



Peramalan Produksi Padi Menggunakan *Extreme Learning Machine* (ELM) Dengan Deteksi *Outlier*

Isnaini Septyana Wanda^{1*}, Fifi Ayu Mufarroha², Achmad Jauhari³, Devie Rosa Anamisa⁴

^{1,2,3,4} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo Madura, Indonesia

Info Artikel

Riwayat Artikel:

Diterima : **16 November 2024**

Direvisi : **25 November 2024**

Diterbitkan : **10 Juni 2025**

Kata Kunci:

Produksi Padi,
Bangkalan,
Peramalan,
Extreme Learning Machine,
Deteksi Outlier,

ABSTRAK

Peramalan produksi padi yang akurat sangat penting dalam mendukung perencanaan pertanian yang efisien. Penelitian ini mengusulkan kombinasi metode *Extreme Learning Machine* (ELM) dengan teknik *deteksi outlier* untuk meningkatkan akurasi peramalan produksi padi. *Deteksi outlier* digunakan untuk mengidentifikasi dan mengatasi anomali dalam data historis dari Dinas Pertanian Kabupaten Bangkalan (1997-2022). Setelah proses *deteksi outlier*, data digunakan untuk melatih model ELM guna memprediksi produksi berdasarkan luas panen. Hasil penelitian menunjukkan bahwa pendekatan ini secara signifikan meningkatkan akurasi dibandingkan dengan model ELM tanpa *deteksi outlier*.

Keywords:

Rice Production,
Bangkalan,
Forecasting,
Extreme Learning Machine,
Outlier Detection,

ABSTRACT

Accurate rice production forecasting is essential for supporting efficient agricultural planning. This study proposes combining the *Extreme Learning Machine* (ELM) method with an outlier detection technique to enhance forecasting accuracy. Outlier detection is applied to identify and manage anomalies within historical data from the Agriculture Office of Bangkalan Regency (1997-2022). After outlier detection, the refined data are used to train the ELM model to predict rice production based on harvested area. The results indicate that this approach significantly improves accuracy compared to the ELM model without outlier detection.

Penulis Korespondensi:

Isnaini Septyana Wanda,
Program Studi Teknik Informatika,
Universitas Trunojoyo Madura
Email: isnainiswandaaa@gmail.com

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



1. PENDAHULUAN

Indonesia dikenal sebagai salah satu negara penghasil beras terbesar di dunia, bersanding dengan Tiongkok dan India. Meskipun demikian, Indonesia masih harus mengimpor beras untuk mencukupi kebutuhan dalam negeri. Tingginya tingkat konsumsi beras di Indonesia, yang mencapai hampir 120 kg per kapita per tahun yang artinya dua kali lipat dari rata-rata konsumsi global menjadi salah satu penyebab utama tingginya permintaan terhadap komoditas ini. Selain itu, pengurangan lahan pertanian akibat alih fungsi menjadi kawasan industri, perumahan, dan perdagangan semakin mempersempit area yang tersedia untuk produksi

pangan [1]. Faktor-faktor ini menunjukkan tantangan besar dalam mencapai swasembada pangan, khususnya dalam hal produksi beras, yang meskipun meningkat, tidak cukup untuk memenuhi konsumsi yang terus berkembang.

Padi, sebagai komoditas utama Indonesia, berperan penting dalam mendukung ketahanan pangan nasional [2]. Namun, sektor pertanian Indonesia, terutama produksi padi, menghadapi berbagai tantangan, termasuk fluktuasi hasil panen yang sering dipengaruhi oleh perubahan iklim, gangguan organisme pengganggu tanaman (OPT), dan alih fungsi lahan. Kabupaten Bangkalan di Provinsi Jawa Timur, meskipun merupakan salah satu daerah penghasil padi utama di Pulau Madura, juga menghadapi masalah serupa dengan penurunan produksi padi yang signifikan dalam beberapa tahun terakhir [3]. Fenomena ini memerlukan perhatian lebih untuk meningkatkan efisiensi produksi dan memprediksi hasil panen secara lebih akurat.

Dalam memprediksi hasil panen, metode tradisional seperti ARIMA (*Autoregressive Integrated Moving Average*) sering dimanfaatkan untuk melakukan pemodelan data deret waktu. Meskipun ARIMA sederhana dan efektif untuk data yang stasioner, metode ini memiliki keterbatasan dalam menangani pola non-linear yang kompleks dan data yang mengalami fluktuasi yang tidak terduga. Penelitian sebelumnya menunjukkan bahwa ARIMA cenderung kesulitan dalam menangani data yang tidak memenuhi asumsi stasionaritas atau dalam mengatasi perubahan yang tidak terduga [4]. Oleh karena itu, metode statistik tradisional ini sering kali kurang optimal dalam menangani data pertanian yang sangat dipengaruhi oleh faktor lingkungan dan perubahan iklim yang tidak terduga.

Sebagai alternatif, *Machine Learning* (ML), khususnya Artificial Neural Networks (ANN), telah menunjukkan kinerja yang lebih baik dalam menangani data non-linear dan kompleks. Penelitian sebelumnya mengungkapkan bahwa metode Machine Learning, terutama Deep Learning, secara signifikan lebih unggul dibandingkan dengan metode statistik tradisional seperti ARIMA dalam hal akurasi prediksi untuk data deret waktu yang kompleks dan dinamis [5]. Extreme Learning Machine (ELM), sebagai salah satu jenis ANN, terbukti lebih efisien dalam menangani data yang mengandung hubungan non-linear dan tidak terstruktur. ELM juga memiliki keuntungan dalam hal waktu pelatihan yang lebih cepat dibandingkan dengan model ANN lainnya [6]. Pendekatan ini dapat diadaptasi untuk sektor pertanian, khususnya dalam peramalan hasil panen padi yang lebih akurat dan efisien.

Selain itu, dalam aplikasi peramalan data pertanian, keberadaan outliers atau data yang menyimpang juga menjadi masalah yang sering mengganggu akurasi model. Outlier dapat menyebabkan model yang tidak akurat dengan menghasilkan estimasi parameter yang bias dan mengarah pada misspesifikasi model. Oleh karena itu, deteksi outlier yang tepat menjadi langkah penting dalam meningkatkan kualitas model prediksi, terutama pada data deret waktu. *Outlier* yang tidak terdeteksi dapat merusak hasil prediksi dan mengurangi efektivitas model dalam menangkap pola data yang relevan. Deteksi outlier yang efektif dapat membantu menyaring data yang tidak relevan atau tidak wajar sebelum digunakan dalam pelatihan model, yang pada gilirannya meningkatkan akurasi dan stabilitas model [7]. Dengan memanfaatkan pendekatan ini, penelitian ini bertujuan untuk mengembangkan model prediksi hasil panen padi di Kabupaten Bangkalan dengan mengintegrasikan ELM dan teknik deteksi outlier, yang diharapkan dapat membantu meningkatkan ketahanan pangan di Indonesia dan mengurangi ketergantungan pada impor beras.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang dapat dilihat pada Gambar 1. Penelitian ini dimulai dengan studi literatur untuk memahami konsep dan metode yang relevan dengan peramalan produksi padi dan penerapan metode ELM. ELM merupakan salah satu metode pembelajaran mesin yang menggunakan pendekatan jaringan saraf tiruan (Artificial Neural Network) dan dirancang untuk menyelesaikan masalah regresi dan klasifikasi dengan efisiensi tinggi. Salah satu keunggulan utama ELM dibandingkan dengan jaringan saraf tradisional adalah bobot di lapisan tersembunyi yang diinisialisasi secara acak dan tidak diperbarui selama proses pelatihan, yang secara signifikan mengurangi waktu pelatihan dan meningkatkan kecepatan komputasi [8]. Model ELM terdiri dari tiga komponen utama, yaitu lapisan input, lapisan tersembunyi, dan lapisan output.



Gambar 1. Tahapan Penelitian Prediksi Produksi Padi Menggunakan Model ELM

Pada lapisan tersembunyi, fungsi aktivasi g digunakan untuk menghasilkan output, yang dirumuskan sebagai berikut:

$$H = g(W_x + b) \quad (1)$$

di mana:

- H adalah output dari lapisan tersembunyi.
- W adalah bobot acak dari lapisan input ke lapisan tersembunyi.
- x adalah input data.
- b adalah bias.
- g adalah fungsi aktivasi, yaitu fungsi radial basis seperti, fungsi Gaussian.

Kemudian, output akhir $\hat{y}$ dihitung dengan:

$$\hat{y} = H \beta \quad (2)$$

di mana β adalah bobot dari lapisan tersembunyi ke lapisan output yang dihitung dengan metode Least Squares.

$$\beta = (H^T H)^{-1} H^T y \quad (3)$$

Pengumpulan data dilakukan dengan mengumpulkan data historis terkait luas panen padi dari Dinas Pertanian Tanaman Pangan, Hortikultura, dan Perkebunan Kabupaten Bangkalan, yang mencakup periode dari tahun 1997 hingga 2022. Data yang digunakan dalam penelitian ini mencakup variabel-variabel seperti luas

panen dan produksi padi. Gambar 2 berikut ini menampilkan sampel data yang digunakan untuk pelatihan model ELM.

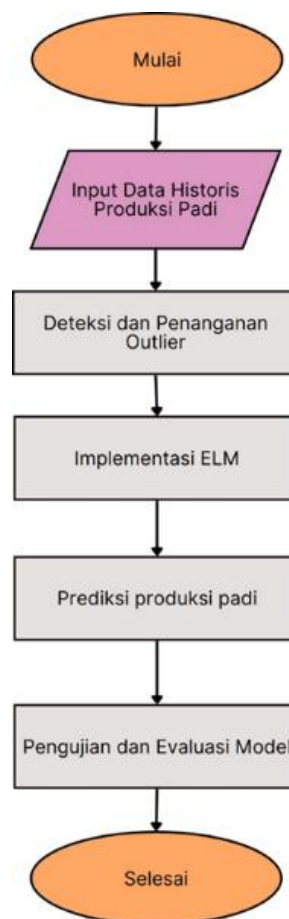
	tahun	luas_panen	produksi	produktivitas
0	1997	39454	1909176.2	48.39
1	1998	41725	1787621.7	42.84
2	1999	40476	1785478.9	44.11
3	2000	39168	1758632.5	44.90
4	2001	39845	1675078.1	42.04

Gambar 2. Sampel Dataset Produksi Padi

Setelah data terkumpul, tahap berikutnya adalah pra-pemrosesan data. Pada tahap ini, data yang telah dikumpulkan diperiksa untuk mengidentifikasi dan menangani outlier yang dapat mempengaruhi kualitas model prediksi. Deteksi dan penanganan outlier dilakukan agar data yang digunakan dalam pelatihan model ELM bebas dari anomali yang dapat merusak hasil analisis [9]. Proses deteksi outlier ini penting karena outlier dapat mengganggu analisis statistik dan mempengaruhi akurasi model prediksi [10]. Metode yang digunakan untuk mendeteksi outlier adalah Local Outlier Factor (LOF), yang merupakan teknik berbasis k-Nearest Neighbors (k-NN) untuk mengidentifikasi data dengan kepadatan lokal yang jauh lebih rendah dibandingkan dengan titik data lainnya [11]. Setelah outlier terdeteksi, data tersebut disesuaikan atau dibersihkan untuk memastikan model bekerja dengan data yang lebih bersih dan konsisten.

Untuk mendeteksi outlier, metode LOF digunakan. LOF menghitung nilai kepadatan lokal dari setiap titik data dalam hubungannya dengan titik-titik di sekitarnya. Rumus LOF untuk titik data p terhadap tetangga-tetangganya dapat ditulis sebagai:

$$LOF(p) = \frac{\sum_{o \in N_k(p)} \frac{L_k(o)}{L_k(p)}}{|N_k(p)|} \quad (4)$$



Gambar 3. Metode Penelitian Prediksi Produksi Padi Menggunakan Model ELM dengan Deteksi Outlier

Proses normalisasi data juga dilakukan untuk memastikan bahwa data berada dalam skala yang seragam. Setelah itu, data dipisahkan menjadi dua kelompok, yaitu data pelatihan dan data pengujian. Data pelatihan dimanfaatkan untuk membangun model ELM, sedangkan data pengujian digunakan untuk menilai kinerja model yang telah dibuat. Model ELM dikembangkan dengan menggunakan data pelatihan yang telah diproses. Model ini bertujuan untuk mempelajari pola yang ada pada data historis luas panen padi dan menghasilkan prediksi yang akurat untuk produksi padi di masa mendatang. ELM memiliki struktur yang terdiri dari tiga lapisan utama, yaitu lapisan input, lapisan tersembunyi, dan lapisan output, yang saling terhubung dengan bobot yang dihitung melalui metode Least Squares [8].

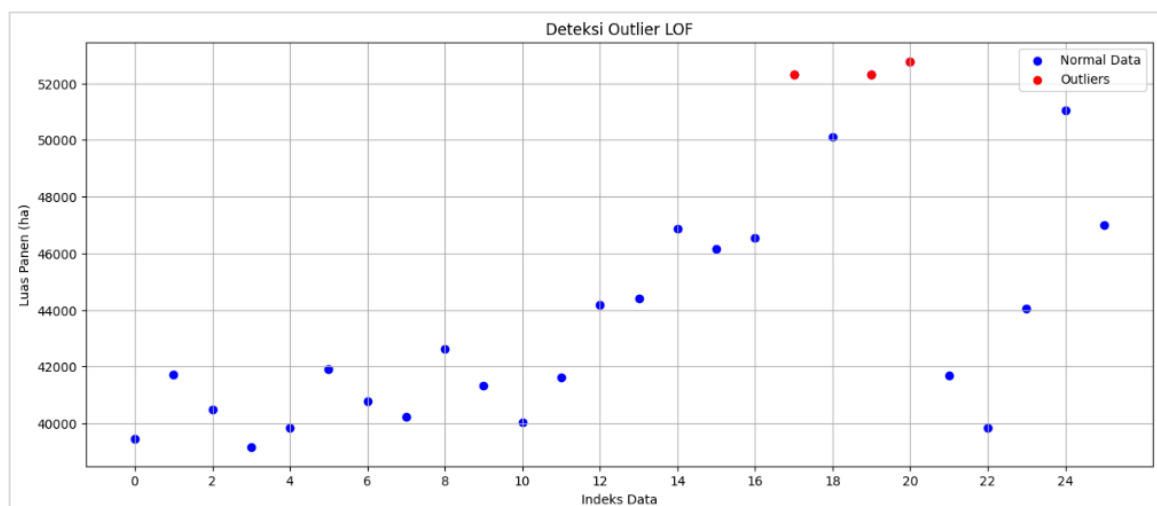
Pada tahap pengujian, model yang telah dilatih diuji menggunakan data pengujian untuk mengukur akurasi prediksi. Evaluasi dilakukan dengan menggunakan Mean Absolute Percentage Error (MAPE), yang digunakan untuk menilai sejauh mana model dapat memprediksi hasil yang sesuai dengan data aktual [12]. Hasil dari pengujian dan evaluasi model kemudian dianalisis untuk menilai efektivitas model ELM dalam memprediksi produksi padi. Berdasarkan analisis tersebut, kesimpulan ditarik mengenai keunggulan dan keterbatasan model yang digunakan, serta rekomendasi untuk penelitian lebih lanjut.

Metode yang digunakan dalam penelitian ini adalah penerapan ELM yang dilengkapi dengan deteksi outlier sebagai langkah pra-pemrosesan data. Deteksi outlier bertujuan untuk mengidentifikasi dan mengurangi pengaruh data yang menyimpang, sehingga model dapat bekerja lebih baik dengan data yang lebih bersih dan konsisten. Setelah data dibersihkan, model ELM dilatih untuk mempelajari pola dari data historis luas panen dan menghasilkan prediksi untuk periode yang akan datang. Diagram alur metode yang digunakan dapat dilihat pada Gambar 3.

3. HASIL DAN ANALISIS

3.1. Deteksi Outlier Menggunakan *Local Outlier Factor* (LOF)

Grafik di bawah ini menunjukkan hasil deteksi outlier pada data luas panen menggunakan metode Local Outlier Factor (LOF). Pada grafik ini, titik-titik yang diidentifikasi sebagai outlier ditandai dengan warna merah untuk menunjukkan data yang berbeda signifikan dari pola umum. Dengan menggunakan LOF, data yang berada dalam kepadatan lokal yang rendah atau signifikan berbeda dari data di sekitarnya dapat diidentifikasi sebagai outlier.



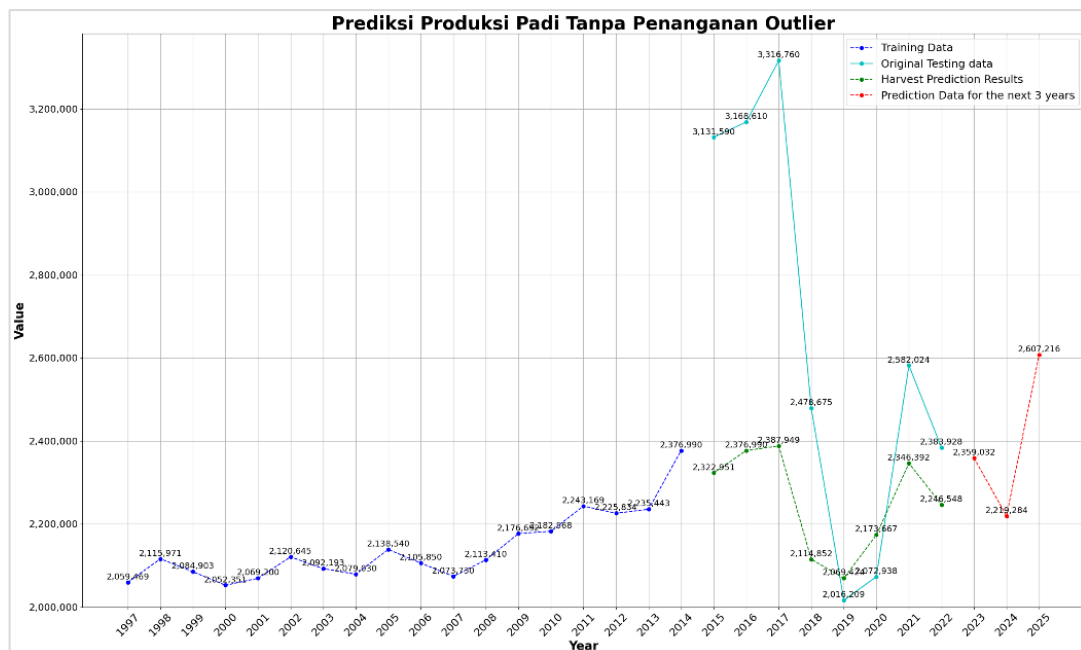
Gambar 4. Deteksi Outlier-Local Outlier Factor (LOF)

Deteksi outlier ini bertujuan untuk mengidentifikasi data yang mungkin akan merusak akurasi model prediksi produksi padi jika tidak ditangani terlebih dahulu. Metode ini digunakan karena efektif dalam mendeteksi outlier pada data yang memiliki distribusi lokal yang bervariasi.

3.2. Prediksi Produksi Padi Tanpa Penanganan Outlier

Grafik berikut ini menampilkan hasil prediksi produksi padi menggunakan model ELM tanpa adanya penanganan outlier. Data yang digunakan dalam model ini belum mengalami proses deteksi atau pembersihan

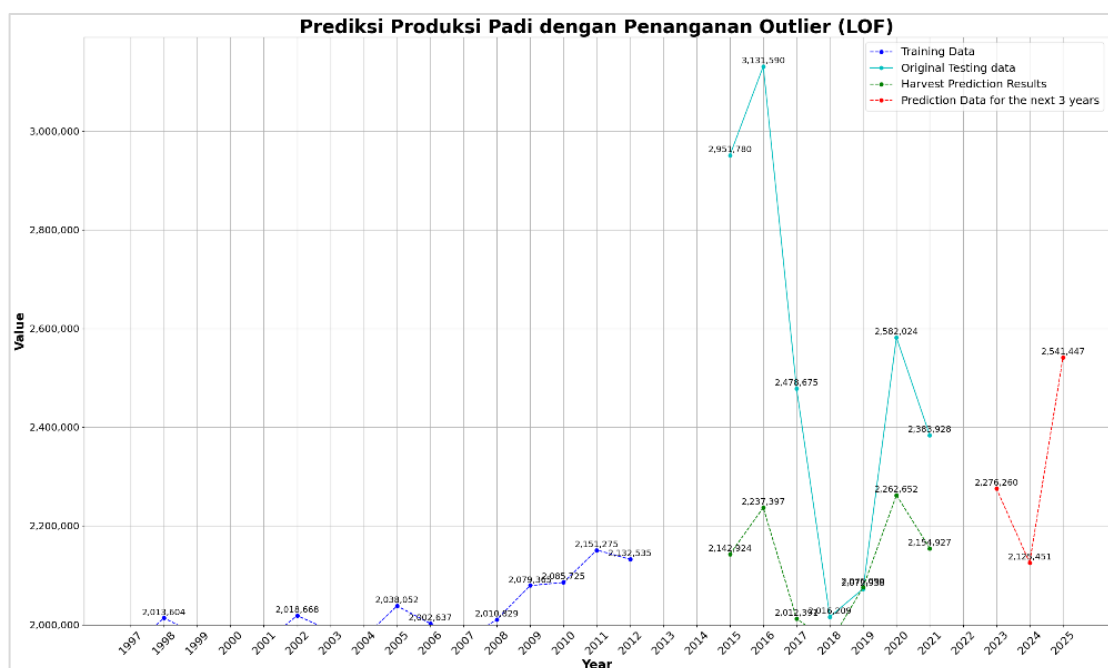
outlier. Grafik menunjukkan bahwa prediksi memiliki beberapa penyimpangan dari data aktual, yang mungkin disebabkan oleh keberadaan outlier dalam data mentah. Tanpa penanganan outlier, model ELM dapat terpengaruh oleh data yang tidak representatif, mengakibatkan akurasi prediksi yang lebih rendah.



Gambar 5. Prediksi Poduksi Padi Tanpa Penanganan Outlier

3.3. Prediksi Produksi Padi dengan Penanganan Outlier Menggunakan LOF

Grafik ini menunjukkan hasil prediksi produksi padi dengan model ELM yang telah dilatih pada data yang sudah melalui proses deteksi dan penanganan outlier menggunakan LOF. Setelah outlier teridentifikasi, data tersebut disesuaikan atau dibersihkan untuk memastikan model ELM menerima data yang lebih konsisten dan bersih.



Gambar 6. Prediksi Poduksi Padi Dengan Penanganan Outlier (LOF)

3.4. Analisis Hasil dan Pengaruh Penanganan Outlier

Dari hasil perbandingan grafik prediksi, terlihat bahwa penanganan outlier menggunakan LOF berdampak positif pada akurasi model ELM. Dengan membersihkan data dari anomali, model lebih mampu menangkap

pola dari data yang konsisten, menghasilkan prediksi yang lebih akurat. Hasil prediksi menunjukkan peningkatan akurasi, yang terlihat dari penurunan MAPE pada data uji, dari 14,4% menjadi 14,2%. Pada data latih, error menurun dari 14% menjadi 11%. Penurunan MAPE menunjukkan bahwa metode ini efektif, meskipun jumlah data yang digunakan relatif sedikit.

Tabel 1. Perbandingan MAPE Model ELM pada Data dengan dan Tanpa Deteksi Outlier

Kondisi Data	MAPE Data Latih	MAPE Data Uji
Tanpa Deteksi Outlier	15.31%	14.48%
Dengan Deteksi Outlier	11.39%	14.22%

4. KESIMPULAN

Penelitian ini bertujuan untuk meningkatkan akurasi prediksi produksi padi di Kabupaten Bangkalan dengan menggunakan model ELM yang dikombinasikan dengan deteksi outlier berbasis LOF sebagai langkah pra-pemrosesan. Berdasarkan hasil analisis, penerapan deteksi outlier berhasil menunjukkan peningkatan akurasi prediksi dibandingkan dengan model yang menggunakan data tanpa proses deteksi outlier. Peningkatan ini terlihat dari nilai MAPE yang lebih rendah setelah dilakukan proses deteksi dan penanganan outlier, yang membuktikan bahwa pembersihan data dari anomali penting dalam memastikan model bekerja dengan data yang bersih dan representatif.

Meskipun demikian, penelitian ini juga menemukan bahwa hasil pengaruh deteksi outlier dapat bervariasi tergantung pada jumlah data yang digunakan. Oleh karena itu, pengembangan lebih lanjut dengan menggunakan dataset yang lebih besar dan variabel yang lebih beragam dapat membantu memperkuat temuan penelitian ini. Prospek penelitian berikutnya juga dapat mempertimbangkan metode outlier deteksi lain atau kombinasi teknik pra-pemrosesan untuk meningkatkan performa model ELM. Dengan begitu, diharapkan model prediksi produksi padi ini dapat diaplikasikan secara lebih luas dan memberikan kontribusi dalam perencanaan produksi pangan secara lebih akurat dan andal.

REFERENSI

- [1] F. Musti Ariska and B. Qurniawan, "PERKEMBANGAN IMPOR BERAS DI INDONESIA," *Journal of Agriculture and Animal Science*, vol. 1, no. 1, pp. 27–34, May 2021, doi: 10.47637/agrimals.v1i1.342.
- [2] M. Sukmayanto, I. Listiana, and T. Hasanuddin, "Analisis Produksi dan Pendapatan Usahatani Padi di Kabupaten Lampung Tengah," *Jurnal Ekonomi Pertanian dan Agribisnis*, vol. 6, no. 2, p. 625, Apr. 2022, doi: 10.21776/ub.jepa.2022.006.02.26.
- [3] S. K. S. A. (KSA) & S. U.-S. I. A. S. F. (ASF) S. & C.-C. S. BPS, "Luas Panen, Produktivitas, dan Produksi Padi Menurut Kabupaten/Kota di Provinsi Jawa Timur, 2022." Accessed: Oct. 12, 2024. [Online]. Available: <https://jatim.bps.go.id/id/statistics-table/3/WmpaNk1YbGFjR0pOUjBKYWFIQIBSU3MwVHhOVWR6MDkjMw==/luas-panen--produktivitas--dan-produksi-padi-menurut-kabupaten-kota-di-provinsi-jawa-timur.html>
- [4] M. F. Rizvi, "ARIMA Model Time Series Forecasting," *Int J Res Appl Sci Eng Technol*, vol. 12, no. 5, pp. 3782–3785, May 2024, doi: 10.22214/ijraset.2024.62416.
- [5] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks," *Future Internet*, vol. 15, no. 8, p. 255, Jul. 2023, doi: 10.3390/fi15080255.
- [6] G. Singh, B. Tripathy, and J. Singh, "An Evaluation of Extreme Learning Machine Algorithm to Forecasting the Gold Price," in *7th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 523–529. doi: 10.1109/ICECA58529.2023.10395223.
- [7] V. Naidoo and S. Du, "A Deep Learning Method for the Detection and Compensation of Outlier Events in Stock Data," *Electronics (Basel)*, vol. 11, no. 21, p. 3465, Oct. 2022, doi: 10.3390/electronics11213465.

- [8] U. Markowska-Kaczmar and M. Kosturek, "Extreme learning machine versus classical feedforward network: Comparison from the usability perspective," Nov. 01, 2021, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-021-06402-y.
- [9] A. Boukerche, L. Zheng, and O. Alfandi, "Outlier Detection," *ACM Comput Surv*, vol. 53, no. 3, pp. 1–37, May 2021, doi: 10.1145/3381028.
- [10] M. Jurusan, M. Uin, M. Malik, and I. Malang, "PENDETEKSIAN OUTLIER PADA REGRESI NONLINIER DENGAN METODE STATISTIK LIKELIHOOD DISPLACEMENT (LD) Siti Tabi'atul Hasanah."
- [11] O. Alghushairy, R. Alsini, T. Soule, and X. Ma, "A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams," *Big Data and Cognitive Computing*, vol. 5, no. 1, p. 1, Dec. 2020, doi: 10.3390/bdcc5010001.
- [12] K. Chen *et al.*, "Prediction of Hydrogen Production from Solid Oxide Electrolytic Cells Based on ANN and SVM Machine Learning Methods," *Atmosphere (Basel)*, vol. 15, no. 11, p. 1344, Nov. 2024, doi: 10.3390/atmos15111344.