

**PREDIKSI CURAH HUJAN MENGGUNAKAN ALGORITMA
REGRESI LINEAR BERGANDA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

AFIFAH NUR LATIFAH

19.11.2889

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**PREDIKSI CURAH HUJAN MENGGUNAKAN ALGORITMA
REGRESI LINEAR BERGANDA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

AFIFAH NUR LATIFAH
19.11.2889

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

HALAMAN PERSETUJUAN

SKRIPSI

**PREDIKSI CURAH HUJAN MENGGUNAKAN ALGORITMA
REGRESI LINEAR BERGANDA**

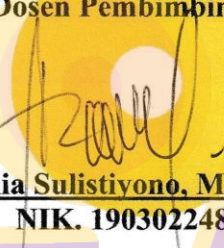
yang disusun dan diajukan oleh

Afifah Nur Latifah

19.11.2889

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 17 Mei 2023

Dosen Pembimbing,


(Mulia Sulistiyono, M.Kom)

NIK. 190302248

HALAMAN PENGESAHAN

SKRIPSI

**PREDIKSI CURAH HUJAN MENGGUNAKAN ALGORITMA
REGRESI LINEAR BERGANDA**

yang disusun dan diajukan oleh

Afifah Nur Latifah

19.11.2889

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Mei 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Arif Dwi Laksito, M.Kom.

NIK. 190302150



Windha Mega Pradnya D, M.Kom

NIK. 190302185



Mulia Sulistiyono, M.Kom

NIK. 190302248



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Mei 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.

NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Afifah Nur Latifah
NIM : 19.11.2889

Menyatakan bahwa Skripsi dengan judul berikut:

Prediksi Curah Hujan Menggunakan Algoritma Regresi Linear Berganda

Dosen Pembimbing : Mulia Sulistiyono, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 17 Mei 2023

Yang Menyatakan,



Afifah Nur Latifah

HALAMAN PERSEMBAHAN

Alhamdulillah robbil 'alamin, segala puji bagi Allah SWT dengan kemurahan dan ridho-Nya, skripsi ini dapat ditulis dengan baik dan lancar hingga selesai. Dengan ini dipersembahkan skripsi ini secara khusus kepada :

1. Kedua orang tuaku, Alim Totoyani (Ayah) dan Setyo Puji Rukmini (ibu) yang tak henti-hentinya selalu memberikan kasih sayang, motivasi, dan cintanya kepadaku, yang selalu memberikan semangat untuk mewujudkan cita-citaku. Terima kasih atas doa dan dukungannya yang selalu diberikan untukku. Terima kasih telah mengupayakan dan memberikan yang terbaik untukku serta menjadi penyemangat dalam hidupku.
2. Kedua adikku, Fitria Nur Aini dan Janu Diega Nur Raihan yang selalu menjadi alasanku untuk tetap berjuang hingga akhir.
3. Jihan Juki Okvianti, sahabat sekaligus keluarga untukku yang telah berkontribusi banyak dalam penulisan skripsi ini, meluangkan banyak waktu dan pikiran. Terimakasih telah menjadi tempat segala cerita saya, semoga tetap bisa menjadi saudara walaupun tidak ada aliran darah.
4. Teman-teman Informatika 05 yang telah memberikan bantuan dan kerjasamanya dari awal hingga akhir perkuliahan.

KATA PENGANTAR



Assalaamu'alaikum Warahmatullaahi Wabarakaatuh

Segala puji dan syukur penulis haturkan kepada Allah SWT yang telah melimpahkan rahmat, taufiq dan hidayah-Nya serta pertolongan-Nya. Sehingga skripsi yang berjudul “Prediksi Curah Hujan Menggunakan Algoritma Regresi Linear Berganda” ini dapat terselesaikan. Shalawat dan salam senantiasa tercurahkan kepada tauladan kita, Nabi Muhammad SAW, keluarganya, serta sahabatnya yang kita nanti-nantikan syafaatnya di hari akhir. Penulisan skripsi ini dilakukan guna memenuhi syarat kelulusan Mata Kuliah Skripsi serta memperoleh gelar Sarjana pada Program Studi S1 Informatika, Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Skripsi ini ditulis berdasarkan penelitian yang telah dilakukan oleh penulis. Dengan terselesainya dan berhasilnya pelaksanaan serta penulisan Skripsi ini, penulis ucapkan terima kasih kepada berbagai pihak yang telah membantu, memberikan bimbingan dan dukungan baik secara material maupun spiritual yang telah diberikan kepada penulis selama penulisan skripsi berlangsung, secara khusus kepada:

1. Bapak Hanif Al Fatta, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
2. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta.
3. Bapak Mulia Sulistiyono, M.Kom. selaku Dosen Pembimbing Akademik Universitas Amikom Yogyakarta yang telah bersedia memberikan berbagai arahan, bantuan, dan bimbingan.

4. Bapak Ibu Dosen pengajar serta penguji Universitas Amikom Yogyakarta yang telah memberikan ilmu dan pengalamannya dengan tulus, mengarahkan saya hingga dapat terselesaikan masa studi penulis sekarang ini.
5. Ibu Dwi Nurani, M.Kom selaku Dosen Wali yang telah memberikan berbagai arahan dan bantuan selama penulis menempuh studi S1 Informatika Universitas Amikom Yogyakarta.
6. Kedua Orang Tua, Keluarga, serta Sahabat yang senantiasa memberikan doa, dukungan, dan bantuan.
7. Teman-teman seperjuangan Informatika 05 yang telah dengan baik menerima dan memberikan berbagai bantuan kepada penulis.
8. Serta semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu dalam penyelesaian skripsi ini.

Penulis menyadari bahwa penulisan skripsi ini masih terdapat kekurangan dan jauh dari kesempurnaan. Oleh karena itu, sangat diharapkan saran dan kritik yang bersifat membangun demi kebaikan dan kemajuan penulis di masa mendatang karena bagaimanapun juga tidak ada gading yang tak retak.

Wassalaamu'alaikum Warahmatullaahi Wabarakaatuh

Yogyakarta, 13 April 2023

Penulis

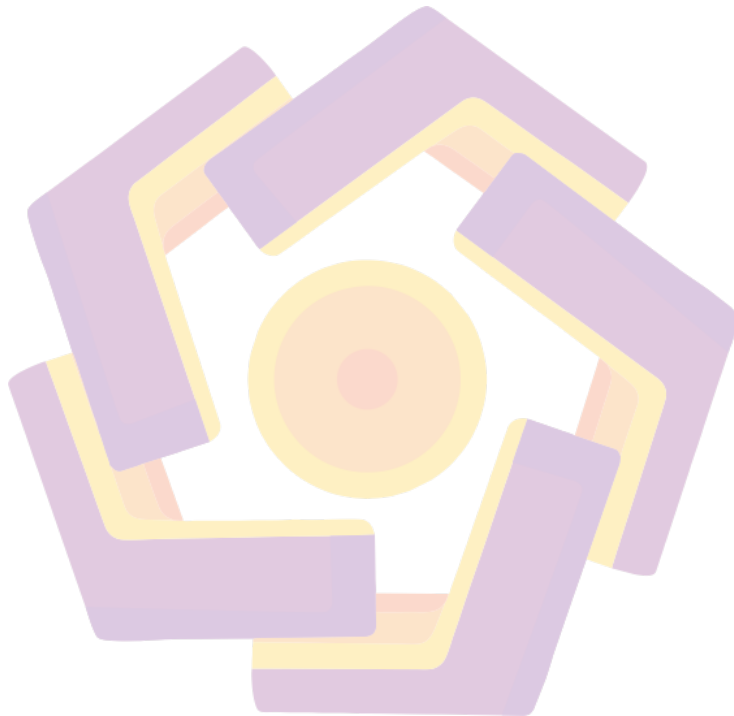
DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xv
INTISARI	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur	7
2.2 Dasar Teori	13
2.2.1 Peramalan.....	13
2.2.2 Curah hujan (mm)	13
2.2.3 Suhu udara (°C).....	13
2.2.4 Kelembaban udara (%).....	14
2.2.5 Kecepatan angin (knot)	15
2.2.6 Lama penyinaran (jam)	15

2.2.7 Probabilistik curah hujan	15
2.2.8 Regresi linear berganda.....	16
2.2.9 Variabel penelitian	17
2.2.10 Root Mean Square Error	17
2.2.11 Korelasi	18
2.2.12 Jenis-jenis hubungan korelasi	19
2.2.13 Koefisien korelasi	20
2.2.14 Interpretasi koefisien korelasi	21
2.2.15 Data Mining	21
2.2.16 Machine Learning	22
2.2.17 Jenis-jenis Data	23
2.2.18 Data Preparation.....	24
2.2.19 Data Preprocessing.....	25
2.2.20 Data Splitting	25
BAB III METODE PENELITIAN	26
3.1 Data Penelitian.....	26
3.2 Alur Penelitian	28
3.3 Alat dan Bahan.....	32
BAB IV HASIL DAN PEMBAHASAN	34
4.1 Hasil Penelitian.....	34
4.2 Pembahasan	43
BAB V PENUTUP	47
5.1 Kesimpulan	47
5.2 Saran	49
REFERENSI	50
LAMPIRAN.....	53

DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	9
Tabel 4.1. Penghitungan Manual untuk Y Predicted	45



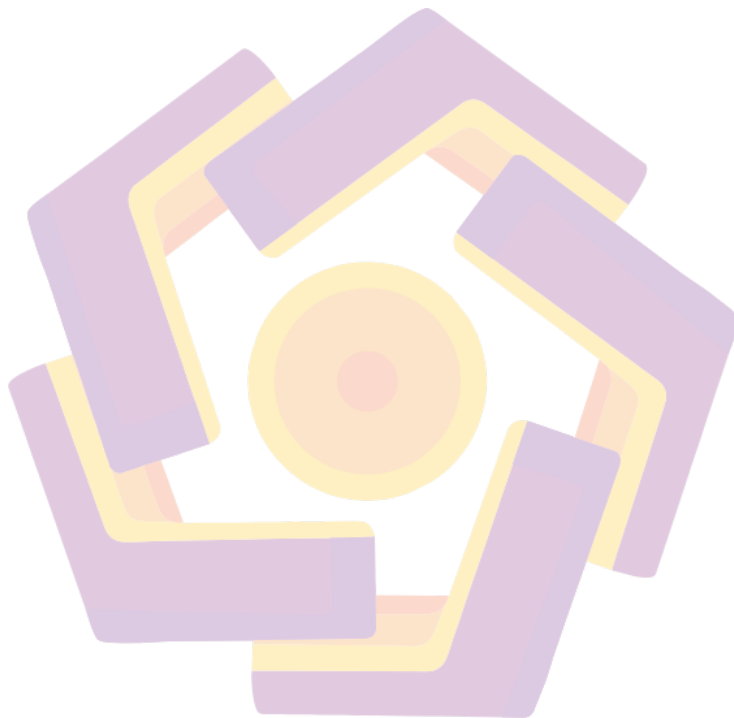
DAFTAR GAMBAR

Gambar 2.1. Jenis-Jenis Hubungan Korelasi	19
Gambar 2.2. Koefisien Korelasi	20
Gambar 3.1. Alur Penelitian	28
Gambar 4.1. Code Program Data Preparation	34
Gambar 4.2. Code Program Menghitung Jumlah Missing Value	35
Gambar 4.3. Code Program Menghapus Nilai NaN	35
Gambar 4.4. Code Program Melihat Detail Kolom dan Data	35
Gambar 4.5. Code Program Menghitung Jumlah Missing Value Kembali	36
Gambar 4.6. Code Program Menghapus Kolom	36
Gambar 4.7. Diagram Boxplot Variabel Suhu Rata-Rata	37
Gambar 4.8. Visualisasi Data Variabel Suhu Rata-Rata	37
Gambar 4.9. Diagram Boxplot Variabel Kelembaban Rata-Rata	38
Gambar 4.10. Diagram Boxplot Variabel Kelembaban Rata-Rata	38
Gambar 4.11. Diagram Boxplot Variabel Lama Peyinarian Matahari	38
Gambar 4.12. Diagram Boxplot Variabel Lama Peyinarian Matahari	39
Gambar 4.13. Diagram Boxplot Variabel Kecepatan Angin Rata-Rata	39
Gambar 4.14. Diagram Boxplot Variabel Kecepatan Angin Rata-Rata	40
Gambar 4.15. Persebaran Data Semua Variabel	40
Gambar 4.16. Korelasi Antar Variabel	40
Gambar 4.17. Code Program Data Splitting	41
Gambar 4.18. Code Program Membuat Object Linear Regresi	41
Gambar 4.19. Code Program Mencari Nilai Koefisien dan Intercept	42
Gambar 4.20. Code Program Mencari Nilai R2 Score	42
Gambar 4.21. Code Program Menghitung Nilai Accuracy	42
Gambar 4.22. Code Program Prediksi Curah Hujan	42
Gambar 4.23. Code Program Evaluasi RMSE	43

DAFTAR LAMPIRAN

Lampiran 1. Code Program Python

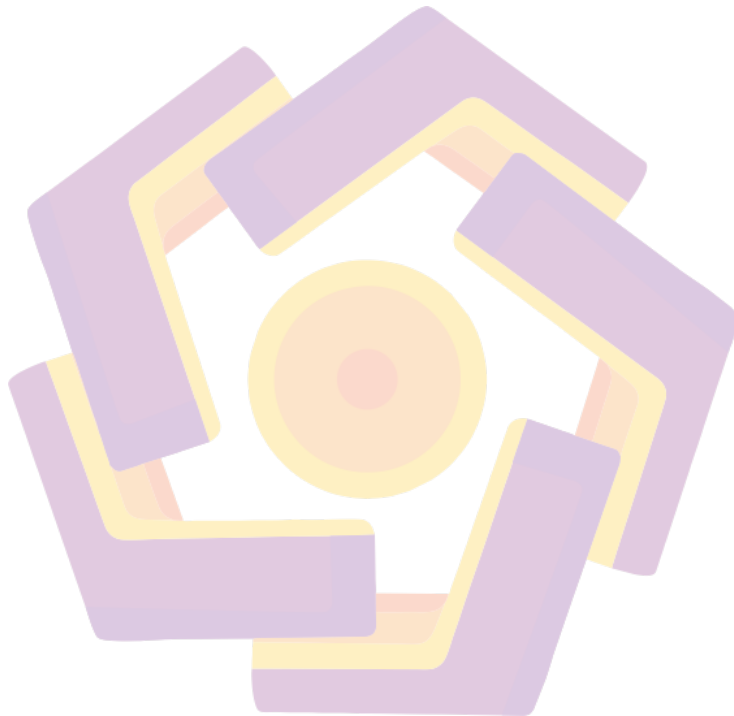
53



DAFTAR LAMBANG DAN SINGKATAN

LS	Lintang Selatan
BT	Bujur Timur
IPPC	International Plant Protection Convention
RMSE	Root Mean Square Error
R^2	Koefisien Determinasi
r	Koefisien Korelasi
WEKA	Waikatu Environment Knowledge Analysis
MAE	Mean Absolute Error
T_x	Suhu rata-rata pada waktu x
T_0	Suhu awal pada suatu tempat
h	Ketinggian pada suatu tempat
RH	Kelembaban relatif
P_A	Tekanan uap air pada saat pengukuran
P_s	Tekanan uap air jenuh
mm	Milimeter
C	Celcius
R	Reamur
F	Fahrenheit
MKGI	Metorologi, Klimatologi, Geofisika, dan Instrumentasi
Y_i	data curah hujan hasil observasi
\hat{Y}_i	data curah hujan hasil prediksi
N	jumlah data
Y	Data curah hujan actual
Y^*	Data curah hujan prediksi
\bar{Y}	Rata-rata data curah hujan actual
T_n	Suhu minimal
T_x	Suhu maksimal
ff_x	Kecepatan angin maksimal
ddd_x	Arah angin dengan kecepatan maksimal

ddd_car	Arah angin
station_id	Detail stasiun
Tavg	Suhu rata-rata
RH_avg	Kelembaban rata-rata
RR	Curah hujan
Ss	Lama penyinaran matahari
ff_avg	Kecepatan angin rata-rata
BMKG	Badan Meteorologi, Klimatologi, dan Geofisika



DAFTAR ISTILAH

RMSE	Evaluasi untuk mengukur tingkat akurasi hasil perkiraan model
MAE	Rata-rata selisih mutlak nilai sebenarnya dengan nilai prediksi
Scatter plot	Diagram pencar
Clustering	Proses mengelompokkan data
Classification	Proses menggolongkan data berdasarkan persamaan atau perbedaan
Assosiation	Proses menemukan suatu kombinasi item yang muncul bersamaan
Prediktor	Variabel yang digunakan untuk memprediksi nilai variabel terkait
Dependen	Terikat dan tidak berdiri sendiri
Independen	Bebas dan tidak terikat
Missing value	Informasi yang tidak tersedia untuk sebuah objek (kasus)
Data Training	Himpunan data yang digunakan untuk melatih model
Data Testing	Himpunan data yang digunakan untuk menguji model
Intensitas	Keadaan tingkatan atau ukuran intensitasnya

INTISARI

Indonesia merupakan wilayah tropis yang memiliki perubahan cuaca tidak menentu. Akibatnya, diperlukan adanya prediksi cuaca sebagai penentu cuaca yang akan terjadi kedepannya. Dengan dapat diperkirakan cuaca yang akan terjadi, dapat segera dilakukan adanya pencegahan supaya tidak terjadi hal-hal yang tidak diinginkan. Mengingat wilayah Yogyakarta merupakan wilayah yang memiliki bentang alam pegunungan dan dataran rendah sehingga dapat menyebabkan terjadinya curah hujan yang tidak sama diberbagai daerah. Kondisi curah hujan yang tidak menentu di berbagai daerah inilah yang membuat sulitnya untuk memprediksi cuaca.

Terdapat berbagai faktor yang berpengaruh terhadap curah hujan. Beberapa parameter yang mempengaruhi terjadinya curah hujan yaitu temperatur suhu, kelembaban, kecepatan angin, lama penyinaran matahari, dan arah. Dengan berbagai parameter data tersebut maka dapat dilakukan penelitian dan dicocokkan dengan prediksi cuaca sebelumnya sehingga menghasilkan prediksi curah hujan yang paling akurat untuk hari ini. Prediksi curah hujan ini dilakukan menggunakan algoritma Regresi Linear Berganda dimana curah hujan sebagai variabel dependen serta suhu rata-rata, kelembaban rata-rata, lama penyinaran matahari, dan kecepatan angin rata-rata sebagai variabel independen.

Pada penelitian ini dengan menggunakan data iklim wilayah Yogyakarta tahun 2010 s.d. 2022 didapatkan hasil berupa nilai *R2 Score* sebesar 12,99%. Prediksi curah hujan pada hari Kamis, 31 Desember 2020 diperoleh nilai prediksi sebesar 14.41778516. Kemudian, pada metode evaluasi *Root Mean Square Error* (RMSE) yang telah dilakukan terjadi adanya penyimpangan antara nilai prediksi curah hujan dengan nilai curah hujan *actual* sebesar 14.783161105087215.

Kata kunci: Regresi Linear Berganda, Root Mean Square Error, Prediksi, Curah Hujan, Data Mining.

ABSTRACT

Indonesia is a tropical region that has erratic weather changes. As a result, it is necessary to have weather predictions as a determinant of the weather that will occur in the future. By predicting the weather that will occur, precautions can be taken immediately so that unwanted things do not occur. Considering that the Yogyakarta region is an area that has mountainous and lowland landscapes that can cause unequal rainfall in various regions. The erratic rainfall conditions in various areas make it difficult to predict the weather.

There are various factors that affect rainfall. Some of the parameters that affect the occurrence of rainfall are temperature, humidity, wind speed, length of sunlight, and direction. With these various data parameters, research can be carried out and matched with previous weather predictions so as to produce the most accurate rainfall prediction for today. Prediction of this rainfall is carried out using the Multiple Linear Regression algorithm where rainfall is the dependent variable and average temperature, average humidity, sunshine duration, and average wind speed are independent variables.

In this study using climate data for the Yogyakarta region in 2010 s.d. 2022 results obtained in the form of an R2 score of 12.99%. Prediction of rainfall on Thursday, December 31, 2020 obtained a prediction value of 14.41778516. Then, in the Root Mean Square Error (RMSE) evaluation method that has been carried out, there is a deviation between the predicted rainfall value and the actual rainfall value of 14.783161105087215.

Keyword: Multiple Linear Regression, Root Mean Square Error, Prediction, Rainfall, Data Mining.