

# Analisis Intensitas Hujan Provinsi Jawa Barat Tahun 2020 Menggunakan Association Rule Apriori dan FP-Growth

Bagus Almahenzar<sup>1\*</sup>, Arie Wahyu Wijayanto<sup>2</sup>

<sup>1</sup>Program Studi D-IV Komputasi Statistik, Politeknik Statistika STIS

<sup>2</sup>Badan Pusat Statistik

<sup>1</sup>221911108@stis.ac.id, <sup>2</sup>ariewahyu@stis.ac.id

## Abstrak

Jawa Barat adalah salah satu daerah di Indonesia dengan curah hujan yang tinggi. Salah satu penyebab banjir yang sering terjadi di wilayah DKI Jakarta disebabkan karena menerima air kiriman dari wilayah Jawa Barat. Metode yang digunakan dalam penelitian ini adalah *Association Rule* dengan menggunakan algoritma *Apriori* dan *FP-Growth*. Association rule adalah aturan pada data mining dalam menentukan semua aturan asosiasi yang memenuhi syarat minimum *support* (*minsup*) dan minimum *confidence* (*minconf*) pada sebuah *dataset*. Pada penelitian ini *minimum support* yang digunakan yaitu sebesar 5 dan *minimum confidence* sebesar 0,9. Tujuan penelitian ini adalah untuk mendapatkan pola intensitas hujan yang sering terjadi pada setiap bulan di wilayah-wilayah stasiun pengukuran. Stasiun pengukuran yang ada di wilayah Jawa Barat yaitu Stasiun Meteorologi Citeko, Pos Meteorologi Penggung, Stasiun Meteorologi Kertajati, Stasiun Klimatologi Bogor, dan Stasiun Geofisika Bandung. Hasil penelitian ini menunjukkan bahwa pada bulan April, Mei, Juni, Oktober, dan November tidak ada pola intensitas hujan yang terjadi antar wilayah stasiun pengukuran. Hujan dengan kategori lebat, sangat lebat, dan ekstrem sangat jarang terjadi. Pada bulan Juli, Agustus, dan September, sebagian besar wilayah tidak terjadi hujan.

Kata Kunci: *Association Rule, Apriori, FP-Growth, Intensitas Hujan*

## Abstract

West Java is one of the areas in Indonesia with high rainfall. One of the causes of frequent flooding in the DKI Jakarta area is due to receiving water from the West Java area. The method used in this study is the Association Rule using the Apriori algorithm and FP-Growth. Association rule is a rule in data mining in determining all association rules that meet the minimum support (*minsup*) and minimum confidence (*minconf*) requirements in a dataset. In this study, the minimum support used is 5 and the minimum confidence is 0.9. The purpose of this study was to obtain a pattern of rainfall intensity that often occurs every month in the measurement station areas. The measurement stations in the West Java region are the Citeko Meteorological Station, the Penggung Meteorological Post, the Kertajati Meteorological Station, the Bogor Climatology Station, and the Bandung Geophysics Station. The results of this study indicate that in April, May, June, October, and November there is no pattern of rain intensity that occurs between the measurement station areas. Heavy, very heavy, and extreme rains are very rare. In July, August, and September, most areas do not experience rain.

Keywords: *Association Rule, Apriori, FP-Growth, Rainfall Intensity*

## 1. Pendahuluan

Indonesia adalah negara yang memiliki dua musim yaitu musim kemarau dan musim hujan. Musim hujan terjadi pada kisaran bulan September hingga bulan Maret sedangkan musim kemarau terjadi pada kisaran bulan Maret hingga September (Diniari, 2020). Berdasarkan Buku “Prakiraan Musim Hujan Tahun 2021/2022 di Indonesia” yang dirilis oleh Badan Meteorologi Klimatologi dan Geofisika, sebagian besar wilayah di Indonesia akan mengalami awal musim hujan 2021/2022 pada kisaran bulan Oktober dan November 2021

(BMKG, 2021). Prakiraan musim hujan penting untuk diketahui sebagai bentuk persiapan dan antisipasi ketika terjadi musibah banjir.

Salah satu daerah dengan curah hujan yang tinggi setiap tahunnya adalah Provinsi Jawa Barat. Berdasarkan data BPS, jumlah curah hujan tahun 2011 hingga 2013 di Provinsi Jawa Barat yaitu 1789 mm, 2510 mm, dan 2682 mm. Curah hujan yang tinggi menjadi salah satu penyebab banjir yang sering terjadi di wilayah DKI Jakarta karena menerima air kiriman dari wilayah Jawa Barat yaitu Bogor (BPS, 2017).

Curah hujan dapat diklasifikasikan menjadi 6 intensitas berdasarkan jumlah curah hujan per hari yaitu 0 mm untuk berawan, 0,5 hingga 20 mm untuk hujan ringan, 20 hingga 50 mm untuk hujan sedang, 50 hingga 100 mm untuk hujan lebat, 100 hingga 150 mm untuk hujan sangat lebat, dan lebih dari 150 mm untuk hujan ekstrem (BMKG, 2021).

Klasifikasi curah hujan menjadi intensitas hujan tentunya lebih mudah dipahami oleh masyarakat. Pola intensitas hujan dari tiap wilayah per harinya di Jawa Barat adalah informasi yang penting untuk kebijakan penanggulangan bencana di wilayah Jawa Barat maupun DKI Jakarta sehingga setiap bulan pengambil keputusan dapat menentukan kebijakan terkait hal-hal yang harus dilakukan (Putri & Wijayanto, 2021).

Pada penelitian sebelumnya dilakukan analisis prediksi hujan di wilayah kota Bandung menggunakan algoritma *Apriori*, didapatkan hasil bahwa semua rule yang dihasilkan memiliki nilai *lift ratio* lebih dari 1.00 sehingga semua rule dapat digunakan sebagai acuan dalam memprediksi hujan pada wilayah Kota Bandung. Nilai *lift ratio* digunakan sebagai validasi aturan asosiasi yang terbentuk (Fauzy et al., 2016).

Algoritma *Apriori* merupakan algoritma yang sering digunakan pada Association Rule (Mansyur & Ridho, 2020) (Mishra et al., 2015) (Lestari, 2015). Pada penelitian sebelumnya hanya menggunakan algoritma *Apriori*, tetapi pada penelitian ini juga akan menggunakan algoritma *FP-Growth* dalam menganalisis pola intensitas hujan yang terjadi di Provinsi Jawa Barat. Kelebihan dari algoritma *FP-Growth* yaitu dapat mengompresi atau memampatkan *dataset* asli karena algoritma *FP-Growth* merupakan pengembangan dari algoritma *Apriori* dengan menggunakan konsep pohon atau *FP-Tree* (*Frequent Pattern Tree*) sehingga struktur data akan menjadi padat.

Tujuan penelitian ini adalah untuk mendapatkan pola intensitas hujan yang sering terjadi pada setiap bulan di wilayah-wilayah stasiun pengukuran dengan menggunakan dua algoritma pada *Association Rule* yaitu *Apriori* dan *FP-Growth*. Tiap daerah di Provinsi Jawa Barat tidak semuanya memiliki stasiun pengukuran sendiri sehingga hanya daerah yang memiliki stasiun pengukuran yang datanya akan digunakan dan dilakukan dianalisis. Stasiun pengukuran yang ada di wilayah Jawa Barat yaitu Stasiun Meteorologi Citeko, Pos Meteorologi Penggung, Stasiun Meteorologi Kertajati, Stasiun Klimatologi Bogor, dan Stasiun Geofisika Bandung. Penggunaan algoritma *Apriori* dan *FP-Growth* pada penelitian ini sesuai dengan penelitian sebelumnya pada bidang lain, yaitu Mansyur (2020), Lestari (2015), dan Fitriyani (2016) tentang keunggulan kedua algoritma tersebut untuk tugas *association rule* pada data.

Kontribusi utama dari penelitian ini yaitu pada perbandingan kinerja algoritma *association rule* yang terpopuler, *Apriori* dan *FP-Growth* untuk permasalahan intensitas hujan yang di dalam literatur umumnya hanya terbatas pada analisis klasifikasi atau prediksi, khususnya di wilayah provinsi Jawa Barat.

## 2. Metode Penelitian

Data yang digunakan adalah data iklim harian sepanjang tahun 2020 dari setiap stasiun pengukuran di Jawa Barat. Sumber data tersebut yaitu dari Layanan Data Online Pusat Database Badan Meteorologi Klimatologi dan Geofisika (BMKG, 2020). Pada data iklim

harian terdapat sepuluh variabel terkait keadaan iklim pada wilayah stasiun pengukuran. Penelitian ini hanya membutuhkan data pada satu variabel yaitu variabel curah hujan. Sepuluh variabel pada data iklim harian seperti ditunjukkan pada tabel 1 berikut:

Tabel 1. Variabel pada Data Iklim Harian

No.	Kode	Variabel	Satuan/Skala
1	ddd_x	Arah angin saat kecepatan maksimum	Derajat (°)
2	ddd_car	Arah angin terbanyak	Derajat (°)
3	RR	Curah hujan	Milimeter (mm)
4	ff_x	Kecepatan angin maksimum	Meter per second (m/s)
5	ff_avg	Kecepatan angin rata- rata	Meter per second (m/s)
6	RH_avg	Kelembapan rata-rata	Persen (%)
7	ss	Lama penyinaran matahari	Jam
8	Tx	Temperatur maksimum	Celcius (°C)
9	Tn	Temperatur minimum	Celcius (°C)
10	Tavg	Temperatur rata-rata	Celcius (°C)

Metode yang digunakan untuk mendapatkan pola intensitas hujan yang sering terjadi pada setiap bulan di wilayah-wilayah stasiun pengukuran yaitu metode *Apriori* dan *FP-Growth*. Dua metode tersebut adalah metode-metode yang digunakan pada *Association Rule*.

*Association rule* adalah aturan pada data mining dalam menentukan semua aturan asosiasi yang memenuhi syarat minimum *support* (*minsup*) dan minimum *confidence* (*minconf*) pada sebuah basis data (Mansyur & Ridho, 2020) (Mishra et al., 2015) (Lestari, 2015). *Association rule mining* merupakan suatu prosedur untuk mendapatkan hubungan antar *item* dalam suatu *dataset*. Dimulai dengan mencari *frequent itemset* yaitu *itemset* yang sering muncul berdasarkan minimum *support* yang ditetapkan (Han & Kamber, 2012).

*Support* didefinisikan sebagai peluang sebuah transaksi yang mengandung *itemset* tertentu pada dataset yang digunakan untuk melihat jumlah kemunculan *itemset* tersebut. Rumus untuk *support* seperti ditunjukkan pada rumus 1 berikut:

$$supp(A) = \frac{\text{transaksi itemset } A}{\text{total transaksi}} \quad (1)$$

*Confidence* adalah indikator yang memperlihatkan kemungkinan suatu item dibeli dengan syarat item tertentu dibeli juga (Mansyur & Ridho, 2020) (Mishra et al., 2015). Rumus untuk *confidence* seperti ditunjukkan pada rumus 2 berikut:

$$conf(A \Rightarrow B) = \frac{supp(A \cup B)}{supp(A)} \quad (2)$$

Algoritma *Apriori* adalah algoritma yang melakukan *scanning* pada item di dalam basis data kemudian dilakukan eliminasi pada item yang tidak memenuhi minimum *support* yang telah ditentukan. Algoritma *Apriori* merupakan algoritma *association rule* yang diperkenalkan oleh Agrawal & Srikant pada tahun 1994 (Mansyur & Ridho, 2020) (Mishra et al., 2015).

Tahapan algoritma Apriori yaitu sebagai berikut:

1. Hitung semua kemunculan *itemset* dengan jumlah *item*-nya 1.
2. Eliminasi semua *itemset* yang tidak memenuhi minimum *support*.
3. Dari *itemset* yang memenuhi minimum *support*, buat semua kombinasi dengan jumlah *item*-nya 2 dan hitung kembali semua kemunculan *itemset*.
4. Eliminasi semua *itemset* yang tidak memenuhi minimum *support*.
5. Begitu seterusnya sampai tidak ada lagi kombinasi *itemset* yang memungkinkan pada basis data.

Algoritma *FP-Growth* (*Frequent Pattern Growth*) merupakan algoritma alternatif yang dapat digunakan untuk menentukan *itemset* yang paling sering muncul atau *frequent itemset* dalam suatu kumpulan data. Pada algoritma *FP-Growth* tidak memerlukan *candidate generation* seperti yang dilakukan pada algoritma *Apriori*. Algoritma *FP-Growth* merupakan pengembangan dari algoritma *Apriori* yaitu dengan menggunakan konsep pohon atau *FP-Tree* (*Frequent Pattern Tree*) sehingga struktur data akan menjadi lebih padat yang berguna untuk mengompresi atau memampatkan *dataset* asli (Asana et al., 2020).

Algoritma *FP-Growth* terdiri dari tiga tahapan sebagai berikut:

1. Tahap pembangkitan *conditional pattern base* - *Conditional pattern base* adalah *subdataset* yang berisi *path prefix* dan *pattern suffix*.
2. Tahap pembangkitan *conditional FP-Tree* - Pada tahap ini dilakukan penjumlahan *support* setiap *item* pada setiap *conditional pattern base*, kemudian setiap *item* yang memiliki jumlah *support* lebih besar atau sama dengan *minimum support* akan dibangkitkan dengan *conditional FP-Tree*.
3. Tahap pencarian *frequent itemset* - Jika *conditional FP-Tree* adalah jalur tunggal, maka *frequent pattern* diperoleh dengan menggabungkan *item* untuk setiap *conditional FP-Tree*. Jika bukan jalur tunggal, maka *FP-Growth* akan dibangkitkan secara rekursif.

Sebagai bagian dari pembelajaran mesin (*machine learning*), model *association rule*, termasuk algoritma *FP-Growth* menarik kesimpulan berdasarkan data latih yang sudah dipelajarinya (Santoso, 2021) (Syukra et al., 2019). Pembelajaran mesin memiliki potensi untuk menghasilkan performa yang unggul seiring dengan semakin banyaknya data yang tersedia, baik untuk prediksi, klasifikasi, *clustering*, *association rule*, dll sehingga dapat diimplementasikan dalam berbagai bidang (Thamrin & Wijayanto, 2021) (Damayanti & Wijayanto, 2021) (Afira & Wijayanto, 2021) (Sikana & Wijayanto, 2021) (Luthfi & Wijayanto, 2021). Penelitian ini membandingkan dua algoritma pada *association rule* yaitu algoritma *Apriori* dan *FP-Growth* yang telah terbukti keunggulannya pada bidang aplikasi lainnya (Fitriyani, 2016) (Arimbawa et al., 2019).

### 3. Hasil dan Pembahasan

#### 3.1. Preprocessing

Tahap *preprocessing* data dilakukan dengan menggunakan Microsoft Excel 2016. Pada setiap stasiun pengukuran, tidak setiap hari data curah hujan dapat diukur sehingga jika data tidak terukur dikodekan dengan “8888” dan jika tidak dilakukan pengukuran dikodekan dengan “9999” atau datanya kosong pada tabel. Tabel 2 berikut adalah laporan iklim harian pada Stasiun Geofisika Bandung pada bulan Maret tahun 2020.

Tabel 2. Laporan Iklim Harian Stasiun Geofisika Bandung Bulan Maret Tahun 2020

Hari ke-	Tn	Tx	Tavg	RH_avg	RR	ss	ff_x	ddd_x	ff_avg	ddd_car
1	21.8	30.4	24.7	84	8888	1.5	5	280	2	W
2	22.1	30	25.1	76	2	3.5	5	300	2	W
3	20	30.8	23.8	78	0.5	5.3	5	270	2	W
4	20	30.6	24.4	76		7.4	8	240	3	W
5	20.8	29	24.1	72	9.9	5.1	3	280	1	W
6	21	31.2	24.5	74	6	1.4	4	240	2	W
7	20.8	29.8	25.3	72	0.4	6.4	5	260	2	W
8	21.9	29.2	24.3	72		6.8	5	270	2	W
9	21.7	29.2	24.6	69	3.4	3.4	6	230	3	W
10	21.6	31.6	25.6	72		3.2	4	240	2	NW
11	22.2	31.4	25.3	77	0.4	5.7	5	290	2	W
12	22.6	27.8	23.8	84	8888	5.6	3	290	1	W
13	21.4	30.6	25.2	78	5.2	0.8	5	270	2	SW
14	21.4	29	24	84	0	7.6	3	290	1	C
15	20.2	30.6	23.8	86	6.4	3.1	3	340	1	C
16	20.6	29.4	23.4	88	38.5	5	3	300	1	C
17	20	31.2	24.3	78	6.4	4.8	3	240	1	C
18	20	27.4	23.3	91	0.5	5.2	2	100	1	C
19	21.2	28.4	23.4	86	20.6	1.2	3	330	1	C
20	20.4	27.8	22.5	91	19.4	3.3	3	240	1	C
21	19.9	29	23.4	86	28.2	3	4	270	2	W
22	22	27.6	24.2	82	0.8	4.4	2	130	1	SE
23	20.6	28	23.1	82	8888	1	3	100	1	C
24	20.2	29.2	23.2	83	2.2	0.7	6	270	1	C
25	20.3	31	24.1	82	2.4	4.3	4	240	2	NW
26	21.8	29.4	24.6	84	21.8	6	3	240	1	C
27	21.4	29.4	24.2	85	26.5	3.4	2	110	1	C
28	21	29	23.8	84	0.8	4.9	3	240	1	C
29	20.9	29.2	23.4	87	12.6	3.7	3	320	2	SE
30	20.2	29.8	23.2	86	16.2	2.8	5	270	2	W
31	20	30	22.9	86	59.7	5	3	100	1	C

Tabel 2 merupakan salah satu dari enam puluh dataset yang diunduh dari Layanan Data Online Pusat Database Badan Meteorologi Klimatologi dan Geofisika.

Data berupa jumlah curah hujan dalam satuan milimeter sehingga harus diklasifikasikan ke dalam intensitas hujan. Berikut kode yang akan diberikan pada rentang jumlah curah hujan:

- Kode “A” untuk 0 mm
- Kode “B” untuk 0,5 – 20 mm
- Kode “C” untuk 20 – 50 mm
- Kode “D” untuk 50 – 100 mm
- Kode “E” untuk 100 – 150 mm
- Kode “F” untuk lebih dari 150 mm

Untuk kode pada setiap stasiun pengukuran, berikan kode berikut setelah kode intensitas hujan:

- Kode “1” untuk Stasiun Meteorologi Citeko
- Kode “2” untuk Pos Meteorologi Penggung
- Kode “3” untuk Stasiun Meteorologi Kertajati
- Kode “4” untuk Stasiun Klimatologi Bogor
- Kode “5” untuk Stasiun Geofisika Bandung

Tabel 3 berikut adalah tabel intensitas hujan pada bulan Januari yang sudah diberikan kode.

Tabel 3. Intensitas Hujan pada Bulan Januari

Hari ke-	Citeko	Penggung	Kertajati	Bogor	Bandung
1	D1	B2	D3	C4	B5
2	D1	B2	B3	C4	B5
3	B1	A2	A3	A4	
4	B1		A3	A4	A5
5	B1		A3	B4	A5
6	B1	B2		B4	A5
7	C1	B2	B3	B4	B5
8	B1	C2	C3	B4	A5
9	C1	C2	B3	B4	B5
10	C1	C2	B3	B4	B5
11	B1		A3		
12	A1	C2	B3		
13	A1	C2	A3		A5
14	B1	E2			A5
15	B1	A2	A3	B4	A5
16	B1	B2	A3	B4	
17	A1	B2	B3	B4	A5
18	B1	C2	D3	D4	B5
19	A1	C2	B3	A4	A5
20	B1	D2	B3	D4	
21	B1	C2	C3	B4	B5
22	A1	A2			B5
23	B1				C5
24	C1	B2	B3	B4	B5
25	B1	C2	B3	B4	C5
26	C1	B2	C3	B4	D5
27	B1		D3	C4	A5
28	C1		B3	B4	B5
29	A1		C3		B5
30	C1	B2	B3	B4	B5
31	A1	A2	A3		A5

Data pada tabel 3 digunakan untuk penerapan Algoritma *Apriori* dengan menggunakan Weka. Penerapan Algoritma *FP-Growth* akan menggunakan bahasa *Python*. Lakukan transformasi dengan menjadikan semua *item* sebagai atribut sehingga akan ada 31 baris dan 30

kolom, selanjutnya berikan nilai 0 untuk kasus tidak terjadi dan 1 untuk kasus terjadi. Tabel 4 merupakan contoh untuk *item* dengan kode “B”.

Tabel 4. *Item* kode “B” untuk Penerapan Algoritma *FP-Growth*

Hari ke-	B1	B2	B3	B4	B5
1	0	1	0	0	1
2	0	1	1	0	1
3	1	0	0	0	0
4	1	0	0	0	0
5	1	0	0	1	0
6	1	1	0	1	0
7	0	1	1	1	1
8	1	0	0	1	0
9	0	0	1	1	1
10	0	0	1	1	1
11	1	0	0	0	0
12	0	0	1	0	0
13	0	0	0	0	0
14	1	0	0	0	0
15	1	0	0	1	0
16	1	1	0	1	0
17	0	1	1	1	0
18	1	0	0	0	1
19	0	0	1	0	0
20	1	0	1	0	0
21	1	0	0	1	1
22	0	0	0	0	1
23	1	0	0	0	0
24	0	1	1	1	1
25	1	0	1	1	0
26	0	1	0	1	0
27	1	0	0	0	0
28	0	0	1	1	1
29	0	0	0	0	1
30	0	1	1	1	1
31	0	0	0	0	0

Tabel 4 merupakan *item* untuk kode “B” yang digunakan sebagai contoh untuk memperlihatkan proses transformasi yang dilakukan. Lakukan hal yang sama seperti tabel 4 untuk lima kode lainnya.

### 3.2. Algoritma Apriori

Setelah proses *preprocessing* telah selesai dilakukan, selanjutnya menentukan minimum *support* yaitu sebesar 5 dan minimum *confidence* sebesar 0,9. Kemudian hitung *support* setiap *item* yaitu jumlah kemunculan *item* pada suatu dataset. Hasil penghitungan *support* setiap *item* seperti pada tabel 5.

Tabel 5. *Support* Setiap *Item*

<i>Item</i>	<i>Support</i>
B1	15

<i>Item</i>	<i>Support</i>
B4	15
B3	12
B5	12
A5	11
B2	9
C2	9
A3	8
A1	7
C1	7
A2	4
C3	4
D3	3
A4	3
C4	3
D1	2
D4	2
C5	2
D2	1
E2	1
D5	1
E1	0
F1	0
F2	0
E3	0
F3	0
E4	0
F4	0
E5	0
F5	0

*Record* dengan warna gelap menunjukkan *item* yang memenuhi minimum *support* dan selanjutnya dapat disebut *frequent*.

Untuk mempermudah pada tahap selanjutnya, gunakan *software machine learning* yaitu Weka. Weka atau *Waikato Environment for Knowledge Analysis* merupakan *software* yang dikembangkan Universitas Waikato di Selandia Baru untuk penerapan algoritma data mining menggunakan bahasa *JAVA*. Hasil yang didapatkan dengan menggunakan Weka ditunjukkan pada tabel 6.

Tabel 6. Sebelas Rule Terbaik yang Ditemukan Menggunakan Weka

No.	Rule	Conf
1	C1 7 → B4 7	1
2	C1 B3 6 → B4 6	1
3	C1 B5 6 → B3 6	1
4	C1 B3 6 → B5 6	1
5	C1 B5 6 → B4 6	1
6	B3 B4 B5 6 → C1 6	1
7	C1 B4 B5 6 → B3 6	1
8	C1 B3 B5 6 → B4 6	1
9	C1 B3 B4 6 → B5 6	1

No.	Rule	Conf
10	C1 B5 6 → B3 B4 6	1
11	C1 B3 6 → B4 B5 6	1

Angka setelah kode menunjukkan jumlah kemunculan *itemset* tersebut. Nilai *confidence* sebesar 1 artinya 100 persen yaitu seperti contoh pada nomor 1 jika C1 muncul sebanyak 7 kali maka B4 juga akan muncul sebanyak 7 kali sehingga *confidence* untuk *rule* tersebut adalah 1 atau 100 persen. Pada tabel 5 nilai *support* B1 sebesar 15, angka ini cukup besar tetapi tidak termasuk 11 *rule* terbaik karena tidak ada *itemset* yang memenuhi *rule* yang mengandung B1 dengan *confidence* yang tinggi yaitu 1.

### 3.3. Algoritma FP-Growth

Penerapan algoritma *FP-Growth* dilakukan dengan menggunakan bahasa *Python*. Persiapan *environment* yang harus dilakukan yaitu menginstall *package mlxtend*. *MLxtend* merupakan *package* yang digunakan pada metode *association rule*. Data yang sudah dilakukan transformasi seperti pada tabel 3 akan digunakan pada penerapan algoritma *FP-Growth* di *Python*. Parameter yang digunakan yaitu data hasil *preprocessing* dan *minimum support*. Hasil yang didapatkan dengan menggunakan *Python* ditunjukkan pada tabel 7.

Tabel 7. Rule yang Ditemukan Menggunakan *Python*

No.	Antecedents	Consequents	Support	Confidence
1	C1	B4	0,2258	1
2	C1, B3	B4	0,1936	1
3	C1, B5	B4	0,1936	1
4	C1, B3	B5	0,1936	1
5	C1, B5	B3	0,1936	1
6	C1, B4, B5	B3	0,1936	1
7	C1, B4, B3	B5	0,1936	1
8	C1, B3, B5	B4	0,1936	1
9	B4, B5, B3	C1	0,1936	1
10	C1, B5	B4, B3	0,1936	1
11	C1, B3	B4, B5	0,1936	1

### 3.4. Hasil Penerapan Setiap Bulan Sepanjang Tahun 2020

Hasil penerapan algoritma *Apriori* pada bulan Februari hingga Desember seperti yang ditunjukkan pada tabel 8 hingga tabel 13 berikut:

Tabel 8. Rule pada Bulan Februari Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	B4 B5 6 → B2 6	1
2	B2 B4 6 → B5 6	1
3	C5 5 → B1 5	1
4	C1 B5 5 → B2 5	1
5	C1 B2 5 → B5 5	1

Tabel 9. Rule pada Bulan Maret Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	B2 B5 5 → B3 5	1
2	B3 B4 5 → B5 5	1

Tabel 10. Rule pada Bulan Juli Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	A1 11 → B3 11	1
2	A1 A5 6 → A3 6	1
3	A1 A2 5 → A3 5	1
4	A2 14 → A3 13	0,93
5	A5 14 → A3 13	0,93

Tabel 11. Rule pada Bulan Agustus Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	A5 15 → A1 15	1
2	A3 A5 13 → A1 13	1
3	A2 A5 9 → A1 9	1
4	A2 A3 A5 7 → A1 7	1
5	A4 6 → A3 6	1
6	A1 A4 5 → A3 5	1

Tabel 12. Rule pada Bulan September Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	A1 10 → A3 10	1
2	A5 10 → A3 10	1
3	A1 A5 9 → A3 9	1
4	A1 A2 6 → A3 6	1
5	A2 A5 6 → A3 6	1
6	A4 A5 5 → A3 5	1
7	A1 A2 A5 5 → A3 5	1
8	A2 11 → A3 10	0,91
9	A5 10 → A1 9	0,9
10	A1 10 → A5 9	0,9
11	A3 A5 10 → A1 9	0,9
12	A1 A3 10 → A5 9	0,9
13	A5 10 → A1 A3 9	0,9
14	A1 10 → A3 A5 9	0,9

Tabel 13. Rule pada Bulan Desember Menggunakan Algoritma *Apriori*

No.	Rule	Confidence
1	B2 B3 7 → B5 7	1
2	A3 5 → B1 5	1
3	C3 5 → B5 5	1
4	B2 B4 B5 5 → B3 5	1
5	B2 B3 B4 5 → B5 5	1

Pada bulan April, Mei, Juni, Oktober, dan November tidak ditemukan rule yang memenuhi minimum *support* yaitu 5. Hasil penerapan algoritma *FP-Growth* pada bulan Februari hingga Desember seperti yang ditunjukkan pada tabel 14 hingga tabel 19 berikut:

Tabel 14. Rule pada Bulan Februari Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	B2, B4	B5	0,2069	1
2	B4, B5	B2	0,2069	1
3	C1, B2	B5	0,1724	1
4	C1, B5	B2	0,1724	1
5	C5	B1	0,1724	1

Tabel 15. Rule pada Bulan Maret Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	B2, B5	B3	0,1613	1
2	B3, B4	B5	0,1613	1

Tabel 16. Rule pada Bulan Juli Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	A5	A3	0,4194	0,9286
2	A2	A3	0,4194	0,9286
3	A1	A3	0,3548	1
4	A1, A5	A3	0,1936	1
5	A1, A2	A3	0,1612	1

Tabel 17. Rule pada Bulan Agustus Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	A5	A1	0,4839	1
2	A3, A5	A1	0,4194	1
3	A2, A5	A1	0,2903	1
4	A2, A3, A5	A1	0,2258	1
5	A4	A3	0,1936	1
6	A1, A4	A3	0,1613	1

Tabel 18. Rule pada Bulan September Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	A2	A3	0,3333	0,9091
2	A5	A3	0,3333	1
3	A2, A5	A3	0,2	1
4	A1	A3	0,3333	1
5	A1	A5	0,3	0,9
6	A5	A1	0,3	0,9
7	A1, A3	A5	0,3	0,9
8	A1, A5	A3	0,3	1
9	A3, A5	A1	0,3	0,9
10	A1	A3, A5	0,3	0,9
11	A5	A1, A3	0,3	0,9
12	A1, A2	A3	0,2	1
13	A1, A2, A5	A3	0,1667	1
14	A5, A4	A3	0,1667	1

Tabel 19. Rule pada Bulan Desember Menggunakan Algoritma *FP-Growth*

No.	Antecedents	Consequents	Support	Confidence
1	B2, B3	B5	0,2258	1
2	B3, B2, B4	B5	0,1613	1
3	B2, B5, B4	B3	0,1613	1
4	C3	B5	0,1613	1
5	A3	B1	0,1613	1

Pada algoritma *FP-Growth*, rule untuk bulan April, Mei, Juni, Oktober, dan November juga tidak ditemukan yang memenuhi *minimum support* yaitu 5.

### 3.5. Interpretasi

Pada bulan Januari, didapatkan sebanyak sebelah rule dengan *confidence* sebesar 1. Rule intensitas hujan yang didapatkan menunjukkan hubungan antara wilayah stasiun pengukuran dengan intensitas hujan ringan dan hujan sedang. Contohnya jika di Stasiun Meteorologi Citeko mengalami hujan sedang dan di Stasiun Meteorologi Kertajati mengalami hujan ringan, maka di Stasiun Klimatologi Bogor akan mengalami hujan ringan.

Untuk rule intensitas hujan pada bulan Februari menunjukkan hubungan antara wilayah stasiun pengukuran dengan intensitas hujan ringan dan hujan sedang. Contohnya jika di Stasiun Klimatologi Bogor dan di Stasiun Geofisika Bandung mengalami hujan ringan, maka di Pos Meteorologi Penggung juga mengalami hujan ringan. Pada bulan Agustus dan September, didapatkan masing-masing enam dan empat belas rule dengan semua rule-nya adalah mengalami berawan.

Pada bulan Desember, didominasi rule dengan intensitas hujan ringan. Contohnya jika di Pos Meteorologi Penggung, Stasiun Meteorologi Kertajati, dan Stasiun Klimatologi Bogor mengalami hujan ringan, maka di Stasiun Geofisika Bandung juga akan mengalami hujan ringan.

## 4. Kesimpulan

Berdasarkan hasil penerapan algoritma *Apriori* dan *FP-Growth* pada data intensitas hujan Provinsi Jawa Barat sepanjang tahun 2020, dapat diambil kesimpulan sebagai berikut:

1. Hasil yang diperoleh dari penerapan algoritma *Apriori* dengan menggunakan *software* Weka dan algoritma *FP-Growth* dengan menggunakan bahasa *Python* menghasilkan *output* yang sama.
2. Pada bulan April, Mei, Juni, Oktober, dan November tidak ada pola intensitas hujan yang terjadi antar wilayah stasiun pengukuran.
3. Hujan dengan kategori lebat, sangat lebat, dan ekstrem sangat jarang terjadi.
4. Pada bulan Juli, Agustus, dan September, sebagian besar wilayah tidak terjadi hujan.

Penelitian ini masih terbatas menggunakan satu variabel bebas yaitu variabel curah hujan yang digunakan untuk menganalisis pola intensitas hujan di Provinsi Jawa Barat. Untuk penelitian berikutnya, penulis menyarankan untuk dapat menggunakan beberapa variabel pada data iklim harian, misalnya lama penyinaran matahari, temperatur, kelembapan, dan kecepatan angin untuk membuat rule yang digunakan dalam melakukan prediksi cuaca pada suatu wilayah. Selanjutnya dapat menggunakan dataset yang lebih besar, misalnya data iklim harian pada suatu negara atau gabungan data tahunan. Pada penelitian selanjutnya juga diharapkan untuk menggunakan *lift ratio* sebagai validasi rule yang terbentuk.

## Daftar Pustaka

- Diniari., E. B., "Pembagian Waktu dan Perubahan Musim di Indonesia," ruangguru, 2020. [Online]. Tersedia: <https://www.ruangguru.com/blog/pembagian-waktu-dan-perubahan-musim-di-indonesia>. [Diakses: 17-Desember-2021]
- Badan Meteorologi Klimatologi dan Geofisika, "Prakiraan Musim Hujan 2021/2022 di Indonesia," BMKG, 2021. [Online]. Tersedia: <https://www.bmkg.go.id/iklim/prakiraan-musim.bmkg>.
- Badan Pusat Statistik, "Jumlah Curah Hujan dan Jumlah Hari Hujan di Stasiun Pengamatan BMKG, 2011-2015," BPS, 2017. [Online]. Tersedia: <https://www.bps.go.id/statictable/2017/02/08/1959/jumlah-curah-hujan-dan-jumlah-hari-hujan-di-stasiun-pengamatan-bmkg-2011-2015.html>
- Badan Meteorologi Klimatologi dan Geofisika, "Probabilitas Curah Hujan 20 mm (tiap 24 jam)," BMKG, 2021. [Online]. Tersedia: <https://www.bmkg.go.id/cuaca/probabilistik-curah-hujan.bmkg>.
- Putri, S. R., & Wijayanto, A. W. (2021). Learning Bayesian Network for Rainfall Prediction Modeling in Urban Area using Remote Sensing Satellite Data (Case Study: Jakarta; Indonesia). Proceedings of The International Conference on Data Science and Official Statistics, 2021, 77-90.
- Fauzy, M., Saleh W, K. R., & Asror, I. (2016). Penerapan Metode Association Rule Menggunakan Algoritma Apriori pada Simulasi Prediksi Hujan Wilayah Kota Bandung. Jurnal Ilmiah Teknologi Informasi Terapan. 2(2), 221-227.
- Mansyur, F., & Ridho, F. (2020). Analisis Pola Permintaan Publikasi Data Badan Pusat Statistik Menggunakan Association Rule Apriori. Klik - Kumpulan Jurnal Ilmu Komputer. 7(2), 187-198.
- Mishra, A. K., Pani, S. K., & Ratha, B. K. (2015). Association Rule Mining with Apriori and Fpgrowth Using Weka. International Journal of Advanced Technology in Engineering and Science. 3(1), 91-99.
- Lestari, Y. D. (2015). Penerapan Data Mining Menggunakan Algoritma FP-Tree dan FP-Growth Pada Data Transaksi Penjualan Obat. Seminar Nasional Teknologi Informasi Dan Komunikasi (SNASTIKOM 2015), 60-65.
- Badan Meteorologi Klimatologi dan Geofisika, "Data Online Pusat Database BMKG: Data Harian," BMKG, 2021. [Online]. Tersedia: [https://dataonline.bmkg.go.id/data\\_iklim](https://dataonline.bmkg.go.id/data_iklim)
- Han, J., & Kamber, M. Data Mining: Concepts and Techniques. Department of Computer Science, University of Illinois. 3rd ed. Waltham: Morgan Kaufmann, 2012.
- Asana, I. M. D. P., Wiguna, I. K. A. G., Atmaja, K. J., & Sanjaya, I. P. A. (2020). FP-Growth Implementation in Frequent Itemset Mining for Consumer Shopping Pattern Analysis Application. Jurnal Mantik, 4(3), 2063-2070.
- Santoso, M. H. (2021). Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom. Brilliance: Research of Artificial Intelligence. 1(2), 54-66.
- Syukra, I., Hidayat, A., & Fauzi, M. Z. (2019). Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendations. Indonesian Journal of Artificial Intelligence and Data Mining, 2(2), 107-115.
- Thamrin, N., & Wijayanto, A. W. (2021). Comparison of Soft and Hard Clustering: A Case Study on Welfare Level in Cities on Java Island. Indonesian Journal of Statistics and Its Applications, 5(1), 141-160.

- Damayanti, A. R., & Wijayanto, A. W. (2021). Comparison of Hierarchical and Non-Hierarchical Methods in Clustering Cities in Java Island using the Human Development Index Indicators year 2018. *Eigen Mathematics Journal*, 4(1), 8-17.
- Afira, N., & Wijayanto, A. W. (2021). Analisis Cluster dengan Metode Partitioning dan Hierarki pada Data Informasi Kemiskinan Provinsi di Indonesia Tahun 2019. *Komputika: Jurnal Sistem Komputer*, 10(2), 101-109.
- Sikana, A. M., & Wijayanto, A. W. (2021). Analisis Perbandingan Pengelompokan Indeks Pembangunan Manusia Indonesia Tahun 2019 dengan Metode Partitioning dan Hierarchical Clustering. *Jurnal Ilmu Komputer*, 14(2), 66-78.
- Luthfi, E., & Wijayanto, A. W. (2021). Analisis Perbandingan Metode Hierarchical, K-Means, dan K-Medoids Clustering Dalam Pengelompokkan Indeks Pembangunan Manusia Indonesia. *INOVASI*, 17(4), 770-782
- Fitriyani, F. (2016). Implementasi Algoritma FP-Growth Menggunakan Association Rule Pada Market Basket Analysis. *Jurnal Informatika*, 2(1), 296-304.
- Arimbawa P, D. A. K., Prabawa, I. N. A., & Mertasana, P. A. (2019). Implementation of Apriori Algorithm in Determining Tourism Visit Patterns to Bali. *International Journal of Engineering and Emerging Technology*, 4(1), 10–14.