

**APLIKASI ANALISIS SENTIMEN BERBASIS
WEBSITE PADA ULASAN *E-COMMERCE* SAYUR
MENGUNAKAN *NAÏVE BAYES* DENGAN *CHI-
SQUARE FEATURE SELECTION***

SKRIPSI

Oleh:

Aditya Alma Krisna

201810225331



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA**

2024



LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Aplikasi Analisis Sentimen Berbasis *Website* Pada Ulasan *E-Commerce* Sayur Menggunakan *Naïve Bayes* Dengan *Chi-Square Feature Selection*

Nama Mahasiswa : Aditya Alma Krisna

Nomor Pokok Mahasiswa : 201810225331

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skrips : 25 Juli 2024

Jakarta, 2 Agustus 2024

MENYETUJUI,

Pembimbing I


(Tri Dharma Putra, ST., M.Sc.)

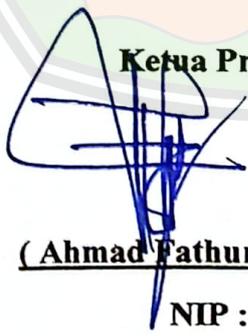
NIDN : 0302117101

Pembimbing II


(Andy Achmad S.T., M.T.)

NIDN : 0317057204

Ketua Program Studi


(Ahmad Fathurrozi, S.E., MMSI)

NIP : 2012786

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya

2024

LEMBAR PENGESAHAN

Judul Tugas Akhir : Aplikasi Analisis Sentimen Berbasis Website Pada Ulasan E-Commerce Sayur Menggunakan Naïve Bayes Dengan Chi-Square Feature Selection

Nama Mahasiswa : Aditya Alma Krisna

Nomor Pokok Mahasiswa : 201810225331

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Tugas akhir : 25 Juli 2024

Jakarta, 2 Agustus 2024

MENGESAHKAN,

Ketua Tim Penguji : Mokhammad Hadi Prayitno, S.Kom.,
M.Kom.

NIDN : 0430087003

Penguji I : Muhammad Yasir, S.Si., M.Kom.

NIDN : 0317129002

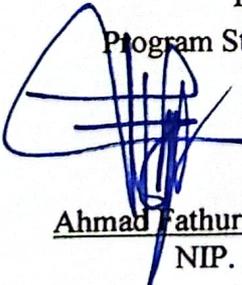
Penguji II : Tri Dharma Putra, ST., M.Sc.

NIDN : 0302117101

MENGETAHUI,

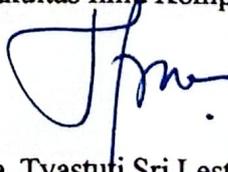
Ketua

Program Studi Informatika


Ahmad Fathurrozi, S.E., M.M.S.I
NIP. 2012486

Dekan

Fakultas Ilmu Komputer


Dr. Dra. Tyastuti Sri Lestari, M.M
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Aditya Alma Krisna
NPM : 201810225331
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Aplikasi Analisis Sentimen Berbasis *Website* Pada Ulasan *E-Commerce* Sayur Menggunakan *Naïve Bayes* Dengan *Chi-Square Feature Selection*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan daripihak manapun.

Jakarta, 13 Juni 2024

Penulis



Aditya Alma Krisna

ABSTRAK

Aditya Alma Krisna. 2024. Aplikasi Analisis Sentimen Berbasis *Website* Pada Ulasan *E-Commerce* Sayur Menggunakan *Naïve Bayes* Dengan *Chi-Square Feature Selection*

E-commerce sayur seperti Happy Fresh, Sayurbox, dan Segari memungkinkan pengguna untuk belanja sayur dan buah tanpa antrai, dengan pengiriman langsung ke lokasi pilihan, serta pilihan produk yang luas termasuk bayam, wortel, dan mangga. *E-commerce* sayur ini populer di kalangan pengguna, seperti yang terlihat dari jumlah unduhan lebih dari 1 juta di Google Play Store. Meskipun demikian, setiap platform memiliki kelebihan dan kelemahan tersendiri, seperti fitur kemitraan yang hanya ada di Segari. Tanggapan pengguna terhadap aplikasi ini dapat bervariasi dari kepuasan hingga kekecewaan, yang sering kali tercermin dalam ulasan mereka di platform ini. Ulasan tersebut bisa berdampak signifikan pada preferensi belanja masyarakat, terutama jika perusahaan tidak merespons dengan cepat untuk memperbaiki layanan berdasarkan umpan balik pengguna. Maka penelitian ini akan membuat sistem untuk melakukan analisis sentimen untuk memahami ulasan pengguna terhadap aplikasi secara lebih mendalam terhadap aplikasi *e-commerce* sayur menggunakan *naïve bayes classifier* (NBC) dan *chi-square feature selection* dengan menggunakan 3 platform aplikasi yaitu Happy Fresh, Sayurbox, dan Segari serta akan dilakukan scrapping ulasan pengguna aplikasi pada *Google Play Store*. Hasil penelitian ini menunjukkan bahwa pengujian Confusion matrix menunjukkan jumlah prediksi yang benar dan salah untuk masing-masing kategori label (positif, netral, negatif), menghasilkan akurasi 81.3%, precision 81.11%, recall 81.3%, dan F1-Score 80.53%. Hasil ini menunjukkan bahwa algoritma Naive Bayes dengan Chi-Square memiliki performa yang cukup baik dalam mengklasifikasikan ulasan pengguna aplikasi "Happy Fresh". Model ini mampu melakukan analisis sentimen dengan cukup akurat, sehingga dapat diaplikasikan secara efektif untuk analisis sentimen pada data ulasan aplikasi.

Kata kunci: Aplikasi, Analisis Sentimen, *Chi-Square*, *E-Commerce*, Naïve Bayes

ABSTRACT

Aditya Alma Krisna. 2024. Website-Based Sentiment Analysis Application on E-Commerce Vegetable Reviews Using Naïve Bayes with Chi-Square Feature Selection

Vegetable e-commerce such as Happy Fresh, Sayurbox, and Segari allow users to shop for vegetables and fruits without queuing, with direct delivery to the location of choice, and a wide selection of products including spinach, carrots, and mangoes. These vegetable e-commerce platforms are popular with users, as seen by the number of downloads of over 1 million on the Google Play Store. Nonetheless, each platform has its own strengths and weaknesses, such as the partnership feature that is only available on Segari. Users' responses to these apps can vary from satisfaction to disappointment, which is often reflected in their reviews on these platforms. Such reviews can have a significant impact on people's shopping preferences, especially if companies do not respond quickly to improve services based on user feedback. So this research will create a system to perform sentiment analysis to understand user reviews of applications in more depth on vegetable e-commerce applications using naïve bayes classifier (NBC) and chi-square feature selection using 3 application platforms namely Happy Fresh, Sayurbox, and Segari and will be scrapping application user reviews on the Google Play Store. The results of this study indicate that Confusion matrix testing shows the number of correct and incorrect predictions for each label category (positive, neutral, negative), resulting in an accuracy of 81.3%, precision 81.11%, recall 81.3%, and F1-Score 80.53%. These results show that the Naive Bayes algorithm with Chi-Square performs quite well in classifying user reviews of the "Happy Fresh" app. This model is able to perform sentiment analysis quite accurately, so it can be applied effectively for sentiment analysis on application review data.

Keywords: *Application, Sentiment Analysis, Chi-Square, E-Commerce, Naïve Bayes*



LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Aditya Alma Krisna
NPM : 201810225331
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

Aplikasi Analisis Sentimen Berbasis Website Pada Ulasan E-Commerce Sayur Menggunakan Naïve Bayes Dengan Chi-Square Feature Selection

Beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 13 Juni 2024

Yang Menyatakan



Aditya Alma Krisna

KATA PENGANTAR

Puji syukur kepada Allah SWT, Tuhan Yang Maha Esa sehingga dapat menyelesaikan proposal skripsi dengan judul “Aplikasi Analisis Sentimen Berbasis Website Pada Ulasan *E-commerce* Sayur Menggunakan *Naïve Bayes* dengan *Chi-Square Feature Selection*”. Skripsi ini disusun sebagai salah satu tahapan dalam menyelesaikan skripsi pada Prodi Informatika, Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.

Oleh karena itu, pada kesempatan ini saya ingin mengucapkan terima kasih kepada semua pihak yang telah mendukung serta membantu saya selama proses penyusunan tugas akhir skripsi, kepada:

1. Orang Tua saya Bapak Parwanto S.H selaku Ayah dan Ibu Ibu Evi Zulfitri Nugraheni selaku Ibu dan Saudara-saudara saya atas do'a, bimbingan serta kasih sayangnya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M selaku Dekan Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.
3. Bapak Acmad Fathurrozi, S.E., MMSI selaku Ketua Program Studi.
4. Bapak Wowon Priatna selaku Dosen Pembimbing Akademik Saya.
5. Bapak Tri Dharma Putra, ST., M.Sc. selaku Dosen Pembimbing I yang selalu memberikan arahan dan masukan terhadap penelitian ini.
6. Bapak Andy Achmad S.T.,M.T. selaku Dosen Pembimbing II atas masukan terhadap penelitian ini.
7. Segenap Dosen Program Studi Ilmu Komputer, yang telah memberikan ilmu pengetahuan selama mengikuti perkuliahan.
8. Rekan-rekan seperjuangan yang terus memberikan dukungan semangat demi terselesainya tugas ini.

Saya menyadari bahwa laporan skripsi ini tidak luput dari berbagai kekurangan. Oleh karena itu, semua kritik dan saran yang bersifat memperbaiki demi kesempurnaan sangat diharapkan.

Jakarta, 2 Agustus 2024



Aditya Alma Krisna

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASIKARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	7
1.3 Rumusan Masalah	8
1.4 Tujuan.....	8
1.5 Manfaat Penelitian.....	9
1.6 Batasan Masalah.....	9
1.7 Sistematika Penulisan.....	10
BAB II TINJAUAN PUSTAKA.....	12
2.1 Penelitian Terdahulu.....	12
2.2 <i>E-Commerce</i> Sayur.....	16
2.2.1 <i>Happy Fresh</i> (https://www.hapyfresh.id/)	17
2.2.2 <i>Sayurbox</i> (https://www.sayurbox.com/)	18

2.2.3	Segari (https://segari.id/products)	20
2.3	Analisis Sentimen	21
2.4	CRISP-DM	22
2.5	<i>Preprocessing</i>	24
2.6	<i>Chi Square</i>	26
2.7	<i>Naïve Bayes</i>	28
2.8	<i>Flowchart</i>	29
2.9	<i>Fishbone</i>	31
2.10	<i>Waterfall</i>	32
2.11	<i>Unified Modelling Language</i>	34
2.12	<i>Website</i>	38
2.13	<i>Python</i>	39
2.14	<i>Flask</i>	41
2.15	<i>XAMPP</i>	42
2.16	MySQL	44
2.17	<i>Confusion Matrix</i>	45
2.18	<i>Blackbox Testing</i>	48
BAB III METODOLOGI PENELITIAN		49
3.1	Tempat dan Waktu Penelitian	49
3.2	Kerangka Pikir Penelitian	50
3.3	Metode Pengumpulan Data	55
3.4	Metode Analisis	56
3.4.1	Analisis Sistem Berjalan	56
3.4.2	Analisis Masalah	58
3.4.3	Analisis Sitem Usulan	60
3.5	Analisis Kebutuhan	63
3.5.1	Analisis Kebutuhan Fungsional	63

3.5.2	Analisis Kebutuhan Non-Fungsional	63
3.5.3	Analisis Kebutuhan Perangkat	64
BAB IV HASIL DAN PEMBAHASAN.....		66
4.1	<i>Business Understanding</i>	66
4.2	<i>Data Understanding</i>	67
4.3	<i>Data Preparation</i>	75
4.3.1	<i>Preprocessing</i>	75
4.3.2	<i>Encoder</i>	80
4.3.3	<i>Split Data</i>	81
4.4	<i>Modelling</i>	83
4.4.1	<i>Chi Square Feature Selection</i>	84
4.4.2	<i>Naïve Bayes</i>	94
4.4.3	<i>Klasifikasi Data Test</i>	98
4.5	<i>Evaluasi</i>	100
4.6	<i>Deployment</i>	103
4.6.1	<i>Perancangan Sistem</i>	103
4.6.2	<i>Implementasi</i>	121
4.6.3	<i>Pengujian Sistem</i>	129
4.7	<i>Pembahasan</i>	130
BAB V KESIMPULAN.....		131
5.1	<i>Kesimpulan</i>	131
5.2	<i>Saran</i>	132
DAFTAR PUSTAKA		133
LAMPIRAN.....		145

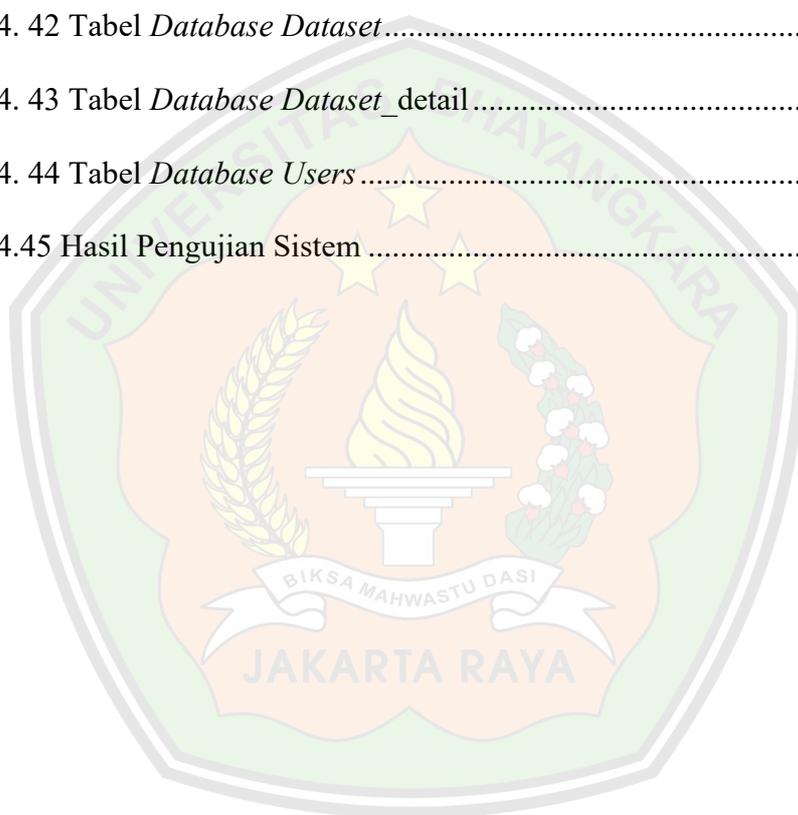


DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	12
Tabel 2.2 Contoh <i>Case Folding</i>	25
Tabel 2.3 Contoh <i>Tokenizing</i>	25
Tabel 2.4 Contoh <i>Stopword Removal</i>	25
Tabel 2.5 Contoh <i>Stemming</i>	26
Tabel 2.6 Notasi <i>Flowchart</i>	30
Tabel 2.7 Notasi <i>Use Case Diagram</i>	35
Tabel 2.8 Notasi <i>Activity Diagram</i>	36
Tabel 2.9 Notasi <i>Class Diagram</i>	36
Tabel 2.10 Notasi <i>Sequence Diagram</i>	37
Tabel 2.11 <i>Confusion Matrix</i>	46
Tabel 3.1 <i>Sample Dataset Happy Fresh</i> (https://www.happyfresh.id/)	52
Tabel 3.2 <i>Sample Dataset Sayurbox</i> (https://www.sayurbox.com/)	53
Tabel 3.3 <i>Sample Dataset Segari</i> (https://segari.id/products)	53
Tabel 3. 4 Tabel Sistem Berjalan	57
Tabel 3.5 Tabel Sistem Usulan	62
Tabel 4. 1 Keterangan <i>Rating Label</i>	68
Tabel 4. 2 <i>Dataset Happy Fresh</i>	70
Tabel 4. 3 <i>Dataset Sayurbox</i>	71
Tabel 4. 4 <i>Dataset Segari</i>	72
Tabel 4. 5 <i>Dataset Sampel</i>	73
Tabel 4. 6 <i>Case Folding</i>	77
Tabel 4. 7 <i>Cleaning</i>	77

Tabel 4. 8 <i>Tokenizing</i>	78
Tabel 4. 9 <i>Stopwords</i>	78
Tabel 4. 10 <i>Stemming</i>	79
Tabel 4. 11 <i>Encoder</i>	80
Tabel 4. 12 Hasil Encoding	80
Tabel 4. 13 <i>Data Train</i>	81
Tabel 4. 14 <i>Data Test</i>	81
Tabel 4. 15 TF <i>Data Train</i>	82
Tabel 4. 16 TF <i>Data Train</i>	82
Tabel 4. 17 Hasil <i>One Hot Encoding</i>	86
Tabel 4. 18 rata-rata <i>one-hot encoder</i>	87
Tabel 4. 19 Total setiap TF pada setiap features	87
Tabel 4. 20 Hasil <i>Observed Features</i>	89
Tabel 4. 21 Nilai <i>Expected</i> setiap features	90
Tabel 4. 22 Hasil <i>Chi-Score</i> setiap Features	91
Tabel 4. 23 Nilai p-value setiap features	92
Tabel 4. 24 Hasil <i>sorting ascending p-values</i>	93
Tabel 4. 25 <i>Data train features selection</i>	93
Tabel 4. 26 <i>Data test features selection</i>	94
Tabel 4. 27 Total kemunculan kelas	96
Tabel 4. 28 Nilai Probabilitas kelas	96
Tabel 4. 29 Nilai probabilitas <i>feature fresh</i> terhadap kelas	97
Tabel 4. 30 Probaibilitas setiap features terhadap kelas	97
Tabel 4. 31 Hasil $P(t 0)$ pada data test ke-1	98
Tabel 4. 32 Hasil $P(t 1)$ pada data test ke-1	98
Tabel 4. 33 Hasil $P(t 2)$ pada data test ke-1	99

Tabel 4. 34 Hasil prediksi seluruh <i>data test</i>	99
Tabel 4. 35 <i>Confusion Matrix</i> 3 Kelas	100
Tabel 4. 36 <i>Confusion Matrix</i> Berdasarkan hasil Prediksi Data Test	101
Tabel 4. 37 Hasil Evaluasi Menggunakan <i>Naive Bayes (Chi Square)</i>	101
Tabel 4. 38 <i>Confusion Matrix</i> Berdasarkan hasil Prediksi <i>Data Test</i>	102
Tabel 4. 39 Hasil Evaluasi Menggunakan <i>Naive Bayes (Chi Square)</i>	102
Tabel 4. 40 Penjelasan Aktor	104
Tabel 4. 41 Penjelasan <i>Use Case</i>	105
Tabel 4. 42 Tabel <i>Database Dataset</i>	114
Tabel 4. 43 Tabel <i>Database Dataset_detail</i>	115
Tabel 4. 44 Tabel <i>Database Users</i>	116
Tabel 4.45 Hasil Pengujian Sistem	129



DAFTAR GAMBAR

Gambar 2.1 Logo <i>Happy Fresh</i>	17
Gambar 2.2 Tampilan Aplikasi <i>Happy Fresh</i>	18
Gambar 2.3 Logo Sayurbox	19
Gambar 2.4 Tampilan Aplikasi Sayurbox.....	19
Gambar 2.5 Logo Segari	20
Gambar 2.6 Tampilan Aplikasi Segari.....	21
Gambar 2.7 Tahapan CRISP-DM	23
Gambar 2.8 Logo <i>Happy Fresh</i>	32
Gambar 2.9 Logo <i>Happy Fresh</i>	33
Gambar 3.1 Kerangka Pikir.....	50
Gambar 3. 2 Alur Sistem Berjalan	57
Gambar 3. 3 Diagram <i>Fishbone</i> Analisis Sentimen Ulasan <i>E-Commerce</i> Sayur	59
Gambar 3.4 Alur Sistem Usulan	61
Gambar 4. 1 <i>Flowchart Model</i>	67
Gambar 4. 2 <i>Text Preprocessing</i>	76
Gambar 4. 3 <i>Chi Square Skelarn</i>	85
Gambar 4. 4 Alur Multinomial <i>Naive Bayes</i>	95
Gambar 4. 5 <i>Use Case Diagram</i>	104
Gambar 4. 6 <i>Activity Diagram Login</i>	106
Gambar 4. 7 <i>Activity Diagram Input Dataset</i>	107
Gambar 4. 8 <i>Activity Diagram Hasil Pelabelan</i>	108

Gambar 4. 9 <i>Activity Diagram Text Preprocessing</i>	109
Gambar 4. 10 <i>Activity Diagram Analisis Sentimen</i>	110
Gambar 4. 11 <i>Sequence Diagram Login</i>	111
Gambar 4. 12 <i>Sequence Diagram Input Dataset</i>	111
Gambar 4. 13 <i>Sequence Diagram Hasil Pelabelan</i>	112
Gambar 4. 14 <i>Sequence Diagram Text Preprocessing</i>	112
Gambar 4. 15 <i>Sequence Diagram Analisis Sentimen</i>	113
Gambar 4. 16 <i>Class Diagram</i>	114
Gambar 4. 17 <i>Login</i>	117
Gambar 4. 18 Rancangan Halaman <i>Home</i>	117
Gambar 4. 19 Rancangan Halaman <i>User</i>	118
Gambar 4. 20 Rancangan Halaman <i>Dataset</i>	119
Gambar 4. 21 Rancangan Halaman <i>Detail Dataset</i>	119
Gambar 4. 22 Rancangan Halaman <i>Analisis Sentimen</i>	120
Gambar 4. 23 <i>Screenshot Database PhpMyAdmin</i>	121
Gambar 4. 24 <i>Database Tabel Dataset</i>	122
Gambar 4. 25 <i>Screenshot Database Tabel Dataset_detail</i>	122
Gambar 4. 26 <i>Screenshot Database Tabel Users</i>	123
Gambar 4. 27 Tampilan Halaman <i>Login</i>	123
Gambar 4. 28 Tampilan Halaman <i>Home</i>	124
Gambar 4. 29 Tampilan Halaman <i>Pengguna</i>	124
Gambar 4.30 Tampilan Halaman <i>Dataset</i>	125
Gambar 4.31 Tampilan Form Upload <i>Dataset</i>	125

Gambar 4.32 Tampilan Halaman Detail *Dataset*..... 126

Gambar 4.33 Tampilan Halaman Analisis Sentimen..... 127

Gambar 4.34 Tampilan Hasil Analisis dalam *Naïve Bayes* dengan *Chi-Square* 128





DAFTAR LAMPIRAN

Lampiran 1.1 Plagiarism Checker.....	145
Lampiran 1.2 Kartu Bimbingan	146
Lampiran 1.3 Surat Rekomendasi Pembimbing.....	148
Lampiran 1.4 Biodata Mahasiswa.....	149
Lampiran 1.5 Database dari google playstore.....	150



