

OPTIMASI MODEL DETEKSI ALERGEN PADA PRODUK PANGAN DENGAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM) DAN *ADAPTIVE BOOSTING* (ADABOOST)

Siska Narulita ^{a,1,*}, Sekarlangit ^{b,2}, Milka Putri Novianingrum ^{c,3}

^{a,b,c} Universitas Nasional Karangturi Semarang, Jl. Raden Patah No. 182-192, Rejomulyo, Semarang Timur, Kota Semarang

¹ sisknarulita84@gmail.com*; ² sekarlangitt@gmail.com; ³ milkaaputrii22@gmail.com

* corresponding author

ARTICLE INFO

Keywords

Allergen
Data Mining
Support Vector Machine
AdaBoost
Hand-Out Validation
Split Validation
Cross Validation

ABSTRACT

One important aspect that needs to be considered in food production is food safety. The implementation of this food safety aspect includes food products that avoid contamination of chemical, physical, and biological substances that can be harmful to human health. In the implementation of the Makan Bergizi Gratis (MBG) program, problems were found related to allergies in the recipients of this assistance program. According to the World Health Organization (WHO), food allergies are ranked as the fourth most serious public health problem, and the only effective treatment for allergy sufferers is to avoid foods that contain allergens. Allergens themselves are compounds or food ingredients that cause allergies and/or intolerances. Laboratory tests of food products for allergen testing that are still carried out traditionally require a lot of time and money, making food producers reluctant to carry out product testing. A way to detect allergen content in food products that is easier, more practical, and more accurate is needed. The research conducted aims to build a prediction model that can be used to detect allergen content in food ingredients through the implementation of the Support Vector Machine (SVM) data mining algorithm optimized with the Adaptive Boosting ensemble learning boosting algorithm (AdaBoost). The research conducted obtained a model that produces the most optimal performance, namely SVM optimized with the AdaBoost algorithm with the split validation method.

1. Pendahuluan

Salah satu kebutuhan dasar manusia adalah pangan, oleh sebab itu pemenuhan akan kebutuhan pangan yang bergizi, aman, dan cukup adalah hak asasi bagi setiap rakyat Indonesia [1]. Aspek yang perlu diperhatikan oleh pengusaha yang bergerak di bidang produksi olahan pangan, baik itu makanan dan juga minuman adalah terkait dengan keamanan pangan [2]. Implementasi aspek keamanan pangan diantaranya dengan menjaga produk olahan makanan dan minuman dari kontaminasi kimia, fisik, maupun biologi yang membahayakan kesehatan manusia sebagai konsumen [3]. Apabila para produsen mengimplementasikan prinsip keamanan pangan dalam melakukan proses produksi, maka konsumen dapat terhindar dari risiko kesehatan sebagai akibat dari produk pangan yang tidak aman [2]. Selain melindungi konsumen dari risiko kesehatan akibat dari konsumsi produk pangan yang tidak aman, aspek keamanan pangan juga menjaga citra dan keberlangsungan atau kesinambungan dari usaha itu sendiri [4].

Seperti yang diketahui bahwa, pelaksanaan program Makan Bergizi Gratis (MBG) ditemukan kendala dimana salah satu diantaranya terkait dengan permasalahan alergi pada penerima program bantuan [5]. Menurut *World Health Organization* (WHO), alergi makanan menduduki peringkat keempat sebagai permasalahan kesehatan masyarakat yang paling serius, dan satu-satunya cara pengobatan yang efektif bagi

penderita alergi adalah dengan menghindari makanan yang mengandung alergen [6]. Di Indonesia, melalui Direktorat Standardisasi Pangan Olahan, Deputi Bidang Pengawasan Pangan Olahan, Badan Pengawas Obat dan Makanan (BPOM) Republik Indonesia, mengeluarkan Pedoman Implementasi Pelabelan Pangan Olahan (Pencantuman Jumlah Bahan Baku dan Informasi Alergen) [7], sebagai implementasi dari Peraturan Badan Pengawas Obat dan Makanan Nomor 31 Tahun 2018 tentang Label Pangan Olahan [8]. Produsen makanan diwajibkan untuk mencantumkan zat-zat yang tergolong alergen pada label kemasan, sehingga konsumen yang mempunyai alergi pada produk pangan tertentu dapat memperoleh informasi terkait alergen yang terkandung di dalamnya [9].

Pengertian alergen itu sendiri adalah senyawa atau bahan pangan yang menjadi penyebab alergi dan/atau intoleransi [9]. Uji laboratorium produk pangan untuk pemeriksaan alergen yang masih dilakukan secara tradisional membutuhkan banyak waktu dan biaya, sehingga membuat para produsen pangan enggan untuk melakukan pengujian produk [10]. Diperlukan suatu cara untuk mendeteksi kandungan alergen pada produk pangan yang lebih mudah, praktis, dan akurat. Penelitian yang dilakukan bertujuan untuk membangun sebuah model prediksi yang dapat digunakan untuk mendeteksi kandungan alergen pada bahan pangan melalui implementasi algoritma *data mining Support Vector Machine (SVM)* yang dioptimasi dengan algoritma *boosting ensemble learning Adaptive Boosting (AdaBoost)*.

Definisi dari *data mining* itu sendiri adalah suatu proses ekstraksi pola dari sekumpulan data besar dengan berbagai teknik atau metode statistik, matematika, serta *artificial intelligence* yang bertujuan untuk mendapatkan informasi tersembunyi pada data yang dapat dipergunakan pada proses pengambilan keputusan [11]. Sedangkan algoritma SVM merupakan satu diantara algoritma klasifikasi yang mempergunakan teknik *machine learning* untuk membuat garis batas sebagai pemisah teks berdasarkan kelasnya [12]. Algoritma SVM mempunyai beberapa kelebihan, diantaranya mampu dalam penanganan data terdistribusi tidak normal, efisiensi memori, dan performansi akurasi yang tinggi [13]. Algoritma AdaBoost satu diantara algoritma *boosting* yang telah banyak digunakan dan juga merupakan metode *ensemble learning* yang dapat meningkatkan performansi model [14]. Berdasarkan kelebihan dari kedua algoritma tersebut, maka dalam penelitian ini dikombinasikan kedua algoritma untuk mendapatkan model deteksi alergen yang paling optimal.

Sebelumnya sudah ada beberapa penelitian tentang deteksi alergen pada produk pangan menggunakan algoritma *data mining*, diantaranya deteksi alergen pada produk pangan mempergunakan algoritma SVM dengan tingkat akurasi model 98,75% [15]. Selanjutnya penelitian deteksi alergen menggunakan algoritma *Convolutional Neural Network (CNN)* yang menghasilkan tingkat akurasi model sebesar 72% [16]. Penelitian untuk klasifikasi alergen protein dengan mengkombinasikan algoritma *deep learning (Extra Tree dan Deep Belief Network)* dan algoritma *ensemble (CatBoost)* menghasilkan akurasi model sebesar 89,16% [17]. Pada penelitian deteksi alergen makanan menggunakan algoritma *Decision Tree* menghasilkan akurasi yang paling optimal sebesar 98% [18]. Pada penelitian deteksi dan analisis alergen makanan menggunakan *hybrid ensemble model* menghasilkan akurasi paling optimal sebesar 96,2% [19].

Penelitian yang dilakukan ini bertujuan untuk mengoptimasi model deteksi alergen pada *dataset* produk pangan yang mempergunakan algoritma SVM dengan algoritma *boosting ensemble learning AdaBoost* untuk mendapatkan performansi model yang optimal.

2. Metode Penelitian

Adapun langkah-langkah atau tahapan penelitian yang digunakan pada penelitian ini mengacu pada metode penelitian yang ditunjukkan pada Gambar 1.

2.1. Data Gathering

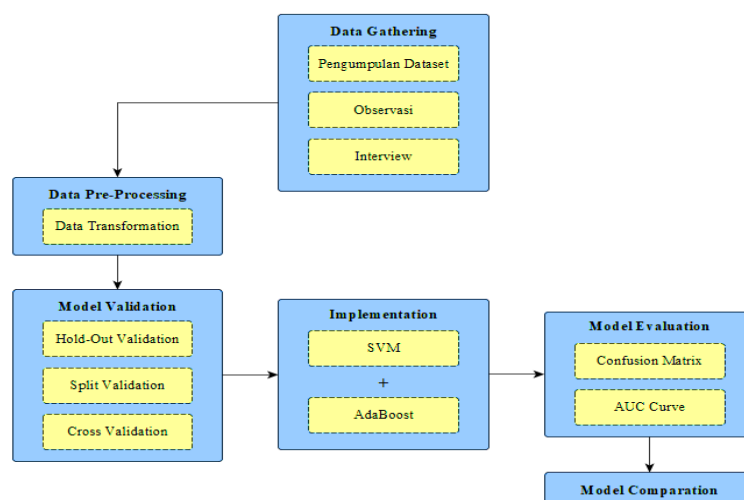
Pada tahap *data gathering* dilakukan pengumpulan *dataset* dari berbagai sumber. *Dataset* ini menjadi komponen utama dalam proses *data mining*. Peneliti melakukan pengumpulan *dataset* dari basis data (*database*) *public* (<https://www.kaggle.com>) maupun *private* (dari hasil pengisian data yang akan dilakukan oleh *user*). Jumlah data pada *dataset* yang digunakan sejumlah 400 *record data*, dengan pembagian jumlah *data training* sebesar 64% atau 256 *record data* dengan *class* mengandung alergen dan sisanya sebesar 36% atau 144 *record data* dengan *class* tidak mengandung alergen. Peneliti juga melakukan observasi terkait dengan penelitian serupa tentang deteksi alergen untuk penentuan penggunaan algoritma yang sesuai. Selain itu juga dilakukan *interview* (wawancara) dengan pihak-pihak terkait, seperti Dinas Kesehatan (Dinkes dan Balai Besar Pengawas Obat dan Makanan (POM), untuk mengkonfirmasi kesesuaian *dataset* yang telah dikumpulkan apakah sudah sesuai dengan fakta di lapangan.

2.2. Data Pre-Processing

Pada tahapan *data pre-processing* atau pra-proses data dilakukan *proses data transformation*, yaitu mengubah *dataset* mentah menjadi sesuai format untuk dapat dipergunakan pada proses *data mining*. Proses *data transformation* yang dilakukan dalam penelitian ini adalah konversi unit untuk mengubah tipe data pada *dataset*.

2.3. Model Validation

Tahap *model validation* merupakan tahapan yang dapat membantu memastikan model dapat bekerja dengan baik pada implementasi data baru. Tahapan ini dapat membantu menghindari *overfitting*, yaitu kondisi dimana model algoritma terlalu kompleks. Metode atau teknik validasi yang dipergunakan di dalam penelitian ini adalah metode atau teknik *hold-out validation* yang membagi *dataset* menjadi set pelatihan (*data training*) dan set tes (*data testing*) yang terpisah dari set pelatihan [20], *split validation* dan *cross validation*. *Split validation* digunakan untuk pengujian performansi model tertentu, sedangkan *cross validation* digunakan untuk melakukan evaluasi terhadap performansi model tertentu [21].



Gambar 1. Metode Penelitian

2.4. Implementation

Pada tahap *implementation* dilakukan pembentukan model dengan algoritma SVM yang dioptimasi dengan algoritma *ensemble boosting* AdaBoost. Konsep SVM secara sederhana digambarkan sebagai suatu upaya mencari *hyperplane* terbaik sebagai pemisah antara dua kelas pada *input space*. Dapat dikatakan bahwa upaya mencari lokasi *hyperplane* ini adalah inti proses pembelajaran pada algoritma SVM [22]. *Adaptive Boosting* (AdaBoost) merupakan salah satu dari algoritma *boosting* yang mempunyai konsep penjumlahan secara bertahap, yang mana beberapa *weak classifier* digunakan untuk mendapatkan *strong classifier* [23].

2.5. Model Evaluation

Evaluasi model pada penelitian ini mempergunakan *confusion matrix* dan *Area Under Curve* (AUC) *curve*. *Confusion matrix* atau tabel kontingensi adalah tabel untuk melakukan evaluasi terhadap performansi model klasifikasi. *Confusion matrix* mengkaji sejauh mana penentuan kategori mengenali *tuple* kelas yang berlawanan [24]. Kurva AUC dipergunakan untuk menentukan area atau daerah yang berada di bawah kurva *Receiver Operating Characteristics* (ROC). Area trapesium yang ada pada metrik AUC bisa ditambahkan untuk penentuan nilai AUC. Semakin bagus model yang dibangun, maka nilai AUC semakin mendekati nilai 1 [25].

2.6. Model Comparison

Pada tahap model comparison, dilakukan perbandingan hasil dari model yang dihasilkan. Metrik yang dibandingkan adalah *accuracy*, *precision*, dan *recall*. *Accuracy* merupakan jumlah hasil prediksi yang benar ataupun salah yang dilakukan oleh model algoritma. *Precision* (presisi) adalah rasio atau perbandingan jumlah terprediksi benar dan positif terhadap seluruh hasil yang terprediksi benar dan positif. Sedangkan *recall* adalah rasio atau perbandingan prediksi yang dinyatakan benar dan positif dibanding keseluruhan data benar dan positif [26].

3. Hasil dan Pembahasan

3.1. Data Gathering

Pada *dataset* yang digunakan terdapat 4 (empat) atribut reguler dan 1 (satu) atribut spesial, serta 399 *record* data. Tabel 1 menunjukkan informasi dari *dataset* yang dipergunakan di dalam penelitian.

Tabel 1. Informasi *Dataset*

No.	Atribut	Role	Tipe Data
1.	Bahan Utama	Atribut	Polinomial
2.	Pemanis	Atribut	Polinomial
3.	Lemak/Minyak	Atribut	Polinomial
4.	Penyedap Rasa	Atribut	Polinomial
5.	Alergen	Atribut	Polinomial
6.	Prediksi	Label	Binomial

3.2. Data Pre-Processing

Pada tahapan *data pre-processing*, dijalankan proses *data transformation* untuk menyesuaikan tipe data atribut sesuai dengan sifat matematis algoritma yang digunakan. Tipe data pada *dataset* yang sebelumnya diset ke tipe data nominal ditransformasikan ke tipe data numerik.

3.3. Model Validation

Model validation yang digunakan dalam penelitian ini adalah *hold-out validation*, *split validation*, dan *cross validation*. Pada metode *hold-out validation* dan *split validation* dilakukan pembagian *dataset* dengan perbandingan antara *data training* sebesar 80% dan *data testing* 20%. Sedangkan untuk *cross validation* digunakan nilai *k* sebesar 10.

3.4. Implementation

Tahap penelitian selanjutnya adalah *implementation*, dimana pada tahap ini dilakukan implementasi algoritma SVM yang dioptimasi menggunakan algoritma *ensemble boosting* AdaBoost.

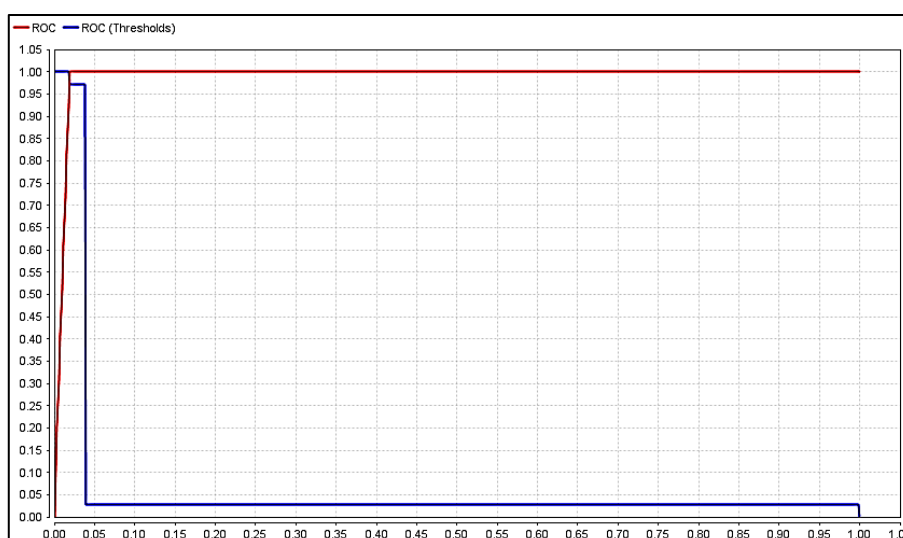
3.5. Model Evaluation

Untuk melakukan evaluasi model pada penelitian ini digunakan *confusion matrix* dan *AUC curve*. Tabel 2 menunjukkan *confusion matrix* dari implementasi algoritma SVM dan AdaBoost dengan metode *hold-out validation*.

Tabel 2. *Confusion Matrix SVM + AdaBoost + Hold-Out Validation*

	True Mengandung Alergen	True Tidak Mengandung Alergen	Class Precision
Pred. Mengandung Alergen	50	0	100,00%
Pred. Tidak Mengandung Alergen	1	29	96,67%
Class Recall	98,04%	100,00%	

Kurva AUC dari model yang terbentuk ini dapat dilihat pada Gambar 2 berikut.



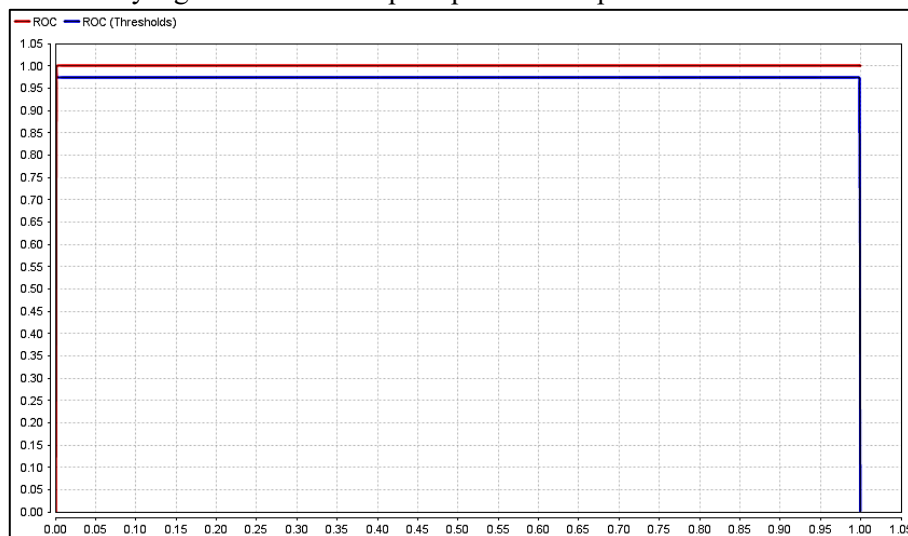
Gambar 2. Kurva AUC Model SVM + AdaBoost + *Hold-Out Validation*

Confusion matrix untuk hasil implementasi algoritma SVM dan AdaBoost dengan metode *split validation* ditunjukkan pada Tabel 3.

Tabel 3. *Confusion Matrix SVM + AdaBoost + Split Validation*

	True Mengandung Alergen	True Tidak Mengandung Alergen	Class Precision
Pred. Mengandung Alergen	51	0	100,00%
Pred. Tidak Mengandung Alergen	0	29	100,00%
Class Recall	100,00%	100,00%	

Kurva AUC untuk model yang telah dihasilkan pada penelitian diperlihatkan Gambar 3.



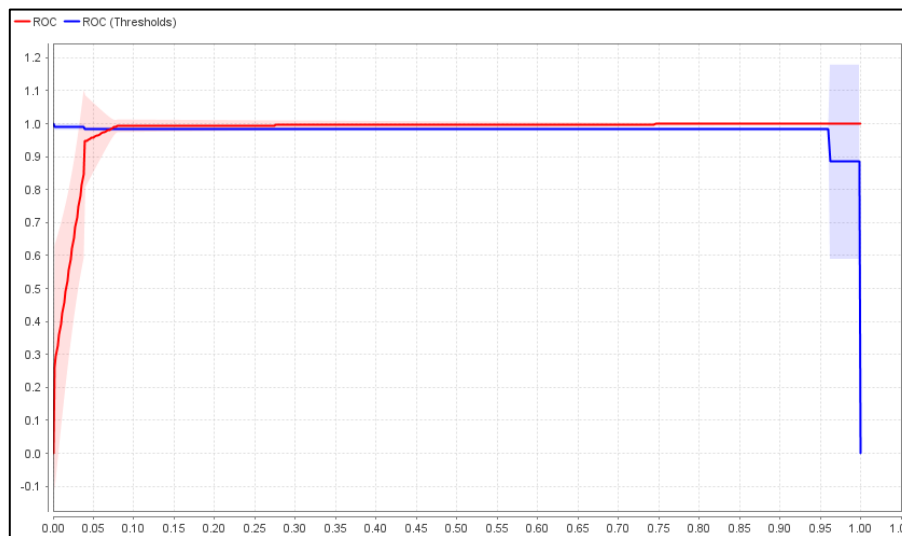
Gambar 3. Kurva AUC Model SVM + AdaBoost + *Split Validation*

Sedangkan *confusion matrix* untuk hasil implementasi algoritma SVM dan AdaBoost dengan metode *cross validation* ditunjukkan Tabel 4.

Tabel 4. *Confusion Matrix SVM + AdaBoost + Cross Validation*

	True Mengandung Alergen	True Tidak Mengandung Alergen	Class Precision
Pred. Mengandung Alergen	247	1	99,60%
Pred. Tidak Mengandung Alergen	9	142	94,04%
Class Recall	96,48%	99,30%	

Kurva AUC untuk model yang terbentuk diperlihatkan Gambar 4.



Gambar 4. Kurva AUC Model SVM + AdaBoost + *Cross Validation*

3.6. Model Comparison

Berdasarkan tabel-tabel *confusion matrix* di atas, dapat dilakukan komparasi terhadap nilai-nilai metrik yang dihasilkan. Tabel 5 di bawah ini menunjukkan komparasi atau perbandingan nilai-nilai metrik model yang dihasilkan.

Tabel 5. Komparasi Metrik Model

Model	Metrik			
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	AUC
SVM + AdaBoost + <i>Hold-Out Validation</i>	98,75%	96,67%	100,00%	0,990
SVM + AdaBoost + <i>Split Validation</i>	100,00%	100,00%	100,00%	1,000
SVM + AdaBoost + <i>Cross Validation</i>	97,50%	94,24%	99,33%	0,978

Dari Tabel 5 di atas, terlihat bahwa nilai *accuracy*, *precision*, dan *recall* pada implementasi algoritma SVM dan AdaBoost menggunakan metode validasi *split validation* menghasilkan nilai maksimal 100%. Pembagian *dataset* menjadi *data training* dan *testing* mempunyai pengaruh terhadap performansi model algoritma. Model yang dilatih dengan jumlah *data training* yang besar dibanding dengan *data testing*, maka model yang dibentuk akan mampu mempelajari pola pada *dataset* dengan lebih baik, hal ini yang menyebabkan model dapat menghasilkan performansi yang lebih baik [27].

4. Kesimpulan

Hasil penelitian yang telah dilakukan menghasilkan beberapa model dimana nilai dari setiap metrik model yang dihasilkan telah ditunjukkan pada tabel komparasi pada Tabel 5 sebelumnya. Berdasarkan tabel komparasi tersebut, nilai *accuracy* terbaik dihasilkan oleh model SVM dan AdaBoost dengan metode *split validation*, sebesar 100%. Untuk model SVM dan AdaBoost dengan metode *hold-out validation* menghasilkan nilai akurasi sebesar 98,75%. Sedangkan model SVM dan AdaBoost dengan metode *cross validation* menghasilkan nilai akurasi sebesar 97,50%.

Untuk penelitian selanjutnya disarankan untuk menggunakan algoritma data mining lainnya, seperti *Artificial Neural Networks* (ANN), *Long Short-Term Memory* (LSTM), *Convolutional Neural Networks* (CNN), dan lain sebagainya dikombinasikan dengan algoritma *ensemble boosting*, seperti *Gradient Boosting*, *XGBoost*, dan lain sebagainya.

Pada penelitian ini, peneliti mengucapkan terima kasih kepada Direktorat Penelitian dan Pengabdian kepada Masyarakat (DPPM), Kementerian Pendidikan Tinggi, Sains, dan Teknologi (Kemdiktisaintek) yang telah memberikan kesempatan kepada para peneliti untuk melakukan penelitian melalui skema hibah Penelitian Dosen Pemula (PDP) Tahun 2025 dan Universitas Nasional Karangturi Semarang yang telah mensupport peneliti dalam melaksanakan penelitian ini hingga selesai.

Daftar Pustaka

- [1] Sabarella, W. B. Komalasari, M. Manurung, M. D. N. Saida, K. Seran, and Y. Supriyati, *Buletin Konsumsi Pangan*, vol. 15, no. 1. Jakarta: Pusat Data dan Sistem Informasi Pertanian - Sekretariat Jenderal, Kementerian Pertanian, 2024.
- [2] N. Latifasari, F. T. Syifa, F. C. Agustia, C. Raharditya, S. A. Cahyani, and P. A. Rini, "Peningkatan Pengetahuan Pentingnya Keamanan Pangan dalam Sistem Produksi Yoghurt Drink Melalui Sosialisasi di UKM Yoghurt Sehati Banyumas," *JURPIKAT: Jurnal Pengabdian Kepada Masyarakat*, vol. 6, no. 1, pp. 92–100, 2025, doi: <https://doi.org/10.37339/jurpikat.v6i1.1981>.
- [3] R. T. Cahyani, Rusmiati, Ngadino, and Narwati, "Kondisi Sanitasi dan Personal Hygiene Industri Tempe di Desa Sambirembe Kecamatan Karangrejo Kabupaten Magetan," *Jurnal Sanitasi Lingkungan*, vol. 2, no. 2, pp. 101–106, 2022, doi: <https://doi.org/10.36086/jsl.v2i2.1398>.
- [4] N. M. Putri, A. R. Gati, S. A. Cahyani, S. Q. Adifaputra, U. Maghfiroh, and N. Latifasari, "Penyuluhan Keamanan Pangan di UKM Produk Manisan Buah Pepaya Desa Karangsalam Kabupaten Banyumas," *IJCOSIN: Indonesia Journal of Community Service and Innovation*, vol. 4, no. 2, pp. 56–65, 2024, doi: <https://doi.org/10.20895/ijcosin.v4i1.1489>.
- [5] Humas KPAI, "KPAI Soroti Kendala dalam Program Makanan Bergizi Gratis, Tegaskan Pentingnya Pengawasan dan Koordinasi," 24 Januari 2025, 2025. <https://www.kpai.go.id> (accessed Jun. 09, 2025).
- [6] S. S. Hamid, "Food Allergy and Physiology: A Review Article," *Indonesian Journal on Health Science and Medicine*, vol. 2, no. 2, pp. 1–9, 2025, doi: <https://doi.org/10.21070/ijhsm.v2i2.108>.
- [7] Direktorat Standardisasi Pangan Olahan, *Pedoman Implementasi Pelabelan Pangan Olahan*. Jakarta: Direktorat Standardisasi Pangan Olahan, Deputi Bidang Pengawasan Pangan Olahan, Badan Pengawas Obat dan Makanan Republik Indonesia, 2019.
- [8] Badan Pengawas Obat dan Makanan Republik Indonesia, *Peraturan Badan Pengawas Obat dan Makanan Nomor 31 Tahun 2018 tentang Label Pangan Olahan*. Indonesia, 2018.
- [9] D. C. Aryani et al., *Panduan Pencantuman Label Pangan Segar*. Jakarta: Badan Pangan Nasional, 2023.
- [10] F. Sarlakifar, H. Malek, N. A. Fard, and Z. Khotanlou, "AllerTrans: A Deep Learning Method for Predicting the Allergenicity of Protein Sequences," *bioRxiv*, vol. 08.09, no. 607419, pp. 1–12, 2024, doi: <https://doi.org/10.1101/2024.08.09.607419>.
- [11] J. K. Wororomi et al., *Data Mining (Memahami Pola di Balik Angka)*. Purbalingga: CV Eureka Media Aksara, 2024.
- [12] Amna et al., *Data Mining*. Padang: PT Global Eksekutif Teknologi, 2023.
- [13] I. S. Aisah, B. Irawan, and T. Suprpti, "Algoritma Support Vector Machine (SVM) untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital," *JATI: Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 6, pp. 3759–3765,

2023, doi: <https://doi.org/10.36040/jati.v7i6.8263>.

- [14] A. Latif and S. K. Wildah, "Analisis Kinerja Algoritma Ensemble dalam Prediksi Perilaku Pembelian Pelanggan," *JATI: Jurnal Mahasiswa Teknik Informatika*, vol. 9, no. 1, pp. 557–563, 2025, doi: <https://doi.org/10.36040/jati.v9i1.12413>.
- [15] S. Narulita, Sekarlangit, and M. P. Novianingrum, "Deteksi Alergen pada Produk Pangan Menggunakan Algoritma Support Vector Machines (SVM)," *BRIDGE: Jurnal publikasi Sistem Informasi dan Telekomunikasi*, vol. 3, no. 1, pp. 64–76, 2025, doi: <https://doi.org/10.62951/bridge.v3i1.393>.
- [16] A. A. G. A. Maheswara, L. Fanani, and A. H. Brata, "Pengembangan Aplikasi Deteksi Alergen pada Makanan Menggunakan Convolutional Neural Network Berbasis Android," *JPTIHK: Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 6, pp. 1–16, 2024, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/13862>.
- [17] A. Kumar and P. S. Rana, "A Deep Learning based Ensemble Approach for Protein Allergen Classification," *PeerJ Computer Science*, vol. 9, no. e1622, pp. 1–26, 2023, doi: <https://doi.org/10.7717/peerj-cs.1622>.
- [18] Z. Ahmad, A. Farooq, M. Fuzail, Y. Aziz, and N. Aslam, "Food Allergy Detection using Machine Learning Approach," *KJMR: Kashf Journal of Multidisciplinary Research*, vol. 2, no. 4, pp. 116–127, 2025, doi: <https://doi.org/10.71146/kjmr402>.
- [19] J. Ahmad *et al.*, "Utilizing Deep Learning Techniques for Detecting and Analyzing Food Allergies," *KJMR: Kashf Journal of Multidisciplinary Research*, vol. 2, no. 3, pp. 46–60, 2025, doi: <https://doi.org/10.71146/kjmr332>.
- [20] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split Data terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," *JSI: Jurnal Sains dan Informatika*, vol. 9, no. 1, pp. 19–28, 2023, doi: <https://doi.org/10.34128/jsi.v9i1.622>.
- [21] M. Amelia and S. Ratnasari, "Analisis Tingkat Kepuasan Pengguna Aplikasi Shopee Menggunakan Algoritma Decision Tree," *Kohesi: Jurnal Multidisiplin Saintek*, vol. 7, no. 4, pp. 81–90, 2025, doi: <https://doi.org/10.3785/kohesi.v7i4.11758>.
- [22] F. Putrawansyah and T. Susanti, "Penerapan Metode Support Vector Machine terhadap Klasifikasi Jenis Jambu Biji," *JIKO: Jurnal Informatika dan Komputer*, vol. 8, no. 1, pp. 193–204, 2024, doi: <http://dx.doi.org/10.26798/jiko.v8i1.988>.
- [23] M. R. Saputro, U. Mahdiah, and D. Swanjaya, "Perbandingan Metode Adaptive Boosting dan Extreme Gradient Boosting untuk Prediksi Hasil Pertandingan Liga Spanyol," *Jurnal Nusantara Engineering*, vol. 7, no. 1, pp. 67–73, 2024, doi: <https://doi.org/10.29407/noe.v7i1.20882>.
- [24] R. Mashitapasha, F. Damayanti, and D. A. Fatah, "Penerapan Metode Decision Tree dalam Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma C4.5," *JATI: Jurnal Mahasiswa Teknik Informatika*, vol. 9, no. 3, pp. 4016–4023, 2025, doi: <https://doi.org/10.36040/jati.v9i3.13532>.
- [25] Amrin, O. Pahlevi, and H. Rianto, "Analisa Komparasi Model Data Mining Algoritma C4.5, CHAID, dan Random Forest Untuk Penilaian Kelayakan Kredit," *CO-SCIENCE: Computer Science*, vol. 5, no. 1, pp. 49–57, 2025, doi: <https://doi.org/10.31294/coscience.v5i1.6208>.
- [26] A. Desiani, A. Amran, Y. Andriani, T. Wahyuni, and F. Rizki, "Perbandingan Algoritma Logistic Regression dan Adaptive Boosting (AdaBoost) dalam Klasifikasi Penyakit Gagal Jantung," *JTI: Jurnal Teknologi Informasi*, vol. 19, no. 1, pp. 71–78, 2025, [Online]. Available: <https://e-journal.upr.ac.id/index.php/JTI/article/view/17173>.
- [27] Y. Septiawan, A. Aglasia, and D. A. Muktiawan, "Peningkatan Performa Analisis Sentimen Ulasan Pelanggan terhadap Layanan Pengiriman Menggunakan Model Naive Bayes yang Dioptimalkan dengan PSO," *ROUTERS: Jurnal Sistem dan Teknologi Informasi*, vol. 3, no. 1, pp. 10–19, 2025, doi: <https://doi.org/10.25181/rt.v3i1.4001>.