

Implementasi Algoritma KNN Untuk Mengetahui Kesuburan Tanah Untuk Tanaman Padi

Arwansyah, Madyana Patasik

STMIK Dipanegara Makassar

arwansyah@dipanegara.ac.id, madyanapatasik@gmail.com

Abstrak

Salah satu factor yang menyebabkan penurunan produksi padi adalah kondisi lahan yang tidak subur yang meliputi ketidakseimbangan antara unsure hara, ph tanah, dan material organik yang dibutuhkan oleh tanaman padi. Komponen ini merupakan unsure penting dalam pertumbuhan tanaman. Kandungan organik yang tidak seimbang dapat menghambat pertumbuhan dan perkembangan tanaman. Apabila hal ini terjadi maka petani akan mengalami penurunan produksi bahkan gagal panen. Penelitian ini mengimplementasikan algoritma yang terdapat pada data mining untuk menghasilkan informasi dan pengetahuan mengenai kesuburan suatu lahan yang akan digunakan untuk tanaman padi. algoritma yang digunakan adalah K-NN, cara kerja dari algoritma ini adalah dengan melakukan perbandingan dari data neighbor yang paling identik dengan data sample. Dengan kata lain, algoritma KNN akan membandingkan data tanah subur dengan data yang dijadikan sample. Hasil penelitian ini menunjukkan bahwa algoritma K-NN dapat menghasilkan informasi mengenai tanah yang subur untuk tanaman padi.

Keywords : *Algoritma, K-NN, Padi, Tanah Subur*

Abstract

One of the factors causing the decline of rice production is the condition of land is that infertile that cover unbalance between elements of nutrients, soil ph, and organic materials. This component is very important element in plant growth. The unbalanced organic content can inhibit the growth and development of the plant. If this happens, then farmers will experience a decrease in production and even harvest failure. This research implement algoritm that exist in data mining to generate useful information and knowledge about fertile land that wiil used for rice plants. The algorithm used is K-Means, the main method of this algorithm is make comparioson from neighbor data that very similar with data sample. In other word, KNN algorithm compare fertile land with data sample. The results of this study showing KNN algorithm can produce information about the fertile soil for rice plants.

Keywords: *Algorithm, K-NN, Soil, Fertile, Rice Plant*

1. Pendahuluan

Integrasi TIK yang efektif di sektor pertanian akan mengarah pada pertanian berkelanjutan melalui penyiapan informasi pertanian yang tepat waktu dan relevan, yang dapat memberikan informasi yang benar kepada petani dalam proses pengambilan keputusan untuk meningkatkan produktivitas. TIK dapat memperbaiki aksesibilitas petani dengan cepat terhadap informasi pasar, input produksi, tren konsumen, yang secara positif berdampak pada kualitas dan kuantitas produksi mereka. Informasi pemasaran, praktek pengelolaan ternak dan tanaman yang baru, penyakit dan hama tanaman atau ternak, ketersediaan transportasi, informasi peluang pasar dan harga pasar input maupun output pertanian sangat penting untuk efisiensi produksi secara ekonomi. Budidaya yang dilakukan oleh masyarakat saat ini di beberapa negara masih sebatas beberapa jenis tanaman, salah satu alasannya adalah kelebihan yang tidak dapat diprediksi, terutama daerah yang memiliki iklim yang tidak menentu sehingga masyarakat lebih memilih tanaman tertentu.

Perkembangan Teknologi Informasi dan Komunikasi (TIK) dari waktu ke waktu sangat pesat dan perannya dalam kehidupan sehari-hari dapat dirasakan di banyak bidang kegiatan kehidupan manusia, termasuk di bidang pertanian. Pembangunan pedesaan dan pertanian berkelanjutan merupakan isu penting yang dibahas strategis saat ini. Di era globalisasi pembangunan pertanian berkelanjutan tidak lepas dari pengaruh pesatnya perkembangan ilmu pengetahuan dan teknologi, termasuk perkembangan di bidang teknologi informasi dan komunikasi. Pemanfaatan TIK di bidang pertanian sering disebut dengan Electronic Agriculture (e-Agriculture). Informasi pertanian merupakan salah satu faktor terpenting dalam produksi dan tidak dapat dipungkiri bahwa informasi pertanian dapat mengarah pada perkembangan yang diharapkan. Informasi pertanian adalah aplikasi terbaik dari pengetahuan yang akan mendorong dan menciptakan peluang untuk pembangunan dan pengurangan kemiskinan.

Di beberapa Negara, berbagai jenis lahan digunakan dalam pertanian dan perkebunan. Masyarakat cenderung mengelola lahan tersebut untuk berbagai jenis tanaman. Sebagian besar tanaman yang dibudidayakan adalah tanaman pangan seperti jagung, padi, dan kedelai. Agar tanaman yang dibudidayakan dapat berkembang dan tumbuh dengan subur dan memberikan hasil panen yang tinggi, ada beberapa hal yang perlu diperhatikan oleh petani. Salah satu hal yang sangat penting dalam budidaya tanaman adalah tanah yang subur. Tanah merupakan komponen pokok yang menentukan tingkat keberhasilan dalam bidang pertanian. Untuk itu petani harus dapat memperhatikan unsure organik yang terkandung dalam tanah. Hal ini penting sebab unsure organik sangat dibutuhkan oleh tumbuhan dan merupakan salah satu hal yang membuat tanah tersebut tergolong ke dalam tanah yang subur.

Saat ini komputer bukan hanya sebagai mesin ketik atau alat komputasi saja yang dapat bekerja lebih cepat dan otomatis melainkan juga dapat digunakan sebagai alat dalam menganalisa dan menyelesaikan suatu permasalahan. Selain itu komputer juga dapat membantu dalam mengambil sebuah keputusan dari suatu permasalahan dengan cepat dengan tingkat keakuratan yang tinggi. Oleh karena itu para ahli dibidang tertentu mencoba memanfaatkan komputer menjadi suatu alat bantu yang dapat menirukan cara kerja otak manusia, sehingga diharapkan akan tercipta komputer yang dapat menyelesaikan suatu permasalahan yang kompleks dan mendukung seluruh tahap keputusan. Dengan demikian komputer dapat memberikan solusi baik dalam menyelesaikan suatu masalah maupun untuk memberikan solusi dalam pemilihan suatu keputusan. Salah satu bidang yang berpotensi memanfaatkan komputer dalam menyelesaikan dan memberikan solusi dari sebuah masalah adalah bidang pertanian dimana dinegara kita bidang ini merupakan aspek pokok dalam mata pencaharian sebagian besar masyarakat. Berdasarkan permasalahan yang ada terkait tanah dan tanaman maka hal yang sebaiknya dilakukan adalah menerapkan algoritma K-NN dalam menghasilkan suatu informasi mengenai lahan yang subur untuk tanaman padi.

2. Metode Penelitian

Dalam rangka keberhasilan penelitian, maka digunakan dua jenis metode penelitian untuk pengumpulan data yaitu :

1. Penelitian pustaka
Penelitian dilakukan melalui buku-buku pustaka dan internet yang dapat memberikan teori-teori mengenai permasalahan yang diteliti, kemudian mencocokkan dengan kemungkinan-kemungkinan yang terjadi dalam usaha penyelesaian masalah.
2. Penelitian lapangan

Penelitian yang dilakukan dengan mengunjungi langsung lokasi penelitian. Di tempat penelitian tersebut penulis melakukan pengamatan langsung terhadap objek penelitian dan melakukan Tanya jawab kepada petani yang terkait.

2.1 Teknik Pengumpulan Data

Pengumpulan data adalah salah satu hal yang penting dilakukan dalam memperoleh data yang diinginkan. Data yang dikumpulkan tersebut akan menjadi sebuah basis data. Dengan adanya data yang diambil tersebut, akan sangat membantu sebagai bahan pertimbangan dalam analisis sistem. Adapun teknik yang digunakan dalam pengumpulan data yaitu :

1. Teknik Wawancara
Teknik ini merupakan suatu teknik pengumpulan data dengan cara melakukan mewawancara terhadap beberapa petani yang terkait yang berada di wilayah objek penelitian.
2. Teknik Observasi
Teknik ini merupakan suatu teknik pengumpulan data dengan cara mengamati dan melihat langsung tanah atau lahan yang digunakan untuk budidaya tanaman padi.

2.2 Alat dan Bahan Penelitian

1. Alat Penelitian:
 - a. Hardware
 1. 1 unit Notebook
 2. Processor inter celeron, ~2.0GHz
 3. Memory RAM DDR 2 GigaByte
 4. Harddisk 500 GB
 - b. Software
 1. Windows 10
 2. Microsoft Office Excel 2007
2. Bahan Penelitian adalah data tanah, data tanaman padi, dan data unsure hara.

2.3 Tinjauan Pustaka

2.3.1 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui. Data mining dapat diartikan sebagai analisa otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola dan relasi-relasi yang tersembunyi dalam sejumlah data yang besar dengan tujuan untuk melakukan klasifikasi, estimasi, prediksi, association rule, clustering, deskripsi dan visualisasi. Secara sederhana data mining bisa dikatakan sebagai proses menyaring atau menambang pengetahuan dari sejumlah data yang besar.[1].

Tujuan Dari Adanya Data Mining adalah:

1. Explanatory, yaitu untuk menjelaskan beberapa kegiatan observasi atau suatu kondisi.
2. Confirmatory, yaitu untuk mengkonfirmasi suatu hipotesis yang telah ada.
3. Exploratory, yaitu untuk menganalisa data baru suatu relasi yang janggal

Proses Data Mining

1. Pembersihan data (Data Cleaning), untuk membersihkan noise dan data yang tidak konsisten.
2. Integrasi Data, penggabungan data dari berbagai sumber.
3. Transformasi data, data diubah menjadi bentuk yang sesuai untuk di mining.
4. Aplikasi teknik data mining, proses dimana teknik data mining diterapkan untuk mengekstrak pola-pola tertentu pada data.
5. Evaluasi pola yang ditentukan.
6. Presentasi pengetahuan, menggunakan teknik visualisasi untuk menampilkan hasil data mining kepada pengguna



Gambar 2.1 Tahapan dalam proses data mining[1]

Tools Data Mining

1. Klasifikasi
Klasifikasi merupakan tools data mining yang paling umum. Ciri-ciri klasifikasi adalah memiliki

definisi yang jelas tentang kelas-kelas dan training set. Klasifikasi bertujuan memprediksi kelas dari suatu data yang belum diketahui kelasnya. Dalam mencapai tujuannya tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data kedalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu.

2. **Estimasi**
Estimasi hampir sama dengan klasifikasi namun estimasi lebih menangani data kontinyu. Contoh estimasi antara lain memperkirakan jumlah anak dalam keluarga, memperkirakan pendapatan keluarga, dan memperkirakan nilai probabilitas pemegang kartu kredit terhadap pada hal yang ditawarkan oleh pihak bank, misalnya tawaran untuk pemasangan iklan dengan tema olah raga ski pada amplop tagihan.
3. **Prediksi**
Prediksi juga hampir sama seperti klasifikasi maupun estimasi, namun prediksi berusaha memprediksikan atau memperkirakan nilai atribut kelas dari suatu data untuk masa yang akan datang
4. **Pengelompokan afinitas**
Pengelompokan afinitas adalah pengelompokan berdasarkan hal – hal yang cenderung dilakukan bersamaan. Misalnya pengelompokan barang – barang yang biasanya dibeli bersamaan dalam suatu supermarket.
5. **Pengelompokan**
Pengelompokan adalah tugas data mining yang menggunakan metode membagi populasi yang heterogen menjadi sejumlah kelompok data yang homogeny. Pengelompokan tidak tergantung pada predefined classes dan training set. Data dikelompokan berdasarkan ciri-ciri yang sama. Pengelompokan sering dijadikan sebagai pendahuluan dalam pemodelan data mining.
6. **Deskripsi**
Deskripsi merupakan tugas sekaligus tujuan dari data mining, yaitu berusaha mendeskripsikan suatu yang sedang terjadi atau terdapat dalam suatu basis data yang rumit. Teknik yang memberikan deskripsi yang jelas misalnya teknik market basket analysis. [2]

2.3.2 Algoritma

Sejarah istilah “algoritma” Kata *algoritma* berasal dari latinisasi nama seorang ahli matematika dari Uzbekistan Al Khaw rizmi (hidup sekitar abad ke- 9), sebagaimana tercantum pada terjemahan karyanya dalam bahasa latin dari abad ke-12 "Algorithmi de numero Indorum". Pada awalnya kata algorisma adalah istilah yang merujuk kepada aturan-aturan aritmetis untuk menyelesaikan persoalan dengan menggunakan bilangan numerik arab (sebenarnya dari India, seperti tertulis pada judul di atas). Pada abad ke-18, istilah ini berkembang menjadi **algoritma**, yang mencakup semua prosedur atau urutan langkah yang jelas dan diperlukan untuk menyelesaikan suatu permasalahan.

Menurut Pranata (2005:8 Pengertian Algoritma adalah suatu prosedur yang tepat untuk memecahkan masalah dengan menggunakan bantuan komputer serta menggunakan suatu bahasa pemrograman tertentu seperti bahasa Pascal, Visual Basic, Java, dan masih banyak lagi bahasa yang lain.

Dalam kehidupan sehari-hari, sebenarnya kita juga menggunakan algoritma untuk melaksanakan sesuatu. Sebagai contoh, ketika kita menulis surat, maka kita perlu melakukan beberapa langkah sebagai berikut:

1. Mempersiapkan kertas dan amplop.
2. Mempersiapkan alat tulis, seperti pena atau pensil.
3. Mulai menulis.
4. Memasukkan kertas ke dalam amplop.
5. Pergi ke kantor pos untuk mengeposkan surat tersebut.

Jenis – Jenis Algoritma

Jenis-jenis Algoritma terdapat beragam klasifikasi algoritma dan setiap klasifikasi mempunyai alasan tersendiri. Salah satu cara untuk melakukan klasifikasi jenis-jenis algoritma adalah dengan memperhatikan paradigma dan metode yang digunakan untuk mendesain algoritma tersebut. Beberapa paradigma yang digunakan dalam menyusun suatu algoritma akan dipaparkan dibagian ini. Masing-masing paradigma dapat digunakan dalam banyak algoritma yang berbeda.

Divide and Conquer, paradigma untuk membagi suatu permasalahan besar menjadi permasalahan-permasalahan yang lebih kecil. Pembagian masalah ini dilakukan terus menerus sampai ditemukan bagian masalah kecil yang mudah untuk dipecahkan. Singkatnya menyelesaikan keseluruhan masalah dengan membagi masalah besar dan kemudian memecahkan permasalahan-permasalahan kecil yang terbentuk.

Dynamic programming, paradigma pemrograman dinamik akan sesuai jika digunakan pada suatu masalah yang mengandung sub-struktur yang optimal dan mengandung beberapa bagian permasalahan yang tumpang tindih. Paradigma ini sekilas terlihat mirip dengan paradigma Divide and Conquer, sama-sama mencoba untuk membagi permasalahan menjadi sub permasalahan yang lebih kecil, tapi secara intrinsik ada perbedaan dari karakter permasalahan yang dihadapi.

Metode serakah, Sebuah algoritma serakah mirip dengan sebuah Pemrograman dinamik, bedanya jawaban dari submasalah tidak perlu diketahui dalam setiap tahap; dan menggunakan pilihan “serakah” apa yang dilihat terbaik pada saat itu.

Adapun jenis-jenis algoritma yang lain adalah :

1. Bahasa Semu (pseudo code)
yaitu dengan menggunakan bahasa sehari-hari, tetapi harus jelas dan struktural.
2. Diagram Alir/Alur (Flowchart)
yaitu dengan membuat suatu penulisan atau penyajian algoritma berupa diagram yang menggambarkan susunan alur logika dari suatu permasalahan.[3]

Penggunaan Algoritma

Dengan algoritma, kita dapat mengatasi masalah dari yang sederhana sampai yang kompleks sekalipun. Namun, seorang user harus mampu membuat suatu program dengan menggunakan bahasa yang difahami oleh komputer. Sebelum disajikan dalam bentuk bahasa pemrograman, sebaiknya kita membuat diagram alir (Flow Chart) dan Pseudocode. Hal ini dimaksudkan agar dapat mempermudah kerja atau mempermudah dalam membuat program. Selain itu, algoritma dapat mengatasi masalah logika dan masalah matematika dengan cara berurutan, tetapi kadang-kadang algoritma tidak selalu berurutan, hal ini dikenal dengan proses percabangan. Kriteria Program Algoritma dalam Bidang Komputer. Pada dasarnya, komputer adalah mesin digital, artinya komputer hanya bisa mengenal kondisi ada arus listrik (biasanya dilambangkan dengan 1) dan tidak ada arus listrik (biasanya dilambangkan dengan 0). Dengan kata lain, kita harus menggunakan sandi 0 dan 1 untuk melakukan pemrograman komputer. Bahasa pemrograman yang menggunakan sandi 0 dan 1 ini disebut bahasa mesin. Karena bahasa mesin sangat susah, maka muncul ide untuk melambangkan untaian sandi 0 dan 1 dengan singkatan kata yang lebih mudah difahami manusia biasa disebut dengan mnemonic code. Bahasa pemrograman yang menggunakan singkatan kata ini disebut bahasa assembly. Program algoritma harus komplit, nyata, dan jelas. Meskipun tugas algoritma tidak menghasilkan solusi, tetapi proses harus berakhir hal ini disebut dengan semi algorithm (prosedur akan berjalan terus atau biasa disebut dengan perulangan). Intinya kita tidak boleh menambah masalah, akan tetapi kita harus mampu menyelesaikan masalah untuk mendapat hasil yang tepat. Adapun contoh algoritma seperti dalam menghitung luas lingkaran dari masukan berupa jari-jari lingkaran. Rumus lingkaran adalah $L = \pi * R * R$ Berikut ini adalah contoh algoritma untuk menghitung luas lingkaran:

1. Masukkan R
2. $\pi \approx 3,14$
3. $L = \pi * R * R$
4. Tulis L

Perhatikan tanda \approx pada baris kedua dan ketiga. Tanda ini berarti nilai di sebelah kanan diberikan pada operan di sebelah kiri. Sebagai contoh, untuk baris kedua, nilai 3,14 diberikan pada variabel π . Berikutnya, nilai $\pi * R * R$ diberikan pada variable L. Baris terakhir menuliskan luas lingkaran tersebut. Seperti yang dikemukakan di atas, bahwa algoritma ada yang tidak berurutan dan biasa disebut dengan perulangan. Adapun contohnya yaitu dalam penghitungan rata-rata dari sekumpulan data yang dimasukkan pengguna. Berikut ini adalah algoritma untuk menghitung rata-rata data yang dimasukkan pengguna:

1. Masukkan N
2. $i = 1$
3. $j = 0$
4. Selama ($i \leq N$) kerjakan baris 4 sampai dengan 7
5. Masukkan dt
6. $i = i + 1$
7. $j = j + dt$
8. Rata $= j / N$
9. Tulis rata

Baris pertama meminta pengguna memasukkan N, yaitu jumlah data. Pada baris kedua, variabel I, yang berguna sebagai pencacah banyaknya data yang telah dimasukkan pengguna, bernilai 1. Pada baris ketiga, variabel j, yang digunakan untuk menyimpan hasil penjumlahan data, diberi nilai 0. Baris keempat

memberikan perintah untuk mengulangi baris keempat sampai dengan baris ketujuh selama I kurang dari sama dengan N. Dengan kata lain, setelah lebih besar dari N, baris kedelapan yang dijalankan. Baris kelima meminta masukkan data yang ke-i. Baris keenam menambah variabel I dengan 1. Perhatikan arti dari perintah $i=i+1$ adalah nilai i ditambah dengan 1 kemudian hasilnya disimpan pada variabel I kembali. Baris ketujuh menambah variabel j dengan data yang dimasukkan pengguna. Sebagaimana dijelaskan di atas, variabel j digunakan untuk menyimpan hasil penjumlahan semua data, jadi untuk setiap masukan data, nilai variabel j harus ditambah dengan data. Baris kedelapan menghitung rata-rata dengan cara membagi hasil penjumlahan dengan banyaknya data. Baris terakhir menuliskan rata-rata tersebut. Tetapi banyak pemrogram yang sudah berpengalaman tidak pernah menuliskan algoritma di atas kertas lagi.. Artinya dia menuliskan algoritma itu di dalam kepalanya. [3]

2.3.3 Algoritma KNN

K-Nearest Neighbor (KNN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query*, akan ditemukan sejumlah *k* obyek atau (titik *training*) yang paling dekat dengan titik *query*. Klasifikasi menggunakan *voting* terbanyak diantara klasifikasi dari *k* obyek.. algoritma KNN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari *query instance* yang baru.

Algoritma metode KNN sangatlah sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke *training sample* untuk menentukan KNN-nya. *Training sample* diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi *training sample*. Sebuah titik pada ruang ini ditandai kelas *c* jika kelas *c* merupakan klasifikasi yang paling banyak ditemui pada *k* buah tetangga terdekat dari titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan *Euclidean Distance* yang direpresentasikan sebagai berikut :

$$D(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2},$$

dimana matriks $D(a, b)$ adalah jarak skalar dari kedua vektor *a* dan *b* dari matriks dengan ukuran *d* dimensi. Pada fase *training*, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data *training sample*. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk *testing data* (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vektor *training sample* dihitung dan sejumlah *k* buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.[4](Sumber : Jurnal vol. 39, pp. 1503-1509, 2011).

Perhitungan Metode K – Nearest Neighbor

Langkah-langkah untuk menghitung metode K-Nearest Neighbor :

1. Menentukan parameter K (jumlah tetangga paling dekat)
2. Menghitung kuadrat jarak euclid (query instance) masing-masing obyek terhadap data sampel yang diberikan
3. Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak euclid terkecil
4. Mengumpulkan kategori Y (Klasifikasi nearest neighbor)
5. Dengan menggunakan kategori nearest neighbor yang paling mayoritas maka diprediksikan nilai query instance yang telah dihitung

Contoh data, terdapat beberpa data yang berasal dari survey questioner tentang klasifikasi kualitas siswa berprestasi apakah Sangat berprestasi, berprestasi atau tidak, dengan objek testing menggunakan dua attribute yaitu prestasi akademik dan prestasi non akademi.[4]

Table 2.1 klasifikasi data

Ni s	X1 = Prestasi akademik	X2 = Prestasi non akademik	Y = klasifikasi
001	9	9	Sangat berprestasi
002	9	8	Sangat berprestasi
003	8	6	berprestasi
004	7	2	?

Kemudian akan dikelompokkan kembali siswa dengan attribute $X1 = 7$ dan $X2 = 2$ tanpa harus mengeluarkan biaya untuk melakukan survey, maka dapat diklasifikasikan siswa tersebut termasuk yang berprestasi, sangat berprestasi atau tidak berprestasi.

Adapun prosedur K-Nearest neighbor secara lengkap adalah sebagai berikut :

1. Menentukan parameter K (jumlah tetangga paling dekat), misalkan kita menggunakan $K = 6$.
2. Menghitung kuadrat jarak euclid (query instance) masing-masing objek terhadap sampel data yang diberikan koordinat query instance adalah $(7,2)$ dimana nilai tersebut berasal dari nilai attribute yang akan diproduksi

Ni s	X1 = Prestasi akademik	X2 = Prestasi non akademik	Y = klasifikasi
001	9	9	$(9-7)^2 + (9-2)^2 = 7,28$
002	9	8	$(9-7)^2 + (8-2)^2 = 6,32$
003	8	6	$(8-7)^2 + (6-2)^2 = 4,12$

3. Kemudian mengurutkan objek-objek termasuk ke dalam kelompok yang mempunyai jarak Euclid terkecil

Ni s	X1 = Prestasi akademik	X2 = Prestasi non akademik	Y = klasifikasi	Jarak terkecil	Apakah termasuk K-Nearest neighbor
001	9	9	$(9-7)^2 + (9-2)^2 = 7,28$	7,28	Ya
002	9	8	$(9-7)^2 + (8-2)^2 = 6,32$	6,32	Ya
003	8	6	$(8-7)^2 + (6-2)^2 = 4,12$	4,12	Ya

4. Mengumpulkan kategori Y (klasifikasi nearest neighbor).

Ni s	X1 = Prestasi akademik	X2 = Prestasi non akademik	Y = klasifikasi	Jarak terkecil	Apakah termasuk K-Nearest neighbor	Y= Kategori
001	9	9	$(9-7)^2 + (9-2)^2 = 7,28$	7,28	Ya	Sangat Berprestasi
002	9	8	$(9-7)^2 + (8-2)^2 = 6,32$	6,32	Ya	Sangat Berprestasi
003	8	6	$(8-7)^2 + (6-2)^2 = 4,12$	4,12	Ya	Berprestasi

5. Dengan menggunakan kategori nearest neighbor yang paling mayoritas, maka dapat diprediksikan nilai query instance yang telah dihitung . kita mempunyai data nis 003 berprestasi dengan nilai 4,12, karena 4,12 memiliki jarak paling dekat maka kita simpulkan bahwa nis 004 yang baru melewati tes prestasi dengan menetapkan klasifikasi attribute $X1 = 7$ dan $X2 = 2$ termasuk kategori Berprestasi

2.3.4 Tanah

Tanaman pangan adalah segala jenis tanaman yang dapat menghasilkan karbohidrat dan protein. Tanaman pangan merupakan sumber makanan yang secara alamiah maupun telah melalui proses, mengandung satu atau lebih senyawa yang berdasarkan kajian-kajian ilmiah dianggap mempunyai fungsi-fungsi fisiologis tertentu yang bermanfaat bagi kesehatan. Serta dikonsumsi sebagaimana layaknya makanan atau minuman, mempunyai karakteristik sensori berupa penampakan, warna, tekstur dan cita rasa yang dapat diterima oleh masyarakat. Selain tidak memberikan kontraindikasi dan tidak memberi efek samping pada jumlah penggunaan yang dianjurkan terhadap metabolisme. Jadi dapat disimpulkan, tanaman pangan berarti segala tanaman yang dapat dikonsumsi oleh masyarakat, sehat, layak dan memiliki kandungan yang bermanfaat.

Tanah merupakan tubuh alam tiga dimensi yang merupakan tempat aktivitas semua makhluk hidup termasuk tempat tumbuhnya tanaman. Tanah mempunyai karakteristik yang mempengaruhi pertumbuhan tanaman yang akan diusahakan. Klasifikasi tanah dan evaluasi lahan merupakan salah satu cara untuk mengetahui kecocokan suatu lahan untuk mengembangkan tanaman pertanian.

Tanah memiliki sifat yang bervariasi, yaitu terdiri dari sifat fisik, kimia dan biologi. Dengan bervariasinya sifat-sifat tersebut, maka tingkat kesuburan pada berbagai jenis tanah berbeda-beda pula,

karena kesuburan suatu tanah tergantung pada sifat-sifat tersebut. Oleh sebab itu diperlukan pemahaman mengenai karakteristik tanah sehingga dapat dimanfaatkan sesuai dengan potensinya.

Alih fungsi lahan pertanian produktif menjadi lahan non pertanian telah berlangsung dan sulit untuk dihindari sebagai akibat pesatnya laju pembangunan yang disertai dengan perubahan iklim. Untuk mendukung swasembada beras maka perlu perluasan areal tanaman padi sawah di daerah yang berpotensi untuk pengembangan sawah irigasi, namun umumnya terkendala oleh kualitas lahan yang rendah dan infrastruktur yang kurang memadai. Usaha yang dilakukan pemerintah untuk mempertahankan swasembada pangan adalah peningkatan mutu program intensifikasi, ekstensifikasi, diversifikasi, dan rehabilitasi lahan pertanian. Perluasan areal tanam (ekstensifikasi) merupakan salah satu pilihan untuk meningkatkan produksi berbagai komoditas termasuk padi sawah. [8]

Jenis Tanah

1. Alfisol

Tanah yang termasuk ordo Alfisol merupakan tanah-tanah yang terdapat penimbunan liat di horizon bawah (terdapat horizon argilik) dan mempunyai kejenuhan basa tinggi yaitu lebih dari 35% pada kedalaman 180 cm dari permukaan tanah. Liat yang tertimbun di horizon bawah ini berasal dari horizon di atasnya dan tercuci ke bawah bersama dengan gerakan air. Padanan dengan sistem klasifikasi yang lama adalah termasuk tanah Mediteran Merah Kuning, Latosol, kadang-kadang juga Podzolik Merah Kuning.

2. Aridisol

Tanah yang termasuk ordo Aridisol merupakan tanah-tanah yang mempunyai kelembaban tanah arid (sangat kering). Mempunyai epipedon ochrik, kadang-kadang dengan horizon penciri lain. Padanan dengan klasifikasi lama adalah termasuk Desert Soil.

3. Entisol

Tanah yang termasuk ordo Entisol merupakan tanah-tanah yang masih sangat muda yaitu baru tingkat permulaan dalam perkembangan. Tidak ada horizon penciri lain kecuali epipedon ochrik, albik atau histik. Kata Ent berarti recent atau baru. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Aluvial atau Regosol.

4. Histosol

Tanah yang termasuk ordo Histosol merupakan tanah-tanah dengan kandungan bahan organik lebih dari 20% (untuk tanah bertekstur pasir) atau lebih dari 30% (untuk tanah bertekstur liat). Lapisan yang mengandung bahan organik tinggi tersebut tebalnya lebih dari 40 cm. Kata Histos berarti jaringan tanaman. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Organik atau Organosol.

5. Inceptisol

Tanah yang termasuk ordo Inceptisol merupakan tanah muda, tetapi lebih berkembang daripada Entisol. Kata Inceptisol berasal dari kata Inceptum yang berarti permulaan. Umumnya mempunyai horizon kambik. Tanah ini belum berkembang lanjut, sehingga kebanyakan dari tanah ini cukup subur. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Aluvial, Andosol, Regosol, Gleihumus, dan lain-lain.

6. Mollisol

Tanah yang termasuk ordo Mollisol merupakan tanah dengan tebal epipedon lebih dari 18 cm yang berwarna hitam (gelap), kandungan bahan organik lebih dari 1%, kejenuhan basa lebih dari 50%. Agregasi tanah baik, sehingga tanah tidak keras bila kering. Kata Mollisol berasal dari kata Mollis yang berarti lunak. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Chernozem, Brunizem, Rendzina, dan lain-lain.

7. Oxisol

Tanah yang termasuk ordo Oxisol merupakan tanah tua sehingga mineral mudah lapuk tinggal sedikit, kandungan liat tinggi tetapi tidak aktif sehingga kapasitas tukar kation (KTK) rendah yaitu kurang dari 16 me/100 g liat, dan banyak mengandung oksida-oksida besi atau oksida Al. Berdasarkan pengamatan di lapangan, tanah ini menunjukkan batas-batas horizon yang tidak jelas. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Latosol (Latosol Merah & Latosol Merah Kuning), Lateritik, atau Podzolik Merah Kuning.

8. Spodosol

Tanah yang termasuk ordo Spodosol merupakan tanah dengan horizon bawah terjadi penimbunan Fe dan Al-oksida dan humus (horizon spodik) sedang, di lapisan atas terdapat horizon eluviasi (pencucian) yang berwarna pucat (albic). Padanan dengan sistem klasifikasi lama adalah termasuk tanah Podzol.

9. Ultisol

Tanah yang termasuk ordo Ultisol merupakan tanah-tanah yang terjadi penimbunan liat di horizon bawah, bersifat masam, kejenuhan basa pada kedalaman 180 cm dari permukaan tanah kurang dari 35%. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Podzolik Merah Kuning, Latosol, dan Hidromorf Kelabu

10. Vertisol

Tanah yang termasuk ordo Vertisol merupakan tanah dengan kandungan liat tinggi (lebih dari 30%) di seluruh horizon, mempunyai sifat mengembang dan mengkerut. Jika kering tanah mengkerut sehingga tanah pecah-pecah dan keras. Jika basah mengembang dan lengket. Padanan dengan sistem klasifikasi lama adalah termasuk tanah Grumusol atau Margalit. [8]

3. Hasil dan Pembahasan

Ada beberapa tahap yang dilakukan dalam analisis algoritma K-Means yang meliputi :

1. Menentukan mengumpulkan informasi mengenai data tanah yang akan dijadikan data sample
2. Mengumpulkan informasi mengenai unsure organik yang dibutuhkan tanaman padi
3. Menentukan jumlah cluster
4. Menghitung jarak setiap data dengan mengambil nilai rata-rata dari unsure organik yang dibutuhkan oleh tanah
5. Mengelompokkan jarak data terdekat dengan centroid

Sample Data Tanah

1. Tanah regosol
Ciri khas: kasar berbutir, berwarna abu-abu dan kuning, dan dibuat dari sedikit organik
Cocok untuk tanaman: Jagung, tembakau, dan buah-buahan
2. Tanah latosol
Ciri: berwarna merah dan kuning, mengandung bahan organik sedang, asam
Cocok untuk tanaman: Beras, kelapa, karet, kopi
3. Tanah humus
Ciri khas: berwarna coklat, subur
Cocok untuk tanaman: Kelapa, nanas, dan nasi
4. Tanah gambut
Ciri khas: Tidak subur
Tidak cocok untuk pertanian
5. Tanah dari alluvium
Ciri khas: berwarna abu-abu, subur
Cocok untuk tanaman: Beras, tebu, kelapa, tembakau, dan buah-buahan
6. Tanah podzol
Ciri khas: Gizi buruk, tidak subur, berwarna merah dan kuning
Cocok untuk tanaman: Kelapa dan jambu mete
7. Tanah laterit
Ciri khas: berwarna merah dan kuning, tidak subur, tidak subur, kehilangan gizi, tandus
Cocok untuk tanaman: Kelapa dan jambu mete
8. Tanah litosol
Characteristic : the minerals still tied at the large grains, less fertile
Suitable for plant : Big plants in the forest
9. Soil of renzina
Characteristic : colored black, less nutrient
Suitable for plant : tobacco, cashew nuts
10. Soil of mediteran
Characteristic : colored red and brown, less fertile
Suitable for plant : Teak wodd, tobacco, and cashew nuts.
11. Soil of sand
Not recommended for agriculture

Sample data tanah merupakan data yang akan di analisis menggunakan algoritma K-NN, membagi ke dalam 2 kategori dan menghasilkan informasi mengenai tanah yang subur dan tidak subur untuk tanaman padi. Tabel hasil analisa algoritma K-NN dapat dilihat pada tabel 3.1 berikut :

Table 3.1 Hasil Perhitungan Algoritma K-NN

No	Data Example	X1 = Ph	X2 = (N)	X3 = (P)	X4 = (K)	Organic	Y = Classification	Information	Suggestion
1	First land	5.00	0.22	23	23	6	10.60085780	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
2	Second land	4.00	0.44	30	30	7	2.56	Fertility	Monitoring the nutrient in the soil
3	Third land	4.50	0.54	40	40	5	13.58806727	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
4	Fourth land	6.00	0.34	34	34	6	3.964088343	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
5	Fifth land	6.5	0.37	32	32	4	2.928213424	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
6	Sixth land	4.4	0.47	35	35	5	6.702610795	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
7	Seventh land	3.70	0.33	38	38	3	11.5133466	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
8	Eighth land	3.50	0.35	36	36	7	1.00719678	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
9	Ninth land	6.20	0.77	44	44	8	19.20090420	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
10	Tenth land	7.00	0.57	35	35	9	7.078751655	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic
11	Eleventh land	4.80	0.45	39	39	3	12.47443486	Net fertility	Fertilization, Fixation of N-biological, Water, Addition of M-organic

4. Kesimpulan

Berdasarkan hasil pembahasan menggunakan algoritma K-NN maka diketahui bahwa:

1. Hasil dari penelitian ini menunjukkan bahwa implemementasi algoritma K-NN dapat menghasilkan informasi mengenai tanah yang subur untuk tanaman padi.
2. Terdapat 1 tanah yang subur sedangkan untuk tanah yang lain terdapat informasi mengenai hal yang harus dilakukan petani agar tanah tersebut menjadi subur.

Daftar Pustaka

- [1] Kusrini, & Emha Taufik Luthfi. (2009). *Algoritma Data mining*. Yogyakarta: Andi.
- [2] Fayyad, Usama. 1996. *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- [3] Munir, Rinaldi. 2005.” **Strategi Algoritmik**”. Bandung. Institut Teknologi Bandung.
- [4] Pranata.2005. “**Strategi Algoritmik**”. Bandung. Institut Teknologi Bandung.

- [5] *Jurnla* vol. 39, pp. 1503-1509, 2011.diakses pada tanggal 25-februari-2014
- [6] Witten, I. H., Frank, E., Hall, M. A. 2011. *Data Mining Practical Machine Learning Tools and Techniques* (3rd ed). USA: Elsevier
- [7] Sevani,Nina.,Marimin. Sukoco,Heru 2009, “Sistem Pakar Penentuan Kesesuaian Lahan Berdasarkan Faktor Penghambat Terbesar (Maximum Limitian Factor) Untuk Tanaman Pangan”, *Jurnal Informatika* Vol.10 No 1.
- [8] Balai Penelitian Tanah. 2005.*Analisis Kimia Tanah, Tanaman, Air, Dan Pupuk*. Badan Penelitian dan Pengembangan Pertanian Depertemen Pertanian. Pulung, M. A. 2009. *Kesuburan tanah*. Penerbit Universitas Lampung. Bandar Lampung. 45 p.