



Klasifikasi Diagnosis Penyakit Alzheimer Menggunakan Algoritma Rotation Forest

Kelompok 4:

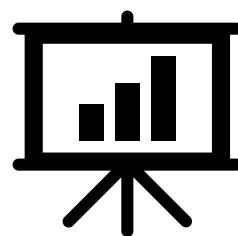
- Sabrina Afifah Putri Utami G1401221035
- R. Mugni Chairil Arbi Asyari G1401221063
- Sherinda Engelina G1401221067
- Ria Yunita G1401221115

Dosen Pembimbing:

Dr. Septian Rahardiantoro, S.Stat., M.Si.

Cici Suhaeni S.Si., M.Si., Ph.D.

Prof. Dr. Anang Kurnia S.Si., M.Si.



OUTLINE



01 Latar

02 Tujuan

03 Data

04 Tinjauan Pustaka

05 Metode

06 Hasil dan Pembahasan

07 Kesimpulan

Latar Belakang



Martha (2008)
Rotation Forest digunakan untuk memprediksi mahasiswa *dropout* & menunjukkan fleksibilitas metode

Saputra(2018)
Rotation Forest lebih akurat dibandingkan CART & Random Forest, tetapi memerlukan waktu komputasi lebih tinggi

Tujuan



1

**Menganalisis efektivitas metode rotation forest
dalam mengklasifikasikan penderita alzheimer**

2

**Mengidentifikasi faktor yang paling berpengaruh
dalam menentukan klasifikasi penderita alzheimer**

DATA



Data ini diperoleh dari platform kaggle yang dapat diakses melalui [Dataset Alzheimer](#)

Fitur data terdiri dari:

- **1 Fitur Y (Kategorik)**
- **15 Fitur X (Numerik)**
- **18 Fitur X (Kategorik)**

Jumlah Amatan **2.149** data pasien

Fitur	Nama Fitur	Keterangan
Y	Status Alzheimer	Kategorik
X1	Usia	Numerik
X2	Gender	Kategorik
:	:	:
X31	<i>Kesulitan Problem Solving</i>	Numerik
X32	Pelupa	Numerik

Tinjauan Pustaka



IPB University
Bogor Indonesia

Rotation Forest



Dicetuskan oleh Juan Rodriguez, Ludmila Kuncheva, dan Christopher J. Hathaway pada tahun 2006 dalam tulisan yang berjudul “Rotation Forest: A New Classifier Ensemble Method”

Metode klasifikasi ensemble untuk membuat klasifikasi yang beragam dan akurat.
2 metode penting dalam rotation forest:

PCA: Ekstraksi peubah

Decision Tree: Dasar klasifikasi



Keunggulan

Pertahankan keberagaman

Akurasi yang baik

Decision Tree yang digunakan relatif sedikit

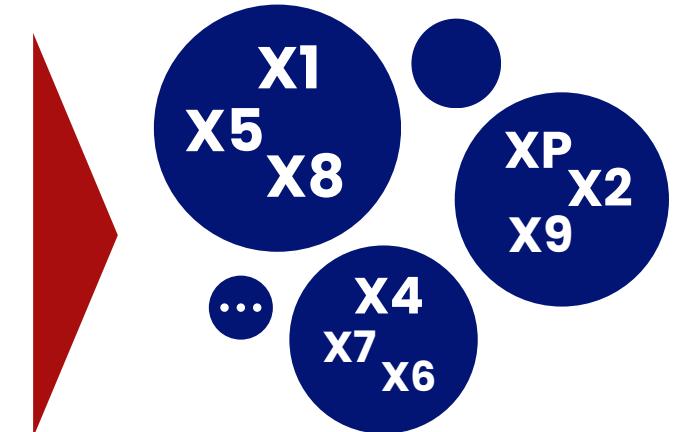
Metode kekar

Tinjauan Pustaka



Rotation Forest

	X1	...	XP	Y
1				
:				
n				



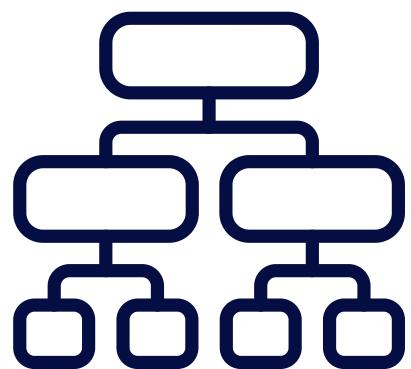
Himpunan bagian acak

	X1	X5	X8
1			
:			
n			

Susun amatan yang berpadanan



Ulang hingga t pohon keputusan



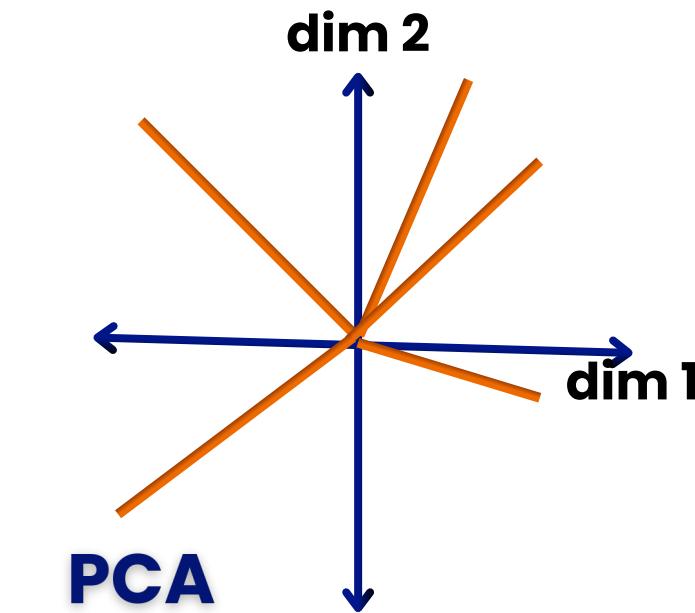
Buat pohon keputusan
 $D_i = X R_i^a Y$

	X1	X2	...	XP
1	a1	.	.	.
p	.	.	.	ap

Susun ulang kolom peubah R_i menjadi R_i^a sehingga bersesuaian dengan susunan gugus asli peubah

$$\begin{bmatrix} a_{i,1}^{(1)}, a_{i,1}^{(2)}, \dots, a_{i,1}^{(M_1)} & \cdots & [0] \\ \vdots & \ddots & \vdots \\ [0] & \cdots & a_{i,k}^{(1)}, a_{i,k}^{(2)}, \dots, a_{i,k}^{(M_k)} \end{bmatrix}$$

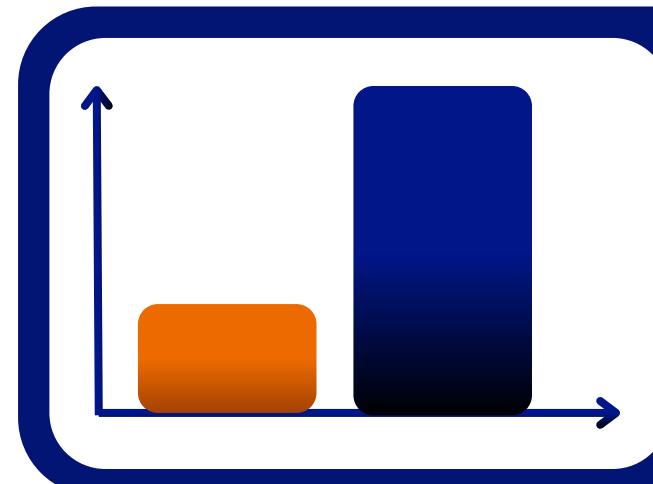
Koefisien PCA disimpan dalam vektor koefisien komponen utama dan disusun ke matriks rotasi R_i



Tinjauan Pustaka



Balancing



data tidak seimbang:

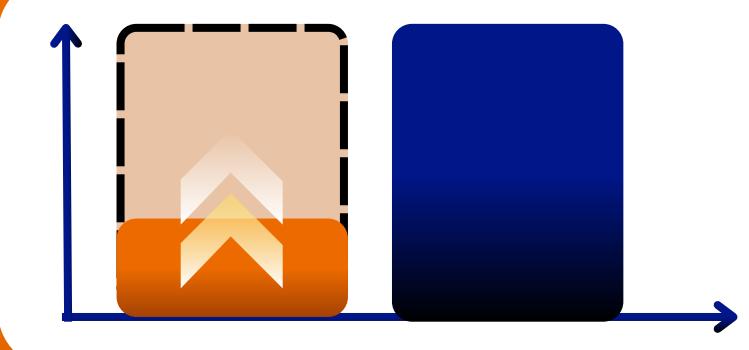
- model condong pada kelas mayoritas
- mengurangi akurasi prediksi untuk kelas minoritas

Undersampling



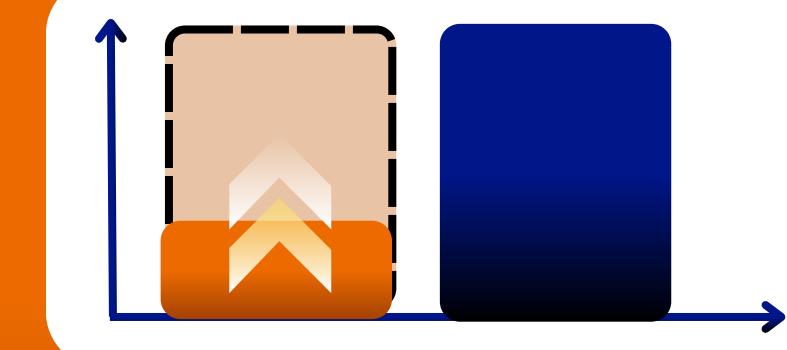
- Menghapus sampel dari kelas mayoritas dengan acak
- Dampak: kehilangan informasi

Oversampling



- Menghasilkan sampel sintetis dari kelas minoritas dengan acak
- Dampak: Overfitting

SMOTE



- Menghasilkan sampel sintetis dari kelas minoritas dengan K-Nearest Neighbor
- Dampak: Overlapping

Tinjauan Pustaka



Metriks Evaluasi

PREDIKSI		REALITAS	TRUE	FALSE
REALITAS	TRUE	TP	FN	
FALSE	FP		TN	

Akurasi

Ketepatan klasifikasi benar

$$\frac{TP + TN}{TP + FP + FN + TN}$$

Presisi

Ketepatan identifikasi kelas positif

$$\frac{TP}{TP + FP}$$

Sensitivitas

Ketepatan identifikasi amatan positif sebenarnya

$$\frac{TP}{TP + FN}$$

Spesifisitas

Ketepatan identifikasi amatan negatif sebenarnya

$$\frac{TN}{FP + TN}$$

F-1 Score

Keseimbangan presisi dan sensitivitas

$$\frac{2 \times \text{sensitivitas} \times \text{presisi}}{\text{sensitivitas} + \text{presisi}}$$

Tinjauan Pustaka



IPB University
Bogor Indonesia

SHapley Additive exPlanations (SHAP)

SHAP

Metode yang digunakan untuk mengidentifikasi fitur yang berpengaruh pada hasil prediksi model

dapat mengevaluasi output model berdasarkan jumlah dampak pada setiap fitur

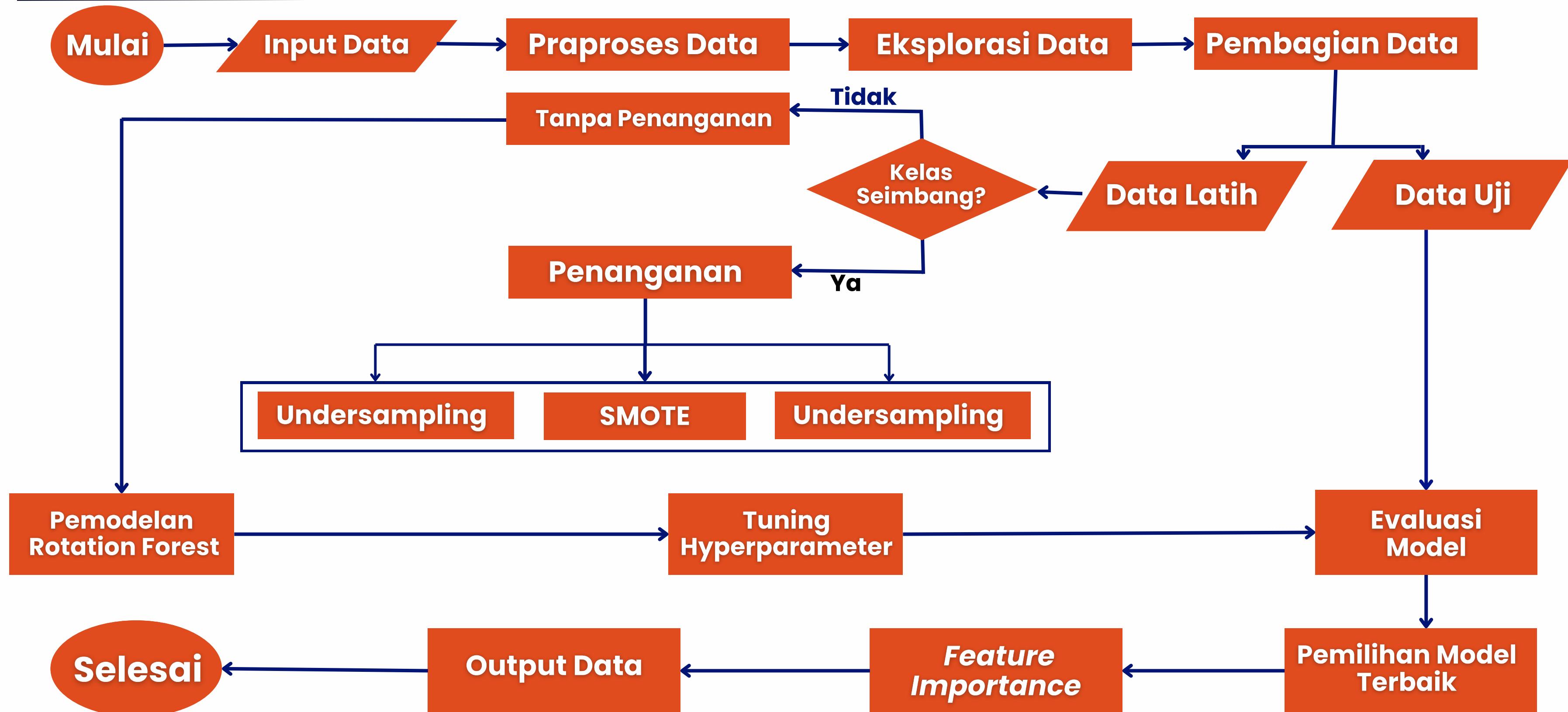
menghitung nilai yang mempresentasikan kontribusi masing-masing fitur terhadap hasil model

untuk memahami bagaimana variabel saling berinteraksi dalam mempengaruhi prediksi model.

Bentuk Model

$$g(z) = \phi_0 + \sum_{j=1}^M \phi_j z_j$$

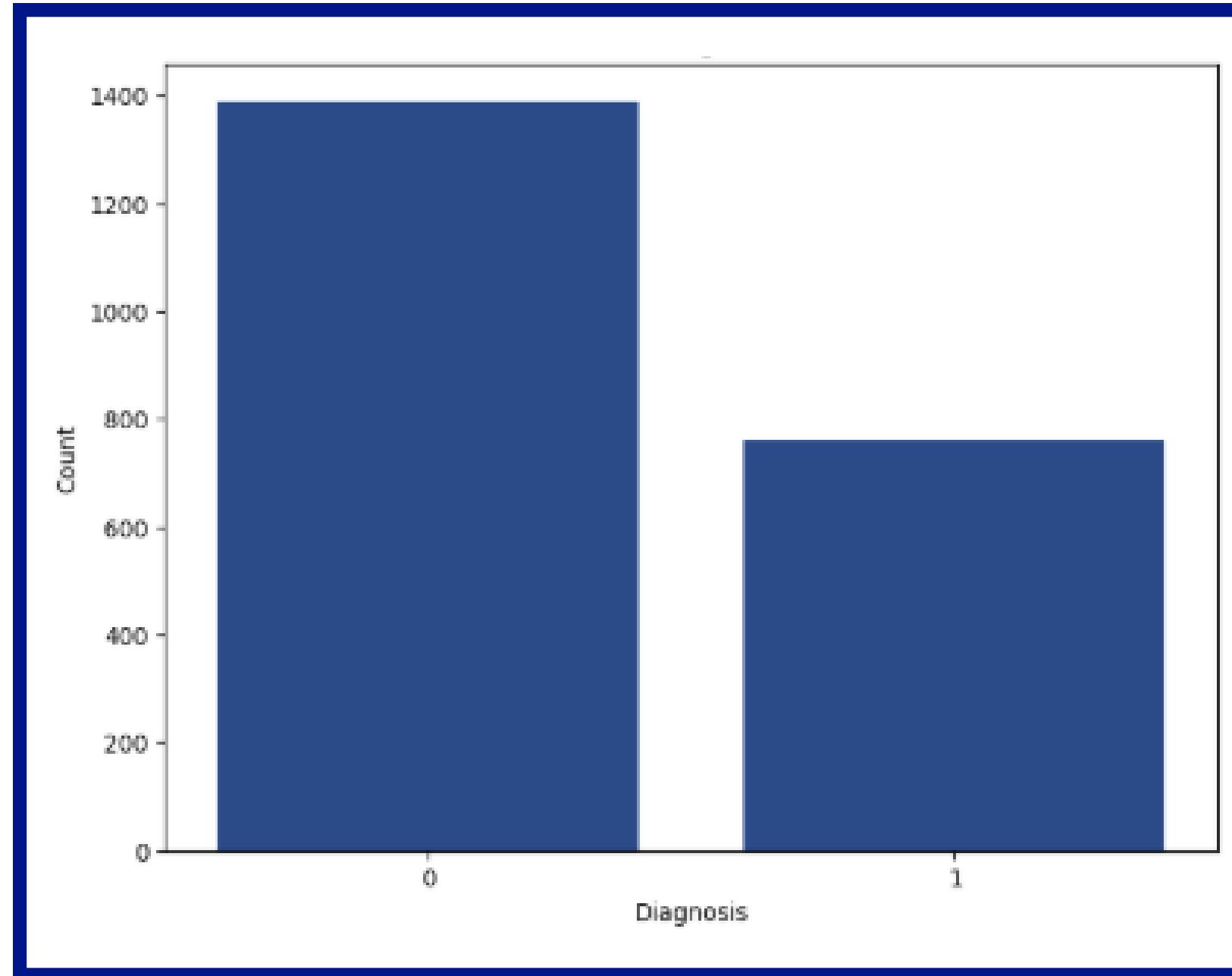
Metode Penelitian



Hasil & Pembahasan



Eksplorasi Peubah Respon



Distribusi data data
tidak seimbang



Penanganan data
tidak seimbang



Kelas 1: 35,36%
Alzheimer

Kelas 0: 64,63%
Non-Alzheimer

(Hasanah et al., 2023)

Algoritma dapat
bekerja secara efisien

SMOTE

Oversampling

Undersampling

Hasil & Pembahasan



Model Tanpa Penanganan

Parameter

21 k 15

Metrik Evaluasi	Data Latih	Data Uji	Perbedaan Metrik Kecil	Tidak Overfitting
Akurasi	0,9983	0,9581	Sensitivitas Data Uji	92,16% prediksi alzheimer benar
Sensitivitas	0,9983	0,9216	Spesifisitas Data Uji	97,83% prediksi nonalzheirmer benar
Spesifisitas	0,9982	0,9783	F1-Score Data Uji	Seimbang dalam klasifikasi antarkelas
F1-score	0,9969	0,9400		

Hasil & Pembahasan



Model Penanganan Keseimbangan Data

SMOTE

Parameter



l 19 k 15

Metrik Evaluasi

Data Latih

Data Uji

Akurasi

0,9987

0,9326

Sensitivitas

0,9982

0,9412

Spesifisitas

0,9278

0,9278

F1-score

0,9987

0,9085

Perbedaan Metrik
Kecil

Tidak Overfitting

Sensitivitas
Data Uji

94,12% prediksi
alzheimer benar

Spesifisitas
Data Uji

92,78% prediksi
nonalzheimer benar

F1-Score
Data Uji

Seimbang dalam
klasifikasi antarkelas

Hasil & Pembahasan



Model Penanganan Keseimbangan Data

Oversampling

Parameter

>>> 20 13

Metrik Evaluasi

Data Latih

0,9996

Data Uji

0,9442

Akurasi

Sensitivitas

Spesifisitas

F1-score

0,9991

0,9603

0,9996

0,9150

0,9603

0,9211

**Perbedaan Metrik
Kecil**

Tidak Overfitting

**Sensitivitas
Data Uji**

**91,50% prediksi
alzheirmer benar**

**Spesifisitas
Data Uji**

**96,03% prediksi
nonalzheirmer benar**

**F1-Score
Data Uji**

**Seimbang dalam
klasifikasi antarkelas**

Hasil & Pembahasan



Model Penanganan Keseimbangan Data

Undersampling

Parameter

***l* 18 *k* 15**

Metrik Evaluasi

Data Latih

0,9975

Data Uji

0,9419

Akurasi

0,9984

0,9412

Sensitivitas

0,9422

0,9422

Spesifisitas

0,9975

0,9201

F1-score

**Perbedaan Metrik
Kecil**

Tidak Overfitting

**Sensitivitas
Data Uji**

**94,12% prediksi
alzheimer benar**

**Spesifisitas
Data Uji**

**94,22% prediksi
nonalzheirmer benar**

**F1-Score
Data Uji**

**Seimbang dalam
klasifikasi antarkelas**

Hasil & Pembahasan



Perbandingan model terbaik

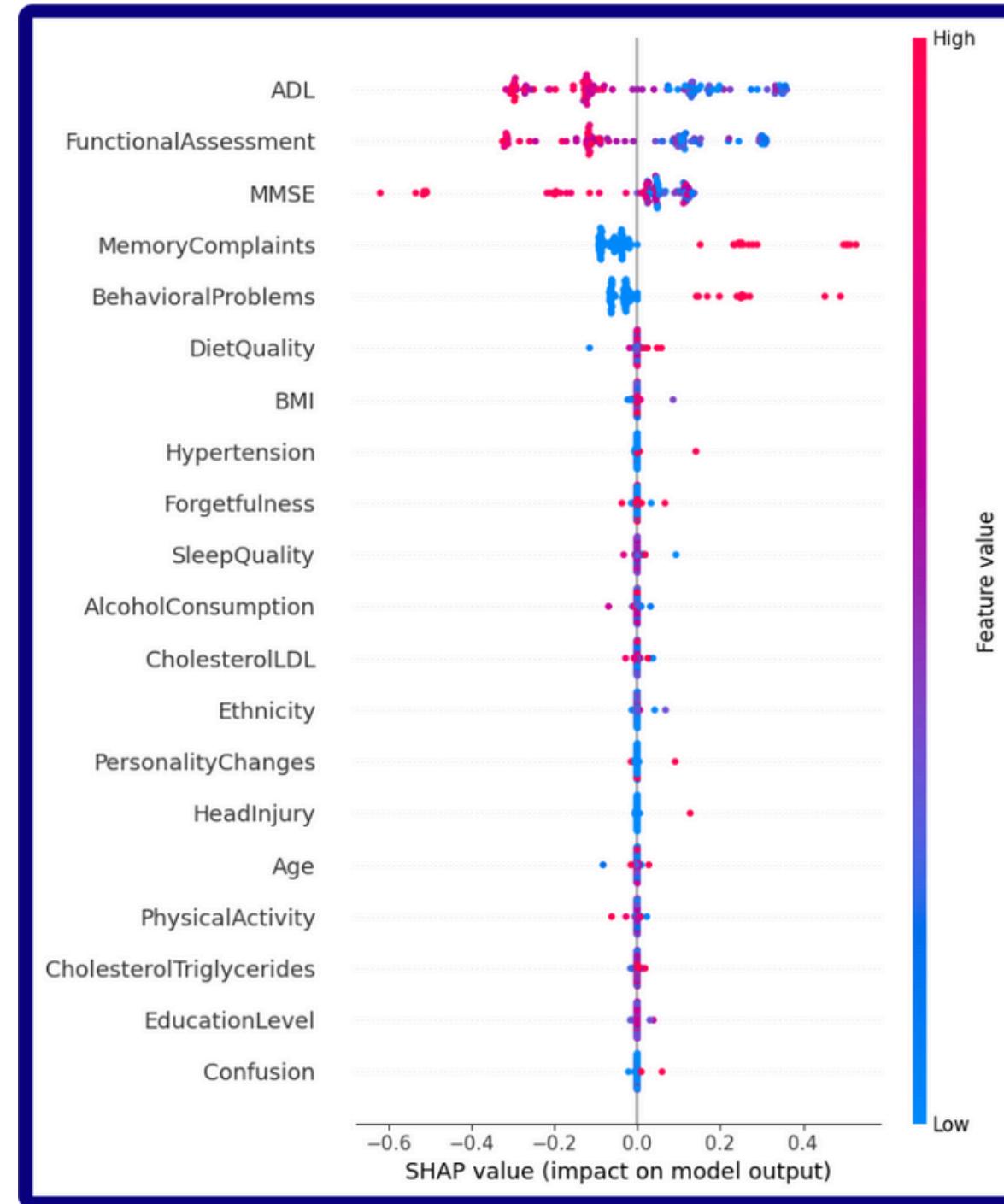
Model	Sensitivitas	Akurasi	
Rotation Forest tanpa Penanganan	0.9216	0.9581	
Rotation Forest SMOTE	0.9412	0.9326	
Rotation Forest Oversampling	0.9150	0.9159	
Rotation Forest Undersampling	0.9412	0.9412	

Rotation Forest dengan penanganan undersampling menjadi model terbaik karena unggul dari segi akurasi dan sensitivitas

Hasil & Pembahasan



Fitur Berpengaruh



5 Fitur paling berpengaruh terhadap diagnosa Alzheimer

Fitur

Koefisien

Pengaruh

Activity of Daily Living

0.186519

Functional Assessment

0.164055

Mini Mental State Examination

0.108771

Memory Complaints

0.099466

Behavioral Problems

0.066298

Kesimpulan



IPB University
Bogor Indonesia

Rotation Forest yang berhasil **mengklasifikasikan status** seseorang dalam mengidap alzheimer dengan **banyak fitur**

Penerapan **rotation forest** dengan penanganan **undersampling** menjadi model terbaik karena memiliki **sensitivitas 94,12%** dan **akurasi 94,12%** dengan parameter terbaik yang diperoleh adalah **L sebesar 18** dan **K sebesar 15.**

lima fitur yang paling berpengaruh:

Menurunkan peluang diagnosa positif

- ADL
- Functional Assessment
- MMSE

Meningkatkan peluang diagnosa positif

- memory complaints
- behavioral problems

Daftar Pustaka



- Aličković, E., & Subasi, A. (2017). Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Computing and Applications*, 28(4), 753–763. <https://doi.org/10.1007/s00521-015-2103-9>
- Apicella, A., Isgrò, F., & Prevete, R. (2024). Don't Push the Button! Exploring Data Leakage Risks in Machine Learning and Transfer Learning. <http://arxiv.org/abs/2401.13796>
- Aziz, M. U., Nasir, N., Malik, M., & Mufeez, M. M. (2025). Psychological effects of early diagnosis: Coping with alzheimer ' s disease and other dementias. *The Critical Review of Social*, 3(2), 1021–1036.
- Dharmendra, I. K., Agus, I. M., Putra, W., & Atmojo, Y. P. (2024). Evaluasi efektivitas SMOTE dan random under sampling pada klasifikasi emosi tweet. *Journal of Informatics*, 9(2), 192–193. <https://doi.org/https://doi.org/10.51211/itbi.v9i2.3183>
- Erlin, E., Desnelita, Y., Nasution, N., Suryati, L., & Zoromi, F. (2022). Dampak SMOTE terhadap kinerja random forest classifier berdasarkan data tidak seimbang. *MATRIX: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(3), 677–690. <https://doi.org/10.30812/matrik.v21i3.1726>
- Fadli, M., & Saputra, R. A. (2023). Klasifikasi dan evaluasi performa model Random Forest untuk prediksi stroke. *JT: Jurnal Teknik*, 12(02), 72–80. <http://jurnal.umt.ac.id/index.php/jt/index>
- Fonda, H., Irawan, Y., Melyanti, R., Wahyuni, R., & Muhammin, A. (2024). A comprehensive stacking ensemble approach for stress level classification in higher education. *Journal of Applied Data Sciences*, 5(4), 1701–1714. <https://doi.org/10.47738/jads.v5i4.388>

Daftar Pustaka



- Aličković, E., & Subasi, A. (2017). Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Computing and Applications*, 28(4), 753–763. <https://doi.org/10.1007/s00521-015-2103-9>
- Apicella, A., Isgrò, F., & Prevete, R. (2024). Don't Push the Button! Exploring Data Leakage Risks in Machine Learning and Transfer Learning. <http://arxiv.org/abs/2401.13796>
- Aziz, M. U., Nasir, N., Malik, M., & Mufeez, M. M. (2025). Psychological effects of early diagnosis: Coping with alzheimer ' s disease and other dementias. *The Critical Review of Social*, 3(2), 1021–1036.
- Dharmendra, I. K., Agus, I. M., Putra, W., & Atmojo, Y. P. (2024). Evaluasi efektivitas SMOTE dan random under sampling pada klasifikasi emosi tweet. *Journal of Informatics*, 9(2), 192–193. <https://doi.org/https://doi.org/10.51211/itbi.v9i2.3183>
- Erlin, E., Desnelita, Y., Nasution, N., Suryati, L., & Zoromi, F. (2022). Dampak SMOTE terhadap kinerja random forest classifier berdasarkan data tidak seimbang. *MATRIX: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(3), 677–690. <https://doi.org/10.30812/matrik.v21i3.1726>
- Fadli, M., & Saputra, R. A. (2023). Klasifikasi dan evaluasi performa model Random Forest untuk prediksi stroke. *JT: Jurnal Teknik*, 12(02), 72–80. <http://jurnal.umt.ac.id/index.php/jt/index>
- Fonda, H., Irawan, Y., Melyanti, R., Wahyuni, R., & Muhammin, A. (2024). A comprehensive stacking ensemble approach for stress level classification in higher education. *Journal of Applied Data Sciences*, 5(4), 1701–1714. <https://doi.org/10.47738/jads.v5i4.388>

Daftar Pustaka



- Hasanah, U., Muhammad Soleh, A., & Sadik, K. (2024). Pengaruh random undersampling, oversampling, dan SMOTE terhadap kinerja model prediksi penyakit kardiovaskular. *Jurnal Matematika, Statistika, Dan Komputasi*, 21(1), 88–102. <https://doi.org/10.20956/j.v21i1.35552>
- Iqbal, M. (2016). *Kajian Sensitivitas Parameter Algoritme pada Metode Rotation Forest [skripsi]*. Bogor(ID): IPB University.
- Irfan, M., Shahrestani, S., & Elkhodr, M. (2023). Enhancing early dementia detection: A machine learning approach leveraging cognitive and neuroimaging features for optimal predictive performance. *Applied Sciences (Switzerland)*, 13(18). <https://doi.org/10.3390/app131810470>
- Kwon, H. S. (2024). Revolutionizing Alzheimer's diagnosis and management: The dawn of biomarker-based precision medicine. 23(4), 188–201.
- Lisdiana, R., & Rofii, M. (2024). Tingkat kemandirian activity of daily living dengan pengukuran Indeks Katz pada pasien dengan gangguan mobilitas fisik di wilayah kerja puskesmas. *Holistic Nursing and Health Science*, 7(1), 16–29. <https://doi.org/10.14710/hnhs.7.1.2024.16-29>
- Liu, K. H., & Huang, D. S. (2008). Cancer classification using Rotation Forest. *Computers in Biology and Medicine*, 38(5), 601–610. <https://doi.org/10.1016/j.combiomed.2008.02.007>

Daftar Pustaka



- Maula Hidayat, F., Kusrini, & Ainul Yaqin. (2024). Penerapan algoritma Naïve Bayes untuk klasifikasi status gizi stunting pada balita. *Dielektrika*, 11(2), 107–118. <https://doi.org/10.29303/dielektrika.v11i2.384>
- Muhammad Hilmy Haidar Aly. (2024). Klasifikasi diabetes menggunakan algoritma support vector machine radial basis function. *Jurnal Teknik Informatika Dan Teknologi Informasi*, 4(1), 28–38. <https://doi.org/10.55606/jutiti.v4i1.3420>
- Mulianto, R. (2016). Kajian Empirik Akurasi Prediksi Klasifikasi Metode Rotation Forest [skripsi]. Bogor(ID): IPB University.
- Mumpuni, E., & Harison. (2024). Implementasi shap pada catboost untuk meningkatkan akurasi prediksi temperatur udara di kota pekanbaru. *Jurnal Studi Multidisipliner*, 8(6), 538–548.
- Naufal, A. R., Satria, R., & Syukur, A. (2015). Penerapan bootstrapping untuk ketidakseimbangan kelas dan weighted information gain untuk feature selection pada algoritma support vector machine untuk prediksi loyalitas pelanggan. *Journal of Intelligent Systems*, 1(2), 98–108.
- Nurfadia, M. F., Harahap, H. S., Rizki, M., Sahidu, G., & Hunaifi, I. (2021). Korelasi skor Mini Mental State Examination (MMSE) dan Montreal Cognitive Assessment versi Indonesia (MoCA-Ina) sebagai instrumen evaluasi fungsi kognitif. *Unram Medical Journal*, 10(1), 371–378. <https://journal.unram.ac.id/index.php/jku/article/view/4293>

Daftar Pustaka



IPB University
Bogor Indonesia

- Quinn, S. (2025). The urgent and global need for democratized blood-based biomarker diagnostics in alzheimer's disease. *The Journal of Precision Medicine: Health and Disease*, 1(March), 1–3. <https://doi.org/10.1016/j.premed.2025.100002>
- Rizky Mubarok, M., Herteno, R., Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat Jalan Ahmad Yani Km, I., & Selatan, K. (2022). Hyper-Parameter tuning pada Xgboost Untuk prediksi keberlangsungan hidup pasien gagal jantung. *Klik - Kumpulan Jurnal Ilmu Komputer*, 9(2), 391–401. <http://klik.ulm.ac.id/index.php/klik/article/view/484>
- Rodríguez-Pérez, R., & Bajorath, J. (2020). Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *Journal of Computer-Aided Molecular Design*, 34(10), 1013–1026. <https://doi.org/10.1007/s10822-020-00314-0>
- Rodríguez, J. J., Kuncheva, L. I., & Alonso, C. J. (2006). Rotation forest: A New classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10), 1619–1630. <https://doi.org/10.1109/TPAMI.2006.211>
- Sha, L., Rakovic, M., Das, A., Gasevic, D., & Chen, G. (2022). Leveraging class balancing techniques to alleviate algorithmic bias for predictive tasks in education. *IEEE Transactions on Learning Technologies*, 15(4), 481–492. <https://doi.org/10.1109/TLT.2022.3196278>

Daftar Pustaka



IPB University
Bogor Indonesia

- Sha, L., Rakovic, M., Das, A., Gasevic, D., & Chen, G. (2022). Leveraging class balancing techniques to alleviate algorithmic bias for predictive tasks in education. *IEEE Transactions on Learning Technologies*, 15(4), 481–492. <https://doi.org/10.1109/TLT.2022.3196278>
- Siagian, Y., Hutahaean, J., Zikra Syah, A., Efendi Hutagalung, J., & Karim, A. (2024). Implementasi metode K-Nearest Neighbours (KNN) untuk klasifikasi penyakit diabetes. *Jurnal Informatika Dan Teknologi Informasi*, 2(3), 253–262. <https://doi.org/10.56854/jt.v2i3.331>
- Syahwaluddin, R., & Alita, D. (2025). The Indonesian Journal of Computer Science. *The Indonesian Journal of Computer Science*, 14(4), 6615–6625. <https://doi.org/https://doi.org/10.33022/ijcs.v13i4.4295>
- Wijaya, J., Soleh, A. M., & Rizki, A. (2018). Penanganan Data Tidak Seimbang pada Pemodelan Rotation Forest Keberhasilan Studi Mahasiswa Program Magister IPB. *Xplore: Journal of Statistics*, 2(2), 32–40. <https://doi.org/10.29244/xplore.v2i2.99>

Terima kasih