

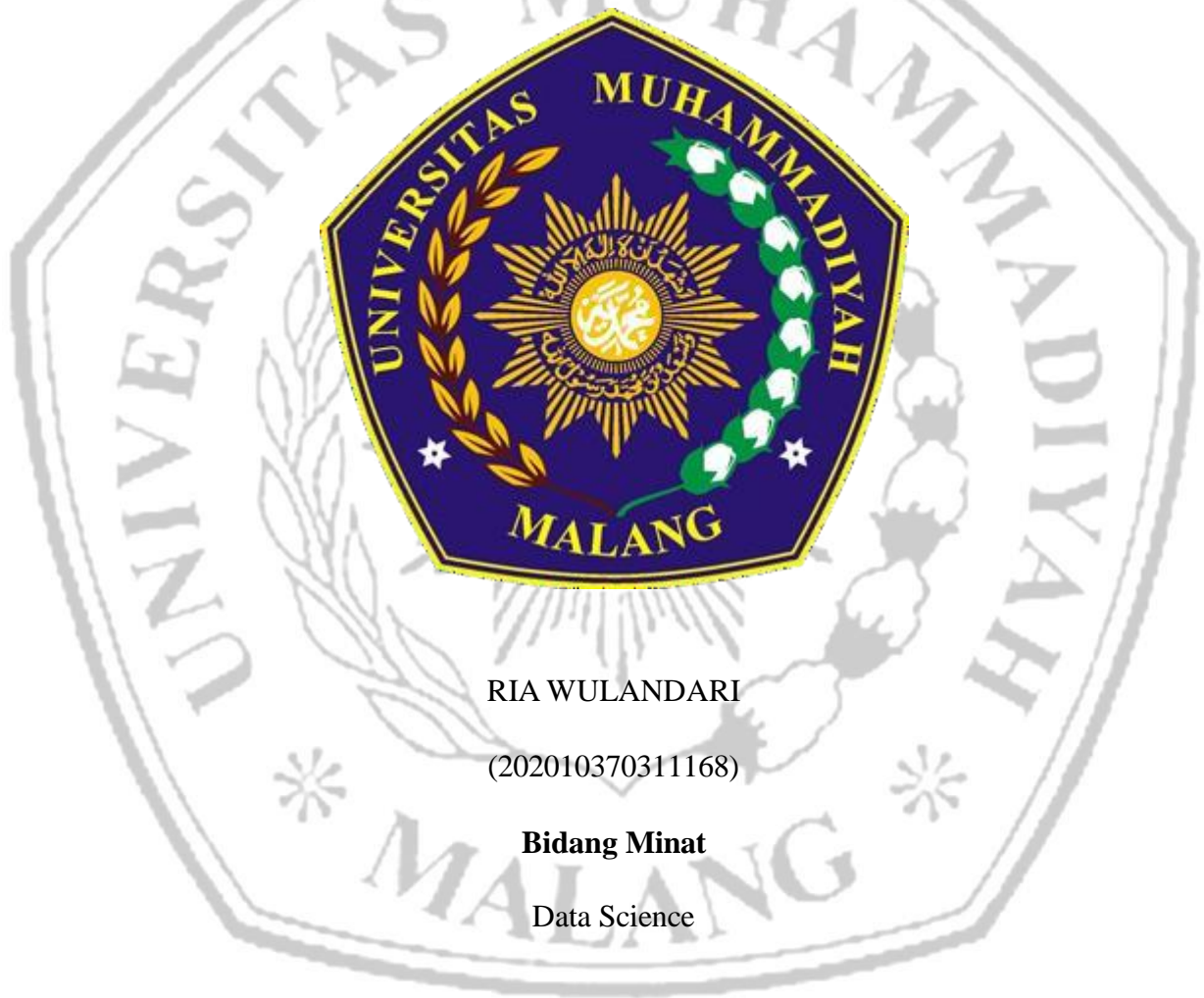
Implementasi XGBoost dan Random Forest pada Penyakit Stunting

Laporan Tugas Akhir

Diajukan Untuk Memenuhi

Persyaratan Guna Meraih Gelar Sarjana

Informatika Universitas Muhammadiyah Malang



RIA WULANDARI

(202010370311168)

Bidang Minat

Data Science

PROGRAM STUDI INFOMARTIKA FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH MALANG

2024

LEMBAR PERSETUJUAN

Implementasi XGBoost dan Random Forest pada Penyakit Stunting

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1

Informatika Universitas Muhammadiyah Malang

Menyetujui,

Malang, 24 September 2024

Dosen Pembimbing 1



Dosen Pembimbing 2



Christian Sri Kusuma Aditya S.Kom., M.Kom

NIP. 180327021991PNS.

Didih Rizki Chandranegara S.kom., M.Kom

NIP. 180302101992PNS.

LEMBAR PENGESAHAN

Implementasi XGBoost dan Random Forest pada Penyakit Stunting

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

RIA WULANDARI

202010370311168

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 24 September 2024

Menyetujui,

Dosen Penguji 1

Dosen Penguji 2



Ir. Yufis Azhar S.Kom., M.Kom.

Vinna Rahmayanti S.Si., M.Si

NIP. 10814100544PNS.

NIP. 180306071990PNS.

Mengetahui,

Ketua Jurusan Informatika



Ir. Galih Wasis Wicaksono S.kom. M.Cs.

NIP. 10814100541PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : RIA WULANDARI

NIM : 202010370311168

FAK./JUR. : Informatika

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul **“Implementasi XGBoost dan Random Forest pada Penyakit Stunting”** beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,
Dosen Pembimbing

Malang, 24 September 2024
Yang Membuat Pernyataan



Christian Sri Kusuma Aditya S.Kom.,
M.Kom

RIA WULANDARI

ABSTRAK

Stunting merupakan masalah serius di Indonesia, mempengaruhi sekitar 21,6% balita akibat kekurangan gizi kronis. Penelitian ini fokus pada implementasi metode XGBoost dan Random Forest untuk mengidentifikasi faktor risiko stunting. Metode XGBoost mencapai akurasi 90%, sedangkan Random Forest mencapai 85%. Hasil ini menunjukkan bahwa XGBoost memiliki performa lebih baik dalam mengklasifikasi risiko stunting. Penelitian ini penting untuk pengembangan strategi pencegahan yang lebih efektif terhadap stunting, mempertimbangkan faktor gizi, ekonomi, dan lingkungan.

Kata Kunci: Stunting, XGBoost, Random Forest, Implementasi, Kesehatan Anak



ABSTRACT

Stunting is a serious problem in Indonesia, affecting around 21.6% of children under five due to chronic malnutrition. This research focuses on implementing the XGBoost and Random Forest methods to identify risk factors for stunting. The XGBoost method achieved 90% accuracy, while Random Forest achieved 85%. These results show that XGBoost has better performance in classifying stunting risk. This research is important for the development of more effective prevention strategies for stunting, considering nutritional, economic and environmental factors.

Keywords: Stunting, XGBoost, Random Forest, Implementation, Children's Health



KATA PENGANTAR

Puji syukur kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “**Implementasi XGBoost dan Random Forest pada Penyakit Stunting**”.

Skripsi ini diajukan untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana pada Program Sarjana Strata 1 Fakultas Teknik Informatika Universitas Muhammadiyah Malang. Skripsi ini merupakan salah satu syarat studi yang harus ditempuh oleh seluruh mahasiswa Universitas Muhammadiyah Malang guna menyelesaikan akhir studi pada jenjang program Sarjana Strata 1.

Penulis menyadari masih terdapat banyak kekurangan dan keterbatasan di dalam penulisan skripsi ini. Oleh sebab itu, penulis mengharapkan adanya saran yang dapat membantu agar tulisan ini dapat berguna untuk perkembangan ilmu pengetahuan teknologi informasi ke depannya.

Malang, 30 agustus 2024

Penulis



Ria Wulandari

LEMBAR PERSEMBAHAN

Melalui kesempatan ini, penulis ingin mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan bantuan selama proses penulisan tugas akhir ini. Tanpa dukungan dan bimbingan kalian, penulis tidak akan mampu menyelesaikan tugas akhir ini dengan baik. Ucapan terima kasih ini penulis persembahkan kepada:

1. Allah SubhanahuWaTa'ala yang maha pemberi petunjuk serta kesehatan. Terima kasih atas segala nikmat dan karunia-Mu, yang telah memberikan penulis kekuatan dan kesabaran dalam menyelesaikan tugas akhir ini.
2. Kedua orang tua tercinta, yang selalu memberikan doa dan dukungan secara moral maupun materi.
3. Bapak Christian Sri Kusuma Aditya, S.Kom, M.Kom. dan Bapak Didih Rizki Chandranegara, S.Kom, M.Kom. selaku Dosen Pembimbing yang selalu memberi arahan dan menyempatkan waktu untuk membimbing saya dalam menyelesaikan tugas akhir.
4. Ir. Galih Wasis Wicaksono, S.Kom, M.Cs. selaku Ketua Jurusan Teknik Informatika Program Studi Teknik Informatika Universitas Muhammadiyah Malang.
5. Kepada Naufal Japa, Gabriella Nathania Manuel, Revido, Rean, Vita, hanifah, Familia Meiyasinta, Friska Ruth, Theo Susanteo, Leo Pradana, Aranazof, Thiya Mariadi, Tasya Putri, Marsha Lenathea yang selalu men-support saya selama pengerjaan skripsi ini.
6. Kepada teman-teman Caramel dan Wallaby yang memberikan bantuan dan Lingkungan pertemanan secara baik serta men-support selama masa perkuliahan saya.
7. Kepada teman magang Cece, Isnainiyahhh, Erlitak yang memberikan bantuan dan semangat mengerjakan tugas akhir ini.
8. Kepada Bu is cimol, Bu evi, Kak ayyak, Kak tika, Pak Bagyo yang men-support selama masa perkuliahan saya.
9. Kepada teman-teman Badminton yang selalu memberikan semangat

Malang, 30 agustus 2024

Penulis



Ria Wulandari

DAFTAR ISI

LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
LEMBAR PERSEMBAHAN	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
BAB II	4
2.1 Peneliti Terdahulu	4
2.1.1 Faktor Penyebab Stunting	5
2.2 Machine learning	6
2.3 XGBoost	6
2.3.1 Mekanisme Kerja XGBoost	6
2.3.2 Kelebihan dan Kekurangan XGBoost	8
2.4 Random Forest	8
2.4.1 Mekanisme Kerja Random Forest	9
2.4.2 Kelebihan dan Kekurangan Random Forest	9
2.5 Implementasi kedua algoritma	9
2.6 Perbandingan Algoritma	9
2.7 Evaluasi Model	10
2.7.1 Metrix Evaluasi	10
2.8 Confusion Matrix	10
2.9 Cross-Validation	11
2.10 Feature Importance	11

2.11 Model Comparison.....	12
2.12 Kemudahan Implementasi.....	12
2.13 Keterbatasan dan Tantangan.....	12
BAB III.....	13
3.1 Skenario Pengujian	13
3.1.1 Pengumpulan Dataset	14
3.2 Pre-Processing	15
3.3 Pembagian Dataset.....	16
3.4 Training Set	17
3.5 Test Set	17
3.6 Pembentukan Model XGBoost.....	18
3.7 Pembentukan Model Random Forest.....	19
3.8 Evaluasi Model Confusion Matrix.....	20
BAB IV	21
4.1 Implementasi.....	21
4.1.1 Pengumpulan Dataset	21
4.2 Preprocessing.....	21
4.3 Hasil pembagian dataset training dan test set.....	24
4.4 Hasil dari Pembentukan model XGBoost	25
4.5 Hasil dari Pembentukan model Random Forest.....	25
4.6 Hasil Evaluasi model Confusion Matrix XGBoost dan Akurasi	26
4.7 Hasil Evaluasi Model Confusion Matrix Random Forest dan akurasi.....	27
4.8 Uji Coba	29
4.8.1 Sebelum melakukan smote.....	29
4.8.2 Percobaan mencari best parameter dengan optuna.....	29
4.9 Analisis Perbandingan	32
4.10 Feature Importances	33
BAB V.....	35
5.1 KESIMPULAN.....	35
5.2 Saran	35
DAFTAR PUSTAKA	36

DAFTAR GAMBAR

Gambar 1 Diagram alur penelitian	13
Gambar 2 source code dan output hasil pembentukan xgboost	25
Gambar 3 source code dan output pembentukan random forest	25
Gambar 4 hasil confusion matrix xgboost	26
Gambar 5 classification Report XGBoost	27
Gambar 6 Confusion Matrix dan pohon keputusan Random Forest	27
Gambar 7 Classification Report Random Forest	28
Gambar 8 sebelum dan sesudah melakukan smote	29
Gambar 9 percobaan pertama mencari best parameter	30
Gambar 10 percobaan kedua mencari best parameter	30
Gambar 11 percobaan ketiga mencari best parameter	30
Gambar 12 percobaan keempat mencari best parameter	31
Gambar 13 Source code GridsearchCV untuk optimasi parameter model XGBoost	31
Gambar 14 nilai feature importance	33



DAFTAR TABEL

Tabel 1 Literatur Review.....	5
Tabel 2 Confusion Matrix	11
Tabel 3 atribut dan tipe data.....	14
Tabel 4 hasil data cleaning.....	22
Tabel 5 hasil tahapan mapping.....	22
Tabel 6 hasil standard scaler	23
Tabel 7 hasil balancing.....	24
Tabel 8 hasil training dan test set.....	24
Tabel 9 Analisis Perbandingan.....	32



DAFTAR PUSTAKA

- [1] A. Husaini, I. Hoeronis, H. H. Lumana, and L. D. Puspareni, "Early Detection of Stunting in Toddlers Based on Ensemble Machine Learning in Purbaratu Tasikmalaya," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 3, p. 487, 2023, doi: 10.26418/justin.v11i3.66465.
- [2] H. Khoiriyah and I. Ismarwati, "Faktor Kejadian Stunting Pada Balita : Systematic Review," *J. Ilmu Kesehat. Masy.*, vol. 12, no. 01, pp. 28–40, 2023, doi: 10.33221/jikm.v12i01.1844.
- [3] J. F. NABABAN, "Klasifikasi Penderita Stunting Dengan Metode Support Vector Machine (Studi Kasus: Lima Puskesmas Di Kota Bandar Lampung)," *Klasifikasi Penderita Stunting Dengan Metod. Support Vector Mach. (Studi Kasus Lima Puskesmas Di Kota Bandar Lampung)*, pp. 1–23, 2021.
- [4] O. S. Bachri and R. M. H. Bhakti, "PENENTUAN STATUS STUNTING PADA ANAK DENGAN MENGGUNAKAN ALGORITMA KNN Stunting Status Determination in Children using KNN ALgorithm," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 3, no. 2, pp. 130–137, 2021.
- [5] G. Chairunisa *et al.*, "Life Expectancy Prediction Using Decision Tree, Random Forest, Gradient Boosting, and XGBoost Regressions," *J. Sintak*, vol. 2, no. 2, pp. 71–82, 2024, doi: 10.62375/jsintak.v2i2.249.
- [6] M. Salsabil, N. Lutvi, and A. Eviyanti, "Implementasi Data Mining Dalam Melakukan Prediksi Penyakit Diabetes Menggunakan Metode Random Forest Dan Xgboost," *J. Ilm. Komputasi*, vol. 23, no. 1, pp. 51–58, 2024, doi: 10.32409/jikstik.23.1.3507.
- [7] M. Banurea, D. Betaria Hutagaol, and O. Sihombing, "Klasifikasi Penyakit Stunting Dengan Menggunakan Algoritma Support Vector Machine Dan Random Forest," *J. TEKINKOM*, vol. 6, no. 2, pp. 540–549, 2023, doi: 10.37600/tekinkom.v6i2.927.
- [8] W. C. Wahyudin *et al.*, "P Rediksi S Tunting P Ada B Alita D I R Umah S Akit K Ota," vol. 2019, pp. 32–36, 2023.
- [9] Y. Yuwanti, F. M. Mulyaningrum, and M. M. Susanti, "Faktor – Faktor Yang Mempengaruhi Stunting Pada Balita Di Kabupaten Grobogan," *J. Keperawatan dan*

- Kesehat. Masy. Cendekia Utama*, vol. 10, no. 1, p. 74, 2021, doi: 10.31596/jcu.v10i1.704.
- [10] E. Juliana, N. Nataliningsih, and I. Aisyah, “Pemenuhan Kebutuhan Gizi dan Perkembangan Anak,” *Sade. J. Pengabd. Kpd. Masy. Univ. Winaya Mukti*, vol. 2, no. 1, pp. 11–19, 2022.
- [11] A. Roihan, P. A. Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [12] P. Septiana Rizky, R. Haiban Hirzi, and U. Hidayaturrohmah, “Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang,” *J Stat. J. Ilm. Teor. dan Apl. Stat.*, vol. 15, no. 2, pp. 228–236, 2022, doi: 10.36456/jstat.vol15.no2.a5548.
- [13] G. Abdurrahman, H. Oktavianto, and M. Sintawati, “Optimasi Algoritma XGBoost Classifier Menggunakan Hyperparameter Gridsearch dan Random Search Pada Klasifikasi Penyakit Diabetes,” *INFORMAL Informatics J.*, vol. 7, no. 3, p. 193, 2022, doi: 10.19184/isj.v7i3.35441.
- [14] R. Supriyadi, W. Gata, N. Maulidah, and A. Fauzi, “Penerapan Algoritma Random Forest Untuk Menentukan Kualitas Anggur Merah,” *E-Bisnis J. Ilm. Ekon. dan Bisnis*, vol. 13, no. 2, pp. 67–75, 2020, doi: 10.51903/e-bisnis.v13i2.247.
- [15] I. Kurniawan, D. C. P. Buani, A. Abdussomad, W. Apriliah, and R. A. Saputra, “Implementasi Algoritma Random Forest Untuk Menentukan Penerima Bantuan Raskin,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 2, pp. 421–428, 2023, doi: 10.25126/jtiik.20231026225.
- [16] G. A. Mursianto, I. M. Falih, M. Irfan, T. Sakinah, and D. S. Prasvita, “Perbandingan Metode Klasifikasi Random Forest dan XGBoost Serta Implementasi Teknik SMOTE pada Kasus Prediksi Hujan,” *J. Senamika*, vol. 2, no. 2, pp. 41–50, 2021.
- [17] Jan Melvin Ayu Soraya Dachi and Pardomuan Sitompul, “Analisis Perbandingan Algoritma XGBoost dan Algoritma Random Forest Ensemble Learning pada Klasifikasi Keputusan Kredit,” *J. Ris. Rumpun Mat. Dan Ilmu Pengetah. Alam*, vol. 2, no. 2, pp. 87–

- 103, 2023, doi: 10.55606/jurrimipa.v2i2.1470.
- [18] T. M. Toma, K. T. Andargie, R. A. Alula, B. M. Kebede, and M. M. Gujo, “Factors associated with wasting and stunting among children aged 06–59 months in South Ari District, Southern Ethiopia: a community-based cross-sectional study,” *BMC Nutr.*, vol. 9, no. 1, pp. 1–16, 2023, doi: 10.1186/s40795-023-00683-3.
- [19] T. Astutiningsih, D. R. S. Saputro, and Sutanto, “Optimasi Algoritme Xtreme Gradient Boosting (XGBoost) pada Harga Saham PT. United Tractors Tbk.,” *SPECTA J. Technol.*, vol. 7, no. 3, pp. 632–641, 2023, doi: 10.35718/specta.v7i3.1031.
- [20] A. Wibowo, I. Darwati, and O. Irnawati, “Prediksi Operasi Sesar Dengan Machine Learning,” *JIMP - J. Inform. Merdeka Pasuruan*, vol. 4, no. 3, pp. 25–29, 2020, doi: 10.37438/jimp.v4i3.228.
- [21] M. Syukron, R. Santoso, and T. Widiharih, “Perbandingan Metode Smote Random Forest Dan Smote Xgboost Untuk Klasifikasi Tingkat Penyakit Hepatitis C Pada Imbalance Class Data,” *J. Gaussian*, vol. 9, no. 3, pp. 227–236, 2020, doi: 10.14710/j.gauss.v9i3.28915.
- [22] R. Siringoringo, R. Perangin-angin, and J. Jamaluddin, “Model Hibrid Genetic-Xgboost Dan Principal Component Analysis Pada Segmentasi Dan Peramalan Pasar,” *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 5, no. 2, pp. 97–103, 2021, doi: 10.46880/jmika.vol5no2.pp97-103.
- [23] Diana Tri Susetianingtias, Eka Patriya, and Rodiah, “Model Random Forest Regression Untuk Peramalan Penyebaran Covid-19 Di Indonesia,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 2, no. 2, pp. 84–95, 2022, doi: 10.51454/decode.v2i2.48.



UNIVERSITAS
MUHAMMADIYAH
MALANG



FAKULTAS TEKNIK

INFORMATIKA

informatika.umm.ac.id | informatika@umm.ac.id

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Ria Wulandari
 NIM : 202010370311168
 Judul TA : Implementasi XGBoost dan Random Forest pada Penyakit Stunting

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	5%
2.	Bab 2 – Daftar Pustaka	25 %	2%
3.	Bab 3 – Analisis dan Perancangan	25 %	3%
4.	Bab 4 – Implementasi dan Pengujian	15 %	2%
5.	Bab 5 – Kesimpulan dan Saran	5 %	0%
6.	Makalah Tugas Akhir	20%	6%

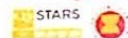
*) Hasil cek plagiarisme diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui
 Pemeriksa (Staf TU)



(.....)



Kampus I
 Jl. Bandung 1 Malang, Jawa Timur
 P. +62 341 551 253 (Hunting)
 F. +62 341 460 435

Kampus II
 Jl. Bendungan Sutarni No 188 Malang, Jawa Timur
 P. +62 341 551 149 (Hunting)
 F. +62 341 582 060

Kampus III
 Jl. Raya Tlogomas No 248 Malang, Jawa Timur
 P. +62 341 464 318 (Hunting)
 F. +62 341 460 435
 E. webmaster@umm.ac.id