

**ANALISIS TINGKAT KECANDUAN MEDIA SOSIAL:
PERBANDINGAN KINERJA ALGORITMA REGRESI LINEAR
dan RANDOM FOREST DALAM MERAMALKAN PERILAKU
PENGGUNA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

MARDIYATUN MARDIAH

20.11.3617

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS TINGKAT KECANDUAN MEDIA SOSIAL:
PERBANDINGAN KINERJA ALGORITMA REGRESI LINEAR
dan RANDOM FOREST DALAM MERAMALKAN PERILAKU
PENGGUNA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi SI Informatika



disusun oleh

MARDIYATUN MARDIAH

20.11.3617

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS TINGKAT KECANDUAN MEDIA SOSIAL: PERBANDINGAN
KINERJA ALGORITMA REGRESI LINEAR DAN RANDOM FOREST
DALAM MERAMALKAN PERILAKU PENGGUNA**

yang disusun dan diajukan oleh

Mardiyatun Mardiah

20.11.3617

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 14 Mei 2024

Dosen Pembimbing,



Kusnawi, S.Kom., M.Eng.
NIK. 190302112

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS TINGKAT KECANDUAN MEDIA SOSIAL: PERBANDINGAN
KINERJA ALGORITMA REGRESI LINEAR DAN RANDOM FOREST
DALAM MERAMALKAN PERILAKU PENGGUNA**

yang disusun dan diajukan oleh

Mardiyatun Mardiah

20.11.3617

Telah dipertahankan di depan Dewan Penguji
pada tanggal 14 Mei 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom, M.Eng
NIK. 190302375

Ike Verawati, M. Kom
NIK. 190302237

Kusnawi, S.Kom, M. Eng.
NIK. 190302112



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 14 Mei 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Mardiyatun Mardiah
NIM : 20.11.3617

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Tingkat Kecanduan Media Sosial: Perbandingan Kinerja Algoritma Regresi Linear dan Random Forest dalam Meramalkan Perilaku Pengguna

Dosen Pembimbing : Kusnawi, S.Kom, M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 14 Mei 2024

Yang Menyatakan,



Mardiyatun Mardiah

HALAMAN PERSEMBAHAN

Alhamdulillah rabbil 'alamin, puji syukur bagi Allah SWT. yang telah memberikan rahmat, karunia dan hidayah-Nya sehingga penulis dapat menyelesaikan tugas skripsi dengan sebaik-baiknya. Dengan ini penulis mempersembahkan skripsi ini kepada semua pihak yang telah membantu dan memberikan motivasi serta dukungan yang besar sehingga dapat diselesaikan dengan sebaik-baiknya.

Melalui skripsi ini saya mempersembahkan terima kasih dan syukur kepada:

1. Rasa Syukur kepada ALLAH SWT yang telah memberikan rahmat berupa karunia untuk terus mengucapkan Syukur dan sabar dalam menyelesaikan tugas akhir skripsi.
2. Orang tua yang sangat saya cintai, bapak Agus Salim, S.H dan ibu Salmah, S.Pd semoga diberikan kesehatan atas segala doa dan motivasi serta dukungannya.
3. Saudara saya Irwansyah, Mayangsari, Ardiansyah, Gino, dan Lubis yang selalu memberikan dukungan dan menghibur kepada saya dalam mengerjakan tugas akhir ini.
4. Pak Kusnawi, S.Kom, M.Eng yang telah membimbing saya dalam menyelesaikan skripsi ini.
5. Ibu dan Bapak dosen yang telah memberikan ilmu dan mengajar selama berkuliah di AMIKOM.
6. Teman-teman "SKIRPSHIT KELAR MAKSIMAL SMSTER 8" yang selalu menemani penulis selama pengerjaan baik di kampus ataupun ditempat lainnya, menghibur dan memberikan dukungan setiap harinya untuk menyelesaikan tugas akhir ini.
7. Teman-teman pengurus periode 2022/2023 di HMIF yang bersama dalam berjuang mencapai kesuksesan yang penuh dengan suka dan duka.
8. Seluruh pihak yang membantu penulis selama mengerjakan tugas akhir, yang tidak dapat disebutkan satu per satu.

KATA PENGANTAR

Alhamdulillah rabbil'alamin, segala puji dan syukur atas segala Rahmat kesehatan dan kekuatan sehingga penulis mampu menyelesaikan skripsi yang berjudul **"Analisis Tingkat Kecanduan Media Sosial: Perbandingan Kinerja Algoritma Regresi Linear dan Random Forest dalam Meramalkan Perilaku Pengguna"**. Oleh sebab itu penulis mengucapkan syukur karena dapat menyelesaikan skripsi ini. Skripsi ini diajukan kepada program studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta.

Selama penyusunan skripsi ini, penulis telah banyak mendapatkan bimbingan dan dukungan dari berbagai pihak, baik berupa kritik, saran, motivasi, serta bantuan lainnya. Oleh sebab itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua Orang tua dan kakak-kakak saya yang telah memberikan dukungan, doa, motivasi, dan pendidikan hingga ke jenjang perguruan tinggi saat ini.
2. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas AMIKOM Yogyakarta
3. Bapak Hanif Al Fatta, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta beserta seluruh jajarannya.
4. Ibu Windha Mega Pradnya Dhuhiita, M.Kom, selaku Kepala Program Studi Informatika Universitas AMIKOM Yogyakarta beserta seluruh jajarannya
5. Bapak Kusnawi, S.Kom, M.Eng., selaku dosen pembimbing saya yang telah memberikan waktu, saran, kritik, ilmu, motivasi, serta bimbingannya dalam menyusun skripsi ini sehingga dapat diselesaikan dengan baik.
6. Bapak dan Ibu Dosen program Studi Informatika Universitas AMIKOM Yogyakarta yang telah mengajarkan dari semester awal hingga akhir sehingga saya memperoleh banyak ilmu dan pengalaman.

Yogyakarta, 14 Mei 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur.....	5
2.2 Dasar Teori	15
2.2.1 Regresi	15
2.2.2 Media Sosial.....	15
2.2.3 Perilaku Pengguna	16
2.2.4 Feature Selection.....	16

2.2.5	Random Forest	18
2.2.6	Regresi Linear	19
2.2.7	Root Mean Squared Error (RMSE)	20
2.2.8	Mean Absolute Error (MAE)	20
2.2.9	Python	21
BAB III METODE PENELITIAN		22
3.1	Objek Penelitian	22
3.2	Alur Penelitian	22
3.3	Alat dan Bahan	27
BAB IV HASIL DAN PEMBAHASAN		29
4.1	<i>Pre-Processing</i>	29
4.2	Label Encoding	35
4.3	Pembagian Data Latih dan Data Uji	35
4.4	Feature Selection	36
4.5	Regresi	42
4.5.1	Random Forest	42
4.5.2	Linear Regression	44
4.6	Evaluasi	45
4.6.1	Random Forest	45
4.6.2	Linear Regression	45
4.7	Visualiasi	46
BAB V PENUTUP		50
5.1	Kesimpulan	50
5.2	Saran	50
REFERENSI		51
LAMPIRAN		55

DAFTAR TABEL

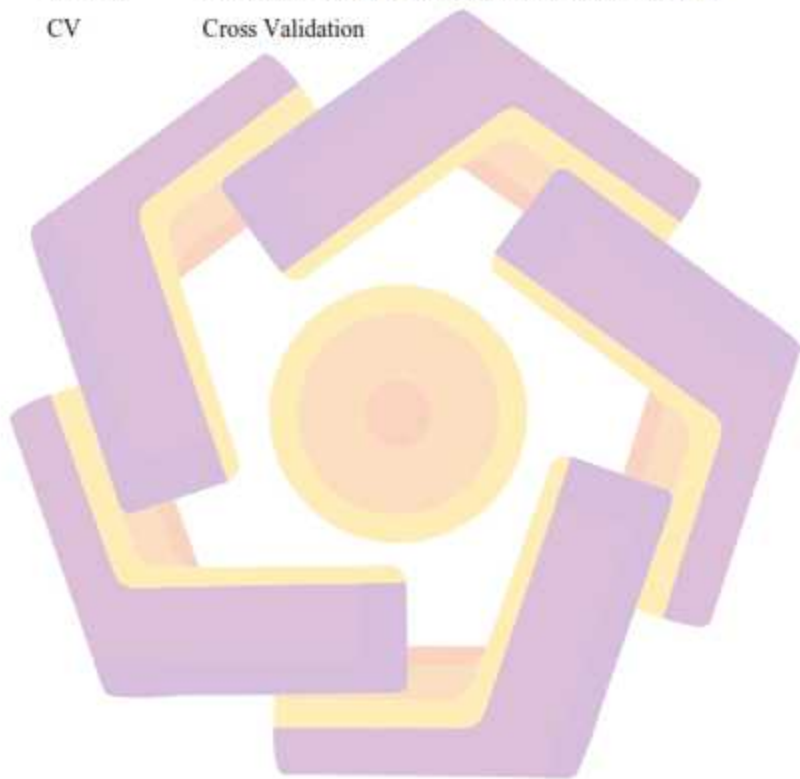
Tabel 2.1 Keaslian Penelitian.....	8
Tabel 3.1 Contoh Perubahan Nama Features.....	23
Tabel 3.2 Contoh Label Encoding pada <i>features</i>	25
Tabel 3.3 Dataset Kaggle.....	28
Tabel 4.1 <i>Output Time Spent</i>	33
Tabel 4.2 Pembagian Data.....	36
Tabel 4.3 <i>Selected Features of Random Forest CV=5</i>	36
Tabel 4.4 <i>Selected Features of Linear Regression CV=5</i>	37
Tabel 4.5 <i>Selected Features of Random Forest CV=15</i>	38
Tabel 4.6 <i>Selected Features of Random Forest CV=15</i>	38
Tabel 4.7 <i>Random Forest CV=5</i> rasio 60:40.....	39
Tabel 4.8 <i>Linear Regression CV=5</i> rasio 60:40.....	40
Tabel 4.9 <i>Random Forest CV=15</i> rasio 60:40.....	40
Tabel 4.10 <i>Selected Features of Linear Regression CV=15</i>	41
Tabel 4.11 Perbandingan <i>Selection Features</i>	42
Tabel 4.12 Evaluasi Random Forest.....	45
Tabel 4.13 Evaluasi Linear Regression.....	46

DAFTAR GAMBAR

Gambar 2.1 Flowchart RFECV	17
Gambar 3.1 Flowchart Penelitian	22
Gambar 3.2 Contoh visualisasi features	24
Gambar 4. 1 <i>Head</i> Dataset	29
Gambar 4.2 <i>Head</i> Dataset-2	29
Gambar 4.3 Daftar <i>Features</i>	30
Gambar 4.4 <i>Bar Plot of Sex</i>	31
Gambar 4.5 <i>Code Transform Gender</i>	32
Gambar 4.6 <i>Bar Plot of Age</i>	32
Gambar 4.7 <i>Code Age to Int</i>	33
Gambar 4.8 <i>Histogram of Time Spent</i>	33
Gambar 4. 9 Features dihapus	34
Gambar 4.10 <i>Sourcecode</i> hapus features	34
Gambar 4.11 Data setelah dilakukan encoding	35
Gambar 4.12 Pohon Keputusan Random Forest	43
Gambar 4.13 Hasil Linear $CV=5$	44
Gambar 4.14 <i>Comparison of RMSE and MAE(80:20)</i>	47
Gambar 4.15 <i>Comparison of RMSE and MAE(60:40)</i>	47
Gambar 4.16 Hasil prediksi rasio 80:20	48
Gambar 4.17 Hasil Prediksi Rasio 60:40	49

DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
RFECV	Recursive Feature Elimination with Cross-Validation
CV	Cross Validation



DAFTAR ISTILAH

Algoritma	Serangkaian langkah yang dirancang dalam menyelesaikan suatu masalah tertentu
Dataset	Data yang tersusun dan disimpan dalam bentuk format tertentu
Rasio	Besarnya perbandingan data
Library	Kumpulan alat dan fungsionalitas yang diakses oleh pengembang dalam merancang perangkat lunak
Data Training	Data yang digunakan untuk melatih model
Data Testing	Data yang digunakan untuk menguji model



INTISARI

Teknologi berkembang semakin pesat, salah satu dampak yang dirasakan dari teknologi yaitu penggunaan media sosial. Media sosial menjadi bagian yang tak terpisahkan dari kehidupan masyarakat. Dalam penggunaan media sosial oleh diberbagai kalangan, berbagai platform pun bermunculan seperti Facebook, Instagram, Twitter, Tiktok dan lain sebagainya. Media sosial sendiri membawa banyak dampak bagi berbagai kalangan, baik dampak negatif ataupun positif. Salah satu isu yang muncul mengenai media yaitu tingginya tingkat kecanduan media sosial. Kecanduan media sosial membuat penggunanya mengalami dampak buruk bagi kehidupan sosialnya. Fokus dalam penelitian ini melakukan perbandingan analisis terhadap tingkat kecanduan media sosial menggunakan algoritma Random Forest dan Linear Regression dalam melihat perilaku pengguna. Sebelum melakukan analisis fitur akan dipilih menggunakan RFECV agar model lebih maksimal dalam melakukan regresi. Hasil akhir yang didapatkan berupa besarnya kesalahan dengan melihat nilai RMSE dan MAE. Nilai RMSE dan MAE yang paling baik didapatkan pada Linear Regresion dengan RMSE 0.99 dan MAE 0.87. Penelitian ini memberikan kontribusi dalam melihat algoritma yang tepat dalam meramalkan perilaku pengguna mengenai tingkat kecanduan media sosial. Penelitian lebih lanjut agar dapat menggunakan data yang lebih banyak dan atau bersumber dari kusioner sendiri agar lebih maksimal dalam melihat tingkat kecanduan media sosial berdasarkan perilaku pengguna serta menggunakan model lainnya.

Kata kunci: Random Forest, Regresi Linear, *Supervised Learning*, RFECV, Media Sosial

ABSTRACT

Technology is developing increasingly rapidly, one of the impacts felt from technology is the use of social media. Social media has become an inseparable part of people's lives. In the use of social media by various people, various platforms have emerged such as Facebook, Instagram, Twitter, TikTok and more. Social media itself has many impacts on various people, both negative and positive impacts. One of the issues that arise is the level of social media addiction. Social media addiction makes its users experience adverse effects on their social life. The focus of this research is to compare the analysis of the level of social media addiction using Random Forest and Linear Regression algorithms in viewing user behaviours. Before analysing the features will be selected using RFECV, so that the model is maximise in performing regression. The final result obtained is the amount of error by looking at the RMSE and MAE values. The best RMSE and MAE values are obtained in Linear Regression with RMSE 0.99 and MAE 0.87. This research contributes to knowing the best algorithm for predicting user behaviour regarding the level of social media addiction. For more research, it can use more data and or sourced from the questionnaire itself in order to optimally observe the level of social media addiction based on user behaviour and use other models.

Keyword: *Random Forest, Linear Regression, Machine Learning, RFECV, Social Media*