

**PENDUGAAN NILAI PRODUKSI HASIL TANGKAPAN
PERIKANAN TANGKAP DI KABUPATEN PATI
MENGUNAKAN ALGORITMA *RANDOM FOREST***

SKRIPSI

**diajukan untuk memenuhi sebagian syarat untuk memperoleh
gelar Sarjana Sains pada program studi Sistem Informasi Kelautan**



Oleh

DINDA FAATIAH RAMADHANI PUTRI

2009938

PROGRAM STUDI SISTEM INFORMASI KELAUTAN

KAMPUS UPI SERANG

UNIVERSITAS PENDIDIKAN INDONESIA

2024

**PENDUGAAN NILAI PRODUKSI HASIL TANGKAPAN
PERIKANAN TANGKAP DI KABUPATEN PATI
MENGUNAKAN ALGORITMA *RANDOM FOREST***

Oleh

Dinda Faatihah Ramadhani Putri

Sebuah skripsi yang ditujukan untuk memenuhi salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi Kelautan

© Dinda Faatihah Ramadhani Putri 2024

Universitas Pendidikan Indonesia

April 2024

Hak cipta dilindungi oleh undang-undang

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

ii

SIK UPI Kampus Serang

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Dinda Faatihah Ramadhani Putri
NIM : 2009938
Program Studi : Sistem Informasi Kelautan
Judul Skripsi : Pendugaan Nilai Produksi Hasil Tangkapan Perikanan Tangkap di Kabupaten Pati Menggunakan Algoritma *Random Forest*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sistem Informasi Kelautan pada Program Studi Sistem Informasi Kelautan Universitas Pendidikan Indonesia Kampus Serang

DEWAN PENGUJI

Penguji I : Luthfi Anzani, S.Pd., M.Si.

tanda tangan 

Penguji II : La Ode Alam Minsaris, S.Pi., M.Si.

tanda tangan 

Penguji III : Ma'ruf, S.T., M.Sc.

tanda tangan 

Ditetapkan di : Serang

Tanggal : 21 Maret 2024

HALAMAN PERSETUJUAN

DINDA FAATIAH RAMADHANI PUTRI

**PENDUGAAN NILAI PRODUKSI HASIL TANGKAPAN PERIKANAN
TANGKAP DI KABUPATEN PATI MENGGUNAKAN ALGORITMA
RANDOM FOREST**

disetujui dan disahkan oleh pembimbing:

Pembimbing I,



**Willdan Aprizal Arifin, S.Pd., M.Kom.
NIPT. 920200819940415101**

Pembimbing II,



**Ayang Armelita Rosalia, S.Pi., M.Si
NIPT. 920200819941203201**

Mengetahui,
Ketua Program Studi Sistem Informasi Kelautan



**Willdan Aprizal Arifin, S.Pd., M.Kom.
NIPT. 9202008199404151**

Pendugaan Nilai Produksi Perikanan Tangkap di Kabupaten Pati Menggunakan Algoritma *Random Forest*

Dinda Faatihah Ramadhani Putri
Program Studi Sistem Informasi Kelautan

ABSTRAK

Wilayah Pati yang berbatasan dengan Laut Jawa memiliki potensi sebagai salah satu penghasil ikan terbesar di wilayah Jawa Tengah, khususnya pada perikanan tangkap yang memberikan kontribusi besar bagi perekonomian wilayah Pati. Diharapkan bahwa kedepannya masyarakat setempat dapat memanfaatkan sumber daya perikanan dengan baik dan benar guna meningkatkan pendapatan mereka. Salah satu cara yang dapat dilakukan yaitu perlu adanya penelitian pendugaan atau prediksi harga jual hasil tangkapan agar dapat tercapainya jumlah produksi dan nilai produksi yang diinginkan dengan memanfaatkan algoritma *machine learning*, salah satunya *Random Forest*. Data yang digunakan diperoleh dari data arsip DKP Kabupaten Pati dan BPS Provinsi Jawa Tengah pada tahun 2013-2022, dengan membandingkan rasio pembagian dataset atau *splitting* 60:40, 70:30, dan 80:20 dan membandingkan metode evaluasi *hyperparameter tuning* menggunakan *RandomizedSearch CV* dan *GridSearch CV*. Hal ini dilakukan agar dapat memaksimalkan kinerja pada algoritma *Random Forest*. Pemilihan metode terbaik dilakukan dengan membandingkan nilai *r-square* (r^2), MSE, dan MAPE dengan rasio *splitting* yang unggul adalah 80:20 menggunakan metode evaluasi *GridSearch CV*, dengan nilai r^2 sebesar 0.85, nilai MSE sebesar 1.3 dan nilai MAPE sebesar 15.38 %. Berdasarkan ketentuan Lewis pada tahun 1982, nilai MAPE antara 10-20% menandakan bahwa model prediksi yang dibentuk dalam kategori baik dan dapat dijadikan sebagai acuan.

Kata kunci: Kabupaten Pati, *Hyperparameter Tuning*, Perikanan Tangkap, *Random Forest*

Estimation of Capture Fisheries Production Value in Pati Regency Using Random Forest Algorithm

Dinda Faatihah Ramadhani Putri
Program Studi Sistem Informasi Kelautan

ABSTRACT

The Pati region is contiguous the Java Sea who has a potential to become one of the largest fisheries production in Central Java, especially in marine fisheries which contributes for the economy of Pati region. Expected that in the future local communities can continue to utilize fishery resources properly and correctly to increase their income. One way that can be done is to conduct estimation or prediction of fisheries catches, research in order to achieve the desired production amount and production value, by utilizing machine learning algorithms, one of them is Random Forest algorithm. The data used was obtained from archival data from the Pati Regency DKP and BPS of Central Java Province in 2013-2022, with splitting ratios of 60:40, 70:30, and 80:20 and comparing hyperparameter tuning evaluation methods using RandomizedSearch CV and GridSearch CV. This action is for maximize the performance of the Random Forest algorithm. The best method was selected by comparing the r-square (r^2), MSE and MAPE values with a superior splitting ratio of 80:20 using the GridSearch CV evaluation method, with an r^2 value of 0.85, an MSE value of 1.3 and a MAPE value of 15.38%. According to Lewis in 1982, MAPE value between 10-20% indicates the prediction is in the good category.

Keywords: Capture Fisheries, Hyperparameter Tuning, Pati Regency, Random Forest

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN HAK CIPTA	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
KATA PENGANTAR	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	vi
SURAT PERNYATAAN.....	vii
ABSTRAK.....	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	9
DAFTAR GAMBAR	xviii
LAMPIRAN.....	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	5
1.5 Ruang Lingkup Peneliti.....	6
BAB II TINJAUAN PUSTAKA.....	6
2.1. Kajian Teori.....	6
2.2. Penelitian yang Relevan	13
BAB III METODE PENELITIAN	15
3.1. Pendekatan/Desain Penelitian	15
3.2. Metode Penelitian.....	16
3.3. Teknik Penelitian.....	17

3.4. Latar/Setting Penelitian	22
3.5. Dataset Penelitian	23
3.6. Prosedur Penelitian.....	23
BAB IV	27
4.1. Cara kerja model algoritma <i>Random Forest</i> dalam mengolah data	27
4.2. Menganalisis tingkat keberhasilan algoritma <i>Random Forest</i>	42
4.3. Mengimplementasikan model pada website melalui <i>framework Flask</i> ..	54
BAB V KESIMPULAN DAN SARAN.....	58
5.1. Kesimpulan.....	58
5.2. Saran.....	59
DAFTAR REFERENSI	12
LAMPIRAN.....	68

DAFTAR TABEL

Tabel 4. 1. Deskripsi Nilai Statistik Pada Dataset.....	27
Tabel 4. 2. Korelasi antar fitur terhadap Nilai Produksi.....	34
Table 4. 3. Pembagian Data Latih Dan Data Uji.....	35
Tabel 4. 4. Parameter Algoritma <i>Random Forest</i>	37
Tabel 4. 5. Perbandingan hasil algoritma <i>Random Forest</i>	39
Tabel 4. 6. Perbandingan akurasi <i>RandomizedSearch CV</i>	46
Tabel 4. 7. Perbandingan hasil algoritma dengan <i>RandomizedSearch CV</i>	47
Tabel 4. 8. Perbandingan akurasi dengan <i>GridSearch CV</i>	51
Tabel 4. 9. Perbandingan hasil algoritma <i>Random Forest</i>	52
Tabel 4. 10. Perbandingan evaluasi <i>Hyperparameter Tuning</i>	53

DAFTAR GAMBAR

Gambar 2.1. Proses Tahapan KDD	7
Gambar 3.1. Flowchart Alur Tahapan Penelitian.....	15
Gambar 3.2. Flowchart Proses Algoritma <i>Random Forest</i>	19
Gambar 3.3. Peta Penelitian Wilayah Kabupaten Pati	22
Gambar 4. 1. Mengecek <i>Missing Value</i> pada <i>Dataset</i>	28
Gambar 4. 2. <i>Boxplot</i> Seluruh Fitur pada <i>Dataset</i>	29
Gambar 4. 3. <i>Boxplot</i> Produksi Setelah <i>Handling Outlier</i>	31
Gambar 4. 4. <i>Boxplot</i> Nilai Produksi Setelah <i>Handling Outlier</i>	31
Gambar 4. 5. Korelasi antar Fitur pada <i>Dataset</i>	32
Gambar 4. 6. Pembentukan Variabel X dan Y	36
Gambar 4. 7. Hasil Normalisasi Fitur menggunakan <i>MinMax Scaler</i>	37
Gambar 4. 8. Grafik Prediksi Splitting 60:40.....	40
Gambar 4. 9. Grafik Prediksi Splitting 70:30.....	41
Gambar 4. 10. Grafik Prediksi Splitting 80:20.....	41
Gambar 4. 11. Grafik Prediksi <i>Splitting</i> 60:40 dengan <i>RandomizedSearch CV</i> ..	44
Gambar 4. 12. Grafik Prediksi <i>Splitting</i> 70:30 dengan <i>RandomizedSearch CV</i> ..	45
Gambar 4. 13. Grafik Prediksi <i>Splitting</i> 80:20 dengan <i>RandomizedSearch CV</i> ..	45
Gambar 4. 14. Grafik Prediksi <i>Splitting</i> 60:40 dengan <i>GridSearch CV</i>	49
Gambar 4. 15. Grafik Prediksi <i>Splitting</i> 70:30 dengan <i>GridSearch CV</i>	49
Gambar 4. 16. Grafik Prediksi <i>Splitting</i> 80:20 dengan <i>GridSearch CV</i>	50
Gambar 4. 17. Tampilan <i>Website</i> Prediksi	55
Gambar 4. 18. Tampilan Hasil Prediksi	56

LAMPIRAN

Lampiran 1. <i>Script Code</i> Pembuatan model Algoritma <i>Random Forest</i>	68
Lampiran 2. <i>Script Code</i> Pembuatan <i>Website</i> Prediksi.....	72
Lampiran 3. Daftar Riwayat Hidup.....	75

DAFTAR REFERENSI

- Adi, S. (2022). *Komparasi Metode Support Vector Machine (Svm), K-Nearest Neighbors (Knn), Dan Random Forest (Rf) Untuk Prediksi Penyakit Gagal Jantung*. 10.
- Ahlburg, D. A. (1995). Simple Versus Complex Models: Evaluation, Accuracy, and Combining. *Mathematical Population Studies*, 5(3), 281–290. <https://doi.org/10.1080/08898489509525406>
- Amiruddin, & Ishak, R. (2022). Implementasi Seleksi Fitur Klasifikasi Waktu Kelulusan Mahasiswa Menggunakan Correlation Matrix With Heatmap. *Jambura Journal of Electrical and Electronics Engineering*, 4(2).
- Anggoro, D. A., & Aziz, N. C. (2021). Implementation of K-Nearest Neighbors Algorithm for Predicting Heart Disease Using Python Flask. *Iraqi Journal of Science*, 62(9), 3196–3219. <https://doi.org/10.24996/ijs.2021.62.9.33>
- Arifin, W. A., Ariawan, I., Rosalia, A. A., Lukman, & Tufailah, N. (2022). Data scaling performance on various machine learning algorithms to identify abalone sex. *Jurnal Teknologi Dan Sistem Komputer*, 10(1), 26–31. <https://doi.org/10.14710/jtsiskom.2022.14105>
- Azmi, U., Hadi, Z. N., & Soraya, S. (2020). ARDL METHOD: Forecasting Data Curah Hujan Harian NTB. *Jurnal Varian*, 3(2), 73–82. <https://doi.org/10.30812/varian.v3i2.627>
- Bergstra, J., Ca, J. B., & Ca, Y. B. (2012). Random Search for Hyper-Parameter Optimization Yoshua Bengio. In *Journal of Machine Learning Research* (Vol. 13).
- Berrar, D. (2018). Cross-validation. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (Vols. 1–3, pp. 542–545). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- Breiman, L. (2001). *Random Forests* (Vol. 45).

- Cahyo, E. N., & Susanti, E. (2023). Analisis Time Series Untuk Deep Learning Dan Prediksi Data Spasial Seismik: Studi Literatur. *Jurnal Teknologi*, 15(2), 124–136. <https://doi.org/10.34151/jurtek.v15i2.3581>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- Choirunisa, P. (2020). *Implementasi Artificial Intelligence Untuk Memprediksi Harga Penjualan Rumah Menggunakan Metode Random Forest Dan Flask*.
- Dalimunthe, D. K., & Hakim, R. B. F. (2023). Application Of Random Forest Algorithm On Watch Price Prediction System Using Framework Flask. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 17(1), 0171–0184. <https://doi.org/10.30598/barekengvol17iss1pp0171-0184>
- Damayanti, H. O. (2018). *Strategi Pengembangan Usaha Penangkapan Ikan Tradisional: Studi Kasus di Desa Pecangaan, Kecamatan Batangan, Kabupaten Pati*.
- Damayanti, H. O. (2020). *The Productivity Of Purse Seine Net Fisheries Business Herna Octivia Damayanti* (Vol. 16, Issue Juni).
- DKP Pati. (2020). *Produksi Perikanan Laut Jumlah Hasil Tangkapan Perairan Laut*.
- Dua, S., & Du, X. (2011). *Data Mining and Machine Learning in Cybersecurity*.
- Dudek, A., & Walesiak, M. (2021). *The Choice of Variable Normalization Method in Cluster Analysis*. <https://wir.ue.wroc.pl/info/article/WUT1860f7e9e0624df8a904aed12638a6ea>
- Faiz, M. N., Somantri, O., Supriyono, A. R., & Muhammad, A. W. (2022). Impact of Feature Selection Methods on Machine Learning-based for Detecting DDoS Attacks : Literature Review. *Journal Of Informatics And Telecommunication Engineering*, 5(2), 305–314. <https://doi.org/10.31289/jite.v5i2.6112>

- Fang, X., Li, X., Zhang, Y., Zhao, Y., Qian, J., Hao, C., Zhou, J., & Wu, Y. (2021). *Random Forest-Based Understanding And Predicting Of The Impacts Of Anthropogenic Nutrient Inputs On The Water Quality Of A Tropical Lagoon*.
- Fathoni, M. Y. (2017). Implementasi Metode Fuzzy Time Series Cheng untuk prediksi Konsentrasi Gas NO₂ Di Udara. *jurnal sistem informasi bisnis*, 7(1), 17. <https://doi.org/10.21456/vol7iss1pp17-23>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *From Data Mining to Knowledge Discovery in Databases* (Vol. 17).
- Fayyad, U., Shapiro, G. P., & Smyth, P. (1996). *From Data Mining to Knowledge Discovery in Databases* (Vol. 17).
- Febrianti, S. S., Boesono, H., Trisnani, D., Hapsari, D., Pemanfaatan, M., Perikanan, S., Perikanan, F., & Kelautan, I. (2013). Analisis Faktor-Faktor Yang Mempengaruhi Harga Ikan Manyung (Arius Thalassinus) Di Tpi Bajomulyo Juwana Pati. In *Journal of Fisheries Resources Utilization Management and Technology* (Vol. 2, Issue 3). <http://www.ejournal-s1.undip.ac.id/index.php/jfrumt>
- Fitriah, N. W. N., Kusumadewi, R., & Farida, A. S. (2022). *Strategi Pengembangan Sektor Perikanan Terhadap Peningkatan Pendapatan Asli Daerah Kabupaten Pangandaran*. 20.
- Ghozali, I. (2016). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23*.
- Harmain, A., Paiman, Kurniawan, H., Kusrini, & Maulina D. (2021). *Data Normalization For K-Means Efficiency On Groups Of Areas With Potential Fires And /Land Fire Based On Heat Spots Distribution*.
- Hastrini, R., Rosyid, A., & Riyadi, H. , P. (2013). *Analisis penanganan (Handling) Hasil Tangkapan Kapal Purse Seine yang Didaratkan di Pelabuhan Perikanan Pantai (PPP) Bajomulyo Kabupaten Pati*. 2, 1–10.
- Hirzi, H. R., Hidayaturrohmah, U., Kertanah, Amaly, M. H., & Satriawan, R. (2023). Prediksi Jumlah Wisatawan Menggunakan Metode Random Forest, Single Exponential Smoothing dan Double Exponential Smoothing. *Jambura Journal*

of *Probability and Statistics*, 4(1), 47–55.
<https://doi.org/10.34312/jjps.v4i1.17088>

Iglesias, G., Talavera, E., Prieto, A. G., Mozo, A., & Canaval, S. G. (2023). *Data Augmentation techniques in time series domain: A survey and taxonomy*.

Islam, M. M., Kashem, M. A., & Uddin, J. (2021). Fish survival prediction in an aquatic environment using random forest model. *IAES International Journal of Artificial Intelligence*, 10(3), 614–622.
<https://doi.org/10.11591/ijai.v10.i3.pp614-622>

Kementerian Kelautan dan Perikanan. (2014). *Peraturan Menteri Kelautan dan Perikanan Nomor 18/PERMEN-KP/2014 Tentang Wilayah Pengelolaan Perikanan Negara Republik Indonesia*.

KKP. (2004). *Keputusan Menteri Kelautan dan Perikanan Republik Indonesia Nomor 12/MEN/2004 tentang Peningkatan Status Pangkalan Pendaratan Ikan menjadi Pelabuhan Perikanan Pantai pada Provinsi Jawa Barat, Jawa Tengah, Jawa Timur dan Lampung*. Jakarta.

Kuvalekar, A., Mahadik, S., Manchewar, S., & Jawale, S. (2020). *HOUSE PRICE FORECASTING USING MACHINE LEARNING*.

Lestari, E. S., & Astuti, I. (2022). Penerapan Random Forest Regression Untuk Memprediksi Harga Jual Rumah Dan Cosine Similarity Untuk Rekomendasi Rumah Pada Provinsi Jawa Barat. *Jurnal Ilmiah FIFO*, 14(2), 131.
<https://doi.org/10.22441/fifo.2022.v14i2.003>

Lewis, C. D. (Colin D. (1982). *Industrial and business forecasting methods*. Butterworth Scientific.

Li, B. (2020). *Random Search Plus: A more effective random search for Random Search Plus: A more effective random search for machine learning hyperparameters optimization machine learning hyperparameters optimization*. https://trace.tennessee.edu/utk_gradthes/5849

Liu, C. H. B., Chamberlain, B. P., Little, D. A., & Cardoso, Â. (2017). Generalising Random Forest Parameter Optimisation to Include Stability and Cost. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial*

Intelligence and Lecture Notes in Bioinformatics), 10536 LNAI, 102–113.
https://doi.org/10.1007/978-3-319-71273-4_9

Muraina, I. O. (2022). *Ideal Dataset Splitting Ratios In Machine Learning Algorithms: General Concerns For Data Scientists And Data Analysts*.
<https://www.researchgate.net/publication/358284895>

Nawawi, M. H. (2020). *Predksi Harga Emas Di Indonesia Menggunakan Algoritma Gradient Boosting Regressor*.

Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Van Le, H., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*, 2021. <https://doi.org/10.1155/2021/4832864>

Ningrum, A. A., & Ihsanudin. (2023). Penerapan Framework Flask Pada Machine Learning Dalam Memprediksi Umur Transformer. *KONVERGENSI*, 19(2), 51–59.

Nti, I. K., Nyarko-Boateng, O., & Aning, J. (2021). Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation. *International Journal of Information Technology and Computer Science*, 13(6), 61–71. <https://doi.org/10.5815/ijitcs.2021.06.05>

Nugroho, R. A., Rahmawati, S., & Eni, D. K. A. (2022). *Perbandingan Prediksi Harga Saham Menggunakan Metode SVR, RFR, dan DTR*. 6.

Nuranisah. (2021). *Analisis Menggunakan Random Forest Dengan Gini Index Algoritma Pada Data*.

Nurhayati, A., Pical, V., Efani, A., Hilyaa, S., Saloko, S., Made, S., & Purnomo, H. A. (2020). *Manajemen Resiko Perikanan Tangkap (Studi Kasus di Tengah Pandemi COVID-19)*. <http://jfmr.ub.ac.id>

Nyitrai, T., & Virág, M. (2019). The Effects Of Handling Outliers On The Performance Of Bankruptcy Prediction Models. *Socio-Economic Planning Sciences*, 67, 34–42. <https://doi.org/10.1016/j.seps.2018.08.004>

- Octivia, H. (2014). Leading Fishery Commodities Of Pati Regency In Central Java Province Scale. *Jurnal Litbang*, *X*(1), 24–34.
- Papadopoulos, T., Kosmas, I., Botsoglou, G., Dourvas, N. I., Maga-Nteve, C., & Michalakelis, C. (2023). Predicting Consumer Service Price Evolution during the COVID-19 Pandemic: An Optimized Machine Learning Approach. *Electronics* (Switzerland), *12*(18). <https://doi.org/10.3390/electronics12183806>
- Penalun, F. E., Hermawan, A., & Avianto, D. (2023). *Perbandingan Random Forest Regression dan Support Vector Regression Pada Prediksi Laju Penguapan*. *13*, 104–111.
- Pradana, G. A. (2023). *A Study of Prediction Model for Capture Fisheries Production in Indonesian Sea Waters Using Machine Learning*.
- Prasojo, B., & Haryatmi, E. (2021). Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *7*(2), 79–89. <https://doi.org/10.25077/teknosi.v7i2.2021.79-89>
- Pratama, Y. W. (2022). *Rancang Bangun Sistem Identifikasi Kepribadian Myers-Briggs Type Indicator (MBTI) Menggunakan Decision tree Berbasis Video Game Untuk Rekomendasi Pemilihan Pekerjaan*.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (Vol. 9, Issue 3). Wiley-Blackwell. <https://doi.org/10.1002/widm.1301>
- Puspasari, G. (2017). Strategi Pengembangan Nelayan Tangkap PPP Bajomulyo Kabupaten Pati. In *Economics Development Analysis Journal* (Vol. 6, Issue 1). <http://journal.unnes.ac.id/sju/index.php/edaj>
- Rácz, A., Bajusz, D., & Héberger, K. (2021). Effect of dataset size and train/test split ratios in qsar/qspr multiclass classification. *Molecules*, *26*(4). <https://doi.org/10.3390/molecules26041111>
- Raharjo, B. (2021). *Pembelajaran Mesin (Machine Learning)*.

- Rahmat, A. W., Ladjamuddin, M. S., & Awaludin, T. D. (2023). Perbandingan Algoritma Decision Tree, Random Forest Dan Naive Bayes Pada Prediksi Penilaian Kepuasan Penumpang Maskapai Pesawat Menggunakan Dataset Kaggle. *Jurnal Rekayasa Informasi*, 12(2)
- Ramayanti, D., & Salamah, U. (2018). Text Classification on Dataset of Marine and Fisheries Sciences Domain using Random Forest Classifier. *International Journal of Computer Techniques*, 5(Sep). <http://www.ijctjournal.org>
- Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang. In *IJCIT (Indonesian Journal on Computer and Information Technology)* (Vol. 5, Issue 1).
- Samudra, Y. A. (2019). *Pendekatan Random Forest Untuk Model Peramalan Harga Tembakau Rajangan Di Kabupaten Temanggung*.
- Santoso, P., Abijono, H., & Anggreini, N. L. (2021). Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data. *Unira Malang* |, 4(2).
- Sari, L., Romadloni, A., & Listyaningrum, R. (2023). Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random Forest. *Infotekmesin*, 14(1), 155–162. <https://doi.org/10.35970/infotekmesin.v14i1.1751>
- Savitri, N. L. P. C., Rahman, R. A., Venyutzky, R., & Rakhmawati, N. A. (2021). Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning. *Jurnal Teknik Informatika Dan Sistem Informasi*, 7(1). <https://doi.org/10.28932/jutisi.v7i1.3216>
- Senan, E. M., Abunadi, I., Jadhav, M. E., & Fati, S. M. (2021). Score and Correlation Coefficient-Based Feature Selection for Predicting Heart Failure Diagnosis by Using Machine Learning Algorithms. *Computational and Mathematical Methods in Medicine*. <https://doi.org/10.1155/2021/8500314>
- Setiawan, H., Bey Pane, A., & Lubis, E. (2019a). *Development Strategy of Bajomulyo Coastal Fishing Port for Improving Port Functions*. 3, 59–72.

- Sholihah, N. N., & Hermawan, A. (2023). *Implementation Of Random Forest And Smote Methods For Economic Status Classification In Cirebon City*. 4(6), 1387–1397. <https://doi.org/10.52436/1.jutif.2023.4.6.1135>
- Sinaga, D. (2014). *Statistika Dasar*.
- Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97. <https://doi.org/10.1016/j.asoc.2019.105524>
- Srimurdianti S, A., & Setiawan, W. (2018). *Peramalan Saham Berdasarkan Data Masa Lalu dengan Pendekatan Fuzzy Time Series*.
- Steven, J., Rigatti, MD, DBIM, & DABFM. (2017). Random Forest. In *Journal Of Insurance Medicine Copyright C 2017 Journal of Insurance Medicine J Insur Med* (Vol. 47).
- Sugiyono. (2018). *Metode Penelitian Kuantitatif Kualitatif Dan R&D*.
- Suliztia, L. M. (2020). *Penerapan Analisis Random Forest Pada Prototype Sistem Prediksi Harga Kamera Bekas Menggunakan Flask*.
- Susanto, A., Hamzah, A., Irnawati, R., Sutrawan Nurdin, H., Nugraheni Supadminingsih, F., Perikanan, J., Pertanian, F., & Sultan Ageng Tirtayasa Jalan Raya Jakarta Km, U. (2020). ISSN The Role of Capture Fisheries Sector in Supporting Food Security of Fishery in Banten Province. In *Journal of Local Food Security* (Vol. 1).
- Susilowati, I., Istiqomah, I., Sukiman, S., & Purnomo, S. (2022). Analisis Margin Pemasaran Ikan Layang di Kabupaten Pati. *Jurnal Ekonomi Pertanian Dan Agribisnis*, 6(1), 238–248. <https://doi.org/10.21776/ub.jepa.2022.006.01.23>
- Syafira, M. N., & Saepudin D. (2023). *Prediksi Return Saham dengan Metode Random Forest dan Penerapannya untuk Seleksi Portofolio*. 10.
- Syakra, Y., Syahril, M., & Yusnidah. (2018). *Implementasi Data Mining Dengan Menggunakan Algoritma Fuzzy Subtractive Clustering Dalam Pengelompokan Nilai Untuk Menentukan Minat Belajar Siswa Smp Primbana Medan*.

- Syahril, M., Erwansyah, K., & Yetri, M. (2020). Penerapan Data Mining Untuk Menentukan Pola Penjualan Peralatan Sekolah Pada Brand Wigglo Dengan Menggunakan Algoritma Apriori., *118*(1), 118–136.
- Tryasmara, H. S., Wijayanto, D., & Budi, B. (2017). Leading Commodities Analysis of Fisheries in Pati Regency. *Journal of Fisheries Resources Utilization Management and Technology*, *6*, 175–179. <http://www.ejournal-s1.undip.ac.id/index.php/jfrumt>
- Verbraecken, J., Wolting, M., Katzy, J., Klop-Penburg, J., Verbelen, T., & Rellermeyer, J. S. (2019). *Survey on Distributed Machine Learning*.
- Vinutha, H. P., Poornima, B., & Sagar, B. M. (2018). Detection of outliers using interquartile range technique from intrusion dataset. *Advances in Intelligent Systems and Computing*, *701*, 511–518. https://doi.org/10.1007/978-981-10-7563-6_53
- Vrigazova, B. (2021). The Proportion for Splitting Data into Training and Test Set for the Bootstrap in Classification Problems. *Business Systems Research*, *12*(1), 228–242. <https://doi.org/10.2478/bsrj-2021-0015>
- Wardhana, S. G., Pakpahan, H. J., Simarmata, K., Pranowo, W., & Purba, H. (2021). *Algoritma Komputasi Machine Learning untuk Aplikasi Prediksi Nilai Total Organic Carbon (TOC)*. *55*, 75–87.
- Wiranda, L., & Sadikin, M. (2019). *Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma* (Vol. 8).
- Yu, C. S., Lin, C. H., Lin, Y. J., Lin, S. Y., Wang, S. Te, Wu, J. L., Tsai, M. H., & Chang, S. S. (2020). Clustering heatmap for visualizing and exploring complex and high-dimensional data related to chronic kidney disease. *Journal of Clinical Medicine*, *9*(2). <https://doi.org/10.3390/jcm9020403>
- Zed, M. (2014). *Metode Penelitian Kepustakaan*.
- Zhang, X., & Liu, A. N. (2023). *Model Averaging Prediction by K-fold Cross-validation*.