

**ANALISIS REGRESI LOGISTIK BINER MENGGUNAKAN
METODE *SPARSE-GROUP* LASSO PADA DATA BERDIMENSI TINGGI
(Studi Kasus: Faktor-Faktor yang Memengaruhi Capaian
Indeks Pembangunan Manusia Kota/Kabupaten di Jawa Barat Tahun 2022)**



SKRIPSI

Diajukan untuk memenuhi sebagian syarat memperoleh
gelar Sarjana Matematika

Oleh:

HASTIALISNA HURUL AENI SETIAWAN

NIM 2102839

**PROGRAM STUDI MATEMATIKA
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2025**

LEMBAR HAK CIPTA

**ANALISIS REGRESI LOGISTIK BINER MENGGUNAKAN
METODE *SPARSE-GROUP* LASSO PADA DATA BERDIMENSI TINGGI
(Studi Kasus: Faktor-Faktor yang Memengaruhi Capaian
Indeks Pembangunan Manusia Kota/Kabupaten di Jawa Barat Tahun 2022)**

Oleh
Hastialisna Hurul Aeni Setiawan

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Matematika pada
Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

© Hastialisna Hurul Aeni Setiawan 2025

Universitas Pendidikan Indonesia

April 2025

Hak Cipta dilindungi undang-undang.

Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
dengan dicetak ulang, difotokopi, atau cara lainnya tanpa izin dari penulis.

LEMBAR PENGESAHAN

**ANALISIS REGRESI LOGISTIK BINER MENGGUNAKAN
METODE *SPARSE-GROUP* LASSO PADA DATA BERDIMENSI TINGGI
(Studi Kasus: Faktor-Faktor yang Memengaruhi Capaian
Indeks Pembangunan Manusia Kota/Kabupaten di Jawa Barat Tahun 2022)**

Hastialisna Hurul Aeni Setiawan
NIM. 2102839

Disetujui dan disahkan oleh:

Pembimbing I,



Drs. Nar Herrhyanto, M.Pd.
NIP. 196106181987031001

Pembimbing II,



Dr. Lukman, S.Si., M.Si.
NIP. 196801281994021001

Mengetahui,

Ketua Program Studi Matematika



Dr. Kartika Yulianti, M.Si.
NIP. 198207282005012001

**ANALISIS REGRESI LOGISTIK BINER MENGGUNAKAN
METODE *SPARSE-GROUP* LASSO PADA DATA BERDIMENSI TINGGI
(Studi Kasus: Faktor-Faktor yang Memengaruhi Capaian
Indeks Pembangunan Manusia Kota/Kabupaten di Jawa Barat Tahun 2022)**

ABSTRAK

Disusun oleh:

Hastialisna Hurul Aeni Setiawan – hastialisnahas@upi.edu

Data berdimensi tinggi, di mana jumlah variabel bebas lebih dari jumlah amatan, kerap menimbulkan permasalahan dalam pemodelan statistik, seperti multikolinearitas, khususnya saat menggunakan metode regresi biasa. Untuk mengatasi hal tersebut, digunakan metode regularisasi, salah satunya adalah *Sparse-Group* LASSO, yaitu pengembangan dari metode LASSO yang mengombinasikan pendekatan LASSO dan *Group* LASSO. Metode ini memungkinkan seleksi variabel secara individu maupun kelompok, dan sesuai diterapkan pada data berdimensi tinggi yang memiliki struktur kelompok antarvariabel. Penelitian ini bertujuan menerapkan regresi logistik biner dengan *Sparse-Group* LASSO pada data capaian IPM 27 kota/kabupaten di Provinsi Jawa Barat tahun 2022, guna mengidentifikasi faktor-faktor yang memengaruhi capaian IPM. Variabel bebas terdiri atas 60 faktor dalam 7 kelompok, meliputi pendidikan, kesehatan, ekonomi, lingkungan, kependudukan, penyandang masalah kesejahteraan sosial (PMKS), dan tata kelola pemerintahan. Model terbaik diperoleh berdasarkan nilai λ optimal dari *k-fold cross-validation* dengan $\alpha = 0,5$, yaitu sebesar 0,0059954. Model ini menghasilkan 23 variabel terpilih dalam 3 kelompok utama: pendidikan, lingkungan, dan PMKS. Variabel dari kelompok pendidikan mencakup angka partisipasi murni, harapan lama sekolah, jumlah guru, jumlah murid, rata-rata lama sekolah, angka melek huruf, dan indeks literasi. Dari kelompok lingkungan, terpilih antara lain luas wilayah, luas hutan, jumlah kecamatan, jumlah kelurahan, tingkat pelayanan sampah, indeks kualitas lingkungan hidup, sanitasi layak, dan akses terhadap sumber air minum layak. Pada kelompok PMKS, faktor seperti jumlah anak terlantar, anak jalanan, gelandangan, pengemis, korban NAPZA, lansia terlantar, korban kekerasan anak, dan jumlah perceraian, turut berpengaruh. Model yang dihasilkan cukup baik dengan akurasi klasifikasi sebesar 96,3% dan nilai AUC mencapai 0,994.

Kata Kunci: Data Berdimensi Tinggi, Regresi Logistik Biner, *Sparse-Group* LASSO, Multikolinearitas, Indeks Pembangunan Manusia.

**BINARY LOGISTIC REGRESSION ANALYSIS USING THE
SPARSE-GROUP LASSO METHOD ON HIGH-DIMENSIONAL DATA
(A Case Study on Factors Affecting the Human Development Index
Achievement of Cities and Regencies in West Java, 2022)**

ABSTRACT

Written by:

Hastialisna Hurul Aeni Setiawan – hastialisnahas@upi.edu

High-dimensional data, where the number of independent variables is more than the number of observations, often causes problems in statistical modeling, such as multicollinearity, especially when using ordinary regression methods. To overcome this, regularization methods are used, one of which is Sparse-Group LASSO, which is a development of the LASSO method that combines the LASSO and Group LASSO approaches. This method allows individual or group selection of variables, and is suitable for high-dimensional data that has a group structure between variables. This study aims to apply binary logistic regression with Sparse-Group LASSO on HDI achievement data of 27 cities/districts in West Java Province in 2022, to identify factors that affect HDI achievement. The independent variables consist of 60 factors in 7 groups, including education, health, economy, environment, population, people with social welfare problems (PMKS), and governance. The best model was obtained based on the optimal λ value from k-fold cross-validation with $\alpha = 0.5$, which was 0.0059954. The model produced 23 selected variables in 3 main groups: education, environment, and PMKS. Variables from the education group include net enrollment rate, expected years of schooling, number of teachers, number of students, average years of schooling, literacy rate, and literacy index. From the environment group, variables such as area, forest area, number of sub-districts, number of villages, level of waste services, environmental quality index, proper sanitation, and access to safe drinking water sources were selected. In the PMKS group, factors such as the number of abandoned children, street children, homeless people, beggars, drug victims, abandoned elderly people, victims of child abuse, and the number of divorces, were also influential. The resulting model is quite good with a classification accuracy of 96.3% and an AUC value of 0.994.

Keywords: High-Dimensional Data, Binary Logistic Regression, Sparse-Group LASSO, Multicollinearity, Human Development Index.

DAFTAR ISI

LEMBAR HAK CIPTA	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
 BAB 1 PENDAHULUAN	 1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	6
 BAB 2 TINJAUAN PUSTAKA	 7
2.1 Data Berdimensi Tinggi	7
2.2 Regresi Logistik	7
2.3 Asumsi Regresi Logistik	8
2.4 Uji Multikolinearitas	9
2.5 Regresi Logistik Biner	10
2.6 Penaksir Parameter Regresi Logistik Biner.....	11
2.7 Interpretasi Parameter Regresi Logistik Biner	13
2.8 LASSO dan <i>Group</i> LASSO	14
2.9 <i>Sparse-Group</i> LASSO.....	17
2.10 Algoritma <i>Block Coordinate Descent</i>	18
2.11 <i>K-Fold Cross-Validation</i>	19

2.12 Ketepatan Klasifikasi Regresi Logistik Biner	20
2.13 Indeks Pembangunan Manusia	21
2.14 Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia ..	24
2.14.1 Aspek Pendidikan.....	25
2.14.2 Aspek Kesehatan	25
2.14.3 Aspek Ekonomi	26
2.14.4 Aspek Lingkungan	27
2.14.5 Aspek Kependudukan	27
2.14.6 Aspek Penyandang Masalah Kesejahteraan Sosial	29
2.14.7 Aspek Tata Kelola Pemerintahan	29
 BAB 3 METODOLOGI PENELITIAN.....	31
3.1 Desain Penelitian	31
3.2 Sumber Data Penelitian	31
3.3 Variabel Penelitian	31
3.3.1 Variabel Terikat.....	31
3.3.2 Variabel Bebas	31
3.3.3 Bentuk Data Penelitian.....	38
3.4 Teknik Analisis Data	38
3.4.1 Pemodelan Data: Regresi Logistik Biner <i>Sparse-Group</i> LASSO.....	38
3.4.2 Metode Analisis Data	40
3.5 Prosedur Penelitian.....	42
 BAB 4 HASIL DAN PEMBAHASAN.....	43
4.1 Gambaran Umum Provinsi Jawa Barat	43
4.2 Analisis Deskriptif.....	43
4.2.1 Deskripsi Kota/Kabupaten Provinsi Jawa Barat Berdasarkan Indeks Pembangunan Manusia	43
4.2.2 Deskripsi Data Kelompok Pendidikan	45
4.2.3 Deskripsi Data Kelompok Kesehatan	48
4.2.4 Deskripsi Data Kelompok Ekonomi	50
4.2.5 Deskripsi Data Kelompok Lingkungan.....	54
4.2.6 Deskripsi Data Kelompok Kependudukan.....	58
4.2.7 Deskripsi Data Kelompok PMKS	60
4.2.8 Deskripsi Data Kelompok Tata Kelola Pemerintahan	63

4.3 Uji Multikolinearitas	67
4.3.1 Matriks Korelasi	67
4.3.2 <i>Variance Inflation Factor</i> (VIF)	68
4.4 Pemodelan Regresi Logistik Biner <i>Sparse-Group</i> LASSO	69
4.4.1 Penyetelan Parameter Penalti λ	70
4.4.2 Model Regresi Logistik Biner <i>Sparse-Group</i> LASSO.....	72
4.5 Interpretasi Model Regresi Logistik Biner <i>Sparse-Group</i> LASSO.....	74
4.5.1 Interpretasi Koefisien Variabel Kelompok Pendidikan	74
4.5.2 Interpretasi Koefisien Variabel Kelompok Lingkungan	76
4.5.3 Interpretasi Koefisien Variabel Kelompok PMKS.....	79
4.6 Ketepatan Klasifikasi Model	81
4.6.1 <i>Apparent Prediction Error Rate</i> (APER).....	81
4.6.2 ROC-AUC	82
 BAB 5 SIMPULAN DAN SARAN	 84
5.1 Kesimpulan.....	84
5.2 Saran.....	85
 DAFTAR PUSTAKA	 86
LAMPIRAN	93

DAFTAR TABEL

Tabel 2.1 Klasifikasi Regresi Logistik Biner	20
Tabel 3.1 Variabel Bebas Kelompok Pendidikan	32
Tabel 3.2 Variabel Bebas Kelompok Kesehatan.....	33
Tabel 3.3 Variabel Bebas Kelompok Ekonomi.....	33
Tabel 3.4 Variabel Bebas Kelompok Lingkungan.....	34
Tabel 3.5 Variabel Bebas Kelompok Kependudukan	35
Tabel 3.6 Variabel Bebas Kelompok PMKS	36
Tabel 3.7 Variabel Bebas Kelompok Tata Kelola Pemerintahan	37
Tabel 3.8 Bentuk Data Penelitian.....	38
Tabel 4.1 Deskripsi Data Kelompok Pendidikan	45
Tabel 4.2 Deskripsi Data Kelompok Pendidikan per Kategori Y	46
Tabel 4.3 Deskripsi Data Kelompok Kesehatan	48
Tabel 4.4 Deskripsi Data Kelompok Kesehatan per Kategori Y	49
Tabel 4.5 Deskripsi Data Kelompok Ekonomi	51
Tabel 4.6 Deskripsi Data Kelompok Ekonomi per Kategori Y	52
Tabel 4.7 Deskripsi Data Kelompok Lingkungan.....	54
Tabel 4.8 Deskripsi Data Kelompok Lingkungan per Kategori Y.....	55
Tabel 4.9 Deskripsi Data Kelompok Kependudukan.....	58
Tabel 4.10 Deskripsi Data Kelompok Kependudukan per Kategori Y	59
Tabel 4.11 Deskripsi Data Kelompok PMKS	60
Tabel 4.12 Deskripsi Data Kelompok PMKS per Kategori Y	62
Tabel 4.13 Deskripsi Data Kelompok Tata Kelola Pemerintahan	64
Tabel 4.14 Deskripsi Data Kelompok Tata Kelola Pemerintahan per Kategori Y	65
Tabel 4.15 <i>Output</i> Uji VIF	69
Tabel 4.16 Koefisien Parameter Saat λ Optimal	72
Tabel 4.17 <i>Odds-Ratio</i> Variabel Kelompok Pendidikan.....	74
Tabel 4.18 <i>Odds-Ratio</i> Variabel Kelompok Lingkungan	76
Tabel 4.19 <i>Odds-Ratio</i> Variabel Kelompok PMKS.....	79
Tabel 4.20 Ketepatan Klasifikasi Model.....	81

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Perbedaan LASSO, Group LASSO, Sparse-Group LASSO	17
Gambar 2.2 Perkembangan Metodologi Perhitungan IPM di UNDP	23
Gambar 3.1 Diagram Alir Penelitian	42
Gambar 4.1 Capaian IPM Jawa Barat Tahun 2012-2022	44
Gambar 4.2 Capaian IPM Kota/Kabupaten Jawa Barat Tahun 2022 per Kategori.....	44
Gambar 4.3 Matriks Korelasi Variabel Bebas	68
Gambar 4.4 Jalur Koefisien Sparse-Group LASSO menggunakan CV.....	70
Gambar 4.5 Kurva Kesalahan Cross-Validation terhadap λ	71
Gambar 4.6 Grafik ROC	82

DAFTAR LAMPIRAN

Lampiran 1. Data Penelitian: Indeks Pembangunan Manusia.	93
Lampiran 2. Data Penelitian: Variabel Kelompok Pendidikan.....	94
Lampiran 3. Data Penelitian: Variabel Kelompok Kesehatan	95
Lampiran 4. Data Penelitian: Variabel Kelompok Ekonomi	96
Lampiran 5. Data Penelitian: Variabel Kelompok Lingkungan.....	97
Lampiran 6. Data Penelitian: Variabel Kelompok Kependudukan.....	98
Lampiran 7. Data Penelitian: Variabel Kelompok PMKS	99
Lampiran 8. Data Penelitian: Variabel Kelompok Tata Kelola Pemerintahan ...	100
Lampiran 9. Program Python: Eksplorasi Data	101
Lampiran 10. Program Python: Transformasi Data (<i>Z-Score</i>).....	104
Lampiran 11. Program Python: Matriks Korelasi	105
Lampiran 12. Program Python: <i>Variance Inflation Factor</i>	106
Lampiran 13. Program R: Regresi Logistik Biner <i>Sparse-Group LASSO</i>	107
Lampiran 14. Program R: Ketepatan Klasifikasi dengan APER	108
Lampiran 15. Program R: Ketepatan Klasifikasi dengan Analisis ROC-AUC ..	109

DAFTAR PUSTAKA

- Abdullah, K., Jannah, M., Aiman, U., Hasda, S., Fadilla, Z., Taqwin, Masita, Ardiawan, K. N., & Sari, M. E. (2021). *METODOLOGI PENELITIAN KUANTITATIF* (N. Saputra, Ed.). Yayasan Penerbit Muhammad Zaini. <http://penerbitzaini.com>
- Alfisyah, Apriati, Y., & Arofah, L. (2016). *Profil Pengemis Di Kawasan Pemakaman Syekh Muhammad Arsyad Albanjari Di Desa Kalampaian Tengah Kecamatan Astambul Kabupaten Banjar Kalimantan Selatan*.
- Alkahfi, C. (2024, Juni 2). *Metrik Evaluasi untuk Model Klasifikasi*. SAINSDATA.ID. <https://sainsdata.id/machine-learning/8871/metrik-evaluasi-untuk-model-klasifikasi/>
- Altman, N., & Krzywinski, M. (2018). The curse(s) of dimensionality. *Nature Methods*, 15(6), 399–400. <https://doi.org/10.1038/s41592-018-0019-x>
- Alwi, W., Irwan, Muh., & Musfirah. (2021). Penerapan Regresi Nonparametrik Spline Dalam Memodelkan Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia (IPM) di Indonesia Tahun 2018. *Jurnal Matematika dan Statistika serta Aplikasinya*, 9(2), 2021.
- Andriansyah, F., Wulandari, C., Hayatri, T. N., & Zahra, A. H. L. (2022). Strategy of Improvement of the Human Development Indeks in Indonesia Through Sustainable Settlement Management. *Dinamika Ekonomi*, 13(2), 259–270. <https://doi.org/10.29313/de.v13i2.9631>
- Anisa, L., & Rifai, N. A. K. (2022). Analisis Regresi Logistik Biner dengan Metode Penalized Maximum Likelihood pada Penyakit Covid-19 di RSUD Pringsewu. *Jurnal Riset Statistika*, 129–136. <https://doi.org/10.29313/jrs.v2i2.1425>
- Antara, I. G. M. Y., & Suryana, I. G. P. E. (2020). Pengaruh Tingkat Kepadatan Penduduk Terhadap Indeks Pembangunan Manusia di Provinsi Bali. *Media Komunikasi Geografi*, 21(1), 63. <https://doi.org/10.23887/mkg.v21i1.22958>
- Arifin, M. Y., Suswandi, P. E., & Adenan, Moh. (2015). Pengaruh Pengeluaran Pemerintah Sektor Kesehatan, Pengeluaran Pemerintah Sektor Pendidikan dan Pertumbuhan Ekonomi terhadap Indeks Pembangunan Manusia Jawa Timur 2006-2013. *Artikel Ilmiah Mahasiswa*.
- Badan Pusat Statistik. (2015). *INDEKS PEMBANGUNAN MANUSIA 2014*.
- Badan Pusat Statistik. (2023). *INDEKS PEMBANGUNAN MANUSIA 2022*.
- Bakin, S. (1999). *Adaptive Regression and Model Selection in Data Mining Problems* [Doctor of Philosophy]. The Australian National University.
- Bastias, D. D. (2010). *ANALISIS PENGARUH PENGELUARAN PEMERINTAH ATAS PENDIDIKAN, KESEHATAN DAN INFRASTRUKTUR TERHADAP PERTUMBUHAN EKONOMI INDONESIA PERIODE 1969-2009*. Universitas Diponegoro.

- Budiman, M. A., & Cahyani, N. (2022). Pemodelan Regresi Logistik Ordinal Pada Indeks Pembangunan Manusia (IPM) Di Jawa Timur Tahun 2020. *Jurnal Statistika dan Komputasi*. <https://api.semanticscholar.org/CorpusID:255430939>
- Bühlmann, P., & van de Geer, S. (2011). *Statistics for High-Dimensional Data*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-20192-9>
- Chatterjee, S., Steinhäuser, K., Banerjee, A., Chatterjee, S., & Ganguly, A. (2012). Sparse Group Lasso: Consistency and Climate Applications. *Proceedings of the 2012 SIAM International Conference on Data Mining*, 47–58. <https://doi.org/10.1137/1.9781611972825.5>
- Chaudhry, I. (2021). *Factors Affecting Good Governance in Pakistan: An Empirical Analysis*. <https://www.researchgate.net/publication/349811395>
- Chen, H., He, Y., Ji, J., & Shi, Y. (2020). The sparse group lasso for high-dimensional integrative linear discriminant analysis with application to alzheimer's disease prediction. *Journal of Statistical Computation and Simulation*, 90(17), 3218–3231. <https://doi.org/10.1080/00949655.2020.1800011>
- Çokluk Bökeoğlu, Ö. (2010). Logistic Regression: Concept and Application. *Kuram ve Uygulamada Eğitim Bilimleri*, 10, 1397–1407.
- Direktorat Analisis dan Pengembangan Statistik. (2022). *Indeks Pembangunan Manusia 2022*. Badan Pusat Statistik.
- Edgar, T. W., & Manz, D. O. (2017). Exploratory Study. Dalam *Research Methods for Cyber Security* (hlm. 95–130). Elsevier. <https://doi.org/10.1016/B978-0-12-805349-2.00004-2>
- Faizin, Moh. (2021). Pengaruh Upah Minimum, Kemiskinan dan Pengangguran pada IPM di Kabupaten/Kota Provinsi Jawa Timur. *Jurnal Samudra Ekonomi dan Bisnis*, 12(2), 214–227. <https://doi.org/10.33059/jseb.v12i2.3027>
- Ferré, J. (2009). Regression Diagnostics. Dalam *Comprehensive Chemometrics* (hlm. 33–89). Elsevier. <https://doi.org/10.1016/B978-044452701-1.00076-4>
- Fransiska, S., & Rizal, Y. (2020). Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Barat dengan Menggunakan Analisis Regresi Data Panel. *UNPjoMath*, 3(3), 134–138.
- Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). *A note on the group lasso and a sparse group lasso*.
- Fu, W., & Knight, K. (2000). Asymptotics for lasso-type estimators. *The Annals of Statistics*, 28(5). <https://doi.org/10.1214/aos/1015957397>
- Garson, G. D. (2014). *LOGISTIC REGRESSION: BINARY AND MULTINOMIAL*. Statistical Associates Publishing. www.statisticalassociates.com
- Gujarati, D. N. (2004). *Basic Econometrics* (4 ed.). McGraw-Hill Companies.

- Hamidah, W. N., Pramesti, W., & Athoillah, M. (2022). *PEMODELAN REGRESI LOGISTIK ORDINAL PADA INDEKS PEMBANGUNAN MANUSIA (IPM) DI JAWA TIMUR*. UNIVERSITAS PGRI ADIBUANA SURABAYA.
- Hardana, A., & Nasution, J. (2022). PENGARUH RASIO KEUANGAN PEMERINTAH DAERAH TERHADAP INDEKS PEMBANGUNAN MANUSIA. *Global Financial Accounting Journal*, 06(01).
- Haris, A. T. L. P. L., Ishak, A. D., Entong, A., & Haris, A. T. E. H. (2023). PENGARUH INDEKS PEMBANGUNAN MANUSIA TERHADAP INDEKS KUALITAS LINGKUNGAN HIDUP DI INDONESIA. *Jurnal Ganec Swara*, 17(3), 1149–1153. <http://journal.unmasmataram.ac.id/index.php/GARA>
- Harryanto, F. F., & Hansun, S. (2017). Penerapan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE. *JATISI: Jurnal Teknik Informatika dan Sistem Informasi*, 3(2), 95–103.
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical Learning with Sparsity*. Chapman and Hall/CRC. <https://doi.org/10.1201/b18401>
- Hayati, N. (2019). FAKTOR-FAKTOR MEMPENGARUHI IPM DI KOTA BANJARMASIN. *JIEP: Jurnal Ilmu Ekonomi dan Pembangunan*, 1(1). <https://doi.org/10.20527/jiep.v1i1.1127>
- Haynes, W. (2013). *Maximum Likelihood Estimation*. <http://www.dorak.info/genetics/>
- Hertjung, W. S. (2009). *THE DINAMYC OF CAUSES OF CHILD SEXUAL ABUSE BASED ON AVAILABILITY OF PERSONAL SPACE AND PRIVACY*.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley. <https://doi.org/10.1002/9781118548387>
- Huang, J., & Zhang, T. (2010). The benefit of group sparsity. *The Annals of Statistics*, 38(4), 1978–2004. <https://doi.org/10.1214/09-AOS778>
- Hutagulung, M. A. I., & Sutarman. (2024). Penalized Maximum Likelihood Estimation dengan Algoritma Gradient descent pada Model Regresi Logistik Multinomial. *IJM: Indonesian Journal of Multidisciplinary*, 2(6), 673–683. <https://journal.csspublishing/index.php/ijm>
- Imro'ah, N., Sartika, I., & Debataraja, N. N. (2020). ANALISIS REGRESI DENGAN METODE LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR (LASSO) DALAM MENGATASI MULTIKOLINEARITAS. *Bimaster : Buletin Ilmiah Matematika, Statistika dan Terapannya*, 9(1). <https://doi.org/10.26418/bbimst.v9i1.38029>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer Publishing Company, Incorporated.

- Johnson, R. Arnold., & Wichern, D. W. . (2007). *Applied Multivariate Statistical Analysis* (6 ed.). Pearson Prentice Hall.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220–246. <https://doi.org/10.1017/S1876404511200046>
- Khoiriah, A. A., Bakri, S., & Santoso, T. (2017). Pengaruh Perubahan Lahan, Tingkat Kemiskinan Dan Pendapatan Beberapa Sektor Perekonomian Terhadap Indeks Pembangunan Manusia : Studi Di Provinsi Lampung. *Jurnal Sylva Lestari*, 5(1), 117. <https://doi.org/10.23960/jsl15117-127>
- Khomsan, A., Dharmawan, A. H., Saharuddin, Alfiasari, Sukandar, D., & Syarief, H. (2015). *Indikator Kemiskinan dan Misklasifikasi Orang Miskin* (1 ed.). Yayasan Pustaka Obor Indonesia.
- Kiha, E. K., Seran, S., & Lau, H. T. (2021). Pengaruh Jumlah Penduduk, Pengangguran, dan Kemiskinan terhadap Indeks Pembangunan Manusia (IPM) di Kabupaten Belu. *Intelektiva: Jurnal Ekonomi, Sosial & Humaniora*, 2(7), 60.
- Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, 72(6), 558–569. <https://doi.org/10.4097/kja.19087>
- Kim, S.-J., Koh, Kwangmoo., Lustig, M., Boyd, S., & Gorinevsky, D. (2007). An Interior-Point Method for Large-Scale -Regularized Least Squares. *IEEE Journal of Selected Topics in Signal Processing*, 1(4), 606–617. <https://doi.org/10.1109/JSTSP.2007.910971>
- Kleinbaum, D. G., & Klein, M. (2010). *Logistic Regression: A Self-Learning Text* (3 ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-1742-3>
- Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2016). Feature Selection. *ACM Computing Surveys*, 50(6), 1–45. <https://doi.org/10.1145/3136625>
- Linawati, Y., Suzantia, H., & Wibowo, M. G. (2021). Dampak Tata Kelola Pemerintahan Terhadap Pertumbuhan Ekonomi dan Indeks Pembangunan Manusia: Studi Kasus Negara Berkembang OKI. *TEMALI: Jurnal Pembangunan Sosial*, 4(2), 133–144. <https://doi.org/10.15575/jt.v4i2.12547>
- Liu, H., & Zhang, J. (2009). Estimation Consistency of the Group Lasso and its Applications. Dalam D. van Dyk & M. Welling (Ed.), *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics* (Vol. 5, hlm. 376–383). PMLR. <https://proceedings.mlr.press/v5/liu09a.html>
- Liu, X. (2016). *Applied Ordinal Logistic Regression Using Stata: From Single-Level to Multilevel Modeling*. SAGE Publications, Inc. <https://doi.org/10.4135/9781071878972>
- Matsui, H. (2019). Sparse group lasso for multiclass functional logistic regression models. *Communications in Statistics - Simulation and Computation*, 48(6), 1784–1797. <https://doi.org/10.1080/03610918.2018.1423693>

- Muhamad, A. R., & Rahmi, D. (2023). Pengaruh Teknologi, Tingkat Kemiskinan, Pengeluaran Pemerintah, dan Pertumbuhan Ekonomi Terhadap Indeks Pembangunan Manusia di Provinsi Jawa Barat 2007-2021. *Jurnal Riset Ilmu Ekonomi dan Bisnis*, 3(1), 45–52. <https://doi.org/10.29313/jrieb.v3i1.1924>
- Muradi, H., Notodiputro, K. A., & Sartono, B. (2020). METODE SPARSE GRUP LASSO UNTUK MEMODELKAN KECENDERUNGAN PERILAKU SEKSUAL REMAJA PUTRI DI PROVINSI BENGKULU. *Euclid*, 7(1), 61. <https://doi.org/10.33603/e.v7i1.2885>
- Nisa, E. M., & Rafikasari, E. F. (2022). PENGARUH JUMLAH TENAGA KERJA, KEMISKINAN DAN IPM TERHADAP PERTUMBUHAN EKONOMI KABUPATEN TRENGGALEK PERIODE 2016-2020. *Juremi: Jurnal Riset Ekonomi*, 1(5), 483–492.
- Noordiansyah, J., Nhita, F., & Triantoro, D. (2016). Prediksi Penyakit Menggunakan Algoritma K-Means dan GA untuk Reduksi Dimensi dengan Mengintegrasikan SVM pada Data Berdimensi Tinggi. *e-Proceeding of Engineering*, 3(2), 3755–3762.
- Padhilah, R., Herrhyanto, N., & Lukman. (2024). Analisis Regresi Logistik Biner dengan Metode Group LASSO dalam Data Berdimensi Tinggi (Studi Kasus : Indeks Pembangunan Manusia Kota/Kabupaten di Jawa Barat). *BIAStatistics Journal of Statistics Theory and Application*, 18(1), 1–14.
- Prasetyo, Y. (2013). KESADARAN MASYARAKAT BEROLAHRAGA UNTUK PENINGKATAN KESEHATAN DAN PEMBANGUNAN NASIONAL. *MEDIKORA*, 11(2), 219–228.
- Provost, F., & Fawcett, T. (2013). *Data Science for Business* (M. Loukides & M. Blanchette, Ed.; 1 ed.). O'Reilly Media, Inc.
- Purnama, M. D. (2024). An Implementation of Ordinal Probit Regression Model on Factor Affecting East Java Human Development Index. *Engineering, MAThematics and Computer Science Journal (EMACS)*, 6(3), 219–225. <https://doi.org/10.21512/emacsjournal.v6i3.12094>
- Rahmi, D. M. (2019). Investasi Modal Manusia Dalam Pembangunan Ekonomi Perspektif Islam. *Jurnal Manajemen Dan Bisnis Indonesia*, 5(1), 55–64.
- Raschka, Sebastian., Liu, Yuxi., Mirjalili, Vahid., & Dzhulgakov, Dmytro. (2022). *Machine learning with Pytorch and Scikit-Learn : Develop Machine Learning and Deep Learning Models with Python*. Packt Publishing, Limited.
- Roflin, E., Liberty, I. A., & Pariyana. (2021). *POPULASI, SAMPEL, VARIABEL DALAM PENELITIAN KEDOKTERAN* (Moh. Nasrudin, Ed.; 1 ed.). NEM.
- Rosyadah, J. A. (2021). Determinan Indeks Pembangunan Manusia (IPM) Provinsi Nusa Tenggara Timur (NTT). *Efficient: Indonesian Journal of Development Economics*, 4(1), 1080–1092. <https://doi.org/10.15294/efficient.v4i1.41076>
- Sangkereng, W., Engka, D. S. M., & Sumual, J. I. (2019). FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN MANUSIA DI

- PROVINSI SULAWESI UTARA. *Jurnal Berkala Ilmiah Efisiensi*, 19(4), 60–71.
- Santika, R. H., Santoso, B., & Mahmudi, H. (2018). Analisis Dampak Pemekaran Daerah Terhadap Kinerja dan Pemerataan Ekonomi di Kabupaten Lombok Utara. *JIP (Jurnal Ilmu Pemerintahan): Kajian Ilmu Pemerintahan dan Politik Daerah*, 3(1), 15–23. <https://doi.org/10.24905/jip.3.1.2018.15-23>
- Schultz, T. W. (1961). Investment in human capital. *American Economic Review*, 51, 1–17. <http://www.jstor.org/stable/1818907?origin=JSTOR-pdf>
- Senaviratna, N. A. M. R., & A. Cooray, T. M. J. (2019). Diagnosing Multicollinearity of Logistic Regression Model. *Asian Journal of Probability and Statistics*, 5(2), 1–9. <https://doi.org/10.9734/ajpas/2019/v5i230132>
- Siagian, Y., Hidayat, N., Gultom, G. A. M., Belcha, D., & Sipatuhar, J. N. (2024). IJEDR: Indonesian Journal of Education and Development Research Pengaruh PDRB dan Jumlah Penduduk Terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Riau Tahun 2010-2022. *IJEDR: Indonesian Journal of Education and Development Research*, 2(2), 1397–1403. www.bps.go.id.
- Simon, N., Friedman, J., Hastie, T., & Tibshirani, R. (2013). A Sparse-Group Lasso. *Journal of Computational and Graphical Statistics*, 22(2), 231–245. <https://doi.org/10.1080/10618600.2012.681250>
- Situngkir, R. herina, & Sembiring, P. (2023). Analisis Regresi Logistik untuk Menentukan Faktor-Faktor yang Mempengaruhi Kesejahteraan Masyarakat Kabupaten/Kota di Pulau Nias. *FARABI: Jurnal Matematika dan Pendidikan Matematika*, 6(1), 25–31. <https://doi.org/10.47662/farabi.v6i1.432>
- Sofiyat, A. I., Tjalla, A., & Mahdiyah. (2023). PEMODELAN REGRESI LOGISTIK BINER TERHADAP PENERIMAAN PEGAWAI DI PT XYZ JAKARTA. *Matematika Sains*, 1(1). <https://doi.org/https://doi.org/10.34005/ms.v1i1.2625>
- Sukadaryati. (2006). Potensi Hutan Rakyat di Indonesia dan Permasalahannya. *Seminar Hasil Litbang Hasil Hutan*, 49–57.
- Sultan, Rahayu, H. C., & Purwiyanta. (2023). Analisis Pengaruh Kesejahteraan Masyarakat Terhadap Pertumbuhan Ekonomi di Indonesia. *Jurnal Informatika Ekonomi Bisnis*, 75–83. <https://doi.org/10.37034/infec.v5i1.198>
- Sunandi, E., Notodoputro, K. A., & Sartono, B. (2021). A study on group lasso for grouped variable selection in regression model. *IOP Conference Series: Materials Science and Engineering*, 1115(1), 012089. <https://doi.org/10.1088/1757-899X/1115/1/012089>
- Sur, P., & Candès, E. J. (2019). A Modern Maximum-Likelihood Theory for High-Dimensional Logistic Regression. *Proceedings of the National Academy of Sciences*, 116(29), 14516–14525. <https://doi.org/10.1073/pnas.1810420116>

- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tibshirani, R. J., & Tibshirani, R. (2009). A bias correction for the minimum error rate in cross-validation. *The Annals of Applied Statistics*, 3(2), 822–829. <https://doi.org/10.1214/08-AOAS224>
- Tristanto, A., & Diartho, H. C. (2018). Strategi Sektor Kesehatan Dalam Meningkatkan Indeks Pembangunan Manusia Di Kabupaten Situbondo. *WIGA: Jurnal Penelitian Ilmu Ekonomi*, 8(2), 1–9.
- Tsagris, M., & Pandis, N. (2021). Multicollinearity. *American Journal of Orthodontics and Dentofacial Orthopedics*, 159(5), 695–696. <https://doi.org/10.1016/j.ajodo.2021.02.005>
- Vitona, D., Erfiani, E., Indahwati, I., Fitrianto, A., & Aliu, M. A. (2024). ANALISIS FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN MANUSIA DI INDONESIA MENGGUNAKAN MODEL REGRESI LOGISTIK BINER. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 5(2), 903–912. <https://doi.org/10.46306/lb.v5i2.634>
- Yektingnsih, E. (2018). ANALISIS INDEKS PEMBANGUNAN MANUSIA (IPM) KABUPATEN PACITAN TAHUN 2018. *Journal Universitas Wijaya Kusuma Surabaya*, 18(2), 32–50.
- Yuan, M., & Lin, Y. (2006). Model Selection and Estimation in Regression with Grouped Variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(1), 49–67. <https://doi.org/10.1111/j.1467-9868.2005.00532.x>
- Zakaria, R. (2018). *Pengaruh Tingkat Jumlah Penduduk, Pengangguran, Kemiskinan, Pertumbuhan Ekonomi, dan Belanja Modal Terhadap Indeks Pembangunan Manusia di Provinsi Jawa Tengah Tahun 2010-2016*. <https://api.semanticscholar.org/CorpusID:158693474>
- Zenda, R. H., & Suparno. (2017). PERANAN SEKTOR INDUSTRI TERHADAP PENYERAPAN TENAGA KERJA DI KOTA SURABAYA. *JEB 17 Jurnal Ekonomi & Bisnis*, 2(1), 371.
- Zuhairoh, Z. A., & Melaniani, S. (2018). PENGARUH ANGKA KEMATIAN BAYI, ANGKA PARTISIPASI MURNI, RASIO KETERGANTUNGAN TERHADAP INDEKS PEMBANGUNAN MANUSIA PROVINSI JAWA TIMUR. *Jurnal Biometrika an Kependudukan*, 7(1), 87–95.