

**PERBANDINGAN MODEL SUPPORT VECTOR REGRESSION,  
RANDOM FOREST REGRESSION, DAN RECURRENT NEURAL  
NETWORK DALAM MEMPREDIKSI HARGA CABAI**

**(STUDI KASUS: PROVINSI SUMATERA SELATAN)**

**SKRIPSI**

Program Studi Sistem Informasi  
Jenjang Sarjana



**Oleh**

**Sarah Desmarina  
NIM 09031282126069**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
DESEMBER 2024**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**PERBANDINGAN MODEL SUPPORT VECTOR REGRESSION,  
RANDOM FOREST REGRESSION, DAN RECURRENT NEURAL  
NETWORK DALAM MEMPREDIKSI HARGA CABAI**

**(STUDI KASUS: PROVINSI SUMATERA SELATAN)**

Sebagai salah satu syarat untuk penyelesaian  
studi di Program Studi Sistem Informasi SI

Oleh


Sarah Desmarina      09031282126069

Palembang, 31 Desember 2024

Mengetahui,  
Ketua Jurusan Sistem Informasi

Pembimbing,

  
  
**Ahmad Rifai, M.T.**  
NIP 197910202010121003

  
**Allsela Meiriza, M.T.**  
NIP 198305132023212026

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Sarah Desmarina

NIM : 09031282126069

Program Studi: Sistem Informasi Reguler

Judul Skripsi : Perbandingan Model Support Vector Regression, Random Forest Regression, dan Recurrent Neural Network dalam Memprediksi Harga Cabai (Studi Kasus: Provinsi Sumatera Selatan)

Hasil Pengecekan *Ithenticate/Turnitin*: 1 %

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 31 Desember 2024



Sarah Desmarina  
NIM 09031282126069

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Selasa

Tanggal : 31 Desember 2024


Nama : Sarah Desmarina

NIM : 09031282126069

Judul : Perbandingan Model Support Vector Regression, Random Forest Regression, dan Recurrent Neural Network dalam Memprediksi Harga Cabai (Studi Kasus: Provinsi Sumatera Selatan)

Komisi Penguji:

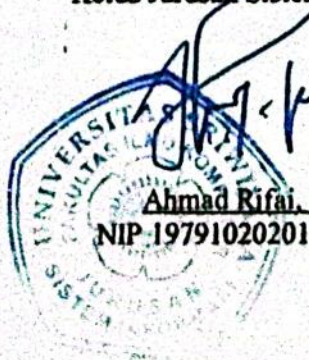
1. Ketua : Ari Wedhasmara, M.TI. 

2. Sekretaris : Putri Eka Sevdiyuni, M.T. 

3. Pembimbing 1: Allsela Meiriza, M.T. 

4. Penguji : Ken Ditha Tania, M.Kom., Ph.D. 

Mengetahui,  
Ketua Jurusan Sistem Informasi



Ahmad Rifai, M.T.  
NIP. 197910202010121003

## KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT, karena atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Perbandingan Model Support Vector Regression, Random Forest Regression, dan Recurrent Neural Network dalam Memprediksi Harga Cabai (Studi Kasus: Provinsi Sumatera Selatan).”** Sebagai salah satu syarat untuk menyelesaikan pendidikan Strata 1 di Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini, penulis mendapatkan dukungan dan bimbingan dari banyak pihak. Maka dari itu, penulis mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Orang tua dan saudara penulis yang selalu berdoa dan memberi dukungan kepada penulis.
2. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Ahmad Rifai, S.T., M.T. selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Ibu Allsela Meiriza, M.T. selaku Dosen Pembimbing Akademik dan Dosen Pembimbing Tugas Akhir yang selalu membimbing dan memberikan arahan selama menyelesaikan Tugas Akhir.
5. Seluruh Dosen dan Staff Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Teman-teman seperjuangan dari Sistem Informasi Angkatan 2021 yang telah banyak membantu penulis selama masa kuliah.

7. Semua pihak yang telah memberikan dukungan dan bantuan kepada penulis selama masa studi dan proses penyusunan Tugas Akhir ini, yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini masih terdapat kekurangan yang disebabkan keterbatasan pengetahuan dan pengalaman. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk penyempurnaan penelitian di masa mendatang. Akhir kata, semoga Tugas Akhir ini dapat memberikan manfaat bagi penulis dan seluruh pembaca.

Palembang, 31 Desember 2024



Sarah Desmarina  
NIM 09031282126069

# **PERBANDINGAN SUPPORT VECTOR REGRESSION, RANDOM FOREST REGRESSION, DAN RECURRENT NEURAL NETWORK DALAM MEMPREDIKSI HARGA CABAI**

**(STUDI KASUS: PROVINSI SUMATERA SELATAN)**

Oleh

Sarah Desmarina  
09031282126069

Cabai adalah salah satu komoditas penyebab inflasi dan sangat vital bagi konsumsi masyarakat Indonesia. Di wilayah Sumatera Selatan, harga cabai seringkali tidak stabil sehingga dibutuhkan analisis prediksi harga cabai guna memitigasi dampak ketidakstabilan harga, terutama di Pasar Tradisional. Penelitian ini membandingkan tiga model penambangan data yaitu Support Vector Regression (SVR), Random Forest Regression (RFR), dan Recurrent Neural Network Long Short Term Memory (RNN LSTM) guna memutuskan model yang terbaik dalam prediksi harga cabai. Dataset mencakup 1240 harga berbagai jenis cabai di Pasar Tradisional Sumatera Selatan dengan memanfaatkan fitur lag dan musiman serta penerapan optimasi parameter guna peningkatan kinerja model. Hasil penelitian mengungkapkan bahwa SVR paling unggul dengan akurasi tertinggi (MAPE 0.03%-0.04% dan  $R^2$  0.90-0.94), stabilitas terbaik (MAE 1187.44-2478.49 dan MSE 11380618.92-16594596.73), waktu komputasi tercepat (23 detik), dan paling tahan melawan outlier. RFR juga berkinerja baik meskipun masih dibawah SVR. RNN LSTM mencatat hasil terendah di semua kategori. Penemuan ini mengkonfirmasi SVR sebagai model yang unggul dalam memprediksi harga cabai dan menyediakan insight berharga guna memitigasi risiko ketidakstabilan harga.

**Kata Kunci:** Cabai, Data Mining, SVR, RFR, RNN LSTM



# COMPARISON OF SUPPORT VECTOR REGRESSION, RANDOM FOREST REGRESSION, AND RECURRENT NEURAL NETWORK IN PREDICTING CHILI PRICES

(CASE STUDY: SOUTH SUMATRA PROVINCE)

By

Sarah Desmarina  
09031282126069

Chili is a key product contributing to inflation and is essential for the Indonesian population. In South Sumatra, chili prices frequently fluctuate, requiring effective price forecasting, especially in Traditional Markets. This study compares three data mining techniques, namely Support Vector Regression, Random Forest Regression, and Long Short Term Memory Recurrent Neural Network, to determine which model is most effective for predicting chili prices. The study utilizes a dataset of 1240 chili prices from various South Sumatra's Traditional Markets, incorporating lag and seasonal features with parameter optimization to enhance model efficacy. The findings indicated that SVR outperformed other models with the highest accuracy (MAPE 0.03%-0.04% and  $R^2$  0.90-0.94), superior stability (MAE 1187.44-2478.49 and MSE 11380618.92-16594596.73), the quickest computation time (23 seconds), and showed the greatest resilience to outliers. Although RFR had strong results, it still fell short of SVR's performance. RNN LSTM ranked lowest across all metrics. These results establish SVR as the most reliable model for predicting chili prices and provide valuable insights to manage price volatility.

**Keyword:** Chili, Data Mining, SVR, RFR, RNN LSTM



## DAFTAR ISI

<b>HALAMAN PENGESAHAN .....</b>	<b>ii</b>
<b>HALAMAN PERNYATAAN.....</b>	<b>iii</b>
<b>HALAMAN PERSETUJUAN.....</b>	<b>iv</b>
<b>KATA PENGANTAR .....</b>	<b>v</b>
<b>DAFTAR ISI.....</b>	<b>ix</b>
<b>DAFTAR TABEL.....</b>	<b>xiii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xiv</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xv</b>
<b>BAB I .....</b>	<b>1</b>
<b>PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	6
<b>BAB II .....</b>	<b>8</b>
<b>TINJAUAN PUSTAKA .....</b>	<b>8</b>
2.1 Penelitian Terdahulu.....	8
2.2 Cabai.....	11
2.3 Data Mining.....	12
2.3.1 Pengertian Data Mining.....	12
2.3.2 Teknik-Teknik Data Mining.....	13
2.3.3 Proses Data Mining.....	15

2.4 Peramalan (Forecasting).....	16
2.4.1 Pengertian Peramalan (Forecasting) .....	16
2.4.2 Metode Peramalan (Forecasting).....	17
2.5 Support Vector Regression (SVR) .....	18
2.6 Random Forest Regression (RFR).....	20
2.7 Recurrent Neural Network (RNN) .....	20
2.8 Evaluasi Model.....	22
2.9 Pustaka Python untuk Model Prediksi.....	23
2.9.1 Pandas .....	23
2.9.2 Scikit-learn.....	23
2.9.3 Tensorflow.....	23
2.9.4 Matplotlib.....	24
<b>BAB III.....</b>	<b>25</b>
<b>METODOLOGI PENELITIAN .....</b>	<b>25</b>
3.1 Objek Penelitian .....	25
3.2 Metode Pengumpulan Data .....	25
3.2.1 Jenis Data.....	25
3.2.2 Sumber Data .....	26
3.2.3 Pengumpulan Data.....	26
3.3 Tahap Penelitian .....	27
3.3.1 Studi Literatur.....	27
3.3.2 Identifikasi Masalah.....	28
3.3.3 Pengumpulan Data.....	28
3.3.4 Pembersihan Data (Data Cleaning).....	28

3.3.5 Seleksi Data (Data Selection) .....	29
3.3.6 Transformasi Data (Data Transformation) .....	29
3.3.7 Proses Penambangan (Mining Process).....	29
3.3.8 Evaluasi Pola (Pattern Evaluation) .....	30
3.3.9 Presentasi Pengetahuan (Knowledge Presentation).....	30
<b>BAB IV .....</b>	<b>31</b>
<b>HASIL DAN PEMBAHASAN .....</b>	<b>31</b>
4.1 Pembersihan Data (Data Cleaning) .....	31
4.2 Seleksi Data (Data Selection).....	31
4.3 Transformasi Data (Data Transformation) .....	32
4.4 Proses Penambangan (Mining Process).....	33
4.4.1 SVR.....	33
4.4.2 RFR.....	34
4.4.3 RNN LSTM .....	34
4.5 Evaluasi Pola (Pattern Evaluation).....	35
4.5.1 Hasil Evaluasi Prediksi Harga Cabai Merah.....	35
4.5.2 Hasil Evaluasi Prediksi Harga Cabai Merah Besar .....	35
4.5.3 Hasil Evaluasi Prediksi Harga Cabai Merah Keriting .....	36
4.5.4 Hasil Evaluasi Prediksi Harga Cabai Rawit .....	37
4.5.5 Hasil Evaluasi Prediksi Harga Cabai Rawit Hijau.....	37
4.5.6 Hasil Evaluasi Prediksi Harga Cabai Rawit Merah.....	38
4.5 Presentasi Pengetahuan (Knowledge Presentation).....	38
4.5.1 Visualisasi Hasil .....	38
4.5.2 Pembahasan .....	41

4.3 Perbandingan Kinerja Model.....	41
<b>BAB V.....</b>	<b>46</b>
<b>PENUTUP.....</b>	<b>46</b>
5.1 Kesimpulan.....	46
5.2 Saran.....	46
<b>DAFTAR PUSTAKA.....</b>	<b>48</b>
<b>LAMPIRAN.....</b>	<b>53</b>

## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	8
Tabel 3.1 Data Harga Cabai Harian.....	26
Tabel 4.1 Hasil Pembersihan Data.....	31
Tabel 4.2 Hasil Seleksi Data.....	32
Tabel 4.3 Hasil Transformasi Data Latih Setelah Standarisasi.....	32
Tabel 4.4. Hasil Transformasi Data Uji Setelah Standarisasi.....	32
Tabel 4.5 Parameter SVR Terbaik Setiap Jenis Cabai.....	33
Tabel 4.6 Hyperparameter RFR Setiap Jenis Cabai.....	34
Tabel 4.7 Pelatihan Model RNN LSTM.....	35
Tabel 4.8 Hasil Evaluasi Prediksi Harga Cabai Merah.....	35
Tabel 4.9 Hasil Evaluasi Prediksi Harga Cabai Merah Besar.....	35
Tabel 4.10 Hasil Evaluasi Prediksi Harga Cabai Merah Keriting.....	36
Tabel 4.11 Hasil Evaluasi Prediksi Harga Cabai Rawit.....	37
Tabel 4.12 Hasil Evaluasi Prediksi Harga Cabai Rawit Hijau.....	37
Tabel 4.13 Hasil Evaluasi Prediksi Harga Cabai Rawit Merah.....	38
Tabel 4.14 Perbandingan Kinerja Model.....	42

## DAFTAR GAMBAR

Gambar 2.1 Pencarian Grid Dua Hiperparameter.....	19
Gambar 2.2 Arsitektur Dasar LSTM.....	21
Gambar 3.1 Diagram Tahap Penelitian.....	27
Gambar 4.1 Visualisasi Hasil Prediksi Harga Cabai Merah.....	39
Gambar 4.2 Visualisasi Hasil Prediksi Harga Cabai Merah Besar.....	39
Gambar 4.3 Visualisasi Hasil Prediksi Harga Cabai Merah Keriting.....	39
Gambar 4.4 Visualisasi Hasil Prediksi Harga Cabai Rawit.....	40
Gambar 4.5 Visualisasi Hasil Prediksi Harga Cavai Rawit Hijau.....	40
Gambar 4.6 Visualisasi Hasil Prediksi Harga Cabai Rawit Merah.....	40

## DAFTAR LAMPIRAN

Lampiran A. Surat Keputusan Pembimbing Skripsi.....	A-1
Lampiran B. Kartu Konsultasi.....	B-1
Lampiran C. Listing Program.....	C-1
Lampiran D. Hasil Turn It In.....	D-1
Lampiran E. Surat Kesiediaan Membimbing.....	E-1
Lampiran F. Surat Pengajuan Topik Skripsi.....	F-1
Lampiran G. Form Perbaikan Ujian Komprehensif.....	G-1
Lampiran H. Surat Keterangan Pengecekan Similarity.....	H-1



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Bahan pokok merupakan komoditas yang sangat krusial dalam menjaga stabilitas ekonomi rumah tangga di Indonesia. Salah satu indikator utama yang mencerminkan stabilitas tersebut adalah inflasi yang diukur melalui Indeks Harga Konsumen (IHK). Menurut Badan Pusat Statistik (BPS), inflasi *year on year* (y-on-y) mencapai 2,75% dengan IHK 105,58 pada Februari 2024, sebagian besar dipicu oleh kenaikan harga sebesar 6,36% di berbagai kelompok pengeluaran, khususnya makanan, minuman, dan tembakau (BPS, 2024). Diantara komoditas penyumbang inflasi dalam kelompok tersebut, cabai menempati posisi penting sebagai komoditas nomor dua setelah beras yang banyak digunakan dalam konsumsi harian masyarakat Indonesia. Beberapa provinsi dengan konsumsi cabai cukup tinggi berada di Pulau Sumatera. Berdasarkan Konsumsi Pangan Tahun 2021, provinsi-provinsi yang berada di Pulau Sumatera termasuk dalam enam besar provinsi dengan tingkat konsumsi cabai besar tertinggi di Indonesia (Kementan, 2021).

Di Provinsi Sumatera Selatan, harga cabai dikenal fluktuatif terutama di pasar-pasar tradisional. Menurut Radio Republik Indonesia di Palembang, harga cabai yang awalnya stabil di awal pekan sering melonjak drastis menjelang akhir pekan. Contohnya pada 26 Juni 2024 di Pasar 26 Ilir Palembang, harga cabai merah dan cabai rawit masing-masing naik dari Rp28.000 per kilogram menjadi Rp38.000 hingga Rp40.000 per kilogram, sementara harga cabai rawit jengki naik drastis menjadi Rp70.000 per kilogram dari sebelumnya Rp44.000 (Sihombing, 2024). Dalam menghadapi tantangan fluktuasi harga cabai di Sumatera Selatan, prediksi

harga cabai yang akurat menjadi sangat penting untuk membantu konsumen, pedagang, dan pembuat kebijakan dalam membuat keputusan yang lebih baik. Salah satu teknik yang dapat digunakan dalam memprediksi harga cabai adalah teknik peramalan (*forecasting*), sehingga pembelian dan penjualan dapat direncanakan secara lebih efektif.

Peramalan (*forecasting*) merupakan proses memperkirakan kejadian yang akan datang berdasarkan data masa lalu maupun saat ini dengan tujuan untuk meminimalkan kesalahan atau selisih antara hasil aktual dan hasil yang diprediksi (Ustadatin et al., 2023). Dalam hal ini, peramalan yang akurat dapat membantu memitigasi risiko kerugian akibat fluktuasi harga, sehingga memungkinkan pengambilan keputusan yang lebih optimal dalam pembelian dan penjualan cabai di pasar tradisional. Oleh karena itu, penerapan teknologi dan metode analitik yang canggih sangat diperlukan untuk mengatasi tantangan ini. Beberapa metode yang dapat digunakan dalam melakukan prediksi harga cabai adalah menggunakan algoritma *data mining*, seperti *Support Vector Regression (SVR)*, *Random Forest Regression (RFR)*, dan *Recurrent Neural Network (RNN)*.

*Support Vector Regression (SVR)* merupakan algoritma yang diperkenalkan oleh Vapnik pada tahun 1955, dikembangkan sebagai bagian dari *Support Vector Machine (SVM)* (Siregar, 2022). SVR ini dirancang untuk regresi dan prediksi dengan meminimalkan risiko melalui pengurangan batas atas kesalahan generalisasi. SVR efektif dalam mengenali pola pada data *time series* dan memberikan hasil prediksi yang akurat jika parameter pentingnya ditentukan dengan tepat. Selain itu, SVR juga mampu menangani kasus non linier dan mengatasi *overfitting*, sehingga menjadikannya relevan untuk prediksi harga

komoditas, seperti cabai (Sriyana et al., 2019). Penelitian sebelumnya yang berjudul “Prediksi Harga cabai Merah Menggunakan Support Vector Regression” menunjukkan bahwa prediksi harga cabai merah dengan SVR menggunakan kernel *Radial Basis Function* (RBF) memberikan hasil yang baik. Model SVR dengan parameter terbaik ( $C=1000$ ,  $\epsilon=5$ ,  $\gamma=1$ ) dari penelitian tersebut menghasilkan prediksi harga cabai merah nasional dengan MAPE pada data latih sebesar 4,07% dan pada data uji sebesar 9,11 %, menunjukkan bahwa akurasi prediksi sangat baik tanpa *overfitting*. Namun, model ini masih bisa diperbaiki dengan *ensemble learning* atau mentransformasikan data ke ruang fitur lain untuk meningkatkan akurasi prediksi harga cabai (Sepri et al., 2020).

*Random Forest Regression* (RFR) merupakan metode untuk pemodelan regresi yang diperkenalkan oleh Breiman pada tahun 2001, yang membangun sejumlah pohon keputusan (*decision trees*) menggunakan sampel data yang diambil secara acak (*bootstrap*) dan memilih subset acak dari variabel untuk setiap pemisahan (*split*) pada node pohon (Fitri & Riana, 2022). Penelitian sebelumnya yang berjudul “Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga” menunjukkan bahwa *Random Forest Regression* menghasilkan *error rate* yang lebih kecil dibandingkan metode lainnya. Dalam penelitian tersebut, nilai *error* pada *Random Forest Regression* tercatat sebesar 0.440 (RMSE), lebih rendah dibandingkan dengan *Linear Regression* yang memiliki *error* 0.550 dan *Gradient Boosted Trees Regression* dengan *error* 0.508 (Fitri, 2023). Hal ini menunjukkan bahwa *Random Forest Regression* dapat menjadi alternatif yang baik untuk prediksi harga cabai yang fluktuatif.

*Recurrent Neural Network* (RNN) merupakan algoritma *deep learning* yang memanfaatkan keadaan internal untuk menangani data berurutan dengan mentransmisikan informasi antar neuron secara berurutan, seperti pengenalan suara dan prediksi deret waktu (Shan et al., 2024). RNN termasuk salah satu jenis dari *Neural Network* (NN), yaitu model pembelajaran mesin yang meniru fungsi otak manusia dalam mengenali pola dan memproses data secara kompleks melalui jaringan neuron yang saling terhubung (Muntaha et al., 2024). RNN terbagi menjadi beberapa varian, antara lain *Long Short-Term Memory* (LSTM), *Bidirectional LSTM*, *Gated Recurrent Units*, *Other Notable Variants*, dan inovasi RNN dengan arsitektur NN lainnya (Mienye et al., 2024). Dalam penelitian ini, penulis akan berfokus pada salah satu varian dari RNN yaitu LSTM. LSTM merupakan algoritma *Deep Learning* tipe RNN yang dirancang untuk prediksi data deret waktu dengan kemampuan menyimpan dan mengelola informasi jangka panjang melalui sel memori dan tiga jenis gerbang yaitu *forget gate*, *input gate*, dan *output gate* (Fahrudin et al., 2023). Berdasarkan penelitian sebelumnya yang berjudul “Evaluation of Cryptocurrency Price Prediction Using LSTM and CNNs Model” didapatkan bahwa model LSTM dan CNN efektif dalam memprediksi harga *cryptocurrency*, dengan LSTM sedikit lebih unggul berdasarkan metrik RMSE, MSE, dan MAE. Namun, volatilitas tinggi pada data *bitcoin* menyulitkan konsistensi evaluasi metrik (Wen & Ling, 2024). Oleh karena itu, model LSTM dipilih karena kemampuannya dalam menangani data deret waktu yang kompleks.

Dengan mempertimbangkan berbagai metode yang tersedia, penelitian ini akan menerapkan metode SVR, RFR, dan RNN LSTM untuk memprediksi harga cabai di Pasar Tradisional Provinsi Sumatera Selatan. Model ini juga akan

memanfaatkan transformasi data dan penambahan fitur, seperti *lag* dan *seasonality* guna menghasilkan prediksi harga cabai yang lebih akurat. Hasil prediksi tersebut diharapkan dapat memberikan manfaat bagi konsumen, pedagang, serta pembuat kebijakan dalam merespon perubahan harga cabai yang fluktuatif. Berdasarkan penjelasan diatas, penulis akan melakukan penelitian yang berjudul **“Perbandingan Model Support Vector Regression, Random Forest Regression, dan Recurrent Neural Network dalam Memprediksi Harga Cabai (Studi Kasus: Provinsi Sumatera Selatan).”**

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dipaparkan, rumusan masalah dari penelitian ini adalah fluktuasi harga cabai yang sering terjadi di Pasar Tradisional Provinsi Sumatera Selatan. Untuk menyelesaikan rumusan masalah tersebut, maka disusun beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimana cara memprediksi harga cabai di Pasar Tradisional Provinsi Sumatera Selatan?
2. Apa model yang paling sesuai dalam memprediksi harga cabai di Pasar Tradisional Provinsi Sumatera Selatan berdasarkan akurasi, stabilitas, waktu optimasi, dan sensitivitas terhadap *outlier*?

## **1.3 Tujuan Penelitian**

Pelaksanaan penelitian ini bertujuan untuk menentukan model terbaik dalam memprediksi harga cabai di Pasar Tradisional Provinsi Sumatera Selatan dengan membandingkan tiga model yaitu *Support Vector Regression* (SVR),

*Random Forest Regression (RFR)*, dan *Recurrent Neural Network Long Short Term Memory (RNN LSTM)*.

#### **1.4 Manfaat Penelitian**

Pelaksanaan penelitian ini diharapkan memberikan beberapa manfaat, diantaranya sebagai berikut:

1. Penelitian menyediakan informasi prediktif mengenai harga cabai di Pasar Tradisional Provinsi Sumatera Selatan, sehingga dapat membantu masyarakat dalam merencanakan pengeluaran mereka dan mengelola anggaran pangan.
2. Penelitian memberikan alat bantu yang dapat digunakan petani maupun pedagang untuk merencanakan waktu tanam dan penjualan, serta mengurangi resiko kerugian akibat fluktuasi harga.
3. Penelitian memberikan data dan analisis yang dapat digunakan oleh pemerintah untuk merumuskan kebijakan yang lebih efektif dalam mengendalikan fluktuasi harga cabai di Pasar Tradisional Provinsi Sumatera Selatan.

#### **1.5 Batasan Masalah**

Agar permasalahan yang dibahas oleh penulis tidak terlalu meluas, maka dilakukan pembatasan masalah diantaranya sebagai berikut:

1. Penelitian ini menggunakan *dataset* yang berasal dari *Website* Pusat Informasi Harga Pangan Strategis Nasional dengan variabel yang digunakan adalah tanggal dan harga cabai.

2. Objek Penelitian di Pasar Tradisional Provinsi Sumatera Selatan.
3. Rentang data penelitian yang digunakan adalah per 1 Januari 2020 sampai 30 September 2024 dengan tipe data harian.
4. Penelitian ini menggunakan model *Support Vector Regression* (SVR), *Random Forest Regression* (RFR), dan *Recurrent Neural Network* varian *Long Short Term memory* (RNN LSTM) untuk memprediksi harga cabai.
5. Dalam melakukan analisis dan pemodelan data, digunakan platform *Jupyter Notebook* dengan bahasa pemrograman *Python*. Tools yang digunakan dalam penelitian ini antara lain:
  - a. *Pandas* untuk pengolahan data.
  - b. *Scikit-learn* untuk penerapan algoritma *machine learning*, seperti *Support Vector Regression* (SVR) dan *Random Forest Regression* (RFR), serta evaluasi model.
  - c. *Tensorflow* untuk membangun dan melatih model *Recurrent Neural Network Long Short Term Memory* (RNN LSTM).
  - d. *Matplotlib* untuk visualisasi hasil prediksi dan analisis data.



## DAFTAR PUSTAKA

- Agustin, F. M. (2024). *High Consumption of Chilies and Shallots Triggers Inflation in South Sumatra*. IDN Times Sumsel. <https://sumsel.idntimes.com/business/economy/feny-agustin/konsumsi-cabai-dan-bawang-merah-yang-tinggi-picu-inflasi-sumsel?page=all>
- Ahdiat, A. (2024). *Konsumsi Cabai per Kapita Indonesia Naik, Rekor Tertinggi pada 2023*. Databoks. <https://databoks.katadata.co.id/demografi/statistik/faf774e4492db0d/konsumsi-cabai-per-kapita-indonesia-naik-rekor-tertinggi-pada-2023>
- Aldi, F., Hadi, F., Rahmi, N. A., & Defit, S. (2023). Standardscaler's Potential In Enhancing Breast Cancer. *Journal of Applied Engineering and Technological Science*, 5(1), 401–413.
- Alshingiti, Z., Alaqel, R., Al-muhtadi, J., Emad, Q., Haq, U., Saleem, K., & Faheem, M. H. (2023). A Deep Learning-Based Phishing Detection System Using CNN, LSTM, and LSTM-CNN. *Electronics*, 12(1), 1–18.
- BPS. (2024). *Inflasi tahun-ke-tahun (y-on-y) pada bulan Februari 2024 adalah 2,75 persen*. Badan Pusat Statistik. <https://www.bps.go.id/id/pressrelease/2024/03/01/2301/inflasi-tahun-ke-tahun--y-on-y--pada-bulan-februari-2024-adalah-2-75-persen.html>
- Fahrudin, R., Informasi, F. T., Catur, U., Cendekia, I., & Ekonomi, F. (2023). Price Prediction System of Basic Commodities Using Long Short-Term Memory Method: Analysis and Implementation for Future Projections. *Journal Research of Social Science, Economics, and Management (JRSSEM)*, 03(07), 1617–1626.
- Falah, R. A., & Rachmaniah, M. (2022). Price Prediction Model for Red and Curly Red Chilies using Long Short Term Memory Method. *Indonesian Journal of Statistics and Its Applications*, 06(1), 143–160. <https://doi.org/https://doi.org/10.29244/ijisa.v6i1p143-160>
- Fitri, E. (2023). Perbandingan Metode Regresi Linier , Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah. *Journal of Applied Computer Science and Technology (JACOST)*, 4(1), 58–64.
- Fitri, E., & Riana, D. (2022). Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer Perceptron. *Methomika: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, 6(1), 69–78.
- Garai, S., Paul, R. K., & Paul, A. K. (2024). Spillover Effects of Covid-19 Induced

- Lockdown on Onion Prices in India. *Journal of Scientific Research and Reports*, 30(3), 21–31. <https://doi.org/10.9734/JSRR/2024/v30i31855>
- Guo, W., & Li, Q. (2021). Construction Cost Prediction for Residential Projects Based on Support Vector Regression. *Springer*, 114–124.
- Huang, G. (2021). Missing Data Filling Method Based on Linear Interpolation and LightGBM. *Journal of Physics: Conference Series*, 1754(1), 1–6.
- Irawati, A. R., Kurniawan, D., Utami, Y. T., & Taufik, R. (2024). An Exploration of TensorFlow-Enabled Convolutional Neural Network Model Development for Facial Recognition: Advancements in Student Attendance System. *Scientific Journal of Informatics*, 11(2), 413–428. <https://doi.org/10.15294/sji.v11i2.3585>
- Jain, S. K., & Gupta, A. K. (2022). Application of Random Forest Regression with Hyper-parameters Tuning to Estimate Reference Evapotranspiration. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 13(5), 742–750.
- Jasim, M. A., & Abdulwahid, S. N. (2021). Data Mining preparation : Process, Techniques and Major Issues in Data Analysis. *IOP Conf. Series: Materials Science and Engineering*, 1090(1), 1–8. <https://doi.org/10.1088/1757-899X/1090/1/012053>
- Kementan. (2021). *Buletin Konsumsi Pangan*. Pusat Data dan Sistem Informasi Pertanian.
- Komaria, V., Maidah, N. El, & Furqon, M. A. (2023). Prediksi Harga Cabai Rawit di Provinsi Jawa Timur Menggunakan Metode Fuzzy Time Series Model Lee. *Komputika: Jurnal Sistem Komputer*, 12(148), 157–167. <https://doi.org/10.34010/komputika.v12i2.10644>
- Manzanares, M. C. S., González, S. G., Rodríguez, Á., Romero, L. A. C., Calderón, V., & Dios, M. Á. Q. (2020). Systematic Review on Inclusive Education , Sustainability in Engineering : An Analysis with Mixed Methods and Data Mining Techniques. *Sustainability*, 12(17), 1–8.
- Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent Neural Networks : A Comprehensive Review of Architectures, Variants, and Applications. *Information*, 15(9), 1–34.
- Miraltamirus, I., Fitri, F., Vionanda, D., & Permana, D. (2023). Stock Price Prediction of PT Bank Syariah Indonesia Tbk Using Support Vector Regression. *UNP Journal of Statistics and Data Science*, 1(3), 112–119.
- Mohamed, M. A., El-henawy, I. M., & Salah, A. (2021). Price Prediction of Seasonal Items Using Machine Learning and Statistical Methods. *Computers*,

*Materials & Continua*, 70(2), 3473–3489.  
<https://doi.org/10.32604/cmc.2022.020782>

- Muntaha, S., Jahan, S., & Bhuiyan, A. I. (2024). Extended Representation Learning Based Neural Network Model for Outlier Detection. *Journal of Artificial Intelligence, Machine Learning and Neural Network*, 04(06), 12–26.
- Nugroho, N. C. T., & Hidayat, E. Y. (2024). Implementation of Adam Optimizer using Recurrent Neural Network (RNN) Architecture for Diabetes Classification. *Jurnal Media Informatika Budidarma*, 8(1), 421–429.
- Prabhu, R. R., & Shetty, S. (2023). Stock Price Prediction Using Random Forest Regression. *International Research Journal of Modernization in Engineering Technology and Science*, 05(05), 3368–3373.  
<https://doi.org/https://www.doi.org/10.56726/IRJMETS39128>
- Rachmawati, F., Jaenudin, J., Ginting, N. B., & Laksono, P. (2024). Machine Learning for the Model Prediction of Final Semester Assessment (FSA) using the Multiple Linear Regression Method. *Jurnal Teknik Informatika*, 17(1), 1–9.
- Raschka, S., Patterson, J., & Nolet, C. (2020). Machine Learning in Python : Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. *Information*, 11(4), 1–44.  
<https://doi.org/10.3390/info11040193>
- Rathakrishnan, V., Beddu, S. B., & Ahmed, A. N. (2022). Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms. *Scientific Reports*, 12(1), 1–16.
- Rofik, Hakim, R. A., Jumanto, Prasetyo, B., & Muslim, M. A. (2024). Optimization of SVM and Gradient Boosting Models Using GridSearchCV in Detecting Fake Job Postings. *Matrik: Jurnal Manajemen, Teknik Informatika, Dan Rekayasa Komputer*, 23(2), 419–430.
- Schuh, G., Reinhart, G., Prote, J., Sauermann, F., Horsthofer, J., Oppolzer, F., & Knoll, D. (2023). Data Mining Definitions and Applications for the Management of Production Complexity. *Procedia CIRP*, 81(2019), 874–879.  
<https://doi.org/10.1016/j.procir.2019.03.217>
- Sepri, D., Fauzi, A., Wandira, R., Riza, O. S., Wahyuni, Y. F., & Hutagaol, H. (2020). Prediksi Harga Cabai Menggunakan Support Vector Regression. *Computer Based Information System Journal*, 02(02), 1–5.
- Setiyani, L., & Utomo, W. H. (2024). Arabica Coffee Price Prediction Using the Long Short Term Memory Network (LSTM) Algorithm. *Scientific Journal of Informatics*, 10(3), 287–296. <https://doi.org/10.15294/sji.v10i3.44162>

- Shan, F., He, X., Armaghani, D. J., & Sheng, D. (2024). Effects of data smoothing and recurrent neural network (RNN) algorithms for real-time forecasting of tunnel boring machine (TBM) performance. *Journal of Rock Mechanics and Geotechnical Engineering*, 16(5), 1538–1551. <https://doi.org/10.1016/j.jrmge.2023.06.015>
- Sharma, D. K., Dharmaraj, A., Ahmed, A. A. A., Kumar, K. S., Phasinam, K., & Naved, M. (2022). A Study on the Relationship Between Cloud Computing and Data Mining in Business Organizations. *Proceedings of Second International Conference in Mechanical and Energy Technology*, 290(10), 91–99.
- Sharma, H., Harsora, H., & Ogunleye, B. (2024). An Optimal House Price Prediction Algorithm : XGBoost. *Analytics*, 3(1), 30–45.
- Sial, A. H., Rashdi, S. Y. S., & Khan, A. H. (2021). Comparative Analysis of Data Visualization Libraries Matplotlib and Seaborn in Python. *International Journal of Advanced Trends in Computer Science and Engineering*, 10(1), 277–281. <https://doi.org/10.30534/ijatcse/2021/391012021>
- Sihombing, A. (2024). *Fluktuasi Harga Cabai dalam Satu Pekan*. Radio Republik Indonesia. <https://rri.co.id/index.php/palembang/berita-foto/5983/fluktuasi-harga-cabai-dalam-satu-pekan>
- Sina, L. B., Secco, C. A., Blazevic, M., & Nazemi, K. (2023). Hybrid Forecasting Methods — A Systematic Review. *Electronics*, 12(9), 1–14.
- Siregar, N. A. (2022). Peramalan Curah Hujan Di Kota Medan menggunakan Metode Support Vector Regression. *Journal of Informatics and Data Science*, 1(1), 7–9.
- Sriyana, Martha, S., & Sulistianingsih, E. (2019). Prediksi Nilai Tukar Dolar Amerika Serikat Terhadap Rupiah dengan Metode Support Vector Regression (SVR). *Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya*, 08(1), 1–10.
- Subha, J., & Saudia, S. (2024). Precipitation forecast using RNN variants by analyzing Optimizers and Hyperparameters for Time-series based Climatological Data. *International Journal of Electrical and Computer Engineering Systems (IJECEs)*, 15(3), 261–274.
- Susanto, A., & Meiryani. (2019). Functions, Processes, Stages And Application Of Data Mining. *International Journal of Scientific & Technology Research*, 8(07), 136–140.
- Thakur, A., Kumar, A., & Awasthi, S. (2024). Familiarization with Pandas for Data Science and Analysis. *International Research Journal of Modernization in Engineering Technology and Science*, 06(04), 11426–11430.

- Ustadatin, F., Muqtadir, A., & Arifia, A. (2023). Implementasi Metode Weighted Moving Average (WMA) Pada Prediksi Harga Bahan Pokok. *Komputika: Jurnal Sistem Komputer*, *12*(2), 203–210.
- Vardhan, B. V. S., Khedkar, M., Srivastava, I., Thakre, P., & Bokde, N. D. (2023). A Comparative Analysis of Hyperparameter Tuned Stochastic Short Term Load Forecasting for Power System Operator. *Energies*, *16*(3), 1–21.
- Wen, N., & Ling, L. (2024). Evaluation of Cryptocurrency Price Prediction Using LSTM and CNNs Models. *JOIV: International Journal On Informatics Visualization*, *7*(3–2), 2016–2024.
- Wibowo, H. E., Novanda, R. R., Khaliqi, M., Sinaga, F. H., Darmansyah, D., Amiruddin, A., & Sari, I. R. M. (2023). Volatility analysis of chili price at Bengkulu Province using ARIMA approach. *IOP Conference Series: Earth and Environmental Science*, *1230*(1), 1–10. <https://doi.org/10.1088/1755-1315/1230/1/012013>
- Winkowski, C. (2019). Classification of forecasting methods in production engineering. *Engineering Management In Product and Services*, *11*(4), 23–33. <https://doi.org/10.2478/emj-2019-0030>
- Yi, D., Ahn, J., & Ji, S. (2020). An Effective Optimization Method for Machine Learning Based on ADAM. *Applied Sciences*, *10*(3), 1–20.