

Analisis Harga Rumah Dengan Metode *K-Means* Dan *Naïve Bayes*

Arman Prasojo Sugiyarto^{*1}, Nur Hayati², Eri Mardiani³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknologi Komunikasi dan Informatika,
Universitas Nasional, Jakarta, Indonesia

e-mail: ^{*1}armansugiyarto12@gmail.com; ²nurh4y@gmail.com; ³erimardiani1@gmail.com

Abstrak

Penelitian ini bertujuan untuk membandingkan kinerja metode K-Means dan Naïve Bayes dalam menganalisis harga rumah. Dataset yang digunakan adalah dataset harga rumah yang diperoleh dari hasil observasi. Penelitian ini dilakukan selama kurang lebih 2 bulan, dengan fokus pada penerapan algoritma K-Means dan Naïve Bayes. Data diolah dan dianalisis dengan menggunakan software Orange dan hasilnya ditampilkan dalam bentuk tabel dan grafik. Hasil analisis menunjukkan bahwa kinerja algoritma K-Means lebih baik daripada algoritma Naïve Bayes dengan nilai akurasi sebesar 30% untuk variabel y jarak fasilitas umum dan 22% untuk variabel y luas tanah serta 82% dengan perhitungan Naïve Bayes. Oleh karena itu, dapat disimpulkan bahwa metode K-Means adalah metode yang lebih efektif dalam menganalisis harga rumah.

Kata kunci— K-Means, Naïve Bayes, analisis harga rumah, software Orange

Abstract

This study aims to compare the performance of the K-Means and Naïve Bayes algorithms in analyzing house prices. The dataset used is a house price dataset obtained from observational results. The study was conducted for approximately 2 months, focusing on the implementation of the K-Means and Naïve Bayes algorithms. The data was processed and analyzed using Orange software, and the results were presented in tables and graphs. The analysis results showed that the K-Means algorithm outperformed the Naïve Bayes algorithm with an accuracy value of 30% for the variable y distance to public facilities and 22% for the variable y land area and 82% with Naïve Bayes calculation. Therefore, it can be concluded that the K-Means method is a more effective method for analyzing house prices.

Keywords— K-Means, Naïve Bayes, house price analysis, Orange software

1. PENDAHULUAN

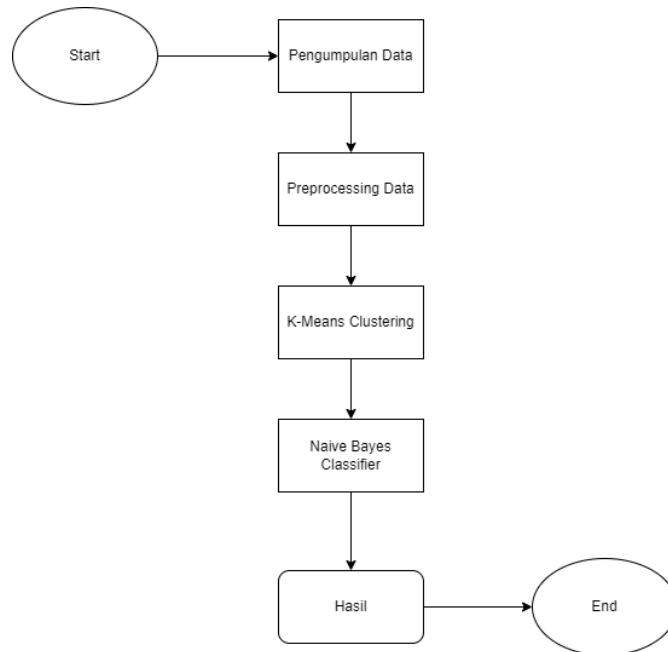
Seiring dengan kemajuan zaman, kebutuhan manusia telah berkembang secara signifikan. Namun ada beberapa kebutuhan yang sangat penting bagi manusia. Kebutuhan manusia yang harus dipenuhi kebutuhan pokok seperti sandang, pangan dan papan. Membeli rumah adalah salah satu keputusan paling penting yang harus diambil seseorang. Bagi calon pembeli rumah, pemahaman harga rumah sangat penting agar mereka dapat membuat keputusan yang bijak dan sesuai dengan kebutuhan serta anggaran mereka [1]. Selain itu, pertumbuhan ekonomi memegang peranan penting karena permintaan perumahan sering dianggap elastis terhadap pendapatan, sehingga menyebabkan peningkatan pendapatan rumah tangga [2]. Setiap orang membutuhkan rumah untuk tempat berlindung dan sebagai lokasi dimana orang berkumpul dan melakukan aktivitas rumah tangga. Rumah tidak hanya berfungsi sebagai tempat perlindungan dan istirahat, tetapi juga tempat interaksi keluarga. Tidak dapat disangkal bahwa membeli rumah adalah salah satu keputusan terpenting yang diambil orang dalam hidup. Harga rumah dapat bergantung pada berbagai faktor mulai dari lokasi rumah, fasilitas, hingga permintaan pasar [3][4]. Beberapa alasan untuk pertumbuhan eksponensial baru-baru ini adalah urbanisasi dan imigrasi yang mengakibatkan tingginya populasi kepadatan penduduk di perkotaan, khususnya yang memiliki industri teknologi[5].

Machine Learning merupakan aplikasi dari ilmu computer yang mampu mengolah banyak informasi dan mempelajari pola untuk memprediksi masa depan dengan menggunakan statistik dan matematika komputasi[6]. Dalam beberapa tahun terakhir, perkembangan teknologi dan ketersediaan data telah membuka peluang baru untuk menerapkan metode analisis data yang canggih, seperti K-Means dan Naïve Bayes, untuk memahami faktor-faktor utama yang mempengaruhi harga perumahan. Dalam industri perumahan, membantu calon pembeli rumah untuk memperkirakan harga rumah sebelum mengambil keputusan dan menjadi pedoman bagi pengembang untuk menentukan yang cocok harga rumah berdasarkan berbagai faktor seperti lokasi, ukuran rumah dan fasilitas sekitar[7].

Metode yang memisahkan sekumpulan data ke dalam beberapa kategori sesuai dengan karakteristik yang sama diidentifikasi sebelumnya adalah bentuk Clustering atau klasifikasi [8]. K-Means Clustering adalah metode clustering data non-hierarki yang mencoba mempartisi data yang ada menjadi satu atau lebih cluster/grup. K-Means dapat digunakan untuk mengidentifikasi tren data yang mengarah pada kisaran harga tertentu berdasarkan karakteristik properti[9]. Analisis cluster adalah suatu metode pengelompokan data (objek) yang didasarkan pada informasi itu ditemukan dalam data yang menggambarkan objek dan relasi di dalamnya[10]. Sedangkan metode Naïve bayes adalah algoritma yang menggabungkan probabilitas sebelumnya dan kondisional ke dalam rumus yang digunakan untuk menghitung probabilitas setiap kategori[11]. Harga rumah di Jakarta Selatan dapat dikelompokkan menjadi 10 cluster sesuai dengan nilai dbi terbaik yaitu 0.129[12]. Penelitian ini membandingkan metode K-Means dan Naïve Bayes dalam analisis harga perumahan untuk memberikan kontribusi positif terhadap pemahaman harga perumahan dan faktor-faktor yang mempengaruhi dengan menggunakan metode analisis data yang inovatif.

Berdasarkan latar belakang permasalahan yang dihadapi yaitu; 1) Bagaimana cara mengelompokkan harga perumahan secara optimal dan efisien dengan metode K-Means? 2) Bagaimana cara menganalisis dan memprediksi pengaruh masing-masing factor dengan metode Naïve Bayes?. Penelitian ini membatasi permasalahan, yaitu; 1) Mencakup kabupaten dan kota di wilayah Tangerang, Tangerang Selatan, Depok dan Bogor 2) Menentukan kriteria segmentasi yang digunakan dalam analisis seperti lokasi, luas tanah, luas bangunan, jumlah kamar tidur, fasilitas atau fitur lain yang relevan.

2. METODE PENELITIAN

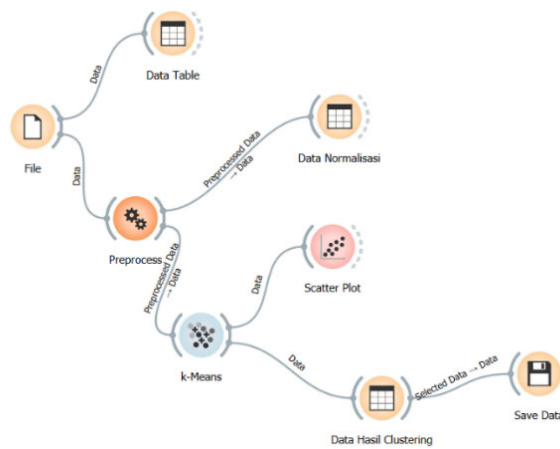


Gambar 1 Tahapan penelitian

A. Pengumpulan Data

Pada penelitian ini, peneliti menggunakan metode pengumpulan data dengan observasi ke masing-masing perumahan. Setiap kunjungan mendapatkan brosur untuk pengumpulan data yang diperlukan. Data primer yang dihasilkan dari observasi sebanyak 152 tipe perumahan.

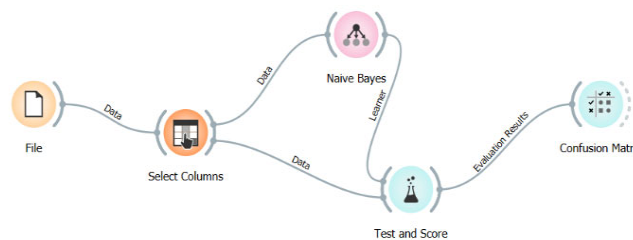
B. K-Means Clustering



Gambar 2 Flowchart K-Means Clustering

Gambar di atas menjelaskan *flowchart K-Means Clustering* dimana tahap pertama yaitu data primer selanjutnya data preprocessing untuk mengubah data yang belum diproses menjadi format yang efisien, scatter plot sebagai pola hubungan antara dua variable di lanjut dengan K-Means Clustering setelah itu data hasil clustering. Analisis pengelompokan atau Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain[13]. K-Means Clustering bekerja dengan menggunakan nilai dari Text Mining yang ada dan akan dihitung jarak terdekat dengan menggunakan Ecludience Distance berdasarkan Centroid yang sudah diatur, yang menjadi acuan pada Cluster atau kelas manakah data yang tersedia[14].

C. Naïve Bayes



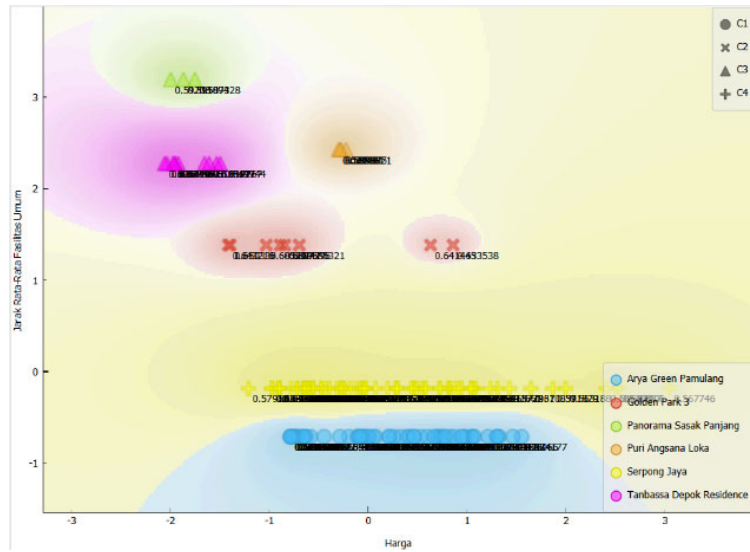
Gambar 3 *Flowchart Naïve Bayes*

Gambar di atas menjelaskan *flowchart Naïve Bayes* Dimana tahap pertama yaitu *select columns* untuk me-target kolom selanjutnya *Naïve Bayes* di lanjut dengan test and score untuk melihat hasil *area under ROC curve*, *classification accuracy*, *F1 Score*, *presicion*, *recall* dan *matthews correlation coefficient*. Setelah itu *confusion matrix* model prediktif yang dapat digunakan untuk memberikan metrik evaluasi yang menunjukkan dan membandingkan nilai riil atau nilai sebenarnya dengan nilai yang diharapkan oleh model. Pemilihan algoritma *Naïve Bayes* ini dapat digunakan untuk melakukan prediksi probabilitas keanggotaan suatu kelas[15].

3. HASIL DAN PEMBAHASAN

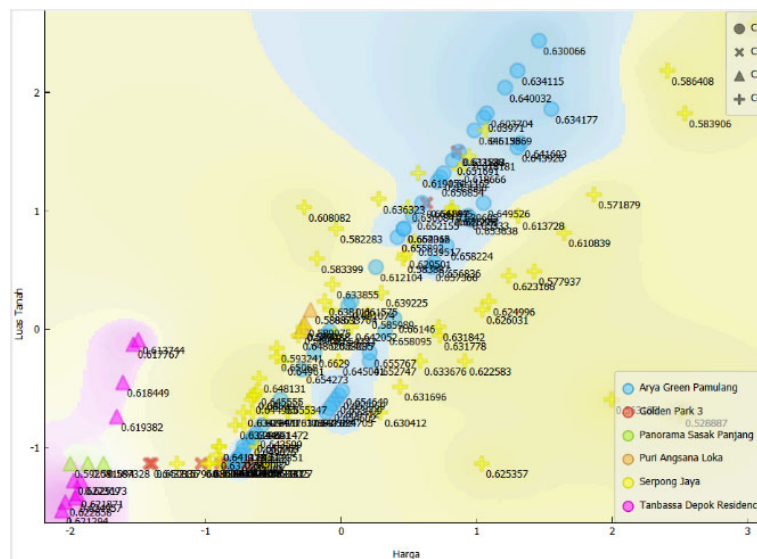
A. Analisa K-Means

Pengelompokan dengan algoritma *K-means* pada penelitian ini menggunakan *software Orange*.



Gambar 4 Scatter Plot Jarak Rata-Rata Fasilitas Umum

Pada gambar di atas menjelaskan dimana *Scatter Plot* tersebut berbentuk plot pencar dengan sumbu x sebagai harga rumah dan sumbu y sebagai jarak rata-rata fasilitas umum. *Scatter plot* memberikan gambaran bahwa semakin mahal harga rumah, maka jarak fasilitas umum semakin dekat. Sedangkan semakin murah harga rumah, maka jarak fasilitas umum semakin jauh.



Gambar 5 Scatter Plot Luas Tanah

Pada gambar di atas menjelaskan dimana *Scatter Plot* tersebut berbentuk plot pencar dengan sumbu x sebagai harga rumah dan sumbu y sebagai luas tanah. *Scatter plot* memberikan gambaran bahwa semakin mahal harga rumah, maka luas tanah semakin besar. Sedangkan semakin murah harga rumah, maka luas tanah semakin kecil.

	Tipe	Kamar Tidur	Kamar Mandi	Cluster (1)	Silhouette	Daerah	Perumahan	Cluster
1	Lotus	?	?	C4	0.636097	Tangerang Sela...	Serpong Jaya	The Garden
2	Lotus	?	?	C4	0.645555	Tangerang Sela...	Serpong Jaya	The Garden
3	Lotus	?	?	C4	0.64454	Tangerang Sela...	Serpong Jaya	The Garden
4	Lotus	?	?	C4	0.644865	Tangerang Sela...	Serpong Jaya	The Garden
5	Lotus	?	?	C4	0.634769	Tangerang Sela...	Serpong Jaya	The Garden
6	Lotus	?	?	C4	0.633445	Tangerang Sela...	Serpong Jaya	The Garden
7	Lotus	?	?	C4	0.637208	Tangerang Sela...	Serpong Jaya	The Garden
8	Lotus	?	?	C4	0.638111	Tangerang Sela...	Serpong Jaya	The Garden
9	Nora Sudut	3	2	C4	0.648862	Tangerang Sela...	Serpong Jaya	The Garden
10	Nora Sudut	3	2	C4	0.648488	Tangerang Sela...	Serpong Jaya	The Garden
11	Nora Sudut	3	2	C4	0.64961	Tangerang Sela...	Serpong Jaya	The Garden
12	Nora Sudut	3	2	C4	0.639572	Tangerang Sela...	Serpong Jaya	The Garden
13	Nora Sudut	3	2	C4	0.636323	Tangerang Sela...	Serpong Jaya	The Garden
14	Nora Khusus	3	2	C4	0.592241	Tangerang Sela...	Serpong Jaya	The Garden
15	Nora Khusus	3	2	C4	0.641373	Tangerang Sela...	Serpong Jaya	The Garden
16	Nora Khusus	3	2	C4	0.648131	Tangerang Sela...	Serpong Jaya	The Garden
17	Nora	3	2	C4	0.638445	Tangerang Sela...	Serpong Jaya	The Garden
18	Nora	3	2	C4	0.65066	Tangerang Sela...	Serpong Jaya	The Garden
19	Nora	3	2	C4	0.641131	Tangerang Sela...	Serpong Jaya	The Garden
20	Canna	2+1	2+1	C4	0.640701	Tangerang Sela...	Serpong Jaya	The Spring
21	Canna	2+1	2+1	C4	0.636084	Tangerang Sela...	Serpong Jaya	The Spring
22	Lantana	3+1	2+1	C4	0.642583	Tangerang Sela...	Serpong Jaya	The Spring
23	Lantana	3+1	2+1	C4	0.630605	Tangerang Sela...	Serpong Jaya	The Spring
24	Maple	3+1	3+1	C4	0.613728	Tangerang Sela...	Serpong Jaya	The Spring
25	Aster	2+1	2+1	C4	0.645041	Tangerang Sela...	Serpong Jaya	The Spring
26	Standart 3 LT	?	?	C4	0.663188	Tangerang Sela...	Serpong Jaya	Orlin Arcade
27	Sudut 3 LT	?	?	C4	0.528887	Tangerang Sela...	Serpong Jaya	Orlin Arcade 2
28	Acacia	?	?	C4	0.629501	Tangerang Sela...	Serpong Jaya	The Terrace
29	Acacia	?	?	C4	0.618666	Tangerang Sela...	Serpong Jaya	The Terrace
30	Acacia	?	?	C4	0.621278	Tangerang Sela...	Serpong Jaya	The Terrace
31	Hazel	3	2	C4	0.637351	Tangerang Sela...	Serpong Jaya	The Height
32	Rose	2	2	C4	0.633855	Tangerang Sela...	Serpong Jaya	The Height
33	Rose	2	2	C4	0.629891	Tangerang Sela...	Serpong Jaya	The Height
34	Rose	2	2	C4	0.624704	Tangerang Sela...	Serpong Jaya	The Height
35	Lily	4+1	2+1	C4	0.631842	Tangerang Sela...	Serpong Jaya	The Height
36	Pointsetia	2	2	C4	0.629473	Tangerang Sela...	Serpong Jaya	The Groov'ist

Gambar 6 Hasil Clustering

Pada gambar di atas merupakan hasil clustering dengan metode *K-Means* menghasilkan 4 cluster. Berikut terdapat hasil DBI pada *software Orange* dengan variabel x yang merupakan harga rumah dan variabel y yang merupakan jarak rata-rata fasilitas umum. Berdasarkan perhitungan SSB dan Rasio, dihasilkan DBI sebesar 0,30091 atau 30%. Jika DBI mendekati 0, maka hasil lebih baik. Sehingga hasil DBI yang merupakan jarak rata-rata fasilitas umum dapat dikatakan sudah valid karena mendekati 0 yaitu 0,30091 atau 30%.

Selanjutnya terdapat hasil DBI pada *software Orange* dengan variabel x yang merupakan harga rumah dan variabel y yang merupakan luas tanah. Berdasarkan perhitungan SSB dan Rasio, dihasilkan DBI sebesar 0,22426 atau 22%. Jika hitungan DBI mendekati 0, maka hasil lebih baik. Sehingga hasil DBI variabel y yang merupakan luas tanah dapat dikatakan sudah valid karena mendekati 0 yaitu 0,22426 atau 22%.

B. Analisa Naïve Bayes

Klasifikasi Analisa dengan algoritma *Naïve Bayes* pada penelitian ini menggunakan *software Orange*, mendapatkan hasil yakni: *area under ROC curve*, *classification accuracy*, *F1 score*, *presicion*, *recall* dan *matthews correlation coefficient*. Pada uji dataset dengan algoritma *Naïve Bayes*, digambarkan berupa *Confusion Matrix*, dari gambar ini didapatkan hasil klasifikasi.

		Predicted							Σ
		RSUD ...	SPBU, ...	Sinarm...	Sinarm...	St.Raw...	Stasiun...	Toll ...	
Actual	RSUD ...	10	0	0	0	0	0	0	10
	SPBU, ...	0	4	0	0	0	0	0	4
	Sinarm...	0	0	6	1	0	0	0	7
	Sinarm...	0	0	1	0	0	0	0	1
	St.Raw...	0	0	0	1	70	0	0	71
	Stasiun...	0	0	0	0	0	3	0	3
	Toll ...	0	0	0	0	0	0	56	56
Σ		10	4	7	2	70	3	56	152

Gambar 7 Confusion Matrix Naïve Bayes

Pada gambar diatas menjelaskan *Confusion Matrix* berbentuk matriks dengan ukuran 7 x 7 yang mengkategorikan masing-masing kelas klasifikasi menjadi RSUD, SPBU, Sinarmas, Sinarmas, St. Rawa, Stasiun dan Toll. Masing-masing kelas yang terdapat pada kondisi kelas sebenarnya, nilai TP sebanyak 70 data untuk kelas St.Rawa dan nilai TN sebanyak 56 data untuk Toll.

Model	AUC	CA	F1	Prec	Recall	MCC
Naive Bayes	0.989	0.980	0.984	0.987	0.980	0.969

Gambar 8 Hasil *Test and Score*

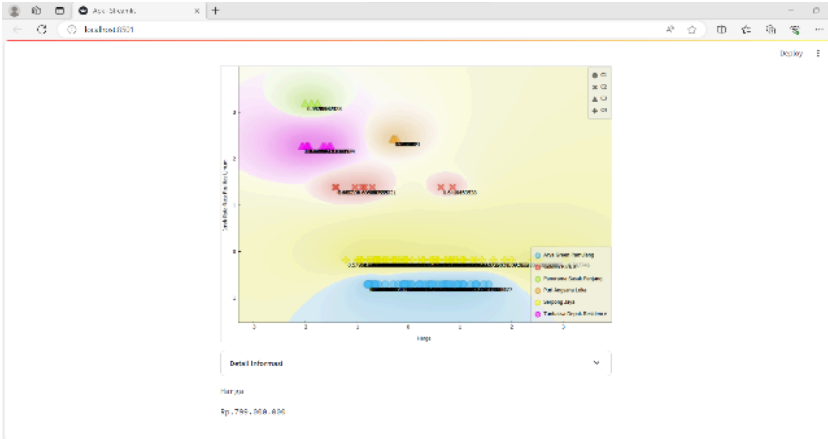
Gambar diatas menunjukkan keakuratan pada algoritma yang digunakan dalam memprediksi dataset yang di uji dengan tepat. Oleh karena itu dapat diketahui keakuratan pada kinerja algoritma *Naïve Bayes* untuk mendapatkan informasi yang dibutuhkan pada nilai yang dihasilkan oleh sistem. Nilai akurasi didapatkan dari *software Orange* sebesar 98%. Nilai akurasi ini menunjukkan keakuratan pada algoritma yang digunakan dalam memprediksi dataset yang di uji dengan tepat.

C. Aplikasi

Hasil dari *deployment* algoritma *K-Means Clustering* dan *Naïve Bayes* untuk analisis harga rumah menggunakan *software Orange* yaitu berupa aplikasi menggunakan *framework* dari Streamlit berupa *website*.

No	Tipe	Lokasi	Luas Lantai	Jumlah Kamar	Jumlah Kamar	Milik	dan Fasilitas
1	Rusun	+	+	+	+	RSU001	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
2	Rusun	+	+	+	+	RSU002	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
3	Rusun	+	+	+	+	RSU003	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
4	Rusun	+	+	+	+	RSU004	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
5	Rusun	+	+	+	+	RSU005	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
6	Rusun	+	+	+	+	RSU006	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
7	Rusun	+	+	+	+	RSU007	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
8	Rusun	+	+	+	+	RSU008	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
9	Rusun	+	+	+	+	RSU009	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
10	Rusun	+	+	+	+	RSU010	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
11	Rusun	+	+	+	+	RSU011	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
12	Rusun	+	+	+	+	RSU012	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
13	Rusun	+	+	+	+	RSU013	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
14	Rusun	+	+	+	+	RSU014	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
15	Rusun	+	+	+	+	RSU015	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
16	Rusun	+	+	+	+	RSU016	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
17	Rusun	+	+	+	+	RSU017	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
18	Rusun	+	+	+	+	RSU018	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
19	Rusun	+	+	+	+	RSU019	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
20	Rusun	+	+	+	+	RSU020	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
21	Rusun	+	+	+	+	RSU021	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
22	Rusun	+	+	+	+	RSU022	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
23	Rusun	+	+	+	+	RSU023	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
24	Rusun	+	+	+	+	RSU024	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
25	Rusun	+	+	+	+	RSU025	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
26	Rusun	+	+	+	+	RSU026	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
27	Rusun	+	+	+	+	RSU027	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
28	Rusun	+	+	+	+	RSU028	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
29	Rusun	+	+	+	+	RSU029	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
30	Rusun	+	+	+	+	RSU030	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
31	Rusun	+	+	+	+	RSU031	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
32	Rusun	+	+	+	+	RSU032	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
33	Rusun	+	+	+	+	RSU033	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
34	Rusun	+	+	+	+	RSU034	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
35	Rusun	+	+	+	+	RSU035	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.
36	Rusun	+	+	+	+	RSU036	Rusun M. Sdk., Terpetak Dg., Terpetak Dg.

Gambar 9 Tampilan Tipe Rumah



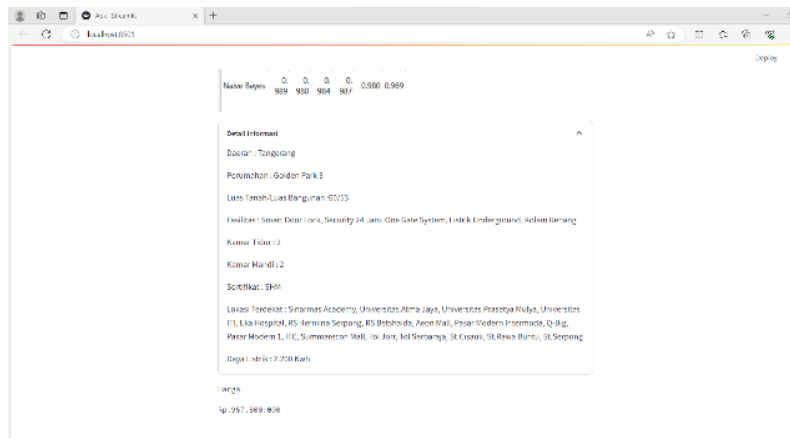
Gambar 10 Tampilan Analisis K-Means

Gambar di atas merupakan *interface* dari pemilihan tipe rumah berupa *dropdown* beserta hasil dari analisis berdasarkan tipe rumah yang sudah melalui tahap algoritma *K-Means* dan *Naïve Bayes*.

Actual \ Predicted	RSUD ...	SPBU ...	Sinarm...	Sinarm...	SLRaw...	Stasiun...	Toll ...	Σ
RSUD ...	10	0	0	0	0	0	0	10
SPBU ...	0	4	0	0	0	0	0	4
Sinarm...	0	0	6	1	0	0	0	7
Sinarm...	0	0	1	0	0	0	0	1
SLRaw...	0	0	0	1	70	0	0	71
Stasiun...	0	0	0	0	0	3	0	3
Toll ...	0	0	0	0	0	0	56	56
Σ	10	4	7	2	70	3	56	152

Model	AUC	CA	F1	Rec	Recall	MCC
Naive Bayes	0.959	0.959	0.954	0.957	0.959	0.959

Gambar 11 Tampilan Analisis Naïve Bayes

Gambar 12 Tampilan *Dropdown* Detail Informasi

Gambar di atas merupakan *interface* dari detail informasi rumah seperti daerah, nama perumahan, luas tanah, luas bangunan, fasilitas, jumlah kamar tidur, jumlah kamar mandi, sertifikat, lokasi terdekat dan daya listrik.

4. KESIMPULAN

Pada penelitian ini dapat di tarik kesimpulan pertama, kinerja algoritma K-Means lebih baik dari pada algoritma Naïve Bayes, dapat di lihat dari tingkat akurasi nilai K-Means dengan variabel x harga rumah mempengaruhi variabel y jarak rata-rata fasilitas umum sebesar 30% dan variabel y luas tanah sebesar 22%. sedangkan Naïve Bayes sebesar 98%.

Kedua, semakin mahal harga rumah maka jarak fasilitas umum semakin dekat. Begitu pula dengan luas tanah yang semakin besar maka harga rumah semakin mahal. Dari penelitian ini, disimpulkan bahwa calon pembeli rumah dapat mempertimbangkan jarak fasilitas umum yang dekat dan luas tanah yang besar akan mempengaruhi harga rumah yang semakin mahal.

Harapan untuk penelitian selanjutnya agar lebih banyak lagi menambahkan jumlah dataset rumah agar hasil jauh lebih baik lagi dan menggunakan algoritma yang lain untuk lebih bervariasi, serta pengembangan aplikasi yang masih sederhana ini dapat dikembangkan.

DAFTAR PUSTAKA

- [1] Wang, L., Wang, G., Yu, H., & Wang, F., "Prediction and analysis of residential house price using a flexible spatiotemporal model", 25, 503-522, 2022. doi:<https://doi.org/10.1080/15140326.2022.2045466>© 2022 The Author(s).
- [2] Reddy, K., Athelli, T., & Kulsum, S., Journal Of Emerging Technologies And Innovative Research (JETIR), 9, 2022. Retrieved from <http://www.jetir.org/>
- [3] Putri, V. A., Prasetijo, A. B., & Eridani, D., "Perbandingan Kinerja Algoritme Naive Bayes dan K-Nearest Neighbor (KNN) Untuk Prediksi Harga Rumah", 162-171, 2022. doi:10.14710/transmisi.24.4.162-171
- [4] Aji, B. G., Sondawa, D. C., Gifari, M. R., & Wijayanto, S., "Penerapan Algoritma K-Means Untuk Clustering Harga Rumah Di Bandung", 14, 17-23, 2023.
- [5] Gude, V. , "A Multi-Level Modeling Approach For Predicting Real-Estate Dynamics", May2023. doi:<http://dx.doi.org/10.1108/IJHMA-02-2023-0024>

-
- [6] Yunardi, R. T., & Dina, N. Z., "Data Mining dan Machine Learning dengan Orange3 Tutorial dan Aplikasinya", In A. Abadi (Ed.). Surabaya, Jawa Timur: Airlangga University Press., 2022.
- [7] Teoh, E. Z., Yau, W. C., Ong, T. S., & Connie, T., "Explainable Housing Price Prediction With Determinant Analysis", 16(5), 1021-1045, July 4, 2022, . doi:10.1108/IJHMA-02-2022-0025
- [8] Sulistiyawati, A., & Supriyanto, E., "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan", 15, 25-36, 2021. Retrieved from <https://doi.org/10.33365/jtk.v15i2.1162>
- [9] Aditya, A., Jovian, I., & Sari, B. N., "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah", 4, 51-88, Januari, 2020. doi:10.30865/mib.v4i1.1784
- [10] Ningrat, Desy Rahmawati., Maruddani, Di Asih I., & Wuryandari, Triastuti, "Analisis Cluster dengan Algoritma K-Means dan Fuzzy C-Means Clustering untuk Pengelompokan Data Obligasi Korporasi", 5(4), 641-650, 2016.
- [11] Rejeki, F., & Ayumi, V., "Analisa Sentimen Mengenai Kenaikan Harga BBM Menggunakan Metode Naive Bayes Dan Support Vector Machine", 6, 1-10, Januari, 2023. doi:10.36085
- [12] Septiani, N., Anwar, S., & Herdiana, R., "Penerapan Algoritma K-Means Clustering Untuk Harga Rumah" , 35-47, 2023.
- [13] Asminah, "Sistem Penentuan Penambahan Koleksi Buku di Perpustakaan menggunakan Metode K-Means Clustering", 4(1), 330-338, 2022.
- [14] Sutikno, W., Palandi, J.F., dan Octavia, C.A., "Klasifikasi E-book Berbahasa Inggris Dengan Menggunakan Metode K-Means Clustering Studi Kasus Perpustakaan STIKI Malang", 7(1), 80-85, 2019.
- [15] Cahyo, Arisa Dwi. , "Metode Naive Bayes untuk Klasifikasi Masa Studi Sarjana", 3(4). 1-19, 2023.