



Klasifikasi penerima bantuan program rehabilitasi rumah tidak layak huni menggunakan algoritme K-Nearest Neighbor

Classification of beneficiaries for the rehabilitation of uninhabitable houses using the K-Nearest Neighbor algorithm

An Naas Shahifatun Na'iem, Harminto Mulyo, Nur Aeni Widiastuti*)

¹⁾Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Nahdlatul Ulama
Jl. Taman Siswa (Pekeng) Tahunan, Jepara 59427, Indonesia

Cara sitasi: A. S. Na'iem, H. Mulyo, and N. A. Widiastuti, "Klasifikasi penerima bantuan program rehabilitasi rumah tidak layak huni menggunakan algoritme K-Nearest Neighbor," *Jurnal Teknologi dan Sistem Komputer*, vol. 10, no. 1, pp. 32-37, 2022. doi: [10.14710/jtsiskom.2022.14110](https://doi.org/10.14710/jtsiskom.2022.14110), [Online].

Abstract – The registrars for rehabilitation programs for uninhabitable settlements are increasing every year. The large data processing of registrants may result in inaccuracies and need a long time to determine livable houses (RTLH) and unfit for habitation (non RTLH). This study aims to apply the K-Nearest Neighbor algorithm in classifying the eligibility of recipients of uninhabitable house rehabilitation assistance. The data used in this study were 1289 data with 13 attributes from the Jepara Regency Public Housing and Settlement Service. Data processing begins with attribute selection, categorization, outlier data cleaning, and data normalization and method application. The proposed system has the best classification at k of 5 with an accuracy of 97.93%, 96.88% precision, 99.53% recall, and an AUC value of 0.964.

Keywords – K-Nearest Neighbor; Euclidean distance; k -fold cross validation; house rehabilitation program

Abstrak – Jumlah pendaftar bantuan program rehabilitasi rumah tidak layak huni (RTLH) terus meningkat tiap tahun. Pengolahan data yang cukup besar mengakibatkan ketidaktepatan dan lamanya dalam menentukan pendaftar yang RTLH maupun Non RTLH. Kajian ini bertujuan untuk menerapkan algoritme K-Nearest Neighbor dalam klasifikasi kelayakan penerima bantuan rehabilitasi rumah tidak layak huni. Data yang digunakan pada penelitian ini sebanyak 1289 data dengan 13 atribut yang diperoleh dari Dinas Perumahan Rakyat dan Kawasan Pemukiman kabupaten Jepara. Pengolahan data dimulai dari seleksi atribut, pengkategorian data, pembersihan data outlier, normalisasi data dan penerapan metode. Sistem yang diusulkan menghasilkan akurasi sebesar 97,93%, presisi 96,88%, sensitivitas 99,53%, dan AUC sebesar 0,964 pada k bernilai 5.

Kata kunci - K-Nearest Neighbor; jarak Euclidean; validasi silang k -fold; program rehabilitasi rumah

I. PENDAHULUAN

Rumah tidak layak huni adalah rumah yang tidak memenuhi persyaratan keselamatan bangunan dari kecukupan minimal dalam unsur ruang dan luas ruangan. Selain itu, rumah juga perlu memperhatikan kualitas material pembangunan seperti dinding hingga memenuhi aspek pendidikan bagi siapa pun menempatnya [1], [2]. Dasar hukum mengenai pelaksanaan kegiatan rehabilitasi rumah tidak layak huni (RTLH) tertuang dalam UUD 1945 pasal 28H ayat (1) dan UU No. 1 Tahun 2011.

Dinas Perumahan Rakyat dan Kawasan Permukiman (Disperkim), terutama daerah Jepara, berupaya untuk memberikan bantuan sosial rehabilitasi rumah tidak layak huni kepada masyarakat miskin melalui program rehabilitasi rumah tidak layak huni. Program ini bertujuan untuk menyejahterakan masyarakat berpenghasilan rendah yang tempat tinggalnya tidak layak huni menjadi layak huni. Program rehabilitasi rumah tidak layak huni diberikan untuk keluarga miskin atau masyarakat berpenghasilan rendah, yang kurang mampu memenuhi kebutuhan melangsungkan hidup, dan tinggal di rumah yang tidak layak untuk di huni yang berdasarkan kriteria yang telah ditentukan oleh dinas.

Penerima bantuan rehabilitasi rumah tidak layak huni ditetapkan melalui beberapa tahapan seleksi dari keluarga miskin yang ada di desa melalui tata cara yang telah ditetapkan. Rumah dikatakan layak huni apabila memenuhi tiga kriteria, yaitu keselamatan bangunan, struktur bangunan, dan kesehatan bangunan.

Berdasarkan data yang diperoleh dari Disperkim Jepara, pendaftar yang mengajukan bantuan semakin meningkat tiap tahun terhitung dari tahun 2017 sampai dengan tahun 2019 seperti dinyatakan dalam Tabel 1. Tahap seleksinya sangat ketat. Masyarakat harus mengajukan proposal beserta berkas yang harus dipenuhi ke Disperkim dan data dimasukkan oleh petugas yang ditunjuk. Jika dinyatakan layak oleh petugas, berkas tersebut akan diberikan ke pihak

*)Penulis korespondensi (Nur Aeni Widiastuti)
Email: nuraeniwidiastuti@unisnu.ac.id

Tabel 1. Pendaftar bantuan rehabilitasi rumah

No	Tahun	Pendaftar	RTLH	Non RTLH
1	2017	1.190	500	690
2	2018	2.039	1.000	1.039
3	2019	2.932	1.000	1.932

lapangan untuk dilakukan survei lapangan. Jika data yang diberikan sesuai dengan data lapangan, maka pendaftar dinyatakan layak dan diajukan untuk pencairan dana. Pendaftar hanya diperbolehkan mendaftar sekali seumur hidup.

Saat ini proses penyeleksian bantuan RTLH di Disperkim Jepara masih dilakukan secara manual, yaitu dengan menyeleksi satu per satu berkas proposal yang sudah dikumpulkan oleh pendaftar. Berkas yang sudah terkumpul direkap dan ditentukan oleh petugas administrasi bidang perumahan untuk kategori RTLH yang bisa mendapatkan bantuan dan yang non RTLH yang masuk golongan layak huni. Proses tersebut tidak efisien dan memerlukan waktu yang lama untuk banyak pendaftar sehingga perlu dikembangkan sistem yang lebih baik, salah satunya sistem pendukung keputusan dengan klasifikasi menggunakan algoritme data mining [3].

Teknik klasifikasi telah banyak diterapkan dalam beragam aplikasi sistem pendukung keputusan [4]. Algoritme *k*-Nearest Neighbor (KNN) mampu memberikan hasil yang kompetitif dan signifikan untuk menyelesaikan masalah klasifikasi [5]. Klasifikasi rumah layak huni di kabupaten Brebes telah dilakukan Simatupang dan Wuryandari [6] dengan algoritme Learning Vector Quantization (LVQ) dan Naïve Bayes (NB). Atribut yang digunakan berjumlah tujuh dengan 838 data. Klasifikasi ini memperoleh ketepatan klasifikasi rumah layak huni 71,43% dengan LVQ dan 95,24% dengan NB.

Selain itu, jaringan syaraf tiruan backpropagation (JST-BP) juga telah diterapkan dalam memberikan bantuan program perbaikan rumah yang tidak layak huni agar tepat sasaran dengan hasil keputusan yang objektif. Putri dkk. [7] menggunakan 15 atribut dengan data sebanyak 160 data. Sistem klasifikasi tersebut menghasilkan dan akurasi sebesar 59% dan pola terbaik dengan layer masukan 15, layer tersembunyi 3, dan laju pembelajaran 0,2 pada *k* bernilai 10. Agustina dkk. [8] menerapkan Support Vector Machine (SVM) untuk mengatasi masalah tidak tepat sasaran dalam mengklasifikasi rumah layak huni dan rumah tidak layak huni. Data yang digunakan adalah sebanyak 160 data yang terbagi menjadi dua jenis, yaitu layak dan tidak layak, dan menggunakan 15 atribut. Kajian ini menghasilkan akurasi yang baik sebesar 95,75%. Lebih lanjut, Permatasari [9] menerapkan algoritme KNN untuk mengatasi masalah pendataan dan klasifikasi tingkat ekonomi warga desa Pacewetan menggunakan 250 data dan 12 atribut. Sistem tersebut dapat mengklasifikasikan status ekonomi keluarga serta didapatkan hasil pengujian dengan validasi silang 5-fold dengan akurasi 97,8%.

Meskipun terdapat berbagai teknik serta metode klasifikasi dalam menentukan kelayakan penerima bantuan rehabilitasi rumah tidak layak huni, namun tingkat akurasi dan penerapan metode untuk kasus secara khusus masih menarik untuk dibahas. Penelitian ini bertujuan menerapkan algoritme KNN untuk menentukan prediksi kelayakan penerima bantuan rehabilitasi rumah tidak layak huni. Algoritme KNN diterapkan untuk mendapatkan hasil akurasi yang tinggi seperti dalam [9], namun digunakan untuk melakukan klasifikasi dalam penyeleksian para pendaftar rehabilitasi rumah yang RTLH dan non RTLH seperti [6]. Kajian ini menggunakan data tahun 2019 dengan menggunakan perhitungan jarak matriks Euclidean dengan dilakukan 10 kali pengujian pada validasi silang *k-fold*. Evaluasi dan validasi hasil dilakukan berdasarkan matriks konfusi dan kurva *receiver operating characteristic* (ROC). Hasil kajian ini diharapkan dapat membantu Disperkim Jepara dalam menentukan klasifikasi kelayakan penerima bantuan RTLH secara tepat dan efisien.

II. METODE PENELITIAN

Desain penelitian ini menggunakan analisis kuantitatif yang menekankan analisis terhadap data yang berupa numerik sebagai alat penunjang dalam capaian penelitian. Tahapan penelitian yang dilakukan meliputi pengumpulan data, pengolahan data awal (prapemrosesan), pelatihan menggunakan KNN, pengujian model dengan validasi silang *k-fold*, dan analisis hasil klasifikasi. Pengumpulan data dilakukan dengan wawancara dan observasi dari Disperkim Jepara.

Data yang digunakan adalah data pendaftar yang mengajukan bantuan program rehabilitasi RTLH pada tahun 2019 sejumlah 2962 dengan masukan 22 atribut. Dataset yang diperoleh kemudian dilakukan pengolahan data awal. Pengolahan data awal menghasilkan 1289 data yang digunakan sebagai data latih dan data uji. Pengujian model dilakukan menggunakan validasi silang *k-fold* pada data uji dan KNN pada data latih. Hasil klasifikasi terhadap data uji dievaluasi berdasarkan matriks konfusi.

A. Prapemrosesan data

Prapemrosesan diperlukan agar data yang diolah memiliki kualitas yang baik. Hal ini dilakukan karena tidak semua data dan atribut dapat digunakan. Tahapan pengolahan awal data yang dilakukan dalam penelitian ini meliputi seleksi atribut, pengkategorian data, pembersihan data *outlier*, dan normalisasi data.

Proses seleksi atribut dilakukan untuk mengambil atribut tertentu dari semua atribut pada data awal yang berpengaruh pada proses perhitungan. Dari 22 atribut yang ditentukan oleh Disperkim, atribut diseleksi menjadi 13 atribut, yaitu nama, kepemilikan rumah dan tanah, ukuran rumah, jumlah anggota keluarga, pondasi, balok, sanitasi, jendela, ventilasi, material atap, material lantai, material dinding dan hasil dengan keluaran RTLH atau non RTLH sesuai Tabel 2. Atribut

ini mewakili variabel yang berpengaruh dan ditentukan oleh Disperkim, yaitu keselamatan (meliputi pondasi, dinding, balok, dan atap), struktur bangunan (luas m²/jiwa, material atap, lantai, dan dinding), dan kesehatan (pencahayaannya dan ventilasi).

Kategorisasi data dilakukan untuk menyusun dan mengategorikan data, yaitu masing-masing atribut, kecuali nama dan hasil, diberi label A1 sampai A11 dan untuk kategorisasi atribut. Pembersihan data *outlier* dilakukan untuk mengidentifikasi dan menghilangkan data yang kosong (*missing value*) dan penghapusan data *outlier* terhadap data penelitian sampai data benar-benar bersih. Pembersihan data *outlier* ini berpengaruh terhadap hasil akurasi yang tinggi seperti dalam [10]. Pada penelitian ini, pembersihan data *outlier* menggunakan SPSS dan memperoleh hasil perhitungan dari jumlah data awal 2962 menjadi 1289.

Normalisasi data dilakukan untuk memperkecil jarak pada data yang menggunakan normalisasi Min-max sehingga dapat menghasilkan keseimbangan nilai antara 0 sampai 1 [11]. Tahapan ini menggunakan perangkat RapidMiner dengan metode transformasi jangkauan. Penghitungan normalisasi Min-max dinyatakan dengan (1). Parameter v' menyatakan nilai dari data baru hasil dari normalisasi, v nilai dari data sebelum dinormalisasi, A'_{max} batas nilai maksimum terbaru bernilai 1, A'_{min} batas nilai minimum terbaru bernilai 0, A_{max} nilai maksimum pada kolom, dan A_{min} menyatakan nilai minimum pada kolom data.

$$v' = \frac{v - A_{min}}{A_{max} - A_{min}} (A'_{max} - A'_{min}) + A'_{min} \quad (1)$$

B. Klasifikasi data

Algoritme klasifikasi yang digunakan adalah KNN. KNN ini dapat melakukan klasifikasi berdasarkan jarak terdekat dari data latih dengan data uji dan dapat menghasilkan nilai akurasi yang tinggi [4]. Kajian ini menggunakan data numerik berdasarkan kategori tiap-tiap parameter. Untuk mencari jarak antara dua titik, yaitu titik pada data latih dan titik pada data uji digunakan matriks jarak Euclidean yang dinyatakan pada (2). Parameter $d(P, Q)$ menyatakan jarak antara data pada titik P dan Q dimana P adalah masukan data pertama dari data latih dan Q adalah masukan data ke-1 dari data uji, dan n adalah jumlah data latih.

$$d(P, Q) = \sqrt{\sum_{i=1}^n (P_i - Q_i)^2} \quad (2)$$

Penelitian ini menggunakan matriks jarak Euclidean sebagai matriks yang memiliki kinerja baik seperti yang dinyatakan dalam [12], [13]. Hal ini dilakukan untuk menunjukkan perbandingan hasil kinerja matriks jarak pada metode yang diusulkan. Setelah jarak dari data yang baru terhadap seluruh data pembelajaran dihitung dan sejumlah k buah yang paling dekat diambil, klasifikasi ditentukan dari titi-titik tersebut [14].

Tabel 2. Kriteria variabel

No	Atribut	Label	Tipe	Deskripsi
1.	Nama		String	Nama Pendaftar
2.	Kepemilikan rumah dan tanah	A1	Nominal	Milik sendiri =1, Milik saudara=2, Bengkok=3, Tanah irigasi=4, Sewa=5
3.	Jumlah anggota keluarga	A2	Numerik	Jumlah penghuni rumah
4.	Ukuran rumah	A3	Nominal	>100m ² = 1, 71m ² -100m ² = 2, 41m ² -70m ² = 3, 21m ² -40m ² = 4, 9m ² -20m ² = 5
5.	Pondasi	A4	Nominal	Layak=1, Ada=2, Tidak Layak= 3, Tidak Ada= 4
6.	Balok	A5	Nominal	Layak=1, Ada=2, Tidak Layak= 3, Tidak Ada= 4
7.	Sanitasi	A6	Nominal	Layak=1, Ada=2, Tidak Layak= 3, Tidak Ada= 4
8.	Jendela	A7	Nominal	Layak=1, Ada=2, Tidak Layak= 3, Tidak Ada= 4
9.	Ventilasi	A8	Nominal	Layak=1, Ada=2, Tidak Layak= 3, Tidak Ada= 4
10.	Material Atap	A9	Nominal	Genteng=1, Beton=2, Asbes=3, Seng=4, Rumbia=5
11.	Material Lantai	A10	Nominal	Keramik=1, Plester=2, Ubin=3, PanggungKayu=4, Tanah=5
12.	Material Dinding	A11	Nominal	Bata Plester=1, Bata Expose=2, Herbel=3, Kalsiboard=4, Asbes=5, Kayu=6, Triplek=7, Bambu=8
13.	Hasil		Nominal	1 = Non RTLH 2 = RTLH

Data diolah menggunakan KNN dengan aplikasi RapidMiner dan menghasilkan model yang digunakan dalam klasifikasi dan pengujian. Ukuran kinerja yang dikaji adalah akurasi, presisi, sensitivitas, dan *area under curve* (AUC). Pengujian pada penelitian ini menggunakan validasi silang *10-fold* dan melakukan percobaan sebanyak 10 kali. Nilai k terbaik diaplikasikan ke dalam metode yang diusulkan.

III. HASIL DAN PEMBAHASAN

Penelitian ini menerapkan KNN untuk klasifikasi kelayakan penerima bantuan rehabilitasi RTLH dengan tahapan meliputi pengumpulan data, pengolahan data

Tabel 3. Data RTLH dan Non RTLH setelah kategorisasi data

No	Nama	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Hasil
1	Abdul alim	1	2	4	3	3	2	2	2	2	5	5	RTLH
2	Siti Khasanah	1	1	4	1	1	4	2	2	2	3	2	RTLH
3	Tumijan	1	2	4	3	1	4	2	2	2	3	7	RTLH
4	Hendrik santoso	1	4	2	1	1	1	2	2	2	3	2	Non RTLH
5	Wakini	1	2	4	1	1	2	2	2	2	5	2	RTLH
6	Jumariyah	1	2	4	1	1	2	2	2	2	3	1	Non RTLH
7	Kemijah	1	1	4	3	3	2	2	4	2	5	5	RTLH
8	Masrurrotun	1	3	4	3	3	2	2	4	2	3	5	RTLH
...
1289	Munawaroh	1	4	2	1	1	2	2	2	2	2	2	Non RTLH
	Nilai data minimum	1	1	2	1	1	1	2	2	2	1	1	
	Nilai data maksimum	1	7	5	3	3	4	2	4	2	5	8	

Tabel 4. Data RTLH dan Non RTLH yang telah dinormalisasi

No	Nama	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Hasil
1	Abdul alim	1,000	0,167	0,667	0,667	0,667	0,333	1,000	0,000	1,000	1,000	0,571	RTLH
2	Siti Khasanah	1,000	0,000	0,667	0,000	0,000	1,000	1,000	0,000	1,000	0,500	0,143	RTLH
3	Tumijan	1,000	0,167	0,667	0,667	0,000	1,000	1,000	0,000	1,000	0,500	0,857	RTLH
4	Hendrik santoso	1,000	0,500	0,000	0,000	0,000	0,000	1,000	0,000	1,000	0,500	0,143	Non RTLH
5	Wakini	1,000	0,167	0,667	0,000	0,000	0,333	1,000	0,000	1,000	1,000	0,143	RTLH
6	Jumariyah	1,000	0,167	0,667	0,000	0,000	0,333	1,000	0,000	1,000	0,500	0,000	Non RTLH
7	Kemijah	1,000	0,000	0,667	0,667	0,667	0,333	1,000	1,000	1,000	1,000	0,571	RTLH
8	Masrurrotun	1,000	0,333	0,667	0,667	0,667	0,333	1,000	1,000	1,000	0,500	0,571	RTLH
...
1289	Munawaroh	1,000	0,500	0,000	0,000	0,000	0,333	1,000	0,000	1,000	0,250	0,143	Non RTLH

awal, pengujian model dengan validasi silang *k-fold*, tahap klasifikasi KNN, serta evaluasi dan validasi.

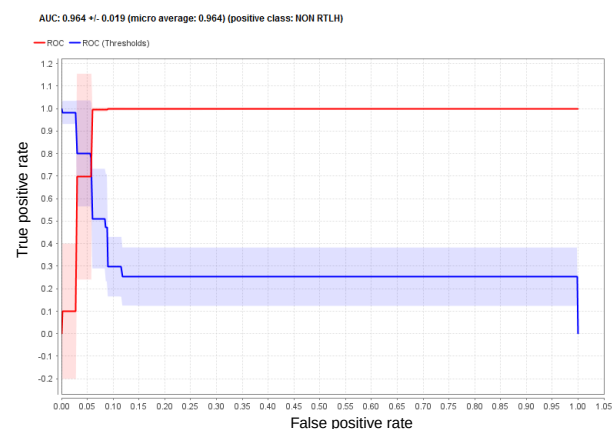
Tahap pengumpulan data dilakukan dengan menggunakan metode wawancara dan observasi diperoleh data pendaftar yang mengajukan bantuan program rehabilitasi rumah tidak layak huni (RTLH) pada tahun 2019 sejumlah 2962. Data diolah terlebih dahulu menggunakan Microsoft Excel, yaitu dilakukan seleksi atribut yang semula 22 atribut menjadi 13 atribut (Tabel 2). Dalam pengkategorian data, data yang semula berupa string diubah menjadi numerik. Data tersebut kemudian dilakukan pembersihan dataset dari *outlier* menjadi sejumlah 1289. Hasil kategorisasi dan pembersihan data dinyatakan dalam Tabel 3.

Data hasil prapemrosesan dinormalisasi dengan membuat data yang sudah ada menjadi nilai yang lebih kecil, yaitu dalam jangkauan 0 sampai 1, sehingga mengoptimalkan proses perhitungan. Nilai minimal dari data menjadi 0 dan nilai maksimal dari data menjadi 1, serta data di antara nilai minimal dan nilai maksimal menyesuaikan dalam jangkauan tersebut. Hasil normalisasi data dinyatakan dalam Tabel 4.

Pengujian untuk mendapatkan model KNN terbaik dilakukan validasi silang *10-fold* dengan menggunakan RapidMiner. Pengujian dilakukan sebanyak 10 kali pada setiap nilai *k* ganjil. Hasil pengujian model memperoleh hasil tertinggi ketika *k* bernilai 5, yaitu sebesar 97,93% seperti yang ditunjukkan pada Tabel 5. Presisi, sensitivitas, dan nilai AUC yang diperoleh berturut-turut 96,88%, 99,53%, dan 0,964. Representasi AUC dari KNN dinyatakan dalam Gambar 1. Nilai AUC ini termasuk kategori sangat baik sehingga bahwa model

Tabel 5. Akurasi KNN berdasarkan nilai *k*

No.	Nilai k	Akurasi (%)	No.	Nilai k	Akurasi (%)
1.	1	97,28	6.	11	97,02
2.	3	97,67	7.	13	96,76
3.	5	97,93	8.	15	96,76
4.	7	97,28	9.	17	96,76
5.	9	97,76	10.	19	96,89



Gambar 1. Representasi AUC dari KNN

KNN ini akurat untuk klasifikasi kelayakan penerima bantuan program rehabilitasi RTLH dengan 13 atribut seperti halnya [9] yang digunakan untuk klasifikasi status ekonomi keluarga.

Dengan model KNN terbaik, jarak masing-masing objek pada data latih terhadap data uji dihitung

Tabel 6. Evaluasi KNN

Kelas Prediksi	Kelas Aktual		Presisi (%)
	RTLH	Non RTLH	
RTLH	327	3	93,39
Non RTLH	14	429	96,84
Sensitivitas (%)	95,89	99,54	97,80

menggunakan jarak Euclidean (2). Kelas mayoritas ditentukan dan hasil evaluasi dinyatakan dalam matriks konfusi dalam Tabel 6. Dari jumlah data 773 kepala keluarga, 327 dikelompokkan ke RTLH, 3 keluarga Non RTLH namun diklasifikasikan ke RTLH, 14 keluarga RTLH namun diklasifikasikan ke non RTLH, dan 429 diklasifikasikan non RTLH. Akurasi yang diperoleh adalah sebesar 97,80% dalam menentukan kelayakan penerima program bantuan rehabilitasi RTLH. Hasil ini menunjukkan bahwa KNN dapat mengenali data numerik, data string, data campuran, dan efektif terhadap jumlah data yang besar. Selain itu, matriks jarak Euclidean memiliki kinerja yang baik seperti dalam [12]-[15].

Dengan perbedaan dataset, jumlah atribut dan metode yang digunakan dalam klasifikasi rumah layak huni, hasil ini lebih baik dari [6] yang menggunakan LVQ dengan akurasi 71,43% dan NB dengan 95,24%, dari [7] yang menggunakan JST-BP dengan akurasi 59%, dan dari [8] yang menggunakan SVM dengan akurasi 95,75%. Hasil ini menunjukkan bahwa KNN paling akurat untuk klasifikasi rumah layak huni seperti klasifikasi status ekonomi keluarga dalam [9], namun bertentangan dengan [16] yang menyatakan bahwa JST dan NB lebih baik dari KNN dalam klasifikasi pendonor darah potensial.

IV. KESIMPULAN

KNN dapat diterapkan untuk klasifikasi kelayakan penerima rehabilitasi RTLH dengan akurasi tertinggi sebesar 97,93% dengan 13 atribut data sebagai variabel yang berpengaruh dan telah ditentukan oleh Disperkim, yaitu aspek keselamatan (meliputi pondasi, dinding, balok dan atap), struktur bangunan (luas rumah, material atap, lantai, dan dinding), dan kesehatan (pencahayaan dan ventilasi).

UCAPAN TERIMA KASIH

Terima kasih ditujukan kepada Kepala Bidang Perumahan Dinas Perumahan dan Kawasan Permukiman Rakyat (Disperkim) Jepara yang telah mengizinkan pengambilan data untuk penelitian ini.

DAFTAR PUSTAKA

- [1] M. Ismowati and A. Subhan, "Implementasi program pemberdayaan sosial kegiatan rehabilitasi sosial rumah tidak layak huni (RTLH) bagi masyarakat miskin di kabupaten Pandeglang," *Transparansi Jurnal Ilmiah Ilmu Administrasi*, vol. 1, no. 2, pp. 194–205, 2018. doi: [10.31334/trans.v1i2.384](https://doi.org/10.31334/trans.v1i2.384)
- [2] M. A. Nurhakim and E. Pandamdari, "Pemenuhan atas sarana dan utilitas pada perumahan subsidi Mutiara Puri Harmoni Rajeg Tangerang menurut undang-undang nomor 1 tahun 2011 tentang perumahan dan permukiman," *Jurnal Hukum Adigama*, vol. 1, no. 1, pp. 437-461, 2018. doi: [10.24912/adigama.v1i1.2152](https://doi.org/10.24912/adigama.v1i1.2152)
- [3] T. Sugihartono, "Implementasi sistem pendukung keputusan penerima bantuan rumah tidak layak huni berbasis web," *Jurnal Sistem Informasi dan Komputer*, vol. 7, no. 1, pp. 52-56, Mar. 2018. doi: [10.32736/sisfokom.v7i1.299](https://doi.org/10.32736/sisfokom.v7i1.299)
- [4] J. Han, M. Kamber, and J. Pei, *Data Mining: Concept and Techniques*. San Francisco: Elsevier, 2012, pp. 1–38.
- [5] D. A. Adeniyi, Z. Wei, and Y. Yongquan, "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method," *Applied Computing and Informatics*, vol. 12, no. 1, pp. 90–108, 2016. doi: [10.1016/j.aci.2014.10.001](https://doi.org/10.1016/j.aci.2014.10.001)
- [6] F. J. Simatupang and T. Wuryandari, "Klasifikasi rumah layak huni di kabupaten Brebes dengan menggunakan metode learning vector quantization dan naive bayes," *Jurnal Gaussian*, vol. 5, no. 1, pp. 99–111, 2016.
- [7] R. R. P. Putri, M. T. Furqon, and B. Rahayudi, "Implementasi metode JST-Backpropagation untuk klasifikasi rumah layak huni (Studi kasus desa Kidul kecamatan Tumpang kabupaten Malang)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, pp. 3360–3365, 2018.
- [8] B. Agustina, W., Furqon, M., & Rahayudi, "Implementasi metode support vector machine (SVM) untuk klasifikasi rumah layak huni (Studi kasus desa Kidul kecamatan Tumpang kabupaten Malang)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, pp. 3366–3372, 2018.
- [9] D. I. Permatasari, "Klasifikasi status ekonomi keluarga dengan menggunakan algoritma k-nearest neighbor di desa Pacewetan kecamatan Pace kabupaten Nganjuk," *Simki-Techsain*, vol. 1, no. 1, pp. 1-10, 2017.
- [10] S. Sugriyono and M. U. Siregar, "Preprocessing kNN algorithm classification using K-means and distance matrix with students' academic performance dataset," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 4, pp. 311–316, 2020. doi: [10.14710/jtsiskom.2020.13874](https://doi.org/10.14710/jtsiskom.2020.13874)
- [11] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma K-NN," *Computer Engineering, Science and System Journal*, vol. 4, no. 1, pp. 78-82, 2019. doi: [10.24114/cess.v4i1.11458](https://doi.org/10.24114/cess.v4i1.11458)
- [12] S. Kapil and M. Chawla, "Performance evaluation of K-means clustering algorithm with various distance metrics," in *International Conference on*

- Power Electronics, Intelligent Control and Energy Systems*, Delhi, India, Jul. 2016, pp. 1–4. doi: [10.1109/ICPEICES.2016.7853264](https://doi.org/10.1109/ICPEICES.2016.7853264)
- [13] S. A. Salihu, I. P. Onyekwere, M. A. Mabayoje, and H. A. Mojeed, “Performance evaluation of manhattan and euclidean distance measures for clustering based automatic text summarization,” *FUOYE Journal of Engineering and Technology*, vol. 4, no. 1, pp. 135-139, 2019. doi: [10.46792/fuoyejt.v4i1.316](https://doi.org/10.46792/fuoyejt.v4i1.316)
- [14] R. Feldman and J. Sanger, *The text mining handbook: advanced approaches in analyzing unstructured data*. England: Cambridge Univ. Press, 2007.
- [15] E. Nasri and A. Selamat, “Aplikasi seleksi penentuan nasabah untuk penjualan barang secara kredit dengan algoritma k-nearest neighbour,” *Jurnal Ilmiah Sains Dan Teknologi*, vol. 4, no. 1, pp. 1–11, 2020.
- [16] M. Lestandy, L. Syafa’ah, and A. Faruq, “Classification of potential blood donors using machine learning algorithms approach,” *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 3, pp. 217–221, 2020. doi: [10.14710/jtsiskom.2020.13619](https://doi.org/10.14710/jtsiskom.2020.13619)



©2022. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).