

TUGAS AKHIR
SENTIMENT ANALYSIS KOMENTAR YOUTUBE SAMSUNG
S20 MENGGUNAKAN METODE MAJORITY VOTE

Disusun untuk Melengkapi dan Memenuhi Syarat Kelulusan
Program Strata S1 Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2021

HALAMAN PERSETUJUAN

SENTIMENT ANALYSIS KOMENTAR YOUTUBE SAMSUNG S20 MENGGUNAKAN METODE MAJORITY VOTE

Oleh:

Risang Putra Pradana

1810651120

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada Sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui oleh:

Pembimbing I

Pembimbing II

Deni Arifianto, M.Kom.

NIDN. 0718068103

Habibatul Azizah Al Faruq, M.Pd.

NIDN. 0718128901

HALAMAN PENGESAHAN
SENTIMENT ANALYSIS KOMENTAR YOUTUBE SAMSUNG
S20 MENGGUNAKAN METODE MAJORITY VOTE

Oleh:

Risang Putra Pradana

1810651120

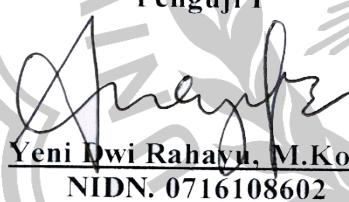
Telah mempertanggung jawabkan Laporan Tugas Akhir pada Sidang Tugas Akhir tanggal 17 April 2021 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh:

Pengaji I


Yeni Dwi Rahayu, M.Kom.
NIDN. 0716108602

Pembimbing I


Deni Arifianto, M.Kom.
NIDN. 0718068103

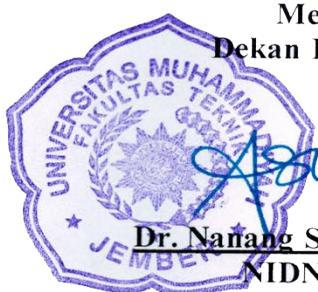
Pengaji II


Victor Wahanggara, S.Kom, M.Kom
NIDN. 0715039003

Pembimbing II


Habibatul Azizah Al Faruq, M.Rd.
NIDN. 0718128901

Mengesahkan,
Dekan Fakultas Teknik




Dr. Nanang Saiful Rizal, S.T., M.T.
NIDN. 0705047806

Mengetahui,
Ketua Program Studi Teknik Informatika




Triawan Adi Cahyanto, S.Kom., M.Kom
NIDN. 0702098804

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

NAMA : RISANG PUTRA PRADANA

NIM : 1810651120

INSTITUSI : Strata-1 Program Studi Teknik Informatika Fakultas
Teknik, Universitas Muhammadiyah Jember.

Menyatakan bahwa Tugas Akhir yang berjudul "**Sentiment Analysis Komentar Youtube Samsung S20 Menggunakan Metode Majority Vote**" bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi akademik.

Jember, Juni 2021

Yang menyatakan,



Risang Putra Pradana

NIM. 1810651120

KATA PENGANTAR

Puji syukurn Kehadirat Allah yang telah melimpahkan Rahmat dan Hidayah-Nya, kemudahan serta kelancaran selama penyusunan Tugas Akhir sehingga dapat terselesaikan. Shalawat serta salam senantiasa tercurahkan kepada Nabi Muhammad SAW beserta keluarga, para sahabat dan para pengikutnya yang tetap istiqamah. Selama penyusunan Tugas Akhir, penulis telah banyak mendapat bimbingan dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis bermaksud menyampaikan ucapan terima kasih kepada:

1. Bapak Dr. Nanang Saiful Rizal, S.T., M.T selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
2. Bapak Triawan Adi Cahyanto, M.Kom selaku Kepala Program Studi Teknik Informatika di Universitas Muhammadiyah Jember.
3. Bapak Deni Arifianto M.Kom selaku dosen pembimbing 1 tugas akhir yang telah memberikan bimbingan dari awal sampai akhir dengan penuh kesabaran sehingga tugas akhir ini dapat terselesaikan.
4. Ibu Habibatul Azizah Al Faruq M.Pd selaku dosen pembimbing 2 tugas akhir yang telah memberikan bimbingan dari awal sampai akhir dengan penuh kesabaran sehingga tugas akhir ini dapat terselesaikan.
5. Ibu Yeni Dwi Rahayu, M.Kom selaku dosen penguji 1 yang telah memberikan kritik, saran dan masukkan yang membangun dalam penelitian ini.
6. Bapak Victor Wahanggara, S.Kom, M.Kom selaku dosen penguji 2 yang telah memberikan kritik, saran dan masukkan yang membangun dalam penelitian ini.
7. Kepada orang tua penulis, Bapak Hariyono,S.Kep.,Ners dan Ibu Ariesta Efiyanti S.E. yang memberikan semangat, doa, motivasi, dan dukungan kepada penulis sehingga dapat menyelesaikan Tugas Akhir ini.
8. Teman-teman seperjuangan yang senasib alumni Politeknik Negeri Jember terimakasih atas doa dan dukungannya selama ini,

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan. Oleh karena itu, segala kritik dan saran yang bersifat membangun selalu penulis harapkan. Semoga Tugas Akhir ini dapat memberikan manfaat. Akhir kata, semoga Allah senantiasa melimpahkan Rahmat dan Hidayah-Nya kepada kita semua. Amin amin ya robbal 'alamin.

Jember, Juni 2021

Penulis



MOTTO

"Dramamu masih memiliki jalan yang panjang. Jadi jangan memaksakan diri untuk pergi dengan cepat. Jika kamu berjalan lambat, kamu bisa melihat lebih banyak dan lebih detail daripada orang-orang yang pergi dengan cepat. Jika kamu bertanya kepadaku siapa yang akan lebih berkembang di antara keduanya, aku katakan itu orang yang berjalan lambat dan melihat banyak."

Kang Oh Hyuk (Dream High 2011)

"There's no better observer than someone who distrusts you with all their heart."

Sora (No Game No Life)

"Lebih baik dibenci apa adanya, daripada dicintai yang tidak apa adanya"

Hikigaya Hachiman(Oregairu)

"Perbedaan antara pemula dan master adalah bahwa master telah gagal lebih dari yang telah dicoba pemula."

Koro-sensei (Ansatsu Kyoushitsu)

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN.....	iv
ABSTRAK.....	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
MOTTO.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Penelitian	3
1.4. Tujuan	4
1.5. Manfaat	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Sentiment Analysis	5
2.2. Penelitian Terkait	5
2.3. Youtube API	8
2.4. Text Mining.....	9
2.5. Preprocessing Text mining.....	9
2.6. Scraping.....	10

2.7.	Term Frequency – Invers Document Frequency (TF-IDF).....	11
2.8.	<i>Support Vector Machine</i>	11
2.9.	<i>K-Nearest Neighbors</i>	13
2.10.	<i>Naïve Bayes</i>	15
2.11.	<i>Majority vote</i>	15
2.12.	Python	16
2.13.	Jupyter Notebook	16
2.14.	<i>Cross validation</i>	17
2.15.	Teknik Imbalance.....	17
2.16.	Evaluasi Sistem.....	17
	BAB III METODOLOGI PENELITIAN	18
3.1.	Alur Penelitian	18
3.2.	Identifikasi Masalah	18
3.3.	Studi Literatur	19
3.4.	Desain Sistem.....	19
3.4.1.	Preprocessing	21
3.4.2.	Preprocessing Text mining.....	23
3.4.3.	TF-IDF	29
3.4.4.	Pembagian Data	31
3.4.5.	<i>Cross Validation</i>	31
3.4.6.	Algoritma <i>Support Vector Machine</i> (SVM).....	34
3.4.7.	Algoritma <i>K-Nearest Neighbors</i> (K-NN).....	37
3.4.8.	Algoritma <i>Naïve Bayes</i>	41
3.4.9.	Algoritma <i>Majority vote</i>	43
3.5.	Skenario pengujian.....	44
	BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM	46
4.1.	Implementasi Sistem <i>Sentiment Analysis</i>	46
4.1.1.	<i>Scraping</i>	46
4.1.2.	Labelisasi.....	48
4.1.3.	Mengatasi <i>imbalance</i>	49

4.1.4. <i>Preprocessing Text Mining</i>	49
4.1.5. TF-IDF	52
4.1.6. Pembagian data	53
4.1.7. <i>Cross validation</i>	53
4.2. Hasil Pengujian	57
4.2.1. Skenario perbandingan akurasi yang dihasilkan Majority vote dibandingkan metode SVM, KNN dan <i>Naïve Bayes</i>	57
4.2.2. Skenario mencari metode yang terpilih dalam metode <i>Majority vote</i>	58
BAB V KESIMPULAN DAN SARAN	59
5.1. Kesimpulan	59
5.2. Saran.....	59
DAFTAR PUSTAKA	60
LAMPIRAN	63

DAFTAR GAMBAR

Gambar 3.1 Alur penelitian	18
Gambar 3.2 Desain Sistem	20
Gambar 3.3 <i>Scraping</i> Komentar	21
Gambar 3.4 <i>Scraping</i> dengan <i>google spreadsheets</i>	21
Gambar 3.5 Alur mengatasi <i>Imbalance</i>	21
Gambar 3.6 Proses <i>Cleaning</i> komentar YouTube	23
Gambar 3.7 Proses <i>Case Folding</i>	24
Gambar 3.8 Proses <i>Tokenizing</i>	25
Gambar 3.9 Proses <i>Stopword Removal</i>	27
Gambar 3.10 Proses <i>Stemming</i>	28
Gambar 3.11 <i>Flowchart Support Vector Machine</i>	35
Gambar 3.12 <i>Flowchart K-Nearest Neighbours</i>	37
Gambar 3.13 <i>Flowchart Majority Vote</i>	43
Gambar 3.14 Skenario pengujian	44
Gambar 4.1 Proses Scraping (1)	46
Gambar 4.2 Proses Scraping (2)	47
Gambar 4.3 Proses Scraping (3)	47
Gambar 4.4 Proses Scraping (4)	48

DAFTAR TABEL

Tabel 2.1 Contoh tahap <i>Tokenizing</i> Teks	10
Tabel 2.2 Contoh tahap Stemming Teks.....	10
Tabel 3.1 Contoh proses <i>Cleaning</i>	24
Tabel 3.2 Contoh proses <i>Case Folding</i>	25
Tabel 3.3 Contoh Proses <i>Tokenizing</i>	26
Tabel 3.4 Contoh proses <i>Stopword Removal</i>	27
Tabel 3.5 Contoh proses <i>Stemming</i> menggunakan <i>library</i> NLTK.	29
Tabel 3.6 Menghitung TF-IDF	29
Tabel 3.7 Penggolongan Kelas teks.....	37
Tabel 3.8 Hasil <i>Cosine Similarity</i> 1	38
Tabel 3.9 Hasil <i>Cosine Similarity</i> 3.....	40
Tabel 3.10 Data k K-NN.....	40
Tabel 3.11 Algoritma <i>Naïve Bayes</i>	41
Tabel 3.12 Penggunaan metode <i>Majority Vote</i>	44
Tabel 4.1 Hasil Labelisasi.....	48
Tabel 4.2 Hasil mengatasi <i>imbalance</i>	49
Tabel 4.3 Preprocessing Text Mining.....	50
Tabel 4.4 Potongan Hasil TF-IDF	52
Tabel 4.5 perhitungan akurasi 2-fold Langkah uji1	53
Tabel 4.6 Akurasi seluruh Langkah uji.....	55
Tabel 4.7 Akurasi setiap metode dengan 50 data uji.....	57