

Perbandingan Kinerja LSTM, Random Forest, dan SVR Berbasis Knowledge Discovery untuk Prediksi Harga Beras Sumatera Selatan

Cheisyah Andini Bahri, Ken Ditha Tania*

Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Sriwijaya, Palembang, Indonesia

Email:¹ 09031182227023@student.unsri.ac.id, ^{2*} kenya.tania@gmail.com

Email Penulis Korespondensi: kenya.tania@gmail.com

Submitted 15-10-2025; Accepted 30-10-2025; Published 31-10-2025

Abstrak

Beras merupakan kebutuhan pokok utama di Indonesia, khususnya Provinsi Sumatera Selatan. Pada februari 2024, BBC News Indonesia melaporkan harga beras premium melonjak hingga Rp18.000 per kilogram, menjadi harga tertinggi dalam sejarah Indonesia. Untuk mengantisipasi dan memprediksi lonjakan serupa di masa depan, penelitian ini menerapkan pendekatan *Knowledge Discovery* serta membandingkan tiga model *machine learning*, yaitu LSTM, *Random Forest*, dan SVR. Pendekatan ini dilakukan melalui tahapan seleksi data, pembersihan, transformasi, pemodelan, dan evaluasi, guna menggali pola tersembunyi dalam data historis. Dataset yang digunakan diperoleh dari situs resmi PIHPS Nasional, berupa data harian harga beras dari Januari 2020 hingga Mei 2025 sebanyak 1.412 data. Evaluasi performa model dilakukan menggunakan metrik MAPE, MAE, dan RMSE sebagai tolak ukur. Hasil penelitian menunjukkan bahwa model SVR memberikan kinerja terbaik dibandingkan LSTM, dan *Random Forest*. Pada kategori beras kualitas super II, model SVR menampilkan nilai MAPE sebesar 0.00 persen, MAE 40.93, dan RMSE 52.54. Model SVR juga konsisten menghasilkan kesalahan prediksi terendah pada kategori lain seperti beras dengan kualitas bawah I MAE 59.39, kualitas medium I MAE 38.92. Penelitian ini diharapkan dapat menjadi dasar pengembangan sistem monitoring harga pangan berbasis *machine learning* untuk mendukung kebijakan yang lebih responsif serta menjaga stabilitas harga beras di masa depan.

Kata Kunci: Harga Beras; Long Short Term Memory; Random Forest; Support Vector Regression; Sumatera Selatan

Abstract

Rice is a primary staple food in Indonesia, particularly in South Sumatra Province. In February 2024, BBC News Indonesia reported that the price of premium rice surged to Rp18,000 per kilogram, marking the highest price in the country's history. To anticipate and predict similar spikes in the future, this study applies a Knowledge Discovery approach and compares three machine learning models: LSTM, Random Forest, and SVR. The approach follows the stages of data selection, cleaning, transformation, modeling, and evaluation to uncover hidden patterns in historical data. The dataset, obtained from the official PIHPS Nasional website, consists of 1,412 daily rice price records from January 2020 to May 2025. Model performance was evaluated using MAPE, MAE, and RMSE metrics. The findings indicate that the SVR model outperformed LSTM and Random Forest, delivering the most accurate results. For the Super Quality II rice category, SVR achieved a MAPE of 0.00 percent, MAE of 40.93, and RMSE of 52.54. SVR also consistently produced the lowest prediction errors in other categories, such as Low Quality I (MAE 59.39) and Medium Quality I (MAE 38.92). This research is expected to serve as a foundation for developing machine learning-based food price monitoring systems to support more responsive policies and maintain rice price stability in the future.

Keywords: Rice Price; Long Short Term Memory; Random Forest; Support Vector Regression; South Sumatra

1. PENDAHULUAN

Beras merupakan salah satu kebutuhan pokok utama negara Indonesia khususnya di Provinsi Sumatera Selatan. Dengan pertumbuhan penduduk yang terus meningkat di setiap wilayah, beras secara signifikan mempengaruhi garis kemiskinan masyarakat Indonesia baik di lingkungan kota maupun desa[1]. Beras sebagai salah satu komoditas bahan pangan paling berpengaruh terhadap kesejahteraan hidup, dan perekonomian negara, hal ini dapat menyebabkan meningkatnya nilai permintaan yang mempengaruhi kebijakan pemerintah dalam menentukan nilai dan harga bahan pokok di Indonesia[2].

Hal ini juga diperkuat oleh pernyataan[3] inflasi bahan pangan merupakan faktor lain selain peningkatan jumlah penduduk, dengan menjadi salah satu tantangan di bidang ekonomi yang cukup krusial, di mana dapat menimbulkan dampak signifikan khususnya pada negara-negara berkembang, dalam ekonomi global perubahan yang muncul sangat mempengaruhi stabilitas pasar. Fenomena inflasi ini menjadi semakin memburuk terjadi ketika pandemi COVID-19 yang membawa pengaruh terhadap distribusi beras dari daerah produsen, di mana hal ini dapat menyebabkan kelangkaan sehingga mempengaruhi harga beras pada suatu pasar[4]. Kemudian konflik negara Rusia dan Ukraina pada Maret 2022 diketahui bahwa Rusia merupakan produsen utama pupuk dunia, pada saat konflik berlangsung negara Rusia membatasi ekspor pupuk yang menyebabkan biaya produksi padi meningkat di berbagai negara termasuk Indonesia, kedua hal tersebut yang menjadi pemicu krisis rantai pasokan global[5].

Berdasarkan pemberitaan yang disampaikan oleh BBC News Indonesia menjelaskan tentang kenaikan harga beras premium terjadi pada bulan Februari 2024 mencapai Rp18.000 per kilogram menjadi harga tertinggi sepanjang sejarah Indonesia[6]. Hal ini mendorong pemerintah Indonesia untuk melakukan kegiatan impor ke beberapa negara produksi beras yaitu pada tahun 2018 negara Thailand menjadi negara pengekspor beras terbesar di Indonesia mencapai 795.600,1 ton setiap tahun, dan pada tahun 2019 negara Pakistan yang mengimpor beras selanjutnya mencapai 182.564,9 ton setiap tahunnya[7]. Berdasarkan peraturan pemerintah Pasal 55 ayat 1 Undang-Undang Pangan Nomor 18 tahun 2012, pemerintah memiliki tanggung jawab untuk menstabilkan ketersediaan dan harga pangan pokok di tingkat produsen maupun konsumen. Hal ini juga dijelaskan dalam Peraturan Presiden Nomor 59 Tahun 2020 tentang pengendalian harga

bahan makanan pokok, yang memberikan tanggung jawab kepada pemerintah untuk mengatur nilai jual standar pembelian dan penjualan komoditas pangan strategis. Dalam pengendalian harga bahan pokok, Dinas Perdagangan Kota Palembang menggunakan data komoditas pangan sebagai referensi utama dalam penetapan harga untuk mendukung kebijakan pemerintah[8].

Berdasarkan permasalahan tersebut, diperlukan alat prediksi berbasis *machine learning* karena dinilai lebih optimal dan akurat dibandingkan metode konvensional. *Machine learning* dapat mempelajari pola dari data historis yang bervariasi, sesuai dengan konteks *data mining*, proses prediksi atau *forecasting* digunakan untuk memperoleh pengetahuan mengenai kejadian di masa mendatang, salah satu pendekatan yang sering diterapkan dalam melakukan prediksi adalah penggunaan model regresi yang bertujuan untuk meminimalkan kesalahan hasil peramalan, serta umum dipakai dalam pemodelan deret waktu[9]. Pendekatan deret waktu digunakan untuk memperkirakan keadaan yang akan datang dengan mengacu pada data masa lalu[10]. Saat ini, peramalan menggunakan deret waktu merupakan hal penting dan cukup kompleks dalam pembelajaran mesin, karena harga selalu cenderung berkorelasi dengan waktu[11]. Penggunaan pembelajaran mesin memungkinkan model prediksi dalam menyesuaikan diri secara lebih fleksibel terhadap pola musiman dan fluktuasi data yang kompleks yang sulit ditangani menggunakan metode statistik biasa. Kegiatan penambangan data sebagai bagian dari proses *knowledge discovery*, yang melibatkan tahapan-tahapan dalam *Knowledge Discovery in Databases* (KDD) berperan penting dalam mengidentifikasi pola tersembunyi dari kumpulan data besar secara otomatis, melalui tahapan yang meliputi pra pemrosesan, pemodelan, dan evaluasi hasil dalam bentuk sistematis[12].

Menurut penelitian yang dilakukan oleh[13] melakukan analisis terhadap dua jenis jaringan saraf, yaitu *Multilayer Perceptron* (MLP) dan *Long Short Term Memory* (LSTM), dengan berfokus memprediksi harga beras dengan jenis IR64 kualitas III, di mana penelitian tersebut menghasilkan model LSTM memberikan hasil lebih optimal dibandingkan model MLP, kinerja ini diukur melalui nilai RMSE yang menunjukkan angka lebih kecil pada data latih (0.49) dan data uji (0.27), serta grafik prediksi sangat mendekati data aktual, temuan tersebut memperlihatkan bahwa algoritma LSTM yang merupakan turunan dari metode RNN juga termasuk ke dalam metode *Deep Learning* yang relevan diterapkan pada prediksi dengan pola deret waktu[14]. Kemudian penelitian[15] menerapkan LSTM dalam kerangka *knowledge discovery* untuk menganalisis indeks kualitas udara, dan melaporkan bahwa LSTM lebih efektif dalam menangkap pola non-linier pada data deret waktu dibandingkan pendekatan konvensional seperti SARIMA.

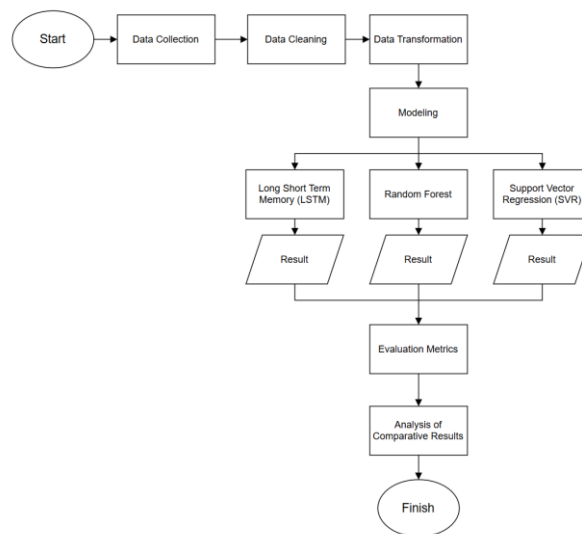
Sementara itu[16] menunjukkan bahwa penggunaan *Random Forest* memiliki kemampuan untuk menyajikan hasil prediksi yang paling akurat dibandingkan metode ARIMA dan MLP dalam memprediksi menggunakan deret waktu pada studi kasus harga logam mulia, hasil penelitian dari metode *Random Forest* menghasilkan MAPE sebesar 2,85 persen, lebih kecil dibandingkan metode ARIMA dan MLP sebesar 18 persen, ini menunjukkan keunggulan dari metode yang menggunakan kombinasi banyak pohon keputusan dengan teknik *bagging*, sehingga sangat baik dalam menghadapi ketidakstabilan data, oleh karena itu metode ini layak digunakan untuk kegiatan peramalan menggunakan deret waktu terutama ketika terdapat data yang bersifat tidak stabil atau menunjukkan tingkat fluktuasi yang tinggi.

Penelitian selanjutnya dilakukan oleh[17] menunjukkan bahwa model *Support Vector Regression* (SVR) mampu memberikan hasil kerja yang baik dalam memprediksi harga dari 17 jenis dari komoditas pokok nasional, di mana penelitian ini menggunakan deret waktu harian tahun 2020 hingga 2022 menunjukkan hasil dengan nilai error rendah seperti MSE sebesar 6.062, MAD sebesar 6.6831, MAPE sebesar 0.0147, RMSE sebesar 7.746, ini mengindikasikan bahwa model *Support Vector Regression* (SVR) efektif dalam menangani peramalan harga pangan dengan kompleksitas tinggi.

Meskipun ketiga algoritma tersebut telah terbukti unggul dalam penelitian masing-masing, belum banyak penelitian yang secara langsung menggabungkan dan membandingkan ketiganya, khususnya untuk kasus prediksi harga beras yang ada di pasar tradisional Provinsi Sumatera Selatan. Selain itu belum banyak penelitian yang mengaitkan penerapan algoritma prediksi dengan pendekatan *knowledge discovery* untuk menganalisis pola harga berdasarkan kategori beras dan faktor waktu. Penelitian ini bertujuan mengisi celah penelitian sebelumnya melalui penerapan tiga model unggulan yaitu LSTM, *Random Forest*, dan SVR ke dalam kerangka *Knowledge Discovery in Databases* (KDD). Ketiga model ini dipilih berdasarkan kebutuhan prediksi yang mampu memproses data dengan pola non-linier dan kompleksitas tinggi dengan mempertimbangkan faktor waktu, dengan mencari nilai prediksi terbaik antara ketiga model tersebut diukur berdasarkan algoritma MAPE, MAE, dan RMSE sebagai tolak ukur. Hasil penelitian ini diharapkan berguna dalam mengevaluasi dan membandingkan kinerja beberapa metode peramalan harga beras dan mendukung pengembangan sistem prediksi berbasis *machine learning* dan *knowledge discovery* untuk pengelolaan pasokan dalam menjaga stabilitas harga pangan di Indonesia khususnya wilayah Provinsi Sumatera Selatan.

2. METODOLOGI PENELITIAN

Metode penelitian ini mengadopsi pendekatan proses *knowledge discovery*, dengan menjadikan tahapan *Knowledge Discovery in Databases* (KDD) sebagai kerangka utama penelitian peramalan harga beras. Pendekatan ini meliputi tahapan seperti pengumpulan dan seleksi data, transformasi, pemodelan, serta evaluasi hasil. Pemanfaatan metode ini memungkinkan penemuan pola yang kompleks, relevan, dan bermanfaat dari data dalam jumlah besar, sekaligus menjembatani pendekatan deduktif dan induktif dalam proses penelitian ilmiah[18][19]. Adapun tahapan dalam penelitian ini ditampilkan pada Gambar 1 alur penelitian digunakan untuk menggambarkan proses analisis mulai dari pengumpulan data hingga evaluasi hasil model.



Gambar 1. Alur Penelitian

Tahapan awal penelitian, sebagaimana ditunjukkan pada Gambar 1, dimulai dengan kegiatan pengumpulan data akan dipakai, kemudian dilakukan analisis prediksi harga beras menggunakan model yang unggul dalam memproses data *time series* sehingga memperoleh model terbaik dalam memprediksi harga beras di pasar tradisional Provinsi Sumatera Selatan.

2.1 Data Collection

Penelitian ini memanfaatkan data sekunder dari portal resmi Pusat Informasi Harga Pangan Strategis Nasional (PIHPS Nasional). Dataset tersebut berupa informasi harga beras harian di Provinsi Sumatera Selatan mulai Januari 2020 hingga Mei 2025, dengan total jumlah 1412 data. Variabel utama yang dipertimbangkan adalah tanggal dan harga untuk mendukung proses analisis dan prediksi. Dataset yang digunakan sebagai dasar analisis penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Dataset Harga Beras Provinsi Sumatera Selatan

Date	Beras Kualitas Bawah I	Beras Kualitas Bawah II	Beras Kualitas Medium I	Beras Kualitas Medium II	Beras Kualitas Super I	Beras Kualitas Super II
01/01/2020	10200.0	10350.0	11300.0	11150.0	13700.0	12350.0
02/01/2020	10200.0	10350.0	11300.0	11150.0	13700.0	12350.0
03/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
04/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
05/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
06/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
07/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
08/01/2020	10200.0	10100.0	11550.0	11000.0	13950.0	12450.0
...	...	...	...	...	...	...
27/05/2025	15100.0	13650.0	13400.0	15650.0	15700.0	16700.0
28/05/2025	15200.0	13650.0	13400.0	15650.0	15750.0	16700.0

Tabel 1 menunjukkan data harga beras harian di pasar tradisional Provinsi Sumatera Selatan dari Januari 2020 hingga Mei 2025 yang terdiri atas enam kategori, yaitu beras kualitas bawah I, bawah II, medium I, medium II, super I, dan super II. Data tersebut menjadi dasar dalam proses analisis dan pelatihan model prediksi untuk mengetahui pola perubahan harga berdasarkan waktu.

2.2 Data Cleaning

Pembersihan dilakukan menggunakan metode interpolasi dengan mengestimasi data yang hilang berdasarkan nilai-nilai terdekat, kemudian dilakukan pengecekan kembali apakah masih ada baris yang memiliki nilai NaN, maka baris tersebut akan dihapus guna menjamin kebersihan dan konsistensi data yang digunakan.

2.3 Data Transformation

Penelitian ini melibatkan proses transformasi data yang ditujukan untuk mempersiapkan data agar dapat digunakan dengan optimal dalam proses pemodelan pada *machine learning*, adapun tahapan transformasi yang digunakan yaitu:

a. Mengubah Tipe Data Tanggal

Pada dataset yang digunakan memiliki kolom tanggal dengan format *object*, tipe data ini belum bisa digunakan untuk analisis menggunakan deret waktu, oleh karena itu dilakukan perubahan pada kolom ini menjadi format *datetime* yang mendukung penggunaan model deret waktu.

b. Mengubah Fitur Lag dan Musim

Penelitian ini menggunakan data historis tanggal dan harga dari masing-masing jenis beras yang tersedia untuk membantu model dalam memahami pola perubahan harga dari waktu ke waktu, maka diperlukan penambahan fitur lag1 yaitu nilai harga satu hari sebelumnya, lag7 yaitu nilai harga tujuh hari sebelumnya, dan *season* yaitu klasifikasi musim ditentukan berdasarkan bulan.

c. Mengatur Ulang Indeks Baris

Setelah dilakukan penambahan fitur lag, dan *season* beberapa baris awal menjadi kosong dan tidak valid karena nilai null, baris tersebut dihapus, kemudian indeks baris diatur ulang agar data tersusun secara rapi dan berurutan, di mana langkah ini penting untuk menjaga konsistensi struktur data dan mencegah kesalahan saat proses pemodelan.

d. Split Data

Data sebanyak 1.412 dibagi menjadi dua bagian, 1.123 untuk pelatihan dan 281 untuk pengujian dengan rasio 80:20. Pembagian ini dilakukan untuk mengevaluasi sejauh mana model dapat bekerja pada data yang belum pernah digunakan selama pelatihan.

e. Standarisasi Data

Setelah dilakukan pembagian data, tahapan berikutnya yaitu melakukan standarisasi data dengan menggunakan *StandardScaler*, proses ini menghasilkan distribusi dengan rata-rata nol dan deviasi standar satu, sehingga varian menjadi satu, karena merupakan hasil kuadrat dari deviasi standar[20]. Mengatur ulang skala data diperlukan supaya semua fitur bisa bekerja secara adil saat membangun model, terutama untuk model yang peka terhadap perbedaan ukuran nilai antar data.

f. Memilih Fitur Dataset yang Digunakan

Tahapan selanjutnya yaitu dilakukan pemilihan fitur (*feature selection*) yang akan digunakan pada saat proses pelatihan model, fitur dari variabel independen (X) meliputi lag1, lag7, dan *season* (musim). Rincian variabel independen yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2.

Tabel 2. Variabel Independen (X)

Jenis Beras	Lag1	Lag7	Season (musim)
Beras Kualitas Bawah I	Beras Kualitas Bawah I_lag1	Beras Kualitas Bawah I_lag7	Beras Kualitas Bawah I_season
Beras Kualitas Bawah II	Beras Kualitas Bawah II_lag1	Beras Kualitas Bawah II_lag7	Beras Kualitas Bawah II_season
Beras Kualitas Medium I	Beras Kualitas Medium I_lag1	Beras Kualitas Medium I_lag7	Beras Kualitas Medium I_season
Beras Kualitas Medium II	Beras Kualitas Medium II_lag1	Beras Kualitas Medium II_lag7	Beras Kualitas Medium II_season
Beras Kualitas Super I	Beras Kualitas Super I_lag1	Beras Kualitas Super I_lag7	Beras Kualitas Super I_season
Beras Kualitas Super II	Beras Kualitas Super II_lag1	Beras Kualitas Super II_lag7	Beras Kualitas Super II_season

Tabel 2 menunjukkan variabel independen yang digunakan sebagai fitur masukan dalam model prediksi. Setiap jenis beras memiliki tiga bentuk variabel turunan, yaitu *lag1* (harga sehari sebelumnya), *lag7* (harga satu minggu sebelumnya), dan *season* (musim), yang digunakan untuk menangkap pola waktu dan musiman pada data harga beras. Variabel-variabel ini membantu model dalam mengenali tren serta fluktuasi harga pada periode tertentu.

Harga aktual dari masing-masing jenis beras digunakan sebagai variabel target (Y). Tabel 3 berikut menampilkan penjelasan lengkapnya.

Tabel 3. Variabel Target (Y)

Jenis Beras	Nama Target (Y)
Beras Kualitas Bawah I	Beras Kualitas Bawah I
Beras Kualitas Bawah II	Beras Kualitas Bawah II
Beras Kualitas Medium I	Beras Kualitas Medium I
Beras Kualitas Medium II	Beras Kualitas Medium II
Beras Kualitas Super I	Beras Kualitas Super I
Beras Kualitas Super II	Beras Kualitas Super II

Tabel 3 menampilkan informasi variabel target (Y) dipakai dalam penelitian yang menggambarkan harga sebenarnya dari setiap jenis beras yang ada di pasar tradisional Provinsi Sumatera Selatan.

g. Menampilkan Ukuran Dataset Untuk Validasi

Untuk menjaga integritas data, dilakukan pemeriksaan jumlah baris dan kolom setelah proses transformasi. Hal ini bertujuan memastikan tidak ada data yang terhapus ataupun terjadi kesalahan dalam tahapan transformasi berlangsung.

2.4 Modeling

a. Long Short Term Memory (LSTM)

Algoritma LSTM dikategorikan sebagai penyempurnaan struktur dari model *Recurrent Neural Network* (RNN), di mana pengembangan ini dilakukan untuk mengatasi permasalahan utama pada jaringan saraf *Recurrent Neural Network* (RNN) konvensional, terutama dalam menyimpan dan memproses informasi jangka panjang pada data sekuensial[21]. LSTM memiliki struktur memori khusus yaitu *cell state*, berguna untuk menyimpan informasi dalam jangka panjang, dengan menggunakan tiga gerbang utama yang bekerja secara bersamaan, yaitu:

Gerbang *forget* (*ft*) berfungsi untuk memilih bagian informasi dari *cell state* sebelumnya yang perlu dihilangkan. Nilai keluarannya berkisar antara 0 dan 1, nilai 0 menandakan bahwa informasi dibuang sepenuhnya, sedangkan nilai 1 berarti informasi tetap dipertahankan.

$$ft = \sigma(Wf \cdot [ht - 1, xt] + bf) \quad (1)$$

Fungsi utama gerbang *input* (*it*) adalah mengendalikan seberapa banyak informasi baru dimasukkan ke memori saat ini. Kandidat memori ($C \sim t$) digunakan bersama-sama untuk memperbarui isi dari *cell state*.

$$it = \sigma(Wi \cdot [ht - 1, xt] + bi) \quad (2)$$

$$C \sim t = \tanh(WC \cdot [ht - 1, xt] + bC) \quad (3)$$

Gerbang *output* (*ot*) mengatur bagian dari *cell state* yang akan digunakan sebagai keluaran tersembunyi (*ht*) pada waktu saat ini, kemudian diteruskan pada tahapan selanjutnya untuk digunakan pada kegiatan prediksi.

$$ot = \sigma(Wo \cdot [ht - 1, xt] + bo) \quad (4)$$

$$ht = ot * \tanh(Ct) \quad (5)$$

b. Random Forest

Model kedua yang digunakan yaitu *Random Forest* merupakan salah satu bentuk algoritma hasil pengembangan sebagai penyempurnaan dari algoritma *Decision Tree*[22]. Metode *Random Forest* dalam *machine learning* adalah metode yang tangguh dan fleksibel, karena mampu menyelesaikan permasalahan klasifikasi maupun regresi, di mana algoritma ini termasuk teknik *ensemble learning* yaitu pendekatan dengan penggabungan hasil peramalan dari sejumlah model untuk memperoleh keakuratan dan kestabilan yang lebih baik[23].

Untuk tahapan pengukuran kestabilan dari hasil prediksi antar pohon dalam model *Random Forest*, menggunakan rumus varians berikut:

$$Var(t) = \frac{1}{N_t} \sum_{i=1}^{N_t} (y_i - \bar{y})^2 \quad (6)$$

Rumus ini digunakan untuk mengukur penyebaran hasil prediksi. Semakin rendah nilai varians, maka hasil prediksi antar pohon semakin konsisten yang menunjukkan model lebih stabil. Pada rumus tersebut (N_t) merujuk pada jumlah sampel dalam node (t), (y_i) merupakan nilai target dari sampel ke- i , sedangkan ($\bar{y}$) menyatakan nilai rata-rata dari target pada jumlah sampel pada node tersebut.

c. Support Vector Regression (SVR)

SVR merupakan teknik prediksi yang optimal dalam menangani prediksi non-linier dengan memanfaatkan fungsi kernel, sehingga cocok digunakan untuk data dengan kompleksitas variabel yang tinggi[24]. Vapnik mengenalkan metode ini pada tahun 1999, dan sejak itu SVR berkembang menjadi salah satu algoritma populer. Dengan menggunakan fungsi ε -insensitive loss sebagai pengganti fungsi $L2$ -norm loss yang umum digunakan, oleh karena itu SVR sudah dianggap sebagai alternatif yang kuat dalam menghadapi *outlier* serta mampu mengurangi risiko *overfitting* dalam konteks regresi linier[25].

$$f(x) = w^T \phi(x) + b \quad (7)$$

Secara umum, model SVR membangun fungsi regresi untuk menghubungkan variabel *input* (x) dan *output*. Nilai (x) menggambarkan vektor masukan, dengan (w^T) berperan sebagai bobot, sedangkan $\phi(x)$ menjalankan peran sebagai fungsi non-linier yang melakukan pemetaan dari ruang berdimensi rendah ke ruang dengan dimensi lebih tinggi.

Parameter (b) berfungsi sebagai komponen bias dalam model. Apabila kernel yang digunakan berbentuk linear, fungsi $\phi(x)$ sama dengan (x) sehingga model bersifat linier secara langsung.

2.5 Evaluation Metrics

Model prediksi dapat dikatakan baik apabila model tersebut memiliki nilai kesalahan (*error*) yang rendah berdasarkan metrik evaluasi. Dalam penelitian ini tingkat kesalahan prediksi diukur berdasarkan metrik evaluasi utama, yaitu:

a. Mean Absolute Percentage Error (MAPE)

Dalam proses evaluasi akurasi prediksi, perhitungan MAPE sering dipakai untuk menilai tingkat keandalan hasil peramalan. Parameter ini menghitung persentase kesalahan rata-rata dari selisih absolut antara nilai aktual dengan hasil estimasi model. Semakin rendah nilai yang diperoleh, semakin tinggi tingkat ketepatan model dalam memprediksi[26].

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (8)$$

b. Mean Absolute Error (MAE)

Perhitungan menggunakan MAE digunakan untuk mengevaluasi kinerja algoritma atau model yang diterapkan dengan menghitung rata-rata selisih absolut antara nilai observasi dan hasil prediksi dalam melakukan peramalan[27]. Nilai MAE yang lebih kecil mengindikasikan hasil peramalan lebih tepat dan mendekati data sebenarnya, sedangkan nilai yang lebih tinggi memperlihatkan adanya penyimpangan prediksi yang cukup besar. Perhitungan MAE dinyatakan dalam persamaan berikut:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (9)$$

c. Root Mean Squared Error (RMSE)

Dalam analisis performa model RMSE digunakan sebagai metrik evaluasi dan tolak ukur kesesuaian hasil prediksi terhadap data sebenarnya[28]. Nilai RMSE yang lebih kecil mengindikasikan bahwa hasil peramalan yang dihasilkan semakin mendekati nilai sebenarnya. RMSE dinyatakan dalam persamaan berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (10)$$

Keterangan:

n = banyaknya data

t = indeks waktu (dari $t = 1$ sampai $t = n$)

y_t = nilai aktual

$\hat{y}_t$ = nilai hasil prediksi

Σ = notasi penjumlahan (sigma)

2.6 Analysis of Comparative Results

Tahap ini merupakan tahap akhir penelitian untuk menggali informasi dan menerapkan konsep *knowledge discovery* dari hasil evaluasi model peramalan. Hasil evaluasi dibandingkan untuk menentukan model dengan performa terbaik, di mana model dengan error terendah dianggap paling akurat meramalkan harga beras di Provinsi Sumatera Selatan periode Januari 2020 hingga Mei 2025.

3. HASIL DAN PEMBAHASAN

3.1 Penerapan dan Hasil Model Long Short Term Memory (LSTM)

Model *Long Short Term Memory* (LSTM) diterapkan untuk memprediksi harga beras berdasarkan data deret waktu di Provinsi Sumatera Selatan. Pada tahap awal, proses normalisasi dilakukan menggunakan *StandardScaler* sehingga setiap fitur memiliki skala nilai yang sebanding dan menjaga stabilitas model. Penerapan model LSTM dilakukan dengan arsitektur jaringan yang terdiri dari 160 unit neuron menggunakan fungsi aktivasi *tanh*, *optimizer Adam*, ukuran *batch* 32, serta pelatihan sebanyak 100 *epoch*. Proses split data dilakukan dengan menerapkan rasio 80:20, dengan pengurutan waktu tetap dijaga (*shuffle=False*) agar pola temporal pada data tidak berubah. Model ini diterapkan untuk setiap jenis beras dengan memanfaatkan fitur *lag1*, *lag7*, dan *season* sebagai variabel masukan, serta harga aktual sebagai variabel target. Setelah model melalui tahap pelatihan, data uji diproses untuk mendapatkan prediksi, yang selanjutnya dikembalikan ke skala semula melalui proses *inverse transform*. Untuk menilai kinerja model, dilakukan evaluasi menggunakan tiga metrik utama, yaitu MAPE, MAE, dan RMSE, ketiga metrik ini digunakan untuk mengukur kesalahan

dan akurasi hasil prediksi. Hasil evaluasi model LSTM untuk masing-masing kategori beras ditampilkan pada Tabel 4 berikut.

Tabel 4. Evaluasi Model LSTM

Kategori Beras	MAPE	MAE	RMSE
Kualitas Bawah I	0.01%	183.99	227.15
Kualitas Bawah II	0.03%	352.18	406.67
Kualitas Medium I	0.02%	334.18	412.56
Kualitas Medium II	0.01%	129.40	146.30
Kualitas Super I	0.01%	124.88	168.72
Kualitas Super II	0.01%	99.38	114.46

Tabel 4 menyajikan hasil penerapan model LSTM, yang menggambarkan adanya variasi kinerja model pada tiap kategori kualitas beras. Hal ini ditunjukkan oleh metrik yang digunakan yaitu nilai MAPE 0.01 persen hingga 0.03 persen, ini menunjukkan tingkat kesalahan rendah, sementara itu nilai MAE tertinggi tercatat pada kategori beras kualitas bawah II sebesar 352.18, sedangkan nilai terendah ada pada kategori beras kualitas super II sebesar 99.38, Secara keseluruhan, model LSTM mampu memprediksi pola perubahan harga beras dengan akurasi baik, meskipun RMSE pada kualitas bawah II dan medium I relatif tinggi.

3.2 Penerapan dan Hasil Model *Random Forest*

Penerapan model *Random Forest* dilakukan untuk memprediksi harga beras dengan menggunakan data historis yang telah dinormalisasi melalui metode *StandardScaler*. Pelatihan model dilakukan dengan pendekatan *Grid Search Cross Validation* untuk menemukan kombinasi parameter terbaik dari dua variabel utama, yaitu jumlah pohon keputusan (*n_estimators*) dan kedalaman maksimum pohon (*max_depth*). Pelatihan dilakukan dengan pengaturan *n_estimators* sebesar 50 dan 100, serta *max_depth* bernilai 10 dan *None*. Data dibagi menjadi 80 persen untuk melatih model, dan 20 persen lainnya dipakai untuk menguji performa model. Model dilatih untuk masing-masing kategori beras dengan memanfaatkan variabel *lag1*, *lag7*, dan *season* sebagai fitur masukan, serta data aktual sebagai target. Model yang telah selesai dilatih kemudian diterapkan pada data uji yang sebelumnya distandarisasi. Prediksi yang diperoleh selanjutnya dikonversi kembali ke skala semula menggunakan *inverse transform*, sehingga hasilnya bisa digunakan untuk evaluasi. Proses evaluasi menggunakan tiga metrik yaitu MAPE, MAE, dan RMSE. Hasil perhitungan metrik evaluasi model *Random Forest* untuk setiap kategori beras ditampilkan pada Tabel 5 berikut.

Tabel 5. Evaluasi Model *Random Forest*

Kategori Beras	MAPE	MAE	RMSE
Kualitas Bawah I	0.01%	130.42	148.52
Kualitas Bawah II	0.04%	474.86	501.41
Kualitas Medium I	0.01%	149.06	167.92
Kualitas Medium II	0.01%	197.55	230.85
Kualitas Super I	0.01%	143.98	171.71
Kualitas Super II	0.02%	266.98	323.62

Tabel 5 menampilkan hasil pengujian dan evaluasi dari model *Random Forest* dalam memprediksi harga beras dari masing-masing kategori di wilayah Sumatera Selatan menunjukkan hasil yang cukup kompetitif dengan MAPE antara 0.01 persen hingga 0.04 persen, kemudian hasil MAE menunjukkan nilai terendah pada kategori beras kualitas bawah I sebesar 130.42, meski nilai RMSE yang dihasilkan cenderung lebih tinggi daripada model lainnya seperti *Support Vector Regression* (SVR), model ini tetap mampu memberikan hasil yang cukup stabil dalam meramalkan harga beras terutama pada kategori beras kualitas menengah ke atas. Oleh karena itu *Random Forest* masih dapat dipertimbangkan sebagai model alternatif atau pelengkap dalam proses peramalan harga beras.

3.3 Penerapan dan Hasil Prediksi *Support Vector Regression* (SVR)

Proses pelatihan model *Support Vector Regression* (SVR) dimulai menggunakan data yang sudah dinormalisasi menggunakan *StandardScaler* kemudian menerapkan metode *Grid Search Cross Validation* untuk menentukan parameter terbaik, meliputi nilai *C*, *epsilon*, dan jenis *kernel*. Parameter *C* diujikan dengan rentang nilai 0.1 hingga 1000, *epsilon* yang digunakan 0.1 dan 1.0, serta penggunaan *kernel linear* yang dipilih karena mampu menangani data dengan hubungan yang bersifat *linear* dan stabil. Model SVR dilatih menggunakan kategori beras dengan variabel masukan *lag1*, *lag7*, dan *season*, serta harga sebenarnya sebagai target. Setelah proses pelatihan, model digunakan untuk memprediksi data uji yang telah diskalakan, kemudian hasilnya dikonversi kembali ke skala semula melalui proses *inverse transform* agar nilai prediksi sesuai dengan satuan harga aktual. Tahap akhir dari proses ini adalah evaluasi model menggunakan tiga metrik pengukuran, yaitu MAPE, MAE, dan RMSE. Hasil perhitungan dari masing-masing metrik disajikan pada Tabel 6.

Tabel 6. Evaluasi Model Support Vector Regression

Kategori Beras	MAPE	MAE	RMSE
Kualitas Bawah I	0.00%	59.39	77.60
Kualitas Bawah II	0.01%	119.45	146.47
Kualitas Medium I	0.00%	38.92	62.10
Kualitas Medium II	0.00%	57.30	84.09
Kualitas Super I	0.00%	37.31	76.48
Kualitas Super II	0.00%	40.93	52.54

Tabel 6 menunjukkan hasil evaluasi model SVR merupakan kinerja paling konsisten dibandingkan model lainnya dengan nilai metrik MAPE berada pada persentase yang rendah di semua kategori beras yaitu 0.00 persen hingga 0.01 persen. Model SVR ini juga menunjukkan nilai MAE dan RMSE yang paling rendah, terutama pada kategori beras dengan kualitas medium I yaitu dengan nilai MAE 38.92 dan nilai RMSE 62.10, kemudian pada kategori beras kualitas super I dengan nilai MAE yaitu 37.31 dan RMSE 76.48. Berdasarkan nilai MAPE, MAE, dan RMSE yang paling rendah, model SVR terbukti menghasilkan prediksi yang paling akurat dan stabil dibandingkan dua model lainnya pada penelitian ini.

3.4 Uji Signifikansi Statistik (*Paired t-test*)

Setelah hasil evaluasi diperoleh, dilakukan analisis lanjutan guna menilai apakah terdapat perbedaan kinerja antar model yang signifikan. Pengujian *paired t-test* diterapkan dengan menggunakan nilai rata-rata semua MAPE dari setiap variabel target (Y) yang telah dihitung sebelumnya, sebagaimana disajikan pada Tabel 7 berikut.

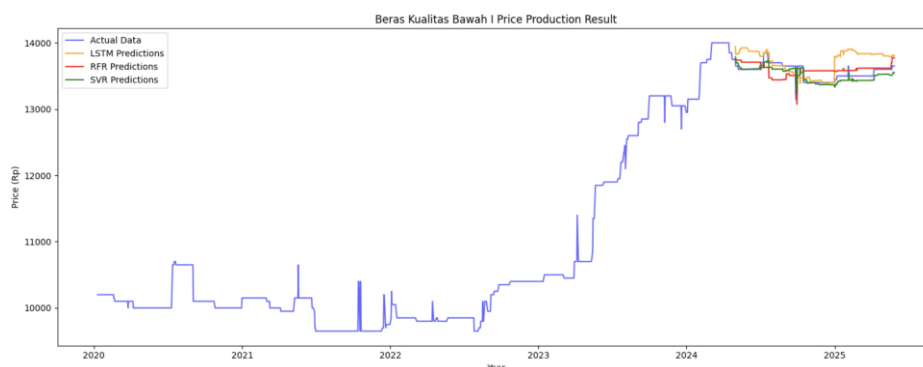
Tabel 7. Hasil Uji t-Test Antar Model

Model	T-Statistik	P-Value	Signifikansi ($\alpha = 0.05$)	Model yang Lebih Unggul
LSTM vs Random Forest	-0.4469	0.6736	Tidak signifikan	-
LSTM vs SVR	3.5689	0.0161	Signifikan	SVR
Random Forest vs SVR	3.4878	0.0175	Signifikan	SVR

Berdasarkan hasil *paired t-test* pada Tabel 7, diketahui bahwa perbandingan antara model LSTM dan *Random Forest* menunjukkan *p-value* sebesar 0.6736 (> 0.05), sehingga perbedaan kinerjanya tidak signifikan secara statistik. Sebaliknya, perbandingan antara LSTM dan SVR *p-value* = 0.0161, serta *Random Forest* dan SVR *p-value* = 0.0175 kedua nilai tersebut (< 0.05), hal ini menunjukkan adanya perbedaan signifikan, dengan nilai MAPE model SVR yang lebih rendah. Hasil ini menegaskan bahwa model SVR memiliki performa paling unggul dibandingkan LSTM dan *Random Forest* dalam memprediksi harga beras.

3.5 Visualisasi Hasil Prediksi

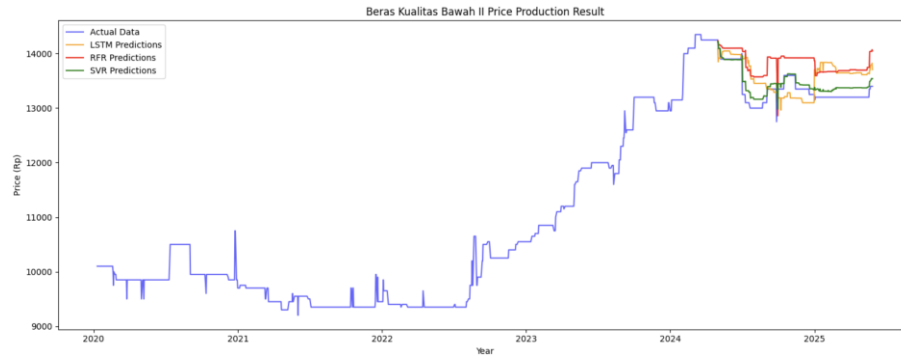
Visualisasi hasil prediksi dari ketiga model untuk setiap kategori kualitas beras bawah I ditampilkan pada gambar 2 berikut, disajikan dalam bentuk grafik yang membandingkan antara nilai aktual dengan hasil prediksi dari setiap model pada kategori kualitas beras bawah I.



Gambar 2. Visualisasi Hasil Prediksi Beras Kualitas Bawah I

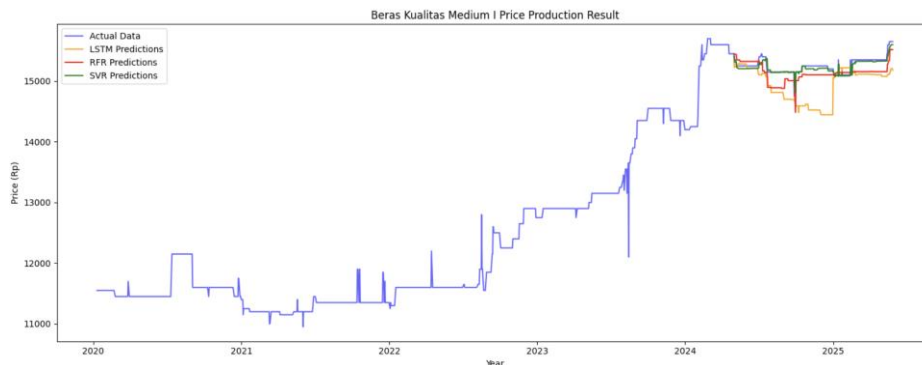
Gambar 2 memperlihatkan prediksi harga beras kualitas Bawah I dibandingkan data aktual. Ketiga model (LSTM, Random Forest, SVR) konsisten menangkap tren harga. LSTM cenderung memberikan proyeksi lebih tinggi pada awal 2025, sedangkan Random Forest dan SVR lebih stabil dan mendekati nilai tengah data aktual.

Untuk menampilkan hasil proyeksi, visualisasi harga dari beras jenis kualitas bawah II ditunjukkan pada Gambar 3.



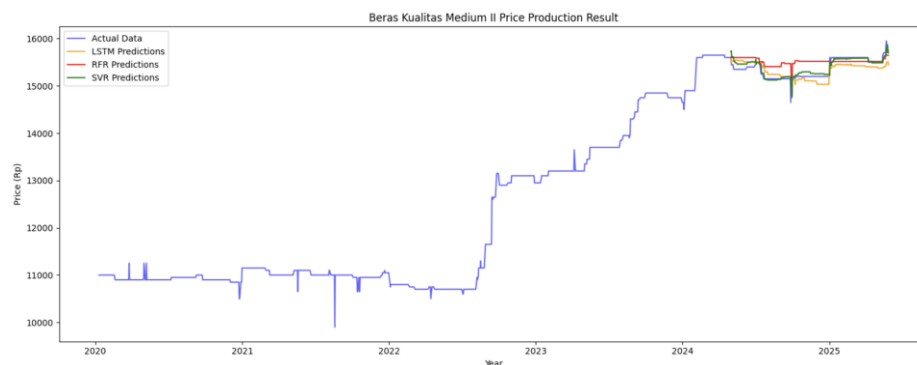
Gambar 3. Visualisasi Hasil Prediksi Beras Kualitas Bawah II

Gambar 3 memperlihatkan hasil visualisasi prediksi harga dari beras dengan kualitas bawah II. Grafik tersebut menunjukkan penggunaan model *Random Forest* dalam memprediksi harga terlihat cenderung menampilkan estimasi perubahan harga yang lebih tinggi, kemudian penggunaan model LSTM menghasilkan prediksi yang lebih seimbang, sedangkan model SVR menunjukkan pola yang paling stabil serta mendekati dengan data sebenarnya. Untuk menggambarkan kinerja model dalam memprediksi harga beras kualitas medium I, hasil visualisasi disajikan pada Gambar 4 berikut.



Gambar 4. Visualisasi Hasil Prediksi Beras Kualitas Medium I

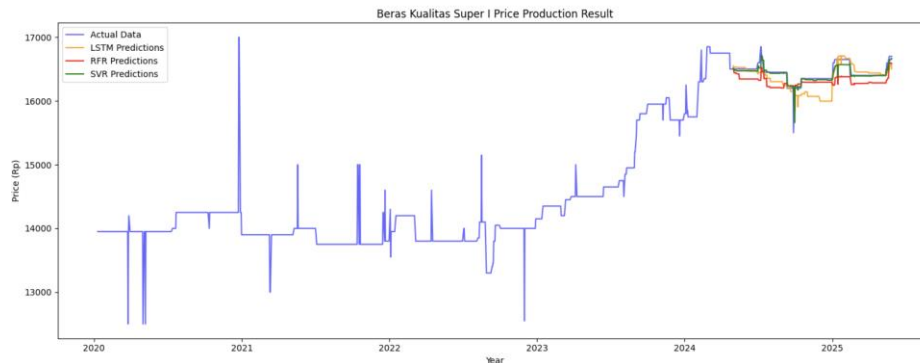
Gambar 4 menunjukkan hasil visualisasi prediksi harga beras kualitas medium I dalam bentuk grafik dengan memperlihatkan data actual yang mengalami kenaikan tajam sejak tahun 2022 hingga awal 2024, kemudian diikuti dengan prediksi yang relatif stabil pada periode 2024 hingga 2025. Terlihat bahwa model *Random Forest* menghasilkan estimasi harga yang cenderung lebih tinggi dengan variasi fluktuatif, model LSTM memberikan proyeksi yang lebih konservatif, sedangkan model SVR tampak paling stabil dengan hasil prediksi yang mendekati nilai tengah dari data sebenarnya. Sebagai gambaran performa dan kinerja model, prediksi harga beras kualitas Medium II dari ketiga model ditampilkan pada Gambar 5.



Gambar 5. Visualisasi Hasil Prediksi Beras Kualitas Medium II

Gambar 5 menunjukkan prediksi harga beras kualitas Medium II dibandingkan data aktual. Data aktual meningkat signifikan sejak akhir 2022 dan stabil pada 2024–2025. SVR paling mendekati data aktual dengan pola konsisten, Random Forest sedikit lebih tinggi dengan fluktuasi, sedangkan LSTM cenderung seimbang meski beberapa titik sedikit lebih rendah dibandingkan dua model lainnya.

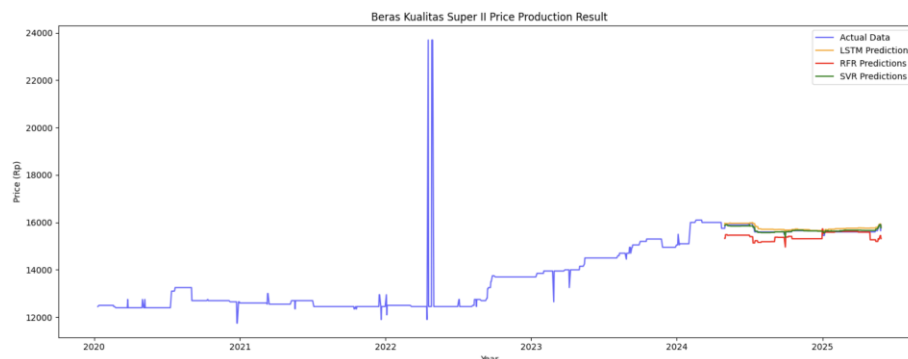
Visualisasi hasil prediksi harga beras kualitas Super I disajikan pada Gambar 6 berikut, untuk menggambarkan kemampuan model dalam mengikuti tren perubahan harga.



Gambar 6. Visualisasi Hasil Prediksi Beras Kualitas Super I

Gambar 6 menampilkan grafik harga beras kualitas super I memperlihatkan fluktuasi dengan kenaikan cukup tajam pada tahun 2021, kemudian cenderung stabil pada rentang harga Rp15.000 sampai Rp16.000 hingga 2025. Model LSTM menghasilkan prediksi dengan pola yang cukup berbeda dan cenderung lebih tajam, sedangkan model *Random Forest* mengikuti tren aktual tetapi masih menunjukkan deviasi di beberapa titik. Sementara itu, SVR memberikan hasil prediksi yang lebih konsisten dan seimbang, sehingga lebih mendekati pergerakan harga aktual dibandingkan dua model lainnya.

Perbandingan hasil prediksi harga beras kualitas Super II diilustrasikan pada Gambar 7 sebagai perbedaan kinerja model.



Gambar 7. Visualisasi Hasil Prediksi Beras Kualitas Super II

Gambar 7 memperlihatkan prediksi harga beras kualitas Super II dibandingkan data aktual. Data aktual menunjukkan fluktuasi tajam pada 2022, kemudian meningkat dan stabil hingga 2025. SVR dan LSTM menunjukkan prediksi lebih konsisten dan dekat dengan data aktual, sedangkan Random Forest sedikit lebih rendah di beberapa titik, namun tetap mengikuti tren kenaikan harga. Secara keseluruhan, ketiga model mampu menangkap kecenderungan harga, dengan SVR dan LSTM lebih mencerminkan pergerakan data aktual.

3.6 Analisis Grafik dan Perbandingan Kinerja Model

Hasil visualisasi dan evaluasi kuantitatif menunjukkan bahwa model SVR memiliki kinerja paling konsisten dalam memprediksi harga beras di seluruh kategori, dengan nilai MAPE, MAE, dan RMSE yang paling rendah dibandingkan LSTM dan Random Forest. Dapat dilihat pada grafik prediksi SVR terlihat paling konsisten mengikuti pola data aktual atau data sebenarnya pada seluruh kategori beras. Dapat dilihat pada kategori beras kualitas bawah I, model SVR menampilkan hasil yang konsisten yaitu MAPE 0.00 persen, MAE 59.39, dan RMSE sebesar 77.60, lebih unggul dibandingkan model LSTM, dan *Random Forest*. Model LSTM mampu menangkap tren jangka panjang, sedangkan Random Forest menunjukkan stabilitas yang baik namun kurang akurat pada data dengan fluktuasi harga tinggi. Selanjutnya, hasil uji *paired t-test* memperkuat temuan tersebut, di mana SVR menunjukkan perbedaan yang signifikan dibandingkan LSTM ($p = 0.0161$) dan Random Forest ($p = 0.0175$), sementara perbedaan antara LSTM dan Random Forest tidak signifikan ($p = 0.6736$). Dengan demikian, model SVR terbukti unggul secara statistik dan menjadi model paling efektif untuk peramalan harga beras di Provinsi Sumatera Selatan periode Januari 2020 hingga Mei 2025.

4. KESIMPULAN

Penelitian ini berhasil menerapkan metode data mining untuk memprediksi harga beras di wilayah Provinsi Sumatera Selatan, dengan menerapkan tiga model *machine learning*, yakni LSTM, *Random Forest*, dan SVR. Berdasarkan hasil analisis dan uji statistik, model SVR menunjukkan kinerja paling unggul dan konsisten dalam mengikuti pola data aktual, sehingga dinilai sebagai algoritma terbaik dalam peramalan harga beras. Hal ini juga diperkuat dengan visualisasi hasil prediksi yang menunjukkan bahwa grafik model SVR paling mendekati data aktual. Penggunaan model LSTM yang dikenal memiliki keunggulan dalam mengenali pola dalam kurun waktu yang lama dengan baik, namun kali ini cenderung menghasilkan tingkat kesalahan yang lebih tinggi pada jenis beras dengan perubahan harga yang cukup tajam. Temuan ini memberikan wawasan penting dalam konteks *knowledge discovery*, yaitu pemilihan model prediksi yang tepat sangat bergantung pada karakteristik data historis dan stabilitas harga. Penelitian ini menyimpulkan bahwa penerapan *knowledge discovery*, menggunakan metode *Knowledge Discovery in Databases* (KDD) berbasis *machine learning*, dengan menggunakan model SVR, memberikan peluang signifikan untuk diterapkan dalam peramalan harga beras sebagai salah satu kebutuhan pokok, sehingga dapat mendukung proses pengambilan keputusan dalam pengelolaan harga beras. Model ini juga dapat dimanfaatkan oleh pemerintah daerah, pedagang, maupun pembuat kebijakan sebagai alat bantu dalam menjaga stabilitas harga pangan dan ketersediaan pasokan bisa dikelola dengan baik di masa mendatang.

REFERENCES

- [1] M. I. Syairozi, "ANALISIS KEMISKINAN DI SEKTOR PERTANIAN (Studi Kasus Komoditas Padi di Kabupaten Malang)," *Media Ekonomi*, vol. 28, no. 2, pp. 113–128, May 2021, doi: 10.25105/me.v28i2.7169.
- [2] Solikhun and T. Yunita, "Quantum Perceptron: A New Approach for Predicting Rice Prices at the Indonesian Wholesale Trade Level," *Jurnal RESTI*, vol. 8, no. 4, pp. 479–485, Aug. 2024, doi: 10.29207/resti.v8i4.5869.
- [3] A. Muñoz-Villamizar, M. Piatti, C. Mejía-Argueta, L. F. Pirabe, J. Namdar, and J. F. Gomez, "Navigating retail inflation in Brazil: A machine learning and web scraping approach to the basic food basket," *Journal of Retailing and Consumer Services*, vol. 79, Jul. 2024, doi: 10.1016/j.jretconser.2024.103875.
- [4] S. Asrin, T. Andita Putri, and A. D. Utami, "Transmisi Harga Beras di Indonesia Pada Masa Pandemi Covid-19," *Jurnal Agribisnis Indonesia*, vol. 10, no. 1, pp. 159–168, Jun. 2022, doi: 10.29244/jai.2022.10.1.159-168.
- [5] R. Ikhwani and S. Suharyo, "Resiliensi Petani Wanita Kepala Keluarga Menghadapi Krisis Pangan," *Forum Penelitian Agro Ekonomi*, vol. 41, no. 1, Jul. 2023.
- [6] R. Hidayat and I. Wibisonya, "Rice Price Prediction with Long Short-Term Memory (LSTM) Neural Network," *Jurnal RESTI*, vol. 8, no. 5, pp. 658–664, Oct. 2024, doi: 10.29207/resti.v8i5.6041.
- [7] R. Nurfalah, Dwiza Riana, and Anton, "Identifikasi Citra Beras Menggunakan Algoritma Multi-SVM Dan Neural Network Pada Segmentasi K-Means," *Jurnal RESTI*, vol. 5, no. 1, pp. 55–62, Feb. 2021, doi: 10.29207/resti.v5i1.2721.
- [8] A. Wijaya and S. O. Kunang, "Implementasi Metode Least Square Pada Sistem Forecasting Harga Bahan Pokok Di Unit Pasar Tradisional Kota Palembang," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 1, pp. 291–301, Feb. 2023, doi: 10.29100/jupi.v8i1.3442.
- [9] A. A. Suyono, K. Kusrini, and M. R. Arief, "Prediksi Indeks Harga Konsumen Komoditas Makanan di Kota Surabaya menggunakan Support Vector Regression," *METIK JURNAL*, vol. 6, no. 1, pp. 45–51, Jul. 2022, doi: 10.47002/metik.v6i1.339.
- [10] I. Mardianto, M. Ichsan Gunawan, D. Sugiarto, and A. Rochman, "Perbandingan Peramalan Harga Beras Menggunakan Metode ARIMA pada Amazon Forecast dan Sagemaker," *Jurnal RESTI*, vol. 4, no. 3, pp. 537–543, Jun. 2020, doi: 10.29207/resti.v4i3.1902.
- [11] L. Domingo, M. Grande, F. Borondo, and J. Borondo, "Anticipating food price crises by reservoir computing," *Chaos Solitons Fractals*, vol. 174, Sep. 2023, doi: 10.1016/j.chaos.2023.113854.
- [12] M. Nilashi *et al.*, "Knowledge discovery for course choice decision in Massive Open Online Courses using machine learning approaches," *Expert Syst Appl*, vol. 199, Aug. 2022, doi: 10.1016/j.eswa.2022.117092.
- [13] H. Azis, P. Purnawansyah, N. Nirwana, and F. A. Dwiyanto, "The Support Vector Regression Method Performance Analysis in Predicting National Staple Commodity Prices," *ILKOM Jurnal Ilmiah*, vol. 15, no. 2, pp. 390–397, Aug. 2023, doi: 10.33096/ilkom.v15i2.1686.390-397.
- [14] R. Febiyane, Y. H. Chrisnanto, and G. Abdillah, "Implementasi Metode Recurrent Neural Network Untuk Prediksi Kejang Pada Penderita Epilepsi Berdasarkan Data Electroencephalogram," *JURIKOM (Jurnal Riset Komputer)*, vol. 12, no. 3, pp. 394–402, Jun. 2025, doi: 10.30865/jurikom.v12i3.8656.
- [15] N. A. Sofiah, K. D. Tania, A. Meiriza, and A. Wedhasmara, "A Comparative Assessment SARIMA and LSTM Models for the Gurugram Air Quality Index's Knowledge Discovery," *International Conference on Electrical Engineering and Computer Science (ICECOS)*, Sep. 2024, doi: 10.1109/ICECOS63900.2024.10791243.
- [16] S. Sen, D. Sugiarto, and A. Rochman, "Komparasi Metode Multilayer Perceptron (MLP) dan Long Short Term Memory (LSTM) dalam Peramalan Harga Beras," *ULTIMATICS*, vol. XII, no. 1, p. 35, 2020, doi: <https://doi.org/10.31937/ti.v12i1.1572>.
- [17] T. Prasetyo, R. A. Putri, D. Ramadhani, Y. Angraini, and K. A. Notodiputro, "Perbandingan Kinerja Metode Arima, Multi-Layer Perceptron, dan Random Forest dalam Peramalan Harga Logam Mulia Berjangka yang Mengandung Pencilan," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 265–274, Apr. 2024, doi: 10.25126/jtiik.20241127392.
- [18] X. Shu and Y. Ye, "Knowledge Discovery: Methods from data mining and machine learning," *Soc Sci Res*, vol. 110, Feb. 2023, doi: 10.1016/j.ssresearch.2022.102817.

- [19] V. Novalia, K. Ditha Tania, A. Meiriza, and A. Wedhasmara, "Knowledge Discovery of Application Review Using Word Embedding's Comparison with CNN-LSTM Model on Sentiment Analysis," *2024 International Conference on Electrical Engineering and Computer Science (ICECOS), IEEE*, Sep. 2024, doi: 10.1109/ICECOS63900.2024.10791113.
- [20] Ainurrochman, D. P. Adi, and A. B. Gumelar, "Deteksi Emosi Wicara pada Media On-Demand menggunakan SVM dan LSTM," *Jurnal RESTI*, vol. 4, no. 5, Oct. 2020, doi: 10.29207/resti.v4i5.2073.
- [21] A. Sujjada and F. Sembiring, "Prediksi Harga Bitcoin Menggunakan Algoritma Long ShortTerm Memory," *Jurnal INOVTEK Polbeng Seri Informatika*, vol. 9, no. 1, p. 2024, Nov. 2024, doi: <https://doi.org/10.35314/isi.v9i1.4247>.
- [22] A. Firmansyah, M. F. Syahidin, and Y. S. Triana, "Prediksi Kebakaran Hutan Berdasarkan Titik Panas dan Iklim Menggunakan Algoritma Random Forest," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 10, no. 2, pp. 145–155, Sep. 2024, doi: 10.25077/TEKNOSI.v10i2.2024.145-155.
- [23] A. Gumilang and S. Agustin, "Deteksi Kepribadian Melalui Margin Pada Tulisan Tangan Menggunakan Random Forest," *Jurnal INOVTEK Polbeng Seri Informatika*, vol. 9, no. 1, p. 2024, Nov. 2024, doi: <https://doi.org/10.35314/isi.v9i1.4176>.
- [24] L. N. Hapsari and N. Rokhman, "Anomaly Detection of Hospital Claim Using Support Vector Regression," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 18, no. 1, p. 1, Jan. 2024, doi: 10.22146/ijccs.91857.
- [25] Z. Liao, S. Dai, and T. Kuosmanen, "Convex support vector regression," *Eur J Oper Res*, vol. 313, no. 3, pp. 858–870, Mar. 2024, doi: 10.1016/j.ejor.2023.05.009.
- [26] D. Hendri, I. Permana, F. Nur Salisah, M. Afdal, M. Megawati, and E. Saputra, "Perbandingan Performa Algoritma SVR, LSTM, dan SARIMA dalam Peramalan Produksi Kelapa Sawit," *Technology and Science (BITS)*, vol. 7, no. 1, 2025, doi: 10.47065/bits.v7i1.7170.
- [27] I. Maulita and A. M. Wahid, "Prediksi Magnitudo Gempa Menggunakan Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron Berdasarkan Data Kedalaman dan Geolokasi," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 5, pp. 221–232, Dec. 2024, doi: 10.52436/1.jpti.470.
- [28] Z. A. Frakusya, R. H. Virgianto, and M. E. Yuggotomo, "Estimasi Konsentrasi PM10 Menggunakan Support Vector Regression," *JRST (Jurnal Riset Sains dan Teknologi)*, vol. 6, no. 1, p. 1, Nov. 2022, doi: 10.30595/jrst.v6i1.8977.