

Metode Clustering Pada Model Algoritma K-Means Untuk Pemilihan Alat Kontrasepsi

Diah Ayu Maulida Wati¹, Diah Puspitasari², Esty Purwaningsih^{2,*}

¹ Sistem Informasi; STMIK Nusa Mandiri; Jl. Damai No.8 Warung Jati Barat, Margasatwa-Jakarta Selatan, telp. (021) 78839502; e-mail: diahayum13@gmail.com

² Sistem Informasi; Universitas Bina Sarana Informatika; Jl. Kamal Raya No.18, Ringroad Barat, Cengkareng, Jakarta Barat, telp.(021) 54378399; e-mail: diah.puspitasari@bsi.ac.id, esty.epw@bsi.ac.id

* Korespondensi: e-mail: esty.epw@bsi.ac.id

Diterima: 29 April 2019 ; Review: 15 Mei 2019 ; Disetujui: 19 Juni 2019

Cara sitasi: Wati DAM, Puspitasari D, Purwaningsih E. 2019. Metode Clustering Pada Model Algoritma K-Means Untuk Pemilihan Alat Kontrasepsi. Informatics For Educators And Professionals. 3 (2): 129 – 138.

Abstrak: Terdapat beberapa faktor yang menjadi pertimbangan bagi pasangan suami istri untuk memutuskan keikutsertaannya dalam suatu program yaitu program keluarga berencana yang telah dicanangkan oleh pemerintah sejak lama. Program keluarga berencana tersebut diperlukan pemilihan alat kontrasepsi yang efektif dan sesuai dengan kondisi tubuhnya. Sehingga penelitian ini menerapkan metode clustering dengan model algoritma K-Means, dimana terdapat keunggulan dalam menyelesaikan masalah clustering pada pemilihan alat kontrasepsi. Data yang digunakan diperoleh dari Rumah Sakit Annisa Citeureup pada tahun 2017. Data diolah dengan menggunakan Rapid Miner dengan menghasilkan perolehan angka 48% untuk *cluster* 1 yang terdiri dari , usia 36 tahun, jumlah anak 2, pendidikan SMA, status kerja tidak kerja. 22% untuk *cluster* 2 yang terdiri dari , usia 46 tahun, jumlah anak 3, pendidikan SMA, status kerja tidak kerja. 30% untuk *cluster* 3 yang terdiri dari , usia 27 tahun, jumlah anak 2, pendidikan SMA, status kerja tidak kerja. Dari hasil pengolahan tersebut, penerapan clustering pada pemilihan alat kontrasepsi dapat bermanfaat bagi Rumah Sakit Annisa Citeureup.

Kata kunci: Alat Kontrasepsi, Clustering, K-Means, Program KB

Abstract: There are several factors that are taken into consideration for married couples to decide their participation in a program, namely a family planning program that has been proclaimed by the government for a long time. The family planning program requires the selection of effective contraceptives and in accordance with their body condition. So that this study applies the clustering method with the K-Means algorithm model, where there are advantages in solving clustering problems in the selection of contraceptives. The data used was obtained from Annisa Citeureup Hospital in 2017. Data was processed using Rapid Miner by generating a figure of 48% for cluster 1 consisting of, age 36 years, number of children 2, high school education, non-employment status. 22% for cluster 2 consisting of, age 46 years, number of children 3, high school education, non-working work status. 30% for cluster 3 consisting of, age 27 years, number of children 2, high school education, non-employment status. From the results of the processing, the application of clustering to the selection of contraceptives can be beneficial for Annisa Citeureup Hospital.

Keywords: Clustering, Contraception Tools, Family Planning Program, K-Means

1. Pendahuluan

Adanya kebijakan pemerintah untuk melakukan program Keluarga Berencana (KB) bersamaan dengan terbentuknya Badan Koordinasi Keluarga Berencana berdasarkan Keppres No. 8 Tahun 1970. Ada beberapa alasan mengapa pasangan tidak mengikuti program keluarga berencana yang dicanangkan oleh pemerintah, diantaranya kurangnya pengetahuan program keluarga berencana (biasanya ini terjadi di daerah terpencil yang kurangnya penyuluhan tentang program keluarga berencana), biaya, efek samping dan mungkin beberapa faktor lainnya. Namun faktor terbesar yang sangat mempengaruhi adalah efek yang ditimbulkan dari penggunaan program KB tersebut, diantaranya apabila pengguna tidak cocok pada alat kontrasepsi yang dipakainya.

Hal tersebut mungkin juga dipengaruhi oleh hormon, apabila semakin bertambah umur makin berubah juga hormon yang ada pada tubuh. Sehingga memungkinkan untuk melakukan penggunaan alat kontrasepsi jenis lainnya. Terdapat 3 fase kelompok usia reproduksi dalam mencapai sasaran yaitu fase menunda, fase menjarangkan, dan menghentikan atau mengakhiri kehamilan. Pertama, biasanya Pasangan Usia Subur (PUS) pada fase “menunda” ditunjukkan memiliki istri berusia kurang dari 20 tahun. Kedua, periode istri berusia 20-30 tahun dimana merupakan usia terbaik untuk mengandung dan melahirkan berada pada fase “menjarangkan”. Ketiga, pada periode ini umur si istri diatas 30 tahun yang merupakan fase “mengentikan atau mengakhiri kehamilan” [Setiawati et al., 2017].

Dalam pemilihan alat kontrasepsi khususnya di Rumah Sakit Annisa, si pemakai biasanya hanya mengukur pada kecocokannya karena sering menggunakan alat kontrasepsi tersebut. Ada beberapa faktor yang dijadikan sebagai penunjang pemakaian alat kontrasepsi. Sehingga dalam melakukan program keluarga berencana dalam pemilihan alat kontrasepsi banyak sekali faktor-faktor yang harus diperhatikan sebelum menggunakan alat kontrasepsi yang akan digunakan tersebut.

Penelitian ini dilakukan menggunakan teknik *data mining* dengan metode *clustering* pada penerapan algoritma *k-means*. Yakni dengan melakukan pengelompokkan berdasarkan pada ciri atau atribut pada variable penentuan penggunaan alat kontrasepsi.

Menurut Wu, Xindong & Kumar, Vipin Algoritma, algoritma *clustering* yang paling populer digunakan adalah *K-means* yang mempunyai keunggulan diantaranya mudah diimplementasikan dan merupakan algoritma yang sederhana [Widiarina, 2015].

Program pemerintah yaitu Keluarga Berencana (KB) dirancang untuk menyeimbangkan antara kebutuhan dan penduduk dengan memakai alat kontrasepsi. Dengan jumlah yang tidak sedikit para akseptor cenderung lebih memilih alat kontrasepsi dengan dasar coba-coba ataupun hanya sekedar menuruti masukan dari orang lain, ini menjadi salah satu dari kekurangan pemahaman dan pengetahuan mereka terhadap alat kontrasepsi. Sehingga untuk dapat mempermudah akseptor dalam memilih alat kontrasepsi yang efektif dan sesuai dengan kondisi tubuhnya maka diperlukan tindakan dan penanganan khusus. Penelitian ini melakukan analisa terhadap pemakaian jenis alat kontrasepsi dengan menggunakan algoritma C4.5. Hasil penelitian didapatkan sebanyak 14 rule dari tree decision algoritma C4.5 dengan jumlah class non hormonal sebanyak 8 rule dan jumlah class hormonal sebanyak 6 rule. Dalam pengujian Cross Validation didapat tingkat akurasi sebesar 85,38% [Wajhillah and Yulianti, 2017].

Penelitian mengenai Sistem Pendukung Keputusan dengan menggunakan algoritma *Naive Bayes* dinilai dapat berpotensi baik dalam sebuah klasifikasi untuk menentukan nilai akurasi dari pemilihan alat kontrasepsi. *Naive Bayes* juga merupakan klasifikasi probabilistik sederhana untuk mengkalkulasi sekelompok probabilitas dengan menjumlahkan frekuensi serta mengkombinasikan nilai. Variabel yang digunakan diantaranya jenis_kelamin, usia_istri, alat_kontrasepsi, jangka_waktu, jumlah_anak, usia_anak_terakhir dan faktor kesehatan lainnya. Diharapkan sistem keputusan ini dapat memberikan solusi bagi pasangan khususnya pada usia subur yang ingin melakukan program KB [Saputra et al., 2018].

Pemilihan dalam pemakaian atau tidak pada alat kontrasepsi di dalam tubuh merupakan salah satu permasalahan klasifikasi. Data diambil dari hasil Survey Prevalensi Kontrasepsi Nasional Indonesia tahun 1987 dengan jumlah responden 1.473 responden. Penelitian ini menggunakan metode klasifikasi salah satunya dengan metode Regresi logistik biner, *Support Vector Machine* (SVM), dan CART (*Classification and Regression Trees*). Pada penelitian ini memiliki 9 (sembilan) variabel bebas. Hasil yang diperoleh dari metode klasifikasi yaitu model SVM lebih baik dibandingkan dengan 2 (dua) metode yang lainnya (Regresi Logistik Biner dan CART). Tetapi bila metode digabung, maka didapat diperoleh bahwa variabel

respon yang diperoleh berdasarkan dari hasil grouping, *cluster k-means* atau *cluster kernel k-means* dapat menghasilkan *Apperant Error Rate* (APER) yang paling kecil [Suwardika, 2016].

Dalam meningkatkan pelayanan KB maka harus dilakukan penyuluhan dan pemberian informasi yang tepat kepada calon pengguna kontrasepsi agar program KB dapat berjalan sesuai harapan pemerintah. Berdasarkan hasil penelitian tersebut didapat bahwa pada kelompok usia 20-30 tahun pemilihan MKJP sebesar (32%) yaitu Implant (25%), AKDR (6%) , MOW (0%) sedangkan untuk pemilihan kontrasepsi *non* MKJP (68%) yaitu pada jenis KB suntik 3 bulan (50%), suntik 1 bulan (17%), pil (2%). Kedua, pada kelompok usia > 30 tahun pemilihan MKJP sebesar (38%) yaitu Implant (26%), AKDR (8%) , MOW (6%) sedangkan untuk pemilihan kontrasepsi *non* MKJP (62%) yaitu pada jenis KB suntik 3 bulan (48%), suntik 1 bulan (6%), pil (6%). Ketiga, tidak terdapat perbedaan pemilihan yang signifikan pada MKJP dan *non* MKJP berdasarkan efek samping pada kelompok usia reproduksi 20-30 tahun dengan pemilihan MKJP maupun *non* MKJP berdasarkan efek samping pada kelompok usia >30 tahun [Setiawati et al., 2017].

Tujuan dari penelitian ini adalah untuk mengetahui jenis alat kontrasepsi yang dapat digunakan dalam mengikuti program keluarga berencana (KB) berdasarkan usia dengan menggunakan metode K-Means.

2. Metode Penelitian

Penelitian ini dilakukan dengan menggunakan data primer, data dan informasi penelitian ini diperoleh dari instansi terkait yaitu data pengguna alat kontrasepsi pada tahun 2017 sebanyak 100 data pada Rumah Sakit Annisa yang digunakan untuk menentukan pemilihan alat kontrasepsi berdasarkan usia pengguna, jumlah anak, pendidikan dan status kerja.

Untuk dapat mengumpulkan data, penelitian ini melakukan observasi dimana penelitian dilakukan dengan pengamatan, pencatatan, dan pengumpulan data-data seperti usia pada istri, jumlah anak, pendidikan dan status pekerjaan sebagai variabel yang berkaitan dengan penggunaan alat kontrasepsi pada Rumah Sakit Annisa sebanyak 100 data pada tahun 2017. Selain itu, melakukan tanya jawab kepada Ka.Poli Bidan /KIA (Klinik Ibu Anak) mengenai hal-hal yang berkaitan dengan pemilihan alat kontrasepsi serta mencari referensi yang bersumber dari jurnal-jurnal dan buku-buku yang berkaitan tentang alat kontrasepsi, data mining clustering, dan algoritma k-means.

Pada penelitian ini mengolah data pemilihan alat kontrasepsi pada masyarakat sekitar khususnya daerah Citeureup, Bogor yang mengikuti program keluarga berencana di Rumah Sakit Annisa. Variabel yang digunakan pada penelitian diantaranya usia pada istri, jumlah anak, pendidikan dan status pekerjaan sesuai dengan kebutuhan penelitian yang diolah dengan metode K-Means menggunakan Rapid Miner.

Menurut Darmawati dalam [Saputra et al.,2018], Cara kerja alat kontrasepsi pada umumnya berfungsi untuk mencegah agar tidak terjadinya ovulasi, melumpuhkan sperma dan menghalangi bertemunya antara sel telur dan sel sperma [Saputra et al.,2018]. Salah satu faktor penentu kesuburan dan prediktor yang paling penting dari transisi fertilitas adalah Kontrasepsi. Untuk analisa pemakaian alat kontrasepsi dipengaruhi oleh beberapa faktor diantaranya sosial, budaya, dan ekonomi sehingga pendekatan secara multidimensional perlu diadopsi. Dalam menganalisis suatu indikator tunggal tidak memungkinkan untuk mengambil semua dimensi dari pemilihan alat kontrasepsi [Priyanti, 2015].

Menurut [Suwardika, 2016], berbagai jenis kontrasepsi diantaranya: 1. Kontrasepsi sterilisasi, seperti tubektomi dan vasektomi, 2. Kontrasepsi teknik, yang terdiri dari *Coitus Interruptus* (senggama terputus), Sistem kalender (pantang berkala), *Prolonged lactation* (menyusui), 3. Kontrasepsi mekanik, terdiri dari kondom, *Spermatisida*, *Vaginal diafragma*, IUD (*Intra Uterine Device*) atau spiral dan 4. Kontrasepsi hormonal seperti pil KB, suntik, susuk (Implan), spiral berhormon.

Menurut Shimonski dalam [Sumarna, 2015] menjelaskan bahwa “*Cluster* adalah sekelompok mesin yang bertindak sebagai sebuah entitas tunggal untuk menyediakan sumber daya dan layanan ke jaringan dengan tujuan untuk menjaga ketersediaan *resource* bagi *client* ketika terjadi kegagalan *software* maupun *hardware* pada *server*.”

Menurut [Aggarwal & Reddy, 2014] *Clustering* berguna dalam beberapa pembelajaran mesin dan tugas penambangan data termasuk segmentasi gambar, pengambilan informasi, pengenalan pola, klasifikasi pola, analisis jaringan dan sebagainya. itu bisa dilihat sebagai

tugas eksplorasi atau langkah sebelum pemrosesan. Jika tujuannya adalah untuk mengekspos dan mengungkapkan pola tersembunyi dalam data, pengelompokan menjadi tugas eksplorasi yang berdiri sendiri. Namun, jika *kluster* yang dihasilkan akan digunakan untuk memfasilitasi penambangan data lain atau tugas pembelajaran mesin, pengelompokan akan menjadi langkah sebelum pemrosesan.

Algoritma *k-means* menetapkan nilai-nilai *cluster* (K) secara acak, hasil sementara dari nilai tersebut menjadi pusat *cluster* atau bisa disebut sebagai *centroid*, *mean*, atau "*means*". Selanjutnya menghitung jarak setiap data dengan masing-masing *centroid* menggunakan rumus *Euclidian* sampai ditemukan jarak terdekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*. Lakukan langkah tersebut sampai nilai *centroid* tidak berubah [Vulandari, 2017:54].

Data *clustering* menggunakan metode *K-Means* ini secara umum dilakukan dengan algoritma dasar sebagai berikut :1. Tentukan jumlah *cluster*, 2. Alokasikan data yang tersedia ke dalam *cluster* secara acak, 3. Hitung rata-rata dari data yang ada di masing-masing *cluster* dan 4. Alokasikan masing-masing data ke dalam rata-rata terdekat. Kembali ke step 3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai rata-rata ada yang di atas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan di atas nilai *threshold* yang ditentukan [Elyana, 2015].

a. Menghitung Jarak Terdekat

Untuk menghitung jarak terdekat, pada penelitian ini digunakan L_2 (*Euclidian*) *distance space*, jarak antara dua titik hitung dengan rumus sebagai berikut :

$$D_{L_2}(x_2, x_1) = \|x_2 - x_1\|_2$$

$$= \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2}$$

Dimana p merupakan dimensi data, adapun data p di sini berjumlah 4 yaitu sesuai dengan variabel yang digunakan sebanyak 4.

b. Pengukuran Nilai Akurasi (*Accuration Rate*)

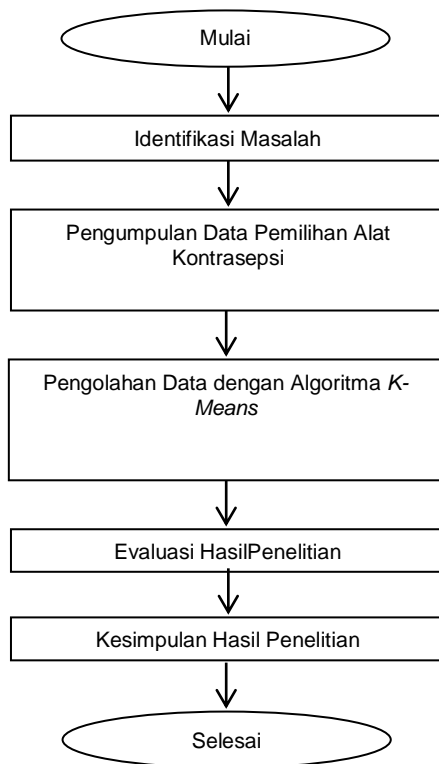
Nilai akurasi merupakan perbandingan antara jumlah *cluster* yang sesuai dengan label pada data asli dibandingkan dengan jumlah data keseluruhan.

$$\text{Accuration Rate} = \frac{\text{jumlah data yang sesuai labe aslinya}}{\text{jumlah data keseluruhan}} \times 100\%$$

3. Hasil dan Pembahasan

Langkah-langkah yang dilakukan dalam tahapan penelitian ini sebagai berikut: **1) identifikasi masalah** bertujuan untuk mengetahui rentang usia dalam melakukan pemilihan alat kontrasepsi di Rumah Sakit Annisa. **2) pengumpulan data**, tahap pengumpulan data dilakukan dengan observasi data laporan pemilihan alat kontrasepsi pada tahun 2017 di Rumah Sakit Annisa. **3) pengolahan data**, tahap pengolahan data dilakukan menggunakan data mining *clustering* dengan algoritma *K-Means* pada data pemilihan alat kontrasepsi di Rumah Sakit Annisa. Sehingga hasilnya dapat berguna untuk dijadikan sebagai pengetahuan oleh pihak Rumah Sakit Annisa agar dapat menunjang sarana maupun prasarana tentang pemilihan alat kontrasepsi. **4) evaluasi hasil penelitian**, tahap evaluasi dilakukan dengan melakukan pengujian dari hasil penelitian menggunakan *RapidMiner* agar hasil olahan yang disajikan dapat mudah dipahami. **5) kesimpulan hasil penelitian**, pada tahap ini penelitian yang telah selesai akan diberikan kesimpulan dari permasalahan yang ada agar dapat digunakan oleh pihak Rumah Sakit Annisa.

Berdasarkan tahapan yang telah dijelaskan dibuat gambaran yang menunjukkan tahapan penelitian yang dilakukan seperti pada Gambar 1.



Sumber: Hasil Penelitian (2019)

Gambar 1. Tahapan Penelitian

Sampel yang digunakan pada penelitian ini adalah laporan data pengguna alat kontrasepsi pada Rumah Sakit Annisa pada tahun 2017 sebanyak 100 data pengguna.

Tabel 1. Populasi dan Sampel

Usia	Jenis Alat Kontrasepsi
22-30	Pil, Suntik
31-41	Spiral (IUD), Implan
42-50	Steril

Sumber: Hasil Pengolahan Data (2019)

Dalam penelitian ini penulis melakukan pengelompokan dengan menggunakan 4 variabel, yaitu :

Tabel 2. Variabel Perhitungan

No.	Variabel
1	Usia Istri
2	Jumlah Anak
3	Pendidikan
4	Status Pekerjaan

Sumber: Hasil Pengolahan Data (2019)

Pada variabel pendidikan data masih berupa teks maka harus dilakukan inisialisasi terlebih dahulu. Yaitu dengan merubah isi data menjadi numerik agar memudahkan dalam proses perhitungannya.

Tabel 3. Data Master Pendidikan

Kriteria	Nilai
SD	1
SMP	2
SMA	3
Sarjana	4

Sumber: Hasil Pengolahan Data (2019)

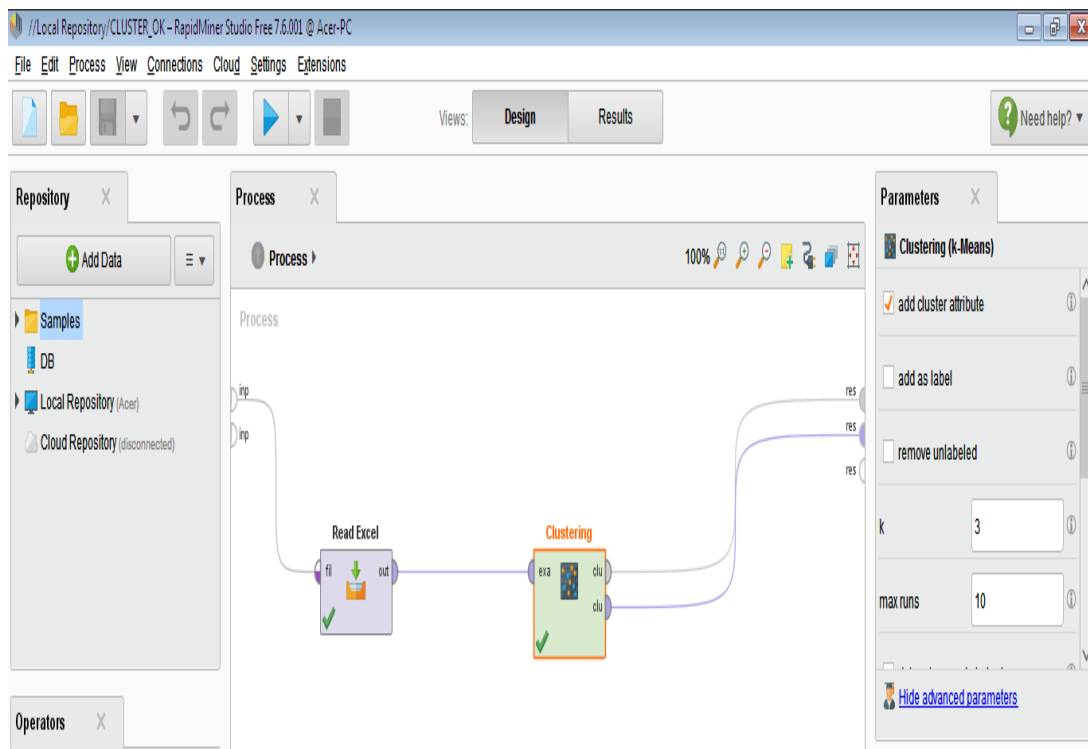
Pada variabel status kerja data masih berupa teks maka harus dilakukan inisialisasi terlebih dahulu. Yaitu dengan merubah isi data menjadi numerik agar memudahkan dalam proses perhitungannya.

Tabel 4. Data Master Status Kerja

Kriteria	Nilai
Kerja	0
Tidak Kerja	1

Sumber: Hasil Pengolahan Data (2019)

Penerapan algoritma *K-Means* pada pemilihan alat kontrasepsi memiliki langkah-langkah sebagai berikut :



Sumber: Hasil Pengolahan Data (2019)

Gambar 2. Menentukan Jumlah K

1. Data cluster

Menghasilkan 100 data yang telah menemukan jarak terdekatnya sehingga menempatkan data ke dalam masing-masing cluster. Di bawah ini adalah data-data yang telah dikelompokkan ke dalam kelompok *cluster* terdekat.

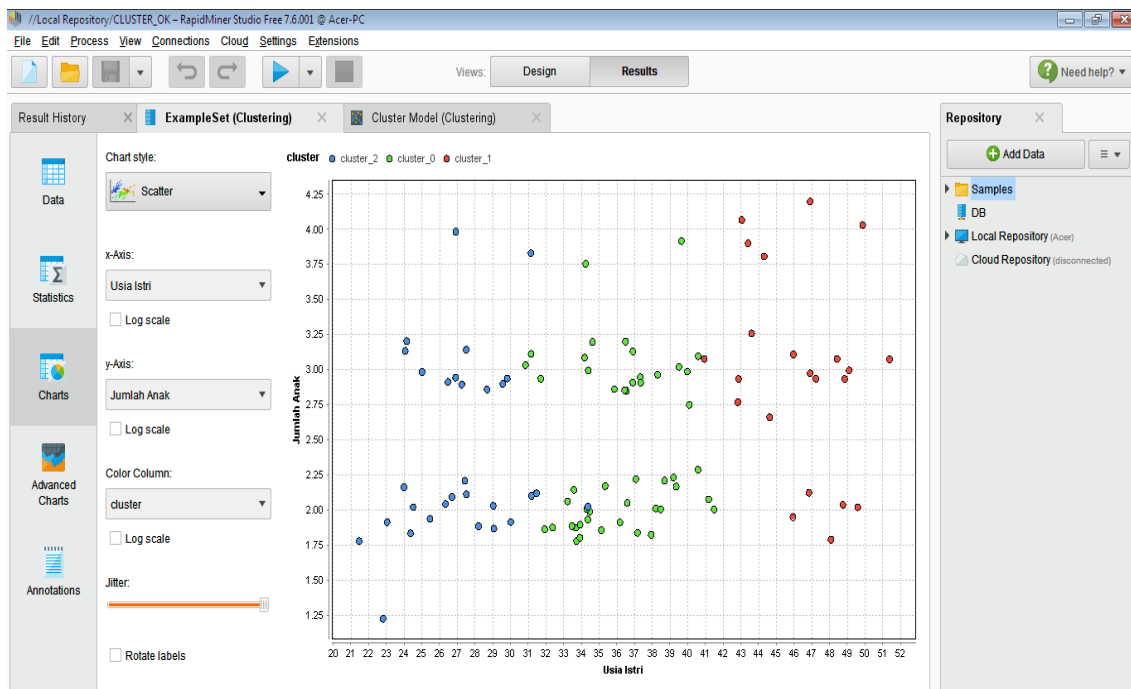
Row No.	id	cluster	Usia Istri	Jumlah Anak	Pendidikan	Status Kerja
1	1	cluster_2	27	3	3	0
2	2	cluster_0	37	2	3	1
3	3	cluster_0	40	4	3	0
4	4	cluster_2	28	2	2	0
5	5	cluster_2	29	4	4	0
6	6	cluster_1	45	2	3	1
7	7	cluster_0	40	2	3	1
8	8	cluster_2	31	2	3	0
9	9	cluster_0	35	2	2	0
10	10	cluster_2	25	2	2	0
11	11	cluster_0	33	3	4	1
12	12	cluster_0	38	2	2	1
13	13	cluster_2	25	2	3	0
14	14	cluster_1	47	4	3	0
15	15	cluster_0	35	3	2	0
16	16	cluster_1	49	2	3	1
17	17	cluster_0	33	2	3	1

Sumber: Hasil Pengolahan Data (2019)

Gambar 3. Pusat Cluster Pada Data 1-17

2. **Chart atau grafik:**

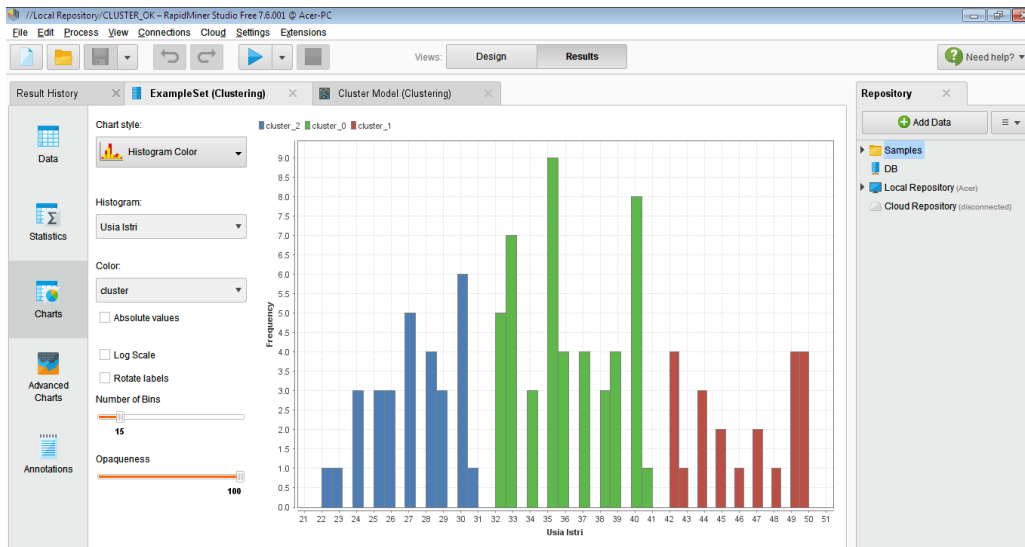
- a. **Scatter** (Grafik Penyebaran), grafik ini menggambarkan penyebaran yang dilakukan masing-masing cluster.



Sumber: Hasil Pengolahan Data (2019)

Gambar 4. Scatter Chart

- b. Histogram, merupakan *chart* atau grafik berbentuk seperti batang yang menjelaskan tentang masing-masing *cluster* dari variabel.

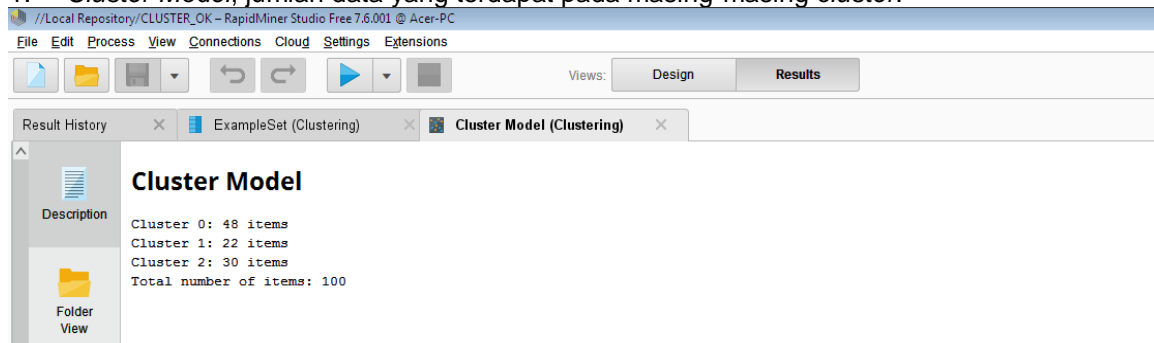


Sumber: Hasil Pengolahan Data (2019)

Gambar 5. Histogram

Hasil yang digunakan pada *Cluster Model* terdiri dari :

1. *Cluster Model*, jumlah data yang terdapat pada masing-masing *cluster*.



Sumber: Hasil Pengolahan Data (2019)

Gambar 6. Hasil *Cluster Model*

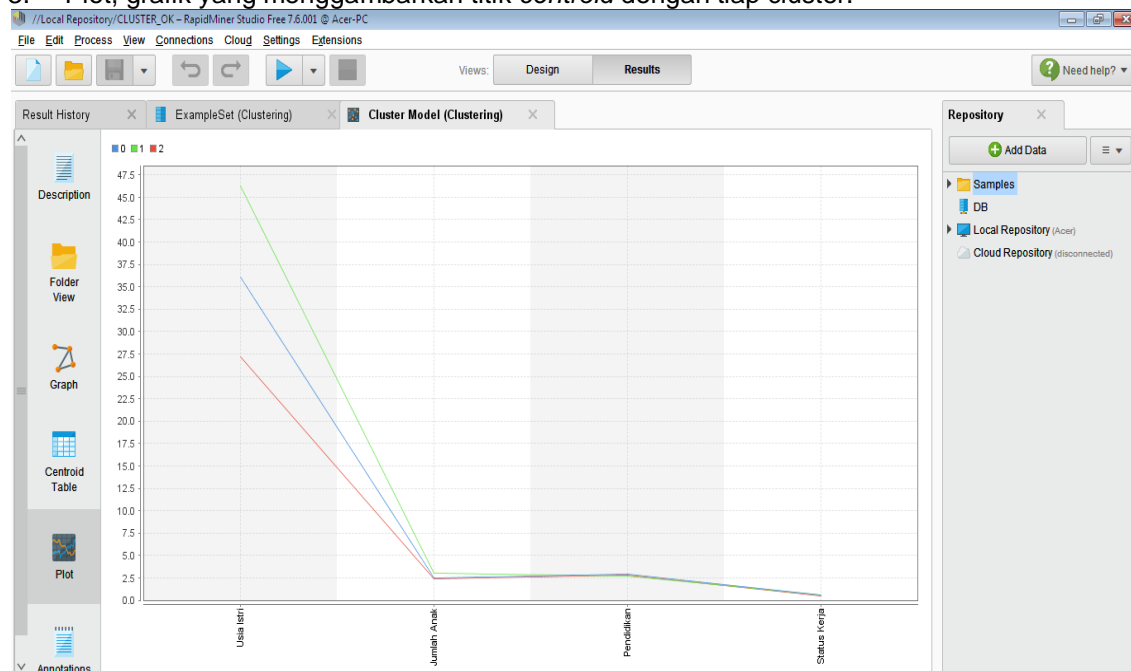
2. *Centroid Table*, table yang menampilkan nilai *centroid* dalam masing-masing *cluster*.

Attribute	cluster_0	cluster_1	cluster_2
Usia Istri	36.062	46.227	27.167
Jumlah Anak	2.479	3	2.433
Pendidikan	2.917	2.727	2.800
Status Kerja	0.542	0.591	0.500

Sumber: Hasil Pengolahan Data (2019)

Gambar 7. Tabel *Centroid*

3. Plot, grafik yang menggambarkan titik *centroid* dengan tiap cluster.



Sumber: Hasil Pengolahan Data (2019)

Gambar 8. Hasil Plot

4. Kesimpulan

Hasil akurasi yang didapatkan yaitu : 48% untuk *cluster* 1 yang terdiri dari , usia 36 tahun, jumlah anak 2, pendidikan SMA, status kerja tidak kerja. 22% untuk *cluster* 2 yang terdiri dari , usia 46 tahun, jumlah anak 3, pendidikan SMA, status kerja tidak kerja. 30% untuk *cluster* 3 yang terdiri dari , usia 27 tahun, jumlah anak 2, pendidikan SMA, status kerja tidak kerja.

Ucapan Terima Kasih

Terima kasih kepada Allah Subhana huwata'ala dan keluarga serta instansi yang telah bekerjasama dan mendukung dalam penyelesaian penelitian ini.

Referensi

- Aggarwal CC, Reddy CK. 2014. Data Clustering Algorithms And Applications. Taylor & Francis Group, LLC. 32 P.
- Elyana I. 2015. Penerapan Algoritma K.Mean Untuk Penentuan Pencocokan Pewarnaan Clustering Secara Otomatis Pada Produk Fashion. J. Paradig. XVII: 85–91.
- Priyanti E. 2015. Peningkatan Backward Elimination Dengan Windowed Momentum Untuk Prediksi Kontrasepsi. J. Paradig. XVII: 34–43.
- Saputra RA, Taufik AR, Ramdhani LS, Oktapian R, Marsusanti E. 2018. Sistem Pendukung Keputusan Dalam Menentukan Metode Kontrasepsi Menggunakan Algoritma Naive Bayes. Semin. Nas. Inov. Dan Tren: 106–111.
- Setiawati E, Handayani OWK, Kuswardinah A. 2017. Pemilihan Kontrasepsi Berdasarkan Efek Samping Pada Dua Kelompok Usia Reproduksi. Unnes J. Ofpublic Heal. 6: 167–173.
- Sumarna. 2015. Perancangan Clustering Ujian Online Studi Kasus Bina Sarana Informatika. J. Techno Nusa Mandiri XII: 35–40.
- Suardika G. 2016. Pengelompokan Dan Klasifikasi Penggunaan Kontrasepsi Di Indonesia. J. Mat. Saint, Dan Teknol. 17: 9–19.
- Vulandari RT. 2017. Data Mining Teori Dan Aplikasi Rapidminer, Cetakan I. Penerbit Gava Media. 1-2, 53-54 P.

- Wajhillah R, Yulianti I. 2017. Penerapan Algoritma C4.5 Untuk Prediksi Penggunaan Jenis Kontrasepsi Berbasis Web. Kumpul. J. Ilmu Komput. 4: 174–187.
- Widiarina. 2015. Klastering Data Menggunakan Algoritma Dynamic K-Means. Tek. Komput. AMIK BSI I: 260–265.