
PERBANDINGAN KINERJA METODE WARD DAN K-MEANS DALAM MENENTUKAN CLUSTER DATA MAHASISWA PEMOHON BEASISWA (STUDI KASUS : STMIK PRINGSEWU)

Fiqih Satria¹ RZ. Abdul Aziz²

¹STMIK Pringsewu ²M.TI Institut Bisnis dan Informatika Darmajaya

¹E-mail: fiqih.satria@gmail.com

²Email: RZ.Aziz@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengetahui langkah-langkah analisis cluster dengan metode metode Ward dan metode K-Means, serta membandingkan hasil analisis kedua metode tersebut untuk mengclusterkan data mahasiswa terkait pengambilan keputusan untuk menentukan mahasiswa yang berhak menerima beasiswa Peningkatan Prestasi Akademik (PPA) dan beasiswa Biaya Bantuan Akademik (BBA) di STMIK PRINGSEWU. Analisis cluster dilakukan menggunakan Aplikasi IBM SPSS Versi 23. Hasil pengclusteran kedua metode tersebut dibandingkan dengan menggunakan rasio simpangan baku dalam kelompok (Sw) dan simpangan baku antar kelompok (Sb). Pada penelitian ini Berdasarkan nilai rasio simpangan baku dalam cluster dan simpangan baku antar cluster menunjukkan bahwa metode Ward dan K-Means memiliki kinerja yang sama ketika digunakan pada clustering beasiswa PPA, kedua metode memiliki nilai rasio (Sw/Sb) yang sama yaitu 0,749959584 %. Dengan demikian maka hasil cluster yang akan digunakan sebagai acuan pada penentuan penerima beasiswa PPA adalah hasil analisis cluster menggunakan metode Ward ataupun K-Means karena keduanya memiliki hasil cluster dan kinerja yang sama. Sedangkan pada clustering beasiswa BBA, metode Ward memiliki kinerja yang lebih baik daripada metode K-Means karena metode Ward memiliki nilai rasio (Sw/Sb) 0.5346668 % lebih kecil dibandingkan dengan nilai rasio (Sw/Sb) metode K-Means yaitu 0,831525302 %. Dengan demikian maka hasil cluster yang dapat digunakan sebagai acuan pada penentuan penerima beasiswa BBA adalah hasil analisis cluster menggunakan metode Ward.

Kata Kunci : Beasiswa, Cluster, K-Means, Simpangan Baku, Ward..

ABSTRACT

This research aims to determine the steps cluster analysis method with Ward method and K-Means method, and compare the results of the analysis of the two methods for clustering student data related decision-making to determine the students are eligible to receive a Peningkatan Prestasi Akademik (PPA) scholarship and Bantuan Biaya Akademik (BBA) scholarship in STMIK Pringsewu. Cluster analysis was performed using IBM SPSS Version 23. Cluster Analysis results of both methods were compared using standard deviation ratio in the group (Sw) and the standard deviation between groups (Sb). Based on this research the ratio of the value of the standard deviation in the cluster and standard deviation between the cluster shows that the method of Ward and K-Means has the same performance when used in clustering scholarship PPA, both methods have a ratio value (Sw/Sb) are the same, 0.749959584 %. Therefore the results of the cluster that will be used as a reference in determining scholarship recipients PPA is the result of cluster analysis using Ward method or K-Means as they both have the results and performance of the same cluster. While on scholarship BBA clustering, Ward method has better performance than the K-Means method because the method of Ward have value of the ratio (Sw / Sb) 0.5346668% smaller than the value of the ratio (Sw/Sb) K-Means method

0.831525302 %. Therefore, the results of the cluster that can be used as a reference in determining BBA scholarship recipients is the result of cluster analysis using Ward method.

Keywords: Cluster, K-Means, Scholarship, Standard Deviation, Ward.

1. PENDAHULUAN

Beasiswa merupakan program kerja yang ada di setiap universitas atau perguruan tinggi. Program beasiswa bertujuan untuk meringankan beban biaya mahasiswa dalam menempuh masa studi kuliah. Pemberian beasiswa kepada mahasiswa dilakukan secara selektif sesuai dengan jenis beasiswa yang diadakan. Sekolah Tinggi Manajemen Informatika dan Komputer Pringsewu (STMIK Pringsewu) adalah salah satu perguruan tinggi di Kabupaten Pringsewu yang memiliki banyak mahasiswa. STMIK Pringsewu menyediakan beberapa program beasiswa untuk meringankan beban biaya pendidikan mahasiswa yang berasal dari keluarga ekonomi menengah dan memberikan apresiasi bagi mahasiswa yang berprestasi, jenis beasiswa yang ada antara lain beasiswa Peningkatan Prestasi Akademik (PPA) dan beasiswa Bantuan Biaya Akademik (BBA).

Clustering adalah suatu teknik analisis yang bertujuan untuk memilah obyek ke dalam beberapa kelompok yang mempunyai sifat berbeda antara kelompok satu dengan yang lain. Pada analisis ini tiap-tiap kelompok bersifat sejenis antar anggota dalam kelompok atau variasi obyek dalam kelompok yang terbentuk sekecil mungkin. Pada bidang pendidikan, clustering dapat diterapkan untuk mengelompokkan data mahasiswa penerima beasiswa untuk mengetahui kelompok mahasiswa yang berhak menerima beasiswa dan kelompok mahasiswa yang tidak berhak menerima beasiswa.

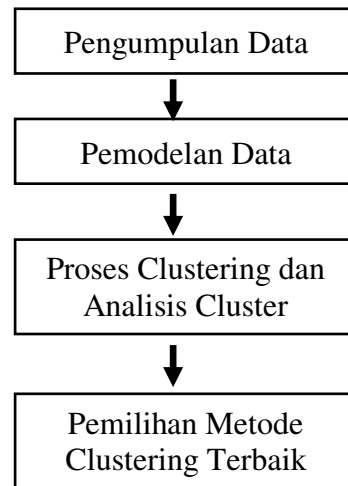
Selama ini proses seleksi penerimaan beasiswa dilakukan secara manual dengan menginputkan data mahasiswa ke dalam file spreadsheet kemudian melakukan sorting data mahasiswa. Dengan cara ini seringkali menimbulkan beberapa permasalahan, antara lain membutuhkan waktu yang lama dan hasil yang kurang efektif karena ketidakjelasan metodologi yang digunakan dalam proses komputasi penerimaan beasiswa, sehingga dibutuhkan suatu metode yang dapat membantu dalam proses pengambilan keputusan mahasiswa yang direkomendasikan menerima beasiswa berdasarkan kriteria-kriteria yang telah ditentukan.

Penelitian ini menggunakan algoritma data mining dalam melakukan clustering (pengelompokan data) menggunakan variabel-variabel yang dapat digunakan sebagai kriteria penentuan seleksi penerimaan beasiswa PPA dan BBA. Metode clustering yang digunakan dalam penelitian ini adalah metode Ward dan metode K-Means. Kemudian setelah melakukan clustering, dilakukan analisis cluster, membandingkan kedua metode tersebut dan memilih metode terbaik yang akan dijadikan acuan dalam pengambilan keputusan penentuan mahasiswa penerima beasiswa. Dari hal tersebut penulis tertarik untuk melakukan penelitian dengan judul Perbandingan Kinerja Metode Ward Dan K-Means Dalam Menentukan Cluster Data Mahasiswa Pemohon Beasiswa dengan studi kasus pada STMIK Pringsewu.

2. METODOLOGI PENELITIAN

Penelitian ini akan dilaksanakan berdasarkan alur penelitian seperti yang ditunjukkan pada Gambar

\



Gambar 1. Alur Penelitian

2.1 Metode Pengumpulan Data

Metode pengumpulan data dapat dilakukan dengan cara antara lain :

a. Studi Pustaka

Studi pustaka dilakukan dengan cara menghimpun informasi yang relevan dengan topik atau masalah yang akan atau sedang diteliti. Informasi diperoleh dari buku-buku ilmiah, laporan penelitian, karangan-karangan ilmiah, tesis dan disertasi, peraturan-peraturan, ketetapan-ketetapan, ensiklopedia, dan sumber-sumber tertulis baik tercetak maupun elektronik.

b. Dokumentasi

Pengambilan data dilakukan melalui dokumen tertulis maupun elektronik dari lembaga/institusi. Dokumen diperlukan untuk mendukung kelengkapan data yang lain.

2.2 Pemodelan Data

Sebelum melakukan proses clustering, dilakukan pemodelan data, karena terdapat perbedaan satuan yang signifikan di antara atribut-atribut yang diteliti.

Tabel 1. Tabel Pemodelan Data Semester

Semester	Kode
1-2	1
3-4	2
5-6	3
7	4

Tabel 2. Tabel Pemodelan Nilai Rata-Rata IP

Kualifikasi	Kode
$IP < (\bar{x} - \sigma)$	1
$(\bar{x} - \sigma) \leq IP < \bar{x}$	2
$\bar{x} \leq IP < (\bar{x} + \sigma)$	3
$IP \geq (\bar{x} + \sigma)$	4

Tabel 3. Tabel Pemodelan Penghasilan Orang Tua

Kualifikasi	Kode
$PO < (\bar{x} - \sigma)$	1
$(\bar{x} - \sigma) \leq PO < \bar{x}$	2
$\bar{x} \leq PO < (\bar{x} + \sigma)$	3
$PO \geq (\bar{x} + \sigma)$	4

Tabel 4. Tabel Pemodelan Jumlah Tanggungan Orang Tua

Jumlah Tanggungan Orang Tua	Kode
1	1
2	2
3	3
≥ 4	4

Tabel 5. Tabel Pemodelan Tagihan Listrik

Kualifikasi	Kode
$TL < (\bar{x} - \sigma)$	1
$(\bar{x} - \sigma) \leq TL < \bar{x}$	2
$\bar{x} \leq TL < (\bar{x} + \sigma)$	3
$TL \geq (\bar{x} + \sigma)$	4

Tabel 6. Tabel Pemodelan PBB

Kualifikasi	Kode
$PBB < (\bar{x} - \sigma)$	1
$(\bar{x} - \sigma) \leq PBB < \bar{x}$	2
$\bar{x} \leq PBB < (\bar{x} + \sigma)$	3
$PBB \geq (\bar{x} + \sigma)$	4

Tabel 7. Tabel Pemodelan Prestasi Akademik/Non-Akademik

Kualifikasi	Kode
Tidak Ada	1
Lomba Tingkat Kecamatan	2
Lomba Tingkat Kabupaten	3
Lomba Tingkat Provinsi/Nasional	4

Tabel 8. Tabel Pemodelan Keanggotaan Dalam Organisasi

Kualifikasi	Kode
Ketua, Organisasi Kampus / UKM	1
Wakil Ketua, Sekretaris, Bendahara Organisasi Kampus / UKM	2
Anggota Organisasi Kampus / UKM	3
Bukan Anggota Organisasi Kampus / UKM	4

3. HASIL DAN PEMBAHASAN

3.1 Data

Sebelum melakukan proses clustering data, terlebih dahulu dilakukan pengumpulan data. Data yang akan diolah dalam pembahasan ini adalah data mahasiswa pemohon beasiswa di STMIK PRINGSEWU. Data mahasiswa yang digunakan adalah 180 data mahasiswa semester 3 sampai dengan semester 7 yang mengajukan permohonan beasiswa tahun 2015. Atribut-atribut data yang digunakan adalah semester, nilai rata-rata IP, rata-rata penghasilan orang tua per bulan, jumlah tanggungan orang tua, nilai tagihan listrik bulan terakhir, nilai Pajak Bumi dan Bangunan (PBB) tahun terakhir, Prestasi Akademik/Non-Akademik dan Keanggotaan dalam Organisasi Kampus. Data diperoleh dengan cara mengumpulkan formulir permohonan beasiswa yang telah diisi oleh mahasiswa pemohon beasiswa.

3. Proses Clustering

Proses clustering pada penelitian ini dilakukan dua kali menggunakan pemodelan data untuk beasiswa PPA dan pemodelan data untuk beasiswa BBA menggunakan dua metode, yaitu metode Ward dan metode K-Means. Clustering data mahasiswa pemohon beasiswa dilakukan menggunakan Software SPSS Versi 23. Cluster yang akan dibentuk yaitu tiga cluster, cluster mahasiswa yang berhak menerima beasiswa, cluster mahasiswa yang dipertimbangkan menerima beasiswa dan cluster mahasiswa yang tidak berhak menerima beasiswa.

4. Perkalian Data Pemodelan dengan Bobot Presentase Variabel

Setelah data pemohon beasiswa dimodelkan, langkah selanjutnya adalah melakukan perkalian bobot presentase variabel dengan data pemohon beasiswa yang telah dimodelkan. Pada bobot presentase variabel beasiswa PPA, IP diberikan bobot tertinggi karena prioritas mahasiswa penerima beasiswa PPA adalah mahasiswa yang memiliki nilai rata-rata IP yang tinggi. Sedangkan pada beasiswa BBA penghasilan orang tua yang diberikan bobot tertinggi.

Tabel 9. Tabel Bobot Variabel Beasiswa PPA

Variabel	Bobot
Semester	10 %
IP	30 %
Jumlah Tanggungan Orang Tua	10 %
Prestasi Akademik/Non-Akademik	10 %
Keanggotaan dalam Organisasi Kampus	10 %
Penghasilan Orang Tua	10 %
Tagihan Listrik	10 %
Pajak Bumi dan Bangunan	10 %

Tabel 10. Tabel Bobot Variabel Beasiswa BBA

Variabel	Bobot
Semester	10 %
IP	10 %
Jumlah Tanggungan Orang Tua	10 %
Prestasi Akademik/Non-Akademik	10 %
Keanggotaan dalam Organisasi Kampus	10 %
Penghasilan Orang Tua	30 %
Tagihan Listrik	10 %
Pajak Bumi dan Bangunan	10

3.2 Analisis Cluster

Analisis cluster dilakukan untuk mendefinisikan cluster ke dalam kategori cluster , yaitu cluster mahasiswa yang berhak menerima beasiswa, cluster mahasiswa yang dipertimbangkan menerima beasiswa dan cluster mahasiswa yang tidak berhak menerima beasiswa. Analisis dilakukan dengan cara melihat nilai rata-rata keseluruhan variabel pada tiap cluster yang terbentuk.

3.2.1 Analisis Cluster Beasiswa PPA Metode Ward

Perincian jumlah cluster dan anggota yang terbentuk pada cluster beasiswa PPA menggunakan metode Ward dapat dilihat pada tabel 11.

Tabel 11. Tabel Keanggotaan Cluster Beasiswa PPA Metode Ward

Nama Cluster	Jumlah Anggota	Rata-Rata Variabel
Cluster 1	83 mahasiswa	0,255
Cluster 2	64 mahasiswa	0,304
Cluster 3	33 mahasiswa	0,334

Dari nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa anggota cluster 3 adalah mahasiswa yang berhak menerima beasiswa, karena memiliki nilai rata-rata keseluruhan data terbesar.

3.2.2 Analisis Cluster Beasiswa PPA Metode K-Means

Perincian jumlah cluster dan anggota yang terbentuk pada cluster beasiswa PPA menggunakan metode K-Means dapat dilihat pada tabel 12.

Tabel 12. Tabel Keanggotaan Cluster Beasiswa PPA Metode K-Means

Nama Cluster	Jumlah Anggota	Rata-Rata Variabel
Cluster 1	33 mahasiswa	0,334
Cluster 2	64 mahasiswa	0,304
Cluster 3	83 mahasiswa	0,255

Dari nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa anggota cluster 1 adalah mahasiswa yang berhak menerima beasiswa, karena memiliki nilai rata-rata keseluruhan data terbesar.

3.2.3 Analisis Cluster Beasiswa BBA Metode Ward

Perincian jumlah cluster dan anggota yang terbentuk pada cluster beasiswa BBA menggunakan metode Ward dapat dilihat pada tabel 13.

Tabel 13. Tabel Keanggotaan Cluster Beasiswa BBA Metode Ward

Nama Cluster	Jumlah Anggota	Rata-Rata Variabel
Cluster 1	40 mahasiswa	0,265
Cluster 2	120 mahasiswa	0,307
Cluster 3	20 mahasiswa	0,216

Dari nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa anggota cluster 2 adalah mahasiswa yang berhak menerima beasiswa, karena memiliki nilai rata-rata keseluruhan data terbesar.

3.2.4 Analisis Cluster Beasiswa BBA Metode K-Means

Perincian jumlah cluster dan anggota yang terbentuk pada cluster beasiswa BBA menggunakan metode K-Means dapat dilihat pada tabel 14.

Tabel 14. Tabel Keanggotaan Cluster Beasiswa BBA Metode K-Means

Nama Cluster	Jumlah Anggota	Rata-Rata Variabel
Cluster 1	96 mahasiswa	0,304
Cluster 2	24 mahasiswa	0,319
Cluster 3	60 mahasiswa	0,249

Dari nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa anggota cluster 2 adalah mahasiswa yang berhak menerima beasiswa, karena memiliki nilai rata-rata keseluruhan data terbesar.

3.3 Menentukan Metode Clustering Terbaik

Untuk mengetahui kinerja kedua metode clustering, digunakan kriteria dua nilai simpangan baku, yaitu rata-rata simpangan baku dalam cluster (S_w) dan simpangan baku antar cluster (S_b). Semakin kecil nilai rasio (S_w/S_b) maka metode tersebut memiliki kinerja yang baik, karena mempunyai tingkat homogenitas yang tinggi.

3.3.1 S_w , S_b dan Rasio Cluster Beasiswa PPA Metode Ward

Simpangan Baku Dalam Cluster (S_w)

Simpangan baku cluster 1 dengan anggota 83 mahasiswa dan rata-rata variabel seluruh mahasiswa dimana nilai $\bar{x}_1 = 0,255422$.

$$S_1 = \sqrt{\frac{(x_1 - \bar{x}_I)^2 + (x_2 - \bar{x}_I)^2 + \dots + (x_{83} - \bar{x}_I)^2}{k-1}}$$

$$S_1 = \sqrt{\frac{(0,2875-0,255422)^2 + (0,2125-0,255422)^2 + \dots + (0,25-0,255422)^2}{83-1}}$$

$$S_1 = 0,032266561$$

Simpangan baku cluster 2, dengan anggota 64 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 3, dimana nilai $\bar{x}_{II} = 0,304296875$.

$$S_2 = \sqrt{\frac{(x_1 - \bar{x}_{II})^2 + (x_2 - \bar{x}_{II})^2 + \dots + (x_{64} - \bar{x}_{II})^2}{k-1}}$$

$$S_2 = \sqrt{\frac{(0,35-0,304296875)^2 + (0,35-0,304296875)^2 + \dots + (0,3-0,304296875)^2}{64-1}}$$

$$S_2 = 0,025802473$$

Simpangan baku cluster 3, dengan anggota 33 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 3, dimana nilai $\bar{x}_{III} = 0,334090909$.

$$S_3 = \sqrt{\frac{(x_1 - \bar{x}_{III})^2 + (x_2 - \bar{x}_{III})^2 + \dots + (x_{33} - \bar{x}_{III})^2}{k-1}}$$

$$S_3 = \sqrt{\frac{(0,3375-0,334090909)^2 + (0,375-0,334090909)^2 + \dots + (0,3125-0,334090909)^2}{33-1}}$$

$$S_3 = 0,031292585$$

Jadi, nilai simpangan baku dalam cluster (S_w) beasiswa PPA dengan menggunakan metode ward adalah:

$$S_w = \frac{S_1 + S_2 + S_3}{3}$$

$$S_w = \frac{0,032266561 + 0,025802473 + 0,031292585}{3}$$

$$S_w = 0,029787206$$

Simpangan Baku Antar Cluster (S_b)

Dapat dilihat pada lampiran 3 bahwa :

$$\bar{x}_I = 0,255422$$

$$\bar{x}_{II} = 0,304296875$$

$$\bar{x}_{III} = 0,334090909$$

$$\bar{X} = \frac{\bar{x}_I + \bar{x}_{II} + \bar{x}_{III}}{3}$$

$$\bar{X} = \frac{0,255422 + 0,304296875 + 0,334090909}{3}$$

$$\bar{X} = 0,29793649$$

$$S_b = \left[(3-1)^{-1} = \sum_{k=1}^3 (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

$$= \left(\frac{(0,255422 - 0,29793649)^2 + (0,304296875 - 0,29793649)^2 + (0,334090909 - 0,29793649)^2}{3 - 1} \right)^{1/2}$$

$$S_b = 0,001577553^{1/2} = 0,039718415$$

Jadi nilai simpangan baku antar cluster (S_b) beasiswa PPA dengan metode ward adalah 0,039718415.

Rasio Antara Simpangan Baku Dalam Cluster dan Simpangan Baku Antar Cluster.

$$\text{Rasio} = \frac{S_w}{S_b} \times 100 \% = \frac{0,029787206}{0,039718415}$$

$$= 0,749959584 \%$$

3.3.2 S_w , S_b dan Rasio Cluster Beasiswa PPA Metode K-Means

Simpangan Baku Dalam Cluster (S_w)

Simpangan baku cluster 1 dengan anggota 33 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 4, dimana nilai $\bar{x}_I = 0,334090909$.

$$S_I = \sqrt{\frac{(x_1 - \bar{x}_I)^2 + (x_2 - \bar{x}_I)^2 + \dots + (x_{33} - \bar{x}_I)^2}{k-1}}$$

$$S_I = \sqrt{\frac{(0,3375 - 0,334090909)^2 + (0,375 - 0,334090909)^2 + \dots + (0,3125 - 0,334090909)^2}{33-1}}$$

$$S_I = 0,031292585$$

Simpangan baku cluster 2 dengan anggota 64 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 4, dimana nilai $\bar{x}_{II} = 0,304296875$.

$$S_2 = \sqrt{\frac{(x_1 - \bar{x}_{II})^2 + (x_2 - \bar{x}_{II})^2 + \dots + (x_{64} - \bar{x}_{II})^2}{k-1}}$$

$$S_2 = \sqrt{\frac{(0,35 - 0,304296875)^2 + (0,35 - 0,304296875)^2 + \dots + (0,3 - 0,304296875)^2}{64-1}}$$

$$S_2 = 0,025802473$$

Simpangan baku cluster 3, dengan anggota 83 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 4, dimana nilai $\bar{x}_{III} = 0,255421687$.

$$S_3 = \sqrt{\frac{(x_1 - \bar{x}_{III})^2 + (x_2 - \bar{x}_{III})^2 + \dots + (x_{83} - \bar{x}_{III})^2}{k-1}}$$

$$S_3 = \sqrt{\frac{(0,2875 - 0,255421687)^2 + (0,2125 - 0,255421687)^2 + \dots + (0,25 - 0,255421687)^2}{83-1}}$$

$$S_3 = 0,032266561$$

Jadi, nilai simpangan baku dalam cluster (S_w) beasiswa PPA dengan menggunakan metode K-means adalah:

$$S_w = \frac{S1+S2+S3}{3}$$

$$S_w = \frac{0,031292585+0,025802473+0,032266561}{3}$$

$$S_w = 0,029787206$$

Simpangan Baku Antar Cluster (Sb)

Dapat dilihat pada lampiran 4 bahwa :

$$\bar{x}_I = 0,334090909$$

$$\bar{x}_{II} = 0,304296875$$

$$\bar{x}_{III} = 0,255421687$$

$$\bar{X} = \frac{\bar{x}_I + \bar{x}_{II} + \bar{x}_{III}}{3}$$

$$\bar{X} = \frac{0,334090909 + 0,304296875 + 0,255421687}{3}$$

$$\bar{X} = 0,29793649$$

$$Sb = \left[(3 - 1)^{-1} = \sum_{k=1}^3 (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

$$\left(\frac{(0,334090909 - 0,29793649)^2 + (0,304296875 - 0,29793649)^2 + (0,255421687 - 0,29793649)^2}{3 - 1} \right)^{1/2}$$

$$Sb = 0,001577553^{1/2} = 0,039718415$$

Jadi nilai simpangan baku antar cluster (Sb) beasiswa PPA dengan metode K-Means adalah 0,039718415.

Rasio Antara Simpangan Baku Dalam Cluster dan Simpangan Baku Antar Cluster.

$$\text{Rasio} = \frac{S_w}{S_b} \times 100 \% = \frac{0,029787206}{0,039718415} \times 100 \%$$

$$= 0,749959584 \%$$

3.3.3 Sw, Sb dan Rasio Cluster Beasiswa BBA Metode Ward

Simpangan Baku Dalam Cluster (Sw)

Simpangan baku cluster 1 dengan anggota 40 mahasiswa dan rata-rata variabel seluruh mahasiswa dapat dilihat pada lampiran 5, dimana nilai $\bar{x}_I = 0,2659375$.

$$S_I = \sqrt{\frac{(x1-\bar{x}_I)^2 + (x2-\bar{x}_I)^2 + \dots + (x56-\bar{x}_I)^2}{k-1}}$$

$$S_I = \sqrt{\frac{(0,2875-0,2659375)^2 + (0,2375-0,2659375)^2 + \dots + (0,25-0,2659375)^2}{40-1}}$$

$$S_I = 0,029686867$$

Simpangan baku cluster 2, dengan anggota 120 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 5, dimana nilai $\bar{x}_{II} = 0,3078125$.

$$S_2 = \sqrt{\frac{(x_1 - \bar{x}_{II})^2 + (x_2 - \bar{x}_{II})^2 + \dots + (x_{120} - \bar{x}_{II})^2}{k-1}}$$

$$S_2 = \sqrt{\frac{(0,35 - 0,3078125)^2 + (0,3125 - 0,3078125)^2 + \dots + (0,35 - 0,3078125)^2}{120-1}}$$

$$S_2 = 0,026581416$$

Simpangan baku cluster 3, dengan anggota 20 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 5, dimana nilai $\bar{x}_{III} = 0,21625$.

$$S_3 = \sqrt{\frac{(x_1 - \bar{x}_{III})^2 + (x_2 - \bar{x}_{III})^2 + \dots + (x_{20} - \bar{x}_{III})^2}{k-1}}$$

$$S_3 = \sqrt{\frac{(0,2 - 0,21625)^2 + (0,2 - 0,21625)^2 + \dots + (0,2375 - 0,21625)^2}{20-1}}$$

$$S_3 = 0,017253909$$

Jadi, nilai simpangan baku dalam cluster (S_w) beasiswa BBA dengan menggunakan metode ward adalah:

$$S_w = \frac{S_1 + S_2 + S_3}{3}$$

$$S_w = \frac{0,029686867 + 0,026581416 + 0,017253909}{3}$$

$$S_w = 0,024507397$$

Simpangan Baku Antar Cluster (S_b)

Dapat dilihat pada lampiran 5 bahwa :

$$\bar{x}_I = 0,2659375$$

$$\bar{x}_{II} = 0,3078125$$

$$\bar{x}_{III} = 0,21625$$

$$\bar{X} = \frac{\bar{x}_I + \bar{x}_{II} + \bar{x}_{III}}{3}$$

$$\bar{X} = \frac{0,2659375 + 0,3078125 + 0,21625}{3}$$

$$\bar{X} = 0,263333333$$

$$S_b = \left[(3-1)^{-1} = \sum_{k=1}^3 (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

$$\left(\frac{(0,2659375 - 0,263333333)^2 + (0,3078125 - 0,263333333)^2 + (0,21625 - 0,263333333)^2}{3-1} \right)^{1/2}$$

$$S_b = 0,002101009^{1/2} = 0,045836766$$

Jadi nilai simpangan baku antar cluster (S_b) beasiswa BBA dengan metode ward adalah 0,045836766.

Rasio Antara Simpangan Baku Dalam Cluster dan Simpangan Baku Antar Cluster.

$$\text{Rasio} = \frac{S_w}{S_b} \times 100 \% = \frac{0,024507397}{0,045836766}$$
$$= 0,5346668 \%$$

3.3.4 S_w , S_b dan Rasio Cluster Beasiswa BBA Metode K-Means

Simpangan Baku Dalam Cluster (S_w)

Simpangan baku cluster 1 dengan anggota 96 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 6, dimana nilai $\bar{x}_I = 0,304817708$

$$S_I = \sqrt{\frac{(x_1 - \bar{x}_I)^2 + (x_2 - \bar{x}_I)^2 + \dots + (x_{96} - \bar{x}_I)^2}{k-1}}$$
$$S_I = \sqrt{\frac{(0,35 - 0,304817708)^2 + (0,3125 - 0,304817708)^2 + \dots + (0,35 - 0,304817708)^2}{96-1}}$$
$$S_I = 0,023811886$$

Simpangan baku cluster 2 dengan anggota 24 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 6, dimana nilai $\bar{x}_{II} = 0,319791667$.

$$S_2 = \sqrt{\frac{(x_1 - \bar{x}_{II})^2 + (x_2 - \bar{x}_{II})^2 + \dots + (x_{24} - \bar{x}_{II})^2}{k-1}}$$
$$S_2 = \sqrt{\frac{(0,375 - 0,319791667)^2 + (0,3375 - 0,319791667)^2 + \dots + (0,3 - 0,319791667)^2}{24-1}}$$
$$S_2 = 0,033564641$$

Simpangan baku cluster 3, dengan anggota 60 mahasiswa dan rata-rata variabel seluruh responden dapat dilihat pada lampiran 6, dimana nilai $\bar{x}_{III} = 0,249375$.

$$S_3 = \sqrt{\frac{(x_1 - \bar{x}_{III})^2 + (x_2 - \bar{x}_{III})^2 + \dots + (x_{60} - \bar{x}_{III})^2}{k-1}}$$
$$S_3 = \sqrt{\frac{(0,2875 - 0,249375)^2 + (0,2375 - 0,249375)^2 + \dots + (0,2375 - 0,249375)^2}{60-1}}$$
$$S_3 = 0,035161929$$

Jadi, nilai simpangan baku dalam cluster (S_w) beasiswa BBA dengan menggunakan metode K-means adalah:

$$S_w = \frac{S_1 + S_2 + S_3}{3}$$
$$S_w = \frac{0,023811886 + 0,033564641 + 0,035161929}{3}$$
$$S_w = 0,030846152$$

Simpangan Baku Antar Cluster (Sb)

Dapat dilihat pada lampiran 6 bahwa :

$$\bar{x}_I = 0,304817708$$

$$\bar{x}_{II} = 0,319791667$$

$$\bar{x}_{III} = 0,249375$$

$$\bar{X} = \frac{\bar{x}_I + \bar{x}_{II} + \bar{x}_{III}}{3}$$

$$\bar{X} = \frac{0,304817708 + 0,319791667 + 0,249375}{3}$$

$$\bar{X} = 0,291328125$$

$$Sb = \left[(3 - 1)^{-1} = \sum_{k=1}^3 (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

$$\left(\frac{(0,304817708 - 0,291328125)^2 + (0,319791667 - 0,291328125)^2 + (0,249375 - 0,291328125)^2}{3 - 1} \right)^{1/2}$$

$$Sb = 0,001376103^{1/2} = 0,037095867$$

Jadi nilai simpangan baku antar cluster (Sb) beasiswa BBA dengan metode K-Means adalah 0,037095867.

Rasio Antara Simpangan Baku Dalam Cluster dan Simpangan Baku Antar Cluster.

$$\text{Rasio} = \frac{S_w}{S_b} \times 100 \% = \frac{0,030846152}{0,037095867} \times 100 \%$$

$$= 0,831525302 \%$$

3.4 Implementasi Metode Clustering Terbaik

Berdasarkan nilai rasio simpangan baku dalam cluster dan simpangan baku antar cluster menunjukkan bahwa metode Ward dan K-Means memiliki kinerja yang sama ketika digunakan pada clustering beasiswa PPA, kedua metode memiliki nilai rasio (Sw/Sb) yang sama yaitu 0,749959584 %. Dengan demikian maka hasil cluster yang akan digunakan sebagai acuan pada penentuan penerima beasiswa PPA adalah hasil analisis cluster menggunakan metode Ward ataupun K-Means karena keduanya memiliki hasil cluster dan kinerja yang sama.

Sedangkan pada clustering beasiswa BBA, metode Ward memiliki kinerja yang lebih baik daripada metode K-Means karena metode Ward memiliki nilai rasio (Sw/Sb) 0,5346668 % lebih kecil dibandingkan dengan nilai rasio (Sw/Sb) metode K-Means yaitu 0,831525302 %. Dengan demikian maka hasil cluster yang dapat digunakan sebagai acuan pada penentuan penerima beasiswa BBA adalah hasil analisis cluster menggunakan metode Ward.

4. KESIMPULAN DAN SARAN

4.1 KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan tentang analisis *cluster* dengan metode *Ward* dan metode *K-Means* pada data mahasiswa pemohon beasiswa di STMIK Pringsewu, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Cluster yang terbentuk dengan metode *Ward* dan *K-Means* pada data mahasiswa pemohon beasiswa Peningkatan Prestasi Akademik (PPA) memiliki jumlah anggota yang sama pada tiap cluster. Sedangkan cluster yang terbentuk dengan metode *Ward* dan *K-Means* beasiswa Biaya Bantuan Akademik (BBA) memiliki jumlah anggota yang berbeda pada tiap cluster.
2. Jika ditinjau dari nilai rasio (S_w/S_b) simpangan baku dalam dan antar cluster pada data mahasiswa pemohon beasiswa Peningkatan Prestasi Akademik (PPA) metode *Ward* dan *K-Means* memiliki kinerja yang sama baik karena memiliki nilai rasio yang sama. Sedangkan pada data mahasiswa pemohon beasiswa Biaya Bantuan Akademik (BBA) metode *Ward* merupakan metode clustering terbaik, karena memiliki nilai rasio yang lebih kecil dari metode *K-Means*.

4.2 SARAN

Adapun karena adanya keterbatasan dalam penelitian ini maka saran yang diajukan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian ini dapat dijadikan sebagai salah satu referensi bagi Perguruan Tinggi untuk digunakan sebagai acuan dalam pengambilan keputusan untuk menentukan mahasiswa penerima beasiswa.
2. Dalam penelitian ini penulis hanya membahas perbandingan analisis *cluster* menggunakan metode *Ward* dan *K-Means*. Oleh karena itu untuk penelitian selanjutnya penulis menyarankan agar dapat mengembangkan metode-metode analisis cluster yang lain karena cakupan metode analisis cluster yang cukup banyak serta dapat dikembangkan dan diaplikasikan pada bidang ilmu dan studi kasus yang berbeda.

REFERENSI

- [1] Agusta Yudi. Februari 2007. *K-Means – Penerapan, Permasalahan dan Metode Terkait*. Jurnal Sistem dan Informatika. Volume 3. <http://www.library.gunadarma.ac.id/journal/view/11303/k-means-penerapan-permasalahan-dan-metode-terkait.html/>. Diakses pada tanggal 2 Desember 2015.
- [2] Connolly, T., Begg, C. 2010. *Database Systems: A Practical Approach To Design, Implementation, And Management*. 5th Edition. America: Pearson Education.
- [3] Dikti. 2014. *Pedoman Umum Beasiswa Dan Bantuan Biaya Pendidikan Peningkatan Prestasi Akademik*.
- [4] Dista Ariyadi, Bagas. 2013. *Sistem Pendukung Keputusan Seleksi Penerima Beasiswa Pada SMA 1 Boja Dengan Menggunakan Metode Analytic Hierarchy Process (AHP)*. <http://eprints.dinus.ac.id/12514/>. Diakses pada tanggal 30 November 2015.

-
- [5] Gordon S. Linoff, Michael J. A. Berry. 2011. *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*, 3rd Edition. Canada: Wiley.
- [6] Jiawei Han, Micheline Kamber and Jian Pei. 2011. *Data Mining: Concepts and Techniques*, 3rd edition. Morgan Kaufmann Publishers.
- [7] Junaeni Ina. Juni 2010. Pengelompokan Wilayah Curah Hujan Kalimantan Barat Berbasis Metode Ward Dan Fuzzy Clustering. *Jurnal Sains Dirgantara*. Volume 7. No.2. http://jurnal.lapan.go.id/index.php/jurnal_sains/article/view/1117. Diakses pada tanggal 5 Desember 2015.
- [8] KBBI. Pengertian Beasiswa. <http://kbbi.web.id/beasiswa> . Diakses pada tanggal 30 November 2015.
- [9] Kusrini, dan Emha Taufik Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: Penerbit Andi.
- [10] Kusumadewi Sri, Rismawan Tedy. 2008. *Aplikasi K-Means Untuk Pengelompokan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka*. SNATI Yogyakarta. ISSN: 1907-5022.
- [11] <http://journal.uui.ac.id/index.php/Snati/article/view/753>. Diakses pada tanggal 3 Desember 2015.
- [12] Larose Daniel T. 2006. *Data Mining Methods and Models*. Canada: Wiley.
- [13] Murniasih Erny. 2009. *Buku Pintar Beasiswa*. Jakarta: Gagas Media.
- [14] Prasetyo Eko. 2012. *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi.
- [15] Rachmatin Dewi. 2014. *Aplikasi Metode-Metode Agglomerative Dalam Analisis Klaster Pada Data Tingkat Polusi Udara*. *Jurnal Infinity*. Volume 3. No.2. <http://e-journal.stkipsiliwangi.ac.id/index.php/infinity/article/view/59>. Diakses pada tanggal 5 Desember 2015.
- [16] Ridho Barakbah, Ali dan Arai Kohei. 2004. *Determining Constraints of Moving Variance to Find Global Optimum and Make Automatic Clustering*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.469.7736&rep=rep1&type=pdf>. Diakses pada tanggal 30 November 2015.
- [17] Simamora, Bilson. 2004. *Riset Pemasaran*. Jakarta: Gramedia Utama.
- [18] Suprihatin. Februari 2011. *Klastering K-Means Untuk Penentuan Nilai Ujian*. *JUSI*. Volume 1. No. 1. ISSN: 2087-8737. <http://is.uad.ac.id/jusi/files/06-JUSI-Vol-1-No-1-Klastering-K-Means-untuk-Penentuan-Nilai-Ujian.pdf>. Diakses pada tanggal 2 Desember 2015.