

SKRIPSI

**IMPLEMENTASI ALGORITMA *CONSTRAINED K-MEANS* DALAM
SISTEM ZONASI SEKOLAH UNTUK OPTIMASI PENERIMAAN
PESERTA DIDIK BARU**

***IMPLEMENTATION OF CONSTRAINED K-MEANS ALGORITHM IN
SCHOOL ZONING SYSTEM TO OPTIMIZE NEW STUDENT ADMISSION***



WIWI LESTIANI

D0221095

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS SULAWESI BARAT

MAJENE

2025

LEMBAR PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA *CONSTRAINED K-MEANS* DALAM SISTEM ZONASI SEKOLAH UNTUK OPTIMASI PENERIMAAN PESERTA DIDIK BARU

Telah dipersiapkan dan disusun oleh :

WIWI LESTIANI

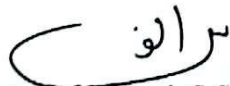
D0221095

Telah dipertahankan di depan Tim Penguji

Pada tanggal 09 Mei 2025

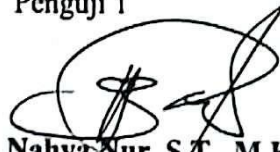
Susunan Tim Penguji

Pembimbing I



Dr. Eng. Sulfayanti, S.Si., M.T
NIP. 198903172020122011

Penguji I



Nahya Nur, S.T., M.Kom.
NIP.199111062019032024

Pembimbing II



Wawan Firgiawan, S.T., M.Kom.
NIDK.8948080023

Penguji II



Nurhilma Arifin, S.Kom., M.T.
NIP. 199304252022032011

Penguji III



A. Amrudi Asnan Cirua, S.T., M.Kom.
NIP.199804022024061001

LEMBAR PENGESAHAN

IMPLEMENTASI ALGORITMA *CONSTRAINED K-MEANS* DALAM
SISTEM ZONASI SEKOLAH UNTUK OPTIMASI PENERIMAAN
PESERTA DIDIK BARU

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Wiwi Lestiani
D0221095

Telah diperiksa dan disetujui oleh:

Pembimbing I

Pembimbing II



Dr. Eng. Sulfayanti, S.Si., M.T
NIP: 198903172020122011



Wawan Firgiawan, S.T., M.Kom
NIDK: 89480880023

Dekan Fakultas Teknik,

Ketua Program Studi Informatika



Dr. Ir. Hafsah Nirwana, M.T
NIP. 1964040519900322002



Mh. Rully Rasyid, S.Kom., M.T
NIP. 198808182022031006

PERNYATAAN ORISINALITAS

Menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah di ajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU N0. 20 Tahun 2003, pasal 25 ayat 2 dan pasal 70)

Majene 09 Mei 2025



Wiwi Lestiani

NIM.D0221095

ABSTRAK

Penerimaan Peserta Didik Baru (PPDB) merupakan proses penting dalam sistem pendidikan Indonesia. Untuk meningkatkan pemerataan akses pendidikan, pemerintah menerapkan sistem zonasi yang mengharuskan calon siswa memilih sekolah terdekat dari tempat tinggal. Namun, implementasinya masih menghadapi tantangan, seperti ketimpangan distribusi siswa dan kurangnya metode zonasi yang optimal. Penelitian ini bertujuan mengatasi masalah tersebut dengan menerapkan algoritma *Constrained K-Means* pada sistem zonasi sekolah dasar di Kabupaten Majene. Dataset yang digunakan terdiri dari 28 lokasi sekolah dan 1.220 data lokasi siswa. *Constrained K-Means* diterapkan untuk mengelompokkan siswa berdasarkan jarak geografis dengan mempertimbangkan kapasitas tiap sekolah. Nilai k ditentukan berdasarkan jumlah sekolah, dan penyusunan anggota tiap kluster disesuaikan dengan kapasitas sekolah menggunakan pendekatan *Linear Programming Algorithm* (LPA). Jarak antara siswa dan sekolah dihitung menggunakan metode *Euclidean Distance*, sedangkan evaluasi kualitas klusterisasi dilakukan menggunakan *Davies-Bouldin Index* (DBI). Hasil penelitian menunjukkan bahwa metode ini mampu mengoptimalkan distribusi siswa ke dalam 28 sekolah dengan jumlah siswa yang hampir merata sesuai kapasitas. Nilai DBI sebesar 1,6683 menunjukkan bahwa pemisahan antar-kluster cukup baik dan distribusi tergolong optimal. Sistem ini memberikan solusi yang lebih efisien, adil, dan berbasis data untuk kebijakan zonasi. Ke depan, sistem ini dapat dikembangkan dengan integrasi web berbasis peta interaktif guna mendukung transparansi dan evaluasi berkala. Penelitian ini diharapkan dapat menjadi acuan bagi pemerintah daerah dalam merancang kebijakan zonasi yang lebih tepat sasaran, mendorong pemerataan pendidikan, dan mengurangi kesenjangan antar sekolah.

Kata Kunci : Penerimaan Peserta Didik Baru, Sistem Zonasi Sekolah, Klusterisasi, *Constrained K-Means*, *Davies-Bouldin Indeks*

ABSTRACT

Student Admission (PPDB) is an important process in Indonesia's education system. To promote equal access to education, the government implements a zoning system that requires prospective students to choose the school closest to their residence. However, its implementation still faces challenges such as uneven student distribution and the lack of an optimal zoning method. This study aims to address these issues by applying the Constrained K-Means algorithm to the elementary school zoning system in Majene Regency. The dataset used consists of 28 school locations and 1.220 student location data points. Constrained K-Means is applied to cluster students based on their geographic distance to schools while considering each school's capacity. The value of k is determined by the number of schools, and the assignment of students to clusters is adjusted according to school capacities using the Linear Programming Algorithm (LPA). The distance between students and schools is calculated using the Euclidean Distance method, while the quality of clustering is evaluated using the Davies-Bouldin Index (DBI). The results show that this method successfully optimizes student distribution across 28 schools, with a nearly equal number of students per school according to their capacities. The obtained DBI score of 1,6683 indicates that the separation between clusters is sufficiently good and the distribution is optimal. This system offers a more efficient, equitable, and data-driven approach to zoning policy. In the future, this system can be further developed by integrating a web-based platform with interactive map visualizations to support transparency and routine policy evaluation. This study is expected to serve as a reference for local governments in designing more accurate zoning policies, promoting educational equity, and reducing disparities between schools.

Keywords: *Student Admission, School Zoning System, Clustering, Constrained K-Means, Davies-Bouldin Index*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan memiliki peran yang sangat vital dalam kemajuan dan perkembangan suatu bangsa. Sebagai salah satu indikator utama pembangunan manusia, pendidikan mencerminkan kualitas sumber daya manusia yang ada di dalam suatu negara. Kualitas hidup suatu bangsa sangat dipengaruhi oleh tingkat pengetahuan dan keterampilan masyarakatnya. Semakin tinggi standar pendidikan di suatu negara, maka semakin tinggi pula tingkat kesejahteraan dan kemajuan yang dapat dicapai oleh masyarakatnya. Pendidikan bukan hanya sekadar proses transfer ilmu pengetahuan, tetapi juga merupakan proses pembentukan karakter dan pengembangan potensi individu untuk melahirkan generasi penerus yang berkualitas dan berdaya saing tinggi. Menjadi bangsa yang maju dan berkompeten adalah cita-cita setiap negara di dunia, termasuk Indonesia (Loho et al., 2022).

Di Indonesia, sistem pendidikan terus mengalami perubahan dan perbaikan seiring dengan perkembangan zaman dan kebutuhan masyarakat. Setiap tahun ajaran baru, lembaga pendidikan di Indonesia menerapkan sistem penerimaan siswa baru. Penerimaan Peserta Didik Baru (PPDB) merupakan program tahunan yang krusial bagi lembaga pendidikan dalam menjaring calon peserta didik. Proses ini tidak hanya bertujuan untuk mengisi kuota siswa di sekolah, tetapi juga sebagai upaya untuk meningkatkan kualitas pendidikan dan pemerataan akses pendidikan bagi seluruh masyarakat (Loho et al., 2022).

Terkait dengan PPDB, pemerintah melalui Kementerian Pendidikan dan Kebudayaan telah mengeluarkan berbagai kebijakan untuk meningkatkan efektivitas dan efisiensi proses penerimaan siswa baru. Salah satu kebijakan penting yang diterapkan adalah Peraturan Menteri Pendidikan dan Kebudayaan Nomor 51 Tahun 2018 tentang Penerimaan Peserta Didik Baru. Kebijakan ini mengatur tentang mekanisme dan prosedur penerimaan peserta didik baru, termasuk penerapan sistem zonasi sekolah sebagai strategi untuk pemerataan akses pendidikan (Kurniawan & Saputra, 2019).

Sistem zonasi merupakan aturan yang mengharuskan calon peserta didik untuk memilih sekolah yang berada dalam radius terdekat dengan tempat tinggalnya. Tujuan utama dari sistem ini adalah untuk menjamin pemerataan akses layanan pendidikan bagi seluruh masyarakat, menghilangkan stigma atau diskriminasi antar sekolah, serta mendekatkan hubungan antara sekolah dengan lingkungan keluarga peserta didik. Dengan sistem zonasi, diharapkan tidak ada lagi istilah sekolah favorit atau unggulan, karena setiap sekolah memiliki kesempatan yang sama dalam menerima peserta didik (Risal et al., 2022).

Namun, dalam pelaksanaannya, sistem zonasi masih menghadapi berbagai tantangan dan permasalahan. Salah satunya adalah adanya kesenjangan sosial ekonomi di masyarakat yang berdampak pada persebaran peserta didik di sekolah-sekolah tertentu. Terdapat perbedaan antara peserta didik dari golongan ekonomi menengah ke atas dengan golongan ekonomi menengah ke bawah, yang menyebabkan perbedaan kasta antar sekolah, yaitu antara sekolah unggulan dan sekolah reguler. Sekolah unggulan cenderung diisi oleh peserta didik dari golongan ekonomi menengah ke atas, sedangkan sekolah reguler diisi oleh peserta didik dari golongan ekonomi menengah ke bawah. Hal ini menimbulkan ketidakadilan dan ketimpangan dalam akses dan kualitas pendidikan (Zainuddin & Risal, 2024).

Selain itu, penerapan sistem zonasi juga menghadapi kendala dalam pengelolaan data dan penentuan wilayah zonasi yang efektif dan efisien. Banyaknya jumlah calon peserta didik dan sekolah yang tersebar di berbagai wilayah memerlukan sistem yang mampu mengelola dan mengelompokkan data secara akurat. Dalam konteks inilah teknologi dan metode data mining dapat berperan penting.

Data *mining* merupakan proses ekstraksi informasi atau pola-pola penting dari data yang besar dan kompleks. Salah satu metode data mining yang relevan dalam konteks zonasi sekolah adalah klasterisasi (*clustering*), yaitu metode pengelompokan data tanpa arahan (*unsupervised*). Klasterisasi bertujuan untuk mengelompokkan data ke dalam klaster-klaster sehingga data dalam satu klaster memiliki karakteristik yang serupa (Ika, 2020).

Ada dua jenis klasterisasi data yang umum digunakan, yaitu *hierarchical clustering* dan *non-hierarchical clustering*. Salah satu metode *non-hierarchical*

clustering yang populer adalah *K-Means*. Metode *K-Means* bekerja dengan mempartisi data ke dalam sejumlah klaster yang ditentukan, dengan tujuan meminimalkan variasi dalam klaster dan memaksimalkan variasi antar klaster. Dalam konteks zonasi sekolah, metode *K-Means* dapat digunakan untuk mengelompokkan wilayah berdasarkan jarak geografis antara tempat tinggal calon peserta didik dan lokasi sekolah (Ika, 2020).

Penerapan algoritma *K-Means* dalam sistem zonasi sekolah memiliki beberapa keuntungan. Pertama, algoritma ini dapat menangani data dalam jumlah besar dengan efisien. Kedua, *K-Means* dapat menghasilkan klaster-klaster yang seimbang dan optimal, sehingga distribusi peserta didik ke sekolah-sekolah dapat merata. Ketiga, metode ini dapat membantu dalam pengambilan keputusan berbasis data (*data-driven decision making*), yang penting dalam perencanaan dan pengelolaan pendidikan (Zainuddin & Risal, 2024). Tetapi masuk dalam beberapa kasus, masalahnya ditemukan pada batasan ukuran *cluster*, yang mengharuskan setiap *cluster* berada dalam rentang yang sama. Dalam kasus tertentu, *K-Means* sering dikombinasikan dengan satu atau lebih *cluster*. Jadi *K-Means* merangkum sejumlah data kecil. Selain itu, algoritma pengelompokan *K-Means* rentan terhadap *noise* pada hasil *clustering* data sehingga mengakibatkan tidak seimbang *cluster* data yang bahkan tidak optimal. Sementara itu, sifat komputasional *Metrik Euclidean* mendefinisikan bahwa algoritma *K-Means* dalam proses komputasi pengelompokan hanya dapat diimplementasikan pada kumpulan data yang berbentuk pola melingkar dengan dimensi dan skala kerapatan data yang tepat (Risal et al., 2022).

Penelitian ini didasarkan pada penelitian sebelumnya oleh Andi Alviadi Nur Risal (2022) yang berjudul “*School Zoning System for Student Admission using Constrained K-Means Algorithms*”. Hasil klasterisasi menunjukkan bahwa penyaluran siswa lebih efektif dengan jumlah anggota klaster yang seimbang. Dengan jumlah sekolah sebanyak 22 sekolah dan jumlah calon siswa baru sebanyak 2248 orang, menghasilkan anggota klaster pada setiap sekolah, yaitu sebanyak 103 siswa pada 21 sekolah dan 85 siswa pada satu sekolah. Sehingga tidak terjadi tumpang tindih anggota klaster, dan tidak ada satu pusat klaster yang tidak memiliki anggota klaster. Sistem ini dapat dijadikan acuan oleh pemerintah dalam membuat

sistem zonasi sekolah yang efisien dan akurat. Di Kabupaten Majene, penerapan sistem zonasi sekolah masih menghadapi tantangan dalam hal pemerataan akses pendidikan dan pengelolaan data calon peserta didik. Kabupaten Majene sebagai salah satu wilayah di Provinsi Sulawesi Barat memiliki karakteristik geografis dan demografis yang unik, dengan sebaran penduduk yang bervariasi. Oleh karena itu, diperlukan pendekatan yang tepat untuk mengoptimalkan sistem zonasi sekolah, khususnya pada jenjang Sekolah Dasar (SD).

Berdasarkan uraian di atas dan beberapa hasil penelitian sebelumnya, penelitian ini bertujuan untuk menerapkan metode *Constrained K-means* dalam menyederhanakan penerimaan peserta didik baru di Kabupaten Majene dengan sistem zonasi. Dengan menggunakan algoritma *Constrained K-means*, diharapkan dapat diperoleh pengelompokan wilayah yang optimal berdasarkan jarak geografis antara tempat tinggal calon peserta didik dan lokasi sekolah. Hal ini diharapkan dapat meningkatkan efisiensi dan keadilan dalam proses penerimaan peserta didik baru.

Oleh karena itu, penulis mengangkat penelitian dengan judul “Implementasi Algoritma *Constrained K-Means* Dalam Sistem Zonasi Sekolah Untuk Menyederhanakan Penerimaan Peserta Didik Baru”. Dalam penelitian ini, pengumpulan data meliputi data tempat tinggal calon peserta didik baru dan koordinat lokasinya, serta data lokasi sekolah yang terdiri dari atribut koordinat lintang dan bujur. Seluruh data tersebut berada di Kabupaten Majene, Sulawesi Barat. Data sekolah yang digunakan mengacu pada jenjang pendidikan Sekolah Dasar (SD).

Dengan penerapan algoritma *Constrained K-means* dalam sistem zonasi sekolah, diharapkan dapat membantu pemerintah daerah dan instansi terkait dalam meningkatkan kualitas dan pemerataan pendidikan. Selain itu, hasil penelitian ini diharapkan dapat menjadi acuan bagi pengembangan sistem zonasi sekolah di daerah lain yang memiliki karakteristik serupa.

1.2 Rumusan Penelitian

Berdasarkan latar belakang yang telah dibahas sebelumnya maka rumusan masalah pada penelitian ini yaitu bagaimana hasil implementasi algoritma

Constrained K-Means dapat digunakan untuk mengoptimalkan sistem zonasi penerimaan peserta didik baru tingkat Sekolah Dasar di Kabupaten Majene ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, adapun tujuan dari penelitian ini yaitu untuk mengoptimalkan sistem zonasi penerimaan peserta didik baru dengan algoritma *Constrained K-Means* guna mengelompokkan wilayah berdasarkan jarak geografis, guna meningkatkan efisiensi dan keadilan dalam penerimaan peserta didik baru di Kabupaten Majene.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat dalam penelitian selanjutnya yang berkaitan dengan algoritma *Constrained K-Means* dalam sistem zonasi sekolah. Hasil penelitian ini juga dapat menjadi acuan bagi pengembangan metode zonasi yang lebih efisien dan efektif di masa depan.

1.5 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah yang ditetapkan untuk menjaga fokus dan kejelasan penelitian, yaitu:

1. Data yang diambil adalah data Sekolah Jenjang SD di Kabupaten Majene Kecamatan Banggae.
2. Algoritma yang digunakan dalam penelitian ini yaitu algoritma *Constrained K-means*.
3. Hasil *clustering* dalam penelitian ini hanya menghitung berdasarkan pada jarak terdekat antara lokasi siswa dan sekolah dengan mempertimbangkan batasan tertentu.

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Pendidikan

Menurut Undang-undang Nomor. 20 Tahun 2003 tentang Sistem Pembelajaran Nasional mengatakan bahwa “Pendidikan merupakan usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran, agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, akhlak mulia, serta ketrampilan yang diperlukan dirinya, masyarakat, bangsa dan negara”. Pendidikan bertujuan agar berkembangnya kemampuan partisipasi peserta didik supaya jadi manusia yang beriman serta bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, serta jadi masyarakat negeri yang demokratis dan bertanggung jawab (Fajri, 2021).

Pendidikan pada dasarnya merupakan proses untuk membantu manusia dalam mengembangkan potensi dirinya sehingga mampu menghadapi setiap perubahan yang terjadi. Dalam rangka pembangunan manusia Indonesia seutuhnya, pembangunan di bidang pendidikan merupakan suatu sarana penting dalam peningkatan kualitas sumber daya manusia. Oleh karena itu, pendidikan perlu mendapatkan perhatian dari berbagai pihak seperti pemerintah, pengelola pendidikan, dan masyarakat. Sesuai dengan perkembangan masyarakat Indonesia, pendidikan banyak menghadapi berbagai hambatan dan tantangan. Salah satu tantangan yang cukup menarik adalah masalah ketimpangan pendidikan. Hal ini diperkuat bahwa ketimpangan pendidikan masih terjadi, seperti ketimpangan pendidikan di perkotaan dan pedesaan. Ketimpangan pendidikan tersebut meliputi akses menuju sekolah, sarana prasarana sekolah seperti bangunan fisik, dan sumber daya guru (tenaga pengajar). Sarana prasarana pendidikan di perkotaan sangat lengkap dan kualitas guru sangat tinggi. Di sisi lain, sarana prasarana di pedesaan sangat terbatas dan kualitas guru masih rendah. Ketimpangan pendidikan pada akhirnya menyebabkan ketimpangan sumber daya manusia (Pangestuti, 2021).

2.1.2 Sekolah

Sekolah merupakan lingkungan pendidikan yang secara sengaja dirancang dan dilaksanakan dengan aturan-aturan yang ketat seperti harus berjenjang dan berkesinambungan, sehingga disebut pendidikan formal dan sekolah adalah lembaga khusus, suatu wahana, suatu tempat untuk menyelenggarakan pendidikan, yang didalamnya terdapat suatu proses belajar mengajar untuk mencapai tujuan pendidikan tertentu. Sekolah juga merupakan salah satu tempat yang efektif bagi pembentukan karakter seorang individu. Sejak dahulu, sekolah telah memiliki tujuan utama di bidang pendidikan, yaitu membentuk manusia yang cerdas dan juga memiliki watak atau karakter yang baik. Sekolah membantu orang tua mengajarkan kebiasaan-kebiasaan yang baik serta menanamkan budi pekerti, dan nilai-nilai religi dan budaya. Sekolah mengembangkan potensi anak dalam ranah *kognitif*, *afektif*, dan *psikomotorik*. Karakter terbentuk dari hasil cara pandang, bersikap dan bertingkah laku yang tampak dalam kehidupan sehari-hari ketika berinteraksi. Karakter dapat menjadi ciri, karakteristik, atau sifat khas seseorang atau sekelompok orang (Shinta & Ain, 2021).

Sekolah memegang peran yang sangat penting dalam pengembangan sumber daya manusia yang unggul dan berkualitas baik secara kognitif maupun budi pekerti dalam menghadapi tantangan abad 21 dengan perkembangan teknologi yang begitu pesat. Adapun kemampuan yang harus dikuasai siswa di abad 21 yang terdiri atas kemampuan berpikir kritis, berpikir kreatif, komunikasi dan kolaborasi dibarengi dengan kecerdasan karakter sehingga pada akhirnya siswa yang memiliki profil pelajar pancasila. Kemampuan siswa abad 21 ini akan dapat dihasilkan dari sekolah yang menerapkan konsep merdeka belajar. Penerapan kebijakan merdeka belajar di sekolah sangat ditentukan oleh kepemimpinan sekolah. Hal ini dikarenakan, kepala sekolah mempunyai peran sentral dalam pengambilan kebijakan untuk kemajuan sekolah. Kepala sekolah harus memahami beban kerjanya sebagai kepala sekolah yang terdiri atas tugas pokok manajerial, pengembangan kewirausahaan, dan supervisi. Pemahaman akan beban kerja kepala sekolah akan mampu menjadi pegangan bagi kepala sekolah untuk menerapkan merdeka belajar di sekolah (Nuryoso et al., 2023).

2.1.3 Sistem Zonasi

Sistem zonasi adalah suatu sistem dimana sekolah harus menerima peserta didik yang memiliki radius tempat tinggal terdekat dengan sekolah. Sistem Zonasi diterapkan sejak adanya Peraturan Menteri Pendidikan dan Kebudayaan Nomor 14 Tahun 2018 Pasal 16 ayat (1) tentang Penerimaan Peserta Didik Baru pada Taman Kanak-Kanak, Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas, Sekolah Menengah Kejuruan, atau bentuk lain yang sederajat, menegaskan bahwa sekolah yang berada di bawah naungan pemerintah wajib menerima calon peserta didik yang berdomisili pada radius zona terdekat dari sekolah paling sedikit sebesar 90% dari total jumlah peserta didik yang diterima (Widyaningtyas et al., 2021). Tujuan utama dari peraturan baru ini adalah untuk mempromosikan peningkatan aksesibilitas layanan pendidikan, seperti yang tercantum dalam Pasal 3 Ayat 1 Permendikbud Nomor 51 Tahun 2018. Dengan demikian, dapat disimpulkan bahwa fokus utama kebijakan zonasi adalah untuk memastikan keterjangkauan yang merata terhadap layanan pendidikan (Pendidikan, 2024).

2.1.4 Peserta Didik

Siswa atau yang biasa disebut dengan peserta didik merupakan salah satu dari komponen pendidikan yang tidak bisa ditinggalkan, karena tanpa adanya peserta didik tidak akan mungkin proses pembelajaran dapat berjalan. Peserta didik merupakan komponen manusiawi yang menempati posisi sentral dalam proses pembelajaran. Di dalam proses pembelajaran, peserta didik sebagai pihak yang ingin meraih cita-cita, memiliki tujuan dan kemudian ingin mencapainya secara optimal. Dengan demikian peserta didik adalah orang yang mempunyai pilihan untuk menempuh ilmu sesuai dengan cita-cita dan harapan masa depan (DR. Hamuni Muhammad Idrus & Dra Aswati, 2022).

2.1.5 *Clustering*

Clustering merupakan metode dalam data *mining* dan penggunaannya tanpa adanya arahan bahkan dapat bekerja dengan berbagai jenis pegelompokkan *clustering* sesuai dengan data *mining* yang dipergunakan seperti *hierarchical clustering* dan *non-hierarchical clustering*. Metode ini selalu menggunakan

partipasi terpisah dengan *cluster* tertentu untuk berjalan sesuai dengan karakteristik data yang digunakan dan kemiripan dengan kelompok satu *cluster* tertentu dan dapat digabungkan dengan kelompok lainnya. Metode data *clustering* juga pastinya memiliki tujuan yang dapat digabungkan dalam observasi objek tertentu sesuai dengan metode awal yang dipergunakan dengan kelompok minimalisasi *objective function* pada dataset dalam proses *clustering* (Nur Aziz & Zuliarso, 2022).

2.1.6 Data Mining

Data *Mining* atau yang bisa disebut *Knowledge Discovery in Databases* (KDD) adalah suatu proses pengerukan atau pengumpulan informasi penting dari suatu data yang besar atau dapat dikatakan sebagai suatu proses pengambilan pola pada data yang akan di proses dengan *output* yang berupa informasi penting (Loho et al., 2022).

Data *mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Dewi, 2020):

a) Deskripsi.

Deskripsi digunakan untuk pengelolaan data mining dengan cara menggambarkan pola kecenderungan dalam suatu data. Deskripsi dalam data mining ini lebih menjelaskan kepada pola-pola yang terbentuk akibat adanya penambahan data.

b) Estimasi.

Estimasi pada data *mining* hampir sama dengan klasifikasi akan tetapi target dari estimasi ini lebih ke hasil yang sifatnya numerik, bukan kategori. Beberapa teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

c) Prediksi.

Prediksi memiliki beberapa kesamaan dengan klasifikasi dan estimasi, hasil yang didapat pada prediksi ini akan ada pada kejadian yang akan datang.

d) Klasifikasi

Klasifikasi pada data *mining* digunakan untuk mengelompokkan sesuatu berdasarkan kategori. Kategori di sini berarti sesuatu yang sudah digolongkan, misal gaji rendah, sedang, dan tinggi, curah hujan rendah, sedang, dan tinggi, dan sebagainya.

e) Pengklusteran

Pengklusteran berbeda dengan klasifikasi, pengklusteran lebih kepada pengelompokan data berdasarkan kemiripan. Pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (*homogen*), dimana kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

f) Asosiasi.

Asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (*Market Basket Analysis*). Contoh asosiasi dalam bisnis adalah mencari barang dalam supermarket yang dibeli secara bersamaan dalam suatu waktu. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu *persentase* kombinasi item tersebut dalam *database* dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiatif.

2.1.7 *Machine Learning*

Machine Learning adalah bagian dari kecerdasan buatan (AI) yang membantu komputer atau mesin pengajaran belajar dari semua data sebelumnya dan membuat keputusan yang cerdas. Kerangka kerja *machine learning* memerlukan penangkapan dan pemeliharaan serangkaian informasi yang kaya dan mengubahnya menjadi basis pengetahuan terstruktur untuk penggunaan yang berbeda di berbagai bidang, salah satunya di bidang pendidikan (Fathurohman, 2021).

Machine learning terbagi menjadi tiga kategori, yaitu: *Supervised Learning*, *Unsupervised Learning*, *Reinforcement Learning*. Skema keterkaitan *artificial*

intelligence dan *machine learning* dapat dijelaskan sebagai berikut (Roihan et al., 2020):

a) *Supervised Learning*

Supervised Learning adalah metode klasifikasi di mana kumpulan data sepenuhnya diberikan label untuk mengklasifikasikan kelas yang tidak dikenal.

b) *Unsupervised Learning*

Unsupervised Learning sering disebut cluster dikarenakan tidak ada kebutuhan untuk pemberian label dalam kumpulan data dan hasilnya tidak mengidentifikasi contoh di kelas yang telah ditentukan.

c) *Reinforcement Learning*

Reinforcement Learning biasanya berada antara *Supervised Learning* dan *Unsupervised Learning*, teknik ini bekerja dalam lingkungan yang dinamis di mana konsepnya harus menyelesaikan tujuan tanpa adanya pemberitahuan dari komputer secara eksplisit jika tujuan tersebut telah tercapai.

2.1.8 Algoritma *Constrained K-Means*

Constrained K-Means adalah pengembangan dari algoritma *K-Means* yang menambahkan batasan tertentu dalam proses klusterisasi untuk meningkatkan keseimbangan dan keakuratan hasil klaster. Algoritma ini mengatasi kelemahan *K-Means*, seperti ketidakseimbangan ukuran klaster dan keterbatasan dalam menangani dataset dengan distribusi yang tidak seragam. Dalam *Constrained K-Means*, pembentukan klaster tidak hanya bergantung pada jarak *Euclidean*, tetapi juga mempertimbangkan batasan tertentu, seperti jumlah minimum objek dalam setiap klaster. Algoritma ini menggunakan *Linear Programming Algorithm* (LPA) untuk mengoptimalkan distribusi data sehingga setiap klaster memiliki jumlah anggota yang sesuai dengan aturan yang telah ditentukan. Dengan pendekatan ini, *Constrained K-Means* mampu menghasilkan klusterisasi yang lebih optimal, terutama dalam skenario di mana distribusi data tidak berbentuk bola atau memiliki kepadatan yang tidak merata (Risal et al., 2022).

2.1.9 Davies-Bouldin index (DBI)

Davies Bouldin Index (DBI) merupakan sebuah metode yang digunakan untuk melakukan validasi terhadap *cluster* secara kuantitatif dari hasil *clustering*. Metode ini memiliki tujuan untuk memaksimalkan jarak *intercluster* antara satu *cluster* terhadap *cluster* yang lain. *Cluster* yang optimal adalah *cluster* yang memiliki nilai separasi yang tinggi dan nilai kohesi yang rendah. Nilai *Davies Bouldin Index* (DBI) yang semakin mendekati nilai 0 menandakan semakin baik *cluster* yang diperoleh. Semakin rendah nilai DBI menunjukkan hasil *cluster* yang optimal (Sholihin & Rudiman, 2022).

Sum of square within cluster (SSW) adalah Persamaan untuk mengetahui matrik kohesi dalam sebuah *cluster* ke-*i* yang dapat dilihat pada Persamaan (2.1).

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (2.1)$$

Keterangan :

m_i : jumlah data dalam *cluster* ke-*i*

c_i : centroid *cluster* ke-*i*

$d(x_j, c_i)$: jarak *euclidean* setiap data ke *centroid*

Sum of square between cluster (SSB) adalah persamaan untuk mengetahui nilai separasi antara *cluster* yang dapat dilihat pada Persamaan (2.2).

$$SSB_{i,j} = d(c_i, c_j) \quad (2.2)$$

Keterangan :

$d(c_i, c_j)$: jarak antar *centroid*

Setelah nilai separasi dan kohesi diperoleh, lalu dilakukan pengukuran rasio ($R_{(i,j)}$) untuk mengetahui nilai perbandingan antara *cluster* ke-*i* dan *cluster* ke-*j*, sesuai dengan Persamaan (2.3).

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (2.3)$$

Persamaan untuk menghitung nilai *Davies Bouldin Index* (DBI) sesuai dengan Persamaan (2.4).

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (2.4)$$

Keterangan :

d : jumlah *cluster* yang digunakan

Semakin rendah nilai *Davies Bouldin Index* (DBI) yang diperoleh, maka semakin baik kualitas *cluster* yang diperoleh dari suatu *clustering* data.

2.2 Penelitian Terkait

Dalam persiapan penelitian ini, penting untuk merujuk pada beberapa penelitian sebelumnya sebagai landasan dan wawasan. Hal ini bertujuan untuk memperoleh informasi dan ide-ide yang dapat dikembangkan dalam penelitian ini. Untuk itu, pada penelitian ini telah menghimpun sejumlah referensi dari penelitian-penelitian terdahulu yang relevan dengan permasalahan yang dibahas dalam penelitian ini.

Tabel 2. 1 Penelitian Terkait

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
1.	Andi Alviadi Nur Risal (2022)	<i>School Zoning System for Student Admission using Constrained K-Means Algorithms</i>	Hasil klasterisasi menunjukkan bahwa penyaluran siswa lebih efektif dengan jumlah anggota klaster yang seimbang. Dengan jumlah sekolah sebanyak 22 sekolah dan jumlah calon siswa baru sebanyak 2248 orang, menghasilkan anggota klaster pada setiap sekolah, yaitu sebanyak 103 siswa pada 21 sekolah dan 85 siswa pada satu sekolah. Sehingga tidak terjadi tumpang tindih anggota klaster, dan tidak ada satu pusat klaster yang tidak memiliki anggota klaster. Sistem ini dapat dijadikan acuan oleh pemerintah dalam membuat sistem zonasi sekolah yang efisien dan akurat.	Perbedaan utama antara penelitian yang akan dibuat dan penelitian Andi terletak pada jenis data yang digunakan, di mana penelitian yang akan saya buat tidak hanya mempertimbangkan koordinat lokasi sekolah dan siswa untuk membentuk klaster zonasi yang optimal, tetapi juga menambahkan batasan jumlah kuota sekolah dalam proses klasterisasi. Selain itu, penelitian yang akan dibuat menerapkan evaluasi klaster menggunakan <i>Davies-Bouldin Index</i> (DBI)	Persamaannya adalah sama-sama menerapkan algoritma <i>Constrained K-Means</i> , dengan berfokus pada pengelompokan data dengan pembatasan atau kendala tertentu, bukan semata-mata pada analisis spasial atau zonasi penerimaan sekolah.

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
				<p>untuk menilai kualitas klusterisasi yang dihasilkan. Evaluasi ini memungkinkan analisis lebih mendalam terhadap efektivitas zonasi dengan menilai pemisahan dan kepadatan antar klaster, sehingga menghasilkan pembagian zonasi yang lebih akurat, seimbang, dan sesuai dengan kapasitas masing-masing sekolah. Sedangkan penelitian Andi lebih menitikberatkan pada koordinat lokasi sekolah dan siswa untuk membentuk klaster zonasi yang lebih optimal tanpa menggunakan evaluasi klaster apapun.</p>	
2.	Ika Pusvitaningrum (2020)	Analisis Pengelompokan Data Argumen Tentang Penerapan Kebijakan Sistem Zonasi pada Pendaftaran Sekolah dengan Menggunakan Metode <i>Clustering</i>	Dari hasil pengolahan data pada kecamatan Tayu dan kecamatan Margoyoso dengan metode <i>K-means</i> , maka dapat ditarik kesimpulan bahwa pada kecamatan Tayu aspek terkuat yang mendasari pandangan masyarakat terhadap sistem zonasi untuk memilih sekolah yaitu aspek "Penerapan Kebijakan Sistem Zonasi" dan aspek "Sistem Pendaftaran Sekolah" sedangkan kecamatan	Penelitian Pusvitaningrum berfokus pada analisis argumen masyarakat mengenai penerapan kebijakan zonasi, menggunakan metode <i>clustering K-Means</i> pada data kuesioner yang berisi pandangan masyarakat tentang zonasi di tingkat SMA. Sedangkan penelitian yang	Penelitian Ika Pusvitaningrum (2020) dan penelitian saya memiliki kesamaan dalam penggunaan metode <i>clustering</i> untuk menganalisis sistem zonasi dalam pendaftaran sekolah. Keduanya berfokus pada pengelompokan data untuk memahami atau

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
			<p>Margoyoso aspek terkuat yang mendasari pandangan masyarakat terhadap sistem zonasi untuk memilih sekolah yaitu aspek "Pandangan Masyarakat Terhadap Label Sekolah" dan "Penerapan Kebijakan Sistem Zonasi"</p> <p>Penelitian ini menganalisis penerapan sistem zonasi dalam pendaftaran sekolah dengan metode clustering menggunakan <i>K-Means</i>. Data diambil dari 380 responden di Kecamatan Pati, yang terdiri dari 199 perempuan dan 181 laki-laki. Hasil analisis menunjukkan adanya lima aspek utama yang memengaruhi pandangan masyarakat terhadap sistem zonasi, yaitu pemilihan sekolah, label sekolah, kebijakan zonasi, sistem pendaftaran, dan dampaknya. Dengan <i>clustering</i>, ditemukan pola-pola tanggapan masyarakat yang bervariasi, yang membantu memahami pro dan kontra terhadap kebijakan tersebut serta memberikan acuan bagi sekolah dalam menghadapi penerimaan peserta didik baru.</p>	<p>akan dibuat berfokus pada implementasi algoritma <i>Constrained K-Means</i> dalam sistem zonasi sekolah dasar (SD) di Kabupaten Majene untuk menyederhanakan proses penerimaan peserta didik baru dengan mempertimbangkan koordinat lokasi sekolah dan siswa untuk membentuk kluster zonasi yang optimal dan menambahkan batasan jumlah kuota sekolah dalam proses klusterisasi.</p>	<p>mengoptimalkan zonasi dalam penerimaan peserta didik baru (PPDB). Selain itu, sama-sama bertujuan untuk mengevaluasi efektivitas sistem zonasi dalam pembagian wilayah sekolah berdasarkan data tertentu.</p>
3.	Wen Gie, Deny Jollyta (2020)	Perbandingan <i>Euclidean</i> dan <i>Manhattan</i> Untuk Optimasi Cluster	Berdasarkan hasil pengolahan 97 data status Covid-19 di provinsi Riau, <i>cluster</i> optimal diperoleh dari perhitungan <i>Euclidean</i>	perbedaannya terletak pada objek penelitian dan metode optimasi jarak yang digunakan.	Kesamaannya yaitu dalam penggunaan metode <i>clustering</i> dan evaluasi kluster

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
		Menggunakan <i>Davies Bouldin Index</i> : Status Covid-19 Wilayah Riau	dan DBI dengan menggunakan formula (3), (1) dan (2). Untuk cluster optimal dari perhitungan <i>Manhattan</i> digunakan formula (4), (1), (2). Pengujian data sebaran Covid-19 pada pengukuran jarak <i>Euclidean</i> dan <i>Manhattan</i> memiliki jumlah anggota <i>cluster</i> yang sama hanya pada $k=3$, namun dengan nilai DBI yang berbeda. <i>Euclidean</i> memiliki nilai DBI terendah pada $k=8$ dengan nilai 0,394. Nilai DBI terendah dari perhitungan <i>Manhattan</i> terdapat pada $k=7$ dengan nilai 0,434. Hasil <i>clustering</i> memperlihatkan bahwa DBI bekerja lebih baik pada <i>Euclidean</i> . Hal tersebut dimungkinkan karena formula <i>Euclidean</i> menggunakan akar kuadrat sehingga menghasilkan nilai DBI yang lebih kecil dibandingkan <i>Manhattan</i> yang perhitungannya menggunakan nilai mutlak.	Penelitian Wen membandingkan metode perhitungan jarak dan <i>Manhattan</i> dalam optima. <i>clustering</i> untuk status Covid-19 di wilayah Riau, sedangkan penelitian yang akan dibuat menggunakan <i>Constrained K-Means</i> dalam sistem zonasi PPDB dengan mempertimbangkan batasan jumlah kuota sekolah. Selain itu, penelitian saya lebih fokus pada pengelompokan siswa berdasarkan lokasi dan evaluasi efektivitas zonasi sekolah, bukan optimasi pemilihan metode jarak dalam <i>clustering</i>	dengan <i>Davies-Bouldin Index</i> (DBI). Keduanya bertujuan untuk mengukur kualitas klasterisasi. dengan DBI, yang membantu menentukan seberapa baik klaster terbentuk berdasarkan pemisahan dan kepadatan data
4.	Muhammad Mahda, Rudi Kurniawan, Tati Suprapti (2025)	Optimalisasi Algoritma <i>K-Means</i> untuk Analisis pengelompokan Data Jurusan Siswa Baru Berbasis <i>Numerical Measure</i>	Hasil penelitian menunjukkan bahwa <i>Euclidean Distance</i> memiliki pemisahan <i>cluster</i> yang lebih baik daripada <i>Manhattan Distance</i> , dengan nilai DBI terendah (0.603) pada $K=2$. Hasil ini menunjukkan bahwa <i>Euclidean Distance</i> lebih efektif dalam mengoptimalkan	Perbedaannya yaitu Penelitian Mahda berfokus pada pengelompokan jurusan siswa baru berbasis <i>Numerical Measure</i> , yang berarti analisisnya lebih condong pada pemilihan jurusan	Kesamaannya yaitu keduanya bertujuan untuk mengoptimalkan proses <i>clustering</i> dalam konteks pendidikan, dengan pendekatan berbasis data untuk mendukung pengambilan keputusan yang

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
			<p>pengelompokan data. Metode ini dapat diterapkan dalam analisis data pendidikan dan bidang lain.</p>	<p>berdasarkan karakteristik numerik tertentu. Sementara itu, penelitian yang akan dibuat lebih berorientasi pada sistem zonasi sekolah dalam PPDB menggunakan <i>Constrained K-Means</i>, yang mempertimbangkan batasan jumlah kuota sekolah dalam proses klusterisasi. Selain itu, saya menggunakan evaluasi kluster dengan <i>Davies-Bouldin Index (DBI)</i> untuk menilai kualitas pengelompokan, yang mungkin tidak menjadi fokus utama dalam penelitian mereka.</p>	<p>lebih efektif.</p>
5.	Zahir Zainuddin, Andi Alviadi Nur Risal (2024)	<i>Balanced clustering for student admission school zoning by parameter tuning of constrained k-means</i>	<p>Penelitian ini menggunakan data dari 2.248 calon siswa dan 22 lokasi sekolah negeri. Hasil pengujian algoritma k-means dalam pengelompokan menunjukkan terbentuknya pola tidak melingkar pada keanggotaan kluster dengan jumlah anggota kluster <i>centroid</i> yang berbeda-beda. Sebaliknya, pengujian algoritma <i>constrained k-means</i> menunjukkan hasil yang seimbang pada keanggotaan kluster dengan nilai keanggotaan sebesar 103 untuk masing-</p>	<p>Perbedaan utama terletak pada pendekatan optimasi dan evaluasi kluster. Penelitian Zahir lebih menitikberatkan pada parameter <i>tuning Constrained K-Means</i> untuk mencapai kluster yang lebih optimal, sementara penelitian yang akan dibuat tidak hanya mempertimbangkan batasan jumlah kuota sekolah, tetapi juga</p>	<p>Persamaannya yaitu dalam penggunaan <i>Constrained K-Means</i> untuk sistem zonasi penerimaan siswa (PPDB). Keduanya berfokus pada <i>balanced clustering</i>, yaitu pembentukan kluster dengan jumlah siswa yang seimbang untuk setiap sekolah. Selain itu, kedua penelitian sama-sama berusaha mengoptimalkan zonasi agar lebih efektif dan adil</p>

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
			masing sekolah sebagai pusat klaster. Hasil penelitian menyatakan bahwa algoritma <i>constrained k-means</i> yang dikembangkan mampu mengatasi masalah pengelompokan data yang tidak seimbang dan masalah tumpang tindih dalam proses penerimaan siswa baru. Dengan kata lain, algoritma <i>k-means</i> terbatas ini dapat menjadi acuan bagi pemerintah dalam mengambil keputusan penerimaan siswa baru.	menambahkan evaluasi kualitas klaster menggunakan <i>Davies-Bouldin Index (DBI)</i> .	bagi calon siswa.
6.	Yassin Achmad Nur Aziza, Eri Zuliarso (2022)	Sistem Penerimaan Siswa Baru Di SMKN3 Pati Berdasar Jalur Prestasi Menggunakan Algoritma <i>Klustering K-Means Berbasis Web</i>	Hasil penelitian Yassin Achmad Nur Aziz menunjukkan bahwa penerapan algoritma K-Means dalam sistem penerimaan siswa baru berbasis web di SMKN 3 Pati berhasil mengelompokkan siswa berdasarkan jalur prestasi. Pengelompokan ini menghasilkan tiga klaster nilai, yaitu M1 (nilai tinggi), M2 (nilai sedang), dan M3 (nilai rendah). Hasil klasterisasi ini membantu sekolah dalam menentukan siswa mana yang layak masuk ke kelas unggulan berdasarkan nilai akademik, sehingga sistem dapat mempercepat proses seleksi dengan lebih efektif dan efisien.	Penelitian Yassin lebih spesifik pada penerimaan siswa di SMKN3 Pati melalui jalur prestasi, sedangkan penelitian yang akan dibuat menggunakan <i>Constrained K-Means</i> untuk sistem zonasi PPDB, dengan mempertimbangkan batasan jumlah kuota sekolah dalam pembentukan klaster. Selain itu, penelitian yang akan dibuat juga menerapkan evaluasi klaster menggunakan DBI untuk mengukur kualitas pengelompokan, yang kemungkinan tidak menjadi fokus utama dalam penelitian Yassin.	Kesamaannya yaitu untuk mengelompokkan siswa berdasarkan kriteria tertentu guna membantu sekolah dalam menentukan penerimaan siswa. Selain itu, kedua penelitian mengimplementasikan sistem berbasis web untuk menampilkan hasil klasterisasi

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
7.	Suiroh, Rini Astuti, Fadhil Muhamad Basysyar (2024)	Implementasi Algoritma <i>K-Means</i> Pada Pengelompokan Data Penerimaan Peserta Didik Baru Di Smkn 1 Balongan	Dari hasil eksperimen dengan menggunakan <i>measure type numerical measure chebychev distance</i> dan <i>corelation similarity</i> menghasilkan kelompok terbaik k=2 dengan nilai DBI 0.508, dengan jumlah <i>cluster</i> 0 sebanyak 421 <i>items</i> sedangkan jumlah <i>Cluster</i> 1 sebanyak 416. Untuk <i>Cluster_0</i> pada atribut kejuruan di dominasi oleh jenis kejuruan TKRO, TITL, DPIB, TKM. Dan untuk atribut asal sekolah pada <i>Cluster_0</i> di dominasi oleh jenis asal sekolah diantaranya seperti Smp Negeri 1 Sliyeg, Smp Muhammadiyah Sukaurip, Mtsn 2 Indramayu, Smp Negeri 1 Balongan, Smp Negeri 3 Balongan, Smp Negeri 1 Karangampel, Smp Negeri 2 Balongan, Smp Negeri 1 Juntinyuat, Smp Negeri 1 Indramayu, Smp PGRI Juntinyuat, Smp Negeri 2 Juntinyuat, Smp Plus Nurul Islam, Smp Negeri 3 Sindang, Smp Negeri 10 Kota Bandung, Mts Muhammadiyah Baruamba, Smp Negeri 1 Ciawigebang, Smp Negeri 2 Cimalaka, Smp Negeri 1 Margasaridan cluster 1 berjumlah 300 item didominasi atribut.	Penelitian Suiroh menggunakan <i>K-Means</i> biasa untuk mengelompokkan data penerimaan siswa di SMKN 1 Balongan tanpa mempertimbangkan batasan kapasitas sekolah. Sementara itu, penelitian yang akan dibuat menerapkan <i>Constrained K-Means</i> , yang membatasi jumlah siswa dalam setiap klaster sesuai dengan kuota sekolah, sehingga distribusi lebih merata. Selain itu, saya menambahkan evaluasi klaster menggunakan <i>Davies-Bouldin Index</i> (DBI) untuk mengukur kualitas pengelompokan, serta menampilkan hasil pemetaan dalam bentuk web menggunakan <i>Streamlit</i> untuk visualisasi yang lebih interaktif.	Kesamaannya untuk mengoptimalkan proses seleksi siswa berdasarkan pengelompokan data, sehingga dapat membantu sekolah dalam menentukan zonasi atau kategori penerimaan siswa.

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
8.	Firgilia Claudiana Manek, Remerta N Naatonis, Yampi R Kaesmeten (2022)	Pengelompokan Sekolah Menurut Kecamatan Di Kabupaten Belu Menggunakan <i>K-Means</i> Berbasis <i>Webgis</i>	Penelitian firgilia Kabupaten Belu menghasilkan sistem pengelompokan sekolah berbasis <i>WebGIS</i> menggunakan algoritma <i>K-Means</i> , yang berhasil membagi sekolah-sekolah ke dalam tiga klaster berdasarkan kecamatan. Sistem ini mempermudah pengguna dalam melihat penyebaran sekolah pada peta digital, sehingga pencarian lokasi sekolah menjadi lebih efisien. Pengujian menggunakan <i>Mean Squared Error</i> (MSE) menunjukkan nilai error sebesar 0,035, yang mengindikasikan tingkat akurasi yang baik dalam pengelompokan data sekolah sesuai lokasi kecamatan.	Penelitian Firgilia berfokus pada pengelompokan sekolah berdasarkan kecamatan di Kabupaten Belu menggunakan <i>K-Means</i> serta menampilkan hasilnya dalam bentuk <i>WebGIS</i> untuk visualisasi spasial. Sementara itu, penelitian yang akan dibuat menerapkan <i>Constrained K-Means</i> dalam sistem zonasi PPDB, dengan tambahan batasan jumlah kuota sekolah untuk memastikan distribusi siswa yang lebih merata. Selain itu, saya menggunakan evaluasi klaster dengan <i>Davies-Bouldin Index</i> (DBI) dan menampilkan hasil pemetaan dalam bentuk web menggunakan <i>Streamlit</i> untuk memberikan tampilan interaktif dalam proses zonasi sekolah.	Kesamaannya untuk pengelompokan sekolah serta penerapan sistem berbasis <i>web</i> untuk menampilkan hasil klusterisasi. Kedua penelitian juga berfokus pada analisis spasial, di mana pengelompokan dilakukan berdasarkan lokasi geografis untuk mendukung proses pengambilan keputusan dalam sektor pendidikan.
9.	Farhan Aditya, Budiyanto Ahaliki, Muhammad Rifai Katili (2024)	Pengembangan Aplikasi Pemetaan Zonasi Sekolah Dasar Di Kota Gorontalo Menggunakan	Hasil penelitian dalam jurnal ini menunjukkan bahwa pengembangan aplikasi pemetaan zonasi sekolah dasar berbasis <i>WebGIS</i> di Kota Gorontalo	Perbedaan utama terletak pada pendekatan dan metode yang digunakan. Penelitian Farhan pada pengembangan	Kesamanya untuk mengelompokkan sekolah atau siswa berdasarkan lokasi geografis guna mendukung

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
		Metode <i>K-Means</i>	dengan metode <i>K-Means</i> mampu meningkatkan efisiensi dan keadilan dalam distribusi peserta didik. Aplikasi ini menampilkan data visual sebaran sekolah serta membantu calon siswa dan orang tua untuk mengakses informasi zonasi penerimaan siswa secara transparan. Melalui pendekatan <i>clustering K-Means</i> , sekolah dapat dikelompokkan berdasarkan jarak dan karakteristik wilayah, sehingga pemerataan jumlah siswa lebih optimal. Sistem ini juga memudahkan Dinas Pendidikan dalam memonitor dan mengatur penerimaan peserta didik baru di Kota Gorontalo.	aplikasi pemetaan zonasi sekolah dasar di Kota Gorontalo menggunakan <i>K-Means</i> , tanpa adanya batasan. jumlah siswa dalam setiap klaster. Sementara itu, penelitian yang akan dibuat menerapkan <i>Constrained K-Means</i> , yang mempertimbangkan batasan jumlah kuota sekolah untuk memastikan distribusi siswa lebih merata. Selain itu, saya menambahkan evaluasi klaster menggunakan <i>Davies-Bouldin Index (DBI)</i> untuk mengukur kualitas pengelompokan, serta menampilkan hasil pemetaan dalam bentuk web menggunakan <i>Streamlit</i> untuk visualisasi yang lebih interaktif	kebijakan penerimaan peserta didik baru (PPDB). Selain itu, kedua penelitian juga mengembangkan aplikasi berbasis web untuk menampilkan hasil zonasi secara interaktif.
10.	Civita C. I. Loho, Vivi P. Rantung, Gladly C. Rorimpandey (2022)	Data Mining Rekomendasi Sekolah Calon Siswa SMA di Kota Tomohon Menggunakan Metode <i>K-Means Clustering</i>		Penelitian civlita lebih berfokus pada rekomendasi sekolah bagi calon siswa SMA di Kota Tomohon, menggunakan <i>K-Means</i> untuk mengelompokkan sekolah berdasarkan karakteristik tertentu.	Kesamaannya untuk pengelompokan sekolah dan siswa. Keduanya bertujuan untuk membantu proses seleksi atau rekomendasi sekolah berdasarkan data <i>clustering</i> , sehingga mendukung

No.	Nama/Tahun	Judul	Hasil	Perbedaan	Persamaan
				Sementara itu, penelitian yang akan dibuat menerapkan <i>Constrained K-Means</i> dalam sistem zonasi PPDB, dengan tambahan batasan jumlah kuota sekolah agar distribusi siswa lebih merata. Selain itu, penelitian yang akan dibuat juga menggunakan evaluasi kluster dengan <i>Davies-Bouldin Index</i> (DBI) dan menampilkan hasil pemetaan dalam bentuk web menggunakan <i>Streamlit</i> untuk visualisasi yang lebih interaktif.	pengambilan keputusan dalam sistem penerimaan peserta didik.

Dari beberapa penelitian terkait penerapan algoritma *K-Means* dalam sistem zonasi sekolah dan klusterisasi data penerimaan siswa, dapat disimpulkan bahwa algoritma ini memiliki beragam implementasi dalam mendukung efisiensi dan pemerataan penerimaan siswa di berbagai jenjang pendidikan. Penelitian yang akan di buat memiliki kesamaan dengan beberapa penelitian sebelumnya dalam penggunaan metode *clustering* untuk sistem zonasi penerimaan siswa (PPDB), terutama dengan algoritma *Constrained K-Means*. Namun, perbedaannya terletak pada pendekatan optimasi dan evaluasi kluster, di mana saya menambahkan batasan jumlah kuota sekolah serta menerapkan evaluasi menggunakan *Davies-Bouldin Index* (DBI) untuk menilai kualitas pengelompokan. Selain itu, hasil klusterisasi akan ditampilkan dalam bentuk web menggunakan *Streamlit* untuk visualisasi yang lebih interaktif dan mudah diakses.

2.3 Kerangka pikir



Gambar 2. 1 Kerangka Berpikir

Gambar tersebut menggambarkan penelitian tentang penerapan algoritma *Constrained K-Means* untuk meningkatkan efisiensi sistem zonasi sekolah dasar dalam penerimaan peserta didik baru di Kabupaten Majene. Rumusan Masalah bertanya bagaimana hasil implementasi algoritma *Constrained K-Means* dapat membantu mengoptimalkan sistem zonasi penerimaan peserta didik baru. Tujuan Penelitian adalah menciptakan zonasi yang lebih efisien dan adil dengan mengelompokkan wilayah berdasarkan jarak geografis, sehingga penerimaan siswa lebih terorganisir. Metode Penyelesaian melibatkan penggunaan algoritma *Constrained K-Means* untuk mengelompokkan lokasi domisili siswa sesuai jarak dari sekolah. Hasil Yang Diharapkan adalah sistem zonasi yang lebih terstruktur, memungkinkan distribusi siswa yang merata dan sesuai dengan kedekatan geografis mereka ke sekolah dengan mempertimbangkan batasan tertentu.

BAB V

PENUTUP

6.1 Kesimpulan

Penelitian ini mengusulkan sistem penerimaan siswa baru tingkat Sekolah Dasar (SD) di Kabupaten Majene, Kecamatan Banggae, Sulawesi Barat dengan menerapkan pendekatan zonasi berbasis algoritma *Constrained K-Means*. Pendekatan ini dirancang untuk menyelesaikan permasalahan ketidakseimbangan data dalam proses zonasi sekolah, sehingga distribusi siswa baru menjadi lebih merata dan efisien. Sistem yang dikembangkan mampu memetakan jumlah siswa baru pada setiap sekolah dengan mempertimbangkan jarak tempat tinggal siswa terhadap lokasi sekolah berdasarkan wilayah zonasi yang telah ditentukan.

Hasil implementasi algoritma *Constrained K-Means* menunjukkan bahwa proses klusterisasi berhasil mengelompokkan calon siswa secara efektif dan seimbang. Dengan jumlah sekolah sebanyak 28 sekolah dan jumlah calon siswa baru sebanyak 1.209 orang, sistem ini membentuk klaster dengan jumlah anggota yang hampir merata. Sebanyak 28 sekolah menerima siswa dengan jumlah yang sesuai dengan kapasitas masing-masing dari sekolah tersebut. Dengan hasil ini, tidak terjadi tumpang tindih anggota klaster, dan setiap pusat klaster memiliki anggota yang sesuai, sehingga tidak ada sekolah yang kelebihan atau kekurangan siswa secara signifikan.

Hasil evaluasi menggunakan *Davies-Bouldin Index (DBI)* menghasilkan nilai 1,6683 untuk kedua skenario, yang mengindikasikan bahwa klaster-klaster yang terbentuk memiliki pemisahan yang baik dan jarak antar klaster cukup optimal. Pada Skenario 1, di mana kapasitas sekolah lebih besar dari jumlah siswa, sistem berhasil mengalokasikan siswa secara merata tanpa ada kelebihan siswa di sekolah-sekolah tertentu. Sementara itu, pada Skenario 2, di mana kapasitas sekolah lebih kecil dari jumlah siswa, algoritma tetap menjaga pemisahan antar klaster dengan baik meskipun ada batasan kapasitas yang lebih ketat. Hasil ini menunjukkan bahwa sistem zonasi berbasis *Constrained K-Means* dapat efektif dalam mengelompokkan siswa sesuai dengan kapasitas sekolah dan kedekatan geografis.

Dengan demikian, penelitian ini menunjukkan bahwa penggunaan algoritma *Constrained K-Means* dalam sistem zonasi sekolah dapat meningkatkan efektivitas dan efisiensi dalam penerimaan peserta didik baru. Penerapan sistem ini di Kabupaten Majene, Kecamatan Banggae, Sulawesi Barat dapat dijadikan model bagi wilayah lain yang menghadapi permasalahan serupa dalam penyaluran siswa ke sekolah sesuai dengan wilayah tempat tinggal mereka.

6.2 Saran

Penelitian ini dapat dikembangkan lebih lanjut dengan mempertimbangkan faktor-faktor tambahan seperti preferensi siswa, dan kualitas pendidikan di setiap sekolah untuk menghasilkan sistem zonasi yang lebih *komprehensif* dan *adaptif* terhadap kebutuhan masyarakat.

DAFTAR PUSTAKA

- Dewi, A. O. P. (2020). Big Data di Perpustakaan dengan Memanfaatkan Data Mining. *Anuva: Jurnal Kajian Budaya, Perpustakaan, Dan Informasi*, 4(2), 223–230. <https://doi.org/10.14710/anuva.4.2.223-230>
- DR. Hamuni Muhammad Idrus, Ms., & Dra Aswati, Mp. M. (2022). Perkembangan Peserta Didik Penerbit Cv.Eureka Media Aksara. *Cv. Eureka Media Aksara*, 122.
- Fajri, N. (2021). at-Tarbiyah al-Mustamirrah : Jurnal Pendidikan Islam Implementasi Penguatan Pendidikan Karakter di Satuan Pendidikan Sekolah Dasar. *Jurnal Pendidikan Islam*, 2, 1–10.
- Fathurohman, A. (2021). Machine Learning Untuk Pendidikan: Mengapa Dan Bagaimana. *Jurnal Informatika Dan Teknologi Komputer (JITEK)*, 1(3), 57–62. <https://journal.amikveteran.ac.id/index.php/jitek/article/view/306>
- Ika, I. P. (2020). Analisis Pengelompokan Data Argumen Tentang Penerapan Kebijakan Sistem Zonasi Pada Pendaftaran Sekolah Dengan Menggunakan Metode Clustering. *Journal of Informatics and Vocational Education*, 3(2), 47–58. <https://doi.org/10.20961/joive.v3i2.43438>
- Kurniawan, D., & Saputra, A. (2019). Penerapan K-Nearest Neighbour dalam Penerimaan Peserta Didik dengan Sistem Zonasi. *Jurnal Sistem Informasi Bisnis*, 9(2), 212. <https://doi.org/10.21456/vol9iss2pp212-219>
- Kusumaningtyas, K., Habibi, M., Dwijayanti, I., & Sumiyarini, R. (2023). Tweet Analysis of Mental Illness Using K-Means Clustering and Support Vector Machine. *Telematika*, 20(3), 295. <https://doi.org/10.31315/telematika.v20i3.9820>
- Loho, C., Rantung, V. P., & Rorimpandey, G. C. (2022). Data Mining Rekomendasi Sekolah Calon Siswa SMA di Kota Tomohon Menggunakan Metode K-Means Clustering. *JOINTER : Journal of Informatics Engineering*, 3(02), 1–9. <https://doi.org/10.53682/jointer.v3i02.34>
- Nur Aziz, Y. A., & Zuliarso, E. (2022). Sistem Penerimaan Siswa Baru Di Smkn 3

- Pati Berdasar Jalur Prestasi Menggunakan Algoritma Klastering K-Means Berbasis Web. *Jurnal Ilmiah Informatika*, 10(02), 86–95. <https://doi.org/10.33884/jif.v10i02.5555>
- Nuryoso, Saputra, D., Qawim, M. M., Hariyadi, A., & Utomo, S. (2023). Kepemimpinan Kepala Sekolah Dalam Menerapkan Konsep Merdeka Belajar. *Equity In Education Journal*, 5(1), 14–22. <https://doi.org/10.37304/ej.v5i1.8256>
- Pangestuti, A. (2021). *Sistem Zonasi dalam Perspektif Pendidikan*. 2(1), 15–21.
- Pendidikan, J. (2024). *Cendikia Cendikia*. 2(3), 454–474.
- Risal, A. A. N., Zainuddin, Z., & Niswar, M. (2022). School Zoning System for Student Admission using Constrained K-Means Algorithms. *Proceeding - IEEE International Conference on Communication, Networks and Satellite, COMNETSAT 2022, June*, 174–178. <https://doi.org/10.1109/COMNETSAT56033.2022.9994366>
- Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 5(1), 75–82. <https://doi.org/10.31294/ijcit.v5i1.7951>
- Shinta, M., & Ain, S. Q. (2021). Strategi Sekolah Dalam Membentuk Karakter Siswa di Sekolah Dasar. *Jurnal Basicedu*, 5(5), 4045–4052. <https://doi.org/10.31004/basicedu.v5i5.1507>
- Sholihin, M. R., & Rudiman. (2022). Pemetaan Sekolah Muhammadiyah di Kabupaten PPU Berdasarkan Fasilitas , Pendidik dan Tenaga Pendidik Menggunakan Metode K-Means Clustering. *Jurnal Keilmuan Dan Aplikasi Teknik Informatika*, 5(36), 45–51. <https://mail.jurnal.yudharta.ac.id/v2/index.php/EXPLORE-IT/article/view/3425>
- Wahyu Pribadi, W., Yunus, A., & Wiguna, A. S. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 493–500. <https://doi.org/10.36040/jati.v6i2.4808>

Widyaningtyas, A., Nugraha, N., & Diana, S. (2021). Antropocene : Jurnal Penelitian Ilmu Humaniora Persepsi Masyarakat Terhadap Sistem Penerimaan Peserta Didik Baru. *Antropocene : Jurnal Penelitian Ilmu Humaniora*, 1(1), 29–37.

Zainuddin, Z., & Risal, A. A. N. (2024). Balanced clustering for student admission school zoning by parameter tuning of constrained k-means. *IAES International Journal of Artificial Intelligence*, 13(2), 2299–2311. <https://doi.org/10.11591/ijai.v13.i2.pp2301-2313>