

APLIKASI MENENTUKAN TINGKAT SOSIAL EKONOMI PENDUDUK MENGUNAKAN ALGORITMA K-MEANS DI PABERASAN SUMENEP INSTITUT TEKNOLOGI NASIONAL MALANG

Risa Rohima

Teknik Informatika, Institut Teknologi Nasional Malang
risarohima@gmail.com

ABSTRAK

Tingkat sosial ekonomi memiliki kedudukan atau posisi, dimana pembawa status tersebut memegang hak dan kewajiban yang harus dimainkan, seperti hak untuk mendapatkan subsidi pemerintah dan berkewajiban memberikan bantuan dana terhadap yang tidak mampu. Namun, pemerataan penyebaran yang dilakukan oleh daerah desa Paberasan Sumenep dilakukan dengan manual sehingga pemerataan yang dilakukan tidak *fair*. Tujuan dari pembuatan program ini ialah untuk monitoring pemerataan penyebaran subsidi pemerintah (Studi Kasus: Daerah Paberasan Sumenep). Aplikasi ini akan menentukan tingkat sosial ekonomi penduduk di desa Paberasan, berupa tingkat sosial ekonomi miskin atau tidak miskin.

Penentuan tingkat ekonominya disini menggunakan metode *K-Means*, yaitu metode *clustering* menggunakan perulangan-perulangan iterasinya guna membandingkan antara iterasi lama dengan iterasi baru. Algoritma *K-Means* dimulai dengan menentukan *K* atau jumlah *cluster* yang ingin dibentuk, dimana *K* yang diambil adalah 2, yaitu miskin dan tidak miskin. Lalu menetapkan *centroid* acak hingga menghitung jarak setiap *centroid*, lalu mengelompokkan data dengan melihat jarak terkecil, lalu melakukan perulangan perhitungan *centroid* dengan membandingkan hasil iterasi perhitungan lama dengan baru hingga tidak mengalami perubahan pengelompokan. *Visual Studio 2010* ialah *software* yang berguna untuk membuat atau mengembangkan aplikasi berbasis *desktop* maupun *website*. Sedangkan *visual basic* ialah bahasa pemrograman yang sederhana untuk *GUI* sehingga pengguna dapat mengatur tampilan halamannya. *SQL Server* ialah database yang digunakan untuk arsitektur *client-server*.

Dari 20 data yang dihitung menghasilkan 14 data yang sesuai dengan hasil dari Paberasan Sumenep, sehingga tingkat keakuratan sistem bernilai 75%. Nilai ini sewaktu-waktu bisa berubah dengan banyaknya data yang dimasukkan. Hal-hal yang dapat mempengaruhi nilai berubah ialah jumlah data yang dihitung, sifat subyektif data yang dapat dari Paberasan Sumenep, serta jumlah *cluster* yang ditentukan.

Kata kunci : *K-Means, Visual Studio, SQL Server, Visual Basic, Tingkat Sosial Ekonomi, Penduduk.*

1. PENDAHULUAN

Suatu penduduk pada setiap desa memiliki tingkat sosial ekonomi yang bermacam-macam. Tingkat sosial ekonomi adalah keadaan ekonomi suatu penduduk diukur dengan jumlah rupiah atau penghasilan rata-rata per bulan berdasarkan upah minimal rata-rata. Kantor balai desa adalah suatu kantor yang mengurus masalah kependudukan, termasuk pendataan perekonomian suatu penduduknya, serta penentuan pembagian subsidi pemerintah seperti yang telah dijelaskan dalam Peraturan Presiden Republik Indonesia Nomor 166 tahun 2014 tentang Program Percepatan Penanggulangan Kemiskinan. Peraturan Presiden Republik Indonesia Nomor 166 tahun 2014 tentang Program Percepatan Penanggulangan Kemiskinan pada Pasal 1 ayat 1 yang menyatakan bahwa penanggulangan kemiskinan adalah kebijakan dan program pemerintah dan pemerintah daerah yang dilakukan untuk mengurangi jumlah penduduk miskin

dalam rangka meningkatkan derajat kesejahteraan rakyat.

Teknologi komputer yang berkembang sangat pesat dan sering kali dimanfaatkan untuk membantu memenuhi pekerjaan manusia, bahkan komputer sering kali berubah fungsi untuk menggantikan pekerjaan manusia yang tidak memerlukan pemikiran dalam bersifat rutinitas. Ada banyak hal yang dapat diterapkan dalam teknologi komputer tersebut, beberapa metode pun sering kali diterapkan dalam pembuatan suatu sistem yang dapat membantu manusia dalam mengerjakan banyak hal, salah satunya adalah metode *K-Means*. Metode ini sering dipakai untuk menyelesaikan memutuskan suatu masalah dengan mengelompokkan data ke dalam suatu kelompok yang biasa disebut dengan “K” berdasarkan jarak tekecil cluster.

Proses pengadministrasian dan pendistribusian subsidi pemerintah atau bantuan dana yang dilakukan oleh Kantor Balai Desa Paberasan Sumenep selama ini dilakukan dengan manual, yaitu dengan melihat

secara kasat mata, yaitu berdasarkan ingatan para petugas saja, sehingga penyebarannya tidak merata kepada yang berhak menerima. Oleh sebab itu, penulis akan menciptakan suatu aplikasi menentukan tingkat sosial ekonomi penduduk menggunakan algoritma *k-means* di paberasan sumenep berbasis *desktop* yang dapat menjawab permasalahan tersebut.

1.1. Rumusan Masalah

Rumusan masalah dalam perancangan aplikasi ini adalah bagaimana merancang aplikasi yang dapat menentukan tingkat sosial ekonomi suatu penduduk di paberasan sumenep-madura dan menerapkan metode *K-Means* ke dalam aplikasi ini?

1.2. Batasan Masalah

Batasan masalah dalam perancangan aplikasi adalah sebagai berikut.

1. Aplikasi yang digunakan berbasis *desktop*, dibuat oleh *visual studio* 2010 dan *SQL Server* 2008.
2. Metode yang digunakan adalah *K-Means*.
3. Kriteria yang digunakan dalam perhitungan status sosial ekonomi adalah pendapatan, pekerjaan, pendidikan, dan kepemilikan barang.
4. Kriteria kepemilikan barang yang digunakan adalah motor, mobil, rumah, status rumah, tipe rumah, dan luas tanah.
5. Aplikasi yang dibuat menampilkan informasi tingkat sosial ekonomi penduduk di Paberasan.

1.3. Tujuan

Tujuan dari penyusunan proposal ini adalah menerapkan algoritma *k-means* ke dalam aplikasi berbasis *desktop* menggunakan *visual studio* 2010.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Pada penelitian oleh Ramadhani, Rima Dias tentang Data Mining Menggunakan Algoritma *K-Means Clustering* untuk menentukan Strategi Promosi Universitas Dian Nuswantoro, disimpulkan bahwa setelah dilakukan pengelompokan data mahasiswa melalui persebaran wilayah berdasarkan potensi akademik menggunakan *KMeans clustering* terbentuk tiga cluster yaitu, cluster satu dengan jumlah 804 mahasiswa dengan rata-rata IPK 3.16, cluster dua dengan jumlah 2792 mahasiswa dengan rata-rata IPK 3.15 dan cluster tiga dengan jumlah 223 mahasiswa dengan rata-rata IPK 3.2. (Ramadhani, 2012)

Penelitian yang dilakukan oleh Ong, Johan Oscar, tentang Implementasi Algoritma *K-Means*

Clustering untuk menentukan Strategi Marketing President University disimpulkan bahwa strategi promosi yang dapat dilakukan oleh pihak marketing President University, yaitu melakukan promosi dengan mengirim tim marketing yang sesuai dengan jurusan yang paling banyak diminati dan melakukan promosi pada kota-kota di Indonesia yang didasarkan pada tingkat kemampuan akademik dari calon mahasiswa. (Ong, 2013)

Penelitian yang dilakukan oleh Rismawan, Tedy untuk mengelompokkan mahasiswa berdasarkan nilai *body mass index (bmi)* & ukuran kerangka menggunakan algoritma *K-Means* berhasil. Dari data yang dilatih, diperoleh 3 kelompok berdasarkan *BMI* dan ukuran kerangka, yaitu *BMI* normal dan kerangka besar, dengan pusat cluster (19,53; 11,52), *BMI* obesitas sedang dan kerangka sedang, dengan pusat cluster (25,44; 10,22), dan *BMI* obesitas berat dan kerangka kecil, dengan pusat cluster (43,25; 8,95). (Rismawan, 2008)

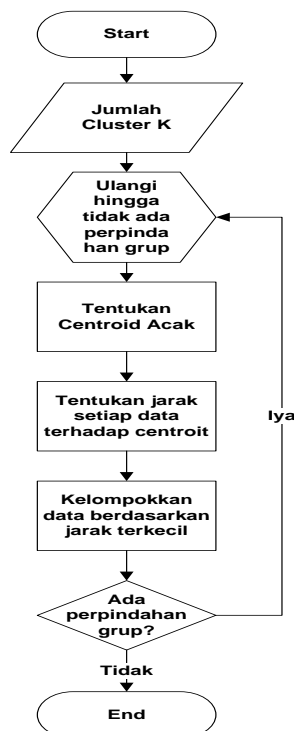
2.2. Sistem Pendukung Keputusan

Sistem Pendukung Keputusan merupakan sistem informasi interaktif yang menyediakan informasi, pemodelan, dan pemanipulasian data. yang membantu pengambilan keputusan dalam situasi semiterstruktur dan tidak terstruktur, di mana tak seorang pun tahu secara pasti bagaimana keputusan seharusnya dibuat. Aplikasi Sistem Pendukung Keputusan menggunakan *CBIS (Computer Based Information Systems)* yang fleksibel, interaktif, dan dapat diadaptasi yang telah dikembangkan untuk mendukung solusi atau masalah manajemen spesifik yang tidak terstruktur. (Prayoko, 2013)

2.3. Algoritma K-Means

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* dimulai dengan pemilihan secara acak *K*, *K* disini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai *cluster*. Hitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus Euclidian hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Pengelompokan data berdasarkan jarak terdekatnya dengan *centroid*. Ulangi langkah tersebut hingga pengelompokan berdasarkan nilai terdekat dengan *centroid* tidak berubah (stabil). (Rismawan, 2008)

Berikut *flowchart* metode *K-Means*. Ditunjukkan pada Gambar 1.



Gambar 1 Flowchart K-Means

Algoritma *Flowchart Metode K-Means*, yaitu sebagai berikut.

1. Proses perhitungan dimulai.
2. Menentukan jumlah K dalam *clustering*.
3. Perulangan perhitungan selama masih ada perubahan grup.
4. Menentukan *Centroid* secara acak, dimana pengacakan yang dimaksud adalah berada diantara data pada setiap variabel terkecil dengan variabel terbesar yang tidak terlampaui jauh dari data yang lain.
5. Menghitung jarak setiap data terhadap *centroid*.
6. Mengelompokkan data berdasarkan jarak terkecil

7. Jika ada perubahan grup dengan grup sebelumnya, maka kembali ke langkah nomor tiga. Jika tidak, maka proses perhitungan selesai.

Berikut rumus untuk perhitungan *K-Means* ditunjukkan pada Persamaan 2.1 (Ramadhani, 2012).

$$\sqrt{(x-y)^2} = |x-y| \dots\dots\dots (2-1)$$

$$d(p,q) = \sqrt{(p1-q1)^2 + (p2-q2)^2} \dots\dots\dots (2-2)$$

$$d(p,q) = \sqrt{(p1-q1)^2 + (p2-q2)^2 + (p3-q3)^2} \dots\dots\dots (2-3)$$

Keterangan:

d : data yang akan dikelompokkan

p : kriteria pada setiap data

q : pusat cluster acak

Untuk menghitung algoritma *k-means*, maka kriteria setiap data dikurangkan dengan pusat cluster acak yang telah dihitung sebelumnya, yaitu nilai tengah dari jumlah kriteria setiap data, proses ini dijumlahkan hingga pada nilai “ p ” dan “ q ” setiap data. Nilai yang dihasilkan oleh “ d ” adalah hasil jarak.

2.4. Sosial Ekonomi

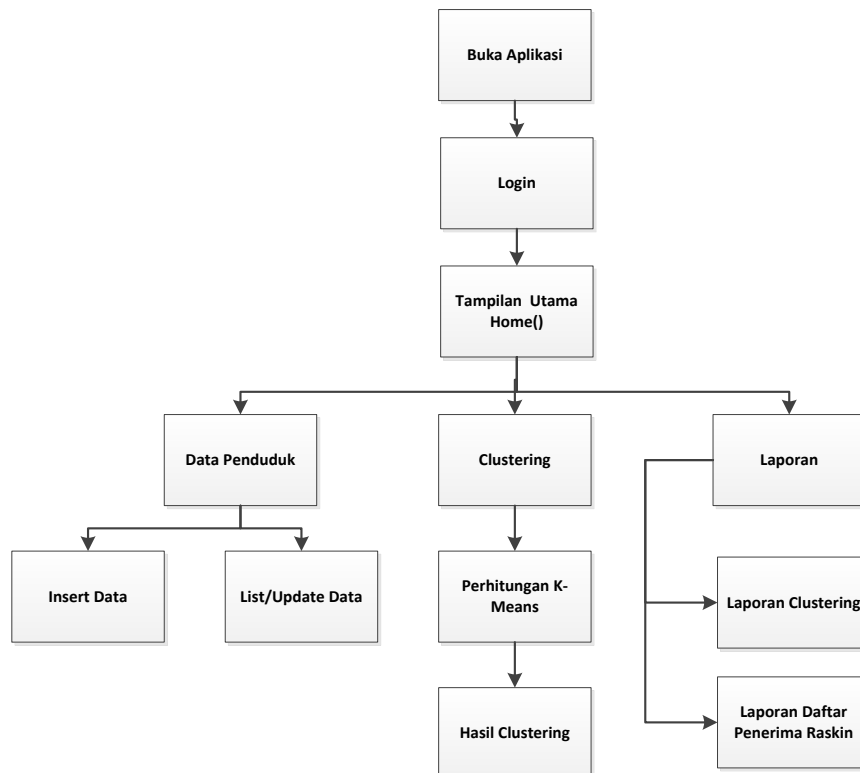
Kondisi sosial ekonomi adalah kedudukan masyarakat atau penempatan posisi masyarakat, dimana pembawa status memegang hak dan kewajiban yang harus dimainkan. (Juariyah, 2010)

Kondisi sosial ekonomi dalam hal ini diposisikan karena beberapa sebab, yaitu pendidikan, pendapatan, pekerjaan, kepemilikan barang, termasuk status rumah, status fisik rumah, dan luas tanah.

3. METODE PENELITIAN

3.1. Diagram Blok Sistem

Diagram blok sistem ditunjukkan pada Gambar



Gambar 2 Diagram Blok Sistem

Penjelasan Blok Diagram Sistem ialah sebagai berikut.

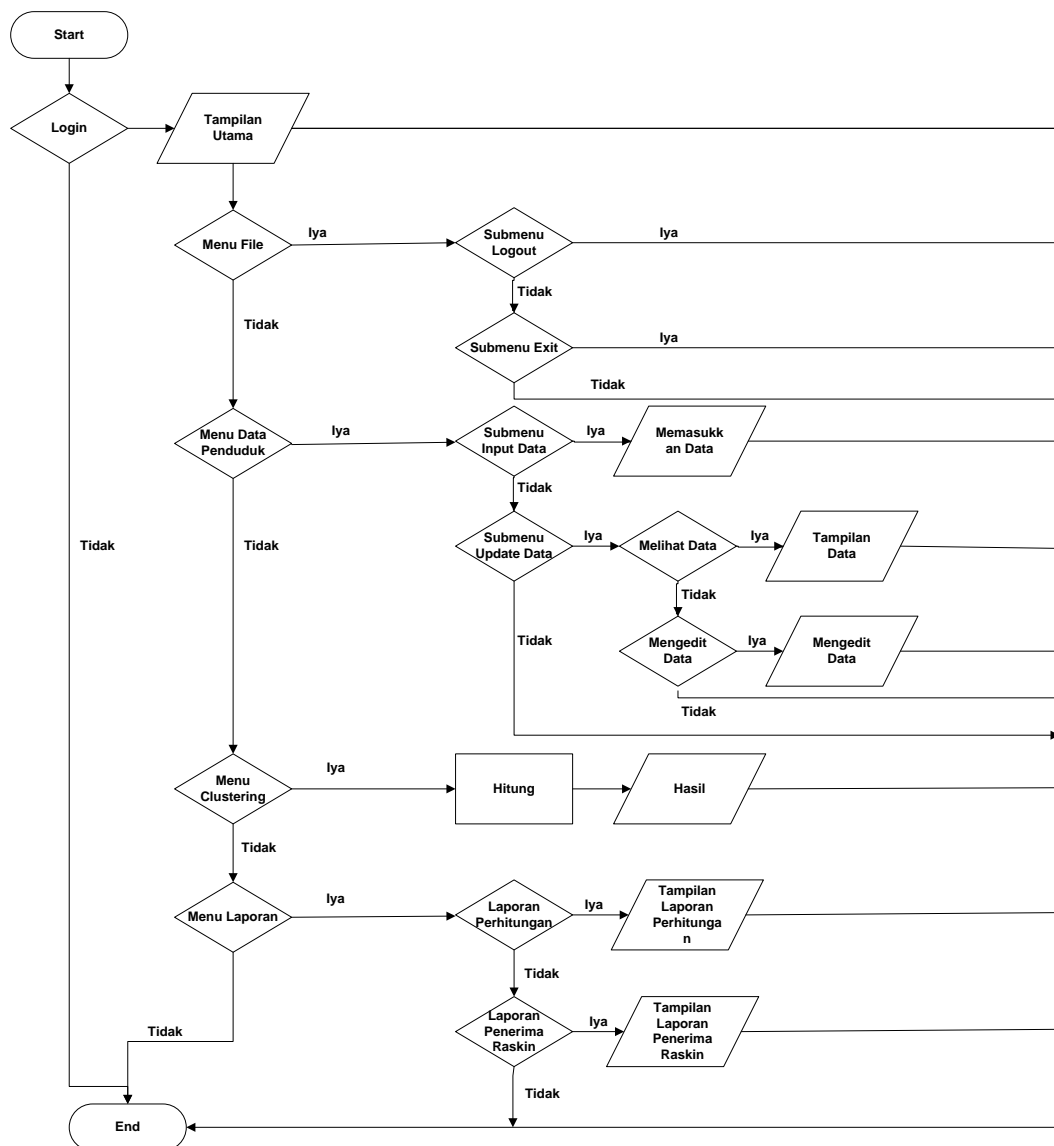
1. *User* membuka aplikasi Penentuan Tingkat Sosial Ekonomi Penduduk Desa Paberasan.
2. *User* memasuki tampilan utama atau Home.
3. *User login* berdasarkan *username* dan *password* yang telah didaftarkan sebelumnya.
4. *User* memilih menu Data Penduduk.
5. *User* memilih submenu Insert Data untuk menambahkan data baru.
6. *User* memilih submenu List Data untuk melihat atau update data lama.
7. Untuk menu *Clustering*, *user* dapat mengelompokkan tingkat sosial ekonomi penduduk Paberasan.

8. Sistem menghitung pengelompokan menggunakan metode *K-Means*.

9. *User* dapat melihat hasil pengelompokan tingkat sosial ekonomi Paberasan Sumenep di menu Laporan.
10. *User* dapat melihat dan mencetak hasil *clustering* penduduk.
11. *User* dapat melihat dan mencetak hasil daftar penerima raskin.

3.2. Flowchart Sistem

Flowchart Sistem ditunjukkan pada Gambar 3.



Gambar 3 Flowchart Sistem

Berikut Penjelasan flowchart sistem dari gambar 3.

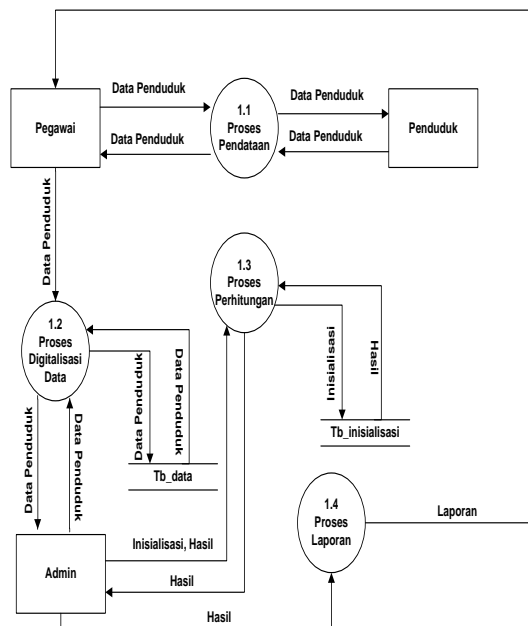
1. Program dimulai.
2. Admin atau pengguna harus login terlebih dahulu untuk mendapat hak akses.
3. Muncul tampilan utama yang berupa pilihan beberapa menu.
4. Jika memilih menu *File*, maka akan diminta untuk memilih beberapa pilihan submenu, yaitu *Logout untuk keluar dari hak akses* dan *Exit untuk keluar dari program*.
5. Jika memilih menu *Data Penduduk*, maka akan diminta memilih submenu *Input Data* untuk menginputkan data baru atau *List Data* untuk melihat data atau mengubah data.
6. Jika memilih menu *Clustering*, maka program akan menampilkan proses perhitungan *k-means*.
7. Jika memilih menu *laporan*, maka pengguna akan mendapat laporan data penduduk.

3.3. Data Flow Diagram Level 1

Data Flow Diagram adalah suatu grafik yang menjelaskan sebuah sistem dengan menggunakan bentuk-bentuk atau simbol untuk menggambarkan aliran data dari proses-proses yang saling berhubungan

Data Flow Diagram Level 1 pada aplikasi ini adalah penjelasan lebih lanjut dari *Data Flow Diagram* Level 0, yaitu dari satu proses menjadi empat proses, yaitu proses pendataan, proses digitalisasi data, proses perhitungan, dan proses laporan. Proses pendataan adalah entitas pegawai

mendata data-data yang diperlukan dalam perhitungan status sosial ekonomi dari penduduk. Proses digitalisasi data adalah proses memberikan data penduduk dari pegawai ke admin untuk disimpan ke dalam basis data. Proses Perhitungan adalah proses normalisasi data nyata yang awalnya non-numerik menjadi numerik agar dapat dihitung status sosial ekonominya menggunakan algoritma *k-means*, proses laporan ialah proses yang akan dicetak menjadi laporan. *Data Flow Diagram Level 1* ditunjukkan pada Gambar 4.

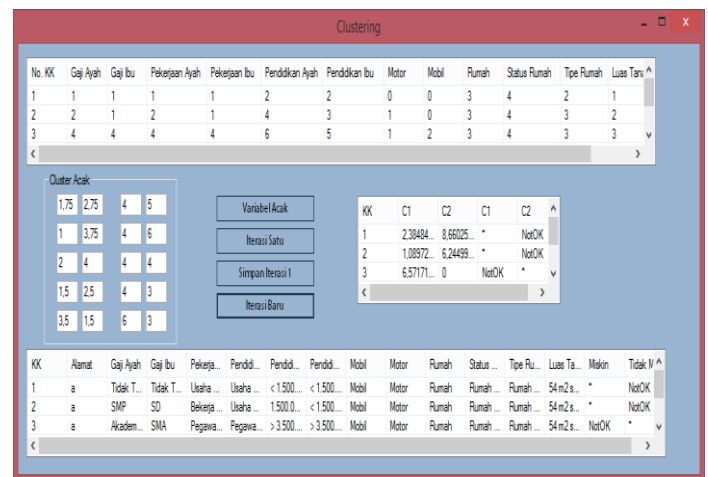


Gambar 4 Data Flow Diagram Level 1

4. HASIL DAN PEMBAHASAN

4.1. Halaman Clustering

Halaman Clustering adalah halaman untuk perhitungan algoritma *k-means*. Pertama kali adalah menentukan cluster acak di iterasi pertama, setelah baru menghitung hasil perhitungan di iterasi satu, karena masih iterasi satu, maka hasil iterasi satu terlebih dahulu disimpan, setelah itu baru menghitung iterasi selanjutnya, dimana hasil iterasi lama dibandingkan dengan hasil iterasi baru, jika telah mendapat nilai atau posisi cluster yang sama diantara keduanya, maka proses perhitungan selesai Seperti ditunjukkan pada Gambar 5.



Gambar 5 Halaman Clustering

4.2. Pengujian Keakuratan Sistem

Pengujian ini adalah pengujian antara kasus lama yang didapat dari kenyataan dengan yang dikerjakan oleh sistem. Berikut ini tabel 4.1.

Tabel 4.1: Tabel Keakuratan Sistem

No	No. KK	Kepala Keluarga	Istri	Miskin	Tidak Miskin	Hasil
1	3529010102090001	Syamsi	HANAFIYAH	*	NotOK	Sesuai
2	3529010104130006	YAYUK LIDI MALIK	-	*	NotOK	Sesuai
3	3529010202090031	SUHARTA	SALAMI	*	NotOK	Sesuai
4	3529010204090012	SANIN	SAADA	*	NotOK	Sesuai
5	3529010204090020	ASMUNI	SUMARSIH	NotOK	*	Sesuai
6	3529010205061265	KOMARITO	MISNATUN	NotOK	*	Sesuai
7	3529010205061640	WELLY ADIYANTO	ENNI WIDIANA	NotOK	*	Tidak Sesuai
8	3529010205062613	BUDI MULYONO	KHAIRUN NISA'	*	NotOK	Tidak Sesuai
9	3529010202080029	SAHAWI	SUTI'	*	NotOK	Sesuai
10	3529010301110015	SINAL	MARYAM	NotOK	*	Sesuai
.
16	3529010305069736	SUHAIRIYANAH	-	NotOK	*	Sesuai
17	3529010305069743	MOHAMMAD IMAM	SUSANTIN	*	NotOK	Sesuai
18	3529010501090016	GIYANTO	SITI AISYAH	*	NotOK	Tidak Sesuai
19	3529010502090038	DUL PANI	SUMARIYA	*	NotOK	Sesuai
20	352901050209005	IBRAHIM	UMRANI	*	NotOK	Sesuai

Hasil dari pengujian keakuratan sistem yang dilakukan pada 20 data adalah sebagai berikut :

Untuk hasil sesuai : $15/20 \times 100 = 75\%$.

Untuk hasil tidak sesuai: $5/20 \times 100 = 25\%$.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Kesimpulan yang didapat ialah dari 20 data yang diuji menghasilkan 15 data yang sesuai dengan hasil dari Paberasan Sumenep, sehingga tingkat keakuratan sistem bernilai 75%. Nilai ini sewaktu-waktu bisa berubah dengan banyaknya data yang dimasukkan. Hal-hal yang dapat mempengaruhi nilai berubah adalah jumlah data yang dihitung, sifat subyektif data yang dapat dari Paberasan Sumenep, serta jumlah cluster yang ditentukan.

5.2. Saran

Saran yang dapat disampaikan adalah sebagai berikut.

1. Meningkatkan aplikasi menggunakan metode KNN untuk penduduk baru yang masuk ke daerah Paberasan Sumenep agar aplikasi tidak mengharuskan menghitung ulang semua data penduduk seperti awal agar memperingan kerja server.
2. Aplikasi ini nantinya dapat dikembangkan agar bisa digunakan untuk review investor yang ingin membangun usahanya di daerah tersebut, bantuan dana yang dilakukan secara individu atau kelompok yang mencari penduduk miskin, dan dapat digunakan untuk meminta dana kepada penduduk tidak miskin.
3. Program ini bisa dibuat menggunakan bahasa pemrograman php sehingga bisa diakses oleh banyak pengguna.

DAFTAR PUSTAKA

- [1] Agusta, Yudi. 2007. *K-Means–Penerapan, Permasalahan dan Metode Terkait*. Jurnal Sistem dan Informatika. 3(1). 47-60.
- [2] Juariyah, S. 2010. *Analisis Kondisi Sosial Ekonomi dan Tingkat Pendidikan Masyarakat Desa Srigading, Kecamatan Labuhan Maringgai, Kabupaten Lampung Timur*. Jurnal Ekonomi dan Pendidikan 7(1)
- [3] Ong, Johan Oscar. 2013. *Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University*.
- [4] Ramadhani, Rima Dias. 2012. *Data Mining Menggunakan Algoritma K-Means Clustering untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro*. Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro.
- [5] Rismawan, Tedy, and Sri Kusumadewi. 2008. *Aplikasi K-Means Untuk Pengelompokan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka*. Seminar Nasional Aplikasi Teknologi Informasi. No. 1907-5022.
- [6] Safik, S., 2013. *Aplikasi Inventori Pada Cv. Arcoma Basco Dengan Menggunakan Visual Basic. net Dan Sql Server 2005*.
- [7] Widya, W., 2013. *Sistem Informasi Manajemen Rumah Sakit Bersalin Ananda Palembang*.
- [8] Prayoko, M.R., 2013. *Sistem Pendukung Keputusan Penentuan Jurusan Pada Sekolah Menengah Atas (SMA) Setia Budi Abadi Perbaungan Dengan Menggunakan Metode Simple Additive Weighting (SAW)*. Jurnal Ilmiah Peneliti Informatika Budi Darma Vol. V, 4(2).