

Implementasi Data Mining Dalam Menentukan Strategi Promosi Program KB Menggunakan Algoritma K-Means Clustering

Fuadz Hasyim, M.Kom.¹, Muafi, M.Kom.²
^{1,2} Universitas Nurul Jadid Probolinggo, Indonesia

Info Artikel

Riwayat Artikel

Diterima: 19-04-2022

Disetujui: 28-06-2022

Kata Kunci

Strategi Promosi;
Program KB;
Data Mining;
Clustering;
K-Means;

muafiumar76@unuja.ac.id

ABSTRAK

Strategi promosi program KB yang dilakukan oleh pihak Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) melalui Balai Penyuluhan Program Keluarga Berencana (BPPKB) masih menyamaratakan seluruh daerah tanpa melihat keberhasilan program KB. Dibutuhkan analisis data agar strategi yang dilakukan bisa tepat sasaran. Data yang akan dianalisis merupakan data rekapitulasi kependudukan yang ada pada BPPKB. Untuk menganalisis data tersebut dibutuhkan metode atau teknik analisis data yang tepat yaitu dengan analisis data mining menggunakan algoritma *k-means clustering*. Dengan cara melihat pola keterkaitan antar data yang digunakan sebagai penentu keberhasilan program KB, pihak BPPKB dapat dengan mudah dalam menentukan strategi promosi program KB hanya melihat pola data dari masing-masing kluster. Data penelitian diambil dari BPPKB Kecamatan Gading untuk data peserta KB aktif tahun 2019-2021, yang mencakup data dari 19 Desa di wilayah Kecamatan Gading Kabupaten Probolinggo. Dari hasil penelitian yang telah dilakukan didapatkan kesimpulan bahwa strategi promosi terbaik bagi calon peserta KB baru agar tepat sasaran adalah dengan mengirim tim admisi BPPKB yang sesuai dengan program KB yang paling banyak diminati dan melakukan promosi berdasarkan wilayah peserta KB baru dengan melakukan penyesuaian menggunakan *promotion mix*.

1. PENDAHULUAN

Program keluarga berencana adalah upaya untuk mewujudkan keluarga berkualitas melalui promosi, perlindungan, dan bantuan dalam mewujudkan hak-hak reproduksi serta penyelenggaraan pelayanan, pengaturan dan dukungan yang diperlukan untuk membentuk keluarga dengan usia kawin yang ideal. Program keluarga berencana juga mengatur jumlah jarak dan usia ideal melahirkan anak yaitu pengaturan kehamilan dan melahirkan anak, pengaturan kehamilan serta membina ketahanan dan kesejahteraan keluarga [1].

Instansi yang mengurus hal ini yaitu Balai Penyuluhan Program Keluarga Berencana (BPPKB) yang ditugaskan oleh pemerintah untuk menyebarkan program ini bagi masyarakat yang ada di wilayah kerja kantor BPPKB, dalam menyebarkan dan mengajak masyarakat untuk mengikuti program KB ini pihak BPPKB melakukan promosi baik itu dengan poster KB yang dipasang di jalan-jalan, sosialisasi langsung ke masyarakat, dan lain-lain. Permasalahan yang sering timbul adalah tidak semua bentuk promosi itu tepat sasaran, pihak BPPKB menyamaratakan promosi yang dilakukan, sehingga masyarakat yang mengikuti program KB ini tidak merata di seluruh daerah, ada yang satu desa sudah mengikuti KB tetapi ada desa lain yang masyarakatnya kurang mengikuti KB. Dalam mengatasi masalah itu perlu adanya analisis data untuk mengidentifikasi program KB sudah berjalan dengan baik atau belum sehingga promosi yang dilakukan bisa tepat sasaran.

Berdasarkan pada latar belakang masalah yang telah dikemukakan, penelitian ini memiliki fokus masalah yang perlu diselesaikan berupa belum adanya strategi promosi yang tepat dalam melakukan promosi program KB. Data penelitian yang akan digunakan adalah data peserta program KB aktif di wilayah kerja BPPKB Kecamatan Gading Kabupaten Probolinggo. Sedangkan tujuan dalam penelitian ini adalah melakukan *clustering* data peserta program KB aktif dengan teknik *data mining* menggunakan algoritma *K-Means*, sehingga dapat dijadikan sebagai referensi dalam menentukan rekomendasi strategi terbaik dalam melakukan promosi program KB yang lebih efektif dan efisien.

Untuk dapat menentukan strategi promosi yang lebih efektif dan efisien, maka dalam penelitian ini dilakukan dengan cara mengolah data-data yang telah didapatkan dari pihak instansi terkait. Data yang telah didapatkan kemudian diolah untuk mengetahui pola yang tepat sehingga informasi yang tersembunyi dapat ditemukan. Metode pengolahan data seperti ini sering disebut sebagai *data mining*. Pada penelitian ini analisa *data mining* dilakukan dengan menggunakan algoritma *K-Means*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan ke dalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama. Dengan adanya pengelompokan-pengelompokan data seperti ini, diharapkan bagi pihak instansi dapat menentukan strategi promosi yang tepat untuk mendapatkan calon anggota KB aktif baru.

K-means merupakan salah satu metode *clustering* non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*. Metode ini mempartisi data ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam *cluster* yang lain [2]. *K-Means* adalah suatu metode penganalisaan data atau metode *Data Mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi [3]. Terdapat dua jenis data *clustering* yang sering dipergunakan dalam proses pengelompokan data yaitu *Hierarchical* dan *Non-Hierarchical*, dan *K-Means* merupakan salah satu metode data *clustering non-hierarchical* atau *Partitional Clustering* [4]. Metode *K-Means Clustering* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Terdapat beberapa penelitian terkait dan menjadi rujukan dalam penelitian ini, yakni: Penelitian pertama dilakukan oleh Randi Rian Putra dan Cendra Wadisman pada tahun 2018, dengan judul “Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means”. Penelitian tersebut dilakukan di salah satu perusahaan bernama MC Laundry. Permasalahan yang terjadi adalah perusahaan membutuhkan metode pengelolaan pelanggan yang tepat, sehingga perusahaan dapat mempertahankan loyalitas dari pelanggan yang dimiliki. Untuk itu perusahaan harus bisa mengenali pelanggan yang potensial dan mempercayainya, hal tersebut bertujuan agar perusahaan dapat memahami kebutuhan pelanggan dan bisa memberikan pelayanan yang prima. Salah satu pelayanan yang akan diberikan perusahaan adalah pemberian hadiah berupa cinderamata yang dilakukan secara *continue* setiap tahun. *Tools* yang digunakan dalam mengimplementasikan algoritma *K-Means Clustering* di dalam penelitian tersebut menggunakan *software* Tanagra. Variabel yang digunakan dalam penelitian tersebut meliputi data nama pelanggan, data jumlah transaksi, dan data total transaksi. Proses pengolahan yang dilakukan adalah dengan cara menyimpan data pelanggan ke dalam *notepad*, kemudian data tersebut dimasukkan ke dalam *software* Tanagra [5].

Penelitian kedua dilakukan oleh Gustientiedina, dkk pada tahun 2019, dengan judul “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru”. Dalam penelitian tersebut dijelaskan bahwa perencanaan kebutuhan obat adalah aspek penting

dalam pengelolaan obat-obatan di rumah sakit. Hal tersebut akan berpengaruh ke dalam proses pengadaan, pendistribusian, dan penggunaan obat-obatan. Dengan perencanaan yang tepat, maka proses pengadaan obat-obatan dapat dilakukan dengan efektif dan efisien. Hal ini dapat menjamin terhadap ketersediaan obat sesuai dengan kebutuhan dan dapat diperoleh dengan lebih cepat. Tujuan dari penelitian tersebut adalah untuk mengetahui klusterisasi data obat dengan menggunakan algoritma *K-Means*. Output dari penerapan algoritma ini akan menghasilkan kelompok data obat dengan kategori pemakaian rendah, sedang, dan tinggi. Dari hasil penelitian tersebut didapatkan kesimpulan bahwa data obat yang masuk ke dalam kelompok rendah adalah data obat yang memiliki rata-rata permintaan setiap tahunnya kurang dari 18.000 buah. Sedangkan permintaan obat yang memiliki permintaan antara 18.000-70.000 buah setiap tahunnya masuk ke dalam kelompok obat dengan pemakaian sedang, dan permintaan obat yang melebihi dari 70.000 buah setiap tahunnya masuk ke dalam kelompok pemakaian tinggi [6].

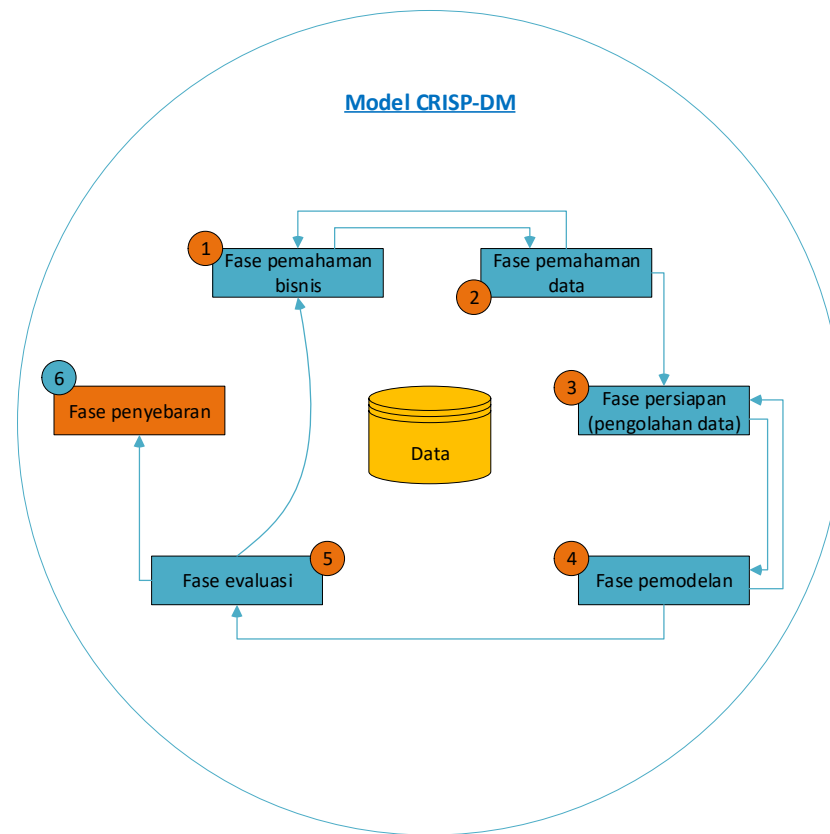
Penelitian ketiga dilakukan oleh Zulfa Nabila, dkk pada tahun 2021, dengan judul “Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means”. Dalam penelitian tersebut dijelaskan bahwa Dinas Kesehatan Provinsi Lampung dan Badan Perencanaan Pembangunan Daerah Provinsi Lampung telah melakukan kajian terkait kasus Covid19 di wilayah Provinsi Lampung dengan membagi 5 kriteria zonasi persebaran kasus Covid19 berdasarkan warna. Dalam proses pengkajian zonasi tersebut belum menggunakan proses perhitungan secara umum dan detail, hanya didasari dari hasil pengumpulan data dan analisis dari tim pakar Gugus Tugas Nasional penentuan zona dengan menggunakan 15 indikator utama. Indikator tersebut terbagi ke dalam 11 indikator epidemiologi, 2 indikator surveilans kesehatan masyarakat, dan 2 indikator pelayanan kesehatan. Tujuan dari penelitian tersebut adalah untuk menganalisa data kasus Covid19 untuk mengetahui pengelompokan kasus Covid19 di Provinsi Lampung dengan menggunakan algoritma *K-Means*. Atribut yang digunakan dalam perhitungan metodenya berupa atribut Kabupaten/Kota, Suspek, Probable, Konfirmasi Positif, Selesai Isolasi, dan Kematian. Sedangkan *cluster* yang digunakan dibagi menjadi 4 yang dikategorikan menjadi Zona Merah, Zona Orange, Zona Kuning, dan Zona Hijau. Validasi yang digunakan adalah Davies-Bouldin Index (DBI), yang mana hasil validasi menggunakan perhitungan manual dan dibandingkan dengan perhitungan menggunakan *tools rapid miner*. Hasil dari penelitian tersebut disimpulkan bahwa meskipun hasil perhitungan manual dan *rapid miner* terdapat perbedaan, namun keduanya sama-sama mendekati 0 yang berarti kluster yang telah ditentukan sudah baik [7].

Penelitian keempat dilakukan oleh Ari Sulistiyawati dan Eko Supriyanto pada tahun 2021, dengan judul “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan”. Dalam penelitian tersebut dijelaskan bahwa penentuan untuk masuk ke dalam kelas unggulan sangat penting dilakukan untuk menghasilkan informasi yang dapat dijadikan acuan sebagai bahan pertimbangan bagi pihak sekolah untuk menentukan langkah terstruktur dalam meningkatkan kualitas dan standar mutu jenjang sekolah menengah pertama kelas unggulan. Langkah-langkah dalam menentukan siswa yang bisa masuk ke dalam kelas unggulan didapatkan berdasarkan 2 aspek penilaian yang dikalkulasi dan telah melewati cakupan tingkat akurasi nilai tertinggi dari standar batasan kriteria ketuntasan minimum (KKM). Hasil identifikasi terhadap nilai rapor kurikulum 2013, menunjukkan atribut penilaian dengan hasil yang beragam. Secara garis besar memiliki persamaan nilai dan memiliki variasi kriteria dengan jumlah yang cenderung setara antar siswa. Selain itu, penempatan kapasitas jumlah siswa kelas unggulan setiap tahun ajaran dapat berubah-ubah seiring dengan pengembangan sekolah dan jumlah penerimaan siswa. Pemecahan permasalahan pengelompokan siswa dengan data yang semakin banyak menjadi kurang efisien. Tujuan dari penelitian tersebut adalah mengembangkan sistem informasi berbasis web untuk mengelola data penilaian yang terpusat dalam menyajikan laporan hasil belajar siswa dan pengelompokan siswa pada kelas unggulan dengan mengimplementasikan algoritma *K-means Clustering*. Pengelompokan data dalam satuan pendidikan dilakukan untuk memudahkan dalam memberikan fasilitas pendidikan siswa

berdasarkan perbedaan kemampuannya dalam belajar dan mengikuti pembelajaran. Hasil Penelitian tersebut berupa aplikasi pengelolaan data penilaian yang terpusat dalam menyajikan laporan hasil belajar siswa menggunakan hasil nilai rapor kurikulum 2013. Penilaian tersebut menjadi dasar atribut yang digunakan dalam proses perhitungan untuk menentukan siswa kelas unggulan, dengan membentuk 2 *cluster* untuk setiap kelasnya [8].

2. METODE

Model yang digunakan untuk menyelesaikan penelitian ini adalah CRISP-DM (Cross Industry Standard Process for Data Mining). Model CRISP-DM diperkenalkan pertengahan tahun 1990 oleh sebuah perusahaan konsorsium Eropa [9]. Dalam CRISP-DM, sebuah proyek Data Mining memiliki siklus hidup yang terbagi dalam 6 fase [10], seperti terlihat pada gambar 1.



Gambar 1. Siklus Hidup CRISP-DM

Adapun penjelasan dari tahapan/langkah di atas yakni sebagai berikut:

- A. Fase Pemahaman Bisnis (Business Understanding Phase)
Meliputi penentuan tujuan bisnis, menilai situasi saat ini, menetapkan tujuan data mining, dan mengembangkan rencana proyek. Tujuan bisnis yang akan dilakukan pada penelitian ini adalah menghasilkan klusterisasi wilayah untuk menentukan strategi promosi program KB yang tepat dengan menggunakan algoritma *K-Means*.
- B. Fase Pemahaman Data (Data Understanding Phase)
Setelah tujuan bisnis dan rencana proyek ditetapkan, langkah selanjutnya melakukan pengumpulan data awal, deskripsi data, eksplorasi data, dan verifikasi kualitas data. Penelitian yang diusulkan ini menggunakan data primer, dengan data yang digunakan adalah data peserta KB aktif periode tahun 2019-2021 di wilayah Kecamatan Gading Kabupaten Probolinggo.

- C. Fase Persiapan Data (Data Preparation Phase)
Pada tahap ini dilakukan identifikasi dan pembangunan jawaban dari data yang telah dikumpulkan untuk bisa melakukan pengelompokan dan pemilahan ke dalam kelompok-kelompok yang telah ditentukan. Atribut yang digunakan dalam penelitian ini berupa atribut Pasangan Usia Subur (PUS), atribut Keluarga Berencana (KB), atribut Keluarga Sejahtera (KS), dan atribut Unmet Need (UN).
- D. Fase Pemodelan (Modeling Phase)
Pada fase ini dilakukan pemilihan model yang akan digunakan untuk melakukan pengelompokan dalam menentukan strategi promosi program KB. Model atau metode yang akan digunakan pada penelitian ini adalah metode *K-Means*. Jumlah data latih yang akan digunakan terdiri dari 19 *record* yang diambil dari 19 desa di wilayah Kecamatan Gading Kabupaten Probolinggo.
- E. Fase Evaluasi (Evaluation Phase)
Pengujian akan dilakukan dengan membandingkan pengelompokan yang dilakukan oleh algoritma *K-Means* dengan pengelompokan yang dilakukan oleh BPPKB Kecamatan Gading Kabupaten Probolinggo.
- F. Fase Penyebaran (Deployment Phase)
Fase ini dilakukan untuk menghasilkan penemuan pengetahuan (identifikasi hubungan yang tak terduga dan berguna) untuk kemudian diterapkan pada operasi bisnis di berbagai tujuan, termasuk *clustering*.

3. HASIL DAN PEMBAHASAN

- A. Pemahaman Bisnis (Business Understanding)
Penerapan data mining pada penelitian ini berhubungan langsung dengan program KB untuk menggali pengetahuan tentang suatu pola dari peserta program KB aktif yang berpotensi terhadap pemilihan strategi yang tepat dalam melakukan promosi program KB.
- B. Pemahaman Data (Data Understanding)
Penelitian ini menggunakan data yang berasal dari BPPKB wilayah Kecamatan Gading Kabupaten Probolinggo. Data tersebut merupakan data peserta KB aktif periode tahun 2019-2021. Parameter yang digunakan seperti Pasangan Usia Subur (PUS), atribut Keluarga Berencana (KB), atribut Keluarga Sejahtera (KS), dan atribut Unmet Need (UN).
- C. Persiapan Data (Data Preparation)
Dataset Peserta KB aktif yang didapatkan dari BPPKB Kecamatan Gading Kabupaten Probolinggo berupa dokumen excel sejumlah 19 record. Atribut yang digunakan adalah Pasangan Usia Subur (PUS), Keluarga Berencana (KB), Keluarga Sejahtera (KS), dan Unmet Need (UN). Lebih detailnya, perhatikan tabel 1 tentang *dataset* peserta KB aktif periode tahun 2019-2021.

Tabel 1. *Dataset* peserta KB aktif periode tahun 2019-2021

NO	Daerah	PUS	KB	KS	UN
1	BATUR	851	639	75.09	16.28
2	BETEK TAMAN	550	415	75.45	40.48
3	SENTUL	492	375	76.22	29.07
4	DANDANG	517	398	76.98	30.15
5	PRASI	843	647	76.75	32.15
6	KERTOSONO	804	722	89.8	20.91
7	DUREN	494	353	71.46	19.26
8	RENTENG	222	196	88.29	24.49
9	KEBEN	229	195	85.15	30.77
10	NOGOSAREN	752	638	84.84	19.91
11	WANGKAL	1559	1272	81.59	26.1
12	GADING WETAN	402	344	85.57	42.73
13	BULU PANDAK	199	173	86.93	24.86

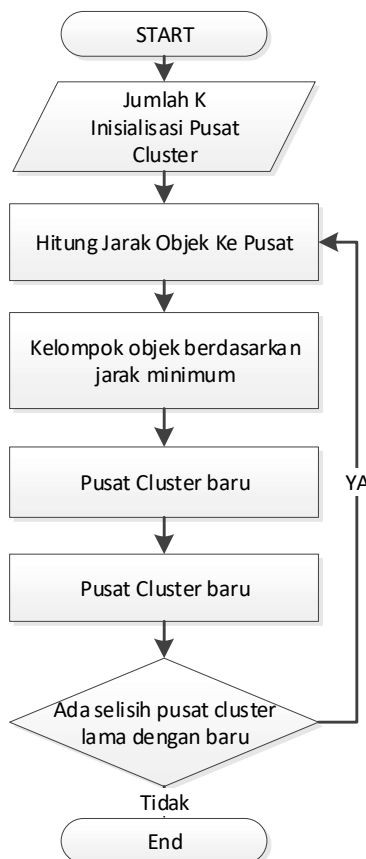
14	MOJOLEGI	574	469	81.71	27.72
15	KALIACAR	971	728	74.97	24.04
16	RANUWURUNG	635	439	69.13	33.94
17	JURANG JERO	726	598	82.37	23.08
18	CONDONG	904	657	72.68	31.2
19	SUMBER SECANG	384	251	65.36	9.16

Data yang berjenis nominal seperti desa asal harus dilakukan proses inialisasi data terlebih dahulu ke dalam bentuk angka/numerikal. Untuk melakukan inialisasi desa asal dapat dilakukan dengan:

- 1) Pada desa asal peserta KB aktif terlebih dahulu dilakukan pembagian wilayah wilayah menjadi beberapa bagian wilayah.
- 2) Kemudian wilayah-wilayah tersebut dilakukan pengurutan angka dari yang terbesar berdasarkan frekuensi peserta KB aktif yang berasal dari wilayah tersebut.
- 3) Wilayah yang memiliki frekuensi terbesar diberi inisial dengan angka 1 dan wilayah yang memiliki frekuensi terbesar kedua diberi inisial dengan angka 2, begitu seterusnya hingga wilayah dengan frekuensi paling sedikit.

D. Pemodelan (Modeling)

Pemodelan adalah fase yang secara langsung melibatkan teknik *data mining* yaitu dengan melakukan pemilihan teknik *data mining* dan menentukan algoritma yang akan digunakan. Setelah semua data desa pada tahun 2019-2021 ditransformasi ke dalam bentuk angka, maka data-data tersebut telah dapat dikelompokkan dengan menggunakan algoritma *K-Means*. Untuk dapat melakukan pengelompokan data-data tersebut menjadi beberapa *cluster* perlu dilakukan beberapa langkah, seperti terlihat pada gambar 2.



Gambar 2. Tahapan Algoritma *K-Means*

- a. Menetapkan jumlah *cluster* (*K*) pada penelitian ini yaitu berjumlah 2 *cluster*.

Tabel 2. Titik Pusat Awal Tiap *Cluster*

	Daerah	PUS	KB	KS	UN
C1	SENTUL	492	375	76.22	29.07
C2	GADING WETAN	402	344	85.57	42.73

- b. Menghitung jarak setiap data ke pusat *cluster* antara objek ke *centroid* dengan perhitungan jarak *Euclidean*.

Sebagai contoh, akan dihitung jarak dari data peserta KB aktif pertama ke pusat *cluster* pertama dengan persamaan:

$$C1. 445.805 = \text{SQRT}(((851 - 492)^2 + (639 - 375)^2 + (75.09 - 76.22)^2 + (16.28 - 29.07)^2))$$

Dari hasil perhitungan di atas di dapatkan hasil bahwa jarak data peserta KB aktif pertama dengan *cluster* pertama adalah 445.805.

Jarak dari data peserta KB aktif pertama ke pusat *cluster* kedua dengan persamaan:

$$C2. 537.992038 = \text{SQRT}(((851 - 402)^2 + (639 - 344)^2 + (75.09 - 85.57)^2 + (16.28 - 42.73)^2))$$

Dari hasil perhitungan di atas didapatkan hasil bahwa jarak data peserta KB aktif pertama dengan *cluster* kedua adalah 537.992038.

Berdasarkan hasil kedua perhitungan di atas dapat disimpulkan bahwa jarak data peserta KB aktif pertama yang paling dekat adalah *cluster* 1, sehingga peserta KB aktif pertama dimasukkan ke dalam *cluster* 1. Hasil perhitungan selengkapnya untuk 19 data peserta KB aktif pertama dapat dilihat pada tabel 3.

- c. Mengelompokkan data ke dalam *cluster* dengan jarak minimal.

Tabel 3. Contoh hasil perhitungan setiap data ke setiap *cluster*

No	C1	C2	MIN	
1	445.8047342	537.992038	445.8047342	1
2	92.3892905	440.7226757	92.3892905	1
3	375	386.0039094	375	1
4	398.7865895	414.5612675	398.7865895	1
5	736.0840762	782.2958065	736.0840762	1
6	785.7709603	825.7959829	785.7709603	1
7	352.1745217	364.8533034	352.1745217	1
8	333.3041573	266.0170972	266.0170972	2
9	326.9367445	260.2080283	260.2080283	2
10	688.1331339	727.1796788	688.1331339	1
11	1659.508258	1718.831695	1659.508258	1
12	354.9972086	342	342	2
13	339.4442932	266.0285445	266.0285445	2
14	474.1781971	498.84687	474.1781971	1
15	870.6301531	923.4455458	870.6301531	1
16	460.8329253	495.5860548	460.8329253	1
17	640.3481105	678.6666063	640.3481105	1
18	773.8126831	826.2197607	773.8126831	1
19	273.2733937	253.6915233	253.6915233	2

- d. Setelah semua data ditempatkan ke dalam *cluster* yang terdekat, kemudian hitung kembali pusat *cluster* yang baru berdasarkan rata-rata anggota yang ada pada *cluster* tersebut.

Tabel 4. hasil hitung rata-rata *cluster* 1

Cluster 1				
1	851	639	75.09	16.28
2	550	415	75.45	40.48
3	492	375	76.22	29.07
4	517	398	76.98	30.15
5	843	647	76.75	32.15
6	804	722	89.8	20.91
7	494	353	71.46	19.26
10	752	638	84.84	19.91
11	1559	1272	81.59	26.1
14	574	469	81.71	27.72
15	971	728	74.97	24.04
16	635	439	69.13	33.94
17	726	598	82.37	23.08
18	904	657	72.68	31.2
Means	762.2857143	596.4285714	77.78857	26.735

Tabel 5. hasil hitung rata-rata *cluster* 2

Cluster 2				
8	222	196	88.29	24.49
9	229	195	85.15	30.77
12	402	344	85.57	42.73
13	199	173	86.93	24.86
19	384	251	65.36	9.16
Means	287.2	231.8	82.26	26.402

Tabel 6. hasil menentukan *centroid* baru

Cluster	centroid baru 1			
C1	762.2857143	596.4285714	77.78857	26.735
C2	287.2	231.8	82.26	26.402

- e. Setelah didapatkan titik pusat yang baru dari setiap *cluster*, lakukan kembali dari langkah ketiga hingga titik pusat dari setiap *cluster* hingga tidak berubah lagi dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.
- f. Pada perhitungan ini iterasi berhenti pada iterasi ke-5 karena kelompok data 1 sama dengan kelompok data 2 dari hasil *clustering*, dan telah mencapai stabil dan konvergen.
- E. Evaluasi (Evaluation)

Dari data hasil implementasi menggunakan algoritma *K-Means* yang telah dilakukan di atas, maka dapat ditentukan beberapa strategi promosi yang dapat dilakukan oleh pihak BPPKB Kecamatan Gading Kabupaten Probolinggo, diantaranya:

- 1) Promosi dengan mengirim tim *marketing* yang sesuai dengan program KB yang paling banyak diminati.
- 2) Promosi pada persebaran wilayah berdasarkan tingkat wilayah peserta KB aktif.

Selanjutnya akan dilakukan analisis strategi promosi berdasarkan *promotion mix* pada masing-masing *cluster* yang terbentuk. Diharapkan dengan strategi yang diberikan pada masing-masing *cluster* dapat sesuai dengan tujuan bisnis yang diharapkan.

Tabel 7. Strategi Promosi berdasarkan *promotion mix*

No.	Strategi Promosi	Cluster Satu	Cluster Dua
1.	Periklanan	√	√
2.	Penjualan Personal		√
3.	Promosi Penjualan	√	√
4.	Hubungan Masyarakat	√	√
5.	Pemasaran Langsung		√

F. Penyebaran (Deployment)

Setelah tahap evaluasi dimana menilai secara detail hasil dari sebuah model maka dilakukan pengimplementasian dari keseluruhan model yang telah dibangun. Selain itu juga dilakukan penyesuaian terhadap model sehingga dapat menghasilkan suatu hasil yang sesuai dengan target awal tahap CRISP-DM ini.

4. KESIMPULAN DAN SARAN

Kesimpulan yang dapat diambil berdasarkan hasil penelitian ini adalah:

- A. Algoritma *K-Means* dapat digunakan untuk mengelompokkan data peserta KB sebagai pendukung keputusan penentuan strategi promosi yang tepat.
- B. Cara penentuan atau inisialisasi *centroid* sangat berpengaruh pada hasil *cluster*.
- C. Strategi promosi bagi calon peserta KB baru yang tepat sasaran untuk setiap wilayah berdasarkan *cluster* yang terbentuk adalah dengan mengirim tim admisi BPPKB yang sesuai dengan program KB yang paling banyak diminati dan melakukan promosi berdasarkan wilayah peserta baru dengan melakukan penyelarasan menggunakan *promotion mix* dan dengan melihat rata-rata pada setiap *cluster*.

5. DAFTAR PUSTAKA

- [1] BKKBN, Keluarga Berencana dan Kontrasepsi, Jakarta: Pustaka Sinar Harapan, 2015.
- [2] INFORMATIKALOGI, "Algoritma K-Means Clustering," 12 November 2016. [Online]. Available: <https://informatikalogi.com/algoritma-k-means-clustering/>. [Accessed 11 April 2022].
- [3] K. P. SINAGA and M.-S. YANG, "Unsupervised K-Means Clustering Algorithm," *IEEE ACCESS*, vol. VIII, no. 1, pp. 80716-80727, 2020.
- [4] V. K. Ayyadevara, Pro Machine Learning Algorithms, India: Apress, 2018.
- [5] R. R. Putra and C. Wadisman, "IMPLEMENTASI DATA MINING PEMILIHAN PELANGGAN POTENSIAL MENGGUNAKAN ALGORITMA K-MEANS," *IntecomS: Journal of Information Technology and Computer Science*, vol. I, no. 1, pp. 72-77, 2018.
- [6] Gustientiedina, M. Adiya and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *JURNAL NASIONAL TEKNOLOGI DAN SISTEM INFORMASI*, vol. V, no. 1, pp. 17-24, 2019.
- [7] Z. Nabila, A. R. Isnain, Permata and Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. II, no. 2, pp. 100-108, 2021.
- [8] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *Jurnal TEKNO KOMPAK*, vol. XV, no. 2, pp. 25-36, 2021.

- [9] W. M. P. Duhita, "CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA," *Jurnal Informatika*, vol. XV, no. 2, pp. 160-174, 2015.
- [10] M. A. Hasanah, S. Soim and A. S. Handayani, "Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir," *Journal of Applied Informatics and Computing (JAIC)*, vol. V, no. 2, pp. 103-108, 2021.