KNN

by Syahminan Syahminan

PAPER · OPEN ACCESS

KNN (K-Nearby Neighbor) for identifying agricultural land

IOP | **ebooks**™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

KNN (K-Nearby Neighbor) for identifying agricultural land

S Syahminan[1,*], J Maknunah[2], R Dijaya[3] and H Hindarto[3]

[1] Universitas Kanjuruhan Malang, Malang, Indonesia
[2] STMIK PPKIA Pradnya paramita, Malang, Indonesia
[3] Universitas Muhammadiyah Sidoarjo, Sidoarjo, Indonesia

*syahminan@unikama.ac.id

Abstract. Agricultural land in Malang Regency is a vital element because most of the population are farmers and working in the agricultural sector. The agricultural land area of Malang Regency is around 15.74% (49,593 hectares) which is paddy fields and 31.31% (98,641 hectares) are fields. Not all types of crops can flourish on each of these farms because every land and plant has different characteristics. Therefore, to help determine the type of food crops that are suitable for a land, a method system for identifying land for food crop cultivation using the KNN (K-Nearest Neighbor) method was developed. With the existence of a land identification decision system for food crop cultivation, this method approach makes it easier to determine which plants are in accordance with the conditions of the land, in order to increase agricultural productivity in the southern part of Malang Regency.

1. Introduction

Agricultural land is typically land concerned for agriculture. Land evaluation is the process of assessing the potential of land for certain uses [1]. Plants that can grow on land are plants that are able to adapt to the surrounding environment. Therefore, not all types of plants can grow on any land [2]. In the Dutch colonial era, the soil classification system was first introduced by Van Mohr in 1910. This soil classification was based on a combination of various parent materials and the weathering process which was emphasized on leaching intensity in relation to climate influences [3], thickness conditions of mud land, a climate which includes temperature and rainfall, as well as the altitude above sea level (topography) [4].

When the land to be used for cultivating plants or food crop is lacking or not suitable, it will produce bad results of agricultural product. Food crop cultivation business often experiences obstacles, one of which is in determining the land to be used as a place of cultivation [5]. The types of food crops that are prioritized in this study is maize plants or commonly known and referred to as corn [6]. Good land characteristics for maize plants or corn are to consider land characteristics such as temperature, rainfall, humidity, drainage, soil texture, crude material, soil depth, base saturation, pH level, and organic c [7]. In general, the characteristics of land suitable for each plant differ according to the characteristics of the plant, but there are some plants that can grow on land with the same characteristics even though the plants are different [8].

Based on the analysis conducted by The Office of Agriculture and Plantation Service Malang Regency, the yields obtained each year are uncertain. In addition, the productivity rate of food crops often decreases. Processing land that is not in accordance with the characteristics of the type of plant can hamper the process of farming and cultivation. In the end, improper processing can be one of the

causes of crop failure [9]. The lack of knowledge and understanding of farmers about the characteristics of the land to be processed and the difficulty of obtaining the right data about the characteristics of the land can make it difficult for farmers to know whether the land is suitable or not for cultivation of food crops.

2. KNN method (K –Nearest Neighbor)

Nearest neighbor is an approach to look for cases by calculating the closeness between new cases and old cases. This approach is based on matching the weights of a number of features [10].

The simple KNN algorithm works based on the shortest distance from the query instance to the training sample to determine its KNN [11]. The sample training is projected into a large dimension space, where each dimension represents the features of the data. This space is divided into sections based on the training sample classification [12]. A point in this space is marked class c if class c is the most common classification in the nearest neighbor k from that point. Near or near neighbors are usually calculated based on Euclidean Distance.

KNN Formula:

$$d_i = \sqrt{\sum_{i=1}^{p} \left(x_{2i} - x_1 \right)^2}$$

With d is the distance between points on training x data and testing data points to be classified, where x = x1, x2, ..., xi and y = y1, y2, ..., you and I represent attribute values and n is attributed dimension [13]

Steps to calculate the K-Nearest Neighbor Algorithm method:

- Determining Parameter K (Number of closest neighbors.
- Calculating the square of the Euclid distance (query instance) of each object against the sample data provided.
- Then, sorting the objects into groups that have the smallest Euclid distance.
- Collecting Y category (Nearest Neighbor Classification) [14].

By using the most majority Nearest Neighbor category, it can be predicted for query instance values [15]. The calculation performed by the formula above with the test data is presented in the following table 1 below.

Table 1. Sample data of the assessed land.

Land Characteristics	Pagelaran Sub-District	Kromengan Sub-District	Dampit Sub-District
Temperature (ºC)	24	22	25
Rainfall (mm/years)	1793	1600	2294
Humidity (%)	87%	78%	24%
Drainage	Average	Average	Good
Soil texture	Average	Refined	Refined
Crude material	12	-	14
Soil depth	>100	>100	76
Base saturation	87.3	87	36
Level of Ph	5-6	7.1	5.9
C organic Percentage (%)	1.2	1.2	1.3

3. Testing results

The following figure 1 presents and explains the results of land assessment performed in each sub-district.

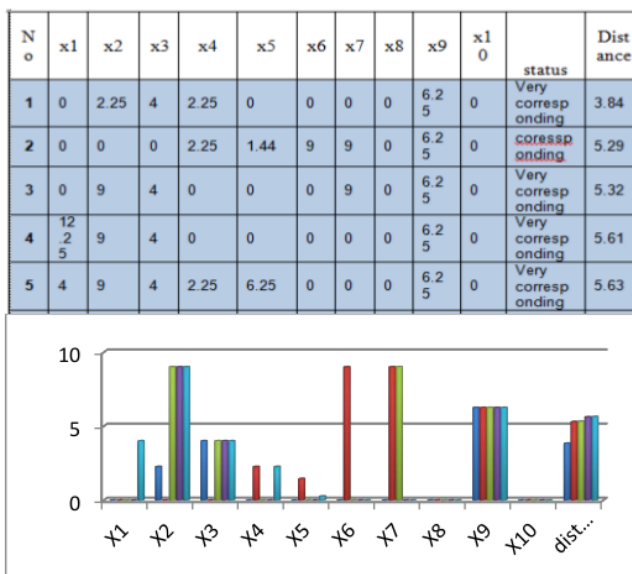| No	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	status	Distance
1	0	2.25	4	2.25	0	0	0	0	6.25	0	Very corresponding	3.84
2	0	0	0	2.25	1.44	9	9	0	6.25	0	coressponding	5.29
3	0	9	4	0	0	0	9	0	6.25	0	Very corresponding	5.32
4	12.25	9	4	0	0	0	0	0	6.25	0	Very corresponding	5.61
5	4	9	4	2.25	6.25	0	0	0	6.25	0	Very corresponding	5.63

Figure 1. The results of the first land assessment.

The following figure 2 presents the results of land assessment performed in Kromengan and Pagelaran sub-districts.
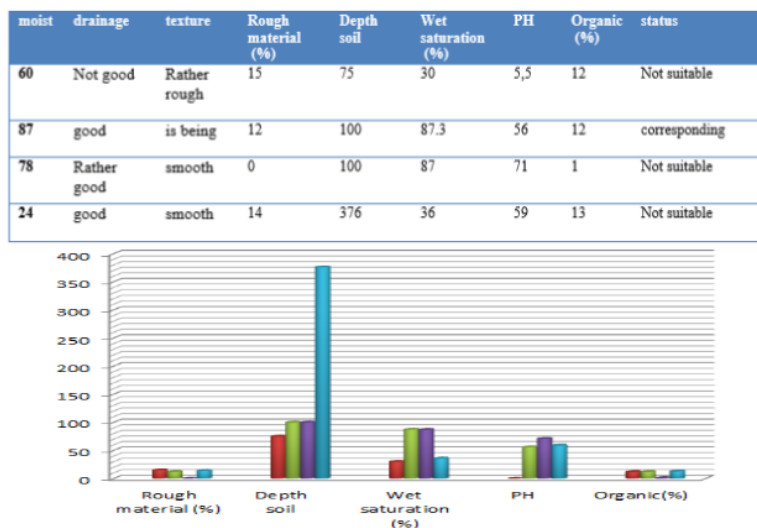
moist	drainage	texture	Rough material (%)	Depth soil	Wet saturation (%)	PH	Organic (%)	status
60	Not good	Rather rough	15	75	30	5,5	12	Not suitable
87	good	is being	12	100	87.3	56	12	corresponding
78	Rather good	smooth	0	100	87	71	1	Not suitable
24	good	smooth	14	376	36	59	13	Not suitable

Figure 2. The results of the second land assessment.

The data presented in the table above is the result of a land assessment trial conducted in Pegelaran and dampit Sub-districts. Based on the results of the land assessment trials obtained, it is known that the land located in the Dampit sub-district is very suitable and good for conducting corn cultivation.

4. Conclusion

Based on the results of tests using the KNN method that has been carried out above in three sub-districts namely Pagelaran Sub district, Kromengan Sub district, and Dampit Sub district, it is known that the K value is 14. From the k value, the following results are obtained: the corn plant has a very suitable level of suitability to be cultivated in Pagelaran Sub district. In Kromengan Sub district, based on the closest k distance value, the corn plant has a very suitable level of suitability to be cultivated in the Kromengan Sub district. In Dampit Sub district, based on the closest k distance value, the corn plant has a very suitable level of suitability to be cultivated in Dampit Sub district.

References

[1]   Ritung S 2007 Guideline for Evaluating Suitability Soil Research Center and World Agroforestry Centre *jurnal penelitian*

[2]   Hasti S 2013 *Sistem Pendukung Keputusan Kesesuaian Lahan Di Kecamatan Percut Sei Tuan* (Medan: Departemen Teknologi Pertanian Fakultas Pertanian USU)

[3]   Mega I M 2010 *Klasifikasi Tanah Dan Kesesuaian Lahan* (Denpasar: Program Studi Agroekoteknologi Fakultas Pertanian Universitas Udayana)

[4]   Reflan N 2018 implementasi metode k-nearest neighbor(knn) untuk memprediksi varietas padi yang cocok untuk lahan pertanian *Jurnal INFORMA Politeknik Indonusa Surakarta* **4**(2)

[5]   Fachrista I A and Sarwendah M 2014 Persepsi dan tingkat adopsi petani terhadap inovasi teknologi pengelolaan tanaman terpadu padi sawah *Agriekonomika* **3**(1) 1-10

[6]   Sirappa M P 2003 prospek pengembangan sorgum di Indonesia Sebagai komoditas alternatif untuk Pangan, pakan, dan industri *Jurnal Litbang Pertanian* **22**(4)

[7]   Prasetyo B H and Suriadikarta D A 2006 Karakteristik, potensi, dan teknologi pengelolaan tanah Ultisol untuk pengembangan pertanian lahan kering di Indonesia *Jurnal Litbang Pertanian* **25**(2) 39-46

[8]   Margolang R D 2015 The Characteristis of Some of The Physical, Chemical, and Biological Properties of Soil in Organic Farming Systems *Jurnal Online Agroekoteaknologi* **3**(2) 717-723

[9]   Suharta N 2010 Karakteristik dan permasalahan tanah Marginal dari batuan sedimen masam Di kalimantan *Jurnal Litbang Pertanian* **29**(4) 2010

[10] Nugraheni O D Palm Kernel Classification Based on Texture and Morphological Image Analysis Using K-Nearest Neighborhood (KNN)

[11] Khotimah N and Istiawan D 2018 Perbandingan Algoritma C4. 5, Naïve Bayes dan K-Nearest Neighbour untuk Prediksi Lahan Kritis di Kabupaten Pemalang *Proceeding of The URECOL* 41-50

[12] Bode A 2017 K-nearest neighbor dengan feature selection menggunakan backward elimination untuk prediksi harga komoditi kopi arabika *ILKOM Jurnal Ilmiah* **9**(2)

[13] Nuari R 2018 Implementasi Metode k-Nearest neighbour (knn) untuk memprediksi Varietas padi yang cocok untuk lahan pertanian *Jurnal INFORMA Politeknik Indonusa Surakarta* **4**(2)

[14] Rudiono R 2019 *Klasifikasi Jenis Tanah Berbasis Website Menggunakan Ekstraksi Ciri Histogram dan Algoritma K-Nearest Neighbor (KNN)* (University of Technology Yogyakarta)

[15] Andrean P 2019 Penerapan metode k-Nn untuk memprediksi Hasil pertanian di Kabupaten malang *Jati (jurnal mahasiswa teknik informatika)* **3**(1)

KNN

PRIMARY SOURCES

1	Lotar Mateus Sinaga, Sawaluddin, Saib Suwilo. "Analysis of classification and Naïve Bayes algorithm k-nearest neighbor in data mining", IOP Conference Series: Materials Science and Engineering, 2020 Publication	6%
2	Nur Rosyid Mubatada'i, Ali Ridho Barakbah, Afrida Helen. "Semantic Information Retrival for Scientific Experimental Papers with Knowlege based Feature Extraction", INOVTEK Polbeng - Seri Informatika, 2019 Publication	6%
3	repository.unika.ac.id Internet Source	5%
4	Intan Nurma Yulita, Mohamad Ivan Fanany, Aniati Murni Arymurthy. "Shallow Classifier with Sampling for Sleep Stage Classification of Autism Patients", International Journal of Engineering & Technology, 2018 Publication	2%