

PENERAPAN ALGORITMA K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA (STUDI KASUS: PUSKESMAS KECAMATAN JAWA MARAJA BAH JAMBI)

Vera Syaputri, Dedy Hartama, Fitri Anggraini, M.Safii, Rafiq Dewa

Program Studi Teknik Informatika, STIKOM Tunas Bangsa Pematangsiantar, Indonesia
 verasyaputri619@gmail.com

ABSTRAK

Gizi pada anak balita merupakan masalah yang sangat penting untuk diperhatikan terutama bagi orang tua dan tenaga kesehatan. Penelitian ini menggunakan teknik data mining yaitu algoritma K-Means untuk mengclustering gizi balita dengan menggunakan 3 cluster yaitu gizi baik, gizi buruk, dan obesitas. Variabel yang digunakan adalah berat badan dan tinggi badan balita. Berdasarkan hasil dari 60 data, jumlah balita yang mengalami status gizi baik pada puskesmas kecamatan jawa maraja pada cluster 0 terdapat 28 balita, pada cluster 1 terdapat 27 balita yang mengalami gizi buruk, dan terdapat 5 balita yang mengalami obesitas pada cluster 2. Diharapkan hasil penelitian ini dapat memberi masukan pada pihak puskesmas agar lebih memperhatikan asupan gizi pada balita sehingga dapat meningkatkan pertumbuhan dan perkembangan balita.

Kata kunci : Status Gizi Balita, Data Mining, Algoritma K-Means

1. PENDAHULUAN

Puskesmas kecamatan jawa maraja merupakan salah satu puskesmas yang ada di desa jawa maraja. Pada puskesmas ini laporan data balita disimpan dalam buku catatan posyandu dan *file digital*. Sehingga pegawai puskesmas sulit untuk mengelompokkan jenis gizi yang dialami balita. Status gizi yang dialami pada balita berbeda-beda setiap bulan nya, maka dari itu diperlukan adanya pengelompokkan status gizi balita. Dalam penelitian ini menggunakan konsep data mining salah satunya yaitu *Algoritma K-Means*. *Algoritma K-Means* merupakan teknik dalam data mining yang dapat mengelompokkan beberapa kelompok berdasarkan jarak, karakteristik, serta kondisi. [1]

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining adalah proses yang menggunakan teknik *statistik*, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat. [2]. *Data mining* adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran untuk menganalisa dan mengekstraks pengetahuan secara otomatis. Adapun tahapan-tahapan proses data mining bersifat *interaktif* dimana user terlibat langsung dengan *knowledge base*. Tahapan tersebut diilustrasikan pada gambar 1 [3].



Gambar 1. Tahapan proses Data Mining

2.2. Algoritma K-Means

K-Means Clustering merupakan salah satu metode *non-hirarki* yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan data yang memiliki karakteristik berbeda akan dikelompokkan dengan kelompok yang lain sehingga data yang berada dalam satu cluster memiliki sifat tingkat variasi yang kecil[4]. *Algoritma K-Means* merupakan teknik yang bekerja dengan berdasarkan pembagia dari beberapa cluster, sehingga prinsip kinerja dari kegiatan pengelompokkan dilakukan secara bertahap [5].

Adapun langkah-langkah dalam algoritma *K-Means* :

1. Tentukan jumlah *cluster*
2. Alokasikan data ke suatu *cluster* secara random dengan rumus sebagai berikut :
3. Hitung *centroid* atau rata-rata dari data yang sudah ada dari masing- masing *cluster*
4. Alokasikan masing-masing data ke *centroid* terdekat
5. Kembali ke step 3, bila ada data yang berpindah cluster maka hitungan centroid di lanjutkan , sedangkan data cluster yang tetap perhitungan di hentikan.

3. METODE PENELITIAN

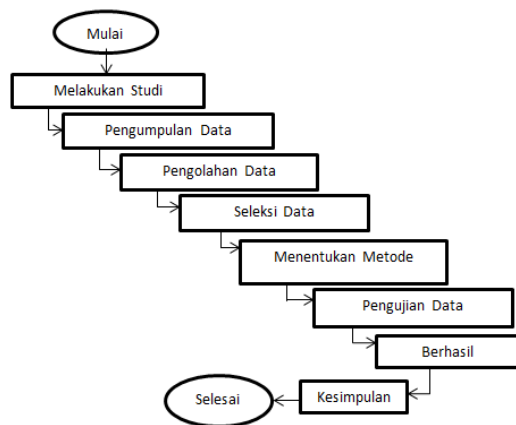
3.1. Identifikasi Masalah

Langkah yang harus dikembangkan dalam algoritma *K-Means* ini adalah dengan mengidentifikasi permasalahan yang akan diteliti yaitu dengan menemukan permasalahan yang akan diteliti sehingga mempermudah langkah selanjutnya.

3.2. Perancangan Penelitian

Rancangan penelitian dilakukan dengan melakukan pengamatan langsung dilapangan untuk mengetahui

pengukuran keakuratan dalam penentuan status gizi balita, kemudian hasil pengamatan akan dilakukan eksperimen. Hasil eksperimen akan dikembangkan menggunakan ilmu pengetahuan yang akan menjadi masukan pemecahan masalah bagi puskesmas kecamatan Jawa Maraja Bah Jambi. Berikut gambar perancangan penelitian dapat dilihat sebagai berikut pada gambar 2.



Gambar 2. Rancangan Penelitian

4. HASIL DAN PEMBAHASAN

4.1. Pengolahan Data

1. menentukan data yang akan diolah.

berikut data balita yang akan diolah dalam penelitian sebagai berikut:

Tabel 1. Data Balita

Balite Ke-	Berat Badan	Tinggi Badan
Balita 1	11,1	85
Balita 2	11	84
Balita 3	16	108
Balita 4	14,1	100
Balita 5	25	110
Balita 6	15	105
Balita 7	20	110
Balita 8	15	102
Balita 9	13	97,5
Balita 10	12	90,5
Balita 11	11	84,4
Balita 12	14,4	95
Balita 13	13	94
Balita 14	18,5	100
Balita 15	13,7	97
Balita 16	14,4	100
Balita 17	16	98
Balita 18	15,8	99
Balita 19	12,3	92,4
Balita 20	12	92
Balita 21	14,2	91,8
Balita 22	10,8	82,9
Balita 23	10,5	82
Balita 24	13,7	88,8
Balita 25	11,1	83,6
Balita 26	14,3	99
Balita 27	14,9	99,1
Balita 28	14,1	97
Balita 29	14	95,6

Balite Ke-	Berat Badan	Tinggi Badan
Balita 30	14,2	94,7
Balita 31	14,9	90,8
Balita 32	11,4	82,1
Balita 33	14,2	101,4
Balita 34	11,8	87,6
Balita 35	12,9	87
Balita 36	15,2	99,4
Balita 37	14,9	100
Balita 38	13,2	98
Balita 39	14,6	99
Balita 40	11,6	92,7
Balita 41	12,5	89,7
Balita 42	13	91,3
Balita 43	13,5	91,7
Balita 44	14,1	91,7
Balita 45	17	98,5
Balita 46	13	89
Balita 47	14	95,7
Balita 48	20,5	101
Balita 49	12	87,6
Balita 50	15	96
Balita 51	12,2	91,4
Balita 52	15	91
Balita 53	15	90,5
Balita 54	14	88,2
Balita 55	13	90
Balita 56	15	94
Balita 57	10,6	82,1
Balita 58	17	99
Balita 59	14	96
Balita 60	16	95

Data balita tersebut belum dapat diolah, maka dari itu data harus diperkecil dengan melakukan normalisasi. Guna normalisasi adalah memperkecil besaran angka dari data tersebut. Berikut rumus normalisasi.

$$= \frac{(\text{Nilai awal} - \text{Nilai minimum})}{(\text{Nilai Maksimum} - \text{Nilai Minimum})}$$

Berdasarkan nilai dari variabel berat badan dan tinggi badan akan dilakukan normalisasi kedalam rentang nilai 0 sampai 1. Berikut hasil perhitungan normalisasi sebagai berikut:

Tabel 2. Normalisasi Data

Balite Ke-	Berat Badan	Tinggi Badan
Balita 1	0,06	0,16
Balita 2	0,05	0,11
Balita 3	0,55	1,37
Balita 4	0,36	0,95
Balita 5	1,45	1,47
Balita 6	0,45	1,21
Balita 7	0,95	1,47
Balita 8	0,45	1,05
Balita 9	0,25	0,82
Balita 10	0,15	0,45
Balita 11	0,05	0,13
Balita 12	0,39	0,68
Balita 13	0,25	0,63

Balite Ke-	Berat Badan	Tinggi Badan
Balita 14	0,8	0,95
Balita 15	0,32	0,79
Balita 16	0,39	0,95
Balita 17	0,55	0,84
Balita 18	0,53	0,89
Balita 19	0,18	0,55
Balita 20	0,15	0,53
Balita 21	0,37	0,52
Balita 22	0,03	0,05
Balita 23	0	0
Balita 24	0,32	0,36
Balita 25	0,06	0,08
Balita 26	0,38	0,89
Balita 27	0,44	0,9
Balita 28	0,36	0,79
Balita 29	0,35	0,72
Balita 30	0,37	0,67
Balita 31	0,44	0,46
Balita 32	0,09	0,01
Balita 33	0,37	1,02
Balita 34	0,13	0,29
Balita 35	0,24	0,26
Balita 36	0,47	0,92
Balita 37	0,44	0,95
Balita 38	0,27	0,84
Balita 39	0,41	0,89
Balita 40	0,11	0,56
Balita 41	0,2	0,41
Balita 42	0,25	0,49
Balita 43	0,3	0,51
Balita 44	0,36	0,51
Balita 45	0,65	0,87
Balita 46	0,25	0,37
Balita 47	0,35	0,72
Balita 48	1	1
Balita 49	0,15	0,29
Balita 50	0,45	0,74
Balita 51	0,17	0,49
Balita 52	0,45	0,47
Balita 53	0,45	0,45
Balita 54	0,35	0,33
Balita 55	0,25	0,42
Balita 56	0,45	0,63
Balita 57	0,01	0,01
Balita 58	0,65	0,89
Balita 59	0,35	0,74
Balita 60	0,55	0,68

2. Menentukan Jumlah Cluster

Pada tahap ini akan diproses dengan memasukkan data kedalam microsoft Excel menggunakan algoritma clustering dan diproses menggunakan tools Rapidminer 5.3 untuk mengelompokkan menjadi 3 cluster.

3. Menentukan Centroid

Untuk menentukan nilai centroid dari data, maka perlu membuat suatu ketentuan bahwa clusterisasi yang dibutuhkan dalam penelitian adalah 3 yaitu dengan memilih nilai titik secara acak. Berikut penetapan nilai centroid awal pada tabel 3 sebagai berikut:

Tabel 3 Nilai Centroid Awal

Cluster	Berat Badan	Tinggi Badan
C1	0,20	0,41
C2	0,36	0,95
C3	0,55	1,37

4. Menentukan Jarak Terdekat

Untuk menghitung jarak antara titik Centroid dengan titik tiap objek menggunakan rumus Eulidian Distance, maka perhitungan untuk jarak terdekat sebagai berikut:

$$\text{Balita1C1} = \sqrt{(0,06 - 0,20)^2 + (0,16 - 0,41)^2} = 0,284238$$

$$\text{Balita2C1} = \sqrt{(0,05 - 0,20)^2 + (0,11 - 0,41)^2} = 0,33541$$

Dan seterusnya sampai data ke-60 pada cluster 1. Hal yang sama dilanjutkan sampai cluster 3. Sehingga didapat jarak terdekat dari masing-masing cluster. Berikut tabel perhitungan pada iterasi 1.

Tabel 4. Perhitungan Iterasi 1

C1	C2	C3	Jarak Terdekat
0,284	0,845	1,31	0,284237815
0,335	0,897	1,36	0,335410197
1,025	0,462	0	0
0,565	0	0,46	0
1,644	1,21	0,91	0,906134831
0,843	0,278	0,19	0,186897694
1,305	0,791	0,41	0,413618583
0,694	0,138	0,33	0,138493077
0,414	0,172	0,63	0,171502243
0,065	0,542	1	0,065367065
0,317	0,878	1,34	0,31672012
0,338	0,265	0,7	0,264862375
0,232	0,334	0,8	0,231773244
0,809	0,44	0,49	0,44
0,403	0,163	0,62	0,16288262
0,574	0,03	0,45	0,03
0,56	0,217	0,53	0,217210341
0,59	0,178	0,47	0,177960903
0,144	0,439	0,9	0,143505769
0,131	0,471	0,93	0,130972286
0,203	0,432	0,87	0,202770971
0,396	0,959	1,42	0,396217923
0,452	1,013	1,47	0,451927237
0,129	0,591	1,04	0,129010726
0,35	0,914	1,37	0,350249614
0,522	0,056	0,5	0,056303491
0,55	0,093	0,48	0,092971863
0,416	0,158	0,61	0,157894737
0,345	0,232	0,68	0,231794756
0,313	0,279	0,72	0,279126556
0,247	0,491	0,91	0,246884184
0,415	0,98	1,44	0,41484937
0,639	0,074	0,39	0,074359686
0,131	0,692	1,15	0,130828386
0,148	0,695	1,15	0,147627592
0,578	0,114	0,46	0,11444313

C1	C2	C3	Jarak Terdekat
0,593	0,08	0,44	0,08
0,442	0,138	0,6	0,138493077
0,533	0,073	0,49	0,072595338
0,182	0,458	0,92	0,181743632
0	0,565	1,02	0
0,098	0,471	0,93	0,097935758
0,145	0,441	0,89	0,145190676
0,192	0,437	0,88	0,191521102
0,646	0,301	0,51	0,300553967
0,062	0,589	1,04	0,062107493
0,35	0,227	0,68	0,226536612
0,997	0,642	0,58	0,581578947
0,121	0,686	1,15	0,121309795
0,415	0,229	0,64	0,228957048
0,094	0,491	0,95	0,09436917
0,259	0,482	0,9	0,259193828
0,254	0,508	0,93	0,253520913
0,17	0,621	1,06	0,169507189
0,052	0,538	0,99	0,052433839
0,337	0,328	0,74	0,328364114
0,443	1,005	1,47	0,442831797
0,665	0,295	0,48	0,294737312
0,364	0,211	0,66	0,210763682
0,448	0,325	0,68	0,324579848

5. Menentukan cluster dari setiap data

Berdasarkan jarak pada pusat cluster. Data yang memiliki jarak terkecil dengan centroid akan menjadi anggota pada kelompok. Terlihat pada tabel 5 dimana posisi data dengan tiap cluster pada iterasi ke-1.

Tabel 5. Data Pengelompokkan Cluster Pada Iterasi 1

Cluster	C1	C2	C3
1	1		
1	1		
3			1
2		1	
3			1
3			1
3			1
2		1	
2		1	
1	1		
1	1		
2		1	
1	1		
2		1	
2		1	
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
2		1	

Cluster	C1	C2	C3
2		1	
2		1	
2		1	
1	1		
1	1		
2		1	
1	1		
1	1		
2		1	
2		1	
2		1	
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
1	1		
2		1	
3			1
1	1		
2		1	
1	1		
1	1		
1	1		
2		1	
1	1		
2		1	
2		1	
2		1	
Jumlah	29	26	5
Total	60		

6. Menghitung Titik Pusat Baru

Menghitung titik pusat baru menggunakan hasil masing-masing cluster. Perhitungan titik pusat pada C1, C2, dan C3.

Untuk menentukan titik pusat baru pada variabel berat badan sebagai berikut :

$$\begin{aligned}
 & \frac{0,6+0,05+0,15+0,05+0,25+0,18+0,15+0,37+0,03+0,00+0,32+0,06+0,44+0,09+0,13+0,24+0,11+0,20+0,25+0,30+0,36+0,25+0,15+0,17+0,45+0,45+0,35+0,25+0,01}{29} \\
 &= 0,20 \\
 & \frac{0,36+0,45+0,25+0,39+0,80+0,20+0,32+0,39+0,55+0,53+0,38+0,44+0,36+0,35+0,37+0,47+0,44+0,27+0,41+0,65+0,35+0,45+0,45+0,65+0,35+0,55}{26} \\
 &= 0,44 \\
 & \frac{0,55+1,45+0,45+0,95+1,00}{5} \\
 &= 0,88
 \end{aligned}$$

Selanjutnya perhitungan iterasi baru untuk nilai tinggi badan.

$$\begin{aligned}
& \frac{0,16+0,05+0,45+0,13+0,63+0,55+0,53+0,52+0,05+0,00+0,36+0,08+0,46+0,01+0,29+0,26+0,56+0,41+0,49+0,51+0,37+0,29+0,49+0,47+0,45+0,33+0,42+0,01}{29} = \\
& 0,34 \\
& \frac{0,95+1,05+0,82+0,68+0,95+0,79+0,95+0,84+0,89+0,89+0,90+0,79+0,72+0,67+1,02+0,92+0,95+0,84+0,89+0,87+0,72+1,00+0,74+0,63+0,89+0,74+0,68}{26} = \\
& 0,84 \\
& = \frac{1,37+1,47+1,21+1,47+1,00}{5} \\
& = 1,31
\end{aligned}$$

Maka data centroid baru iterasi 2 pada dapat dilihat pada tabel 5 sebagai berikut:

Tabel 5. Data Centroid Baru Iterasi 2

Cluster	Berat Badan	Tinggi Badan
C1	0,20	0,34
C2	0,44	0,84
C3	0,88	1,31

Dengan menggunakan langkah-langkah yang sama seperti sebelumnya untuk menentukan jarak centroid dengan menggunakan centroid baru pada iterasi 2. Berikut hasil jarak centroid iterasi 2 pada tabel 6 dibawah ini:

Tabel 6. Hasil Perhitungan Iterasi 2

C1	C2	C3	Jarak Terdekat
0,23184	0,77726	1,410267	0,231837814
0,28058	0,82831	1,459075	0,280577947
1,08478	0,54256	0,335989	0,335989464
0,62667	0,13361	0,63126	0,133609466
1,68518	1,1964	0,594362	0,594361549
0,90425	0,37292	0,440312	0,372915112
1,35729	0,81727	0,182389	0,182388736
0,75363	0,2152	0,498721	0,215198777
0,47733	0,18784	0,7978	0,187838899
0,11873	0,48434	1,126447	0,118729139
0,26315	0,80975	1,44181	0,263150888
0,39127	0,16054	0,791079	0,160537293
0,29461	0,27811	0,922361	0,278111666
0,85146	0,3796	0,366727	0,366726932
0,46379	0,12618	0,76134	0,126181981
0,63488	0,11899	0,606785	0,11899236
0,60999	0,11354	0,568696	0,113541146
0,64352	0,10941	0,539474	0,109410641
0,20775	0,38755	1,0317	0,207746229
0,19274	0,42327	1,067548	0,192744355
0,24227	0,32887	0,939877	0,24226898
0,34037	0,8889	1,518157	0,340366065
0,39641	0,94476	1,574202	0,396408546
0,11882	0,49391	1,100503	0,118817322
0,29349	0,84247	1,47084	0,293492299
0,58167	0,0802	0,64694	0,080201165
0,60755	0,06224	0,598196	0,062242079
0,47551	0,09055	0,73242	0,09054731
0,40295	0,14963	0,792704	0,149628591

C1	C2	C3	Jarak Terdekat
0,36796	0,18203	0,815885	0,18203015
0,26723	0,37471	0,950127	0,267226746
0,3539	0,90183	1,521217	0,353900012
0,70056	0,19491	0,583846	0,194907705
0,08584	0,62365	1,258437	0,085841618
0,08629	0,60737	1,22294	0,086292709
0,63417	0,08481	0,5655	0,084814937
0,65141	0,10957	0,567176	0,109568863
0,50581	0,16659	0,765908	0,166592708
0,59152	0,06277	0,624045	0,062768775
0,24077	0,42671	1,069402	0,240765129
0,06447	0,49304	1,128007	0,064473512
0,15607	0,39518	1,030734	0,156070341
0,19575	0,35466	0,983873	0,195750512
0,23158	0,33616	0,94974	0,231578513
0,69187	0,21564	0,493691	0,215638952
0,055	0,50514	1,12897	0,055004053
0,40786	0,14537	0,788798	0,145366859
1,03472	0,58633	0,328002	0,328002432
0,0698	0,61407	1,246621	0,069801373
0,46703	0,10191	0,712743	0,101905184
0,15728	0,43448	1,077522	0,157278347
0,28098	0,36442	0,936175	0,28097666
0,26953	0,39072	0,959627	0,269533546
0,1483	0,51881	1,113211	0,148298656
0,09327	0,45664	1,085693	0,093270254
0,38188	0,20671	0,799219	0,206714087
0,38682	0,93549	1,564257	0,386822214
0,71214	0,22091	0,470565	0,220910519
0,42261	0,13301	0,777176	0,133012627
0,48859	0,191	0,703283	0,190997159

Pembahasan:

$$\text{Balita1 C1} = \sqrt{(0,06 - 0,20)^2 + (0,16 - 0,34)^2} = 0,23184$$

$$\text{Balita2 C1} = \sqrt{(0,05 - 0,20)^2 + (0,11 - 0,34)^2} = 0,28058$$

Dilanjutkan sampai data ke-60 pada nilai cluster C1, C2, dan C3.

Selanjutnya menentukan cluster masing-masing data balita berdasarkan jarak titik pusat. Berikut hasil pengelompokan posisi tiap data iterasi 2 pada tabel 7 sebagai berikut:

Tabel 7. Hasil pengelompokan Iterasi 2

Cluster	C1	C2	C3
1	1		
1	1		
3			1
2		1	
3			1
2		1	
3			1
2		1	
2		1	
1	1		
1	1		

Cluster	C1	C2	C3
2		1	
2		1	
3			1
2		1	
2		1	
2		1	
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
2		1	
2		1	
2		1	
1	1		
1	1		
2		1	
1	1		
1	1		
2		1	
2		1	
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
1	1		
2		1	
2		1	
2		1	
Jumlah	28	27	5
Total	60		

Selanjutnya menghitung nilai titik pusat baru pada iterasi 3 menggunakan hasil dari setiap anggota masing-masing cluster. Berikut tabel 8 data centroid baru pada iterasi 3:

Tabel 8. Data centroid Baru Iterasi 3

Cluster	Berat Badan	Tinggi Badan
C1	0,20	0,33

C2	0,42	0,84
C3	0,95	1,25

Pembahasan:

Menghitung nilai cluster baru pada variabel berat badan:

$$= \frac{0,06+0,15+0,05+0,18+0,15+0,37+0,03+0,00+0,32+0,06+0,04+0,09+0,13+0,24+0,11+0,20+0,25+0,30+0,36+0,25}{28} = 0,20$$

$$= \frac{0,36+0,45+0,45+0,25+0,39+0,25+0,32+0,39+0,55+0,53+0,38+0,44+0,36+0,35+0,37+0,37+0,47+0,44+0,27+0,41}{26} = 0,42$$

Selanjutnya menghitung nilai cluster baru pada variabel tinggi badan:

$$C1 = \frac{0,16+0,11+0,45+0,13+0,55+0,53+0,52+0,05+0,00+0,36+0,08+0,46+0,01+0,29+0,26+0,56+0,41+0,49+0,51+0,37+0,29+0,49+0,45+0,33+0,42+0,01}{28} = 0,33$$

$$C2 = \frac{0,95+1,21+1,05+0,82+0,68+0,63+0,79+0,95+0,84+0,89+0,89+0,90+0,79+0,72+0,67+1,02+0,92+0,95+0,84+0,89+0,87+0,72+0,74+0,63+0,89+0,74+0,68}{26} = 0,84$$

$$C3 = \frac{1,37+1,47+1,47+0,95+1,00}{5} = 1,25$$

Selanjutnya dilakukan kembali seperti langkah 11 sampai 3. Jika nilai *centroid* tidak sama atau belum *balanced* serta posisi data masih berubah maka proses iterasi berlanjut pada iterasi selanjutnya. Namun jika nilai *centroid* bernilai sama dengan iterasi sebelumnya sudah optimal serta posisi *cluster* data balita tidak mengalami perubahan lagi maka proses iterasi berhenti atau selesai. Berikut tabel 9 data centroid baru iterasi 3:

Tabel 9. Data Centroid Baru Iterasi 3

Cluster	Berat Badan	Tinggi Badan
C1	0,20	0,33
C2	0,42	0,84
C3	0,95	1,25

Menggunakan langkah yang sama seperti sebelumnya untuk menentukan jarak dari centroid dengan menggunakan centroid iterasi 3. Berikut hasil perhitungan iterasi 3 pada tabel 10 sebagai berikut :

Tabel 10. Hasil Perhitungan iterasi 3

C1	C2	C3	Jarak Terdekat
0,344	0,896	1,515	0,343518368

C1	C2	C3	Jarak Terdekat
0,711	0,187	0,625	0,187007893
0,079	0,616	1,261	0,079221338
0,078	0,603	1,218	0,077921413
0,644	0,093	0,586	0,092705986
0,662	0,11	0,594	0,109912673
0,516	0,147	0,794	0,14668234
0,602	0,055	0,648	0,055180031
0,25	0,413	1,087	0,24976293
0,075	0,486	1,132	0,07481544
0,167	0,388	1,036	0,166485
0,206	0,35	0,987	0,205632503
0,24	0,334	0,948	0,240414251
0,701	0,235	0,487	0,235062584
0,062	0,5	1,128	0,062215728
0,418	0,136	0,802	0,136321854
1,043	0,605	0,258	0,257531968
0,062	0,607	1,248	0,062027808
0,477	0,108	0,718	0,1083726
0,167	0,424	1,088	0,167132173
0,288	0,368	0,926	0,287504927
0,275	0,394	0,948	0,275341644
0,149	0,518	1,104	0,149342979
0,103	0,451	1,087	0,103139295
0,391	0,211	0,797	0,211031269
0,377	0,929	1,562	0,376986942
0,721	0,24	0,467	0,239676523
0,433	0,123	0,791	0,122792319
0,497	0,205	0,695	0,205026808

Pembahasan :

$$\text{Balita1C1} = \sqrt{(0,06 - 0,20)^2 + (0,16 - 0,33)^2} \\ = 0,2227$$

Dilanjutkan sampai data ke-60 pada nilai cluster C1, C2, dan C3.

Kemudian menentukan cluster dari masing-masing data pada pusat cluster. Berikut hasil pengelompokan nilai cluster dari setiap masing-masing data pada tabel 11 berikut ini:

Tabel 11 Hasil cluter Pada Iterasi 3

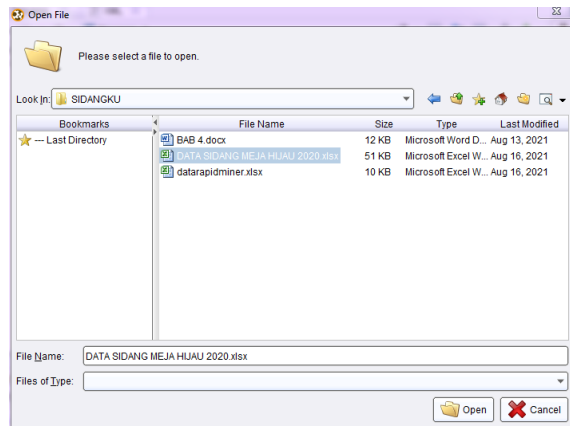
Cluster	C1	C2	C3
1	1		
1	1		
3			1
2		1	
3			1
2		1	
3			1
2		1	
2		1	
1	1		
1	1		
2		1	
2		1	
3			1
2		1	
2		1	
2		1	
2		1	
1	1		

Cluster	C1	C2	C3
1	1		
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
2		1	
2		1	
2		1	
1	1		
1	1		
2		1	
2		1	
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
1	1		
2		1	
3			1
1	1		
2		1	
1	1		
1	1		
1	1		
1	1		
1	1		
2		1	
1	1		
2		1	
2		1	
2		1	
Jumlah	28	27	5
Total	60		

Nilai *centroid* sudah bernilai sama dengan iterasi sebelumnya, serta posisi *cluster* data balita tidak mengalami perubahan lagi maka proses iterasi berhenti atau selesai.

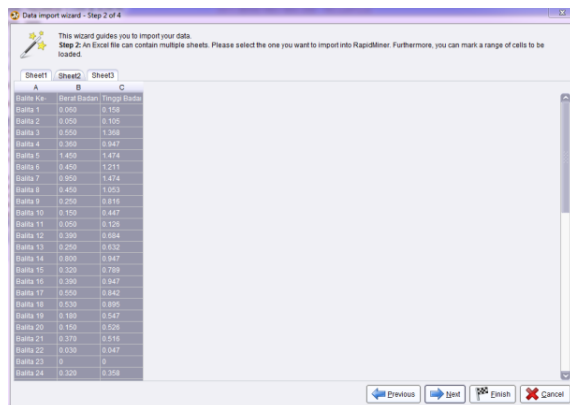
4.2. Penerapan Pada Rapidminer

Pilih data yang akan digunakan untuk diimport. Lakukan *drag and drop* data yang dipilih. Berikut proses import data yang dilakukan.



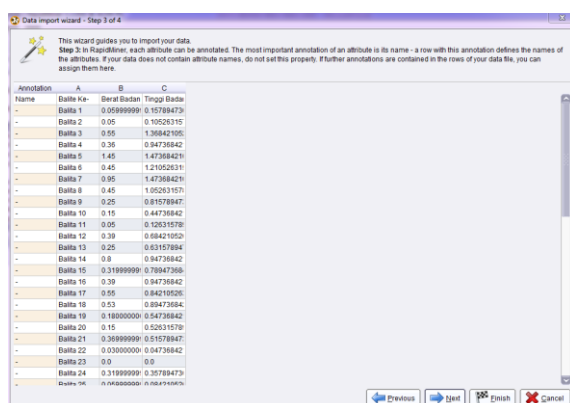
Gambar 3. Import Data Excel

Pada import data *wizard* step 1, pilih data pada folder yang akan di import.



Gambar 4. Alur Proses Import Data Step 2

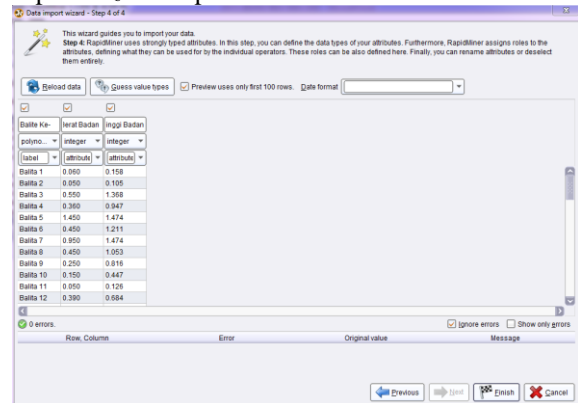
Pada data import *wizard* step 2, data sidang meja hijau 2020.xlsx dengan import excel dimasukkan file excel dapat berisi beberapa lembar, cukup menandai yang akan dimuat. pada proses kali ini peneliti memilih set 2 untuk dilanjutkan prosesnya kemudian klik *next*.



Gambar 5. Alur Proses Import Data Step 3

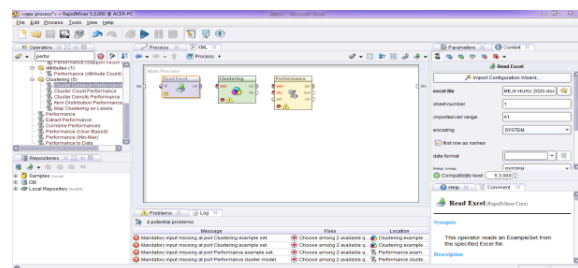
Pada step 3 tidak ada yang dilakukan apapun, setiap *atribut* dapat dianotasi. Yang paling penting pada step 3 adalah namanya mengandung atribut maka dari itu langsung ke data *import wizard* step 4

dengan cara klik *next* maka akan muncul from data import *wizard* step 4.



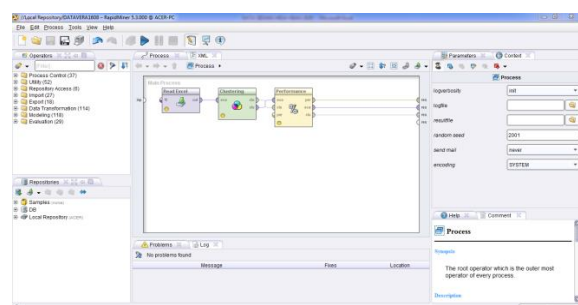
Gambar 6. Alur Proses Import Data Step 4

Setelah ditentukan label yang dipilih, kemudian klik *finish*. Tentukan sebagai tempat untuk menyimpan data pada folder. Selanjutnya *drag and drop* kedalam main proses. Fungsi *validation* ini adalah memaksimalkan nilai akurasi pengolahan data seperti pada gambar 7 dibawah ini:



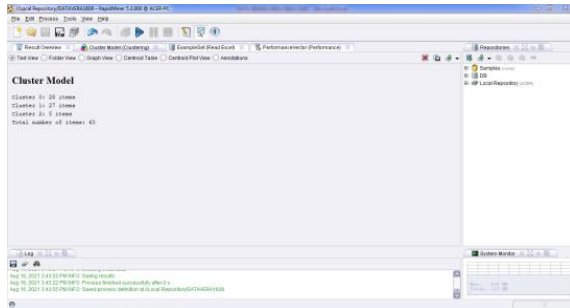
Gambar 7. Drag and Drop data Balita

Langkah selanjutnya *drag and drop* operator *clustering* dan operator *performance* lalu hubungkan pada *output*. Selanjutnya klik *run* yang berada diposisi tengah atas.



Gambar 8. Drag and Drop data Clustering dan Performance

Maka akan menampilkan hasil akhir serta langkah terakhir dalam menggunakan tools Rapidminer. Dapat dilihat pada gambar 9 dibawah ini:



Gambar 9. Data Hasil Clustering

5. KESIMPULAN DAN SARAN

Berdasarkan pembahasan pada penelitian adalah dapat disimpulkan bahwa penerapan algoritma K-means yang telah diterapkan kedalam Rapidmine memiliki hasil yang sama. Yaitu menghasilkan 3 cluster yaitu C1 (Gizi Baik) terdapat 28 data balita, C2 (Gizi Buruk) terdapat 27 data balita, dan C3 (Obesitas) terdapat 5 data balita. Hasil dari penelitian ini diharapkan dapat memberi masukan pada pihak puskesmas agar lebih memperhatikan asupan gizi pada balita sehingga dapat meningkatkan pertumbuhan dan perkembangan balita yang ada di Indonesia.

DAFTAR PUSTAKA

- [1] Buulolo, E. (2020). *Data Mining Untuk Perguruan Tinggi*. Yogyakarta: CV BUDI UTAMA.
- [2] Chandra, M. D., Irawan, E., Saragih, I. S., Windarto, A. P., & Suhendro, D. (2021). Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang. *BIOS : Jurnal Teknologi Informasi dan Rekayasa Komputer*, 1-9.
- [3] Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining Dan Penerapan*. Yogyakarta: CV BUDI UTAMA.
- [4] Siregar, A. M., & Pushpabhuana, A. (2017). *Data Mining Pengolahan Data mining dengan Rapidminer*. Surakarta: CV KEKATA GROUP.
- [5] Toresa, D. (2020). *Implementasi K-Means Terhadap Penyebaran Penyakit Tbc di Riau Menggunakan Rapidminer*. *JUTIM (Jurnal Teknik Informatika MUSIRAWAS) Vol 05 No 01 Juni 2020*, 1-8.