

PREDIKSI POLA PEMESANAN KAMAR PALING DIMINATI DI HOTEL SINGARAJA MENGGUNAKAN METODE KNN

M. Mahaputra Hidayat¹⁾, R. Dimas Adityo²⁾, Satrio Cahyo Agung Wibowo³⁾,
Yerryant Lituhayu Apta Widyadana⁴⁾, Ryammizard Zulfano P.N⁵⁾, Rizqi Nur
Rifa'i⁶⁾

E-mail : ¹⁾mahaputra@ubhara.ac.id , ²⁾dimas@ubhara.ac.id , ³⁾satriocyber07@gmail.com ,
⁴⁾yerrylaw21@gmail.com , ⁵⁾ryamizardzpn@gmail.com , ⁶⁾rosiped123@gmail.com

^{1,2,3,4,5,6} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Bhayangkara Surabaya

(Naskah masuk: 14 Januari 2025, diterima untuk diterbitkan: 30 April 2025)

Abstrak

Hotel Singaraja menghadapi tantangan kedepannya untuk memaksimalkan kamar hunian dan menentukan harga yang paling diminati oleh para pelanggan. Penelitian ini bertujuan untuk memprediksi pola pemesanan kamar dan harga yang paling diminati menggunakan metode *K-Nearest Neighbors* (KNN). Data yang digunakan mencakup informasi seperti pemesanan tipe kamar dan harga kamar sesuai tipe. Metode KNN dipilih karena kemampuannya dalam mengklasifikasikan data berdasarkan kedekatan karakteristik. Hasil analisis menunjukkan bahwa metode KNN mampu memprediksi pola pemesanan dan harga kamar yang populer yang baik, Algoritma KNN berhasil mengidentifikasi pola ini dengan tingkat akurasi mencapai 99.6%. Prediksi ini dapat membantu manajerial hotel yang baik dalam menentukan harga dan merancang strategi promosi sesuai dengan tipe kamar yang diminati pelanggan, serta meningkatkan efisiensi pengelolaan kamar. Dengan demikian, penerapan metode KNN memberikan manfaat signifikan dalam pengambilan keputusan terkait pemasaran dan perencanaan harga kamar di Hotel Singaraja.

Kata Kunci: pola pemesanan kamar, *k-nearest Neighbor*, perencanaan harga.

1. PENDAHULUAN

Seiring dengan meningkatnya jumlah hotel dan penginapan di berbagai destinasi wisata, persaingan dalam industri perhotelan semakin ketat. Hotel-hotel harus menghadapi tantangan untuk tetap relevan di pasar yang dinamis dengan terus meningkatkan layanan, strategi pemasaran, dan efisiensi operasional. Salah satu aspek penting dalam pengelolaan hotel adalah memahami pola pemesanan kamar dan preferensi harga pelanggan. Informasi ini tidak hanya membantu hotel untuk mempertahankan tingkat hunian yang optimal sepanjang tahun, tetapi juga memungkinkan mereka untuk menyesuaikan harga dan strategi promosi sesuai kebutuhan pasar.

Di era digital saat ini, data menjadi aset yang sangat berharga bagi perusahaan, termasuk dalam industri perhotelan. Dengan analisis data yang tepat, hotel dapat mengidentifikasi tren dan pola yang tersembunyi dalam data pemesanan mereka. Metode *K-Nearest Neighbor* adalah suatu metode yang melakukan klasifikasi terhadap data baru dan mengelompokkan data tersebut dengan data lama berdasarkan jarak paling dekat dengan data baru tersebut [1]. Dalam hal ini, metode *K-Nearest Neighbors* (KNN) menawarkan pendekatan yang sederhana namun efektif untuk memanfaatkan data historis guna memprediksi kebutuhan pelanggan di masa depan. KNN, sebagai salah satu algoritma *machine learning* berbasis *instance*, memungkinkan analisis yang mendalam berdasarkan kedekatan karakteristik antar data. Metode KNN juga sangat tangguh terhadap training data yang memiliki banyak *noise* dan efektif apabila training datanya besar [2]. Hal ini sangat relevan untuk kasus prediksi pola pemesanan dan preferensi harga karena pola-pola ini sering kali dipengaruhi oleh faktor kompleks seperti musim, hari libur, dan perilaku pelanggan. *Data mining* melibatkan penggunaan teknik – teknik analisis data yang canggih untuk mengekstraksi informasi dari data yang sangat besar, beragam, dan kompleks [3].

Penelitian ini bertujuan untuk mengimplementasikan metode KNN dalam memprediksi pola pemesanan kamar dan menentukan harga yang paling diminati di Hotel Singaraja. Dengan memanfaatkan data historis pemesanan, penelitian ini berusaha memberikan wawasan yang dapat diterapkan dalam pengambilan keputusan strategis, seperti pengelolaan inventaris kamar, penetapan harga yang kompetitif, dan desain promosi yang lebih efektif. Hasil dari penelitian ini diharapkan dapat membantu Hotel Singaraja untuk meningkatkan kepuasan pelanggan, mengoptimalkan pendapatan, dan mempertahankan daya saing di industri yang terus berkembang. Penggunaan metode K-NN sudah banyak digunakan, diantaranya K-NN dalam aktifitas manusia dalam prediksi kelulusan mahasiswa [4]. K-NN sebagai metode untuk melakukan Prediksi Kualitas Udara [5].

2. METODOLOGI

Pada Penelitian ini dilakukan untuk memprediksi pola pemesanan kamar dan harga yang paling diminati di Hotel Singaraja menggunakan metode *K-Nearest Neighbors* (KNN). Proses penelitian melibatkan beberapa tahapan yaitu metode pengumpulan data yang dilakukan dengan data penjualan kamar di Hotel Singaraja, sedangkan metode untuk memprediksi Tipe kamar yang diminati adalah dengan menggunakan metode *K-Nearest Neighbors* (KNN). Runtutan tahapan metode *K-Nearest Neighbors* (KNN) dijelaskan pada gambar 1 *flowchart* Prediksi KNN.



Gambar 1. Flowchart

2.1 Pencarian Data

Data yang didapatkan adalah data penjualan hotel di bulan Januari 2024 sampai November 2024. Di dalam data penjualan tersebut terdapat 7 kolom yang terdiri dari id_transaksi, tipe_kamar, Harga kamar, tanggal user_id, status, id_bill. data tersebut akan diolah lagi dengan mengurangi dan hanya menggunakan kolom yang diperlukan untuk melakukan prediksi bulanan dan tahunan dari penjualan tipe kamar di hotel tersebut.

id_transaksi	tipe_kamar	total	tgl	user_id	status	id_bill
H03329	AC DOBLE	180000	01/01/2024 15:12	2	1	1019
H0162	AC SINGLE	160000	01/01/2024 15:13	2	1	1020
H04905	AC DOBLE	180000	01/01/2024 15:13	2	1	1021
H01540	AC DOBLE	180000	01/01/2024 15:15	2	1	1022
H09200	AC SINGLE	160000	01/01/2024 15:16	2	1	1023
H08221	AC SINGLE	160000	01/01/2024 15:16	2	1	1024
H03210	STANDAT	120000	01/01/2024 15:17	2	1	1025
H09624	STANDAT	120000	01/01/2024 15:18	2	1	1026
H01729	AC SINGLE	160000	01/01/2024 15:19	2	1	1027

2.1 Tabel Data Hotel Singaraja

2.2 Normalisasi Data

Merupakan suatu kumpulan file data yang saling terhubung dan juga dikendalikan secara terpusat [6]. Karena KNN menggunakan jarak (misalnya, jarak Euclidean), penting untuk menormalkan data agar semua fitur memiliki skala yang sama, berikut adalah rumus dari Formula Normalisasi:

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}}$$

Gambar 2.2 Rumus Normalisasi

2.3 Menentukan Nilai K

Pilih nilai k (jumlah tetangga terdekat). Biasanya, nilai k kecil memberikan prediksi yang lebih sensitif, tetapi terlalu kecil dapat membuat model terlalu peka terhadap *noise*. Nilai k optimal sering dicari menggunakan teknik validasi silang.

2.4 Menghitung Jarak

Euclidean Distance merupakan metode perhitungan jarak antara dua titik dalam ruang. Euclidean yang diperkenalkan oleh matematikawan Yunani bernama Euclid [7]. Menggunakan Metrik Jarak untuk mengukur kedekatan data, rumusnya seperti ini:

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Gambar 2.3 Rumus Euclidian

2.5 Menentukan tetangga terdekat

Untuk mengurutkan seluruh data berdasarkan jarak terdekat ke data uji, sehingga data dengan jarak terpendek berada di urutan paling atas, setelah mengurutkan data, pilih K data teratas sebagai tetangga terdekat. Nilai K ditentukan berdasarkan kebutuhan spesifik dan dapat mempengaruhi akurasi model.

2.6 Prediksi

2.6.1 Klasifikasi

Mengambil mayoritas kelas dari tetangga terdekat berarti menentukan kelas suatu data dengan cara mengamati kelas dari sejumlah data lain yang berada paling dekat dengan data tersebut. Metode ini sering digunakan dalam algoritma *K-Nearest Neighbors* (KNN), di mana kelas yang paling banyak muncul di antara tetangga-tetangga terdekat data akan dipilih sebagai kelas data tersebut. Jumlah tetangga yang dipertimbangkan biasanya ditentukan oleh parameter K, dan proses ini melibatkan penghitungan jarak antara data yang ingin diklasifikasikan dengan data-data lainnya dalam ruang fitur. Sebuah matriks dari prediksi yang akan dibandingkan dengan kelas yang asli dari inputan atau dengan kata lain berisi informasi nilai aktual dan prediksi pada klasifikasi [8].

2.6.2 Regresi

Menghitung rata-rata nilai target dari tetangga terdekat adalah proses menentukan nilai prediksi untuk suatu data dengan cara menghitung nilai rata-rata dari nilai target data-data lain yang berada paling dekat dengan data tersebut dalam ruang fitur. Analisis regresi berusaha untuk memperkirakan hubungan antara variabel output dan satu set independen variabel input dengan secara otomatis belajar dari sejumlah akurasi sampel [9]. Prosedur ini sering digunakan dalam algoritma seperti *K-Nearest Neighbors* (KNN) untuk masalah regresi, di mana nilai prediksi dihasilkan berdasarkan rata-rata dari K data terdekat yang dipilih berdasarkan metrik jarak tertentu, seperti jarak Euclidean, Manhattan, atau lainnya. Mesin pembelajaran dapat menghasilkan prediksi atau regresi dengan menggunakan data citra yang diproses berdasarkan data yang telah dipelajari dalam proses training [10]. Jumlah tetangga yang digunakan, yang dikenal sebagai parameter K, memainkan peran penting dalam menentukan akurasi dan generalisasi model. Metode ini bekerja dengan asumsi bahwa data yang berdekatan dalam ruang fitur memiliki karakteristik atau nilai target yang mirip.

3. HASIL DAN PEMBAHASAN

3.1 Preprocessing

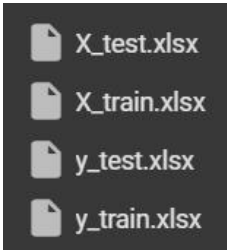
Tahap awal sebelum mengolah data tersebut adalah dengan preprocessing, proses tersebut merupakan tahapan yang harus dilakukan sebelum melakukan *data mining*. Tahap *Cleaning* dilakukan agar dapat menghasilkan data dengan variabel yang tertata dengan baik dan rapi ketika melakukan tahapan *data mining*. Proses yang dilalui adalah dengan menghapus kolom yang dianggap tidak relevan untuk analisis, lalu Mengonversi kolom tanggal menjadi format datetime agar bisa diekstrak informasi seperti tanggal dan bulan dan membagi bulan menjadi 3 periode.

tanggal	bulan	periode_bulanan	total	tipe_kamar_asli	tipe_kamar_prediksi
4	2	1	160000	AC SINGLE	AC SINGLE
26	3	0	160000	AC SINGLE	AC SINGLE
13	8	2	160000	AC SINGLE	AC SINGLE
21	8	0	160000	AC SINGLE	AC SINGLE
26	4	0	120000	STANDAT	STANDAT
27	7	0	180000	AC DOBLE	AC DOBLE
3	5	1	120000	STANDAT	STANDAT
6	7	1	120000	STANDAT	STANDAT
1	10	1	160000	AC SINGLE	AC SINGLE

Gambar 3.1 Tabel Data Yang Sudah Di Preprocessing

3.2 Split Data

Setelah data selesai melalui proses *preprocessing* tahapan berikutnya adalah dengan membagi data yang telah di *cleaning* menjadi 2 tipe file data yang akan digunakan untuk melatih model di *data mining* yaitu data uji dan data latih. Kedua data tersebut menggunakan parameter 20% data uji dan 80% data latih. Kedua data Tersebut dibagi lagi menjadi 4 data yang berbeda seperti pada gambar 3.2, data x menunjukkan data independen (fitur) yang berisi semua kolom kecuali kolom tipe kamar dan pada data y menunjukkan data target (label) yang berisi hanya kolom tipe kamar. Pembagian ini bertujuan untuk memastikan bahwa model yang dikembangkan dapat mempelajari pola dari data latih dan kemudian dievaluasi kinerjanya menggunakan data uji yang terpisah.



Gambar 3.2 File Split Data Uji dan Data Latih

3.3 Training Data

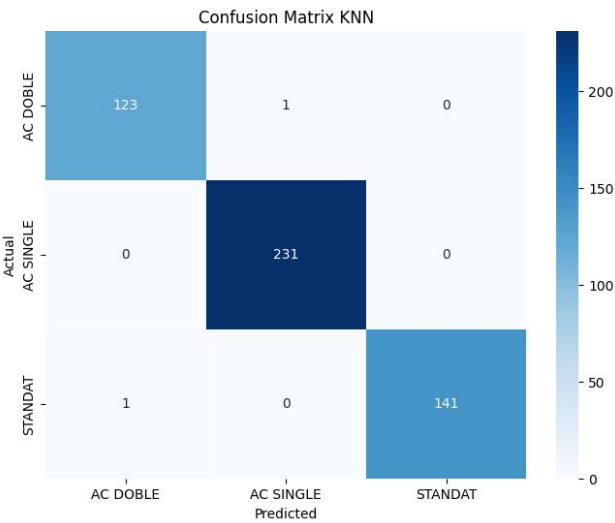
Setelah membagi dataset yang telah dibersihkan menjadi dua subset, yaitu data latih (80%) dan data uji (20%), langkah berikutnya adalah melatih model *K-Nearest Neighbors* (KNN). Training data adalah proses pelatihan data untuk mencari model yang tepat. Setelah mendapatkan 2 tipe data proses selanjutnya adalah dengan melatih model KNN menggunakan data latih. Model yang digunakan adalah model Euclidean agar dapat memprediksi atau membuat Keputusan dengan akurasi yang tepat.

tanggal	bulan	periode_bulanan	total	tipe_kamar_asli	tipe_kamar_prediksi
4	2	1	160000	AC SINGLE	AC SINGLE
26	3	0	160000	AC SINGLE	AC SINGLE
13	8	2	160000	AC SINGLE	AC SINGLE
21	8	0	160000	AC SINGLE	AC SINGLE
26	4	0	120000	STANDAT	STANDAT
27	7	0	180000	AC DOBLE	AC DOBLE
3	5	1	120000	STANDAT	STANDAT
6	7	1	120000	STANDAT	STANDAT
1	10	1	160000	AC SINGLE	AC SINGLE

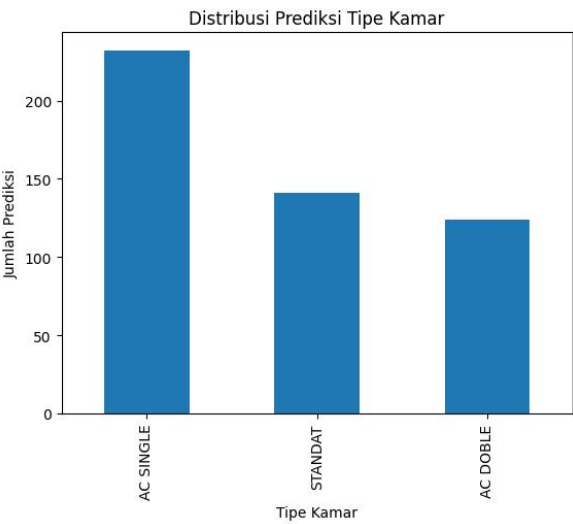
Gambar 3.3 Tabel Data Yang Sudah Di Training

3.4 Prediksi

Setelah model dilatih, algoritma digunakan untuk memprediksi pola pemesanan dan preferensi harga berdasarkan data historis. Prediksi menunjukkan bahwa tipe kamar tertentu dengan harga tertentu memiliki popularitas yang lebih tinggi pada periode tertentu. Algoritma KNN berhasil mengidentifikasi pola ini dengan tingkat akurasi mencapai 99.6%, yang dianggap cukup baik untuk diterapkan dalam pengambilan keputusan manajerial di hotel.



Gambar 3.41 Diagram Confusion Matrix



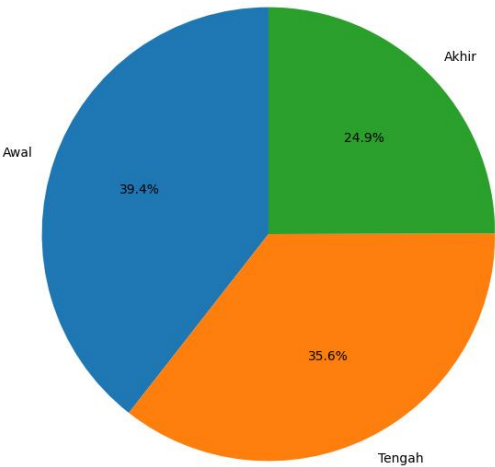
Gambar 3.42 Diagram Distribusi Prediksi Tipe Kamar

3.5 Analisis

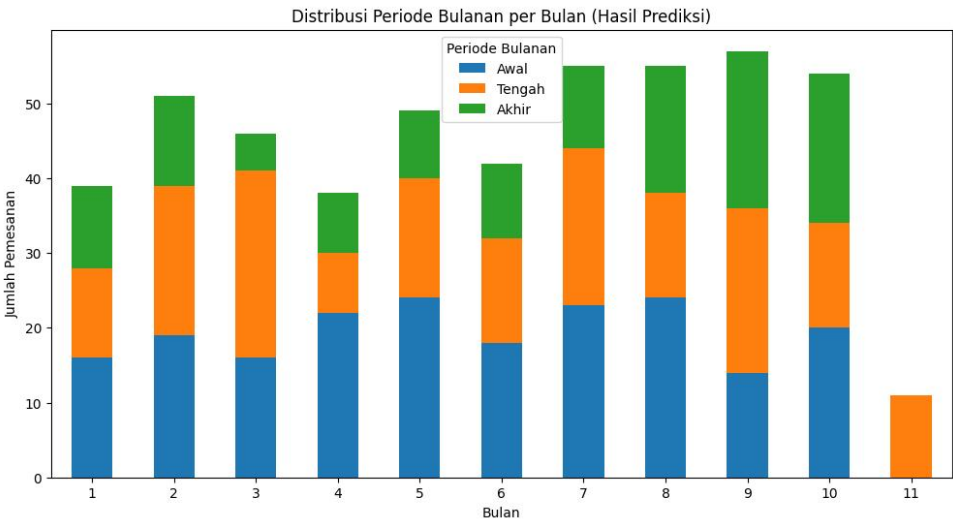
Proses terakhir setelah melakukan Prediksi terhadap tipe hotel adalah melakukan analisis hasil dari prediksi tersebut, analisis ini dibuat bedasarkan 2 waktu yaitu bulanan dan tahunan. Analisis bulanan menggunakan periode di dalam bulan seperti awal bulan, tengah bulan, dan akhir bulan. Setelah mendapatkan analisa terhadap tipe kamar yang laris di setiap bulan didapat hasil analisis tahunan dengan menampilkan grafik – grafik peningkatan maupun penurunan di semua tipe kamar bedasarkan bulanan dan tahunan. Dari gambar 3.51 dan gambar 3.52 dapat disimpulkan bahwa periode awal adalah periode yang sering untuk memesan kamar hotel dengan rincian setiap periode yaitu pada periode awal tipe kamar paling laris adalah AC SINGLE, periode tengah tipe kamar paling laris

adalah AC SINGLE, dan pada periode terakhir tipe kamar paling laris adalah AC SINGLE.

Distribusi Periode Bulanan Selama Setahun (Hasil Prediksi)



Gambar 3.51 Diagram Distribusi Periode Bulanan Selama Setahun



Gambar 3.52 Diagram Distribusi Periode Bulanan per Bulan

4. KESIMPULAN

Penelitian ini menggunakan metode K-Nearest Neighbors (KNN) untuk memprediksi dan menganalisis pola pemesanan tipe kamar yang paling diminati di Hotel Singaraja. Dengan memanfaatkan data historis, metode ini mampu mengidentifikasi pola yang paling diminati atau populer dengan tingkat akurasi yang baik. Proses penelitian melibatkan langkah-langkah seperti normalisasi data, penentuan nilai K, penghitungan jarak, pemilihan tetangga terdekat, serta klasifikasi atau regresi berdasarkan data tersebut. Hasil prediksi menunjukkan bahwa metode KNN memberikan manfaat signifikan, seperti membantu manajemen hotel dalam menentukan harga yang kompetitif, menyusun strategi promosi, serta meningkatkan efisiensi pengelolaan kamar. Analisis hasil prediksi dilakukan secara bulanan dan tahunan untuk mengidentifikasi tren pemesanan yang lebih rinci. Dengan demikian, metode ini menjadi alat yang efektif dalam pengambilan keputusan strategis di industri perhotelan.

5. DAFTAR RUJUKAN

[1] Abdul Ghani Muttaqin, Karina Auliasari, Febriana Santi Wahyuni., 2010. PENERAPAN METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI PENJUALAN BERBASIS WEB PADA PT.WIKA INDUSTRI ENERGY. *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 4 No. 2, pp.1-6

[2] Indahsari, D. K., & Kurniawan, Y. I., 2019. APLIKASI PREDIKSI USIA KELAHIRAN DENGAN METODE K-NEAREST NEIGHBOR. *Jurnal Kebidanan*, 11(01), pp.1 - 14.

[3] Yuli Mardi., 2019. Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *J. Edik Inform*, Vol. 2 No. 2, pp. 213–219.

[4] L. A. R. Hakim, A. A. Rizal, and D. Ratnasari., 2019. Aplikasi Prediksi Kelulusan Mahasiswa Berbasis K-Nearest Neighbor (K-NN). *jtim*, Vol. 1 No. 1, pp. 30-36.

[5] Adinda Amalia, Ati Zaidiah, Ika Nurlaili Isnainiyah., 2022. Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, Vol. 7 No. 2, pp. 496-507

[6] Muhammad Rafli Kusnaldi, Timotius Gulo, Soeb Aripin., 2022. Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal. *Journal of Computer System and Informatics (JoSYC)*, Vol. 3 No. 4, pp. 330-338

[7] Raihan Alya Shafira, Yahfizham Yahfizham, Aninda Muliani Harahap., 2023. MENENTUKAN JARAK TERPENDEK DALAM PENGIRIMAN BARANG DENGAN PERBANDINGAN EUCLIDEAN DISTANCE DAN MANHATTAN DISTANCE. *JOURNAL OF SCIENCE AND SOCIAL RESEARCH*, Vol. 6 No. 3, pp. 1-8

- [8] Durmuş Özkan Şahin, Erdal Kılıç., 2019. Two new feature selection metrics for text classification. *Automatika*, Vol. 60 No. 7, pp. 162-171
- [9] Muhammad Fadhlil Aziz, Sofi Defiyanti, Betha Nurina Sari., 2018. PERBANDINGAN ALGORITMA CART DAN K-NEAREST NEIGHBOR UNTUK PREDIKSI LUAS LAHAN PANEN TANAMAN PADI DI KABUPATEN KARAWANG. *Jurnal TAM (Technology Acceptance Model)*, Vol. 9 No. 2, pp. 74-78
- [10] Muhammad Alfathan Harriz, Harlis Setiyowati., 2023. KOMPARASI ALGORITMA DECISION TREE DAN KNN DALAM MENGLASIFIKASI DAERAH BERDASARKAN PRODUKSI LISTRIK. *JIKO (Jurnal Informatika dan Komputer)*, Vol. 7 No. 2, pp. 167-171