
PENERAPAN METODE K-MEANS PEMETAAN CALON PENERIMA JAMKESDA

¹M. Nanda Variestha Waworuntu, ²Muhammad Faisal Amin

Prodi Teknik Informatika STMIK Banjarbaru

Jl. A. Yani KM 33,3 Loktabat, Banjarbaru, 0511 4782882

nanda.variestha@gmail.com, faisal.amin@stmik-banjarbaru.ac.id

Abstract

Kelurahan Kemuning, one of the Social Welfare Section, there is poor community service to receive Regional Health Insurance. During this section of Social Welfare Section in Kelurahan Kemuning, there is no method that can classify the level of poverty so that the beneficiaries on target, so the Kelurahan can't prevent the inaccuracies. Therefore, poverty grouping can assist Kelurahan in making the right decision to prevent the inaccuracies of recipients of Regional Health Insurance. In this research, the application of the k-means method is implemented in an application made with 2 clusters. This study uses as many as 440 data samples. From result of calculation of Davies Bouldin Index obtained value determination of cluster amount with value 2 cluster (0,243), 3 cluster (0,256), 4 cluster (0,275). The value used is 2 clusters because the value is close to 0.

Keywords: : data mining, k-means, poverty, davies bouldin index

Abstrak

Pada Kelurahan Kemuning salah satunya Seksi Kesejahteraan Sosial (KESSOS) terdapat pelayanan masyarakat miskin untuk menerima bantuan Jaminan Kesehatan Daerah (JAMKESDA). Selama ini bagian Seksi KESSOS pada Kelurahan Kemuning belum ada metode yang dapat mengelompokkan tingkat kemiskinan agar penerima bantuan tepat sasaran, sehingga pihak Kelurahan tidak dapat mencegah ketidaktepat sasaran tersebut. Oleh sebab itu, pengelompokan kemiskinan dapat membantu pihak Kelurahan dalam mengambil keputusan yang tepat untuk mencegah ketidaktepat sasaran penerima JAMKESDA. Pada penelitian ini, penerapan metode K-Means diimplementasikan pada aplikasi yang dibuat dengan 2 klaster. Penelitian ini menggunakan sebanyak 440 sampel data. Dari hasil perhitungan Davies Bouldin Index diperoleh nilai penentuan jumlah cluster dengan nilai 2 klaster (0.243), 3 klaster (0.256), 4 klaster (0.275). Nilai yang digunakan adalah 2 klaster karena nilai tersebut mendekati 0.

Kata kunci: data mining, k-means, kemiskinan, davies bouldin index

1. PENDAHULUAN

Beberapa seksi yang terdapat pada Kelurahan Kemuning salah satunya Seksi Kesejahteraan Sosial (KESSOS) terdapat pelayanan masyarakat miskin untuk menerima bantuan Jaminan Kesehatan Daerah (JAMKESDA). Namun selama ini belum ada aplikasi yang diimplementasikan untuk pengelompokan tingkat kemiskinan untuk penerima bantuan JAMKESDA.

Menurut data dari Seksi KESSOS [1] 440 orang yang diajukan bantuan JAMKESDA pada tahun 2015 yang menerima berjumlah 375 orang (85,23%) tepat sasaran, dan yang tidak tepat sasaran berjumlah 65 orang (14,77%). Ketidaktepat sasaran ini akan mempengaruhi pengeluaran dana pemerintah.

Selama ini bagian Seksi KESSOS pada Kelurahan Kemuning belum ada metode yang dapat mengelompokkan tingkat kemiskinan agar penerima bantuan tepat sasaran, sehingga pihak Kelurahan tidak dapat mencegah ketidaktepat sasaran tersebut. Untuk mencegah kondisi tersebut terjadi, maka diperlukan penelitian untuk mengolah data tersebut yang dimiliki oleh Seksi KESSOS, data yang dimanfaatkan adalah data penerima bantuan JAMKESDA pada tahun 2016 untuk pengelompokan tersebut dengan atribut jumlah penghasilan dan jumlah individu. Penelitian ini diperlukan karena jika tingkat kemiskinan dapat dikelompokkan lebih dini, maka pihak yang bersangkutan untuk mengajukan bantuan tersebut mampu mencegah ketidaktepat sasaran untuk penerima bantuan JAMKESDA.

Beberapa penelitian terdahulu yang menerapkan pengelompokan menggunakan metoda *K-Means*. Penelitian pertama yang dilakukan oleh Cahyo Aji Nugroho, Rully A. Hendrawan dan Irmasari Hafidz [2]. Aplikasi yang mereka terapkan berhasil menyelesaikan masalah dan dapat disimpulkan bahwa dapat membantu Bapemas Kota Surabaya dalam melakukan kebijakan pemilihan pelatihan yang tepat berdasarkan karakteristik KSM.

Kemudian penelitian yang dilakukan oleh peneliti Zainul Aras Z dan Sarjono [3]. Aplikasi yang mereka terapkan berhasil menyelesaikan masalah, dapat disimpulkan bahwa dapat dipergunakan sebagai salah satu alat sistem yang membantu keputusan di dalam menentukan kelompok Penduduk Prioritas yang mendapatkan bantuan bedah rumah, sedangkan pada penelitian ini juga menggunakan metode *K-Means* dalam pemetaan JAMKESDA, karena algoritma *K-Means Clustering* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi serta umum penggunaannya dalam praktek [4].

Penelitian yang dilakukan oleh Warnia Nengsih, Program Studi Sistem Informasi Komputer, Politeknik Caltex Riau, Pekanbaru tentang "*Descriptive Modelling Menggunakan K-Means Untuk Pengklasteran Tingkat Kemiskinan Di Propinsi Riau*", permasalahan yang terjadi adalah kesulitan dalam menentukan wilayah mana yang mengalami tingkat kemiskinan yang paling tinggi dan normal serta wilayah dengan tingkat kemiskinan rendah menjadi sebuah alasan untuk melakukan pengklasteran ini. Hasil *cluster* yang diperoleh dimana *record* 3 dan *record* 9 berada pada *cluster* 2. *Record* 1,2,4,5,6,7,8,10,11,12 berada pada *cluster* 3. Tidak ada kota atau kabupaten yang berada pada *cluster* 1. Metode ini memberikan pengetahuan pengklusteran wilayah yang tersebar di Propinsi Riau, sehingga gambaran yang diperoleh dapat menjadi acuan bagi Pemerintah dalam pelaksanaan dan pengaturan sebuah kebijakan [5].

Penelitian yang dilakukan oleh Febi Fajrianti, Muhammad Nadjib Bustan dan Muhammad Arif Tiro Fakultas Matematika Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Makassar tentang "*Penggunaan Analisis Cluster K-Means Dan Analisis Diskriminan Dalam Pengelompokan Desa Miskin Di Kabupaten Pangkep*", permasalahan yang terjadi adalah terkhusus di Sulawesi Selatan, salah satu daerah yang menjadi sorotan fokus penanggulangan kemiskinan yaitu di Kabupaten

Pangkep, karena memiliki tingkat persentase kemiskinan yang cukup tinggi dibandingkan dengan kabupaten lainnya. Kebijakan kemiskinan oleh Pemerintah Kabupaten Pangkep, selain harus mampu memperkecil jumlah penduduk miskin, juga sekaligus harus bisa mengurangi tingkat kedalaman dan keparahan dari kemiskinan. Tahapan analisis data yang dilakukan yaitu melakukan pengelompokan awal dengan menggunakan analisis *cluster K-Means*, dari beberapa *cluster* yang dibentuk selanjutnya akan dilakukan analisis diskriminan untuk melihat akurasi pengelompokan. Berdasarkan analisis diskriminan diperoleh 98,06% (3 *cluster*), 94,17% (4 *cluster*), 91,26% (5 *cluster*), 94,17% (6 *cluster*), 86,41% (7 *cluster*), dan 88,35% (8 *cluster*). Dari uraian di atas diperoleh bahwa dengan membentuk pengelompokan dengan 3 *cluster*, maka hasil akurasi pengelompokan yang diperoleh lebih baik dari cluster lain yang dibentuk [6].

Penelitian dilakukan oleh Aniq Noviciatie Ulfah dan Shofwatul Uyun Program Studi Teknik Informatika Universitas Islam Negeri Sunan Kalijaga tentang “Analisis Kinerja Algoritma Fuzzy C-Means dan K-Means pada Data Kemiskinan”, Salah satu upaya untuk mewujudkan program pemerintah Kabupaten Gunungkidul dalam rangka untuk pengentasan kemiskinan adalah dengan melakukan pendataan data kemiskinan warganya. Pemerintah telah merumuskan pendataan dengan melakukan pembobotan terhadap 15 indikator kedalam 3 kelompok. Banyaknya data dan indikator yang harus digunakan tentunya akan menimbulkan kesulitan dalam pelaksanaannya, tidak efektif dan kurang obyektif. Oleh karena itu diperlukan otomatisasi dalam proses clustering data kemiskinan. Penelitian ini bertujuan untuk menganalisis kinerja antara algoritma FCM dan K-means yang diimplementasikan pada data kemiskinan di Desa Girijati Purwosari menjadi 3 cluster. Beberapa tahapan yang harus dilakukan sebelum dilakukan clustering, terlebih dahulu dilakukan prapengolahan yaitu data cleaning dan data transformation untuk selanjutnya dilakukan clustering menggunakan kedua algoritma tersebut. Hasil perhitungan digunakan untuk membandingkan antara algoritma FCM dengan K-Means. Kesesuaian data antara algoritma FCM dengan perhitungan indikator kemiskinan di Desa Girijati sebesar 50% dan untuk algoritma K-Means sebesar 83,33%. Algoritma K-Means lebih tepat digunakan pada pengelompokan data kemiskinan berdasarkan ketiga kriteria kemiskinan dibandingkan algoritma FCM [7].

2. METODOLOGI PENELITIAN

Penelitian ini adalah penelitian eksperimen dengan langkah-langkah penelitian sebagai berikut:

a. Pengumpulan data

Data yang digunakan adalah data penduduk seperti NIK, Nama, Tempat Tanggal Lahir, Jenis Kelamin, Status Perkawinan, Alamat, Pekerjaan. Data ini diambil dari Kelurahan Kemuning. Data lain yang dibutuhkan adalah data Jumlah Penghasilan dan Jumlah Anggota Keluarga dari Seksi Kesejahteraan Sosial dan Seksi Ekonomi Pembangunan.

- b. Pre-processing
Pada tahap pre-processing dilakukan data integration dan. Data integration yaitu penggabungan data dari data penduduk dan data jumlah penghasilan serta jumlah anggota keluarga sehingga menjadi 1 tabel.
- c. Clustering
Pada tahap ini, algoritma yang digunakan adalah algoritma *K-means*, sebanyak 3 kali percobaan dengan nilai 2, 3, dan 4 cluster. Percobaan dilakukan berulang-ulang dengan menggunakan software RapidMiner hingga diperoleh nilai optimal, yaitu nilai dimana nilai k mendekati 0 dengan menggunakan *cluster distance performance* dengan *main criterion* Davies Bouldin. Nilai optimal ini diperoleh pada k berjumlah 2 dengan nilai 0.243.

2.1 JAMKESMAS dan JAMKESDA

JAMKESMAS (Jaminan Kesehatan Masyarakat) dan JAMKESDA (Jaminan Kesehatan Daerah) sebenarnya sama-sama program jaminan kesehatan yang diperuntukkan untuk fakir miskin namun ruang lingkupnya berbeda. JAMKESMAS adalah program yang ruang lingkupnya lebih luas yang diperuntukkan untuk warga miskin di seluruh Indonesia, sedangkan JAMKESDA adalah program jaminan bantuan pembayaran biaya pelayanan kesehatan yang diberikan Pemerintah Daerah Kabupaten atau Provinsi untuk menyasar warga fakir miskin yang belum memiliki JAMKESMAS. Adanya JAMKESDA dikarenakan tidak semua warga miskin yang tinggal di daerah terjaring program JAMKESMAS oleh karena itu untuk mengatasi masalah tersebut diluncurkan program JAMKESDA [11].

2.2 Data Mining

Data mining adalah proses menemukan pola yang menarik dan pengetahuan dari data yang berjumlah besar [8]. *Data mining* adalah aktivitas yang menggambarkan sebuah proses analisis yang terjadi secara iteratif pada *database* yang besar, dengan tujuan mengekstrak informasi dan *knowledge* yang akurat dan berpotensi berguna untuk *knowledge workers* yang berhubungan dengan pengambilan keputusan dan pemecahan masalah [9].

Aktivitas *data mining* dapat dipisahkan menjadi 6, berdasarkan tujuan dari analisis yaitu [10]:

- a. Estimasi
Estimasi mirip dengan klasifikasi, kecuali bahwa variabel target bersifat numerik, bukan kategorikal. Model dibangun menggunakan "lengkap" catatan, yang memberikan nilai dari variabel target serta prediktor. Kemudian, untuk pengamatan baru, perkiraan nilai variabel target dibuat, berdasarkan nilai-nilai prediktor.
- b. Deskripsi
Terkadang, para peneliti dan analis hanya mencoba menemukan cara untuk mendeskripsikan pola dan tren yang ada dalam data. Sebagai contoh, seorang jajak pendapat dapat mengungkap bukti bahwa mereka yang diberhentikan cenderung tidak mendukung calon petahana saat ini dalam pemilihan presiden. Deskripsi pola dan tren sering menyarankan penjelasan yang mungkin untuk pola dan tren tersebut.

- c. **Prediksi**
Prediksi serupa dengan klasifikasi dan estimasi, kecuali untuk prediksi, hasil dari kesalahan ini. Contoh dari laporan kinerja dalam bisnis dan penelitian yaitu :
- 1) Memprediksi harga saham tiga bulan ke depan
 - 2) Memprediksi persentase peningkatan trafik kematian tahun depan jika batas kecepatan dinaikkan.
 - 3) Memprediksikan yang lebih baik dari yang terbaik adalah bola *WorldSeries*, berdasarkan perbandingan statistik tim.
- d. **Klasifikasi**
Dalam klasifikasi, ada target variabel kategori, seperti braket pendapatan, misalnya dapat dipartisi menjadi tiga kelas atau kategori: pendapatan tinggi, pendapatan menengah, dan pendapatan rendah. Model penambangan data memeriksa sejumlah besar catatan, setiap catatan yang berisi informasi tentang variabel target serta aset variabel *input* atau prediktor.
- e. **Clustering**
Pengelompokan mengacu pada pengelompokan catatan, pengamatan, atau kasus ke dalam kelas-kelas objek yang serupa. Pernyataan interaktif untuk merekam apa yang cocok untuk orang lain, dan tidak sama dengan catatan dalam kelompok lain. *Clustering* berbeda dari klasifikasi dimana tidak ada variabel target untuk pengelompokan. Tugas pengelompokan tidak mencoba mengklasifikasikan, memperkirakan, atau memprediksi nilai variabel target. Algoritma pengelompokan berusaha mengelompokkan seluruh kumpulan data ke dalam sub kelompok atau gugus yang relatif homogen, dimana kesamaan catatan dalam *cluster* dimaksimalkan dan kemiripan dengan rekaman di luar *cluster* diminimalkan.
- f. **Asosiasi**
Tugas asosiasi untuk penggalian data adalah pekerjaan menemukan atribut yang “*go together*.” Paling umum dalam dunia bisnis, dimana ia dikenal sebagai *affinity analysis or market basket analysis*, tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

2.3 **K-Means**

K-Means merupakan metode klasterisasi yang paling terkenal dan banyak digunakan diberbagai bidang karena sederhana, mudah diimplementasikan, memiliki kemampuan untuk mengklaster data yang besar dan *kompleksitas* waktunya *linear* $O(nKT)$ dengan n adalah jumlah dokumen, K adalah jumlah klaster, dan T adalah jumlah iterasi. *K-means* merupakan metode pengklasteran secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. Dengan *partitioning* secara iterasi, *K-Means* mampu meminimalkan rata-rata jarak setiap data ke klasternya. Metode ini dikembangkan oleh Mac Queen pada tahun 1967.

Dasar algoritma *K-means* adalah sebagai berikut :

- a. Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
- b. Bangkitkan k *centroid* (titik pusat klaster) awal secara random/acak, kemudian untuk menghitung *centroid cluster* k -i berikutnya, digunakan rumus sebagai berikut:

(2.1)

$$v = \sqrt{\frac{\sum_{i=1}^n x_i}{n}}$$

Dimana:

i : 1,2,3,...n

v : *centroid* pada *cluster*

x_i : objek ke- i

n : banyaknya objek/jumlah objek yang menjadi anggota *cluster*

- c. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.2)$$

Ket :

d_{ij} : Jarak objek antara objek x dan y

n : Jumlah Atribut

X_i : Objek Data

Y_i : Data *Cluster*

- d. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*-nya (C).
e. Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

Adapun karakteristik dari algoritma *K-Means* salah satunya adalah sangat sensitif dalam penentuan titik pusat awal *cluster* karena *K-Means* membangkitkan titik pusat *cluster* awal secara random. Pada saat pembangkitan awal titik pusat yang random tersebut mendekati solusi akhir pusat *cluster*, *K-Means* mempunyai posibilitas yang tinggi untuk menemukan titik pusat *cluster* yang tepat. Sebaliknya, jika awal titik pusat tersebut jauh dari solusi akhir pusat *cluster*, maka besar kemungkinan ini menyebabkan hasil pengklasteran yang tidak tepat. Akibatnya *K-Means* tidak menjamin hasil pengklasteran yang unik. Inilah yang menyebabkan metode *K-Means* sulit untuk mencapai optimum global, akan tetapi hanya minimum lokal. Selain itu, algoritma *K-Means* hanya bisa digunakan untuk data yang atributnya bernilai *numeric*. [12]

2.3 Davies-Bouldin Index

Davies-bouldin index merupakan salah satu metode evaluasi internal yang mengukur evaluasi cluster pada suatu metode pengelompokan yang didasarkan pada nilai kohesi dan separasi. Dalam suatu pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap centroid dari cluster yang diikuti. Sedangkan separasi didasarkan pada jarak antar centroid dari clusternya. Sum of square within cluster (SSW) merupakan persamaan yang digunakan untuk mengetahui matrik kohesi dalam sebuah cluster ke- i yang dirumuskan sebagai berikut :

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (2.3)$$

Dari persamaan tersebut, m_i merupakan jumlah data dalam cluster ke- i , c_i adalah centroid cluster ke- i , dan $d()$ merupakan jarak setiap data kecentroid yang dihitung menggunakan jarak euclidean.

Sum of square between cluster (SSB) merupakan persamaan yang digunakan untuk mengetahui separasi antar cluster yang dihitung menggunakan persamaan :

$$SSB_{i,j} = d(c_i, c_j) \quad (2.4)$$

Setelah nilai kohesi dan separasi diperoleh, kemudian dilakukan pengukuran rasio (R_{ij}) untuk mengetahui nilai perbandingan antara cluster ke- i dan cluster ke- j . Cluster yang baik adalah cluster yang memiliki nilai kohesi sekecil mungkin dan separasi yang sebesar mungkin. Nilai rasio dihitung menggunakan persamaan sebagai berikut :

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (2.5)$$

Nilai rasio yang diperoleh tersebut digunakan untuk mencari nilai daviesbouldin index (DBI) dari persamaan berikut :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (2.6)$$

Dari persamaan tersebut, k merupakan jumlah cluster yang digunakan. Semakin kecil nilai DBI yang diperoleh (non-negatif ≥ 0), maka semakin baik cluster yang diperoleh dari pengelompokan K-means yang digunakan [4]

3. HASIL DAN PEMBAHASAN

3.1 Hasil Clustering

Cluster Model yang dihasilkan dari pengelompokan data penduduk dengan algoritma *K-Means* ditampilkan dalam Tabel 1, Tabel 2 dan Tabel 3. Pada Tabel 1 tampak bahwa jumlah *cluster* yang digunakan sebanyak 2 dengan keanggotaan *Cluster 0* terdiri dari 334 *item*, *Cluster 1* terdiri dari 106 *item* dengan jumlah keseluruhan adalah sebanyak 440 *item*. Pada Tabel 2 tampak bahwa jumlah *cluster* yang digunakan sebanyak 3 dengan keanggotaan *Cluster 0* terdiri dari 43 *item*, *Cluster 1* terdiri dari 87 *item*, dan *Cluster 2* terdiri dari 310 *item* dengan jumlah keseluruhan adalah sebanyak 440 *item*. Pada Tabel 3 tampak bahwa jumlah *cluster* yang digunakan sebanyak 4 dengan keanggotaan *Cluster 0* terdiri dari 60 *item*, *Cluster 1* terdiri dari 52 *item*, *Cluster 2* terdiri dari 302, dan *Cluster 3* terdiri dari 26

item dengan jumlah keseluruhan adalah sebanyak 440 item. Percobaan pertama dengan 2 cluster yang lebih optimal dikarenakan nilai *Performance Vektor (Davies Bouldin Index)* mendekati nilai 0 dengan nilai 0.243.

Tabel 1 Percobaan Dengan 2 Cluster

Cluster 0	Cluster 1	Performance Vektor (Davies Bouldin Index)
334	106	0.243

Tabel 2 Percobaan Dengan 3 Cluster

Cluster 0	Cluster 1	Cluster 2	Performance Vektor (Davies Bouldin Index)
43	87	310	0.256

Tabel 3 Percobaan Dengan 3 Cluster

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Performance Vektor (Davies Bouldin Index)
60	52	302	26	0.275

Pada Tabel 4 menampilkan nilai *centroid* pada masing-masing cluster dengan atribut yang digunakan adalah jumlah penghasilan dan jumlah anggota keluarga. Pada nilai cluster yang berjumlah 2 cluster dengan cluster 0 dengan nilai *centroid* Jumlah Penghasilan adalah 570794.83889 dan Jumlah Anggota Keluarga adalah 1.15269 dan sedangkan cluster 1 dengan nilai *centroid* Jumlah Penghasilan adalah 2405188.67925 dan Jumlah Anggota Keluarga adalah 3.424528.

Tabel 4 Nilai Centroid Dengan 2 Cluster

Attribute	Cluster 0	Cluster 1
Jumlah Penghasilan	570794.83889	2405188.67925
Jumlah Anggota Keluarga	1.15269	3.424528

Pada Tabel 5 menampilkan nilai *centroid* pada masing-masing cluster dengan atribut yang digunakan adalah jumlah penghasilan dan jumlah anggota keluarga. Pada nilai cluster yang berjumlah 3 cluster dengan cluster 0 dengan nilai *centroid* Jumlah Penghasilan adalah 3290697.67442 dan Jumlah Anggota Keluarga adalah 3.83721, cluster 1 dengan nilai *centroid* Jumlah Penghasilan adalah 1651915.70881 dan Jumlah Anggota Keluarga adalah 2.88506, dan cluster 2 dengan nilai *centroid* Jumlah Penghasilan adalah 517350.99846 dan Jumlah Anggota Keluarga adalah 1.07097.

Tabel 5 Nilai Centroid Dengan 3 Cluster

Attribute	Cluster 0	Cluster 1	Cluster 2
Jumlah Penghasilan	3290697.67442	1651915.70881	517350.99846

Attribute	Cluster 0	Cluster 1	Cluster 2
Jumlah Anggota Keluarga	3.83721	2.88506	1.07097

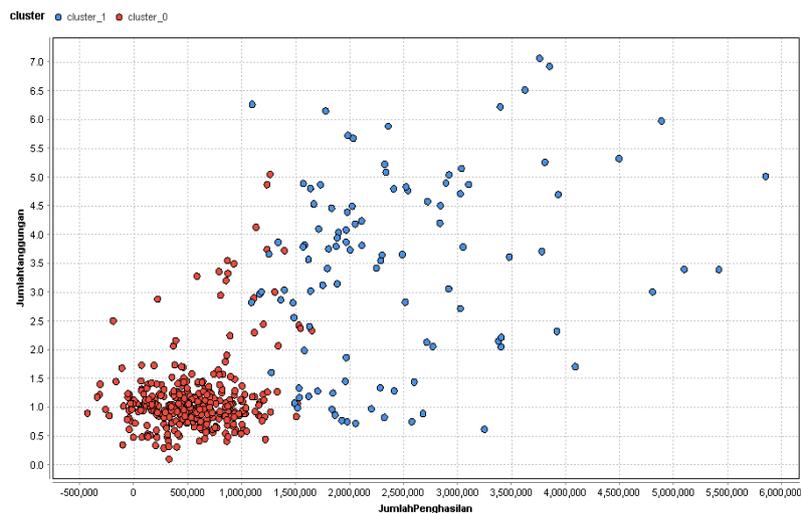
Pada Tabel 6 menampilkan nilai *centroid* pada masing-masing *cluster* dengan atribut yang digunakan adalah jumlah penghasilan dan jumlah anggota keluarga. Pada nilai *cluster* yang berjumlah 4 *cluster* dengan *cluster* 0 dengan nilai *centroid* Jumlah Penghasilan adalah 1369999.99999 dan Jumlah Anggota Keluarga adalah 2.43333, *cluster* 1 dengan nilai *centroid* Jumlah Penghasilan adalah 2183012.82051 dan Jumlah Anggota Keluarga adalah 3.44231, *cluster* 2 dengan nilai *centroid* Jumlah Penghasilan adalah 504565.59445 dan Jumlah Anggota Keluarga adalah 1.04305, dan *cluster* 3 dengan nilai *centroid* Jumlah Penghasilan adalah 3750000.0 dan Jumlah Anggota Keluarga adalah 4.15385.

Tabel 6 Nilai *Centroid* Dengan 4 *Cluster*

Attribute	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Jumlah Penghasilan	1369999.99999	2183012.82051	504565.59445	3750000.0
Jumlah Anggota Keluarga	2.43333	3.44231	1.04305	4.15385

3.2 Analisis Cluster

Keanggotaan cluster dikelompokkan berdasarkan kemiripannya percobaan yang digunakan hanya dengan nilai 2 *cluster*. Dari pengamatan hasil clustering diperoleh bahwa data penduduk yang menempati *cluster* 0 memiliki jumlah penghasilan rendah dan jumlah anggota keluarga sedikit sedangkan yang menempati *cluster* 1 memiliki jumlah penghasilan tinggi dan jumlah anggota keluarga sedikit. Dapat disimpulkan bahwa *cluster* 0 bisa disebut dengan *cluster* “tidak mampu” dan *cluster* 1 bisa disebut dengan *cluster* “mampu”. Komposisi keanggotaan *cluster* berdasarkan jumlah penghasilan dan jumlah anggota keluarga dapat dilihat pada gambar 1.



Gambar 1 Hasil *Clustering* Menggunakan *Chart style Scatter*

4. SIMPULAN

Dari hasil penelitian dapat diambil kesimpulan bahwa :

- Penerapan Metode *K-Means* untuk pemetaan masyarakat berdasarkan tingkat kemiskinan kelurahan Kemuning berhasil diterapkan ke dalam aplikasi sehingga penerima bantuan JAMKESDA tepat sasaran.
- Masih ada faktor-faktor selain jumlah penghasilan dan jumlah anggota keluarga yang dapat menyebabkan seseorang termasuk golongan tidak mampu atau mampu, untuk mengetahuinya dapat dilakukan penelitian lebih lanjut lagi.
- Pada proses pengelompokan dengan aplikasi yang telah dibuat, pengelompokan tingkat kemiskinan dengan menggunakan 440 sampel data yang di dapat dari hasil penerima tahun 2015 didapatkan jumlah *cluster* menggunakan Uji Validitas Davies Bouldin Index (DBI) sebanyak 2 *cluster* dengan nilai 0.243 dan nilai tersebut paling optimal digunakan karena mendekati 0 (nol).
- Hasil Pengelompokan terdapat 2 kelompok, Mampu dan Tidak Mampu dengan keanggotaan kelompok tidak mampu sebanyak 334 anggota dan kelompok mampu sebanyak 106 anggota.

DAFTAR PUSTAKA

- [1] Kelurahan Kemuning, “**Data Penerima Jamkesda**” Banjarbaru, 2016.
- [2] C. A. Nugroho, R. A. Hendrawan dan I. Hafidz, “**Clustering Kelompok Swadaya Masyarakat (KSM) dalam Menentukan Kebijakan Bantuan Badan Pemberdayaan Masyarakat di Kota Surabaya dengan Menggunakan Metode Self-Organizing Map (SOM) dan K-Means**”, *Teknik ITS*, vol. Vol. 1 No. 1, pp. A-368, 2012.
- [3] Z. Aras Z dan S. , “**Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means**”, *Jurnal Manajemen Sistem Informasi*, vol. Vol 1 No. 2, pp. 159-170, 2016.

- [4] E. Prasetyo, "**Data Mining Mengolah Data menjadi Informasi dengan Matlab**", Yogyakarta: Andi, 2016.
- [5] W. Nengsih, "**DESCRIPTIVE MODELLING MENGGUNAKAN K-MEANS UNTUK PENGCLUSTERAN TINGKAT KEMISKINAN DI PROPINSI RIAU**", 2016.
- [6] F. Fajrianti, M. N. Bustan dan M. A. Tiro, "**PENGGUNAAN ANALISIS CLUSTER K-MEANS DAN ANALISIS DISKRIMINAN DALAM PENGELOMPOKAN DESA MISKIN DI KABUPATEN PANGKEP**", 2018.
- [7] A. N. Ulfah dan S. Uyun, "**Analisis Kinerja Algoritma Fuzzy C-Means dan K-Means pada Data Kemiskinan**" *Jatisi*, vol. Vol. 1 No. 2, pp. 139 -148, 2015.
- [8] J. Han dan M. Kamber, "**Data Mining Concepts and Techniques, San Fransisco: Morgan Kaufmann**", 2011.
- [9] C. Verrellis, Business Intelligence: "**Data Mining and Optimization for Decision Making**", Italy: Wiley, 2009.
- [10] D. T. Larose, "**DISCOVERING KNOWLEDGE IN DATA**", America: Wiley, 2005.
- [11] A. E. Putri dan M. Budiharto, "**Paham JKN (Jaminan Kesehatan Nasional)**", Jakarta: Friedrich-Ebert-Stiftung, 2014.
- [12] Agusta dan Yudi, K-Means, "**Web Blog For Data Mining and Clustering**", Yogyakarta: Andi, 2001.