

BAB II

TINJAUAN PUSTAKA

Bab ini menyajikan penelitian-penelitian terdahulu yang relevan dengan topik prediksi hasil panen padi. Pembahasan difokuskan pada dua kelompok utama, yaitu penelitian yang mengembangkan model prediksi hasil panen padi dan penelitian yang mengkaji faktor-faktor yang mempengaruhi hasil panen padi. Pada bagian akhir bab ini, disajikan kesenjangan penelitian yang menjadi dasar perumusan tujuan dan pendekatan penelitian.

2.1 Penelitian Terdahulu Terkait Model Prediksi Hasil Panen Padi

Penelitian-penelitian terdahulu menunjukkan bahwa pendekatan statistik dan *machine learning* banyak digunakan untuk memprediksi hasil panen padi. Metode statistik seperti regresi linier digunakan untuk memahami hubungan dasar antar variabel, sementara algoritma *machine learning*, khususnya algoritma *ensemble* digunakan untuk menangani data pertanian yang bersifat kompleks dan non-linear. Secara umum, penelitian-penelitian tersebut berfokus pada pencapaian akurasi prediksi yang tinggi dengan memanfaatkan data historis, iklim, dan lingkungan pada berbagai skala wilayah.

Meskipun algoritma *ensemble* seperti *Random Forest*, *XGBoost* dan *Gradient Boosting Tree Regression*. sering menunjukkan performa yang unggul, analisis faktor yang mempengaruhi hasil panen pada sebagian besar penelitian masih digunakan sebagai analisis pendukung untuk menjelaskan kinerja model terbaik, dan belum menjadi tujuan utama penelitian. Ringkasan penelitian terdahulu berbasis model prediksi hasil panen padi disajikan pada Tabel 2.1.

Tabel 2. 1 Penelitian Terdahulu Terkait Model Prediksi

Penulis Penelitian	Teknik prediksi	Model /Algoritma yang digunakan	Jenis Prediksi/ Fokus studi	Jenis dan sumber data	Hasil Penelitian	Implikasi terhadap Penelitian Ini
Shah et al., (2020)	Statistik	<i>Multiple Linear Regression (MLR)</i>	Identifikasi faktor penentu hasil panen padi	Data <i>Time Series</i> : 10 tahun terakhir (2005 – 2014) .	Model MLR berbasis rata-rata masa tanam lebih stabil.	Dasar penggunaan Regresi Linier Berganda sebagai model interpretatif
Asamoah et al., (2024)	<i>Machine Learning</i>	<i>Random Forest</i>	Prediksi hasil panen dan efisiensi agronomi.	Data <i>Time Series</i> (1991 – 2020) .	<i>Random Forest</i> menunjukkan kinerja prediksi yang baik	Mendukung penggunaan RF untuk hubungan non-linear

Abate et al., (2025)	<i>Machine Learning</i>	<i>Random Forest, Gradient Boosting, KNN, dan Decision Tree regressors</i>	Prediksi hasil panen berbasis data lingkungan dan citra satelit.	Data Time Series (2004 – 2017)	<i>Random Forest Regressor</i> memiliki performa terbaik..	Menunjukkan keunggulan algoritma <i>ensemble</i>
Islam et al., (2024)	<i>Machine Learning</i>	<i>Random Forest, Neural Network, Decision Tree, Linear Regression, and Gradient Boosting</i>	Prediksi hasil panen padi Aman.	Data <i>Historis</i> (1970–2022)	<i>Random Forest</i> paling akurat dan stabil.	Referensi pemilihan RF dan matrik evaluasi
Mugema ngango et al., (2024)	<i>Machine learning</i>	<i>Linear Regression, Random Forest, Gradient Boosting, Support Vector Machine, Arti?icial Neural Network, eXtreme Gradient Boosting Tree, and AdaBoost.</i>	Prediksi hasil panen padi berbasis input dan praktik budidaya	dataset sekunder dari Institut Statistik Nasional Rwanda (NISR).	<i>Adaptive Boosting Tree</i> menunjukkan kinerja terbaik.	Dasar perbandingan multi-algoritma
Mahesh & Soundra	<i>Machine learning</i>	<i>Categorical Boosting (CatBoost), Light Gradient-Boosting Machine</i>	Evaluasi algoritma <i>boosting</i> untuk	Data Time Series (1961 – 2016)	<i>CatBoost</i> memiliki kinerja terbaik.	Menguatkan penggunaan

pandiyana, (2024)		<i>(LightGBM), dan eXtreme Gradient Boosting (XGBoost)</i>	prediksi hasil panen			algoritma <i>boosting</i>
Lou et al., (2024)	<i>Ensemble learning</i>	<i>ridge regression, gradient boosting, and random forest</i>	Prediksi hasil panen di berbagai tahap pertumbuhan	Data Time Series (2012 – 2021)	<i>regresi voting ensemble</i> memiliki kinerja sangat baik.	<i>Menunjukkan potensi ensemble</i>
Mahdizadeh Gharakhani & Perez, (2024)	<i>Ensemble learning</i>	<i>Adaptive Boosting (AdaBoost), Gradient Boosting Machine (GBM), XGBoost, LightGBM, dan Random Forest (RF)</i>	Prediksi hasil panen	Data pertanian dan iklim regional Kanada.	Model <i>XGBoost</i> menunjukkan kinerja prediksi terbaik.	Referensi penggunaan <i>XGBoost</i>
Ravikumar et al., (2025)	<i>Ensemble machine learning</i>	<i>Long Short-Term Memory (LSTM) networks, Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost) (SERWI Ensemble)</i>	Prediksi hasil panen tanam pangan.	Dataset komprehensif multi-dekade 2023 - 2024	Penelitian ini mengusulkan <i>SERWI</i> sebagai model ensemble hibrida yang unggul.	Referensi penggunaan <i>XGBoost</i>

Erlin et al., (2024)	<i>Machine Learning</i>	<i>Random Forest, Gradient Boosting, Decision Tree, Support Vector Machine, and Artificial Neural Network</i>	Prediksi produksi padi di Indonesia	Data sekunder BPS & BMKG (2018–2023)	RF terbaik (R^2 hingga 0,91)	Konteks Indonesia, masih fokus akurasi & data sekunder
----------------------	-------------------------	---	-------------------------------------	--------------------------------------	---------------------------------	--

2.2 Penelitian Terdahulu Terkait Faktor-Faktor Yang Mempengaruhi Hasil Panen Padi

Penelitian terdahulu juga banyak mengkaji faktor-faktor yang mempengaruhi hasil panen padi. Faktor-faktor tersebut umumnya dikelompokkan menjadi faktor internal dan faktor eksternal. Faktor internal berkaitan dengan praktik dan keputusan budidaya petani, seperti penggunaan pupuk, pestisida, jenis benih, dan pengelolaan lahan. Faktor eksternal berkaitan dengan kondisi lingkungan, seperti curah hujan dan iklim.

Hasil-hasil penelitian menunjukkan bahwa baik faktor internal maupun eksternal berperan penting dalam menentukan produktivitas padi. Namun, sebagian besar penelitian masih membahas faktor-faktor tersebut secara terpisah atau menggunakan pendekatan analisis tertentu, sehingga belum mengintegrasikan seluruh faktor dalam satu kerangka analisis prediktif yang utuh. Ringkasan penelitian terdahulu berbasis faktor-faktor disajikan pada

Tabel 2. 2 Penelitian Terkait Hasil Panen Padi

Penulis	Fokus Faktor	Kategori Faktor	Temuan Utama Penelitian Terdahulu	Implikasi terhadap Penelitian Ini
Susanti et al., (2019)	Pupuk dan pestisida	Internal	Kontribusi input $\pm 56\%$ terhadap hasil panen	Variabel pupuk dan pestisida dianalisis berdasarkan data petani
Fathur et al., (2023)	Penggunaan pestisida	Internal	Overuse berdampak lingkungan & kesehatan	Penggunaan pestisida dikaitkan dengan hasil panen
Ansari et al., (2023)	Curah hujan & suhu	Eksternal	Variasi suhu dan curah hujan mempengaruhi	Faktor iklim dimasukkan

			fluktuasi hasil panen padi di Indonesia	sebagai variabel eksternal
Ma et al., (2025)	Iklim & efisiensi nitrogen	Internal & Eksternal	Interaksi iklim dan nitrogen menentukan hasil	Iklim dan penggunaan pupuk dianalisis secara bersamaan
Salam et al., (2021)	Praktik budidaya petani	Internal	Manajemen lahan berpengaruh signifikan	Data petani mencerminkan praktik budidaya yang sebenarnya di lapangan
Mugemangango et al., (2024)	Pupuk, musim tanam, benih	Internal & Eksternal	Faktor pupuk & musim dominan	Faktor dianalisis lebih rinci secara lokal

2.3 Kesenjangan Penelitian

Berdasarkan kajian pada Tabel 2.1 dan Tabel 2.2, dapat diidentifikasi beberapa kesenjangan penelitian. Dari aspek metode, sebagian besar penelitian prediksi hasil panen padi masih berfokus pada peningkatan akurasi model tanpa menjadikan analisis faktor-faktor yang mempengaruhi hasil panen sebagai tujuan utama penelitian. Dari aspek data, penelitian masih banyak menggunakan data sekunder atau data agregat berskala regional hingga nasional sehingga belum sepenuhnya merepresentasikan variasi praktik budidaya di tingkat petani.

Dari aspek analisis, sebagian penelitian belum mengintegrasikan faktor internal dan eksternal secara bersamaan dalam satu kerangka analisis. Selain itu, dari aspek konteks lokal, penelitian berbasis data primer petani yang menggambarkan kondisi nyata di lapangan, khususnya pada wilayah seperti Kabupaten Nagekeo, masih relatif terbatas. Oleh karena itu, penelitian ini dilakukan untuk membandingkan beberapa algoritma regresi serta

mengidentifikasi faktor-faktor internal dan eksternal yang paling berpengaruh terhadap hasil panen padi berbasis data petani di Kabupaten Nagekeo.