



**PENERAPAN KLASIFIKASI *MULTI-LABEL* PADA RESEP
MASAKAN INDONESIA UNTUK DETEKSI ALERGEN
MAKANAN MENGGUNAKAN ALGORITMA *K-NEAREST
NEIGHBOR*, *RANDOM FOREST*, DAN *XGBOOST***

LAPORAN TUGAS AKHIR

**RAMADHANI NUR SARJITO
41521010097**

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**PENERAPAN KLASIFIKASI *MULTI-LABEL* PADA RESEP
MASAKAN INDONESIA UNTUK DETEKSI ALERGEN
MAKANAN MENGGUNAKAN ALGORITMA *K-NEAREST
NEIGHBOR, RANDOM FOREST, DAN XGBOOST***

PROPOSAL PENELITIAN

RAMADHANI NUR SARJITO

41521010097

**UNIVERSITAS
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

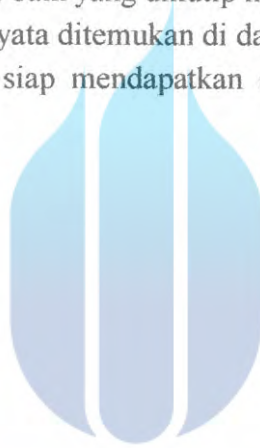
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ramadhani Nur Sarjito
NIM : 41521010097
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Klasifikasi Multi-Label pada Resep Masakan Indonesia untuk Deteksi Alergen Makanan Menggunakan Algoritma K-Nearest Neighbor, Random Forest, dan XGBoost

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 24 Juli 2025



Ramadhani Nur Sarjito

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

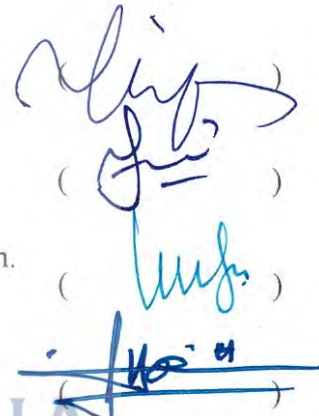
Laporan Skripsi ini diajukan oleh:

Nama : Ramadhani Nur Sarjito
NIM : 41521010097
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Klasifikasi Multi-Label pada Resep Masakan Indonesia untuk Deteksi Alergen Makanan Menggunakan Algoritma K-Nearest Neighbor, Random Forest, dan XGBoost

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Dr. Ir. Eliyani.
NIDN : 0321026901
Ketua Penguji : Dr. Afiyati, S.Si., M.T.
NIDN : 0316106908
Penguji 1 : Inna Sabily Karima, S.Kom., M.Kom.
NIDN : 0324018902
Penguji 2 : Lukman Hakim, S.T., M.Kom.
NIDN : 0327107701



Jakarta, 19 Juli 2025

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Dr. Ir. Eliyani selaku dosen pembimbing MPTI yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensuport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 10 Juli 2025



Ramadhani Nur Sarjito

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Ramadhani Nur Sarjito
NIM : 41521010097
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Penerapan Klasifikasi Multi-Label pada Resep Masakan Indonesia untuk Deteksi Alergen Makanan Menggunakan Algoritma K-Nearest Neighbor, Random Forest, dan XGBoost

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 24 Juli 2025

Yang menyatakan,



Ramadhani Nur Sarjito

ABSTRAK

Nama : Ramadhani Nur Sarjito
NIM : 41521010097
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Penerapan Klasifikasi *Multi-Label* pada Resep Masakan Indonesia untuk Deteksi Alergen Makanan Menggunakan Algoritma *K-Nearest Neighbor*, *Random Forest*, dan *XGBoost*
Dosen Pembimbing : Eliyani, Dr., Ir.

Alergi makanan merupakan masalah kesehatan yang berdampak signifikan terhadap kualitas hidup dan dapat menyebabkan reaksi serius, bahkan mengancam jiwa. Di Indonesia, sekitar 1,6 juta anak-anak mengalami alergi susu sapi, sehingga dibutuhkan sistem yang mampu mendeteksi kandungan alergen secara otomatis. Penelitian ini bertujuan untuk mengembangkan sistem deteksi alergen pada resep masakan Indonesia menggunakan pendekatan klasifikasi *multilabel* berbasis *machine learning*. Sebanyak 7.840 resep dikumpulkan dari situs Cookpad.com melalui teknik *web scraping*. Proses pelabelan dilakukan secara otomatis, berdasarkan daftar kata kunci dari lima kategori alergen utama yaitu susu, kacang tanah, telur, makanan laut, dan gandum. Data yang diperoleh kemudian diproses melalui tahap *preprocessing* yang meliputi pembersihan teks, penghapusan tanda baca dan angka, konversi huruf kecil, tokenisasi, penghapusan *stopword*, serta *stemming* untuk menghasilkan representasi data yang bersih dan konsisten. Selanjutnya, tiga algoritma *machine learning* yaitu *K-Nearest Neighbors* (KNN), *Random Forest* (RF), dan *Extreme Gradient Boosting* (XGB) diterapkan untuk membangun model klasifikasi. Evaluasi dilakukan menggunakan metrik akurasi, presisi, *recall*, dan F1-score. Hasil menunjukkan bahwa XGB dengan tuning *hyperparameter* melalui *GridSearchCV* memberikan performa terbaik, dengan nilai *recall* tertinggi sebesar 0,9794 untuk deteksi alergen susu. Sistem ini diimplementasikan ke dalam aplikasi web berbasis Streamlit, sehingga memudahkan pengguna dalam mendeteksi kandungan alergen pada resep makanan secara praktis.

Kata kunci: Deteksi Alergen, *K-Nearest Neighbor* (KNN), *Random Forest* (RF), dan *Extreme Gradient Boosting* (XGB), Resep Masakan Indonesia.

ABSTRACT

Nama : Ramadhani Nur Sarjito
NIM : 41521010097
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Implementation of *Multi-Label* Classification on Indonesian Food Recipes for Food Allergen Detection Using the *K-Nearest Neighbor*, *Random Forest*, and XGBoost Algorithms
Dosen Pembimbing : Eliyani, Dr., Ir.

Food allergies are a health issue that significantly impacts quality of life and can cause severe, even life-threatening reactions. In Indonesia, approximately 1.6 million children suffer from cow's milk allergy, highlighting the need for a system capable of automatically detecting allergen content. This study aims to develop an allergen detection system for Indonesian food recipes using a multi-label classification approach based on machine learning. A total of 7,840 recipes were collected from the website Cookpad.com using web scraping techniques. Labeling was conducted automatically, based on a keyword list from five main allergen categories: milk, peanuts, eggs, seafood, and wheat. The collected data then underwent preprocessing steps, including text cleaning, removal of punctuation and numbers, lowercasing, tokenization, stopword removal, and stemming, to produce clean and consistent data representation. Next, three machine learning algorithms—K-Nearest Neighbors (KNN), Random Forest (RF), and Extreme Gradient Boosting (XGB)—were applied to build the classification model. Evaluation was performed using metrics such as accuracy, precision, recall, and F1-score. The results showed that XGB, with hyperparameter tuning via GridSearchCV, delivered the best performance, achieving the highest recall value of 0.9794 for milk allergen detection. This system was implemented as a web application using Streamlit, making it easy for users to practically detect allergen content in food recipes.

Keywords: *Allergen Detection, K-Nearest Neighbor (KNN), Random Forest (RF), Extreme Gradient Boosting (XGB), Indonesian Recipes.*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
 BAB I PENDAHULUAN.....	 1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
 BAB II TINJAUAN PUSTAKA	 5
2.1 Penelitian Terdahulu	5
2.2 Gap Penelitian	17
2.3 Teori Pendukung.....	18
2.3.1 Alergi Makanan	18
2.3.2 Alergen Makanan	18
2.3.3 TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>).....	19
2.3.4 <i>K-Nearest Neighbor</i>	20
2.3.5 <i>Random Forest</i>	21
2.3.6 <i>XGBoost</i>	22
2.3.7 <i>Multilabel Classification</i>	23
2.3.8 <i>Cross Validation</i>	24
2.3.9 <i>Data Preprocessing</i>	25
2.3.10 <i>Machine learning</i>	25
2.3.11 <i>Confussion matrix</i>	26
 BAB III METODE PENELITIAN.....	 28
3.1 Jenis Penelitian.....	28
3.2 Tahapan Penelitian.....	28
3.2.1. Pengumpulan Data	29
3.2.2. Pelabelan Dataset	30

3.2.3. Dataset Resep Makanan	31
3.2.4. Preprocessing Data	32
3.2.5. Ekstraksi Fitur	36
3.2.6. <i>Data Train-Test Splitting</i>	39
3.2.7. Modelling	39
3.2.8. Evaluasi dan Analisis Hasil	40
3.3 Perancangan GUI	41
3.4 Penerapan	42
3.3.1 Penerapan <i>Preprocessing</i> Data	42
3.3.2 Penerapan TF-IDF	44
3.3.3 Penerapan Data <i>Train-Test Splitting</i>	45
3.3.4 Penerapan Klasifikasi <i>Multilabel</i> KNN	45
3.3.5 Penerapan Klasifikasi <i>Multilabel</i> RF	47
3.3.6 Penerapan Klasifikasi <i>Multilabel</i> XGB	49
3.3.7 Penerapan Testing dan Evaluasi Performa	50
3.3.8 Penerapan GUI	54
BAB IV HASIL DAN PEMBAHASAN	58
4.1. Skenario Uji Coba dan Analisis	58
4.1.1. Hasil Uji Coba Pelabelan pada Dataset	58
4.1.2. Hasil Uji Coba Pada <i>Hyperparameter Default</i> (Non-GridsearchCV) ...	60
4.1.3. Hasil Uji Coba Pada <i>Hyperparameter</i> dengan <i>GridSearchCV</i>	62
4.1.4. Hasil Uji <i>Interface</i>	65
4.2. Pembahasan	66
4.2.1. Pembahasan Uji Pelabelan pada Dataset	67
4.2.2. Pembahasan Hasil Klasifikasi <i>Multilabel Hyperparameter Default</i>	67
4.2.3. Pembahasan Hasil Klasifikasi <i>Multilabel Hyperparameter Tuning</i>	69
4.2.4. Pembahasan Performa Model	71
4.2.5. Pembahasan Uji Coba <i>Interface</i>	79
BAB V KESIMPULAN DAN SARAN	81
5.1 Kesimpulan	81
5.2 Saran	82
DAFTAR PUSTAKA	84
LAMPIRAN	91

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	5
Tabel 3.1 Jumlah Data Resep Makanan.....	30
Tabel 3.2 Fitur Dataset.....	31
Tabel 3.3 Contoh Hasil Setiap Proses pada Preprocessing Data	33
Tabel 3.4 Contoh Hasil Data Clean	36
Tabel 3.5 Contoh Hasil Pembobotan Term dengan TF-IDF.....	37
Tabel 4.1 Rangkuman Skenario Uji Coba yang Dilakukan	58
Tabel 4.2 Contoh Hasil Dataset Pelabelan Otomatis	58
Tabel 4.3 Hasil Uji Dataset pada Model dengan Hyperparameter Default	61
Tabel 4.4 Hyperparameter KNN yang Digunakan.....	62
Tabel 4.5 Hyperparameter RF yang Digunakan	62
Tabel 4.6 Hyperparameters XGBoost yang Digunakan.....	63
Tabel 4.7 Hasil Hyperparameters Tuning Setiap Model pada Dataset	64
Tabel 4.8 Hasil Uji Dataset pada Model dengan Hyperparameter Tuning.....	64
Tabel 4.9 Hasil Uji Coba Skenario	71



DAFTAR GAMBAR

Gambar 2.1 Alergen Makanan	18
Gambar 2.2 Pohon Keputusan XGBoost [40].....	23
Gambar 2.3 Proses Cross Validation	24
Gambar 3.1 Diagram Alur Penelitian	29
Gambar 3.2 Proses Pengumpulan Data.....	29
Gambar 3.3 Diagram Alur Model Klasifikasi Multilabel	40
Gambar 3.4 Sketsa GUI Deteksi Alergen	42
Gambar 3.5 Kode Semu Penerapan Preprocessing Data	43
Gambar 3.6 Kode Semu Penerapan TF-IDF	44
Gambar 3.7 Kode Semu Data Train-Test Splitting.....	45
Gambar 3.8 Penerapan Klasifikasi Multilabel KNN	46
Gambar 3.9 Kode Semu Penerapan Klasifikasi Multilabel RF	48
Gambar 3.10 Kode Semu Penerapan Klasifikasi Multilabel XGB.....	49
Gambar 3.11 Kode Semu Perhitungan Metriks Multilabel Klasifikasi	51
Gambar 3.12 Kode Semu Visualisasi Confussion Matrix	52
Gambar 3.13 Kode Semu Visualisasi Matrix per Alergen Model Default dan Tuning	53
Gambar 3.14 Potongan Kode Semu Penerapan GUI (Bagian 1)	55
Gambar 3.15 Potongan Kode Semu Penerapan GUI (Bagian 2)	55
Gambar 3.16 Potongan Kode Semu Penerapan GUI (Bagian 3)	56
Gambar 3.17 Potongan Kode Semu Penerapan GUI (Bagian 4)	56
Gambar 3.18 Hasil Implementasi GUI	57
Gambar 4.1 Persebaran Label Dataset	60
Gambar 4.2 Hasil Uji Interface	65
Gambar 4.3 Confusion Matrix Alergen Susu Model XGB Hyperparameter Default	68
Gambar 4.4 Confusion Matrix Alergen Susu Model XGB Hyperparameter Tuning	70
Gambar 4.5 Perbandingan Accuracy untuk Hyperparameter Default dan Tuning	73
Gambar 4.6 Perbandingan Precision untuk Hyperparameter Default dan Tuning	75
Gambar 4.7 Perbandingan Recall untuk Hyperparameter Default dan Tuning	76
Gambar 4.8 Perbandingan F1-Score untuk Hyperparameter Default dan Tuning	78

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	91
Lampiran 2 Curriculum Vitae	92
Lampiran 3 Surat Pernyataan HAKI.....	94
Lampiran 4 Sertifikat BNSP	96
Lampiran 5 Form Revisi Dosen Penguji.....	97
Lampiran 6 Hasil Cek Turnitin	99
Lampiran 7 Halaman Persetujuan	100

