

Perbandingan Metrik Euclidean dan Metrik Manhattan untuk *K-Nearest Neighbors* dalam Klasifikasi Kismis

Moch. Anjas Aprihartha^{*1}, Taufik Nur Alam², M. Husniyadi³

^{1,2,3}Program Studi PJJ Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Indonesia
Email: ¹anjas.aprihartha@dsn.dinus.ac.id, ²118202300068@mhs.dinus.ac.id,
³118202200012@mhs.dinus.ac.id

Abstrak

Pada konferensi negara-negara penghasil kismis tanpa biji ke-63, Negara Turki dinobatkan sebagai pemimpin dunia dalam produksi dan ekspor kismis musim 2022/2023. Dalam menghasilkan kismis yang berkualitas sangat bergantung pada varietas anggur dan proses pengolahannya seperti pengeringan, grading, dll. Pada beberapa tempat, penyortiran kismis masih dilakukan secara manual berdasarkan indeks mutu. Namun, proses penyortiran dengan manual memiliki beberapa kelemahan seperti membutuhkan biaya yang besar dan pemilihan kismis oleh manusia cenderung tidak konsisten serta kurang efisien. Untuk mengatasi masalah tersebut maka dilakukan penelitian untuk mengembangkan metode alternatif dalam mengidentifikasi jenis kismis. *K-Nearest Neighbors* (KNN) merupakan algoritma *machine learning* nonparametrik yang dapat menyelesaikan masalah klasifikasi atau regresi. Metode KNN memprediksi kategori dengan memanfaatkan hubungan jarak antartetangga terdekat. Dalam analisis data, mengetahui jarak antartetangga terdekat dapat dihitung menggunakan dua metrik utama seperti jarak Manhattan dan jarak Euclidean. Studi kali ini akan mengembangkan metode KNN dalam mendeteksi jenis kismis Besni dan Kecimen. Setiap pelatihan dan pengujian menggunakan parameter k yang berbeda, yaitu 3, 5, 7, dan 9. Hasil penelitian diperoleh KNN pada metrik Euclidean menunjukkan nilai akurasi, presisi, sensitivitas, dan spesifisitas untuk setiap k hampir tidak memiliki perbedaan yang signifikan. Akurasi tertinggi sebesar 0,850 ketika $k=7$ dan $k=9$. Sedangkan KNN pada metrik Manhattan menunjukkan Akurasi tertinggi sebesar 0,867 ketika $k=5$ dan $k=7$. Algoritma KNN dengan metrik Manhattan dan metrik Euclidean, keduanya memberikan akurasi yang kuat dalam mengklasifikasikan data. Berdasarkan hasil uji kedua algoritma maka KNN pada metrik Manhattan lebih baik dibandingkan KNN pada metrik Euclidean.

Kata kunci: *Euclidean, Kismis, KNN, Manhattan, Metrik*

Abstract

At the 63rd conference of seedless raisin producing countries, Turkey was named the world leader in raisin production and export for the 2022/2023 season. In producing quality raisins, it really depends on the grape variety and processing processes such as drying, grading, etc. In some places, sorting raisins is still done manually based on the quality index. However, the manual sorting process has several disadvantages such as requiring large costs and selecting raisins by humans tends to be inconsistent and less efficient. To overcome this problem, research was carried out to develop alternative methods for identifying types of raisins. *K-Nearest Neighbors* (KNN) is a nonparametric machine learning algorithm that can solve classification or regression problems. The KNN method predicts categories by utilizing the distance relationship between nearest neighbors. In data analysis, knowing the distance between nearest neighbors can be calculated using two main metrics such as Manhattan distance and Euclidean distance. This study will develop the KNN method for detecting Besni and Kecimen raisins. Each training and testing use different k parameters, namely 3, 5, 7, and 9. The research results obtained by KNN on the Euclidean metric show that the accuracy, precision, sensitivity, and specificity values for each k have almost no significant differences. The highest accuracy is 0.850 when $k=7$ and $k=9$. Meanwhile, KNN on the Manhattan metric shows the highest accuracy of 0.867 when $k=5$ and $k=7$. The KNN algorithm with Manhattan metrics and Euclidean metrics both provides strong accuracy in classifying data. Based on the test results of the two algorithms, KNN on the Manhattan metric is better than KNN on the Euclidean metric.

Keywords: *Euclidean, KNN, Manhattan, Metric, Raisin*

1. PENDAHULUAN

Pada konferensi negara-negara penghasil kismis tanpa biji ke-63, Negara Turki dinobatkan sebagai pemimpin dunia dalam produksi dan ekspor kismis musim 2022/2023, dimulai pada 1 September 2022 disusul oleh Amerika menempati urutan kedua dan Cina diposisi ketiga [1].

Dalam menghasilkan kismis yang berkualitas sangat bergantung pada varietas anggur dan proses pengolahannya seperti pengeringan, grading, dll . Pada beberapa tempat, penyortiran kismis masih dilakukan secara manual berdasarkan indeks mutu [2]. Namun, proses penyortiran dengan manual memiliki beberapa kelemahan seperti membutuhkan biaya yang besar dan pemilihan kismis oleh manusia cenderung tidak konsisten serta kurang efisien.

Untuk mengatasi masalah tersebut maka dilakukan penelitian untuk mengembangkan metode alternatif dalam mengidentifikasi kualitas kismis. Beberapa penelitian terdahulu telah melakukan studi terkait masalah klasifikasi kismis dengan menerapkan metode *machine learning*. Penelitian oleh Guo, *et al.* [3] yang mengklasifikasikan kismis berdasarkan keasliannya dengan metode SVM dan *multiscale fusion convolutional neural network*. Penelitian oleh Zhang, *et. al* [4] yang mengklasifikasikan kismis dalam tiga kategori, Hongxiangfei, Manaiti, dan Munage dengan algoritma *k-nearest neighbors* (KNN), *linear discriminant analysis*, dan SVM. Penelitian oleh Tarakci dan Ozkan [5] yang mengklasifikasikan jenis kismis Osmancık dan Cammeo menggunakan metode KNN dan *weighted k-nearest neighbors* (WKNN). Terakhir penelitian oleh Feng, *et. al* [6] dengan algoritma KNN, SVM, dan *radial basis function neural network* dalam mengidentifikasi jenis kismis Wuhebai dan Xiangfei.

Pada penelitian sebelumnya, metode KNN cukup sering digunakan dalam mendeteksi jenis kismis. KNN merupakan metode yang bersifat nonparametrik [7]. Oleh karena itu, metode KNN tidak melibatkan asumsi spesifik dengan distribusi data tertentu sehingga dapat menangani data apapun secara fleksibel. Metode ini bekerja dengan memperhatikan kesamaan karakteristik pada data terdekatnya. Metode KNN dapat diaplikasikan pada dua metrik yang berbeda, yaitu metrik Euclidean dan metrik Manhattan [8,9]. Dalam penelitian Dinata, *et. al* [8] yang membandingkan algoritma KNN metrik Euclidean dengan KNN metrik Manhattan untuk klasifikasi transportasi bus, diperoleh akurasi tertinggi pada metrik Manhattan, yaitu 84%. Penelitian oleh Salsabila, *et. al* [9] yang membandingkan metrik Euclidean dan metrik Manhattan pada algoritma KNN dalam studi kasus klasifikasi *stunting* balita, diperoleh akurasi metrik Manhattan lebih tinggi daripada metrik Euclidean. Meskipun dari dua sampel kasus tersebut KNN pada metrik Manhattan lebih unggul dibandingkan metrik Euclidean, namun belum cukup bukti bahwa hasil serupa dapat terjadi pada studi kasus yang berbeda. Dalam literatur yang ada, algoritma KNN metrik Euclidean sangat populer dikalangan peneliti karena kemudahan penggunaannya dan adanya pustaka (*library*) pada *software R*.

Berdasarkan paparan yang telah dijelaskan, penelitian ini berpusat pada perbandingan efektifitas antara metode KNN pada metrik Euclidean dan metrik Manhattan dalam klasifikasi kismis. Kedua metode memiliki perbedaan dalam mengukur jarak antar dua atau lebih objek. Setelah dilakukan analisis komprehensif pada dua metode tersebut kemudian menentukan metode yang terbaik berdasarkan penilaian akurasi, sensitivitas, spesifisitas, dan presisi. Hasil penelitian ini diharapkan dapat menjadi acuan dalam pengembangan sistem seleksi kismis untuk memperkuat efisiensi di industri.

2. METODE PENELITIAN

2.1. Data dan Variabel Penelitian

Data pada penelitian ini diperoleh dari sumber terbuka UCI *Machine Learning Respository* [10]. Dataset terdiri dari 8 variabel dengan jumlah sampel sebanyak 900 amatan.

Tabel 1. Variabel dan Jenis Data Penelitian

Variabel	Jenis Data
<i>Class</i>	Kategorik 1. Besni 2. Kecimen
<i>Area</i>	Numerik
<i>Major Axis Length</i>	Numerik
<i>Minor Axis Length</i>	Numerik
<i>Eccentricity</i>	Numerik
<i>Convex Area</i>	Numerik
<i>Extent</i>	Numerik
<i>Perimeter</i>	Numerik

2.2. Preprocessing Data

Data yang dikumpulkan dalam keadaan mentah harus menjalani serangkaian pra-pemrosesan agar menjadi data yang siap diproses dengan algoritma data mining. Menurut *preprocessing data* merupakan serangkaian operasi yang mengubah data mentah menjadi data yang berkualitas [11]. Tahap *preprocessing data* tidak hanya diperlukan untuk memenuhi kebutuhan, tetapi juga efektifitas dalam meningkatkan performa model klasifikasi. Ini mencakup reduksi dimensi, deteksi *outlier*, dan transformasi. Salah satu teknik transformasi data adalah transformasi *Z-Score Normalization*.

$$Z = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (1)$$

dengan x_{ij} merupakan amatan ke- i pada variabel ke- j , \bar{x}_j merupakan rata-rata pada variabel ke- j , dan σ_j merupakan standar deviasi pada variabel ke- j .

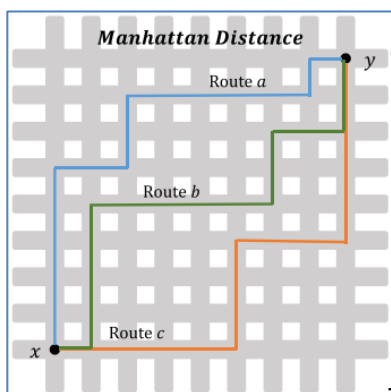
2.3. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) merupakan algoritma *machine learning* yang dapat menyelesaikan masalah klasifikasi atau regresi. Selain itu KNN dapat digunakan dalam mendeteksi adanya *outlier* [12]. Metode KNN memprediksi kategori dengan memanfaatkan hubungan jarak antartetangga terdekat. Dalam analisis data, mengetahui jarak antartetangga terdekat dapat dihitung menggunakan dua metrik utama seperti jarak Manhattan dan jarak Euclidean.

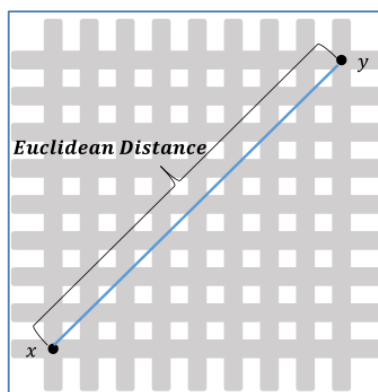
Jarak Manhattan merupakan penjumlahan panjang yang diproyeksikan ke sumbu oleh jalur terpendek yang harus dilalui selangkah demi selangkah ke titik lain dalam sistem koordinat seperti yang ditunjukkan pada Gambar 1 [13]. Sedangkan Jarak Euclidean merupakan jarak garis lurus antara dua titik seperti ditunjukkan pada Gambar 2 dan dapat mengukur data dari struktur hiperbola sirkuler dalam ruang berdimensi tinggi [7].

$$Dist_{MH}(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2)$$

$$Dist_{EL}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$



Gambar 1. Jarak Manhattan



Gambar 2. Jarak Euclidean

2.4. Ukuran Performa Model

Confusion matrix merupakan sebuah tabel yang berisi hasil uji model klasifikasi. Tabel tersebut dapat mempermudah dalam melakukan pengukuran kinerja model klasifikasi.

Tabel 2. *Confusion Matrix*

	Kelas Prediksi	
	Ya	Tidak
Kelas Aktual	Ya	a
	Tidak	c

Beberapa ukuran kinerja model klasifikasi yang sering digunakan untuk menilai kemampuan suatu algoritma klasifikasi disajikan pada Tabel 3 [14].

Tabel 3. Ukuran Kinerja Model

Pengukuran	Rumus
Akurasi	$\frac{a + d}{a + b + c + d}$
Sensivifitas	$\frac{a}{a + b}$
Spesifisitas	$\frac{d}{c + d}$
Presisi	$\frac{a}{a + c}$

3. HASIL DAN PEMBAHASAN

Proses diawali dengan memeriksa variabel-variabel independen yang berjenis numerik memiliki skala seragam, ini dilakukan agar tidak ada variabel yang mendominasi variabel lainnya. Apabila variabel belum seragam maka dilakukan transformasi *Z-Score Normalization* pada persamaan (1).

Setelah data ditransformasi kemudian dataset dibagi menjadi dua bagian yaitu 80% data *training* dan 20% data *testing*. Proses dilakukan dilakukan dengan *simple random sampling*. Ini bertujuan agar setiap amatan memiliki peluang yang sama untuk masuk dalam dataset *training* ataupun dataset *testing*. Dataset *training* sebanyak 720 amatan dan dataset *testing* sebanyak 180 amatan.

Kemudian kedua algoritma yaitu KNN Manhattan dan KNN Euclidean dilatih dengan dataset *training* menggunakan persamaan (2) dan (3). Setelah proses pelatihan selesai, algoritma diuji menggunakan data *testing*. Tujuan uji ini untuk memastikan agar kedua algoritma dapat bekerja dengan baik pada data baru. Setiap pelatihan dan pengujian menggunakan parameter k yang berbeda, yaitu 3, 5, 7, dan 9.

3.1. Kinerja Algoritma KNN pada Metrik Euclidean

Hasil uji klasifikasi KNN jarak Euclidean dengan parameter $k = 3$ disajikan pada Tabel 4

Tabel 4. *Confusion Matrix* 3-NN Metrik Euclidean

		Kelas Prediksi	
		Besni	Kecimen
Kelas Aktual	Besni	68	12
	Kecimen	16	84

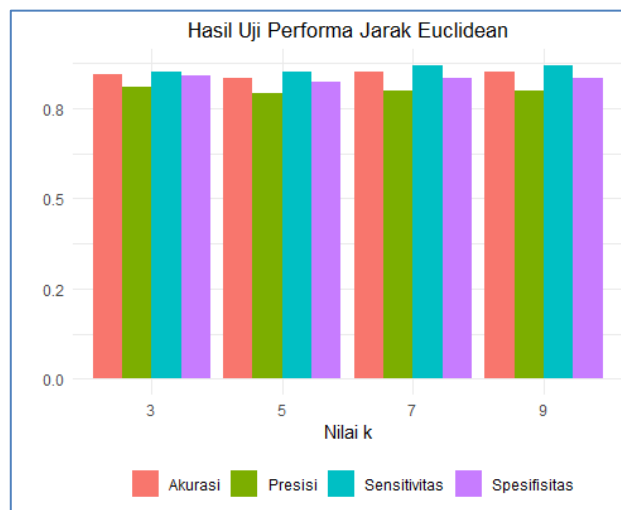
Berdasarkan Tabel 4 diperlihatkan bahwa 68 amatan tepat terklasifikasi dalam kelompok Besni dan 84 amatan tepat terklasifikasi dalam kelompok Kecimen. Sedangkan terdapat 28 amatan yang misklasifikasi.

Untuk mengetahui seberapa baik atau buruknya kemampuan algoritma dalam mengklasifikasi data dapat dihitung menggunakan persamaan pada Tabel 3. Hasil perhitungan ditunjukkan pada Tabel 5.

Tabel 5. Ukuran Performa 3-NN Metrik Euclidean

Pengukuran	Hasil
Akurasi	0,844
Sensitivitas	0,850
Spesifisitas	0,840
Presisi	0,809

Pada Tabel 5, hasil akurasi 0,844 berarti algoritma berhasil mengklasifikasikan kedua kelas dengan benar sebesar 84,4%. Sensitivitas 0,850 berarti algoritma berhasil mengklasifikasikan kelas Besni dengan benar sebesar 85%. Spesifisitas 0,840 berarti algoritma berhasil mengklasifikasikan kelas Kecimen dengan benar sebesar 84%. Presisi 0,809 berarti semua hasil yang algoritma identifikasi sebagai kelas Besni, 80,9% dari mereka benar-benar dari kelas Besni. Keempat pengukuran menghasilkan nilai diatas 0,80 yang menunjukkan kemampuan algoritma memiliki tingkat keberhasilan 80% dalam mengklasifikasi data. Hasil uji performa KNN metrik Euclidean untuk setiap k ditampilkan pada Gambar 3.



Gambar 3. Hasil Uji Performa KNN pada Metrik Euclidean

Pada Gambar 3, menunjukkan hasil akurasi, presisi, sensitivitas, dan spesifisitas untuk setiap k hampir tidak memiliki perbedaan yang signifikan. Akurasi tertinggi sebesar 0,850 ketika $k = 7$ dan $k = 9$ dan terendah sebesar 0,835 ketika $k = 5$. Presisi tertinggi sebesar 0,809 ketika $k = 3$ dan terendah

sebesar 0,790 ketika $k = 5$. Sensitivitas tertinggi sebesar 0,870 ketika $k = 7$ dan $k = 9$ dan terendah sebesar 0,850 ketika $k = 3$ dan $k = 5$. Spesifisitas tertinggi sebesar 0,840 ketika $k = 3$ dan terendah sebesar 0,823 ketika $k = 5$.

3.2. Kinerja Algoritma KNN pada Metrik Manhattan

Hasil uji klasifikasi KNN jarak Manhattan dengan parameter $k = 7$ disajikan pada Tabel 6.

Tabel 6. *Confusion Matrix* 7-NN Metrik Manhattan

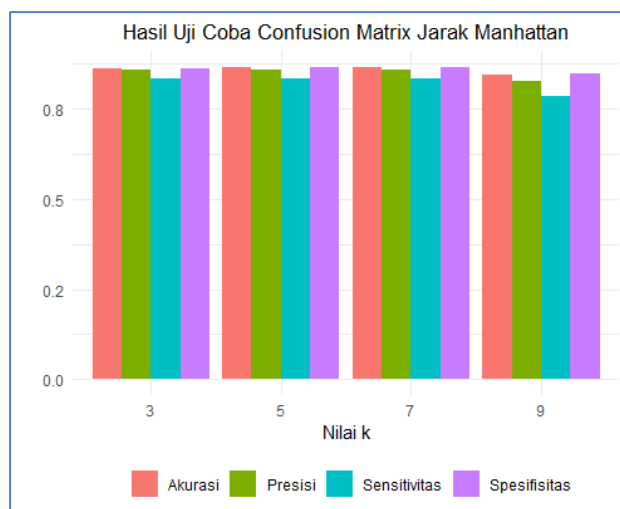
		Kelas Prediksi	
		Besni	Kecimen
Kelas Aktual	Besni	70	10
	Kecimen	14	86

Berdasarkan Tabel 6 diperlihatkan bahwa 70 amatan tepat terklasifikasi dalam kelompok Besni dan 86 amatan tepat terklasifikasi dalam kelompok Kecimen. Sedangkan terdapat 24 amatan lainnya yang misklasifikasi. Hasil pengukuran performa model yang ditunjukkan pada Tabel 7.

Tabel 7. Ukuran Performa 7-NN Metrik Manhattan

Pengukuran	Hasil
Akurasi	0,867
Sensitivitas	0,875
Spesifisitas	0,860
Presisi	0,833

Pada Tabel 7, hasil akurasi 0,867 berarti algoritma berhasil memprediksi dengan tepat kedua kelas sebesar 86,7%. Sensitivitas 0,850 berarti algoritma berhasil mengklasifikasikan kelas Besni dengan benar sebesar 87,5%. Spesifisitas 0,860 berarti algoritma berhasil mengklasifikasikan kelas Kecimen dengan benar sebesar 86%. Presisi 0,833 berarti semua hasil yang algoritma identifikasi sebagai kelas Besni, 83,3% dari mereka benar-benar dari kelas Besni. Keempat pengukuran menghasilkan nilai diatas 0,83 yang menunjukkan kemampuan algoritma memiliki tingkat keberhasilan 83% dalam mengklasifikasi data. Hasil uji performa KNN metrik Manhattan untuk setiap k ditampilkan pada Gambar 4.

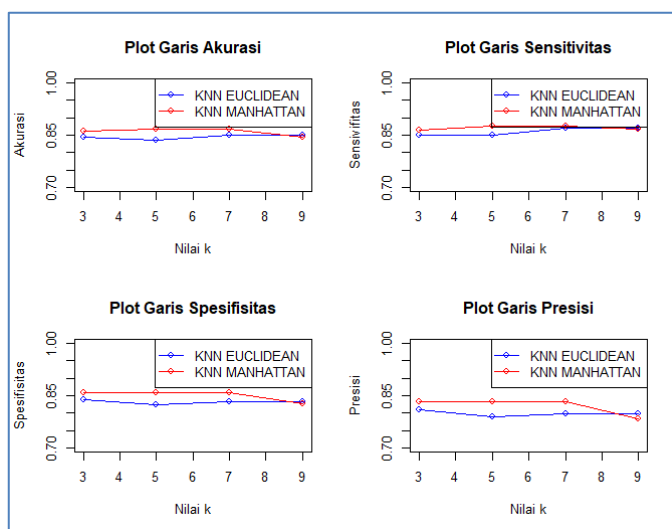


Gambar 4. Hasil Uji Performa KNN pada Metrik Manhattan

Pada Gambar 4, menunjukkan hasil akurasi, presisi, sensitivitas, dan spesifisitas untuk setiap k hampir tidak memiliki perbedaan yang signifikan. Ketika $k = 5$ dan $k = 7$ nilai akurasi, presisi, sensitivitas, dan spesifisitas tertinggi secara berturut-turut adalah 0,867; 0,860; 0,833; 0,867. Ketika $k = 9$ nilai akurasi, presisi, sensitivitas, dan spesifisitas terendah secara berturut-turut adalah 0,844; 0,826; 0,875; 0,847.

3.3. Perbandingan Performa KNN Euclidean dan KNN Manhattan

Perbandingan kinerja algoritma KNN metrik Euclidean dan metrik Manhattan yang meliputi akurasi, sensitivitas, spesifisitas, dan presisi disajikan dalam plot garis pada Gambar 5.



Gambar 5. Perbandingan Performa KNN Metrik Euclidean dan Metrik Manhattan

Pada Gambar 5 ditunjukkan plot garis akurasi menunjukkan KNN dengan metrik Manhattan cenderung memberikan akurasi sedikit lebih tinggi dibandingkan KNN dengan metrik Euclidean pada k yang diuji. Kedua algoritma tidak menunjukkan ada perubahan akurasi yang signifikan dalam rentang k yang diuji, akurasi berkisar antara 0,84 hingga 0,86. Plot garis sensitivitas dan spesifisitas menunjukkan KNN dengan metrik Manhattan memiliki sensitivitas dan spesifisitas lebih tinggi daripada KNN dengan metrik Euclidean untuk setiap k yang diuji. Plot garis sensitivitas dan spesifisitas menunjukkan perbedaan sensitivitas dan spesifisitas dalam dua algoritma tidak terlalu signifikan ketika k semakin tinggi. Plot garis presisi menunjukkan KNN dengan metrik Manhattan lebih tinggi daripada KNN dengan metrik Euclidean pada $k = 3, 5, 7$. Perubahan terlihat ketika $k = 9$, presisi KNN dengan metrik Euclidean lebih tinggi dibandingkan KNN dengan metrik Manhattan.

Berdasarkan hasil uji kedua metrik maka KNN dengan metrik Manhattan memberikan performa lebih baik dibandingkan KNN dengan metrik Euclidean. Meskipun demikian, hampir tidak ada perbedaan signifikan pada ukuran performa algoritma keduanya. Algoritma KNN dengan metrik Manhattan dan KNN dengan metrik Euclidean, keduanya memberikan akurasi yang kuat dalam mengklasifikasikan data.

Terdapat penelitian pada data yang sama dengan metode yang berbeda. Rangkuman hasil penelitian disajikan pada Tabel 8. Penelitian yang dilakukan oleh Wismadi, *et al.* [15] dalam mengklasifikasikan varietas kismis jenis Kecimen dan Besni dengan algoritma *artificial neural network* (ANN). Hasil penelitian dengan metode tersebut diperoleh akurasi sebesar 86,7%. Penelitian oleh Cinar, *et al.* [16] yang mengklasifikasikan jenis kismis Kecimen dan Besni menggunakan algoritma regresi logistik, *multilayer perceptron*, dan *support vector machine* (SVM). Hasil penelitian diperoleh metode SVM memberikan akurasi yang lebih tinggi dibandingkan metode lainnya, yaitu 86,44%. Penelitian lainnya oleh Anggraini *et al.* [17] menggunakan algoritma *naive bayes*, *decision tree*, *random forest*, *neural network* serta SVM dalam mengidentifikasi jenis kismis Kecimen dan Besni. Hasil penelitian diperoleh

efektifitas algoritma bekerja lebih baik pada *neural network* dibandingkan metode lainnya yang dibuktikan dengan akurasi 86,61%. Berdasarkan hasil pengukuran performa dengan metode yang berbeda, hampir tidak ada perbedaan signifikan akurasi dalam klasifikasi jenis kismis Kecimen dan Besni, akurasi berkisar antara 83,3% hingga 86,7%.

Tabel 8. Rangkuman Penelitian pada Kategori Kismis Kecimen dan Besni

Peneliti	Metode	Akurasi
Penelitian ini	KNN Metrik Euclidean	83,3% - 85%
	KNN Metrik Manhattan	84,4% - 86,7%
Wismadi, <i>et al.</i> [13]	<i>Artificial Neural Network</i> (ANN)	86,7%
Cinar, <i>et al.</i> [14]	Regresi Logistik	85,22%
	<i>Multilayer Persepsion</i>	86,33%
	<i>Support Vector Machine</i> (SVM)	86,44%
Anggraini <i>et al.</i> [15]	<i>Naive Bayes</i>	83,61%
	<i>Decision Tree</i>	86,25%
	<i>Random Forest</i>	83,33%
	<i>Neural Network</i>	86,61%
	<i>Support Vector Machine</i> (SVM)	86,39%

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan maka dapat ditarik beberapa kesimpulan. Hasil performa algoritma KNN pada metrik Euclidean menunjukkan nilai akurasi, presisi, sensitivitas, dan spesifisitas untuk setiap k hampir tidak memiliki perbedaan yang signifikan. Akurasi tertinggi sebesar 0,850 ketika $k = 7$ dan $k = 9$ dan terendah sebesar 0,835 ketika $k = 5$. Hal serupa terjadi pada algoritma KNN pada metrik Manhattan menunjukkan akurasi, presisi, sensitivitas, dan spesifisitas untuk setiap k juga hampir tidak memiliki perbedaan yang signifikan. Akurasi tertinggi sebesar 0,867 ketika $k = 5$ dan $k = 7$ dan terendah sebesar 0,844 ketika $k = 9$. Algoritma KNN dengan metrik Manhattan dan KNN dengan metrik Euclidean, keduanya memberikan akurasi yang kuat dalam mengklasifikasikan data. Berdasarkan hasil uji kedua algoritma maka KNN pada metrik Manhattan lebih baik dibandingkan KNN pada metrik Euclidean.

DAFTAR PUSTAKA

- [1] I. Ugural, "63rd International Raisin Conference held in Izmir," *Turkish Agri News*, Oct. 25, 2023. [Online]. Available: <https://www.turkishagrinenews.com/63-international-raisin-conference-held-in-izmir/> (accessed June 1, 2024).
- [2] K. Mollazade, M. Omid, and A. Arefi, "Comparing data mining classifiers for grading raisins based on visual features," *Computers and Electronics in Agriculture*, vol. 84, pp. 124-131, Oct. 2012. <https://doi.org/10.1016/j.compag.2012.03.004>.
- [3] J. Guo, C. Chen, and C. Chen, "Near-infrared spectroscopy combined with pattern recognition algorithms to quickly classify raisins," *Scientific Reports*, vol. 12, no. 7928, 2022. <https://doi.org/10.1038/s41598-022-12001-1>.
- [4] Y. Zhang, Y. Yang, C. Ma, and L. Jiang, "Identification of multiple raisins by feature fusion combined with NIR spectroscopy," *PloS One*, vol. 17, no. 7, p. e0268979, 2022. <https://doi.org/10.1371/journal.pone.0268979>.
- [5] F. Tarakeci and I. A. Ozkan, "Comparison of classification performance of kNN and WKNN algorithms," *Selcuk University Journal of Engineering Sciences*, vol. 20, no. 2, pp. 32-37, 2021.
- [6] L. Feng, S. Zhu, C. Zhang, Y. Bao, P. Gao, and Y. He, "Variety identification of raisins using near-infrared hyperspectral imaging," *Molecules*, vol. 23, no. 11, p. 2907, 2018. <https://doi.org/10.3390/molecules23112907>.

- [7] Y. Liang, Y. Pan, X. Yuan, W. Jia, and Z. Huang, "Surrogate modeling for long-term and high-resolution prediction of building thermal load with a metric-optimized KNN algorithm," *Energy and Built Environment*, vol. 4, no. 6, pp. 709-724, 2023. <https://doi.org/10.1016/j.enbenv.2022.06.008>.
- [8] R. K. Dinata, H. Akbar, dan N. Hasdyna, "Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 104-111, 2020. <https://doi.org/10.33096/ilkom.v12i2.539.104-111>
- [9] S. Salsabila, S. Martha, dan W. Andani, "Komparasi Algoritma K-Nearest Neighbor Dengan Euclidean Distance Dan Manhattan Distance Untuk Klasifikasi Stunting Balita (Studi Kasus: Puskesmas Kelurahan Parit Mayor)," *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 13, no. 2, 2024. <http://dx.doi.org/10.26418/bbimst.v13i2.77245>
- [10] UCI Machine Learning Repository, "Raisin Dataset," [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Raisin+Dataset>. (accessed May 6, 2024).
- [11] A. Tawakuli, B. Havers, V. Gulisano, D. Kaiser, and T. Engel, "Survey: Time-series data preprocessing: A survey and an empirical analysis," *Journal of Engineering Research*, 2024. <https://doi.org/10.1016/j.jer.2024.02.018>.
- [12] J. Yang, X. Tan, and S. Rahardja, "Outlier detection: How to Select k for k-nearest-neighbors-based outlier detectors," *Pattern Recognition Letters*, vol. 174, pp. 112-117, 2023. <https://doi.org/10.1016/j.patrec.2023.08.020>.
- [13] G. J. Tielen, T. Lulek, M. R. M. J. Traa, M. Kuzma, and W. J. Caspers, "The role of the Manhattan distance in antiferromagnetic ordering," *Physica A Statistical Mechanics & Its Applications*, vol. 246, pp. 199–220, 1997. [https://doi.org/10.1016/S0378-4371\(97\)00341-5](https://doi.org/10.1016/S0378-4371(97)00341-5).
- [14] M. A. Aprihartha, J. Prasetya, and S. I. Fallo, "Implementasi CART-Real Adaboost dalam Memprediksi Minat Pelanggan Membeli Sepatu," *Jurnal EurekaMatika*, vol. 12, no. 1, pp. 35-46. <https://doi.org/10.17509/jem.v12i1.67808>.
- [15] Wismadi, H. N., Yofi, C., Manumono, T. F., Hendrawan, F. A., Hilmy, M. R., Puspitasari, A., ... and Rifai, A. P., "Klasifikasi Varietas Biji Kismis dengan Artificial Neural Network," *Jurnal Optimasi Teknik Industri (JOTI)*, vol. 5, no. 1, pp. 8-13, 2023. <http://dx.doi.org/10.30998/joti.v5i1.13951>.
- [16] İ. Çınar, M. Koklu, and P. D. Ş. Taşdemir, "Classification of Raisin Grains Using Machine Vision and Artificial Intelligence Methods," *GJES*, vol. 6, no. 3, pp. 200–209, 2020.
- [17] Anggraini, R. A., Apriyani, Y., Pertiwi, M. W., Kusmira, M., and Bahri, S. , "Klasifikasi Jenis Kismis Menggunakan Teknik Data Mining," *Jurnal Kajian Ilmiah*, vol. 24, no. 1, pp. 45-56, 2024. <https://doi.org/10.31599/ryvqk945>.

Halaman Ini Dikosongkan