

**PERBANDINGAN *SUPPORT VECTOR MACHINE* DAN *K-NEAREST NEIGHBOR* DALAM ANALISIS SENTIMEN
KEBIJAKAN PPN 12% PADA BARANG DAN JASA MEWAH**

TUGAS AKHIR



Muhammad Rizki Fachriza Bailey

NIM. 2101020071

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG**

2026

**PERBANDINGAN *SUPPORT VECTOR MACHINE* DAN *K-NEAREST NEIGHBOR* DALAM ANALISIS SENTIMEN
KEBIJAKAN PPN 12% PADA BARANG DAN JASA MEWAH**

Tugas Akhir

Skema Skripsi

*Tugas Akhir diajukan sebagai salah satu persyaratan memperoleh gelar Sarjana
Teknik pada Program Studi Teknik Informatika*



Muhammad Rizki Fachriza Bailey

NIM. 2101020071

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG**

2026

PERNYATAAN MENGENAI TUGAS AKHIR DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tugas akhir berjudul “Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah” adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan melanggar peraturan yang berlaku dalam karya tulis dan hak intelektual, maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Maritim Raja Ali Haji dan menerima sanksi lainnya sesuai dengan peraturan yang berlaku. Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Maritim Raja Ali Haji.

Tanjungpinang, 14 Januari 2026



Muhammad Rizki Fachriza Bailey

NIM. 2101020071



© Hak Cipta Milik Universitas Maritim Raja Ali Haji, Tahun 2026

Hak Cipta Dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah; dan pengutipan tersebut tidak merugikan kepentingan Universitas Maritim Raja Ali Haji.

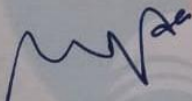
Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Universitas Maritim Raja Ali Haji.

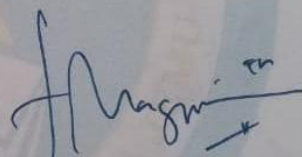


HALAMAN PERSETUJUAN

Judul : Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah
Nama : Muhammad Rizki Fachriza Bailey
NIM : 2101020071
Program Studi : Teknik Informatika
Tanggal Persetujuan : 11 Desember 2025

Disetujui oleh,
Komisi Pembimbing


Nurfalinda, S.T., M.Cs
(Ketua)


Fortia Magfira, M.Kom
(Anggota)

HALAMAN PENGESAHAN

Judul : Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah

Nama : Muhammad Rizki Fachriza Bailey

NIM : 2101020071

Telah berhasil dipertahankan di hadapan Komisi Penguji sebagai bagian persyaratan yang diperlukan dalam memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Jurusan Teknik Elektro dan Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

KOMISI PENGUJI

Ketua	: Nerfita Nikentari, S.T., M.Cs. Ph.D	(.....)
Anggota I	: Nolan Efranda, M.Kom	(.....)
Anggota II	: Novrizal Fattah Fahmitra, S.Kom., M.Kom	(.....)
Anggota III	: Nurfalinda, S.T., M.Cs	(.....)
Anggota IV	: Fortia Magfira, M.Kom	(.....)

Mengetahui,
Ketua Jurusan Teknik Elektro dan Informatika

Hollanda Arief Kusuma, S.IK., M.Si
NIP. 198004012019031016

Tanggal Ujian: 22 Desember 2025

Tanggal Lulus:

PRAKATA

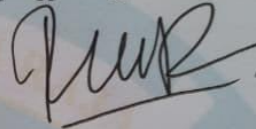
Puji dan syukur dipanjatkan kepada Allah *subhanahu wa ta'ala* atas segala karunia-Nya sehingga laporan tugas akhir ini berhasil diselesaikan. Tema yang dipilih dalam tugas akhir yang akan dilaksanakan pada bulan Desember 2025 ini ialah Analisis Sentimen, dengan judul **“Perbandingan *Support Vector Machine* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Kebijakan PPN 12% Pada Barang dan Jasa Mewah”**.

Dalam penulisan skripsi ini, penulis menyadari bahwa banyak pihak yang telah memberikan dukungan, bantuan serta bimbingan, baik secara langsung atau tidak langsung. Oleh karena itu, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

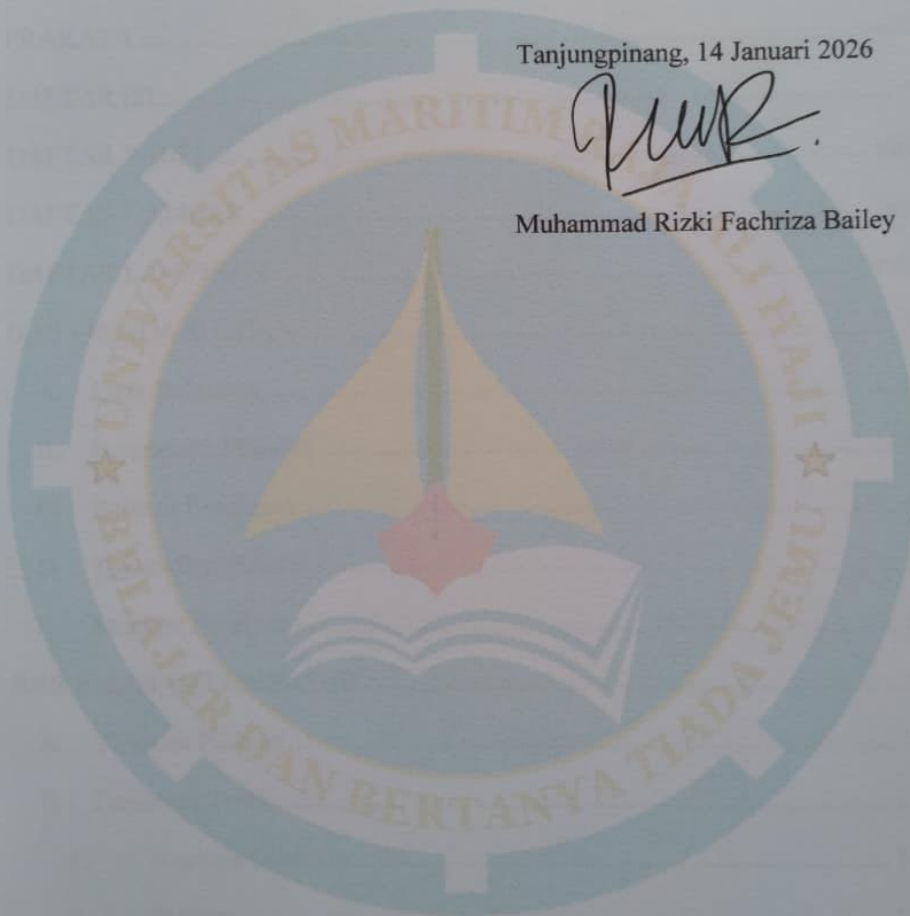
1. Ibu Martaleli Bettiza, S.Si., M.Sc. selaku Dekan Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.
2. Bapak Hollanda Arief Kusuma, S.IK., M.Si. selaku Ketua Jurusan Teknik Elektro dan Informatika.
3. Bapak Feri Irawan, S.Kom., M.Kom selaku Plt. Koordinator Program Studi Teknik Informatika.
4. Ibu Nurfalinda, S.T., M.Cs dan Ibu Fortia Magfira, M.Kom selaku Komisi Pembimbing, yang telah memberi saran, pendapat serta bimbingan kepada penulis.
5. Ibu Nerfita Nikentari, S.T., M.Cs. Ph.D., Bapak Nolan Efranda, M.Kom., dan Bapak Novrizal Fattah Fahmitra, S.Kom., M.Kom., selaku Komisi Penguji yang telah memberikan saran dan pendapat terkait penelitian yang penulis lakukan.
6. Staff Administrasi Fakultas Teknik dan Teknologi Kemaritiman yang selalu membantu kelancaran proses administrasi bagi penulis selama menempuh pendidikan di Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.
7. Kedua orang tua, kakak dan rekan-rekan sejawat Teknik Informatika yang telah mendukung dan mendoakan kelancaran penelitian penulis.

Semoga laporan tugas akhir ini bermanfaat dan dapat menjadi sumber acuan atau referensi untuk penelitian-penelitian di masa depan, khususnya bagi Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

Tanjungpinang, 14 Januari 2026



Muhammad Rizki Fachriza Bailey



DAFTAR ISI

HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
ABSTRAK	v
ABSTRACT	vii
PRAKATA.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Perumusan Masalah	4
C. Batasan Penelitian	4
D. Tujuan Penelitian	5
E. Manfaat Penelitian	5
BAB II KAJIAN LITERATUR	7
A. Tinjauan Pustaka	7
B. Landasan Teori.....	11
1. Analisis Sentimen	11
2. Python	11
3. Google Colab	11
4. Indonesian Sentiment Lexicon (Inset Lexicon)	12
5. Grid Search.....	13
6. Preprocessing Data.....	14

7.	Feature Extraction	15
8.	Support Vector Machine	15
9.	K-Nearest Neighbor	18
10.	Confusion Matrix	19
BAB III METODE PENELITIAN.....		23
A.	Waktu dan Tempat Penelitian	23
B.	Jenis Penelitian.....	23
C.	Instrumen Penelitian.....	23
D.	Prosedur Penelitian	24
E.	Studi Pustaka.....	28
F.	Pengumpulan Data	29
G.	Analisis Data	38
H.	Perhitungan Manual	39
I.	Evaluasi	61
J.	Perancangan Aplikasi.....	63
BAB IV HASIL DAN PEMBAHASAN		68
A.	Preprocessing Data.....	68
B.	Pelabelan Sentimen	72
C.	Pembagian Dataset	73
D.	Pencarian Parameter Terbaik	74
E.	Model Support Vector Machine.....	75
F.	Model K-Nearest Neighbor.....	77
G.	Perbandingan Akurasi SVM dan KNN	78
H.	Visualisasi Wordcloud	80
BAB V SIMPULAN DAN SARAN		93

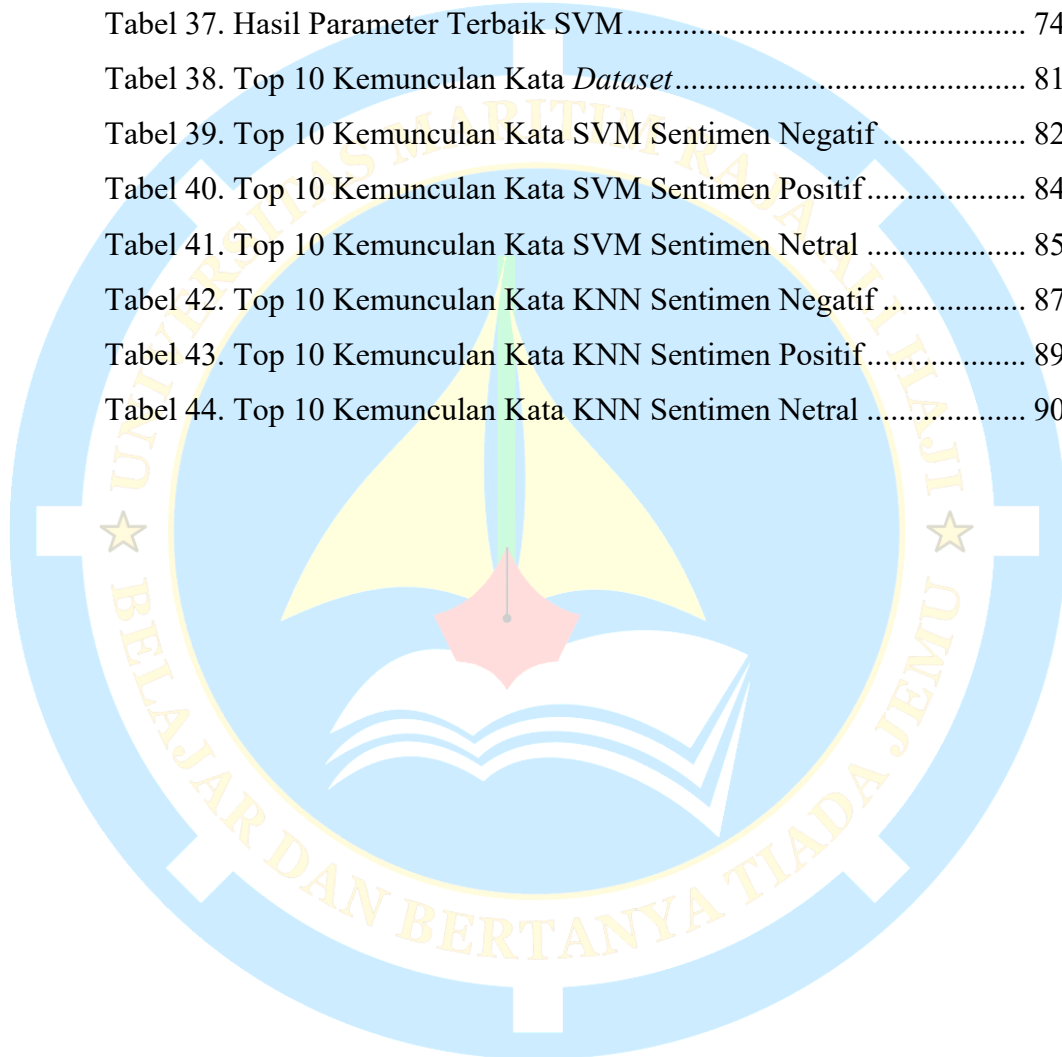
A. Simpulan	93
B. Saran.....	93
DAFTAR PUSTAKA	95
LAMPIRAN.....	102



DAFTAR TABEL

Tabel 1. <i>Confusion Matrix</i>	20
Tabel 2. Daftar Video <i>Youtube</i>	29
Tabel 3. Sampel Untuk Data Latih.....	38
Tabel 4. Sampel Untuk Data Uji	39
Tabel 5. Sampel Data Latih Proses <i>Data Cleaning</i>	39
Tabel 6. Sampel Data Uji Proses <i>Data Cleaning</i>	39
Tabel 7. Sampel Data Latih Proses <i>Case Folding</i>	40
Tabel 8. Sampel Data Uji Proses <i>Case Folding</i>	40
Tabel 9. Sampel Data Latih Proses <i>Tokenizing</i>	41
Tabel 10. Sampel Data Uji Proses <i>Tokenizing</i>	41
Tabel 11. Sampel Data Latih Proses <i>Slang Words</i>	42
Tabel 12. Sampel Data Uji Proses <i>Slang Words</i>	42
Tabel 13. Sampel Data Latih Proses <i>Stop Words Removal</i>	43
Tabel 14. Sampel Data Uji Proses <i>Stop Words Removal</i>	43
Tabel 15. Sampel Data Latih Proses <i>Stemming</i>	44
Tabel 16. Sampel Data Uji Proses <i>Stemming</i>	44
Tabel 17. Sampel Data Latih Setelah <i>Preprocessing Data</i>	45
Tabel 18. Sampel Data Uji Setelah <i>Preprocessing Data</i>	45
Tabel 19. Sampel Data Latih dengan Label	46
Tabel 20. Sampel Data Uji dengan Label	46
Tabel 21. Kemunculan Kata Data Latih.....	48
Tabel 22. Kemunculan Kata Data Uji	49
Tabel 23. Fitur Ekstraksi Data Latih	50
Tabel 24. Fitur Ekstraksi Data Uji	50
Tabel 25. Inisialisasi Parameter	52
Tabel 26. Data Latih Untuk SVM.....	52
Tabel 27. Data Uji Untuk SVM	52
Tabel 28. Data Hasil Fitur Ekstraksi Setelah Diubah	53
Tabel 29. Hasil Klasifikasi SVM	58
Tabel 30. Perbandingan Hasil Pelabelan Inset Lexicon dan SVM	58

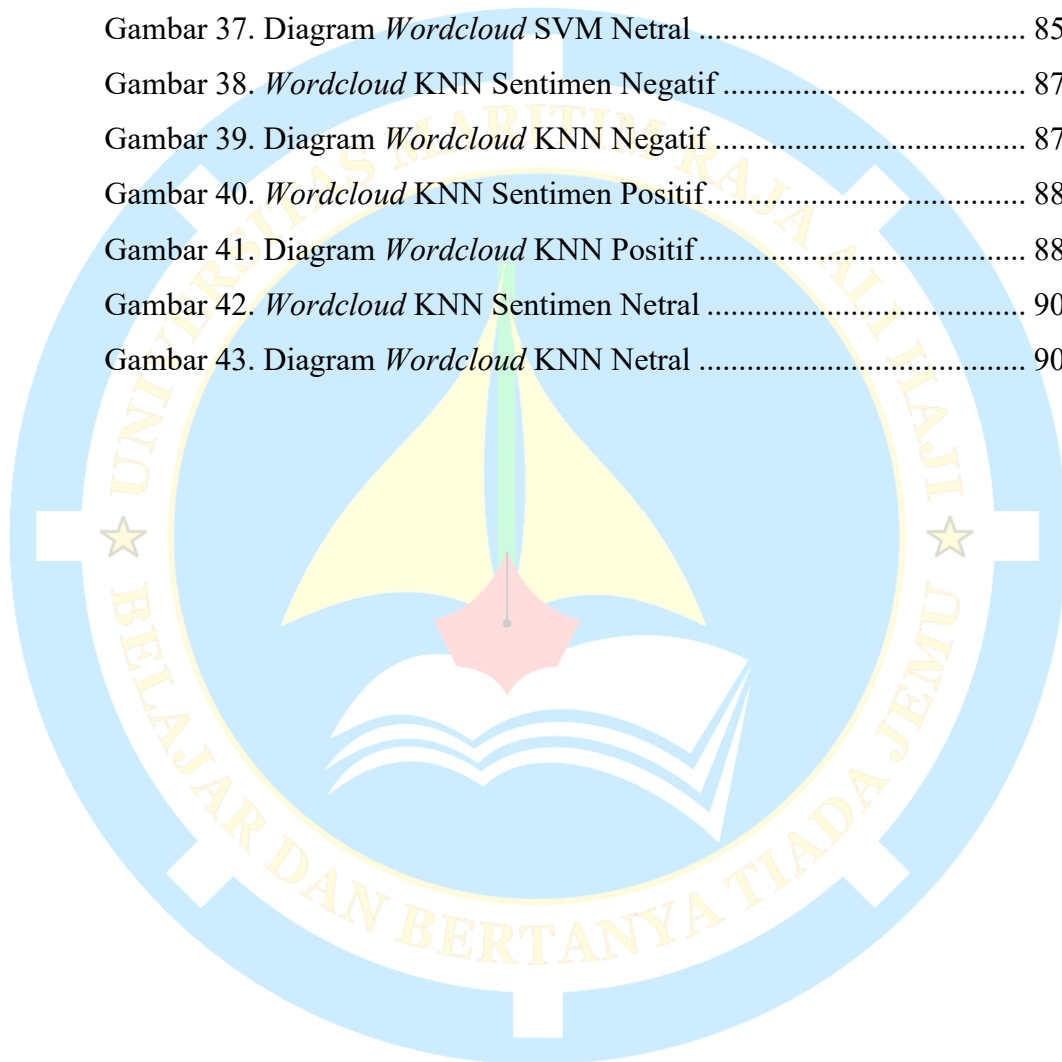
Tabel 31. Urutkan Data KNN	60
Tabel 32. Hasil Label Data Uji	61
Tabel 33. Perbandingan Hasil Pelabelan Inset Lexicon dan KNN	61
Tabel 34. <i>Confusion Matrix Support Vector Machine</i>	62
Tabel 35. <i>Confusion Matrix K-Nearest Neighbor</i>	62
Tabel 36. Parameter SVM.....	74
Tabel 37. Hasil Parameter Terbaik SVM.....	74
Tabel 38. Top 10 Kemunculan Kata <i>Dataset</i>	81
Tabel 39. Top 10 Kemunculan Kata SVM Sentimen Negatif	82
Tabel 40. Top 10 Kemunculan Kata SVM Sentimen Positif.....	84
Tabel 41. Top 10 Kemunculan Kata SVM Sentimen Netral	85
Tabel 42. Top 10 Kemunculan Kata KNN Sentimen Negatif	87
Tabel 43. Top 10 Kemunculan Kata KNN Sentimen Positif.....	89
Tabel 44. Top 10 Kemunculan Kata KNN Sentimen Netral	90



DAFTAR GAMBAR

Gambar 1. Logo <i>Python</i>	11
Gambar 2. Logo <i>Goole Colab</i>	12
Gambar 3. <i>Hyperplane Support Vector Machine</i>	16
Gambar 4. Prosedur Penelitian	25
Gambar 5. Proses Fitur Ekstraksi	47
Gambar 6. Diagram Proses SVM	51
Gambar 7. Diagram Proses KNN	59
Gambar 8. <i>Data Flow Diagram Level 0</i>	64
Gambar 9. <i>Data Flow Diagram Level 1</i>	65
Gambar 10. Desain <i>Website</i> Halaman Beranda	65
Gambar 11. Desain <i>Website</i> Halaman Analisis Sentimen	66
Gambar 12. Hasil Implementasi Halaman Beranda	67
Gambar 13. Hasil Implementasi Halaman Analisis Sentimen	67
Gambar 14. Data Sebelum <i>Data Cleaning</i> dan <i>Case Folding</i>	68
Gambar 15. Data Setelah <i>Data Cleaning</i> dan <i>Case Folding</i>	68
Gambar 16. Hasil <i>Tokenizing</i>	69
Gambar 17. Hasil <i>Slang Words</i>	70
Gambar 18. Hasil <i>Stop Word Removal</i>	71
Gambar 19. Hasil <i>Stemming</i>	71
Gambar 20. Hasil Pelabelan Sentimen	72
Gambar 21. Persentase Hasil Pelabelan Sentimen	73
Gambar 22. Pembagian Dataset	73
Gambar 23. Hasil Pencarian Parameter	75
Gambar 24. <i>Confusion Matrix</i> SVM	75
Gambar 25. Laporan Klasifikasi SVM Parameter <i>Gridsearch</i>	76
Gambar 26. <i>Confusion Matrix</i> KNN	77
Gambar 27. Rincian <i>Confusion Matrix</i> KNN	78
Gambar 28. Perbandingan Akurasi SVM dan KNN	78
Gambar 29. Perbandingan Waktu Klasifikasi Model SVM dan KNN	79
Gambar 30. <i>Wordcloud Global</i>	80

Gambar 31. Diagram Frekuensi Kata Terbanyak	80
Gambar 32. <i>Wordcloud</i> SVM Sentimen Negatif	82
Gambar 33. Diagram <i>Wordcloud</i> SVM Negatif	82
Gambar 34. <i>Wordcloud</i> SVM Sentimen Positif.....	83
Gambar 35. Diagram <i>Wordcloud</i> SVM Positif	84
Gambar 36. <i>Wordcloud</i> SVM Sentimen Netral	85
Gambar 37. Diagram <i>Wordcloud</i> SVM Netral	85
Gambar 38. <i>Wordcloud</i> KNN Sentimen Negatif	87
Gambar 39. Diagram <i>Wordcloud</i> KNN Negatif	87
Gambar 40. <i>Wordcloud</i> KNN Sentimen Positif.....	88
Gambar 41. Diagram <i>Wordcloud</i> KNN Positif.....	88
Gambar 42. <i>Wordcloud</i> KNN Sentimen Netral	90
Gambar 43. Diagram <i>Wordcloud</i> KNN Netral	90



DAFTAR LAMPIRAN

Lampiran 1. Hasil Validasi 50 Sampel Data Oleh Ahli Bahasa 102

