

Perbandingan Kinerja Metode *Average Linkage* dan Metode *Ward*

Dalam Menentukan Hasil Analisis *Cluster*

(Studi kasus: Data Jumlah Kasus Penyakit pada 12 Kecamatan di Kabupaten

Kulon Progo Tahun 2015)

SKRIPSI

untuk memenuhi sebagian persyaratan

mencapai derajat Sarjana S-1



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Riski Yuli Artanti

14610019

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UIN SUNAN KALIJAGA

YOGYAKARTA

2018



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi/Tugas akhir
Lamp :-

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Riski Yuli Artanti
NIM : 14610019
Judul Skripsi : Perbandingan Kinerja Metode *Average Linkage* dan Metode *Ward* Dalam Menentukan Hasil Analisis *Cluster* (Studi Kasus Data Jumlah Penyakit pada 12 Kecamatan di Kabupaten Kulon Progo)

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam bidang matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 8 Mei 2018

Pembimbing II

Pipit Pratiwi Rahayu, M.Si
NIP. 19861208 201503 2 006

Pembimbing I

Dr. Epha Diana Supandi, M.Sc
NIP. 19750912 200801 2 015



PENGESAHAN SKRIPSI/TUGAS AKHIR

Nomor : B-215/Un.02/DST/PP.05.3/05/2018

Skripsi/Tugas Akhir dengan judul : Perbandingan Kinerja Metode Average Linkage dan Metode Ward dalam Menentukan Hasil Analisis Cluster (Studi Kasus : Data Jumlah Kasus Penyakit pada 12 Kecamatan di Kabupaten Kulon Progo Tahun 2015)

Yang dipersiapkan dan disusun oleh :

Nama : Riski Yuli Artanti

NIM : 14610019

Telah dimunaqasyahkan pada : 25 Mei 2018

Nilai Munaqasyah : A / B

Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga

TIM MUNAQASYAH :

Ketua Sidang

Dr. Epha Diana Supandi, M.Sc
NIP. 19750912 200801 2 015

Penguji I

Pipit Pratiwi Rahayu, M.Sc
NIP.19861208 201503 2 006

Penguji II

Moh. Farhan Quadratullah, M.Si
NIP.19790922 200801 1 011

Yogyakarta, 31 Mei 2018

UIN Sunan Kalijaga

Fakultas Sains dan Teknologi

Dekan



Dr. Murtono, M.Si

NIP. 19691212 200003 1 001

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini, saya:

Nama : Riski Yuli Artanti

NIM : 14610019

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Perbandingan Kinerja Metode *Average Linkage* dan
Metode *Ward* Dalam Menentukan Hasil Analisis *Cluster*
(Studi Kasus Data Jumlah Kasus Penyakit pada 12
Kecamatan di Kabupaten Kulon Progo)

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya orang lain, kecuali tertulis diacu dalam naskah dan disebutkan dalam daftar pustaka.

Apabila kelak di kemudian hari terbukti ada ketidakbenaran dalam pernyataan saya di atas, maka saya bertanggungjawab sepenuhnya.

Yogyakarta, 19 April 2018



RISKI YULI ARTANTI

NIM: 14610019

KATA PENGANTAR

Bismillahirrohmanirrohim,

Alhamdulillah, penulis panjatkan puji syukur kehadiran Allah SWT yang telah memberikan petunjuk dan kekuatan serta melimpahkan rahmat dan hidayahNya sehingga penyusunan skripsi ini dapat berjalan dengan lancar. Shalawat dan salam semoga senantiasa tercurahkan kepada junjungan kita Nabi Muhammad SAW beserta keluarga, para sahabat dan para pengikutnya seluruh umat Islam hingga akhir zaman, Insyaallah termasuk kita, Amin.

Penyusunan skripsi ini dimaksudkan untuk memenuhi sebagian persyaratan guna memperoleh gelar sarjana Program Studi Matematika. Skripsi ini berisi mengenai pembahasan perbandingan analisis *cluster* menggunakan metode *average linkage* dan *ward*. Penulis menyadari bahwa tanpa bantuan, bimbingan, dan motivasi dari berbagai pihak, laporan skripsi ini tidak dapat selesai dengan baik. Oleh karena itu ucapan terima kasih disampaikan sebesar-besarnya dan semoga Allah memberikan ridho-Nya kepada:

1. Bapak Dr. Murtono, M.Si selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
2. Bapak Muh. Wakhid Mustofa, M.Si, selaku Kaprodi Matematika UIN Sunan Kalijaga Yogyakarta.
3. Ibu Malahayati, M.Si selaku Dosen Pembimbing Akademik Matematika 2014 yang telah membimbing selama kami selama 4 tahun dan juga telah memberikan arahan, bantuan, nasehat yang sangat berarti sekali buat saya.

4. Ibu DrEpha Diana Supandi, M.Sc, selaku Dosen Pembimbing I dan Ibu Pipit Pratiwi Rahayu, M.Si selaku pembimbing II yang telah meluangkan waktu memberikan bimbingan, arahan, bantuan, dan ilmu sehingga skripsi ini dapat terselesaikan.
5. Ibu Pipit Pratiwi Rahayu, M.Si, selaku dosen penguji I yang telah meluangkan waktu memberikan bimbingan, arahan, bantuan, dan ilmu sehingga skripsi ini dapat terselesaikan.
6. Bapak Moh. Farhan Qudratulloh, M.Si selaku dosen penguji II yang telah meluangkan waktu memberikan bimbingan, arahan, bantuan, dan ilmu sehingga skripsi ini dapat terselesaikan.
7. Bapak/Ibu Dosen dan Staf Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta atas ilmu, bimbingan dan pelayanan selama perkuliahan sehingga penyusunan skripsi ini dapat terselesaikan.
8. Ibu Marningsih dan Bapak Surahno tercinta yang disetiap saat selalu memberikanku motivasi dan semangat kehidupan untuk meraih kesuksesan. Terimakasih atas kasih sayang yang kalian berikan selama ini dan telah menjadi orang tua yang luar biasa.
9. Kakakku Rizky Setiawan dan Kakak Iparku Rizky Nur Anisa serta segenap keluarga besarku yang tak pernah putus memberikan doa dan dukungan kepadaku untuk meraih kesuksesan
10. Teruntuk keluarga di Kos Babe dan alumni kos babe yaitu Marisa, Uyun, Esta, Riana, Mbak puput, Almarhumah Eva, Mbak Via, Mbak Haliemah , Mbak Rinda, Mbak Firqoh dan semuanya terimakasih untuk dukungan, hiburan, kegilaan

yang pernah kita ciptakan dan semangat yang diberikan selama ini. Sukses buat kita semua. Love you ALL...

11. Teman-temanku satu Prodi Matematika angkatan 2014 yang telah memberi warna, bantuan, pertanyaan “ kapan seminar” , “kapan munaqosyah” dan dukungan selama ini. Semoga sukses buat kita semua.
12. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu dalam penyusunan skripsi ini.

Semoga Allah SWT berkenan membalas kebaikan mereka dengan pahala yang berlipat ganda. Hanya kepada Allah penulis menyembah dan memohon ampunan atas segala kekurangan dan kekhilafan. Semoga skripsi ini bermanfaat bagi pembaca pada umumnya dan penulis pada khususnya.

Yogyakarta, 10 Mei 2018

Penulis

Riski Yuli Artanti

NIM. 14610019

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERSEMBAHAN

Skripsi ini penulis persembahkan kepada:

- *Allah SWT yang telah memberikan kesempatan hamba-Nya untuk mencari ilmu dan belajar arti kehidupan*
- *Junjungan Nabi Besar, Nabi Muhammad SAW yang mengentaskan umat Nya dari jahiliyah*
- *Ibu dan Bapakku yang telah membesarkan, mendidik, mendo'akan dan mendukung lahir dan batin*
- *Kakakku yang senantiasa mendukung lahir dan batin*
- *Semua teman-teman yang selalu memberikan semangat kepadaku*
- *Guru-guruku yang telah memberikan ilmu kepadaku*
- *Almamater Prodi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga*

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

*Succes is walking from failure to failure
with no loss of enthusias (Winston Churchill)*

*Sukses berjalan dari kegagalan satu menuju kegagalan
lain tanpa kehilangan semangat dan antusiasme
(Winston Churchill)*

*Banyak hal yang rumit dikerjakan, sulit dipikirkan tapi ternyata
mudah jika diselesaikan. Kesulitan sebenarnya ada karena tidak
mau memulainya dan menyelesaikannya
(Motivator Muda Muslim)*

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
SURAT PERSETUJUAN	ii
PENGESAHAN SKRIPSI	iii
PERNYATAAN KEASLIAN	iv
KATA PENGANTAR	v
HALAMAN PERSEMBAHAN	viii
HALAMAN MOTTO	ix
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR SIMBOL	xvii
DAFTAR LAMPIRAN.....	xviii
ABSTRAKSI	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Batasan Masalah	4
1.3 Rumusan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	5
1.6 Tinjauan Pustaka.....	5
1.7 Sistematika Penulisan	8
BAB II LANDASAN TEORI.....	10

2.1	Analisis Multivariat	10
2.2	Matriks dan Operasi Matrik	13
2.2.1	Definisi Matriks	13
2.2.2	Jenis Matrik	14
2.2.3	Penjumlahan Matriks.....	15
2.2.4	Perkalian Matriks dengan Skalar	16
2.2.5	Perkalian Dua Matriks.....	16
2.2.6	<i>Transpose</i> Matriks.....	17
2.2.7	Invers Matriks	18
2.2.8	Determinan Matriks.....	18
2.3	Matrik Data Multivariat	18
2.4	Skala Pengukuran Data	20
2.5	Analisis <i>Cluster</i>	21
2.5.1	Tujuan Analisis <i>Cluster</i>	22
2.5.2	Prosedur Analisis <i>Cluster</i>	24
2.6	Pemilihan Metode Terbaik dengan Simpangan Baku.....	35
2.7	Interpretasi <i>Cluster</i>	36
2.8	KESEHATAN & PENYAKIT.....	40
2.8.1	Pengertian Kesehatan.....	40
2.8.2	Usaha Kesehatan Perorangan, Masyarakat dan Lingkungan.....	41
2.8.3	Macam-macam penyakit dan Cara Mencegahnya.....	43
BAB III METODE PENELITIAN		50
3.1	Jenis Penelitian.....	50
3.2	Objek Penelitian	50
3.3	Variabel Penelitian	51

3.4	Sumber Penelitian.....	51
3.5	Teknik Analisis Data	52
3.6	Alur Penelitian.....	52
BAB IV PEMBAHASAN		54
4.1	Analisis <i>Cluster</i> Hierarki	54
4.2	Metode <i>Average Linkage</i>	55
4.3	Metode <i>Ward</i>	63
4.4	Pemilihan Metode Terbaik dengan Simpangan Baku	71
BAB V STUDI KASUS		76
5.1	Analisis <i>Cluster</i> dengan Menggunakan Metode <i>Average Linkage</i>	78
5.2	Analisis <i>Cluster</i> dengan Menggunakan Metode <i>Ward</i>	88
5.3	Pengklasteran II (Pengelompokkan Tanpa menggunakan Variabel HIV/AIDS, DBD, dan IMS)	97
5.3.1	Analisis <i>Cluster</i> dengan Metode <i>Average Linkage</i>	97
5.3.2	Analisis <i>Cluster</i> dengan Metode <i>Ward</i>	106
5.4	Analisis Hasil Pengolahan	114
5.5	Pemilihan Metode Terbaik.....	115
BAB VI PENUTUP		116
6.1	Kesimpulan	116
6.2	Saran	117
DAFTAR PUSTAKA		119
CURRICULUM VITAE		121
LAMPIRAN		122

DAFTAR TABEL

	Halaman
Tabel 1.1 Perbandingan Penelitian	7
Tabel 2.1 Matriks Data Multivariat	19
Tabel 2.2 Nilai Variabel di dalam <i>Cluster</i> ke- <i>i</i>	38
Tabel 2.3 Perhitungan Nilai F Analisis Diskriminan.....	39
Tabel 4.1 Data Nilai X_1 dan X_2 dari 5 objek	57
Tabel 4.2 Matriks Jarak (Tahap I)	58
Tabel 4.3 Matriks Jarak (Tahap II)	58
Tabel 4.4 Matriks Jarak (Tahap III).....	59
Tabel 4.5 <i>Test of Equality of Group Means</i> (Sampel) <i>Average Linkage</i>	61
Tabel 4.6 <i>Standardized Coefficient</i> (Sampel) <i>Average Linkage</i>	62
Tabel 4.7 <i>Structure Matrix</i> (Sampel) <i>Average Linkage</i>	62
Tabel 4.8 <i>Cannonical Diskriminan Coefficient</i> (Sampel) <i>Average Linkage</i> ..	63
Tabel 4.9 Metode Ward (Tahap I)	66
Tabel 4.10 Metode Ward (Tahap II).....	67
Tabel 4.11 <i>Test of Equality of Group Means</i> (Sampel) <i>Ward</i>	69
Tabel 4.12 <i>Standardized Coefficient</i> (Sampel) <i>Ward</i>	70
Tabel 4.13 <i>Structure Matrix</i> (Sampel) <i>Ward</i>	70
Tabel 4.14 <i>Cannonical Diskriminan Coefficient</i> (Sampel) <i>Ward</i>	71
Tabel 4.15 Nilai rata-rata <i>cluster</i> I.....	72
Tabel 4.16 Nilai rata-rata <i>cluster</i> II	72
Tabel 4.17 Nilai Simpangan baku ke- <i>k</i>	73
Tabel 5.1 Data hasil laporan publikasi Badan Pusat Statistik(BPS) DIY	78

Tabel 5.2 <i>Case Processing Summary</i> Metode <i>Average</i> (Pengklasteran I) ..	79
Tabel 5.3 <i>Test of Equality of Group Means</i> (Pengklasteran I) <i>Average</i>	83
Tabel 5.4 <i>Standardized Coefficient</i> (Pengklasteran I) <i>Average</i>	85
Tabel 5.5 <i>Structure Matrix</i> (Pengklasteran I) <i>Average</i>	86
Tabel 5.6 <i>Cannonical Discriminant function</i> (Pengklasteran I) <i>Average</i>	86
Tabel 5.7 Hasil rata-rata perkuster Metode <i>Average</i> (Pengklasteran I)	87
Tabel 5.8 <i>Case Processing Summary</i> Metode <i>Ward</i> (Pengklasteran I)	88
Tabel 5.9 <i>Test of Equality of Group Means</i> (Pengklasteran I) <i>Ward</i>	92
Tabel 5.10 <i>Standardized Coefficient</i> (Pengklasteran I) <i>Ward</i>	94
Tabel 5.11 <i>Structure Matrix</i> (Pengklasteran I) <i>Ward</i>	95
Tabel 5.12 <i>Cannonical Discriminant function</i> (Pengklasteran I) <i>Ward</i>	95
Tabel 5.13 Hasil rata-rata perkuster Metode <i>Ward</i> (Pengklasteran I)	96
Tabel 5.14 <i>Case Processing Summary</i> Metode <i>Average</i> (Pengklasteran II).	97
Tabel 5.15 <i>Test of Equality of Group Means</i> (Pengklasteran II) <i>Average</i> ...	101
Tabel 5.16 <i>Standardized Coefficient</i> (Pengklasteran II) <i>Average</i>	103
Tabel 5.17 <i>Structure Matrix</i> (Pengklasteran II) <i>Average</i>	103
Tabel 5.18 <i>Cannonical Discriminant function</i> (Pengklasteran II) <i>Average</i> .	104
Tabel 5.19 Hasil rata-rata perkuster Metode <i>Average</i> (Pengklasteran II).	105
Tabel 5.20 <i>Case Processing Summary</i> Metode <i>Ward</i> (Pengklasteran II)...	106
Tabel 5.21 <i>Test of Equality of Group Means</i> (Pengklasteran II) <i>Ward</i>	109
Tabel 5.22 <i>Standardized Coefficient</i> (Pengklasteran II) <i>Ward</i>	111
Tabel 5.23 <i>Structure Matrix</i> (Pengklasteran II) <i>Ward</i>	112
Tabel 5.24 <i>Cannonical Discriminant function</i> (Pengklasteran II) <i>Ward</i>	112
Tabel 5.25 Hasil rata-rata perkuster Metode <i>Ward</i> (Pengklasteran II)	113

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh Dendogram.....	22
Gambar 2.2 Klasifikasi Prosedur Pengklasteran	34
Gambar 3.1 Alur Penelitian.....	53
Gambar 4.1 <i>Dendogram Average Linkage</i> untuk tabel 4.1.....	60
Gambar 4.2 <i>Dendogram Ward</i> untuk tabel 4.1	68
Gambar 5.1 <i>Dendogram</i> Pengklasteran <i>Average Linkage</i> (Pengklasteran I)	82
Gambar 5.2 <i>Dendogram</i> Pengklasteran <i>Ward</i> (Pengklasteran I).....	91
Gambar 5.3 <i>Dendogram</i> Pengklasteran <i>Average Linkage</i> (Pengklasteran II)	100
Gambar 5.4 <i>Dendogram</i> Pengklasteran <i>Ward</i> (Pengklasteran II).....	108

DAFTAR SIMBOL

n	: banyaknya objek
p	: banyaknya variabel
d_{ik}	: jarak antara obyek i dalam cluster (UV) dan obyek k dalam cluster
$N_{(UV)}$	: jumlah obyek dalam (UV)
N_W	: jumlah obyek dalam (W)
Y_{ij}	: data untuk cluster ke- i objek ke- j
Y_i	: jumlah total data pada cluster ke- i
Y	: jumlah seluruh data
D	: skor diskriminan
b	: Koefisien diskriminan atau bobot
d_{ij}	: jarak euclidean (ukuran kemiripan) antara objek ke- i dan objek ke- j
x_{ik}	: data dari objek ke- i pada variabel ke- k
x_{jk}	: data dari objek ke- j pada variabel ke- k
x_i	: data ke- i
$\bar{x}$	: rata-rata data
s	: simpangan baku
s_w	: simpangan baku dalam kelompok
s_b	: simpangan baku antar kelompok
K	: banyaknya kelompok yang terbentuk
s_i	: Simpangan baku kelompok ke- i
$\bar{x}_i$	: Rataan kelompok ke- i

x_1 : data ke-I untuk kelompok ke-1

$\bar{x}_1$: rata-rata kelompok ke-I



DAFTAR LAMPIRAN

	Halaman
Lampiran 1. <i>Agglomeration Schedule</i> Metode <i>Average Linkage</i> (Pengklasteran I).....	122
Lampiran 2. <i>ClusterMembership</i> Metode <i>Average Linkage</i> (Pengklasteran I).....	122
Lampiran 3. <i>Agglomeration Schedule</i> Metode <i>Ward</i> (Pengklasteran I)	123
Lampiran 4. <i>Cluster Membership</i> Metode <i>Ward</i> (Pengklasteran I).....	123
Lampiran 5. <i>Agglomeration Schedule</i> Metode <i>Average Linkage</i> (Pengklasteran II)	124
Lampiran 6. <i>ClusterMembership</i> Metode <i>Average Linkage</i> (Pengklasteran II).....	124
Lampiran 7. <i>Agglomeration Schedule</i> Metode <i>Ward</i> (Pengklasteran II)....	125
Lampiran 8. <i>Cluster Membership</i> Metode <i>Ward</i> (Pengklasteran II)	125
Lampiran 9. <i>Proximity</i> Matriks (Pengklasteran I)	126
Lampiran 10. <i>Proximity</i> Matriks (Pengklasteran II).....	127
Lampiran 11. Data hasil Pengelompokkan dengan Metode <i>Average Linkage</i> (Pengklasteran I).....	127
Lampiran 12. Data hasil Pengelompokkan dengan Metode <i>Ward</i> (P I)	128
Lampiran 13. Data hasil Pengelompokkan Metode <i>Average</i> (P II).....	128
Lampiran 14. Data hasil Pengelompokkan Metode <i>Ward</i> (P II)	129
Lampiran 15. Perhitungan Nilai F pada variabel HIV	129
Lampiran 16. Perhitungan simpangan baku pada kedua metode pada data sesungguhnya.....	131

Perbandingan Kinerja Metode *Average Linkage* dan Metode *Ward* Dalam Menentukan Hasil Analisis *Cluster*

(Studi kasus: Data Jumlah Kasus Penyakit pada 12 Kecamatan di Kabupaten

Kulon Progo Tahun 2015)

Oleh: Riski Yuli Artanti (14610019)

ABSTRAK

Analisis *cluster* merupakan suatu teknik analisis dengan tujuan untuk memilah obyek ke dalam beberapa kelompok yang mempunyai sifat berbeda antara kelompok satu dengan yang lain. Dalam analisis ini tiap-tiap kelompok bersifat homogen antar anggota dalam kelompok atau variasi obyek kelompok yang terbentuk sekecil mungkin

Penelitian ini bertujuan untuk membentuk *cluster* dengan data jumlah kasus penyakit pada 12 kecamatan di kabupaten Kulon Progo tahun 2015 dengan metode *cluster* hierarki *Average Linkage* dan *Ward*, serta menjelaskan perbandingan hasil yang diperoleh jika ditinjau dari simpangan baku dalam dan antar kelompok.

Langkah-langkah dalam analisis *cluster* hierarki yaitu melakukan standarisasi data, menentukan kemiripan atau ketidakmiripan antar data, proses pengklasteran dengan menggunakan matriks jarak dan *agglomeration schedule*, menentukan jumlah *cluster* dan anggotanya, menginterpretasikan hasil *cluster* yang dibentuk. Selanjutnya untuk pemilihan metode terbaik menggunakan nilai minimum simpangan baku dalam kelompok terhadap simpangan baku antar kelompok.

Hasil menunjukkan bahwa pada metode *average linkage* membentuk 4 *cluster*, *cluster* pertama 8 kecamatan, *cluster* kedua 2 kecamatan, *cluster* ketiga 1 kecamatan, *cluster* keempat 1 kecamatan. Pada metode *ward* terbentuk *cluster* pertama 3 kecamatan, *cluster* kedua 5 kecamatan, *cluster* ketiga 2 kecamatan, *cluster* keempat 2 kecamatan. Untuk metode *average* diperoleh kelompok kecamatan *cluster* 1 dengan tingkat kesehatan sangat baik, *cluster* 2 dengan tingkat kesehatan baik, *cluster* 3 dengan tingkat kesehatan kurang baik, *cluster* 4 dengan tingkat kesehatan buruk/rawan penyakit. Untuk metode *ward* diperoleh kelompok kecamatan *cluster* 1 dengan tingkat kesehatan sangat baik, *cluster* 2 dengan tingkat kesehatan baik, *cluster* 3 dengan tingkat kesehatan kurang baik, *cluster* 4 dengan tingkat kesehatan buruk/rawan penyakit. Pada perbandingan nilai simpangan baku (s), nilai terkecil dimiliki oleh metode *average linkage* yang berarti metode *average* adalah metode terbaik diantara kedua metode tersebut.

Kata Kunci: Analisis *cluster*, *Average Linkage*, *Ward*, Simpangan baku Dalam dan Antar Kelompok.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis *cluster* merupakan suatu teknik analisis dengan tujuan untuk memilah obyek ke dalam beberapa kelompok yang mempunyai sifat berbeda antara kelompok satu dengan yang lain. Dalam analisis ini tiap-tiap kelompok bersifat homogen antar anggota dalam kelompok atau variasi obyek kelompok yang terbentuk sekecil mungkin.

Secara umum terdapat dua metode pengelompokan data dalam analisis *cluster* yaitu metode hierarki dan metode non-hierarki (Supranto, 2010). Analisis *cluster* dengan metode hierarki adalah analisis yang pengklasteran datanya dilakukan dengan cara mengukur jarak kedekatan pada setiap obyek yang kemudian membentuk sebuah dendogram. Sedangkan analisis *cluster* dengan menggunakan metode non hierarki adalah metode pengelompokkan dengan menentukan berapa kelas terlebih dahulu. Jenis metode hierarki ada beberapa metode yaitu metode Pautan Tunggal(*single linkage*), metode pautan lengkap(*completelinkage*), metode antar pusat(*centroid linkage*), metode pautan rata-rata(*average linkage*) dan metode Ward (*ward's method*), sedangkan pada metode non-hierarki adalah metode *K-Means* dan *K-Median*.

Langkah-langkah dalam analisis *cluster* hierarki yaitu melakukan standarisasi data, menentukan ukuran kemiripan atau ketidakmiripan antar data, proses pengklasteran dengan matriks jarak dan *agglomeration schedule*, menentukan jumlah *cluster* dan anggotanya, menginterpretasikan hasil *cluster* yang dibentuk. Setelah diperoleh hasil pengklasteran, kemudian dilakukan perhitungan rasio simpangan baku antar *cluster* berguna untuk mengetahui metode mana yang kinerjanya paling baik. Metode pengelompokan yang baik merupakan metode yang mempunyai nilai simpangan baku dalam kelompok (s_w) yang minimum dan nilai simpangan baku antar kelompok (s_b) yang maksimum.

Pada skripsi ini metode yang digunakan adalah metode keterkaitan rata-rata (*average linkage*) dan metode *Ward*. Salah satu alasan menggunakan metode keterkaitan rata-rata (*average linkage*) karena metode ini sudah banyak dibahas pada skripsi dan juga merupakan metode terbaik dari pada metode hierarki yang lainnya. Alasan menggunakan metode *Ward* karena metode *Ward* juga merupakan metode terbaik pada analisis *cluster* dengan metode hierarki karena metode ini dapat meminimumkan jumlah kuadrat (ESS). Jadi dari kedua metode paling baik ini diantara metode hierarki lainnya, penulis ingin membandingkan diantara kedua metode terbaik tersebut.

Konsep analisis *cluster* serta penerapannya dalam kehidupan sehari-hari sangat luas sehingga mencakup berbagai bidang psikiatri contohnya identifikasi jenis depresi, pada bidang sosiologi contohnya pengelompokan kabupaten berdasarkan tingkat perceraian, di bidang antropologi contohnya pengelompokan tingkat kepadatan penduduk, bidang penelitian pasar contohnya pengelompokan

memahami perilaku pembeli, bidang ekonomi contohnya mengidentifikasi produk baru dan lain sebagainya.

Penerapan analisis *cluster* pada bidang kesehatan pada skripsi ini yaitu dengan mengelompokkan kecamatan untuk mengetahui tingkat kesehatan menurut jumlah kasus beberapa penyakit yang diderita oleh masyarakat. Setiap tahun terdapat kasus masyarakat yang menderita suatu penyakit, pada skripsi ini kasus penyakit yang dipilih adalah jumlah kasus penyakit dan masalah kesehatan yang cukup sering terjadi pada wilayah kabupaten Kulon Progo yaitu penyakit HIV/AIDS, IMS(Infeksi Menular Sexual), DBD(Demam berdarah *Dengue*), Diare, serta masalah gizi buruk pada balita. Penelitian ini bertujuan untuk membentuk *cluster* dengan data jumlah kasus penyakit pada 12 kecamatan di kabupaten Kulon Progo tahun 2015 dengan metode *cluster* hierarki *average linkage* dan metode *ward*, serta membandingkan metode *average* dan metode *ward* jika ditinjau dari simpangan baku dalam kelompok(s_w) dan nilai simpangan antar kelompok(s_b) sehingga nanti dapat diketahui metode terbaik.

Dengan pemilihan kedua metode tersebut selanjutnya dilakukan perhitungan untuk memperoleh metode terbaik. Metode pengelompokan yang baik merupakan metode yang mempunyai nilai simpangan baku dalam kelompok (s_w) yang minimum dan nilai simpangan baku antar kelompok (s_b) yang maksimum (Barakbah dan Arai, 2004). Hal tersebut menjadi latar belakang penulis tertarik untuk melakukan penelitian pada skripsi ini dengan judul “Perbandingan Kinerja Metode *Average Linkage* dan Metode *Ward* Dalam Menentukan Hasil Analisis

Cluster”. Adapun maksud dari kinerja metode dalam judul skripsi tersebut merupakan suatu hasil kerja(proses) dari masing-masing metode.

1.2 Pembatasan Masalah

Penulis membatasi penelitian dengan menggunakan jarak *SquareEuclidean* dalam perhitungan kemiripan antar obyek pada pemilihan metode terbaik dari dua metode yaitu metode *Average Linkage* dan metode *Ward* dengan data jumlah kasus penyakit di kabupaten Kulon progo pada tahun 2015.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat dirumuskan permasalahan sebagai berikut:

1. Bagaimana prosedur analisis *cluster* metode *Average Linkage* dan *Ward*.
2. Bagaimana hasil *cluster* yang terbentuk dengan metode *Average Linkage*, metode *Ward* pada data jumlah kasus penyakit dalam kecamatan di kabupaten Kulon Progo tahun 2015?
3. Manakah metode yang terbaik dari kedua metode *Average Linkage*, metode *Ward* jika ditinjau dari s_w dan s_b pada data jumlah kasus penyakit di kabupaten Kulon Progo tahun 2015?

1.4 Tujuan Penelitian

Tujuan dari penulisan tugas akhir ini adalah:

1. Mengetahui prosedur analisis *cluster* metode *Average Linkage* dan *Ward*.

2. Mengetahui hasil *cluster* yang terbentuk dengan metode *Average Linkage* dan metode *Ward* data jumlah kasus penyakit dalam kecamatan di kabupaten Kulon Progo tahun 2015.
3. Mengetahui pengelompokkan terbaik antara metode *Average Linkage* dan metode *Ward* jika ditinjau dari simpangan dalam kelompok (s_w) dan nilai simpangan antar kelompok (s_b).

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat bermanfaat untuk:

1. Bagi penulis sendiri, dapat memperdalam ilmu tentang analisis *cluster*.
2. Bagi para pembaca, dapat menambah pengetahuan tentang aplikasi pada ilmu statistik yaitu bidang kesehatan khususnya dengan menggunakan analisis *cluster*.
3. Bagi pihak dinas kesehatan Kabupaten Kulon Progo, dapat membantu melihat tingkat kesehatan daerah/kecamatan Kabupaten Kulon Progo.

1.6 Tinjauan Pustaka

Dalam tinjauan pustaka ini penelitian akan mendeskripsikan dan menelaah buku buku dan yang mempunyai relevansi dengan obyek pembahasan.

1. Skripsi yang berjudul “Analisis *cluster* Kabupaten/Kota di Jawa Tengah berdasarkan indikator Kemiskinan Menggunakan Metode *Complete Linkage* dan *K-Means*” yang ditulis oleh Tika Widya Wardani(2012), mahasiswa jurusan matematika Universitas Islam Negeri Sunan Kalijaga. Hasil dari penelitian dari metode

CompleteLinkage terbentuk 3 cluster, *cluster* pertama beranggotakan 5 kota,*cluster* kedua beranggotakan 17 kota,*cluster* ketiga beranggotakan 13 kota.Pada metode *K-Means* terbentuk 3 *cluster*, *cluster* pertama beranggotakan 5 kota, *cluster* kedua beranggotakan 13 kota, *cluster* ketiga beranggotakan 17 kota.

2. Skripsi yang berjudul “Perbandingan Analisis*Cluster* Menggunakan Metode *Single Linkage*, *Complete linkage*,*Average Linkage* dan *K-Means* Untuk Mengelompokkan Kecamatan Berdasarkan Variabel Jenis Ternak di kabupaten Semarang”, yang ditulis oleh Azma Sholiha (2015) mahasiswa jurusan Matematika di Universitas Negeri Semarang. Hasil kinerja terbaik dari klaster yang terbentuk menggunakan metode *Single Linkage* didapat enam pengelompokkan kecamatan. Untuk metode *Complete Linkage* terbentuk enam pengelompokan kecamatan, untuk metode *Average Linkage* didapat dua pengelompokan. Dan dengan menggunakan metode *K-means* di dapat enam pengelompokkan kecamatan.
3. Skripsi yang berjudul “Penerapan metode *Ward* pada analisis gerombol hieraki” yang ditulis oleh Sarfia (2016) mahasiswa jurusan Matematika Universitas Halu Oleo. Hasil dari penelitian metode *ward* terbentuk 3 *cluster*. Hasil penelitian *cluster* I, terdiri dari 10 kabupaten, *cluster* II terdiri dari 3 kabupaten, *cluster* III terdiri 1 kabupaten.

Tabel 1.1 Perbandingan Penelitian

No	Peneliti	Objek Penelitian	Metode	Variabel
1	Tika Widya Wardani	35 kabupaten/kota di provinsi Jawa Tengah	Analisis <i>Cluster Complete Linkage</i> dan <i>K-Means</i>	Luas lantai per kapita, pengeluaran per kapita untuk non makanan, pengguna alat KB, proses persalinan, ketersediaan air bersih, ketersediaan jamban sendiri/ bersama
2	Azma Sholiha	19 Kecamatan di kabupaten Semarang pada tahun 2012	Analisis <i>Cluster Single Linkage, Complete Linkage, Average Linkage</i> , dan <i>K-Means</i>	Banyak ternak kambing, domba, sapi potong, sapi perah, ayam buras, itik mentok
3	Sarfia	14 Kabupaten di Sulawesi Tenggara	Analisis <i>Cluster Metode Ward</i>	Populasi sapi, kerbau, kuda, kambing, domba, babi, Ayam kampung, ayam petelur, itik/itik manila
4	Riski Yuli Artanti	12 kecamatan di Kabupaten Kulon Progo	Analisis <i>Cluster Average Linkage</i> , dan <i>Metode Ward</i>	Jenis penyakit : HIV/AIDS, IMS (Infeksi Menular Sexual), DBD (Demam berdarah <i>Dangue</i>), Diare dan Gizi buruk pada balita

Perbedaan skripsi yang saya teliti dengan ketiga skripsi yang menjadi bahan referensi yaitu dilakukannya pemilihan metode terbaik pada tahap akhir sedangkan dari ketiga referensi skripsi tersebut hanya menampilkan hasil analisis *cluster* saja. Kelemahan pada skripsi saya objek penelitian yang diteliti sedikit yaitu hanya 12 objek, sedangkan dari ketiga referensi skripsi yang saya gunakan jumlah objek yang digunakan lebih dari 12.

1.7 Sistematika Penulisan

Dalam rangka mempermudah pemahaman dan pembahasan terhadap permasalahan yang diangkat dan diteliti, maka pembahasannya akan disusun secara sistematis sesuai tata urutan dari permasalahan yang ada, dengan uraiannya sebagai berikut:

Bab I : Pendahuluan

Bab ini terdiri dari latar belakang, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka dan sistematika penulisan.

Bab II : Landasan Teori

Bab ini memuat tentang teori teori serta definisi-definisi dalam analisis *Cluster*.

Bab III : Metode Penelitian

Bab ini menjelaskan tentang jenis penelitian, variabel penelitian, sumber data, teknik analisis data dan alur penelitian.

Bab IV : Pembahasan

Bab ini menjelaskan tentang konsep yang digunakan dalam tahapan – tahapan melakukan analisis *Cluster* metode *Average Linkage* dan Metode *Ward*.

Bab V : Studi Kasus

Bab ini merupakan studi kasus dari pembahasan skripsi, yang berisi tentang hasil dari penerapan analisis *cluster* kecamatan di Kulon Progo berdasarkan jenis penyakit menggunakan metode *Average Linkage* dan Metode *Ward*.

Bab VI: Penutup

Bab ini merupakan bagian penutup yang terdiri dari kesimpulan atas jawaban masalah yang ada dan saran-saran yang berkaitan dengan penelitian sejenis di masa yang akan datang.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan, maka hasil penelitian dalam skripsi ini dapat penulis simpulkan sebagai berikut:

1. Analisis *cluster* adalah suatu alat untuk mengelompokkan sejumlah n obyek berdasarkan p variat yang secara relatif mempunyai kesamaan karakteristik diantara obyek-obyek tersebut, sehingga keragaman di dalam suatu kelompok tersebut lebih kecil dibandingkan keragaman antar kelompok. Prosedur analisis *cluster* dengan metode *Average Linkage* dan *Ward* yaitu dimulai dengan deksripsi masalah, pemilihan variabel penelitian, pengambilan data, pengukuran kemiripan, pembentukan *cluster* dengan metode *Average Linkage* dan *Ward*, interpretasi *cluster* dan langkah terakhir adalah dilakukan uji diskriminan untuk mengetahui apakah ada perbedaan dari keempat *cluster* yang terbentuk.
2. Hasil *cluster* yang terbentuk dengan metode *Average Linkage* dan *Ward* untuk mengelompokkan 12 kecamatan di Kulon Progo sebagai berikut:
 - a. Untuk metode *average linkage* diperoleh:
 - *cluster*-1 adalah Wates, Panjatan, Galur, Lendah, Girimulyo, Nanggulan, Kalibawang, Samigaluh tingkat kesehatannya sangat baik.
 - *cluster*-2 adalah Pengasih, Kokap tingkat kesehatannya baik

- *cluster-3* adalah Temon, tingkat kesehatannya kurang baik
- *cluster-4* adalah Sentolo tingkat kesehatannya buruk/rawan penyakit.

b. Untuk Metode *Ward* diperoleh:

- *cluster-1* adalah Wates, Panjatan, Girimulyo, tingkat kesehatannya sangat baik
- *cluster-2* adalah Galur, Lendah, Nanggulan, Kalibawang, Samigaluh, tingkat kesehatannya baik
- *cluster-3* adalah Pengasih, Kokap, tingkat kesehatannya kurang baik
- *cluster-4* adalah Temon, Sentolo, tingkat kesehatannya buruk/rawan penyakit

3. Jika ditinjau dari nilai rasio simpangan s_w terhadap s_b , menunjukkan bahwa nilai rasio simpangan baku (s) pada metode *Average Linkage* yaitu 7.5 % lebih kecil dibandingkan dengan metode *Ward* 20.6 %, sehingga metode *average linkage* merupakan metode paling baik diantara kedua metode yang diteliti.

6.2 Saran

Pada skripsi ini penulis hanya mengkaji tentang dua metode *cluster* yaitu metode *average linkage* dan metode *ward*, yang diterapkan pada bidang kesehatan. Penelitian ini hanya menggunakan dua metode tersebut, oleh sebab itu, masih ada peluang bagi penelitian selanjutnya untuk

menggunakan metode-metode lain yang ada dalam analisis *cluster* dengan mengaplikasikan pada bidang ilmu lain. Seperti metode *single linkage*, *centroid linkage*, *complete linkage* dan masih banyak lagi.



DAFTAR PUSTAKA

- Anton, Howard.2005. *Aljabar Linear Elementer*. Jakarta: Erlangga.
- Ariyanto.2009. Pengolahan data statistika SPSS 16. Jakarta: Salemba Infotek.
- Azma Sholiha. 2015. *Perbandingan Analisis Klaster Menggunakan Metode Single Linkage, Complete Linkage, Average Linkagedan K-Means Untuk Pengelompokkan Kecamatan Berdasarkan Variabel Jenis Ternak di Kabupaten Semarang*.Skripsi.Univeristas Negeri Semarang
- Badan Pusat Statistik (BPS). 2016. *Kabupaten Kulon Progo Dalam Angka Tahun2016*. Yogyakarta: Badan Pusat Statistik.
- Barakbah Ali & Arai Kohei. 2004. *Determining Constraints of Moving Variance to Find Global Optimum and Make Automatic Clustering*.Diakses dari http://lecturer.eepis-its.edu/~ridho/papers/Barakbah_IES_2004.pdf. Pada tanggal 11 November 2017 jam 22.00 WIB
- Ichsan,M. Yulianti & Sri Rejeki.1992.*Ilmu Kesehatan dan Gizi*. Jakarta: Universitas Terbuka, Departmen Pendidikan dan Kebudayaan.
- Johnson, Richard.A. & Dean Wichern. 1996.*Applied Multivariate StastisticalAnalysis*.3th. New Delhi: Prentice-Hall

Jonathan Sarwono.2007.*Teori Analisis Multivariat*. Diakses dari <http://www.jonathansarwono.info/mvariat/multivariat.htm>. Pada tanggal 31 Oktober 2017, jam 09.00 WIB.

Sarfia. 2016. *Penerapan Metode Ward Pada Analisis Gerombol Hirarki*. Skripsi.Universitas Halu Oleo.

Simamora, Bilson. 2005. *Analisis Multivariat Pemasaran Edisi Pertama*. Jakarta: PT. Gramedia Pustakatama.

Supranto, J. 2004. *Analisis Multivariat: Arti dan Interpretasi*. Jakarta : PT. Asdi Mahasatya

Supandi, Epha Diana.2011.*Handout Analisis Multivariat Terapan*. Matematika, UIN SUKA,Yogyakarta.

Supranto, Jonathan. 2010. *Analisis Multivariate Arti & Interpretasi*. Jakarta: Rineka Cipta.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA