

METODE GENERALIZED MIXED EFFECT SUPPORT VECTOR MACHINE (GMESVM) UNTUK MODEL MULTINOMIAL: PENGEMBANGAN DAN APLIKASINYA

HENGKI MURADI



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

@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.

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.



@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.

PERNYATAAN MENGENAI DISERTASI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa disertasi dengan judul “**Metode Generalized Mixed Effect Support Vector Machine (GMESVM) untuk Model Multinomial: Pengembangan dan Aplikasinya**” 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 disertasi ini.

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

Bogor, Maret 2025

Hengki Muradi
G161180031

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

IPB University





@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.

RINGKASAN

HENGKI MURADI. Metode Generalized Mixed Effect Support Vector Machine (GMESVM) untuk Model Multinomial: Pengembangan dan Aplikasinya. Dibimbing oleh ASEP SAEFUDDIN, I MADE SUMERTAJAYA, AGUS MOHAMAD SOLEH dan DEDE DIRGAHAYU DOMIRI.

Support Vector Machine (SVM) adalah metode pembelajaran mesin yang dapat digunakan untuk klasifikasi dan regresi. Metode ini dikenal mampu menangani hubungan nonlinier antar peubah, robust terhadap multikolinieritas dan autokorelasi, serta bebas overfitting. Penelitian ini memiliki tiga tujuan utama: (1) mengevaluasi kinerja SVM dalam kasus regresi dengan peubah respon kontinu yang disebut dengan *Support Vector Regression* (SVR), (2) mengevaluasi kinerja SVM dalam kasus klasifikasi dengan peubah respon multinomial, dan (3) mengembangkan serta mengevaluasi metode *Generalized Mixed Effect Support Vector Machine* (GMESVM).

Kajian pertama bertujuan untuk menilai linieritas hubungan antara Peubah prediktor dan peubah respon kontinu, serta menguji kemampuan SVR dalam menangani hubungan nonlinier. Data yang digunakan berupa data citra Sentinel-1 yang meliputi polarisasi *vertical-vertical* (VV) dan *vertical-horizontal* (VH) dan indeks polarisasi; *ratio polarization index* (RPI), *normalized different polarization index* (NDPI), dan *average polarization index* (API), serta umur padi pada blok sawah di PT. Sang Hyang Seri (SHS) Subang pada musim tanam pertama tahun 2022. Data teracak di bagi menjadi data training dan data testing dengan komposisi 70:30 dan diulang sebanyak 10 kali. Metode SVR dievaluasi dengan beberapa tipe kernel yaitu linier, polinomial, dan radial basis function (RBF) dan dibandingkan dengan metode regresi linier (LM). Model umur padi dari metode SVR lebih baik dibandingkan dengan model umur padi dari metode LR, dimana metode SVR menghasilkan model umur padi terbaik pada kondisi 4 (empat) prediktor yaitu VH, RPI, NDPI, dan API dengan rata-rata RMSE sebesar 11,03 dan rata-rata koefisien determinasi (adjust) sebesar 89,69%. Hasil penelitian menunjukkan bahwa SVR dengan kernel RBF memberikan performa terbaik dalam menangani pola hubungan nonlinier antara peubah prediktor dengan peubah respon.

Kajian kedua bertujuan untuk mengevaluasi metode klasifikasi multiklas SVM One-versus-One (OvO) dan Generalized SVM (GenSVM) dengan berbagai tipe kernel dan parameter. Data simulasi merupakan data glass yang diperoleh dari package kkn R sedangkan data empirik menggunakan data citra Sentinel-1 dan fase padi. Akurasi SVM OvO mencapai 100% ketika $\text{Cost} = 2^{14}$ dan akurasi model juga meningkat seiring kenaikan gamma dan mencapai optimal saat $\gamma = 2^4$. Kinerja GenSVM mencapai optimal saat parameter κ 2,5 atau 4,0, dan semakin optimal ketika parameter λ semakin kecil, mencapai puncaknya pada $\lambda = 2^{-16}$. Selain itu, GenSVM juga menunjukkan kinerja terbaik saat parameter p disetel ke 1,7. Dengan demikian, kinerja SVM OvO dan GenSVM sangat bergantung pada pengaturan parameter namun terdapat resiko overfitting ketika parameter diatur terlalu ekstrem sehingga diperlukan tuning parameter. Pada studi empiris, metode SVM OvO dan GenSVM dievaluasi dan dibandingkan dengan metode *multinomial logistic regression* (MLR). Hasil menunjukkan bahwa SVM OvO memberikan akurasi tertinggi pada skenario empat kelas dengan akurasi $79,20 \pm 0,21\%$. Prediksi model menunjukkan kesalahan terutama pada fase air dan bera, akibat jumlah contoh yang tidak seimbang dan faktor-faktor eksternal.

Kajian ketiga merupakan kajian utama dalam penelitian ini. Dalam kajian ini dikembangkan metode GMESVM untuk klasifikasi multinomial dengan menggabungkan SVM OvO dan *Generalized Linear Mixed Model* (GLMM) melalui pendekatan *combination forecasting* dengan pembobotan mean aritmatik, dan metode *variance-covariance* (vaco). Data simulasi dibangkitkan dengan mempertimbangkan faktor jumlah kelas peubah respon, jumlah Peubah prediktor, ukuran contoh, dan level korelasi antar peubah prediktor. Jumlah kelas peubah respon ditetapkan sebanyak 6 kelas, peubah prediktor terdiri dari 3 (tiga) peubah efek tetap dan 1 (satu) peubah efek acak, ukuran contoh terdiri dari $n = 100$, $n = 400$, dan $n = 1000$, dan level korelasi diatur dengan ketentuan; Lemah (L) yang berarti korelasi kurang dari 0,4; Moderate (M) yang berarti korelasi 0,41-0,60; Kuat (K) yang berarti korelasi 0,61-0,80, dan Sangat Kuat (SK) yang berarti korelasi 0,81-1,00. Skenario level korelasi disusun menjadi 8 (delapan) skenario yaitu; L,L,L; L,L,M; L,L,K; L,L,SK; M,M,M; M,M,K; K,K,K; dan K,K,SK. Skenario L,LM bermakna $r_{x_1,x_2} = L$, $r_{x_1,x_3} = L$ dan $r_{x_2,x_3} = M$.

Hasil simulasi menunjukkan bahwa peningkatan ukuran contoh berpengaruh positif terhadap kinerja metode SVM, di mana pada contoh $n = 100$, kinerja optimal tercapai pada skenario korelasi L,L,L hingga L,L,SK. Untuk ukuran contoh $n = 400$, SVM dapat berfungsi optimal di hampir semua skenario kecuali K,K,K, sedangkan pada $n = 1000$, kinerja terbaik ditemukan pada L,L,L dan K,K,K, meskipun menurun pada K,K,SK; model SVM tetap konsisten tanpa masalah overfitting. Sebaliknya, GLMM pada $n = 100$ tidak mencapai kinerja optimal, tetapi saat ukuran contoh meningkat menjadi $n = 400$, akurasi GLMM meningkat signifikan, mencapai di atas 90%, dengan kinerja terbaik pada skenario L,L,SK, meski terdeteksi masalah overfitting. Ketika ukuran contoh ditingkatkan lagi hingga $n = 1000$, GLMM menunjukkan kinerja yang sangat baik dan konsisten, serta mengurangi risiko overfitting.

Hasil simulasi metode GMESVM menunjukkan bahwa metode ini lebih efektif ketika dibangun dari kombinasi prediksi peluang yang diperoleh dari SVM dan GLMM, serta dipengaruhi oleh ukuran contoh dan level korelasi Peubah prediktor. Pada ukuran contoh 100, GMESVM belum optimal tetapi mampu memperbaiki kesalahan prediksi SVM dan GLMM, dengan akurasi lebih dari 90% pada data training dan lebih dari 83% pada data testing. Ketika ukuran contoh meningkat menjadi 400 dan 1000, kinerja GMESVM meningkat signifikan, konsisten di atas 90% pada semua skenario, dan lebih baik dibandingkan dengan SVM dan GLMM karena dapat mengakomodasi hubungan non-linear dan efek lainnya. Model GMESVM terbaik diperoleh melalui pemilihan metode kombinasi yang optimal dengan cara ensemble.

Pada studi empiris, GMESVM meningkatkan akurasi model fase pertumbuhan padi, dengan akurasi tertinggi pada GMESVM Mean Aritmatik rata-rata 80,95%, lebih baik dari SVM (80,38%) dan GLMM (79,01%). Kendala yang dihadapi adalah masih adanya kesalahan klasifikasi hingga 20%, terutama karena model belum mengakomodasi tren indeks polarisasi secara penuh. Kombinasi SVM dan GLMM dengan pendekatan kombinasi memberikan akurasi model yang lebih tinggi dan sangat potensial untuk aplikasi klasifikasi. Namun, penelitian lanjutan masih diperlukan, terutama untuk mengintegrasikan efek waktu dan spasial dalam model GMESVM guna meningkatkan kinerjanya.

Kata Kunci : SVM, GLMM, GMESVM, Multinomial, Kombinasi, Peramalan

SUMMARY

HENGKI MURADI. Generalized Mixed Effect Support Vector Machine (GMESVM) Method for Multinomial Models: Development and Application. Supervised by ASEP SAEFUDDIN, I MADE SUMERTAJAYA, AGUS MOHAMAD SOLEH and DEDE DIRGAHAYU DOMIRI.

Support Vector Machine (SVM) is a machine learning method that can be used for classification and regression. This method is known to be able to handle nonlinear relationships between variables, robust to multicollinearity and autocorrelation, and free of overfitting. This study has three main objectives: (1) evaluate the performance of SVM in the case of regression with continuous response variables called Support Vector Regression (SVR), (2) evaluate the performance of SVM in the case of classification with multinomial response variables, and (3) develop and evaluate the Generalized Mixed Effect Support Vector Machine (GMESVM) method.

The first study aims to assess the linearity of the relationship between predictor variables and continuous response variables, and test the ability of SVR to handle nonlinear relationships. The data used is Sentinel-1 image data which includes vertical-vertical (VV) and vertical-horizontal (VH) polarisation and polarisation indices; ratio polarization index (RPI), normalised different polarization index (NDPI), and average polarization index (API), as well as the age of rice in the rice paddy block at PT Sang Hyang Seri (SHS) Subang in the first planting season of 2022. The data were divided into training data and testing data with a composition of 70:30 and repeated 10 times. The SVR method was evaluated with several kernel types namely linear, polynomial, and radial basis function (RBF) and compared with the linear regression (LM) method. The rice paddy age model from the SVR method is better than the rice paddy age model from the LR method, where the SVR method produces the best rice paddy age model in the condition of 4 (four) predictors namely VH, RPI, NDPI, and API with an average RMSE of 11.03 and an average coefficient of determination (adjust) of 89.69%. The results showed that SVR with RBF kernel gave the best performance in handling nonlinear relationship patterns between variables.

The second study aims to evaluate the SVM One-versus-One (OvO) and Generalized SVM (GenSVM) multiclass classification methods with various kernel types and parameters. Simulated data is glass data obtained from kkn R package while empirical data uses Sentinel-1 image data and rice phase. The accuracy of SVM OvO reached 100% when $\text{Cost} = 2^{14}$ and the accuracy of the model also increased as gamma increased and reached the optimum when $\text{gamma} = 2^4$. The performance of GenSVM reached the optimum when the kappa parameter was 2.5 or 4.0, and became more optimal when the lambda parameter was smaller, reaching its peak at $\lambda = 2^{(-16)}$. In addition, GenSVM also shows the best performance when the p parameter is set to 1.7. Thus, the performance of SVM OvO and GenSVM is highly dependent on the parameter settings but there is a risk of overfitting when the parameters are set too extreme so parameter tuning is required. In the empirical study, SVM OvO and GenSVM methods were evaluated and compared with multinomial logistic regression (MLR). Results showed that SVM OvO gave the highest accuracy in the four-class scenario with $79.20 \pm 0.21\%$ accuracy. Model predictions showed errors especially in the water and fallow phases, due to unbalanced sample size and external factors.

The third study is the main study in this research. In this study, the GMESVM method for multinomial classification is developed by combining SVM OvO and Generalized Linear Mixed Model (GLMM) through a combination forecasting approach with arithmetic mean weighting, geometric mean, and variance-covariance method (vaco). Simulation data is generated by considering the number of response variable classes, the number of estimating variables, sample size, and the level of correlation between estimating variables. The number of response variable classes was set at 6 classes, the estimating variables consisted of 3 (three) fixed effect variables and 1 (one) random effect variable, the sample size consisted of $n = 100$, $n = 400$, and $n = 1000$, and the correlation level was set with the following provisions; Weak (L) which means a correlation of less than 0.4; Moderate (M) which means a correlation of 0.41-0.60; Strong (K) which means a correlation of 0.61-0.80, and Very Strong (SK) which means a correlation of 0.81-1.00. The correlation level scenario is organized into 8 (eight) scenarios, namely; L,L,L; L,L,M; L,L,K; L,L,SK; M,M,M; M,M,K; K,K,K; and K,K,SK. The scenario L,LM means $r_{x_1,x_2} = L$, $r_{x_1,x_3} = L$ and $r_{x_2,x_3} = M$.

The simulation results show that increasing the sample size has a positive effect on the performance of the SVM method, where at sample $n = 100$, optimal performance is achieved in the L,L,L to L,L,SK correlation scenarios. For sample size $n = 400$, SVM can function optimally in almost all scenarios except K,K,K, while at $n = 1000$, the best performance is found in L,L,L and K,K,K, although it decreases in K,K,SK; the SVM model remains consistent without overfitting issues. In contrast, the GLMM at $n = 100$ did not achieve optimal performance, but as the sample size increased to $n = 400$, the accuracy of the GLMM improved significantly, reaching above 90%, with the best performance in the L,L,SK scenario, although an overfitting problem was detected. When the sample size was further increased to $n = 1000$, the GLMM showed excellent and consistent performance, and reduced the risk of overfitting.

The simulation results of the GMESVM method show that it is more effective when built from a combination of probability predictions obtained from SVM and GLMM, and is influenced by the sample size and the correlation level of the estimating variables. At a sample size of 100, GMESVM is not optimal but is able to correct the prediction errors of SVM and GLMM, with an accuracy of more than 90% on training data and more than 83% on testing data. When the sample size increases to 400 and 1000, the performance of GMESVM improves significantly, consistently above 90% in all scenarios, and compares favorably with SVM and GLMM because it can accommodate non-linear relationships and other effects. The best GMESVM model is obtained through selecting the optimal combination method in an ensemble manner.

In the empirical study, GMESVM improved the accuracy of the rice growth phase model, with the highest accuracy in GMESVM Mean Arithmetic averaging 80.95%, better than SVM (80.38%) and GLMM (79.01%). The obstacle is that there is still a misclassification of up to 20%, mainly because the model does not fully accommodate the trend of the polarization index. The combination of SVM and GLMM with the combination approach provides higher model accuracy and has great potential for classification applications. However, further research is still needed, especially to integrate time and spatial effects in the GMESVM model to improve its performance.

Keywords: SVM, GLMM, GMESVM, Multinomial, Combination, Forecasting



METODE GENERALIZED MIXED EFFECT SUPPORT VECTOR MACHINE (GMESVM) UNTUK MODEL MULTINOMIAL: PENGEMBANGAN DAN APLIKASINYA

HENGKI MURADI

Disertasi
sebagai salah satu syarat untuk memperoleh gelar
Doktor
Pada Program Studi Statistika dan Sains Data
© Hak Cipta milik IPB, tahun 2025¹
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.

**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.

Penguji Luar Komisi Pembimbing pada Ujian Tertutup Disertasi:

1. Dr. Bagus Sartono, S.Si., M.Si. 
2. Dr. Budi Susetyo, S.Si., M.Si



Promotor Luar Komisi Pembimbing pada Sidang Promosi Terbuka Disertasi:

1. Dr. Budi Susetyo, S.Si., M.Si
2. Dr. Esa Prakasa, M.T. 



Judul Disertasi: METODE GENERALIZED MIXED EFFECT SUPPORT VECTOR MACHINE (GMESVM) UNTUK MODEL MULTINOMIAL: PENGEMBANGAN DAN APLIKASINYA (Generalized Mixed Effect Support Vector Machine (Gmesvm) Method For Multinomial Models: Development And Application)

Nama : Hengki Muradi
NIM : G161180031

Disetujui oleh

Pembimbing 1
Prof. Dr. Ir. Asep Saefuddin, M.Sc



Pembimbing 2
Prof. Dr. Ir. I Made Sumartejaya, M.Si



Pembimbing 3
Dr. Agus Mohamad Soleh, S.Si., MT



Pembimbing 4
Dr. Dede Dirgahayu Domiri, M.Si



Diketahui oleh

Ketua Program Studi:
Dr. Kusman Sadik, M.Si
NIP. 196909121997021001



Dekan Sekolah Sains Data, Matematika,
Dan Informatika:
Prof. Dr. Ir. Agus Bueno, M.Si., M.Kom.
NIP. 197807232007011001



Tanggal Ujian:
(07 Desember 2024)

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala nikmat sehat dan kesempatan sehingga karya ilmiah ini berhasil diselesaikan. Karya ilmiah ini diberi judul “Metode Generalized Mixed Effect Support Vector Machine (GMESVM) Untuk Model Multinomial: Pengembangan Dan Aplikasinya”. Karya ilmiah ini penulis susun sebagai salah satu syarat untuk mendapatkan gelar Doktor Statistika dan Sains Data pada Departemen Statistika, Sekolah Sains Data, Matematika, dan Informatika, Institut Pertanian Bogor.

Penulisan karya ilmiah ini dapat diselesaikan berkat bimbingan, masukan, saran dan bimbingan dari banyak pihak. Oleh karena itu, penulis mengucapkan rasa terima kasih yang sedalam-dalamnya kepada semua pihak yang telah terlibat dalam penyusunan karya ilmiah ini, antara lain:

1. Komisi pembimbing yang telah banyak memberikan bimbingan, arahan dan dorongan semangat kepada penulis dengan penuh kesabaran: Bapak Prof. Dr. Ir. Asep Saefuddin, M.Sc., Bapak Prof. Dr. Ir. I Made Sumertajaya, M.Si., Bapak Dr. Agus Mohamad Soleh, S.Si, MT., dan Bapak Dr. Dede Dirgahayu Domiri, M.Si.
2. Penguji luar komisi ujian yaitu Dr. Budi Susetyo, S.Si, M.Si., Dr. Bagus Sartono, M.Si, dan Dr. Esa Prakasa, M.T yang telah memberikan masukan sehingga karya ilmiah ini dapat lebih baik lagi.
3. Jajaran pimpinan Pusat Riset Penginderaan Jauh dan Pusat Riset Sains Data dan Informasi BRIN, yang telah memberikan kesempatan kepada penulis untuk melanjutkan pendidikan S3 di Institut Pertanian Bogor (IPB).
4. Sekolah Pascasarja IPB beserta jajarannya yang telah membantu memfasilitasi studi penulis.
5. Ketua Program Studi serta Bapak/Ibu Dosen Departemen Statistika dan Sains Data yang telah banyak memberikan ilmu, pencerahan, dan arahan selama masa perkuliahan.
6. Sekretariat Program Studi Statistika dan Sains Data IPB yang dengan sabar selalu yang membantu segala proses administrasi.
7. Mak dan Bapak terkasih yang selalu menjadi psikolog terbaik dalam hidup penulis.
8. Istri tercinta Karina Eka Damayanti dan anak-anak terkasih Anindya Clarissa Muradi & Akhtar Laksana Muradi yang selalu setia mendampingi setiap langkah perjuangan ayah dan bunda.
9. Keluarga besar penulis; Uti, Kakung, Mama dan Papa, Ayuk Meli & Kak Masyur, Cik Hen & Cik Lia, Cikti dan om Wisnu, eri & anton, yogi & lilin, nanda, cinta dan seluruh keponakanku tersayang (Fadhil, Zura, Adiba, Fawas, Aidan) serta keluarga besar yang senantiasa memotivasi, mendukung dan mendo'akan kelancaran studi penulis.
10. Teman-teman S3 Statistika IPB Angkatan 2018 dan teman-teman S3 lainnya yang selalu memberikan semangat kepada penulis.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan dan mohon maaf atas segala kekurangan dan kesalahan yang terdapat dalam karya ilmiah ini.

Bogor, Maret 2025
Hengki Muradi

DAFTAR ISI

DAFTAR TABEL	xvi
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvii
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan	5
1.4 Manfaat	5
1.5 Ruang Lingkup	5
1.6 Kebaruan (<i>novelty</i>)	6
1.7 Sistematika Disertasi	8
II TINJAUAN PUSTAKA	9
2.1 Distribusi Multinomial	9
2.2 Support Vector Machine (SVM)	11
2.3 <i>Mixed Effect Support Vector Machine</i>	17
2.4 <i>Generalized Linear Mixed Effect Model</i> (GLMM)	18
2.5 Akurasi Model	22
2.6 Penggunaan Machine Learning Untuk Pemodelan Fase Pertumbuhan Padi	22
III SUPPORT VECTOR REGRESSION (SVR) UNTUK PEMODELAN FASE PERTUMBUHAN PADI MENGGUNAKAN DATA SATELIT SENTINEL-	
1	25
3.1 Pendahuluan	25
3.2 Regresi Linear	26
3.3 Support Vector Regression (SVR)	26
3.4 Metode	28
3.5 Hasil Penelitian	30
IV ANALISIS KINERJA SUPPORT VECTOR MACHINES: SIMULASI DAN APLIKASI.	35
4.1 Pendahuluan	35
4.2 Support Vector Machine (SVM)	36
4.3 Multinomial Logistic Regression (MLR)	38
4.4 Akurasi Model	38
4.5 Metode	39
V GENERALIZED MIXED EFFECT SUPPORT VECTOR MACHINE (GMESVM) UNTUK PEUBAH RESPON MULTINOMIAL	48
5.1 Pendahuluan	48
5.2 Metode Support Vector Machine One vs One (SVM OvO)	49
5.3 Metode <i>Generalized Linear Mixed Effect Model</i> (GLMM)	50
5.4 Metode <i>Generalised Mixed Effect Support Vector Machine</i>	50
5.4.1. Skenario Pemodelan GMESVM	53
5.5 Metode	58
5.6 Pembahasan	79

5.7 Kesimpulan	82
VI PEMBAHASAN UMUM	83
VII SIMPULAN UMUM DAN SARAN	88
7.1. Simpulan	88
7.2 Saran	88
DAFTAR PUSTAKA	89
Lampiran 1 Script GEMSVM	94
Lampiran 2 : Hasil Pemodelan	111
RIWAYAT HIDUP	118



DAFTAR TABEL

Tabel 1. Jenis Kernel	11
Tabel 2. Karakteristik GLMM	18
Tabel 3. Data Satelit Untuk Pemodelan Fase Pertumbuhan Padi	22
Tabel 4. Indeks Vegetasi Dari Data Satelit	23
Tabel 5. Tipe Kernel SVM	27
Tabel 6. Statistik Deskriptif Polarisasi, Indeks Polarisasi, dan Umur Padi	30
Tabel 7. Akurasi Model SVR dan LR	32
Tabel 8. Skenario Pemodelan Fase Pertumbuhan Padi	44
Tabel 9. Hasil Pemodelan Fase Pertumbuhan Padi	46
Tabel 10. Kinerja Model SVM, GLMM, GMESVM Mean Aritmatik (MA), GMESVM Mean Geometrik (MG), dan GMESVM Vaco n=100	62
Tabel 11. Uji Anova n = 100	64
Tabel 12. Uji Beda Akurasi Model pada Data Training dan Data Testing	65
Tabel 13. Kinerja Model SVM, GLMM, GMESVM Mean Aritmatik, GMESVM Mean Geometrik, dan GMESVM Vaco n=400	66
Tabel 14. Uji Anova n = 400	68
Tabel 15. Uji Beda Akurasi Model pada Data Training dan Data Testing	68
Tabel 16. Kinerja Model SVM, GLMM, GMESVM Mean Aritmatik, GMESVM Mean Geometrik, dan GMESVM Vaco n=1000	69
Tabel 17. Uji Anova n = 400	67
Tabel 18. Uji Beda Akurasi Model pada Data Training dan Data Testing	72
Tabel 19. Definisi Operasional	74
Tabel 20. Statistik Deskriptif Indeks Polarisasi di Setiap Musim Tanam dan Fase Pertumbuhan Padi	76
Tabel 21. Korelasi Antar Peubah prediktor Berdasarkan Musim Tanam	77
Tabel 22. Hasil Implementasi Metode SVM, GLMM, dan GMESVM pada Model Fase Pertumbuhan Padi	78
Tabel 23. Uji Anova Akurasi Model Fase Pertumbuhan Padi	78
Tabel 24. Hasil Uji Beda Akurasi Metode SVM, GLMM, dan GMESVM pada Model Fase Pertumbuhan Padi	79

DAFTAR GAMBAR

Gambar 1. Peta Jalan Penelitian	7
Gambar 2. Hyperplane, Support Vector, dan Maximum Margin	12
Gambar 3. Loss Function	12
Gambar 4. Ilustrasi GenSVM 2 dimensi dan $K = 3$	16
Gambar 5. Ilustrasi q_i^{kj} pada kondisi $K = 3$ dan $m = 2$	16
Gambar 6. Area Penelitian	28
Gambar 7. Langkah-Langkah Analisis Data	29
Gambar 8. Scatterplot Antara Polarisasi/Indeks Polarisasi Terhadap Umur Padi	30
Gambar 9. Akurasi Model LM dan SVR; (a) RMSE, (b). Adjust R^2	33



Gambar 10. Langkah-Langkah Simulasi SVM, GenSVM, dan MLR	39
Gambar 11. Hasil Simulasi Metode SVM OvO: Parameter <i>Cost</i>	40
Gambar 12. Hasil Simulasi Metode SVM OvO: Parameter γ	40
Gambar 13. Hasil Simulasi Metode GenSVM: Parameter Kappa	40
Gambar 14. Hasil Simulasi Metode GenSVM: Parameter Lambda	41
Gambar 15. Hasil Simulasi Metode GenSVM: parameter p	41
Gambar 16 Langkah-Langkah Studi Empirik Metode SVM, GenSVM, dan MLR	42
Gambar 17. Area Penelitian	43
Gambar 18 (a).Boxplot of VV, VH, and RPI, (b). Boxplot of VV Based On Paddy Phase, (c). Boxplot of VH Based on Paddy Phase, and (d). Boxplot of RPI Based on Paddy Phase	45
Gambar 19 (a).Overall Accuration (OA) dan Kappa untuk 7-kelas, (b). Overall Accuration (OA) dan Kappa 6-kelas, (c). Overall Accuration (OA) dan Kappa 5-kelas, dan (d). Overall Accuration (OA) dan Kappa 4-kelas.	47
Gambar 20. Langkah-Langkah Simulasi	58
Gambar 21. Langkah-Langkah Membangkitkan Data	60
Gambar 22. Akurasi Model SVM, GLMM, dan GMESVM saat $n = 100$	63
Gambar 23. Boxplot Akurasi Model pada Skenario $n = 100$	64
Gambar 24. Akurasi Model SVM, GLMM, dan GMESVM saat $n = 400$	67
Gambar 25. Boxplot Akurasi Model pada Skenario $n = 400$	67
Gambar 26. Akurasi Model SVM, GLMM, dan GMESVM saat $n = 1000$	70
Gambar 27. Kinerja SVM pada Skenario $n = 1000$	71
Gambar 28. Langkah-langkah Studi Empirik	73
Gambar 29. Area Penelitian	73
Gambar 30. Fase Pertumbuhan Padi dan Indeks Polarisasi citra Satelit Sentinel-1 (RPI)	74
Gambar 31. Pola Hubungan Umur Padi dengan Indeks Polarisasi; (a). RPI, (b). NDPI, (c). API.	75
Gambar 32. Boxplot Akurasi Model Fase Pertumbuhan Padi	77

DAFTAR LAMPIRAN

Lampiran 1. Script GEMSVM	93
Lampiran 2. Hasil Pemodelan	110
Riwayat Hidup Penulis	117



@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.