

- 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.

EVALUASI METODE SELEKSI *EIGENVECTOR* PADA *EIGENVECTOR SPATIAL FILTERING* DENGAN MODEL *BOOSTING* UNTUK PREDIKSI KONSENTRASI PM_{2.5}

PUTRI NISRINA AZ-ZAHRA



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



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Evaluasi Metode Seleksi *Eigenvector* pada *Eigenvector Spatial Filtering* dengan Model *Boosting* untuk Prediksi Konsentrasi PM_{2.5}” 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, 1 Juli 2026

Putri Nisrina Az-Zahra
M0501241050

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan 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.

RINGKASAN

PUTRI NISRINA AZ-ZAHRA. Evaluasi Metode Seleksi *Eigenvector* pada *Eigenvector Spatial Filtering* dengan Model *Boosting* untuk Prediksi Konsentrasi PM_{2.5}. Dibimbing oleh ANIK DJURAI DAH dan ERFIANI.

Autokorelasi spasial merupakan fenomena ketika pola ketergantungan spasial menunjukkan hubungan yang sistematis antara nilai suatu peubah di suatu lokasi dengan nilai peubah yang sama di lokasi sekitarnya. Keberadaan karakteristik tersebut pada data spasial perlu diperhatikan karena dapat menyebabkan bias dalam model statistik apabila tidak ditangani dengan tepat. Salah satu metode yang efektif dalam menangani permasalahan tersebut adalah *Eigenvector Spatial Filtering* (ESF). ESF memanfaatkan vektor ciri dari matriks pembobot spasial untuk menangkap pola ketergantungan spasial dalam data dan mengikutsertakan pola tersebut ke dalam model sebagai peubah bebas tambahan sehingga dapat meningkatkan kualitas model prediksi.

Pada penerapan ESF, pemilihan vektor ciri yang optimal menjadi salah satu aspek yang perlu diperhatikan. Tahapan ini penting dalam pemodelan untuk menghindari *overfitting*, meningkatkan interpretabilitas model, dan menurunkan beban komputasi. Meski efektif dalam seleksi parsial, berbagai metode pemilihan vektor ciri yang umum digunakan belum memberikan informasi eksplisit mengenai kontribusi individual masing-masing vektor ciri terhadap prediksi model. Sebagai alternatif, metode SHAP dapat diterapkan untuk melakukan seleksi fitur yang dalam hal ini yaitu vektor ciri, dengan basis yang kuat dalam teori permainan. Vektor ciri dengan nilai SHAP tertinggi dapat diprioritaskan untuk dipilih dalam pemodelan ESF, sementara vektor ciri dengan kontribusi rendah dapat dieliminasi.

Metode ESF masih terbatas dalam menangkap hubungan nonlinier maupun interaksi yang kompleks antar peubah. Algoritma pembelajaran mesin telah menjadi pendekatan yang efektif karena kemampuannya dalam menangkap pola yang kompleks dan nonlinier. Oleh karena itu, integrasi ESF dengan algoritma pembelajaran mesin telah diterapkan seiring dengan berkembangnya data geospasial skala besar. Pada penelitian ini, akan dilakukan evaluasi metode pemilihan vektor ciri pada ESF dengan integrasi model *boosting* pada kasus kualitas udara dengan PM_{2.5} sebagai indikatornya. *Particulate Matter* (PM_{2.5}) merupakan partikel halus berdiameter kurang dari 2,5 μm yang digunakan sebagai indikator utama kualitas udara. Ukurannya yang sangat kecil dapat menembus saluran pernafasan terdalam manusia dan menimbulkan gangguan pernapasan, penyakit kardiovaskular, serta peningkatan risiko kematian dini. Provinsi DKI Jakarta sebagai ibu kota Indonesia menghadapi permasalahan polusi udara yang semakin serius akibat tingginya kepadatan penduduk dan intensitas penggunaan transportasi yang tinggi sehingga konsentrasi PM_{2.5} kerap melampaui ambang batas aman.

Penelitian ini mengkaji evaluasi metode seleksi *eigenvector* pada ESF yang dikombinasikan dengan model pembelajaran mesin berbasis *Gradient Boosting Machine* (GBM) untuk prediksi konsentrasi PM_{2.5} dengan tujuan mengatasi autokorelasi spasial dan meningkatkan akurasi model prediksi data spasial. Metode seleksi *eigenvector* yang dibandingkan meliputi kriteria *eigenvalue* positif, Indeks Moran, LASSO, dan SHAP (*Shapley Additive exPlanations*). Data simulasi dibangkitkan menggunakan tiga model dasar spasial (SEM, SAR, GSM) dengan

berbagai tingkat autokorelasi spasial. Sedangkan data empiris meliputi konsentrasi PM_{2.5} diperoleh melalui stasiun pemantauan di Provinsi DKI Jakarta, peubah lingkungan dan meteorologi melalui *Google Earth Engine*, serta data lokasi jalan raya dan industri melalui *OpenStreetMap*.

Hasil simulasi menunjukkan bahwa metode seleksi berbasis SHAP secara konsisten memberikan performa terbaik dengan nilai akurasi pemodelan tertinggi dan distribusi nilai yang paling sempit di seluruh skenario autokorelasi spasial dan model pembangkitan data. Metode ini mampu memilih jumlah *eigenvector* yang lebih sedikit dan lebih efektif sehingga menghasilkan model yang lebih stabil dan akurat dibandingkan metode lain. Pada data empiris, penambahan *eigenvector* melalui ESF meningkatkan kinerja model GBM dalam memprediksi konsentrasi PM_{2.5}, dengan metode SHAP menunjukkan hasil terbaik pada skenario *10-fold cross-validation*. Namun, pada pemisahan data berbasis *spatial blocked cross-validation* yang lebih ketat secara spasial, peningkatan kinerja model menjadi kurang signifikan.

Penerapan evaluasi metode melalui data empiris pada kasus PM_{2.5} menunjukkan bahwa penambahan *eigenvector* secara signifikan meningkatkan kinerja model dibandingkan dengan model tanpa komponen spasial. Berdasarkan skenario *10-fold cross-validation*, metode seleksi berbasis SHAP menghasilkan akurasi prediksi tertinggi dengan nilai, serta mampu menangkap ketergantungan spasial dan hubungan nonlinier secara efektif. Metode SHAP menunjukkan tingkat ketahanan yang baik melalui pemilihan komponen spasial yang stabil dan konsisten pada berbagai wilayah. Temuan ini menunjukkan keunggulan metodologis dari integrasi ESF dengan pembelajaran mesin dan seleksi fitur berbasis SHAP sehingga menghasilkan kerangka pemodelan spasial yang lebih interpretatif dan *robust*.

Temuan baru dari penelitian ini adalah penerapan SHAP sebagai metode seleksi *eigenvector* yang tidak hanya meningkatkan akurasi prediksi tetapi juga memberikan interpretasi yang lebih jelas terhadap kontribusi pola spasial dalam model. Pendekatan ini mengatasi keterbatasan metode seleksi tradisional yang kurang eksplisit dalam menjelaskan peran masing-masing *eigenvector*. Implikasi penelitian ini mendukung pengembangan model prediksi kualitas udara yang lebih akurat dan interpretatif. Saran untuk penelitian selanjutnya adalah peningkatan kinerja model pada wilayah yang terpisah secara spasial dengan integrasi pemodelan spasiotemporal dan pengembangan matriks pembobot spasial yang lebih representatif. Temuan ini membuka peluang pengembangan lebih lanjut dengan menggabungkan pemodelan spasiotemporal dan penyempurnaan matriks pembobot spasial untuk meningkatkan representativitas dan kinerja model pada skala yang lebih luas. Implikasi praktis dari penelitian ini dapat digunakan pada pengembangan sistem peringatan dini kualitas udara di kawasan perkotaan yang memerlukan model prediksi yang akurat dan mudah diinterpretasikan untuk mendukung pengambilan keputusan pengendalian polusi udara.

Kata kunci: *Eigenvector Spatial Filtering*, Seleksi Vektor Ciri, *Gradient Boosting Machine*, Konsentrasi PM_{2.5}, *Shapley Additive exPlanations*.

SUMMARY

PUTRI NISRINA AZ-ZAHRA. Evaluation of Eigenvector Selection Method in Eigenvector Spatial Filtering with Boosting Model for PM_{2.5} Concentration Prediction. Supervised by ANIK DJURAI DAH dan ERFIANI.

Spatial autocorrelation is a phenomenon in which spatial dependence patterns indicate a systematic relationship between the value of a variable at a given location and the value of the same variable at neighboring locations. The presence of this characteristic in spatial data must be considered because it can introduce bias in statistical models if not properly addressed. One effective method for handling this issue is Eigenvector Spatial Filtering (ESF). ESF utilizes the eigenvectors of a spatial weights matrix to capture spatial dependence patterns in the data and incorporates these patterns into the model as additional independent variables, thereby improving the quality of predictive models.

In the implementation of ESF, selecting the optimal eigenvectors is an important aspect to be considered. This step is crucial in modeling to avoid overfitting, enhance model interpretability, and reduce computational burden. Although commonly used eigenvector selection methods are effective in partial selection, they do not yet provide explicit information regarding the individual contribution of each eigenvector to model prediction. Alternatively, the SHAP method may be applied to feature selection offering a robust basis in game theory. Eigenvectors with the highest SHAP values can be prioritized in ESF modeling, whereas those with low contributions can be eliminated.

The ESF method remains limited in capturing nonlinear relationships and complex interactions between variables. Machine learning algorithms have become effective approaches because of their ability to model complex and nonlinear patterns. Thus, the integration of ESF with machine learning algorithms has been implemented alongside the growing availability of large-scale geospatial data. In this study, we evaluate eigenvector selection methods in ESF with the integration of boosting models in the case of air quality, using PM_{2.5} as the indicator. Particulate Matter (PM_{2.5}) refers to fine particles with diameters less than 2.5 µm that are used as a primary indicator of air quality. Their extremely small size allows them to penetrate deep into the human respiratory tract and cause respiratory disorders, cardiovascular disease, and increased risk of premature death. The province of DKI Jakarta, as the capital of Indonesia, faces increasingly serious air pollution problems due to high population density and intense transportation usage, so that PM_{2.5} concentrations frequently exceed safe thresholds.

This study examines the evaluation of eigenvector selection methods in ESF combined with a machine learning model based on Gradient Boosting Machine (GBM) for predicting PM_{2.5} concentrations, with the aim of addressing spatial autocorrelation and improving the accuracy of spatial data predictive models. The eigenvector selection methods compared include the positive eigenvalue criterion, Moran's Index, LASSO, and SHAP (Shapley Additive exPlanations). Simulation data is generated using three basic spatial models (SEM, SAR, GSM) with various degrees of spatial autocorrelation, while empirical data covers PM_{2.5} concentrations obtained from monitoring stations in DKI Jakarta Province, environmental and

meteorological variables via Google Earth Engine, and road network and industry location data via OpenStreetMap.

Simulation results show that SHAP-based selection methods consistently deliver the best performance, with the highest modeling accuracy values and the narrowest value distributions across all spatial autocorrelation scenarios and data generation models. This method can select a smaller and more effective set of eigenvectors, resulting in more stable and accurate models compared to other methods. In empirical data, the addition of eigenvectors through ESF improves GBM model performance in predicting PM_{2.5} concentrations, with the SHAP method achieving the best results in the 10-fold cross-validation scenario. However, with spatially stricter data splitting using spatial blocked cross-validation, improvements in model performance have become less significant.

Application of the method evaluation using empirical data in the PM_{2.5} case demonstrates that the addition of eigenvectors significantly enhances model performance compared to models without spatial components. Based on the 10-fold cross-validation scenario, SHAP-based selection methods yield the highest prediction accuracy, as well as effectively capturing spatial dependence and nonlinear relationships. The SHAP method displays strong robustness through stable and consistent selection of spatial components across regions.

These findings demonstrate a methodological advantage from the integration of ESF with machine learning and SHAP-based feature selection, resulting in a more interpretable and robust spatial modeling framework. The novelty of this study is the application of SHAP as an eigenvector selection method, which not only improves prediction accuracy but also provides clearer interpretation of the contribution of spatial patterns in the model. This approach overcomes the limitations of traditional selection methods that are less explicit in explaining the role of each eigenvector. The implications of this research support the development of more accurate and interpretable air quality prediction models. Recommendations for future research include improving model performance in spatially separated regions by integrating spatiotemporal modeling and developing more representative spatial weights matrices. These findings open further opportunities for development by combining spatiotemporal modeling and refining the spatial weights matrix (SWM) to enhance representativeness and model performance on a larger scale. The practical implications of this research can be used in the development of early warning systems for urban air quality, which require accurate and easily interpreted predictive models to support decision-making in air pollution control.

Keywords: Eigenvector Spatial Filtering, Eigenvector Selection, Gradient Boosting Machine, PM_{2.5} Concentration, Shapley Additive exPlanations.



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.

**EVALUASI METODE SELEKSI *EIGENVECTOR* PADA
EIGENVECTOR SPATIAL FILTERING DENGAN MODEL
BOOSTING UNTUK PREDIKSI KONSENTRASI PM_{2.5}**

PUTRI NISRINA AZ-ZAHRA

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

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



@Hak cipta milik IPB University

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:
1 Dr. Yenni Angraini, M.Si



Judul Tesis : Evaluasi Metode Seleksi *Eigenvector* pada *Eigenvector Spatial Filtering* dengan Model *Boosting* untuk Prediksi Konsentrasi PM_{2.5}
Nama : Putri Nisrina Az-Zahra
NIM : M0501241050

Disetujui oleh

Pembimbing 1:
Prof. Dr. Ir. Anik Djuraidah, MS



Pembimbing 2:
Dr. Ir. Erfiani, M.Si.



Diketahui oleh

Ketua Program Studi:
Dr. Agus Mohamad Soleh, S.Si., M.T.
NIP N19750315 199903 1 004



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



Tanggal Ujian: 16 Juli 2026

Tanggal Lulus:



@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.

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Penelitian yang dilaksanakan sejak Bulan Februari 2025 sampai Bulan Juni 2026 ini disusun sebagai salah satu syarat untuk menyelesaikan studi Magister Sains pada Program Studi Statistika dan Sains Data. Proses penyusunan tesis ini tidak terlepas bantuan serta dukungan dari berbagai pihak, oleh karena itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu dan Bapak yang selalu mendoakan, memberi dukungan, memfasilitasi, dan memberi kasih sayang yang tidak pernah putus selama proses penulis menempuh studi pascasarjana.
2. Ibu Prof. Dr. Ir. Anik Djuraidah, MS dan Bunda Dr. Ir. Erfiani, M.Si. selaku komisi pembimbing yang telah memberikan arahan, masukan, motivasi, serta mendukung penuh proses penyelesaian tesis dari awal sampai akhir.
3. Ibu Dr. Yenni Angraini, M.Si. dan Ibu Sachnaz Desta Oktarnia S.Stat., M.Agr.Sc., Ph.D. selaku dosen penguji dan dosen moderator sidang ujian tesis yang telah memberikan evaluasi, kritik, serta saran yang membangun dalam proses penyempurnaan tesis.
4. Seluruh dosen Program Studi Statistika dan Sains Data IPB University yang telah memberikan ilmu yang bermanfaat serta pengalaman yang berharga selama penulis menempuh studi pascasarjana.
5. Aa, Hanifah, Mubdi, Kakek, Nenek, Aki, Umi, serta keluarga besar penulis yang turut memberikan doa, dukungan, serta kasih sayang sehingga memberikan semangat bagi penulis dalam menyelesaikan studi pascasarjana.
6. Sahabat selama proses perkuliahan, Kak Meavi, Abil, Zamrah, Aini, dan Windi yang selalu hadir untuk saling memberikan dukungan dan motivasi dalam menyelesaikan seluruh rangkaian studi pascasarjana.
7. Sahabat bimbingan, Nisa dan Nabil yang secara konsisten saling memberikan informasi, motivasi, dan aktif dalam bertukar pikiran untuk dapat menyelesaikan tesis dengan maksimal.
8. Sahabat penulis, Erja, Ghina, dan Nidya yang telah memberikan berbagai afirmasi, motivasi, dan turut serta mengikuti proses penulis selama studi pascasarjana dan seluruh rangkaian penyelesaian tesis.
9. Semua pihak yang telah mendukung penulis untuk dapat menjalankan proses studi pascasarjana dan melakukan penyusunan tesis dengan baik yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa penyusunan Karya Ilmiah ini masih jauh dari sempurna, oleh karena itu kritik dan saran yang membangun akan senantiasa penulis harapkan demi kesempurnaan di masa mendatang. Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, 1 Agustus 2026

Putri Nisrina Az-Zahra



@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.

DAFTAR ISI

DAFTAR ISI	xv
DAFTAR TABEL	xvi
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvii
I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
II. TINJAUAN PUSTAKA	4
2.1 Ketergantungan Spasial	4
2.2 Vektor Ciri Moran (<i>Moran Eigenvector</i>)	4
2.3 <i>Eigenvector Spatial Filtering</i> (ESF)	5
2.4 Metode Seleksi Vektor Ciri	6
2.5 <i>Gradient Boosting Machine</i> (GBM)	6
2.6 <i>Shapley Additive exPlanations</i> (SHAP)	7
2.7 Konsentrasi PM _{2.5}	8
2.8 Variabel Lingkungan dari Penginderaan Jauh dan Reanalisis	9
2.9 Variabel Sumber Emisi Polutan	11
III. KAJIAN SIMULASI	12
3.1 Data Simulasi	12
3.2 Tahapan Penelitian Data Simulasi	13
3.3 Hasil Kajian Simulasi	17
3.3.1 Pembangkitan Data	17
3.3.2 Analisis Komponen Spasial	18
3.3.3 Pemisahan Data	19
3.3.4 Seleksi <i>Eigenvector</i>	21
3.3.5 Evaluasi Model	24
IV. KAJIAN EMPIRIS	31
4.1 Data Empiris	31
4.3 Hasil Kajian Empiris	35
4.3.1 Eksplorasi Data	35

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.

4.3.2 Analisis Komponen Spasial	37
4.3.3 Seleksi <i>Eigenvector</i>	38
4.3.4 Evaluasi Model	42
V. SIMPULAN DAN SARAN	45
DAFTAR PUSTAKA	46
RIWAYAT HIDUP	49
LAMPIRAN	50

DAFTAR TABEL

1 Kategori AQI untuk PM _{2.5}	9
2 Performa model pada setiap tingkat autokorelasi spasial	27
3 Daftar peubah	31
4 Statistik deskriptif peubah respon dan peubah penjelas	36
5 Nilai VIF peubah penjelas	37
6 Uji aspek spasial	38
7 Hyperparameter pada GBM	39
8 Evaluasi performa model	43

DAFTAR GAMBAR

1 Visualisasi kontribusi fitur	8
2 Diagram alir kajian simulasi	16
3 Diagram pencar antara peubah respon dan peubah bebas (a) X1, (b) X2, (c) X1X2, (d) X12, dan (e) X22	17
4 Distribusi statistik LM berdasarkan (a) SEM, (b) SAR, (c) GSM (ρ), dan (d) GSM (λ)	18
5 Ilustrasi pemisahan data dengan metode (a) 10-fold cross-validation dan (b) spatial blocked cross-validation	19
6 Koefisien determinasi perbandingan metode validasi silang berdasarkan (a) SEM dan (b) SAR, dan (c) GSM	20
7 Rata-rata jumlah vektor ciri terpilih untuk setiap metode seleksi berdasarkan (a) SEM, (b) SAR, dan (c) GSM	21
8 Spektrum nilai eigen dengan vektor ciri terpilih untuk setiap metode seleksi (a) Nilai eigen Positif, (b) Indeks Moran, (c) LASSO, (d) SHAP – SEM, (e) SHAP – SAR, dan (f) SHAP – GSM	23

9	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan SEM berdasarkan (a) $\lambda = 0,8$ dan (b) $\lambda = -0,8$	24
10	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan SAR berdasarkan (a) $\rho = 0,8$ dan (b) $\rho = -0,8$	25
11	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan GSM berdasarkan (a) $\rho = 0,8$; $\lambda = 0,8$ dan (b) $\rho = -0,8$; $\lambda = -0,8$	26
12	Waktu komputasi pemodelan untuk setiap metode seleksi berdasarkan (a) SEM, (b) SAR, dan (c) GSM	29
13	Diagram pencar banyaknya vektor ciri terpilih dengan RMSE	30
14	Proses pengambilan data empiris	32
15	Diagram alir analisis data empiris	34
16	Sebaran $PM_{2.5}$ di DKI Jakarta Bulan Januari sampai Juni tahun 2025	35
17	Korelasi antara peubah respon dan peubah bebas	36
18	Ketertagaan pada titik SKPU Provinsi DKI Jakarta	37
19	Jumlah vektor ciri terpilih untuk setiap metode seleksi	39
20	Spektrum nilai eigen dengan vektor ciri terpilih untuk setiap metode seleksi (a) Nilai eigen Positif, (b) Indeks Moran, (c) LASSO, (d) SHAP – random CV, dan (e) SHAP – spatial CV	40
21	Kontribusi fitur berdasarkan SHAP pada model GBM dengan (a) 10-fold <i>cross-validation</i> dan (b) <i>spatial blocked cross-validation</i>	42
22	Nilai R^2 model GBM untuk setiap metode seleksi	43

DAFTAR LAMPIRAN

1	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan SEM berdasarkan (a) $\lambda = 0.5$, (b) $\lambda = -0.5$, (c) $\lambda = 0.2$, dan (d) $\lambda = -0.2$	50
2	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan SAR berdasarkan (a) $\rho = 0.5$, (b) $\rho = -0.5$, (c) $\rho = 0.2$, dan (d) $\rho = -0.2$	51
3	Nilai R^2 pada setiap metode seleksi pada pembangkitan data dengan GSM berdasarkan (a) $\rho = 0.5$; $\lambda = 0.5$, (b) $\rho = -0.5$; $\lambda = -0.5$, (c) $\rho = 0.2$; $\lambda = 0.2$, dan (d) $\rho = -0.2$; $\lambda = -0.2$	52

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.