

PENERAPAN MODEL *GEOGRAPHICALLY WEIGHTED RANDOM FOREST* DAN *SHAPLEY ADDITIVE EXPLANATIONS* (SHAP) TERHADAP PRODUKTIVITAS PADI DI INDONESIA

DEDEN AHMAD RABANI



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan Model *Geographically Weighted Random Forest* dan *Shapley Additive Explanations* (SHAP) terhadap Produktivitas Padi di Indonesia” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Deden Ahmad Rabani
G1401221016



ABSTRAK

DEDEN AHMAD RABANI. Penerapan Model *Geographically Weighted Random Forest* dan *Shapley Additive Explanations* (SHAP) terhadap Produktivitas Padi di Indonesia. Dibimbing oleh AAM ALAMUDI dan RAHMA ANISA.

Produktivitas padi di Indonesia menunjukkan variasi spasial antarwilayah akibat interaksi non-linear yang kompleks antara faktor iklim dan sosioekonomi. Kondisi ini menuntut pendekatan pemodelan yang tidak hanya mampu menangkap hubungan non-linear antarpeubah, tetapi juga heterogenitas spasial antarwilayah. Penelitian ini bertujuan untuk membandingkan penerapan model *Geographically Weighted Random Forest* (GWRF) dengan model *Random Forest* global dalam memodelkan produktivitas padi di Indonesia, dan juga mengidentifikasi peubah penting lokal beserta arah pengaruhnya melalui pendekatan interpretasi berbasis *Shapley Additive Explanations* (SHAP). GWRF dipilih karena mampu menggabungkan keunggulan *Random Forest* dalam menangani hubungan non-linear dengan pembobotan geografis untuk menangkap variasi karakteristik antarwilayah, sedangkan SHAP digunakan untuk mengatasi sifat *black box* dari model berbasis *Random Forest* dengan menjelaskan kontribusi spesifik setiap peubah pada setiap lokasi. Penelitian menggunakan amatan 471 kabupaten/kota di Indonesia pada tahun 2024 dengan delapan peubah penjelas. Model GWRF dengan *bandwidth* optimal 20 tetangga terdekat (*nearest neighbors*) lebih baik dibandingkan *Random Forest* global, dengan nilai OOB R^2 yang lebih tinggi yaitu sebesar 96,23% serta nilai OOB *RMSE* dan OOB *MAE* yang lebih rendah dengan nilai masing-masing sebesar 2,08 dan 1,64. Analisis SHAP menunjukkan bahwa jumlah penyuluh pertanian merupakan faktor paling dominan terhadap prediksi produktivitas padi di Indonesia, diikuti oleh IPM, curah hujan, kelembapan tanah, dan radiasi matahari secara global, dengan pola kontribusi yang bervariasi secara signifikan antarwilayah.

Kata kunci: GWRF, peubah penting, produktivitas padi, random forest, SHAP

ABSTRACT

DEDEN AHMAD RABANI. Application of Geographically Weighted Random Forest and Shapley Additive Explanations (SHAP) to Rice Productivity in Indonesia. Supervised by AAM ALAMUDI and RAHMA ANISA.

Rice productivity in Indonesia exhibits spatial variation across regions due to complex non-linear interactions between climatic and socioeconomic factors. This condition calls for a modeling approach capable of capturing not only non-linear relationships among variables but also spatial heterogeneity across regions. This study aims to compare the application of the Geographically Weighted Random Forest (GWRF) model with the global Random Forest model in modeling rice productivity in Indonesia, as well as to identify locally important variables and their directions of influence through an interpretation approach based on Shapley Additive Explanations (SHAP). GWRF was selected for its ability to combine the strength of Random Forest in handling non-linear relationships with geographic weighting to capture variation in regional characteristics, while SHAP was employed to address the black-box nature of Random Forest-based models by explaining the specific contribution of each variable at each location. This study used observations from 471 districts/cities in Indonesia in 2024 with eight explanatory variables. The GWRF model with an optimal bandwidth of 20 nearest neighbors outperformed the global Random Forest, with a higher OOB R^2 of 96.23% and lower OOB $RMSE$ and OOB MAE values of 2.08 and 1.64, respectively. SHAP analysis revealed that the number of agricultural extension workers is the most dominant factor predicting rice productivity in Indonesia, followed globally by the Human Development Index (HDI), rainfall, soil moisture, and solar radiation, with contribution patterns varying considerably across regions.

Keywords: Feature importance, GWRF, random forest, rice productivity, SHAP



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENERAPAN MODEL *GEOGRAPHICALLY WEIGHTED RANDOM FOREST* DAN *SHAPLEY ADDITIVE EXPLANATIONS* (SHAP) TERHADAP PRODUKTIVITAS PADI DI INDONESIA

DEDEN AHMAD RABANI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Dr. Farit Mochamad Afendi, S.Si., M.Si.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Penerapan Model *Geographically Weighted Random Forest* dan
Shapley Additive Explanations (SHAP) terhadap Produktivitas
Padi di Indonesia

Nama : Deden Ahmad Rabani
NIM : G1401221016

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Ir. Aam Alamudi, M.Si.



Pembimbing 2:
Rahma Anisa, S.Stat., M.Si., M.Act.Sc



Diketahui oleh

Ketua Program Studi:
Dr. Bagus Sartono, S.Si., M.Si
NIP. 197804112005011002



Tanggal Ujian:
(25 Juli 2026)

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Judul yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2025 sampai bulan Juli 2026 ini ialah “Penerapan Model *Geographically Weighted Random Forest* dan *Shapley Additive Explanations* (SHAP) terhadap Produktivitas Padi di Indonesia”. Pada proses pembuatan karya ilmiah ini, penulis tak terlepas mendapatkan dukungan, arahan, bimbingan, dan do'a dari berbagai pihak. Oleh karena itu, penulis ingin memberikan ucapan terimakasih setinggi-tingginya kepada:

1. Ibu Elis Lisnawati dan Bapak Sunarya selaku keluarga penulis yang senantiasa memberikan doa, dukungan, motivasi, dan kasih sayang kepada penulis.
2. Bapak Ir. Aam Alamudi, M.Si. dan Ibu Rahma Anisa, S.Stat., M.Si., M.Act.Sc selaku dosen pembimbing yang telah memberikan bimbingan, arahan, nasihat, saran, dan semangat kepada penulis selama proses penyusunan karya ilmiah ini.
3. Keluarga besar Program Studi Statistika dan Sains Data yang telah memberikan ilmu, pengalaman, dan pembelajaran yang bermanfaat selama masa perkuliahan.
4. Teman-teman dan sahabat Statistika IPB 59, terutama R. Mugni Chairil Arbi Asyari, Hariyol, Haidar Ramdhani, Muhammad Abdan Rofi, Delita Nur Hasanah, Sabrina Afifah Putri Utami, Shalma Kaisya Candradewi, dan Raihana Asma Amani, yang telah membersamai, memberikan dukungan, serta berjuang bersama selama masa perkuliahan.
5. Seluruh pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan, dukungan, dan doa dalam penyelesaian karya ilmiah ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2026

Deden Ahmad Rabani



DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
TINJAUAN PUSTAKA	3
2.1 <i>Random Forest</i>	3
2.2 <i>Geographically Weighted Random Forest</i>	4
2.3 <i>Shapley Additive Explanations</i> (SHAP)	5
III METODE	6
3.1 Data	6
3.2 Prosedur Analisis Data	6
IV HASIL DAN PEMBAHASAN	9
4.1 Eksplorasi Data	9
4.2 Pemodelan <i>Geographically Weighted Random Forest</i> (GWRF)	11
4.3 <i>Local R²</i> Model GWRF	12
4.4 Identifikasi Peubah Penting Menggunakan SHAP	13
V SIMPULAN DAN SARAN	21
5.1 Simpulan	21
5.2 Saran	21
DAFTAR PUSTAKA	22
LAMPIRAN	25
RIWAYAT HIDUP	28



Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1	Daftar peubah yang digunakan	6
2	Kombinasi <i>hyperparameter</i> model <i>Random Forest</i> global	7
3	Statistik deskriptif peubah respon dan peubah penjelas	9
4	Evaluasi model <i>Random Forest</i> dan <i>GWR</i>	12

DAFTAR GAMBAR

1	Ilustrasi algoritma <i>Random Forest</i> (Mengüç dan Aydin 2022)	3
2	Peta sebaran produktivitas padi Indonesia	10
3	Diagram pencar peubah penjelas terhadap produktivitas padi: (a) X_1 , (b) X_2 , (c) X_3 , (d) X_4 , (e) X_5 , (f) X_6 , (g) X_7 , (h) X_8	10
4	Diagram garis hasil evaluasi <i>hyperparameter</i> <i>n_estimator</i> serta <i>max_feature</i> berdasarkan R^2 pada model global	11
5	Diagram garis hasil evaluasi parameter <i>bandwidth</i> berdasarkan R^2 pada model lokal	12
6	Distribusi R^2 lokal: (a) Peta sebaran R^2 lokal (b) Histogram sebaran R^2 lokal	13
7	Tingkat kontribusi peubah secara global: (a) <i>mean</i> SHAP value (b) SHAP value	13
8	Peta sebaran kontribusi SHAP jumlah penyuluh pertanian terhadap produktivitas padi	14
9	Peta sebaran kontribusi SHAP IPM terhadap produktivitas padi	15
10	Peta sebaran kontribusi SHAP curah hujan terhadap produktivitas padi	16
11	Peta sebaran kontribusi SHAP kelembapan tanah terhadap produktivitas padi	16
12	Peta sebaran kontribusi SHAP radiasi matahari terhadap produktivitas padi	17
13	Peta sebaran kontribusi SHAP persentase kepala rumah tangga petani terhadap produktivitas padi	18
14	Peta sebaran kontribusi SHAP rata-rata suhu maksimum terhadap produktivitas padi	18
15	Peta sebaran kontribusi SHAP rata-rata suhu minimum terhadap produktivitas padi	19

DAFTAR LAMPIRAN

1	<i>Local R^2</i> model <i>GWR</i>	26
2	Kontribusi <i>shap value</i> tiap peubah di tiap wilayah	27