

LAPORAN SKRIPSI

***SENTIMENT ANALYSIS* MENGGUNAKAN METODE *TEXT MINING*
DAN RANDOM FOREST UNTUK KLASIFIKASI KELUHAN PENGGUNA
(STUDI KASUS: APLIKASI *KAI ACCESS*)**



Disusun Oleh :

Agung Gunawan

2019230012

PROGRAM STUDI TEKNOLOGI INFORMASI

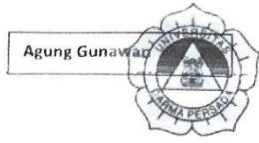
FAKULTAS TEKNIK

UNIVERSITAS DARMA PERSADA

JAKARTA

2023

LEMBAR PERBAIKAN



UNIVERSITAS DARMA PERSADA

Jl. Taman Malaka Selatan, Pondok Kelapa, Jakarta Timur, Indonesia 13450

Telp. (021) 8649051, 8649053, 8649057 Fax. (021) 8649052

E-mail : humas@unsada.ac.id Home page : <http://www.unsada.ac.id>

LEMBAR PERBAIKAN SEMINAR TUGAS AKHIR

Nama : Agung Gunawan
NIM : 2019230012
Fakultas/Jurusan : Teknik / Teknologi Informasi

No.	Keterangan	Dosen
	Gambar di cetak ketuk. Silahkan di zip panahan	Serdin

Mengetahui,

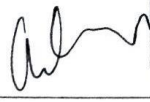






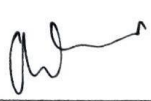
Kajur Teknologi Informasi

Adam Arif Budiman, M.Kom.

MONDZUKURI • TRILINGUAL • ENERGI TERBARUKAN



LEMBAR BIMBINGAN

NO.	Tanggal Bimbingan	Kegiatan, Lokasi, Hasil	Paraf Verifikasi
1	16 Juni 2023	Menyerahkan BAB I - II	
2	16 Juni 2023	Istilah Asing Ditulis <i>Italic</i>	
3	19 Juni 2023	Uraian Landasan Teori Ditambah	
4	23 Juni 2023	Menyerahkan Revisi BAB I - II	
5	7 Juli 2023	Menyerahkan BAB III - V	
6	10 Juli 2023	Mengikuti Bimbingan <i>Online</i> , Revisi BAB I - V	
7	10 Juli 2023	Mengikuti Bimbingan <i>Online</i> , Menjelaskan 100% <i>Web Streamlit</i>	
8	13 Juli 2023	Menyerahkan Revisi BAB I - V	
9	14 Juli 2023	ale seminar	18.

LEMBAR PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : Agung Gunawan

NIM : 2019230012

Fakultas : Teknik

Jurusan : Teknologi Informasi

Menyatakan bahwa Laporan Skripsi ini saya susun sendiri berdasarkan hasil *filtering data* ulasan pengguna *KAI Access* di *Google Play Store* serta memadukannya dengan studi literatur atau bahan-bahan referensi lain yang terkait dan relevan di dalam penyelesaian Laporan Skripsi ini.

Demikian pernyataan ini penulis buat dengan sesungguhnya

Jakarta, 10 Juli 2023



Agung Gunawan

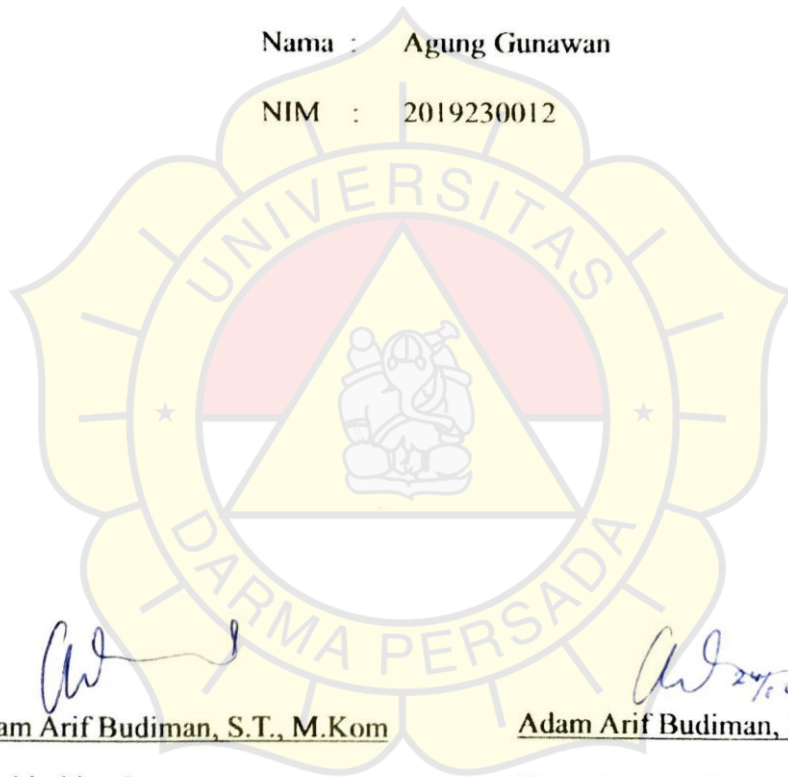
LEMBAR PENGESAHAN

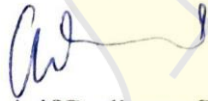
SENTIMENT ANALYSIS MENGGUNAKAN METODE *TEXT MINING* DAN
RANDOM FOREST UNTUK KLASIFIKASI KELUHAN PENGGUNA (STUDI
KASUS: APLIKASI *KAI ACCESS*)

Disusun oleh :

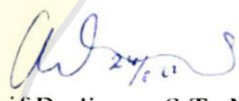
Nama : Agung Gunawan

NIM : 2019230012




Adam Arif Budiman, S.T., M.Kom

Pembimbing Laporan


Adam Arif Budiman, S.T., M.Kom

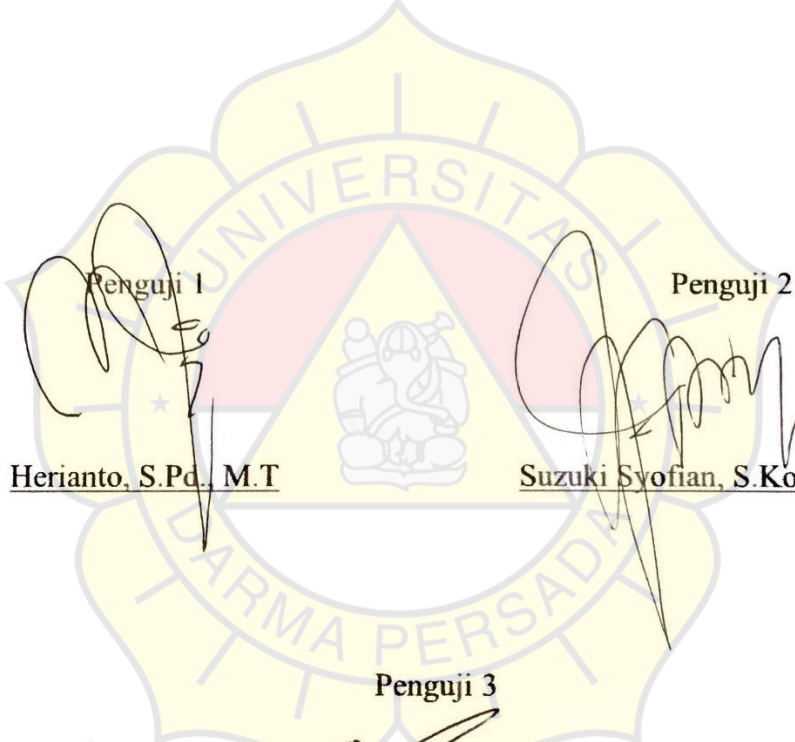
Ketua Jurusan Teknologi Informasi


LEMBAR PENGUJI SKRIPSI

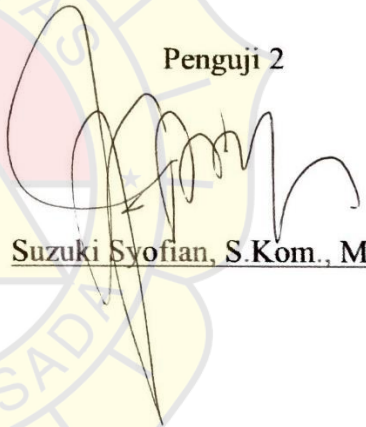
Laporan Skripsi yang berjudul :

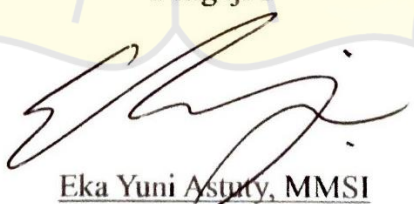
“*SENTIMENT ANALYSIS* MENGGUNAKAN METODE *TEXT MINING* DAN
RANDOM FOREST UNTUK KLASIFIKASI KELUHAN PENGGUNA (STUDI
KASUS: APLIKASI *KAI ACCESS*)” ini telah ujikan pada tanggal

14 Agustus 2023



Penguji 1

Herianto, S.Pd., M.T

Penguji 2

Suzuki Syofian, S.Kom., M.Kom

Penguji 3

Eka Yuni Astuty, MMSI

KATA PENGANTAR

Puji syukur kepada Allah SWT, atas limpahan rahmah dan hidayah-Nya, sehingga penulis dapat menyelesaikan Laporan Skripsi ini. Tak lupa shalawat serta salam semoga tercurah limpahkan kepada nabi besar kita yaitu Nabi Muhammad SAW, beserta keluarga, sahabat dan kita sebagai umatnya mudah-mudahan bisa mendapatkan syafaat beliau di yaumul akhir kelak.

Laporan Skripsi dengan judul “*Sentiment Analysis Menggunakan Metode Text Mining dan Random Forest Untuk Klasifikasi Keluhan Pengguna (Studi Kasus: Aplikasi KAI Access)*” penulis susun guna memenuhi persyaratan dalam menyelesaikan Tugas Akhir, pada Program Studi Teknologi Informasi di Universitas Darma Persada.

Dan pada kesempatan kali ini penulis ingin mengucapkan banyak terima kasih kepada pihak-pihak yang telah membantu secara langsung maupun tidak secara langsung di dalam penulisan tugas akhir ini. Oleh karena itu, penulis mengucapkan banyak terima kasih kepada:

1. Dekan Fakultas Teknik Universitas Darma Persada: Bapak Dr. Ade Supriatna S.T., M.T.
2. Ketua Jurusan Teknologi Informasi dan Dosen Pembimbing saya: Bapak Adam Arif Budiman, S.T., M.Kom.
3. Seluruh Dosen Teknologi Informasi di Universitas Darma Persada yang tidak bisa saya sebutkan satu persatu namun tidak mengurangi rasa hormat saya, yang sudah mengajarkan saya segudang ilmu yang tidak ternilai kebajikannya, terlebih untuk saya di masa kini dan masa depan kelak.

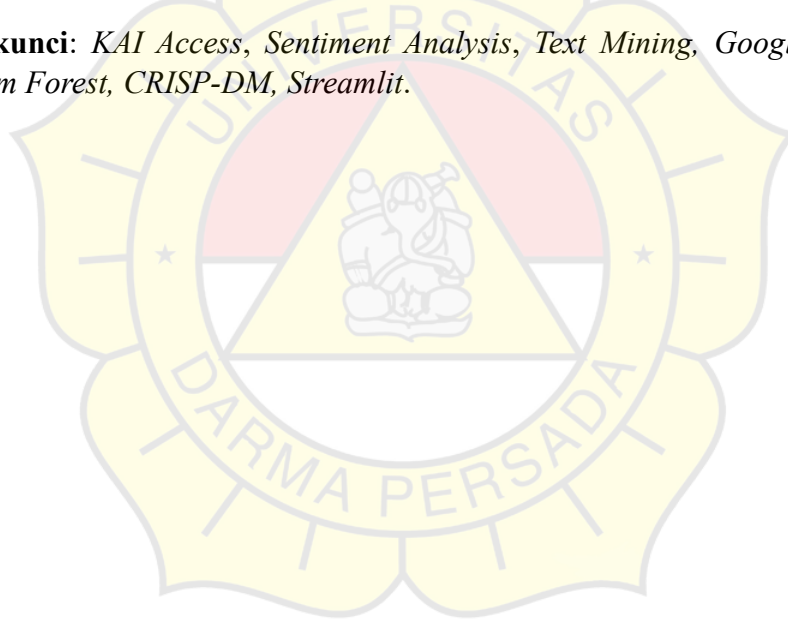
4. Kedua orang tua saya: Ibu Ngaisah dan Bapak Parimin, atas kasih sayangnya yang selalu memberikan dukungan moral dan materiil kepada penulis, khususnya dalam menjalani masa perkuliahan.
5. Teman - teman Aliansi Daring: Dhino Rahmad Kusuma, Imam Andrianto, Muhammad Sadam Mahendra, Malik Abdul Aziz, Firmansah, serta Dwi Putro Sulisty, yang telah menemani dan membantu penulis dalam menjalani masa perkuliahan.
6. Teman lainnya: M. Syamsul Muhyidin dan Dafit Krismunanto, yang telah menyediakan waktunya dalam memberikan semangat, ide dan kontribusi kepada penulis.

Penulis menyadari bahwa Laporan Skripsi ini masih belum sempurna. Oleh karena itu, kritik dan saran yang membangun selalu penulis harapkan, demi penyusunan laporan yang lebih baik kedepannya. Penulis berharap, semoga Laporan Skripsi ini dapat bermanfaat bagi penulis dan juga para pembaca.

ABSTRAK

KAI Access merupakan aplikasi yang dimiliki oleh PT. Kereta Api Indonesia untuk pelayanan *mobile ticketing* secara *online*. Keluhan pengguna merupakan aspek penting yang perlu dipertimbangkan ketika mengembangkan aplikasi *KAI Access* untuk layanan *mobile ticketing*. Pengguna *KAI Access* menimbulkan banyak opini terhadap aplikasi *KAI Access*, salah satunya yaitu pada *Google Play Store*. *Sentiment analysis* merupakan bidang ilmu yang meneliti bagaimana menyampaikan *sentiment* atau opini yang diungkapkan di dalam teks, seperti melakukan klasifikasi *sentiment positive* dan *sentiment negative* dari pengguna. Oleh karena itu, dibuatlah *Sentiment Analysis* Menggunakan Metode *Text Mining* dan *Random Forest* Untuk Klasifikasi Keluhan Pengguna (Studi Kasus: Aplikasi *KAI Access*) yang di *hosting* menggunakan *web apps streamlit* pada <https://kai-sen.streamlit.app/>. Penelitian ini dibuat menggunakan metode pengembangan *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) agar dapat memenuhi kebutuhan implementasi pada kasus Ilmu Data (*Data Science*).

Kata kunci: *KAI Access, Sentiment Analysis, Text Mining, Google Play Store, Random Forest, CRISP-DM, Streamlit.*



DAFTAR ISI

LEMBAR PERBAIKAN	i
LEMBAR BIMBINGAN	ii
LEMBAR PERNYATAAN	iii
LEMBAR PENGESAHAN	iv
LEMBAR PENGUJI SKRIPSI.....	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	5
1.6 Metodologi Penelitian	5
1.6.1 Metode Pengumpulan Data.....	5
1.6.2 Metode Implementasi	6
1.7 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI	9
2.1 <i>Data Mining</i>	9
2.2 <i>Python</i>	9
2.3 <i>Jupyter Lab</i>	9
2.4 <i>Sentiment Analysis</i>	10
2.5 <i>CRISP-DM</i>	10
2.6 <i>Text Mining</i>	11

2.7	<i>Text Preprocessing</i>	11
2.8	Pelabelan Data (<i>Labeling</i>).....	12
2.9	Pembobotan Kata (<i>TF-IDF</i>).....	12
2.10	<i>Random Forest</i>	13
2.11	<i>Confusion Matrix</i>	15
2.12	<i>Streamlit</i>	16
2.13	<i>Cloud Database (Data Space)</i>	17
BAB III ANALISIS DAN RANCANGAN SISTEM.....		19
3.1	<i>Flowchart CRISP-DM</i>	19
3.2	<i>CRISP-DM</i>	19
3.2.1	Business Understanding.....	19
3.2.2	Data Understanding.....	20
3.2.3	Data Preparation.....	22
3.2.3.1	Data Checking.....	22
3.2.3.2	Data Cleaning.....	23
3.2.3.3	Data Normalize.....	25
3.2.3.4	Words Removal.....	26
3.2.3.5	Tokenizing.....	27
3.2.4	Modeling.....	31
3.2.5	Evaluation.....	44
3.2.6	Deployment.....	47
3.3	<i>Use Case Diagram</i>	47
3.4	<i>Activity Diagram</i>	48
3.4.1	Activity Diagram Register.....	48
3.4.2	Activity Diagram Home (<i>Admin</i>).....	49
3.4.3	Activity Diagram Sentiment Analysis (<i>Admin</i>).....	50
3.4.4	Activity Diagram Sentiment Predictor (<i>User</i>).....	50
3.4.5	Activity Diagram Report (<i>Admin</i>).....	51
3.4.6	Activity Diagram Account Management (<i>Admin</i>).....	52
3.4.7	Activity Diagram Account Management (<i>User</i>).....	53
3.5	<i>Cloud Database</i>	54
3.6	Perancangan Tampilan.....	54
3.6.1	Halaman <i>Login dan Register</i>	54

3.6.2	Halaman <i>Admin</i>	55
3.6.3	Halaman <i>User</i>	56
BAB IV IMPLEMENTASI HASIL		57
4.1	Spesifikasi Analisis Dan Implementasi	57
4.2	<i>Requirements</i> Pada <i>Web Apps Streamlit</i>	57
4.3	Implementasi <i>Web Apps Streamlit</i>	58
4.3.1	Halaman <i>Login</i>	58
4.3.2	Halaman <i>Register</i>	59
4.3.3	Halaman <i>Dashboard Home (Admin)</i>	60
4.3.4	Halaman <i>Sentiment Analysis (Admin)</i>	60
4.3.5	Halaman <i>Sentiment Predictor (User)</i>	61
4.3.6	Halaman <i>Report (Admin)</i>	62
4.3.7	Halaman <i>Account Management (Admin)</i>	62
4.3.8	Halaman <i>Account Management (User)</i>	63
BAB V PENUTUP		65
5.1	Kesimpulan.....	65
5.2	Saran	66
DAFTAR PUSTAKA.....		67

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix.....	15
Tabel 3. 1 Rumus Menghitung TF-IDF	33
Tabel 3. 2 Sample Dokumen.....	33
Tabel 3. 3 Menghitung TF pada tiap dokumen.....	34
Tabel 3. 4 Menghitung IDF	34
Tabel 3. 5 Menghitung TF-IDF	34
Tabel 3. 6 Menghitung Entropy “layan”	35
Tabel 3. 7 Menghitung Entropy “mudah”	35
Tabel 3. 8 Menghitung Entropy “error”	35
Tabel 3. 9 Menghitung Entropy “kai”	35
Tabel 3. 10 Menghitung Entropy “beli”	36
Tabel 3. 11 Menghitung Entropy “tiket”	36
Tabel 3. 12 Menghitung Entropy “aplikasi”	36
Tabel 3. 13 Menghitung Entropy “susah”	36
Tabel 3. 14 Menghitung Entropy “akses”.....	36
Tabel 3. 15 Menghitung Gain(t)	37
Tabel 3. 16 Menghitung Gain untuk Root Node	37
Tabel 3. 17 Menghitung Entropy “layan”.....	39
Tabel 3. 18 Menghitung Entropy “mudah”	39
Tabel 3. 19 Menghitung Entropy “error”.....	39
Tabel 3. 20 Menghitung Entropy “kai”	39
Tabel 3. 21 Menghitung Entropy “beli”	39
Tabel 3. 22 Menghitung Entropy “tiket”	40

Tabel 3. 23 Menghitung Entropy “aplikasi”	40
Tabel 3. 24 Menghitung Entropy “akses”	40
Tabel 3. 25 Menghitung Gain (t)	40
Tabel 3. 26 Menghitung Gain untuk Node	41
Tabel 3. 27 Menghitung Gain untuk Node	42
Tabel 4. 1 Spesifikasi Analisis dan Implementasi	57



DAFTAR GAMBAR

Gambar 3. 1 Flowchart CRISP-DM	19
Gambar 3. 2 Scraping Google Play Store (Discraping 6 Juli 2023).....	20
Gambar 3. 3 Filter data	21
Gambar 3. 4 Menampilkan hasil filter.....	21
Gambar 3. 5 Meyimpan dataset ke folder lokal.....	21
Gambar 3. 6 Labeling dengan 3 orang reviewer	22
Gambar 3. 7 Menampilkan dataset	22
Gambar 3. 8 Memisahkan dan cek kondisi fitur.....	23
Gambar 3. 9 Pie Chart Keluhan Pengguna KAI Access.....	23
Gambar 3. 10 Proses Case Folding.....	24
Gambar 3. 11 Proses Cleansing	24
Gambar 3. 12 Proses Stemming	25
Gambar 3. 13 Proses Slang Word Normalization	25
Gambar 3. 14 Proses Stopword Removal.....	26
Gambar 3. 15 Proses Unwanted Word Removal	26
Gambar 3. 16 Proses Short Word Removal	27
Gambar 3. 17 Proses Split Words	27
Gambar 3. 18 Proses pembuatan Pie Chart	28
Gambar 3. 19 Pie Chart kata positive dan negative.....	28
Gambar 3. 20 Proses pembuatan Wordcloud.....	29
Gambar 3. 21 Wordcloud seluruh kata positive dan negative	29
Gambar 3. 22 Filtering 1200 Positive dan 1200 Negative	30
Gambar 3. 23 Data uji sebanyak 2400.....	30

Gambar 3. 24 Proses TF-IDF.....	31
Gambar 3. 25 Proses Split data latih dan uji.....	32
Gambar 3. 26 Proses Modeling dengan Random Forest Classifier.....	32
Gambar 3. 27 Menghitung Entropy Total.....	35
Gambar 3. 28 Hasil perhitungan Node “susah”.....	38
Gambar 3. 29 Menghitung Entropy Total (Kedua).....	38
Gambar 3. 30 Hasil perhitungan Node “error”.....	41
Gambar 3. 31 Menghitung Entropy Total (Ketiga).....	42
Gambar 3. 32 Hasil perhitungan Node.....	43
Gambar 3. 33 Cara Kerja Random Forest.....	44
Gambar 3. 34 Proses Classification Report.....	45
Gambar 3. 35 Proses Confusion Matrix.....	46
Gambar 3. 36 Confusion Matrix.....	46
Gambar 3. 37 Standar Klasifikasi.....	47
Gambar 3. 38 Use Case Diagram pada web apps Streamlit.....	48
Gambar 3. 39 Activity Diagram Register.....	49
Gambar 3. 40 Activity Diagram Home (Admin dan User).....	49
Gambar 3. 41 Activity Diagram Sentiment Analysis (Admin).....	50
Gambar 3. 42 Activity Diagram Sentiment Predictor (Admin dan User).....	51
Gambar 3. 43 Activity Diagram Report (Admin).....	51
Gambar 3. 44 Activity Diagram Account Management (Admin) - Register.....	52
Gambar 3. 45 Activity Diagram Account Management (Admin) - Update.....	52
Gambar 3. 46 Activity Diagram Account Management (Admin) - Delete.....	53
Gambar 3. 47 Activity Diagram Account Management (User).....	53

Gambar 3. 48 Cloud Database untuk Pengelolaan Pengguna	54
Gambar 3. 49 Rancangan Halaman Login.....	55
Gambar 3. 50 Rancangan Halaman Register.....	55
Gambar 3. 51 Rancangan Dashboard Admin	56
Gambar 3. 52 Rancangan Dashboard User.....	56
Gambar 4. 1 Requirements pada Web Apps Streamlit.....	58
Gambar 4. 2 Halaman Login	59
Gambar 4. 3 Halaman Register.....	59
Gambar 4. 4 Halaman Dashboard Home (Admin dan User).....	60
Gambar 4. 5 Halaman Sentiment Analysis (Admin)	60
Gambar 4. 6 Halaman Sentiment Predictor (User).....	61
Gambar 4. 7 Contoh hasil Sentiment Predictor	61
Gambar 4. 8 Halaman Report (Admin)	62
Gambar 4. 9 Halaman Account Management (Admin) - Register	63
Gambar 4. 10 Halaman Account Management (Admin) - Update	63
Gambar 4. 11 Halaman Account Management (Admin) - Delete	63
Gambar 4. 12 Halaman Account Management (User)	64



TEKNOLOGI INFORMASI

UNIVERSITAS DARMA PERSADA