

DATA MINING DALAM PENGELOMPOKAN JENIS DAN JUMLAH PEMBAGIAN ZAKAT DENGAN MENGGUNAKAN METODE CLUSTERING K-MEANS (STUDI KASUS: BADAN AMIL ZAKAT KOTA BENGKULU)

Prahasti

Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Dehasen Bengkulu
Jl. Meranti Raya No. 32 Kota Bengkulu. 38228 Telp.(0736) 22027, 265957 Fax.(0736) 341139
prahasti.mona82@gmail.com

ABSTRAK - Penelitian ini menerapkan data mining dengan dalam pengelompokan jenis dan penerima zakat. Penerapannya dilakukan dengan algoritma k-means clustering dimana data yang akan diinputkan dikelompokkan berdasarkan pendidikan dan jenis pekerjaan dalam pembagian zakat. Kemudian dibentuk cluster dengan menggunakan nilai centroid untuk menentukan titik pusat terdekat jarak antar data. Dalam algoritma k-means clustering pengolahan data dihentikan dalam hitungan iterasi data tidak mengalami perubahan (data tetap) dari data yang telah dikelompokkan tersebut. Pengujiannya dilakukan dengan menggunakan perangkat lunak RapidMiner percobaan yang dilakukan dengan metode k-means clustering yang terdiri dari unit masukan, unit pengolahan data dan unit keluarannya, k-means clustering pengelompokan data 1-2-1-1, 1-2-1-2 dan 3-4-3-4. Hasil yang didapat dari pengujian tersebut adalah pengelompokan pembagian zakat dengan setiap cluster tidak sama. Hasil pengujiannya ditampilkan dalam bentuk grafik slatter.

Kata Kunci : Data Mining, K-Means Clustering, Zakat

ABSTRACT - This research applies data mining with in grouping type and receiver of zakat. The application is done by k-means clustering algorithm where the data to be input is grouped based on education and the type of work in the distribution of zakat. Then formed cluster by using centroid value to determine the nearest center point of distance between data. In the k-means clustering algorithm the data processing is stopped in a matter of data iteration has not changed (fixed data) from the data that has been grouped. Testing is done by using RapidMiner software experiments conducted with k-means clustering method consisting of input units, data processing units and output units, k-means clustering grouping data 1-2-1-1, 1-2-1-2 and 3-4-3-4. The results obtained from the test is the grouping of zalcat sharing with each cluster is not the same. Test results are displayed in the form of slatter graphs.

Keywords: Data Mining, K-Means Clustering, Zakat

I. PENDAHULUAN

Setiap lembaga mempunyai system operasional yang setia transaksi kegiatan operasinya selalu dicatat dan didokumentasikan. Setiap transaksi sangat berguna bagi lembaga tersebut untuk segala keperluan.

Data mining merupakan salah satu cabang ilmu komputer yang relative baru. Dimana data mining mengekstraksi informasi atau pola yang

penting atau menarik dari data yang ada di *database* yang besar.

Data mining diterapkan di banyak bidang seperti dikemukakan oleh Suprihatin (2011) menyatakan bahwa teknik *clustering k-means* digunakan untuk menentukan nilai huruf ujian akhir pada Universitas Ahmad Dahlan (UAD). Tutik (2014) mengelompokkan surat-surat dalam Al-Qur'an berdasarkan ayat, jumlah ruku' dan lama membaca surat dengan menggunakan algoritma *clustering k-means*

untuk menemukan hasil yang optimal. Deka (2014) menerangkan bahwa teknik *clustering k-means* digunakan untuk mengelompokkan judul buku sesuai dengan kategorinya sehingga memudahkan pustakawan dalam pengelompokan peletakan buku dan merancang strategi dalam meningkatkan minat baca mahasiswa pada Universitas Islam Indonesia. Andika (2008) mengemukakan bahwa algoritma *k-means clustering* dapat diimplementasikan dengan metode perangkat lunak untuk verifikasi citra sidik jari *pin minutiae*. Tahta (2012) menjelaskan tentang gabungan metode *single linkage clustering* dan *k-means* dalam menyelesaikan problem kerja praktek Jurusan Teknik Industri ITS sehingga memberikan hasil *cluster* yang lebih baik dengan parameter uji *cluster variance* dan metode *silhouette coefisien*.

Penelitian menerapkan data mining untuk mengolah data jenis dan jumlah pembagian zakat pada Badan Amil Zakat Kota Bengkulu. Zakat yang dibayarkan merupakan salah satu tuntunan bagi setiap umat muslim menurut syariat Islam, untuk mengolah jenis dan jumlah pembagian zakat, dibutuhkan sebuah metode yang bisa digunakan untuk menggali informasi tersembunyi dari data tersebut. Metode tersebut dilakukan dengan menggunakan metode data mining.

Penelitian dilakukan untuk menganalisa data mining dengan metode *clustering k-Means* yang kemudian diterjemahkan dalam sebuah perangkat lunak yang digunakan untuk

mengelompokkan jenis zakat dan pembagian zakat berdasarkan data yang dilakukan sehingga didapat kelompok-kelompok jenis dan jumlah pembagian zakat menurut syariat Islam khususnya bagi mereka (pemilik harta) yang telah memenuhi persyaratan syar'i (nishab dan haul) sebagai wajib zakat (muzaki).

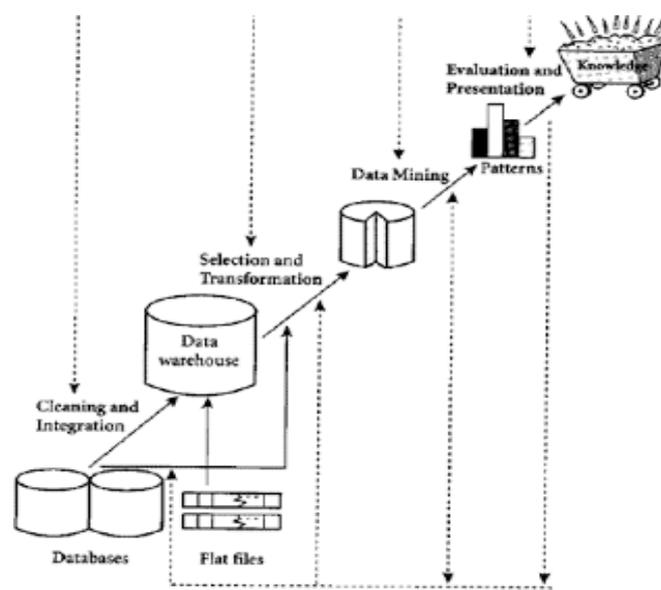
II. LANDASAN TEORI

A. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Deka, 2014)

B. Data Mining

Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database yang besar (Tachir Hendro Pudjiantoro, 2011). Menurut Santosa data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan dimasa depan. Data mining ini juga dikenal dengan istilah *pattern recognition* (Johan Oscar Ong, 2013).



Gambar 1. Tahap-Tahap Data Mining

C. K-Means Clustering

Metode *k-means clustering* merupakan metode pengelompokan *partitioned clustering*. Metode *k-means* mengelompokkan dokumen didasarkan pada jarak terdekat dengan *centroid*. *K-means* merupakan metode pengelompokan yang sederhana dan dapat digunakan dengan mudah. Tetapi pada jenis data tertentu, K-means tidak dapat memberikan segmentasi data dengan baik, sehingga kelompok yang terbentuk tidak murni data yang sama (Rendi Handoyo, 2014).

Pengelompokan data dengan metode KMeans ini secara umum dilakukan dengan algoritma.

1. Tentukan jumlah kelompok.
2. Alokasikan data ke dalam kelompok secara acak.
3. Hitung pusat kelompok (centroid/ratarata) dari data yang ada di masing-masing kelompok.
4. Alokasikan masing-masing data ke centroid/rata-rata terdekat.
5. Kembali ke langkah c, apabila masih ada data yang berpindah kelompok, atau apabila ada perubahan nilai centroid diatas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan

D. ZAKAT

Zakat adalah salah satu rukun Islam yang memiliki dimensi ibadah dan muamalah sekaligus. Menunaikan zakat adalah kewajiban ibadah yang merupakan konsekuensi ketaatan kepada perintah Allah SWT, dan pemanfaatannya dapat membantu menyelesaikan permasalahan ekonomi (muamalah) yang dihadapi oleh masyarakat (Gamsir Bachmid, 2012).

Syarat-syarat wajib zakat yaitu:

1. Islam : Zakat hanya diwajibkan bagi orang Islam saja.
2. Merdeka: hamba sahaya tidak wajib mengeluarkan zakat kecuali zakat fitrah, sedangkan tuannya wajib mengeluarkannya.
3. Milik sepenuhnya: harta yang akan dizakati hendaknya milik sepenuhnya orang yang beragama Islam dan harus merdeka.
4. Cukup Haul: maksudnya harta tersebut dimiliki genap setahun selama 354 hari menurut tanggal hijriah atau 365 hari menurut tanggalan Masehi.
5. Cukup Nisab: nisab adalah nilai minimal sesuatu harta yang wajib dikeluarkan zakatnya.

III. ANALISA DAN PERANCANGAN SISTEM

Tabel 1. Sampel Data Warga Penerima Zakat

No	Nama Warga	Pendidikan Tertinggi	Pekerjaan	Penghasilan/Bulan (Ribuan)
1	Aminudin	SD	Buruh	700
2	Awaludin	SD	Buruh	750
3	Yuhir Ajoyo	SMP	Nelayan	550
4	Iskandar	SD	Buruh	750
5	Yanuar	SMA	Nelayan	650
6	Azman	SMP	Pedagang	900
7	Heri Gunawan	SMA	Pedagang	900
8	Harta Gunawan	SMA	Nelayan	550
9	Apriyandi	SMP	Buruh	740
10	Hermawansyah	SMP	Nelayan	580
11	Sahirman	SMA	Petani	840
12	Ninsarman	Tidak Sekolah	Petani	800
13	Sucipto	SMP	Nelayan	560
14	Mulyanto	SMA	Pedagang	900
15	Mulyadi	SD	Nelayan	560
16	Ansori Saleh	SMP	Buruh	500
17	Zamzami	SMA	Buruh	700
18	Suratno	SMP	Buruh	700

No	Nama Warga	Pendidikan Tertinggi	Pekerjaan	Penghasilan/Bulan (Ribuan)
19	Kustomo	SMP	Buruh	750
20	Rohim Nopriandi	SMA	Petani	850
21	Sugianto	SMA	Pedagang	950
22	Aliwar	SMP	Petani	840
23	Pahrorozi	SMP	Petani	800
24	Supriyanto	Tidak Sekolah	Petani	800
25	Kurniawan	SD	Nelayan	550

A. Tahap Tranformasi Data

Dengan menggunakan metode k-means clustering, dilakukan trnsformasi data yang berjenis data nominal yaitu pekerjaan dan pendidikan. Data berjenis nominal tersebut diinisialisasikan ke dalam bentuk angka melalui beberapa langkah agar dapat diolah dengan menggunakan metode algoritma k-means clustering. Hasil masing-masing inisialisasi sampel data:

Tabel 2. Inisialisasi Data Pendidikan Tertinggi

Pendidikan	Frekuensi	Inisiliasi
SMP	10	1
SMA	8	2
SD	5	3
Tidak Sekolah	2	4

Tabel 3. Inisialisasi Data Pekerjaan

Pekerjaan	Frekuensi	Inisiliasi
Buruh	8	1
Nelayan	7	2
Petani	6	3
Pedagang	4	4

Untuk dapat melakukan pengelompokan data-data tersebut menjadi beberapa cluster dengan langkah:

1. Tentukan jumlah cluster yang diinginkan.
2. Tentukan titik pusat awal cluster (*centroid*), sehingga diperoleh titik pusat dari setiap cluster

Data ke- i	Centroid	Pendidikan Tertinggi	Pekerjaan	Penghasilan/Bulan (Ribuan)
5	Yanuar	2	2	650
7	Heri Gunawan	2	4	900
16	Ansori Saleh	1	1	500

Setelah menentukan centroid awal, maka setiap data akan menemukan centroid terdekatnya dengan menghitung jarak setiap data ke masing-masing centroid sehingga diperoleh hasil penghitungan jarak setiap data untuk masing-masing *cluster*.

Tabel 4. Hasil Penghitungan Jarak Setiap Data Untuk Masing-Masing Cluster (Iterasi 1)

No	Nama Warga	Pendidikan	Pekerjaan	Penghasilan	Jarak Ke centroid			Jarak Terdekat	Cluster
					C0	C1	C2		
1	Aminudin	3	1	700	50.020	200.125	200.010	50.020	C0
2	Awaludin	3	1	750	100.010	150.033	250.008	100.010	C0
3	Yuhir Ajoyo	1	2	550	100.005	350.007	50.010	50.010	C2
4	Iskandar	3	1	750	100.010	150.033	250.008	100.010	C0
5	Yanuar	2	2	650	0	250.008	150.007	0	C0
6	Azman	1	4	900	250.010	1	400.011	1	C1
7	Heri Gunawan	2	4	900	250.008	0	400.012	0	C1
8	Harta Gunawan	2	2	550	100	350.006	50.020	50.020	C2
9	Apriyandi	1	1	740	90.011	160.031	240	90.011	C0
10	Hermawansyah	1	2	580	70.007	320.008	80.006	70.007	C2
11	Sahirman	2	3	840	190.003	60.00	340.007	60.00	C1
12	Ninsarman	4	3	800	150.017	100.025	300.022	100.025	C1

No	Nama Warga	Pendidikan	Pekerjaan	Penghasilan	Jarak Ke centroid			Jarak Terdekat	Cluster
					C0	C1	C2		
13	Sucipto	1	2	560	90.006	340.007	60.008	60.008	C2
14	Mulyanto	2	4	900	250.008	0	400.012	0	C1
15	Mulyadi	3	2	560	90.006	340.007	60.042	60.042	C2
16	Ansori Saleh	1	1	500	150.007	400.012	0	0	C2
17	Zamzami	2	1	700	50.010	200.022	200.002	50.010	C0
18	Suratno	1	1	700	50.020	200.025	200	50.020	C0
19	Kustomo	1	1	750	100.010	150.033	250	100.010	C1
20	Rohim Nopriandi	2	3	850	200.002	50.010	350.007	50.010	C1
21	Sugianto	2	4	950	300.007	50	450.011	50	C1
22	Aliwar	1	3	840	190.005	60.017	340.006	60.017	C1
23	Pahrorozi	1	3	800	150.007	100.010	300.007	100.010	C1
24	Supriyanto	4	3	800	150.017	100.025	300.022	100.025	C1
25	Kurniawan	3	1	550	100	350.007	50.050	50.050	C2

Tabel 5. Hasil Penghitungan Jarak Setiap Data Untuk Masing-Masing Cluster (Iterasi 2)

No	Nama Warga	Pendidikan	Pekerjaan	Penghasilan	Jarak Ke centroid			Jarak Terdekat	Cluster
					C0	C1	C2		
1	Aminudin	3	1	700	2.494	153.019	155.007	2.494	C0
2	Awaludin	3	1	750	47.791	103.028	205.005	47.791	C0
3	Yuhir Ajoyo	1	2	550	152.227	303.004	5.072	5.072	C2
4	Iskandar	3	1	750	47.791	103.028	205.005	47.791	C0
5	Yanuar	2	2	650	52.228	203.004	105	52.228	C0
6	Azman	1	4	900	197.799	47.017	355.008	42.017	C1
7	Heri Gunawan	2	4	900	197.797	47.007	355.007	47.007	C1
8	Harta Gunawan	2	2	550	152.224	303.002	5.006	5.006	C2
9	Apriyandi	1	1	740	37.789	113.026	195.004	37.789	C0
10	Hermawansyah	1	2	580	122.228	273.004	35.010	35.010	C2
11	Sahirman	2	3	840	137.799	13.002	295.002	13.003	C1
12	Ninsarman	4	3	800	97.817	53.038	255.012	58.032	C1
13	Sucipto	1	2	560	142.227	293.004	15.024	15.024	C2
14	Mulyanto	2	4	900	197.797	47.007	355.007	42.007	C1
15	Mulyadi	3	2	560	142.229	293.004	15.046	15.046	C2
16	Ansori Saleh	1	1	500	202.224	353.008	45.015	45.015	C2
17	Zamzami	2	1	700	2.236	153.016	155.002	2.236	C0
18	Suratno	1	1	700	2.404	153.019	155.004	2.404	C0
19	Kustomo	1	1	750	47.787	103.028	205.003	47.787	C1
20	Rohim Nopriandi	2	3	850	147.789	3.007	305.002	3.007	C1
21	Sugianto	2	4	950	247.793	97.003	405.006	97.003	C1
22	Aliwar	1	3	840	137.792	13.040	295.003	13.040	C1
23	Pahrorozi	1	3	800	97.798	53.010	255.004	53.010	C1
24	Supriyanto	4	3	800	97.817	53.038	255.012	53.038	C1
25	Kurniawan	3	1	550	152.228	303.004	5.137	5.137	C2

Tabel 6. Hasil Penghitungan Jarak Setiap Data Untuk Masing-Masing Cluster (Iterasi 1)

No	Nama Warga	Pendidikan	Pekerjaan	Penghasilan	Jarak Ke centroid			Jarak Terdekat	Cluster
					C0	C1	C2		
1	Aminudin	3	1	700	17.529	158.021	150.008	17.529	C0
2	Awaludin	3	1	750	32.516	108.030	200.006	32.516	C0
3	Yuhir Ajoyo	1	2	550	167.505	308.005	0.728	0.728	C2
4	Iskandar	3	1	750	32.516	108.030	200.006	32.516	C0
5	Yanuar	2	2	650	67.506	208.005	100.001	67.506	C0
6	Azman	1	4	900	182.525	42.019	350.007	42.019	C1
7	Heri Gunawan	2	4	900	182.523	42.004	350.007	42.004	C1
8	Harta Gunawan	2	2	550	167.502	308.003	0.319	0.319	C2
9	Apriyandi	1	1	740	22.523	118.030	190.003	22.523	C0
10	Hermawansyah	1	2	580	137.506	278.006	30.009	30.009	C2
11	Sahirman	2	3	840	122.514	18.005	290.002	18.005	C1
12	Ninsarman	4	3	800	82.546	58.032	250.013	58.032	C1
13	Sucipto	1	2	560	157.506	298.005	10.026	10.026	C2
14	Mulyanto	2	4	900	182.523	42.004	350.007	42.004	C1
15	Mulyadi	3	2	560	157.506	298.005	10.083	10.083	C2
16	Ansori Saleh	1	1	500	217.502	358.010	50.012	50.012	C2
17	Zamzami	2	1	700	17.500	158.018	150.003	17.500	C0
18	Suratno	1	1	700	17.529	158.022	150.004	17.529	C0
19	Kustomo	1	1	750	32.526	108.032	200.003	32.516	C1
20	Rohim Nopriandi	2	3	850	132.513	8.011	300.002	8.011	C1
21	Sugianto	2	4	950	232.518	92.002	400.006	92.002	C1
22	Aliwar	1	3	840	122.528	18.038	290.003	18.038	C1
23	Pahrorozi	1	3	800	82.527	58.012	100.001	58.012	C1
24	Supriyanto	4	3	800	82.546	58.032	250.013	58.032	C1
25	Kurniawan	3	1	550	167.505	308.004	1.294	1.294	C2

- Proses perhitungan tela dilakukan dan pada iterasi ke-2 dan ke -3 ternyata posisi cluster tidak mengalami perubahan disamping itu, tidak adanya data lagi yang berpindah dari satu cluster ke cluster yang lain. Untuk itu proses iterasi dihentikan dan hasil dari perhitungan diperoleh sebanyak 3 cluster dengan 3 iterasi.
- Disamping itu, berdasarkan data perhitungan iterasi diperoleh hasil pengelompokan data sebagai berikut:

Tabel 7 Hasil Pengelompokan

Kel	Anggota Kelompok	Jumlah Anggota
0	[1,2,4,5,9,17,18,19]	8 anggota
1	[6,7,11,12,14,20,21,22,23,24]	10 anggota
2	[3,8,10,13,15,16,25]	7 anggota

Dari proses clustering atau pengelompokan diatas maka dihasilkan pengelompokan berdasarkan kedekatan jarak antar titik pusat dengan data warga dari setiap atributnya. Adapun hasil dari penelitian seperti terlihat pada table dibawah ini:

Tabel 8. Hasil Penelitian

Cluster	Hasil Penelitian	
	Centroid Akhir	Anggota
Cluster 0	2, 1.125, 717.5	Jumlah anggota = 8 orang. Terdiri dari: 1. Pendidikan SD dan Pekerjaan Buruh sebanyak 3 warga. a. Anggota. b. Aminudin c. Awwaludin d. Iskandar

Cluster	Hasil Penelitian	
	Centroid Akhir	Anggota
		2. Pendidikan SMP dan Pekerjaan Buruh sebanyak 3 warga. Anggota: a. Apiyadi b. Suratno c. Kustomo 3. Pendidikan SMA dan Pekerjaan Buruh sebanyak 1 warga Anggota: a. Zamzami. 4. Pendidikan SMA dan pekerjaan Nelayan sebanyak 1 warga Anggota: a. Yanuar

IV. PENUTUP

Kesimpulan

1. Menentukan cluster dari masing-masing setiap data sangat penting dilakukan sehingga didapat suatu inisiasi data dari nomial kedalam data numeric.
2. Penerapan clusteing k-means mampu menghasilkan pengelompokan jenis dan pembagian zakat menggunakan data riil dan baik digunakan sebagai cara untuk pengelompokan jenis dan penerima zakat.
3. Clustering K-means mampu memberikan hasil yang baik dan dapat digunakan untuk pengolahan data pengelompokan jenis dan pembagian zakat pada pada tahun-tahun berikutnya.
4. Hasil pengujian yang didapat dapat memberikan hasil yang baik sehingga dapat digunakan untuk pengolahan pengelompokan jenis dan pembagian zakat untuk tahun-tahun berikutnya.

Saran

1. Perbaiki prosedur kerja yang berhubungan dengan proses pemasukan data-data transaksi kedalam system basis data perlu mendapat perhatian yang khusus

sehingga ketersediaan data padaa saat diperlukan dapat lebih terjamin.

2. Hasil penelitian ini diharapkan dapat digunakan Badan Amil Zakat Kota Bengkulu, sebagai rujukan atau menganalisa untuk lebih meningkatkan akurasi analisa kelayakan menentukan pengelompokan jenis dan pembagian zakat lebih optimal.
3. Melakukan pengembangan dengan menggunakan metode optimasi yang lain atau degnan seleksi fitur untuk ketepatan penyelesaian atribut.

DAFTAR PUSTAKA

- [1] Goldie Gunadi,et.al (2012). “Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (Fp-Growth)” Ed. Jurnal Telematika Mkom Vol 4 No.1
- [2] Afrisawati (2013). “Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means” Ed. Pelita Informatika Budi Darma, Volume: V No.3.
- [3] Gamsir Bachdim (2011) “Perilaku Muzaaki dalam membayar zakat mal (Studi Fenomenologi Pengaki di Kota Kediri” Ed. Jurnal Aplikasi Manajemen Vol.10..
- [4] Johan Oscar Ong (2003) “Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Presiden University” Ed. Jurnal Ilmiah Nuswantoro. Vol. 12 No.1.
- [5] Ni Ketut Dewi Ari Jayanti (2014) “Analisa Pengelompokan Konstrasi Program Studi Menggunakan K-Means Clustering” Ed. Konferensi Nasional Sistem Informasi.
- [6] Rendi Handoyo, et.al (2014) “Perbandingan metode clustering menggunakan metode Single Linkage dan K-means pada Pengelompokan Dokumen” Ed. Issn Vol. 15.
- [7] Rima Dias Ramadhani (2013) “Data Mining Menggunakan Algoritma K-means Clustering Untuk menentukan Strategi Promosi Univeristas Dian Nuswantoro” Ed. Jurnal.

[8] Tacbir Hendro Pudjiantoro, et.al(2011) “Penerapan Data Mining Untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa Baru” Ed. Konferensi Nasional Sistem dan Informatika.

[9] Yudi Wibisono (2011) “Perbandingan Partition Around Medoids (PAM) dan K-Means Clustering Untuk Tweets” Ed. Jurnal Knsi.