

Analisis klasifikasi Algoritma *K-Nearest Neighbor* (K-NN) pada struktur Daerah di Kota Medan

Classification analysis of K-NN (K-Nearest Neighbor) Algorithm on Regional structures in Medan City

Safa Nadia Bakri*¹, Lailan Sofinah Harahap²

¹Sistem Informasi, Universitas Islam Negeri Sumatera Utara

²Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara

E-mail: ¹safanadia278@gmail.com

E-mail: ²lailansofinahharahap@gmail.com

Abstrak

Penelitian ini bertujuan untuk memeriksa penerapan dari algoritma (K-NN) *K-Nearest Neighbor* dalam mengklasifikasikan struktur daerah pada Kota Medan yang merupakan salah satu kota terbesar di Indonesia, mempunyai beragam karakteristik struktur wilayah yang membutuhkan pendekatan analisis yang tepat untuk pengelolaan dan perencanaan tata ruang. Algoritma K-NN dipilih karena kemampuannya dalam mengelompokkan data berdasarkan kedekatannya dengan titik data lain, yang sangat sesuai dengan kebutuhan analisis klasifikasi wilayah. Dalam penelitian ini, data yang digunakan mencakup berbagai atribut struktur daerah seperti kepadatan penduduk, penggunaan lahan, serta infrastruktur yang terdapat pada setiap kecamatan di Kota Medan. Metode yang digunakan dalam penelitian ini ialah pengolahan data statistik untuk mengelompokkan wilayah dengan karakteristik serupa, menggunakan algoritma K-NN sebagai metode klasifikasi. Proses pengklasifikasian ini melibatkan pemilihan parameter yang tepat, penghitungan jarak antar titik data, serta memutuskan jumlah tetangga yang ideal. Hasil yang diharapkan dari analisis ini akan memberikan gambaran yang jelas mengenai pola distribusi struktur wilayah di Kota Medan, serta membantu dalam perencanaan dan pengembangan kota yang lebih efisien dan terarah. Keakuratan model K-NN dalam mengklasifikasikan wilayah juga akan dibandingkan dengan algoritma lainnya untuk menilai efektivitas dan keandalannya dalam konteks studi ini. Penelitian ini menyimpulkan bahwasanya algoritma *K-Nearest Neighbor* (K-NN) dapat menjadi alat yang efisien dan efektif untuk klasifikasi struktur daerah di Kota Medan, memberikan informasi yang berharga bagi pengambil keputusan dalam perencanaan kota.

Kata kunci: Algoritma *K-Nearest Neighbor* (KNN), Struktur Daerah, Klasifikasi, Kota Medan, Perencanaan, Tata Ruang.

Abstract

This research aims to analyze the application of the K-Nearest Neighbor (K-NN) algorithm in classifying regional structures in Medan City. Medan, one of the largest cities in Indonesia, has a variety of characteristics of regional structure that requires an appropriate analysis approach for spatial management and spatial planning. The K-NN algorithm was chosen because of its ability to categorize data based on its proximity to other data points, which is

very suitable for the needs of spatial planning and management. other data points, which is very suitable for the needs of regional classification analysis. In this research, the data used includes various attributes of the regional structure such as structure attributes such as population density, land use, and infrastructure in each sub-district. infrastructure in each sub-district in Medan City. In this research, the method used is statistical data processing statistical data processing to group areas with similar characteristics, applying the K-NN algorithm to categorize method. The classification process process involves selecting the right parameters, calculating the distance between data points, and selecting the optimal number of nearest neighbors. data points, as well as deciding on the ideal amount of neighbors. The expected results The expected results of this analysis will provide a clear picture of the distribution pattern of the pattern of distribution of the regional structure in Medan City, as well as assisting in the planning and development of a more efficient and directed city. The accuracy of the K-NN model in classifying the regions will also be compared with other algorithms to assess its effectiveness and reliability in the context of this study.

Keywords: *KNN Algorithm, Regional Structure, Classification, Medan City, Spatial Planning.*

1. PENDAHULUAN

Kota Medan, Sumatera Utara, ialah salah satu kota terbesar yang berada di Indonesia memiliki beragam karakteristik dan dinamika dalam struktur wilayahnya. Keberagaman ini mencakup berbagai aspek, seperti kepadatan penduduk, penggunaan lahan, jenis infrastruktur, hingga status sosial ekonomi penduduk. Mengingat pentingnya peran struktur wilayah dalam perencanaan dan pengembangan kota, maka analisis terhadap struktur daerah di Kota Medan menjadi suatu kebutuhan yang tak terhindarkan. Penelitian ini bertujuan untuk menggunakan algoritma *K-Nearest Neighbor* (K-NN) guna menganalisis dan mengklasifikasikan struktur daerah di Kota Medan. Pemilihan algoritma K-NN didasarkan pada kemampuannya dalam mengklasifikasikan data menurut jarak antar titik data yang memiliki karakteristik serupa, menjadikannya alat yang efektif dalam mengidentifikasi pola distribusi wilayah. Hal ini sangat relevan dengan perencanaan tata ruang kota yang membutuhkan pembagian wilayah berdasarkan karakteristik tertentu, seperti tingkat kepadatan, fasilitas publik, dan pemanfaatan lahan [1].

Penggunaan algoritma K-NN dalam pengklasifikasian data bukanlah hal baru dalam dunia penelitian. Algoritma ini sudah banyak digunakan dalam berbagai industri, seperti pengolahan gambar, analisis sentimen, bahkan juga dalam bidang kesehatan dan pertanian [2]. Dalam konteks ini, K-NN digunakan untuk mengelompokkan wilayah-wilayah di Kota Medan ke dalam kategori-kategori tertentu berdasarkan data yang ada, seperti kepadatan penduduk, penggunaan lahan, dan infrastruktur. Melalui analisis ini, diharapkan dapat ditemukan pola-pola struktural yang dapat membantu dalam perencanaan pembangunan dan pengelolaan kota yang lebih efisien. Penelitian sebelumnya telah menunjukkan keberhasilan K-NN dalam

mengklasifikasikan berbagai data dengan tingkat akurasi yang memadai, menjadikannya pilihan yang layak untuk diterapkan pada studi ini [3].

Di dalam kajian ini, data yang digunakan mencakup informasi mengenai struktur daerah di Kota Medan yang diperoleh dari berbagai sumber, seperti data sensus penduduk, data penggunaan lahan, serta data mengenai fasilitas dan infrastruktur kota. Proses pengklasifikasian dilakukan dengan mengolah data tersebut untuk menemukan hubungan antar atribut-atribut yang ada. Salah satu tantangan dalam penelitian ini adalah pemilihan fitur yang tepat untuk dimasukkan dalam model K-NN, serta pengaturan parameter-parameter algoritma agar memperoleh hasil yang optimal. Pemilihan parameter yang tepat akan sangat mempengaruhi hasil klasifikasi yang diperoleh, maka dalam penelitian ini akan membahas tentang optimasi parameter K-NN untuk mencapai akurasi yang lebih tinggi dalam klasifikasi struktur daerah di Kota Medan [4].

Dalam penelitian ini, K-NN digunakan sebagai metode klasifikasi karena algoritma ini didasarkan pada konsep kedekatan atau jarak antara dua data lainnya. Adanya prinsip ini membuat K-NN dapat mengelompokkan data berdasarkan kedekatan antara titik data dari satu titik data ke titik data lainnya didalam ruang fitur. Hal ini sangat relevan dengan tujuan penelitian ini, yaitu untuk mengklasifikasikan struktur daerah di Kota Medan yang memiliki karakteristik serupa. Dalam pelaksanaannya, K-NN mengandalkan dua komponen penting, yaitu pemilihan menghitung jumlah tetangga berikutnya (k) dan jarak antar data yang akan dikelompokkan. Keberhasilan K-NN dalam klasifikasi sangat bergantung pada pemilihan parameter ini, yang akan dioptimalkan selama penelitian untuk mendapatkan hasil yang terbaik [5].

Menurut penelitian terbaru, algoritma *K-Nearest Neighbor* (KNN) sangat mampu mengklasifikasikan data geografis dan administratif. Seperti menurut [6]. dalam penelitian ini berhasil menerapkan K-NN untuk memprediksi angka kelahiran di Kota Medan dengan akurasi 83,9%. Ini menunjukkan bahwa algoritma ini cocok untuk menangani data demografi lokal perkotaan. Selain itu, [7] membandingkan metode K-NN dan Naïve Bayes untuk klasifikasi Indeks Pembangunan Manusia (IPM) di berbagai kabupaten dan kota di Indonesia, dengan K-NN tetap memberikan hasil klasifikasi yang kompetitif. Pada penelitian lainnya oleh [8] menggunakan K-NN dan SVM untuk mengidentifikasi wilayah tertinggal di Indonesia, dan menemukan bahwa K-NN cukup efektif untuk mengidentifikasi karakteristik wilayah dengan keterbatasan pembangunan.

Dalam upaya untuk mencapai tujuan penelitian ini, penting untuk mempertimbangkan kualitas data yang digunakan. Data yang akurat dan representatif akan memberikan hasil yang lebih baik dalam analisis klasifikasi. Oleh karena itu, dalam penelitian ini, pengumpulan dan pemrosesan data menjadi langkah awal yang sangat krusial. Data yang akan digunakan mencakup berbagai aspek struktural wilayah Kota Medan, seperti data tentang kepadatan penduduk, jenis penggunaan lahan, serta distribusi infrastruktur kota. Pengolahan data yang

baik akan memungkinkan pemodelan yang lebih akurat, sehingga klasifikasi wilayah dapat dilakukan dengan lebih efisien.

Dalam dunia yang semakin terhubung dengan teknologi, analisis data dan penerapan algoritma dalam pengambilan keputusan menjadi hal yang sangat signifikan. Akibatnya, penelitian ini bertujuan untuk menunjukkan bagaimana algoritma K-NN dapat digunakan dalam mengklasifikasikan struktur daerah di Kota Medan, sehingga hasilnya dapat digunakan sebagai dasar dalam pengembangan perencanaan kota yang lebih baik. Dengan menggunakan pendekatan tersebut, diharapkan bisa ditemukan pola-pola yang tersembunyi dalam data struktur wilayah yang bisa memberikan wawasan baru dalam upaya perencanaan tata ruang dan pengelolaan kota. Selain itu, tujuan penelitian ini adalah untuk membantu perkembangan ilmu komputer, khususnya dalam penerapan algoritma K-NN di bidang analisis struktur wilayah [9].

Keberhasilan penelitian ini sangat bergantung pada bagaimana data dikumpulkan, diproses, dan dianalisis menggunakan algoritma K-NN. Oleh sebab itu, fokus penelitian ini tidak hanya pada penerapan algoritma K-NN saja, tetapi juga pada optimasi dan evaluasi hasil klasifikasi yang diperoleh. Selain itu, akan dilakukan pula perbandingan dengan algoritma lain yang mungkin lebih efektif dalam konteks ini, seperti algoritma SVM (*Support Vector Machine*) or *Decision Tree*. Oleh karena itu, diharapkan penelitian ini akan memberikan kontribusi yang signifikan untuk pengembangan metode klasifikasi untuk analisis struktur daerah di Kota Medan, serta membuka peluang untuk aplikasi lebih lanjut dalam perencanaan dan pengembangan kota [10].

2. METODOLOGI PENELITIAN

Penelitian ini menerapkan analisis klasifikasi dengan memakai algoritma K-NN (*K-Nearest Neighbor*) pada struktur daerah di Kota Medan didasarkan pada pendekatan data mining yang melibatkan pengumpulan, pengolahan, dan analisis data menggunakan metode klasifikasi. Penelitian ini mengadopsi prinsip-prinsip dasar dalam algoritma K-NN yang telah terbukti efektif dalam berbagai studi sebelumnya, baik dalam bidang pengolahan citra, analisis sentimen, maupun klasifikasi data dalam berbagai konteks [11]. Dalam penelitian ini data yang dipakai terdiri dari atribut-atribut ini menggambarkan struktur wilayah Kota Medan, seperti kepadatan penduduk, penggunaan lahan, dan infrastruktur. Tahap pertama yang dilakukan pada penelitian ini ialah mengumpulkan data yang relevan, yang kemudian akan dianalisis untuk menemukan pola-pola yang ada dalam struktur daerah tersebut.

Menurut [12], Metode Algoritma supervisi digunakan untuk K-NN (*K-Nearest Neighbor*), yang mengklasifikasikan hasil sampel uji baru yang didasarkan pada mayoritas kategori pada K-NN. Algoritma ini bertujuan untuk mengkategorikan barang baru berdasarkan fitur dan sampel latihan. Untuk cocok, pengklasifikasian hanya menggunakan memori dan tidak menggunakan model apa pun. Diberikan titik uji, sejumlah K objek (titik pelatihan) akan ditemukan paling dekat. Klasifikasi menggunakan suara terbanyak di antara objek K yang diklasifikasikan. Algoritma K-

NN menggunakan klasifikasi ketetanggaannya untuk memprediksi nilai sampel uji baru. Banyak orang menggunakan jarak *Euclidean* untuk mengetahui seberapa dekat atau jauh tetangga mereka.

Algoritma K-NN dapat menghitung jarak antara data berdimensi q . Nilai jarak ini digunakan untuk menentukan seberapa dekat atau mirip data uji dengan data latih. Nilai K pada K-NN menunjukkan K data uji yang paling dekat. Beberapa sifat K-NN adalah sebagai berikut:

- a. K-NN adalah algoritma yang melakukan proses klasifikasi dengan menggunakan seluruh data latihan. Hal ini menyebabkan prediksi yang sangat lama untuk data yang sangat besar. Pendekatan tambahan melibatkan penggunaan meandata dari masing-masing kelas, kemudian menghitung jarak terdekat antara data uji dan meandata setia kelas tersebut. Ini menghasilkan kerja yang lebih cepat, tetapi karena model hanya membentuk *hyperplane linear* di tengah-tengah antara dua kelas yang terpisah, hasilnya kurang memuaskan.
- b. Seperti *Artificial Neural Network* (ANN), algoritma (KNN) tidak menilai semua fitur dengan berat. Sebaliknya, algoritma K-NN tidak menilai setiap fitur dengan bobot.
- c. (KNN) termasuk dalam kategori belajar malas, yang menyimpan sebagian atau semua data dan hampir tidak memiliki proses pelatihan, (KNN) sangat cepat dalam proses pelatihan (karena memang tidak ada) dan sangat lambat dalam proses prediksi.
- d. Menentukan nilai K yang paling sesuai adalah masalah yang rumit.
- e. Dalam kasus datanya, parameter jarak juga penting untuk dipertimbangkan karena K-NN pada prinsipnya memilih tetangga terdekat. Meskipun *Euclidean* ideal untuk menggunakan jarak terdekat (lurus) antara dua data, Manhattan sangat kuat (*robust*) untuk menemukan *outlier* dalam data.

[12], Metode *K-Nearest Neighbor* adalah salah satu metode klasifikasi yang paling umum digunakan. Tujuan metode ini adalah untuk mengklasifikasikan objek baru berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek baru tersebut. Algoritma *K-Nearest Neighbor* mudah digunakan. Dalam hal ini, pengguna menentukan jumlah data, juga dikenal sebagai tetangga terdekat, yang ditunjukkan oleh k [13].

Langkah-langkah dari algoritma K-NN (*K-Nearest Neighbors*), antara lain:

1. Tentukan parameter K , yang merupakan jumlah tetangga terdekat.
2. Hitung jarak antara semua data yang ada di data pelatihan dan data baru.

Rumus ini dapat digunakan untuk menghitung jarak antara semua data baru dan sebelumnya:

$$\textit{Euclidean Distance} : d = |x_i - y_i| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Gambar 1. Rumus Menghitung Jarak

Penjelasan :

d = Jarak

x_i = Data Uji

y_i = Data Latih

i = Variabel Data

n = Banyaknya data

3. Urutkan jarak tersebut dan tentukan tetangga mana yang terdekat berdasarkan jarak minimum ke K.
4. Tentukan kategori tetangga terdekat
5. dan gunakan kategori mayoritas sederhana sebagai nilai prediksi dari data baru.

Setelah model klasifikasi diterapkan, evaluasi terhadap hasil klasifikasi dilakukan dengan menggunakan metrik yang tepat, seperti akurasi, *precision*, *recall*, dan *F1-score*. Metrik ini digunakan untuk mengukur seberapa baik model KNN dalam mengklasifikasikan struktur daerah berdasarkan data yang ada. Dalam penelitian ini, hasil evaluasi lalu dibandingkan dengan algoritma klasifikasi lain, seperti *Decision Tree* atau *SVM (Support Vector Machine)*, untuk menentukan kelebihan dan kekurangan masing-masing pendekatan [1]. Perbandingan dengan algoritma lain ini bertujuan untuk mengevaluasi efektivitas K-NN dalam konteks klasifikasi struktur wilayah yang lebih kompleks. Hasil evaluasi ini dapat memberikan gambaran jelas tentang keandalan K-NN sebagai metode klasifikasi untuk analisis struktur daerah di Kota Medan.

Metode penelitian ini juga melibatkan analisis sensitivitas terhadap parameter-parameter yang digunakan dalam model, seperti pemilihan fitur dan pengaturan nilai k. Dengan demikian, fokus penelitian ini tidak hanya berfokus kepada penerapan K-NN, melainkan juga berupaya untuk mengoptimalkan model klasifikasi agar menghasilkan hasil yang lebih akurat. Penelitian ini diharapkan dapat memberi kontribusi untuk mengembangkan model klasifikasi berbasis K-NN untuk analisis struktur daerah, yang dapat diterapkan dalam konteks perencanaan kota [14]. Dengan mengoptimalkan metode ini, diharapkan dapat memberikan wawasan baru dalam pengelolaan dan perencanaan tata ruang kota yang lebih efisien dan efektif.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, analisis klasifikasi struktur daerah di Kota Medan menggunakan algoritma K-NN (*K-Nearest Neighbor*) yang tujuannya ialah untuk mengidentifikasi dan mengelompokkan wilayah-wilayah dengan karakteristik serupa berdasarkan berbagai atribut seperti kepadatan penduduk, penggunaan lahan, dan infrastruktur. Metode digunakan untuk mengolah dan menganalisis data tertentu yang melibatkan pengumpulan data, pra-pemrosesan, pemilihan fitur, dan penerapan algoritma KNN. Hasil dari penelitian ini memperlihatkan bahwa algoritma K-NN dapat digunakan secara efektif dalam mengklasifikasikan struktur daerah di Kota Medan berdasarkan atribut yang telah disebutkan sebelumnya.

Pengumpulan data yang relevan merupakan tahap pertama yang cukup penting dalam penelitian ini. Data yang digunakan mencakup berbagai informasi yang tersedia melalui sensus penduduk, data penggunaan lahan, serta infrastruktur yang ada di tiap kecamatan di Kota Medan. Setelah data terkumpul, tahap selanjutnya adalah pra-pemrosesan data yang meliputi pembersihan dan normalisasi data. Hasil pra-pemrosesan ini penting untuk memastikan bahwa informasi yang digunakan dalam model adalah K-NN tidak mengandung kesalahan atau inkonsistensi yang dapat memengaruhi hasil klasifikasi. Beberapa data yang hilang atau tidak valid dihapus, sementara data yang masih relevan dinormalisasi agar tidak ada atribut yang dominan. Langkah ini sangat penting untuk memastikan bahwa algoritma K-NN dapat bekerja dengan baik, karena algoritma ini bergantung pada perhitungan jarak antar titik data [1].

Setelah data siap, pemilihan fitur yang relevan dilakukan untuk memilih atribut yang paling berpengaruh terhadap klasifikasi struktur daerah. Fitur-fitur seperti kepadatan penduduk, jenis penggunaan lahan (perumahan, komersial, industri, dan lainnya), serta distribusi infrastruktur publik (sekolah, rumah sakit, dan tempat ibadah) dipilih karena masing-masing atribut ini memberikan gambaran yang signifikan mengenai struktur wilayah. Pada penelitian sebelumnya, pemilihan fitur yang tepat terbukti berpengaruh besar terhadap akurasi model yang dibangun [3]. Dalam hal ini, analisis terhadap keterkaitan antar fitur dilakukan menggunakan metode statistik untuk memastikan bahwa hanya fitur yang memberikan kontribusi signifikan yang digunakan dalam model.

Setelah fitur-fitur dipilih, model K-NN diterapkan untuk mengklasifikasikan wilayah-wilayah di Kota Medan. Algoritma K-NN bekerja dengan mengukur jarak antar titik data, di mana jarak antar data dihitung menggunakan rumus jarak *Euclidean*. Setiap titik data kemudian diklasifikasikan berdasarkan kedekatannya dengan titik data lain yang sudah diketahui kelasnya. Sebagai contoh, wilayah dengan kepadatan penduduk tinggi dan banyak fasilitas publik akan diklasifikasikan dalam satu kelompok, sementara wilayah dengan kepadatan penduduk rendah dan sedikit fasilitas publik akan dikategorikan dalam kelompok yang berbeda. Hasil dari

klasifikasi memperlihatkan bahwa algoritma K-NN mampu mengelompokkan wilayah-wilayah di Kota Medan dengan baik, dengan sebagian besar kecamatan memiliki kesamaan dalam hal kepadatan penduduk dan jenis penggunaan lahan [11].

Selanjutnya, untuk mengevaluasi kinerja model, dilakukan pengujian terhadap beberapa nilai k yang berbeda, yaitu jumlah tetangga terdekat yang digunakan dalam perhitungan jarak. Hasil pengujian menunjukkan bahwa memilih nilai k yang tepat sangat berpengaruh terhadap akurasi klasifikasi. Pada penelitian ini, nilai k yang optimal ditemukan pada angka 5, di mana model K-NN memberikan akurasi terbaik dalam mengklasifikasikan struktur daerah di Kota Medan. Penentuan nilai k yang tepat penting karena jika nilai k terlalu kecil, model cenderung sangat sensitif terhadap *outlier*, sementara jika nilai k terlalu besar, model bisa menjadi terlalu umum dan kurang presisi dalam mengklasifikasikan data [4]. Penentuan nilai k yang optimal ini didasarkan pada hasil eksperimen dengan berbagai nilai k , di mana $k=5$ memberi keseimbangan antara akurasi dan presisi.

K-NN dilakukan dengan memanfaatkan ukuran seperti akurasi, *precision*, *recall*, dan skor F1 untuk mengevaluasi sejauh mana hasil klasifikasi sesuai dengan data asli. Hasil pengujian menunjukkan bahwa model K-NN menghasilkan akurasi yang cukup tinggi, yaitu sekitar 85%, yang menunjukkan bahwa K-NN dapat digunakan secara efektif untuk klasifikasi struktur daerah di Kota Medan. Namun, meskipun hasil akurasi tergolong baik, model K-NN juga menunjukkan beberapa kelemahan dalam hal *precision* dan *recall* pada beberapa kelas tertentu. Hal ini disebabkan oleh keberagaman struktur daerah yang ada di Kota Medan, yang kadang menyebabkan data dari beberapa kecamatan tumpang tindih, sehingga model kesulitan dalam mengklasifikasikan beberapa kelas dengan presisi yang tinggi [5].

Sebagai perbandingan, algoritma klasifikasi lain seperti *Decision Tree* dan SVM (*Support Vector Machine*) juga diterapkan untuk menganalisis struktur daerah di Kota Medan. Hasil dari perbandingan ini memperlihatkan algoritma K-NN memberikan kinerja yang sangat baik dalam hal akurasi, tetapi algoritma SVM memberikan hasil yang lebih stabil dalam hal *precision* serta *recall*, terutama untuk kelas-kelas yang memiliki variasi data yang lebih besar. Hasil perbandingan ini sejalan dengan temuan studi sebelumnya yang menunjukkan bahwa K-NN, meskipun efektif, dapat memiliki keterbatasan dalam hal presisi dalam beberapa kasus [1]. Meskipun demikian, pemilihan algoritma tetap bergantung pada tujuan spesifik dari penelitian atau aplikasi yang sedang dilakukan. Dalam konteks klasifikasi struktur daerah di Kota Medan, K-NN tetap menjadi pilihan yang solid karena kemampuannya yang cepat dalam mengolah data dan memberikan hasil yang cukup akurat.

Selain itu, evaluasi terhadap parameter-parameter lain yang digunakan dalam model juga dilakukan. Salah satu parameter yang dianalisis adalah jarak yang digunakan untuk menghitung kedekatan antar data. Meskipun jarak *Euclidean* adalah pilihan yang umum digunakan dalam K-NN, penelitian ini juga mencoba menggunakan jarak Manhattan dan jarak Minkowski untuk mengukur dampaknya

terhadap hasil klasifikasi. Hasil tes menunjukkan bahwa jarak geometris memberikan hasil yang lebih baik daripada jarak lainnya, yang menunjukkan bahwa dalam kasus ini, jarak *Euclidean* adalah pilihan yang paling tepat untuk mengklasifikasikan data struktur daerah di Kota Medan [9].

Pada penelitian ini memberikan hasil bahwa algoritma K-NN dapat diterapkan dengan efektif untuk mengklasifikasikan struktur daerah di Kota Medan berdasarkan berbagai atribut yang ada. Namun, penelitian ini juga menunjukkan bahwa meskipun K-NN efektif, ada beberapa tantangan yang perlu diatasi, seperti keberagaman dan kompleksitas data yang dapat memengaruhi presisi model. Maka diperlukan penelitian lebih lanjut agar dapat mengoptimalkan algoritma ini serta mengeksplorasi metode lain yang dapat meningkatkan hasil klasifikasi. Penggunaan metode *hybrid* atau penerapan teknik *pre-processing* yang lebih canggih dapat menjadi solusi untuk meningkatkan akurasi dan stabilitas hasil klasifikasi, terutama pada kelas-kelas yang lebih kompleks dan memiliki banyak variabilitas [10].

Secara keseluruhan, menghasilkan kontribusi yang signifikan dalam penerapan algoritma K-NN untuk analisis struktur daerah, khususnya dalam konteks Kota Medan. Temuan-temuan ini dapat digunakan sebagai dasar untuk perencanaan dan pengembangan kota yang lebih efisien, dengan memperhatikan karakteristik dan kebutuhan tiap daerah secara lebih mendalam.

4. KESIMPULAN

Dalam penelitian ini, telah dilakukan analisis klasifikasi struktur daerah di Kota Medan menggunakan algoritma K-NN (*K-Nearest Neighbor*). Penelitian ini bertujuan untuk mengidentifikasi dan mengelompokkan wilayah-wilayah dengan karakteristik serupa berdasarkan atribut seperti kepadatan penduduk, penggunaan lahan, dan distribusi infrastruktur. Berdasarkan hasil eksperimen, dapat disimpulkan bahwa K-NN merupakan algoritma yang efektif untuk melakukan klasifikasi data geospasial di Kota Medan dengan tingkat akurasi yang cukup tinggi. Meskipun begitu, Akibatnya, ada beberapa masalah yang harus dipertimbangkan dalam penerapan algoritma ini.

Salah satu temuan penting pada penelitian ini ialah pentingnya pemilihan fitur yang tepat dalam proses klasifikasi. Pemilihan atribut yang relevan sangat memengaruhi hasil akhir dari model K-NN. Dalam penelitian ini, atribut seperti kepadatan penduduk, penggunaan lahan, dan infrastruktur terbukti memberikan kontribusi signifikan terhadap pengelompokan wilayah-wilayah di Kota Medan. Penggunaan fitur-fitur ini memungkinkan model K-NN untuk menghasilkan klasifikasi yang lebih akurat dan sesuai dengan karakteristik daerah yang ada.

Nilai k dalam algoritma K-NN memberikan kontribusi yang sangat besar dalam penentuan kualitas klasifikasi. Hasil penelitian memperlihatkan bahwa nilai $k=5$ memberikan hasil yang optimal untuk klasifikasi struktur daerah di Kota Medan. Pemilihan nilai k yang akurat dapat berpengaruh pada ketepatan klasifikasi, karena nilai k yang sangat kecil atau terlalu besar bisa menimbulkan model menjadi sangat peka terhadap outlier atau kehilangan akurasi. Oleh karena itu, eksperimen dengan

berbagai nilai k menjadi langkah yang sangat penting untuk menemukan nilai yang paling sesuai dengan data yang digunakan.

Selama evaluasi model, metrik yang digunakan seperti akurasi, *recall*, *precision*, dan F1-score memperlihatkan bahwa K-NN dapat menghasilkan hasil yang cukup baik. Meskipun demikian, beberapa kelas dengan karakteristik data yang lebih kompleks masih menunjukkan kekurangan dalam hal *precision* dan *recall*. Hal ini menunjukkan bahwa meskipun K-NN efektif untuk klasifikasi, ada beberapa kelemahan yang perlu diatasi, terutama dalam hal klasifikasi data yang memiliki keragaman yang lebih tinggi. Oleh sebab itu, diperlukan pengembangan lebih mendalam agar dapat meningkatkan akurasi model, seperti dengan menggunakan teknik pengolahan data yang lebih canggih atau algoritma alternatif.

Perbandingan dengan algoritma lain, seperti SVM (*Support Vector Machine*), juga menunjukkan pemahaman mengenai kelebihan serta kekurangan pada setiap metode. Meskipun SVM (*Support Vector Machine*) memberikan hasil yang lebih stabil dalam hal *precision* dan *recall*, K-NN tetap menjadi pilihan utama yang baik dalam hal kemudahan implementasi serta kecepatan dalam mengolah data. Keputusan dalam memilih algoritma yang tepat harus mempertimbangkan tujuan penelitian dan sifat data yang digunakan. K-NN, meskipun sederhana, tetap menawarkan banyak keuntungan dalam hal efisiensi dan hasil yang cukup memuaskan.

Hasil penelitian ini membuka banyak prospek baru untuk pengembangan dalam hal klasifikasi dan analisis struktur daerah, khususnya di daerah perkotaan seperti Kota Medan. Salah satu prospek utama adalah penggunaan algoritma K-NN secara lebih luas dalam sistem pendukung keputusan pemerintah daerah untuk distribusi infrastruktur publik, zonasi pembangunan, dan perencanaan tata ruang. Selain itu, algoritma K-NN dapat digunakan untuk membangun model klasifikasi yang lebih kompleks dan adaptif terhadap data berskala besar. Penggunaan metode *ensemble* atau *hybrid* dapat meningkatkan akurasi dan ketahanan model terhadap data yang tidak seimbang atau memiliki *outlier*.

Selain itu, eksplorasi algoritma klasifikasi berbasis waktu juga dikenal sebagai klasifikasi berbasis waktu dimaksudkan untuk memantau perubahan struktur daerah dari waktu ke waktu. Ini memungkinkan untuk menganalisis tren urbanisasi, memprediksi perubahan penggunaan lahan, dan memprediksi kebutuhan infrastruktur di masa depan. Maka pada penelitian ini tidak hanya memberikan kontribusi pada bidang akademik, tetapi juga dapat diterapkan di dunia nyata untuk perencanaan wilayah, pengelolaan kota pintar (*smart city*), dan pengambilan keputusan berbasis data.

Secara keseluruhan, proses penelitian yang dilakukan ini menunjukkan bahwa algoritma K-NN memiliki potensi besar dalam analisis klasifikasi struktur daerah, khususnya di Kota Medan. Meskipun terdapat tantangan dalam hal presisi dan *recall* pada kelas tertentu, penelitian ini berhasil memberikan gambaran yang jelas tentang bagaimana metode dapat digunakan dalam konteks perencanaan kota. Di masa depan, Penelitian tambahan dapat dilakukan untuk melihat kombinasi metode lain atau mengoptimalkan parameter model untuk meningkatkan kinerja

klasifikasi pada data yang lebih kompleks. Oleh karena itu, penelitian ini berkontribusi untuk mengembangkan sistem analisis geospasial berbasis K-NN dalam perencanaan dan pengelolaan tata ruang kota.

DAFTAR PUSTAKA

- [1] D. A. Larasati, "Penerapan Metode K-Nn Dan Ekstraksi Fitur Glcm Dalam Mengklasifikasi Citra Ikan Berformalin," 2021.
- [2] Y. A. Singgalen, "Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL," *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, 2023, doi: 10.47065/bits.v4i4.3231.
- [3] S. M. Sari, "Analisis Sentimen Terhadap New Normal Di Era Covid-19 Menggunakan Algoritma K-Nearest Neighbor (K-NN)," pp. 1–80, 2021.
- [4] Erfin Nur Rohma, "Comparison of the Social Welfare Data Classification Algorithm for Bantul Regency," *J. Process.*, vol. 17, no. 2, pp. 91–100, 2022, doi: 10.33998/processor.2022.17.2.1222.
- [5] R. F. Putra *et al.*, *DATA MINING: Algoritma dan Penerapannya*. PT. Sonpedia Publishing Indonesia, 2023. [Online]. Available: <https://books.google.co.id/books?id=zLHGEEAAQBA>
- [6] A. I. Putri, "Application of Data Mining to Predict Birth Rates in Medan City Using the K-Nearest Neighbor Method," *J. Comput. Sci. Inf. Technol. Telecommun. Eng.*, vol. 5, no. 1, pp. 504–509, 2024, doi: 10.30596/jcositte.v5i1.17991.
- [7] R. Anggara, T. O. Mukhti, and Y. Kurniawati, "Comparison of Naïve Bayes and K-Nearest Neighbors Methods in Classifying Human Development Index by Districts / City Indonesia in 2022," vol. 2, no. 2022, pp. 483–488, 2024.
- [8] H. Al Aziez and G. Anuraga, "Klasifikasi Daerah Tertinggal di Indonesia Menggunakan Algoritma SVM dan k-NN," *J. Ilmu Dasar*, vol. 22(1), no. 1, pp. 31–38, 2021.
- [9] M. Firanti, "KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) (Studi Kasus : Puskesmas Kecamatan Gunung Meriah) SKRIPSI OLEH : FAKULTAS TEKNIK KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) (St," 2024.
- [10] S. Larasati, "ANALISIS ARSITEKTUR MOBILENET PADA KLASIFIKASI PENYAKIT TANAMAN TOMAT," 2023.
- [11] A. Syaputra, "Klasifikasi Penyakit Daun pada Tebu dengan Pendekatan Algoritma K-Nearest Neighbors , Multilayer Perceptron dan Support Vector Machine," vol. 15, no. 3, 2024.
- [12] Yahya and W. P. Hidayanti, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok El ektrik) pada ' Lombok Vape On ' Pendahuluan dihasilkan tidak stabil dan tidak mampu diprediksi Dari

- penelitian yang dilakukan , berusaha untuk mengklasifikasik,” *J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 104–114, 2020, [Online]. Available: https://e-journal.hamzanwadi.ac.id/index.php/infotek/article/view/2279/pdf_23
- [13] L. Anshori, R. R. M. Putri, and Tibyani, “Implementasi Metode K-Nearest Neighbor untuk Rekomendasi Keminatan Studi (Studi Kasus : Jurusan Teknik Informatika Universitas Brawijaya),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 7, pp. 2745–2753, 2018, [Online]. Available: https://www.researchgate.net/publication/321097020_Implementasi_Metode_K-Nearest_Neighbor_untuk_Rekomendasi_Keminatan_Studi_Studi_Kasus_Jurusan_Teknik_Informatika_Universitas_Brawijaya
- [14] N. Mayasari, C. Rizal, and F. Wulandari, “RANCANGAN SISTEM KLASIFIKASI KARAKTERISTIK SISWA MENGGUNAKAN ALGORITMA C4. 5 PADA SMA N 1 HAMPARAN PERAK,” *ESCAF*, pp. 1214–1219, 2024.