Identifikasi Penyakit Tanaman Kopi Arabika dengan Metode K-Nearest Neighbor (K-NN)

Anggi Yhurinda Perdana Putri[1], Anwar Sodik[2]
Institut Teknologi Adhi Tama Suarabaya[1,2]
e-mail: anggi@itats.ac.id

ABSTRACT

offee is a mainstay of plantation commodities that has a significant contribution to the Indonesian economy. The coffee plant in particular is Arabica coffee (coffea arabica) which has a higher selling price than Robusta coffee with a market share of 72%. However, Arabica coffee is more susceptible to certain diseases, which causes farmers to experience difficulties in maintaining Arabica coffee and lack of knowledge about how to deal with the disease. One solution to this problem is to identify early in detecting diseases in Arabica coffee plants and how to overcome them. Therefore, the researchers aimed to identify Arabica coffee plant diseases using the K-Nearest Neighbor (K-NN) method. KNN method is a method applied to overcome the problem of identifying a disease that is measured qualitatively and quantitatively. Based on the test data used in this study, the method can diagnose Arabica coffee plant diseases with an accuracy level of 80%.

Keywords: *K-Nearest Neighbor, K-NN, Plant disease, Arabica Coffee Plant*

ABSTRAK

Kopi merupakan komoditas andalan perkebunan yang mempunyai kontribusi cukup besar dalam perekonomian Indonesia. Tanaman kopi khususnya adalah kopi arabika (coffea arabica) memiliki harga jual yang lebih tinggi dari kopi robusta dengan pangsa pasar 72%. Namun kopi arabika lebih peka terinfeksi penyakit tertentu, hal itu menyebabkan petani mengalami kesulitan dalam memelihara kopi jenis arabika serta kurangnya pengetahuan tentang cara penanganan penyakit. Salah satu solusi permasalahan ini adalah dengan melakukan identifikasi awal dalam mendeteksi penyakit pada tanaman kopi Arabika beserta cara penanggulangannya. Oleh karena itu, peneliti bertujuan untuk melakukan identifikasi penyakit tanaman kopi Arabika dengan metode K-Nearest Neighbor (K-NN). Metode KNN merupakan metode yang diterapkan untuk mengatasi permasalahan identifikasi suatu penyakit yang diukur secara kualitatif dan kuantitatif. Berdasarkan data uji yang digunakan pada penelitian ini, metode KNN dapat melakukan diagnosa penyakit tanaman kopi Arabika dengan tingkat akurasi sebesar 80%.

Kata Kunci: K-Nearest Neighbor, KNN, Penyakit Tanaman, Tanaman Kopi Arabika

PENDAHULUAN

Kopi sebagai komoditas unggulan dalam subsektor perkebunan, kopi memiliki kontribusi cukup besar dalam perekonomian Indonesia karena memiliki peluang pasar yang baik didalam maupun diluar negeri. Pengaruh tingkat konsumsi kopi yang cukup tinggi menyebabkan komoditi tersebut menjadi prioritas di masyarakat Indonesia. Tanaman kopi termasuk dalam Genus Coffea terdiri dari beberapa jenis yaitu Coffea Arabica (Kopi Arabika), Coffea robusta (Kopi Robusta) dan Coffea liberica (Kopi Liberika) [1]. Dari tiga jenis tanaman kopi tersebut, kopi arabika (coffea arabica) yang diprioritaskan di Jawa Timur, karena memiliki nilai jual yang paling tinggi dengan pangsa pasar 72% [2]. Selain itu, kopi jenis Arabika mempunyai cita rasa terbaik dari 2 jenis kopi lainnya. Namun, tanaman kopi jenis arabika lebih rentan terserang penyakit. Akibatnya, petani kopi mengalami kesulitan dalam pemeliharaannya. Oleh karena itu, petani kopi perlu untuk melakukan diagnosa awal penyakit yang menyerang tanaman kopi arabika, supaya dapat dilakukan pengendalian penyakit yang berpotensi meningkatkan produktifitas kopi [3].

Identifikasi penyakit pada tanaman kopi arabika dengan menggunakan metode K-Nearest Neighbor, dimanfaatkan untuk memberikan solusi dalam melakukan diagnosa awal berdasarkan jenis penyakit yang menyerang tanaman kopi arabika. Metode KNN dipilih karena memberikn hasil identifikasi penyakit yang lebih baik [4] serta nilai akurasi yang tinggi [5].

Penelitian dengan metode K-Nearest Neighbor (K-NN) telah dilakukan sebelumnya oleh Abdul Rohman dengan judul penelitian "Model Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Mahasiswa" dengan hasil uji akurasi sebesar 85,15%. Sehingga dapat disimpulkan bahwa metode K-NN mampu memberikan hasil Identifikasi yang cukup akurat [5]. Berdasarkan uraian diatas, peneliti membuat suatu penelitian dengan judul "Identifikasi Penyakit Tanaman Kopi Arabika dengan Metode K-Nearest Neighbor (K-NN)". Penelitian ini diharapkan dapat memberikan informasi diagnosa yang akurat mengenai penyakit tanaman Kopi Arabika berdasarkan gejala-gejala yang ada dan membantu keterbatasan pengetahuan petani atau masyarakat mengenai penyakit tanaman Kopi Arabika.

TINJAUAN PUSTAKA

Tanaman Kopi Arabika

Kopi (*Coffea sp*.) merupakan spesies tanaman pohon yang tergolong dalam famili *Rublaceae* dan *genus coffea*, tumbuh tegak, bercabang, dan dapat tumbuh mencapai tinggi 12 meter. Jenis kopi terdiri dari kopi Arabika, Robusta dan Liberika [1]. Kopi jenis Arabika memiliki ukuran yang lebih pendek dibanding kopi jenis lainnya. Dengan cita rasa yang paling baik diantara Robusta dan Liberika, kelemahan dari tanaman kopi jenis ini yaitu sangat rentan terserang penyakit.

Penyakit Tanaman Kopi

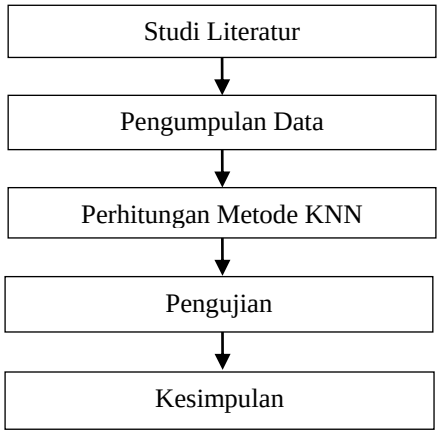
Penyakit yang menyerang tanaman kopi dapat disebabkan oleh jamur maupun virus. Jenis-jenis penyakit tanaman kopi Arabika digolongkan dalam 5 kelas penyakit, yaitu: 1). Nematoda, 2). Karat Daun, 3). Bercak Daun, 4). Antraknos, dan 5) Jamur Upas.

K-nearest Neighbor

Metode K-*Nearest Neighbor* (KNN) merupakan metode klasifikasi terhadap obyek berdasarkan data pembelajaran dengan jarak yang terdekat. Metode KNN diterapkan untuk mengatasi permasalahan identifikasi yang diukur secara kualitatif maupun kuantitatif. Sebelum mencari jarak data ke tetangga adalah menentukan nilai K tetangga (neighbor). Lalu, untuk mendefinisikan jarak antara dua titik pada data training dan titik pada data testing, yaitu menggunakan rumus *Euclidean Distance.* Hasil dari perhitungan jarak, kemudian diurutkan dari data dengan nilai terkecil hingga terbesar sesuai dengan nilai k yang telah ditentukan sebelumnya.

METODE

Metode yang dilakukan dalam penelitian ini terdiri dari studi literatur, pengumpulan data, pengujian, dan kesimpulan. Tahapan-tahapan dalam penelitian ini, dapat diilustrasikan dalam bentuk bagan metode penelitian. Pada gambar 1 ditunjukkan alur metode penelitian.



Gambar 1. Metode Penelitian

Studi literatur yang dibutuhkan untuk merealisasikan penelitian ini yaitu 1). Teori mengenai KNN dan 2). Teori mengenai penyakit tanaman kopi Arabika. Kemudian dilakukan proses pengumpulan data yang menghasilkan 16 data gejala penyakit dan bobot masing-masing gejala yang diperoleh dari wawancara pakar. Gejala dan bobot ditunjukkan pada Tabel 1

Tabel 1. Gejala dan Bobot Penyakit

No.	Gejala	Nama Gejala	Opsi	Nilai
1	G1	Daun menguning	Tidak	0
			sedikit	50
			banyak	90
2	G2	Terdapat tepung berwarna jingga pada sisi bawah daun	Tidak	0
			sedikit	50
			banyak	90
3	G3	Terdapat bercak kuning pada sisi atas daun	Tidak	0
			sedikit	50
			banyak	90
4	G4	Bercak coklat pada sisi atas daun	Tidak	0
			sedikit	30
			banyak	50
5	G5	Daun mengering	Tidak	0
			sedikit	30
			banyak	50
6	G6	Daun gugur	Tidak	0
			sedikit	40
			banyak	60
7	G7	Tanaman tampak gundul	Tidak	0
			sedikit	30
			banyak	50
8	G8	Bercak dengan inti putih kelabu dan lingkar kuning	Tidak	0
			sedikit	50
			banyak	90

9	G9	Bawah daun terdapat jamur putih	Tidak	0
			sedikit	50
			banyak	90
10	G10	Jala putih atau perak pada batang	Tidak	0
			sedikit	50
			banyak	90
11	G11	Gumpalan putih pada celah kulit pohon	Tidak	0
			sedikit	50
			banyak	90
12	G12	Kerak warna merah jambu pada cabang	Tidak	0
			sedikit	50
			banyak	90
13	G13	Terdapat bintil kecil kemerahan pada cabang	Tidak	0
			banyak	50
			sedikit	90
14	G14	Tanaman terlihat kerdil	Tidak	0
			sedikit	50
			banyak	90
15	G15	Bunga sedikit	Tidak	0
			sedikit	50
			banyak	90
16	G16	Buah banyak yang kosong	Tidak	0
			sedikit	50
			banyak	90

Perhitungan Metode KNN

Menentukan data input

Data input diperoleh dari inputan gejala yang ditunjukkan pada Tabel 2

Tabel 2. Data Input Gejala Penyakit Kopi Arabika

Gejala	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16
bobot	0	0	0	30	0	0	0	50	0	0	0	90	0	90	0	0

Menentukan *Range* dan Normalisasi data latih

Data latih ditunjukkan pada Tabel 3

Tabel 3. Range Data Latih pada Gejala Penyakit Kopi Arabika

Gejala	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16
Min	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Max	90	90	90	50	50	60	50	90	90	90	90	90	90	90	90	90
Range	90	90	90	50	50	60	50	90	90	90	90	90	90	90	90	90

Kemudian data input dinormalisasi, sehingga sebaran data pada *range* yang sama yaitu [0,1]. Cara menghitung normalisasi yaitu dengan menghitung data input dikurangi range data minimal dan dibagi dengan range data latih, dihitung dengan persamaan 1

$$\frac{Data\ Input - Min}{Range} \quad \text{.....} \quad (1)$$

Pada tabel 4 ditunjukkan hasil normalisasi data input

Tabel 4. Hasil Normalisasi Data Input Gejala Penyakit Kopi Arabika

Gejala	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16
bobot	0	0	0	0.6	0	0	0	0.56	0	0	0	1	0	0.56	0	0

Menghitung jarak Euclidean

Tahap selanjutnya yaitu menghitung jdata latih dengan menggunakan Jarak Euclidean dengan menggunakan persamaan 2

$$d(a,b) = \sum_{i=0}^{n} (Xi - Yi)^2 \quad \ldots\ldots\ldots\ldots\ldots (2)$$

Dimana:

d(a,b) : Jarak Euclidean

x : data 1

y : data 2

i,j : fitur i,j

n : jumlah fitur

Hasil perhitungan jarak Euclidean ditunjukkan pada Tabel 5

Tabel 5. Hasil perhitungan Jarak Euclidean

NO	Jarak	Kelas	NO	Jarak	Kelas
10	1.74271	5	5	2.148787	1
9	1.930905	5	4	2.224998	3
7	1.981395	1	2	2.328461	2
8	1.981395	2	3	2.328461	4
6	2.076203	3	1	2.462058	4

Hasil perhitungan jarak *Euclidean* kemudian diurutkan dari nilai jarak terkecil ke terbesar. Nilai K yang digunakan pada penelitian ini yaitu K=5, sehingga diambil 5 jarak terkecil dari hasil perhitungan jarak pada Tabel 5 dengan *highlight* kuning. Hasil jarak terkecil yang diperoleh yaitu pada data ke-10 dengan nilai 1.74271 dan termasuk dalam kelas 5 (penyakit Upas).

HASIL DAN PEMBAHASAN

Pengujian dilakukan dengan menguji pengaruh nilai k terhadap akurasi dengan variasi yaitu 1-5, dengan sebaran data latih 10, 15, dan 20, serta 10 data uji. Hasil pengujian k menghasilkan akurasi dengan rata-rata 65.99%. Akurasi tertinggi yaitu pada k=1 dengan jumlah data latih 15, K=4 dan K=5 pada data latih 10 dengan akurasi sebesar 80%. Hal tersebut dikarenakan sebaran data latih yang tidak seimbang, sehingga pengaruh k berbeda-beda pada setiap jumlah data latih. Hasil pengujian ditunjukkan pada Tabel 6

Tabel 6. Hasil Pengujian

K	Akurasi(%)			Rata-rata Akurasi
	10 data uji	15 data uji	20 data uji	
1	70%	80%	70%	73.33%
2	60%	60%	70%	63.33%
3	60%	60%	60%	60%
4	80%	60%	60%	66.66%
5	80%	60%	60%	66.66%
TOTAL RATA-RATA AKURASI				**65.99%**

KESIMPULAN

Hasil dari penelitian ini menghasilkan kesimpulan bahwa hasil pengujian K 1-5 menghasilkan akurasi dengan rata-rata nilai 65.99%. Sedangkan akurasi tertinggi yaitu pada k=1 dengan jumlah data latih 15, K=4 dan K=5 pada data latih 10 dengan akurasi sebesar 80%.

DAFTAR PUSTAKA

[1] Danarti dan Najiyanti. 1999. Palawija Budidaya dan Analisis Usaha Tani. Jakarta: Penebar Swadaya.
[2] Moh. Cholil Mahfud, Luki Rosmahani, dan Diding Rachmawati. 1998. Pengendalian Hama dan Penyakit Tanaman Kopi Arabika. Malang: Balai Pengkajian Teknologi Pertanian Karangploso.
[3] Mahfud, Moh. Cholil. 2012. Teknologi dan Strategi Pengendalian Penyakit Karat Daun untuk Meningkatkan Produksi Kopi Nasional. Malang: Balai Pengkajian Teknologi Pertanian Jawa Timur.
[4] Redjeki, Sri. 2013. Perbandingan Algoritma Backpropagation dan K-Nearest Neighbor (K-NN) untuk Identifikasi Penyakit. Program Studi Teknik Informatika. STMIK AMIKOM. Yogyakarta.
[5] Rohman, Abdul. 2015. Model Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Mahasiswa. Jurusan Elektronika. Fakultas Teknik Universitas Pandanaran. Semarang
[6] Amiratus, Rahmi. 2013. Penerapan Metode Fuzzy K-NN untuk Menentukan Kualitas Hasil Rendeman Tebu. Program Teknologi Informasi dan Ilmu Komputer. Universitas Brawijaya. Malang.
[7] Anggraeni, Daria. 2015. Sistem Pakar untuk Identifikasi Hama dan Penyakit Tanaman Tebu dengan Metode Fuzzy AHP. Malang : Universitas Brawijaya.