

**IMPLEMENTASI SMOTE UNTUK MENINGKATKAN KINERJA
KLASIFIKASI *RANDOM FOREST* DALAM PENILAIAN RISIKO KREDIT**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
NAFA NUR ADIFIA NANDA
09040220060

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2024

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : NAFA NUR ADIFIA NANDA

NIM : 09040220060

Program Studi : Matematika

Angkatan : 2020

Menyatakan bahwa Saya tidak melakukan plagiat dalam penulisan skripsi Saya yang berjudul "IMPLEMENTASI SMOTE UNTUK MENINGKATKAN KINERJA KLASIFIKASI *RANDOM FOREST* DALAM PENILAIAN RISIKO KREDIT". Apabila suatu saat nanti terbukti Saya melakukan tindakan plagiat, maka Saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini Saya buat dengan sebenar-benarnya.

Surabaya, 28 Februari 2024

Yang menyatakan,



NAFA NUR ADIFIA NANDA
NIM. 09040220060

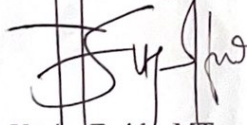
LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : NAFA NUR ADIFIA NANDA
NIM : 09040220060
Judul Skripsi : IMPLEMENTASI SMOTE UNTUK MENINGKATKAN
KINERJA KLASIFIKASI *RANDOM FOREST* DALAM
PENILAIAN RISIKO KREDIT

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I



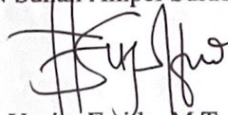
Yuniar Farida, MT
NIP. 197905272014032002

Pembimbing II



Wika Dianita Utami, M.Sc.
NIP. 199206102018012003

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya



Yuniar Farida, M.T.
NIP. 197905272014032002

PENGESAHAN TIM PENGUJI SKRIPSI

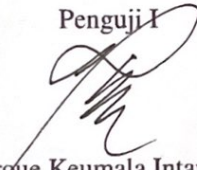
Skripsi oleh

Nama : NAFA NUR ADIFIA NANDA
NIM : 09040220060
Judul Skripsi : IMPLEMENTASI SMOTE UNTUK MENINGKATKAN
KINERJA KLASIFIKASI *RANDOM FOREST* DALAM
PENILAIAN RISIKO KREDIT

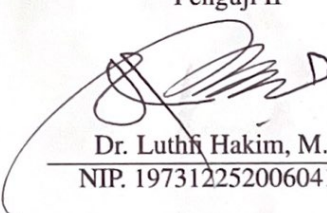
Telah dipertahankan di depan Tim Penguji
pada tanggal 7 Maret 2024

Mengesahkan,
Tim Penguji

Penguji I


Putroue Keumala Intan, M.Si
NIP. 198805282018012001

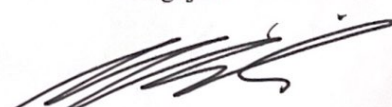
Penguji II


Dr. Luthfi Hakim, M.Ag
NIP. 197312252006041001

Penguji III


Yuniar Farida, MT
NIP. 197905272014032002


Penguji IV


Wika Dianita Utami, M.Sc.
NIP. 199206102018012003

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya




Dr. A. Saepul Hamdani, M.Pd.
NIP. 196507312000031002



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpustakaan@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : NAFA NUR ADIFIA NANDA
NIM : 09040220060
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / MATEMATIKA
E-mail address : adifiana13@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
 Skripsi Tesis Desertasi Lain-lain (.....)
yang berjudul :

IMPLEMENTASI SMOTE UNTUK MENINGKATKAN KINERJA KLASIFIKASI
RANDOM FOREST DALAM PENILAIAN RISIKO KREDIT

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 26 Maret 2024

Penulis

(NAFA NUR ADIFIA N)
nama terang dan tanda tangan

ABSTRAK

IMPLEMENTASI SMOTE UNTUK MENINGKATKAN KINERJA KLASIFIKASI *RANDOM FOREST* DALAM PENILAIAN RISIKO KREDIT

Kredit memegang peran penting dalam aktivitas perbankan, mendukung investasi, pertumbuhan usaha, dan pemenuhan keuangan. Namun, risiko kredit dapat timbul karena peminjam mungkin gagal memenuhi kewajiban pembayaran, sehingga diperlukan penilaian kelayakan seseorang dalam pemberian kredit, dengan melakukan klasifikasi untuk menentukan atau mempertimbangkan pemberian kredit tersebut. Penelitian ini bertujuan mengklasifikasi penilaian risiko kredit menggunakan metode *Random Forest*. *Random Forest* adalah algoritma pohon keputusan yang telah terbukti memiliki tingkat akurasi tinggi dalam klasifikasi data dengan pendekatan *ensemble* yang terdiri dari sejumlah pohon keputusan. Sebelum melakukan proses klasifikasi, seringkali dihadapkan pada tantangan ketika data tidak dapat diolah secara langsung karena mengalami ketidakseimbangan kelas. Untuk mengatasi ketidakseimbangan data kelas tersebut, penelitian ini menggunakan algoritma SMOTE (*Synthetic Minority Over-sampling Technique*). Algoritma SMOTE adalah pendekatan yang berfokus pada *oversampling*, yang bertujuan untuk meningkatkan jumlah data pada kelas minor dengan cara membuat data sintesis yang seimbang dengan data pada kelas minor. Hasil penelitian menunjukkan bahwa model optimal pada pembagian data *training* dan *testing* 80:20 serta penerapan metode SMOTE pada *Random Forest* berpengaruh dalam meningkatkan evaluasi kinerja. Nilai *accuracy* mengalami peningkatan dari 91.54% menjadi 94.41%. Nilai *precision* meningkat dari 90.83% menjadi 97.03%, dan nilai *recall* meningkat sangat signifikan dari 60.26% menjadi 91.55%. Hasil dari proses klasifikasi ini diharapkan akan memberikan kontribusi penting dalam pengembangan metode alternatif untuk mengklasifikasikan penilaian risiko kredit.

Kata kunci: Klasifikasi, *Random Forest*, SMOTE, Risiko Kredit.

ABSTRACT

IMPLEMENTATION OF SMOTE TO IMPROVE THE PERFORMANCE OF RANDOM FOREST CLASSIFICATION IN CREDIT RISK ASSESSMENT

Credit plays an important role in banking activities, supporting investment, business growth, and financial fulfillment. However, credit risk can arise because the borrower may fail to fulfill payment obligations, so it is necessary to assess the eligibility of a person in granting credit, by classifying to determine or consider granting the credit. This research aims to classify credit risk assessment using the Random Forest method. Random Forest is a decision tree algorithm that has been proven to have a high level of accuracy in data classification with an ensemble approach consisting of a number of decision trees. Before carrying out the classification process, it is often faced with challenges when data cannot be processed directly because it experiences class imbalance. To overcome the class imbalance, this research uses the SMOTE (Synthetic Minority Over-sampling Technique) algorithm. The SMOTE algorithm is an approach that focuses on oversampling, which aims to increase the amount of data in the minor class by creating synthetic data that is balanced with the data in the minor class. The results showed that the optimal model on the division of training and testing data 80:20 and the application of the SMOTE method in Random Forest had an effect in improving performance evaluation. The accuracy value increased from 91.54% to 94.41%. The precision value increased from 90.83% to 97.03%, and the recall value increased significantly from 60.26% to 91.55%. The results of this classification process are expected to make an important contribution to the development of alternative methods for classifying credit risk assessments.

Keywords: Classification, Random Forest, SMOTE, Credit Risk.

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI	iii
HALAMAN PERNYATAAN KEASLIAN	iv
MOTTO	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
ABSTRAK	xii
ABSTRACT	xiii
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	10
1.3. Tujuan Penelitian	10
1.4. Manfaat Penelitian	10
1.5. Batasan Masalah	11
1.6. Sistematika Penulisan	12
II TINJAUAN PUSTAKA	13
2.1. Kredit	13
2.2. Penilaian Risiko Kredit	14
2.3. Klasifikasi	19
2.4. Preprocessing	20
2.4.1. Normalisasi	21
2.4.2. SMOTE	22
2.5. Decision Tree	23
2.6. Algoritma Random Forest	25

2.7. Confusion Matrix	28
2.8. Integrasi Keilmuan	29
III METODE PENELITIAN	34
3.1. Jenis Penelitian	34
3.2. Sumber Data dan Variabel Penelitian	34
3.3. Tahapan Penelitian	36
IV HASIL DAN PEMBAHASAN	40
4.1. Deskripsi Data	40
4.1.1. Analisis Deskriptif	40
4.1.2. Pembobotan Data Kategorik	45
4.2. Pre-Processing	46
4.2.1. Normalisasi Data	46
4.2.2. Mengatasi Ketidakseimbangan Data dengan Metode SMOTE	48
4.3. Klasifikasi Risiko Kredit Dengan Algoritma <i>Random Forest</i>	54
4.3.1. Pembagian data	54
4.3.2. Menentukan Pohon	55
4.3.3. Pembangunan Pohon Keputusan Model <i>Random Forest</i>	57
4.4. Evaluasi Hasil	71
4.4.1. Evaluasi Hasil Berdasarkan Pembagian Data	71
4.4.2. Perbandingan Evaluasi Hasil Klasifikasi <i>Random Forest</i> Tanpa SMOTE dan Dengan SMOTE	73
4.5. Integrasi Keilmuan	76
V PENUTUP	80
5.1. Kesimpulan	80
5.2. Saran	81
DAFTAR PUSTAKA	81

DAFTAR TABEL

2.1 Confusion Matrix	28
3.1 Deskripsi Variabel Data	35
4.1 Analisis Deskriptif Data Numerik	40
4.2 Tabel Karakteristik Variabel <i>Home</i> Terhadap Status	41
4.3 Tabel Karakteristik Variabel <i>Intent</i> Terhadap Status	43
4.4 Tabel Karakteristik Variabel <i>Default</i> Terhadap Status	44
4.5 Hasil Pembobotan Data Kategorik	46
4.6 Data Setelah Dinormalisasi	47
4.7 Data Minoritas Kelas 1	49
4.8 Data Ke-164 Kelas 1	50
4.9 Data Kelas 1 dengan 1 Data Sintesis	51
4.10 Data Ke-3.792 Kelas 1	52
4.11 Data Kelas 1 dengan 2 Data Sintesis	53
4.12 Data Setelah SMOTE	54
4.13 Pembagian Data <i>Training</i> dan Data <i>Testing</i>	55
4.14 Hasil Penentuan Pohon	56
4.15 Sampel Data Acak Pembentukan Pohon Secara Manual	57
4.16 Data Cabang Kedua Kiri (<i>True</i>)	63
4.17 Data Cabang Kedua Kanan (<i>False</i>)	65
4.18 Data Cabang Ketiga pada <i>Intent</i>, ≤ 0.005 : <i>True</i>	68
4.19 Perbandingan Hasil Evaluasi Model	71
4.20 Confussion Matrix <i>Random Forest</i> Rasio 80:20	72
4.21 Perbandingan Hasil Kinerja <i>Random Forest</i> Tanpa SMOTE dan Dengan SMOTE	74

DAFTAR GAMBAR

2.1 Struktur <i>Decision Tree</i>	24
2.2 Arsitektur <i>Random Forest</i>	25
3.1 <i>Flowchart</i> Penelitian	39
4.1 Persebaran Data <i>Home</i>	41
4.2 Persebaran Data <i>Intent</i>	42
4.3 Persebaran Data <i>Default</i>	44
4.4 Persebaran Data <i>Status</i>	45
4.5 Persebaran Data <i>Status</i>	48
4.6 Pohon Keputusan <i>Random Forest</i> Cabang Utama	62
4.7 Pohon Keputusan <i>Random Forest</i> dengan kedalaman pohon $\equiv 2$	67
4.8 Pohon Keputusan <i>Random Forest</i> dengan Kedalaman Pohon $\equiv 3$	70
4.9 Pohon Keputusan <i>Random Forest</i> dengan Data Sampel Sebanyak 10	70
4.10 Grafik Perbandingan Hasil Evaluasi	74

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Abd. Manab, A. R. R. (2022). Vol.4 No.1 Februari 2022. *PELAKSANAAN THE FIVE C'S OF CREDIT ANALYSIS DALAM PEMBERIAN KREDIT PEMILIKAN RUMAH DI BANK TABUNGAN NEGARA CABANG PAMEKASAN*, 4(1).
- Alfani W.P.R., A., Rozi, F., and Sukmana, F. (2021). Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 6(1):155–160.
- Anisha, F., Dodi Vionanda, amalita, N., and Zilrahmi (2023). Application of Random Forest for The Classification Diabetes Mellitus Disease in RSUP Dr. M. Jamil Padang. *UNP Journal of Statistics and Data Science*, 1(2):45–52.
- Apriyanti, N. P. R., Putra, I. K. G. D., and Putra, I. M. S. (2020). Peramalan Jumlah Kecelakaan Lalu Lintas Menggunakan Metode Support Vector Regression. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 8(2):72.
- Armonica, A. (2022). Klasifikasi Jenis Persalinan pada Ibu Hamil dengan Metode Random Forest. *PHP Rosa - Prosiding Seminar Nasional*, pages 184–188.
- Arora, N. and Kaur, P. D. (2020). A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Applied Soft Computing Journal*, 86:105936.
- Arvirianty, A. (2018). Kasus Kredit Macet Tirta Amarta, Negara Rugi Rp 1,83 T. Technical report.

- Baskoro, B. B., Susanto, I., and Khomsah, S. (2021). Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR). *INISTA (Journal of Informatics Information System Software Engineering and Applications)*, 3(2):21–29.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. (1993). *Classification and Regression Trees (CART)*. Chapman And Hall, New York.
- Budianto, E. W. H. (2023). Bibliometric And Literature Review Of Financing Risk In Islamic Banking. *JPS (Jurnal Perbankan Syariah)*, 4(1):79–97.
- Butarbutar, S. A. G. (2023). TINJAUAN YURIDIS TERHADAP TINDAK PIDANA PENGGELAPAN PADA KASUS JAMINAN FIDUSIA DI INDONESIA (Studi Kasus Putusan PN Purwakarta Nomor 227/Pid.B/2012/PN.Pwk). *JURNAL RECTUM: Tinjauan Yuridis Penanganan Tindak Pidana*, 5(1):16.
- Cahyaningtyas, C., Nataliani, Y., and Widiyari, I. R. (2021). Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE. *Aiti*, 18(2):173–184.
- Cakra Pratama, T. (2018). Penerapan Metode K-Nearest Neighbour Dalam Menentukan Kelayakan Calon Nasabah Yang Layak Untuk Kredit Mobil (Studi Kasus : Pt. Astra International, Tbk-Toyota). *Jurnal Riset Komputer (JURIKOM)*, 5(4):402–408.
- Chicco, D., Starovoitov, V., and Jurman, G. (2021). The Benefits of the Matthews Correlation Coefficient (MCC) over the Diagnostic Odds Ratio (DOR) in Binary Classification Assessment. *IEEE Access*, 9(Mcc):47112–47124.

- Chyntia, M. H., Ratnawati, D. E., and Arwani, I. (2022). Analisis Sentimen berbasis Aspek terhadap Ulasan Hotel Tentrem Yogyakarta menggunakan Algoritma Random Forest Classifier. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(4):1702–1708.
- Damuri, A., Riyanto, U., Rusdianto, H., and Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6):219.
- DEMİR, S. and ŞAHİN, E. K. (2022). Evaluation of Oversampling Methods (OVER, SMOTE, and ROSE) in Classifying Soil Liquefaction Dataset based on SVM, RF, and Naïve Bayes. *European Journal of Science and Technology*, (34):142–147.
- Depari, D. H., Widiastiwi, Y., and Santoni, M. M. (2022). Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *Informatik : Jurnal Ilmu Komputer*, 18(3):239.
- Devika, R., Avilala, S. V., and Subramaniaswamy, V. (2019). Comparative study of classifier for chronic kidney disease prediction using naive bayes, KNN and random forest. *Proceedings of the 3rd International Conference on Computing Methodologies and Communication, ICCMC 2019*, (Iccmc):679–684.
- DISKOMINFO (2023). Berani dan Tegas, Bupati Indramayu Bongkar Kasus Kredit Macet Rp230 Miliar BPR KR. Technical report.
- Euis Saraswati, Yuyun Umaidah, and Apriade Voutama (2021). Penerapan Algoritma Artificial Neural Network untuk Klasifikasi Opini Publik Terhadap Covid-19. *Generation Journal*, 5(2):109–118.

- Farida, Y., Nurfadila, M. R., and Yuliati, D. (2022). Identifying Significant Factors Affecting the Human Development Index in East Java Using Ordinal Logistic Regression Model. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 6(3):476.
- Femi Dwi Astuti and Febri Nova Lenti (2021). Implementasi SMOTE untuk Mengatasi Imbalance Class pada Klasifikasi Car Evolution menggunakan K-NN. *Jurnal JUPITER*, 13:89–98.
- Gde Agung Brahma Suryanegara, Adiwijaya, and Mahendra Dwifebri Purbolaksono (2021). Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1):114–122.
- Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., and Budiarto, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEE Access*, 8:90847–90861.
- Indrawati, A. (2021). Penerapan Teknik Kombinasi Oversampling Dan Undersampling Untuk Mengatasi Permasalahan Imbalanced Dataset. *JIKO (Jurnal Informatika dan Komputer)*, 4(1):38–43.
- Ismamudi, I., Hartati, N., and Sakum, S. (2023). Peran Bank dan Lembaga Keuangan dalam Pengembangan Ekonomi: Tinjauan Literatur. *Jurnal Akuntansi Neraca*, 1(2):35–44.
- Jackins, V., Vimal, S., Kaliappan, M., and Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *Journal of Supercomputing*, 77(5):5198–5219.

- Kelvin, K., Mulyawan, B., and Sutrisno, T. (2020). Penerapan Algoritma C4.5 Untuk Penerimaan Kelayakan Kredit Pada Koperasi (Studi Kasus:Koperasi Xyz). *Jurnal Ilmu Komputer dan Sistem Informasi*, 8(1):109.
- Martha, S., Andani, W., and Rizki, S. W. (2022). Perbandingan Metode k-Nearest Neighbor, Regresi Logistik Biner, dan Pohon Klasifikasi pada Analisis Kelayakan Pemberian Kredit. *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*, 10(2):262–273.
- Nainggolan, I. O. and Christina (2022). Sistem Pendukung Keputusan Penerimaan Kelayakan Kredit Menggunakan Model 5C7P Di Bank Perkreditan Rakyat Berbasis Web. *Jurnal Teknologi Kesehatan Dan Ilmu Sosial (Tekesos)*, 4(1):22–35.
- Nirwana, A., Siregar, A., and Rahmat, R. (2022). Klasifikasi Permasalahan Kredit Macet Pada Bank Menggunakan Algoritma Decision Tree C4. 5. *Scientific Student Journal for Information, Technology and Science*, 3(1):43–50.
- Novinna, V. (2020). Perlindungan Konsumen dari Penyebarluasan Data Pribadi oleh Pihak Ketiga: Kasus Fintech Peer”To Peer Lending”. *Jurnal Magister Hukum Udayana (Udayana Master Law Journal)*, 9(1):92.
- Novitasari, D. C., Foeady, A. Z., Nariswari, R., Asyhar, A. H., Ulinuha, N., Farida, Y., Santi, D. R., Ilham, and Setiawan, F. (2020). Whirlwind Classification with Imbalanced Upper Air Data Handling using SMOTE Algorithm and SVM Classifier. *Journal of Physics: Conference Series*, 1501(1).
- Nurdeni, D. A. (2022). Extracting Information From Twitter Data To Identify Types of Assistance for Victims of Natural Disasters: an Indonesian Case Study. *Journal of Management Information and Decision Sciences*, 25(S1):1–14.

- Oon Wira Yuda, Darmawan Tuti, Lim Sheih Yee, and Susanti (2022). Penerapan Penerapan Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Random Forest. *SATIN - Sains dan Teknologi Informasi*, 8(2):122–131.
- Polat, K. (2019). A hybrid approach to parkinson disease classification using speech signal: The combination of smote and random forests. In *2019 Scientific Meeting on Electrical-Electronics Biomedical Engineering and Computer Science (EBBT)*, pages 1–3.
- Prabowo, I. A., Remawati, D., and Wardana, A. P. W. (2020). Klasifikasi Tingkat Gangguan Tidur Menggunakan Algoritma Naïve Bayes. *Jurnal Teknologi Informasi dan Komunikasi (TIKomSiN)*, 8(2).
- Prasetya, J. and Abdurakhman, A. (2023). Comparison of Smote Random Forest and Smote K-Nearest Neighbors Classification Analysis on Imbalanced Data. *Media Statistika*, 15(2):198–208.
- Prasetyo, P. D., Suta Wijaya, I. G. P., and Yudo Husodo, A. (2019). Klasifikasi Genre Musik Menggunakan Metode Mel-Frequency Cepstrum Coefficients dan K-Nearest Neighbors Classifier. *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)*, 1(2):189–197.
- Puspitasari, D., Al Khautsar, S. S., and Mustika, W. P. (2019). Algoritma Naïve Bayes Untuk Memprediksi Kredit Macet Pada Koperasi Simpan Pinjam. *Jurnal Informatika Upgris*, 4(2).
- Rahim, A. M. A., Pratiwi, I. Y. R., and Fikri, M. A. (2023). Klasifikasi penyakit jantung menggunakan metode synthetic minority over-sampling technique dan random forest classifier. *Indonesian Journal of Computer Science*, 12(5).

- Ramadhan, N. G. and Khoirunnisa, A. (2021). Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine. *Jurnal Media Informatika Budidarma*, 5(4):1580.
- Ramesh, T. R., Lilhore, U. K., Poongodi, M., Simaiya, S., Kaur, A., and Hamdi, M. (2022). Predictive Analysis of Heart Diseases With Machine Learning Approaches. *Malaysian Journal of Computer Science*, 2022(Special Issue 1):132–148.
- Reynaldo, Mulyawan, B., and Sutrisno, T. (2020). Rekomendasi Pemilihan Program Studi Tarumangara Menggunakan Metode. *Jurnal Komputer dan Informatika*, 15(1):326–333.
- Riska Chairunisa, Adiwijaya, and Widi Astuti (2020). Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(5):805–812.
- Rizal, M., Syahaf, M. Z., Priyambodo, S. R., and Rhamdani, Y. (2023). Optimasi Algoritma Naïve Bayes Menggunakan Forward Selection Untuk Klasifikasi Penyakit Ginjal Kronis. *Naratif : Jurnal Nasional Riset, Aplikasi dan Teknik Informatika*, 5(1):71–80.
- Rohimah, W., Budianto, E. W. H., and Dewi, N. D. T. (2023). Pemetaan Penelitian seputar Bank CIMB Niaga Syariah dan Konvensional: Studi Bibliometrik VOSviewer dan Literature Review. *JEMPER (Jurnal Ekonomi Manajemen Perbankan)*, 5(1):30.
- Sajjadnia, Z., Khayami, R., and Moosavi, M. R. (2020). Preprocessing Breast Cancer Data to Improve the Data Quality, Diagnosis Procedure, and Medical Care Services. *Cancer Informatics*, 19:7–12.

- Shaban, W. M., Rabie, A. H., Saleh, A. I., and Abo-Elsoud, M. A. (2020). A new COVID-19 Patients Detection Strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier. *Knowledge-Based Systems*, 205:106270.
- Sidik, A. D. and Ansawarman, A. (2022). Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine Learning. *Formosa Journal of Multidisciplinary Research*, 1(3):559–568.
- Sitepu, R. and Manohar, M. (2022). Implementasi Algoritma K-Nearest Neighbor Untuk Klasifikasi Pengajuan Kredit. *Jurnal Sistem Informasi, Teknik Informatika dan Teknologi Pendidikan*, 1(2):49–56.
- Sulistiyono, M., Pristyanto, Y., Adi, S., and Gumelar, G. (2021). Implementasi Algoritma Synthetic Minority Over-Sampling Technique untuk Menangani Ketidakseimbangan Kelas pada Dataset Klasifikasi. *Sistemasi*, 10(2):445.
- Sutoyo, E. and Fadlurrahman, M. A. (2020). Penerapan SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Television Advertisement Performance Rating Menggunakan Artificial Neural Network. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 6(3):379.
- Syariah, B. U., Syahriza, R., Ahmadi, N., and Rahmani, B. (2023). Faktor Eksternal dan Internal yang Mempengaruhi Risiko Pembiayaan Pada. 9(02):2294–2301.
- Syukron, A., Sardiarinto, S., Saputro, E., and Widodo, P. (2023). Penerapan Metode Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Gagal Jantung. *Jurnal Teknologi Informasi dan Terapan*, 10(1):47–50.
- Wijaya, A. E., Bani, R., Sukarni, S., and Weighting, S. A. (2019). *Jurnal Teknologi*

- Informasi dan Komunikasi STMIK Subang, Oktober 2019 ISSN: 2252-4517. (April):100–110.
- Wijayanti, N. N. T. and Purnamawati, I. G. A. (2022). Analisis Penerapan Sistem Informasi Akuntansi Dalam Pemberian Kredit Untuk Menekan Angka Kredit Macet pada LPD Seririt. *Jurnal Akuntansi Profesi*, 13(3):671–680.
- Wulandari, L., Farida, Y., Fanani, A., Ulinuha, N., and Intan, P. K. (2020). Evaluation of disadvantaged regions in east java based-on the 33 indicators of the ministry of villages, development of disadvantaged regions, and transmigration using the ensemble ROCK (Robust clustering using link) method. *Advances in Science, Technology and Engineering Systems*, 5(5):193–200.
- Yang, J., Gong, J., Tang, W., Shen, Y., Liu, C., and Gao, J. (2019). Delineation of urban growth boundaries using a patch-based cellular automata model under multiple spatial and socio-economic scenarios. *Sustainability (Switzerland)*, 11(21).
- Yunitaningtyas, K. and Yolanda, A. M. (2022). TIMUR BERDASARKAN INDIKATOR STATUS KESEHATAN Penyedia Data Statistik Berkualitas untuk Indonesia Maju. *Jurnal Statistika Terapan*, 2(1):19.
- Akbulut, Y. (2023). Automated Pneumonia Based Lung Diseases Classification with Robust Technique Based on a Customized Deep Learning Approach. *Diagnostics*, 13(2).
- Lasniari, S., Jasil, J., Sanjaya, S., Yanto, F., & Affandes, M. (2022). Klasifikasi Citra Daging Babi dan Daging Sapi Menggunakan Deep Learning Arsitektur ResNet-50 dengan Augmentasi Citra. *Jurnal Sistem Komputer dan Informatika (JSON)*, 3(4), 450.

Candra, R. M., & Rozana, A. (2020). Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor. *IT Journal Research and Development*, 5(1), 45–52.



UIN SUNAN AMPEL
S U R A B A Y A