

TUGAS AKHIR

**PERBANDINGAN MODEL REGRESI
UNTUK MEMPREDIKSI HARGA JUAL CABAI RAWIT
BERDASARKAN KEADAAN IKLIM HARIAN
(Studi Kasus : Kota Semarang)**



MIKO ARDIAN

20110011

**PROGRAM STUDI S1 SAINS DATA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

TUGAS AKHIR

**PERBANDINGAN MODEL REGRESI
UNTUK MEMPREDIKSI HARGA JUAL CABAI RAWIT
BERDASARKAN KEADAAN IKLIM HARIAN
(Studi Kasus : Kota Semarang)**

**COMPARISON OF REGRESSION MODELS
FOR PREDICTION THE SELLING PRICES OF CAYENNE
PEPPER BASED ON DAILY CLIMATE CONDITIONS
(Case Study: Semarang City)**

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains Data



MIKO ARDIAN

20110011

**PROGRAM STUDI S1 SAINS DATA
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2024**

HALAMAN PERSETUJUAN PEMBIMBING

**PERBANDINGAN MODEL REGRESI
UNTUK MEMPREDIKSI HARGA JUAL CABAI RAWIT
BERDASARKAN KEADAAN IKLIM HARIAN
(Studi Kasus : Kota Semarang)**

**COMPARISON OF REGRESSION MODELS
FOR PREDICTION THE SELLING PRICES OF CAYENNE
PEPPER BASED ON DAILY CLIMATE CONDITIONS
(Case Study: Semarang City)**

Dipersiapkan dan Disusun Oleh

MIKO ARDIAN

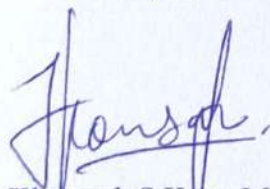
20110011

Fakultas Informatika

Institut Teknologi Telkom Purwokerto

Pada Tanggal : 13 Mei 2024

Pembimbing Utama,



Siti Khomsah S.Kom.,M.Cs
NIDN. 0517108101

Pembimbing Pendamping,



Dr. Ridwan Pandiyya, S.Si.,M.Sc
NIDN. 0625088202

**PERBANDINGAN MODEL REGRESI
UNTUK MEMPREDIKSI HARGA JUAL CABAI RAWIT
BERDASARKAN KEADAAN IKLIM HARIAN
(Studi Kasus : Kota Semarang)**

**COMPARISON OF REGRESSION MODELS
FOR PREDICTION THE SELLING PRICES OF CAYENNE
PEPPER BASED ON DAILY CLIMATE CONDITIONS
(Case Study: Semarang City)**

Disusun Oleh
MIKO ARDIAN
20110011

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir
Pada Selasa, 28 Mei 2024

Penguji 1,



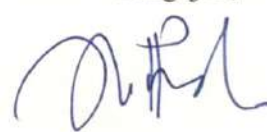
Atika Ratna Dewi, S.Si., M.Sc.
NIDN. 0615128703

Penguji 2



Yesy Diah Rosita, S.Kom., M.Kom
NIDN. 0705108503

Penguji 3,



Dr. Yogo Dwi Prasetyo, S.Si., M.Si.
NIDN. 0101088702

Pembimbing Utama,



Siti Khomsah S.Kom., M.Cs
NIDN. 0517108101

Pembimbing Pendamping,



Dr. Ridwan Pandiya, S.Si., M.Sc
NIDN. 0625088202

Dekan,



Auliya Burhanuddin S.Si., M.Kom
NIK. 19820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Miko Ardian
NIM : 20110011
Program Studi : S1 Sains Data

Menyatakan bahwa Tugas Akhir dengan judul berikut:

**PERBANDINGAN MODEL REGRESI UNTUK MEMPREDIKSI HARGA
JUAL CABAI RAWIT BERDASARKAN KEADAAN IKLIM HARIAN
(Studi Kasus : Kota Semarang)**

Dosen Pembimbing Utama : Siti Khomsah S.Kom.,M.Cs
Dosen Pembimbing Pendamping : Dr. Ridwan Pandiya, S.Si.,M.Sc

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim 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 tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian 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.

Purwokerto, 13 Mei 2024

Yang Menyatakan,



(Miko Ardian)

KATA PENGANTAR

Puji syukur senantiasa penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat, taufiq, hidayah serta ridha-Nya sehingga penulis dapat menyelesaikan dan melaksanakan penelitian tugas akhir ini yang berjudul : “Perbandingan Model Regresi Untuk Memprediksi Harga Jual Cabai Rawit Berdasarkan Keadaan Iklim Harian (Studi Kasus : Kota Semarang)” tepat waktu dan tanpa hambatan apapun. Penulisan tugas akhir ini diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains Data pada program studi Sains Data Fakultas Informatika, Institut Teknologi Telkom Purwokerto.

Penulis menyadari bahwa dalam penulisan dan penyusunan tugas akhir ini tidak terlepas dari bantuan, dukungan, motivasi, saran dan bimbingan yang penulis terima dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan segenap ungkapan terimakasih yang tiada batas kepada:

1. Kedua orang tua, adik-adik dan seluruh anggota keluarga penulis yang selalu mendoakan, mendukung, memberikan dorongan, motivasi serta memberi nasihat dan semangat selama pengerjaan tugas akhir.
2. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Bapak Auliya Burhanuddin S.Si., M.Kom. selaku Dekan Fakultas Informatika Institut Teknologi Telkom Purwokerto.
4. Ibu Siti Khomsah S.Kom., M.Cs. selaku Kaprodi Sains Data, Dosen Wali sekaligus Dosen Pembimbing Utama yang telah meluangkan waktu ditengah kesibukan beliau, memberikan kritik, saran, masukan, dorongan dan arahan kepada penulis dalam menempuh studi di Prodi Sains Data hingga proses penyelesaian tugas akhir.
5. Bapak Dr. Ridwan Pandiya, S.Si.,M.Sc selaku Dosen Pembimbing Pendamping yang telah meluangkan waktu ditengah kesibukan beliau, memberikan kritik, saran, masukan, dorongan dan arahan kepada penulis dalam proses penyelesaian tugas akhir.

6. Ibu Atika Ratna Dewi, S.Si., M.Sc, ibu Yesy Diah Rosita, S.Kom., M.Kom dan bapak Dr. Yogo Dwi Prasetyo, S.Si., M.Si yang telah berkenan meluangkan waktu ditengah kesibukannya untuk menguji dan membantu menyempurnakan tugas akhir ini.
7. Seluruh Dosen, Staff, Laboran beserta jajarannya di Fakultas Informatika Institut Teknologi Telkom Purwokerto yang telah memberikan ilmu, dukungan dan pengalaman yang bermanfaat bagi penulis selama masa perkuliahan.
8. Seluruh teman-teman Sains Data 2020 yang tidak mampu penulis sebutkan satu persatu, menjadi saudara di tanah rantau dan semua pengalaman bermanfaat yang penulis rasakan selama masa perkuliahan.
9. Seluruh teman-teman BEM FIF 2022/2023 yang tidak mampu penulis sebutkan satu persatu, untuk semua pengalaman berorganisasi yang penulis rasakan selama masa perkuliahan.
10. Lagu-lagu Ariana Grande, Billie Eilish dan Feby Putri yang menemani penulis selama mengerjakan tugas akhir.
11. Semua pihak yang telah membantu penulis dalam pembuatan tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Di luar itu, penulis sebagai manusia biasa yang menyadari sepenuhnya bahwa masih banyak kekurangan dalam penulisan tugas akhir ini, baik dari segi tata bahasa, susunan kalimat maupun isi. Penulis sangat berharap agar para pembaca yang telah membaca tugas akhir ini dapat memberikan kritik yang konstruktif sehingga penulis dapat meningkatkan hasil penulisannya di lain kesempatan. Demikian pengantar yang bisa penulis sampaikan, semoga tugas akhir ini dapat menambah wawasan ilmu pengetahuan dan memberikan manfaat nyata bagi pembaca dan masyarakat luas.

Purwokerto, 13 Mei 2024

Penulis,



Miko Ardian

DAFTAR ISI

HALAMAN PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN TUGAS AKHIR	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
ABSTRAK	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN	16
1.1 Latar Belakang	16
1.2 Rumusan Masalah	18
1.3 Pertanyaan Penelitian	18
1.4 Batasan Masalah	19
1.5 Tujuan Penelitian	19
1.6 Manfaat Penelitian	19
BAB II TINJAUAN PUSTAKA	20
2.1 Tinjauan Pustaka.....	20
2.2 Landasan Teori.....	31
2.2.1 Fluktuasi Harga Jual	31
2.2.2 Iklim.....	31
2.2.3 Transformasi <i>BoxCox</i>	32
2.2.4 Standarisasi <i>StandardScaler</i>	34
2.2.5 Prediksi Menggunakan Model Regresi Non-Linear	35
2.2.6 Matriks Evaluasi.....	49
2.2.7 <i>SHapley Additive explanation (SHAP)</i>	51
BAB III METODOLOGI PENELITIAN.....	54
3.1 Subjek dan Objek Penelitian.....	54

3.2	Alat dan Bahan Penelitian	54
3.2.1	Alat Penelitian	54
3.2.2	Bahan Penelitian	55
3.3	Diagram Alir Penelitian.....	55
3.3.1	Tahap Identifikasi Masalah	57
3.3.2	Tahap Pengumpulan Data	57
3.3.3	Tahap <i>Exploratory Data Analysis</i> (EDA).....	59
3.3.4	Tahap Pre-processing Data	61
3.3.5	Tahap Pemodelan dan Evaluasi.....	66
3.3.6	Tahap Interpretasi Hasil	70
BAB IV HASIL DAN PEMBAHASAN		72
4.1	<i>Exploratory Data Analysis</i> (EDA)	72
4.2	<i>Pre-processing Data</i>	83
4.3	Pemodelan dan Evaluasi.....	89
4.3.1	<i>XGBoost Regression</i>	89
4.3.2	<i>KNN Regression</i>	90
4.3.3	<i>Random Forest Regression</i>	92
4.4	Perbandingan Performa Model dan Hasil Prediksi	93
4.5	Pengaruh Variabel-Variabel Iklim Terhadap Prediksi Harga Jual.....	96
BAB V KESIMPULAN DAN SARAN.....		100
5.1	Kesimpulan	100
5.2	Saran.....	101
DAFTAR PUSTAKA		102
LAMPIRAN		108

DAFTAR TABEL

Tabel 2. 1 Literatur <i>Review</i>	24
Tabel 2. 2 Data Sampel untuk Perhitungan Manual	33
Tabel 2. 3 Perhitungan Manual Transformasi <i>BoxCox</i> menggunakan Data Sampel.	33
Tabel 2. 4 Perhitungan Manual Standarisasi <i>StandardScaler</i> menggunakan Data Sampel.....	34
Tabel 2. 5 Dataset Untuk Perhitungan Manual	38
Tabel 2. 6 Perhitungan Jarak <i>Euclidean</i>	44
Tabel 2. 7 Pemilihan K Terdekat Dari Data Uji	44
Tabel 2. 8 <i>Bootstrapping</i> Fitur-Fitur Dataset	47
Tabel 2. 9 Kriteria Nilai <i>MAPE</i>	50
Tabel 2. 10 Data Sampel untuk Perhitungan Manual Nilai <i>Shapely</i>	52
Tabel 3. 1 Data Harga Jual Cabai Rawit	58
Tabel 3. 2 Data Keadaan Iklim.....	58
Tabel 3. 3 Integrasi Data Keadaan Iklim Dan Harga Jual.....	59
Tabel 3. 4 Kombinasi Nilai Pada Parameter <i>XGBoost Regression</i>	67
Tabel 3. 5 Kombinasi Nilai Pada Parameter <i>KNN Regression</i>	68
Tabel 3. 6 Kombinasi Nilai Pada Parameter <i>Random Forest Regression</i>	68
Tabel 3. 7 Perbandingan Performa Antar Model.....	70
Tabel 4. 1. Spesifikasi Variabel Penelitian	72
Tabel 4. 2 Persebaran <i>Missing Value</i> Pada Dataset.....	73
Tabel 4. 3 Persebaran Data 8888 dan Data 0 <i>Absolut</i>	74
Tabel 4. 4 Persebaran Data <i>Outlier</i>	75
Tabel 4. 5 Statistika Deskriptif Masing-Masing Variabel	75
Tabel 4. 6 Cuplikan Data Sebelum dan Setelah Penanganan <i>Missing Value</i>	85
Tabel 4. 7 Uji-Uji Asumsi Klasik Pada Dataset	86
Tabel 4. 8 Cuplikan Data Sebelum dan Setelah Transformasi <i>BoxCox</i>	87
Tabel 4. 9 Persebaran Data <i>Outlier</i> Sebelum dan Setelah Transformasi <i>BoxCox</i>	87
Tabel 4. 10 Evaluasi Model <i>XGboost Regression</i>	90
Tabel 4. 11 Evaluasi Model <i>KNN Regression</i>	91

Tabel 4. 12 Evaluasi Model <i>Random Forest Regression</i>	93
Tabel 4. 13 Perbandingan Matriks Evaluasi Setiap Model.....	95
Tabel 4. 14 Data Baru	96

DAFTAR GAMBAR

Gambar 2. 1 Histogram Data Sebelum Transformasi (kiri) dan Setelah Transformasi (Kanan).....	33
Gambar 2. 2 Histogram Data Sebelum Standarisasi (kiri) dan Setelah Standarisasi (kanan)	35
Gambar 2. 3 Ilustrasi Kerja Algoritma <i>XGBoost Regression</i>	36
Gambar 2. 4 <i>Flowchart</i> Model <i>XGBoost Regression</i>	38
Gambar 2. 5 Pembangunan <i>Tree</i> Awal	39
Gambar 2. 6 Pembangunan <i>Tree-Tree XGBoost Regression</i>	41
Gambar 2. 7 Ilustrasi Kerja Algoritma <i>KNN</i>	42
Gambar 2. 8 <i>Flowchart</i> Model <i>KNN Regression</i>	43
Gambar 2. 9 Ilustrasi Kerja Algoritma <i>Random Forest Regression</i>	45
Gambar 2. 10 <i>Flowchart</i> Model <i>Random Forest Regression</i>	46
Gambar 2. 11 Pembangunan <i>Tree-Tree</i> Awal <i>Random Forest Regression</i>	48
Gambar 2. 12 Pembangunan <i>Tree-Tree Random Forest Regression</i>	49
Gambar 3. 1 Diagram Alir Penelitian	56
Gambar 4. 1 Grafik Garis Variabel Suhu	77
Gambar 4. 2 Grafik Garis Variabel Kelembapan	77
Gambar 4. 3 Grafik Garis Variabel Curah Hujan	77
Gambar 4. 4 Grafik Garis Variabel Lama Penyinaran.....	78
Gambar 4. 5 Grafik Garis Variabel Kecepatan Angin	78
Gambar 4. 6 Grafik Garis Variabel Harga Jual	78
Gambar 4. 7 Histogram Variabel Suhu (kiri) dan Kelembapan (kanan)	79
Gambar 4. 8 Histogram Variabel Curah Hujan (kiri) dan Lama Penyinaran (kanan)	80
Gambar 4. 9 Histogram Variabel Kecepatan Angin (kiri) dan Harga Jual (kanan) ..	80
Gambar 4. 10. Persebaran <i>Outlier</i> Variabel Suhu (kanan) dan Kelembapan (kiri) ...	81

Gambar 4. 11 Persebaran <i>Outlier</i> Variabel Curah Hujan (kanan) dan Lama Penyinaran (kiri)	81
Gambar 4. 12 Persebaran <i>Outlier</i> Variabel Kecepatan Angin (kanan) dan Harga Jual (kiri)	82
Gambar 4. 13 <i>Heatmap Correlation</i> menggunakan Metode Spearman	83
Gambar 4. 14 Persebaran <i>Outlier</i> Variabel Curah Hujan Setelah Menghapus Data 8888	84
Gambar 4. 15 Cuplikan Data Setelah Standarisasi <i>StandardScaler</i>	88
Gambar 4. 16 Grafik <i>R2-Score</i> Setiap Model dengan 5 <i>Splitting</i> yang Berbeda	94
Gambar 4. 17 Perbandingan Data Aktual dan Data Hasil Prediksi menggunakan Data Uji	95
Gambar 4. 18 Grafik Hasil Prediksi Pada Data Baru	96
Gambar 4. 19 Diagram <i>SHAP</i> untuk Variabel yang Berpengaruh Signifikan Terhadap Harga Jual.....	97
Gambar 4. 20 Diagram Nilai <i>Shapely</i> untuk Variabel yang Berkontribusi Terhadap Harga Jual.....	98