

PENERAPAN METODE CLUSTERING PADA ANALISIS REALISASI PENDAPATAN ASLI DAERAH DENGAN ALGORITMA K-MEANS

Farell Apuilino Iman Seno Aji, Sentot Achmadi, FX. Ariwibisono

Program Studi Teknik Informatika S1, Fakultas Teknologi Industri
Institut Teknologi Nasional Malang, Jalan Raya Karanglo km 2 Malang, Indonesia
smithywerbenjaggermanjenssen@gmail.com

ABSTRAK

Retribusi Daerah adalah pungutan daerah sebagai pembayaran atas jasa atau pemberian izin tertentu yang khusus disediakan dan atau diberikan oleh Pemerintah Daerah untuk kepentingan orang pribadi atau badan menurut UU No. 28 tahun 2009. Pemasukan Pajak Daerah dan Retribusi Daerah tersebut lalu dikumpulkan oleh para Pejabat Pajak untuk digolongkan serta dianalisis. Dalam proses tersebut, data retribusi pendapatan asli daerah yang masuk akan diinputkan ke dalam Microsoft Excel, sehingga saat dilakukan penganalisisan serta pengelompokan secara manual menggunakan Microsoft Excel, menjadi tidak efektif karena data yang diolah sangat banyak sehingga menyebabkan tampilan sedikit membingungkan dan rawan terhadap kesalahan input pada proses perhitungan. Selain memakan waktu untuk melakukan perhitungan, terjadinya kesalahan perhitungan dapat mengakibatkan terjadinya kesalahan analisis oleh para Pegawai Pajak. Pada penelitian ini, akan dibuat sebuah sistem berbasis web dengan menggunakan metode *K-Means*, yang dapat membantu Pegawai Pajak dalam menganalisis dan juga pengelompokan hasil pemasukan seluruh retribusi PAD. Sehingga sebanyak apapun data yang diinputkan, hasil perhitungan diharapkan akan konsisten sesuai dengan metode *K-Means* dan kesalahan perhitungan dapat diminimalisir karena perhitungan telah dilakukan secara otomatis. Dalam proses analisis dan penggolongan data retribusi pendapatan asli daerah, perhitungan dilakukan berdasarkan kriteria yang telah ditentukan oleh instansi terkait, Kriteria - kriteria tersebut nantinya akan diproses dengan perhitungan metode *K-Means* berbasis web yang menggunakan framework *Laravel* dan bahasa pemrograman *PHP (Hypertext Preprocessor)* dengan *Database MySQL*. Hasil dari pengujian metode pada 115 data yang telah diuji dengan sistem ini, sebanyak 37 data masuk kedalam cluster C8 sebagai golongan restoran dan juga 31 data masuk kedalam cluster C7 sebagai golongan cafe, yang artinya pemasukan pajak restoran dan cafe memiliki pengaruh yang cukup besar dalam retribusi pendapatan asli daerah di Kabupaten Gresik. Dari hasil perhitungan antara manual menggunakan Ms. Excel dibandingkan perhitungan di sistem yang telah dibuat memiliki rata – rata error sebesar 0.32% dengan tingkat keakuratan sebesar 99.68%. Berdasarkan hasil pengujian fungsional menunjukkan jika sistem yang telah dibuat dapat berjalan dengan baik dan sesuai dengan fungsi yang tersedia. Sistem yang telah dibuat diuji coba dengan menggunakan 3 aplikasi browser yaitu Microsoft Edge 91.0.864.54 (Official build) (64-bit), Google Chrome 91.0.4472.114 (Official Build) (64-bit) dan Mozilla Firefox 89.0.1 (64-bit).

Kata Kunci : *K-Means, Pendapatan Asli Daerah, Laravel, Php, MySQL*

1. PENDAHULUAN

1.1 Latar Belakang

Retribusi Daerah adalah pungutan daerah sebagai pembayaran atas jasa atau pemberian izin tertentu yang khusus disediakan dan atau diberikan oleh Pemerintah Daerah untuk kepentingan orang pribadi atau badan menurut UU No. 28 tahun 2009. PAD Retribusi Daerah merupakan pemasukan dari suatu kabupaten atau kota, dimana dalam jenis objek pajak tersebut tidak hanya meliputi 1 bidang saja, melainkan berbagai perijinan yang memerlukan fasilitas umum, dan sebagainya.

Kota/kabupaten yang memiliki pemasukan objek pajak dalam jumlah besar, tentu akan sangat menguntungkan kota/kabupaten tersebut. Pemasukan Pajak Daerah dan Retribusi Daerah yang tinggi membuat kestabilan ekonomi di daerah tersebut menjadi sangat baik. Bila melihat dari seberapa besar pembangunan, perijinan, dan juga proyek – proyek

yang berjalan di wilayah kabupaten yang dikehendaki, maka dalam hal ini tentunya besaran pajak yang dikeluarkan calon objek pajak juga berbeda tergantung dengan jenis objek apakah yang akan atau telah mereka dirikan di kawasan milik daerah tersebut. Karena dalam melakukan pembayaran pajak, tiap jenis pajak memiliki besaran biaya yang berbeda beda.

Pemasukan Pajak Daerah dan Retribusi Daerah tersebut lalu dikumpulkan oleh para Pejabat Pajak untuk digolongkan serta dianalisis. Dalam proses tersebut, data pemasukan pajak dan retribusi daerah yang masuk akan diinputkan ke dalam Microsoft Excel, sehingga saat dilakukan penganalisisan serta pengelompokan secara manual menggunakan Microsoft excel, menjadi tidak efektif karena data yang diolah sangat banyak sehingga menyebabkan tampilan sedikit membingungkan dan rawan terhadap kesalahan input pada proses perhitungan. Selain

memakan waktu untuk melakukan perhitungan, terjadinya kesalahan perhitungan dapat mengakibatkan terjadinya kesalahan analisis oleh para Pegawai Pajak.

Pada penelitian ini, akan dibuat sebuah sistem berbasis web dengan menggunakan metode K-Means, yang dapat membantu Pegawai Pajak dalam menganalisis dan juga pengelompokan hasil pemasukan seluruh retribusi PAD. Sehingga sebanyak apapun data yang diinputkan, hasil perhitungan diharapkan akan konsisten sesuai dengan metode K-Means dan kesalahan perhitungan dapat diminimalisir karena perhitungan telah dilakukan secara otomatis. Dalam proses analisis dan penggolongan data retribusi pendapatan asli daerah, perhitungan dilakukan berdasarkan kriteria yang telah ditentukan oleh instansi terkait, Kriteria - kriteria tersebut nantinya akan diproses dengan perhitungan metode K-Means berbasis web yang menggunakan framework Laravel dan bahasa pemrograman PHP (Hypertext Preprocessor) dengan Database MySQL.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian tentang PAD telah dilakukan oleh beberapa peneliti dan menuliskan hasilnya dalam bentuk jurnal, antara lain Rudi Prasetyo, 2017, yang membahas tentang pengaruh pendapatan pajak daerah dan retribusi daerah terhadap peningkatan PAD. Dalam penelitiannya, Rudi Prasetyo menganalisis data pajak daerah dan retribusi daerah dengan metode Kolmogorov-Smirnov yang digunakan untuk uji normalitas data, selanjutnya dengan Uji Asumsi Klasik, Uji Multikolinearitas, Uji Autokorelasi, dan Uji Heteroskedastisitas.

Selanjutnya, Rahayu Mayang Sari, 2015, dalam artikelnya yang berjudul "Prediksi Data Anggaran Pendapatan Belanja Daerah Menggunakan Algoritma K-Means", membahas tentang pengelompokan Anggaran Pendapatan Belanja Daerah di tahun yang akan datang, dengan mengumpulkan setiap jenis pemasukan pajak daerah dan juga jenis anggaran belanja daerah, lalu melakukan analisis pemasukan pajak daerah dengan metode K-Means Clustering. Penelitian ini dilakukan dengan menggunakan software Tanagra.

Lalu, Reynaldo Ferdian Saputra, 2020, yang membahas tentang penggunaan metode clustering k-means untuk menentukan hasil dari tes gaya belajar siswa. Dalam penelitiannya, penempatan gaya belajar dibagi menjadi 3 cluster yaitu auditori, visual, dan kinestetik. Penggunaan metode K-Means Clustering digunakan pada saat user test memilih jawaban, maka akan ditentukan mengarah ke kategori apakah setiap jawaban yang diinputkan oleh user pada saat menginputkan jawaban.

Sedangkan, Hijrah Aziz, 2016, dengan judul ANALISIS DAN PERANCANGAN SISTEM INFORMASI PENGELOLAAN PAJAK DAERAH BERBASIS WEB yang menjelaskan tentang

rancangan pembuatan website pengelolaan pajak daerah. Dalam penelitiannya, Aziz membuat sistem untuk melakukan entry data pendaftaran bagi calon pembayar wajib pajak daerah dan juga hanya untuk menganalisis data pemasukan objek pajak oleh Pegawai Pajak. Penelitian tersebut diharapkan dapat membantu tiap calon wajib pajak dalam melakukan pendaftaran formulir pembayaran PAD secara online dan juga membantu pegawai pajak dalam mengelola basis data objek pajak.

2.2. Dasar Teori

2.2.1 Data Clustering

(Irwansyah, 2015) Clustering adalah sebuah proses untuk mengelompokkan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum.

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma clustering. Oleh karena itu, clustering sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. Clustering banyak digunakan dalam berbagai aplikasi seperti misalnya pada business intelligence, pengenalan pola Citra, web search, bidang ilmu biologi, dan untuk keamanan (security).

2.2.2 Algoritma K-Means - Clustering

(MNM Ediyanto, 2013) K-Means Cluster Analysis merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain. Tujuan pengelompokan adalah untuk meminimalkan objective function yang di set dalam proses clustering, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu cluster dan memaksimalkan variasi antar cluster.

(Agusta, 2007) Proses penentuan centroid dan penempatan objek dalam cluster diulangi sampai nilai centroid konvergen (centroid dari semua cluster tidak berubah lagi). Secara umum metode K-Means Cluster Analysis menggunakan algoritma sebagai berikut :

1. Tentukan k sebagai jumlah cluster yang di bentuk.
Untuk menentukan banyaknya cluster k dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak cluster.

- Bangkitkan k Centroid (titik pusat cluster) awal secara random. Penentuan centroid awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k cluster, kemudian untuk menghitung centroid cluster ke-i berikutnya, digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^n x^2}{n} ; i=1,2,3,\dots,n \quad (1)$$

Dimana

v : centroid pada cluster

objek ke-i: banyak objek/jumlah objek yang menjadi anggota cluster

- Hitung jarak setiap objek ke masing-masing centroid dari masing-masing cluster. Untuk menghitung jarak antara objek dengan centroid penulis menggunakan Euclidian Distance.

$$D(a, b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2} ; i=1,2,3,\dots,n \quad (2)$$

Dimana :

D adalah jarak

a adalah atribut

n adalah jumlah dimensi (atribut)

a_k dan b_k adalah atribut ke-k dari objek data a dan b

- Alokasikan masing-masing objek ke dalam centroid yang paling terdekat. Untuk melakukan pengalokasian objek kedalam masing-masing cluster pada saat iterasi secara umum dapat dilakukan dengan dua cara yaitu dengan hard k-means, dimana secara tegas setiap objek dinyatakan sebagai anggota cluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat cluster tersebut, cara lain dapat dilakukan dengan fuzzy C-Means
- Lakukan iterasi, kemudian tentukan posisi centroid baru dengan menggunakan persamaan (1)
- Ulangi langkah 3 jika posisi centroid baru tidak sama. Pengecekan konvergensi dilakukan dengan membandingkan matriks group assignment pada iterasi sebelumnya dengan matriks group assignment pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma k-means cluster analysis sudah konvergen, tetapi jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya.

2.2.3 Data Mining

(Turban, dkk.2005). Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data Mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat pengetahuan yang terkait dari berbagai database besar (Turban, dkk.2005).

(Larose, 2005), Data Mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dengan sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik

pengenalan pola seperti teknik stastik dan matematika.

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2005) :

- Deskripsi**
Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola dan kecenderungan.
- Estimasi**
Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik daripada kearah kategori. Model dibangun dengan menggunakan record lengkap yang menyediakan nilai dari variabel prediksi.
- Prediksi**
Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang.

Beberapa metode dari teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

- Klasifikasi**
Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, pendapatan rendah.
- Pengklusteran**
Pengklusteran merupakan pengelompokan record, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan record-record dalm kluster lain.
- Asosiasi**
Tugas asosiasi dalam Data Mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. (Kusrini, Emha Taufiq Luthfi, Algoritma Data Mining, 2009).

3. METODE PENELITIAN

3.1. Analisa Masalah

Analisis masalah merupakan usaha untuk memperjelas pokok-pokok permasalahan sehingga lebih spesifik lagi dari masalah yang diteliti, pada bagian ini penulis menyimpulkan permasalahan yang justru jarang dipikirkan, namun sangat riskan sekali bila luput dari penglihatan para pejabat daerah. Hal tersebut membuat Para pejabat Daerah setempat maupaun ASN pajak kadang kala tidak mengetahui tidak mengetahui dari sector manakah pemasukan paling besar dalam pembayaran pajak asli daerah kategori retribusi.

BPPKAD dan BAPPEDA tingkat kabupaten saat ini tentunya memiliki beraneka ragam asset daerah dan juga

berbagai perijinan penggunaan aset milik daerah itu sendiri, yang telah banyak dikelola, terutama oleh pihak swasta, dengan membayar pajak yang telah ditetapkan oleh pemerintah daerah tersebut. Banyaknya objek pajak yang masuk ke database tim BPPKAD pun membuat data pemasukan pajak asli daerah menjadi cukup sulit dianalisis, terutama menyangkut prospek masa depan sebagai pemasukan terbanyak dalam segi bidang apakah daerah tersebut yang dapat menghasilkan nilai realisasi pajak tertinggi. Dalam hal ini data mining clustering mempunyai peranan dalam mengolah dan mengelompokkan data pada objek pajak tersebut.

Data Objek Pajak yang digunakan untuk mencari potensi terbaik di masa depan, manakah sector pajak yang paling banyak masuk ke dalam keuangan daerah, nantinya juga dapat dijadikan pertimbangan oleh pejabat daerah tersebut sebagai rencana kedepan bagaimana memaksimalkan lagi pemasukan pajak asli daerah. Hal tersebut juga diperlukan algoritma K-Means agar hasil akhirnya lebih akurat dalam memprediksi pendapatan di masa depan.

3.2. Analisis Kebutuhan

Dalam pembuatan sistem ini, terdapat beberapa hal kebutuhan sebagai analisis data, meliputi seluruh data objek Pajak Asli Daerah tahun 2020 yang meliputi semua jenisnya, meliputi pajak perhotelan, rumah makan, parkir, perijinan dan lain – lain. Apabila data tersebut sudah terpenuhi maka analisis kebutuhan data sudah terpenuhi dan proses pengelompokan data bisa dilakukan berdasarkan cluster.

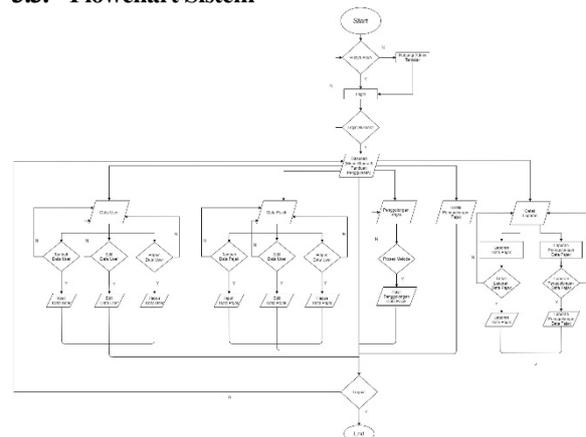
Adapun kebutuhan fungsional yang dibutuhkan dalam system inventaris objek pajak adalah sebagai berikut:

1. Sistem dapat melakukan input data objek pajak.
2. Sistem dapat melakukan input data admin.
3. Sistem dapat melakukan pengelompokan dan mengidentifikasi data objek pajak.

Sedangkan kebutuhan non fungsional yang diperlukan adalah sebagai berikut:

1. Sistem dapat dijalankan dengan baik pada desktop menggunakan web browser seperti Microsoft Edge, Google Chrome dan Mozilla Firefox.
2. Sistem harus dapat memastikan bahwa data yang digunakan dalam sistem harus terlindung dari akses yang tidak berwenang dan hanya dapat diakses oleh internal Pejabat Pajak
3. Sistem memiliki tampilan (antar muka) yang mudah dipahami oleh Pejabat Pajak.

3.3. Flowchart Sistem

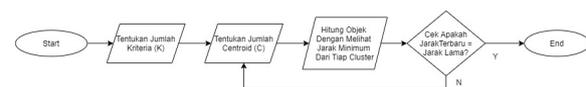


Gambar 1. Tampilan Flowchart Alur Sistem

Untuk menggunakan system ini pertama admin harus login terlebih dahulu kemudian sistem mendeteksi apakah pengguna terdaftar sebagai admin atau bukan, jika pengguna adalah orang lain maka akan diminta login lagi dengan benar, sedangkan apabila pengguna adalah admin (Pejabat Pajak) maka admin dapat menginputkan data objek pajak, dan data akan diotomatis di simpan di database, kemudian dalam proses pengolahan data objek pajak akan di ambil dari database dan akan di teruskan kedalam proses perhitungan menggunakan metode Clustering K-Means untuk mendapatkan hasil akhir yang diinginkan oleh Pejabat Pajak

3.4. Flowchart Metode Clustering K-Means

Adapun Metode yang akan digunakan pada pengerjaan program ini adalah sebagai berikut :

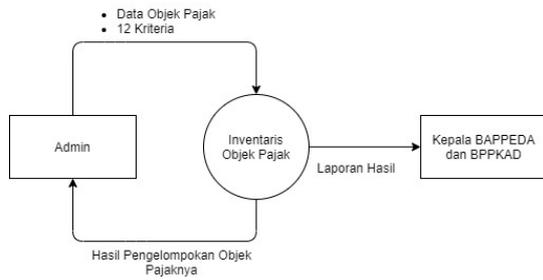


Gambar 2. Tampilan Flowchart Alur Sistem

Pada gambar diatas, saat data telah masuk ke database, system akan menentukan jumlah cluster yang akan digunakan untuk clustering, lalu menentukan pusat clusternya. Selanjutnya, menghitung tiap jarak antar data. Hasil dari proses perhitungan data awal hingga akhir, akan membentuk sebuah keluaran hasil cluster, yang mana data yang telah dihitung tadi dikelompokkan berdasarkan jarak terkecil. Untuk mengecek apakah proses clustering telah selesai, perlu melihat kesamaan hasil cluster yang baru dengan hasil literasi cluster sebelumnya, jika iya maka proses berakhir dan menghasilkan pengelompokan data yang diinginkan

3.5. Diagram DFD

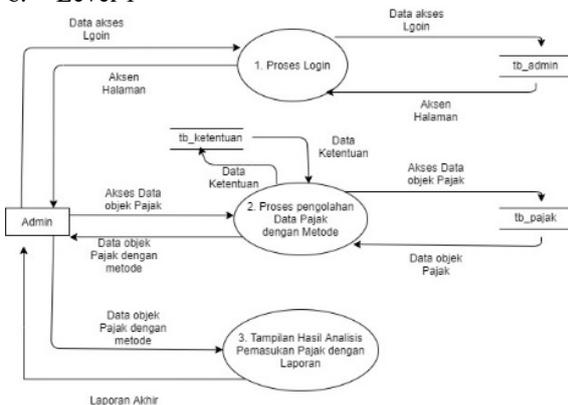
a. Level 0



Gambar 3. Tampilan DFD Level 0 Alur Sistem

Menjelaskan alur diagram level 0, dimana admin dapat menginputkan data objek pajak yang kemudian di proses oleh sistem menggunakan perhitungan metode. Setelah proses selesai hasil perhitungan berupa data objek pajak beserta golongan akan ditampilkan di data objek pajak

b. Level 1

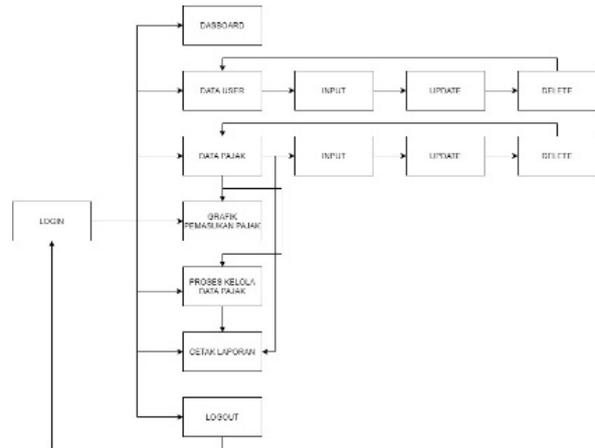


Gambar 4. Tampilan DFD Level 1 Alur Sistem

Gambar diatas menjelaskan alur diagram level 1, untuk memudahkan dalam membaca alur dari sistem yang akan dibuat. Pada sistem, admin sebagai hak akses dapat login ke halaman admin. Admin juga dapat menginputkan, mengubah, dan menghapus data pajak lalu akan masuk ke tabel pajak. Setelah admin menginputkan data objek pajak, data akan di proses menggunakan perhitungan metode, lalu akan menghasilkan output data pajak beserta jenisnya. Lalu data pajak yang telah diproses dengan metode akan di cetak ke laporan.

3.6. Struktur Menu

Struktur menu yang akan dirancang di dalam sistem ini adalah sebagai berikut



Gambar 5. Tampilan Struktur Menu

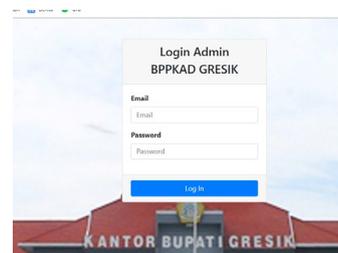
Pada gambar diatas, user akan login terlebih dahulu. Setelah Login, maka akan dihadapkan beberapa menu seperti pada gambar diatas. Diawali dari dashboard yang isinya merupakan ucapan selamat datang kepada admin. Lalu ada menu data user yang isinya nanti merupakan user yang dapat mengakses system ini. Nantinya yang bisa menambahkan user baru hanyalah admin yang telah terdaftar di system ini. Selanjutnya ada menu Data Pajak yang berisikan semua data Pendapatan Asli Daerah, yang nantinya bisa dilakukan olah data di menu tersebut. Pada menu berikutnya yaitu Grafik Pemasukan Pajak, yaitu untuk melihat seberapa besar pemasukan dari tiap bulannya, dan juga tahun. Dan pada menu Proses Kelola Data Pajak, yaitu proses clustering k-means terhadap data - data yang telah masuk ke dalam database system. Lalu ada menu cetak laporan yang nantinya dapat mencetak data pemasukan pajak, mulai dari data mentah, hingga data setelah di proses clustering k-means.

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Sistem

1. Halaman Login

Pada bagian ini, admin / pegawai bppkad harus login terlebih dahulu agar dapat mengakses sistem seperti gambar dibawah ini.



Gambar 6. Tampilan Halaman Login

2. Halaman Penggolongan Pajak

Pada halaman ini, akan ditampilkan golongan data pajak yang di proses k-means. Bila ingin melakukan proses k-means ulang, user dapat

menekan tombol proses Kmeans berwarna biru seperti pada gambar dibawah ini

No Pajak	Cluster	Keterangan
3525020003716	C6	Golongan Pajak Rumah Kos / Villa
35250200016814	C4	Golongan Pajak Perhotelan / Apartemen
35250200019214	C7	Golongan Pajak Homestay / Losmen
35250200026114	C3	Golongan Pajak Hiburan Lain-Lain

Gambar 7. Tampilan Halaman Penggolongan Pajak

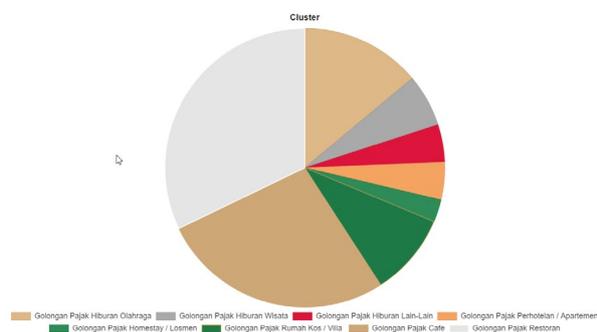
Tampilan Proses K-Means dengan perhitungan lengkapnya dapat dilihat pada gambar dibawah ini.

Luas lahan	Daya tampung	Jumlah pembangkit	Kapasitas pemakaian	Sumber daya	Jumlah kamar	Jumlah meja	Jumlah sarana ibadah	Jumlah lantai	Kebutuhan keamanan tambahan	Potensi kecelakaan	Kebutuhan tenaga medis darurat
c1 11	1.1	1	13.2	1	0	0	5	1	2	2	2
c2 30	15	2	82.5	2	0	0	20	1	4	4	5
c3 0.5	0.7	0	7.7	1	0	0	10	2	1	2	1
c4 15	0.5	2	66	2	100	0	25	10	5	2	2

Gambar 8. Tampilan Halaman Perhitungan K-Means

3. Halaman Grafik Pemasukan Pajak

Pada halaman ini, akan ditunjukkan grafik hasil penggolongan data pajak yang telah dihitung dengan metode K-Means.



Gambar 9. Tampilan Halaman Grafik

4.2. Pengujian Sistem

1. Pengujian Fungsional

Pada Pengujian ini, dilakukan pengujian fitur-fitur yang ada pada system, serta penerapan metode clustering dengan algoritma k-means pada pengelompokan data pemasukan pajak asli daerah. Hasil pengujian fungsional sistem ditunjukkan dalam Tabel 1.

Tabel 1. Tabel Pengujian Fungsional

No	Fungsi Yang Diuji	B	G
1	Login	✓	-
2	Halaman Dashboard	✓	-
3	Halaman Data User	✓	-
	a. Tambah Data User	✓	-
	b. Ubah Data User	✓	-
	c. Hapus Data User	✓	-
4	Halaman Data Pajak	✓	-
	a. Tambah Data Pajak	✓	-
	b. Ubah Data Pajak	✓	-
	c. Hapus Data Pajak	✓	-
5	Halaman Penggolongan Pajak	✓	-

	a. Proses K-Means	✓	-
	b. Hasil	✓	-
6	Halaman Grafik Pemasukan Pajak	✓	-
7	Cetak Laporan	✓	-
	a. Laporan Data Pajak	✓	-
	b. Laporan Data Pajak yang sudah di proses K-Means	✓	-
8	Logout	✓	-

Keterangan :

- B : Berhasil
- G : Gagal

Dengan hasil pengujian fungsional diatas, didapatkan hasil bahwa semua halaman, button dan laporan pada aplikasi penerapan metode clustering dengan algoritma k-means pada pengelompokan data pemasukan pajak asli daerah yang dibuat telah berhasil dan berjalan dengan baik.

2. Pengujian Browser

Tabel 2. Tabel Pengujian Browser

No	Fungsi Yang Diuji	Mozilla Firefox	Microsoft Edge	Google Chrome
1	Login	✓	✓	✓
2	Halaman Dashboard	✓	✓	✓
3	Halaman Data User	✓	✓	✓
	Tambah Data User	✓	✓	✓
	Ubah Data User	✓	✓	✓
	Hapus Data User	✓	✓	✓
4	Halaman Data Pajak	✓	✓	✓
	Tambah Data Pajak	✓	✓	✓
	Ubah Data Pajak	✓	✓	✓
	Hapus Data Pajak	✓	✓	✓
5	Halaman Penggolongan Pajak	✓	✓	✓
	Proses K-Means	✓	✓	✓
	Hasil	✓	✓	✓
6	Halaman Grafik Pemasukan Pajak	✓	✓	✓
7	Cetak Laporan	✓	✓	✓
	Laporan Data Pajak	✓	✓	✓
	Laporan Data Pajak yang sudah di proses K-Means	✓	✓	✓
8	Logout	✓	✓	✓

Keterangan :

- ✓ : Berhasil
- X : Gagal

3. Pengujian User Admin

Pada Pengujian ini, diperlukan untuk mengetahui tingkat kepuasan pengguna terhadap pengalaman pengguna dan tampilan antarmuka aplikasi. Penggunaan Sistem ini hanya akan digunakan oleh Pihak BAPPEDA dan BPPKAD serta Pemda Setempat. Pengujian dilakukan terhadap 6 pegawai BPPKAD yang bertugas dalam merekap Data Pajak. Masing-masing kuesioner berisi tujuh pertanyaan mengenai pengalaman pengguna ketika mengoperasikan aplikasi.

Berdasarkan hasil dari pengisian kuesioner yang didapatkan penulis dari responden dan dihitung jumlah keseluruhannya mendapatkan persentase hasil seperti berikut, dapat dilihat pada Tabel 3.

Tabel 3. Tabel Pengujian User Admin

No	Pertanyaan	Penilaian		
		Baik (B)	Cukup(C)	Kurang / Perlu Dirubah (K)
1	Apakah ukuran dan warna <i>font</i> tulisan sudah terlihat jelas?	(6) 100%	(0) 0%	(0) 0%
2	Apakah warna tombol, <i>background</i> , dan juga grafik pada sistem sudah sesuai dengan yang diharapkan oleh pihak BPPKAD?	(4) 66.7%	(2) 33.3%	(0) 0%
3	Apakah alur menu pada sistem mudah dipahami?	(5) 83.3%	(1) 16.7%	(0) 0%
4	Apakah <i>form-form</i> yang ada pada sistem sudah sesuai dan mudah digunakan?	(2) 33.3%	(4) 66.7%	(0) 0%
5	Apakah kriteria data yang digunakan sesuai dengan kriteria – kriteria data pajak PAD yang telah ditentukan dari BPPKAD?	(6) 100%	(0) 0%	(0) 0%
6	Apakah sistem ini lebih mempermudah dalam penggolongan Data Pendapatan Asli Daerah untuk dianalisis?	(4) 66.7%	(2) 33.3%	(0) 0%
7	Apakah sistem ini dapat lebih mempermudah Pegawai BPPKAD dalam menyiapkan bahan RAPBD untuk periode berikutnya?	(6) 100%	(0) 0%	(0) 0%

Dari hasil pengujian diatas, dapat disimpulkan jika pengujian pada user dari 6 Responden menghasilkan jumlah jawaban B:33, C:9, K:0. Pemilih B: $33 / 42 \times 100 = 78.57\%$. Pemilihan KS: $9/42 \times 100 = 21,42\%$. Pemilihan TS: $0/42 \times 100 = 0\%$.

4. Pengujian Metode

Pada pengujian ini, perhitungan metode K-Means dilakukan pada studi kasus Kantor BAPPEDA selaku pemilik data penelitian dan BPPKAD sebagai tempat penyimpanan data dalam menentukan data pajak yang akan dijadikan referensi sebagai pembahasan APBD oleh pemda setempat dengan penerapan metode clustering dengan algoritma k-means pada pengelompokan data pemasukan pajak asli daerah:

a. Penentuan Data Proses

Dalam menentukan data yang akan dianalisis dengan metode K-Means maka langkah pertama yang dilakukan adalah menentukan data yang akan di proses clustering. Pada kasus ini akan digunakan sampel data sejumlah 115 data pajak proses. Adapun data pajak yang digunakan dapat dilihat pada Tabel 4. dibawah ini

Tabel 4. Tabel Data Pajak

Data NPWPD Ke -	Tanggal Transaksi	Alamat Pajak	Kecamatan
1	2019-02-07	JL. DR. WAHIDIN SH ICON MALL 2F	Kebomas
2	2019-02-10	JL. KALIMANTAN 192, GKB	Manyar
3	2019-02-18	Jl. Raya Bunder No. 1A Gresik	Duduk Sampeyan
4	2019-01-13	ICON MALL Lt LG 678A, GRESIK	Kebomas
5	2019-02-10	JL. A. YANI	Gresik

Kemudian ada 12 kriteria yang akan dilakukan proses perhitungan selanjutnya. Adapun data kriteria tersebut dapat dilihat pada Tabel 5.

Tabel 5. Tabel Kriteria Tiap Data Pajak

Data ke -	K5	K10	K11	K12
1	Tidak Ada	Rendah	Sangat Rendah	Sangat Rendah
2	Bantuan Swasta	Rendah	Rendah	Sangat Rendah
3	Bantuan Swasta	Rendah	Rendah	Sangat Rendah
4	Tidak Ada	Rendah	Sangat Rendah	Sangat Rendah
5	Bantuan Swasta	Rendah	Rendah	Sangat Rendah

b. Memberikan bobot pada setiap kriteria

Dengan memberi bobot pada setiap kriteria, memiliki tujuan agar data kriteria yang berbentuk non angka dapat dirubah menjadi angka sehingga dapat dilakukan perhitungan dengan mudah. Berikut pemberian bobot pada setiap kriteria dapat dilihat pada Tabel 6.

Tabel 6. Tabel Bobot Konversi Angka Pada Tiap Kriteria

No	Konversi SDL		Konversi Tinggi Rendah Sedang		
	SDL	Nilai	No	Nilai Kata	Nilai
1	Gardu Induk PLN	3	1	Sangat Tinggi	5
2	Bantuan Swasta	2	2	Tinggi	4
3	Tidak Ada	1	3	Sedang	3
			4	Rendah	2
			5	Sangat Rendah	1

c. Menentukan jumlah cluster dan nilai awal centroid

Pada proses ini, perhitungan penentuan golongan data pajak akan digolongkan menjadi 8 golongan, sehingga memerlukan 8 *cluster* dan 8 *centroid* untuk proses perhitungan berikutnya. Didalam Tabel *Centroid* awal ini juga nantinya setiap nilainya akan dapat ditentukan dengan mencari nilai tertinggi, nilai tengah dan nilai terendah.

Tabel 7. Tabel Centroid Awal

	C1	C2	C3	C4	C5	C6	C7	C8
K5	1	2	1	2	1	1	1	2
K10	2	4	1	5	2	2	2	2
K11	2	4	2	2	1	1	1	2
K12	2	5	1	2	1	1	1	1

Keterangan :

Cluster 1 : Hiburan Olahraga

Cluster 5 : Home Stay / Losmen

- Cluster 2 : Hiburan Wisata
- Cluster 3 : H. Lain – Lain
- Cluster 4 : Perhotelan / Apartemen
- Cluster 6 : Rumah Kos / Villa
- Cluster 7 : Kafe
- Cluster 8 : Restoran

d. Menghitung jarak setiap data ke pusat cluster menggunakan perhitungan jarak Euclidean.

Proses ini bertujuan untuk mencari jarak terpendek antara objek ke centroid dengan rumus perhitungan jarak Euclidean yang ada di Bab 2 bagian 2.2.2 tentang Algoritma *K-Means – Clustering*.

Iterasi ke-1 :

Berikut Contoh perhitungan data pertama untuk *centroid* 1 menggunakan data nilai tertinggi, nilai tengah dan nilai terkecil pada tabel sebagai centroid awal dan dihitung menggunakan rumus jarak atau *Euclidean*. Cara ini bisa diterapkan hingga centroid 8 :

Jarak data dengan *Centroid* 1 :

$$\sum_{k=1}^n (a_k - b_k)^2 = (11 - 11)^2 + (10 - 1.1)^2 + (0 - 1)^2 + (13.2 - 13.2)^2 + (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (1 - 5)^2 + (1 - 1)^2 + (2 - 2)^2 + (2 - 2)^2 + (2 - 2)^2$$

= 96,21

$$D(a, b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2} = \sqrt{96,21} = 9.80866963$$

Lakukan hingga *Centroid* 8 untuk penentuan jarak data pertama. Lalu hitung jarak pada setiap baris data seperti perhitungan contoh data Ke-1 diatas sebanyak data pada Tabel 4.4, hasil perhitungan dapat dilihat pada Tabel 8. dibawah ini.

Data Ke-	Jarak Terpendek	Cluster Baru
1	5.102570039	C7
2	12.30902514	C7
3	13.39738963	C7
4	2.591969136	C7
5	13.2277782	C8

e. Mencari nilai centroid baru

Setelah mendapatkan label *cluster* untuk masing-masing data, maka dicari *centroid* baru untuk literasi berikutnya dengan menjumlahkan seluruh data masing-masing *cluster* dibagi jumlah anggota *cluster*.

Hitung Lagi Pusat cluster dengan menghitung rata rata dari nilai iterasi ke 1. Pusat cluster akhir ditunjukkan pada Tabel 9

Tabel 9. Tabel centroid akhir

	C1	C2	C3	C4	C5	C6	C7	C8
K5	1.625	3	1	3	2	1	1	2
K10	2.1875	4	2.6	4	2	1	2	2
K11	2.0625	4.142857143	2	2	1	1	1	2
K12	2.0625	3.714285714	2.4	2	1	1	1	1

Setelah menentukan titik pusat baru dari setiap cluster, lalu hitung lagi dengan titik pusat baru. Bila nilai pusat cluster masih berbeda, hitung lagi hingga

pusat cluster baru sama dengan pusat cluster yang lama, sehingga perhitungan berhenti di titik pusat terakhir.

Pada penelitian ini iterasi berhenti pada iterasi ke 4 dimana titik pusat cluster tetap dan sama dengan iterasi ke 3. Hasilnya ditunjukkan pada Tabel 10

Tabel 10. Tabel Literasi 4 / Literasi Akhir

Data Ke-	Jarak Terpendek	Cluster Baru
1	3.73614089	C7
2	4.45822172	C8
3	2.64368995	C8
4	3.6407631	C7
5	4.52362416	C8

Dengan hasil *Clustering* tersebut di dapatkan bahwa dari data pajak yang telah diuji, sebanyak 37 data pajak masuk kedalam cluster paling dominan dalam perhitungan sebelumnya yaitu masuk dalam cluster C8 sebagai Pajak Restoran dimana sebagian besar pengusaha berkecimpung di dunia kuliner, dan juga Golongan C7 yaitu kafe dengan menghasilkan 31 data.

5. Pengujian Presentase Kemungkinan Error

Tahap berikutnya yaitu pengujian hasil pengelompokan menggunakan Metode *K-Means* menggunakan system yang telah dibuat, dengan perhitungan manual pada Microsoft Excel.

Tabel 11. Pengujian Persentase Kemungkinan Error

Data Ke-	Ms. Excel	Sistem	Presentase Error
1	3.736140891	3.671033045	1.77%
2	4.458221722	4.458221722	5.57%
3	2.643689953	2.643689953	5.71%
4	3.6407631	3.68067234	1.08%
5	4.523624157	4.523624157	3.33%

Dari hasil perbandingan diatas, dapat diketahui bahwa hasil perhitungan antara manual menggunakan Ms. Excel dibandingkan perhitungan di sistem yang telah dibuat memiliki rata – rata error sebesar 0.32% dengan tingkat keakuratan sebesar 99.68%.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari hasil pembuatan sistem dan juga penggolongan data pajak yang telah dilakukan, terdapat kesimpulan yang menyatakan bahwa:

1. Hasil dari pengujian metode pada 115 data yang telah diuji dengan sistem ini, sebanyak 37 data masuk kedalam cluster C8 sebagai golongan restoran dan juga 31 data masuk kedalam cluster C7 sebagai golongan cafe, yang artinya pemasukan pajak restoran dan café memiliki pengaruh yang cukup besar dalam retribusi pendapatan asli daerah di Kabupaten Gresik.
2. Hasil penggolongan data pemasukan pajak yang telah diuji dapat dijadikan bahan pertimbangan evaluasi oleh PEMDA setempat untuk dibahas dalam rapat RAPBD periode berikutnya

3. Dari hasil pengujian fungsional, sistem yang telah dibuat dapat berjalan dengan baik dan sesuai dengan fungsi yang tersedia. Sistem telah diuji coba dengan menggunakan 3 aplikasi browser yaitu Microsoft Edge 91.0.864.54 (Official build) (64-bit), Google Chrome 91.0.4472.114 (Official Build) (64-bit) dan Mozilla Firefox 89.0.1 (64-bit).

5.2. Saran

Agar pengembangan sistem yang telah dibuat menjadi lebih baik kedepannya, maka ada beberapa saran yang bisa diberikan, yaitu:

1. Sistem yang telah dibuat untuk kedepannya agar dapat menggolongkan seluruh jenis pendapatan asli daerah yang ada di BPPKAD KAB. Gresik
2. Sistem yang telah dibuat dapat dikembangkan lagi dengan metode selain k-means seperti K-NN, GUSTAFSON-KESSEL, Fuzzy C-Means, Double Exponential Smothing Brown dan metode lainnya

DAFTAR PUSTAKA

- [1] Agusta, Y. "K-Means-Penerapan, Permasalahan dan Metode Terkait". Denpasar, Bali: *Jurnal Sistem dan Informatika (Februari 2007)* Vol. 3: 47-60; 2007.
- [2] Aziz, Hijrah; Akib, Faisal. "ANALISIS DAN PERANCANGAN SISTEM INFORMASI PENGELOLAAN PAJAK DAERAH BERBASIS WEB (Studi Kasus: Kantor DPPKAD Kota Palopo)" *Jurusan Sistem Informasi, Fakultas Sains dan Teknologi, UIN Alauddin*
- [3] Mayang Sari, Rahayu. "Prediksi Data Anggaran Pendapatan Belanja Daerah Menggunakan Algoritma K-Means" *Jurnal Smik Smik Riau* Vol.1 No.2 (2015).
- [4] Prasetyo, Rudi. "ANALISIS PENGARUH PENERIMAAN PAJAK DAERAH DAN RETRIBUSI DAERAH TERHADAP PENINGKATAN PENDAPATAN ASLI DAERAH" *Jurnal Ilmu dan Riset Akuntansi* Vol.6 No.3 (Maret 2017)
- [5] Saputra, Reynaldo Ferdian. "IMPLEMENTASI METODE K-MEANS CLUSTERING PADA TES PSIKOLOGI UNTUK MENENTUKAN KELOMPOK BELAJAR SISWA BERBASIS MOBILE". *JATI (Jurnal Mahasiswa Teknik Informatika) ITN Malang* Vol.5 No.1 (2020)
- [6] Yuli Firmansah, Ridho. "ANALISIS RFM (RECENCY, FREQUENCY AND MONETARY) PRODUK MENGGUNAKAN METODE K-MEANS". *JATI (Jurnal Mahasiswa Teknik Informatika)* Vol. 5 No. 1, Maret (2021)
- [7] Irwansyah, Edi. Faisal, Muhammad. "Advanced Clustering: Teori dan Aplikasi". 1 Jan 2015; Deepublish - books.google.com
- [8] MNM Ediyanto, N Satyahadewi. "PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTER ANALYSIS" *jurnal.untan.ac.id* Vol 2. No.2, hal 133-136 (2013).