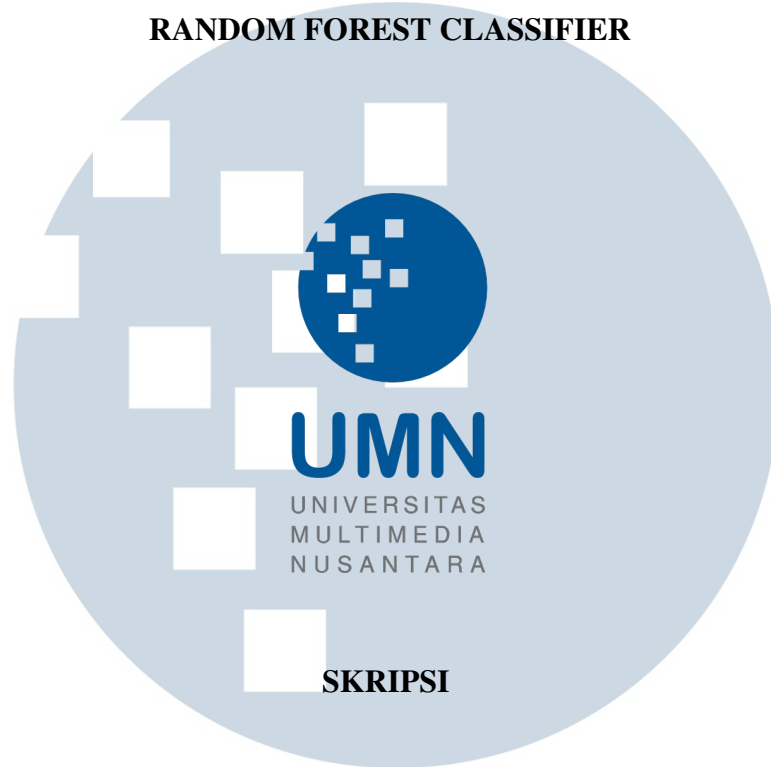


**ANALISIS SENTIMEN MASYARAKAT TERHADAP PERALIHAN
TELEVISI ANALOG KE DIGITAL MENGGUNAKAN ALGORITMA
RANDOM FOREST CLASSIFIER**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Elfajar Bintang Samudera
00000037235

UMN

**UNIVERSITAS
MULTIMEDIA
NUSANTARA**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

**ANALISIS SENTIMEN MASYARAKAT TERHADAP PERALIHAN
TELEVISI ANALOG KE DIGITAL MENGGUNAKAN ALGORITMA
RANDOM FOREST CLASSIFIER**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Elfajar Bintang Samudera

00000037235

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Elfajar Bintang Samudera
Nomor Induk Mahasiswa : 00000037235
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Masyarakat terhadap Peralihan Televisi Analog ke Digital Menggunakan Algoritma Random Forest Classifier

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

UMM

Tangerang, 16 Juni 2023



(Elfajar Bintang Samudera)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN MASYARAKAT TERHADAP PERALIHAN TELEVISI ANALOG KE DIGITAL MENGGUNAKAN ALGORITMA RANDOM FOREST CLASSIFIER

oleh

Nama : Elfajar Bintang Samudera
NIM : 00000037235
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika


Telah diujikan pada hari Selasa, 20 Juni 2023

Pukul 15:00 s/d 17:00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang


(Suwito Pomalingo, S.Kom., M.Kom.)
NIDN: 0911098201

Penguji


(Farica Perdana Putri, S.Kom., M.Sc.)
NIDN: 0331019301

Pembimbing


(Alexander Waworuntu, S.Kom., M.T.I.)

NIDN: 0309068503

Ketua Program Studi Informatika,


(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Elfajar Bintang Samudera
NIM : 00000037235
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN MASYARAKAT TERHADAP PERALIHAN
TELEVISI ANALOG KE DIGITAL MENGGUNAKAN ALGORITMA
RANDOM FOREST CLASSIFIER**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 16 Juni 2023

Yang menyatakan



Elfajar Bintang Samudera

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Halaman Persembahan / Motto

”It might be that this idea is only the beginning of Wisdom, and not its final form.”

Edward James Kenway (1693 – 1735)

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Masyarakat terhadap Peralihan Televisi Analog ke Digital Menggunakan Algoritma Random Forest Classifier dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Alexander Waworuntu, S.Kom., M.T.I., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini. memberikan bimbingan atas terselesainya Skripsi/Tesis ini.
5. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.
6. Norbertus Dewa Rucci, sebagai sahabat yang selalu membantu dalam keadaan susah maupun senang.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 16 Juni 2023



Elfajar Bintang Samudera

ANALISIS SENTIMEN MASYARAKAT TERHADAP PERALIHAN TELEVISI ANALOG KE DIGITAL MENGGUNAKAN ALGORITMA RANDOM FOREST CLASSIFIER

Elfajar Bintang Samudera

ABSTRAK

Televisi merupakan salah satu media hiburan dan informasi yang digemari masyarakat. Jenis televisi yang banyak beredar di masyarakat yaitu televisi analog. Seiring berkembangnya zaman, televisi analog semakin tertinggal dan mulai tergantikan oleh jenis televisi baru yakni televisi digital. Televisi digital dinilai memberikan kinerja yang lebih baik dibanding televisi analog. Tepat pada tanggal 2 November 2022 silam, keputusan Pemerintah dalam memastikan bahwa migrasi siaran televisi analog ke digital telah dilaksanakan. Seiring dilaksanakannya program Analog Switch Off di berbagai daerah, terdapat berbagai macam pandangan pro dan kontra yang terlontar di kalangan masyarakat. Salah satu media sosial yang dipakai oleh masyarakat dalam berkomunikasi dan penyebaran informasi secara cepat adalah Twitter. Oleh karena itu, penelitian ini dilakukan dengan tujuan untuk melakukan klasifikasi apakah sentimen yang dilontarkan oleh masyarakat terkait kebijakan *alaog switch off* termasuk dalam sentimen positif atau negatif. Model klasifikasi yang dipakai adalah algoritma *random forest* dengan bantuan Lexicon Inset dalam melabeli data, Count Vectorizer dan TF IDF Vectorizer untuk proses vektorisasi data serta memberikan bobot pada kata, juga proses pembagian data latih dan data uji dengan rasio yang beragam. Dari penelitian ini didapatkan klasifikasi terbaik dengan menggunakan metode vektorisasi Count Vectorizer, rasio data latih dan data uji sebesar 80%:20% dengan nilai *accuracy* 88%, *precision* sebesar 88%, *recall* 88%, dan f1-score 88%.

Kata kunci: Televisi Analog, Televisi Digital, Sentimen, *Twitter*, *Random Forest*

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Analysis of Public Sentiment on the Transition of Analogue to Digital Television using a Random Forest Classifier Algorithm

Elfajar Bintang Samudera

ABSTRACT

Television is one of the most popular entertainment and information media public. The type of television that is widely circulating in society is analog television. Along with the development of the times, analog television is getting left behind and is starting to be replaced by a new type of television, namely digital television. Digital television is considered to provide better performance than analog television. Right on November 2, 2022, the Government's decision to ensure that the migration of analogue to digital television broadcasts has been carried out. As the Analog Switch Off program was implemented in various regions, there were various pro and con views raised by the community. One of the social media used by the public in communicating and disseminating information quickly is Twitter. Therefore, this research was conducted with the aim of classifying whether the sentiments expressed by the public regarding the analog switch off policy included positive or negative sentiments. The classification model used is the random forest algorithm with the help of Lexicon Inset in labeling data, the Count Vectorizer and TF IDF Vectorizer for the process of vectorizing data and giving weights to words, as well as the process of dividing training data and test data with various ratios. From this study, the best classification was obtained using the Count Vectorizer vectorization method, the ratio of training data and test data was 80%:20% with accuracy value of 88%, precision of 88%, recall value 88%, and f1-score 88%.

Keywords: Analogue Television, Digital Television, Sentiment, Twitter, Random Forest

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Analisis Sentimen	6
2.2 Analog Switch Off	6
2.3 Twitter	6
2.4 Text Preprocessing	7
2.5 Decision Tree	7
2.6 Random Forest Classifier	9
2.7 Term Frequency-Inverse Document Frequency (TF-IDF)	9
2.8 Hyperparameter Tuning	10
2.9 Confusion Matrix	11
BAB 3 METODOLOGI PENELITIAN	13
3.1 Gambaran Umum Penelitian	13
3.2 Metodologi Penelitian	14
3.2.1 Studi Literatur	14
3.2.2 Pengumpulan Data	14
3.2.3 Pengolahan Data	17
3.2.4 Menerapkan fitur TF-IDF	20
3.2.5 Implementasi Algoritma Random Forest Classifier	21
3.2.6 Uji Evaluasi	22
BAB 4 HASIL DAN DISKUSI	23
4.1 Hasil Implementasi Pembelajaran Mesin	23
4.2 Alur Skenario Uji Coba	37
4.3 Uji Skenario	39
4.3.1 Skenario 1-8	41
4.3.2 Skenario 9-16	48
4.4 Uji Evaluasi	51
4.5 Uji Prediksi	55
BAB 5 SIMPULAN DAN SARAN	59
5.1 Simpulan	59

5.2 Saran	59
DAFTAR PUSTAKA	61



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 3.1	Diagram Alur Penelitian	13
Gambar 3.2	Jumlah <i>tweet</i> televisi analog dan televisi digital sejak September 2022 hingga April 2023	15
Gambar 3.3	Diagram alur data scrapping	16
Gambar 3.4	Data hasil scrapping	16
Gambar 3.5	Diagram Alur Preprocessing	17
Gambar 3.6	Diagram Alur Penerapan Fitur TF-IDF	21
Gambar 4.1	Pengambilan <i>tweet</i> dengan kata kunci "televisi analog"	23
Gambar 4.2	Pengambilan <i>tweet</i> dengan kata kunci "televisi digital"	24
Gambar 4.3	Proses <i>drop column date</i>	24
Gambar 4.4	Proses <i>Tahap casefolding</i>	24
Gambar 4.5	Proses penghapusan tanda tagar dan <i>username</i>	25
Gambar 4.6	Proses <i>Tahap penghapusan URL</i>	25
Gambar 4.7	Tahap penghapusan karakter non-numerik dan bukan alfabet	26
Gambar 4.8	Tahap penghapusan <i>multiple spaces tweet</i>	26
Gambar 4.9	Tahap penghapusan <i>missing values</i> dan <i>duplicated values</i>	27
Gambar 4.10	Tahap penggabungan dua <i>dataframe</i> menjadi satu	27
Gambar 4.11	Tahap <i>tokenizing</i>	28
Gambar 4.12	Tahap <i>stopword removal</i>	29
Gambar 4.13	Tahap <i>normalization</i>	30
Gambar 4.14	Tahap <i>Stemming</i>	31
Gambar 4.15	Contoh hasil <i>stemming</i>	31
Gambar 4.16	Tahap <i>labelling</i>	32
Gambar 4.17	Proses pengecekan kalimat positif, negatif, dan nol	33
Gambar 4.18	<i>Pie Chart</i> data positif dan negatif	34
Gambar 4.19	Proses penambahan hasil <i>labelling</i> pada <i>dataframe</i>	34
Gambar 4.20	Tahap penggabungan kata-kata selama tahap <i>preprocessing</i> menjadi sebuah kalimat	35
Gambar 4.21	<i>Wordcloud</i> sentimen yang dilabeli positif	36
Gambar 4.22	<i>Wordcloud</i> sentimen yang dilabeli negatif	37
Gambar 4.23	Alur skenario uji coba penelitian	38
Gambar 4.24	Proses vektorisasi data teks menggunakan <i>CountVectorizer</i>	41
Gambar 4.25	Proses <i>train test split</i> pada skenario pertama dan kedua dengan ukuran <i>test_size</i> 20%	41
Gambar 4.26	Proses <i>train test split</i> pada skenario ketiga dan keempat dengan ukuran <i>test_size</i> 30%	42
Gambar 4.27	Proses <i>train test split</i> pada skenario kelima dan keenam dengan ukuran <i>test_size</i> 40%	42
Gambar 4.28	Proses <i>train test split</i> pada skenario ketujuh dan kedelapan dengan ukuran <i>test_size</i> 50%	42
Gambar 4.29	Implementasi algoritma <i>random forest classifier</i> untuk melatih data latih.	43
Gambar 4.30	Pengujian performa model menggunakan fungsi <i>predict</i>	43
Gambar 4.31	Parameter <i>default</i> sebelum dilakukan proses <i>hyperparameter tuning</i>	44
Gambar 4.32	Proses implementasi metode <i>Random Search</i> untuk mencari parameter terbaik	45

Gambar 4.33	Hasil pencarian <i>random search</i> pada skenario kedua untuk mendapatkan parameter terbaik	46
Gambar 4.34	Proses implementasi algoritma <i>random forest</i> menggunakan parameter optimal yang diperoleh melalui hasil iterasi <i>random search</i>	46
Gambar 4.35	Proses vektorisasi data <i>train</i> menggunakan <i>Tfidfvectorizer</i>	49
Gambar 4.36	<i>dataframe</i> yang berisi kolom <i>features</i> beserta skor IDF	49
Gambar 4.37	<i>dataframe</i> yang berisi kolom <i>features</i> beserta skor TF-IDF	50
Gambar 4.38	Proses <i>train test split</i> pada skenario kesembilan dan kesepuluh dengan ukuran <i>test_size</i> 20%	50
Gambar 4.39	Proses <i>train test split</i> pada skenario kesebelas dan keduabelas dengan ukuran <i>test_size</i> 30%	51
Gambar 4.40	Proses <i>train test split</i> pada skenario ketigabelas dan keempatbelas dengan ukuran <i>test_size</i> 40%	51
Gambar 4.41	Proses <i>train test split</i> pada skenario kelimabelas dan keenambelas dengan ukuran <i>test_size</i> 50%	51
Gambar 4.42	Hasil pengujian prediksi pada percobaan pertama	56
Gambar 4.43	Hasil pengujian prediksi pada percobaan kedua	56
Gambar 4.44	<i>Text area</i> untuk menerima <i>input</i> sentimen yang akan diklasifikasi	57
Gambar 4.45	Hasil prediksi sentimen pada percobaan pertama	57
Gambar 4.46	Percobaan prediksi pada percobaan kedua	58
Gambar 4.47	Hasil prediksi sentimen pada percobaan kedua	58



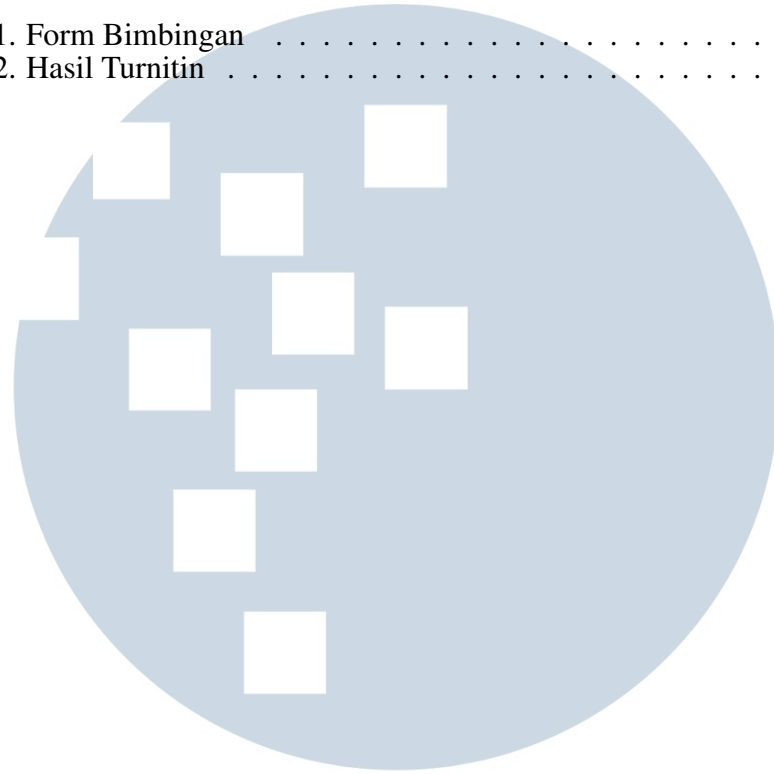
DAFTAR TABEL

Tabel 2.1	Tabel <i>Confusion Matrix</i>	11
Tabel 3.1	Contoh <i>Data Cleaning</i>	18
Tabel 3.2	Contoh <i>Casefolding</i>	18
Tabel 3.3	Contoh <i>Tokenizing</i>	18
Tabel 3.4	Contoh <i>Stowpord Removal</i>	19
Tabel 3.5	Contoh <i>Normalizing</i>	19
Tabel 3.6	Contoh <i>Stemming</i>	20
Tabel 4.1	Tabel Uji Skenario Pembelajaran Mesin	40
Tabel 4.2	Tabel rekap skenario terbaik setiap skenario	47
Tabel 4.2	Tabel rekap skenario terbaik setiap skenario	48
Tabel 4.3	Tabel Hasil Uji Skenario	52
Tabel 4.3	Tabel Hasil Uji Skenario	53
Tabel 4.3	Tabel Hasil Uji Skenario	54



DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	63
Lampiran 2. Hasil Turnitin	65



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA