

**Feature Selection of Student Performance using Random Forest
Algorithm**

Laporan Tugas Akhir

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang



Muhammad Ainur Rofiq
(201810370311146)

**Bidang Minat
Data Science**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2023**

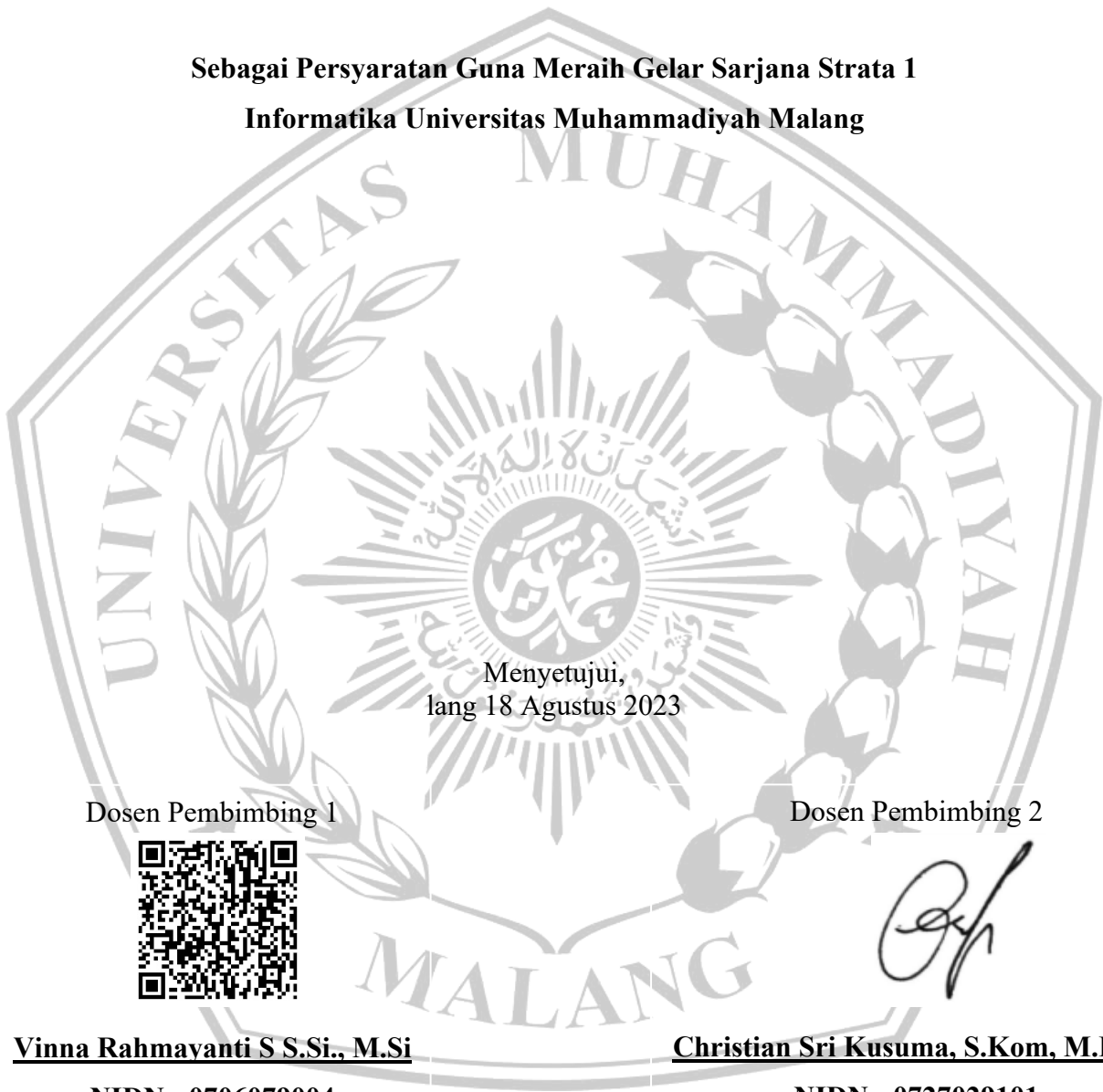
LEMBAR PERSETUJUAN

Feature Selection of Student Performance using Random Forest Algorithm

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1

Informatika Universitas Muhammadiyah Malang



Menyetujui,
tanggal 18 Agustus 2023

Dosen Pembimbing 1



Dosen Pembimbing 2

A handwritten signature in black ink is located below the text 'Dosen Pembimbing 2'. The signature is stylized and appears to be 'Christian Sri Kusuma'.

Vinna Rahmayanti S.Si., M.Si

NIDN : 0706079004

Christian Sri Kusuma, S.Kom, M.Kom

NIDN : 0727029101

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

evaluasi pembelajaran siswa menggunakan pemilihan fitur pada
model random forest

TUGAS AKHIR

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

Disusun Oleh :

Muhammad Ainur Rofiq

201810370311146

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 29 September 2023

Menyetujui,

Dosen Penguji 1



Wildan Suharso S.Kom., M.Kom

NIP. 10817030596PNS.

Dosen Penguji 2



Briansyah Setio Wiyono S.Kom.,

M.Kom

NIP. 190913071987PNS.

Mengetahui,
Ketua Jurusan Informatika



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

NIP. 10814100541PNS.

LEMBAR PERNYATAAN

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

NAMA : Muhammad Ainur Rofiq
NIM : 201810370311146
FAK./JUR. : Informatika

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul "**Feature Selection of Student Performance using Random Forest Algorithm**" 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 dikemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya say aini, atau ada klaim dari pihak lain terhadap keaslian karya say aini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,
Dosen pembimbing



Vinna Rahmayanti S, S.Si., M.Si

Malang, 18 Agustus 2023
Yang Membuat Pernyataan



Muhammad Ainur Rofiq

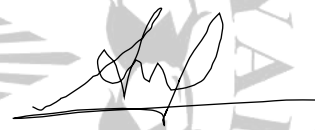
KATA PENGANTAR

Dengan memanjatkan puji syukur kehadiran Allah SWT. Atas limpahan rahmat dan hidayah-NYA sehingga peneliti dapat menyelesaikan tugas akhir yang berjudul :

“FEATURE SELECTION OF STUDENT PERFORMANCE USING RANDOM FOREST ALGORITHM”

Di dalam tulisan ini disajikan pokok – pokok bahasan yang meliputi pendahuluan, studi literatur, metode penelitian dan hasil penelitian yang telah didapatkan dari hasil penelitian dan disimpulkan berdasarkan hasil yang telah di dapat oleh peneliti. Peneliti menyadari sepenuhnya bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan keterbatasan. Oleh karena itu, peneliti mengharapkan saran yang membangun agar tulisan ini bermanfaat bagi perkembangan ilmu pengetahuan.

Malang, 18 Agustus 2022



Muhammad Ainur Rofiq



DAFTAR ISI

LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	1
BAB I PENDAHULUAN	2
1.1. Latar Belakang	2
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	4
1.4. Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu	5
2.2. Studi Pustaka	5
2.2.1. Filter Method	5
2.2.2. Wrapper Method	5
2.2.3. Embeded Method	5
2.2.4. PCA	6
2.2.5. Univariate Feature Selection	6
2.2.6. Regularization Method	6
BAB III METODE PENELITIAN	7
3.1. Dataset	7
3.2. Data Preprocessing	8
3.3. Modelling	8
BAB IV HASIL DAN PEMBAHASAN	12
BAB V KESIMPULAN	14
DAFTAR PUSTAKA	15

DAFTAR PUSTAKA

- [1] K. Hosny and A. El-korany, "Applying adaptive learning by integrating semantic and machine learning in proposing student assessment model," *Int. J. Electr. Comput. Eng. IJECE*, vol. 12, no. 2, p. 2014, Apr. 2022, doi: 10.11591/ijece.v12i2.pp2014-2025.
- [2] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, p. 11, Dec. 2022, doi: 10.1186/s40561-022-00192-z.
- [3] N. R. Yadav and S. S. Deshmukh, "Prediction of Student Performance Using Machine Learning Techniques: A Review," in *Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)*, S. Tamane, S. Ghosh, and S. Deshmukh, Eds., in *Advances in Computer Science Research*, vol. 105. Dordrecht: Atlantis Press International BV, 2023, pp. 735–741. doi: 10.2991/978-94-6463-136-4_63.
- [4] A. Tarik, H. Aissa, and F. Yousef, "Artificial Intelligence and Machine Learning to Predict Student Performance during the COVID-19," *Procedia Comput. Sci.*, vol. 184, pp. 835–840, 2021, doi: 10.1016/j.procs.2021.03.104.
- [5] Software Engineering, Wolkite University, Ethiopia, E. B. Belachew, F. A. Gobena, and Computer Sciences, Wolkite University, Ethiopia, "Student Performance Prediction Model using Machine Learning Approach: The Case of Wolkite University," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 7, no. 2, pp. 46–50, Feb. 2017, doi: 10.23956/ijarcsse/V7I2/01219.
- [6] A. Rivas, J. M. Fraile, P. Chamoso, A. González-Briones, S. Rodríguez, and J. M. Corchado, "Students Performance Analysis Based on Machine Learning Techniques," in *Learning Technology for Education Challenges*, L. Uden, D. Liberona, G. Sanchez, and S. Rodríguez-González, Eds., in *Communications in Computer and Information Science*, vol. 1011. Cham: Springer International Publishing, 2019, pp. 428–438. doi: 10.1007/978-3-030-20798-4_37.
- [7] A. Rivas, A. González-Briones, G. Hernández, J. Prieto, and P. Chamoso, "Artificial neural network analysis of the academic performance of students in virtual learning environments," *Neurocomputing*, vol. 423, pp. 713–720, Jan. 2021, doi: 10.1016/j.neucom.2020.02.125.
- [8] M. B. Imani, M. R. Keyvanpour, and R. Azmi, "A NOVEL EMBEDDED FEATURE SELECTION METHOD: A COMPARATIVE STUDY IN THE APPLICATION OF TEXT CATEGORIZATION," *Appl. Artif. Intell.*, vol. 27, no. 5, pp. 408–427, May 2013, doi: 10.1080/08839514.2013.774211.
- [9] A. Alshantqi and A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," *IEEE Access*, vol. 8, pp. 203827–203844, 2020, doi: 10.1109/ACCESS.2020.3036572.
- [10] Y. A. Alsariera, Y. Baashar, G. Alkaws, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and Evaluation of Different Machine Learning Algorithms for Predicting Student Performance," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, May 2022, doi: 10.1155/2022/4151487.
- [11] N. Sánchez-Marño, A. Alonso-Betanzos, and M. Tombilla-Sanromán, "Filter Methods for Feature Selection – A Comparative Study," in *Intelligent Data Engineering and Automated Learning - IDEAL 2007*, H. Yin, P. Tino, E. Corchado, W. Byrne, and X. Yao, Eds., in *Lecture Notes in Computer Science*, vol. 4881. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 178–187. doi: 10.1007/978-3-540-77226-2_19.
- [12] A. Bhusal, "Predicting Student's Performance Through Data Mining".
- [13] N. Pudjihartono, T. Fadason, A. W. Kempa-Liehr, and J. M. O'Sullivan, "A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction," *Front. Bioinforma.*, vol. 2, p. 927312, Jun. 2022, doi: 10.3389/fbinf.2022.927312.
- [14] A. Paramita, "Implementing Machine Learning Techniques for Predicting Student Performance in an E-Learning Environment," *IJIS Int. J. Inform. Inf. Syst.*, vol. 4, no. 2, pp. 149–156, Sep. 2021, doi: 10.47738/ijis.v4i2.112.
- [15] D. P. M. Abellana and D. M. Lao, "A new univariate feature selection algorithm based on the best–worst multi-attribute decision-making method," *Decis. Anal. J.*, vol. 7, p. 100240, Jun. 2023, doi: 10.1016/j.dajour.2023.100240.
- [16] F. Song, Z. Guo, and D. Mei, "Feature Selection Using Principal Component Analysis," in *2010 International Conference on System Science, Engineering Design and Manufacturing Informatization*, Yichang, China: IEEE, Nov. 2010, pp. 27–30. doi: 10.1109/ICSEM.2010.14.
- [17] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithm: A review," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1060–1073, Apr. 2022, doi: 10.1016/j.jksuci.2019.06.012.
- [18] D. Thakur and S. Biswas, "Guided regularized random forest feature selection for smartphone based human activity recognition," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 7, pp. 9767–9779, Jul. 2023, doi: 10.1007/s12652-022-03862-5.

- [19] L. Sandra, F. Lumbangaol, and T. Matsuo, "Machine Learning Algorithm to Predict Student's Performance: A Systematic Literature Review," *TEM J.*, pp. 1919–1927, Nov. 2021, doi: 10.18421/TEM104-56.
- [20] K. Mthimunye and F. M. Daniels, "Predictors of academic performance, success and retention amongst undergraduate nursing students: A systematic review," *South Afr. J. High. Educ.*, vol. 33, no. 1, Apr. 2019, doi: 10.20853/33-1-2631.





UNIVERSITAS
MUHAMMADIYAH
MALANG



FAKULTAS TEKNIK

INFORMATIKA

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

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Muhammad Ainur Rofiq
 NIM : 201810370311146
 Judul TA : Feature Selection of Student Performance using Random Forest Algorithm

Hasil Cek Plagiarisme dengan Turnitin

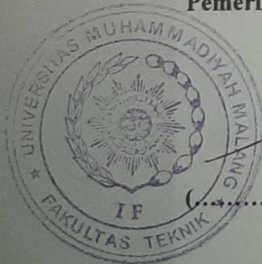
No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	6%
2.	Bab 2 – Daftar Pustaka	25 %	6%
3.	Bab 3 – Analisis dan Perancangan	25 %	1%
4.	Bab 4 – Implementasi dan Pengujian	15 %	4%
5.	Bab 5 – Kesimpulan dan Saran	5 %	0%
6.	Makalah Tugas Akhir	20%	16%

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

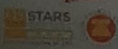
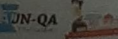
*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staff TU)



[Handwritten Signature]
 Mps an



Kampus I
 Jl. Bandung 1 Malang, Jawa Timur
 P. +62 341 951 253 (Hunting)
 F. +62 341 460 435

Kampus II
 Jl. Bendungan Sutarni No 188 Malang, Jawa Timur
 P. +62 341 951 149 (Hunting)
 F. +62 341 682 080

Kampus III
 Jl. Raya Tlocomas No 246 Malang, Jawa Timur
 P. +62 341 484 318 (Hunting)
 F. +62 341 480 435
 E. webmaster@umm.ac.id