

**PERBANDINGAN METODE SELEKSI FITUR DALAM ANALISIS
SENTIMEN KEBIJAKAN BAHAN BAKAR SUBSIDI DI *PLATFORM*
*YOUTUBE***

TUGAS AKHIR

Oleh:

AHMAD RAMADHANIEL IHSAN

1610817210002



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, JUNI 2023**

**PERBANDINGAN METODE SELEKSI FITUR DALAM ANALISIS
SENTIMEN KEBIJAKAN BAHAN BAKAR SUBSIDI DI *PLATFORM*
*YOUTUBE***

TUGAS AKHIR

Oleh:

AHMAD RAMADHANIEL IHSAN

1610817210002



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, JUNI 2023**

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini,

Nama : Ahmad Ramadhaniel Ihsan
NIM : 1610817210002
Fakultas : Teknik
Prodi : Teknologi Informasi
Judul Tugas Akhir : Perbandingan Metode Seleksi Fitur Dalam Analisis Sentimen Kebijakan Bahan Bakar Subsidi Di *Platform Youtube*
Pembimbing Utama : Nurul Fathanah Mustamin, S.Pd., M.T
Pembimbing Pendamping : Andreyan Rizky Bakara, S.Kom., M.Kom

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, Juni 2023

Penulis,

Ahmad Ramadhaniel Ihsan

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI

Perbandingan Metode Seleksi Fitur Dalam Analisis Sentimen Kebijakan Bahan Bakar
Subsidi Di Platform Youtube

Oleh

Ahmad Ramadhaniel Ihsan (1610817210002)

Telah dipertahankan di depan Tim Penguji pada 16 Juni 2023 dan dinyatakan

LULUS

Komite Penguji :

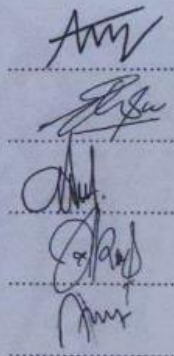
Ketua : Andry Fajar Zulkarnain, S.ST., M.T.
NIP 199007272019031018

Anggota 1 : Eka Setya Wijaya, S.T., M.Kom.
NIP 198205082008011010

Anggota 2 : Muti'a Maulida, S.Kom., M.T.I.
NIP 198810272019032013

Pembimbing Utama : Nurul Fathanah Mustamin, S.Pd., M.T.
NIP 199110252019032018

Pembimbing Pendamping : Andreyan Rizky Baskara, S.Kom., M.Kom.
NIP 199307032019031011



Banjarbaru, 18 JUL 2023
diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,



Dr. Ir. Yuliana Sari, S.Kom.,
M.Kom.
NIP 197401071998021001

Koordinator Program Studi
S-1 Teknologi Informasi,



Dr. Ir. Yuliana Sari, S.Kom.,
M.Kom.
NIP 198411202015042002

PERSETUJUAN TUGAS AKHIR

**PERBANDINGAN METODE SELEKSI FITUR DALAM ANALISIS SENTIMEN
KEBIJAKAN BAHAN BAKAR SUBSIDI DI *PLATFORM YOUTUBE***

OLEH

AHMAD RAMADHANIEL IHSAN

NIM. 1610817210016

Telah terpenuhi semua persyaratan akademik dan administrasi,

Telah diperiksa dan disetujui untuk diseminarkan

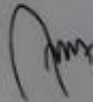
Banjarmasin, 03 Juni 2023

Pembimbing Utama,



Nurul Fathanah Mustamin, S.Pd., M.T.
NIP. 19911025 201903 2 018

Pembimbing Pendamping,



Andreyan Rizky Baskara, S.Kom., M.Kom
NIP. 19930703 201903 1 011

ABSTRAK

Pemberitaan terkait kebijakan digitalisasi pembayaran BBM bersubsidi tidak hanya dimuat di media cetak dan elektronik, tetapi juga di kanal YouTube. YouTube saat ini menjadi media informasi yang sangat banyak diakses oleh banyak orang. Menurut hasil riset 92% pengguna internet Indonesia menyatakan YouTube adalah tujuan pertama mereka ketika mencari video. Dari segi kuantitas penonton, YouTube sudah menyaingi televisi sebagai sarana media yang paling sering diakses orang Indonesia. Telah dilakukan penelitian terkait perbandingan 3 metode seleksi fitur yaitu *chi square*, *mutual information*, dan *information gain* dalam analisis sentimen kebijakan bahan bakar subsidi di platform youtube. Adapun Model terbaik *machine learning* dari penelitian ini adalah dengan menggunakan seleksi fitur *chi square* dengan metode pengklasifikasian menggunakan MKNN yang memperoleh nilai *accuracy* 86.40%, *precision* mencapai 84.48%, *recall* mencapai 90.74% dan *f1-score* mencapai 87.5%. Pada evaluasi pertama menggunakan data *testing* model ini mampu mengklasifikasi 89 data *testing* dengan benar dari 103 total data *testing* secara keseluruhan. Model mampu mengenali *true positive* sebanyak 49 data dan mampu mengenali *true negative* sebanyak 40 data. Sedangkan pada evaluasi kedua menggunakan seluruh data model ini mampu mengklasifikasi 461 data dengan benar dari 511 total data secara keseluruhan. Model mampu mengenali *true positive* sebanyak 230 data dan mampu mengenali *true negative* sebanyak 231 data.

Kata Kunci : Analisis Sentimen, Chi Square, Information Gain, Mutual Information, Modified K-Nearest Neighbor, Seleksi Fitur

ABSTRACT

News regarding the policy of digitizing subsidized fuel payments is not only published in print and electronic media, but also on the YouTube channel. YouTube is currently an information medium that is widely accessed by many people. According to research results, 92% of Indonesian internet users say YouTube is their first destination when looking for videos. In terms of viewership, YouTube has rivaled television as the media platform most frequently accessed by Indonesians. Research has been carried out regarding the comparison of 3 feature selection methods such as *chi square*, *mutual information*, and *information gain* in sentiment analysis of subsidized fuel policies on the YouTube platform. The best machine learning model from this research is to use chi square feature selection with the classification method using MKNN which obtains an accuracy value of 86.40%, precision reaches 84.48%, recall reaches 90.74% and f1-score reaches 87.5%. In the first evaluation using the data testing model, it was able to classify 89 data tests correctly out of 103 total data tests as a whole. The model is able to recognize 49 data of true positives and is able to recognize true negatives of 40 data. Whereas in the second evaluation using all the data this model was able to classify 461 data correctly out of 511 total data as a whole. The model is able to recognize 230 data of true positives and is able to recognize true negatives of 231 data.

Keyword: Chi-Square, Feature Selection, Information Gain, Modified K-Nearest Neighbor, Mutual Information, Sentiment Analysis

HALAMAN PERSEMBAHAN

Penulis mempersembahkan Tugas Akhir ini kepada:

1. Ayah, Ibu, Almarhum Nenek saya yang sudah membantu memberikan semangat, dukungan, dan materi serta doa yang tidak henti-hentinya dalam penyelesaian Tugas Akhir ini.
2. Ibu Nurul Fathanah Mustamin, S.Pd., M.T. selaku Dosen Pembimbing Utama yang selalu menyempatkan waktu untuk memberikan bimbingan, arahan dan dukungan kepada penulis dari awal sampai penyelesaian Tugas Akhir ini.
3. Bapak Andreyan Rizky Bakara, S.Kom., M.Kom. selaku Dosen Pembimbing Pendamping yang selalu menyempatkan waktu untuk memberikan bimbingan, arahan dan dukungan kepada penulis dari awal sampai penyelesaian Tugas Akhir ini.
4. Ibu Ir. Yuslena Sari, S.Kom., M.Kom. selaku Koordinator Program Studi Teknologi Informasi dan Dosen Pembimbing Akademik yang telah memberikan banyak sekali dukungan terkait penyelenggaraan perkuliahan hingga Tugas Akhir ini dapat diselesaikan dengan baik.
5. Seluruh Dosen beserta Staf Program Studi Teknologi informasi yang turut membantu dan mengarahkan dalam penyelesaian Tugas Akhir ini.
6. Teman – teman penulis Fatra, Cakra, Rizaldi, Akbar, El, Adit, Farid dan Nafis yang telah mendengarkan keluh kesah penulis serta selalu memberikan saran dan dukungan dalam menyelesaikan Tugas Akhir ini.
7. Gym, Teman-Teman Komunitas Badass yang telah mendengarkan keluh kesah penulis serta selalu memberikan saran dan dukungan dalam menyelesaikan Tugas Akhir ini.
8. Saudari Khairunnisa yang telah mendengarkan keluh kesah penulis serta selalu memberikan saran dan dukungan dalam menyelesaikan Tugas Akhir ini.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT. yang telah melimpahkan taufik, rahmat, dan hidayah-NYA kepada penulis sehingga dapat menyelesaikan Tugas Akhir dengan judul "Perbandingan Metode Seleksi Fitur Dalam Analisis Sentimen Kebijakan Bahan Bakar Subsidi Di *Platform Youtube*". Tugas Akhir ini disusun untuk memenuhi salah satu syarat untuk memperoleh gelar Sarjana Srata-1 Program Studi Teknologi Informasi di Fakultas Teknik Unuversitas Lambung Mangkurat, Banjarmasin. Dalam kesempatan ini penulis sangat bersyukur dan mengucapkan banyak terimakasih kepada semua pihak yang telah terlibat dan membantu, terutama:

1. Rektor Universitas Lambung Mangkurat, Bapak Prof. Dr. Ahmad, S.E., M.Si. yang memimpin dan memajemen jalannya seluruh perkuliahan yang ada di Universitas Lambung Mangkurat.
2. Dekan Fakultas Teknik, Bapak Prof. Dr. Ir. Irphan Fitriani Radam, S.T., M.T., IPU yang telah memberikan layanan terbaik dalam perkuliahan, terkhusus pada pelaksanaan Tugas Akhir di Lingkungan Fakultas Teknik.
3. Ketua Program Studi Teknologi Informasi. Ibu Ir. Yuslena Sari, S.Kom., M. Kom. Yang telah memberi arahan dan solusi dalam penyelesaian Tugas Akhir.
4. Pembimbing Utama, Ibu Nurul Fathanah Mustamin, S.Pd., M.T. yang telah memberikan waktu, bimbingan, dan arahan serta dukungan motivasi kepada Penulis selama dalam proses penelitian hingga penyelesaian Tugas Akhir.
5. Pembimbing Pendamping, Bapak Andreyan Rizky Bakara, S.Kom., M.Kom. yang telah memberikan waktu, bimbingan, dan arahan dalam proses penelitian dan penulisan laporan Tugas Akhir.
6. Dosen-dosen dan seluruh Staf program Studi Teknologi Informasi yang telah mengarahkan dan membantu dalam proses penyelesaian Tugas Akhir ini.

Akhir kata, penulis menyampaikan terimakasih kepada semua pihak yang turut membantu dalam penyelesaian laporan Tugas Akhir ini. Harapan dari penyusunan laporan ini adalah semoga apa yang penulis susun dapat memberikan manfaat bagi semua pihak yang memerlukan dan sebagai acuan informasi untuk pengembangan penelitian-penelitian terkait dimasa yang akan datang.

Banjarmasin, Juni 2023

Penulis,

Ahmad Ramadhaniel Ihsan

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN SAMPUL DALAM.....	ii
LEMBAR PERNYATAAN.....	iii
LEMBAR PENGESAHAN.....	iv
PERSETUJUAN TUGAS AKHIR.....	v
ABSTRAK.....	vi
ABSTRACT.....	vii
HALAMAN PERSEMBAHAN.....	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xviii
DAFTAR RIWAYAT HIDUP.....	xviii
BAB I PENDAHULUAN.....	Error! Bookmark not defined.
1.1 Latar Belakang	Error! Bookmark not defined.
1.2 Rumusan Masalah	Error! Bookmark not defined.
1.3 Batasan Masalah.....	Error! Bookmark not defined.
1.4 Tujuan Penelitian.....	Error! Bookmark not defined.
1.5 Manfaat Penelitian.....	Error! Bookmark not defined.
BAB II TINJAUAN PUSTAKA	Error! Bookmark not defined.
2.1 Landasan Teori.....	Error! Bookmark not defined.
2.1.1 Digitalisasi Pembelian Peralite dan Solar Bersubsidi.....	Error! Bookmark not defined.
2.1.2 YouTube	Error! Bookmark not defined.
2.1.3 Analisis Sentimen	Error! Bookmark not defined.
2.1.4 Machine learning	Error! Bookmark not defined.
2.1.5 Supervised Machine learning.....	Error! Bookmark not defined.

2.1.6	K-Nearest Neighbor	Error! Bookmark not defined.
2.1.7	Modified K-Nearest Neighbor (MKNN)	Error! Bookmark not defined.
2.1.8	Seleksi Fitur	Error! Bookmark not defined.
2.1.9	Information Gain.....	Error! Bookmark not defined.
2.1.10	Mutual Information.....	Error! Bookmark not defined.
2.1.11	Chi Square.....	Error! Bookmark not defined.
2.1.12	Text Preprocessing.....	Error! Bookmark not defined.
2.1.13	Pembobotan Teks.....	Error! Bookmark not defined.
2.1.14	Confusion Matrix	Error! Bookmark not defined.
2.2	Penelitian Terkait	Error! Bookmark not defined.
2.2.1	Analisis Sentimen Terhadap Ulasan Pengguna MRT Jakarta Menggunakan Information Gain dan Modified K-Nearest Neighbor	Error! Bookmark not defined.
2.2.2	Implementation of K-Nearest Neighbor (KNN) Algorithm For Public Sentiment Analysis of Online Learning	Error! Bookmark not defined.
2.2.3	Sentiment Analysis Peringkasan Review Film Menggunakan Metode Information Gain dan K-Nearest Neighbor	Error! Bookmark not defined.
2.2.4	Analisis Sentimen Masyarakat Indonesia Tentang Vaksin Covid-19 Di Twitter Dengan Menggunakan Metode K-Nearest Neighbor Dan Seleksi Fitur Chi Square	Error! Bookmark not defined.
2.2.5	Analisis Sentimen Twitter Menggunakan Naive Bayes Classifier dengan Seleksi Fitur Mutual Information.....	Error! Bookmark not defined.
2.3	Kerangka Pemikiran	Error! Bookmark not defined.
BAB III METODOLOGI PENELITIAN		Error! Bookmark not defined.
3.1.	Alat dan Bahan Penelitian	Error! Bookmark not defined.
3.1.1.	Alat Penelitian.....	Error! Bookmark not defined.
3.1.2.	Bahan Penelitian	Error! Bookmark not defined.
3.2.	Prosedur Penelitian.....	Error! Bookmark not defined.
3.2.1	Identifikasi Masalah.....	Error! Bookmark not defined.
3.2.2	Studi Literatur.....	Error! Bookmark not defined.
3.2.3	Pengumpulan Data.....	Error! Bookmark not defined.

3.2.4	Melakukan Eksperimen	Error! Bookmark not defined.
3.2.5	Implementasi Sistem.....	Error! Bookmark not defined.
3.2.6	Pengujian Model	Error! Bookmark not defined.
BAB IV HASIL DAN PEMBAHASAN		43
4.1	Pengumpulan Data	43
4.2	Perlabelan Data	43
4.3	<i>Text Pre-Processing</i>	44
4.3.1	<i>Case Folding</i>	45
4.3.2	<i>Filtering</i>	45
4.3.3	<i>Tokenzing</i>	46
4.3.4	<i>Normalization</i>	47
4.3.5	<i>Stopwards</i>	48
4.3.6	<i>Stemming</i>	49
4.4	Pembobotan Teks (TF-IDF)	50
4.5	Implementasi Model <i>Machine Learning</i>	53
4.5.1	<i>Import Library</i>	54
4.5.2	<i>Import Dataset</i>	55
4.5.3	<i>Split Dataset</i>	55
4.5.4	Implementasi Metode <i>Chi Square</i>	57
4.5.5	Implementasi Metode <i>Mutual Information</i>	57
4.5.6	Implementasi Metode <i>Information Gain</i>	57
4.5.7	Implementasi Metode Klasifikasi MKNN.....	57
4.5.8	Menyimpan Model <i>Machine Learning</i>	60
4.5.9	Menampilkan Model <i>Machine Learning</i>	60
4.5.10	Menjalankan Model <i>Machine Learning</i>	60
4.6	Eksperimen Model	61
4.6.1	Eksperimen Model <i>Machine Learning</i> dengan.....	61
4.6.2	Eksperimen Model <i>Machine Learning</i> dengan.....	64
4.6.3	Eksperimen Model <i>Machine Learning</i> dengan.....	67

4.7	Pengujian Model	70
4.8	Evaluasi Model Terbaik	70
BAB V KESIMPULAN DAN SARAN		73
5.1	Kesimpulan.....	73
5.2	Saran.....	73
DAFTAR PUSTAKA		Error! Bookmark not defined.
LAMPIRAN.....		Error! Bookmark not defined.
	Lampiran 1. Lampiran Dataset dan Hasil	78
	Lampiran 2. Contoh Perhitungan Manual.....	100
	Lampiran 3. Lembar Konsultasi.....	105

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait Pembobotan Teks	ERROR! BOOKMARK NOT DEFINED.
Tabel 2.2 Model Confusion Matrix....	ERROR! BOOKMARK NOT DEFINED.
Tabel 2.3 Ringkasan Penelitian Terkait	19
Tabel 3.1 Nama Library Dan Fungsinya.....	24
Tabel 3.2 Daftar Data Judul Video (Data Training)	28
Tabel 3.3 Daftar Data Judul Video (Data Testing)	29
Tabel 3.4 Acuan Perlabelan	30
Tabel 3.5 Contoh Komentar Dengan Label Positif Dan Negatif	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.6 Proses Case Folding.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.7 Proses Normalization	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.8 Proses Cleaning.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.9 Proses Stopword.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.10 Proses Stemming.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.11 Contoh Dokumen	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.12 Tabel Frekuensi Term.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.13 Contoh Hasil Pembobotan Text	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.14 Contoh Kalimat.....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.15 Contoh Pemisahan Kalimat	37
Tabel 3.16 Hasil Entropy SetiapTerm	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.17 Hasil Gain Setiap Term....	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.18 Contoh Hasil dari Mutual Information	ERROR! BOOKMARK NOT DEFINED.
Tabel 3.19 Contoh Hasil dari Chi-Square	ERROR! BOOKMARK NOT DEFINED.
Tabel 4.1 Contoh Komentar Dengan Label Positif Dan Negatif	44
Tabel 4.2 Acuan Perlabelan	44

Tabel 4.3 Hasil Case Folding.....	45
Tabel 4.4 Hasil Filtering	46
Tabel 4.5 Hasil Tokenizing.....	46
Tabel 4.6 Hasil Normalization	47
Tabel 4.7 Hasil Stopwords	48
Tabel 4.8 Hasil Stemming.....	50
Tabel 4.9 Hasil Pembobotan Teks	52
Tabel 4.10 Hasil Eksperimen Pertama Chi Square	61
Tabel 4.11 Nilai Confusion Matrix dari Eksperimen Pertama.....	62
Tabel 4.12 Hasil Eksperimen Kedua Chi Square.....	62
Tabel 4.13 Nilai Confusion Matrix dari Eksperimen Kedua	63
Tabel 4.14 Hasil Eksperimen Ketiga Chi-Square	63
Tabel 4.15 Nilai Confusion Matrix dari Eksperimen Ketiga	64
Tabel 4.16 Hasil Eksperimen Pertama MI.....	64
Tabel 4.17 Nilai Confusion Matrix dari Eksperimen Pertama MI.....	65
Tabel 4.18 Hasil Eksperimen Kedua MI.....	65
Tabel 4.19 Nilai Confusion Matrix dari Eksperimen Kedua MI	66
Tabel 4.20 Hasil Eksperimen Ketiga MI	66
Tabel 4.21 Nilai Confusion Matrix dari Eksperimen Ketiga MI	67
Tabel 4.22 Hasil Eksperimen Pertama IG.....	67
Tabel 4.23 Nilai Confusion Matrix dari Eksperimen Pertama IG	68
Tabel 4.24 Hasil Eksperimen Kedua IG	68
Tabel 4.25 Nilai Confusion Matrix dari Eksperimen Kedua IG	69
Tabel 4.26 Hasil Eksperimen Ketiga IG	69
Tabel 4.27 Hasil Eksperimen Matrix dari Eksperimen Ketiga IG	70
Tabel 4.28 Hasil Pengujian Model.....	70
Tabel 4.29 Hasil Pengujian Model Terbaik	71

DAFTAR GAMBAR

Gambar 2.1 Kerangka Pemikiran.....	ERROR! BOOKMARK NOT DEFINED.
Gambar 3.1 Alur Penelitian	ERROR! BOOKMARK NOT DEFINED.
Gambar 3.2 Alur <i>Texting Processing</i>	31
Gambar 3.3 Alur Sistem	41
Gambar 4.1 Implementasi Model Machine Learning	53
Gambar 4.2 Confusion Matrix	71
Gambar 4.3 Confusion Matrix Evaluasi Kedua.....	72