PENERAPAN FUZZY K-NEAREST NEIGHBOR (FKNN) UNTUK DIAGNOSA PENYAKIT LAYU PADA HUTAN BERDASARKAN WILT DATASET

Lintang Anginseto[1], Edy Santoso[2], Sutrisno[3]

[1]Mahasiswa Program Studi Teknik Informatika Universitas Brawijaya

[2,3]Staff Pengajar Program Studi Teknik Informatika Universitas Brawijaya

Program Studi Teknik Informatika

Fakultas Ilmu Komputer

Universitas Brawijaya

Jalan Veteran Malang 65145, Indonesia

Email : anginseto@gmail.com, .edy144@ub.ac.id, trisno@ub.ac.id

ABSTRAK

Penelitian ini membahas tentang penerapan metode algoritma FKNN (Fuzzy K-Nearest Neighbour) untuk klasifikasi penyakit layu (wilt disease). Sistem yang di rancang dan dibangun berdasarkan wilt dataset yang diambil dari website UCI Machine Learning Repository yang diambil pada tahun 2014. Metode algoritma FKNN adalah algoritma yang memberikan nilai pada keanggotaan kelas pada vector sampel dan bukan menempatkan vector pada kelas tertenmtu. Pengujian ini dilakukan terhadapt 6 jumlah data latih yang berbeda yaitu 60, 90, 150, 220, 300 dan 350 data dengan data uji sebanyak 30 dan 60 data. Hasil dari pengujian ini untuk mengetahui pengaruh data latih, nilai k, terhadap tingkat akurasi dari sistem ini dalam bentuk persentase.

Kata kunci : wilt disease, FKNN, klasifikasi

ABSTRACT

This study discusses the application of the algorithm method FKNN (Fuzzy K-Nearest Neighbour) for the classification of wilt diseases (wilt disease). The system was designed and built by Wilt dataset taken from the website of the UCI Machine Learning Repository taken in 2014. Methods FKNN algorithm is an algorithm that gives value to the class membership in the sample vector and not put vector in a particular class. This testing is done terhadapt 6 different amount of training data that is 60, 90, 150, 220, 300 and 350 of data with test data as much as 30 and 60 data. The results of this test to determine the effect of training data and k values of the level of accuracy of this system in the form of a percentage.

Keywords: wilt disease, FKNN, classification

1. Pendahuluan

1.1 Latar Belakang

Hutan sangat penting bagi kehidupan di muka bumi, terutama bagi kehidupan generasi mendatang. Kesalahan dalam pengelolaan hutan berarti menyiksa kehidupan generasi kita mendatang. Hutan merupakan kumpulan pepohonan yang tumbuh rapat beserta tumbuh-tumbuhan memanjat dengan bunga yang beraneka warna yang berperan sangat penting bagi kehidupan di bumi ini. Dari sudut pandang ekonomi, hutan merupakan tempat menanam modal jangka panjang yang sangat menguntungkan dalam bentuk Hak Pengusahaan Hutan (HPH)[IAA-01].

Salah satu jenis hutan dan merupakan yang terluas di Indonesia adalah hutan hujan tropis. Hutan hujan tropis merujuk pada tipe hutan di kawasan tropis yang selalu diguyur hujan sepanjang tahun. Hutan hujan tropis ditumbuhi beragam jenis pohon yang membentuk lapisan tajuk. Secara umum terdapat pohon bertajuk tinggi yang membentuk kanopi menaungi tanaman lainnya, kemudian pohon

menengah seperti tanaman merambat dan perdu, dan terakhir tanaman permukaan tanah seperti rumput dan lumut. Di hutan hujan tropis tidak ada jenis pohon tertentu yang mendominasi kawasan. Semua berbagi tempat dalam ekosistem dengan jumlah yang sedikit-sedikit tapi keragamannya tinggi [CRR-15].

Hutan hujan tropis juga dikenal sebagai paru-paru dunia. Diperkirakan sekitar 40% produksi oksigen dunia dihasilkan dari tempat ini. Hutan ini juga merupakan penyimpan cadangan karbon dunia. Setiap kerusakan yang terjadi di hutan ini menyebabkan berdampak serius terhadap perubahan iklim global[CRR-15].

Dengan mengetahui pentingnya hutan berdasarkan penjelasan di atas, maka sangat penting untuk menjaga kelestarian hutan. Beberapa hal yang dapat merusak kelestarian hutan, seperti perkembangan populasi penduduk atau perluasan area usaha atau perusahaan di sekitar hutan yang mengakibatkan penebangan hutan yang kadang tidak terkontrol, kebakaran hutan dan penyebab alami seperti penyakit pada tanaman hutan. Penyebab alami ini dapat dikaitkan oleh perubahan iklim yang terjadi saat ini.

Salah satu hal yang merusak hutan adalah penyakit layu (Wilt Disease) pada pepohonan. Wilt Disease atau penyakit layu adalah gejala alami dari penyakit tanaman yang diakibatkan kekurangan air pada daun dan batang. Bagian tanaman yang terkena akan kehilangan ketebalannya dan menjadi layu, jika hal ini terjadi terus menerus akan menyebabkan tanaman menjadi mati. Gejala ini dapat disebabkan oleh kekurangan air, bakteri atau virus[IAS-12].

Berdasarkan uraian di atas maka penelitian ini diberi judul "Penerapan Fuzzy K-Nearest Neighbour (FKNN) Untuk Diagnosa Penyakit Pada Hutan Berdasarkan Wilt Dataset" dengan klasifikasi berdasarkan parameter GLCM mean texture, Mean green value, Mean red value , Mean NIR value , dan Standard deviation.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dirumuskan permasalahan sebagai berikut :

1. Bagaimana penerapan metode Fk-NN pada klasifikasi penyakit berdasarkan Wilt Data Set dengan parameter GLCM mean texture, Mean green value, Mean red value , Mean NIR value , dan Standard deviation.
2. Berapa besar tingkat akurasi dari hasil klasifikasi Wilt Data Set menggunakan metode Fk-NN yang dipengaruhi sejumlah nilai k (tetangga).

1.3 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah:

1. Batasan subjek penelitian berupa dataset dari UCI Machine Learning Repository, yaitu Wilt Data Set yang memiliki dua klasifikasi (sakit dan tutupan lahan lain nya).
2. Parameter yang digunakan untuk menentukan klasifikasi, yaitu GLCM mean texture, Mean green value, Mean red value , Mean NIR value , dan Standard deviation dengan jumlah total data sebanyak 320 data.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam pembuatan skripsi ini adalah:

1. Menerapkan metode Fk-NN pada Wilt Dataset.
2. Memperoleh hasil tingkat akurasi klasifikasi yang dipengaruhi sejumlah nilai k pada Wilt Dataset menggunakan metode Fk-NN.

1.5 Manfaat Penelitian

Manfaat yang dapat diambil dari penelitian ini yaitu dapat membantu dalam proses deteksi penyakit wilt pada hutan dengan mengguanakan Fk-NN. Sehingga nantinya dapat digunakan untuk penjaga hutan atau peneliti untuk mendeteksi wilt pada pepohonan hutan lebih dini.

1.6 Sistematika Penulisan

Sistematika penulisan dalam skripsi ini sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, serta sistematika penulisan skripsi.

BAB 2 TINJAUAN PUSTAKA

Bab ini berisi tentang dasar teori dan referensi yang mendasari pembuatan klasifikasi data pada

penyakit wilt yaitu kajian pustaka, hutan, penyakit layu (wilt disease), dan Fk-NN.

BAB 3 METODOLOGI

Bab ini berisi tentang metode–metode yang digunakan untuk memprediksi tingkat resiko menggunakan klasifikasi Fk-NN, pada studi kasus diagnosis penyakit layu.

BAB 4 PERANCANGAN

Bab ini membahas implementasi dari klasifikasi Fk-NN penyakit layu pada tanaman hutan yang digunakan dalam penelitian ini.

BAB 5 HASIL DAN PEMBAHASAN

Bab ini berisi pembahasan dari implementasi klasifikasi Fk-NN pada penyakit Wilt yang sesuai dengan sistem dan hasil pengujian yang dilakukan.

BAB 6 PENUTUP

Bab ini berisi kesimpulan yang diperoleh dari hasil pembuatan dan pengujian yang dikembangkan dalam skripsi ini serta saran untuk pengembangan lebih lanjut.

2. Landasan Kepustakaan

Landasan kepustakaan berisi uraian dan pembahasan tentang teori, konsep, model, metode, atau sistem dari literatur ilmiah, yang berkaitan dengan tema, masalah, atau pertanyaan penelitian. Dalam landasan kepustakaan terdapat landasan teori dari berbagai sumber pustaka yang terkait dengan teori dan metode yang digunakan dalam penelitian. Jika dibutuhkan sesuai dengan karakteristik penelitiannya dan syarat kecukupan khusus keminatan tertentu, bisa juga terdapat kajian pustaka yang menjelaskan secara umum penelitian-penelitian terdahulu yang berhubungan dengan topik skripsi dan menunjukkan persamaan dan perbedaan skripsi tersebut terhadap penelitian terdahulu yang dituliskan.

2.1 Penyakit Layu (Wilt Disease)

Kelayuan pada tanaman terutama pada bagian daun, tunas atau tanaman secara keseluruhan, dapat disebabkan karena hilangnya turgor pada bagian-bagian tersebut. Hilangnya turgor tersebut dapat disebabkan karena adanya gangguan di dalam berkas pembuluh/pengangkutan atau adanya kerusakan pada susunan akar, yang menyebabkan tidak seimbangnya penguapan dengan pengangkutan air. Penyakit layu (wilt disease) pada tanaman dapat disebabkan oleh faktor biotik yaitu bakteri sehingga disebut layu bakteri (Pseudomonas solanacearum) atau oleh jamur/cendawan yang disebut penyakit layu Fusarium (Fusarium oxysporum). Selain karena penyakit biotik, kelayuan pada tanaman juga dapat disebabkan karena faktor abiotik (kekurangan air). Pengenalan gejala kelayuan pada tanaman dan ciri-ciri khususnya harus diketahui para petani, supaya dalam pengendaliannya menjadi lebih efektif dan efisien. Berikut adalah karakteristik/ciri-ciri gejala kelayuan pada tanaman dan penyebabnya.

2.2 Hutan

Menurut ahli silvika, hutan merupakan suatu asosiasi tumbuh-tumbuhan yang sebagian besar terdiri atas pohon-pohon atau vegetasi berkayu yang menempati areal luas. Sedangkan ahli ekologi mengartikan hutan sebagai suatu masyarakat tumbuh-tumbuhan yang dikuasai oleh pohon-pohon dan mempunyai keadaan lingkungan yang berbeda dengan keadaan di luar hutan.

Pada dasarnya, semua variasi tersebut akan mempunyai kesamaan persepsi apabila ditarik suatu kesimpulan, yakni suatu asosiasi kehidupan, baik tumbuh-tumbuhan (flora) maupun binatang (fauna) dari yang sederhana sampai yang bertingkat tinggi dan dengan luas sedemikian rupa serta mempunyai kerapatan tertentu dan menutupi areal,sehinga dapat membentuk iklim mikro tertentu. Assosiasi adalah suatu komunitas tumbuhan yang mempunyai komposisi tumbuhan berbunga di dalam suatu formasi. Kerapatan hutan disebabkan oleh adanya semak belukar, tanaman penutup tanah, dan adanya pohon-pohon pemanjat. Hal ini mendorong tercipatanya habitat berbagai jasad renik dan fauna yang disebabkan oleh adanya ketersediaan pakan. Hutan yang tumbuh dan berkembang tidak lepas dari faktor-faktor yang mempengaruhinya, terutama lingkungan[IAA-01].

2.5 Logika Fuzzy

Logika fuzzy adalah suatu cara yang tepat untuk memetakan ruang input kedalam suatu ruang output [KUS-03]. Konsep ini diperkenalkan dan dipublikasikan pertama kali oleh Lotfi A. Zadeh, seorang profesor dari University of California di Berkeley pada tahun 1965. Logika fuzzy menggunakan ungkapan bahasa untuk menggambarkan nilai variabel. Logika fuzzy bekerja dengan menggunakan

derajat keanggotaan dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang telah ditentukan. Telah disebutkan sebelumnya bahwa logika fuzzy memetakan ruang input ke ruang output. Antara input dan output ada suatu kotak hitam yang harus memetakan input ke output yang sesuai. Alasan mengapa orang menggunakan logika fuzzy, yaitu [KUS-10]:

1. Konsep logika fuzzy mudah dimengerti karena menggunakan konsep matematis yang sederhana dan mudah dimengerti.
2. Logika fuzzy sangat fleksibel.
3. Logika fuzzy memiliki toleransi terhadap data yang tidak tepat.
4. Logika fuzzy mampu memodelkan fungsi-fungsi nonlinier yang sangat kompleks.
5. Logika fuzzy dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan.
6. Logika fuzzy dapat bekerjasama dengan teknik-teknik kendali secara konvensional.
7. Logika fuzzy didasarkan pada bahasa alami.

2.6 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data Mining merupakan suatu proses yang menggunakan teknik statistika, matematika, kecerdasan buatan dan machine learning untuk mengidentifikasikan informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar [TUR-05]. Data Mining merupakan teknologi yang sangat berguna untuk membantu perusahaan-perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Dengan data mining dapat meramalkan trend dan sifat-sifat perilaku bisnis yang sangat berguna untuk mendukung pengambilan keputusan penting. Analisis yang diotomatisasi yang dilakukan oleh data mining melebihi yang dilakukan oleh sistem pendukung keputusan tradisional yang sudah banyak digunakan. Data Mining mengeksplorasi basis data untuk menemukan pola-pola yang tersembunyi, mencari informasi untuk memprediksi yang mungkin saja terlupakan oleh para pelaku bisnis karena terletak di luar ekspektasi mereka [KHU-07].

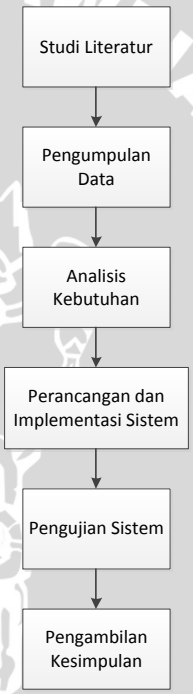
Metode data mining ini dapat dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, diantaranya yaitu metode deskripsi, estimasi, prediksi, klasifikasi, pengklusteran dan asosiasi [LAR-05].

Pada penelitian ini digunakan metode data mining dengan prediksi menggunakan metode gabungan klasifikasi k-Nearest neighbor dengan metode fuzzy.

3. Metodologi

Dalam bab ini akan dibahas metode dan langkah-langkah yang dilakukan dalam penelitian menggunakan metode Fuzzy k-Nearest Neighbor untuk diagnosa penyakit layu (Wilt Disease).

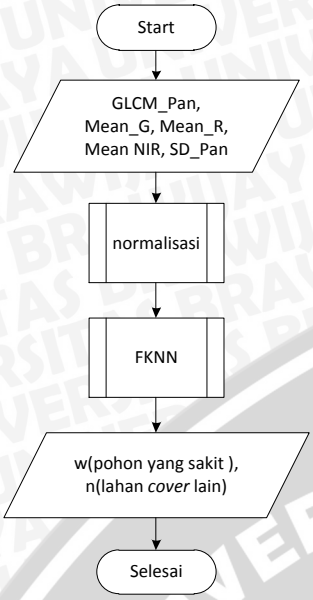
Penelitian dilakukan dengan tahap-tahap sebagai berikut :



Gambar 3.1 Tahap-tahap penelitian

3.1 Deskripsi Sistem

Secara umum, sistem yang dibuat merupakan perangkat lunak yang mengimplementasikan metode algoritma Fuzzy k-Nearest Neighbor (Fk-NN) untuk mendiagnosa penyakit layu. Lalu tingkat akurasi klasifikasi sistem ini akan diuji berdasarkan *wilt dataset*. Alur proses klasifikasi dijelaskan pada gambar.



Gambar 3.2 Alur Proses Pada Sistem

IV. Implementasi

Dalam implementasi metode Fuzzy K-nearest Neighbor (FK-NN) untuk diagnosa penyakit wilt pada hutan dibutuhkan beberapa aspek yang perlu diperhatikan yaitu segi perangkat keras dan perangkat lunak.

4.1 Lingkungan Implementasi Perangkat Keras

Dalam mengembangkan sistem dan penerapan metode penelitian ini digunakan beberapa komponen perangkat keras sebagai berikut :

Processor : Intel ® Core i3-4005u 1,7ghz

Memory : 6 GB

GPU : NVIDIA GT 930m 2 GB

Hard disk: 500 GB

4.2 Lingkungan Implementasi Perangkat Lunak

Pengembangan sistem dan penelitian ini dibutuhkan beberapa perangkat lunak yang digunakan sebagai berikut ini :

1. Sistem Operasi yang digunakan adalah Windows 8.1 64bit
2. Aplikasi pembangunan GUI dan code menggunakan Netbeans
3. Bahasa pemrograman yang dipakai yaitu bahasa pemrograman java.
4. Komponen untuk untuk menjalankan Netbeans yaitu Java Development Kit (JDK)

V. Pengujian Dan Analsisis

Pada bab ini berisi pengujian dan analisis dari hasil implementasi Fuzzy k-Nearest Neighbor untuk diagnosa penyakit layu berdasarkan wilt dataset.

5.1 Pengujian

Pengujian bertujuan untuk mengetahui tingkat akurasi klasifikasi oleh system. Pengujian dilakukan dengan 2 macam pengujian, yaitu : pengujian tingkat akurasi terhadap data latih dan pengujian tingkat akurasi terhadap nilai k.

5.1.1 Pengujian Tingkat Akurasi Terhadap Jumlah Data Latih

Pada pengujian ini menggunakan beberapa jumlah data latih yaitu : 60, 80, 120, 200, dan 350 data. Pada tabel dilakukan pengujian terhadap 30 data uji dengan nilai k = 4.

Tabel 5.1 Tingkat akurasi terhadap jumlah data latih

jumlah data latih	akurasi
60	76.67%
80	73.33%
120	70%
200	66.67%
350	66.67%

5.1.2 Pengujian Tingkat Akurasi Terhadap Nilai k

Pada pengujian tingkat akurasi terhadap pengaruh nilai k, digunakan sebanyak 200 data latih dan data uji yang digunakan adalah sebanyak 60 data. Untuk nilai k yang digunakan adalah k = 2 sampai dengan k = n. Hasil dari pengujian tingkat akurasi terhadap nilai k dapat dilihat pada tabel 5.2

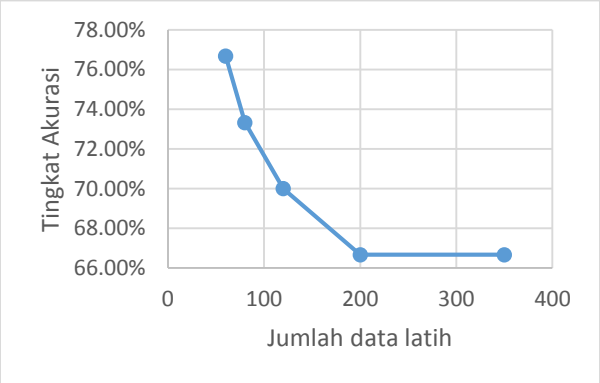
Tabel 5.2 Pengujian tingkat akurasi terhadap nilai k

k	akurasi
1	71.67%

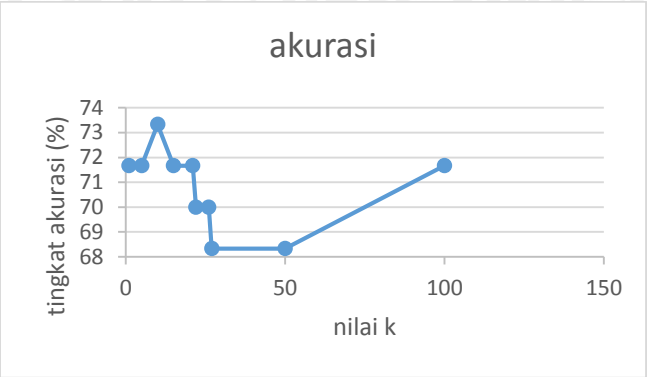
Tabel 5.2 Pengujian tingkat akurasi terhadap nilai k (lanjutan)

2	71.67%
3	73.33%
4	71.67%
5	71.67%
6	71.67%
7	71.67%
8	71.67%
9	73.33%
10	73.33%
15	71.67%
21	71.67%
22	70%
26	70%
27	68.33%
50	68.33%
100	71.67%

Berdasarkan hasil pengujian tingkat akurasi terhadap data latih, nilai k dan sebaran kelas, maka dibuatlah grafik hubungan tingkat akurasi Fuzzy k-Nearest Neighbor dengan tingkat akurasi k-Nearest Neighbor untuk melihat keakuratan sistem. Grafik ditunjukkan pada gambar dibawah.



Gambar 5.1 Grafik Perbandingan Jumlah Data Latih Terhadap Tingkat Akurasi



Gambar 5.2 Grafik Perbandingan Nilai k Terhadap Tingkat Akurasi

5.2 Analisis

Pada sub bab ini akan dilakukan analisis dari hasil pengujian dan penerapan Fuzzy k-Nearest Neighbor (Fk-NN) untuk diagnosa penyakit layu berdasarkan wilt dataset. Analisis ini dilakukan untuk mengetahui karakteristik data dari pengujian yang telah dilakukan.

5.2.1 Analisis hasil pengujian tingkat akurasi terhadap data latih

Pada pengujian tingkat akurasi terhadap jumlah data latih, pada grafik terlihat bahwa jumlah data latih berpengaruh cukup besar terhadap tingkat akurasi. Semakin banyak jumlah data latih, semakin menurun tingkat akurasi klasifikasi oleh system. Namun tingkat akurasi terlihat tidak berpengaruh saat menggunakan data latih sebanyak 350 data.

5.2.2 Analisis hasil pengujian tingkat akurasi terhadap nilai k

Berdasarkan grafik perbandingan tingkat akurasi diatas, terlihat nilai k memberikan pengaruh terhadap tingkat akurasi data, namun mulai menurun saat nilai k ≥ 10. Akan tetapi tingkat akurasi klasifikasi mulai menunjukkan peningkatan saat nilai k bernilai ≥ 50.

VI. PENUTUP

6.1 kesimpulan

Berdasarkan hasil penelitian dengan metode Fuzzy K-NN untuk mengetahui tingkat akurasi untuk klasifikasi Wilt Dataset pada hutan, dapat diambil kesimpulan sebagai berikut :

Metode Fuzzy k-Nearest Neighbor (FK-NN) bisa diterapkan untuk penyakit layu. Langkah pertama adalah menyiapkan data latih dan data uji yang akan digunakan lalu diproses menggunakan min-max normalization. Kemudian data latih dan data uji yang telah dinormalisasi dihitung menggunakan persamaan euclidean distance. Setelah itu output data penyakit layu ditransformasikan kedalam bentuk fuzzy. Pada akhirnya didapatkan hasil diagnose penyakit layu yang dihitung menggunakan persamaan Fuzzy k-Nearest Neighbor (FK-NN).

Jumlah data latih yang semakin banyak cukup berpengaruh pada penurunan tingkat akurasi klasifikasi penyakit layu dengan menggunakan metode Fuzzy k-Nearest Neighbor. Namun tingkat akurasi terlihat stabil saat menggunakan data latih tertentu.

Penentuan nilai k sedikit berpengaruh terhadap tingkat akurasi sistem, tingkat akurasi mulai menurun saat nilai k bernilai 1 hingga 27. Namun mulai mengalami peningkatan akurasi saat nilai k ≥ 50.

6.2 Saran

Ada beberapa saran yang dapat dijadikan pertimbangan agar penelitian selanjut nya menjadi lebih baik, yaitu :

1. Diharapkan penelitian berikut nya menggunakan tiga klasifikasi kelas berbeda agar hasil tersebut lebih akurat.
2. Dengan menggunakan metode lain atau menggunakan metode F-KNN yang telah dimodikasi diharapkan dapat memperoleh hasil akurasi klasifikasi yang lebih baik.
3. Menggunakan data latih dan data training dengan klasifikasi yang seimbang.
4. Menggunakan data training dan data latih yang disesuaikan dengan perbandingan jumlah data latih dan data uji dari sumber.

VII. Daftar Pustaka

[1] Adnan. 2013. Penentuan Kualitas Kopi Menggunakan Teknologi Near-Infra Red. German. Tersedia di:<https://postharvestnotes.wordpress.com/2013/06/11/penentuan-kualitas-kopi-menggunakan-teknologi-near-infra-red/> [diakses 1 Maret 2016]

[2] Aisyah, Imas SP., M.Si, 2012. Mengenal Gejala Penyakit Layu pada Tanaman dan Cara Menanganinya. Cianjur. Widyaiswara PPPPTK Pertanian Cianjur.

[3] Ir. Arifin Arief, M.P.,2001. Hutan dan Kehutanan. Kanisius. Yogyakarta.

[4] Matworks, 2016. Gray Level Co-Occurence Matrix (GLCM) Tersedia di : <http://www.mathworks.com/help/images/gray-level-co-occurrence-matrix-glcm.html?requestedDomain=www.mathworks.com> [diakses 1 Maret 2016]

[5] Risnandar, Cecep., Hutan Hujan Tropis. Tersedia di : <https://jurnalbumi.com/hutan-hujan-tropis/> [di akses 1 Maret 2016]

[6] Yuistianto, Adi. 2016. Penerapan Algoritma Fuzzy K-Nearest Neighbor (FK-NN) pada Klasifikasi Penyakit Kelainan Tulang Belakang Berdasarkan Vertebral Column Dataset. S1. Universitas Brawijaya.