

**“ANALISIS *CLUSTERING* KRIMINALITAS BERDASARKAN
TINGKAT PENDIDIKAN WARGA BINAAN PEMASYARAKATAN
(WBP) MENGGUNAKAN METODE *AGGLOMERATIVE
HIERARCHICAL CLUSTERING* (AHC)“**

PROPOSAL



Disusun oleh :

MARHAMA M.BABURAHMAN SAIFUDDIN
NIM : 121055520124118

**PRODI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALUKU UTARA
TAHUN 2025**

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Puji syukur saya panjatkan ke hadirat Allah SWT atas segala rahmat, karunia, serta petunjuk- Nya sehingga saya dapat menyelesaikan penyusunan proposal ini dengan judul “Analisis Clustering Kriminalitas Berdasarkan Tingkat Pendidikan Warga Binaan Pemasyarakatan (WBP) Menggunakan Metode *Agglomerative Hierarchical Clustering* (AHC) “. Tak lupa shalawat serta salam senantiasa tercurah kepada baginda Nabi Muhammad SAW, yang telah menjadi suri teladan bagi umat manusia.

Proposal ini disusun sebagai langkah awal dalam penyusunan skripsi yang akan menjadi salah satu syarat kelulusan pada program studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Maluku Utara. Dalam perjalanan penyusunan proposal ini, saya mengalami berbagai macam tantangan dan hambatan, namun dengan bimbingan, dukungan, dan dorongan dari berbagai pihak, saya berhasil menyelesaikannya.

Saya mengucapkan terima kasih yang tak terhingga kepada Bapak **Sahriar Hamza, S.Kom, M.T** selaku Pembimbing I, dan Bapak **Agil Assagaf, S.T., M.Kom**, selaku Pembimbing II yang telah memberikan masukan, serta bimbingan yang sangat berharga dalam menyusun proposal ini.

Ucapan terima kasih juga saya sampaikan kepada keluarga saya yang senantiasa memberikan doa, dukungan, serta motivasi dalam menyelesaikan studi dengan baik. Saya juga ingin mengucapkan terima kasih kepada sahabat

seperjuangan yang selalu memberikan semangat dan dukungan.

Akhir kata, saya menyadari bahwa proposal ini jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang membangun sangat saya harapkan untuk perbaikan di masa yang akan datang. Semoga proposal ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan di bidang Teknologi Informasi
Wassalamualaikum Warahmatullahi Wabarakatu.

Ternate, 04 Oktober 2025

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN	ii
LEMBAR PERNYATAAN KEASLIAN	ii
HALAMAN PERSEMBERAHAN	Error! Bookmark not defined.
KATA PENGANTAR.....	Error! Bookmark not defined.
DAFTAR ISI	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
ABSTRAK.....	Error! Bookmark not defined.
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	
2.1. Penelitian Terkait.....	6
2.2. Data <i>Mining</i>	14
2.3. Analisis <i>Clustering</i>	14
2.4. <i>K-Means</i>	15
2.5. Pengujian <i>Silhouette Coeffisient</i>	17
2.6. Tingkat Pendidikan.....	19
2.7. Kriminalitas dan Pemasyarakatan.....	19
2.8. Warga Binaan Pemasyarakatan (WBP).....	21
2.9. Rumah Tahanan Negara (Rutan) Kelas II Ternate	21
2.10. <i>Website</i>	22
2.11. <i>Hyper Text Mark Up Language (HTML)</i>	Error! Bookmark not defined.
2.12. <i>Hypertext Preprocessor (PHP)</i>	Error! Bookmark not defined.

2.14. <i>Prototype</i>	Error! Bookmark not defined.
2.15. <i>Unifield Modeling language (UML)</i>	Error! Bookmark not defined.
2.15.1. <i>Use Case Diagram</i>	Error! Bookmark not defined.
2.15.2. <i>Diagram Aktivitas (Activity Diagram)</i> ...	Error! Bookmark not defined.
2.16. Metode Penelitian	Error! Bookmark not defined.
2.17. <i>White Box</i>	23
2.18. <i>Flowchart</i>	24
BAB III METODE PENELITIAN	26
3.1. Objek dan Waktu Penelitian	26
3.2. Sifat, Jenis dan Pendekatan Penelitian.....	26
3.3. Metode Pengumpulan Data.....	Error! Bookmark not defined.
3.4. Alat dan Bahan Penelitian	Error! Bookmark not defined.
3.4.1. Perangkat Keras (<i>Hardware</i>)	Error! Bookmark not defined.
3.4.2. Perangkat Lunak (<i>Software</i>)	Error! Bookmark not defined.
3.5. Pengambangan Perangkat Lunak.....	Error! Bookmark not defined.
3.6. Analisis Masalah.....	Error! Bookmark not defined.
3.8. Sistem yang Diusulkan	Error! Bookmark not defined.
3.9. <i>Use Case Diagram</i>	Error! Bookmark not defined.
3.10. Perancangan <i>Database</i>	Error! Bookmark not defined.
3.11. Desain Antarmuka	Error! Bookmark not defined.
3.12. Diagram Alir.....	26
3.13. Uraian Diagram Alir	Error! Bookmark not defined.
3.14. Implementasi Metode <i>K-Means</i>	Error! Bookmark not defined.
3.15. Pengujian <i>Silhouette Coefficient</i>	Error! Bookmark not defined.
3.16. Perencanaan Pengujian Sistem	Error! Bookmark not defined.

DAFTAR PUSTAKA

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Model <i>Prototype</i> (Titania).....	25
Gambar 3.1 Pemodelan <i>Prototype clustering</i> kriminalitas	33
Gambar 3.2 Sistem yang diusulkan.....	35
Gambar 3.3 <i>Usecase</i> diagram	36
Gambar 3.4 Tampilan <i>login</i>	37
Gambar 3.5 Tampilan tambah data kriteria.....	37
Gambar 3.6 Tampilan perhitungan dan hasil	38
Gambar 3.7 Diagram Alir Penelitian	38

DAFTAR TABEL

	Halaman
Tabel 2.1 Penelitian terkait	6
Tabel 2.2 Simbol-simbol <i>Use case</i>	26
Tabel 2.3 Simbol-simbol <i>Activity</i> diagram.....	27
Tabel 2.4 Simbol <i>Flowchart</i> (Khesya, 2021).....	29
Tabel 3.1 Detail Spesifikasi <i>Hardware</i>	32
Tabel 3.2 Spesifikasi <i>Software</i>	33
Tabel 3.3 Data yang akan diolah.....	39
Tabel 3.4 Inisialisasi Kriteria tingkat pendidikan	40
Tabel 3.5 Inisialisasi Kriteria jenis kejahatan	40
Tabel 3.6 Data yang sudah ditransformasi.....	40
Tabel 3.7 Pusat <i>cluster</i> (<i>centroid</i>) awal.....	42
Tabel 3.8 Hasil perhitungan jarak	43
Tabel 3.9 Interpretasi Nilai <i>Silhouette Coefficient</i>	45

BAB I

PENDAHULUAN

1.1. Latar Belakang

Tindak kriminalitas berpengaruh terhadap keamanan masyarakat serta mengancam ketenteraman lahir dan batin. Apabila masyarakat merasa keamanannya terancam, besar kemungkinan hal tersebut berdampak pada kesejahteraan dan ketenteraman mereka. Kriminalitas merupakan segala bentuk tindakan yang merugikan secara ekonomis maupun psikologis serta melanggar hukum yang berlaku di Indonesia dan norma-norma sosial serta agama (Suriani, 2020).

Kriminalitas merupakan permasalahan yang kerap terjadi dalam kehidupan sehari-hari. Tingkat kriminalitas yang tinggi dapat mengancam keamanan dan ketertiban masyarakat (Pakdosen, 2023). Karena itu, penting untuk melakukan analisis dan pengelompokan (clustering) terhadap data kriminalitas guna memahami pola dan karakteristik kejahatan yang terjadi.

Salah satu faktor yang dapat memengaruhi tingkat kriminalitas adalah tingkat pendidikan. Tingkat pendidikan berkaitan dengan pembentukan perilaku dan kesadaran masyarakat terhadap hukum dan aturan (Putra, 2021). Berangkat dari hal tersebut, muncul pertanyaan: sejauh mana hubungan tingkat pendidikan dengan perilaku kriminal yang dilakukan warga binaan pemasyarakatan (WBP), Melalui penelitian ini diharapkan muncul wawasan yang lebih mendalam tentang pola kriminalitas di lingkungan lembaga pemasyarakatan, dengan fokus pada tingkat pendidikan para narapidana.

Analisis pengelompokan dengan Agglomerative Hierarchical Clustering (AHC) digunakan sebagai alat utama untuk menelaah apakah ada pola perilaku tertentu yang dapat diidentifikasi berdasarkan tingkat pendidikan WBP. AHC merupakan metode pengelompokan hierarkis bottom-up: setiap observasi mulanya dianggap sebagai satu cluster, lalu pasangan cluster yang paling mirip digabungkan secara bertahap hingga terbentuk struktur hierarki. Hasilnya dapat divisualisasikan melalui dendrogram, sehingga peneliti dapat menentukan jumlah cluster yang tepat dengan memotong dendrogram pada tingkat tertentu. Proses penggabungan ditentukan oleh ukuran jarak dan kriteria pengaitan (linkage) yang relevan (misalnya single, complete, average, atau Ward).

Selain itu, pengujian kualitas clustering penting dilakukan untuk memastikan hasil yang valid. Salah satu metrik yang digunakan adalah Silhouette Coefficient, yang memadukan konsep cohesion dan separation untuk mengukur seberapa baik suatu data berada pada cluster yang tepat. Dengan Silhouette, dapat dievaluasi kualitas dan ketegasan batas antarkluster yang dihasilkan AHC.

Sejumlah penelitian terdahulu menganalisis kriminalitas dengan pendekatan K-Means (misalnya Imam Zuhdi Muzakkiy, 2023; Nadya Septiani, 2020). Namun, kajian yang secara khusus menganalisis pengelompokan kriminalitas berdasarkan tingkat pendidikan WBP di Rutan Kelas II Ternate dengan pendekatan AHC masih sangat terbatas. Pemanfaatan AHC yang menyediakan struktur hierarki dan visualisasi dendrogram diharapkan memberi sudut pandang baru sekaligus melengkapi temuan dari pendekatan non-hierarkis sebelumnya. Penggunaan *Silhouette Coefficient* sebagai uji kualitas akan semakin memastikan kevalidan hasil

analisis.

Dengan melakukan analisis clustering kriminalitas berdasarkan tingkat pendidikan WBP di Rutan Kelas II Ternate menggunakan *Agglomerative Hierarchical Clustering* dan pengujian menggunakan *Silhouette Coefficient*, penelitian ini diharapkan memberikan pemahaman yang lebih mendalam mengenai hubungan antara tingkat pendidikan dan kejahatan. Hasilnya juga diharapkan menjadi acuan bagi pengembangan kebijakan dan program pencegahan kriminalitas di lingkungan pemasyarakatan.

Penelitian akan mengambil contoh kasus dari Rutan Kelas II Ternate, sebuah lembaga pemasyarakatan yang mengelola narapidana di wilayah Kota Ternate. Data yang dianalisis mencakup tingkat pendidikan narapidana dan catatan kriminalitas mereka.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, berikut rumusan masalahnya:

1. Bagaimana melakukan analisis clustering kriminalitas berdasarkan tingkat pendidikan WBP di Rutan Kelas II Ternate menggunakan metode Agglomerative Hierarchical Clustering (AHC)?
2. Bagaimana menguji kualitas clustering yang dihasilkan metode AHC menggunakan Silhouette Coefficient?

1.3. Batasan Masalah

1. Data penelitian ini adalah data Warga Binaan Pemasyarakatan (WBP) khususnya narapidana dari tahun 2019 sampai 2023 yang bersumber dari Rutan Kelas II Ternate.

2. Data tingkat pendidikan WBP yang di ambil dari tingkat pendidikan SD sampai S1 dan tidak sekolah

1.4. Tujuan Penelitian

Adapun tujuan penelitian yang diangkat. Yaitu:

1. Melakukan analisis clustering kriminalitas berdasarkan tingkat pendidikan WBP di Rutan Kelas II Ternate menggunakan metode *Agglomerative Hierarchical Clustering* (AHC).
2. Menguji kualitas hasil clustering yang diperoleh menggunakan *Silhouette Coefficient*.

1.5. Manfaat Penelitian

Penelitian ini diharapkan akan memberikan manfaat sebagai berikut:

1. Meningkatkan pemahaman tentang bagaimana tingkat pendidikan WBP berkorelasi dengan perilaku kejahatan, sehingga dapat menjadi dasar pengembangan program rehabilitasi dan resosialisasi di Rutan Kelas II Ternate.
2. Memberikan pemahaman tentang analisis *Clustering* kriminalitas berdasarkan tingkat pendidikan WBP di Rutan Kelas II Ternate.
3. Memberikan informasi tentang kualitas *Clustering* yang telah dilakukan menggunakan metode *Silhouette Coefficient*.

1.6. Sistematika Penulisan

Untuk memudahkan pembahasan dalam proposal ini, sistematika penulisan dibagi menjadi 3 (tiga) bab yang terdiri dari:

BAB I PENDAHULUAN

Terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Memaparkan teori-teori yang didapat dari sumber-sumber relevan untuk digunakan sebagai panduan dalam penelitian serta penyusunan proposal ini.

BAB III METODE PENELITIAN

Bab ini membahas tentang metode penelitian yang telah dilakukan oleh penulis dengan permasalahan yang diangkat.

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terkait

Pada bab ini, penelitian yang dilakukan penulis bukanlah yang pertama kali dalam konteks *Clustering K-Means* kriminal berdasarkan tingkat pendidikan. Sebelumnya membahas temuan penelitian ini penting untuk mencatat bahwa penelitian ini tidak berdiri sendiri, tetapi merupakan kontribusi terhadap kerangka kerja yang sudah ada dalam *Clustering K-Means* kriminal berdasarkan tingkat pendidikan. Sejumlah penelitian terkait telah mengulas topik yang sama sebelumnya, dan hasil-hasil mereka memberikan wawasan berharga yang membentuk dasar untuk penelitian. Oleh karena itu, dalam bab ini, akan disajikan landasan teori yang relevan dengan penelitian yang akan dilakukan. Daftar penelitian terkait dapat ditemukan dalam Tabel 2.1.

Tabel 2.1 Penelitian terkait

No	Judul dan Penulis	Metode Penelitian	Hasil Penelitian
1	(Arifin, 2020)	Klasterisasi Genre Cerpen Kompas Menggunakan <i>Agglomerative Hierarchical Clustering- Single Linkage</i>	Hasil penelitian ini menunjukkan bahwa Metode <i>Agglomerative Hierarchical Clustering - Single Linkage</i> memiliki performansi yang lebih baik dibandingkan dengan metode k-Means. Dari 127 dataset cerpen yang telah diujicobakan didapatkan nilai akurasi dari metode <i>Agglomerative Hierarchical</i>

			Clustering - Single Lingkage 72,2441 %, sedangkan metode k-Means hanya 67,7953 %. Selain itu juga dapat disimpulkan dokumen yang membahas topik yang sama cenderung untuk mengelompok menjadi satu klaster, dan klaster dapat membantu mendapatkan dokumen yang relevan. Sedangkan Purity (kemurnian) suatu klaster direpresentasikan sebagai anggota klaster yang paling banyak sesuai (cocok) di suatu kelas
2	(Iyan Yulianti 2023)	Analisis Clustering Donor Darah dengan Metode Agglomerative Hierarchical Clustering	Hasil dari penelitian ini dinyatakan berhasil, dengan tingkat akurasi sebesar 0.713346292. dengan standarisasi hasil di inteprestasi sangat baik (<i>Strong Structur</i>) antara data dan cluster yang terbentuk. Dengan ini maka metode Agglomerative Hierarchical Clustering dapat diterapkan ke dalam orange dengan nilai validasi yang sama. Sumber data yang digunakan pada penelitian ini adalah data yang diperoleh dari Unit Donor Darah PMI Purwakarta.
3	(Justitia et al., 2021)	Implementasi Metode Agglomerative Hierarchical Clustering Pada	Penerapan metode ditarik Agglomerative Hierarchical Clustering

		Segmentasi Pelanggan Barbershop (Studi Kasus : RichDjoe Barbershop Malang)	pada penelitian ini menghasilkan 2 parameter jarak terbaik yaitu <i>single linkage</i> dan <i>average linkage</i> . Kedua parameter jarak ini dalam proses pengelompokan data memberikan hasil paling optimal pada titik potong 391 dengan jumlah 8 cluster. 2. Tingkat akurasi metode <i>Agglomerative Hierarchical Clustering</i> yang diuji menggunakan <i>Silhouette Coefficient</i> pada <i>single linkage</i> dan <i>average linkage</i> yang merupakan parameter jarak terbaik mendapatkan nilai rata-rata paling tinggi akurasi yaitu 0.968850698. Pada pengujian pada setiap data mendapatkan hasil 376 dari 399 data berada tepat pada cluster-nya dan pengujian pada setiap cluster-nya mendapatkan hasil 6 dari 8 cluster yang terbentuk tepat.
4.	<i>Data Mining Clustering</i> Tingkat Kejahatan Dengan Metode Algoritma <i>K-Means</i> (Studi Kasus: Kejaksaan Negeri Binjai) Yosi Mayona, Relita Buaton, Magdalena Simanjutak. (Agustus 2022)	Metode yang digunakan dalam penelitian ini meliputi penambangan data dan algoritma <i>Clustering K-Means</i> . Data yang digunakan dalam penelitian adalah data sekunder yang diperoleh dari Kejaksaan Negeri Binjai.	dijelaskan, masing-masing dengan karakteristik rentang usia dan jenis kejahatannya. Makalah ini juga menyediakan tabel dan grafik untuk memvisualisasikan hasil pengelompokan.
5.	Analisis <i>Data Mining</i> Pengelompokan	Penelitian menggunakan <i>Data Mining</i> sebagai	Setelah melalui beberapa iterasi, hasil

	Kasus Kejahatan Terjadi Di Kecamatan Medan Polonia Dengan Menggunakan Metode <i>K-Means Clustering</i> . Nadya Septiani, Kamil Erwansyah, S.Kom, M.Kom., Mhd.Gilang Suryanata, S.Kom, M.Kom. (Februari 2020)	Tindak Yang metode. <i>Data Mining</i> adalah proses ekstraksi informasi berharga dari data dengan menggunakan berbagai teknik seperti statistik, matematika, dan <i>machine learning</i> . Algoritma <i>K-Means Clustering</i> dipilih untuk pengelompokkan kasus tindak kejahatan karena kelebihan dalam kemudahan implementasi, kecepatan, dan adaptabilitas.	penelitian membagi kasus tindak kejahatan menjadi tiga <i>cluster</i> dengan pusat <i>cluster</i> masing-masing. <i>Cluster</i> Narkoba merupakan <i>cluster</i> terbesar dengan anggota <i>cluster</i> 36 data. Mayoritas tindak pidana yang terjadi adalah Narkoba dengan lokasi kejadian mayoritas di Kelurahan Sari Rejo dan Polonia. <i>Cluster</i> Pencurian merupakan <i>cluster</i> dengan anggota <i>cluster</i> 38 data. Mayoritas tindak pidana yang terjadi adalah Pencurian, Perjudian dan Pemerkosaan dengan lokasi kejadian mayoritas di Kelurahan Madras Hulu dan Sari Rejo. <i>Cluster</i> Perjudian merupakan <i>cluster</i> dengan anggota <i>cluster</i> 26 data. Mayoritas tindak pidana yang terjadi adalah Premanisme dan Perjudian dengan lokasi kejadian mayoritas di Kelurahan Polonia dan Madras Hulu.
6.	Implementasi <i>Data Mining</i> Dalam Pengelompokkan Data Potensi Kejahatan	Penelitian melibatkan penggunaan <i>K-Means Clustering</i> untuk mengelompokkan potensi tingkat kejahatan	Berdasarkan perhitungan <i>K-Means</i> , data wilayah di Kecamatan Beringin dibagi menjadi dua

	Menggunakan Algoritma <i>K-Means Clustering</i> Pada Kapolsek Beringin. Riska Putria Saragih, Muhammad Zunaidi, Hafizah (Juli 2023)	di Kecamatan Beringin. Langkah pertama adalah menentukan jumlah <i>cluster</i> (k), yang dihitung berdasarkan analisis data yang ada. Kemudian, <i>centroid</i> awal dipilih secara acak. Selanjutnya, jarak setiap data <i>input</i> dihitung terhadap masing-masing <i>centroid</i> menggunakan rumus jarak <i>Euclidean</i> . Data dikelompokkan ke dalam <i>cluster</i> dengan jarak terdekat. Nilai <i>centroid</i> baru dihitung dengan mengambil rata-rata data dalam masing-masing <i>cluster</i> . Proses ini diulang hingga anggota setiap <i>cluster</i> tidak berubah. Hasil akhir adalah <i>cluster-cluster</i> yang mencerminkan tingkat kejahatan yang berbeda di berbagai wilayah di Kecamatan Beringin.	<i>cluster</i> utama (C1 dan C2) berdasarkan tingkat kejahatan. Dalam <i>cluster</i> 1, terdapat wilayah dengan tingkat kejahatan yang lebih tinggi, sementara dalam <i>cluster</i> 2, terdapat wilayah dengan tingkat kejahatan yang lebih rendah.
7.	Penggunaan Algoritma <i>K-Means</i> Pada Metode <i>Clustering</i> Untuk Menganalisa Tindak Kriminal. Imam Zuhdi Muzakkiy, Khoirul Husein, Kelfin Antonius, Kevin Raihan Hidayat, El Emir Di Haryanto, Iman Paryudi (Mei 2023)	Metode penelitian yang digunakan dalam penelitian ini adalah metode algoritma <i>K-Means Clustering</i> . Penelitian ini menggunakan data kriminal dari Kepolisian Daerah Sumatera Utara (POLDASU) yang terdiri dari 28 Polres. Data tersebut diolah dengan <i>software</i> RapidMiner 7.3.	Hasil penelitian menunjukkan bahwa pengelompokan daerah rawan tindak kriminal pada POLDASU menggunakan metode <i>K-Means Clustering</i> menghasilkan 3 kelompok. Kelompok pertama terdiri dari daerah rawan tindak kriminal.
8.	Sistem Informasi Geografis Pengelompokan Tingkat Kriminalitas Kota Malang	Metode yang digunakan dalam penelitian ini adalah metode <i>K-Means</i> dengan menggunakan pengujian keakuratan	Penerapan dari hasil <i>Clustering</i> yang dilakukan menunjukkan bahwa hasil <i>Clustering</i> menggunakan KMeans

	Menggunakan Metode <i>K-Means</i> . Danang Sutejo, Yosep Agus Pranoto, Hani Zulfia Zahro (Maret 2020).	menggunakan <i>Davies-Bouldin Index</i>	menghasilkan percobaan 3 <i>cluster</i> yang memiliki hasil analisis keakuratan terkecil menggunakan Davies Bouldin Index yaitu sebesar 2,401 dengan pembagian daerah intensitas aman(C1) yaitu Kecamatan Kedungkandang, dalam <i>cluster</i> (C2) cukup rawan yaitu Kecamatan Blimbings, dalam <i>cluster</i> (C3) rawan yaitu kecamatan Klojen, Sukun dan dalam <i>cluster</i> sangat rawan (C4) yaitu kecamatan Lowokwaru. Pada pengujian user yang telah dilakukan, user memilih sangat baik adalah 32%, memilih baik 55%, memilih cukup baik 13% dan memilih kurang baik 0% kemudian hasil pengujian fungsionalitas aplikasi pada <i>browser</i> yang berbeda, sudah berjalan dengan baik
9.	Analisa Metode <i>K-Means</i> pada Pengelompokan Kriminalitas Menurut Wilayah (Dewi, 2019).	Metode penelitian yang digunakan dalam studi ini adalah metode analisis menggunakan metode <i>K-Means</i> pada pengelompokan kriminalitas menurut provinsi. Penelitian ini menggunakan data kriminalitas dari tahun 2012 yang terdiri dari 31	Hasil penelitian menunjukkan bahwa terdapat dua <i>cluster</i> yang terbentuk, yaitu <i>cluster</i> tinggi (C1) dan <i>cluster</i> rendah (C2). <i>Cluster</i> tinggi terdiri dari provinsi-provinsi dengan tingkat kriminalitas yang tinggi, seperti

		<p>provinsi di Indonesia. Tahapan penelitian meliputi pengumpulan data, pengolahan data, <i>Clustering</i>, dan analisis.</p>	<p>Sumatera Utara, Sumatera Selatan, Jawa Barat, Sulawesi Tengah, Sulawesi Selatan, dan Papua. Sedangkan <i>cluster</i> rendah terdiri dari provinsi-provinsi dengan tingkat kriminalitas yang rendah, seperti Aceh, Sumatera Barat, Riau, Jambi, Bengkulu, dan provinsi lainnya. Dalam kesimpulan penelitian, disimpulkan bahwa metode <i>K-Means</i> dapat digunakan untuk pengelompokan kriminalitas menurut provinsi. Hasil penelitian ini memberikan informasi penting bagi pemerintah untuk memberikan perhatian lebih pada provinsi-provinsi dengan tingkat kriminalitas yang tinggi.</p>
10.	Pemetaan Wilayah Rawan Kriminalitas Di Kabupaten Batu Bara Menggunakan Metoda <i>K-Means</i> , Rizki Fadillah (September 2022)	<p>Metode penelitian yang digunakan dalam penelitian ini adalah metode kualitatif. Metode ini menggunakan data berbentuk kata, skema, dan gambar. Tahapan dalam metode penelitian ini meliputi identifikasi masalah, studi literatur, analisis data, perancangan sistem, dan implementasi.</p>	<p>Hasil penelitian ini adalah pengembangan suatu sistem informasi yang mampu menampilkan dan memetakan jumlah tindak kejahatan yang terjadi di Kabupaten Batu Bara menggunakan metode <i>K-Means Clustering</i>. Hasil penelitian menyatakan bahwa keamanan di Kabupaten Batu Bara</p>

		cenderung rawan di sekitar pusat kota, sedangkan wilayah pinggir kota dengan kepadatan penduduk relatif rendah cenderung lebih aman. Selain itu, tindak kejahatan dengan intensitas tertinggi terjadi pada larut malam. Dalam analisis, faktor-faktor seperti kinerja polsek, kepadatan penduduk, dan angka kemiskinan mempengaruhi kriminalitas sebesar 55%, sedangkan 45% dipengaruhi oleh faktor lain. Implementasi sistem yang telah dibuat dapat memudahkan dan menyelesaikan persoalan daerah rawan kriminalitas. Program tersebut mudah dan cepat serta mendapatkan hasil yang diharapkan.
--	--	---

Penelitian saya berjudul "Analisis *Clustering* Kriminalitas Berdasarkan Tingkat Pendidikan Warga Binaan Pemasyarakatan (WBP) Menggunakan Metode *Agglomerative Hierrarcy Clustering* (AHC)" Penelitian ini memfokuskan analisis *clustering* kriminalitas dengan mempertimbangkan tingkat pendidikan dari warga binaan pemasyarakatan (WBP) di Rutan Kelas II Ternate. Sebagai perbandingan, penelitian sebelumnya yang terdapat dalam tabel berkaitan dengan penggunaan

algoritma *K-Means Clustering* untuk mengelompokkan daerah rawan tindak kriminal di Kepolisian Daerah Sumatera Utara (POLDASU) dengan data dari 28 Polres. Penelitian sebelumnya tidak merinci pengujian tertentu atau tujuan spesifik. Berbeda dengan itu, penelitian saya mencakup penggunaan metode AHC *Clustering* untuk menganalisis tingkat pendidikan sebagai faktor pengelompokan kriminalitas di lingkungan Rutan Kelas II Ternate. Perbedaan fokus penelitian, subjek, metode pengujian, dan tujuan penelitian menjadi ciri khas yang membedakan penelitian saya dari penelitian terdahulu.

2.2. Data Mining

Data *Mining* (DM) merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata *mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu *Data Mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan database (Septiani, 2020).

Data Mining adalah proses penggalian informasi dari data set yang besar melalui penggunaan algoritma dan teknik yang diambil dari bidang statistik, *machine learning* dan database sistem manajemen (Suryadi, 2019)

2.3. Analisis *Clustering*

Clustering atau klasterisasi adalah metode pengelompokan data merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Objek yang di dalam *cluster* memiliki kemiripan karakteristik antar satu

sama lainnya dan berbeda dengan *cluster* yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *Clustering* (Mayona, 2022).

Clustering merupakan salah satu teknik *Data Mining* yang digunakan untuk mendapatkan kelompok-kelompok dari obyek-obyek yang mempunyai karakteristik yang umum didata yang cukup besar. Tujuannya menemukan *cluster* yang berkualitas dalam waktu yang layak. *Clustering* dalam *Data Mining* berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data (Suriani, 2020)

2.4. *Agglomerative Hierarchical Clustering*

AHC merupakan sebuah algoritma atau metode dalam analisis klaster yang bertujuan untuk mengelompokkan data ke dalam klaster-klaster hierarkis. Pendekatan ini dimulai dengan memperlakukan setiap data sebagai sebuah klaster tunggal, dan kemudian secara bertahap menggabungkan klaster-klaster yang memiliki kesamaan tertinggi berdasarkan suatu metrik atau jarak yang ditentukan. Proses AHC dimulai dengan mengukur jarak antara setiap pasangan data, lalu menggabungkan dua klaster dengan jarak terdekat untuk membentuk satu klaster baru. Proses ini berlanjut dengan menggabungkan klaster-klaster yang semakin besar hingga semua data tergabung dalam satu klaster utama. Selama proses ini, terbentuklah sebuah struktur hirarkis yang menggambarkan hubungan antar klaster, biasanya dalam bentuk pohon dendrogram atau grafik (Sundari et al., 2023)

Salah satu kelebihan dari metode AHC adalah fleksibilitasnya dalam memilih metrik atau jarak yang digunakan untuk mengukur kesamaan antar data. Metrik tersebut dapat berupa jarak Euclidean, jarak Manhattan, atau jarak Mahalanobis,

tergantung pada karakteristik data dan tujuan analisis. Selain itu, penggunaan AHC juga dapat dikontrol dengan memilih jumlah klaster yang diinginkan atau menggunakan kriteria pemotongan (*cut-off*) untuk menghentikan proses penggabungan klaster. Jarak antara dua cluster dapat dihitung menggunakan Euclidean Distance ataupun manhattan distance

1. Untuk menghitung matriks jarak Euclidean Distance dengan rumus 2.2

2. Untuk menghitung matriks jarak Manhattan Distance dengan rumus 2.1

di mana x dan y adalah dua titik, D adalah jumlah dimensi, dan x_j dan y_j adalah nilai pada dimensi ke j .

Untuk Proses pengclustering AHC dapat dilakukan dengan Beberapa metode pengelompokkan secara Hierarki Aglomeratif :

1. Single Linkage (Jarak Terdekat) single linkage yang digunakan dalam pengelompokan berdasarkan jarak terdekat atau terkecil antara satu cluster dengan cluster yang lain dengan menggunakan rumus 2.3:

2. Complete Linkage (Jarak Terjauh) complete linkage digunakan untuk mengelompokkan data cluster satu dengan yang lain berdasarkan jarak terjauh atau jarak maksimum antara elemen-elemen yang berbeda di dalam kluster, dengan menggunakan rumus 2.4:

3. Average Linkage (Jarak Rata-rata) Pendekatan ini memusatkan perhatian pada mengukur jarak terbesar yang terjadi antara dua klaster, sehingga menghasilkan klaster yang memiliki perbedaan maksimum. Berikut adalah rumus perhitungan jarak Average Linkage :

2.5. Pengujian *Silhouette Coeffisient*

Salah satu indikator evaluasinya adalah *koefisien siluet* (SC) yang mengevaluasi *cluster* berdasarkan ukuran rata-rata jarak antara satu titik data dengan titik data lainnya dalam *cluster* yang sama (kohesi) dan jarak rata-rata antar *cluster-cluster* yang berbeda (pemisahan). Kelebihan SC saja bergantung pada partisi kumpulan data tetapi tidak pada algoritma pengelompokan itu sendiri.(Tambunan, 2020)

Metode *koefisien silhouette* adalah cara untuk mengevaluasi seberapa baik sebuah *cluster* diatur. Evaluasi ini penting untuk menentukan seberapa akurat data dikelompokkan. Untuk menghitung *koefisien silhouette*, langkah-langkah spesifik harus diikuti (Muzakkiy, 2023):

1. Temukan jarak rata-rata antara objek i dan semua objek lain dalam grup yang sama.

Keterangan 2.3:

$\alpha(i)$ adalah nilai atau skalar yang ingin dihitung untuk elemen i .

A adalah himpunan indeks yang tidak sama dengan i .

$|A|$ adalah jumlah elemen dalam himpunan A .

$d(i, j)$ adalah jarak atau perbedaan antara elemen i dan elemen j .

Keterangan 2.4:

$d(i, C)$ adalah nilai atau skalar yang menunjukkan rata-rata jarak elemen i ke semua elemen dalam himpunan C .

$|C|$ adalah jumlah elemen dalam himpunan C , yaitu panjang himpunan C .

2. Hitung jarak rata-rata antara setiap objek dan semua data dalam berbagai kelompok, dan pilih nilai minimum untuk setiap objek.

3. Setelah itu, tentukan pengukuran numerik dari *koefisien silhouette*.

Keterangan 2.6:

$S(i)$: Nilai Silhouette Coefficient untuk objek i .

$a(i)$: Rata-rata jarak objek i ke objek-objek lain dalam kelompok yang sama (intra-cluster distance).

$b(i)$: Rata-rata jarak objek i ke objek-objek dari kelompok lain yang paling dekat

$\max\{q(i), b(i)\}$: Maksimum antara $q(i)$ dan $b(i)$.

Nilai koefisien siluet dapat berkisar dari -1 hingga 1, dan nilai 1 menunjukkan bahwa objek ke- i berada di $cluster$ yang benar, sehingga hasil $Clustering$ akurat.

Ketika nilai *koefisien siluet* adalah 0, tidak pasti apakah objek ke-*i* milik *cluster U* atau *cluster V* karena terletak di antara dua *cluster* (Muzakkiiy, 2023).

Nilai *koefisien siluet* menunjukkan kualitas struktur *cluster*, dan nilai -1 berarti *cluster* yang berbeda mungkin lebih cocok untuk objek tertentu. Kauffman

dan Rousseuw (1990) membuat tabel dengan kriteria subyektif untuk pengelompokan kualitas berdasarkan skor siluet (Muzakkiy, 2023).

2.6. Tingkat Pendidikan

Tingkat pendidikan adalah tahapan pendidikan yang ditetapkan berdasarkan tingkat perkembangan peserta didik, tujuan yang akan dicapai dan kemauan yang dikembangkan. Tingkat pendidikan berpengaruh terhadap perubahan sikap dan perilaku hidup sehat. Tingkat pendidikan yang lebih tinggi akan memudahkan seseorang atau masyarakat untuk menyerap informasi dan mengimplementasikannya dalam perilaku dan gaya hidup sehari-hari, khususnya dalam hal kesehatan. Pendidikan formal membentuk nilai bagi seseorang terutama dalam menerima hal baru (Rusmanto, 2019).

Jenjang merupakan tingkat atau level dari bawah ke atas yang terbagi dalam beberapa tahap. Sementara pendidikan merupakan pembelajaran yang mengarahkan suatu kelompok individu untuk memperoleh sebuah kebiasaan. Jenjang pendidikan adalah tingkatan pendidikan yang telah dikukuhkan berlandaskan strata atau hirarki dan level perkembangan siswa, misi yang akan diraih dan keterampilan yang akan dikembangkan. Di Indonesia jenjang pendidikan diklasifikasikan dalam beberapa tahap, diantaranya adalah pendidikan usia dini, pendidikan dasar, pendidikan menengah dan pendidikan tinggi. Jenjang pendidikan yang ada diklasifikasikan berlandaskan keahlian siswa, usia dan tingkat kecakapannya (Anis, 2021).

2.7. Kriminalitas dan Pemasyarakatan

Kriminalitas atau tindak kejahatan bukanlah suatu peristiwa herediter yang

merupakan bawaan sejak lahir, bukan juga sesuatu yang dapat diwariskan secara biologis. Perilaku tindak kriminal dapat dilakukan oleh siapapun, baik itu pria maupun wanita, dapat berlangsung dari usia anak-anak, dewasa, maupun usia lanjut. Tindak kriminal dapat terjadi secara sadar yaitu dipikirkan terlebih dahulu, direncanakan, dan diarahkan pada satu maksud tertentu dalam keadaan yang benar-benar sadar. Selain itu dapat pula dilakukan dalam keadaan setengah sadar, dipengaruhi oleh impuls-impuls yang kuat dari dalam diri seseorang, dorongan-dorongan paksaan yang kuat (kompulsi-kompulsi), dan obsesi-obsesi. Suatu kejahatan dapat dilakukan secara tidak sadar sama sekali, misalkan karena terpaksa untuk tetap bertahan hidup dengan cara melawan dan membalas menyerang yang pada akhirnya terjadi peristiwa pembunuhan (Saragih, 2023).

Pemasyarakatan merupakan bagian yang tidak terpisahkan dari sistem peradilan pidana terpadu yang diselenggarakan oleh pemerintah sebagai bagian dari proses penegakan hukum dalam rangka pelayanan serta pembinaan dan pembimbingan untuk reintegrasi sosial. Pemasyarakatan dalam Undang-Undang Nomor 22 tahun 2022 tentang Pemasyarakatan, pemasyarakatan adalah subsistem peradilan pidana yang menyelenggarakan penegakan hukum di bidang perlakuan terhadap tahanan, anak, dan warga binaan. Tahanan adalah tersangka atau terdakwa yang sedang menjalani proses peradilan dan ditahan di rumah tahanan negara. Anak yang dimaksud dalam pengertian pemasyarakatan tersebut adalah anak yang berkonflik dengan hukum yang selanjutnya disebut anak adalah anak yang telah berumur 12 (dua belas) tahun, tetapi belum berumur 18 (delapan belas) tahun yang diduga melakukan tindak pidana (Usman, 2022).

2.8. Warga Binaan Pemasyarakatan (WBP)

Warga binaan pemasyarakatan adalah sebagian masyarakat yang mendapatkan sanksi atas tindakan kriminalitas yang dilakukannya dan berada dalam suatu lembaga pemasyarakatan, seperti penjara atau rumah tahanan, karena telah melakukan tindakan pidana dan sedang menjalani hukuman. Pembinaan warga binaan bertujuan untuk menyadarkan mereka agar menyesali perbuatannya, mengembalikan mereka menjadi warga masyarakat yang baik, taat kepada hukum, menjunjung tinggi nilai-nilai moral, sosial, dan keagamaan, serta meningkatkan kualitas sumber daya manusia warga binaan pemasyarakatan. Pembinaan warga binaan dilakukan melalui berbagai program, seperti pelatihan keterampilan, kesadaran beragama, dan kemandirian (Wulandari, 2015).

Warga Binaan Pemasyarakatan atau yang dapat disingkat dengan WBP, pada Pasal 1 angka 5 Undang-Undang Republik Indonesia Nomor 12 Tahun 1995 tentang Pemasyarakatan menjelaskan bahwa “Warga Binaan Pemasyarakatan adalah Narapidana, Anak Didik Pemasyarakatan, dan Klien Pemasyarakatan. Pada Undang-Undang Republik Indonesia Nomor 12 Tahun 1995 tentang Pemasyarakatan Pasal 1 angka 7 yang menyatakan dengan jelas bahwa “Narapidana adalah terpidana yang sedang menjalani masa pidana hilang kemerdekaan di Lembaga Pemasyarakatan” (123dok.com, 2023).

2.9. Rumah Tahanan Negara (Rutan) Kelas II Ternate

Rumah Tahanan Negara (Rutan) adalah tempat tersangka atau terdakwa ditahan selama proses penyidikan, penuntutan, dan pemeriksaan di sidang pengadilan di Indonesia. Rumah Tahanan Negara merupakan unit pelaksana teknis

di bawah Kementerian Hukum dan Hak Asasi Manusia (dahulu Departemen Kehakiman). Selain Rutan yang dikelola oleh Direktorat Jenderal Pemasyarakatan Kementerian Hukum dan Hak Asasi Manusia, beberapa Instansi memiliki Rumah Tahanan sendiri. Sebut saja Kepolisian RI, Komisi Pemberantasan Korupsi, dan Badan Narkotika Nasional. Rutan didirikan pada setiap ibu kota kabupaten atau kota, dan apabila perlu dapat dibentuk pula Cabang Rutan. Di dalam Rutan, ditempatkan tahanan yang masih dalam proses penyidikan, penuntutan, dan pemeriksaan di Pengadilan Negeri, Pengadilan Tinggi, dan Mahkamah Agung (Wikipedia, 2022).

2.10. Streamlit

Pada penelitian ini digunakan graphical user interface (GUI) Streamlit. Streamlit adalah kerangka kerja web yang ditujukan untuk menyebarkan model dan visualisasi dengan mudah menggunakan bahasa Python, yang cepat dan minimalis tetapi juga memiliki tampilan yang cukup baik serta ramah pengguna. Tersedia widget bawaan untuk masukan pengguna, seperti pengunggahan gambar, penggeser, masukan teks, dan elemen hypertext markup language (HTML) lain yang sudah dikenal, seperti checkboxes dan radio buttons. Setiap kali pengguna berinteraksi dengan aplikasi Streamlit, skrip Python dijalankan kembali dari atas ke bawah.

Hal ini merupakan sebuah konsep penting yang perlu diingat saat mempertimbangkan berbagai status aplikasi yang akan dipilih. Streamlit merupakan aplikasi yang tidak berbayar dan pengguna tidak perlu memiliki pengetahuan pengembangan front-end yang mahir untuk mengoperasikannya.

Streamlit dapat dijalankan pada editor Anaconda serta bahasa Python seri 3.7 ke atas, tetapi tidak mendukung pada editor Jupyter Notebook, sehingga harus dikonversi ke editor Pycharm atau Visual Code. Tampilan beranda pada aplikasi Streamlit dapat dipisahkan menjadi dua bagian, yaitu buttons, untuk pemilihan menu, serta tampilan visual chart (Widi Hastomo et al., 2022).

2.11. Python

Python merupakan bahasa pemrograman yang berorientasi obyek dinamis, dapat digunakan untuk bermacam macam pengembangan perangkat lunak. Python menyediakan dukungan yang kuat untuk integrasi dengan bahasa pemrograman lain dan alat-alat bantu lainnya. Python hadir dengan pustaka-pustaka standar yang dapat diperluas serta dapat dipelajari hanya dalam beberapa hari. Bahasa pemrograman yang interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif (Nugroho, 2020).

2.12. *Black Box*

Pengujian black box adalah metode pengujian perangkat lunak yang berfokus pada pemeriksaan fungsionalitas sistem dari sudut pandang pengguna, tanpa memperhatikan struktur internal atau kode sumber program. Dalam pendekatan ini, penguji tidak memiliki pengetahuan tentang bagaimana sistem diimplementasikan secara teknis, tetapi hanya mengetahui spesifikasi dan persyaratan yang harus dipenuhi oleh perangkat lunak. Oleh karena itu, pengujian black box sering juga

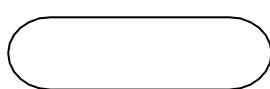
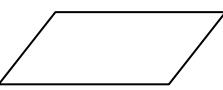
disebut sebagai pengujian berbasis spesifikasi.

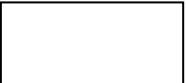
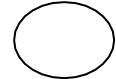
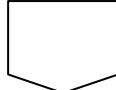
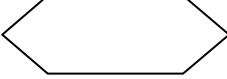
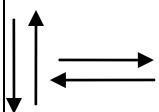
Tujuan utama dari pengujian black box adalah untuk memastikan bahwa sistem bekerja sesuai dengan fungsionalitas yang diharapkan oleh pengguna. Pengujian akan memberikan input pada perangkat lunak dan memeriksa apakah output yang dihasilkan sesuai dengan hasil yang diharapkan. Melakukan pengujian fungsional terhadap setiap halaman dan fitur, seperti tampilan hasil clustering, dan visualisasi dendrogram. Pengujian ini memastikan bahwa input yang dimasukkan menghasilkan output yang sesuai dan setiap halaman berfungsi dengan benar. Hasil Sistem yang telah teruji, dengan hasil pengujian menunjukkan bahwa aplikasi berjalan sesuai harapan dan bebas dari kesalahan fungsional.

2.13. *Flowchart*

Flowchart adalah langkah-langkah pemecahan masalah yang ditulis atau dilambangkan dengan simbol-simbol tertentu. *Flowchart* akan menunjukkan alur program secara realistik dan logis, sehingga *Flowchart* dapat dipahami sebagai langkah-langkah pemecahan masalah yang ditulis dalam simbol-simbol tertentu (Khesya, 2021). Berikut ini simbol-simbol dan kegunaan dari simbol-simbol yang sering digunakan untuk menggambarkan suatu algoritma dalam bentuk diagram alir, dapat dilihat pada tabel 2.4.

Tabel 2.4 Simbol *Flowchart* (Khesya, 2021)

No	Simbol	Nama	Fungsi Simbol
1.		“Terminal”	Awal atau akhir suatu program (Prosedur).
2.		“Output/Input”	Proses <i>input</i> atau <i>output</i> terlepas dari jenis perangkat.

3.		<i>“Process”</i>	Proses operasional komputer.
4.		<i>“Decision”</i>	Untuk menunjukkan bahwa suatu kondisi tertentu mengarah pada dua kemungkinan, ya/tidak.
5.		<i>“Connector”</i>	Koneksi penghubung proses ke proses lain pada halaman yang sama.
6.		<i>“Offline Connector”</i>	Koneksi penghubung dari satu proses ke proses lain di halaman lain.
7.		<i>“Predefined Process”</i>	Mewakili ketentuan penyimpanan untuk diproses untuk memberikan awal harga.
8.		<i>“Punched Card”</i>	Input berasal dari kartu atau output ditulis ke kartu.
9.		<i>“Documen”</i>	Mencetak output dalam format dokumen (melalui printer).
10.		<i>“Flow”</i>	Menyatakan jalannya arus suatu proses.

BAB III

METODE PENELITIAN

3.1. Objek dan Waktu Penelitian

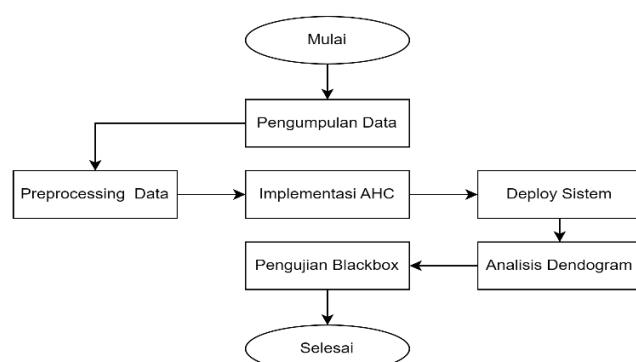
Objek yang diteliti pada penelitian ini yaitu kantor Rumah Tahanan Milik Negara (Rutan) Kelas II Ternate yang dimana data yang diteliti adalah data para narapidana yang kemudian di kelompokkan menggunakan metode AHC. Waktu penelitian yang dilakukan pada semester ganjil tahun ajaran 2025/2026.

3.2. Sifat, Jenis dan Pendekatan Penelitian

Jenis penelitian yang akan dilakukan menerapkan konsep kuantitatif, yang merupakan salah satu metode penelitian yang banyak digunakan dalam ilmu pengetahuan dan penelitian *social*. Penelitian ini memiliki sifat deskriptif, karena akan menggambarkan pola kriminalitas berdasarkan tingkat pendidikan warga binaan pemasarakatan (WBP). Dengan menggunakan metode AHC dan pengujian *silhouette coefficient*, penelitian ini menandakan pendekatan kuantitatif dalam menganalisis dan mengelompokkan data kriminalitas.

3.3 Alur Penelitian

Tahapan penelitian dapat di lihat pada flowchart di bawah ini:



Gambar 3.1. Tahapan Penelitian

Berikut ini adalah penjelasan dari tahapan-tahapan alur penelitian dari gambar 3.1

1. Pengumpulan Data

Menghimpun data sekunder resmi dari Rutan Kelas II(B) Ternate yang mencakup profil WBP, tingkat pendidikan, serta atribut pendukung (misal jenis kejahatan, tahun perkara). Apabila diperlukan, dilakukan wawancara terstruktur dengan petugas Rutan guna klarifikasi variabel dan konteks data. Seluruh data bersifat anonim (tanpa identitas personal).

2. *Preprocessing* Data

Preprocessing Data merupakan tahapan dalam memproses dataset agar menjadi dataset yang siap digunakan. tahap *Preprocessing* yang dilakukan pada penelitian ini adalah data *cleaning* Dimana Memastikan kualitas data sebelum diimplementasikan pada algoritma clustering. Melakukan proses seperti pembersihan data (menghapus duplikasi, menangani *missing values*), normalisasi data, dan *encoding* variabel kategorikal jika ada.

3. Implementasi *Agglomerative Hierarchical Clustering* (AHC)

Pada tahap ini, Rutan Kelas II(B) Ternate dan data WBP dikelompokkan menggunakan *algoritma Agglomerative Hierarchical Clustering* (AHC), yaitu metode *clustering* yang bekerja dari bawah ke atas (*bottom-up*). Setiap data awalnya dianggap sebagai satu cluster, kemudian secara bertahap digabungkan berdasarkan tingkat kemiripan hingga membentuk satu cluster besar. Penggabungan dilakukan berdasarkan jarak terdekat antar cluster, dengan menggunakan metode *linkage* seperti *single*, *complete*, atau *average linkage*. Proses

ini menghasilkan dendrogram, yaitu grafik berbentuk pohon yang menunjukkan hubungan hierarkis antar data. Dendrogram digunakan untuk menentukan jumlah cluster optimal dengan memotongnya pada titik jarak tertentu.

4. Deploy Sistem

Penelitian ini menggunakan *Streamlit* untuk Membuat antarmuka yang *user-friendly* untuk sistem *clustering*, sehingga *admin* dapat dengan mudah mengakses dan menganalisis hasil *clustering*. Mengembangkan aplikasi berbasis web menggunakan *Streamlit* yang menampilkan halaman seperti Dashboard, Hasil *Clustering*, dan Visualisasi Dendrogram. Hasil Aplikasi web yang siap digunakan oleh admin Puskesmas untuk mengelola data dan melihat hasil clustering secara real-time.

5. Analisis Dendogram

Menentukan jumlah *cluster* optimal berdasarkan dendrogram. Mengamati dendrogram untuk menemukan titik-titik potong yang tepat agar mendapatkan jumlah *cluster* optimal. Biasanya, titik potong ditentukan pada jarak maksimum antara node yang berbeda. Jumlah cluster yang optimal untuk analisis Rutan Kelas II(B) Ternate dan data WBP di Puskesmas, yang nantinya akan memudahkan dalam interpretasi hasil *clustering*.

6. Pengujian *Blackbox*

Memastikan bahwa semua fitur dalam aplikasi berfungsi sesuai dengan spesifikasi yang direncanakan. Melakukan pengujian fungsional terhadap setiap halaman dan fitur, seperti tampilan penjelasan AHC dan tampilan hasil *clustering* serta *visualisasi* dendrogram. Pengujian ini memastikan bahwa *input* yang

dimasukkan menghasilkan *output* yang sesuai dan setiap halaman berfungsi dengan benar. Hasil Sistem yang telah teruji, dengan hasil pengujian menunjukkan bahwa aplikasi berjalan sesuai harapan dan bebas dari kesalahan fungsional.

3.4 Metode Pengumpulan Data

Metode pengumpulan data yang peniliti terapkan dalam penelitian ini, yaitu metode data sekunder dan wawancara. Berikut penjelasan penerapan metode pengumpulan data:

1. Data Sekunder

Data sekunder merujuk kepada data yang telah dikumpulkan dan ada sebelumnya oleh pihak lain atau dalam konteks lain untuk tujuan yang mungkin berbeda dengan tujuan penelitian atau analisis yang Anda lakukan saat ini. Adapun data yang diambil terkait Rutan Kelas II Ternate, termasuk kapasitas tahanan, jumlah narapidana, latar belakang pendidikan, dan jenis kejahatan yang dilakukan oleh WBP yang dimana data tersebut didapatkan dalam bentuk *work sheet*.

2. Wawancara

Wawancara dilakukan untuk mendapatkan data yang berhubungan dengan para Warga Binaan Pemasyarakatan (WBP), yaitu dengan mewawancarai para pegawai untuk mendapatkan informasi mengenai WBP, yang dimana informasi tersebut mencakup jumlah tahanan dan narapidana yang ada didalam Rutan Kelas II Ternate dan informasi terkait lainnya.

3.5 Metode *Preprocessing* Data

Metode *preprocessing* data dilakukan untuk memastikan Rutan Kelas II(B) Ternate dan data WBP siap digunakan dalam algoritma *Agglomerative*

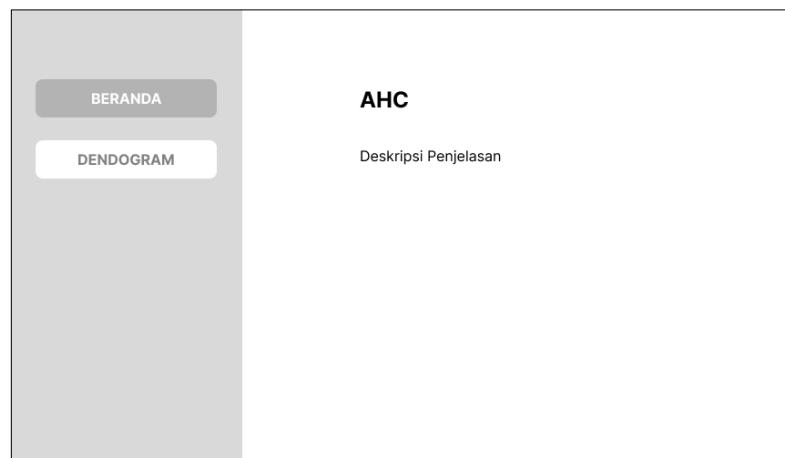
Hierarchical Clustering (AHC). Tahap pertama adalah data *cleaning*, yaitu membersihkan data dari duplikasi, menangani data yang hilang, dan menyelaraskan format. Seanjutnya tahap data *transformation*, data dinormalisasi agar atribut memiliki skala yang seragam, dan atribut baru dibuat jika diperlukan.

3.6 Rancangan Antarmuka

Rancangan antar muka (*Interface*) merupakan bagian visual dan interaktif dari sebuah sistem, perangkat lunak yang menjadi penghubung antara pengguna dan teknologi yang digunakan. Interface mencakup tampilan visual, tata letak, dan mekanisme interaksi yang membuat pengguna dapat berkomunikasi dengan sistem secara efektif.

1. Tampilan beranda

Tampilan beranda merupakan tampilan menu bagian pertama yang menampilkan penjelasan clustering AHC. Tampilan halaman beranda seperti pada gambar 3.3

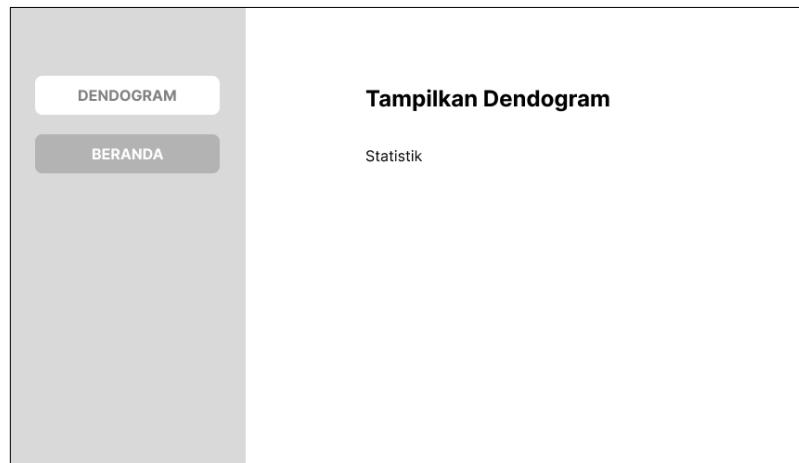


Gambar 3.3 Tampilan Menu Beranda

2. Tampilan lihat dendogram

Tampilan lihat dendogram merupakan tampilan menu bagian kedua yang

menampilkan hasil clustering data dan hasil dendogramnya. Tampilan menu lihat dendogram seperti pada gambar 3.4



Gambar 3.4 Tampilan Menu Lihat Dendogram

3.7 Rencana Pengujian Hasil

Pada metode pengujian sistem, peneliti menggunakan metode *black box testing*. Peneliti menggunakan black box testing karena metode ini adalah metode pengujian yang tidak melihat dan menguji *source code* program melainkan fokus pada detail aplikasi seperti tampilan aplikasi, fungsi pada aplikasi dan alur fungsi seperti yang diinginkan pelanggan. Berikut tabel pengujinya dapat dilihat pada tabel 3.3.

Tabel 3.3 Pengujian *Black Box*

No	Deskripsi Skenario	Skenario Uji	Output yang di harapkan	Hasil Pengujian
1	Navigasi ke Beranda	Pengguna mengklik tombol "Beranda" pada menu navigasi sebelah kiri	Sistem menampilkan halaman berisi penjelasan tentang metode clustering HC (<i>Hierarchical Clustering</i>)	

2	Navigasi ke Dendogram	Pengguna mengklik tombol "Lihat Dendogram" pada menu kiri	Sistem menampilkan visualisasi dendogram hasil proses clustering terhadap data	
---	-----------------------	---	--	--

DAFTAR PUSTAKA

- 123dok.com, 2023, Pengertian dan Macam-Macam Warga Binaan Pemasyarakatan, <https://123dok.com/article/pengertian-dan-macam-macam-warga-binaan-pemasyarakatan>.
- Anis, H., 2021, Jenjang Pendidikan, <https://www.tripven.com/jenjang-pendidikan/>
- Buaton, R., Zarlis, M., Mawengkang, H., & Effendi, S. 2019. *Clustering Time Series Data Mining* dengan Jarak Kedekatan *Manhattan City*. Prosiding Seminar Nasional Riset *Information Science* (SENARIS), ISSN:2686-0260, Vol. 1 Issue.1 September, 2019
- Dewi, S. M., Windarto, A. P., Damanik, I. S., & Satria, H. 2019. Analisa Metode *K-Means* pada Pengelompokan Kriminalitas Menurut Wilayah. Seminar Nasional Sains & Teknologi Informasi (SENSASI), Vol. 1 Issue.1 Juli, 2019
- Dicoding, 2021, Contoh *Use Case Diagram* Lengkap dengan Penjelasannya, <https://www.dicoding.com/blog/contoh-use-case-diagram/>
- Habrum, R. J., Khairan, A., Studi, P., Informatika, T., Teknik, F., Khairun, U., & Selatan, K. T. 2019, Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (PPN) Ternate Menggunakan Algoritma *K-Means*, Vol. 2 Issue.1 April, 2019
- Hidayati, R., Zubair, A., Pratama, A. H., & Indiana, L. 2021. Analisis *Silhouette Coefficient* pada 6 Perhitungan Jarak *K-Means Clustering Silhouette Coefficient Analysis in 6 Measuring Distancesof K-Means Clustering*. Techno.com, Vol. 20 Issue.2 Mei, 2021
- Khesya, N. 2021. Mengenal *Flowchart* dan *Pseudocode* Dalam Algoritma dan Pemrograman. *Preprints*, Vol. 1 Issue.1 Januari, 2021
- Mayona, Y., Buaton, R., & Simanjutak, M. 2022. Data *Mining Clustering* Tingkat Kejahatan Dengan Metode Algoritma *K-Means* (Studi Kasus: Kejaksaan Negeri Binjai). Jurnal Informatika Kaputama (JIK), E-ISSN:2685-5240, Vol. 6 Issue.3 Agustus, 2022
- Muzakkij, I. Z., Husein, K., Antonius, K., Hidayat, K. R., Emir, E., Haryanto, D., & Paryudi, I. 2023. Penggunaan Algoritma *K-Means* Pada Metode *Clustering*

- Untuk Menganalisa Tindak Kriminal. *Journal of Informatics and Advanced Computing* (JIAC), E-ISSN:2722-0346, Vol. 4 Issue.1 Mei, 2023
- pakdosen. (2023). Kriminalitas, <https://pakdosen.co.id/kriminalitas-adalah/>
- Praniffa, A. C., Syahri, A., Sandes, F., Fariha, U., Giansyah, Q. A., & Hamzah, M. L. (2023). Pengujian *Black Box* Dan *White Box* Sistem Informasi Parkir Berbasis Web *Black Box and White Box Testing of Web-Based Parking Information System*. *Jurnal Testing Dan Implementasi Sistem Informasi*, Vol. 1 Issue.1 Januari, 2023
- Putra, A. D., Martha, G. S., Fikram, M., & Yuhan, R. J. 2021. Faktor-Faktor yang Memengaruhi Tingkat Kriminalitas di Indonesia Tahun 2018. *Indonesian Journal of Applied Statistics*, Vol. 3 Issue.2 November, 2020
- Noviantoro, A., Silviana, A. B., Fitriani, R. R., & Permatasari, H. P. 2022. Rancangan Dan Implementasi Aplikasi Sewa Lapangan Badminton Wilayah Depok Berbasis Web. *Jurnal Teknik Dan Science*, Vol. 1 Issue.2 Juli, 2022
- Rusmanto, Saputra, I., & Makhriati. 2019. Pengaruh Tingkat Pendidikan, Latar Belakang Pendidikan, Pengalaman Dan Faktor Sosial Terhadap Pemanfaatan Catatan Atas Laporan Keuangan (CALK) Pada Skpd Kota Banjarmasin. *Dinamika Ekonomi Jurnal Ekonomi Dan Bisnis*, Vol. 12 Issue.2 September, 2019
- Saragih, R. P., Zinaidi, M., & Hafizah, H. 2023. Implementasi Data *Mining* Dalam Pengelompokkan Data Potensi Kejahatan Menggunakan Algoritma *K-Means Clustering* Pada Kapolsek Beringin. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, Vol. 2 Issue.4 Juli, 2023
- Septiani, N., & Erwansyah, K. 2020. Analisis Data *Mining* Pengelompokkan Kasus Tindak Kejahatan Yang Terjadi Di Kecamatan Medan Polonia Dengan Menggunakan Metode *K-Means Clustering*. *Jurnal Cyber Tech*, E-ISSN:2675-9802, Vol. 3 Issue.2 Februari, 2020
- Sri Fastaf, C. A., & Yamasari, Y. 2022. Analisa Pemetaan Kriminalitas Kabupaten Bangkalan Menggunakan Metode *K-Means* dan *K-Means++*. *Journal of Informatics and Computer Science* (JINACS), ISSN:2686-2220, Vol. 3 Issue.4 April, 2022

- Struyf, A., Hubert, M., & Rousseeuw, P. J. (1996). *Clustering in an object-oriented environment*. *Journal of Statistical Software*, Vol. 1 Issue.1 March 1996
- Suriani, L. 2020. Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma *K-Means Clustering*. *Jurnal Sistem Komputer Dan Informatika (JSON)*, e-ISSN:2685-998X, Vol. 1 Issue.2 Januari, 2020
- Suryadi, U. T., & Supriatna, Y. 2019. Sistem *Clustering* Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma *K-Means*. *Jurnal Teknologi Dan Komunikasi STMIK Subang*, ISSN:2252-4517, Vol. 12 Issue.1 April, 2019
- Sutejo, D., Agus Pranoto, Y., & Zulfia Zahro', H. 2020. Sistem Informasi Geografis Pengelompokan Tingkat Kriminalitas Kota Malang Menggunakan Metode. *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 4 Issue.1 Juli, 2020
- Tambunan, H. B., Barus, D. H., Hartono, J., Alam, A. S., Nugraha, D. A., & Usman, H. H. H. 2020. *Electrical peak load clustering analysis using K-means algorithm and silhouette coefficient*. *Proceeding - 2nd International Conference on Technology and Policy in Electric Power and Energy*, Vol. 1 Issue.4 Oktober, 2020
- Tampubolon, H. D., Suhada, S., Safii, M., Solikhun, S., & Suhendro, D. 2021. Penerapan Algoritma *K-Means* dan *K-Medoids Clustering* untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi. *Jurnal Ilmu Komputer Dan Teknologi*, ISSN:2746-4237, Vol. 2 Issue.2 November, 2021
- Titania Pricillia, & Zulfachmi. 2021. *Survey Paper*: Perbandingan Metode Pengembangan Perangkat Lunak (*Waterfall*, *Prototype*, RAD). *Jurnal Komputasi*, Vol. 1 Issue.1 Maret, 2021
- Usman, A. 2022. Fungsi Pemasyarakatan Bagian Satu, <https://bpsdm.kemenkumham.go.id/informasi-publik/publikasi/pojok-penyaluhan-hukum/fungsi-pemasyarakatan-bagian-satu>
- Wikipedia, 2022, Rumah Tahanan Negara, https://id.wikipedia.org/wiki/Rumah_Tahanan_Negara
- Wulandari, S. (2015). Fungsi Sistem Pemasyarakatan Dalam Merehabilitasi Dan

Mereintegrasi Sosial Warga Binaan Pemasyarakatan. Serat Acitya – Jurnal Ilmiah UNTAG Semarang, Vol. 4 *Issue.2* Februari, 2019