

Muhammad Yusuf¹, Arief Setyanto², Komang Aryasa³

^{1,2,3}Program Studi PJJ Magister Teknik Informatika, Universitas Amikom Yogyakarta

¹yusuf.1175@students.amikom.ac.id, ²arief_s@amikom.ac.id,

³komang.aryasa@dipanegara.ac.id

Abstract

Currently, climate change in Indonesia, which is a tropical region, is always uncertain and makes it difficult to predict weather conditions. Weather conditions can be influenced by temperature, air pressure, wind speed, humidity and rainfall. Rainfall is a climate parameter that has a high level of diversity due to climate anomalies. There are several factors that influence the characteristics of the diversity of rainfall, namely geographical, orographic, topographical, orientation and structure of the islands. These factors cause the distribution pattern of rainfall to be uneven between one area and another. For that we need a method that can solve the problem of predicting rainfall both daily, monthly and yearly. Prediction of rainfall with a statistical approach can be done through the Multiple Linear Regression method. Where in this study, rainfall is the dependent variable, while temperature and humidity are independent variables. The results obtained from the WEKA Application with a total of 60 data from 2017 to 2021, the correlation coefficient value is 0.8175, and from the evaluation results using Linear Regression, the MAE error rate is 78.8695 and the RMSE is 95.1982. It can be concluded that the effect of temperature and air on the occurrence of rainfall is 81.75%.

Keywords: Rainfall, Prediction, Multiple Linear Regression, WEKA, Correlation Coefficient

Abstrak

Saat ini perubahan iklim di Indonesia yang merupakan wilayah tropis selalu tidak menentu dan mengakibatkan sulitnya dalam memprediksi kondisi cuaca. Kondisi cuaca dapat dipengaruhi oleh suhu, tekanan udara, kecepatan angin, kelembapan udara serta curah hujan. Curah hujan merupakan parameter iklim yang memiliki tingkat keragaman yang tinggi akibat anomali iklim. Ada beberapa faktor yang mempengaruhi karakteristik keragaman curah hujan yaitu faktor geografis, orografis, topografis, orientasi serta struktur kepulauannya. Faktor-faktor tersebut mengakibatkan pola sebaran dari curah hujan menjadi tidak merata antara suatu daerah dengan daerah lainnya. Untuk itu diperlukan suatu metode yang dapat menyelesaikan masalah dalam memprediksi curah hujan baik harian, bulanan maupun tahunan. Prediksi curah hujan dengan pendekatan statistik dapat dilakukan melalui metode Regresi Linear Berganda. Dimana dalam penelitian ini curah hujan menjadi variabel dependen sedangkan suhu dan kelembapan udara menjadi variabel independen. Didapatkan hasil dari Aplikasi WEKA dengan jumlah data 60 dari tahun 2017 sampai dengan tahun 2021, diperoleh nilai koefisien korelasi sebesar 0,8175, dan dari hasil Evaluasi menggunakan Regresi Linear menunjukkan tingkat galat MAE sebesar 78,8695 dan RMSE sebesar 95,1982. Dapat ditarik kesimpulan bahwa pengaruh suhu dan udara terhadap terjadinya curah hujan adalah 81,75%.

Kata kunci: Curah Hujan, Prediksi, Regresi Linear Berganda, WEKA, Koefisien Korelasi.

1. PENDAHULUAN

Negara Indonesia berada di daerah tropis serta tergolong sebagai negara yang kebanyakan daerahnya memiliki iklim tropis basah. Saat ini



perubahan iklim di Indonesia menjadi tidak menentu. Saat cuaca di daerah terlihat cerah, tiba-tiba dapat berubah menjadi hujan atau bahkan badai. Perubahan iklim yang tidak menentu tersebut yang mengakibatkan sulitnya dalam memprediksi kondisi cuaca [1]. Kondisi cuaca dipengaruhi oleh suhu, tekanan, kecepatan angin, kelembapan udaran serta curah hujan.

Curah hujan adalah parameter iklim yang memiliki tingkat keragaman yang tinggi akibat anomali iklim. Anomali iklim menjadikan curah hujan sangat sulit untuk diketahui. Ada beberapa faktor yang mempengaruhi karakteristik keragaman curah hujan tersebut, yaitu faktor geografis, orografis, topografis, orientasi serta struktur kepulauannya [2]. Faktor-faktor tersebut mengakibatkan pola sebaran dari curah hujan menjadi tidak merata antara suatu daerah dengan daerah lainnya [3].

Kondisi curah hujan yang tidak menentu dan fluktuatif membuat prediksi cuaca di Indonesia menjadi sulit, termasuk di Kota Sorong yang berada pada koordinat Garis Lintang $00^{\circ}51'$ Selatan dan Garis Bujur $131^{\circ}15'$ Timur. Adanya curah hujan berlebihan akan mengakibatkan banjir dan tanah longsor. Dampak terparah pernah terjadi di Kota Sorong pada tahun 2020 lalu, dimana terjadi bencana longsor yang mengakibatkan 3 (tiga) orang meninggal dunia dan bencana banjir yang merendam 3 (tiga) kecamatan di Kota Sorong yaitu Kecamatan Sorong (Kelurahan Klademak, Kelurahan Remu Utara), Kecamatan Sorong Utara (Kelurahan Malangke di) dan Sorong Manoi (Kelurahan Malabutor) - sumber : bnpb.go.id [4].

Untuk itu diperlukan suatu inovatif untuk menyelesaikan masalah dalam memprediksi curah hujan. Diantaranya melalui pendekatan statistik [5]. Prediksi curah hujan dapat menggunakan pendekatan statistik melalui metode regresi linear berganda. Metode Regresi Linear Berganda merupakan metode yang banyak diaplikasikan dalam proses prediksi yang menggunakan data dalam jumlah besar serta jumlah variabel yang banyak sehingga dapat digunakan untuk proses prediksi seperti metode regresi pada pendekatan statistik.

Berbagai penelitian terkait prediksi cuaca maupun curah hujan telah banyak dilakukan oleh penelitian sebelumnya, diantaranya oleh [6] yang melakukan analisis dan prediksi cuaca menggunakan parameter temperatur, kecepatan angin, arah angin, tekanan atmosphere, kelembapan dan curah hujan. Hasil akurasi prediksi yang didapatkan menggunakan metode Multi Linear Regression sebesar 88%.

Selain itu oleh [7], yang melakukan analisis curah hujan di Semarang dengan dataset yang diperoleh dari BMKG stasiun Meteorologi Ahmad Yani Semarang periode tahun 2015 hingga 2017. Variabel independent yang digunakan terdiri atas suhu, kelembapan, tekanan, dan kecepatan angin, sedangkan variabel dependent yang digunakan adalah curah hujan. Didapatkan hasil uji hipotesis korelasi keempat variabel independent terhadap variabel dependent yaitu nilai koefisien determinasi (R^2) sebesar 25.5% sedangkan 74.5% dipengaruhi oleh variabel lainnya.

Penelitian [2] [3] [8] menggunakan dua metode persamaan regresi yaitu model regresi linear sederhana dan regresi linear berganda dalam melakukan prediksi curah hujan bulanan. Parameter yang digunakan adalah suhu udara (T) dan kelembapan udara (RH). Untuk penelitian [2] didapatkan hasil perbandingan perhitungan nilai korelasi pearson dan penyimpangan prediksi total hujan bulanan terhadap total hujan aktual, yaitu Nilai korelasi $r = 0.72$ dan $RMSE = 77.42$ mm/bulan untuk variable predictor suhu udara, Nilai korelasi $r = 0.73$ dan $RMSE = 77.13$ mm/bulan untuk variable predictor kelembapan udara dan Nilai korelasi $r = 0.70$ dan $RMSE = 80.53$ mm/bulan untuk variable predictor suhu udara dan kelembapan udara.

Sedangkan penelitian [3] didapatkan hasil evaluasi simulasi prediksi total hujan bulanan yaitu $r = 0.66$ dengan menggunakan predictor suhu udara, $r = 0.58$ menggunakan predictor kelembapan udara, dan $r = 0.49$ untuk predictor suhu dan kelembapan udara sekaligus. Untuk hasil evaluasi menggunakan RMSE didapatkan predictor suhu memiliki nilai RMSE sebesar 14.41 mm/bulan, predictor kelembapan udara sebesar 14.35 mm/bulan dan predictor suhu serta kelembapan sekaligus menghasilkan nilai sebesar 14.55 mm/bulan.

Penelitian [8] mendapatkan hasil prediksi berupa nilai koefisien korelasi yang kuat sebesar $r = 0.90$ dan RMSE yang baik sebesar 40.96% antara curah hujan dan kelembapan udara terhadap prediksi curah hujan. Berdasarkan beberapa penelitian diatas, maka penulis mencoba melakukan Analisis Prediksi Curah Hujan di Kota Sorong dengan menggunakan metode metode *Multiple Linear Regression*.

2. METODOLOGI PENELITIAN

2.1. Peramalan

Peramalan adalah tindakan memprediksi secara metodis sesuatu yang paling mungkin terjadi di masa depan berdasarkan pengetahuan sebelumnya dan saat ini sehingga kesalahan (selisih antara apa yang terjadi dan apa yang diharapkan) dapat dikurangi. Prediksi tidak harus menawarkan jawaban secara pasti melainkan berusaha untuk mencari jawaban sedekat mungkin terhadap sesuatu yang akan terjadi [9]. Prediksi atau peramalan dalam data mining bisa dilakukan dengan metode time series dimana data dikumpulkan selama periode waktu tertentu seperti harian, mingguan, bulanan, triwulan, tahunan [10], sehingga data ini bisa memprediksi berbagai domain termasuk bidang klimatologi diantaranya prediksi curah hujan. Prediksi curah hujan merupakan aspek penting dari prakiraan iklim, dimana hasil akurat dan tepat waktu akan membantu untuk perencanaan dan pengelolaan sumber daya air, peringatan banjir, kegiatan konstruksi, operasi penerbangan, dan lain sebagainya.

2.2. Curah Hujan

Curah hujan merupakan parameter iklim yang perilaku atau gejalanya tampak jelas akibat anomali iklim. Anomali iklim seringkali terjadi berulang berdasarkan ketergantungan dari dinamika atmosfer. Indikator dominan yang dapat digunakan untuk melihat gejala terjadinya suatu anomali iklim yaitu kelembapan udara dan Suhu Udara [11]. Curah hujan dapat diukur menggunakan alat penakar hujan dengan satuan millimeter (mm). Curah hujan dengan ukuran 1 mm setara dengan jumlah 1 liter/ m^2 [12].

2.3. Suhu Udara

Suhu udara merupakan suatu keadaan derajat panas atau dinginnya udara. Dapat diukur dalam skala Celcius ($^{\circ}C$), Reamur ($^{\circ}R$), atau Fahrenheit ($^{\circ}F$). Suhu udara suatu wilayah dapat diukur berdasarkan dua kondisi, yaitu saat suhu udara minimum dan saat suhu udara maksimum [3]. Suhu rata-rata pada suatu wilayah di permukaan bumi dapat dihitung menggunakan rumus berikut ini:

$$T_x = T_0 - 0.6 \frac{h}{100} \quad (1)$$

Dimana

T_x = suhu rata-rata pada waktu x

T_0 = suhu awal pada suatu tempat

h = ketinggian pada suatu tempat dan 0.6 serta 100 adalah konstanta vertikal gradient temperatur (*lapce rate*), artinya setiap kenaikan 100 meter, suhu udara turun menjadi rata-rata 0.6 $^{\circ}C$.

Sedangkan untuk menghitung suhu udara permukaan rata-rata harian dapat menggunakan untuk data suhu udara rata-rata pengamatan pada pukul 07.00 LT, pukul 13.00 LT dan pukul 18.00 LT [3]. Untuk rumusnya dapat dilihat dibawah ini berikut:

$$T_{mean} = \frac{2 \times T_7 + T_{13} + T_{18}}{4} \quad (2)$$

Dimana:

T_{mean} = suhu udara permukaan rata-rata harian ($^{\circ}$)

T_7 = Suhu udara pengamatan jam 07.00 LT

T_{13} = Suhu udara pengamatan jam 13.00 LT

T_{18} = Suhu udara pengamatan jam 18.00 LT

Tinggi rendahnya suhu suatu permukaan dapat mempengaruhi laju penguapan, sehingga jumlah air yang menguap akan semakin banyak [13]. Suhu permukaan berbanding terbalik terhadap kelembapan. Semakin tinggi suhu suatu tempat maka kelembapan di daerah tersebut semakin rendah dan sebaliknya, apabila suhu daerah tersebut rendah maka kelembapan daerah tersebut menjadi tinggi. Antara suhu dan kelembapan sama-sama memiliki pengaruh terhadap curah hujan [14].

2.4. Kelembaban Udara

Kelembaban udara merupakan jumlah uap air yang terdapat dalam udara atau atmosfer. Uap air merupakan sumber bentuk kondensasi yang dapat menyerap dengan baik radiasi sinar matahari dan radiasi bumi. Uap air dapat berubah bentuk menjadi cair atau pun pada yang kemudian jatuh ke bumi sebagai hujan [15]. Kelembaban terbagi menjadi dua jenis yaitu :

- a) Kelembaban udara absolut, yaitu jumlah uap air yang terkandung dalam udara dan dinyatakan dengan satuan gram uap air untuk $1 m^3$ udara
- b) Kelembaban udara relatif (nisbi), yaitu perbandingan banyaknya jumlah uap air dengan banyaknya jumlah air maksimum dalam udara untuk suhu yang sama dan dinyatakan dalam persentase (%). Berikut adalah persamaan kelembaban udara

$$R_H = \frac{P_A}{P_s} \times 100\% \tag{3}$$

Dimana:

R_H = kelembaban relatif (%)

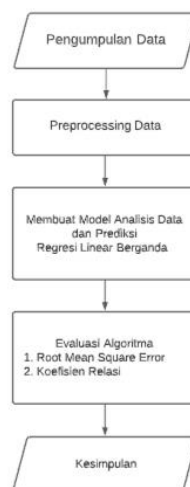
P_A = tekanan uap air pada saat pengukuran (N/m^2)

P_s = tekanan uap air jenuh (N/m^2)

Kelembaban relatif (nisbi) adalah parameter penting yang patut diperhatikan dan untuk mengukurnya secara akurat dapat digunakan psychrometer.

2.5. Tahap Penelitian

Tahapan penelitian yang dilakukan dimulai dari pengumpulan dataset. Selanjutnya dilakukan preprocessing dataset. Data-data yang tidak diperlukan dalam penelitian dihilangkan. Kemudian dilanjutkan dengan pembuatan model analisis data serta evaluasi hasil dari model yang dibuat. Terakhir adalah penarikan kesimpulan dan saran dari hasil penelitian yang dilakukan. Berikut adalah Gambar 1. terkait alur Tahapan Penelitian:



Gambar 1. Tahapan Penelitian

2.5.1. Pengumpulan Dataset

Dataset yang digunakan dalam penelitian ini diperoleh dari Kantor BMKG DEO Sorong, yang terdiri dari: (1) Data rerata kelembapan udara bulanan, (2) data rerata suhu udara bulanan, dan (3) data total curah hujan bulanan. Data rerata bulanan merupakan hasil dari jumlah data rata-rata perhari selama satu bulan dan kemudian dibagi dengan banyaknya data pada bulan yang bersangkutan [3]. Yang menjadi variabel dependent adalah data total curah hujan bulanan. Dan yang lainnya menjadi variabel independent. Data suhu udara dan kelembapan udara merupakan dataset BMKG DEO Sorong yang didapat dari tahun 2017 hingga 2021 (5 Tahun). Berikut Tabel 1. terkait Kode dan Keterangan Variabel.

Tabel 1. Kode dan Keterangan Variabel

No	Keterangan	Variabel	Kode
1	Rerata kelembapan udara bulanan	Variabel Independent	X_1
2	Rerata suhu udara bulanan	Variabel Independent	X_2
3	Total curah hujan bulanan	Variabel Dependent	Y

2.5.2. Preprocessing Data

File data yang didapatkan dari BMKG berbentuk xls yang kemudian diubah menjadi bentuk CSV (Comma Separate Value). Data yang didapatkan dari BMKG berbentuk file terpisah untuk masing-masing variabel. Kemudian file-file data tersebut disatukan menjadi satu file. Selanjutnya proses cleaning data terhadap missing value dan normalisasi untuk menyetarakan nilai untuk setiap attribut dan kelas. Hasil proses ini, didapatkan data sebanyak 60 data.

2.5.3. Permodelan Data

Setelah preprocessing data, selanjutnya data dianalisis menggunakan model Regresi linier berganda (*Multiple Linear Regression*). MLR merupakan salah metode prediksi yang dilakukan melalui pembentukan persamaan regresi untuk simulasi prediksi curah hujan bulanan dengan menggunakan lebih dari satu variabel independen [8]. Metode regresi linear berganda menggunakan dua jenis variabel, yaitu variabel bebas atau variabel prediktor (independent) dan variabel tidak bebas atau variabel respon (dependent). Adapun persamaan umum dari metode ini [16] yaitu:

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_kX_k \quad (4)$$

Dimana:

B_0 = konstanta

B = koefisien variabel X_1, X_2, \dots, X_k

Y = variabel yang diduga (variabel dependen);

X_i = variabel penduga (variabel independen).

Variabel bebas dapat memprediksi nilai variabel tidak bebas dan dinyatakan dengan X_1, X_2, \dots, X_k . Sedangkan variabel tidak bebas merupakan variabel yang terjadi karena respon dari variabel bebas dan dinyatakan

dengan Y [17]. Dalam penelitian ini, yang menjadi variabel prediktor atau variabel bebas yaitu suhu udara, kelembaban udara, kecepatan angin, dan tekanan udara. Sedangkan variable tidak bebas yaitu total hujan.

2.5.4. Evaluasi Data

Setelah didapatkan model atau persamaan hasil pengolahan data, kemudian dilakukan evaluasi. Evaluasi digunakan untuk menguji kelayakan serta relevansi keterkaitan antara variable independent dan variable dependent. Evaluasi algoritma yang digunakan dalam penelitian ini adalah Root Mean Square Error (RMSE) dan Koefisien Relasi. RMSE merupakan metode evaluasi yang digunakan untuk mengetahui perbedaan antara nilai-nilai prediksi antara model dengan nilai kebenaran yang diamati [8].

Melalui RMSE dapat diketahui besarnya penyimpangan yang terjadi antara nilai prediksi curah hujan dengan nilai curah hujan actual. Semakin besar nilai RMSE maka semakin jauh nilai data total hujan prediksi terhadap total hujan actual. Semakin kecil nilai RMSE maka semakin baik prediksi total hujan terhadap total hujan actual. Tingkat kesalahan yang dapat diminimalisir dapat meningkatkan akurasi prediksi [18]. Berikut adalah persamaan Root Mean Squar Error

$$RMSE = \left[\frac{1}{N} \sum_{n=1}^N (f_n - r_n)^2 \right]^{\frac{1}{2}} \quad (5)$$

Dimana:

- f_n = Data curah hujan hasil Obeservasi
- r_n = Data curah hujan hasil prediksi

Nilai Koefisien Korelasi digunakan untuk menentukan besarnya hubungan kedekatan antara total prediksi curah hujan terhadap total huja actual. Kedekatan tersebut didapatkan dari hasil prediksi menggunakan variabel prediktor diantara suhu dengan kelembapan udara. Nilai koefisien memiliki rentang nilai paling sedikit -1 dan paling besar 1, dan dapat dinyatakan sebagai berikut:

- a) Jika nilai r mendekati +1, artinya hubungan antara prediksi total hujan dengan total hujan actual sangat kuat dan positif
- b) Jika nilai r mendekati - 1, artinya hubungan antara prediksi total hujan dengan total hujan actual sangat sangat kuat dan negatif
- c) Jika nilai r +0.5 hingga 0,5, artinya hubungan antara prediksi total hujan dengan total hujan actual cukup kuat
- d) Jika nilai r lebih kecil +0.5 atau lebih besar -0.5, artinya hubungan antara prediksi total hujan dengan total hujan actual dianggap lemah

Validasi nilai koefisien korelasi yaitu semakin kuat korelasi maka hasil validasi semakin baik sehingga tingkat akurasi perkiraan akan semakin kecil. Berikut adalah persamaan untuk menghitung koefisien korelasi

$$r = \sqrt{1 - \frac{\Sigma(Y-\bar{Y})^2}{\Sigma(Y-Y^*)^2}} \quad (6)$$

Dengan:

r = nilai korelasi

Y = data curah hujan bulanan

Y^* = data curah hujan bulanan hasil prediksi

\bar{Y} = rata-rata data curah hujan actual

3. HASIL DAN PEMBAHASAN

Pada Penelitian ini Penulis melakukan pengujian menggunakan alat bantu atau tools data mining yaitu WEKA 3.8.4 (Waikatu Environment for Knowledge Analysis). Weka merupakan sistem penting dalam sejarah penambangan data dan komunitas penelitian machine learning, karena sistem ini adalah satu-satunya toolkit yang telah diadopsi secara luas dan bertahan selama jangka waktu yang lama [19]. Disini Weka digunakan sebagai alat bantu untuk melakukan analisis prediksi curah hujan menggunakan algoritma Linear Regression, sehingga dapat mengidentifikasi klasifikasi iklim di wilayah Kota Sorong, proses prediksi dilakukan untuk mendapatkan seberapa kuat korelasi antara variabel-variabel independen dengan variabel dependennya.

3.1. Pengumpulan Dataset

Dataset diperoleh dari BMKG DEO Sorong, dimana datanya dari tahun 2017 hingga 2021 (5 Tahun) berupa (1) Data rerata suhu udara bulanan, (2) data rerata kelembapan udara bulanan, dan (3) data total curah hujan bulanan. Gambar 2. berikut memperlihatkan dataset yang akan digunakan dalam proses pengolahan data di aplikasi WEKA:

	A	B	C	D	E
1	No.	Bulan	Suhu	Kelembapan	Curah Hujan
2	1	Januari 2017	27	88	128
3	2	Februari 2017	27	85	372
4	3	Maret 2017	27	85	338
5	4	April 2017	28	85	241
6	5	Mei 2017	27	87	528
7
58	56	Agustus 2021	27	88	511
59	57	September 2021	27	89	519
60	58	Oktober 2021	28	85	180
61	59	November 2021	27	85	287
62	60	Desember 2021	27	87	330

Gambar 2. Tabel Dataset

Tabel Kolom Suhu adalah X_1 dan Kelembapan adalah X_2 yang akan menjadi Variabel Independen, sedangkan Tabel Kolom Curah_Hujan adalah Y yang akan menjadi Variabel Dependen.

3.2. Preprocessing Dataset

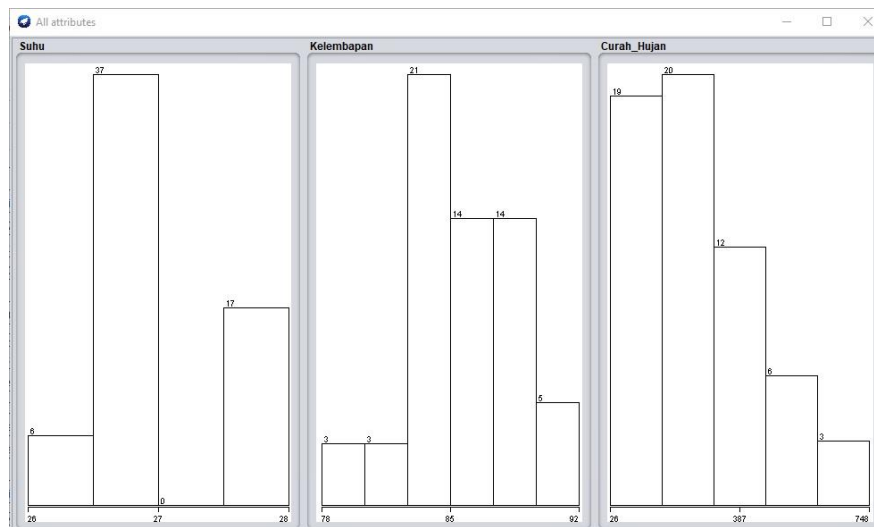
Selanjutnya dilakukan *Preprocessing Data* yaitu merubah data mentah menjadi data yang bisa diolah, dimana pada proses ini file format xls diubah menjadi csv, dimana sebelumnya terdapa kolom 'No.' dan 'Bulan'

dihilangkan, sehingga file csv dapat dilihat seperti pada gambar 3. dibawah ini:

Suhu,Kelembapan,Curah_Hujan			
27,88,128			
27,85,372			
27,85,338			
28,85,241			
27,87,528			
...			
27,88,511			
27,89,519			
28,85,180			
27,85,287			
27,87,330			

Gambar 3. Dataset diubah ke csv

Data CSV yang telah dibuat kemudian dibuka pada aplikasi WEKA dan didapatkan hasil visualisasi grafik seperti pada gambar 4. berikut:



Gambar 4. Hasil Visualisasi Dataset

3.3. Permodelan Dataset

Kemudian dilakukan permodelan data pada aplikasi WEKA menggunakan Algoritma Regresi Linear Berganda (*Multiple Linear Regression*) dengan melakukan test mode dimana sebanyak 80% data dilakukan *split training* (Rasio 80:20) dan dapat dilihat seperti pada gambar 5. berikut:

```

=== Run information ===

Scheme:      weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8 -num-decimal-places 4
Relation:    DatasetCuaca_5 tahun_suhu dan kelembaban
Instances:   60
Attributes:  3
              Suhu
              Kelembapan
              Curah_Hujan
Test mode:   split 80.0% train, remainder test

=== Classifier model (full training set) ===
    
```

Gambar 5. Informasi Permodelan Data

Maka hasil yang diperoleh dari permodelan data menggunakan algoritma Linear Regression dimana training split menggunakan 48 jumlah data dari 60 data yang ada, didapatkan persamaan regresi linear dan hasil prediksi curah hujan seperti yang terlihat pada gambar 6. berikut:

```

=== Classifier model for training split (48 instances) ===

Linear Regression Model

Curah_Hujan =

    35.3748 * Kelembapan +
    -2775.7634
=== Predictions on test split ===
    
```

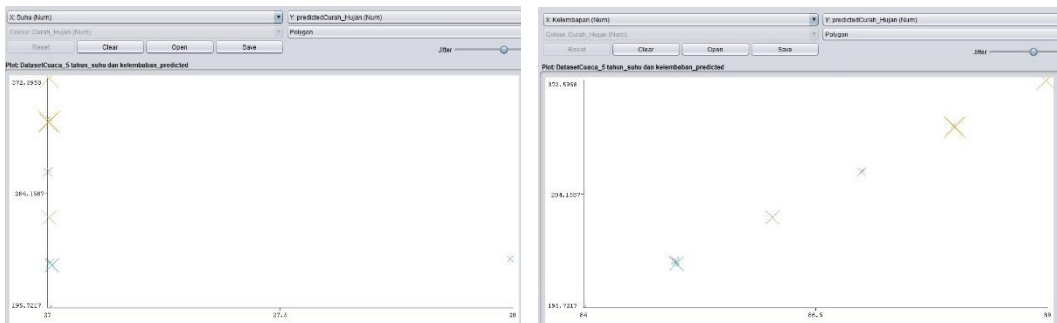
} Persamaan Regresi Linier
untuk

inst#	actual	predicted	error
1	377	266.471	-110.529
2	175	195.722	20.722
3	236	231.097	-4.903
4	121	231.097	110.097
5	233	301.846	68.846
6	478	337.221	-140.779
7	302	337.221	35.221
8	511	337.221	-173.779
9	519	372.596	-146.404
10	180	231.097	51.097
11	287	231.097	-55.903
12	330	301.846	-28.154

} Hasil Prediksi untuk 12 data sisa yang diuji dari 48 data yang di training.

Gambar 6. Hasil Permodelan Data dan Prediksi

Gambar 6. berikut merupakan data grafik yang dihasilkan Prediksi Curah Hujan terhadap hubungan korelasi dengan variabel Suhu dan Kelembapan:



Gambar 6. Plot Grafik Hubung

3.4. Evaluasi Dataset

Dari hasil permodelan data pada gambar 5. diatas maka didapatkan Hasil Ringkasan Evaluasi seperti yang terlihat pada gambar 7. berikut:

```

=== Summary ===
Correlation coefficient           0.8175
Mean absolute error              78.8695
Root mean squared error          95.1982
Relative absolute error          70.821 %
Root relative squared error      68.5683 %
Total Number of Instances        12
  
```

Gambar 7. Hasil Ringkasan

Gambar 7. diatas memperlihatkan bahwa didapatkan Nilai Korelasi Koefisien sebesar 0,8175 dengan tingkat galat MAE (*Mean Absloute Error*) sebesar 78,8695 dan RMSE (*Root Mean Squared Error*) mencapai 95,1982.

4. SIMPULAN

Penelitian ini melakukan prediksi curah hujan di wilayah Kota Sorong menggunakan algoritma Regresi Linear Berganda (*Multiple Linear Regression*) Dari data yang didapatkan berupa Suhu dan Kelembapan wilayah Kota Sorong selama 5 tahun dari tahun 2017 sampai dengan tahun 2021. Analisa kinerja data menggunakan bantuan *tools* aplikasi WEKA didapatkan 3 (tiga) ukuran akurasi yaitu Nilai Korelasi Koefisien, Nilai MAE dan Nilai RMSE. Teknik Klasifikasi Linear Regression pada dataset sebanyak 60 data menggunakan 80% data latih dan 20% data uji. Didapatkan hasil Nilai Korelasi Koefisien sebesar 0,8175, Nilai MAE (*Mean Absolute Error*) sebesar 78,8695 dan Nilai RMSE (*Root Mean Squared Error*) sebesar 95,1982. Besarnya nilai RMSE yang mencapai 95,1982 tidak bagus terhadap proses prediksi curah hujan karena besarnya penyimpangan (*error*) antara nilai prediksi curah hujan dan nilai curah hujan actual, namun dari hasil besarnya nilai korelasi koefisien (*r*) sebesar 0,8175 menandakan hubungan yang sangat kuat dan positif antara prediksi total hujan dengan total hujan actual, hal ini bisa dilihat dari grafik plot di gambar 6 (hasil dan pembahasan) diatas, dimana suhu dan kelembapan memiliki pengaruh yang signifikan atau kuat dalam memprediksi terjadinya curah hujan walaupun dalam persamaannya kelembapan udara paling signifikan atau kuat, dimana pengaruhnya sebesar 81,75%

DAFTAR PUSTAKA

- [1] B. Suma, "Implementasi Machine Learning Di Dalam Prediksi Cuaca Disusun sebagai salah satu syarat untuk kelulusan Program Strata 1 di Program Studi Teknik Informatika Universitas Pasundan Bandung oleh : BANDUNG SEPTEMBER 2020," no. September 2020, 2021, doi: 10.13140/RG.2.2.16086.47680.

- [2] I. Jhonson Arizona Saragih *et al.*, "Prediksi Curah Hujan Bulanan Di Deli Serdang Menggunakan Persamaan Regresi Dengan Prediktor Data Suhu Dan Kelembapan Udara," *J. Meteorol. Klimatologi dan Geofis.*, vol. 7, no. 2, pp. 6–14, 2020.
- [3] A. Fadholi, "Persamaan regresi prediksi curah hujan bulanan menggunakan data suhu dan kelembapan udara di Ternate," *J. Stat.*, vol. 13, no. 1, pp. 7–16, 2013.
- [4] R. Jati, "Banjir dan Longsor Kota Sorong Akibatkan Tiga Orang Meninggal Dunia," 2020. [Online]. Available: [https://bnpb.go.id/berita/Banjir dan Longsor Kota Sorong Akibatkan Tiga Orang Meninggal Dunia](https://bnpb.go.id/berita/Banjir%20dan%20Longsor%20Kota%20Sorong%20Akibatkan%20Tiga%20Orang%20Meninggal%20Dunia). [Accessed: 01-Jan-2022].
- [5] H. Muhamad, I. Cholissodin, and B. D. Setiawan, "Optimasi Support Vector Regression (SVR) Menggunakan Algoritma Improved-Particle Swarm Optimization (IPSO) untuk Peramalan Curah Hujan," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 11, pp. 2548–964, 2017.
- [6] N. Anusha, M. Sai Chaithanya, and G. Jithendranath Reddy, "Weather Prediction Using Multi Linear Regression Algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 590, no. 1, 2019, doi: 10.1088/1757-899X/590/1/012034.
- [7] A. Luthfiarta, A. Febriyanto, H. Lestiawan, and W. Wicaksono, "Analisa Prakiraan Cuaca dengan Parameter Suhu, Kelembaban, Tekanan Udara, dan Kecepatan Angin Menggunakan Regresi Linear Berganda," *JOINS (Journal Inf. Syst.*, vol. 5, no. 1, pp. 10–17, 2020, doi: 10.33633/joins.v5i1.2760.
- [8] E. Dewi *et al.*, "Prediksi Curah Hujan Di Kabupaten Majalengka Dengan Menggunakan Algoritma Regresi Rainfall Prediction in Majalengka District Using Regression Algorithm," *Inf. [?]*, vol. 67, no. 1, 2019.
- [9] M. Kafil, "Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 3, no. 2, pp. 59–66, 2019, doi: 10.36040/jati.v3i2.860.
- [10] S. Aftab, M. Ahmad, N. Hameed, M. S. Bashir, I. Ali, and Z. Nawaz, "Rainfall prediction in Lahore City using data mining techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 4, pp. 254–260, 2018, doi: 10.14569/IJACSA.2018.090439.
- [11] S. Putramulyo and S. Alaa, "Prediksi Curah Hujan Bulanan Di Kota Samarinda Menggunakan Persamaan Regresi Dengan Prediktor Data Suhu dan Kelembapan Udara," *Eig. Math. J.*, vol. 2, no. 2, p. 13, 2018, doi: 10.29303/emj.v2i2.20.
- [12] N. Binternagel, "Adaptasi dan Mitigasi Perubahan Iklim Global," no. October, 2009.
- [13] A. Ardhitama, "Simulasi Prakiraan Jumlah Curah Hujan Dengan Menggunakan Data Parameter Cuaca (Study Kasus Di Kota Pekanbaru Tahun 2012)," *J. Sains Teknol. Modif. Cuaca*, vol. 14, no. 2, p. 111, 2013,

doi: 10.29122/jstmc.v14i2.2690.

- [14] D. Prakoso, "Analisis Pengaruh Tekanan Udara, Kelembapan Udara Dan Suhu Udara Terhadap Tingkat Curah Hujan Di Kota Semarang," pp. 1-77, 2018.
- [15] Y. S. Swarinoto and S. Sugiyono, "Pemanfaatan Suhu Udara Dan Kelembapan Udara Dalam Persamaan Regresi Untuk Simulasi Prediksi Total Hujan Bulanan Di Bandar Lampung," *J. Meteorol. dan Geofis.*, vol. 12, no. 3, pp. 271-281, 2011, doi: 10.31172/jmg.v12i3.109.
- [16] P. S. A. Husaini Usman, *Pengantar statistika*, Ed. 2. Jakarta: Bumi Aksara, 2006., 2006.
- [17] Sudjana, *Metode statistika*. Bandung : Tarsito, 2002.
- [18] I. R. Rumahorbo, "Persamaan Regresi Prediksi Curah Hujan Bulanan Menggunakan Data Suhu dan Kelembapan Udara di Bengkulu," *Pros. Semin. Pendidik. Mat. dan ...*, vol. 2, no. 2721, pp. 1-11, 2020.
- [19] I. Russell and Z. Markov, "An Introduction to the Weka Data Mining System (Abstract Only)," vol. 18, no. 3, pp. 742-742, 2017, doi: 10.1145/3017680.3017821.