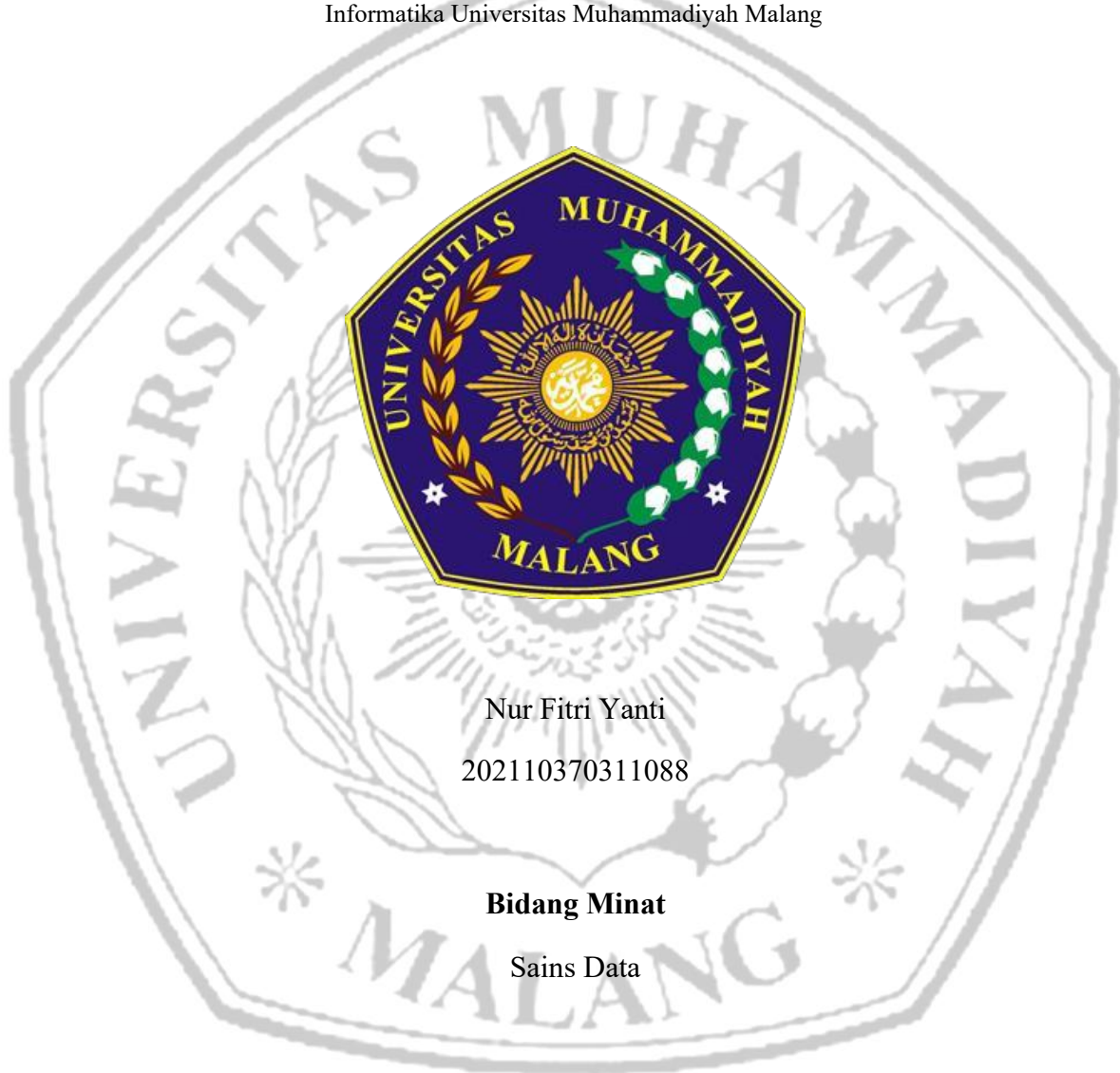


Analisis Fitur Dan Prediksi Loyalitas Pelanggan Menggunakan Algoritma Pembelajaran Mesin

Proposal Tugas Akhir

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



Nur Fitri Yanti

202110370311088

Bidang Minat

Sains Data

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

2025

LEMBAR PERSETUJUAN

**Analisis Fitur Dan Prediksi Loyalitas Pelanggan Menggunakan
Algoritma Pembelajaran Mesin**

TUGAS AKHIR

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

Menyetujui,

Malang, 7 Oktober 2025

Dosen Pembimbing 1

Dosen Pembimbing 2



NIP.

Christian Sri Kusuma Aditya S.Kom.,

M.Kom

NIP. 180327021991PNS.

LEMBAR PENGESAHAN

**Analisis Fitur Dan Prediksi Loyalitas Pelanggan Menggunakan
Algoritma Pembelajaran Mesin**

TUGAS AKHIR

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

Disusun Oleh :

Nur Fitri Yanti

202110370311088

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis pengujian
pada tanggal 7 Oktober 2025

Menyetujui,

Dosen Penguji 1



Vinna Rahmayanti S S.Si., M.Si

NIP. 180306071990PNS.

Dosen Penguji 2



Nur Hayatin S. ST., M.Kom

NIP. 10809070476PNS.

Mengetahui,

Ketua Jurusan Informatika



Agus Eko Minarno S.Kom., M.Kom. IPM.

NIP. 10814100540PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : Nur Fitri Yanti

NIM : 202110370311088

FAK./JUR. : Informatika

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul “Analisis Fitur Dan Prediksi Loyalitas Pelanggan Menggunakan Algoritma Pembelajaran Mesin” 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



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

Malang, 7 Oktober 2025
Yang Membuat Pernyataan



SPESIAL BUKU KEPAN
10000
METERAI
TEMPEL
1D177ANX116745955
NUR FITRI YANTI
Nur Fitri Yanti

ABSTRAK

Industri telekomunikasi memiliki tingkat persaingan tinggi, sehingga memahami perilaku pelanggan menjadi kunci dalam mengurangi *customer churn*, yaitu kondisi ketika pelanggan berhenti berlangganan. Penelitian ini bertujuan untuk mengidentifikasi fitur-fitur yang berpengaruh terhadap churn serta membangun model prediksi menggunakan dua algoritma pembelajaran mesin, yaitu XGBoost dan Random Forest. Dataset yang digunakan berasal dari *Telco Customer Churn* (Kaggle) dengan total 7.043 data pelanggan. Proses penelitian mencakup *preprocessing data*, seleksi fitur menggunakan Sequential Feature Selection (SFS), dan interpretasi model menggunakan SHAP (Shapley Additive Explanations) untuk menjelaskan kontribusi setiap fitur terhadap prediksi churn.

Hasil penelitian menunjukkan bahwa fitur Contract, Tenure, dan MonthlyCharges menjadi faktor paling berpengaruh terhadap churn. Metode SFS meningkatkan performa model, di mana Random Forest mencapai akurasi 0.7915, sedangkan XGBoost memperoleh akurasi 0.7863. Analisis SHAP memperkuat hasil tersebut dengan menunjukkan bahwa pelanggan dengan kontrak jangka panjang dan masa langganan lama memiliki kecenderungan churn lebih rendah. Kombinasi metode SFS dan SHAP tidak hanya meningkatkan akurasi model, tetapi juga memberikan interpretasi yang lebih mendalam terhadap faktor-faktor penentu churn pelanggan.

Kata Kunci: Customer Churn, XGBoost, Random Forest, Sequential Feature Selection (SFS), SHAP

ABSTRACT

The telecommunications industry faces intense competition, making it crucial to understand customer behavior in order to reduce *customer churn* a condition where customers discontinue their subscriptions. This study aims to identify the most influential features affecting churn and to build a predictive model using two machine learning algorithms: XGBoost and Random Forest. The dataset used in this research is the *Telco Customer Churn* dataset from Kaggle, consisting of 7,043 customer records. The research process includes data preprocessing, feature selection using Sequential Feature Selection (SFS), and model interpretation with SHAP (Shapley Additive Explanations) to explain the contribution of each feature to churn prediction.

The results show that the features Contract, Tenure, and MonthlyCharges are the most influential factors in predicting churn. The application of SFS improved the model performance, with Random Forest achieving an accuracy of 0.7915, while XGBoost obtained 0.7863. The SHAP analysis reinforced these findings by revealing that customers with long-term contracts and longer subscription periods tend to have a lower likelihood of churn. The combination of SFS and SHAP not only enhances model accuracy but also provides deeper interpretability regarding the key factors influencing customer churn.

Keywords: Customer Churn, XGBoost, Random Forest, Sequential Feature Selection, SHAP.

LEMBAR PERSEMBAHAN

Puji syukur kepada Allah SWT yang senantiasa melimpahkan kasih sayang- Nya dan memberikan kemudahan. Skripsi ini merupakan persembahan kecil tanda cinta dan terima kasih yang saya berikan untuk orang-orang yang telah membuat hidup saya berharga dan bermakna. Saya persembahkan kepada:

1. Kedua orang tua yang sangat penulis sayangi dan hormati Ayah (Muzaki) dan Ibu(Nor Hayati) yang senantiasa mendoakan , memberi nasehat dan selalu menyemangati penulis. Ayah dan ibu yang selalu memberi dukungan penuh terhadap apa yang penulis kerjakan. penulis mengucapkan terima kasih atas semua hal yang sudah diberikan, semua doa, kasih sayang dan pengorbanan untuk penulis menjadi semangat untuk melakukan yang terbaik.
2. Adik yang penulis sayangi Selviana Rahmah dan Nirmala Nur Salshabilla yang selalu membuat penulis tersenyum dan semangat disaat sedang merasa sedih. Terimakasih banyak atas semua hal yang telah kalian lakukan selalu menjadi semangat penulis hingga saat ini dan seterusnya.
3. Bapak Christian Sri Kusuma Aditya, S.Kom., M.Kom, selaku dosen pembimbing dan juga dosen wali penulis, yang telah meluangkan waktu dan tenaga untuk membimbing penulis, dan segala arahan yang diberikan selama penulis menempuh studi dan menyelesaikan tugas akhir hingga selesai dengan baik.
4. Bapak/Ibu Dekan Fakultas Teknik Universitas Muhammadiyah Malang dan Bapak/Ibu Dosen Jurusan Informatika Universitas Muhammadiyah Malang, yang telah memberikan ilmu yang bermanfaat selama perkuliahan.
5. Kepada teman-teman yang selalu membantu serta menemani penulis dari awal perkuliahan hingga akhir. Terimakasih banyak atas segala hal yang telah kalian berikan seperti motivasi, semangat, saran dan sebagainya sehingga membuat penulis tidak merasa sendirian saat perkuliahan.
6. Terakhir untuk diri saya sendiri, Nur Fitri Yanti. Terimakasih sudah bertahan sejauh ini meski sering menyalahkan diri sendiri atau merasa selalu tidak yakin atas apa yang dijalani, tapi pada akhirnya mampu mengatasi hal tersebut hingga saat ini. Semoga kedepannya ada banyak hal baik yang menyertai penulis.

KATA PENGANTAR

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

” Analisis Fitur Dan Prediksi Loyalitas Pelanggan Menggunakan Algoritma Pembelajaran Mesin”

Di dalam tulisan ini disajikan pokok-pokok bahasan yang meliputi latar belakang, metode penelitian serta hasil pembahasan dari proses penelitian dan telah disimpulkan berdasarkan hasil yang telah diperoleh pada proses penelitian ini.

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, 12 September 2025



Nur Fitri Yanti

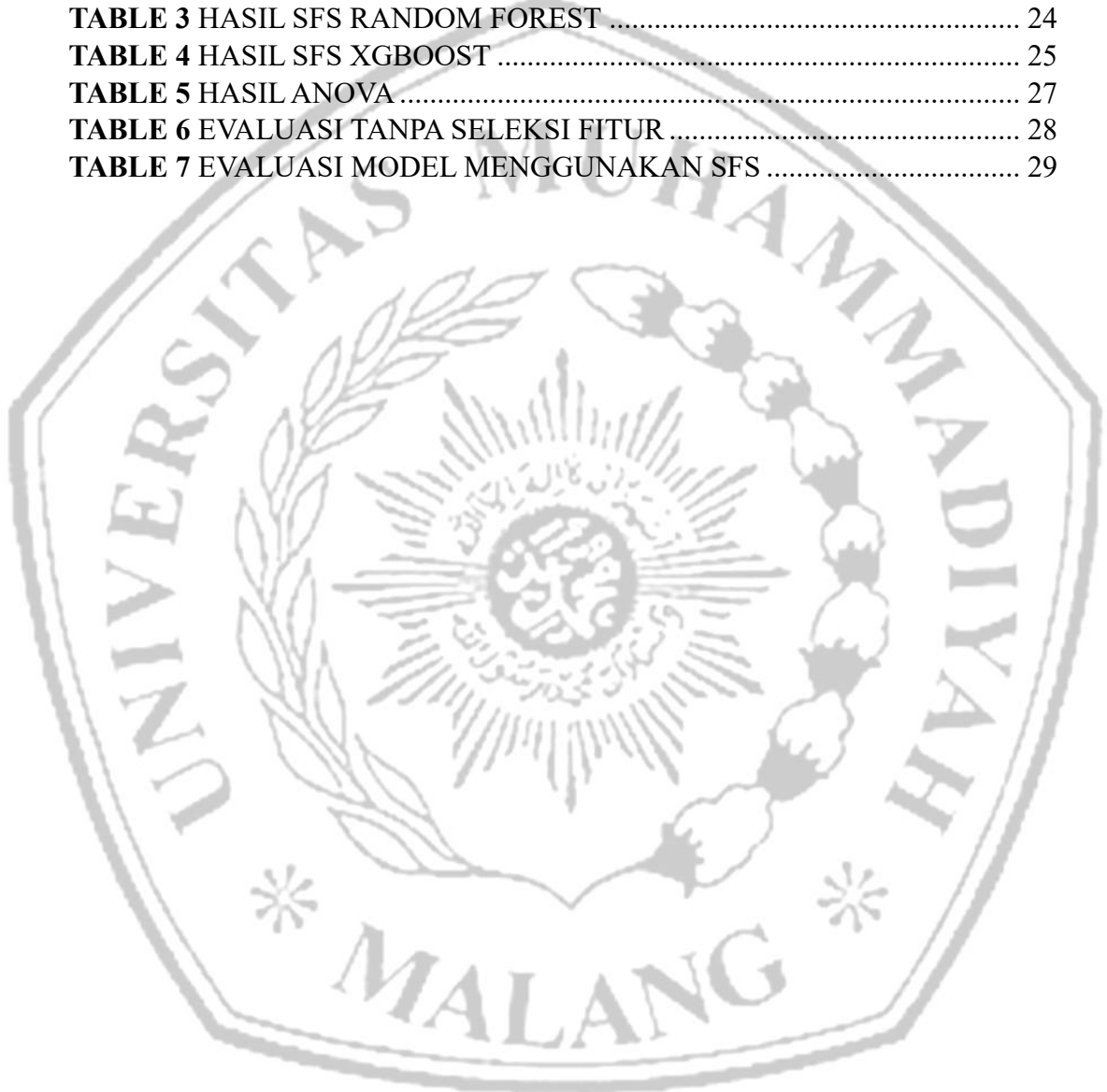
DAFTAR ISI

LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAA	iii
ABSTRAK.....	iv
ABSTRACT.....	v
LEMBAR PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
BAB 1	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah.....	3
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu.....	5
2.2 Loyalitas Pelanggan.....	7
2.3 Pemilihan Fitur dan Teknik Interpretasi.....	7
2.3.1 SFS.....	7
2.3.2 SHAP	8
2.3.3 ANOVA.....	9
2.4 XGBoost	10
2.5 Random Forest.....	11
2.6 Bahasa Python.....	11
BAB III	13
METODOLOGI PENELITIAN	13
3.1 Identifikasi Masalah	13
3.2 Alur Penelitian.....	13

3.3 Exploratory Data Analysis.....	14
3.4 Dataset	15
3.5 Data Preprocessing	16
3.6 Penerapan Teknik Seleksi Fitur.....	18
3.7 Model.....	19
3.8 Evaluasi Model.....	20
3.9 interpretasi hasil	21
3.10 Skenario Uji Coba.....	22
BAB IV.....	24
HASIL DAN PEMBAHASAN.....	24
4.1 Analisi dan Seleksi Fitur	24
4.1.1 SFS	24
4.1.2 ANOVA.....	27
4.2 Evaluasi Kinerja Model.....	28
4.2.1 Evaluasi model.....	28
4.2.2 Evaluasi model seleksi fitur	28
4.3 Interpretasi dan Implikasi Bisnis	29
4.3.1 Interpretasi model SHAP	30
4.3.2 Analisis Hubungan dan Karakteristik Pelanggan yang Churn	32
4.3.3 Rekomendasi Strategi Bisnis	33
BAB V.....	36
KESIMPULAN DAN SARAN.....	36
5.1 Kesimpulan.....	36
5.2 Saran	36
DAFTAR PUSTAKA	37

DAFTAR TABEL

TABLE 1 PENELITIAN TERDAHULU.....	5
TABLE 2 NORMALISASI DATA.....	17
TABLE 3 HASIL SFS RANDOM FOREST	24
TABLE 4 HASIL SFS XGBOOST	25
TABLE 5 HASIL ANOVA	27
TABLE 6 EVALUASI TANPA SELEKSI FITUR.....	28
TABLE 7 EVALUASI MODEL MENGGUNAKAN SFS.....	29



DAFTAR GAMBAR

GAMBAR 1 ALUR PENELITIAN	14
GAMBAR 2 TREN AKURASI SFS	26
GAMBAR 3 SHAP SUMMARY PLOT	30
GAMBAR 4 SHAP DEPENDENCE PLOT	31



DAFTAR PUSTAKA

- [1] P. Senthana, R. M. K. T. Rathnayaka, B. Kuhaneswaran, and B. T. G. S. Kumara, "Development of churn prediction model using XGBoost - Telecommunication Industry in Sri Lanka," in *2021 IEEE International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Apr. 2021. doi: 10.1109/IEMTRONICS52119.2021.9422657.
- [2] T. Tasmalaila Hanifa and S. Al-Faraby, "Analisis Churn Prediction pada Data Pelanggan PT. Telekomunikasi dengan Logistic Regression dan Underbagging."
- [3] M. Óskarsdóttir, C. Bravo, W. Verbeke, C. Sarraute, B. Baesens, and J. Vanthienen, "Social network analytics for churn prediction in telco: Model building, evaluation and network architecture," *Expert Syst Appl*, vol. 85, pp. 204–220, Nov. 2017, doi: 10.1016/j.eswa.2017.05.028.
- [4] B. Prabadevi, R. Shalini, and B. R. Kavitha, "Customer churning analysis using machine learning algorithms," *International Journal of Intelligent Networks*, vol. 4, pp. 145–154, Jan. 2023, doi: 10.1016/j.ijin.2023.05.005.
- [5] S. Wu, W. C. Yau, T. S. Ong, and S. C. Chong, "Integrated Churn Prediction and Customer Segmentation Framework for Telco Business," *IEEE Access*, vol. 9, pp. 62118–62136, 2021, doi: 10.1109/ACCESS.2021.3073776.
- [6] Y. A. Kristanto and F. Ekonomi, "Pengaruh Kualitas Layanan Terhadap Loyalitas Pelanggan Yang Dimediasi Oleh Kepuasan Konsumen (Studi Kasus Pada Usaha Susu Murni Siliwangi A-Karsan Bandung)," *Jurnal Ilmu Manajemen*, vol. 19, pp. 13–22, 2022.
- [7] J. Maan and H. Maan, "Customer Churn Prediction Model using Explainable Machine learning," *International Journal of Computer Science Trends and Technology*, vol. 11, [Online]. Available: www.ijestjournal.org
- [8] A. B. Alpriansah and Y. Ramdhani, "SISTEMASI: Jurnal Sistem Informasi Optimasi Fitur dengan Forward Selection pada Estimasi Tingkat Obesitas menggunakan Random Forest Feature Optimization with Forward Selection on Obesity Rate Estimation using Random Forest." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [9] "PERBANDINGAN METODE SELEKSI FITUR FILTER."
- [10] M. F. Thoriq, W. J. Pranoto, and F. Faldi, "Penerapan Seleksi Fitur Analysis of Variance Pada Algoritma Random Forest Classifier Dalam Klasifikasi Nilai Mahasiswa," *Explore: Jurnal Sistem Informasi dan Telematika*, vol. 14, no. 2, p. 185, Dec. 2023, doi: 10.36448/jsit.v14i2.3187.
- [11] "IMPLEMENTASI METODE RANDOM FOREST DAN."
- [12] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," May 2017, [Online]. Available: <http://arxiv.org/abs/1705.07874>

- [13] H. Kim, M. Khomidov, and J. H. Lee, "XGBoost and SHAP-Based Analysis of Risk Factors for Hypertension Classification in Korean Postmenopausal Women," *Bioengineering*, vol. 12, no. 6, Jun. 2025, doi: 10.3390/bioengineering12060659.
- [14] J. Zhang and Z. Zhao, "Corporate ESG rating prediction based on XGBoost-SHAP interpretable machine learning model," *Expert Syst Appl*, vol. 295, Jan. 2026, doi: 10.1016/j.eswa.2025.128809.
- [15] E. Zdravevski, P. Lameski, C. Apanowicz, and D. Ślęzak, "From Big Data to business analytics: The case study of churn prediction," *Applied Soft Computing Journal*, vol. 90, May 2020, doi: 10.1016/j.asoc.2020.106164.
- [16] Ety Sutanty, Maukar, Dina Kusuma Astuti, and Handayani, "Penerapan Model Arsitektur VGG16 Untuk Klasifikasi Jenis Sampah," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 407–419, Sep. 2023, doi: 10.51454/decode.v3i2.331.
- [17] "PENGENALAN BAHASA PEMROGRAMAN PYTHON MENGGUNAKAN APLIKASI GAMES UNTUK SISWA/I DI WILAYAH KEMBANGAN UTARA." [Online]. Available: <https://codecombat.com/>.
- [18] Q. H. Nguyen *et al.*, "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil," *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/4832864.
- [19] I. Gede, I. Sudipa, and M. Darmawiguna, "BUKU AJAR DATA MINING." [Online]. Available: <https://www.researchgate.net/publication/377415198>
- [20] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, "A survey on missing data in machine learning," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00516-9.
- [21] S. Dan *et al.*, "Analisis Korelasi," *Jurnal Ilmiah Wahana Pendidikan*, vol. 2024, no. 16, pp. 357–365, doi: 10.5281/zenodo.13763658.



UNIVERSITAS
MUHAMMADIYAH
MALANG



FAKULTAS TEKNIK

INFORMATIKA

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

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Nur Fitri Yanti
NIM : 202110370311088
Judul TA : Analisis fitur dan prediksi loyalitas pelanggan menggunakan algoritma pembelajaran mesin

Hasil Cek Plagiarisme dengan Turnitin

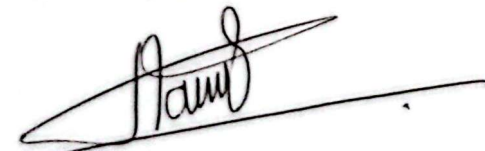
No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	0%
2.	Bab 2 – Daftar Pustaka	25 %	9%
3.	Bab 3 – Analisis dan Perancangan	25 %	0%
4.	Bab 4 – Implementasi dan Pengujian	15 %	2%
5.	Bab 5 – Kesimpulan dan Saran	5 %	5%
6.	Makalah Tugas Akhir	20%	0%

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

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

Mengetahui,

Pemeriksa (Staff TU)


 (.....)



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

Kampus II
 Jl. Bendungan Sutani 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