

PERBANDINGAN METODE SELANG PREDIKSI DALAM PEMODELAN REGRESI *RANDOM FOREST* UNTUK RESPON BERSEBARAN TIDAK SIMETRIK

INDAH MAGFIRRAH



PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025

@Hak cipta milik IPB University

IPB University



IPB University
Bogor Indonesia

- 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.

Perpustakaan IPB University



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Perbandingan Metode Selang Prediksi dalam Pemodelan Regresi *Random Forest* untuk Respon Bersebaran Tidak Simetrik” 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 tesis ini.

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

Bogor, November 2025

Indah Magfirrah
G1501231005

RINGKASAN

INDAH MAGFIRRAH. Perbandingan Metode Selang Prediksi dalam Pemodelan Regresi *Random Forest* untuk Respon Bersebaran Tidak Simetrik. Dibimbing oleh BAGUS SARTONO dan KHAIRIL ANWAR NOTODIPUTRO.

Random forest merupakan pendekatan prediktif yang handal dengan akurasi dan stabilitas yang tinggi. Meskipun demikian, metode ini memiliki keterbatasan dalam mengukur ketidakpastian hasil prediksi. Oleh karena itu, berbagai pengembangan metode selang prediksi berbasis *random forest* telah banyak dilakukan, di antaranya *Quantile Regression Forests* (QRF), *Split Conformal* (SC), dan *Out-of-Bag Prediction Interval* (OOB-PI). Metode QRF dapat dipandang sebagai perluasan dari *random forest*, karena memanfaatkan sebaran kumulatif bersyarat dari peubah respon untuk membentuk selang prediksi. Di sisi lain, metode SC membangun selang prediksi dengan membagi data menjadi dua bagian, yakni satu bagian untuk melatih model dan bagian lainnya untuk menduga kuantil dari sebaran galat. Sementara itu, metode OOB-PI membentuk selang prediksi berdasarkan kuantil empiris dari sebaran galat prediksi *out-of-bag*.

Selang prediksi adalah rentang nilai yang berpeluang besar untuk mencakup nilai sebenarnya sehingga dapat digunakan untuk pengambilan keputusan dengan tingkat kepercayaan tertentu. Penelitian yang telah dilakukan sebelumnya menunjukkan metode OOB-PI lebih unggul dibandingkan metode QRF dan SC dengan temuan penting jika sebaran peubah respon simetrik, maka akan dihasilkan sebaran galat prediksi *out-of-bag* yang juga bersifat simetrik. Sebaran galat berbentuk simetrik akan menghasilkan selang prediksi yang lebih akurat sehingga tidak bias ke satu sisi serta konsisten karena memiliki ragam yang kecil. Dalam praktiknya, sering dijumpai peubah respon yang bersebaran tidak simetrik sehingga menimbulkan tantangan tersendiri dalam mengembangkan model prediksi. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi kinerja metode QRF, SC, dan OOB-PI pada peubah respon yang tidak simetrik menggunakan kajian simulasi. Pada kajian simulasi, akan dibangkitkan sebaran galat yang menjulur ke kanan untuk mencerminkan peubah respon pada berbagai tingkat kemiringan. Selain kajian simulasi, penelitian ini juga akan mengevaluasi kinerja ketiga metode menggunakan data empiris untuk menganalisis dan memprediksi pengeluaran per kapita rumah tangga yang diperoleh dari SUSENAS 2023 BPS Provinsi Jawa Barat.

Penelitian ini difokuskan pada Kabupaten/Kota dengan tingkat Produk Domestik Regional Bruto (PDRB) yang tinggi, sedang (median/menengah), dan rendah di Provinsi Jawa Barat. Daerah dengan tingkat PDRB yang tinggi cenderung menunjukkan tingkat pengeluaran per kapita yang lebih besar, sementara daerah dengan PDRB yang rendah umumnya memiliki tingkat pengeluaran per kapita yang lebih rendah. Wilayah dengan PDRB tinggi yang dipilih dalam penelitian ini adalah Kabupaten Bekasi ($n = 1157$) dan Kota Bandung ($n = 1105$). Wilayah dengan PDRB median yang dipilih adalah Kabupaten Cianjur ($n = 1044$) dan Kota Bogor ($n = 808$). Sementara wilayah dengan PDRB rendah yang dipilih adalah Kabupaten Pangandaran ($n = 730$) dan Kota Banjar ($n = 564$) sehingga total contoh yang digunakan untuk analisis empiris sebanyak 5408 rumah tangga.

Tahapan simulasi utama dimulai dengan pembangkitan data simulasi sebanyak $n + n_0$, dengan $n = 500, 1.000, 2.000, 3.000$ merupakan ukuran contoh



pelatihan, dan $n_0 = 1.000$ ditetapkan sebagai contoh pengujian pada setiap skenario. Evaluasi kinerja metode QRF, SC, dan OOB-PI dilakukan menggunakan contoh pengujian dengan ukuran tetap, sehingga pengaruh variasi ukuran contoh pelatihan terhadap kinerja model dapat diamati secara lebih objektif, adil, dan tidak dipengaruhi oleh keragaman contoh pengujian. Setelah data dibangkitkan, dilakukan pembagian secara acak menjadi contoh pelatihan berukuran n dan contoh pengujian n_0 . Selanjutnya, model QRF, SC, dan OOB-PI dibangun menggunakan contoh pelatihan dengan *hyperparameter* terbaik. Evaluasi kinerja masing-masing model dilakukan pada contoh pengujian untuk menghasilkan selang prediksi dengan tingkat signifikansi $\alpha = 0,05$. Kinerja model diukur berdasarkan dua metrik, yaitu tingkat cakupan (*coverage rate*) dan lebar selang (*interval width*). Seluruh prosedur simulasi diulang sebanyak 1.000 replikasi, kemudian hasil dugaan *coverage rate* dan *interval width* dianalisis lebih lanjut menggunakan *Analysis of Variance* (ANOVA). Prosedur analisis data empiris diawali dengan eksplorasi data, kemudian dilanjutkan dengan tahap praproses data. Setelah itu dilakukan *tuning hyperparameter* pada *random forest* menggunakan *10-fold cross validation*, dengan evaluasi berdasarkan nilai MAE pada data uji hasil validasi silang. Selanjutnya, membangun model selang prediksi dengan *hyperparameter* terbaik serta mengevaluasinya menggunakan *10-fold cross validation*. Seluruh proses ini diulang sebanyak 200 replikasi dengan $\alpha = 0,05$ sehingga menghasilkan dugaan kinerja model yaitu tingkat cakupan dan lebar selang.

Hasil kajian simulasi menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara metode SC dan OOB-PI pada dugaan tingkat cakupan di berbagai skenario. Metode OOB-PI menghasilkan tingkat cakupan yang konsisten dan stabil mendekati 0,95 serta lebar selang yang relatif sempit. Sedangkan metode SC menunjukkan kinerja yang sebanding tetapi cenderung menghasilkan selang yang lebih lebar dengan ragam yang relatif besar. Sementara metode QRF menunjukkan kinerja yang kompetitif, namun konsistensinya dalam mempertahankan tingkat cakupan masih kurang stabil.

Hasil kajian empiris menunjukkan bahwa metode SC dan OOB-PI memiliki kinerja yang kompetitif dengan tingkat cakupan sebesar 95%. Namun, metode OOB-PI menghasilkan selang prediksi yang relatif lebih sempit serta simpangan baku yang lebih rendah. Temuan ini mengindikasikan bahwa metode OOB-PI menunjukkan efisiensi metode yang stabil dan konsisten dalam menghasilkan selang prediksi pengeluaran per kapita meskipun pada wilayah dengan tingkat kemajuluran tinggi dan keragaman pengeluaran per kapita yang besar. Sebaliknya, metode SC kurang stabil karena cenderung memperlebar selang prediksi untuk mempertahankan cakupan 95% sehingga efisiensi kerjanya sedikit menurun.

Kata kunci: pengeluaran per kapita rumah tangga, tanpa kemiskinan, selang prediksi Out-of-Bag, quantile regression forests, split conformal

SUMMARY

INDAH MAGFIRRAH. Comparison of Prediction Interval Methods in Random Forest Regression Modeling for Asymmetrically Distributed Response. Supervised by BAGUS SARTONO and KHAIRIL ANWAR NOTODIPUTRO.

Random forest is a powerful predictive method known for its high accuracy and stability. However, it cannot directly quantify prediction uncertainty. To handle the problem, several methods for constructing prediction intervals have been developed, including Quantile Regression Forests (QRF), Split Conformal (SC), and Out-of-Bag Prediction Intervals (OOB-PI). The QRF method is a generalization of random forest where the conditional cumulative distribution of the response variables is used to construct the prediction interval. In contrast, SC builds prediction intervals by splitting the dataset into two parts, with one part used for model training and the other for estimating the quantiles of the residual distribution. Meanwhile, the OOB-PI method constructs prediction intervals based on the empirical quantiles of out-of-bag prediction errors.

A prediction interval is a range within which an observation is expected to fall with a specified probability, given a certain confidence level. Previous studies have shown that the OOB-PI method is more powerful than the QRF and SC methods, with the important finding that if the response variable distribution is symmetric, the out-of-bag prediction error distribution will also be symmetric. Symmetric error distribution will result in a more accurate prediction interval that is not biased to one side and is consistent because it has a small variance. In practice, it is usually found that the response variable is not symmetrically distributed, which poses a challenge in developing a prediction model. Therefore, this study used a simulation approach to evaluate the performance of three methods for asymmetrically distributed responses. The simulation generated right-skewed error distributions to reflect response variables with varying degrees of skewness. In addition, the methods were applied to empirical data to analyze and predict household per capita expenditure obtained from the 2023 SUSENAS conducted by BPS in West Java Province.

This study focuses on regions with high, medium, and low Gross Regional Domestic Product (GRDP) levels in West Java Province. In general, an increase in GRDP tends to raise per capita expenditure. In contrast, a decline in GRDP tends to reduce it, as GRDP serves as a key indicator of economic well-being. The selected high-GRDP regions are Bekasi Regency ($n = 1157$) and Bandung City ($n = 1105$); the medium-GRDP regions are Cianjur Regency ($n = 1044$) and Bogor City ($n = 808$); and the low-GRDP regions are Pangandaran Regency ($n = 730$) and Banjar City ($n = 564$). Thus, the total sample used for the empirical data analysis is 5408 households.

The main simulation began by generating data of size $n + n_0$, where $n = 500$; 1.000; 2.000; and 3.000, representing the training sample sizes, and $n_0 = 1.000$ was fixed as the test sample size in each scenario. The performance of the QRF, SC, and OOB-PI methods was evaluated using this fixed-size test set, enabling objective comparisons across different training sample sizes. This setup helped minimize the impact of variability in the test data. Subsequently, the data was randomly split into a training set of size n and a test set of size n_0 . The QRF,



SC, and OOB-PI models were trained on the training set using optimal hyperparameters, and their performance was evaluated on the test set at a significance level of $\alpha = 0.05$. The evaluation metrics for the prediction intervals included coverage rate and interval width. The simulation procedure was repeated 1.000 times, and the estimated coverage rate and interval width were further analyzed using Analysis of Variance (ANOVA). The empirical data analysis procedure began with exploratory data analysis, followed by data preprocessing. Hyperparameter tuning for random forest was then conducted using 10-fold cross validation, with evaluation based on the Mean Absolute Error (MAE) from the validation folds. Using the optimal hyperparameters, prediction interval models were subsequently constructed and evaluated through 10-fold cross validation. This process was repeated 200 times at a significance level of $\alpha = 0.05$, resulting in estimated model performance in coverage rate and interval width.

The simulation study results showed, there was no significant difference between the SC and OOB-PI methods on estimating the coverage rate in various scenarios. The OOB-PI method resulted in a consistent and stable coverage rate close to 0.95 and a relatively narrow interval width. The SC method performed similarly but resulted in larger interval widths with greater variance. In contrast, the QRF method also performed competitively, but its coverage rate was less stable.

The empirical analysis showed that the SC and OOB-PI methods had competitive performance with a coverage rate of 95%. However, the OOB-PI method resulted in a relatively narrower prediction interval and a lower standard deviation. These findings indicate that the OOB-PI method has stable and consistent efficiency in estimating per capita expenditure, even in regions with high levels of skewness and large variations in per capita expenditure. In contrast, the SC method was less stable because it tended to widen the prediction interval to achieve 95% coverage, thereby slightly reducing its efficiency performance.

Keywords: household per capita expenditure, no poverty, out-of-bag prediction interval, quantile regression forests, split conformal

- 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 mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



© Hak Cipta milik IPB, tahun 2025
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.

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.



PERBANDINGAN METODE SELANG PREDIKSI DALAM PEMODELAN REGRESI *RANDOM FOREST* UNTUK RESPON BERSEBARAN TIDAK SIMETRIK

INDAH MAGFIRRAH

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

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 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 Tesis:
Prof. Dr. Ir. Muhammad Nur Aidi, M.S.**



Digitally signed by
Muhammad Nur Aidi
Date: 8 Dec 2025 05:36:06 WIB
Verify at: [digipb.ipb.ac.id](#)



@Hak cipta milik IPB University

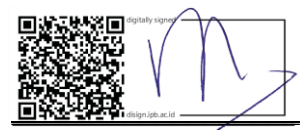
Judul Tesis : Perbandingan Metode Selang Prediksi dalam Pemodelan Regresi
Random Forest untuk Respon Bersebaran Tidak Simetrik
Nama : Indah Magfirrah
NIM : G1501231005

Disetujui oleh

Pembimbing 1:
Dr. Bagus Sartono, S.Si., M.Si.



Pembimbing 2:
Prof. Dr. Ir. Khairil Anwar Notodiputro, M.S.



Diketahui oleh

Ketua Program Studi:
Dr. Agus Mohamad Soleh, S.Si., M.T.
NIP 19750315 199903 1 004



Dekan Sekolah Sains Data, Matematika, dan Informatika:
Prof. Dr. Ir. Agus Bueno, M.Si., M.Kom.
NIP 19660702 199302 1 001



Tanggal Ujian:
(17 November 2025)

Tanggal Lulus:

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.



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhanaahu wa ta'ala* atas segala karunia-Nya sehingga penelitian ini dapat terselesaikan. Tema yang dipilih dalam penelitian ini adalah metode selang prediksi dalam algoritma pembelajaran mesin statistika, dengan judul “Perbandingan Metode Selang Prediksi dalam Pemodelan Regresi *Random Forest* untuk Respon Bersebaran Tidak Simetrik”. Karya ilmiah ini tidak lepas dari bantuan, bimbingan dan dukungan dari berbagai pihak, untuk itu penulis menyampaikan rasa terima kasih kepada:

1. Bapak Dr. Bagus Sartono, S.Si., M.Si. dan Bapak Prof. Dr. Ir. Khairil Anwar Notodiputro, M.S. selaku komisi pembimbing yang secara konsisten telah memberikan bimbingan, ilmu, arahan, saran, dan masukan dalam menyelesaikan karya ilmiah ini;
2. Bapak Prof. Dr. Ir. Muhammad Nur Aidi, M.S. selaku penguji luar komisi ujian tesis yang telah memberikan masukan, saran, dan arahan yang sangat berarti dalam penyusunan karya ilmiah ini;
3. Ibu Dr. Ir. Indahwati, M.Si. selaku pimpinan sidang tesis yang telah memberikan saran, dan arahan selama proses sidang tesis;
4. Bapak Jamaludin, Ibu Salmiah, Adik Wahda Dwi Sari dan Adik Muhammad Akmal Gantari serta seluruh keluarga yang telah memberikan dukungan, doa, dan motivasi demi kelancaran studi penulis;
5. Rekan-rekan mahasiswa pascasarjana program Magister Statistika dan Sains Data angkatan 2023 khususnya Meisyatul Ilma, Irdayanti, Yana Wulandari, Asyifah Qalbi dan Silvia Tri Wahyuni yang telah memberikan dukungan, dan semangat;
6. Rekan-rekan mahasiswa pascasarjana program Magister Kost Bu Roma khususnya Farida Kurniasari, Rizqia Nurul A, dan Rizqi Aulia Fajarwati H yang telah memberikan dukungan, motivasi, dan semangat;
7. Seluruh dosen dan tenaga pendidik Program Studi Statistika dan Sains Data IPB University yang telah banyak memberikan ilmu, kenyamanan dan kemudahan dalam proses pembelajaran selama menempuh jenjang kuliah terutama dalam membantu menyediakan data empiris yang digunakan dalam penelitian ini.

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

Bogor, November 2025

Indah Magfirrah

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 ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
TINJAUAN PUSTAKA	4
2.1 <i>Random Forest</i>	4
2.2 <i>Quantile Regression Forests</i>	6
2.3 <i>Split Conformal</i>	7
2.4 <i>Out-of-Bag Prediction Interval</i>	8
2.5 <i>Cross Validation</i>	10
2.6 <i>Mean Absolute Error</i>	10
2.7 <i>Factorial Randomized Complete Block Design</i>	11
III METODE PENELITIAN	13
3.1 Data	13
3.2 Model	17
3.3 Prosedur Analisis Data	20
IV HASIL DAN PEMBAHASAN	23
4.1 Kajian Simulasi	23
4.2 Kajian Empiris	29
V SIMPULAN DAN SARAN	37
5.1 Simpulan	37
5.2 Saran	37
DAFTAR PUSTAKA	38
LAMPIRAN	41
RIWAYAT HIDUP	64

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	Algoritma dasar pembentukan <i>random forest</i>	5
2	Skenario simulasi	15
3	Peubah penjelas	15
4	<i>Hyperparameter</i> simulasi	19
5	<i>Hyperparameter</i> data empiris	20
6	Hasil ANOVA dari tingkat cakupan dan lebar selang	26
7	Rangkuman kinerja metode terbaik pada berbagai kondisi simulasi	28
8	Statistik deskriptif pengeluaran per kapita rumah tangga (x Rp 100.000)	30
9	Perbandingan rata-rata dan simpangan baku dugaan lebar selang	34

DAFTAR GAMBAR

1	Ilustrasi model QRF	6
2	Ilustrasi model SC	7
3	Ilustrasi model OOB-PI	9
4	Ilustrasi <i>10-fold cross validation</i>	10
5	Diagram alir proses pembangkitan data	14
6	PDRB menurut Kab/Kota di Provinsi Jawa Barat tahun 2023	17
7	Diagram alir proses analisis simulasi utama	20
8	Diagram alir proses analisis data empiris	22
9	Boxplot dugaan tingkat cakupan pada peubah penjelas saling berkorelasi	24
10	Boxplot dugaan lebar selang pada peubah penjelas saling berkorelasi	25
11	Plot interaksi antara metode, fungsi $m(x)$ dan tingkat kemenjuluran terhadap tingkat cakupan	27
12	Plot interaksi antara metode, fungsi $m(x)$ dan tingkat kemenjuluran terhadap lebar selang	27
13	Boxplot pengeluaran per kapita rumah tangga (x Rp 100.000) berdasarkan wilayah	30
14	Nilai korelasi antar peubah kontinu	31
15	Hasil rata-rata MAE dari proses <i>tuning random forest</i>	32
16	Boxplot dugaan tingkat cakupan pada 6 wilayah di Provinsi Jawa Barat	33
17	Boxplot dugaan lebar selang pada 6 wilayah di Provinsi Jawa Barat	34
18	Perbandingan selang prediksi pada wilayah dengan PDRB (a) Tinggi; (b) Median; dan (c) Rendah	36



DAFTAR LAMPIRAN

1	Lampiran 1 Hasil dugaan dari 1.000 replikasi terhadap tingkat cakupan (<i>coverage rate</i>) pada peubah penjelas tidak berkorelasi	42
2	Lampiran 2 Hasil dugaan dari 1.000 replikasi terhadap lebar selang (<i>interval width</i>) pada peubah penjelas tidak berkorelasi	43
3	Lampiran 3 Hasil <i>Analysis of variance</i> dari <i>coverage rate</i> dan <i>interval width</i>	44
4	Lampiran 4 Hasil uji perbandingan ganda Tukey pada <i>coverage rate</i> dan <i>interval width</i>	46
5	Lampiran 5 Gambaran umum pengeluaran per kapita rumah tangga di Provinsi Jawa Barat pada Kabupaten Bekasi, Kabupaten Cianjur, Kabupaten Pangandaran, Kota Bandung, Kota Banjar dan Kota Bogor menurut setiap peubah penjelas	48
6	Lampiran 6 Hasil rata-rata MAE dari proses <i>tuning random forest</i> dengan 10-fold	57
7	Lampiran 7 Hasil dugaan pengeluaran per kapita rumah tangga metode SC dan OOB-PI untuk 20 pengamatan pada masing-masing Kabupaten/Kota	58

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.