

Pengelompokan Remaja Berdasarkan Segmentasi Usia Menggunakan Metode K-Means *Clustering* (Studi Kasus : Desa Sindangsari)

Rini Rahmawati

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Agus Bahtiar

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Alamat: Cirebon, Jawa Barat, Indonesia 45131

Korespondensi penulis: rinirhmwti05@gmail.com

Abstract. Data mining is processing information from a database that can be used for various needs. One of the methods in data mining, namely Clustering which aims to find groupings from a series of patterns, points, objects and documents. The K-Means clustering algorithm is an algorithm that plays an important role in the field of data mining and is simple to implement and run. The K-Means Clustering method attempts to group existing data into several groups, where the data in one group have the same characteristics. By conducting clustering research youth based on age segmentation using the k-means clustering method is expected to be able to contribute especially to PIK R colleagues in dividing the segmentation of PIK R members easily and systematically without using manual methods. This age segmentation can be used to determine the level of development, needs, and preferences of adolescents in various aspects of life. This study aims to process the number of adolescents for members of the PIK-R organization, it is hoped that it will make it easier for secretaries in the Pik-R organization to manage new membership recruitment data based on age and knowing which hamlet has the most teenage population. In each cluster it is classified based on which criteria are prioritized. System testing was carried out 4 times with data consisting of 24 attributes 1789 records of new PIK-R members to get precision implementation results K-Means Clustering method.

Keywords: Data Mining, K-Means, Clustering, and Adolescent Phase.

Abstrak. Pemrosesan sebuah informasi dari suatu database yang dapat digunakan untuk berbagai kebutuhan disebut dengan Data Mining. Metode dalam data mining salah satunya yaitu *Clustering* yang memiliki tujuan untuk menemukan pengelompokan dari serangkaian pola, titik, objek maupun dokumen. Algoritma K-Means *clustering* merupakan salah satu algoritma yang sederhana untuk diimplementasikan atau dijalankan serta berperan penting dalam bidang data mining. Metode K-Means *Clustering* berusaha mengelompokkan data ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama. Dengan dilakukannya penelitian pengelompokan remaja berdasarkan segmentasi usia menggunakan metode k-means *clustering* ini diharapkan dapat memberikan kontribusi khususnya untuk rekan-rekan PIK R dalam membagi segmentasi anggota PIK R dengan mudah dan sistematis

Received Maret 30, 2023; Revised April 25, 2023; Accepted Mei 09, 2023

* Rini Rahmawati, rinirhmwti05@gmail.com

tanpa menggunakan metode manual. Segmentasi usia ini dapat digunakan untuk menentukan tingkat perkembangan, kebutuhan, dan preferensi remaja dalam berbagai aspek kehidupan. Penelitian ini bertujuan mengolah jumlah penduduk usia remaja untuk anggota organisasi PIK R, harapannya supaya bisa memudahkan sekertaris dalam organisasi Pik-R untuk mengelola data rekrutmen keanggotaan baru berdasarkan usia sertamengetahui di Dusun mana yang memiliki penduduk remaja paling banyak. Pada setiap cluster diklasifikasikan sesuai dengan kriteria mana yang lebih diprioritaskan. Pengujian pada sistem ini dilakukan sebanyak 4 kali percobaan terhadap data yang terdiri dari 24 atribut 1789 record anggota baru PIK-R untuk mendapatkan hasil implementasi metode K-Means Clustering.

Kata kunci: Data Mining, K-Means, Clustering, dan Fase Remaja.

LATAR BELAKANG

Menurut WHO (World Health Organization) Remaja adalah penduduk yang berusia 10-19 tahun, adapun pengertian Remaja menurut Badan Kependudukan dan Keluarga Berencana (BKKBN) adalah penduduk yang berusia 10-24 tahun dan belum menikah. Sehingga dapat disimpulkan bahwa pengertian Remaja adalah masa peralihan dari kanak-kanak menuju dewasa..

Pengolahan Jumlah penduduk dalam penelitian ini data diperoleh dari Desa Sindangsari tahun 2022. Melalui data ini peneliti akan mengelompokkan para remaja sesuai dengan segmentasi usia. Segmentasi ini dilakukan untuk memudahkan dalam menyampaikan informasi yang sesuai dengan kebutuhan dan perkembangan remaja pada setiap kategori usia tersebut.

Dengan judul mengelompokkan jumlah penduduk menurut kategori umur menggunakan pendekatan, penelitian Sahat Sonang, Arifin Tua Purba, dan Ferri Ojak Ymanuel Pardede dimuat dalam Jurnal Tekinkom, volume dua, nomor dua, Desember 2019, e-issn: 2621 -3079, dan doi: 0.37600/tekinkom.v2i2.115. Pengelolaan kependudukan merupakan salah satu tanggung jawab pemerintah Indonesia, menurut Kmeans. Kedudukan ekonomi suatu wilayah dapat dipengaruhi oleh pertumbuhan dan perpindahan penduduk karena menghasilkan zona eksklusif yang dapat menahan perpindahan penduduk. Jumlah penduduk Kota Pematangsiantar tahun 2018 menjadi fokus penelitian ini. Ini dipisahkan menjadi dua kategori usia, termasuk balita, anak kecil, remaja awal, remaja akhir, dewasa awal, tua, dan lanjut usia. Pengolahan data ini menyesuaikan menggunakan jumlah penduduk kota pematangsiantar perkecamatan yang pengelompokannya dilakukan menggunakan memakai metode k-means clustering.

menggunakan pengelompokan ini, kita mampu lihat bahwa jumlah penduduk pada setiap kecamatan sesuai grup umur masing-masing sebagai akibatnya dapat kita laksanakan program yang lebih sempurna dalam menaikkan sumber daya insan, yang termasuk ke pada grup kecamatan yang memiliki jumlah penduduknya paling tinggi buat setiap kategori usia yaitu kecamatan siantar martoba, siantar timur, serta siantar utara menggunakan pusat cluster (4.072, 4.203, 4.019, 8.759, 6.221, 5.882, 4.923, 3.255, 2.101). Sedangkan yang termasuk ke pada grup kecamatan yang mempunyai jumlah penduduknya rendah pada setiap kategori usia antara lain kecamatan siantar selatan, siantar barat, siantar marimbun, siantar marihat, dan siantar sitalajari yang mempunyai jumlah sentra cluster (2.101, 2.224, 2.191, 4.279, 3.234, 3.290, 3.012, 2.149, 1.360). (Sonang et al., 2019)

NO_KEC	NAMA_KENO_KEL	NAMA_KENO_KK	NK	NAMA	SHDK	JNSKLMINTMPT	LHFTGL_LHR	UMUR	GDR	STATUS	AGAMA	PENDIDIK PEKERJA	NAMA_BI	NAMA_AYDUSUN	ALAMAT	NO_RT	NO_RW	KODE_PO					
2	CIKONENI	2.003	SINDANG	3207020101	320732270	FAUZAAN	Anak	Laki-laki	CIAMIS	27-06-2000	16	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI MEN HERSAFARI A, JETAK	DUSUN JE	1	1	0
2	CIKONENI	2.003	SINDANG	3207020101	320731080	FAHAZ NUJI	Anak	Laki-laki	CIAMIS	08-07-2000	13	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI MEN HERSAFARI A, JETAK	DUSUN JE	1	1	0
2	CIKONENI	2.003	SINDANG	3207020111	320702571	KEISHYA	Anak	Perempuan	CIAMIS	17-12-2000	13	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI RIMAYAN TAUFIK H, JETAK	JETAK	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207020201	320702210	FADLYAN	Anak	Laki-laki	CIAMIS	21-07-2010	12	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI SRI SUPA GANDI SU, JETAK	DUSUN JE	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207020301	320702150	OKI MUHJ	Anak	Laki-laki	CIAMIS	15-01-2000	23	Tahun	-	Belum	Kat	Islam	Tamat	SD	BELUM/TTI MAEMUN HAERUM, JETAK	DUSUN JE	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207020501	320702640	ALMA FITI	Anak	Perempuan	CIAMIS	24-09-2000	13	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI HENI OKT JAJAT MU, JETAK	DUSUN JE	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207020601	320702690	CALLYS TI	Anak	Perempuan	JAKARTA	29-05-2010	11	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI NUR SETI GEMA SA, JETAK	DUSUN JE	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207020901	320702470	RAISYA P	Anak	Perempuan	CIAMIS	07-08-2010	12	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI TITA DEWI DINDIN AF, JETAK	DUSUN JE	1	1	46.261
2	CIKONENI	2.003	SINDANG	3207021101	320702581	NABILA M	Anak	Perempuan	TASKIMAL	18-12-2000	13	Tahun	-	Belum	Kat	Islam	Tidak	Belu	BELUM/TTI RATNA NIDANI MU, JETAK	DUSUN JE	1	1	46.261

Gambar 1 Data Remaja

Data di atas merupakan data remaja penduduk Desa Sindangsari yang akan diteliti melalui metode clustering dengan algoritma K-Means, yang didalamnya memiliki beberapa class.

Masalah yang ditemui yaitu Di Desa Sindangsari terutama pada organisasi PIK R Kecamatan Cikoneng Kabupaten Ciamis dalam hal pengelompokan remaja berdasarkan kategori usia serta untuk melihat di Dusun mana yang memiliki penduduk remaja paling banyak ada keterbatasan dalam mengelola data yaitu masih memakai cara manual sehingga memerlukan waktu relatif lama Berdasarkan pemaparan latar belakang diatas maka penelitian ini merusmukan beberapa problem sebagai berikut : Bagaimana cara pengelompokan remaja berdasarkan segmentasi usia menggunakan algoritma K-means dalam menggunakan aplikasi Orange serta Bagaimana cara mengetahui dusun mana yang memiliki jumlah penduduk remaja paling banyak.

Penerapan algoritma K-means pada pengelompokan jumlah penduduk Remaja berdasarkan kategori usia ini bertujuan untuk memudahkan dalam penyampaian informasi, memudahkan dalam penelitian dan analisis yang sesuai dengan karakteristik

remaja, memudahkan dalam mengambil keputusan dalam pengembangan sosial, pendidikan serta kesehatan. Oleh karena itu diusulkan penelitian dengan judul Pengelompokan Remaja Berdasarkan Segmentasi Usia. Adapun beberapa manfaat dari penelitian ini diantaranya adalah : Memiliki Manfaat Teoritis dari penelitian ini dapat menjadi pembelajaran dan pemahaman tentang metode K-Means sehingga dapat diaplikasikan ke Analisa sentiment menggunakan aplikasi Orange adapula manfaat praktisnya adalah bisa memberikan kontribusi terhadap organisasi PIK R yang akan mengrekrutmen anggota baru.

KAJIAN TEORITIS

A. Data Mining

Data mining adalah sekelompok metode untuk menggambarkan pengetahuan yang sebelumnya belum ditemukan dari basis data besar. Pengambilan keputusan dapat dibantu oleh pola yang ditemukan. Penambangan data tidak hanya dapat digunakan untuk menggali informasi atau fakta baru, tetapi juga dapat digunakan untuk memperdalam pemahaman kita tentang apa yang sudah diketahui. Data mining adalah proses yang menarik pengetahuan dari database atau gudang data yang cukup besar dengan mengekstraksi dan mengidentifikasi informasi yang dapat digunakan menggunakan metode statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. (Muttaqin & Defriani, 2020)

Data mining adalah sebuah proses dengan memakai Teknik serta indera analisis data buat suatu mencari pola serta korelasi yang tersembunyi yang berasal dari sebuah data. Meringkas serta mengekstrak berita mudah serta lumrah yang sebelumnya tak diketahui merupakan taktik penambangan data primer. (Surya et al., 2021)

B. K-Means

K-Means yaitu sebuah metode pengelompokan data non-hierarkis yang akan mencoba membagi data yang terdapat menjadi satu atau lebih cluster/grup. Metode ini mempartisi data sebagai cluster/kelompok sebagai akibatnya data menggunakan ciri yang sama dikelompokkan dalam cluster yang sama. Metode KMeans merupakan metode clustering non hirarkis yang bertujuan untuk mendukung pengelompokan

variabel untuk dimasukkan ke dalam kelaskelas yang ada pada hasil akhir perhitungan.(Sonang et al., 2019)

K-means ialah algoritma clustering klasik berbasis partisi, prosedur pemecahan k-means mempunyai banyak keuntungan praktis dipahami, sederhana, konvergensi cepat serta sebagainya, namun mempunyai kelemahan: sensitif buat menentukan sentra clustering awal serta praktis buat konvergen optimalisasi local.(Rustam et al., 2018)

C. Clustering

Clustering ialah galat satu metode penambangan data tanpa pengawasan yang paling umum. Objek dengan atribut karakteristik serupa ditempatkan dalam kategori, dan atribut karakteristik objek dalam kategori berbeda berbeda. Pengelompokan hasil lebih mendekati keadaan sebenarnya, sehingga analisis pengelompokan banyak digunakan dalam berbagai bidang. Algoritme K-means, juga dikenal sebagai sejenis metode pengelompokan cepat, diusulkan oleh Macqueen pada tahun 1967. K-means dapat mempertahankan skalabilitas dan efisiensi yang baik saat menangani batch. data, dengan kemampuan pencarian lokal yang kuat dan kecepatan konvergensi yang cepat. Prosedur Algoritma K-means: memilih k sampel dari n sampel sebagai pusat pengelompokan awal, dan pusat pengelompokan awal ditentukan secara acak.(Feng et al., 2022)

D. Fase Reemaja

Menurut World Health Organization (WHO), remaja adalah orang yang tinggal di suatu negara yang berusia 10 sampai 19 tahun; menurut Peraturan Menteri Kesehatan Republik Indonesia Nomor 25 Tahun 2014, remaja adalah orang yang tinggal di suatu negara yang berusia 10 sampai dengan 18 tahun; dan menurut Badan Kependudukan dan Keluarga Berencana (BKKBN), remaja adalah orang yang tinggal di suatu negara yang berusia 10 sampai 24 tahun dan belum menikah. Masa transisi antara masa kanak-kanak dan masa dewasa adalah masa remaja. Anda akan melalui fase perkembangan fisik dan mental yang cukup pesat selama ini. Oleh karena itu, remaja dapat dibagi ke dalam tahapan berikut:

a) Pra Remaja (11 - 14 tahun)

Untuk anak laki-laki berusia 12 hingga 14 tahun, rentang waktu ini cukup singkat—sekitar satu tahun. Karena kecenderungan perilaku negatif, periode ini terkadang disebut sebagai fase negatif. Tahap di mana komunikasi antara orang tua dan anak paling menantang. Perubahan pada tubuh, termasuk perubahan hormonal yang dapat mengakibatkan perubahan suasana hati yang tidak rasional, juga mengganggu perkembangan fungsi tubuh.

b) Remaja Awal (13 atau 14 tahun - 17 tahun)

Perubahan terjadi dengan cepat dan mencapai klimaksnya selama fase ini. Di usia ini, banyak terjadi ketidakstabilan dan ketidakstabilan emosi di banyak bidang. Mereka mencari jati diri karena tidak jelas statusnya saat ini. Pola interaksi sosial mulai bergeser. Mereka sering menampilkan perilaku seperti orang dewasa dan merasa diberdayakan untuk membuat penilaian sendiri. Selama tahap pertumbuhan ini, kemandirian dan kualitas yang sangat berbeda tercapai, pemikiran menjadi lebih rasional, tidak berbentuk, dan idealis, dan jumlah waktu yang dihabiskan jauh dari keluarga meningkat.

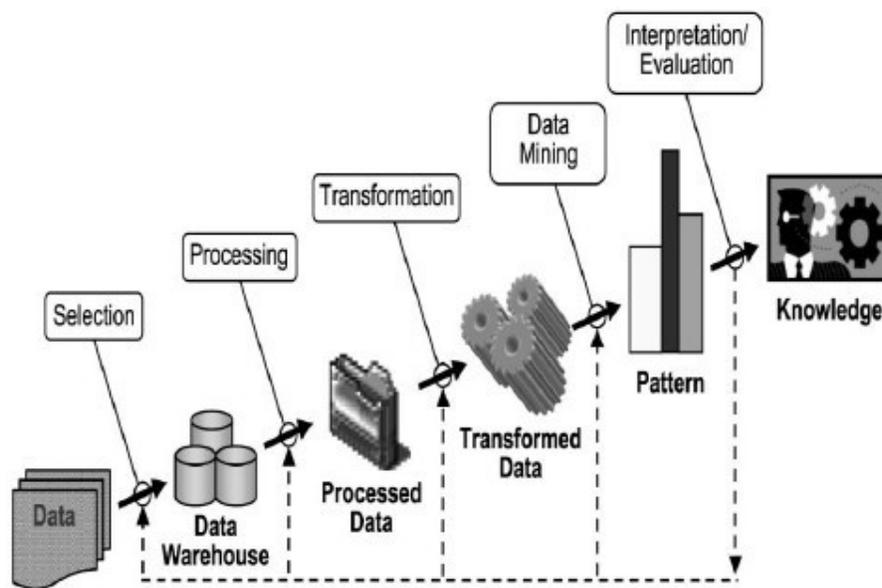
c) Remaja Lanjut (17 - 21 tahun)

Dia saat ini ingin menjadi pusat perhatian dan menekankan siapa dirinya, yang merupakan taktik remaja awal lainnya. Dia idealis, memiliki aspirasi yang tinggi, antusias, dan memancarkan banyak energi. Dia ingin menjadi mandiri secara emosional dan membangun rasa dirinya sendiri. Ada perubahan fisik yang terjadi pada masa remaja ini yang terjadi begitu cepat, seperti perubahan ciri-ciri seksual seperti pembesaran payudara, perkembangan pinggang untuk anak perempuan, dan tumbuhnya kumis, janggut, dan suara yang lebih dalam untuk anak laki-laki. Ada juga perubahan mental yang terjadi selama ini. Pada tahap ini, pencapaian identitas juga akan sangat terlihat; berpikir masih logis, tidak berbentuk, dan idealis; dan waktu yang dihabiskan jauh dari keluarga terus meningkat. Perubahan-perubahan tersebut di atas juga dikategorikan dalam fase pubertas, yang ditandai dengan pematangan fungsi seksual yang cepat, perubahan berat dan tinggi badan, serta pertumbuhan kerangka atau tubuh, terutama pada tahap awal masa remaja. Pubertas, bagaimanapun, tidak terjadi sekaligus. Tahap proses yang berkembang secara bertahap adalah pubertas. Selama tahap ini, adalah hal yang umum untuk mengamati anak-anak muda menghabiskan waktu berjama-

jam di depan cermin untuk mencoba tampil ideal dan menarik. pakaian dan asesoris yang selalu bergaya, tindakan lain yang terkadang kita anggap tidak pantas, dan lain sebagainya. (Diananda, 2018)

METODE PENELITIAN

Metode penelitian yang dipergunakan pada penelitian ini memakai Knowledge Discovery in Database (KDD) ialah proses yg bertujuan untuk menggali serta menganalisis data yang sangat besar sebagai berita yang bermanfaat untuk pengetahuan. Langkah krusial pada proses KDD terdiri dari tahapan-tahapan sebagai berikut:



Gambar 2 Tahapan KDD

a. Data Selection

Pengumpulan data terjadi pada langkah ini. Data yang diperoleh kemudian dipilih berdasarkan karakteristiknya, dan temuan seleksi ditambahkan ke dataset. Karena mempelajari data mining dan pengembangan model baru bergantung pada data yang dihasilkan, maka proses pembuatan data menjadi sangat penting.

b. Pre-Processing

Sebelum data mining dapat digunakan, prosedur untuk membersihkan data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kekurangan data, seperti kesalahan ketik, harus dilakukan. Memanfaatkan data atau informasi lain yang terkait dengan dan diperlukan untuk KDD, seperti data atau informasi eksternal, adalah metode lain dari augmentasi data yang ada.

c. Transformation

Pekerjaan pengembangan data dilakukan selama tahap ini untuk mempersiapkan data dengan lebih baik untuk pemodelan data mining. Pengurangan dimensi, seperti pemilihan fitur dan pengambilan sampel data, dapat dilakukan untuk menyiapkan data dengan lebih baik. Selain itu, transformasi fungsional dan transformasi atribut seperti mengubah nilai numerik menjadi atribut diskrit keduanya dimungkinkan.

d. Data Mining

Data ini dapat disimpan dalam database, gudang data, atau penyimpanan informasi sebagai rangkaian prosedur untuk mengungkap pola yang menarik dalam volume data yang sangat besar.

e. Interpretation/Evaluation

Penambangan data disajikan dalam format yang dapat dengan mudah dipahami oleh pihak yang berkepentingan. Ungkapan itu adalah bagian dari proses interpretasi KDD, yang melibatkan pemeriksaan apakah pola yang ditemukan atau kebaruan bertentangan dengan informasi atau hipotesis yang tersedia sebelumnya.

HASIL DAN PEMBAHASAN

Sumber data pada penelitian ini berasal dari Desa Sindangsari. Sumber Data ini meliputi data penduduk Desa Sindangsari, waktu pengambilan data ini dilaksanakan pada tanggal 29 Desember 2022. Data penelitian ini termasuk pada data sekunder karna data yang diperoleh dari laporan pemerintah desa.

Teknik pengumpulan data pada penelitian ini menggunakan proses pengamatan. Seperti observasi serta dokumentasi. Observasi dilakukan dengan cara pengamatan dari tabel yang berisi tentang penduduk Desa Sindangsari dari NO_Kec, Nama_Kec, Nama_Kel, NIK, Nama, SHDK, JNSKLMN, TMPT_LHR, TGL_LHR, Umur, GDR, Status, Agama, Pendidikan, Pekerjaan, Nama_Ibu, Nama_Ayah, Dusun, Ala

mat, No_RT, No_RW dan Kode Pos. Berikut ini merupakan teknik pengumpulan data dan informasi yang dilakukan ialah sebagai berikut:

a. Observasi

Observasi ialah menggambarkan sesuatu dengan memakai mata. Observasi atau yang sering disebut dengan pengamatan mencakup aktivitas pemusatan penelitian terhadap suatu objek menggunakan semua alat indra. (Dan et al., 2021)

b. Dokumentasi

Dokumentasi adalah pengamatan eksklusif untuk memperoleh data yang sesuai dengan pembahasan. Hal ini dilakukan untuk mengetahui ilustrasi konkret di lapangan. (Prawiyogi et al., 2021)

c. Wawancara

Wawancara adalah salah satu teknik yang dilakukan untuk mengumpulkan data atau informasi yang berkaitan dengan masalah penelitian di lapangan. Dengan wawancara peneliti dapat mengetahui data dan informasi yang dibutuhkan. Penulis mengajukan pertanyaan kepada narasumber untuk mengumpulkan data atau informasi yang diperlukan pada penelitian ini.

Hasil Penggunaan Algoritma K-Means Menggunakan Aplikasi Orange

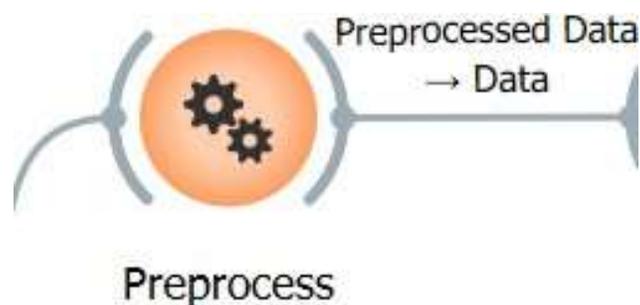
Hasil penelitian disajikan sebagai langkah implementasi data mining pada data remaja desa Sindangsari dengan menggunakan pendekatan algoritma K-Means yang nantinya akan didukung oleh aplikasi Orange.

1. Data Selection

Pada tahap selection penulis menyeleksi atau memilih data yang diperlukan data tersebut kemudian disesuaikan dengan kebutuhan. Data dalam penelitian ini terdiri dari 24 atribut dengan jumlah data 1789 record.

2. Preprocessing

Langkah pengolahan data penelitian ini menggunakan widget pengolahan data dengan cara menormalkan fitur dengan tool Orange kemudian memilih Normalize to [-1,1].



Gambar 3 Widget Proses preprocessing data

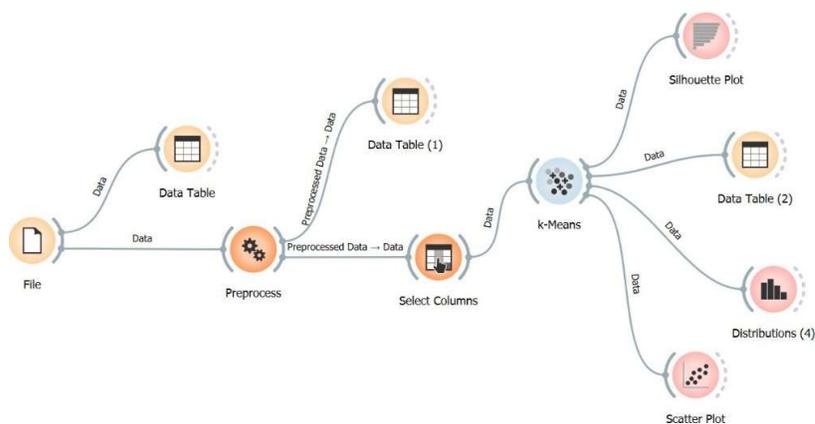
Gambar 4 Saat menggunakan alat oranye untuk penambangan data, widget pemrosesan data digunakan sebagai preprosesor. Widget ini melakukan transformasi yang memberikan setiap atribut nilai -1 dan standar deviasi 1.

3. Transformasi

Transformasi data merupakan proses untuk merubah type atribut data akan diubah sebagai bentuk yang sinkron untuk proses data mining. sebab pada penelitian ini akan dilakukan uji coba secara teoritis serta menggunakan perangkat lunak tool orange, setelah itu data yang diproses sebelumnya akan diubah agar kompatibel dengan teknik K-Means yang digunakan. Data berikut ini ditransformasikan, seperti terlihat pada tabel data image dari orange, dan terdiri dari 24 atribut dan 1789 record.

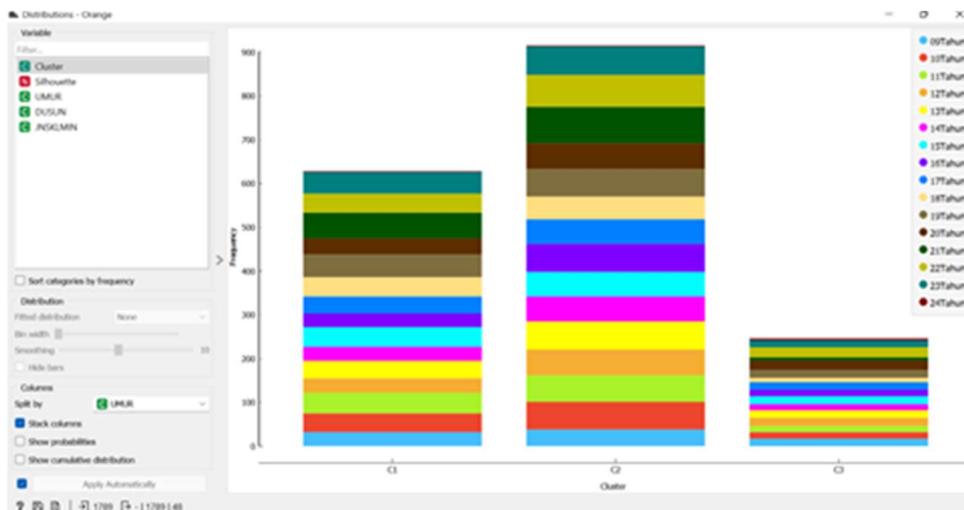
4. Data Mining

Metode penambangan data sinkron dipilih saat ini. Untuk menggunakan dan mengolah data yang menggunakan orange menggunakan pendekatan algoritma K-Means, data tersebut kemudian dibakukan sesuai dengan langkah-langkah proses data mining.



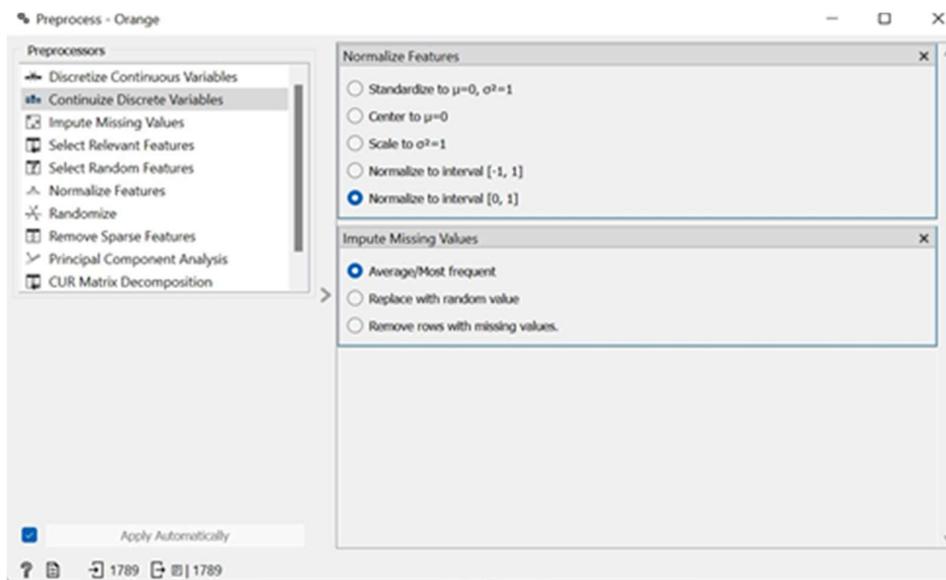
Gambar 4 Model Proses Algoritma K-Means Menggunakan Orange Model

Algoritma K-Means digunakan pada Gambar 4 untuk mengelompokkan data remaja dusun Sindangsari dengan menggunakan teknik orange process model.



Gambar 5 Hasil Distributions-Orange

Sebaran cluster 1, 2, dan 3 berdasarkan usia digambarkan pada Gambar 5. Gambar 4 menampilkan hasil penggunaan alat berwarna orange pada widget distro untuk melihat sebaran usia remaja di Desa Sindangsari.



Gambar 6 Hasil preprocesing-Orange

Gambar 6 menampilkan hasil penggunaan widget orange pretreatment untuk menyiapkan data anak muda desa Sindangsari sebelum dilakukan clustering menggunakan teknik K-Means.

Pengelompokan Remaja Berdasarkan Segmentasi Usia Menggunakan Metode K-Means Clustering (Studi Kasus : Desa Sindangarsi)

	NAMA	TGL_LHR	NAMA_IBU	NAMA_BAPAK	NO_KEC	NAMA_KEC	NO_KEL	NAMA_KEL	NO_KK	NK
1	FALDIAN NABIL	27-06-2006	MEN HERMANA	SAFARI ABDUL	2	OKENENG	0	SINDANGSARI	0.00198999	0.81176
2	RAAZ NUR AZKA	08-07-2009	MEN HERMANA	SAFARI ABDUL	2	OKENENG	0	SINDANGSARI	0.00198999	0.81176
3	KESHA DEKA	17-12-2009	RIHANNATI	SAIFUL HADYAT	2	OKENENG	0	SINDANGSARI	0.00264778	0.81176
4	RAZWANDAH	25-07-2010	DE SURPHATI	GANDI SUGANDI	2	OKENENG	0	SINDANGSARI	0.0732395	0.81166
5	DREI MURJANI	15-01-2009	MAMIRANI	HERMANI	2	OKENENG	0	SINDANGSARI	0.2670981	0.81166
6	ADIA FIBRINI	24-09-2008	HEM OKTAVIYI	SAIT MARJAIT	2	OKENENG	0	SINDANGSARI	0.138871	0.81176
7	CALYSTA NALA	29-05-2011	NUR SETIABATI	EDMA SANTOSA	2	OKENENG	0	SINDANGSARI	0.264768	0.81167
8	SALDA PUTRI A.	07-08-2010	ISA DEWI PUSPL	SINDIN APRISA	2	OKENENG	0	SINDANGSARI	0.268422	0.8116
9	NABILA NAWA	18-12-2009	RATNA NINGSI	DANI MAHADI	2	OKENENG	0	SINDANGSARI	0.33527	0.8116
10	NAURAL KEYSA	26-11-2005	ADA ADI NAGET	TOTO SUGARDO	2	OKENENG	0	SINDANGSARI	0.431989	0.81166
11	DWI WINDATI	01-07-2006	AJUN ARIANSI	AONG	2	OKENENG	0	SINDANGSARI	0.533226	0.81166
12	SISA MELANI	09-05-2007	RIAS KRUSMADI	WIA KLOWANA	2	OKENENG	0	SINDANGSARI	0.533226	0.8116
13	MISA APRILLIANI	09-04-2003	RIAS KRUSMADI	WIA KLOWANA	2	OKENENG	0	SINDANGSARI	0.533226	0.8116
14	BUNGA SEPTIA	25-09-2006	WAW ROSLIWANI	DADI DARINAK	2	OKENENG	0	SINDANGSARI	0.566748	0.8116
15	MAULANA FEB.	19-02-2012	WAW ROSLIWANI	DADI DARINAK	2	OKENENG	0	SINDANGSARI	0.566748	0.81166
16	WINDA HASTUTI	03-06-2006	ETI ROHWATI	ASEP MAHAR	2	OKENENG	0	SINDANGSARI	0.566748	0.8116
17	DEVI PUTRI PER.	16-11-2001	SRI SUGIARTI	ONDO SUWAN	2	OKENENG	0	SINDANGSARI	0.566748	0.8116
18	ADYA RAHMA	10-10-2002	LELA SELAMATI	NADI KURNADI	2	OKENENG	0	SINDANGSARI	0.566748	0.8116
19	DHEA DWI PU.	24-09-2000	OKIRWATI	YADI SETIABANI	2	OKENENG	0	SINDANGSARI	0.699309	0.8116
20	DITA NADIA PER.	04-02-2007	WIDANINGSIH	ASEP DARGO	2	OKENENG	0	SINDANGSARI	0.9677625	0.8116
21	DEVI FEBRIANTI	06-02-2003	WIDANINGSIH	ASEP DARGO	2	OKENENG	0	SINDANGSARI	0.9677625	0.8116
22	ZENIRA AZENIA	15-10-2008	TRIANI	ASEP KUMARU	2	OKENENG	0	SINDANGSARI	0.9996721	0.8116
23	FEBRIAN KURNI	24-02-2004	LARINA	CANDIA	2	OKENENG	0	SINDANGSARI	0.107432	0.81166
24	WISELY CHRISTI	13-06-2007	LARINA	CANDIA	2	OKENENG	0	SINDANGSARI	0.107432	0.81166
25	SAZURDI RA.	10-10-2007	MINDU ROSITA	DANI APRIANA	2	OKENENG	0	SINDANGSARI	0.135542	0.81166
26	RENITA ARIANI	11-11-2004	WINDU KURNI	ANDRI NABILA	2	OKENENG	0	SINDANGSARI	0.16796	0.8116
27	SANDIA PUTRI	23-10-2012	UNIK HELINA	KUSUMANDI EP.	2	OKENENG	0	SINDANGSARI	0.199899	0.8116
28	SARITA PUTRA	16-12-2010	RIYANING	PERMANA	2	OKENENG	0	SINDANGSARI	0.308647	0.81166

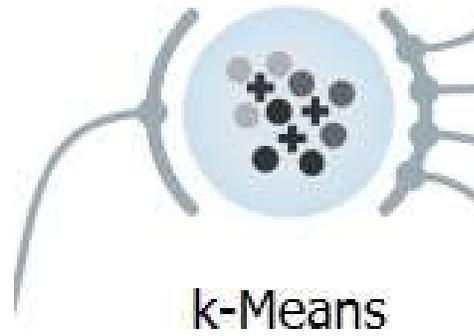
Gambar 7 Hasil dari widget Data Tabel-Orange

Keluaran dari widget Tebel-Orange Data yang digunakan untuk pre-processing data menggunakan widget pre-processing data young ditunjukkan pada Gambar 7.

Ignored (23)	Selected (3)
GEOR	UMUR
NO_KEC	DUSUN
NO_RT	JNSKLIMN
NO_RW	
NAMA_KEC	
NO_KEL	
NAMA_KEL	
NK	
SHCK	
TMP_LHR	
STATUS	
AGAMA	
PENDUKUN	
PEKERJAAN	
ALAMAT	
KODE_POS	
NO_KK	
NAMA_IBU	
NAMA_BAPAK	

Gambar 8 Hasil widget select coloum

Gambar 8 menunjukkan tahapan Umur, No_RT, No_RW, Dusun, JnsKlImn, Nama dan Tanggal_Lhr yang digunakan untuk pemilihan atribut dari data remaja, diklaster dengan algoritma K-Means menggunakan tools orange.



Gambar 9 Hasil Widget Algoritma K-Means

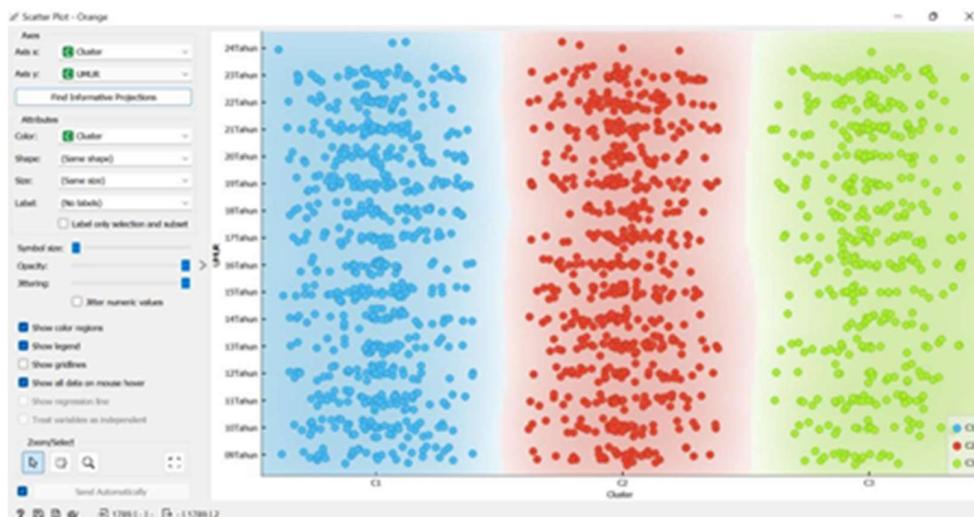
Gambar 9 didasarkan pada pengelompokan data berdasarkan usia yang menggunakan alat oranye dan teknik K-Means.

	NAME	TGL_LHR	Cluster	Sibwekto	UMR	NO_KT	NO_RN	DUSUN	PEKERJAAN
1	BAGDAN NAR...	27-06-2006	C3	0.621465	16Tahun	0.00	0.00	ETAK	Laki-laki
2	RAIZ NUR AZNA	08-07-2009	C3	0.619537	17Tahun	0.00	0.00	ETAK	Laki-laki
3	KESHA SOKA...	13-12-2009	C3	0.614892	17Tahun	0.00	0.00	ETAK	Perempuan
4	SALWINDHIA ...	21-07-2010	C3	0.616271	17Tahun	0.00	0.00	ETAK	Laki-laki
5	DIO MUHAMMAD	15-01-2000	C3	0.621006	23Tahun	0.00	0.00	ETAK	Laki-laki
6	ALMA FITRIAN	24-09-2009	C3	0.614892	17Tahun	0.00	0.00	ETAK	Perempuan
7	CAHYA NALA...	29-05-2011	C3	0.568457	11Tahun	0.00	0.00	-	Perempuan
8	SALYA PUTRI A...	07-06-2010	C3	0.614892	17Tahun	0.00	0.00	ETAK	Perempuan
9	NABILA MAHA...	18-12-2009	C3	0.614892	17Tahun	0.00	0.00	ETAK	Perempuan
10	NAKAL KEYSA...	26-11-2005	C3	0.621007	17Tahun	0.00	0.00	ETAK	Laki-laki
11	DIN WINANTI	01-07-2006	C3	0.614702	16Tahun	0.00	0.00	ETAK	Perempuan
12	ELSA MELANI	09-05-2007	C3	0.614702	15Tahun	0.00	0.00	ETAK	Perempuan
13	MISA APRIJANI	09-04-2003	C3	0.615489	19Tahun	0.00	0.00	ETAK	Perempuan
14	BUNGA SEPSTA...	25-09-2006	C3	0.616781	16Tahun	0.00	0.00	ETAK	Perempuan
15	MAULANA FER...	19-02-2012	C3	0.618935	10Tahun	0.00	0.00	ETAK	Laki-laki
16	WENIA HASTUTI	03-06-2006	C3	0.616781	16Tahun	0.00	0.00	ETAK	Perempuan
17	DEVI PUTRI PER...	16-11-2001	C3	0.618002	21Tahun	0.00	0.00	ETAK	Perempuan
18	ADYA RAHMA...	10-10-2002	C3	0.615231	20Tahun	0.00	0.00	ETAK	Perempuan
19	DHEA DMI PU...	24-09-2000	C3	0.615664	22Tahun	0.00	0.00	ETAK	Perempuan
20	DITA NADIA PU...	04-02-2007	C3	0.613958	17Tahun	0.0233	0.00	ETAK	Perempuan
21	DEVI FEBRIANTI	06-02-2003	C3	0.614706	19Tahun	0.0233	0.00	ETAK	Perempuan
22	ZHAFRA AZNIA	15-10-2008	C3	0.538977	14Tahun	0.0233	0.00	COLENDORA	Perempuan
23	FERRIS KURNI...	24-02-2004	C3	0.620769	18Tahun	0.0233	0.00	ETAK	Laki-laki
24	WISSEY CHRISTI	13-06-2007	C3	0.616781	17Tahun	0.0233	0.00	ETAK	Laki-laki
25	LAZARDY RA...	10-10-2007	C3	0.568740	17Tahun	0.0233	0.00	-	Laki-laki
26	NONYA ARIAN...	15-11-2004	C3	0.615376	18Tahun	0.0233	0.00	ETAK	Perempuan
27	SAZKA PUTRI...	23-10-2012	C3	0.613866	10Tahun	0.0233	0.00	ETAK	Perempuan
28	SALYA PUTRI...	16-12-2010	C3	0.618052	12Tahun	0.0233	0.00	ETAK	Laki-laki
29	AMANGA PUTRI...	23-11-2011	C3	0.615622	11Tahun	0.0233	0.00	ETAK	Perempuan

Gambar 10 Hasil dari clustering

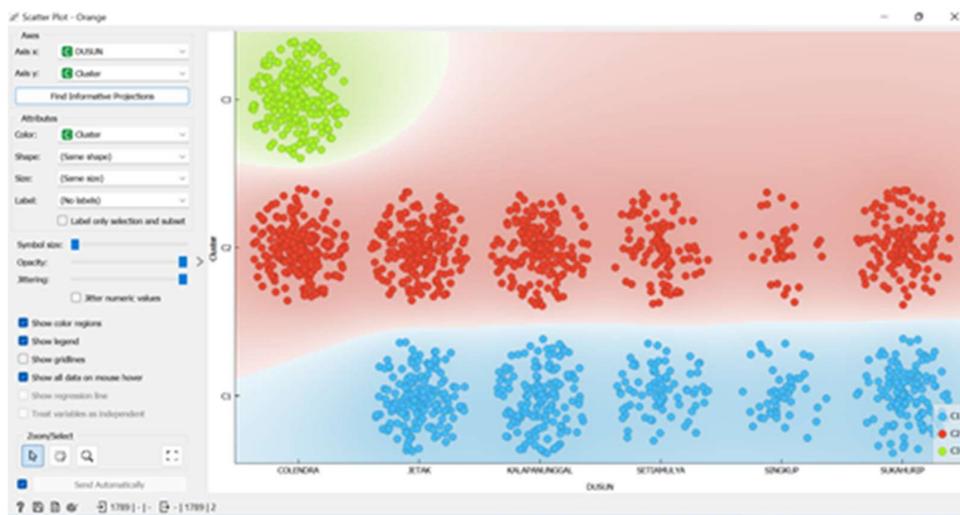
Gambar 10 menampilkan hasil penggunaan teknik K-Means.

*Pengelompokan Remaja Berdasarkan Segmentasi Usia Menggunakan Metode K-Means Clustering
(Studi Kasus : Desa Sindangsari)*



Gambar 11 Hasil visualisasi Scatter Plot

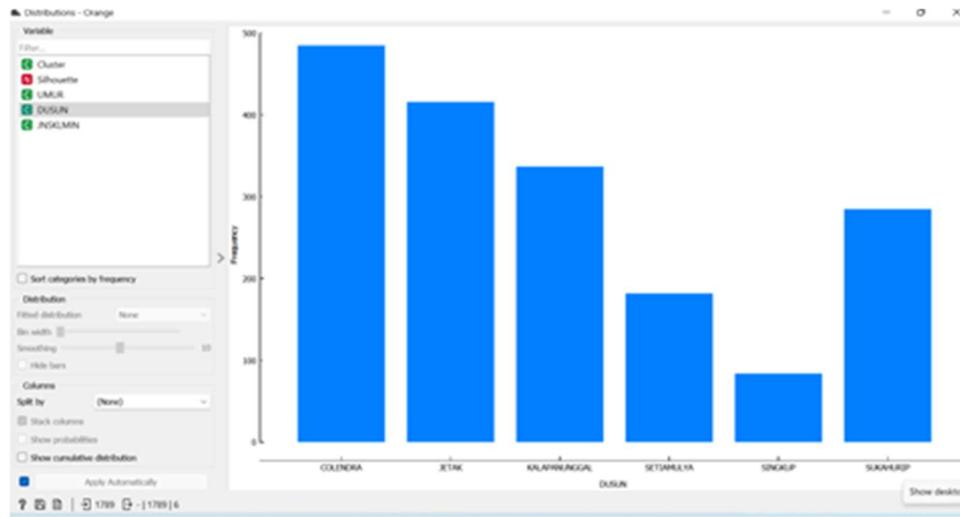
Hasil pemanfaatan widget Scatter Plot untuk mengilustrasikan hasil clustering menggunakan algoritma KMeans berdasarkan Data Usia Pemuda Desa Sindangsari pada alat berwarna orange ditunjukkan pada Gambar 11.



Gambar 12 Hasil Visualisasi Scatter Plot Dusun

Gambar 12 memperlihatkan cluster 1 sampai 3 dan dusun, pada cluster 3 hanya terdapat dusun colendra saja sedangkan pada cluster 3 sebaliknya dusun colendra tidak termasuk kedalam cluster 3.

Hasil Mengetahui Dusun yang paling banyak penduduk Remaja



Gambar 13 Distributions-Orange

Gambar 13 merupakan hasil dari tujuan kedua yaitu untuk mengetahui Dusun manayang memiliki jumlah remaja paling banyak.Dapat dilihat dari gambar tersebut bahwa Dusun Colendra yang memiliki Remaja paling banyak sejumlah 485 orang.

KESIMPULAN DAN SARAN

Hasil dari penelitian tugas akhir yang berjudul Pengelompokan Remaja Berdasarkan Segmentasi Usia dapat disimpulkan bahwa pengelompokan dengan menggunakan aplikasi orange lebih mudah dan dapat mengefesienkan waktu dibandingkan dengan pengelompokan secara manual.Dari penggunaan aplikasi tersebut dapat terbagi menjadi 3 cluster dan dapat dilihat pula bahwa dusun yang memiliki jumlah remaja yang banyak yaitu terdapat di Dusun Colendra.

Mengingat masih banyak hal yang belum bisa diimplementasikan dari penelitian ini, maka masih banyak lagi beberapa saran adalah:

1. Untuk penelitian selanjutnya diharapkan dapat melakukan langkahlangkah uji coba seperti: pengujian angka klaster, pengujian validitas klaster. Sehingga cluster yang diperoleh lebih optimal
2. Aplikasi harus diperbarui secara berkala, agar dapat memberikan rekomendasi tepat
3. Menggabungkan dengan metode atau pendekatan lain untuk mendapatkan hasil penelitian lebih baik

DAFTAR REFERENSI

- Dan, T., Yudianto, A., & Artikel, I. (2021). PENGARUH MOTIVASI BELAJAR TERHADAP HASIL BELAJAR SISWA KELAS VII SMP NEGERI 2 KEDOKAN BUNDER KABUPATEN INDRAMAYU. *Jurnal Pendidikan Indonesia*, 2(1).
- Diananda, A. (2018). PSIKOLOGI REMAJA DAN PERMASALAHANNYA. In *ISTIGHNA* (Vol. 1, Issue 1). www.depkes.go.id
- Feng, G., Fan, M., & Chen, Y. (2022). Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining. *IEEE Access*, 10, 19558–19571. <https://doi.org/10.1109/ACCESS.2022.3151652>
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*, 12(2), 121–129. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- Prawiyogi, A. G., Sadiyah, T. L., Purwanugraha, A., & Elisa, P. N. (2021). Penggunaan Media Big Book untuk Menumbuhkan Minat Membaca di Sekolah Dasar. *Jurnal Basicedu*, 5(1), 446–452. <https://doi.org/10.31004/basicedu.v5i1.787>
- Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). OPTIMASI K-MEANS CLUSTERING UNTUK IDENTIFIKASI DAERAH ENDEMIK PENYAKIT MENULAR DENGAN ALGORITMA PARTICLE SWARM OPTIMIZATION DI KOTA SEMARANG. *ILKOM Jurnal Ilmiah*, 10(3), 251–259. <https://doi.org/10.33096/ilkom.v10i3.342.251-259>
- Sonang, S., Purba, A. T., & Pardede, F. O. I. (2019). PENGELOMPOKAN JUMLAH PENDUDUK BERDASARKAN KATEGORI USIA DENGAN METODE K-MEANS. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 2(2), 166. <https://doi.org/10.37600/tekinkom.v2i2.115>
- Surya, A. D., M.Sapriyaldi, Wanto, A., Windarto, A. P., & Damanik, I. S. (2021). Komparasi Algoritma Machine Learning untuk Penentuan Performance Terbaik Pada Prediksi Produksi Tanaman Jahe di Indonesia. *Seminar Nasional Ilmu Sosial Dan Teknologi (SANISTEK)*, 276–284.
- Dan, T., Yudianto, A., & Artikel, I. (2021). PENGARUH MOTIVASI BELAJAR TERHADAP HASIL BELAJAR SISWA KELAS VII SMP NEGERI 2 KEDOKAN BUNDER KABUPATEN INDRAMAYU. *Jurnal Pendidikan Indonesia*, 2(1).
- Diananda, A. (2018). PSIKOLOGI REMAJA DAN PERMASALAHANNYA. In *ISTIGHNA* (Vol. 1, Issue 1). www.depkes.go.id
- Feng, G., Fan, M., & Chen, Y. (2022). Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining. *IEEE Access*, 10, 19558–19571. <https://doi.org/10.1109/ACCESS.2022.3151652>
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*, 12(2), 121–129. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>

- Prawiyogi, A. G., Sadiyah, T. L., Purwanugraha, A., & Elisa, P. N. (2021). Penggunaan Media Big Book untuk Menumbuhkan Minat Membaca di Sekolah Dasar. *Jurnal Basicedu*, 5(1), 446–452. <https://doi.org/10.31004/basicedu.v5i1.787>
- Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). OPTIMASI K-MEANS CLUSTERING UNTUK IDENTIFIKASI DAERAH ENDEMIK PENYAKIT MENULAR DENGAN ALGORITMA PARTICLE SWARM OPTIMIZATION DI KOTA SEMARANG. *ILKOM Jurnal Ilmiah*, 10(3), 251–259. <https://doi.org/10.33096/ilkom.v10i3.342.251-259>
- Sonang, S., Purba, A. T., & Pardede, F. O. I. (2019). PENGELOMPOKAN JUMLAH PENDUDUK BERDASARKAN KATEGORI USIA DENGAN METODE K-MEANS. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 2(2), 166. <https://doi.org/10.37600/tekinkom.v2i2.115>
- Surya, A. D., M.Sapriyaldi, Wanto, A., Windarto, A. P., & Damanik, I. S. (2021). Komparasi Algoritma Machine Learning untuk Penentuan Performance Terbaik Pada Prediksi Produksi Tanaman Jahe di Indonesia. *Seminar Nasional Ilmu Sosial Dan Teknologi (SANISTEK)*, 276–284.