Akreditasi Sinta 5 KEMRISTEKDIKBUD No. 225/E/KPT/2022; E-ISSN 2723-7958

Sistem Prediksi Produksi Beras Menggunakan *Multiple Linear Regression* untuk Optimalisasi Ketahanan Pangan di Kabupaten Minahasa

Oliver Simon Hardianto Raharusun[1], Alfiansyah Hasibuan[2]

1,2 Program Studi Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado
Email: oliverraharusun@gmail.com

***Abstract*— Food security is a strategic issue requiring a data-driven approach, especially for Minahasa Regency, which relies heavily on the agricultural sector. Predicting rice production is essential to ensure adequate food availability amidst environmental and agricultural challenges. This study aims to develop a rice production prediction model using multiple linear regression by leveraging variables such as harvested area, rainfall, number of rainy days, average temperature, and air humidity. Historical data were used to train the multiple linear regression model. Evaluation was conducted using metrics such as R-squared (R^2) and Mean Absolute Percentage Error (MAPE) to measure prediction performance. Statistical analysis was utilized to identify the most influential variables on rice production. The model demonstrated good performance with an R^2 value of 0.667, indicating that 66.7% of the variability in rice production data could be explained by the model. A MAPE value of 10.61% reflects an acceptable level of prediction error. The results of the study revealed that harvested area is the most significant variable influencing production, followed by rainfall and the number of rainy days. This prediction system has significant potential to support data-driven food security planning. Policies such as stock distribution, land use planning, and climate risk mitigation can be optimized using this system. Further development is required to enhance prediction accuracy by incorporating additional variables.**

***Keywords*— Prediction System, Rice Production, Multiple Linear Regression, Food Security**

***Abstrak*— Ketahanan pangan merupakan isu strategis yang memerlukan pendekatan berbasis data, terutama bagi Kabupaten Minahasa yang bergantung pada sektor agraris. Prediksi produksi beras menjadi penting untuk memastikan ketersediaan pangan yang memadai di tengah tantangan lingkungan dan agraris. Penelitian ini bertujuan untuk mengembangkan model prediksi produksi beras menggunakan regresi linear berganda, dengan memanfaatkan variabel-variabel seperti luas panen, curah hujan, jumlah hari hujan, suhu rata-rata, dan kelembapan udara. Data historis digunakan untuk melatih model regresi linear berganda. Evaluasi dilakukan menggunakan metrik seperti *R-squared* (R^2) dan *Mean Absolute Percentage Error* (MAPE) untuk mengukur performa prediksi. Analisis statistik digunakan untuk mengidentifikasi variabel yang paling berpengaruh terhadap produksi beras. Model menunjukkan performa yang baik dengan nilai R^2 sebesar 0.667, yang mengindikasikan bahwa 66,7% variabilitas data produksi beras dapat dijelaskan oleh model. Nilai MAPE sebesar 10.61% menunjukkan tingkat kesalahan prediksi yang dapat diterima. Hasil penelitian menunjukkan bahwa luas panen merupakan variabel paling signifikan dalam memengaruhi produksi, diikuti oleh curah hujan dan jumlah hari hujan. Sistem prediksi ini memiliki potensi besar untuk digunakan dalam mendukung perencanaan ketahanan pangan berbasis data. Kebijakan seperti distribusi stok, perencanaan kebutuhan lahan, dan mitigasi risiko iklim dapat dioptimalkan dengan menggunakan sistem ini. Pengembangan lebih lanjut diperlukan untuk meningkatkan akurasi prediksi dengan menambahkan variabel lain.**

***Kata Kunci*— Sistem Prediksi, Produksi Beras, Regresi Linear Berganda, Ketahanan Pangan**

I. PENDAHULUAN

Ketahanan pangan merupakan salah satu isu strategis global yang terus mendapatkan perhatian besar, terutama di negara berkembang seperti Indonesia. Di Indonesia, sektor pertanian, khususnya produksi beras, memainkan peran vital dalam memenuhi kebutuhan pangan lebih dari 270 juta penduduknya. Beras sebagai makanan pokok utama, sangat mempengaruhi stabilitas sosial dan ekonomi. Namun, dalam beberapa tahun terakhir, sektor pertanian Indonesia mengalami tantangan besar, termasuk ketidakstabilan produksi beras yang disebabkan oleh faktor cuaca, bencana alam, dan perubahan pola konsumsi masyarakat.

Kabupaten Minahasa merupakan salah satu wilayah di Sulawesi Utara, Indonesia yang memiliki ketergantungan tinggi terhadap produksi beras sebagai kebutuhan pangan utama masyarakat. Dalam beberapa tahun terakhir, terjadi ketidakstabilan pada tingkat produksi dan distribusi beras di daerah ini. Kondisi ini berdampak pada kenaikan harga beras, yang kemudian memicu masalah ketersediaan pangan bagi masyarakat. Seperti yang dilaporkan oleh Ibu Sandra Karlos Arsjad, Kepala Bidang Dinas Pangan Kabupaten Minahasa dalam wawancaranya mengatakan bahwa fluktuasi harga beras menjadi salah satu indikator permasalahan stok dan produksi yang belum sepenuhnya stabil. Semakin tinggi stok yang tersedia, maka harga beras cenderung lebih stabil dan rendah, namun ketika stok

Akreditasi Sinta 5 KEMRISTEKDIKBUD No. 225/E/KPT/2022; E-ISSN 2723-7958

berkurang, harga beras meningkat secara signifikan. Permasalahan ini mengindikasikan adanya ketergantungan pada manajemen produksi dan distribusi yang lebih baik agar ketahanan pangan dapat terjaga.

Dalam upaya menjaga ketahanan pangan di daerah Minahasa, peneliti merancang dan membangun sistem prediksi produksi beras menggunakan algoritma *multiple linear regression. Multiple linear regression* dipilih karena keunggulannya dalam prediksi, terbukti dengan nilai *R-squared* yang tinggi dalam berbagai penelitan sebelumnya. Misalnya *multiple linear regression* mencapai nilai *R-squared* 0.99 atau 99% dalam prediksi hasil panen wortel [1] dan 1 atau 100% dalam memprediksi hasil panen kacang kedelai di Kecamatan Surade [2].

Implementasi sistem ini diharapkan mampu memberikan estimasi produksi beras yang akurat sehingga diharapkan mampu menjadi acuan dalam pengambilan keputusan strategis terkait ketahanan pangan. Sistem ini tidak hanya diharapkan dapat menjaga ketahanan pangan di Kabupaten Minahasa, tetapi juga diharapkan meningkatkan stabilitas harga beras.

II. TINJAUAN PUSTAKA

A. Sistem

Sistem, menurut Azhar Susanto, adalah kumpulan komponen fisik dan bukan fisik yang terhubung satu sama lain dan digabungkan untuk mencapai tujuan tertentu [3].

Menurut Hutahaean, sistem dapat didefinisikan sebagai suatu rangkaian kerja dari berbagai prosedur yang berhubungan satu sama lain dan bekerja sama untuk mencapai tujuan atau melakukan kegiatan tertentu [4].

Berdasarkan definisi di atas, sistem dapat didefinisikan sebagai kumpulan elemen atau komponen yang saling terhubung untuk mencapai tujuan tertentu.

B. Prediksi

Prediksi adalah proses sistematis untuk memperkirakan sesuatu yang paling mungkin terjadi di masa depan dengan menggunakan informasi masa lalu dan saat ini. Ini dilakukan untuk mengurangi kesalahan (selisih antara sesuatu yang terjadi dengan hasil perkiraan) [5]. Menurut Kamus Besar Bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi, meramal, atau memperkirakan. Peramalan, juga dikenal sebagai prediksi, adalah suatu prosedur untuk mengumpulkan informasi faktual tentang situasi sosial masa depan berdasarkan informasi tentang masalah kebijakan yang telah ada [6].

C. Produksi

Kata produksi merupakan kata serapan dari bahasa Inggris, yaitu *production* dalam kamus besar Bahasa Indonesia, kata produksi diartikan sebagai proses mengeluarkan hasil penghasilan. Produksi terjadi karena suatu proses atau aktivitas ekonomi yang menggunakan beberapa masukan (*input*). Oleh karena itu, aktivitas produksi dikombinasikan sebagai input dan menghasilkan output [7].

D. Padi

Padi (*Oryza sativa L.)* merupakan tanaman pangan yang sangat penting di dunia setelah gandum dan jagung. Padi merupakan tanaman pangan yang sangat penting karena beras masih digunakan sebagai makanan pokok bagi sebagian besar penduduk dunia terutama Asia sampai sekarang. Beras merupakan komoditas strategis di Indonesia karena beras mempunyai pengaruh yang besar terhadap kestabilan ekonomi dan politik. Saat ini, Indonesia masih sering menghadapi masalah pangan seperti adanya alih fungsi lahan pertanian menjadi kawasan industri dan pemukiman yang menyebabkan penurunan produktivitas beras [8].

E. Ketahanan Pangan

Ketahanan pangan adalah kondisi ketika semua orang, kapan saja, memiliki akses fisik, sosial, dan ekonomi terhadap pangan yang cukup, aman, dan bergizi untuk memenuhi kebutuhan diet mereka, serta mampu menjalani kehidupan yang sehat dan aktif. Ketahanan pangan di Indonesia sering kali terpengaruh oleh faktor produksi pertanian yang tidak stabil, termasuk produksi beras. Pemerintah Indonesia telah melakukan berbagai upaya untuk mencapai ketahanan pangan yang lebih baik, salah satunya dengan memastikan stabilitas produksi beras melalui kebijakan pertanian yang berbasis data [9]. Dengan demikian, ketahanan pangan yang kuat memerlukan sistem yang dapat memprediksi hasil produksi secara akurat untuk memastikan ketersediaan pangan yang berkelanjutan.

F. Streamlit

Streamlit adalah sebuah *framework open-source* yang membantu pengembang membangun aplikasi web yang dapat diakses menggunakan bahasa pemrograman Python. Dengan menggunakan *framework* ini, pengembang dapat membangun aplikasi dengan mudah menggunakan skrip Python yang sederhana tanpa perlu memiliki pengetahuan mendalam tentang pengembangan *web frontend*. Dengan menggunakan *Streamlit*, Anda dapat membuat antarmuka pengguna (UI) responsif dengan cepat yang menampilkan visualisasi data, tabel, grafik, dan elemen interaktif lainnya. Selain itu, fitur-fiturnya mempermudah penyebaran aplikasi sehingga pengguna lain dapat dengan mudah mengaksesnya secara online. Karena itu, pengembang *data science* dan machine learning telah memilih *Streamlit* untuk mempercepat pembuatan dan distribusi aplikasi berbasis web [10].

G. Multiple Linear Regression

Analisis regresi merupakan suatu model matematis yang dapat digunakan untuk mengetahui pola hubunga nantara dua atau lebih variabel [11]. Analisis regresi lebih akurat dalam melakukan analisis korelasi. Menggunakan analisis regresi, peramalanatau perkiraan nilai variabel terikat pada nilai variabel bebas lebih akurat. Karena hasil regresi merupakan nilai prediksi, maka nilai tersebut belum tentu tepat dengan nilai riilnya, semakin sedikit penyimpangan nilai prediksi terhadap nilai riilnya, maka persamaan regresi yang dihasilkan semakin tepat dengan kondisi riilnya. Analisis regresi merupakan suatu metode statistika yang

dipakai untuk mengetahui dan menentukan bentuk hubungan yang terjadi antara variabel-variabel, yang bertujuan untuk memerkirakan dan meramalkan nilai dari variabel lain yang telah diketahui. Ada dua jenis persamaan regresi linear, yaitu [12]:

1. Analisis regresi sederhana (*simple linear regression*)
2. Analisis regresi berganda (*multiple linear regression*)

Model persamaannya dapat digambarkan sebagai berikut:

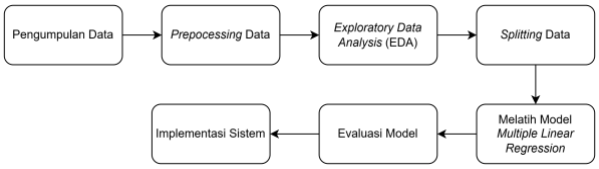
$$Y = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n \qquad (1)$$

Dimana:
Y : Variabel dependen (variabel terikat)
X : Variabel independen (variabel bebas)
a : Nilai konstanta
b : Nilai koefisien regresi
X_1 : Variabel independen pertama
X_2 : Variabel independen kedua
X_n : Variabel independen ke-n

III. METODOLOGI

Penelitian mengumpulkan dataset yang berkaitan dengan topik, yakni histori produksi beras dan variabel-variabel data lainnya. Langkah selanjutnya adalah *preprocessing* data untuk memeriksa dan penanganan terhadap duplikasi data, *missing values*, dan *outlier*. Selanjutnya, dilakukan EDA (*Exploratory Data Analysis)* pada dataset untuk memahami pola, distribusi, korelasi antar-variabel dan lain-lain.

Setelah *splitting* data yaitu membagi dataset menjadi data latih (*training*) dan data validasi (*testing*), langkah selanjutnya melakukan pemodelan data menggunakan metode *Multiple Linear Regression* atau Regresi Liniar Berganda. Setelah melatih model, dilakukan evaluasi terhadap kinerja model. Langkah selanjutnya implementasi model ke dalam sistem yang dibangun.



Gbr. 1 Metodologi Penelitian

IV. *HASIL DAN PEMBAHASAN*

A. *Pengumpulan Data*

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 19 entries, 0 to 18
Data columns (total 7 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Tahun                     19 non-null     int64
 1   Produksi Beras (ton)      19 non-null     float64
 2   Luas Panen (Ha)           19 non-null     float64
 3   Jumlah Curah Hujan (mm)   12 non-null     float64
 4   Jumlah Hari Hujan (Hari)  5 non-null      float64
 5   Rata-rata Suhu (Celcius)  5 non-null      float64
 6   Rata-rata Kelembapan Udara  4 non-null    float64
dtypes: float64(6), int64(1)
memory usage: 1.2 KB
```

Gbr. 2 Dataset Penelitian

Penelitian dimulai dari pengumpulan data (dapat dilihat pada Gambar 2) dengan cara wawancara. Selain itu, pengumpulan data juga dilakukan dengan pengambilan data dari situs web resmi Badan Pusat Statistik (BPS) Kabupaten Minahasa dan Provinsi Sulawesi Utara. Data yang dikumpulkan berupa informasi histori produksi beras, tahun, luas lahan panen, curah hujan, suhu, dan kelembapan udara di Kabupaten Minahasa dari tahun 2003-2023.

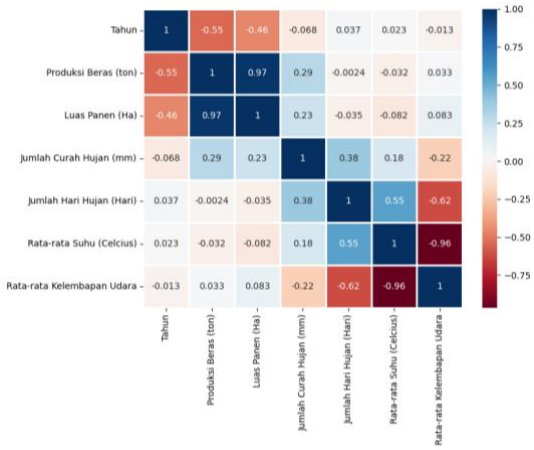
B. *Prepocessing Data*

Dalam tahap *prepocessing* data, langkah awal yang di lakukan adalah pengecekan terhadap duplikasi data. Kemudian, dilakukan pengecekan terhadap keberadaan nilai kosong di setiap variabel dalam dataset. Untuk menangani nilai yang kosong atau *missing value* dalam dataset, dilakukan metode imputasi yaitu mengisi nilai kosong dengan *mean* atau nilai rata-rata.

Tahapan selanjutnya adalah pengecekan terhadap *outlier* dalam dataset. *Outlier* adalah objek yang memiliki nilai ekstrim, baik terlalu rendah atau terlalu tinggi, sehingga sangat berbeda dari objek lainnya [13]. Penanganan terhadap nilai *outlier* dapat dilakukan dengan beberapa metode, di antaranya adalah menghapus, membiarkan, imputasi dan lain-lain. Dalam studi kasus prediksi produksi beras di Kabupaten Minahasa, metode penanganan nilai *outlier* yaitu dengan membiarkan (tidak diubah), yang ditinjau dengan berbagai pertimbangan.

C. *Exploratory Data Analysis (EDA)*

EDA atau *Exploratory Data Analysis* adalah salah satu tahapan dari ilmu data yang bertujuan untuk mengenal dan memahami kumpulan data [14]. Salah satu jenis dari analisi data eksplorasi adalah analisis korelasi. Analisis korelasi merupakan tahapan penting dalam pemahaman hubungan antara berbagai variabel dalam dalam kumpulan data. Pada gambar *correlation heatmap* (Gambar 3), nilai-nilai korelasi antar setiap pasang variabel ditampilkan dalam bentuk warna dengan rentang nilai dari -1 hingga 1. Korelasi positif dinyatakan dalam warna terang hingga biru gelap, sedangkan untuk korelasi negatif dinyatakan dalam warna merah terang hingga merah gelap.



Gbr. 3 *Corellation Heatmap*

Dalam dataset ini, beberapa variabel memiliki korelasi positif yang cukup kuat, ditandai dengan warna biru cerah

hingga biru gelap pada visualisasi korelasi. Salah satu contohnya adalah hubungan antara "Produksi Beras (ton)" dan "Luas Panen (Ha)", yang memiliki korelasi positif yang signifikan dengan nilai korelasi sebesar 0.97.

Hal ini menunjukkan bahwa terdapat hubungan positif yang sangat kuat dan hampir sempurna antara jumlah produksi beras dan luas panen. Artinya dapat diasumsikan bahwa ketika luas panen meningkat, jumlah produksi beras cenderung meningkat, dan sebaliknya. Korelasi yang tinggi ini juga mengindikasikan bahwa luas panen merupakan salah satu faktor utama yang memengaruhi jumlah produksi beras dalam dataset ini.

D. *Splitting Data*

Splitting data merupakan tahapan membagi dataset menjadi dua bagian yaitu data latih (*training*) dan data validasi (*testing*). Dataset dibagi dengan komposisi 80% untuk data *training* dan 20% persen data *testing.* Data *training* digunakan untuk melatih model agar mempelajari pola dari data, Sedangkan data *testing* digunakan untuk mengevaluasi kinerja model pada data baru.

Di dalam dataset terdapat 19 data, sehingga diperoleh sebanyak 15 data untuk data *training* dan diperoleh 4 data untuk data *testing.*

```
from sklearn.model_selection import train_test_split
X = df[['Luas Panen (Ha)','Jumlah Curah Hujan (mm)','Jumlah Hari Hujan (Hari)',
        'Rata-rata Suhu (Celcius)','Rata-rata Kelembapan Udara']]
y = df['Produksi Beras (ton)']

# Membagi data menjadi training dan testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Gbr. 3 *Splitting* Data

E. *Pemodelan*

Di tahap pemodelan, peneliti mengimplementasikan algoritma *multiple linear regression* dan membuat model regresi dengan variabel "model" untuk memanggil kelas (Gambar 4)LinearRegression().

```
from sklearn.linear_model import LinearRegression

# Membuat model regresi linier
model = LinearRegression()

# Melatih model
model.fit(X_train, y_train)
```

```
LinearRegression
LinearRegression()
```

Gbr. 4 Algoritma *Multiple Linear Regression*

Kemudian memprediksikan produksi beras menggunakan algoritma *multiple linear regression* dengan cara menguji data *training* menggunakan data *testing*, kemudian menampilkan hasil prediksi (Gambar 5).

```
# Membuat prediksi
y_pred = model.predict(X_test)

print("Prediksi Jumlah Produksi Beras (Ton):", y_pred)
```

```
Prediksi Jumlah Produksi Beras (Ton): [75130.15539922 62617.37574455 89815.67741501 42824.9483904 ]
```

Gbr. 5 Hasil Prediksi

F. *Evaluasi Model*

Dalam *R-squared,* semakin nilai mendekati 1 maka model semakin bagus. Begitupun sebaliknya, jika model menjauhi nilai 1 atau mendekati 0 maka model semakin buruk. Dengan nilai *R-squared* sebesar 0,667, model *multiple linear regression* mampu menjelaskan 66,7% variasi dalam data produksi beras (gambar 6). Hal ini menunjukan model cukup baik dalam melakukan prediksi.

Dengan *Mean Absolute Percentage Error* (MAPE) sebesar 10,61% menunjukkan bahwa rata-rata kesalahan prediksi hanya sekitar 10,61% dari nilai aktual. Nilai MAPE yang berada diantara 10% - 20% menunjukan bahwa model

```
R-squared (R²): 0.6674186892591518
Mean Absolute Percentage Error (MAPE): 10.612478271715517 %
```

Gbr. 6 Hasil Evaluasi Model

memiliki kemampuan prediksi yang baik [15].

G. *Implementasi*

Setelah melakukan pelatihan dan evaluasi pada model regresi, tahapan selanjutnya adalah menyimpan model dalam format .pkl yang kemudian akan diimplementasikan ke dalam sistem berbasis *website* menggunakan *framework Streamlit.*

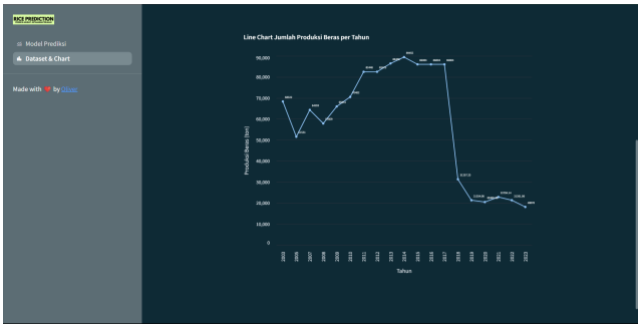


Gbr. 7 Halaman Model Prediksi

Pada halaman model prediksi sistem ini, pengguna/*user* dapat langsung melakukan prediksi dengan menginput jumlah luas panen (Ha), curah hujan (mm), hari hujan, suhu (°C), dan kelembapan udara (%) seperti pada Gambar 7.



Gbr. 8 Halaman *Dataset & Chart* (Bagian Dataset)

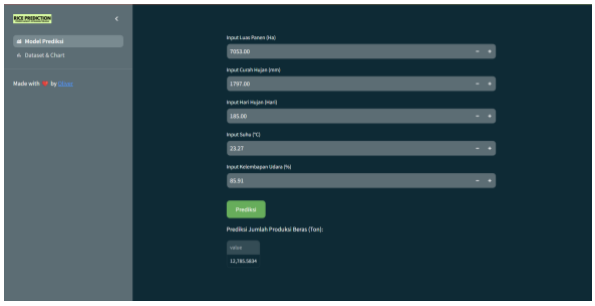


Gbr. 9 Halaman *Dataset & Chart* (Bagian *Line Chart*)

Pada halaman *Dataset & Chart* sistem ini pengguna/*user* bisa melihat dataset histori produksi padi dari tahun 2003 – 2023(Gambar 8). Selain itu, *user*/pengguna juga bisa melihat tren produksi padi dari tahun ke tahun, 2003-2023 (Gambar 9).

H. Uji Coba Prediksi Menggunakan Sistem

Sebagai uji coba, variabel-variabel prediktor diisi dengan angka 7053.00 sebagai jumlah luas panen (Ha), 1797.00 sebagai jumlah curah hujan (mm), 185.00 sebagai jumlah hari hujan (hari), 23.27 sebagai jumlah suhu (°C), dan 85.91 sebagai jumlah persentase kelembapan udara (%). Setelah variabel-variabel prediktor diisi, selanjutnya menekan tombol "Prediksi" maka hasil prediksi akan muncul seperti pada Gambar 10.



Gbr. 10 Hasil Uji Coba Prediksi Menggunakan Sistem

V. KESIMPULAN

Penelitian ini menghasilkan model *multiple linear regression* dengan nilai *R-Squared* sebesar 0.667 yang berarti 66.7% faktor-faktor produksi beras dapat dijelaskan oleh luas panen, curah hujan, hari hujan, suhu, dan kelembapan udara. Sedangkan sisanya 33.3 % dapat dijelaskan oleh faktor -faktor lain yang tidak diteliti dalam penelitian ini. Presentase rata-rata kesalahan relatif (MAPE) yang diperoleh yaitu 10.61% yang mana menunjukan bahwa model ini baik dalam melakukan prediksi. Penelitian ini juga menghasilkan model prediksi yang telah diintegrasikan ke dalam sistem berbasis website yang diharapkan dapat membantu pemerintah dan pemangku kepentingan dalam merancang kebijakan ketahanan pangan berbasis data. Adapun saran untuk penelitian selanjutnya yaitu sertakan faktor-faktor tambahan seperti penggunaan pupuk, jenis tanah, penyakit tanaman, dan metode irigasi yang mungkin berdampak pada hasil produksi.

DAFTAR PUSTAKA

[1] F. F. Az-zahra, T. Informatika, U. M. Sukabumi, K. Sukabumi, R. L. Berganda, and A. Regresi, "Prediksi hasil panen wortel menggunakan algoritma regresi linear berganda," *J. Mhs. Tek. Inform.*, vol. 8, no. 5, pp. 10255–10262, 2024.

[2] M. F. I. Raehan, A. B. Kusdinar, and D. Indrayana, "Penerapan Regresi Linier Berganda Untuk Memprediksi Hasil Panen Kacang Kedelai (Studi Kasus: Kecamatan Surade)," *J. Mhs. Tek. Inform.*, vol. 8, no. 5, pp. 10572–10579, 2024.

[3] A. Susanto, *Sistem Informasi Akuntansi*, 1st ed. Bandung: Lingga Jaya, 2017.

[4] J. Hutahaean, *Konsep Sistem Informasi*, 1st ed. Deepublish, 2016.

[5] F. Rohmawati, M. G. Rohman, and S. Mujilahwati, "Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec.Sugio Kab.Lamongan Menggunakan Metode Fuzzy Time Series," *Jouticla*, vol. 2, no. 2, 2017, doi: 10.30736/jti.v2i2.66.

[6] S. Wardah and I. Iskandar, "Forecasting Analysis of Packaged Banana Chips Product Sales (Case Study: Home Industry Arwana Food Tembilahan)," *Ind. Eng. J.*, vol. 11, no. 3, p. 135, 2017.

[7] A. A. I. Alitawan and I. K. Sutrisna, "Faktor-Faktor yang Mempengaruhi Pendapatan Petani Jeruk pada Desa Gunung Bau Kecamatan Kintamani Kabupaten Bangli," *E-Jurnal Ekon. Pembang. Univ. Udayana*, vol. 6, no. 5, p. 165350, 2017.

[8] A. A. T. Susilo, R. Kurniawan, and G. P. Julyansa, "Penerapan Model Regresi Linier Berganda Pada Sistem Prediksi Produksi Hasil Pertanian(Padi) Di Kota Lubuklinggau Dengan Bahasa Pemrograman Phyton," *J. Inform. dan Komput.* , vol. 15, no. 1, pp. 24–30, 2024.

[9] FAO, *The State of Food and Agriculture 2019*. Rome, 2019.

[10] P. A. Azhar, M. A. Pratama, and R. Fitriani, "Prediksi Harga Mobil Audi Bekas Menggunakan Model Regresi Linear dengan Framework Streamlit," vol. 6, no. 1, 2024, doi: 10.37802/joti.v6i1.763.

[11] T. N. Padilah and R. I. Adam, "Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang," *FIBONACCI J. Pendidik. Mat. dan Mat.*, vol. 5, no. 2, p. 117, 2019, doi: 10.24853/fbc.5.2.117-128.

[12] D. S. Syilfi, Dwi Ispriyanti, "Analsis Regresi Linier Piecewise Dua Segmen," *J. Gaussin*, vol. 1, 2017.

[13] P. R. Fitrayana and D. R. S. Saputro, "Algoritme Clustering Large Application (CLARA) untuk Menangani Data Outlier," in *PRISMA, Prosiding Seminar Nasional Matematika*, 2022, vol. 5, pp. 721–725.

[14] F. V. P. Samosir, L. P. Mustamu, E. D. Anggara, A. I. Wiyogo, and A. Widjaja, "Exploratory Data Analysis terhadap Kepadatan Penumpang Kereta Rel Listrik," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 2, pp. 449–467, 2021, doi: 10.28932/jutisi.v7i2.3700.

[15] Arwin Datumaya Wahyudi Sumari, Muhammad Bisri Musthafa, Ngatmari, and Dimas Rossiawan Hendra Putra, "Perbandingan Kinerja Metode-Metode Prediksi pada Transaksi Dompet Digital di Masa Pandemi," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 4, pp. 642–647, 2020, doi: 10.29207/resti.v4i4.2024.