**KLASIFIKASI JENIS TANAMAN IRIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR (KNN)**

**Ema Ainun Novia \* Farhan Maulana\*\*  Reghina Putri Anjani \*\*\* Seta Permana\*\*\*\***

**M. Nurkamal Fauzan\*\*\*\*\***

Mahasiswa Teknik Informatika Politeknik Pos Indonesia \*\*Jurusan Teknik Informatika Politeknik Pos Indonesia

Jl. Sari Asih No. 54 – Bandung 40151, Indonesia Telp. +6222 2009570, Fax. +6222 200 9568

\*) [emaainunnovia@gmail.com](mailto:emaainunnovia@gmail.com), \*\*)[farhanmaulana88@gmail.com](mailto:farhanmaulana88@gmail.com), \*\*\*) [reghinaputria@gmail.com](mailto:reghinaputria@gmail.com), \*\*\*\*)setapermana21@gmail.com,

\*\*\*\*\*\*) m.nurkamal.f@poltekpos.ac.id

**ABSTRAK**

Tanaman iris biasa digunakan untuk mempercantik halaman pekarangan rumah dengan aneka warna yang indah pada tanaman hias ini menjadi bahan perhatian khususnya bagi pecinta tanaman. Namun dengan banyaknya jenis tanaman iris membuat kita sulit untuk mengetahui nama tumbuhan yang kita minati.Sistem klasifikasi tanaman iris bekerja dengan cara membandingkan data petal dan sepal yang telah tersimpan pada database terhadap data petal dan sepal yang akan diuji. Data petal dan sepal akan diklasifikasikan dengan menggunakan penerapan metode K-Nearest Neighboor yaitu berfungsi untuk menghitung jarak terdekat. Data sampel pada penelitian ini menggunakan 3 klasifikasi yaitu virginaca, versi color, setosa. Data latih dengan 15 data.

**Kata Kunci**: Knn, klasifikasi, iris.

**I. PENDAHULUAN**

K-Nearest Neighboor (K-NN) merupakan metode yang dapat digunakan untuk proses klasifikasi atau mengelompokan data citra uji yang sudah ditentukan kelasnya pada data latih yang sudah disimpan pada datastore.

Tanaman iris biasa digunakan untuk mempercantik halaman pekarangan rumah dengan aneka warna yang indah pada tanaman hias ini menjadi bahan perhatian khususnya bagi pecinta tanaman. Namun dengan banyaknya jenis tanaman iris membuat kita sulit untuk mengetahui nama tumbuhan yang kita minati.

Sistem klasifikasi jenis tanaman iris ini beroperasi dengan melakukan perhitungan pada sepal dan petal pada data, lalu di hitung jarak kemiripan menggunakan metode KNN, setelah didapat nilai yang paling terdekat lalu sistem melakukan perhitungan vote kemiripan yang paling banyak muncul yang akan dijadikan output dari sistem ini.

**II. LANDASAN TEORI**

**2.1 Klasifikasi**

Klasifikasi sendiri berfokus pada metode untuk penyusunan data yang secara rapi mengacu kepada beberapa aturan atau kaidah yang sudah ditetapkan, dengan kata lain klasifikasi merupakan pengelompokkan terhadap sesuatu berdasarkan kelas. Pengertian klasifikasi secara umum adalah merupakan sebuah usaha yang menata pengetahuan atau juga benda ke dalam tata urutan yang pada umumnya sistematis. Klasifikasi sendiri memiliki proses seperti mendefinisikan kelas, menentukan hubungan antara kelas satu dan juga kelas lainnya, menentukan hubungan satu dan hubungan lainnya. Selain itu juga menetapkan elemen untuk kelas dalam suatu sistem klasifikasi.

**2.2 Algoritma KNN**

K-nearest neighbors atau knn adalah algoritma yang berfungsi untuk melakukan klasifikasi suatu data berdasarkan data pembelajaran (train data sets), yang diambil dari k tetangga terdekatnya (nearest neighbors). Dengan k merupakan banyaknya tetangga terdekat.

**A. Cara Kerja Algoritma K-Nearest Neighbors (KNN)**

K-nearest neighbors melakukan klasifikasi dengan proyeksi data pembelajaran pada ruang berdimensi banyak. Ruang ini dibagi menjadi bagian-bagian yang merepresentasikan kriteria data pembelajaran. Setiap data pembelajaran direpresentasikan menjadi titik-titik c pada ruang dimensi banyak.

**Klasifikasi Terdekat (Nearest Neighbor Classification)**

Data baru yang diklasifikasi selanjutnya diproyeksikan pada ruang dimensi banyak yang telah memuat titik-titik c data pembelajaran. Proses klasifikasi dilakukan dengan mencari titik c terdekat dari c-baru (nearest neighbor). Teknik pencarian tetangga terdekat yang umum dilakukan dengan menggunakan formula jarak euclidean. Berikut beberapa formula yang digunakan dalam algoritma knn.

**B. Langkah-Langkah Perhitungan**

a) Tentukan k bilangan bulat positif berdasarkan ketersediaan data pembelajaran.

b) Pilih tetangga terdekat dari data baru sebanyak k.

c) Tentukan klasifikasi paling umum pada langkah (ii), dengan menggunakan frekuensi terbanyak.

d) Keluaran klasifikasi dari data sampel baru.

Perancangan adalah suatu kegiatan yang memiliki tujuan untuk mendesign sistem baru yang dapat menyelesaikan masalah-masalah yang dihadapi perusahaan yang diperoleh dari pemilihan alternatif sistem yang terbaik.

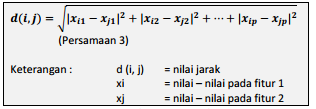
**2.3 Aplikasi**

Secara istilah pengertian aplikasi adalah suatu program yang siap untuk digunakan yang dibuat untuk melaksanankan suatu fungsi bagi pengguna jasa aplikasi serta penggunaan aplikasi lain yang dapat digunakan oleh suatu sasaran yang akan dituju. Menurut kamus computer eksekutif, aplikasi mempunyai arti yaitu pemecahan masalah yang menggunakan salah satu tehnik pemrosesan data aplikasi yang biasanya berpacu pada sebuah komputansi yang diinginkan atau diharapkan maupun pemrosesan data yang di harapkan.

**III. METODOLOGI PENELITIAN**

Data yang digunakan dalam penelitian ini yaiu  
didapat dari web <https://archive.ics.uci.edu/ml/datasets/iris> yang digunakan untuk menentukan klasifikasi dari tanaman iris.  
  
A. Metode *K-Nearest Neighboor*(KNN)  
Pengertian klasifikasi yaitu proses  
pengelompokan, artinya mengumpulkan benda / entitas  
yang sama serta memisahkan benda/entitas yang tidak  
sama[1]. KNN (*K-Nearest Neighboor*) adalah metode  
untuk melakukan klasifikasi terhadap objek  
berdasarkan data pembelajaran yang jaraknya paling  
dekat atau memiliki persamaan ciri paling banyak  
dengan objek tersebut. Dekat atau jauhnya tetangga  
biasanya dihitung dengan jarak *Euclidean*. Teknik ini  
sederhana dan dapat memberikan akurasi yang baik  
terhadap hasil klasifikasi[1].

sederhana dan dapat memberikan akurasi yang baik  
terhadap hasil klasifikasi[1].



Dalam algoritma ini, nilai k yang terbaik itu  
tergantung pada jumlah data. Ukuran nilai k yang besar  
belum tentu menjadi nilai k yang terbaik begitu juga  
sebaliknya[1].  
Langkah – langkah untuk menghitung algoritma  
KNN :  
1. Menentukan nilai k.2. Menghitung kuadrat jarak euclid (queryinstance) masing – masing objek terhadap  
training data.  
3. Mengurutkan objek – objek tersebut ke  
dalam kelompok yang mempunyai jarak  
euclid terkecil.  
4. Mengumpulkan label class Y (klasifikasi  
Nearest Neighborhood).

B. Data Klasifikasi Tanaman Iris

Dibawah ini adalah 6 dari 15 dataset yang akan digunakan sebagai perhitungan KNN terhadap data uji:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sepal Length (cm) | Sepal Width (cm) | Petal Length (cm) | Petal Width (cm) | Klasifikasi |
| 5,1 | 3,5 | 1,4 | 0,2 | Setosa |
| 4,9 | 3 | 1,4 | 0,2 | Setosa |
| 5,1 | 2,5 | 3 | 1,1 | Versicolor |
| 5,7 | 2,8 | 4,1 | 1,3 | Versicolor |
| 6,3 | 3,3 | 6 | 2,5 | Virginica |
| 5,8 | 2,7 | 5,1 | 1,9 | Virginica |

**IV. PEMBAHASAN**

Proses sistem pendeteksian klasifikasi jenis  
tanaman iris ini dibangun untuk mengklasifikasi 3  
jenis tanaman dengan 6 dari 15 data tanaman iris yang  
akan diujikan pada data uji dengan nilai k = 3 dengan data sebagai berikut:

Data Uji

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sepal Length (cm) | Sepal Width (cm) | Petal Length (cm) | Petal Width (cm) | Klasifikasi |
| 5 | 2 | 3 | 5 | ? |

Pada proses pengujian dari 6 dataset yang terdiri dari :  
a. 2 data tanaman iris Setosa  
b. 2 data tanaman iris Virginica  
c. 2 data tanaman iris Versicolor  
Pengujian kedekatan hasil data dilakukan dengan menghitung  
data uji yang akan di implementasikan kedalam dataset pada proses  
pengklasifikasian di setiap pengujian per-kategori jenis  
data tanaman irirs. Adapun data detail pengujiansistem secara  
keseluruhan, sebagai berikut :

Hasil Perhitungan :

1. √(5-5,1)2 + (2-3.5) 2 + (3-1.4) 2 + (5-0.2) 2 = 5,31 (Setosa)
2. √(5-4.9) 2 + (2-3) 2 + (3-1.4) 2 + (5-0.2) 2 = 5,15 (Setosa)
3. √(5-5.1) 2 + (2-2.5) 2 + (3-3) 2 + (5-1.1) 2 = 2,6 (Versicolor)
4. √(5-5.7) 2 + (2-2.8) 2 + (3-4.1) 2 + (5-1.3) = 4,00 (Versicolor)
5. √(5-6.3) 2 + (2-3.3) 2 + (3-6) 2 + (5-2.5) 2 = 4,31 (Virginica)
6. √(5-5.8) 2 + (2-2.7) 2 + (3-5.1) 2 + (5-1.9) 2 = 3,89 (Virginica)

K = 3 yang diambil adalah data terdekat ke 0 yaitu data 3,6, dan 4 yang berarti data tersebut memiliki perbandingan 2 versicolor:1 virginica.

V. KESIMPULAN

Berdasarkan dataset klasifikasi tanaman iris yang tersedia dari <https://archive.ics.uci.edu/ml/datasets/iris>, diambil 6 data yang masing-masing data merepresentasikan hasil dari klasifikasinya. Dengan data uji yang berisi Sepal Length (cm) = 5, Sepal Width (cm) = 2, Petal Length (cm) = 3, Petal Width (cm) = 5 didapat hasil dengan K=3 adalah 2 versicolor:1 virginica yang berarti data yang uji termasuk kedalam klasifikasi versicolor.

**REFERENSI**

[1] Prasetyo, E. (2012). Mengolah Data menjadi Informasi menggunakan Matlab. Yogyakarta.

[2] Syahid Dani, Jumadi, Nursantika Dian. (2016). SISTEM KLASIFIKASI JENIS TANAMAN HIAS DAUN PHILODENDRON MENGGUNAKAN METODE K-NEAREST NEIGHBOOR (KNN) BERDASARKAN NILAI HUE, SATURATION, VALUE (HSV), Jurnal JOIN. 20-23.

[3] Munir Rinaldi, Junaidi, Johanis. (2018). SISTEM KLASIFIKASI PENGGUNA NARKOBA MENGGUNAKAN METODE K-NEAREST NEIGHBOR. *JURNAL REALTECH*, Vol. 14.

[4] Supriyanto, Wahanan Nursinta Adi, Whidhiasih Retno Nugroho. (2013). KLASIFIKASI BUAH BELIMBING BERDASARKAN CITRA RED-GREEN-BLUE MENGGUNAKAN KNN DAN LDA. *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic.*