

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 penuaan 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 pengelolaannya agar dapat diperoleh menjadi pengetahuan, menggunakan data mining. Pengelompokan 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 di antara 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 pasien penderita Demensia sebagai contoh tingkat pendidikan serta tingkat sosial ekonomi. Dengan semakin banyak data yang dikumpulkan, diharapkan dapat diketahui penyebab Demensia 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, berbasis 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 Access 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 penelitian. 64 dari subyek yang dimasukkan ditandai 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 dikarakteristikan sebagai nondemented pada saat kunjungan awal mereka dan kemudian dikarakterisasi sebagai demensia pada kunjungan berikutnya.

Data klinis dan data demografis pada OASIS 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].
7. 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 abu-abu atau putih oleh proses segmentasi jaringan otomatis [13].

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
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	M	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	M	R	80	12	4	28	0	1689	0.712	1.039
10	OAS2_0005	OAS2_0005_MR2	Nondemented	2		1010	M	R	83	12	4	29	0.5	1701	0.711	1.032
11	OAS2_0005	OAS2_0005_MR3	Nondemented	3		1603	M	R	85	12	4	30	0	1699	0.705	1.033
12	OAS2_0007	OAS2_0007_MR1	Demented	1		0	M	R	71	16		28	0.5	1357	0.748	1.293
13	OAS2_0007	OAS2_0007_MR3	Demented	3		518	M	R	73	16		27	1	1365	0.727	1.286
14	OAS2_0007	OAS2_0007_MR4	Demented	4		1281	M	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	M	R	68	12	2	27	0.5	1457	0.806	1.205
18	OAS2_0009	OAS2_0009_MR2	Demented	2		576	M	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

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:

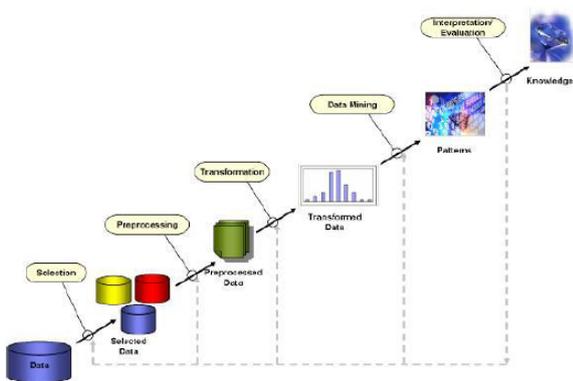
1. Sebelum tahap penambangan informasi dimulai perlu dilakukan seleksi data op-

erasional . 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 kluster 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:

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \dots\dots(1)$$

dimana: D(I,j)=Jarak data ke i ke pusat kluster j
 X_{ki} = Data ke i atribut data ke k
 X_{kj} =Titik pusat ke j pada atribut ke k

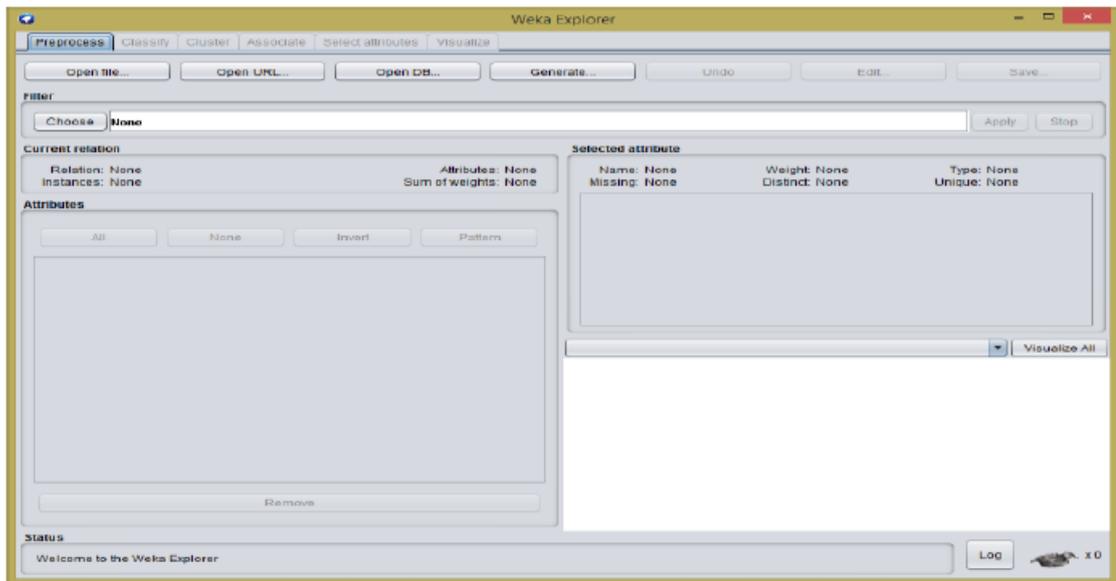
4. Mengkalkulasi kembali pusat kluster dengan keanggotaan kluster yang sekarang. Pusat kluster adalah mean dari semua data/objek dalam kluster tertentu. Jika dikehendaki bisa juga menggunakan nilai tengah dari kluster tersebut, tidak hanya rata-rata yang bisa dipakai sebagai ukuran.

$$R_k = \frac{1}{N_k} (X_{1k} + X_{2k} + \dots + X_{nk}) \dots\dots\dots(2)$$

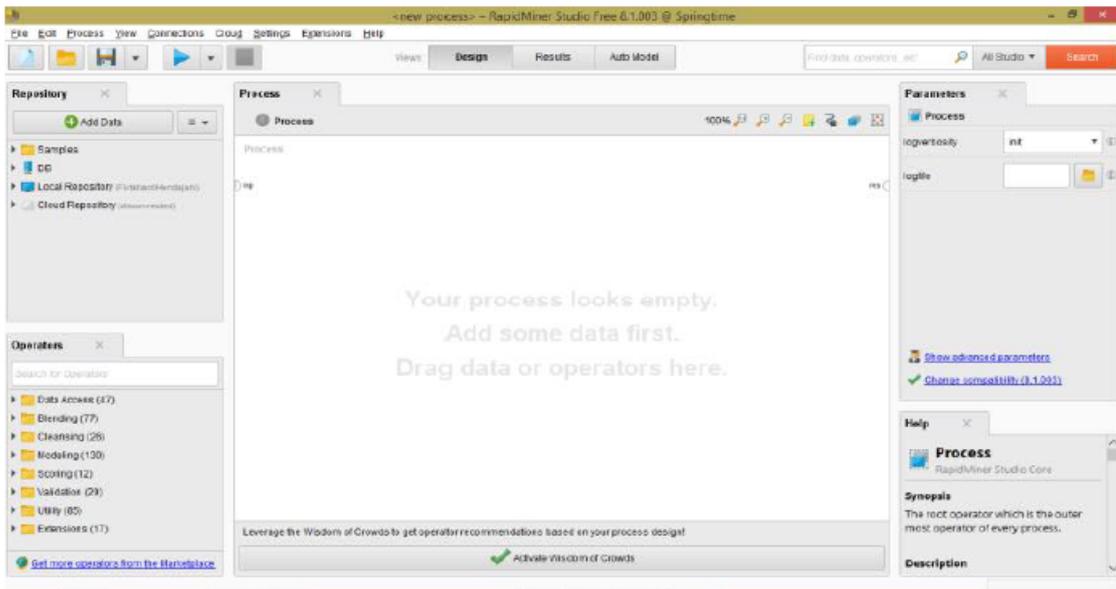
dimana : R_k = Rata-rata baru. N_k = Jumlah data training pattern pada kluster (k). X_{nk} = Pola ke (n) yang menjadi bagian kluster (k).

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 di jalankan teknik K-Means Clustering.



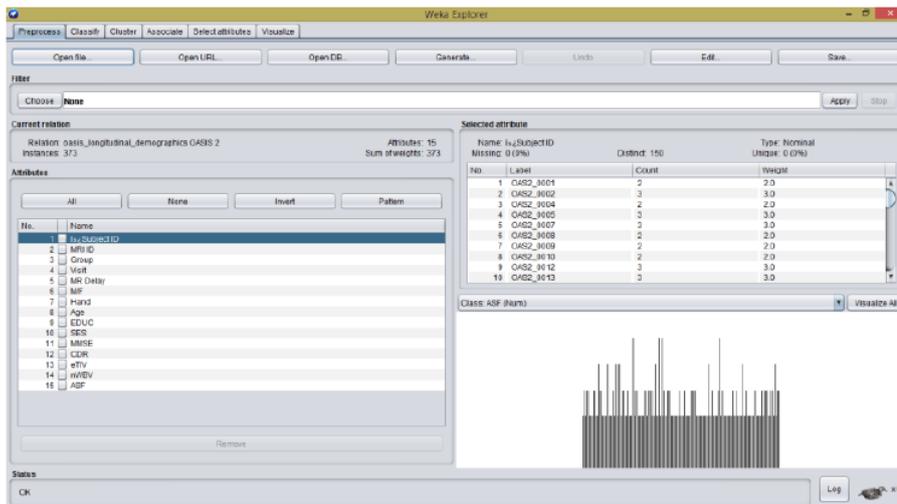
Gambar 3: Halaman muka Weka 3.8.3



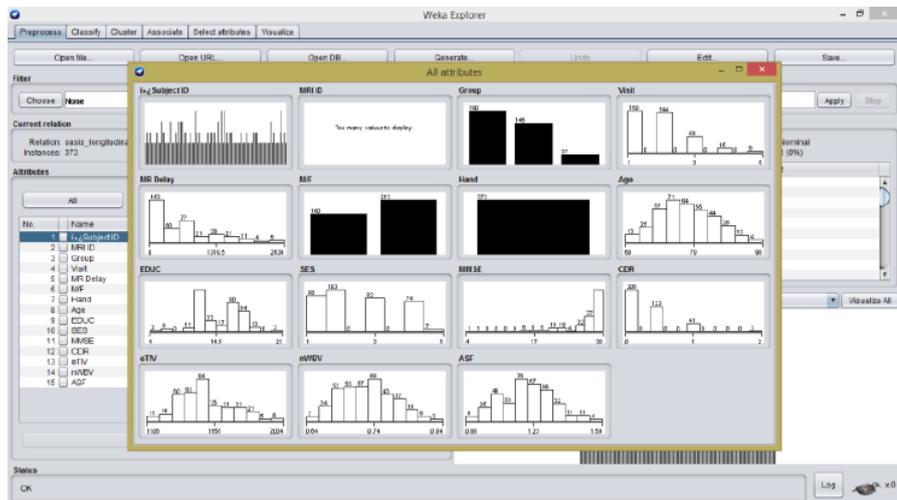
Gambar 4: Halaman muka Rapid Miner Studio

Pada aplikasi Weka, pilih Explorer lalu Open Dataset tentang OASIS 2 yang memiliki ekstension .artf dan hasilnya seperti Gambar 5. Tampak pilihan Preprocessor, Classify, Cluster, Associate, Select Attributes, Visualize. Jika

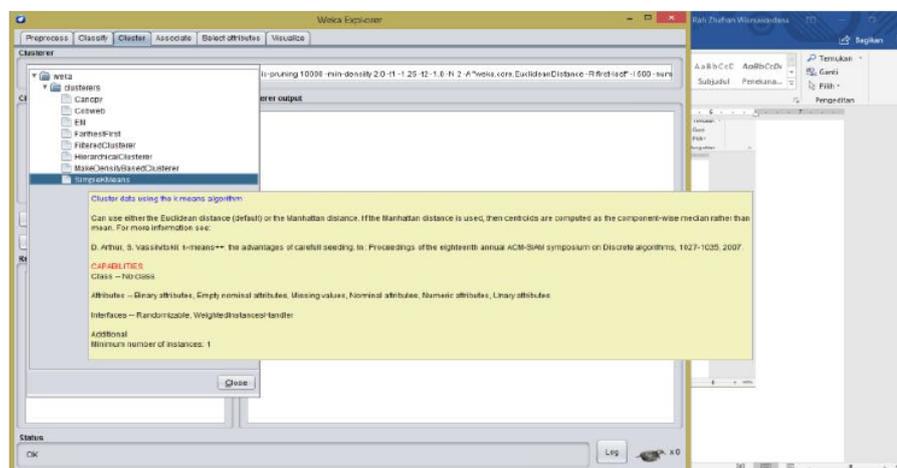
dilihat hasil pada pilihan Visualize All pada Gambar 5 maka tampilannya menjadi seperti Gambar 6. Kemudian menu Cluster dan pilih SimpleKMeans pada Gambar 7.



Gambar 5: File OASIS 2 pada Weka 3.8.3



Gambar 6: File OASIS 2 pada Weka 3.8.3



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

Clusterer output

=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-dens
Relation:    oasis_longitudinal_demographics OASIS 2
Instances:   373
Attributes:  15
             i>Subject ID
             MRI ID
             Group
             Visit
             MR Delay
             M/F
             Hand
             Age
             EDUC
             SES
             MMSE
             CDR
             eTIV
             nWBV
             ASF
Test mode:   evaluate on training data

=== Clustering model (full training set) ===

KMeans

Number of iterations: 5
Within cluster sum of squared errors: 1035.2358867508558

Initial starting points (random):

Cluster 0: OAS2_0067,OAS2_0067_MR2,Nondemented,2,451,M,R,68,12,4,29,0,1438,0.738,1.
Cluster 1: OAS2_0032,OAS2_0032_MR1,Demented,1,0,M,R,80,12,3,21,0.5,1307,0.679,1.342

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute      Full Data      Cluster#
                (373.0)        (208.0)        (165.0)
=====
i>Subject ID   OAS2_0048      OAS2_0070      OAS2_0048
MRI ID         OAS2_0001_MR1 OAS2_0001_MR1 OAS2_0001_MR2
Group          Nondemented    Nondemented    Demented
Visit          1.082          1.9712         1.7697
MR Delay       595.1046       682.375        485.0909
M/F            F              F              M
Hand           R              R              R
Age            77.0134        76.9375        77.1091
EDUC           14.5979        15.2356        13.7939
SES            2.4605         2.2692         2.7015
MMSE           27.3423        29.1106        25.1132
CDR             0.2909         0.0385         0.6091
eTIV           1488.1287      1464.9904      1517.297
nWBV           0.7286         0.7422         0.7137
ASF            1.1955         1.214          1.1721

Time taken to build model (full training data) : 0.09 seconds

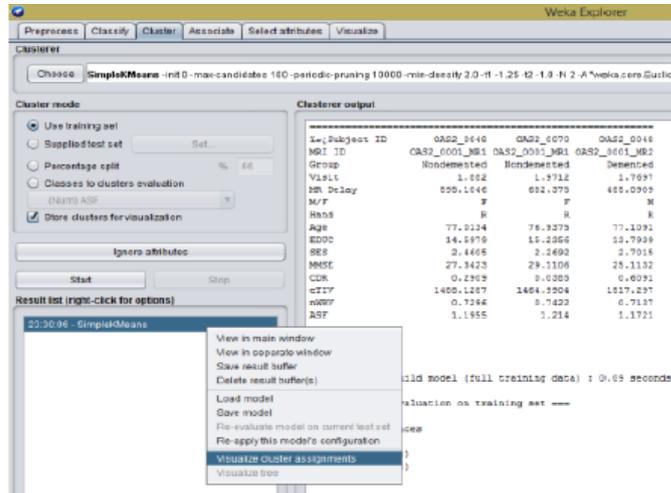
--- Model and evaluation on training set ---

Clustered Instances

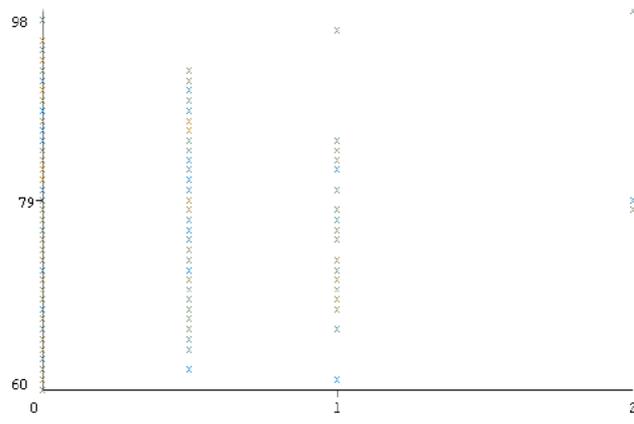
0      208 ( 56%)
1      165 ( 44%)
    
```

Ada 2 kluster yaitu kluster0 untuk non demensia dan kluster1 untuk demensia. Agar dapat melihat visualisasi hasil prosesnya pada Gambar 8. Dan hasilnya di Gambar 9, dimana sumbu X untuk tingkat CDR dan sumbu Y

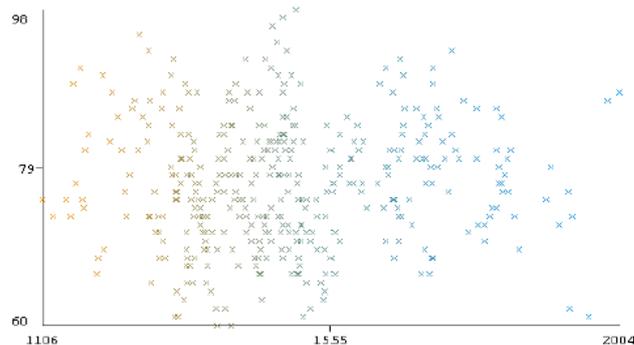
nya untuk umur. Hasilnya dapat dipilih atribut yang diinginkan untuk sumbu X dan sumbu Y. Gambar 10 sumbu X e-TIV dan Y nya untuk umur.



Gambar 8: Menu visualisasi hasil K-Means Cluster

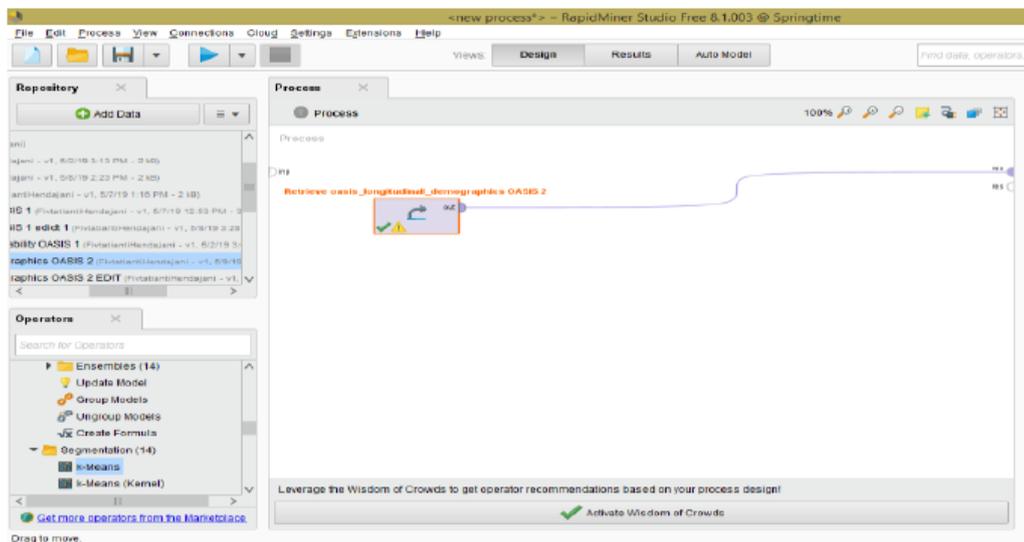


Gambar 9: Visualisasi hasil KMeans Cluster OASIS2



Gambar 10: Visualisasi lainnya

Sedangkan pada Rapid Miner, dataset OASIS 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.



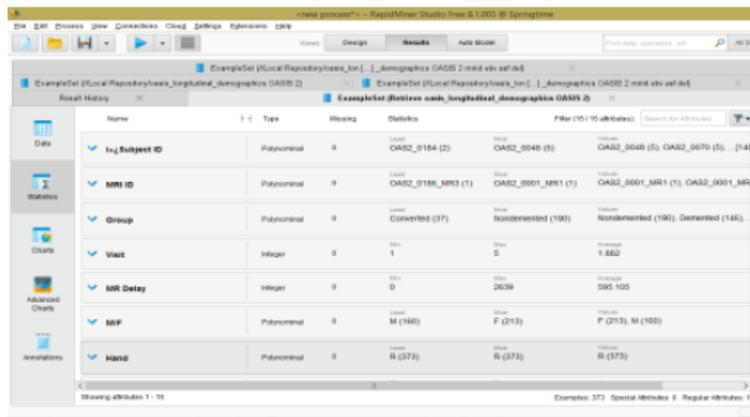
Gambar 11: Data OASIS2 dimasukkan ke Process

Row No.	k33Object ID	BRID	Group	Util	BR Delay	BR	Head	Age	EDUC	SES
1	OAS2_3961	OAS2_3961...	Nondermented	1	0	M	R	87	14	2
2	OAS2_3961	OAS2_3961...	Nondermented	2	457	M	R	88	14	2
3	OAS2_3962	OAS2_3962...	Demented	1	0	M	R	75	12	7
4	OAS2_3962	OAS2_3962...	Demented	2	559	M	R	75	12	7
5	OAS2_3962	OAS2_3962...	Demented	3	1895	M	R	89	12	7
6	OAS2_3964	OAS2_3964...	Nondermented	1	0	F	R	88	10	3
7	OAS2_3964	OAS2_3964...	Nondermented	2	538	F	R	90	18	3
8	OAS2_3965	OAS2_3965...	Nondermented	1	0	M	R	89	12	4
9	OAS2_3965	OAS2_3965...	Nondermented	2	1010	M	R	83	12	4
10	OAS2_3965	OAS2_3965...	Nondermented	3	1093	M	R	85	12	4
11	OAS2_3967	OAS2_3967...	Demented	1	0	M	R	71	10	7
12	OAS2_3967	OAS2_3967...	Demented	3	518	M	R	73	16	7
13	OAS2_3967	OAS2_3967...	Demented	4	1281	M	R	75	16	7
14	OAS2_3968	OAS2_3968...	Nondermented	1	0	F	R	93	14	2
15	OAS2_3968	OAS2_3968...	Nondermented	2	742	F	R	95	14	2
16	OAS2_3968	OAS2_3968...	Demented	1	0	M	R	88	12	7

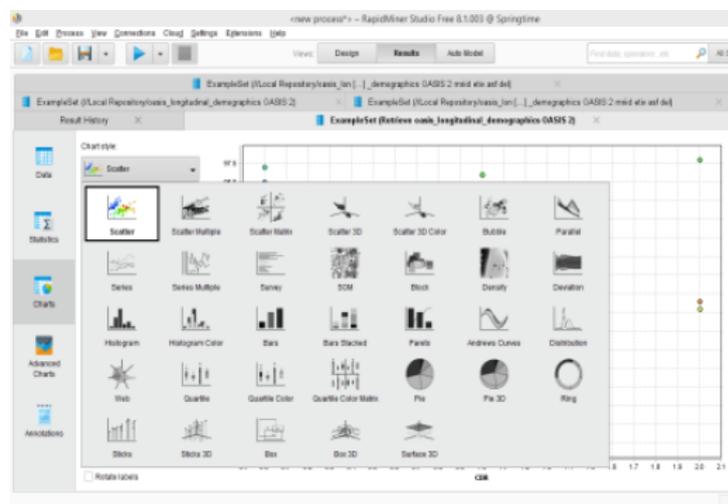
Gambar 12: Hasil data yang sudah dijalankan

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.

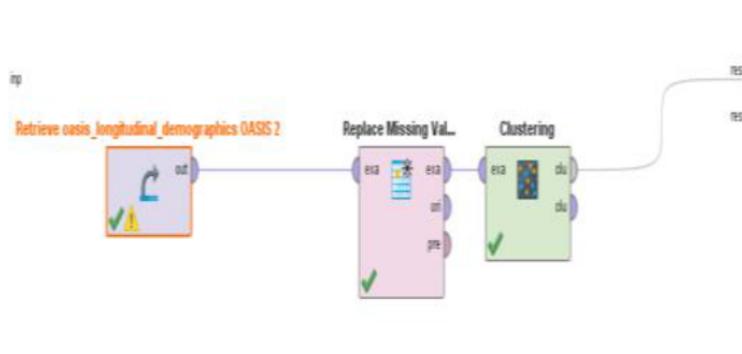


Gambar 13: Tampilan statistik setelah data dijalankan



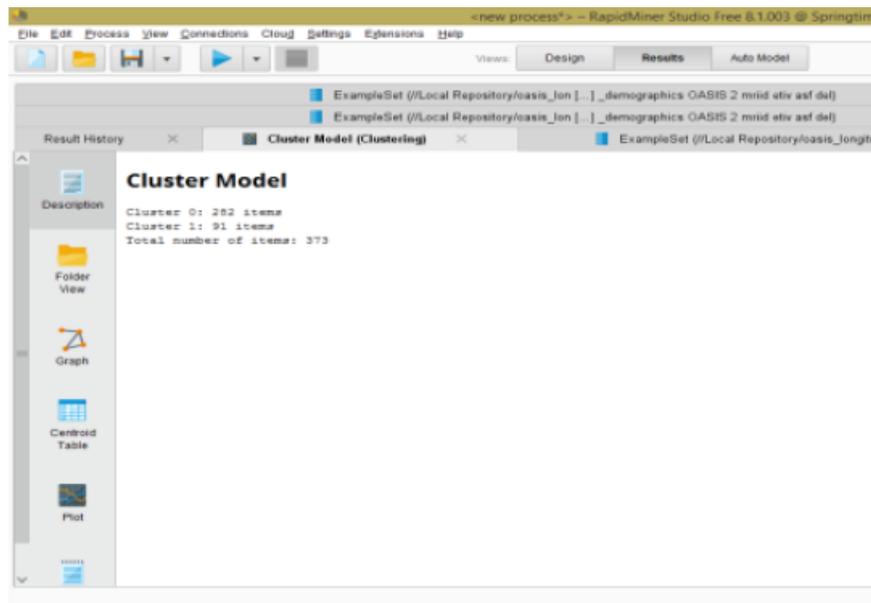
Gambar 14: Tampilan chart

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.



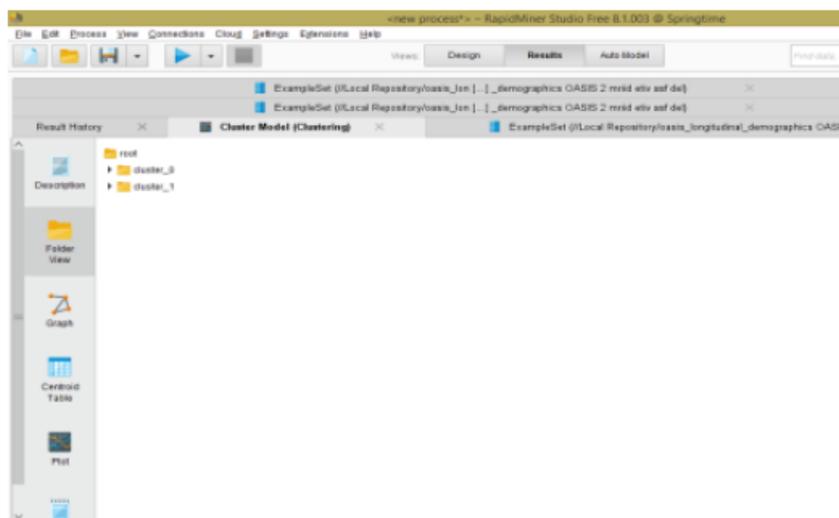
Gambar 15: Model K- Means pada RapidMiner

Hasilnya dapat dilihat pada Gambar 16 untuk pilihan deskripsi.

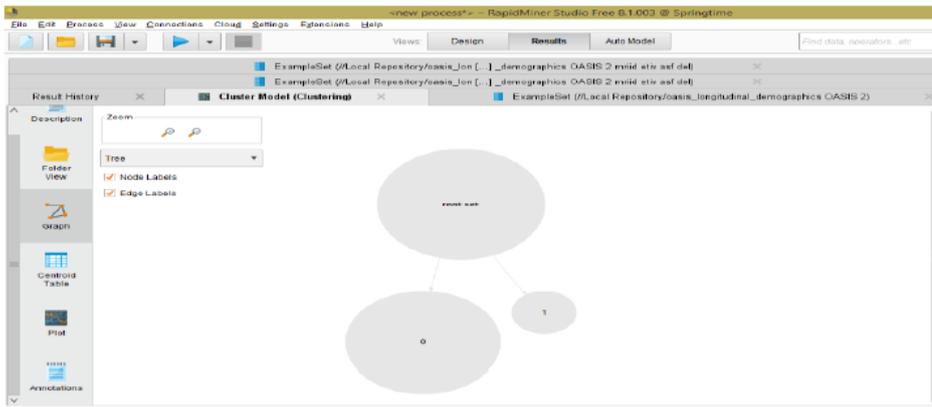


Gambar 16: Hasil pilihan deskripsi

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.



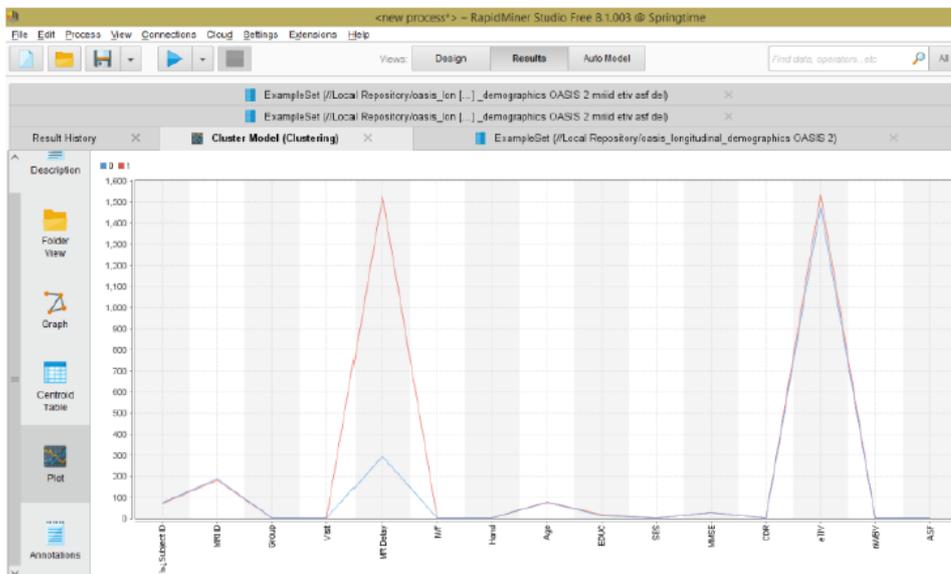
Gambar 17: Hasil pilihan folder view



Gambar 18: Tampilan hasil dalam bentuk graph

Attribute	cluster_0	cluster_1
id_Subject ID	78.484	70.001
MRI ID	108.428	179.473
Group	1.821	1.485
Visit	1.594	2.065
MRI Delay	294.899	1025.033
MIS	1.585	1.627
Hand	1	1
Age	76.394	78.934
EDUC	14.485	15.011
SES	2.457	2.374
MMSE	27.124	28.011
CDR	0.317	0.209
eTV	1474.404	1530.059
nINBN	0.731	0.725
ASF	1.296	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 sederhana, 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 design. Dari sisi waktu proses tidak terdapat perbedaan diantara keduanya, Namun RapidMiner lebih baik karena dapat memproses sampai dengan skala big data dan dengan tampilan lebih baik.

Daftar Pustaka

- [1] 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

- patients for the clinician”, *Journal of Psychiatric 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. Fatenos, 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 Pengelompokan Penyebaran Diare Di Kabupaten Langkat”, *Cogito Smart Journal*, vol. 2, no. 2, pp. 108–119, Desember 2016.
- [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. J. J. 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.