

SKRIPSI

**ANALISIS SENTIMEN KEBIJAKAN PENGHAPUSAN
KEWAJIBAN SKRIPSI PERGURUAN TINGGI PENGGUNA
INSTAGRAM MENGGUNAKAN METODE *LOGISTIC
REGRESSION***



*Mencerdaskan dan
Memartabatkan Bangsa*

FATHIA DWI ASTUTI

1512618073

PROGRAM STUDI

PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER

FAKULTAS TEKNIK



UNIVERSITAS NEGERI JAKARTA

2024

HALAMAN PENGESAHAN SKRIPSI

ANALISIS SENTIMEN KEBIJAKAN PENGHAPUSAN KEWAJIBAN SKRIPSI PERGURUAN TINGGI PENGGUNA *INSTAGRAM* MENGUNAKAN METODE *LOGISTIC REGRESSION*

Fathia Dwi Astuti, NIM. 1512618073

NAMA DOSEN	TANDA TANGAN	TANGGAL
Dr. Widodo, S.Kom, M.Kom Dosen Pembimbing 1		07 Juli 2024
Bambang Prasetya Adhi, S.Pd., M.Kom Dosen Pembimbing 2		07 Juli 2024

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
Muchammad Ficky Duskarnaen, M.Sc Ketua Penguji		15 Juli 2024
Diat Nurhidayat, M.T.I Dosen Penguji 1		10 Juli 2024
Ali Idrus, S.Kom.,M.Kom. Dosen Penguji 2		12 Juli 2024

LEMBAR PERNYATAAN

Dengan ini menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 03 Juli 2024

Yang membuat pernyataan,



Fathia Dwi Astuti
1512618073



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Fathia Dwi Astuti
NIM : 1512618073
Fakultas/Prodi : Fakultas Teknik/Pendidikan Teknik Informatika dan Komputer
Alamat email : fathiaastuti@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (.....)

yang berjudul :

ANALISIS SENTIMEN KEBIJAKAN PENGHAPUSAN KEWAJIBAN SKRIPSI
PERGURUAN TINGGI PENGGUNA *INSTAGRAM* MENGGUNAKAN METODE
LOGISTIC REGRESSION

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Juli 2024

Penulis

(Fathia Dwi Astuti)

KATA PENGANTAR

Segala puji dan syukur kehadiran Allah *Subhanahu Wa Ta'ala* atas segala nikmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Sentimen Kebijakan Penghapusan Kewajiban Skripsi Perguruan Tinggi Pengguna *Instagram* Menggunakan Metode *Logistic Regression*” yang merupakan persyaratan untuk meraih gelar Sarjana Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta. Shalawat serta salam selalu tercurah kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam*, beserta keluarga dan para sahabatnya yang telah membimbing manusia dalam kebaikan.

Penulis menyadari sepenuhnya bahwa skripsi ini selesai berkat bantuan, petunjuk, saran, bimbingan dan dorongan dari berbagai pihak. Untuk itu pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Allah *Subhanahu Wa Ta'ala* atas nikmat, rahmat dan karunia-Nya kepada penulis,
2. Orang tua, kakak, dan keluarga penulis yang senantiasa memberikan doa, dukungan dan motivasi kepada penulis,
3. Bapak Muchammad Ficky Duskarnaen, M.Sc selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer yang telah memberikan bantuan selama perkuliahan,
4. Bapak Dr. Widodo, S.Kom, M.Kom, selaku dosen pembimbing pertama yang telah membagikan ilmunya dan memberikan saran dalam penulisan skripsi,
5. Bapak Bambang Prasetya Adhi, S.Pd., M.Kom, selaku dosen pembimbing kedua yang juga telah memberikan koreksi serta arahan dalam penulisan skripsi,
6. Bapak Diat Nurhidayat, M.T.I., selaku dosen pembimbing Praktik Keterampilan Mengajar, dan pembimbing akademik penulis yang telah memberikan rekomendasi dan motivasi selama perkuliahan,
7. Seluruh Bapak dan Ibu dosen Pendidikan Teknik Informatika dan Komputer yang telah mengajar dengan sabar dan membagikan ilmunya kepada penulis,

8. Seluruh mahasiswa/i Pendidikan Teknik Informatika dan Komputer Angkatan 2018,
9. Seluruh pihak yang terlibat dan/atau berkontribusi dalam hidup penulis sehingga penulis dapat menyelesaikan skripsi dan studi di Universitas Negeri Jakarta.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, maka dari itu penulis memohon maaf apabila terdapat kekurangan dan kesalahan, baik dari isi maupun tulisan karena keterbatasan pengetahuan yang dimiliki. Akhir kata, penulis berharap agar penulisan dan penyusunan skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak yang terkait.

Jakarta, 03 Juli 2024

Penulis,



Fathia Dwi Astuti



ABSTRAK

Fathia Dwi Astuti. Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta, Analisis Sentimen Kebijakan Penghapusan Kewajiban Skripsi Perguruan Tinggi Pengguna *Instagram* Menggunakan Metode *Logistic Regression*. 2024. Dosen Pembimbing: Dr. Widodo, S.Kom., M. Kom., Bambang Prasetya Adhi, S.Pd., M.Kom.

Sistem pendidikan Indonesia menjadikan skripsi sebagai salah satu syarat kelulusan perguruan tinggi untuk memperoleh gelar. Namun pada tanggal 29 Agustus 2023, Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi menyampaikan kebijakan baru mengenai skripsi yang tidak lagi diwajibkan dan dapat diganti dalam bentuk lain, kebijakan ini diusut dalam Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 53 Tahun 2023 tentang Penjaminan Mutu Pendidikan Tinggi. Penelitian ini bertujuan untuk melakukan analisis sentimen kebijakan penghapusan kewajiban skripsi perguruan tinggi pada komentar *Instagram* dan mengevaluasi kinerja klasifikasinya. Analisis sentimen dilakukan dengan dua skenario menggunakan metode *Logistic Regression* untuk klasifikasi, pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)*, validasi data *10-fold Cross Validation*, *oversampling* data *Adaptive Synthetic (ADASYN)* dan evaluasi kinerja *Confusion Matrix*. Skenario pertama menghasilkan *accuracy* 87,5%, *precision* 88,2%, *recall* 90,5%, dan *f-1 score* 89,4% dan skenario kedua menghasilkan *accuracy* 89,6%, *precision* 91,5%, *recall* 87,7%, dan *f-1 score* 89,5%. Dari evaluasi kedua skenario dapat disimpulkan bahwa model klasifikasi dengan metode *Logistic Regression* memiliki kinerja yang baik pada tugas analisis sentimen kebijakan penghapusan kewajiban skripsi, terlebih lagi jika dikombinasikan dengan *ADASYN*.

Kata kunci: ***Analisis Sentimen. Logistic Regression. Adaptive Synthetic (ADASYN). Kebijakan Skripsi. Instagram.***

ABSTRACT

Fathia Dwi Astuti. *Informatics and Computer Engineering Education, Faculty of Engineering, Jakarta State University, Sentiment Analysis of University Thesis Requirements Policy on Instagram Using the Logistic Regression. 2024. Advisors: Dr. Widodo, S. Kom., M. Kom., Bambang Prasetya Adhi, S.Pd., M.Kom.*

The Indonesian education system requires a thesis as one of the requirements for university graduation to obtain a degree. However, on August 29 2023, the Minister of Education, Culture, Research and Technology announced a new policy regarding theses which are no longer required and can be replaced in another form, this policy is regulated in Minister of Education, Culture, Research and Technology Regulation Number 53 of 2023 concerning Quality Assurance of Higher Education. This study aims to analyze the sentiment of the policy of eliminating university thesis obligations on Instagram comments and evaluate its classification performance. Sentiment analysis was carried out with two scenarios using the Logistic Regression method for classification, Term Frequency-Inverse Document Frequency (TF-IDF) for weighting, 10-fold Cross Validation for data validation, Adaptive Synthetic (ADASYN) for oversampling data and Confusion Matrix for performance evaluation. The first scenario produces an accuracy of 87.5%, precision 88.2%, recall 90.5%, and f-1 score 89.4% and the second scenario produces an accuracy of 89.6%, precision 91.5%, recall 87.7 %, and f-1 score 89.5%. From the evaluation of the two scenarios, it can be concluded that the classification model using the Logistic Regression method has good performance in the task of sentiment analysis for the policy of eliminating thesis obligations, especially when combined with ADASYN.

Keywords: *Sentiment Analysis. Logistic Regression. Adaptive Synthetic (ADASYN). Thesis Policy. Instagram.*

DAFTAR ISI

HALAMAN PENGESAHAN SKRIPSI.....	2
LEMBAR PERNYATAAN	3
KATA PENGANTAR.....	5
ABSTRAK	7
ABSTRACT	8
DAFTAR TABEL	11
DAFTAR GAMBAR.....	12
BAB I PENDAHULUAN.....	Error! Bookmark not defined.
1.1. Latar Belakang.....	13
1.2. Identifikasi Masalah.....	5
1.3. Batasan Masalah	5
1.4. Rumusan Masalah.....	6
1.5. Tujuan Penelitian	6
1.6. Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA.....	8
2.1. <i>Text Mining</i>	8
2.2. Analisis Sentimen	8
2.3. Skripsi	9
2.4. <i>Instagram</i>	10
2.5. <i>Logistic Regression</i>	10
2.6. <i>Support Vector Machine</i>	12
2.7. <i>Preprocessing</i>	13
2.7.1. <i>Case Folding</i>	14
2.7.2. <i>Cleansing</i>	14
2.7.3. <i>Tokenizing</i>	15
2.7.4. <i>Normalization</i>	15
2.7.5. <i>Stopword Removal</i>	15
2.7.6. <i>Stemming</i>	16
2.8. <i>Term Frequency – Inverse Document Frequency (TF – IDF)</i>	16
2.9. <i>Adaptive Synthetic (ADASYN)</i>	17
2.10. <i>K-fold Cross Validation</i>	17
2.11. <i>Confusion Matrix</i>	19

2.12. Penelitian Relevan	20
2.13. Kerangka Berpikir.....	25
BAB III METODOLOGI PENELITIAN	27
3.1. Tempat dan Waktu Penelitian.....	27
3.2. Alat dan Bahan Penelitian.....	27
3.2.1. Alat Penelitian	27
3.2.2. Bahan Penelitian.....	28
3.3. Diagram Alir Penelitian	29
3.4. Teknik dan Prosedur Pengumpulan Data.....	29
3.5. Pengolahan Data Penelitian	31
3.5.1. <i>Case Folding</i>	32
3.5.2. <i>Cleansing</i>	33
3.5.3. <i>Tokenizing</i>	34
3.5.4. <i>Normalization</i>	35
3.5.5. <i>Stopword Removal</i>	36
3.5.6. <i>Stemming</i>	37
3.6. Pembobotan <i>TF-IDF</i>	39
3.7. Pengklasifikasian <i>Logistic Regression</i>	40
3.8. Teknik Analisis Data.....	41
3.8.1. <i>10-fold Cross Validation</i>	41
3.8.2. <i>Confusion Matrix</i>	42
BAB IV HASIL PENELITIAN.....	46
4.1. Pengumpulan Data Penelitian.....	46
4.2. Pengolahan Data Penelitian	47
4.3. Hasil Penelitian	49
4.4. Analisis Hasil Penelitian.....	56
4.5. Pembahasan.....	57
4.6. Aplikasi Hasil Penelitian.....	59
BAB V KESIMPULAN DAN SARAN	60
5.1. Kesimpulan	60
5.2. Saran	60
DAFTAR PUSTAKA	61
LAMPIRAN.....	67

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	19
Tabel 2. 2 Penelitian Relevan.....	22
Tabel 3. 1 Jadwal Pelaksanaan Penelitian.....	27
Tabel 3. 2 Daftar Perangkat Keras	28
Tabel 3. 3 Daftar Perangkat Lunak	28
Tabel 3. 4 Contoh Dataset.....	31
Tabel 3. 5 Contoh Proses Pelabelan Data	31
Tabel 3. 6 Contoh Hasil Proses <i>Case Folding</i>	33
Tabel 3. 7 Contoh Hasil Proses <i>Cleansing</i>	33
Tabel 3. 8 Contoh Hasil Proses <i>Tokenizing</i>	34
Tabel 3. 9 Contoh Kata Kamus <i>Normalization</i>	35
Tabel 3. 10 Contoh Hasil Proses <i>Normalization</i>	36
Tabel 3. 11 Contoh Hasil Proses <i>Stopword Removal</i>	37
Tabel 3. 12 Contoh Hasil Proses <i>Stemming</i>	37
Tabel 3. 13 Contoh Hasil <i>Preprocessing</i>	38
Tabel 3. 14 Contoh Perhitungan <i>TF</i> dan <i>IDF</i>	39
Tabel 3. 15 Contoh Perhitungan <i>TF-IDF</i>	39
Tabel 3. 16 Detail Pengujian <i>10-fold Cross Validation</i>	42
Tabel 3. 17 Detail <i>Confusion Matrix</i>	43
Tabel 3. 18 Evaluasi Skenario Pertama.....	44
Tabel 3. 19 Evaluasi Skenario Kedua	45
Tabel 4. 1 Hasil Evaluasi Skenario Pertama.....	56
Tabel 4. 2 Hasil Evaluasi Skenario Kedua.....	57

DAFTAR GAMBAR

Gambar 2. 1 Konsep <i>Logistic Regression</i>	11
Gambar 2. 2 Konsep <i>Support Vector Machine</i>	13
Gambar 2. 3 Konsep Pembagian Data <i>K-fold Cross Validation</i>	18
Gambar 2. 4 Kerangka Berpikir	26
Gambar 3. 1 Diagram Alir Penelitian	29
Gambar 3. 2 Postingan <i>Instagram</i> Folkative	30
Gambar 3. 3 Tahapan <i>Preprocessing</i>	32
Gambar 3. 4 Diagram Alir Proses Klasifikasi <i>Logistic Regression</i>	40
Gambar 3. 5 Tahapan Evaluasi Kinerja Model Klasifikasi.....	43
Gambar 4. 1 Pengambilan Data Komentar <i>Instagram</i> Folkative.....	46
Gambar 4. 2 Tampilan Data Hasil <i>Scraping</i>	47
Gambar 4. 3 Dataset Komentar	48
Gambar 4. 4 Hasil <i>Preprocessing</i> Pada <i>Jupyter Notebook</i>	48
Gambar 4. 5 Hasil Pembobotan <i>TF-IDF</i>	49
Gambar 4. 6 Hasil Klasifikasi Menggunakan <i>Logistic Regression</i>	50
Gambar 4. 7 Visualisasi Dataset	50
Gambar 4. 8 Hasil <i>10-Cross Validation Support Vector Machine</i> Pertama	51
Gambar 4. 9 Hasil <i>10-Cross Validation Logistic Regression</i> Pertama	51
Gambar 4. 10 <i>Confusion Matrix Support Vector Machine</i> Skenario Pertama.....	52
Gambar 4. 11 <i>Confusion Matrix Logistic Regression</i> Skenario Pertama.....	52
Gambar 4. 12 Visualisasi Dataset Setelah <i>Oversampling</i>	53
Gambar 4. 13 Hasil <i>10-Cross Validation Support Vector Machine</i> Kedua	54
Gambar 4. 14 Hasil Hasil <i>10-Cross Validation Logistic Regression</i> Kedua	54
Gambar 4. 15 <i>Confusion Matrix Support Vector Machine</i> Skenario Kedua.....	55
Gambar 4. 16 <i>Confusion Matrix Logistic Regression</i> Skenario Kedua	55