

The Application of the K-Nearest Neighbor (KNN) Method to Determine House Locations in the Batuphat and Tambon Tunong Areas, Aceh

Abil Khairi^{1*}, Irgi Fahrezi², Irfan Sahputra³, Said Fadlan Anshari⁴

^{1,2,3,4} Teknik Informatika, Universitas Malikussaleh, Indonesia

*Corresponding Author Email: *abillzein792@gmail.com

ABSTRAK

Received: 29 December 2023
Revised: 31 December 2023
Accepted: 31 December 2023
Available online: 1 January 2024

Kata Kunci:

Metode K-Nearest Neighbors (KNN), lokasi rumah, Batuphat, Tambon Tunong, pengklasifikasian.

Penelitian ini bertujuan untuk menerapkan metode K-Nearest Neighbors (KNN) dalam mencari lokasi sebuah rumah yang berada tepat di tengah perbatasan antara Batuphat dan Tambon Tunong. Permasalahan yang dihadapi oleh teman-teman kampus adalah kesulitan dalam menentukan apakah rumah tersebut termasuk dalam wilayah Batuphat atau Tambon Tunong. Metode KNN digunakan karena kemampuannya dalam melakukan klasifikasi berdasarkan jarak terdekat dengan tetangga terdekat. Data yang digunakan dalam penelitian ini mencakup informasi lokasi rumah dan wilayah Batuphat serta Tambon Tunong. Proses training dilakukan untuk membentuk model KNN berdasarkan data lokasi yang telah diketahui, sedangkan proses testing digunakan untuk mengklasifikasikan lokasi rumah yang tidak diketahui ke dalam wilayah Batuphat atau Tambon Tunong. Hasil penelitian menunjukkan bahwa metode KNN dapat digunakan untuk menentukan lokasi sebuah rumah yang berada di perbatasan Batuphat dan Tambon Tunong. Dengan mempertimbangkan jarak terdekat dengan tetangga terdekat, rumah dapat diklasifikasikan ke dalam salah satu wilayah dengan tingkat akurasi yang tinggi. Penelitian ini memberikan kontribusi dalam memberikan solusi bagi teman-teman kampus yang menghadapi kesulitan dalam menentukan lokasi rumah di perbatasan Batuphat dan Tambon Tunong. Metode KNN dapat digunakan sebagai alat bantu yang efektif dalam menyelesaikan masalah ini. Penelitian ini juga dapat menjadi dasar untuk pengembangan lebih lanjut dalam bidang pengklasifikasian lokasi berdasarkan metode KNN.

ABSTRACT

Keywords:

K-Nearest Neighbors (KNN) method, house location, Batuphat, Tambon Tunong, classification

This study aims to apply the K-Nearest Neighbors (KNN) method to find the location of a house situated precisely on the border between Batuphat and Tambon Tunong. The issue faced by the college friends is the difficulty in determining whether the house falls within the Batuphat or Tambon Tunong area. The KNN method is used due to its ability to classify based on the nearest neighbors' distance. The data used in this research includes information on the house's location and the Batuphat and Tambon Tunong areas. The training process is conducted to form the KNN model based on the known location data, while the testing process is employed to classify the unknown house location into either the Batuphat or Tambon Tunong area. The results of the study demonstrate that the KNN method can be utilized to determine the location of a house situated on the border between Batuphat and Tambon Tunong. By considering the nearest neighbors' distance, the house can be classified into one of the areas with a high level of accuracy. This research contributes to providing a solution for college friends who face difficulties in determining the house location on the Batuphat and Tambon Tunong border. The KNN method can serve as an effective tool in addressing this problem. Moreover, this study can serve as a basis for further development in the field of location classification based on the KNN method.

1. INTRODUCTION

Dalam pencarian lokasi rumah di wilayah Batuphat atau Tambon Tunong, klasifikasi data menjadi suatu hal yang penting. Salah satu metode klasifikasi yang sering digunakan adalah metode K-Nearest Neighbors (KNN). Metode KNN telah banyak diterapkan dalam berbagai bidang seperti pengenalan pola, pengenalan teks, dan pengolahan objek. Metode KNN memiliki keunggulan dalam pengolahan data training dan testing yang berjumlah besar.

Selain itu, KNN juga telah diterapkan dalam berbagai penelitian, seperti pengklasifikasian penyakit diabetes dan penerima beasiswa PPA dan BBM. Metode KNN merupakan metode klasifikasi berbasis contoh atau non-parametrik yang dianggap sebagai metode yang sederhana namun efektif dalam proses data mining. Namun, algoritma KNN memiliki kelemahan dalam pengolahan dataset yang besar dan kompleks serta tidak efisien dalam hal waktu.

Metode KNN menggunakan pengukuran jarak yang sesuai untuk mengklasifikasikan data baru. Dalam proses ini, jarak dengan tetangga terdekat sebanyak K dihitung, dan label kelas tetangga terdekat diprediksi sebagai label kelas untuk data baru. Pemilihan nilai K memiliki pengaruh signifikan terhadap akurasi KNN. Jika nilai K terlalu kecil, KNN dapat menjadi sensitif terhadap noise, sedangkan jika nilai K terlalu besar, dapat menyebabkan bias pada model.

Dalam konteks penerapan metode KNN untuk mencari lokasi rumah di wilayah Batuphat atau Tambon Tunong, metode KNN dapat menjadi alat yang berguna. Dengan mempertimbangkan jarak terdekat dengan tetangga terdekat, rumah-rumah dapat diklasifikasikan ke dalam wilayah Batuphat atau Tambon Tunong. Penerapan metode KNN ini memiliki keunikan dalam memetakan atau mengklasifikasikan data rumah ke dalam kelas-kelas yang telah ditentukan sebelumnya.

Penelitian ini bertujuan untuk menerapkan metode K-Nearest Neighbors (KNN) dalam mencari lokasi rumah di wilayah Batuphat atau Tambon Tunong. Melalui penelitian ini, diharapkan dapat memberikan kontribusi dalam memudahkan proses pencarian lokasi rumah di wilayah yang memiliki batasan yang kabur.

2. RESEARCH METHODS

Beberapa cara digunakan untuk pengukuran jarak antara data baru (*data testing*) dengan data lama (*data training*), antara lain *manhattan distance* (*city block distance*) dan *euclidean distance*. Pengukuran jarak yang paling sering digunakan adalah *euclidean distance*. Jarak *euclidean distance* didefinisikan seperti pada persamaan 1:

$$d_i = \sqrt{\sum (x1 - x2)^2 pi} \tag{1}$$

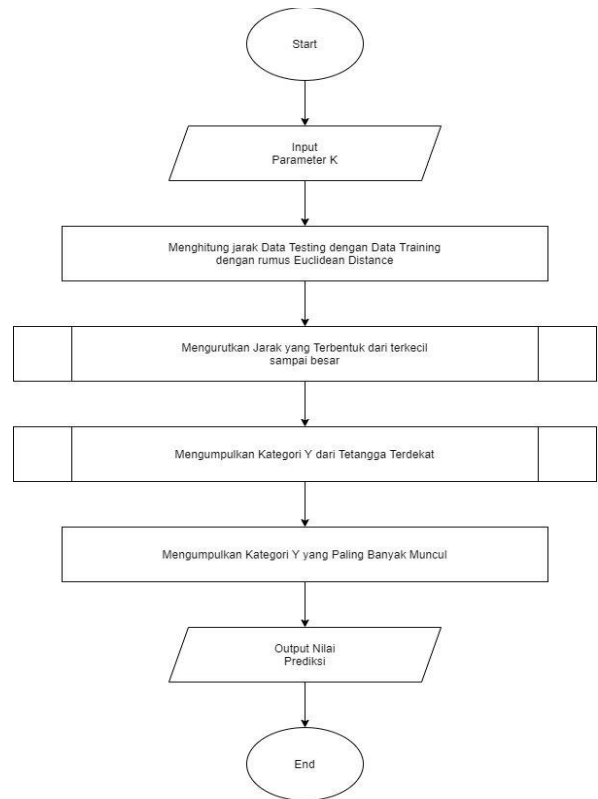
Keterangan:

- d_i = Jarak
- i = Variabel data
- p = Dimensi data
- $x1$ = Sampel data
- $x2$ = Data Uji

Algoritma KNN memiliki beberapa kelebihan diantaranya ketangguhan terhadap data training yang memiliki banyak noise dan data dalam jumlah yang besar. Kelemahan dari algoritma KNN adalah perlu menentukan jumlah tetangga terdekat dari target data, yang disimbolkan dengan nilai parameter K, data training yang didasarkan pada hasil perhitungan jarak kurang akurat karena harus memilih, mencoba dan menentukan jenis jarak yang digunakan dan atribut mana yang dipakai untuk mendapatkan hasil yang perhitungan jarak yang terbaik, serta biaya komputasi yang tinggi karena membutuhkan perhitungan jarak dari tiap *query instance* pada keseluruhan contoh data training.

Algoritma KNN merupakan salah satu metode *supervised learning* karena hasil *query instance* perlu diklasifikasikan berdasarkan mayoritas dari kategori KNN. Pengumpulan data menggunakan metode observasi. Data primer diperoleh dari bagian desa Batuphat dan Tambon Tunong. Data primer yang dijadikan contoh dalam penelitian ini adalah data longitude

dan latitude. Dalam penelitian ini menggunakan data training yang berjumlah 30 data. Dan data testing berjumlah 10 data. Untuk metode pengolahan data menggunakan metode KNN untuk klasifikasi. Algoritma KNN dapat dilihat pada Gambar 1 berikut ini:



Gambar 1. Alur Algoritma K-NN

Berikut penjelasan setiap tahapan dari algoritma KNN:

1. Menentukan jumlah tetangga terdekat yang disimbolkan dengan nilai parameter K. Nilai pada parameter K yang akurat untuk algoritma ini tergantung pada data training yang digunakan;
2. Menghitung dan menentukan nilai kuadrat jarak data testing objek terhadap data training yang diberikan dengan menggunakan persamaan 1;
3. Melakukan pengurutan dari hasil perhitungan no 2 secara ascending (urut dari nilai rendah ke nilai tinggi);
4. Mengumpulkan data pada kategori Y (Klasifikasi tetangga terdekat berdasarkan nilai K);
5. Kategori Y yang paling banyak muncul menjadi hasil akhir dari klasifikasi.

3. RESULT AND DISCUSSION

3.1 Pengolahan Data Awal

Tabel 1. Data Training

NO.	RUMAH	LATITUDE	LONGITUDE	LOKASI
1	BUDUY	11	26	batuphat
2	FARHAN	15	27	batuphat
3	AFIF	14	28	batuphat
4	IRGI	18	30	batuphat
5	ANDRE	17	26	batuphat
6	MAJID	12	28	batuphat
7	WAK ABBAS	13	27	batuphat
8	ANDRY	19	26	batuphat
9	CECEP	10	28	batuphat
10	UJANK	16	29	batuphat
11	ADI	21	25	tambon tunong
12	QIAL	25	22	tambon tunong
13	REZA	26	24	tambon tunong
14	ABIL	23	25	tambon tunong
15	IYAN	29	25	tambon tunong
16	SUKMA	22	20	tambon tunong
17	FAJAR	26	21	tambon tunong
18	NEPO	24	25	tambon tunong
19	ALDY	28	26	tambon tunong
20	ADAM	27	22	tambon tunong

Tabel 2. Sampel Data Testing

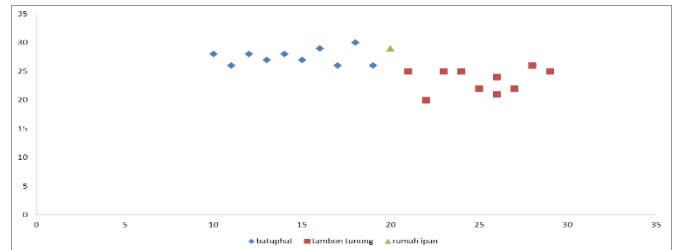
NO.	RUMAH	LATITUDE	LONGITUDE	LOKASI
1	IPAN	20	29	?
2	HARIS	19	26	?
3	RAIN	25	27	?

Data lokasi rumah yang terlihat seperti Tabel 1 adalah data mentah yang belum diolah dengan metode KNN. Data lokasi rumah Masyarakat Kecamatan Batuphat dan Gampong Tambon Tunong dijadikan sebagai data training sedangkan data lokasi salah satu masyarakat seperti yang terlihat pada Tabel 2 dijadikan sebagai data testing.

Tabel 3. Perhitungan Jarak

RANKING	EUCLIDIAN DISTANCE	RUMAH	LOKASI
1	2.236067977	IRGI	batuphat
2	3.16227766	ANDRY	batuphat
3	4	UJANK	batuphat
4	4.123105626	ADI	tambon tunong
5	4.242640687	ANDRE	batuphat
6	5	ABIL	tambon tunong
7	5.385164807	FARHAN	batuphat
8	5.656854249	NEPO	tambon tunong
9	6.08276253	AFIF	batuphat
10	7.280109889	WAK ABBAS	batuphat

Data lokasi rumah masyarakat Batuphat dan Tambon Tunong dijadikan sebagai data training sedangkan data lokasi Masyarakat dijadikan sebagai data testing. Proses perhitungan euclidean distance dilakukan dengan menggunakan rumus Euclidian seperti contoh perhitungan di tabel 3 dengan nilai K=10. Data testing yang digunakan sebagai contoh perhitungan adalah data dengan nomor 1, atas nama Ipan.



Gambar 2. Sampel pengujian atas nama Ipan

Perhitungan dari nilai *euclidean distance* selanjutnya diurutkan berdasarkan tetangga terdekat dari nilai K. Berikutnya dilakukan klasifikasi apakah termasuk mendapatkan beasiswa atau tidak berdasarkan nilai K. Dengan menggunakan data testing nomor 1 nama Ipan yang dikategorikan berlokasi di Batuphat.

4. CONCLUSION

Hasil dari penelitian ini menunjukkan bahwa pengujian dengan menggunakan data testing yang diukur terhadap longitude dan latitude yang berasal dari dua desa yang berbeda yaitu Batuphat dan Tambon Tunong dapat memberikan hasil yang baik yang berasal dari perhitungan jarak menggunakan Euclidean Distance. Namun untuk kedepannya mungkin diperlukan pengembangan lebih lanjut lagi terhadap penambahan atribut lain sehingga hasil klasifikasi lebih akurat.

REFERENCES

- [1] Anshori, L., Regasari, R., & Putri, M. (2018). Implementasi Metode K-Nearest Neighbor untuk Rekomendasi Keminatan Studi (Studi Kasus : Jurusan Teknik Informatika Univ Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya, 2(7), 2745–2753.
- [2] Dinata, R.K., Adek, R.T., Hasdyna, N., Retno, S. (2023). K-nearest neighbor classifier optimization using purity. AIP Conference Proceedings. 2431(1).
- [3] Hasdyna, N., Retno, S. (2022). Machine Learning Approach to Determine the Drug-Prone Areas in Lhokseumawe City, Indonesia. International Journal of Multidisciplinary Research and Analysis. 5(9): 2354-2464.
- [4] Hutami, R., & Astuti, E. Z. (2016). Implementasi Metode K-Nearest Neighbor Untuk Prediksi Penjualan Furniture Pada CV. Octo Agung Jepara. Universitas Dian Nuswantoro Semarang.
- [5] Mustakim, G. O. F. (2016). Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa, 13(2), 195–202.
- [6] Rahardja, C.A., Juardi, T., Agung, H. (2019). “Implementasi Algoritma K-Nearest Neighbor Pada Website Rekomendasi Laptop,” J. Buana Inform., vol. 10, no. 1, p. 75, 2019, doi: 10.24002/jbi.v10i1.184.
- [7] Retno, S., Dinata, R.K., Hasdyna, N. (2023). Evaluasi model data chatbot dalam natural language processing menggunakan k-nearest neighbor. Jurnal CoSciTech (Computer Science and Information Technology. 4(1): 146-153.

- [8] Retno, S., Rosnita, L., Anshari, S.F. (2023). Sistem Informasi Pelayanan Cuti Berbasis Web Pada PT Pupuk Iskandar Muda Menggunakan PHP dan MySQL. *TECHSI-Jurnal Teknik Informatika*, 14(1), 33-41.
- [9] Sesilia, N.R, Harsani, P. (2018) .“Penerapan K-Nearest Neighbor (KNN) untuk Klasifikasi Anggrek Berdasarkan Karakter Morfologi Daun dan Bunga,” vol. 15, no. 1, pp. 118–125.
- [10] Sumarlin, S. (2016). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM. *Jurnal Sistem Informasi Bisnis*, 5(1), 52–62. <https://doi.org/10.21456/vol5iss1pp52-62>
- [11] Suwirmayanti, N. L. G. P. (2017). Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil. *Techno. Com*, 16(2), 120–131.
- [12] Tang, Y., Jing, L., Li, H., & Atkinson, P. M. (2016). A multiple-point spatially weighted k-NN method for object-based classification. *International Journal of Applied Earth Observation and Geoinformation*, 52, 263–274. <https://doi.org/10.1016/j.jag.2016.06.017>
- [13] Tharwat, A., Mahdi, H., Elhoseny, M., & Hassanien, A. E. (2018). Recognizing human activity in mobile crowdsensing environment using optimized k-NN algorithm. *Expert Systems with Applications*, 107, 32–44. <https://doi.org/10.1016/j.eswa.2018.04.017>
- [14] Yahya, Y and Hidayanti, W.P. (2020). “Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada ‘Lombok Vape On,’” *Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 104–114, 2020, doi: 10.29408/jit.v3i2.2279.