

ANALISIS SENTIMEN PROFIL INSTAGRAM

NON SKRIPSI MAGANG

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



disusun oleh

ANDRE

19.11.3102

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

ANALISIS SENTIMEN PROFIL INSTAGRAM

NON SKRIPSI MAGANG

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ANDRE

19.11.3102

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

HALAMAN PERSETUJUAN

NON SKRIPSI

ANALISIS SENTIMEN PROFIL INSTAGRAM

yang disusun dan diajukan oleh

Andre

19.11.3102

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 17 Mei 2023

Dosen Pembimbing,



Arifiyanto Hadinegoro, S.Kom, MT
NIK. 1903022289

HALAMAN PENGESAHAN
NON SKRIPSI
ANALISIS SENTIMEN PROFIL INSTAGRAM

yang disusun dan diajukan oleh

Andre

19.11.3102

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Mei 2023

Susunan Dewan Penguji

Nama Penguji

Arif Dwi Laksito, M. Kom
NIK. 190302150

Anna Baita, M. Kom
NIK. 190302290

Bayu Setiaji, M. Kom
NIK. 190302216

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Mei 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
DAFTAR ISI.....	iv
DAFTAR TABEL.....	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN.....	x
RINGKASAN	xi
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Tujuan	2
C. Manfaat	3
D. Batasan	3
E. Identitas Tempat Magang	3
F. Sistematika Laporan.....	4
BAB II LANDASAN TEORI.....	6
A. Tinjauan Pustaka	6
B. Dasar Teori.....	12
1. Analisis Sentimen	12
2. Text Mining	12
3. Machine Learning	13
4. Pengumpulan Data	13
a. Wawancara.....	13
b. Dokumentasi	14
c. Observasi.....	14
5. Akuisisi Data.....	14
a. Scraping	15
b. Labeling	15
6. Text Preprocessing.....	15

a.	Cleaning	16
b.	Case Folding	16
c.	Tokenizing	16
d.	Normalization	16
e.	Stopword Removal.....	17
f.	Stemming	17
g.	TF-IDF	18
7.	Pengujian.....	18
a.	Holdout Validation.....	18
b.	White Box Testing	19
8.	Evaluasi.....	20
a.	Confusion Matrix	20
b.	Accuracy	21
c.	Runtime.....	21
9.	Modeling.....	22
a.	Supervised Learning	22
b.	Multi-Model.....	24
c.	N-gram	24
d.	Hyperparameter Tuning	25
e.	Resampling Data.....	25
10.	Algoritma Klasifikasi.....	26
11.	Scrum	27
12.	Media Sosial Instagram.....	28
13.	API.....	29
14.	Python Flask.....	30
BAB III METODOLOGI.....		31
A.	Alur Magang	31
B.	Analisis Kegiatan	33
1.	Proyek Magang Industri.....	34
2.	Pelatihan Softskill	34
3.	Pembelajaran Mandiri	35

4. Sertifikasi Nasional/Internasional	35
5. Portofolio	36
C. Alur dan Analisis Perancangan Produk.....	36
1. Pengumpulan Data	36
2. Pengembangan Produk.....	37
a. Pengembangan Model Produk	37
b. Pengembangan Sistem Produk.....	62
3. Pengujian Produk	70
a. Pengujian Model Produk.....	70
b. Pengujian Sistem Produk	70
BAB IV PEMBAHASAN.....	71
A. Pembahasan Hasil Kegiatan.....	71
B. Pembahasan Hasil Produk.....	72
1. Scrum	72
2. Scraper	74
3. Akuisisi Data.....	81
4. Preprocessing	83
5. Model	90
6. API.....	92
a. API Prediction.....	92
b. API Create.....	96
c. API Update.....	101
BAB V PENUTUP	106
A. Kesimpulan	106
B. Saran	107
REFERENSI	108
LAMPIRAN.....	113

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	8
Tabel 2.3 Confusion Matrix	20
Tabel 4.1 Dataset preprocessing	83
Tabel 4.2 Dataset hasil cleaning	84
Tabel 4.3 Dataset hasil case folding	84
Tabel 4.4 Dataset hasil tokenizing	85
Tabel 4.5 Dataset hasil normalization	86
Tabel 4. 6 Dataset hasil stopwords removal	87
Tabel 4.7 Dataset hasil stemming	87
Tabel 4.8 Dataset hasil perhitungan TF-IDF	88
Tabel 4.9 Perbandingan metode preprocessing	89
Tabel 4.10 Hasil random oversampling	90
Tabel 4.11 Perbandingan performa model	91

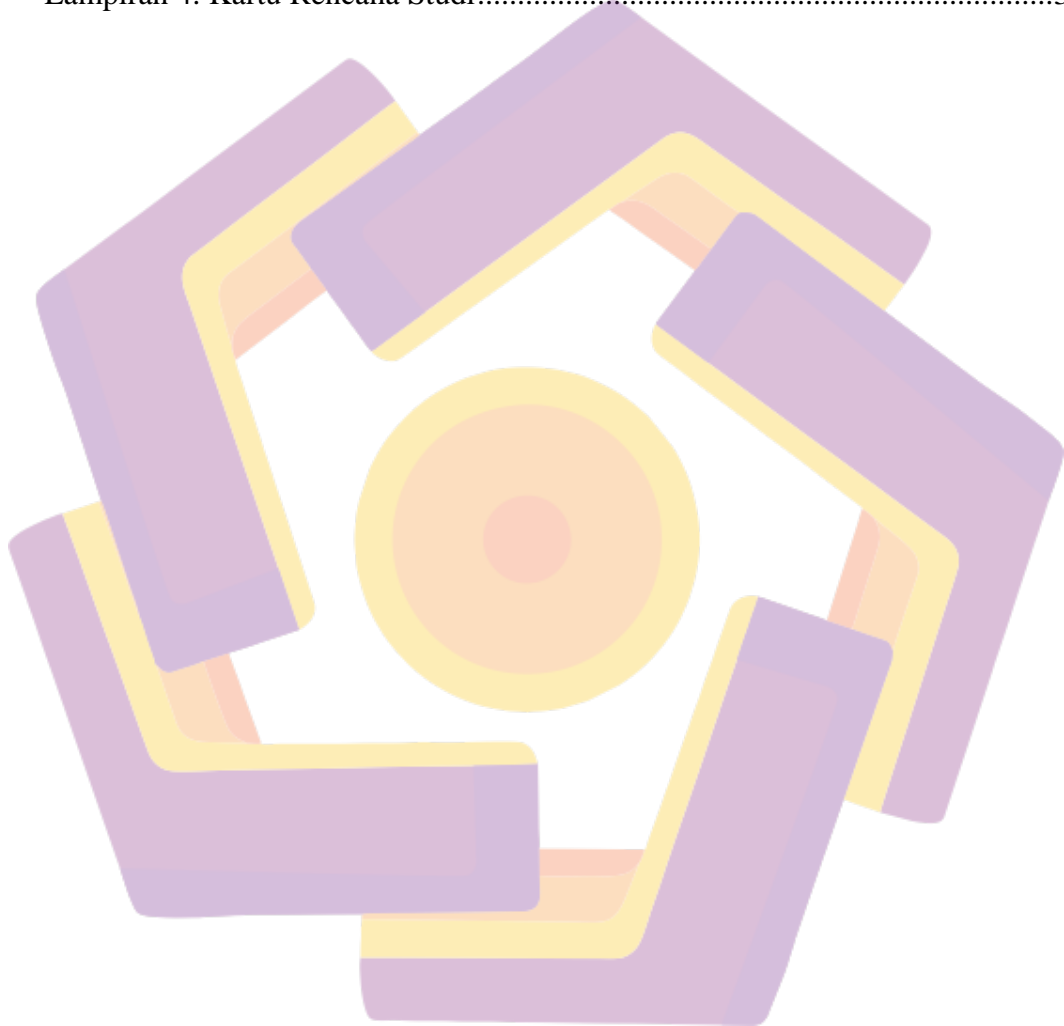
DAFTAR GAMBAR

Gambar 1.1 Struktur organisasi Central AI	4
Gambar 2.1 Delapan negara pengguna Instagram terbanyak [31].....	28
Gambar 3.1 Alur magang MBKM	31
Gambar 3.2 Alur pengembangan model produk.....	38
Gambar 3.3 Alur akuisisi data	39
Gambar 3.4 Alur scaping	41
Gambar 3.5 Alur labeling	42
Gambar 3.6 Alur Preprocessing	44
Gambar 3.7 Alur cleaning.....	46
Gambar 3.8 Alur case folding.....	47
Gambar 3.9 Alur tokenizing	49
Gambar 3.10 Alur normalization	50
Gambar 3.11 Alur stopword removal	52
Gambar 3.12 Alur stemming.....	54
Gambar 3.13 Alur perhitungan TF-IDF.....	56
Gambar 3.14 Alur modeling	57
Gambar 3.15 Alur multi-model.....	59
Gambar 3.16 Alur resampling data RandomOverSampler	61
Gambar 3.17 Alur dan pembagian pengembangan sistem produk	63
Gambar 3.18 Alur API prediction.....	65
Gambar 3.19 Alur API create dan update	68
Gambar 4.1 Sertifikat DSBIZ	72
Gambar 4.2 Hasil scraping (data) profile.....	75
Gambar 4.3 Hasil scraping (media) stories.....	76
Gambar 4.4 Hasil scraping (media) highlight.....	76
Gambar 4.5 Hasil scraping (data) post.....	77
Gambar 4.6 Hasil scraping (media) post.....	77
Gambar 4.7 Hasil scraping (data) tagged.....	78
Gambar 4.8 Hasil scraping (media) tagged.....	78

Gambar 4.9 Hasil scraping (data) IGTV	79
Gambar 4.10 Hasil scraping (media) IGTV	79
Gambar 4.11 Gambar fitur scraper	81
Gambar 4.12 Perbandingan tiga dan dua label pada model Indo-Bert	82
Gambar 4.13 Hasil pengujian pertama API prediction	93
Gambar 4.14 Hasil pengujian kedua API prediction	93
Gambar 4.15 Hasil pengujian ketiga API prediction	94
Gambar 4.16 Hasil pengujian keempat API prediction	95
Gambar 4.17 Hasil pengujian kelima API prediction	96
Gambar 4.18 Hasil pengujian pertama API create	97
Gambar 4.19 Hasil pengujian kedua API create	97
Gambar 4.20 Hasil pengujian tiga API create	98
Gambar 4.21 Hasil pengujian keempat API create	99
Gambar 4.22 Hasil pengujian kelima API create	100
Gambar 4.23 Hasil pengujian pertama API create	101
Gambar 4.24 Hasil pengujian pertama API update	102
Gambar 4.25 Hasil pengujian kedua API update	102
Gambar 4.26 Hasil pengujian ketiga API update	103
Gambar 4.27 Hasil pengujian keempat API update	104
Gambar 4.28 Hasil pengujian kelima API update	104
Gambar 4.29 Hasil pengujian keenam API update	105

DAFTAR LAMPIRAN

Lampiran 1. Logbook Selama Magang.....	111
Lampiran 2. Sertifikat MSIB	161
Lampiran 3. Laporan Magang.....	164
Lampiran 4. Kartu Rencana Studi.....	314



RINGKASAN

Media sosial mempengaruhi perubahan cara berkomunikasi pada masyarakat. Salah satu media sosial yang sering digunakan yaitu Instagram. Instagram menjadi media untuk bersosialisasi dan mengekspresikan diri. Seiring perkembangan teknologi, fungsinya bergeser menjadi sarana untuk melakukan kegiatan bisnis. Dalam proses bisnis opini dari pelanggan merupakan hal yang penting untuk kemajuan dan perkembangan perusahaan. Oleh karena itu, untuk mengetahui apakah akun instagram yang digunakan digemari atau tidak oleh publik dapat dilakukan analisis sentimen dari komentar yang ada pada akun. Analisis sentimen merupakan proses untuk memahami isi teks dan dikelompokkan kedalam beberapa jenis kategori sentimen. Untuk mempermudah dan mempersingkat waktu pengumpulan data dibuat scraper yang secara otomatis akan mengambil data berdasarkan username akun instagram. Untuk melakukan analisis sentimen dengan cepat dilakukan menggunakan model machine learning menggunakan metode klasifikasi. Agar model dapat mengetahui makna setiap data teks dilakukan proses ekstraksi fitur TF-IDF. Sebelum data diolah oleh model data dibersihkan menggunakan metode preprocessing cleaning, case folding, tokenizing, dan normalization. Data berisi 4066 data dengan pembagian label sentimen positif dan negatif, serta pembagian 80% untuk data latih dan 20% untuk data uji. Model yang digunakan memiliki akurasi 85.26% dan runtime 0.8 detik yang dihasilkan dari multi-model terhadap tiga model machine learning yaitu Support Vector Machine, Naive Bayes, dan Random Forest dengan menerapkan *hyperparameter tuning*. Dikarenakan analisis sentimen pada dasarnya merupakan hal subjektif maka dibuat sistem yang memfasilitasi pengguna untuk membuat model sesuai dengan preferensi mereka.

Kata kunci: analisis sentimen, preprocessing, Instagram, model machine learning.