Perbandingan Penggunaan Aplikasi Rapid Miner Dengan Weka Untuk Penerapan Metode K-Means Clustering Pada Pengelompokan Penderita Demensia

Fivtatianti Hendajani¹ dan Anita²

Sistem Informasi, STMIK Jakarta STI&K¹, Sistem Informasi, Universitas Gunadarma² Jl. BRI Radio Dalam, Jakarta Selatan¹, Jl. Margonda Raya 100, Depok² E-mail : fivtatiantihendajani@gmail.com, anita_sugianto@staff.gunadarma.ac.id

Abstrak

Demensia merupakan gangguan neurodegenerative yang terkait dengan penuan yang menyebabkan penurunan kemampuan kognitif manusia karena degenerasi struktur otak. Banyak cara untuk mendeteksi penderita demensia namun masih belum diketahui penyebabnya. Oleh karena itu data pasien yang berkaitan dengan data demografis ditambahkan selain data klinis. Terdapat banyak data yang diharapkan kemudian dapat digunakan untuk mengetahui penyebab penyakit demensia. Banyak data yang ada akan lebih mudah pengelolaanya agar dapat diperoleh menjadi pengetahuan, menggunakan data mining. Pengelompokkan banyak datanya menggunakan K-Means Clustering. Banyak aplikasi komputer untuk menerapkan metode data mining diantaranya Rapid Miner dan Weka. 2 aplikasi ini akan dibandingkan untuk penerapan pengelompokan penderita demensia menggunakan K-Means Clustering. Hasil yang di dapat adalah memulai aplikasi Weka lebih mudah namun untuk visualisasi akan lebih baik menggunakan Rapid Miner. Terdapat sedikit perbedaan hasil dalam pembentukan kluster penderita demensia diatara 2 aplikasi tersebut.

Kata Kunci : demensia ,data mining, k-means clustering, rapid miner, weka

Pendahuluan

Demensia mempengaruhi sebagian besar orang di seluruh dunia. Demensia adalah nama dari sekelompok gejala yang menyebabkan penurunan memori seseorang yang dapat mempengaruhi kemampuannya untuk melakukan kegiatan sehari-hari. Hal ini ditandai dengan hilangnya sinapsis dan neuron di daerah kortikal kortikal dan sub otak. Penyebab pasti dari penyakit ini tidak diketahui. Sebagian besar pasien yang menderita Demensia sudah tua, tetapi juga bisa terjadi pada setiap orang. Perkembangan penyakit akan menyebabkan hilangnya memori jangka panjang [1].

Diperkirakan bahwa sekitar 25 juta orang menderita demensia saat ini dan sebagai konsekuensi dari penuaan penduduk, jumlah orang yang terkena diharapkan dua kali lipat setiap 20 tahun [2][3]. Kehadiran keluhan kognitif sangat umum pada orang tua dan mungkin merupakan tanda pertama dari gangguan demensia berlangsung seperti penyakit Alzheimer. Hal ini dimungkinkan untuk mengidentifikasi orang-orang dengan keluhan kognitif yang beresiko untuk perkembangan demensia, yang mengatakan, yang memiliki Penurunan Mild Cognitive (MCI) [4][5]. Sejak berdirinya MCI diperlukan demonstrasi kognitif penurunan lebih besar dari yang diharapkan bagi usia seorang individu dan tingkat pendidikan, tes neuropsikologis adalah elemen kunci dalam prosedur diagnostik [6].

Penderita Demensia dapat diketahui dari hasil cek Magnetig Resonance Imaging (MRI). Namun menurut beberapa penelitian [1], penyebabnya belum pasti. Oleh karena itu dikumpulkan banyak data yang berhubungan dengan pasein penderita Demensia sebagai contoh tingkat pendidikan serta tingkat sosial ekonomi. Dengan semakin banyak data yang dikumpulkan, diharapkan dapat diketahui penyebab Demensiaa dengan bantuan aplikasi computer.

Data Mining atau penambangan data adalah analisis dari set data pengamatan, ser-

ingkali dalam skala besar untuk menemukan hubungan diantaranya dan merangkum data dengan cara baru yang lebih baik dimengerti dan bermanfaat bagi pemilik data [7][8]. Data mining bisa diterapkan pada bidang interdisipliner yang membawa teknik lain dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk mengatasi masalah ekstraksi informasi dari basis data [9].

Penggunaan perangkat lunak aplikasi data mining banyak diterapkan di berbagai bidang. Diharapkan aplikasi data mining dapat digunakan untuk dapat mengubah data menjadi suatu pengetahuan. Banyak pula perangkat lunak untuk data mining yang dapat membantu pekerjaan untuk mendapatkan pengetahuan yaitu Orange, Rapid miner, Weka, Rattle GU dan R Studio. Perbandingan aplikasi yang akan digunakan adalah Weka dan Rapid Miner. Weka adalah aplikasi data mining dengan tampilan yang sederhana, beerbasis Java. Weka memiliki 2 versi, versi GUI dan library. Dikembangkan pertama kali oleh Universitas Waikato. Simbol aplikasi menggunakan burung Weka yang hanya ada di Selandia Baru. Sedangkan Rapid Miner pertama kali dinamai Yet Another Learning Environment atau disingkat YALE. Tahun 2007 diganti nama menjadi RapidMiner. Rapid Miner mampu menganalisa data yang berukuran sangat besar sampai ke level big data.

Metode Penelitian

Data set yang digunakan adalah Open Acess Series of Imaging Studies (OASIS 2) [10] seperti terlihat pada Gambar 1. Dataset OASIS diselenggarakan oleh central.xnat.org, menyediakan akses terbuka ke database yang signifikan dari neuroimaging dan data pencitraan diproses selain data demografi, kognitif, dan spektrum genetik platform mudah diakses untuk digunakan dalam neuroimaging, klinis, dan penelitian kognitif pada penuaan normal dan penurunan kognitif. Tujuan proyek ini adalah untuk membuat dataset neuroimaging otak yang tersedia secara bebas untuk komunitas ilmiah. Menyusun dan mendistribusikan dataset neuroimaging secara bebas, diharapkan dapat memfasilitasi penemuan di masa depan dalam ilmu saraf dasar dan klinis.

OASIS2 terdiri dari koleksi longitudinal dari 150 subjek berusia 60 hingga 96 tahun.

Setiap subjek dipindai pada satu, dua atau lebih kunjungan, dipisahkan oleh setidaknya satu tahun untuk total 373 sesi pencitraan. Untuk setiap subjek, 3 atau 4 individu scan MRI-weighted T1 yang diperoleh dalam sesi pemindaian tunggal disertakan. Subjek semuanya menggunakan tangan kanan dan termasuk pria dan wanita. 72 dari subyek ditandai sebagai tidak terikat selama peneli-64 dari subyek yang dimasukkan ditian. tandai sebagai gila pada saat kunjungan awal mereka dan tetap demikian untuk pemindaian berikutnya, termasuk 51 orang dengan penyakit Alzheimer ringan sampai sedang. 14 poin lain dikarakteristikkan sebagai nondemented pada saat kunjungan awal mereka dan kemudian dikarakterisasi sebagai demensia pada kunjungan berikutnya.

Data klinis dan data demografis pada OA-SIS 2 [11] adalah :

- 1. Usia saat akuisisi gambar dalam tahun.
- 2. Jenis kelamin (M atau F)
- 3. Pendidikan ,tahun pendidikan
- 4. Status sosial ekonomi SES yang dinilai oleh Hollingshead Index of Social Position dan diklasifikasikan ke dalam kategori dari 1 (status tertinggi) hingga 5 (status terendah) [12].
- 5. Skor Pemeriksaan Mini-Mental State MMSE (kisaran dari 0 = terburuk hingga 30 = terbaik) [13].
- 6. CDR Clinical Dementia Rating (0 = tanpa demensia, 0,5 = AD sangat ringan, 1 = AD ringan, 2 = AD sedang) [14].
- Faktor penskalaan ASF Atlas (unitless).Faktor penskalaan terkomputasi yang mengubah otak ruang asli dan tengkorak ke target atlas (mis., penentu matriks transformasi) (Buckner et al., 2004)
- 8. eTIV Diperkirakan volume intrakranial total (cm3) [15].
- 9. nWBV Menormalkan volume seluruh otak, dinyatakan sebagai persen dari semua voxel dalam gambar bertopeng atlas yangdiberi label sebagai materi abuabu atau putih oleh proses segmentasi jaringan otomatis [13].

Jurnal Ilmiah KOMPUTASI, Volume 18 No : 2, Juni 2019, p-ISSN 1412-9434/e-ISSN 2549-7227

	А	В	С	D	E	F	G	н	1	J	к	L	м	N	0	Р
1	Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF	
2	OAS2_0001	OAS2_0001_MR1	Nondemented	1	0	м	R	87	14	2	27	0	1987	0.696	0.883	
3	OAS2_0001	OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2	30	0	2004	0.681	0.876	
4	OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12		23	0.5	1678	0.736	1.046	
5	OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12		28	0.5	1738	0.713	1.010	
6	OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12		22	0.5	1698	0.701	1.034	
7	OAS2_0004	OAS2_0004_MR1	Nondemented	1	0	F	R	88	18	3	28	0	1215	0.710	1.444	
8	OAS2_0004	OAS2_0004_MR2	Nondemented	2	538	F	R	90	18	3	27	0	1200	0.718	1.462	
9	OAS2_0005	OAS2_0005_MR1	Nondemented	1	0	М	R	80	12	4	28	0	1689	0.712	1.039	
10	OAS2_0005	OAS2_0005_MR2	Nondemented	2	1010	М	R	83	12	4	29	0.5	1701	0.711	1.032	
11	OAS2_0005	OAS2_0005_MR3	Nondemented	3	1603	М	R	85	12	4	30	0	1699	0.705	1.033	
12	OAS2_0007	OAS2_0007_MR1	Demented	1	0	М	R	71	16		28	0.5	1357	0.748	1.293	
13	OAS2_0007	OAS2_0007_MR3	Demented	3	518	М	R	73	16		27	1	1365	0.727	1.286	
14	OAS2_0007	OAS2_0007_MR4	Demented	4	1281	М	R	75	16		27	1	1372	0.710	1.279	
15	OAS2_0008	OAS2_0008_MR1	Nondemented	1	0	F	R	93	14	2	30	0	1272	0.698	1.380	
16	OAS2_0008	OAS2_0008_MR2	Nondemented	2	742	F	R	95	14	2	29	0	1257	0.703	1.396	
17	OAS2_0009	OAS2_0009_MR1	Demented	1	0	М	R	68	12	2	27	0.5	1457	0.806	1.205	
18	OAS2_0009	OAS2_0009_MR2	Demented	2	576	М	R	69	12	2	24	0.5	1480	0.791	1.186	
19	OAS2_0010	OAS2_0010_MR1	Demented	1	0	F	R	66	12	3	30	0.5	1447	0.769	1.213	
20	OAS2_0010	OAS2_0010_MR2	Demented	2	854	F	R	68	12	3	29	0.5	1482	0.752	1.184	
21	OAS2_0012	OAS2_0012_MR1	Nondemented	1	0	F	R	78	16	2	29	0	1333	0.748	1.316	
22	OAS2_0012	OAS2_0012_MR2	Nondemented	2	730	F	R	80	16	2	29	0	1323	0.738	1.326	
23	OAS2_0012	OAS2_0012_MR3	Nondemented	3	1598	F	R	83	16	2	29	0	1323	0.718	1.327	
24	OAS2_0013	OAS2 0013 MR1	Nondemented	1	0	F	R	81	12	4	30	0	1230	0.715	1.427	
	Shee	t1 (+)													4	
Sia	0															

Gambar 1: Dataset OASIS2 [10]

Data mining adalah proses menemukan hubungan, pola dan tren baru yang bermakna dengan menyaring banyak data yang disimpan dalam penyimpanan, menggunakan teknologi pengenalan pola, perhitungan statistik dan teknik matematika [16][17]. Data mining disebut juga dengan Knowledge Discovery in Database (KDD) atau pengenalan pola [18]. Ada 4 kelompok Data Mining yaitu model prediksi (prediction modelling), analisis kelompok (cluster analysis), analisis asosiasi (association analysis) dan deteksi anomaly (anomaly detection) [19].

Proses data mining dimulai dari penyeleksian data sumber ke data target, dilanjutkan kemudian ke tahap pemrosesan awal untuk memperbaiki kualitas data, transformasi, data mining diikuti tahap penyajian kembali dan evaluasi yang menghasilkan keluaran berupa pengetahuan baru yang diharapkan memberikan arti yang lebih baik [20].

Berikut ini adalah penjelasan tahapan data mining berdasarkan Gambar 2:

1. Sebelum tahap penambangan informasi dimulai perlu dilakukan seleksi data operasional . Data tersebut digunakan untuk proses data mining, disimpan terpisah dari data operasional dalam suatu berkas.

- 2. Pemrosesan awal /cleaning. Proses pembersihan data perlu dilakukan . Proses pembersihan yang perlu dilakukan adalah menghilangkan duplikasi data, pengecekan data yang inkonsisten, dan mengkoreksi kesalahan data.
- 3. Coding adalah proses perubahan data yang telah dipilih, sampai akhirnya data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses penemuan hal baru dari sesuatu yang sudah ada dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.
- 4. Data mining adalah proses pencarian pola atau informasi yang lebih baik dalam data terpilih dengan menggunakan teknik atau metode atau algoritma yang sangat bervariasi. Metode atau algoritma

yang tepat yang akan dipilih sangat tergantung pada tujuan dan proses pada KDD secara keseluruhan.

5. Perlu ditampilkan pola informasi yang dihasilkan dari proses data mining dalam bentuk yang mudah dimengerti oleh pihak yang membutuhkan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretation. Dimulai dari pemeriksaan apakah pola atau informasi yang ditemukan berlawanan dengan fakta atau hipotesis yang ada sebelumnya.



Gambar 2: Proses dalam Data Mining

Klustering atau klasterisasi adalah metode pengelompokan data. Menurut Tan [21], klustering adalah sebuah proses pengelompokan data ke dalam beberapa bagian atau kluster sehingga data dalam satu kluster memiliki level kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum.

Teknik pengelompokan yang bekerja berdasarkan klustering partisi disebut dengan K-means. Prinsip kerjanya adalah pengelompokan klustering secara hirarki yang dilakukan secara bertahap. Hanya ada satu pemilihan penggabungan suatu item terhadap item lainnya disetiap perulangan dari pengelompokan klustering hirarki [22]. Langkah-langkah melakukan klustering dengan metode K-Means adalah [13]:

- 1. Pemilihan jumlah kluster k.
- 2. Memberi harga awal k pusat kluster bisa dilakukan dengan berbagai cara. Cara

random merupakan cara yang g sering dilakukan. Angka random diberikan sebagai nilai awal pusat kluster.

3. Mengalokasikan semua data/ objek ke kluster terdekat. Jarak kedua obyek menjadi dasar penentuan kedekatan dua objek tersebut. Teori jarak Euclidean digunakan untuk menghitung jarak semua data ke setiap titik pusat kluster yang dirumuskan sebagai berikut:

dimana: D(I,j)=Jarak data ke i ke pusat kluster j Xki = Data ke i atribut data ke k Xkj=Titik pusat ke j pada atribut ke k

4. Mengkalkulasi kembali pusat kluster dengan keanggotaan kluster yang sekarang. Pusat kluster adalah mean dari semua data/objek dalam kluster tertentu. Jika dikehendaki bisa juga menggunakan nilai tengah dari kluster tersebut, tidak hanya rata-rata yang bisa dipakai sebagai ukuran.

$$R_{k=\frac{1}{N_{k}}(X_{1k}+X_{2k}+\dots+X_{nk})}$$
.....(2)

dimana : Rk = Rata-rata baru. Nk = Jumlah data training pattern pada kluster (k). $Xnk = Pola \ ke \ (n) \ yang \ menjadi bagian \ kluster \ (k).$

5. Melakukan lagi tiap objek memakai pusat kluster yang baru. Jika pusat kluster tidak berubah lagi maka proses klustering selesai. Jika belum selesai maka kembali ke langkah nomer 3.

Aplikasi data mining yang digunakan sebagai perbandingan adalah Rapid Miner Studio dan Weka 3.8.3. Pada Gambar 3 terdapat halaman muka Weka 3.8.3 dan Rapid Miner Studio pada Gambar 4. Di masing masing aplikasi akan dijalankan teknik K-Means Clustering. Jurnal Ilmiah KOMPUTASI, Volume 18 No : 2, Juni 2019, p-ISSN 1412-9434/e-ISSN 2549-7227

9	Weka E	xplorer		×
Preprocess Classify Cluster Associat	e Tiselect attributes Tivisualize			
Open tile Open URL	Gene	rate Un	do Edit	Save
Choose None				Apply Stop
Current relation		Selected attribute		
Relation: None Instances: None	Attributes: None Sum of weights: None	Name: None Missing: None	Weight: None Distinct: None	Type: None Unique: None
Attributes				
All None	Invert Pattern			Visuelize All
Status				
Welcome to the Welka Explorer				L00 - X

Gambar 3: Halaman muka Weka 3.8.3

3	<new process=""> - RepidMiner Studio Free & 1.003 @ Springtime</new>	- 8 ×			
Elle Edit Brocess Yew Connections C	ug getings Egensions Help				
🔰 🚍 🖬 • 🕨 •	Views Design Results Auto Model	n 🛷 🔎 All Studio 🔹 Search			
Repository ×	Process ×	Parameters X			
🗘 Add Data 🛛 = 👻	🔮 Process 100% 🖉 🖉 🛃 👹 🔛	Process			
🕨 🛄 Samples	Process	logvertosity nt 🔻 🗊			
 B DG Local Repository (Financial Annalysis) 	Dat wo	logife 📃 🙆 🗈			
Operators ×		_			
dearch for Operation		Change compatibility (3.1.003)			
 Data Access (17) Blending (77) 		Help ×			
Cleansing (26) Medaling (130)		Process			
Scoling (12)		RapidMiner Studio Core			
 Validation (29) Utility (85) 		Synopsis The root operator which is the outer most operator of every process.			
Extensions (17)	Leverage the Wisdom of Crowds to get operator recommendations based on your process design!				
Get more operators from the Harketplace	Activate Vilscolm of Growds	Description			

Gambar 4: Halaman muka Rapid Miner Studio

Pada aplikasi Weka, pilih Explorer lalu Open Dataset tentang OASIS 2 yang memiliki ekstension .artf dan hasilnya seperti Gambar 5. Tampak pilihan Preprocesor, Classify, Cluster, Associate, Select Attributes, Visualize. Jika dilihat hasil pada pilihan Visualize All pada Gambar 5 maka tampilannya menjadi seperti Gambar 6. Kemudian menu Cluster dan pilih SimpleKMeans pada Gambar 7.

Jurnal Ilmiah KOMPUTASI, Volume 18 No : 2, Juni 2019, p-ISSN 1412-9434/e-ISSN 2549-7227

Weka	Explorer		- 8
Preprocess Classify Cluster Associate Select attributes Visualize			
Open file	ierste	Undo Ed	2
ller			
Choose None			Apply Stop
urrent relation	Selected attribute		
Relation: casis_longitudinal_demographics OASIS 2 Attributes: 15 Instances: 373 Sum of weights: 373	Name: is¿SubjectID Nissing: 0 (9%)	Distinct 150	Type: Nominal Unique: 0 (0%)
W-lbades	No Label	Count	Weight
Allabetes	1 0492 0001	2	20
	2 0452 0002	3	30
All None Invert Pattern	3 0452 0004	2	20
	4 0462 0005	3	30
No. Norma	5 0492 0007	2	3.0
HV. Prane	5 0482 0008	2	20
1 Is/S00jectiD	7 0492 0009	2	20
2 🔲 MP0 ID	8 0482 0010	2	20
3 Group	9 0492 0012	2	30
4 Weft	10 0492 0012	2	30
5 MR Delay	10 0102_0010	-	5.5
6 U MP			
7	Class: ASF (Num)		 Visualize
8 Age			
a EDOC			
10 54.5			1
11 L MASE			
12 CDR			
14			
16 ASP		and black with a second	ole a le conde
Remove			
failure .			
OK			Log
UT1			

Gambar 5: File OASIS 2 pada Weka3.8.3

Preprocess Classify Clush	er Associate Select atributes Visualice	Weka	Explorer		- 8 ×
Cpen file	Open URL	Open DB Gen All at	eratoUndo	Edt 🗆 🗙	Save
Current relation Relation: pasis_longitudina Instances: 373	is ¿ Subject ID	MRI ID Too many value to display.	Croup		Apply Otop
All All All Name 1 1.4 (Closed) 2 Mill (Closed) 3 Mill (Closed) 4 Value 5 Mill (Closed) 6 Mill (Closed) 7 Mill (Closed) 8 All 9 Closed) 10 Closed) 11 Closed) 12 All 13 elfV 15 All	100 baday Image: Contract of the second s		Hand		Veralize M
Status CK					Log @ x0

Gambar 6: File OASIS 2 pada Weka 3.8.3

2		Wolca Explicitor	- 🖻 🔜 Falizhatan Witzuwaidana 🔟 - O//
Preprocess	Classify Claster Associate Select attris	utos (Viscalize)	📑 Bagkan
lusterer			P Tenukar
T in wetz	a fusterers	Is-pruning 10008 -min-density 2.0-11 -1 25-12-1.0-N 2-A "weke.core.EuclidearDi	Distance - R first-iset - 1600 - sure Subjudul Peneluna. V D Pith -
	Canopy	erer output	「G Pengeditan ・ 6 ・ ・ - 5 ・ ・ 7 ・ ・
EN			rences fores
	FiberadClusterer HierarchicalClusterer		Perspektur n. n.
	MakeOensilyBasedClusterer SimplekMeans		
	Cluster data using the kimeans algorithm	p -	
	Can use either the Euclidean distance (de	fault) or the Manhatlan distance. If the Manhatlan distance is used, then centroids are o	computed as the component-wise median rather than
	D. Arthur, S. Vassevtski, t-means++ the	advantages of carefull seeding. In: Proceedings of the eighteenth annual ACM-SIAM syn	mposium on Discrete alconitings, 1927-1035, 2007.
2	CAPABILITIES		
	Class - ND class		
	Altibules - Brary attributes, Empty nomin	ial affinitutes, Ulessing values, Nominal affinitutes, Numeric affinitutes, Unary affinitutes	
	Additional	and by harring to	
	Minimum number of instances: 1		
	gos		
tatus			
OK			Log x0

Gambar 7: Pilihan SimpleKMeans

Pada SimpleKMeans, klik kanan pada pil- sesuai kebutuhan. Hasilnya seperti di bawah ihan Choose untuk merubah konfigurasinya ini :

```
riodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10
usterer output
=== Run information ===
                weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-dens
Scheme:
Relation:
                oasis_longitudinal_demographics OASIS 2
                373
Instances:
Attributes:
               15
                 Subject ID
                MRI ID
                Group
                 Visit
                MR Delav
                 M/F
                 Hand
                 Age
                 EDUC
                 SES
                 MMSE
                 CDR
                 eTIV
                 nWBV
                ASF
                evaluate on training data
Test mode:
=== Clustering model (full training set) ===
kMeans
 Number of iterations: 5
 Within cluster sum of squared errors: 1035.2358867508558
 Initial starting points (random):
Cluster 0: 0A52 0D67,0A52 0D67 NR2,Nondemented,2,451,M,R,68,12,4,29,D,1438,0.738,1.
Cluster 1: 0AS2_0032, 0AS2_0032_MR1, Demented, 1, 0, M, R, 90, 12, 3, 21, 0.5, 1307, 0.679, 1.342
Missing values globally replaced with mean/mode
Final cluster centroids:
                                 Cluster#
                 Full Data
Attribute
                                       0
                                                     1
                                  (208.0)
                                               (165.0)
                     (373.0)
 Subject ID
                   OA52_0048
                               OAS2_0070
                                              OAS2_0048
               0A52_0001_MR1_0A52_0001_MR1_0A52_0001_MR2
 MRI ID
                 Nondemented Nondemented
Group
                                              Demented
                       1.882
                                                1.7697
 Visit
                                   1.9712
 MR Delay
                     595.1046
                                  682.375
                                               485.0909
                        r
                                      F
M/E
                                                     м
                                        R
Hand
                           R
                                                     R
                     77.0134
                                  76.9375
                                               77.1091
 Age
 EDUC
                     14.5979
                                  15.2356
                                               13.7939
 SES
                     2.4605
27.3423
                                   2.2692
                                                2.7015
                                  29,1106
 MMSE
                                               25.1132
 CDR
                       0.2909
                                   0.0385
                                                 0.6091
 eTIV
                    1488.1287
                                1464.9904
                                               1517.297
 DWEV
                       0.7296
                                   0.7422
                                                0.7137
 ASE
                       1.1955
                                    1.214
                                                1.1721
Time taken to build model (full training data) : 0.09 seconds
```

--- Model and evaluation on training set ----

Clustered Instances

0 208 (56%) 1 165 (44%) mensia dan kluster1 untuk demensia. Agar yang diinginkan untuk sumbu X dan sumbu Y. dapat melihat visualisai hasil prosesnya pada Gambar 10 sumbu X e-TIV dan Y nya untuk Gambar 8. Dan hasilnya di Gambar 9, dimana sumbu X untuk tingkat CDR dan sumbu Y

Ada 2 kluster yaitu kluster0 untuk non de- nya untuk umur. Hasilnya dapat dipilih atribut umur.

Preprocess Classify Chaster Associate Select				a support a support
	attributes Visualize			
Justerer				
Choose SimpleKMeans -init 0 -max-candidates 1	00-periodic-pruning 1000	0 -min-density 2.0 -†	-1.25 -12 - 1.0 -N	2 -A "weika.core.Eusli
Cluster mode	Clusterer output			
Use training set				
O Supplied test set Set	X=;Subject ID MRI 1D	0AS2_0048 CAS2_0001_NR1	GAG2_0070 DAS2_0001_MR1	0A52_0048 0A52_0001_MR2
O Percentage split % 66	Group	Nondemented	Nondemented	Demented
 Classes to clusters evaluation 	MR. Drlav	555,1046	612.375	400.0000
(Num) ASF T	N/T	7	F	N
Store clusters for visualization	Hand	R	R	R
	Age	77.0134	76.9375	77.1091
	EDOC	14.5979	15.2356	11.7939
Ignero attributes	SES	2.4605	2.2692	2.7015
	MMSE	27.3423	29.1106	25.1132
Start Stop	CDR	0.2505	0.0365	0.6091
Parult list (right click for actions)	eIIV	1408.1267	1454.5504	1817.297
result list (right-click for options)	DAR?	0.7296	0.7422	0.7137
23:30:06 - SimpleKMeans	ASE	1.1955	1,214	1.1721
View in main View in sopa Save reaut Delete neut	window rate window uffer buffer(s)	ild model (ful	L training dat	a) : 01.09 second
Load model Save model Par-realistic Re-apply this	model on current lest set model's configuration	uluation on two	aining set	
Visualize due Visualize bree	ster as signments	3		

Gambar 8: Menu visualisasi hasil K-Means Cluster



Gambar 9: Visualisasi hasil KMeans Cluster OASIS2



Gambar 10: Visualisasi lainnya

Sedangkan pada Rapid Miner, dataset OA-SIS 2 dipanggil dengan menggunakan instruksi Add data pada ikon Design kemudian klik and drag data yang diinginkan pada area Process seperti pada Gambar 11 atau dapat dilakukan dengan klik 2 kali data yang mau diproses.

-3	<new process*=""> – RapidMiner Studio Free 8.1.003 @ Springtime</new>	
File Edit Process Yiew Connections Clou	d Gettings Extensions Help	
	Views: Design Results Auto Nodel	tata, operators
Repository ×	Process ×	
Add Data = 👻	🖲 Process 100% 🖉 🖉 📮 🧟	i 🖝 🗵
anii ^	Precess	
lajani - v1, 8/0/10 3/13 PM - 0 60)	1 in	
ajani - v1, 5/5/18 2:23 PM - 2 k8)		115
antiHendajani - v1, 5/7/19 1:15 PM - 2 kB)	Retrieve onsis_JongAuclinel_demographics OASIS 2	
18 1 @livtetianti-tendejani - v1, 6/7/19 12:63 PM - 3		
HS 1 edict 1 (Histophic Hendajani - VI, Brante 2:28		
sbilly OASIS 1 (FivtatiantiHandaiani - v1. 5/2/19 3:		
rophics OASIS 2 (Elvistiantiliands) and - v1, 8/9/10		
raphics OASIS 2 EDIT (rivitatiantifiendajani - vi, 🗸		
< >		
Operators 🖂		
Search for Operators		
Ensembles (14)		
💡 Update Model		
Croup Models		
8 th Ungroup Models		
√x Create Formula		
- Egmentation (14)		
K-Means		
k-Means (Kernel)	Leverage the Wisdom of Crowds to get operator recommendations based on your process design/	
<	contrage the matching of chords to get operation recommendational based on your process design	
Get more operators from the Marketolace	Activate Wiedom of Crowds	
Drag to move.		

Gambar 11: Data OASIS2 dimasukkan ke Process

8				snew p	process* = R	apidMiner Studi	io Free 8.1.003 (9 Springtime			
Elle Edit Proc	iess Yiew Con	rections Cloud &	ettings Egtension	o <u>H</u> elp		_		2			
🚺 🛅	H -	•		Viewo.	Design	Repuits	Auto Medel		Pindo	data, operatora . o	oto 🔎 Ali s
			ExampleSet (//	Local Repository	/oasis_lon []	demographics 🔾	ASIS 2 mriid etiv a	rsf del)			
Example 6	Set (MLacal Rapor	itory/ossis_longitud	inal_demographic	· CASIS 2)	× 1	ExampleSet (//Lo	cal Repeatory/eas	is_len []_dem	ographics OASIS	2 mrid ativ asf	da) ×
Re	sult History	×			ExampleS	et (Retrieve oasi	is_longitudinal_d	lemegraphics 0	ASIS 2) ×		
	ExampleSet (373 examples, 0 special attributes, 15 regular attributes)								Filter (373/373	examples); al	
Della	NOW NO.	Tag Subject ID	MINI ID	Group	Visit	MR Delay	B/WT	Hand	Age	EDUC	St S
	1	OA32_0001	0A32_0001	Nondemented	1	0	м	R	87	14	2
_	2	OA62_0001	OA62_0001	Nondemonted	2	457	м	R	00	14	2
Σ	3	OAS2_0002	OAS2_0002	Demented	1	0	м	R	75	12	2
Statistics	4	OAS2_0002	OAS2_0002	Demented	2	559	м	R	76	12	?
	5	GAS2_0902	GAB2_0002	Demented	3	1895	м	PR .	80	12	7
	0	0402_0004	0402_0004	Nondemented	1	0	r	R	00	10	3
Charte	7	0482_0004	0482_0004	Nondemonted	2	638	F	R	90	18	3
	8	OAS2_0005	OAS2_0005	Nondemented	1	0	м	R	80	12	4
	8	OAS2_0905	OAS2_0005	Nondemented	2	1010	м	R	83	12	4
Advanced	10	OAS2_0905	0A82_0005	Nondemented	з	1093	м	R	85	12	4
Charts	- 11	0482_0997	0482_0007	Demented	1	0	м	rt.	71	16	7
	12	CAS2_0007	OAS2_0007	Demented	3	512	м	R	72	16	2
1	13	O482_0007	O4S2_0007	Demented	4	1281	м	R	75	16	?
Annotations	14	0482_0008	0482_0008	Nondemented	1	0	r .	R	93	14	z
	15	0402_0000	0402_0000	Nondemented	2	742	<i>n</i>	R	95	14	2
	16	0482 0009	0462 0005	Demented	1	0	м	R	88	12	2

Gambar 12: Hasil data yang sudah dijalankan

Hubungkan bagian keluaran data ke proses kemudian jalankan dengan men klik tanda berikut dan didapat hasil pada Gambar 12. Tampilan data yang dijalankan dapat dipilih pada bagian sebelah kiri berupa statistik, chart dan advanced chart. Gambar 13 adalah pilihan untuk statistik. Sedangkan untuk pilihan chart dapat dilihat tampilannya di Gambar 14. Sumbu X,Y dan Z dapat diset sesuai keinginan.

		< THEN I	ргосица*> - 1	RepidMiner Studio Free 8.1.	003 @ Springtime		
Est Dee	ess year Donestices Cloud I	prings Eglensions Link					
à 😑	H • 🕨 • 🔳	Views	Design	Results Auto In	odeli .	Prod dala. (perators . alt 🛛 👂	AU 55
	1	ExampleSet (ALocal Repository	loasis_lon []	_demographics OAS85 2 mild	etivasf delj 🛛 🖂		
ExampleS	et (M.ocal Repository/cesis_longitu	(inal_demographics ()ABB 2)	× 1	ExampleDet (VLocal Reposito	ny louasia_lon []_dumographi	ics OASIS 2 militiativ asfidet	
Res	alt History 🖂		Example	Set (Rottieve casis_longitudi	nal_demographics OASIS 2	6 ×	
-	Name	H Tope	Mosing	Olatorico	Piller (15	115 all balance Search for Advitudes	7.
Data	 Ind Subject ID 	Polynominal	0	CA52_0184 (2)	OA82_0048 (5)	CAS2_0048 (5), CAS2_0070 (5),	[14
Σ. Tatistics	Y MALID	Polynominal	0	Last OAS2_0186_MR3 (1)	CA82_0001_MR1 (1)	CAS2_6001_MR1 (1). CAS2_000	1,18
-	V Group	Polynominal		Converted (37)	Nondemented (190)	Volume Nondersented (190), Demented (1	146), -
Charle	Visit	Integer	0	1	5	1.862	
Advanced	Y MR Delay	Integer	0	0	2639	595.105	
Charts	₩ MF	Polynominal	0	M (160)	F (213)	P (213), M (160)	
kenelations	¥ Hand	Polynominal	0	R (373)	R (373)	131-46 R (373)	
	4						>
	Creating and adds 1 - 15				Examples	2.272 Openia Altones El Regular Altre	ures 10

Gambar 13: Tampilan statistik setelah data dijalankan

()	an Very Consultant	Churd Salitana Edu	chiji	w process*> – Rapid	Miner Studio Frei	e 8.1.003 @ Springtin	e		
	H • 🕨	•	View	Design	Results A	uto Model			🔎 Ali Shid
ExampleSe Res	d (i/Local Repository/oa	Example sis_kngtudnal_demog	Set (/Local Repusit ophics (JASIS 2)	rykanis_lan()_der × 📑 Ene	regraphice DASES 2 repleSet (NLocal Rep Netfleve casis Law	rniid ein anf delj praitory/sanis_lon[] situdinal demospanisi	× demographics 04 re-04585 28	1883 2 mini atin asıl dağı X	
Deta	Chart style:	. "							•
Σ Statistics	scatter	tube Natyre	$\frac{ \vec{k} \vec{k} }{ \vec{k} \vec{k} }$ Boather Hattice	Sutter 30	Sulter 3D Caller	555 0.404	No.		
Ciurs	Series	Eeres Multiple	BIT Dervey	CO MARKED SOM	Deck.	Denafy	Devidion		
Adapted	Ha. Hatiyan	<u>ила.</u> наоран сам	nen Lein	La Data Datad	HL 	Andrews Duries			
	The last	Coartie Coartie	Casetine College	Duarthe Color Matter	ž.		N		
Anotations	Bildua Distria	58cta 30	Des Co	Dox 3D	Sarteor 30	CD8			20 21
									V

Gambar 14: Tampilan chart

Setelah mencari model yang diinginkan di bagian kiri bawah Operator yaitu Segmentasi sub K-Means. Pilih dengan klik dan drag ke bagian proses atau dengan klik 2 kali seperti Gambar 15. Jalankan model tersebut sampai semua bagian terdapat tanda ceklist.



Gambar 15: Model K- Means pada RapidMiner

Hasilnya dapat dilihat pada Gambar 16 untuk pilihan deskripsi.



Gambar 16: Hasil pilihan deskripsi

Terdapat ada 2 cluster sama seperti jika dijalankan pada Weka. Pada Gambar 17 tampilan pilihan folder, Gambar 18 tampilan pilihan graph, Gambar 19 tampilan Centroid Table dan Gambar 20 tampilan pilihan Plot.

-3					<new pro<="" th=""><th>ocess*> - Ra</th><th>pidMiner Studie</th><th>o Free 8.1.003 @ 5p</th><th>vingtime</th><th></th></new>	ocess*> - Ra	pidMiner Studie	o Free 8.1.003 @ 5p	vingtime	
Ð	ie Edit Proces	a Yee Oos	nections Cloud Se	lings Eglensions	Help					
	2 📁 🛛	- 1	 • 		Views:	Design	Results	Auto Model		
				ExampleSet (//Loc	cal Repairtory/o	nis Jon ()	demographics CA	SIS 2 mrid etiv asf d	4 ×	
				ExampleSet (//Loc	cal Repository/or	nis_les []_	demographics OA	SIS 2 mrid etiv auf d	4 ×	
	Result History	×	Cluster M	odel (Clustering)	×		ExampleSet (//	Local Repository/loss	is_longitudinal_demo	graphics OASIS
^	Description Polder View	P reat	9 3							
ł	Graph									
	Centroid Table									
l	-									
~	1									

Gambar 17: Hasil pilihan folder view

٠.		<n< th=""><th>ew process*> - Rapi</th><th>idMiner Studio Free</th><th>8.1.003 @ Springtime</th><th></th></n<>	ew process*> - Rapi	idMiner Studio Free	8.1.003 @ Springtime						
Eile Ed≵ Broce	ee ⊻ew <u>C</u> onnections Clou	g Settings Extensions Help									
	-	Vi	ews: Design	Results Aut	to Model	Find data. operatorsetc					
📕 ExampleSet (//Local Repository/casis_lon (]_demographics OASIS 2 mild et/z asf del) 🛛 🖂											
ExampleSet (//Local Repository/casis_ion (]_demographics OASIS 2 millid stir asf delj ×											
Result Histor	y 🛛 🖬 Clu	ster Model (Clustering) $ imes$		ExampleSet (//Local R	Repository/casis_longitudinal_demo	graphics OASIS 2) ×					
A	Zeem										
Description	a a										
Folder View	Tree	*									
	Node Labels										
Z	🛃 Edge Labela										
			rest sat								
oraph											
						1					
Table											
				1							
Plot											
			Ŭ								
Annotationa											
V											

Gambar 18: Tampilan hasil dalam bentuk graph

Edd Brocks	is these Consections Cloud Follogs Edge	<new process*=""> -</new>	RapidMiner Studio Free 8.1.003	@ Springtime		
		Views: Design	Results Auto Model	Find data, operators.	. oto 🔎 Ali Sh	
	ExampleSe	(//Local Repository/oasis_lon []_demographics OASIS 3 mrid ativ	asf del) 🛛 🖂		
	ExampleSe	(//Local Repository/oasis_lon [_demographics OASIS 3 mrid etiv	asfdel) ×		
Result History	/ X 📓 Cluster Model (Cluste	ning) ×	Exampleset (vrocal Kepositor)	/oasis_longEumral_demographics (LASIS 2)	X	
Description						
	Attribute	cluster_0		cluster_1		
Folder View	T=2Subject ID	75.454		70.901	70.001	
	MRIID	109.429		179.473	179.473	
	Group	1.821		1.495	1.495	
Creph	Viait	1.504		3.055	3.055	
	MR Delay	294.699		1525.033	1525.033	
	M/F	1.585		1.627		
Centroid Table	Hand	1		4		
	Age	76.394		78.934	78.934	
	EDUC	14.465		15.011	15.011	
Piot	000	2.457		2.374	2.374	
	MMSE	27.124		28.011	28.011	
	CDR	0.317		0.209	0.209	
	eTN	1474.404		1530.559	1530.059	
	nWBV	0.731		0.725	_	
Annotations	ASF	1.206		1.162		

Gambar 19: Tampilan hasil dalam bentuk Centroid Table



Gambar 20: Tampilan hasil untuk Plot

Perbandingan hasil pembentukkan kluster penderita demensia menggunakan aplikasi Weka dan Rapid Miner dapat dilihat pada hasil pemrosesan dengan Weka dan Gambar 19. Dapat diperhatikan nilai-nilai yang tertera pada kolom kluster 0 dan kluster 1, terdapat perbedaan nilai.

Penutup

Setelah melakukan percobaan penerapan K-Means Clustering untuk dataset OASIS 2 pada 2 aplikasi data mining didapatkan bahwa dengan aplikasi Weka yang memiliki tampilan sedehana, lebih mudah untuk memulai menggunakan aplikasi. Jika terdapat banyak persentase error maka perbaikan dilakukan secara manual. Berbeda dengan aplikasi RapidMiner, dimana untuk memulai menggunakan aplikasi terdapat sedikit kesulitan namun untuk hasilnya lebih mudah penggunaannya. Terdapat perbaikan proses secara otomatis pada ikon de-Dari sisi waktu proses tidak terdapat sign. perbedaan diantara keduanya, Namun Rapid-Miner lebih baik karena dapat memproses sampai dengan skala big data dan dengan tampilan lebih baik.

Daftar Pustaka

- Akhila J A, Christin Markose, Aneesh R P, "Feature Extraction and Classification of Dementia with Neural Network", International Conference on Intelligent Computing, Instrumentation and Control Technologies, 2017.
- [2] João Maroco, Dina Silva and etc , "Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification, trees and random forests", BMC Research Notes, 4:299, 2011.
- [3] Ferri CPM, Brayne C, "Global prevalence of dementia, a Delphi consensus study", Lancet Neurology, 366:2112-2117, 2005,.
- [4] Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST:

Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review) - Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001, 56:1133-1142.

- [5] Portet F, Ousset PJ, Visser PJ, Frisoni GB, Nobili F, Scheltens P, Vellas B, Touchon J: Mild cognitive impairment (MCI) in medical practice: a critical review of the concept and new diagnostic procedure. Report of the MCI Working Group of the European Consortium on Alzheimer's Disease, J Neurol Neurosurg Psychiatry, 77:714-718, 2006.
- [6] de Mendonca A, Guerreiro M, Ribeiro F, Mendes T, Garcia C, "Mild cognitive impairment - Focus on diagnosis", Journal of Molecular Neuroscience, 23:143-147, 2004.
- [7] Larose, Daniel T., "Discovering Knowledge in Data An Introduction to Data Mining, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.
- [8] David Hand, Heikki Mannila and Padhraic Smyth, Principles of Data Mining, MIT Press, Cambridge, MA, 2001
- [9] Peter Cabena, Pablo Hadjinian, Rolf Stadler, JaapVerhees, and Alessandro Zanasi, "Discovering Data Mining: From Concept to Implementation", Prentice Hall, Upper Saddle River, NJ, 1998.
- [10] https://www.oasis-brains.org/ , Diakses tanggal 20 April 2019
- [11] Marcus, Daniel S and etc, "Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults", Journal of Cognitive Neuroscience 22:12, pp. 2677–2684, Massachusetts Institute of Technology, 2010
- [12] A. Hollingshead, "Two factor index of social position New Haven", CT: Yale University Press, 1957.
- [13] M.F. Folstein, S.E. Folstein, and P.R. McHung, "Mini-mental state, A practical method for grading the cognitive state of

chiatric Research, 12, 189-198., 1975

- [14] J. C. Morris, "The Clinical Dementia Rating (CDR): Current version and scoring rules", Neurology, 43, 2412–2414, 1993.
- [15] R. L. Buckner, D. Head, J. Parker, A. F. Fotenos, D. S. Marcus and J. C. Morris, et al, "A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: Reliability and validation against manual measurement of total intracranial volume", Neuroimage, 23, 724–738, 2004
- [16] Kiki Fatmawati, Agus Perdana Windarto, "Data Mining : Penerapan Rapidminer dengan K-Means Cluster pada Daerah Terjangkit Demam Berdarah Dengue (DBD) Berdasarkan Provinsi CESS", Journal of Computer Engineering System and Science, p-ISSN :2502-7131, Vol. 3 No. 2, h-173, Juli 2018.
- [17] Charles Jhony Manto Sianturi dan Fina Nasari, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Penyebaran Diare Di Kabupaten Langkat", Cogito Smart Journal, vol. 2, no. 2, pp. 108–119, Desember 2016.

- patients for the clinician", Journal of Psy- [18] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method", Int. J. Artif. Intell. Res., vol. 1, no. 2, pp. 26–33, 2017.
 - [19] S. Mulyati, "Penerapan Data Mining Dengan Metode Clustering Untuk Pengelompokan Data Pengiriman Burung", vol. 1, no. Senatkom, 2015.
 - [20] A. P. Windarto, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering", Techno.COM, vol. 16, no. 4, pp. 348-357, 2017.
 - [21] N. Aggarwal, K. Aggarwal, and K. Gupta, "Comparative Analysis of K-means and Enhanced K-means Clustering Algorithm for Data Mining", Int. Jouranl Sci. Eng. Res., vol. 3, no. 3, 2012.
 - [22] Yuli Siyamto, Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor Di PT. Winstar Batam, Jurnal Media Informatika Budidarma, vol. 1, no. 2, pp. 28-31, Juni 2017.