkan uji coba yang menghasilkan performa model yang tertinggi. Hasil uji coba menghasilkan nilai parameter yang digunakan pada masing-masing algoritma yaitu Bayesnet (*default*), KNN ( $k = 3$ ), ANN (*learning rate* = 0,3, *momentum* = 0,2), RF (*seed* = 4), dan SVM ( $c = 1$ , *kernel* = *polykernel*).

Performa masing-masing algoritma pembelajaran mesin dalam mengidentifikasi otomatis jenis resak

berdasarkan empat indikator evaluasi model yaitu *overall accuracy*, koefisien *Kappa*, rata-rata *precision*, dan rata-rata *recall*. Makin tinggi nilai indikator evaluasi model, maka makin baik model yang dihasilkan. Seleksi variabel, analisis model identifikasi (klasifikasi), dan uji performa model dilakukan menggunakan perangkat lunak WEKA versi 3.9.6 (Frank et al. 2016).

## HASIL DAN PEMBAHASAN

### Karakter morfologi daun resak

Berdasarkan hasil analisis sidik ragam diketahui bahwa karakter morfologi berupa warna, ukuran, bentuk, dan tekstur daun pada kelima jenis resak mempunyai perbedaan yang signifikan ( $F_{(4,245)} = 6,61-240,40$ ;  $P < 0,001$ ) (Tabel 2). Daun *Vatica pauciflora* mempunyai warna merah dan hijau yang tertinggi dengan nilai masing-masing sebesar  $60,79 \pm 16,12$  dan  $94,46 \pm 21,36$ . Dari hasil analisis uji Tukey diketahui warna merah *V. pauciflora* berbeda signifikan dengan *V. bantamensis* ( $P = 0,000$ ; 95% CI = 14,62; 27,63), *V. rassak* ( $P = 0,000$ ; 95% CI = 10,27; 23,27), dan *V. granulata* ( $P = 0,000$ ; 95% CI = 4,93; 17,94) serta tidak berbeda nyata dengan *V. venulosa* ( $P = 0,064$ ; 95% CI = -0,22; 12,78). Adapun warna biru relatif sama pada keempat resak, kecuali *V. venulosa* yang berbeda signifikan dengan jenis lain dan mempunyai nilai terendah  $17,98 \pm 12,56$ . Dari hasil kombinasi warna merah, hijau, dan biru dapat diketahui bahwa warna daun resak dapat dibedakan menjadi tiga warna, yaitu hijau gelap (*V. bantamensis* dan *V. rassak*), hijau (*V. granulata* dan *V. venulosa*), dan hijau terang (*V. pauciflora*) (Gambar 1).

*Vatica rassak* mempunyai panjang, lebar, luas, dan keliling daun yang terbesar dengan nilai masing-masing berurutan sebesar  $27,85 \pm 4,68$  cm,  $9,82 \pm 1,66$  cm,  $206,53 \pm 64,13$  cm<sup>2</sup>, dan  $66,53 \pm 11,04$  cm. Hasil pengukuran tersebut sesuai dengan hasil Lemmens et al. (1994) dan Kalima & Wardani (2017), di mana *V. rassak* memiliki daun dengan panjang 13–32 cm dan lebar 5–11 cm. *V. bantamensis* mempunyai ukuran daun yang terkecil dengan nilai panjang, lebar, luas, dan keliling berurutan sebesar  $11,54 \pm 2,71$  cm,  $3,88 \pm 0,93$  cm,  $31,37 \pm 14,16$  cm<sup>2</sup>, dan  $27,08 \pm 6,63$  cm. Menurut Kalima & Wardani (2017), *V. bantamensis* memiliki panjang dan lebar daun masing-masing sebesar 7,5–18,0 cm dan 3,5–7,5 cm. Dari hasil uji Tukey diketahui bahwa *V. bantamensis* tidak berbeda nyata dengan *V. venulosa* pada karakter panjang ( $P = 0,784$ ; 95% CI = -1,05, 2,254), lebar ( $P = 0,729$ ; 95% CI = -0,88, 0,34), luas ( $P = 0,852$ ; 95% CI = -25,12, 11,64), dan keliling ( $P = 0,603$ ; 95% CI = -7,05, 2,20). Hasil analisis menunjukkan urutan ukuran daun resak dari yang terbesar ke terkecil, yaitu *V. rassak* > *V. granulata* > *V. pauciflora* > *V. venulosa* > *V. bantamensis*.

**Tabel 2.** Rata-rata dan standar deviasi dari karakter morfologi pada lima jenis resak koleksi Kebun Raya Bogor

Karakter	<i>V. bantamensis</i>	<i>V. granulata</i>	<i>V. pauciflora</i>	<i>V. rassak</i>	<i>V. venulosa</i>	Nilai F <sub>hitung</sub>
Merah	39,66±4,72 <sup>a</sup>	49,35±8,33 <sup>bc</sup>	60,79±16,12 <sup>d</sup>	44,01±10,32 <sup>ab</sup>	54,51±15,55 <sup>cd</sup>	24,93
Hijau	61,07±6,53 <sup>a</sup>	78,39±10,66 <sup>b</sup>	94,46±21,36 <sup>c</sup>	67,05±12,76 <sup>a</sup>	81,27±19,48 <sup>b</sup>	36,64
Biru	22,81±5,49 <sup>b</sup>	25,04±6,68 <sup>b</sup>	24,49±11,57 <sup>b</sup>	26,61±6,90 <sup>b</sup>	17,98±12,56 <sup>a</sup>	6,61
Panjang (cm)	11,54±2,71 <sup>a</sup>	23,37±3,15 <sup>c</sup>	16,65±3,42 <sup>b</sup>	27,85±4,68 <sup>d</sup>	12,29±1,56 <sup>a</sup>	240,4
Lebar (cm)	3,88±0,93 <sup>a</sup>	6,14±0,89 <sup>c</sup>	4,80±1,06 <sup>b</sup>	9,82±1,66 <sup>d</sup>	4,16±0,79 <sup>a</sup>	236,1
Luas (cm <sup>2</sup> )	31,37±14,16 <sup>a</sup>	103,08±26,44 <sup>c</sup>	57,53±21,73 <sup>b</sup>	206,53±64,13 <sup>d</sup>	38,11±10,32 <sup>a</sup>	233,6
Keliling (cm)	27,08±6,63 <sup>a</sup>	53,38±6,97 <sup>c</sup>	38,88±8,13 <sup>b</sup>	66,53±11,04 <sup>d</sup>	29,51±4,37 <sup>a</sup>	192,0
Circularity	0,52±0,05 <sup>b</sup>	0,45±0,06 <sup>a</sup>	0,47±0,06 <sup>a</sup>	0,57±0,05 <sup>c</sup>	0,55±0,08 <sup>c</sup>	33,96
Aspect ratio	2,78±0,21 <sup>a</sup>	3,59±0,43 <sup>c</sup>	3,21±0,41 <sup>b</sup>	2,72±0,22 <sup>a</sup>	2,89±0,44 <sup>a</sup>	51,12
Round	0,36±0,03 <sup>c</sup>	0,28±0,04 <sup>a</sup>	0,32±0,04 <sup>b</sup>	0,37±0,03 <sup>c</sup>	0,35±0,06 <sup>c</sup>	40,86
Solidity	0,97±0,01 <sup>ab</sup>	0,98±0,01 <sup>bc</sup>	0,97±0,01 <sup>a</sup>	0,98±0,01 <sup>d</sup>	0,98±0,02 <sup>cd</sup>	15,86
Kontras	88,14±54,14 <sup>c</sup>	29,60±19,91 <sup>ab</sup>	21,68±18,07 <sup>a</sup>	42,01±20,08 <sup>b</sup>	34,51±22,26 <sup>ab</sup>	37,68
Korelasi	0,17±0,01 <sup>a</sup>	0,20±0,03 <sup>bc</sup>	0,17±0,03 <sup>a</sup>	0,19±0,04 <sup>b</sup>	0,21±0,03 <sup>c</sup>	14,07
IDM	0,49±0,04 <sup>a</sup>	0,57±0,05 <sup>c</sup>	0,52±0,05 <sup>ab</sup>	0,54±0,04 <sup>b</sup>	0,50±0,05 <sup>a</sup>	22,49
Entropi	6,60±0,37 <sup>c</sup>	6,22±0,32 <sup>a</sup>	6,44±0,39 <sup>bc</sup>	6,34±0,27 <sup>ab</sup>	6,52±0,40 <sup>c</sup>	8,88

Keterangan : IDM = Inverse Difference Moment, angka yang diikuti oleh huruf yang sama pada baris yang sama menunjukkan tidak berbeda nyata berdasarkan uji Tukey,  $df_{\text{between groups}} = 4$ ,  $df_{\text{within groups}} = 245$ ;  $p \text{ value} < 0,01$ .



**Gambar 1.** Citra daun resak koleksi Kebun Raya Bogor: (a) *V. bantamensis*, (b) *V. granulata*, (c) *V. pauciflora*, (d) *V. rassak*, dan (e) *V. venulosa*

Jika nilai *circularity* makin mendekati 1, maka bentuk daun makin mendekati bulat (lingkaran). Sebaliknya, makin tinggi nilai *aspect ratio* mengindikasikan bentuk daun makin melanset (Ferreira & Rasband 2012). Nilai *round* mirip dengan nilai *circularity*, namun lebih dipengaruhi oleh nilai AR, sehingga makin mendekati 1, maka bentuk daun makin mendekati bulat (Zdilla et al. 2016). Sedangkan nilai *solidity* makin mendekati 1, maka bentuk daun makin mendekati bidang cembung atau cekung (Ferreira & Rasband 2012; Zdilla et al. 2016).

Berdasarkan hasil uji Tukey diketahui bahwa karakter *aspect ratio* dan *round* pada *V. bantamensis*, *V. rassak*, dan *V. venulosa* tidak berbeda nyata ( $P > 0,05$ ). Ketiga jenis tersebut cenderung berbentuk bulat memanjang dengan nilai *circularity* = 0,52–0,57, *aspect*

*ratio* = 2,72–2,89, dan *round* = 0,35–0,37. Adapun bentuk daun *V. pauciflora* dan *V. granulata* cenderung memanjang sampai melanset dengan nilai *circularity* = 0,45–0,47, *aspect ratio* = 3,21–3,59, dan *round* = 0,28–0,32. Menurut Lemmens et al. (1994) dan Kalima & Wardani (2017), bentuk daun resak sebagai berikut: elips atau melanset (*V. bantamensis*), memanjang sampai elips menyempit (*V. rassak*), elips sampai bulat telur-melanset (*V. venulosa*), elips sampai melanset (*V. pauciflora*), dan bulat telur sungsang menyempit (*V. granulata*). Kelima jenis resak yang diamati mempunyai nilai *solidity* yang hampir sama (0,97–0,98), yang mengindikasikan tepi daun berbentuk cembung.

Berdasarkan hasil uji Tukey diketahui bahwa *V. bantamensis*, *V. pauciflora*, dan *V. venulosa* mempunyai

karakter entropi mempunyai nilai yang tidak berbeda nyata. Berdasarkan Tabel 2 diketahui bahwa *V. bantamensis* memiliki kontras dan entropi tertinggi, namun nilai IDM terendah dengan nilai masing-masing  $88,14 \pm 54,00$ ;  $6,60 \pm 0,37$ ; dan  $0,49 \pm 0,04$ . Ketiga nilai tersebut menunjukkan bahwa citra daun *V. bantamensis* mempunyai nilai elemen piksel yang tidak seragam, perbedaan nilai elemen piksel yang tinggi, dan cenderung tidak teratur (acak). Adapun citra daun *V. granulata* cenderung mempunyai nilai elemen piksel yang relatif seragam dan lebih teratur yang terlihat dari nilai IDM yang tinggi ( $0,57 \pm 0,05$ ) dan entropi yang rendah ( $6,22 \pm 0,32$ ). Berdasarkan nilai tekstur tersebut dapat diketahui bahwa *V. bantamensis* cenderung mempunyai tekstur citra daun yang kasar, sedangkan *V. granulata* mempunyai tekstur citra daun yang halus.

### Kemiripan jenis resak

Pemilihan komponen utama dapat dilakukan berdasarkan nilai *eigen value*. *Eigen value* menunjukkan seberapa besar varian data yang dapat dijelaskan oleh komponen utama (Jolliffe 2011). Makin tinggi nilai *eigen value*, makin tinggi pula varian sepanjang arah matriks kovarians/korelasi. Berdasarkan hasil analisis komponen utama diketahui terdapat empat komponen utama (KU) dengan *eigen value* >1 (Tabel 3). KU 1 dan KU 2 dapat menjelaskan varian data yang tinggi masing-masing sebesar 28,86% dan 22,69% dari varian karakter morfologi resak. Keempat komponen utama tersebut mampu menjelaskan 78,28% dari varian karakter morfologi resak.

**Tabel 3.** *Eigen value* dan varian pada masing-masing komponen utama

Komponen Utama	<i>Eigen value</i>	Varian (%)	Akumulasi varian (%)
1	4,33	28,86	28,86
2	3,40	22,69	51,55
3	2,13	14,21	65,76
4	1,88	12,52	78,28

Hasil analisis komponen utama juga menghasilkan nilai *loadings*. Nilai *loadings* menunjukkan besarnya korelasi dari masing-masing variabel asal dengan variabel baru (komponen utama) yang terbentuk (Hammer 2022). Nilai *loadings* dapat bernilai positif maupun negatif, di mana nilai positif menunjukkan bahwa variabel tersebut mempunyai hubungan positif dengan komponen utama, sedangkan nilai negatif menunjukkan hubungan negatif dengan komponen utama. Signifikansi nilai *loadings* dipengaruhi oleh jumlah sampel yang digunakan (Hair et al. 2010). Dengan jumlah sampel sebanyak 250 sampel, nilai *loadings* lebih dari 0,35 mengindikasikan variabel tersebut

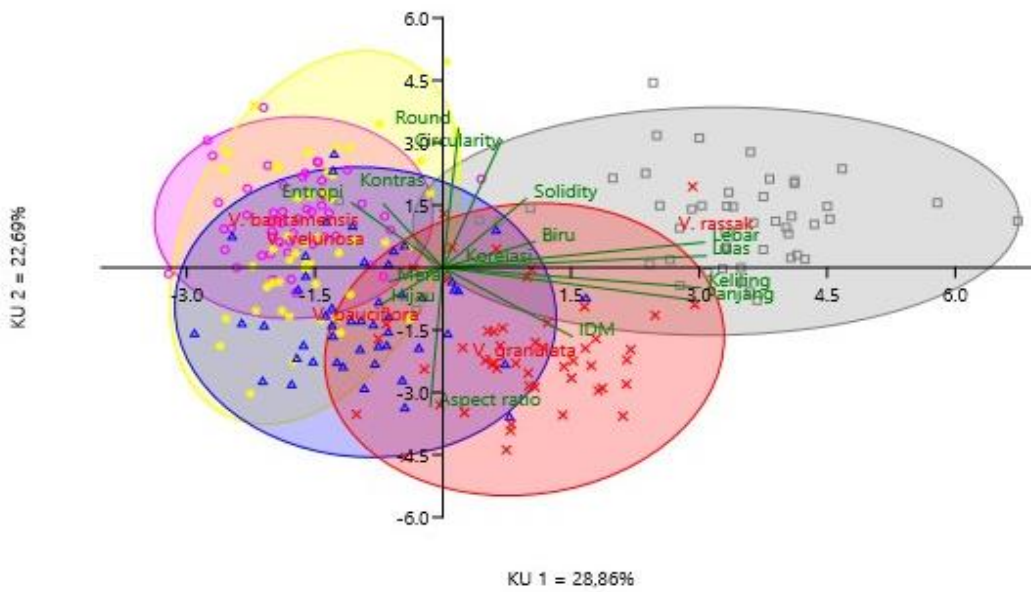
secara signifikan berpengaruh terhadap komponen utama (Hair et al. 2010). Berdasarkan hasil analisis komponen utama diketahui bahwa semua karakter morfologi mempunyai nilai *loadings* lebih dari 0,35 pada komponen utama (Tabel 4). Nilai tersebut mengindikasikan bahwa semua karakter morfologi secara signifikan mempengaruhi perbedaan penciri morfologi daun resak. Komponen utama 1 (KU 1) terdiri dari lima karakter, yaitu panjang, lebar, luas, keliling daun, dan IDM. Komponen utama 2 (KU 2) terdiri dari empat karakter, yaitu *circularity*, *aspect ratio*, *round*, dan *solidity*. Komponen utama 3 (KU 3) terdiri dari dua karakter, yaitu warna merah dan hijau, sedangkan komponen utama 4 (KU 4) terdiri dari empat karakter, yaitu warna biru, kontras, korelasi, dan entropi. KU 1 mewakili ukuran daun, KU 2 mewakili karakter bentuk daun, KU 3 mewakili karakter warna daun, dan KU 4 mewakili karakter tekstur daun.

Gambar 2 menunjukkan hubungan yang positif dan negatif. Karakter panjang, lebar, luas, dan keliling daun mempunyai hubungan yang positif antar karakter morfologi lima jenis resak yang diamati, di mana keempat karakter tersebut membentuk sudut lancip (sudut < 90°). Adapun antara karakter warna merah dan hijau mempunyai korelasi yang positif, namun kedua karakter tersebut berkorelasi negatif terhadap warna biru (sudut > 90°). Karakter bentuk juga mengindikasikan hal yang serupa, di mana *circularity*, *solidity*, dan *round* mempunyai korelasi yang positif, namun ketiga karakter tersebut berkorelasi negatif dengan *aspect ratio*. Pada karakter tekstur diketahui bahwa entropi dan kontras berkorelasi positif, namun keduanya berkorelasi negatif dengan IDM. Berdasarkan Gambar 2 juga dapat diketahui bahwa *V. bantamensis* mempunyai kemiripan dengan *V. venulosa* berdasarkan karakter warna merah, warna hijau, kontras, dan entropi. Adapun *V. rassak* cenderung berbeda dengan jenis lainnya, di mana jenis ini mempunyai korelasi yang positif dengan karakter warna biru, panjang, lebar, luas, keliling, *sircularity*, *solidity*, *round*, korelasi, dan IDM.

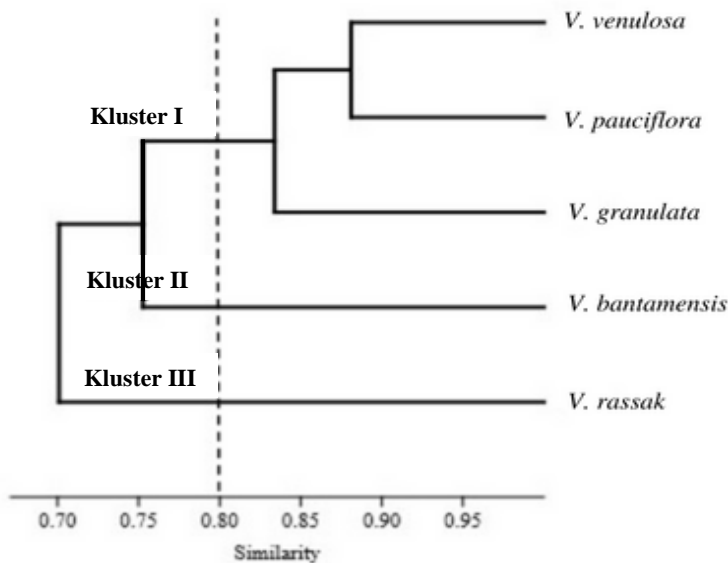
Hasil analisis kluster berupa dendrogram kemiripan jenis resak disajikan pada Gambar 3. Pada tingkat kemiripan 80%, kelima jenis resak dikelompokkan menjadi tiga kluster. Kluster I terdiri dari *V. venulosa*, *V. pauciflora*, dan *V. granulata*. Ketiga jenis resak tersebut secara morfologi memiliki ciri warna hijau sampai hijau terang, ukuran daun di antara *V. bantamensis* dan *V. rassak*, bentuk memanjang sampai lanset, dan tekstur yang lebih seragam. Kluster II yaitu *V. bantamensis* yang memiliki ciri warna hijau gelap, ukuran daun yang kecil, bentuk bulat memanjang, dan tekstur tidak seragam (acak). Kluster III yaitu *V. rassak* dengan ciri warna hijau, ukuran daun besar, bentuk bulat memanjang, dan tekstur cukup seragam.

**Tabel 4.** Nilai *loadings* karakter morfologi pada masing-masing komponen utama (KU)

Karakter	KU 1	KU 2	KU 3	KU 4
Merah	-0,20	-0,09	<b>0,93</b>	0,01
Hijau	-0,22	-0,24	<b>0,87</b>	-0,02
Biru	0,34	0,17	0,17	<b>0,57</b>
Panjang	<b>0,93</b>	-0,22	0,11	0,20
Lebar	<b>0,96</b>	0,17	0,11	0,11
Luas	<b>0,96</b>	0,08	0,08	0,11
Keliling	<b>0,94</b>	-0,13	0,11	0,18
Circularity	0,21	<b>0,85</b>	0,10	-0,38
Aspect ratio	-0,05	<b>-0,92</b>	-0,01	0,15
Round	0,06	<b>0,92</b>	0,03	-0,15
Solidity	0,30	<b>0,46</b>	0,09	-0,34
Kontras	-0,22	0,42	-0,30	<b>0,52</b>
Korelasi	0,06	0,03	0,44	<b>-0,51</b>
Inverse Difference Moment	<b>0,47</b>	-0,46	-0,21	-0,45
Entropi	-0,33	0,43	0,32	<b>0,64</b>



**Gambar 2.** Hubungan antar karakter morfologi berdasarkan komponen utama 1 dan 2



**Gambar 3.** Dendrogram kemiripan lima jenis resak koleksi Kebun Raya Bogor

### Identifikasi otomatis jenis resak

Seleksi variabel yang digunakan dalam identifikasi otomatis menggunakan pembelajaran mesin dilakukan dalam rangka mengurangi dimensi data dan waktu proses serta dapat meningkatkan kinerja algoritma pembelajaran mesin (Al Janabi & Kadhim 2018). Metode seleksi CFS menggunakan korelasi Pearson dalam menghitung hubungan antar variabel prediksi dan variabel prediksi dengan target klasifikasi (Hall 2000). Secara umum, korelasi antar karakter warna, bentuk, dan ukuran daun mempunyai korelasi yang tinggi ( $r > 0,80$ ), sedangkan antar karakter tekstur mempunyai korelasi yang lebih rendah ( $r < 0,4$ ) (Tabel 5). Dilihat dari korelasi antara karakter morfologi dan jenis resak diketahui karakter *aspect ratio*,

*round*, dan kontras mempunyai korelasi yang paling tinggi ( $r > 0,5$ ).

Hasil seleksi variabel metode seleksi CFS dengan metode pencarian berupa *Best First* diperoleh nilai *merit* tertinggi, yaitu 0,66. Nilai tersebut mengindikasikan bahwa kumpulan karakter morfologi yang terpilih mempunyai korelasi dengan jenis resak sebesar 0,66. Kumpulan variabel hasil seleksi tersebut terdiri atas 11 karakter morfologi daun yaitu warna hijau, warna biru, panjang, lebar, luas, *round*, *solidity*, kontras, korelasi, IDM, dan entropi. Kesebelas karakter tersebut digunakan dalam identifikasi otomatis jenis resak menggunakan lima algoritma pembelajaran mesin.

**Tabel 5.** Korelasi Pearson antar karakter morfologi dan karakter morfologi dengan jenis resak

Karakter	Hijau	Biru	Panjang	Lebar	Luas	Keliling	Cir.	AR	Round	Solidity	Kontras	Korelasi	IDM	Entropi	Jenis <i>Vatica</i>
Merah	0,92**	0,14*	-0,08	-0,11	-0,14*	-0,09	-0,01	0,06	-0,03	-0,05	-0,19**	0,26**	-0,13*	0,23**	0,37**
Hijau		-0,09	-0,05	-0,13*	-0,16*	-0,07	-0,15*	0,16*	-0,16*	-0,13*	-0,23**	0,18**	-0,11	0,14*	0,46**
Biru			0,33**	0,33**	0,31**	0,33**	0,09	-0,05	0,08	0,11	0,15*	-0,18**	-0,02	0,33**	0,04
Panjang				0,91**	0,93**	0,96**	-0,06	0,20**	-0,18**	0,12	-0,19**	0,01	0,39**	-0,23**	0,25**
Lebar					0,98**	0,92**	0,30**	-0,21**	0,22**	0,27**	-0,10	0,05	0,26**	-0,14*	-0,03
Luas						0,93**	0,23**	-0,11	0,11	0,25**	-0,14*	0,05	0,28**	-0,18**	-0,01
Keliling							0,00	0,09	-0,08	0,13*	-0,18**	0,04	0,36**	-0,20**	0,19**
Cir.								-0,82**	0,82**	0,62**	0,07	0,19**	-0,11	0,05	-0,43**
AR									-0,98**	-0,28**	-0,27**	-0,00	0,28**	-0,23**	0,62**
Round										0,28**	0,27**	0,04	-0,27**	0,24**	-0,58**
Solidity											0,00	0,24**	0,04	-0,02	-0,12
Kontras												-0,33**	-0,37**	0,38**	-0,51**
Korelasi													0,01	-0,05	0,17**
IDM														-0,79**	0,34**
Entropi															-0,25**

Keterangan : Tanda \* menunjukkan signifikan pada taraf 5%, tanda \*\* menunjukkan signifikan pada taraf 1%, dan tanpa tanda menunjukkan tidak signifikan, Cir. = Circularity, AR = Aspect ratio, IDM = Inverse Difference Moment

Performa model klasifikasi pada pembelajaran mesin dapat dilihat dari beberapa indikator evaluasi. *Accuracy* menunjukkan tingkat ketepatan antara hasil prediksi dan aktual (Musfiroh et al. 2021), sedangkan koefisien Kappa mengukur tingkat kesesuaian antara nilai prediksi dan nilai yang diharapkan (Delgado & Tibau 2019). Menurut Pratiwi & Ulama (2016), *precision* menggambarkan tingkat ketepatan antara informasi yang diminta dengan hasil prediksi, sedangkan *recall* menggambarkan ketepatan model untuk mengklasifikasikan data dengan benar.

Berdasarkan Tabel 6 dapat diketahui bahwa identifikasi jenis resak menggunakan lima algoritma menghasilkan performa yang cukup baik. Algoritma BayesNet menghasilkan performa model yang paling rendah (*overall accuracy* = 0,80, koefisien Kappa = 0,75, rata-rata *precision* = 0,81, dan rata-rata *recall* = 0,80). Algoritma yang terbaik dalam identifikasi jenis resak menggunakan karakter morfologi daun adalah KNN (*overall accuracy* = 0,92, koefisien Kappa = 0,90, rata-rata *precision* = 0,93, dan rata-rata *recall* = 0,92). Algoritma SVM mempunyai performa terbaik kedua (*overall accuracy* =

0,88, koefisien Kappa = 0,85, rata-rata *precision* = 0,89, dan rata-rata *recall* = 0,88).

**Tabel 6.** Performa lima algoritma pembelajaran mesin dalam identifikasi lima jenis resak koleksi Kebun Raya Bogor

Algoritma	Overall Accuracy	Koefisien Kappa	Rata-rata Precision	Rata-rata Recall
BayesNet	0,80	0,75	0,81	0,80
KNN	<b>0,92</b>	<b>0,90</b>	<b>0,93</b>	<b>0,92</b>
ANN	0,86	0,82	0,87	0,86
RF	0,86	0,82	0,87	0,86
SVM	0,88	0,85	0,89	0,88

*Confusion matrix* hasil identifikasi lima jenis resak menggunakan algoritma KNN disajikan pada Tabel 7. Algoritma KNN mampu memprediksi dengan benar semua data uji ketiga jenis resak (*V. bantamensis*, *V. granulata*, dan *V. rassak*) dengan nilai *recall* sebesar 1, namun hanya *V. rassak* yang mempunyai nilai *precision* sempurna, yaitu 1. Pada jenis *V. venulosa*, terdapat satu data uji yang



diprediksi sebagai *V. bantamensis*, sehingga menghasilkan nilai *recall* sebesar 0,92. Jenis *V. pauciflora* merupakan jenis yang kurang dapat diprediksi dengan baik oleh algoritma KNN dengan nilai *recall* terendah sebesar 0,63.

Rendahnya nilai *recall* tersebut dikarenakan *V. pauciflora* mempunyai kemiripan karakter morfologi daun yang tinggi dengan jenis *V. granulata* dan *V. venulosa*.

**Tabel 7.** *Confussion matrix, precision, dan recall* hasil identifikasi lima jenis resak koleksi Kebun Raya Bogor menggunakan algoritma KNN

	Prediksi					Precision	Recall
	<i>V. bantamensis</i>	<i>V. granulata</i>	<i>V. pauciflora</i>	<i>V. rassak</i>	<i>V. venulosa</i>		
Aktual							
<i>V. bantamensis</i>	11	0	0	0	0	0,92	1,00
<i>V. granulata</i>	0	10	0	0	0	0,91	1,00
<i>V. pauciflora</i>	0	1	5	0	2	1,00	0,63
<i>V. rassak</i>	0	0	0	9	0	1,00	1,00
<i>V. venulosa</i>	1	0	0	0	11	0,85	0,92

KNN banyak digunakan dalam menyelesaikan masalah klasifikasi karena relatif sederhana, dapat mengolah data yang besar, dapat dimodifikasi, dan menghasilkan performa klasifikasi yang tinggi (Imandoust & Bolandraftar 2013; Uddin et al. 2022). Performa suatu model tergantung karakter (fitur) dan algoritma yang digunakan dalam membangun model. Performa algoritma KNN dalam identifikasi tumbuhan pada penelitian ini dengan parameter model berupa morfologi daun (warna, ukuran, bentuk, dan tekstur) menghasilkan performa yang lebih baik dibandingkan dengan penelitian serupa. Yusniar & Kustiyo (2014) menggunakan 2 *Dimensional Principal Component Analysis* (2D-PCA) untuk mengidentifikasi *Shorea* spp. dan menghasilkan *overall accuracy* sebesar 0,75. Ni'mah et al. (2018) melakukan identifikasi tumbuhan obat herbal berdasarkan tekstur GLCM menghasilkan akurasi sebesar 0,83, sedangkan Furqan et al. (2020) melakukan identifikasi varietas bugenvil menggunakan tekstur GLCM dan menghasilkan akurasi sebesar 0,87.

## KESIMPULAN

Lima jenis resak (*Vatica* spp.) yang diamati mempunyai karakter morfologi berupa warna, ukuran, bentuk, dan tekstur yang relatif berbeda. *Vatica granulata*, *V. pauciflora*, dan *V. venulosa* mempunyai kemiripan yang lebih tinggi dibandingkan dengan jenis *V. bantamensis* dan *V. rassak*. Algoritma pembelajaran mesin *K-Nearest Neighbor* mampu melakukan identifikasi jenis resak secara otomatis dengan performa yang sangat baik. Penggunaan karakter morfologi dan algoritma pembelajaran mesin dapat menjadi alternatif dalam pengidentifikasian jenis-jenis tumbuhan.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Bapak Agus Suhatman dan Bapak Harto dari Direktorat Pengelolaan Koleksi Ilmiah, Badan Riset dan Inovasi Nasional (BRIN) atas dukungannya dalam menyediakan data tanaman koleksi dan membantu pengambilan sampel daun.

## DAFTAR PUSTAKA

- Achmalidi R. 2001. Potret Keadaan Hutan Indonesia. Forest Watch Indonesia, Bogor.
- Agmalaro MA, Kustiyo A, Akbar AR. 2013. Identifikasi tanaman buah tropika berdasarkan tekstur permukaan daun menggunakan jaringan syaraf tiruan. *Jurnal Ilmu Komputer dan Agri-Informatika* 2(2): 73–82.
- Al Janabi, Kadhim R. 2018. Data reduction techniques: A comparative study for attribute selection methods. *International Journal of Advanced Computer Science and Technology* 8(1): 1-13.
- Ashton P. 1998. *Vatica venulosa*. *The IUCN Red List of Threatened Species* 1998: e.T33458A9785745. <https://dx.doi.org/10.2305/IUCN.UK.1998.RLTS.T33458A9785745.en> (diakses 1 Desember 2022).
- Azadnia R, Kheiralipour K. 2021. Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier. *Journal of Applied Research on Medicinal and Aromatic Plants* 25: 100327. doi:10.1016/j.jarmap.2021.100327.
- Begue A, Kowlessur V, Singh U, Mahomoodally F, Pudaruth S. 2017. Automatic recognition of medicinal plants using machine learning techniques. *International Journal of Advanced Computer Science and Applications* 8(4): 166–175.

- Bouckaert RR, Frank E, Hall M, Kirkby R, Reutemann PP, Seewald A, Scuse D. 2016. WEKA Manual for Version 3-9-1. University of Waikato, Hamilton.
- BPS - Badan Pusat Statistik. 2022. Kota Bogor dalam Angka. Badan Pusat Statistik Kota Bogor, Bogor.
- Delgado R, Tibau X-A. 2019. Why Cohen's Kappa should be avoided as performance measure in classification. *PLoS ONE* 14(9): e0222916. doi:10.1371/journal.pone.0222916.
- Esa NM, Jumari, Murningsi, Arifiani D. 2016. Sebaran dan karakter morfologi *Endiandra* (Lauraceae) dari Sumatra, koleksi Herbarium Bogoriense, Pusat Penelitian Biologi-LIPI. *Jurnal Akademika Biologi* 5(4): 32–38.
- Fajri M. 2008. Pengenalan umum Dipterocarpaceae, kelompok jenis bernilai ekonomi tinggi. *Info Teknis Dipterokarpa* 2(1): 9–21.
- Ferreira T, Rasband W. 2012. ImageJ User Guide: IJ1.46r. <https://imagej.nih.gov/ij/docs/guide/146.html> (diakses 10 Januari 2023).
- Frank E, Hall MA, Witten IH. 2016. The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques. University of Waikato, Hamilton.
- Furqan M, Sriani, Harahap LS. 2020. Klasifikasi daun bugenvil menggunakan gray level co-occurrence matrix dan K-Nearest Neighbor. *Jurnal CoreIT* 6(1): 22–29.
- Govaerts R, Nic Lughadha E, Black N, Turner R, Pato A. 2021. The world checklist of vascular plants, a continuously updated resource for exploring global plant diversity. *Scientific Data* 8: 215. doi:10.1038/s41597-021-00997-6.
- Hair JF, Black WC, Babin BJ, Anderson RE. 2010. *Multivariate Data Analysis* (7 th. Edition). Prentice Hall International, New York.
- Hall MA. 1998. Correlation-based feature subset selection for machine learning. PhD Dissertation, Department of Computer Science, The University of Waikato, Hamilton.
- Hall MA. 2000. Correlation-based feature selection of discrete and numeric class machine learning. University of Waikato, Hamilton.
- Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: PAleontological STatistics software package for education and data analysis. *Palaeontologia Electronica* 4(1): 1–9.
- Hammer Ø. 2022. PAST - PAleontological STatistics Version 4.11. Reference manual. University of Oslo, Oslo.
- Haralick RM, Shanmugam K, Dinstein I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 3(6): 610–621.
- Hussein BR, Malik OA, Ong WH, Slik JWF. 2020. Automated classification of tropical plant species data based on machine learning techniques and leaf trait measurements. In: Alfred R, Lim Y, Haviluddin H, On CK (eds.). *Computational Science and Technology*. Springer, Singapore.
- Ibrizzah AA, Setyaningsih A, Prasetyo B, Amalia R. 2018. Mengenal Keunikan Dipterocarpaceae. Universitas Gadjah Mada, Yogyakarta.
- Imandoust SB, Bolandraftar M. 2013. Application of K-Nearest Neighbor (KNN) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications* 3(5): 605–610.
- Jalonen R, Hong LT, Lee SL, Loo J, Snook L. 2014. Integrating genetic factors into management of tropical Asian production forests: A review of current knowledge. *Forest Ecology and Management* 315: 191–201. doi:10.1016/j.foreco.2013.12.011.
- Jolliffe I. 2011. Principal Component Analysis. *International Encyclopedia of Statistical Science*: 1094–1096. doi:10.1007/978-3-642-04898-2\_455.
- Kalima T, Wardani W. 2017. *Vatica bantamensis* & *Vatica rassak*. In: Yulita KS, Partomihardjo T, Wardani W (eds.). *Prekursor Buku Daftar Merah Indonesia 1:50 Jenis Pohon Kayu Komersial*. Indonesian Plant Red List Authority, Bidang Botani, Pusat Penelitian Biologi-LIPI, Bogor.
- Khalid M, Lee ELY, Yusof R, Nadaraj M. 2008. Design of an intelligent wood species recognition system. *International Journal of Simulation: Systems, Science & Technology* 9(3): 9–18.
- Kho SJ, Manickam S, Malek S, Mosleh M, Dhillon SK. 2017. Automated plant identification using artificial neural network and support vector machine. *Frontiers in Life Science* 10(1): 98–107. doi:10.1080/21553769.2017.1412361.
- Kumar CA, Sooraj MP, Ramakrishnan S. 2017. A comparative performance evaluation of supervised feature selection algorithms on microarray datasets. *Procedia Computer Science* 115: 209–217.
- Lemmens LHMJ, Soerianegara I, Keating WG, Wong WC, Ilic J. 1994. *Vatica*. In: Soerianegara I, Lemmens RHMJ (eds.). *Plant Resources of South-East Asia No 5(1): Timber Trees; Major Commercial Timbers*. PROSEA Foundation, Bogor.
- Mattjik AA, Sumertajaya IM. 2011. *Sidik Peubah Ganda*. IPB Press, Bogor.
- Mclellan T. 2000. Geographic variation and plasticity of leaf shape and size in *Begonia dregei* and *B. homonyma* (*Begoniaceae*). *Botanical Journal Of the Linnean Society* 132(1): 79–95.
- McHugh ML. 2011. Multiple comparison analysis testing in ANOVA. *Biochemia Medica* 21(3): 203–209.

- Meekiong K, Latiff A, Tawan CS, Miraadila MI. 2012. Leaf micro-morphological characteristics of selected *Vatica* species (Dipterocarpaceae) from Kubah National Park, Sarawak. *Malaysian Applied Biology* 41(2): 41–44.
- Musfiroh D, Khaira U, Utomo PEP, Suratno T. 2021. Analisis sentimen terhadap perkuliahan daring di Indonesia dari twitter dataset menggunakan InSet Lexicon. *MALCOM* 1(1) : 24–33.
- Nasution HN. 2016. Identifikasi jenis tanaman anggrek menggunakan jaringan saraf tiruan dengan metode backpropagation. *Jurnal Education and Development* 1(1): 12–19.
- Ngatiman, Saridan A. 2012. Eksplorasi jenis-jenis Dipterokarpa. *Jurnal Penelitian Ekosistem Dipterokarpa* 6(1): 1–10.
- Ni'mah FS, Sutojo T, Setiadi DRIM. 2018. Identifikasi tumbuhan obat herbal berdasarkan citra daun menggunakan algoritma gray level co-occurrence matrix dan K-Nearest Neighbor. *Jurnal Teknologi dan Sistem Komputer* 6(2): 51–56.
- Petrus S, Togar FM, Siti MK. 2021. Identifikasi jenis pohon family Dipterocarpaceae pada hutan rawa gambut di Kawasan Hutan Dengan Tujuan Khusus (KHDTK) Universitas Tanjungpura Kecamatan Mandor Kabupaten Landak Kalimantan Barat. *Jurnal Hutan Lestari* 9(4): 584–598.
- Pooma R, Newman MF. 2017. *Vatica pauciflora*. The IUCN Red List of Threatened Species 2017:e.T33151A2833358. <https://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T33151A2833358.en> (diakses 1 Desember 2022).
- Pratiwi SND, Ulama BSS. 2016. Klasifikasi email spam dengan menggunakan metode Support Vector Machine dan K-Nearest Neighbor. *Jurnal Sains dan Seni ITS* 5(2): 2337–3520.
- Pujiati R, Rochmawati N. 2022. Identifikasi citra daun tanaman herbal menggunakan metode Convolutional Neural Network (CNN). *Journal of Informatics and Computer Science* 3(3): 351–357.
- Robiansyah I. 2018. *Vatica bantamensis*. The IUCN Red List of Threatened Species 2018:e.T31319A135558901. <https://dx.doi.org/10.2305/IUCN.UK.2018-1.RLTS.T31319A125626167.en> (diakses 1 Desember 2022).
- Rosdayanti H, Siregar UJ, Siregar IZ. 2019. Karakter penciri morfologi daun meranti (*Shorea* spp.) pada area budidaya ex-situ KHDTK Haurbentes. *Media Konservasi* 24(2): 207–215.
- Rosinta DR, Yoza D, Sribudiani E. 2019. Karakteristik habitat dan pola distribusi kulim (*Scorodocarpus Borneensis* Becc.) di sepanjang jalur utama patroli Taman Hutan Raya Sultan Syarif Hasyim Provinsi Riau. *Jurnal Ilmu-Ilmu Kehutanan* 3(2): 32–44.
- Saridan A. 2012. Keragaman jenis dipterokarpa dan potensi pohon penghasil minyak keruing di hutan dataran rendah Kabupaten Berau, Kalimantan Timur. *Jurnal penelitian Dipterokarpa* 6(2): 75–83.
- Schmeller DS, Bauch B, Gruber B, Juskaitis R, Budrys E, Babij V, Lanno K, Sammul M, Varga Z, Henle K. 2008. Determination of conservation priorities in region with multiple political jurisdictions. *Biodiversity and Conservation* 17: 3623–3630.
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9: 671–675.
- Setiawati T. 2013. Diversitas genetik kerabat liar ubi jalar *Ipomoea trifida* berumbi asal Citatah Jawa Barat berdasarkan karakter kromosom. *Indonesian Journal of Applied Sciences* 3(3): 84–87
- Singh A, Ganapathysubramanian B, Singh AK, Sarkar S. 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science* 21(2): 110–124. doi:10.1016/j.tplants.2015.10.015.
- Srinual A, Thammathaworn A. 2008. Leaf anatomy of *Vatica* L. (Dipterocarpaceae) in Thailand. *Tropical Natural History* 8(2): 121–134.
- Theera S, Sawitree T, Charun M. 2012. A preliminary study of leaf morphology and flow cytometry in the *Vatica diospyroides* Symington, endangered medicinal plant of Peninsular Thailand. *Journal of Medicinal Plants Research* 6(20): 3681–3688.
- Uddin S, Haque I, Lu H, Moni MA, Gide E. 2022. Comparative performance analysis of K-Nearest Neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports* 12: 3623–3630. doi:10.1038/s41598-022-10358-x.
- Wardani M, Astuti IP, Heriyanto N. 2017. Analisis vegetasi jenis-jenis Dipterocarpaceae di Kawasan Hutan Seksi 1 Way Kanan, Taman Nasional Way Kambas, Lampung. *Buletin Kebun Raya* 20(1): 51–64.
- Widiati KY, Sri AH, Karmini. 2023. Pemanfaatan kayu resak (*Vatica* sp.) dari lahan terbiarkan sebagai bahan konstruksi bangunan. *Ulin – Jurnal Hutan Tropis* 7(1): 38–44.
- Widiyono W. 2021. Biological and economic value of Dipterocarpaceae, the main timber forest product of Indonesia. *Indonesian Journal of Applied Environmental Studies* 2(2): 104–112. doi:10.33751/injast.v2i2.4016.
- Yusniar E, Kustiyo A. 2014. Identifikasi daun *Shorea* menggunakan KNN dengan ekstraksi fitur 2DPCA. *Jurnal Ilmu Komputer dan Agri-Informatika* 3(1): 19–27.

Zdilla MJ, Hatfield SA, McLean KA, Cyrus LM, Laslo JM, Lambert HW. 2016. Circularity, solidity, axes of a best fit ellipse, aspect ratio, and roundness of the foramen ovale. *Journal of Craniofacial Surgery*

27(1):

228. doi:10.1097/scs.0000000000002285.

222–