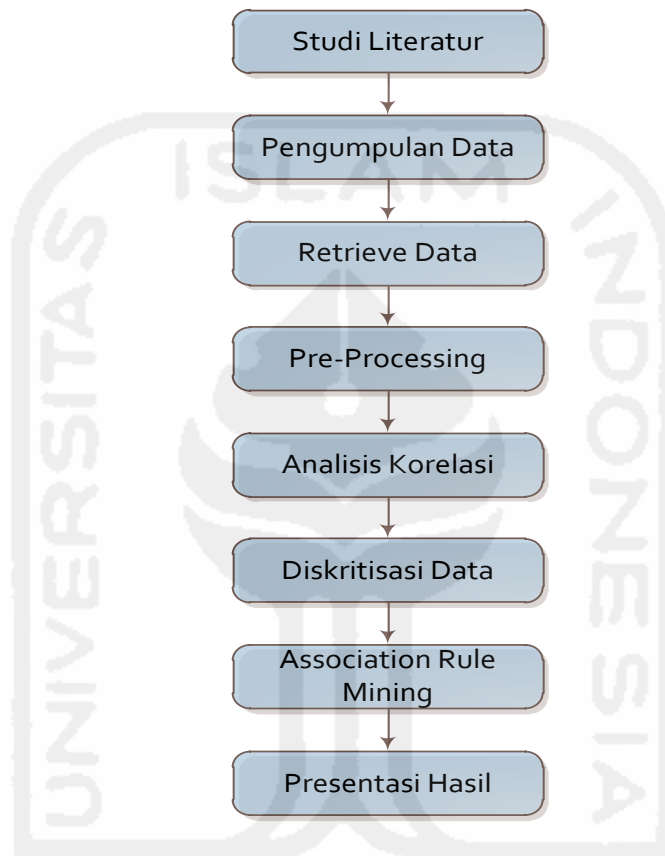


BAB III

METODELOGI PENELITIAN

Dalam menyelesaikan penelitian ini dibuat beberapa tahapan. Tahap-tahap kegiatan dijelaskan dalam Gambar 3.1



Gambar 3. 1. Tahap-Tahap Penelitian

3.1. Studi Literatur

Pada tahap awal dilakukan studi literatur yang bertujuan untuk menjelaskan kajian pustaka dari teori-teori penunjang yang mendukung konstruksi penelitian. Kegiatan ini dilakukan dengan membaca buku, jurnal, artikel laporan penelitian, dan situs-situs di internet. Keluaran dari studi literatur ini adalah terkoleksinya referensi yang relevan dengan rumusan masalah. Tujuannya adalah untuk memperkuat permasalahan serta sebagai dasar teori dalam melakukan studi dan juga menjadi dasar untuk melakukan penelitian.

3.2. Pengumpulan Data

Pengumpulan data adalah prosedur yang sistematis dan standar untuk memperoleh data yang diperlukan. Selalu ada hubungan antara metode atau teknik pengumpulan data dengan masalah, tujuan dan hipotesis penelitian. Pengumpulan data dilakukan dengan observasi dokumen dari laporan yang ada di Dinas Kesehatan Kota Banjarbaru, Badan Pusat Statistik Kota Banjarbaru, Badan Meteorologi Klimatologi dan Geofisika (BMKG) Kelas I Stasiun Klimatologi Banjarbaru dan Badan Meteorologi Klimatologi dan Geofisika (BMKG) Kelas II Stasiun Klimatologi Bandara Syamsudin Noor Banjarmasin. Data yang dikumpulkan yaitu curah hujan, kelembaban, suhu udara, kepadatan penduduk, fasilitas kesehatan, usia dan kejadian DBD di Kota Banjarbaru.

3.3. Retrieve Data

Tahap retrieve data merupakan tahap pengambilan data atau tahap penempatan data. Pada tahap pemutakhiran, data yang tersimpan diperbaharui dan disesuaikan dengan peristiwa terbaru. Kemudian pada tahap retrieving, data yang tersimpan diakses dan diringkas kembali untuk diproses lebih lanjut atau untuk keperluan pembuatan laporan. Data yang disimpan pada tahap ini yaitu data curah hujan, kelembaban, suhu udara, kepadatan penduduk, fasilitas kesehatan, usia dan kejadian DBD di 5 Kecamatan yang terdapat pada Kota Banjarbaru.

3.4. Pre-Processing

Tahap *pre-processing* merupakan tahap mentransformasikan data mentah ke format yang sesuai untuk analisis. Proses ini bertujuan untuk pembersihan data, dimana informasi yang tidak dibutuhkan dibuang. Teknik atau metode yang digunakan dalam data preprocessing, diantaranya:

- **Data cleaning** : Menghilangkan nilai-nilai data yang salah, memperbaiki kekacauan data dan memeriksa data yang tidak konsisten.
- **Data integrasi** : Menggabungkan data dari beberapa sumber (database, data cube, atau file) ke dalam penyimpanan data yang sesuai.

- **Data transformasi** : Normalisasi dan pengumpulan data sehingga menjadi sama.
- **Data reduksi** : Menguraikan data ke dalam bentuk yang lebih kecil ukurannya tetapi tetap menghasilkan hasil analitis yang sama.

3.5. Analisis Korelasi

Analisis korelasi merupakan tahap untuk mengetahui ada tidaknya hubungan antar 2 variabel atau lebih. Dalam melakukan analisis korelasi langkah awal yang perlu dilakukan adalah mengidentifikasi variabel yang hendak dikorelasikan. Identifikasi dilakukan untuk menentukan posisi variabel apakah sebagai variabel X (pengaruh) dan variabel Y (terpengaruh). Pengujian analisis korelasi biasanya menggunakan nilai signifikan dengan kriteria sebagai berikut :

- Jika angka signifikansi hasil riset $< 0,05$, maka hubungan kedua variabel signifikan.
- Jika angka signifikansi hasil riset $> 0,05$, maka hubungan kedua variabel tidak signifikan.

Pada analisis korelasi dilakukan untuk mengetahui distribusi kejadian DBD secara spasial dan temporal. Analisis akan dilakukan setiap parameter yang dapat mempengaruhi distribusi kejadian DBD. Parameter-parameter tersebut berupa curah hujan, kelembaban udara, suhu, kepadatan penduduk, fasilitas kesehatan, usia penderita DBD dan angka kejadian DBD. Analisis spasial dan temporal dilakukan dengan menggunakan teknik dimensi waktu dan wilayah. Dimensi waktu dalam menganalisis distribusi kejadian DBD akan dibagi menjadi 2 jenis yaitu periode Bulan dan periode tahun. Sedangkan untuk wilayah akan dilakukan analisis per kecamatan dan Kotamadya Banjarbaru.

Dengan menggunakan data sementara dimana curah hujan, kelembaban udara dan suhu sebagai variabel X (pengaruh), kemudian kejadian DBD sebagai variabel Y (terpengaruh) yang didapatkan dari Dinas Kesehatan Kota Banjarbaru dapat di analisis korelasi hubungan antar variabel (lihat gambar 3.2).

Correlations

		Curah Hujan_X1	Kelembaban_Udara_X2	Suhu_X3	DBD_Y	
Spearman's rho	Curah_Hujan_X1	Correlation Coefficient	1.000	.900 ^{**}	.333	.700 [*]
		Sig. (2-tailed)		.001	.381	.036
		N	9	9	9	9
	Kelembaban_Udara_X2	Correlation Coefficient	.900 ^{**}	1.000	-.017	.733 [*]
		Sig. (2-tailed)	.001		.966	.025
		N	9	9	9	9
	Suhu_X3	Correlation Coefficient	.333	-.017	1.000	.200
		Sig. (2-tailed)	.381	.966		.606
		N	9	9	9	9
	DBD_Y	Correlation Coefficient	.700 [*]	.733 [*]	.200	1.000
		Sig. (2-tailed)	.036	.025	.606	
		N	9	9	9	9

** Correlation is significant at the 0.01 level (2-tailed).

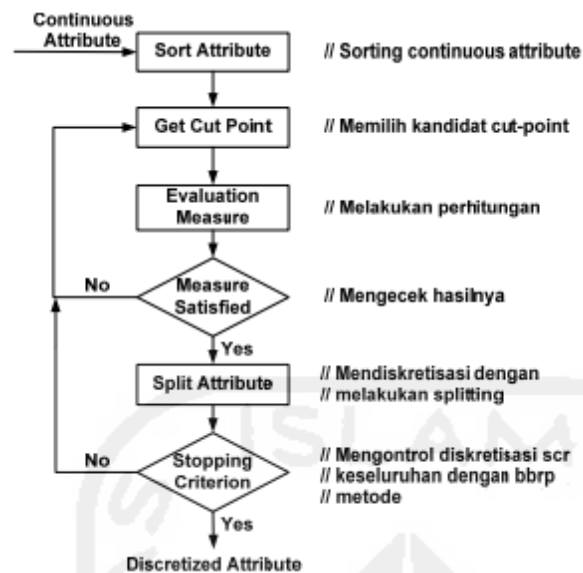
* Correlation is significant at the 0.05 level (2-tailed).

Gambar 3. 2. Analisis Korelasi dengan SPSS

Pengujian dengan menggunakan program SPSS pada gambar 3.2. menunjukkan output hubungan antar variabel X (pengaruh) dan variabel Y (terpengaruh). Variabel curah hujan dengan variabel kejadian DBD memiliki nilai signifikan $0,036 < 0,05$ dan variabel kelembaban udara dengan variabel kejadian DBD memiliki nilai signifikan $0,025 < 0,05$. Maka dapat disimpulkan bahwa curah hujan dan kelembaban udara mempunyai hubungan atau korelasi yang signifikan terhadap kejadian DBD. Sedangkan variabel suhu dengan variabel kejadian DBD memiliki nilai signifikan $0,6 > 0,05$. Hal ini menyatakan bahwa variabel suhu tidak ada korelasi terhadap kejadian DBD. Hasil analisis korelasi ini menunjukkan faktor iklim memberikan pengaruh terhadap penyebaran penyakit DBD di Kota Banjarbaru.

3.6. Diskretisasi Data

Tahap diskretisasi data merupakan teknik untuk mereduksi sekumpulan nilai yang terdapat pada atribut continuous, dengan membagi range dari atribut ke dalam interval.



Gambar 3. 3. Proses Diskretisasi (Junaedi dkk, 2011).

1. Sorting, melakukan sorting nilai atribut continuous yang mau didiskretisasi.
2. Memilih “cut-point”, banyak fungsi evaluasi yang dapat digunakan seperti binning dan pengukuran entropy.
3. Splitting, dilakukan evaluasi cut-point yang ada dan pilih satu yang terbaik dan lakukan split range nilai atribut continuous ke dalam dua partisi. Diskretisasi berlanjut untuk tiap partisi sampai kondisi berhenti tercapai.
4. Stopping criterion, diperlukan untuk menghentikan proses diskretisasi. (Junaedi dkk, 2011).

Pada penelitian Mukhlash (2011), sebagai contoh hasil diskretisasi data dari parameter yang mempengaruhi kejadian penyakit DBD sebagai berikut :

Tahun	Kode Kelura...	Nama Kelurahan	Kepadatan	Penderita Pe...	Fasilitas Kese...	Penderita Tet...
2000	3578090002	PutatJaya	Tinggi	Rendah	Ada	Sedang
2000	3578090003	Banyu Urip	Sedang	Rendah	Ada	Sedang
2000	3578090004	Kupang Krajan	Sedang	Rendah	Ada	Sedang
2000	3578090005	Petemon	Rendah	Rendah	Ada	Sedang
2000	3578090006	Sawahan	Rendah	Sedang	Ada	Sedang
2000	3578100001	Tembok Dukuh	Sedang	Rendah	Ada	Sedang
2000	3578100002	Bubutan	Rendah	Rendah	Tidak Ada	Sedang
2000	3578100003	Alun-alunContong	Rendah	Sedang	Tidak Ada	Rendah
2000	3578100004	Gundih	Sedang	Rendah	Tidak Ada	Rendah
2000	3578100005	Jepara	Sedang	Rendah	Ada	Rendah
2000	3578110001	EmbongKallasin	Rendah	Rendah	Ada	Tinggi
2000	3578110002	Ketabana	Rendah	Sedang	Ada	Tinggi

Gambar 3. 4. Contoh hasil diskretisasi data (Mukhlash, 2011).

Pada penelitian ini, diskretisasi akan dilakukan menggunakan teknik statistik dengan mencari jumlah kelas dan interval atau lebar kelas. Pada setiap parameter akan mempunyai jumlah kelas dan interval atau lebar kelas masing-masing. Adapun langkah-langkah mencari jumlah kelas dan lebar kelas dapat dilakukan dengan menggunakan rumus berikut :

1. Menentukan range (R)

$$R = \text{Data terbesar} - \text{Data terkecil} \quad (3.1)$$

2. Menentukan banyak kelas yang akan dibuat (K)

$$K = 1 + 3,33 \log N \quad (3.2)$$

Dimana : K = Banyak Kelas

N = Banyaknya Data

3. Menentukan panjang interval

$$I = R / K \quad (3.3)$$

Dimana : I = Interval Kelas

R = Range

K = Banyak kelas

3.7. Association Rule Mining

Tahap *association rule mining* merupakan teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Aturan asosiatif antara suatu kombinasi item akan diperoleh dari parameter curah hujan, kelembaban, suhu udara, kepadatan penduduk, fasilitas kesehatan, usia dan kejadian DBD di Kota Banjarbaru dengan data yang sudah di diskretisasi. Fungsi *association rule mining* digunakan untuk menemukan relasi atau korelasi diantara himpunan item.

Association rule mining dalam penyelesaiannya dibagi menjadi 2 tahap yaitu :

1. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support sebuah item diperoleh dengan rumus berikut:

$$\text{support}(A) = \frac{\text{Jumlah Transaksi mengandung } A}{\text{Total Transaksi}} \quad (3.4)$$

Sedangkan untuk nilai support dari 2 item diperoleh dari rumus berikut:

$$\text{support}(A \cap B) = \frac{\text{Jumlah Transaksi mengandung } A \text{ dan } B}{\text{Total Transaksi}} \quad (3.5)$$

2. Pembentukan aturan asosiatif

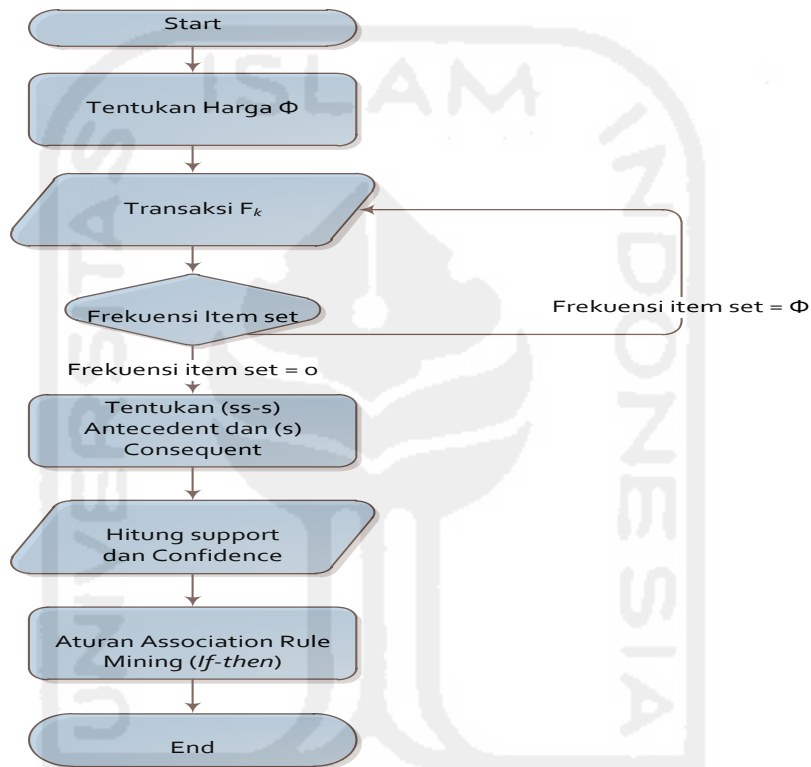
Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan asosiatif A _B Nilai confidence dari aturan A _B diperoleh dari rumus berikut:

$$\text{Confidence} = P(B | A) = \frac{\text{Jumlah Transaksi mengandung } A \text{ dan } B}{\text{Jumlah Transaksi mengandung } A} \quad (3.6)$$

3.7.1. Algoritma Apriori pada Association Rule Mining

Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Algoritma Apriori yang bertujuan untuk menemukan frequent itemsets dijalankan pada sekumpulan data dengan jumlah yang besar. Kaidah asosiasi apriori digunakan

untuk menggambarkan hubungan antar item pada tabel data transaksional ataupun data relasional. Paradigma dasar asosiasi apriori adalah dengan mencari himpunan kandidat dengan panjang $(k+1)$ dari sekumpulan pola frequent dengan panjang k , lalu mencocokkan jumlah kemunculan pola tersebut dengan informasi yang terdapat dalam database. Sehingga analisis apriori akan menemukan semua aturan apriori yang memenuhi syarat minimum untuk support dan syarat minimum untuk confidence.



Gambar 3. 5. Algoritma Apriori pada Association Rule Mining

Alur algoritma *association rule mining* diatas menunjukkan bahwa dalam memulai proses terlebih dahulu menentukan harga Φ . Tujuan dari menentukan harga Φ untuk menghitung nilai *frekuensi* item set. Selanjutnya melakukan transaksi F_k . F_k merupakan Himpunan semua *frekuensi* Item Set yang terdiri dari K item. Proses transaksi F_k akan berhenti jika nilai *frekuensi* item set = 0, tetapi jika nilai *frekuensi* item set $\geq \Phi$ maka proses transaksi F_k akan terus berlanjut. Sebagai contoh pendekatan *association rule* terhadap kasus DBD pada penelitian Mukhlash (2011) sebagai berikut:

Tahun	Kode Kelura...	Nama Kelurahan	Kepadatan	Penderita Pe...	Fasilitas Kese...	Penderita Tel...
2000	3578090002	Putat Jaya	Tinggi	Rendah	Ada	Sedang
2000	3578090003	Banyu Urip	Sedang	Rendah	Ada	Sedang
2000	3578090004	Kupang Krajan	Sedang	Rendah	Ada	Sedang
2000	3578090005	Petemon	Rendah	Rendah	Ada	Sedang
2000	3578090006	Sawahana	Rendah	Sedang	Ada	Sedang
2000	3578100001	Tembok Dukuh	Sedang	Rendah	Ada	Sedang
2000	3578100002	Bubutan	Rendah	Rendah	Tidak Ada	Sedang
2000	3578100003	Alun-alun Contong	Rendah	Sedang	Tidak Ada	Rendah
2000	3578100004	Gundih	Sedang	Rendah	Tidak Ada	Rendah
2000	3578100005	Jepara	Sedang	Rendah	Ada	Rendah
2000	3578110001	Embong Kallasin	Rendah	Rendah	Ada	Tinggi
2000	3578110002	Ketabana	Rendah	Sedang	Ada	Tinggi

Gambar 3. 6. Contoh Hasil Prapemrosesan Data (Mukhlash, 2011)

Dengan menggunakan gambar 3.6 dapat mencari kombinasi item set. Algoritma akan secara langsung memeriksa semua kombinasi yang ada untuk dapat menghitung kejadian munculnya setiap item.

Tabel 3. 1. Kombinasi Kepadatan Penduduk Terhadap Penderita DBD

No	Kombinasi	Jumlah
1.	Kepadatan rendah, Penderita DBD rendah	3
2.	Kepadatan rendah, Penderita DBD sedang	3
3.	Kepadatan rendah, Penderita DBD Tinggi	0
4.	Kepadatan sedang, Penderita DBD rendah	5
5.	Kepadatan sedang, Penderita DBD sedang	0
6.	Kepadatan sedang, Penderita DBD tinggi	0
7.	Kepadatan tinggi, Penderita DBD rendah	1
8.	Kepadatan tinggi, Penderita DBD sedang	0
9.	Kepadatan tinggi, Penderita DBD tinggi	0

Dari tabel 3.1 kombinasi 2 item set mempunyai jumlah munculnya kejadian setiap item set. Misalkan kita tentukan harga $\Phi = 3$ maka didapatkan himpunan yaitu :

$$F_2 = \{ \{ \text{Kepadatan rendah, Penderita DBD rendah} \}, \{ \text{Kepadatan rendah, Penderita DBD sedang} \}, \{ \text{Kepadatan sedang, Penderita DBD rendah} \} \}$$

Setelah transaksi F_k telah selesai, tahap selanjutnya yaitu menentukan (ss-s) *Antecedent* dan (s) *Consequent*. *Rule* yang dipakai adalah If x Then y, dimana x adalah *Antecedent* dan y adalah *Consequent*.

Untuk {Kepadatan rendah, Penderita DBD rendah}

- Jika (ss-s) = Kepadatan rendah
- Jika s = Penderita DBD rendah
- Maka => If Kepadatan rendah then Penderita DBD rendah
- Jika (ss-s) = Penderita DBD rendah
- Jika s = Kepadatan rendah
- Maka => If Penderita DBD rendah then Kepadatan rendah

Setelah mendapatkan *rule* yang telah ditentukan, maka *rule* tersebut dapat dipakai untuk menghitung nilai *support* dan *confidence*.

Tabel 3. 2. Support dan Confident untuk 2 Antecedent

No	Rule	Support	Confidence
1.	If Kepadatan rendah then Penderita DBD rendah	$3/12 = 25\%$	$3/6 = 50\%$
2.	If Penderita DBD rendah then Kepadatan rendah	$3/12 = 25\%$	$3/9 = 33\%$
3.	If Kepadatan rendah then Penderita DBD sedang	$3/12 = 25\%$	$3/6 = 50\%$
4.	If Penderita DBD sedang then Kepadatan rendah	$3/12 = 25\%$	$3/3 = 100\%$
5.	If Kepadatan sedang then Penderita DBD rendah	$5/12 = 42\%$	$5/5 = 100\%$
6.	If Penderita DBD rendah then Kepadatan sedang	$5/12 = 42\%$	$5/9 = 55\%$

Setelah didapat *Support* dan *Confidence* untuk masing-masing kandidat, maka hasil yang memenuhi nilai *confidence* dari perhitungan tersebut merupakan *rule* yang memenuhi aturan *association rule mining*. Misalkan nilai *confidence* minimal 80%, maka *rule* yang memenuhi aturan *association rule mining* yaitu :

If Penderita DBD sedang then Kepadatan rendah

If Kepadatan sedang then Penderita DBD rendah

3.8. Presentasi Hasil

Tahap presentasi hasil ini merupakan tahapan akhir dari pelaksanaan penelitian, yaitu penyampaian kesimpulan atas hasil analisis spasial dan temporal untuk pencarian aturan *association rule* terhadap penyakit demam berdarah dengue di Kota Banjarbaru. Tujuan aturan *association rule* untuk mengetahui hubungan antara

parameter curah hujan, kelembaban udara, suhu, usia penderita DBD, fasilitas kesehatan, dan kepadatan penduduk terhadap kejadian DBD. Untuk mencari aturan *association rule* tersebut, terlebih dahulu data-data dari parameter akan didiskretisasi. Diskretisasi akan dilakukan menggunakan teknik statistik dengan mencari jumlah kelas dan interval atau lebar kelas. Pada setiap parameter akan mempunyai jumlah kelas dan interval atau lebar kelas masing-masing. Aturan *association rule* yang diharapkan yaitu :

Tabel 3. 3. Contoh Hasil Aturan *Association Rule*

NO	Rule	Support	Confidence
1.	<i>IF curah hujan kelas 1 then kasus DBD kelas 1</i>	18,33	78.6%
2.	<i>IF kasus DBD kelas 2 then curah hujan kelas 2</i>	18,33	70.1%
3.	<i>IF kasus DBD kelas 3 then curah hujan kelas 3</i>	15,83	75.5%

Dari tabel 3.3. memberikan contoh gambaran hasil aturan *association rule* untuk curah hujan terhadap kasus DBD. Pada contoh hasil tersebut menjelaskan hubungan asosiasi curah hujan terhadap kasus DBD terjadi ketika curah hujan berada pada kelas 1 dengan kasus DBD pada kelas 1. Ketika kasus DBD meningkat pada kelas 2, curah hujan mengalami kenaikan pada kelas 2. Hal itu juga terjadi ketika kasus DBD menjadi kelas 3, maka curah hujan juga mengalami kenaikan pada kelas 3. Hasil yang dapat disimpulkan bahwa ketika curah hujan tinggi akan memberikan pengaruh yang signifikan dalam penularan penyakit DBD.