

BAB V

PENUTUP

5.1 KESIMPULAN

Berdasarkan hasil dari penelitian yang dilakukan pada data penderita diabetes melitus maka diambil kesimpulan sebagai berikut:

1. Penelitian ini menggunakan dataset penderita *diabetes melitus* yang bersumber dari website kaggle yang berasal dari Amerika Serikat yang berupa data survei telpon terkait kesehatan terutama *diabetes melitus* yang diberinama BRFSS (*Behavioral Risk Factor Surveillance System*) pada tahun 2014 dengan jumlah data 70.692 data dan 22 atribut.
2. Peneliti menggunakan metode *K-Means Clustering* dengan merekomendasikan pusat *cluster*-nya 4 yang dimana berdasarkan tipe diabetes melitus menurut KemKes (Kementrian Kesehatan Republik Indonesia).
3. Pada Perhitungan manual untuk pengambilan sampel menggunakan tabel Monion & Marrison dengan jumlah 592 data dan atribut yang digunakan 15 atribut sehingga mendapatkan jumlah iterasinya 9, sedangkan pada perhitungan menggunakan *tools* data yang digunakan dataset yaitu 70.692 data dan atribut yang digunakan juga sebanyak 15 dengan jumlah iterasi digunakan yaitu 9 dan max.
4. Dalam melakukan komparasi peneliti menggunakan beberapa perhitungan yang terdapat perbandingan hasil perhitungan manual dengan *tools* data

mining, hal ini terjadi karena adanya perbedaan dalam pemilihan pusat *cluster* yang dipilih secara acak.

Tabel 5. 1 Perhitungan Clustering K-Means

Perbandingan	Data	Jumlah Iterasi	Jumlah Cluster	Jumlah Data tiap Cluster	Centroid	Nilai Ratio	DBI
Perhitungan Manual (Excel)	Data Sampel (593)	9	C1	432	4	73%	0,892
			C2	35		6%	
			C3	43		7%	
			C4	83		14%	
Perhitungan RapidMiner	Data Set (70.692)	9	C1	10.389	4	15%	1,027
			C2	4.749		7%	
			C3	41.155		58%	
			C4	14.399		20%	
		Max	C1	38.623		55%	1,038
			C2	16.832		24%	
			C3	8.672		12%	
			C4	6.565		9%	
Perhitungan WEKA	Data Set (70.692)	9	C1	8.557	4	12%	-
			C2	20.248		29%	
			C3	13.175		19%	
			C4	28.712		41%	
		Max	C1	8.647		12%	-
			C2	20.230		29%	
			C3	13.120		19%	
			C4	28.695		41%	
Perhitungan SPSS	Data Set (70.692)	9	C1	12.034	4	17%	3,892
			C2	24.147		34%	
			C3	20.269		29%	
			C4	14.242		20%	
		Max	C1	12.034		17%	3,892
			C2	24.147		34%	
			C3	20.269		29%	
			C4	14.242		20%	
Perhitungan Python	Data Set (70.692)	9	C1	38.619	4	55%	1,308
			C2	6.566		9%	
			C3	16.832		24%	
			C4	8.675		12%	
		Max	C1	38.632		55%	1,309
			C2	8.669		12%	
			C3	16.834		24%	

			C4	6.566		9%	
--	--	--	----	-------	--	----	--

Pada tabel 5.1 menunjukkan perbedaan perhitungan secara manual dan tools data mining. Perhitungan manual memiliki DBI terbaik dari pada perhitungan dengan manual namun data yang dipakai hanya sampel sedangkan untuk DBI terbaik adalah rapid miner. Perhitungan manual yang memiliki nilai ratio paling tinggi pada cluster 1, pada rapidminer yang memiliki nilai ratio paling tinggi 9 iterasi terletak pada cluster 3 dan maximal terletak pada cluster, sedangkan weka yang memiliki nilai ratio paling tinggi pada 9 iterasi dan maximal sama yaitu cluster 4, seperti weka nilai ratio yang paling tinggi di spss sama baik dari 9 iterasi maupun max iterasi yaitu cluster 2 dan pada phyton nilai ratio tertinggi pada 9 iterasi dan max iterasi sama yaitu cluster. Jadi dapat ditarik kesimpulan bahwa nilai iterasi pada tools weka, spss, phyton tidak memiliki perubahan yang signifikan dari segi jumlah maupun ratio yang dihasilkan namun pada rapidminer terjadi perubahan baik dari segi ratio dan jumlah cluster yang dihasilkan.

5.2 SARAN

Analisa perhitungan penyakit diabetes melitus ini masih jauh dari kata sempurna sehingga diperlukan perbaikan dan pengembangan, maka terdapat beberapa saran yang dapat diberikan untuk mengembangkan lebih lanjut, yaitu:

1. Diperlukan adanya kegiatan pengumpulan data evaluasi secara berkala sebagai bahan input dari sistem pengelompokkan data penderita *diabetes melitus*
2. Diharapkan dalam penelitian selanjutnya dapat menggunakan metode dan algoritma lain dalam menganalisa data.

3. Diharapkan penelitian selanjutnya dapat menggunakan data penderita *diabetes melitus* yang lebih banyak lagi serta mencakup faktor lainnya agar memiliki persentasi akurasi yang lebih baik.
4. Diharapkan untk selanjutnya peneliti mambuat sistem untuk perhitungan *clustering K-means*.