

**KLASIFIKASI CYBERBULLYING PADA KOMENTAR INSTAGRAM
AKUN PLESBOL.INC MENGGUNAKAN NAÏVE BAYES
SKRIPSI**

**Diajukan untuk memenuhi sebagian persyaratan dalam memperoleh gelar
Sarjana Komputer Program Studi Sistem Informasi**



Disusun Oleh :

AKIRA PERMATA RAMADHANI

19082010088

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

SURABAYA

2024

SKRIPSI
KLASIFIKASI CYBERBULLYING PADA KOMENTAR INSTAGRAM
AKUN PLESBOL.INC MENGGUNAKAN NAÏVE BAYES

Disusun Oleh:
AKIRA PERMATA RAMADHANI
19082010088

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi
Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Pada Tanggal, 28 Mei 2024

Pembimbing:

1.



Eka Dyar Wahyuni, S.Kom., M.Kom
NIP. 19841201 2021212 005


2.



Amalia Anjani Arifiyanti, S.Kom., M.Kom
NIP. 19920812 2018032 001

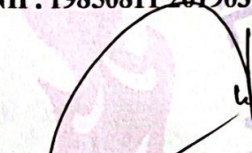
Tim Penguji:

1.



Dr. Eng. Agussalim, MT
NIP. 19850811 2019031 005

2.



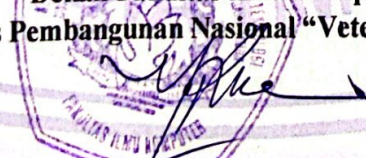
Dhiah Satria Yudha Kartika, S.Kom., M.Kom
NPT. 201198 60 522249

3.



Abdul Rezha Efrat Najaf, S.Kom., M.Kom
NIP. 19940929 202203 1008

Mengetahui,
Dekan Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur


Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PENGESAHAN
KLASIFIKASI CYBERBULLYING PADA KOMENTAR INSTAGRAM
AKUN PLESBOL.INC MENGGUNAKAN NAÏVE BAYES

Disusun Oleh:
AKIRA PERMATA RAMADHANI
19082010088

Telah disetujui mengikuti Ujian Negara Lisan Gelombang Mei
Periode Mei pada Tanggal, 28 Mei 2024

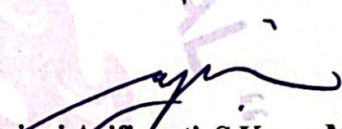
Menyetujui,

Dosen Pembimbing 1



Eka Dvar Wahyuni, S.Kom., M.Kom
NIP. 19841201 2021212 005

Dosen Pembimbing 2



Amalia Anjani Arifiyanti, S.Kom., M.Kom
NIP. 19920812 2018032 001

Mengetahui,

Ketua Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur



Agung Brastama Putra, S.Kom, M.Kom.
NIP. 19851124 20211211 003

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR
FAKULTAS ILMU KOMPUTER

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Akira Permata Ramadhani

NPM : 19082010088

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 28 Mei 2024 dengan judul:

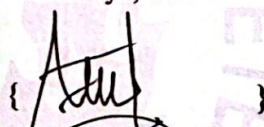
KLASIFIKASI CYBERBULLYING PADA KOMENTAR INSTAGRAM AKUN PLESBOL.INC MENGGUNAKAN NAÏVE BAYES

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.


Surabaya, 10 Juni 2024

Dosen penguji yang memeriksa revisi:

1. Dr. Eng. Agussalim, MT
NIP. 19850811 2019031 005
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom
NPT. 201198 60 522249
3. Abdul Rezha Efrat Najaf, S.Kom., M.Kom
NIP. 19940929 202203 1008

{  }

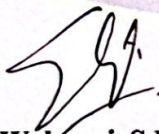
{  }

{  }

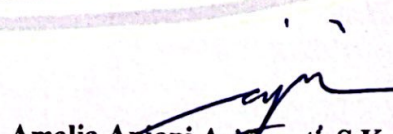
Mengetahui,

Dosen Pembimbing 1

Dosen Pembimbing 2



Eka Dyar Wahyuni, S.Kom., M.Kom
NIP. 19841201 2021212 005



Amalia Anjani Ardiyanti, S.Kom., M.Kom
NIP. 19920812 2018032 001



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL
"VETERAN" JAWA TIMUR
FAKULTAS ILMU KOMPUTER

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Akira Permata Ramadhani
NPM : 19082010088
Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi/Tugas Akhir sebagai berikut :

**KLASIFIKASI CYBERBULLYING PADA KOMENTAR INSTAGRAM
AKUN PLESBOL.INC MENGGUNAKAN NAÏVE BAYES**

Bukan merupakan plagiat dari Skripsi/Tugas Akhir/Penelitian orang lain dan juga bukan Produk/Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Instansi Pendidikan lain. Jika dinyatakan dikemudian hari pernyataan terbukti benar, maka saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 19 Juni 2024

Hormat Saya,



Akira Permata Ramadhani
NPM. 19082010088

Judul : KLASIFIKASI CYBERBULLYING PADA KOMENTAR
INSTAGRAM AKUN PLESBOL.INC DENGAN MENGGUNAKAN NAÏVE
BAYES

Pembimbing 1 : Eka Dyar Wahyuni, S.Kom., M.Kom.

Pembimbing 2 : Amalia Anjani Arifiyanti S.Kom., M.Kom.

ABSTRAK

Cyberbullying merupakan tindakan yang berhubungan pada penggunaan teknologi digital untuk sengaja menyakiti, merendahkan, ataupun merundung orang lain secara *online*. Penelitian ini berfokus pada klasifikasi komentar *cyberbullying* di media sosial, khususnya komentar Instagram, dimana banyak pihak yang kemudian menjadi sekumpulan orang yang tidak menyukai sesuatu akan bersatu untuk memberikan pendapat dan komentar negatif, yang bisa menyebabkan menurunkan kepercayaan diri dan dampak buruk lainnya bagi pengguna lain maupun pemilik akun. Berdasarkan pada masalah tersebut, dilakukan klasifikasi komentar Instagram mengenai *cyberbullying* sebagai upaya pencegahan tindakan tersebut. Data yang digunakan pada penelitian ini sebanyak 2000 data yang diambil dari media sosial Instagram, dimana data ini akan melalui berbagai macam proses agar dapat dieksekusi. Proses yang dilalui diantaranya, mulai dari pelabelan data, *text preprocessing*, visualisasi *wordcloud*, pembagian data, TF-IDF, klasifikasi *naïve bayes*, dan evaluasi model. Pada penelitian ini, digunakan metode *Naïve Bayes* yang menggunakan semua jenis *naïve bayes* yaitu, *Multinomial*, *Bernoulli*, dan *Gaussian*. Penelitian ini menggunakan dua kelas yaitu kelas *Bully* yang sebanyak 898 data dan kelas Tidak *Bully* sebanyak 1102 data. Pembagian data menggunakan dua cara yakni, *cross-validation* (90:10) dan *holdout* (70:30). Penelitian ini menggunakan 12 skenario dan berdasarkan hasil pengujian yang sudah dilakukan, skenario yang terpilih adalah skenario *holdout* (70:30) dengan nilai akurasi 84%, nilai *precision* 84%, *recall* 84%, dan *f1-score* 84%.

Kata kunci :

Cyberbullying, Instagram, Klasifikasi, *Naïve Bayes*

KATA PENGANTAR

Segala Puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan segala rahmat-Nya serta sholawat dan salam tercurahkan kepada junjungan Nabi Muhammad SAW, sehingga penulis dapat menyelesaikan skripsi ini dengan judul “KLASIFIKASI CYBERBULLYING PADA KOMENTAR AKUN PLESBOL.INC INSTAGRAM MENGGUNAKAN NAÏVE BAYES” sebagai salah satu syarat dalam menyelesaikan Program Studi S1/Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Terselesainya skripsi ini tentunya tak lepas dari dorongan dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan rasa terima kasih yang tak terhingga kepada semua pihak yang membantu kelancaran penulisan tugas akhir ini. Untuk itu pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Sutriyo dan Ibu Henny Srie Suhesti beserta keluarga saya yang telah memberikan dukungan materi dan doa agar skripsi ini bisa selesai serta mendoakan saya agar selalu menjadi orang yang lebih baik dan bermanfaat.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom. selaku dosen pembimbing 1 yang selalu memberikan waktu dan arahan dalam menyelesaikan skripsi.
3. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom. selaku dosen pembimbing 2 yang selalu memberikan waktu dan arahan dalam menyelesaikan skripsi.
4. Koordinator Program Studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur Bapak Agung Brastama Putra, S.Kom., M.Kom.

5. Seluruh dosen Program Studi Sistem Informasi atas ilmu-ilmu yang diberikan selama masa perkuliahan.
6. Teman-teman Sistem Informasi angkatan 2019 yang selalu memberikan dukungan dalam mengerjakan tugas sampai skripsi ini. Dan juga kepada teman-teman Tim Prasmanan yang selalu membantu saya dalam pengerjaan skripsi ini.
7. Teman-teman nongkrong saya yang tidak dapat saya sebutkan satu-persatu yang telah membantu saya dalam pengerjaan skripsi ini.
8. Semua pihak yang tidak bisa saya sebutkan satu-persatu. Terima kasih telah memberikan dukungan dan bantuan dari hal kecil hingga hal yang tidak bisa saya lakukan sendiri.
9. Kepada diri saya pribadi yang telah berjuang sampai pada tahap ini. “Tetap sehat dan tetap semangat”.

Semoga Allah membalas dengan balasan sebaik-baiknya. Dalam pengerjaan skripsi ini saya menyadari bahwa skripsi ini masih ada kekurangan. Semoga skripsi ini dapat bermanfaat bagi semua pihak khususnya pada perkembangan ilmu pengetahuan di program studi Sistem Informasi.

DAFTAR ISI

ABSTRAK	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1 Penelitian Terdahulu	8
2.2 Dasar Teori	11
2.2.1 Cyberbullying	12
2.2.2 Text Mining	15
2.2.3 Text Preprocessing.....	16
2.2.4 VADER.....	17
2.2.5 SMOTE.....	18
2.2.6 Cross-validaton	18
2.2.7 Holdout	19
2.2.8 Naïve Bayes	19

2.2.9 TF-IDF	20
2.2.10 Confusion Matrix	21
2.2.11 Flask.....	22
BAB III METODOLOGI PENELITIAN	23
3.1 Studi Literatur	23
3.2 Analisis Kebutuhan.....	24
3.3 Pembangunan Model	24
3.4 Visualisasi	28
BAB IV HASIL DAN PEMBAHASAN	29
4.1 Implementasi Kebutuhan	29
4.2 Pembangunan Model	29
4.2.1 Pengumpulan Data	29
4.2.2 Penyaringan Data	31
4.2.3 Pelabelan Data	32
4.2.4 Text Preprocessing.....	35
4.2.5 Wordcloud	51
4.2.6 Pembagian Data	53
4.2.7 Pembobotan TF-IDF	57
4.2.8 Klasifikasi Naïve Bayes	58
4.2.9 Evaluasi Model	65
4.3 Visualisasi	78
BAB V PENUTUP.....	81
5.1 Kesimpulan	81
5.2 Saran	82
DAFTAR PUSTAKA	83
LAMPIRAN.....	83

DAFTAR GAMBAR

Gambar 1. 1 Diagram persentase sosial media yang paling banyak terjadi cyberbullying.....	2
Gambar 3.1 Alur penelitian.....	23
Gambar 3.2 Perancangan model	25
Gambar 4. 1 Tampilan Halaman Web Scraper Apify.....	30
Gambar 4. 2 Tampilan Contoh Hasil Pengumpulan Data.....	30
Gambar 4. 3 Contoh komentar Instagram yang didapat	31
Gambar 4. 4 Source code menampilkan kolom	31
Gambar 4. 5 Source code penghapusan kolom	31
Gambar 4. 6 Source code labelling data secara otomatis menggunakan VADER	33
Gambar 4. 7 Pie chart pelabelan data.....	34
Gambar 4. 8 Source code emot handling	35
Gambar 4. 9 Source code mengubah ke lowercase.....	37
Gambar 4. 10 Source code dari clean URL.....	38
Gambar 4. 11 Source code dari clean symbol.....	39
Gambar 4. 12 Source code melakukan clean number	41
Gambar 4. 13 Source code clean huruf berulang	42
Gambar 4. 14 Source code melakukan clean spasi lebih	44
Gambar 4. 15 Source code melakukan clean row	45
Gambar 4. 16 Source code melakukan normalisasi	46
Gambar 4. 17 Source code menginstall library tokenizing, stopword removal, dan stemming.....	47
Gambar 4. 18 Source code melakukan tokenizing dan stopword removal	49
Gambar 4. 19 Source code melakukan stemming.....	50
Gambar 4. 20 Wordcloud sebelum text preprocessing	51

Gambar 4. 21 Wordcloud sesudah text preprocessing	52
Gambar 4. 22 Wordcloud dari kelas Bully.....	52
Gambar 4. 23 Wordcloud dari kelas Tidak Bully	53
Gambar 4. 24 Source code library SMOTE dan vektor TF-IDF.....	54
Gambar 4. 25 Jumlah data sebelum dilakukan SMOTE	54
Gambar 4. 26 Source code mengubah teks menjadi vektor lalu menerapkan teknik SMOTE	55
Gambar 4. 27 Jumlah data setelah dilakukan SMOTE	55
Gambar 4. 28 Source code pembagian data Cross validation.....	56
Gambar 4. 29 Source code pembagian data Hold-out (70:30).....	57
Gambar 4. 30 Pembobotan TF-IDF untuk Cross validation	57
Gambar 4. 31 Source code TF-IDF untuk hold-out	57
Gambar 4. 32 Source code dan hasil Cross validation data normal dengan multinomial naive bayes	59
Gambar 4. 33 Source code dan hasil cross validation data normal dengan bernoulli naive bayes.....	59
Gambar 4. 34 Source code dan hasil Cross validation data normal dengan gaussian naive bayes	60
Gambar 4. 35 Source code dan hasil Cross validation data SMOTE dengan multinomial naive bayes	60
Gambar 4. 36 Source code dan hasil Cross validation data SMOTE dengan bernoulli naive bayes.....	61
Gambar 4. 37 Source code dan hasil Cross validation data SMOTE dengan gaussian naive bayes	61
Gambar 4. 38 Source code dan hasil klasifikasi hold-out (70:30) data normal dengan multinomial naive bayes.....	62
Gambar 4. 39 Source code dan hasil hold-out (70:30) data normal menggunakan bernoulli naive bayes.....	63
Gambar 4. 40 Source code dan hasil hold-out (70:30) data normal menggunakan gaussian naive bayes	63
Gambar 4. 41 Source code dan hasil hold-out (70:30) data SMOTE menggunakan	

multinomial naive bayes	64
Gambar 4. 42 Source code dan hasil hold-out (70:30) data SMOTE menggunakan bernoulli naive bayes.....	64
Gambar 4. 43 Source code dan hasil hold-out (70:30) data SMOTE menggunakan gaussian naive bayes	65
Gambar 4. 44 Source code dan hasil confusion matrix dan f1-score cross validation data normal menggunakan multinomial naive bayes	66
Gambar 4. 45 source code dan hasil confusion matrix dan f1-score cross validation data normal menggunakan bernoulli naive bayes	66
Gambar 4. 46 source code dan hasil confusion matrix dan f1-score cross validation data normal menggunakan gaussian naive bayes.....	67
Gambar 4. 47 source code dan hasil confusion matrix dan f1-score cross validation data SMOTE menggunakan multinomial naive bayes.....	68
Gambar 4. 48 source code dan hasil confusion matrix dan f1-score cross validation data SMOTE menggunakan bernoulli naive bayes.....	69
Gambar 4. 49 source code dan hasil confusion matrix dan f1-score cross validation data SMOTE menggunakan gaussian naive bayes.....	69
Gambar 4. 50 source code dan hasil confusion matrix holdout (70:30) data normal menggunakan multinomial naive bayes	70
Gambar 4. 51 source code dan hasil confusion matrix holdout (70:30) data normal menggunakan bernoulli naive bayes	71
Gambar 4. 52 source code dan hasil confusion matrix holdout (70:30) data normal menggunakan gaussian naive bayes.....	72
Gambar 4. 53 source code dan hasil confusion matrix holdout (70:30) data SMOTE menggunakan multinomial naive bayes	73
Gambar 4. 54 source code dan hasil confusion matrix holdout (70:30) data SMOTE menggunakan bernoulli naive bayes.....	74
Gambar 4. 55 source code dan hasil confusion matrix holdout (70:30) data SMOTE menggunakan gaussian naive bayes	75
Gambar 4. 56 confusion matrix data SMOTE holdout bernoulli naive bayes (I).	75
Gambar 4. 57 confusion matrix data SMOTE holdout bernoulli naive bayes (II)	76
Gambar 4. 58 Contoh cek overfitting model holdout(70:30) SMOTE bernoulli..	76
Gambar 4. 59 Source code Python	78

Gambar 4. 60 Source code HTML.....	79
Gambar 4. 61 Tampilan web.....	79
Gambar 4. 62 Tampilan web dan penggunaannya.....	80

DAFTAR TABEL

Tabel 4. 1 Tabel contoh hasil menghapus kolom.....	32
Tabel 4. 2 Contoh tabel hasil pelabelan VADER dan manual.....	33
Tabel 4. 3 Tabel contoh hasil emot handling	36
Tabel 4. 4 Tabel contoh hasil case folding.....	37
Tabel 4. 5 Tabel contoh hasil URL cleaning	38
Tabel 4. 6 Tabel contoh hasil clean symbol.....	40
Tabel 4. 7 Tabel contoh hasil clean number	41
Tabel 4. 8 Tabel contoh hasil pembersihan huruf berulang.....	43
Tabel 4. 9 Tabel contoh hasil pembersihan spasi berlebih	44
Tabel 4. 10 Tabel contoh hasil normalisasi.....	46
Tabel 4. 11 Tabel contoh hasil tokenizing dan stopword removal	49
Tabel 4. 12 Tabel contoh hasil stemming	51
Tabel 4. 13 Tabel hasil evaluasi model.....	77