

**KLASIFIKASI RISIKO KETAHANAN PANGAN MENURUT
KABUPATEN/KOTA DI INDONESIA MENGGUNAKAN
ALGORITMA *RANDOM FOREST***



**Oleh
M. FADEL
NIM 22037027**

**PROGRAM STUDI DIII STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2026**

**KLASIFIKASI RISIKO KETAHANAN PANGAN MENURUT
KABUPATEN/KOTA DI INDONESIA MENGGUNAKAN
ALGORITMA *RANDOM FOREST***

TUGAS AKHIR

Diajukan memenuhi sebagian persyaratan memperoleh gelar Ahli Madya



**Oleh
M. FADEL
NIM 22037027**

**PROGRAM STUDI DIII STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2026**

PERSETUJUAN TUGAS AKHIR

KLASIFIKASI RISIKO KETAHANAN PANGAN MENURUT KABUPATEN/KOTA DI INDONESIA MENGGUNAKAN ALGORITMA *RANDOM FOREST*

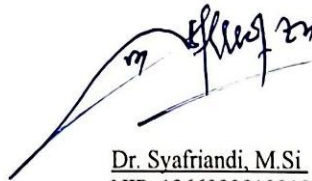
Nama : M. Fadel
NIM : 22037027
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Mengetahui:
Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005

Padang, 28 Januari 2026
Disetujui Oleh:
Pembimbing



Dr. Syafriandi, M.Si
NIP. 196609081991031003

PENGESAHAN LULUS UJIAN TUGAS AKHIR

Nama : M. Fadel
NIM : 22037027
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

KLASIFIKASI RISIKO KETAHANAN PANGAN MENURUT KABUPATEN/KOTA DI INDONESIA MENGGUNAKAN ALGORITMA *RANDOM FOREST*

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji
Departemen Statistika, Fakultas Matematika dan Ilmu
Pengetahuan Alam Universitas Negeri Padang

Padang, 28 Januari 2026

Tim Penguji,

	Nama	Tanda Tangan
1. Ketua	: Dr. Syafriandi, M.Si	1. 
2. Anggota	: Dr. Yenni Kurniawati, S.Si., M.Si	2. 
3. Anggota	: Fadhilah Fitri, S.Si., M.Stat	3. 

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : M. Fadel
NIM : 22037027
Program Studi : DIII Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, tugas akhir saya dengan judul **“Klasifikasi Risiko Ketahanan Pangan Menurut Kabupaten/Kota Di Indonesia Menggunakan Algoritma *Random Forest*”** adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 19840223201012205

Saya yang menyatakan,



M. Fadel
NIM. 22037027

**M. Fadel : Klasifikasi Risiko Ketahanan Pangan Menurut
Kabupaten/Kota di Indonesia Menggunakan Algoritma
*Random Forest***

ABSTRAK

Indonesia menghadapi tantangan serius dalam mewujudkan ketahanan pangan yang berkelanjutan. Permasalahan utama terletak pada kerentanan multidimensi yang dipicu oleh faktor struktural seperti ketimpangan sosial ekonomi, perubahan iklim, dan keterbatasan infrastruktur yang memperparah kerawanan pangan di wilayah dengan tingkat kemiskinan tinggi, rendahnya pendidikan, dan masalah kesehatan. Kompleksitas ini dipertegas oleh Indeks Ketahanan Pangan antar provinsi dengan kinerja terbaik (Bali) dan terburuk (Papua). Rendahnya capaian pada wilayah tertentu mencerminkan keterbatasan akses terhadap pangan bergizi, infrastruktur dasar, dan permasalahan sosial ekonomi. Penelitian ini bertujuan untuk mengetahui hasil klasifikasi dan variabel indikator yang paling berkontribusi dalam menghadapi risiko ketahanan pangan di Indonesia.

Penelitian ini menggunakan data sekunder tahun 2024 dari Badan Pangan Nasional dan Badan Pusat Statistik dengan tujuh variabel prediktor. Analisis dilakukan menggunakan algoritma *Random Forest* dengan penyeimbangan data melalui *Synthetic Minority Oversampling Technique* (SMOTE) pada data *training*. Data dibagi menjadi data *training* (70%) dan data *testing* (30%), lalu evaluasi model dilakukan menggunakan *confusion matrix*, sedangkan kontribusi variabel dianalisis menggunakan *Variable Importance Measure* (VIM) berdasarkan *Mean Decrease Gini* (MDG).

Hasil penelitian menunjukkan bahwa SMOTE efektif meningkatkan pemerataan kelas dan kemampuan model dalam mengenali wilayah rentan pangan. Model *Random Forest* dengan SMOTE menghasilkan performa klasifikasi yang lebih baik, ditunjukkan oleh peningkatan nilai *recall* menjadi 68,42% dan *F1-Score* sebesar 60,47% dibandingkan model tanpa SMOTE. Selain itu, hasil MDG menunjukkan Pola Pangan Harapan dan Prevalensi Ketidakcukupan Konsumsi Pangan sebagai variabel paling berkontribusi. Temuan ini menjadi dasar penting bagi pemerintah dalam menyusun kebijakan ketahanan pangan yang lebih terarah dan tepat sasaran.

Kata Kunci: Ketahanan pangan, Klasifikasi, *Random Forest*, SMOTE, *Variable Importance Measure*

**M. Fadel : Classification of Food Security Risk Across Districts/Cities
in Indonesia Using the *Random Forest* Algorithm**

ABSTRACT

Indonesia faces serious challenges in achieving sustainable food security. The main problem lies in multidimensional vulnerability triggered by structural factors such as socioeconomic inequality, climate change, and limited infrastructure, which exacerbate food insecurity in areas with high poverty rates, low education, and health problems. This complexity is emphasized by the Food Security Index between the best-performing provinces (Bali) and the worst-performing provinces (Papua). Low performance in certain regions reflects limited access to nutritious food, basic infrastructure, and socioeconomic issues. This study aims to determine the classification results and indicator variables that contribute most to addressing food security risks in Indonesia.

This study used secondary data from 2024 from the National Food Agency and the Central Statistics Agency (BPS) with seven predictor variables. The analysis was conducted using the Random Forest algorithm with data balancing using the Synthetic Minority Oversampling Technique (SMOTE) on the training data. The data were divided into training data (70%) and testing data (30%). Model evaluation was conducted using a confusion matrix, while variable contributions were analyzed using the Variable Importance Measure (VIM) based on the Mean Decrease Gini (MDG).

The results showed that SMOTE effectively improved class distribution and the model's ability to recognize food-vulnerable areas. The Random Forest model with SMOTE produced better classification performance, as indicated by an increase in recall to 68.42% and an F1-Score of 60.47% compared to the model without SMOTE. Furthermore, the MDG results indicated that Expected Food Patterns and Prevalence of Insufficient Food Consumption were the most contributing variables. These findings provide an important basis for the government in formulating more targeted and targeted food security policies.

Keywords: food security, classification, *Random Forest*, SMOTE, variable importance measure.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga peneliti dapat menyelesaikan tugas akhir ini dengan judul **“Klasifikasi Risiko Ketahanan Pangan Menurut Kabupaten/Kota Di Indonesia Menggunakan Algoritma *Random Forest* ”**. Tugas akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar Ahli Madya di Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang.

Penyusunan tugas akhir ini tidak lepas dari dukungan, bimbingan, serta bantuan dari berbagai pihak. Oleh karena itu, peneliti ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Syafriandi, M.Si., pembimbing Tugas Akhir sekaligus dosen penasehat akademik yang telah memberikan arahan, bimbingan, motivasi, dan nasehat yang bermanfaat kepada peneliti.
2. Ibu Dr. Yenni Kurniawati, S.Si., M.Si., Kepala Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang sekaligus dosen pembahas yang telah memberikan kritik dan saran demi kesempurnaan Tugas Akhir ini.
3. Ibu Fadhilah Fitri, S.Si., M.Stat, dosen pembahas yang telah memberikan kritik dan saran demi kesempurnaan Tugas Akhir ini.
4. Bapak Dodi Vionanda, M.Si., Ph.D., Koordinator Diploma III Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.

5. Bapak/Ibu dosen Departemen Statistika, yang telah mengajar, membimbing, dan memberikan ilmu kepada peneliti.
6. Kedua orang tua tercinta, serta keluarga yang telah memberikan dukungan moral, spiritual, dan material yang tiada henti selama peneliti menempuh pendidikan hingga terselesaikan tugas akhir ini.
7. Teman-teman seperjuangan yang selalu memberikan semangat, kerja sama, dan kebersamaan yang sangat berarti bagi peneliti.

Peneliti menyadari bahwa tugas akhir ini masih memiliki kekurangan dan keterbatasan. Oleh karena itu, peneliti sangat mengharapkan kritik dan saran yang membangun untuk dijadikan sebagai bahan evaluasi. Semoga penyusunan tugas akhir ini dapat memberikan manfaat bagi pembaca.

Padang, Januari 2026

M. Fadel

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	vii
DAFTAR TABEL	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	8
C. Rumusan Masalah	9
D. Tujuan Penelitian.....	9
E. Manfaat Penelitian.....	9
BAB II KERANGKA TEORITIS	10
A. Kajian Teori	10
1. Ketahanan Pangan.....	10
2. Indeks Ketahanan Pangan	11
3. Faktor yang Diduga Mempengaruhi Ketahanan Pangan	14
4. <i>Machine learning</i>	21
5. Klasifikasi	23
6. Data Tidak Seimbang.....	24
7. <i>Synthetic Minority Oversampling Technique</i>	26
8. <i>Random Forest</i>	28
9. Evaluasi Model.....	34
B. Penelitian Relevan.....	36
BAB III METODOLOGI PENELITIAN	38
A. Jenis Penelitian.....	38
B. Jenis dan Sumber Data	38
C. Variabel Penelitian	38
D. Teknik Analisis Data.....	39
E. Diagram Alir (<i>Flowchart</i>)	41
BAB IV HASIL DAN PEMBAHASAN	43
A. Deskripsi Data.....	43

B. Hasil Analisis	51
C. Pembahasan.....	58
BAB V PENUTUP	63
A. Kesimpulan	63
B. Saran.....	64
DAFTAR PUSTAKA	66
LAMPIRAN.....	71

DAFTAR GAMBAR

Gambar	Halaman
1. Indeks Ketahanan Pangan Menurut Provinsi di Indonesia Tahun 2024	3
2. Ilustrasi Metode <i>Undersampling</i> dan <i>Oversampling</i>	25
3. Ilustrasi Metode SMOTE	27
4. Struktur <i>Random Forest</i>	29
5. <i>Flowchart</i> Alur Penelitian	42
6. Persentase Kategori Ketahanan Pangan	43
7. <i>Boxplot</i> SPPH Berdasarkan Status Ketahanan Pangan	44
8. <i>Boxplot</i> Air Layak Berdasarkan Status Ketahanan Pangan	45
9. <i>Boxplot</i> PKKP Berdasarkan Status Ketahanan Pangan	46
10. <i>Boxplot</i> Sanitasi Layak Berdasarkan Status Ketahanan Pangan	47
11. Rata – Rata RLS Berdasarkan Status Ketahanan Pangan	48
12. Rata – Rata Produk Domestik Regional Bruto per Kapita Berdasarkan Status Ketahanan Pangan	49
13. Rata – Rata IPM Berdasarkan Status Ketahanan Pangan	50
14. VIM Berdasarkan MDG	54

DAFTAR TABEL

Tabel	Halaman
1. <i>Cut off Point</i> Skor Komposit.....	12
2. Penyederhanaan Kategori Risiko Ketahanan Pangan	13
3. <i>Confusion Matrix</i>	35
4. Variabel Penelitian	39
5. Struktur Data	39
6. Perbandingan Proporsi Pada Kelas	51
7. Perbandingan Proporsi Pada Kelas Setelah SMOTE	52
8. <i>Tuning</i> Parameter Random Forest.....	53
9. Nilai <i>Confusion Matrix</i> Random Forest.....	56
10. Perbandingan Nilai <i>Accuracy, Precision, Recall, Specificity</i> , dan <i>F1- Score</i> . 57	

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Penelitian	71
2. <i>Syntax</i> Penelitian	71
3. <i>Output Syntax</i>	77
4. <i>Link Python Syntax</i> Penelitian.....	79

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Indonesia dikenal sebagai negara agraris dengan sumber daya alam yang melimpah. Namun, kondisi ini belum sepenuhnya menjamin kemudahan akses pangan bagi seluruh masyarakat (Rachman *et al.*, 2021). Sebagian besar penduduk menggantungkan hidup pada sektor pertanian, yang didukung oleh iklim tropis serta ketersediaan lahan subur yang luas. Menurut Tian *et al.* (2021), lahan yang subur memiliki potensi besar untuk ditanami tanaman pangan sebagai sumber utama makanan pokok. Oleh karena itu, ketersediaan pangan perlu dijaga secara berkelanjutan guna memenuhi kebutuhan gizi masyarakat.

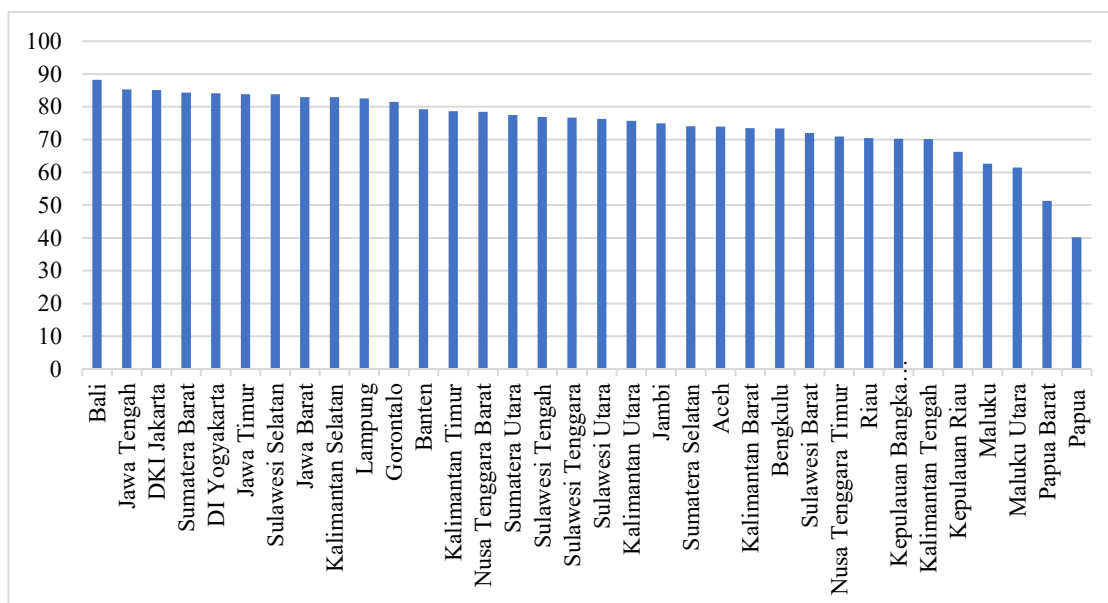
Menurut Badan Pangan Nasional, pangan merupakan kebutuhan dasar manusia yang dipenuhi setiap saat, dan hak ini dijamin oleh Undang-Undang Dasar Negara Republik Indonesia sebagai elemen fundamental dalam upaya menciptakan manusia yang berkualitas. Menurut Undang-Undang Republik Indonesia Nomor 18 Tahun 2012, pangan mencakup semua produk yang berasal dari sumber daya hewani dan hayati, baik yang sudah diolah maupun masih mentah, yang diperuntukkan bagi konsumsi manusia. Pangan juga meliputi bahan tambahan, bahan baku, serta unsur lain yang berperan dalam proses pengolahan makanan dan minuman untuk konsumsi manusia

Chaereni *et al.* (2020) menyatakan bahwa ketahanan pangan merupakan kondisi di mana setiap individu, baik secara fisik, sosial, maupun ekonomi, memiliki akses yang memadai terhadap pangan yang cukup, aman, bergizi, dan terjangkau. Berdasarkan Undang-Undang Nomor 7 Tahun 1996 tentang pangan,

ketahanan pangan mencakup ketersediaan pangan yang cukup dalam jumlah dan mutu, serta distribusi yang merata untuk memenuhi kebutuhan penduduk. Fokus dari ketahanan pangan adalah untuk meningkatkan ketersediaan pangan, pemantapan distribusi pangan, percepatan penganeekaragaman pangan, dan pengawasan keamanan pangan segar. Pembangunan ketahanan pangan dilaksanakan sebagai upaya untuk meningkatkan pertumbuhan ekonomi dan penurunan kemiskinan sebagai perwujudan pembangunan sosial, budaya, dan ekonomi sebagai bagian pembangunan secara keseluruhan.

Di Indonesia, persoalan ketahanan pangan masih menjadi tantangan serius, terutama di wilayah-wilayah yang memiliki tingkat kemiskinan tinggi, keterbatasan akses air bersih, rendahnya tingkat pendidikan, serta masalah kesehatan seperti stunting. Berbagai tantangan seperti ketimpangan sosial ekonomi, perubahan iklim, dan infrastruktur yang belum merata memperparah kerawanan pangan, terutama di wilayah terpencil dan miskin (Badan Pangan Nasional, 2023). Kompleksitas ini menuntut adanya pendekatan analitis yang mampu menangkap keragaman faktor secara menyeluruh dan akurat.

Salah satu alat penting yang digunakan untuk memantau kondisi ketahanan pangan di Indonesia adalah *Food Security and Vulnerability Atlas* (FSVA). FSVA merupakan peta tematik yang menggambarkan tingkat ketahanan dan kerentanan pangan di setiap kabupaten/kota, disusun oleh Badan Pangan Nasional bersama *World Food Programme* (WFP). FSVA mengukur ketahanan pangan berdasarkan tiga aspek utama, yaitu ketersediaan, akses, dan pemanfaatan pangan. Untuk melihat kondisi ketahanan pangan di Indonesia, Gambar 1 menyajikan indeks ketahanan pangan per provinsi.



Sumber : Badan Pangan Nasional (2024)

Gambar 1. Indeks Ketahanan Pangan Menurut Provinsi di Indonesia Tahun 2024

Berdasarkan Gambar 1, terlihat bahwa provinsi di Indonesia memiliki indeks ketahanan pangan yang beragam. Secara umum, nilai indeks berkisar antara 40 hingga lebih dari 85. Jika mengacu pada standar kategori Indeks Ketahanan Pangan nasional, nilai 80 atau lebih termasuk kategori tinggi, nilai antara 65 hingga kurang dari 80 termasuk kategori sedang, dan nilai di bawah 65 termasuk kategori rendah. Provinsi dengan skor tertinggi adalah Bali, yang menggambarkan kondisi ketahanan pangan yang baik karena didukung oleh ketersediaan pangan yang memadai, infrastruktur distribusi yang baik, serta layanan kesehatan yang merata. Sementara itu, provinsi dengan nilai indeks rendah seperti Papua Barat dan Papua menunjukkan masih adanya tantangan dalam mewujudkan ketahanan pangan yang inklusif dan berkelanjutan. Rendahnya capaian tersebut mencerminkan keterbatasan akses terhadap pangan bergizi, rendahnya infrastruktur dasar, serta permasalahan sosial ekonomi dan kesehatan masyarakat.

Kesenjangan nilai indeks ini berkaitan erat dengan Tujuan Pembangunan Berkelanjutan (*Sustainable Development Goals/SDGs*) ke-2, yaitu *Zero Hunger*, yang menargetkan penghapusan kelaparan, pencapaian ketahanan pangan, perbaikan gizi, dan peningkatan pertanian berkelanjutan. Oleh karena itu, ketimpangan indeks ketahanan pangan antarprovinsi di Indonesia menjadi isu strategis dalam pencapaian SDGs 2030, khususnya dalam memastikan tidak ada daerah yang tertinggal (*no one left behind*) dalam akses terhadap pangan dan gizi.

Perbedaan nilai indeks ini menegaskan adanya kesenjangan ketahanan pangan antarwilayah di Indonesia. Hal ini sejalan dengan permasalahan struktural seperti ketimpangan pembangunan, perbedaan kondisi geografis, dan distribusi sumber daya yang belum merata. Oleh karena itu, kebijakan ketahanan pangan perlu diarahkan secara lebih fokus pada daerah dengan indeks rendah agar tercapai pemerataan akses pangan, peningkatan kualitas gizi, dan penguatan daya tahan masyarakat terhadap kerentanan pangan.

Dalam penelitian ini, kategori risiko ketahanan pangan diperoleh dari pemetaan yang dilakukan oleh Badan Pangan Nasional melalui FSVA (*Food Security and Vulnerability Atlas*) yang membagi wilayah ke dalam beberapa tingkat risiko ketahanan pangan (Badan Pangan Nasional, 2024). Kategori risiko tersebut digunakan sebagai variabel dependen. Sementara itu, variabel independen yang digunakan terdiri atas Skor Pola Pangan Harapan (SPPH), Prevalensi Ketidacukupan Konsumsi Pangan (PKKP), akses air minum layak, akses sanitasi layak, rata-rata lama sekolah, Produk Domestik Regional Bruto (PDRB) per kapita, dan Indeks Pembangunan Manusia (IPM). Pemilihan variabel-variabel tersebut didasarkan pada kajian empiris sebelumnya yang menunjukkan keterkaitannya

dengan kondisi ketahanan pangan di suatu wilayah (Firdaus & Cahyono, 2017; Rachmaningsih & Priyarsono, 2012; Young *et al.*, 2021; Sartika & Cahyono, 2025).

Menurut Badan Pangan Nasional, risiko ketahanan pangan dikelompokkan ke dalam enam kategori, yaitu Sangat Rentan, Rentan, Agak Rentan, Agak Tahan, Tahan, dan Sangat Tahan. Pada penelitian ini, kategori tersebut disederhanakan menjadi dua kelas utama, yaitu Risiko Rentan (Sangat Rentan, Rentan, dan Agak Rentan) dan Risiko Tahan (Agak Tahan, Tahan, dan Sangat Tahan). Penyederhanaan ini dilakukan untuk mempermudah proses analisis dan interpretasi hasil klasifikasi, serta membantu mengidentifikasi wilayah prioritas yang membutuhkan perhatian lebih dalam peningkatan ketahanan pangan di Indonesia.

Distribusi wilayah pada tiap kategori risiko ketahanan pangan tidak merata dan variabel-variabel penjelas yang digunakan saling berkaitan serta tidak bersifat linier, sehingga sulit dianalisis hanya dengan deskripsi sederhana. Kondisi ini menunjukkan perlunya metode klasifikasi yang mampu mengolah banyak variabel sekaligus dan menangani ketidakseimbangan kelas, agar risiko ketahanan pangan setiap wilayah dapat dipetakan secara lebih akurat. Hasil klasifikasi tersebut penting untuk membantu pemerintah menentukan wilayah prioritas serta merumuskan intervensi kebijakan ketahanan pangan yang lebih tepat sasaran.

Klasifikasi merupakan salah satu metode dalam *machine learning* yang bertujuan mengelompokkan data ke dalam kategori tertentu berdasarkan pola yang ditemukan dalam data. Berdasarkan teknik pembelajarannya, *machine learning* dibedakan menjadi *supervised learning* dan *unsupervised learning* (Mailund, 2017; Goldstein *et al.*, 2017; Nelli, 2015). *Unsupervised learning* adalah pembelajaran tanpa pengawasan, dimana algoritma mencoba untuk memahami struktur atau pola

dalam data tanpa memerlukan informasi label. *Unsupervised learning* terbagi atas cluster dan reduksi dimensi. Sementara itu, *supervised learning* adalah metode pembelajaran yang menggunakan data berlabel untuk melatih algoritma dalam melakukan prediksi atau klasifikasi. Jenis supervised learning terbagi menjadi dua yaitu regresi dan klasifikasi. Contoh regresi seperti *Support Vector Regression* (SVR), sedangkan metode klasifikasi antara lain *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine* (SVM).

Random Forest merupakan algoritma *machine learning* yang banyak digunakan dalam tugas klasifikasi karena kemampuannya dalam menangani data berdimensi tinggi, bersifat non-linier, serta data yang tidak seimbang (Liaw & Wiener, 2002). Algoritma ini bekerja dengan membangun sejumlah pohon keputusan (*decision trees*) secara acak, kemudian menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan prediksi akhir yang lebih akurat (Khalilia *et al.*, 2011). Salah satu keunggulan *Random Forest* adalah kemampuan dalam mengukur pentingnya setiap variabel melalui *variable importance*, yang membantu dalam memahami faktor-faktor utama yang memengaruhi hasil klasifikasi. Misalnya, dalam penelitian Ramayanti *et al.*, (2023) menggunakan algoritma ini untuk mengklasifikasikan kemiskinan dan menemukan variabel kunci di Sumatera Barat. Selain itu, Komang *et al.*, (2024) membuktikan efektivitas *Random Forest* dalam memprediksi produksi padi sebagai bagian dari ketahanan pangan nasional.

Meskipun *Random Forest* cukup tangguh dalam menangani data tidak seimbang, distribusi kelas yang timpang tetap dapat memengaruhi performa model. Ketidakseimbangan ini biasanya terjadi ketika jumlah data pada kelas mayoritas jauh lebih besar dibandingkan kelas minoritas seperti distribusi indeks ketahanan

pangan pada Gambar 1. Menurut Jian *et al.* (2016), model yang dilatih pada data tidak seimbang cenderung memiliki akurasi rendah dalam memprediksi kelas minoritas, karena informasi dari kelas mayoritas mendominasi proses pembelajaran. Akibatnya, batas keputusan yang dihasilkan menjadi bias terhadap kelas mayoritas.

Oleh karena itu, jika ditemukan ketidakseimbangan dalam data, diperlukan teknik penyeimbangan kelas untuk meningkatkan kinerja model klasifikasi. Berbagai teknik dapat digunakan, seperti *Random Oversampling* (ROS), *Synthetic Minority Oversampling Technique* (SMOTE), *Adaptive Synthetic Sampling* (ADASYN), *Random Undersampling*, serta metode penyeimbangan lainnya. Di antara berbagai metode tersebut, penelitian ini menggunakan SMOTE karena teknik ini tidak hanya menambah jumlah sampel minoritas, tetapi juga menghasilkan data sintetis baru melalui interpolasi antar sampel yang serupa sehingga mampu merepresentasikan pola kelas minoritas secara lebih baik tanpa menduplikasi data asli. Dengan menerapkan SMOTE sebelum proses pelatihan model, distribusi kelas menjadi lebih seimbang dan memungkinkan *Random Forest* mempelajari karakteristik setiap kelas secara lebih adil, sehingga hasil klasifikasi menjadi lebih optimal dan tidak bias terhadap kelas mayoritas.

Beberapa studi juga membuktikan bahwa kombinasi algoritma *Random Forest* dengan SMOTE efektif dalam mengatasi ketidakseimbangan kelas pada data klasifikasi. Penelitian yang dilakukan oleh Sholihah & Hermawan (2023), di Kota Cirebon menggunakan *Random Forest* dan SMOTE untuk mengklasifikasikan status ekonomi rumah tangga, di mana kelas rumah tangga berpendapatan tinggi jauh lebih sedikit dibanding kelas rendah dan menengah. Hasil penelitian

menunjukkan bahwa penerapan SMOTE mampu meningkatkan performa *Random Forest* secara signifikan, dengan akurasi 93% dan nilai *recall* yang tinggi, khususnya dalam mendeteksi kelas minoritas yang sebelumnya sulit dideteksi.

Namun, hingga kini belum banyak penelitian yang menerapkan metode ini dalam studi ketahanan pangan di Indonesia. Oleh karena itu, penelitian ini diharapkan dapat mengisi kesenjangan tersebut dengan menyediakan model klasifikasi yang mendukung perencanaan dan pengambilan kebijakan yang lebih presisi, terlebih di wilayah status ketahanan pangan. Kelebihan *Random Forest* untuk mengidentifikasi variabel paling berpengaruh menjadikannya pilihan yang tepat untuk kondisi ketahanan pangan. Hasil klasifikasi yang dihasilkan dapat dimanfaatkan sebagai sistem peringatan dini dan dasar pengalokasian sumber daya pangan secara lebih tepat sasaran.

Berdasarkan uraian latar belakang masalah di atas, maka dilakukan penelitian dengan judul **“Klasifikasi Risiko Ketahanan Pangan Menurut Kabupaten/Kota Di Indonesia Menggunakan Algoritma *Random Forest* ”**

B. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Data yang di analisis adalah Indeks Ketahanan Pangan Kabupaten/Kota di Indonesia yang bersumber dari Badan Pangan Indonesia tahun 2024.
2. Metode klasifikasi yang diterapkan adalah *Random Forest* menggunakan SMOTE.
3. Klasifikasi yang digunakan dibagi 2 kategori yaitu rentan pangan dan tahan pangan.

C. Rumusan Masalah

Rumusan masalah pada penelitian ini sebagai berikut.

1. Bagaimana klasifikasi risiko ketahanan pangan Kabupaten/Kota di Indonesia menggunakan algoritma *Random Forest* sebelum dan sesudah SMOTE?
2. Variabel indikator mana saja yang paling berkontribusi dalam klasifikasi risiko ketahanan pangan menggunakan analisis *Random Forest*?

D. Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut.

1. Mengetahui klasifikasi risiko ketahanan pangan Kabupaten/Kota di Indonesia menggunakan algoritma *Random Forest* sebelum dan sesudah SMOTE.
2. Mengidentifikasi variabel indikator yang paling berkontribusi dalam klasifikasi risiko ketahanan pangan di Indonesia.

E. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut.

1. Bagi Peneliti, penelitian ini dapat menambah pemahaman dan keahlian dalam machine learning terutama metode *Random Forest*.
2. Bagi Pemerintah, penelitian ini dapat membantu pemerintah mengidentifikasi penyebab utama yang berkontribusi terhadap ketahanan pangan. Dengan memahami karakteristik dan pola yang muncul dari data, pemerintah dapat merancang kebijakan dan program intervensi yang lebih efektif dan tepat sasaran untuk mencegah ketimpangan ketahanan pangan.
3. Bagi Pembaca, penelitian ini diharapkan dapat menjadi tambahan referensi dan memberikan pembaca wawasan terkait keunggulan dan kelemahan metode dan menyesuaikan pendekatan yang lebih tepat.

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Penerapan algoritma *Random Forest* yang dikombinasikan dengan metode *Synthetic Minority Oversampling Technique* (SMOTE) terbukti lebih efektif dalam mengklasifikasikan risiko ketahanan pangan dibandingkan model tanpa penyeimbangan data. Penggunaan SMOTE berhasil mengoreksi bias model terhadap kelas mayoritas (Tahan), ditunjukkan dengan meningkatnya kemampuan deteksi wilayah rentan (*True Positive*) dari 9 wilayah menjadi 13 wilayah. Selain itu, model ini berhasil menekan risiko kegagalan deteksi wilayah berisiko (*False Negative*) dari 10 wilayah menjadi 6 wilayah. Secara keseluruhan, kualitas klasifikasi mengalami peningkatan kualitas prediksi yang lebih adil dan representatif. Meskipun akurasi mengalami sedikit penurunan dari 91,61% menjadi 89,03%, namun performa model dalam menangkap sinyal kerawanan pangan meningkat drastis, tercermin dari kenaikan nilai *Recall* sebesar 68,42% dan *F1-Score* sebesar 60,47%. Hal ini membuktikan bahwa model yang dihasilkan lebih sensitif terhadap karakteristik wilayah berisiko tinggi sehingga lebih handal digunakan sebagai instrumen deteksi dini (*early warning system*). Dengan demikian, penggunaan SMOTE dalam algoritma *Random Forest* di Indonesia dapat meningkatkan performa *Random Forest* dalam mengklasifikasikan risiko ketahanan pangan kabupaten/kota di Indonesia Tahun 2024.

2. Berdasarkan analisis menggunakan algoritma *Random Forest* dengan rata-rata nilai *Mean Decrease Gini* (MDG), dapat disimpulkan bahwa atribut utama yang paling mempengaruhi risiko ketahanan pangan di Indonesia tahun 2024 adalah Prevalensi Ketidacukupan Konsumsi Pangan (PKKP) dan Skor Pola Pangan Harapan (SPPH). Data menunjukkan bahwa wilayah yang memiliki PKKP yang rendah dan SPPH tinggi memiliki kemungkinan yang lebih kecil untuk rentan pangan. Atribut ini menjadi sebagai faktor pembeda utama dibandingkan dengan atribut-atribut lainnya. Oleh karena itu, atribut PKKP dan SPPH menjadi pertimbangan paling signifikan dalam menentukan klasifikasi status ketahanan pangan suatu wilayah.

B. Saran

1. Bagi Pemerintah, disarankan untuk meningkatkan kebijakan dan program yang berfokus pada penguatan ketahanan pangan di daerah dengan angka Prevalensi Ketidacukupan Konsumsi Pangan (PKKP) yang tinggi. Program dapat berupa peningkatan akses terhadap sumber pangan, pengembangan infrastruktur pertanian, serta dukungan terhadap petani lokal untuk menjaga ketersediaan pangan yang berkelanjutan. Selain itu, perlu dilakukan integrasi kebijakan lintas sektor antara bidang pertanian, kesehatan, dan sosial untuk memastikan bahwa peningkatan ketahanan pangan berjalan secara menyeluruh. Pemerintah juga diharapkan memperkuat sistem pemantauan data pangan di tingkat kabupaten/kota agar identifikasi wilayah rawan pangan dapat dilakukan lebih cepat dan tepat sasaran.
2. Untuk penelitian selanjutnya, disarankan agar memperluas variabel penelitian dengan menambahkan faktor-faktor lain yang berpotensi mempengaruhi

ketahanan pangan, seperti tingkat pendidikan masyarakat, infrastruktur transportasi, atau kondisi iklim dan lingkungan. Peneliti selanjutnya juga dapat membandingkan kinerja algoritma *Random Forest* dengan metode klasifikasi lainnya seperti *XGBoost*, *Support Vector Machine*, atau *Neural Network* untuk mendapatkan model prediksi yang lebih optimal. Selain itu, penggunaan data panel antar tahun dapat menjadi alternatif untuk melihat dinamika perubahan ketahanan pangan dari waktu ke waktu.

DAFTAR PUSTAKA

- Akbar, K., & Hayaty, M. (2020). Data Balancing untuk Mengatasi Imbalance Dataset pada Prediksi Produksi Padi Balancing Data to Overcome Imbalance Dataset on Rice Production Prediction. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 2(02), 1–14.
- Amaliah, S., Nusrang, M., & Aswi, A. (2022). Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, 4(3), 121-127.
- Aziz, A., & Muharni. (2016). Analisis Ketahanan Rumah Tangga Pada Keluarga Pra Sejahtera dengan Status Gizi Balita di Kelurahan Sri Meranti Kecamatan Rumbai Kota Pekanbaru. *Jurnal Proteksi Kesehatan*, 5(1), 60–66.
- Azmi, B. N., Hermawan, A., & Avianto, D. (2023). Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver. *Jurnal Teknologi Informasi Dan Multimedia*, 4(4).
- Badan Pangan Nasional. (2023). *Indeks Ketahanan Pangan*.
- Badan Pangan Nasional. (2024). *Peta Ketahanan dan Kerentanan Pangan FSVA tahun 2024*.
- Badan Pusat Statistik. (2023). *Statistik Indonesia 2023*.
- Blagus, R., & Lusa, L. (2013). *SMOTE for high-dimensional class-imbalanced data* (Vol. 14). <http://www.biomedcentral.com/1471-2105/14/106>
- Boulesteix, A. L., Janitza, S., Kruppa, J., & R. Konig, I. (2012). Overview of *Random Forest* Methodology and Practical Guidance With Emphasis on Computational Biology and Bioinformatics. *Journal Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6), 493–507.
- Breiman, L. (2001). *Random Forests, Machine Learning* (Vol. 45). Kluwer Academic.
- Breiman, L., Friedman, J., Olshen, R. A., & J.Stone, C. (2017). *Classification and Regression Trees*. Routledge.
- Budiawati, Y. (2024). Hubungan Antara Karakteristik Sosial Ekonomi dan Status Kesehatan Balita Terhadap Ketahanan Pangan di Provinsi Banten. *Jurnal Ilmu Pertanian Tirtayasa*, 6(2), 407.
- Chaereni, R., Agustanto, D., Wahyu, R. A., & Nainggolan, P. (2020). Ketahanan Pangan Berkelanjutan. *Jurnal Kependudukan Dan Pembangunan Lingkungan*, 2.

- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Oversampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Erlin, E., Desnelita, Y., Nasution, N., Suryati, L., & Zoromi, F. (2022). Dampak SMOTE Terhadap Kinerja Random Forest Classifier Berdasarkan Data Tidak Seimbang. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 21(3), 677–690. <https://doi.org/10.30812/matrik.v21i3.1726>
- Firdaus, N., & Cahyono, B. D. (2017). How food consumption pattern and dietary diversity describe food security: Evidence from Yogyakarta and East Nusa Tenggara. *Jurnal Ekonomi dan Pembangunan*, 25(1), 27–38.
- García, V., Sánchez, J. S., Martín-Félez, R., & Mollineda, R. A. (2012). Surrounding neighborhood-based SMOTE for learning from imbalanced data sets. *Progress in Artificial Intelligence*, 1(4), 347–362. <https://doi.org/10.1007/s13748-012-0027-5>
- Genuer, R., Poggi, J. M., Tuleau-Malot, C., & Villa-Vialaneix, N. (2017). *Random Forests for Big Data*. *Big Data Research*, 9, 28–46. <https://doi.org/10.1016/j.bdr.2017.07.003>
- Genuer, R., & Poggi, J.-M. (2020). *Random Forest With R*. Springer.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison- Wesley.
- Goldstein, B. A., Navar, A. M., & Carter, R. E. (2017). Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges. *European Heart Journal*, 38(23), 1805–1814. <https://doi.org/10.1093/eurheartj/ehw302>
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- Hannida, C. P., & Sambodo, H. (2025). Analysis Of The Effect Of Rice Production, Poverty, And Prevalence Of Undernourishment (Pou) On Food Security In Indonesia. *SEPA: Jurnal Sosial Ekonomi Pertanian dan Agribisnis*, 22(1), 11–19.
- Hapsari, N. I., & Rudiarto, I. (2017). Faktor-Faktor yang Mempengaruhi Kerawanan dan Ketahanan Pangan dan Implikasi Kebijakan di Kabupaten Rembang. *Jurnal Wilayah Dan Lingkungan*, 5(2), 125. <https://doi.org/10.14710/jwl.5.2.125-140>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer.
- Igal, L., & Segui, S. (2017). *Introduction to Data Science*. Springer Cham.

- Janitza, S., & Hornung, R. (2018). On the overestimation of *Random Forest*'s out-of-bag error. *PLoS ONE*, 13(8). <https://doi.org/10.1371/journal.pone.0201904>
- Jian, C., Gao, J., & Ao, Y. (2016). A new sampling method for classifying imbalanced data based on support vector machine ensemble. *Neurocomputing*, 193, 115–122. <https://doi.org/10.1016/j.neucom.2016.02.006>
- Ketaren, A., & Nasution, P. P. P. A. (2018). Konflik Gam-Ri Dan Kerentanan Pangan Masyarakat Transmigran. *Jurnal Antropologi: Isu-Isu Sosial Budaya*, 20(1), 103–117. <https://doi.org/10.25077/jantro.v20.n1.p103-117.2018>
- Khalilia, M., Chakraborty, S., & Popescu, M. (2011). Predicting disease risks from highly imbalanced data using *Random Forest*. *BMC Medical Informatics and Decision Making*, 11(1). <https://doi.org/10.1186/1472-6947-11-51>
- Komang, N., Mita, A., Fajar Siddiq, M., Laurnt, A., Erviana, R., & Kurniawan, R. (2024). Optimalisasi Ketahanan Pangan: Perbandingan Metode Machine Learning dan Time Series dalam Memprediksi Produksi Padi di Jawa Tengah. *Seminar Nasional Sains Data*, 2024.
- Lemons, K. (2020). A Comparison Between Naïve Bayes and *Random Forest* to Predict Breast Cancer. *International Journal of Undergraduate Research and Creative Activities*, 12(1), 1. <https://doi.org/10.7710/2168-0620.0287>
- Liaw, A., & Wiener, M. (2002). *Classification and Regression by Random Forest*. 2. <https://www.researchgate.net/publication/228451484>
- Mailund, T. (2017). *Beginning Data Science in R: Data Analysis, Visualization, and Modelling for the Data Scientist* (Apress, Ed.).
- Mishina, Y., Murata, R., Yamauchi, Y., Yamashita, T., & Fujiyoshi, H. (2015). Boosted *Random Forest*. *IEICE Transactions on Information and Systems*, E98D(9), 1630–1636. <https://doi.org/10.1587/transinf.2014OPP0004>
- Mukhopadhyay, S. (2018). *Advanced Data Analytics Using Phyton*. Apress Berkeley.
- Mulugeta, G., Zewotir, T., Tegegne, A. S., Juhar, L. H., & Muleta, M. B. (2023). Classification of imbalanced data using machine learning algorithms to predict the risk of renal graft failures in Ethiopia. *BMC Medical Informatics and Decision Making*, 23(1). <https://doi.org/10.1186/s12911-023-02185-5>
- Nelli, F. (2015). *Phyton Data Analytics: Data ANlysis and Science Using Pandas, matplotlib, and the Phyton Programming Language*. Apress.
- Özuysal, M., Calonder, M., Lepetit, V., & Fua, P. (2010). Fast Keypoint Recognition using Random Ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 448–461. <http://cvlab.epfl.ch>
- Permana, A. A., Wahyussin, S., Santoso, L. W., Wibowo, G. W. N., Wardhani, A. K., Rahmadden, Wahidin, A. J., Yulastuti, G. E., Elisawati, Wijayanti, R. R.,

- & Abdurrasyid. (2023). *Machine Learning*. Global Eksekutif Teknologi. <https://www.researchgate.net/publication/382851316>
- Pratiwi, F. E., & Zain, I. (2014). Klasifikasi Pengangguran Terbuka Menggunakan CART (Classification and Regression Tree) di Provinsi Sulawesi Utara. *Jurnal Sains Dan Semi Pomits*, 3(1), 2337–3520.
- Rachman, A., Badan, S., Fiskal, K., & Keuangan, K. (2021). Paradigma dan Dimensi Strategi Ketahanan Pangan Indonesia Paradigm and Dimensions of Indonesia's Food Security Strategy. *Jejaring Administrasi Publik*, 13(1), 35–48.
- Rachmaningsih, T., & Priyarsono, D. S. (2012). Ketahanan pangan di kawasan timur Indonesia. *Jurnal Ekonomi dan Pembangunan Indonesia*, 13(1), 1.
- Ramayanti, F., Dodi Vionanda, Dony Permana, & Zilrahmi. (2023). Application of *Random Forest* to Identify for Poor Households in West Sumatera Province. *UNP Journal of Statistics and Data Science*, 1(2), 97–104. <https://doi.org/10.24036/ujsds/vol1-iss2/31>
- Ramentol, E., Caballero, Y., & Herrera, F. (2012). *SMOTE-RSB*: A Hybrid Preprocessing Approach Based on Oversampling and Undersampling for High Imbalanced Data-Sets Using SMOTE and Rough Sets Theory* (2nd ed., Vol. 33).
- Riski, H., Mundiastutik, L., & Catur Adi, A. (2019). Ketahanan Pangan Rumah Tangga, Kejadian Sakit dan Sanitasi Lingkungan Berhubungan dengan Status Gizi Balita Usia 1-5 Tahun di Surabaya. *Jurnal Kesehatan Masyarakat*, 130–134. <https://doi.org/10.2473/amnt.v3i3.2019.130-134>
- Rizki Ardiningrum, L., Junaidi, & Umiyati, E. (2021). Pengaruh Indeks Ketahanan Pangan, Pengeluaran Pemerintah di Bidang Pendidikan dan Kesehatan Terhadap Indeks Pembangunan Manusia pada 10 Provinsi di Pulau Sumatera. *Jurnal Ekonomi Sumberdaya Dan Lingkungan*, 10(2).
- Santoso, B. (2018). Kajian Simulasi *Over-Sampling* K-Tetangga Terdekat pada Regresi Logistik Terboboti dan Penerapannya untuk Klasifikasi Rumahtangga Miskin di Provinsi Daerah Istimewa Yogyakarta. IPB University.
- Sholihah, N. N., & Hermawan, A. (2023). Implementation Of *Random Forest* And *Smote* Methods For Economic Status Classification In Cirebon City. *Jurnal Teknik Informatika (Jutif)*, 4(6), 1387–1397. <https://doi.org/10.52436/1.jutif.2023.4.6.1135>
- Sofyan, S., & Prasetyo, A. (2021). Penerapan Synthetic Minority *Oversampling* Technique (SMOTE) Terhadap Data Tidak Seimbang Pada Tingkat Pendapatan Pekerja Informal Di Provinsi D.I. Yogyakarta Tahun 2019. *Seminar Nasional Official Statistics*.

- Syita Noordianty, A., Najma, S., & Siti Nurlaela, R. (2024). Kajian Literatur : Penerapan Aspek Sanitasi Terhadap Mutu dan Produk Pangan. *Karimah Tauhid*, 3(7).
- Tambe, A. B., Akeh, M. L., Tendongfor, N., Dhlamini, T., Chipili, G., & Mbhenyane, X. (2023). The predictors of food security and dietary diversity among internally displaced persons' children (6–59 months) in Bamenda health district, Cameroon. *Conflict and Health*, 17(1). <https://doi.org/10.1186/s13031-023-00511-2>
- Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2006). *Introduction to Data Mining*. Pearson Addison Wesley.
- Tian, X., Engel, B. A., Qian, H., Hua, E., Sun, S., & Wang, Y. (2021). Will reaching the maximum achievable yield potential meet future global food demand. *Journal of Cleaner Production*, 294. <https://doi.org/10.1016/j.jclepro.2021.126285>
- Undang-Undang Nomor 7 Tahun (1996).
- Undang-Undang Republik Indonesia Nomor 18 Tahun, Pub. L. No. 18, Lembaran Negara Republik Indonesia 227 (2012).
- Wijaya, J., Soleh, A. M., & Rizki, A. (2018). Penanganan Data Tidak Seimbang pada Pemodelan *Rotation Forest* Keberhasilan Studi Mahasiswa Program Magister IPB (Vol. 2, Issue 2).
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.). Elseiver.
- Yahya, M. G., Utami, I. P., & Ariansyah, S. (2023). Pemodelan Spasial Prevalensi Ketidacukupan Konsumsi Pangan Menggunakan Pendekatan Ketahanan Pangan di Indonesia Tahun 2022.