

SKRIPSI

**ASPECT BASED SENTIMENT ANALYSIS UNTUK
MENGIDENTIFIKASI KELUHAN PENAGIHAN PINJAMAN
ONLINE DI MEDIA SOSIAL X MENGGUNAKAN
ALGORITMA BAYESIAN NETWORK DAN C4.5**



FRASETIA ADI KUSUMA

1512622038

**PROGRAM STUDI
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA**

2026



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Frasetia Adi Kusuma
NIM : 1512622038
Fakultas/Prodi : Teknik/Pendidikan Teknik Informatika dan Komputer
Alamat email : fak.addikus@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Aspect Based Sentiment Analysis untuk Mengidentifikasi Keluhan Penagihan Pinjaman Online
di Media Sosial X Menggunakan Algoritma Bayesian Network dan C4.5.

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 8 Juni 2026

Penulis,

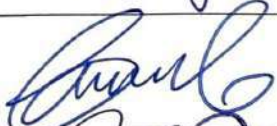
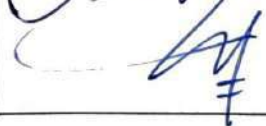
Frasetia Adi Kusuma

LEMBAR PENGESAHAN TUGAS AKHIR BENTUK: SKRIPSI

Nama Mahasiswa : Frasetia Adi Kusuma
NIM : 1512622038
Program Studi/Fakultas : Pendidikan Teknik Informatika dan Komputer / Teknik
Judul Penelitian : *Aspect Based Sentiment Analysis* untuk Mengidentifikasi
Keluhan Penagihan Pinjaman Online di Media Sosial X
Menggunakan Algoritma Bayesian Network dan C4.5.

ini telah dipertahankan dihadapan dewan penguji dan dinyatakan **Lulus** / ~~Tidak Lulus~~)*
Ujian Tugas Akhir oleh Tim Penguji pada tanggal 8 Juni 2026.

Tim Penguji,

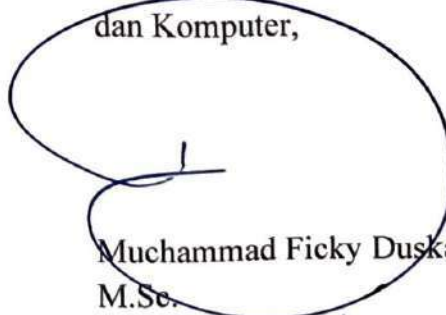
Ketua Penguji	Ressy Dwitias Sari, S.T., M.T.I. NIDN 0115098902	
Anggota Penguji (Penguji Ahli 1)	Via Tuhamah Fauziastuti, S.Si., M.Ed. NIDN 0010019110	
Anggota Penguji (Penguji Ahli 2)	Bambang Satrio, S.Pd., M.Pd. NUPTK 474777567613022	
Anggota Penguji (Pembimbing 1)	Dr. Widodo, S. Kom., M.Kom. NIDN 0025037206	
Anggota Penguji (Pembimbing 2)	Neng Ayu Herawati, S.Pd., M.T. NUPTK 125277667723015	

Mengetahui,
Dekan Fakultas Teknik



Prof. Dr. Neneng Siti Silfi Ambarwati,
S.Si., Apt., M.Si
NIDN 0029027204

Koordinator Program Studi
Pendidikan Teknik Informatika
dan Komputer,

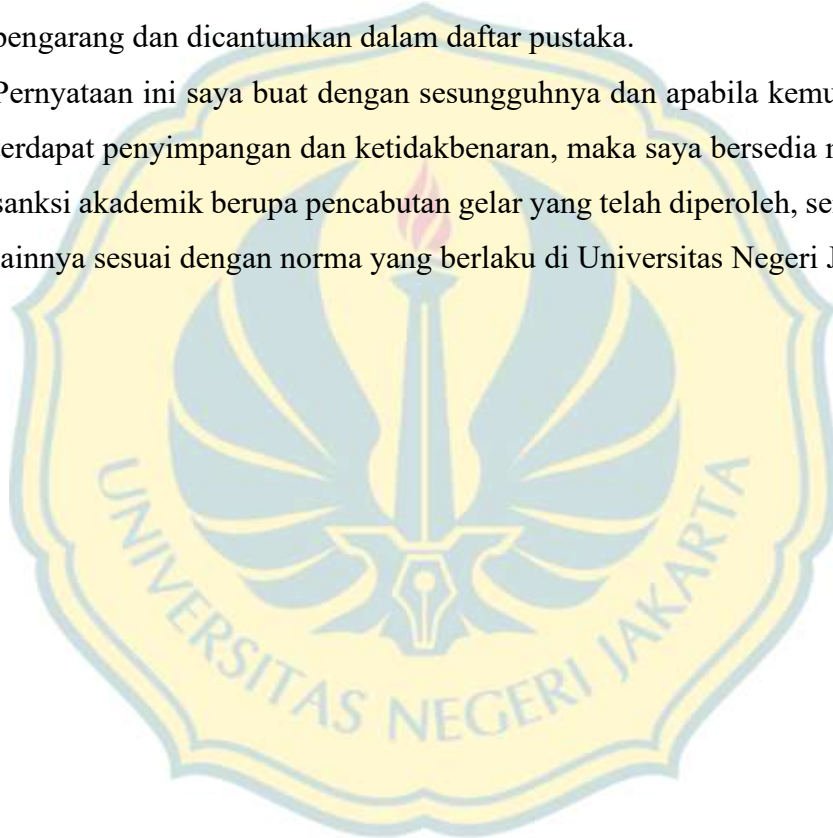


Muchammad Ficky Duskarnaen,
M.Sc.
NIDN 0024097304

LEMBAR PERNYATAAN

Dengan ini, saya menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.



Intelligentia - Dignitas

Jakarta, 1 Juni 2026

Yang membuat pernyataan,



Frasetia Adi Kusuma

No. Registrasi: 1512622038

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "*Aspect Based Sentiment Analysis* untuk Mengidentifikasi Keluhan Penagihan Pinjaman Online di Media Sosial X Menggunakan Algoritma Bayesian Network dan C4.5" dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Pendidikan pada Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta.

Penulis menyadari bahwa terselesaikannya skripsi ini tidak lepas dari bantuan, bimbingan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Komarudin, M. Si., selaku Rektor Universitas Negeri Jakarta. Terima kasih atas kesempatan dan fasilitas yang diberikan sehingga penulis dapat menempuh dan menyelesaikan pendidikan di lingkungan Universitas Negeri Jakarta.
2. Ibu Prof. Dr. Neneng Siti Silfi Ambarwati, S. Si., Apt., M. Si., selaku Dekan Fakultas Teknik Universitas Negeri Jakarta. Terima kasih atas kebijakan dan dukungan akademik yang telah diberikan selama penulis menempuh studinya hingga penyusunan skripsi ini.
3. Bapak Muhammad Ficky Duskarnaen, ST, M. Sc., selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer serta Dosen Pembimbing Akademik. Terima kasih atas arahan dan perhatian yang telah bapak berikan kepada seluruh mahasiswanya untuk menyelesaikan pendidikan tepat waktu.
4. Bapak Dr. Widodo, S. Kom., M. Kom., selaku Dosen Pembimbing I yang dengan penuh keikhlasan, ketelitian, dan kesabaran telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, serta masukan yang sangat berharga bagi penulis selama proses penyusunan skripsi ini.

5. Ibu Neng Ayu Herawati, S. Pd., M. T., selaku Dosen Pembimbing II yang dengan penuh perhatian, ketulusan, dan dedikasi telah meluangkan waktu, tenaga, dan pikiran dalam memberikan bimbingan, arahan, serta motivasi yang sangat berarti bagi penulis selama proses penyusunan skripsi ini.
6. Seluruh Dosen dan Staf Program Studi Pendidikan Teknik Informatika dan Komputer, atas ilmu pengetahuan dan bantuan yang telah diberikan selama masa perkuliahan.
7. Ayah Feri dan Ibu Rosnati, terima kasih atas doa, kasih sayang, dukungan moral, materi, dan semangat yang tiada henti diberikan kepada penulis.
8. Kakak Indra, Ayuk Dea, serta Kala, yang sering mengajak penulis makan AYCE selama masa perkuliahan hingga terselesaikannya skripsi ini.
9. (Alm.) Riza Sukma, M. Hum., selaku pembimbing lapangan penulis selama melaksanakan PKL. Terima kasih atas ilmu, teknik wicara publik, serta segala bimbingan yang telah diberikan. Semoga segala ilmu yang beliau amalkan menjadi pahala jariyah.
10. Feryadi dan Arif selaku teman yang masih akrab aja sampai sekarang.
11. Teman-teman PKM di SMAN 53 Jakarta dan rekan PKL di Badan Pengembangan dan Pembinaan Bahasa, terima kasih atas lingkungan kerja yang suportif, keseruan, serta pengalaman profesional yang sangat mendidik.
12. Keluarga besar, terima kasih yang sebesar-besarnya atas doa, dukungan moril maupun materiil, serta perhatian yang senantiasa mengalir dari seluruh sanak saudara sepanjang perjalanan akademik penulis.
13. Podcast Seminggu, terima kasih karena asem-asem seger selalu hadir menemani dan menghibur penulis di setiap hari Senin.
14. Teman-teman Program Studi Pendidikan Teknik Informatika dan Komputer angkatan 2022, yang telah memberikan semangat, bantuan, dan hiburan selama masa studi.
15. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah turut berkontribusi dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari

semua pihak demi perbaikan dan penyempurnaan skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya di bidang analisis sentimen, kecerdasan buatan, dan perlindungan konsumen digital di Indonesia.

Jakarta, 1 Juni 2026



Penulis

Frasetia Adi Kusuma



Intelligentia - Dignitas

ABSTRAK

Frasetia Adi Kusuma, *Aspect Based Sentiment Analysis* untuk Mengidentifikasi Keluhan Penagihan Pinjaman *Online* di Media Sosial X Menggunakan Algoritma *Bayesian Network* dan C4.5. Dosen Pembimbing: Dr. Widodo, S. Kom., M. Kom. dan Neng Ayu Herawati, S.Pd., M.T. Program Studi Pendidikan Teknik Informatika dan Komputer. Fakultas Teknik. Universitas Negeri Jakarta. 2026.

Penelitian ini mengimplementasikan *Aspect-Based Sentiment Analysis* (ABSA) keluhan penagihan pinjaman *online* di media sosial X dengan membandingkan algoritma *Bayesian Network* dan C4.5. Data diproses mengacu pada tahapan *Knowledge Discovery in Databases* (KDD) hingga tahap representasi vektor numerik menggunakan FastText. Teknik *Synthetic Minority Over-sampling Technique* (SMOTE) diterapkan pada keseluruhan data untuk menangani ketidakseimbangan kelas, kemudian dievaluasi menggunakan skema *Stratified 5-Fold Cross Validation*. Hasil dari delapan skenario pengujian menunjukkan bahwa intervensi SMOTE justru menurunkan kinerja *Bayesian Network* akibat distorsi probabilitas. Hal ini ditandai dengan turunnya akurasi aspek dari 75,50% menjadi 75,20%, serta meningkatnya angka *false positive* pada sentimen, meskipun akurasi bergeser dari 73,10% ke 76,50%. Sebaliknya, performa algoritma C4.5 meningkat secara masif setelah penggunaan SMOTE, dengan lonjakan akurasi pada kelas aspek 76,30% menjadi 84,00% dan target sentimen dari 74,80% menjadi 81,90%. Kesimpulannya, algoritma C4.5 terbukti lebih adaptif dalam memotong batas keputusan pada ruang vektor *dense* FastText. Integrasi FastText, SMOTE, *Stratified 5-Fold CV*, dan C4.5 merupakan arsitektur pemodelan terbaik untuk mengklasifikasikan teks keluhan praktik penagihan pada penelitian ini.

Kata kunci: ABSA, *Bayesian Network*, C4.5, FastText, SMOTE, Pinjaman Online

ABSTRACT

Frasetia Adi Kusuma, *Aspect Based Sentiment Analysis to Identify Online Loan Debt Collection Complaints on Media Social X Using Bayesian Network and C4.5 Algorithms*. Supervisors: Dr. Widodo, S.Kom., M.Kom. and Neng Ayu Herawati, S.Pd., M.T. Study Program of Informatics and Computer Engineering Education, Faculty of Engineering, Universitas Negeri Jakarta, 2026.

This study implements Aspect-Based Sentiment Analysis (ABSA) on online loan collection complaints on social media X by comparing Bayesian Network and C4.5 algorithms. Data processing follows the Knowledge Discovery in Databases (KDD) stages up to the numerical vector representation stage using FastText. The Synthetic Minority Over-sampling Technique (SMOTE) was applied to the entire dataset to handle class imbalance, followed by an evaluation using a Stratified 5-Fold Cross Validation scheme. The results from eight testing scenarios indicate that SMOTE intervention actually degraded the performance of the Bayesian Network due to probability distortion. This was marked by a decline in aspect target accuracy from 75.50% to 75.20%, along with an increase in false positive rates for the sentiment target, despite the overall accuracy shifting from 73.10% to 76.50%. Conversely, the performance of the C4.5 algorithm improved massively following the application of SMOTE, with a surge in accuracy for the aspect target from 76.30% to 84.00% and the sentiment target from 74.80% to 81.90%. In conclusion, the C4.5 algorithm proved to be more adaptive in establishing decision boundaries within the FastText dense vector space. The integration of FastText, SMOTE, Stratified 5-Fold CV, and C4.5 constitutes the optimal modeling architecture for classifying debt collection complaint texts in this study.

Keywords: ABSA, Bayesian Network, C4.5, FastText, SMOTE, Online Loan

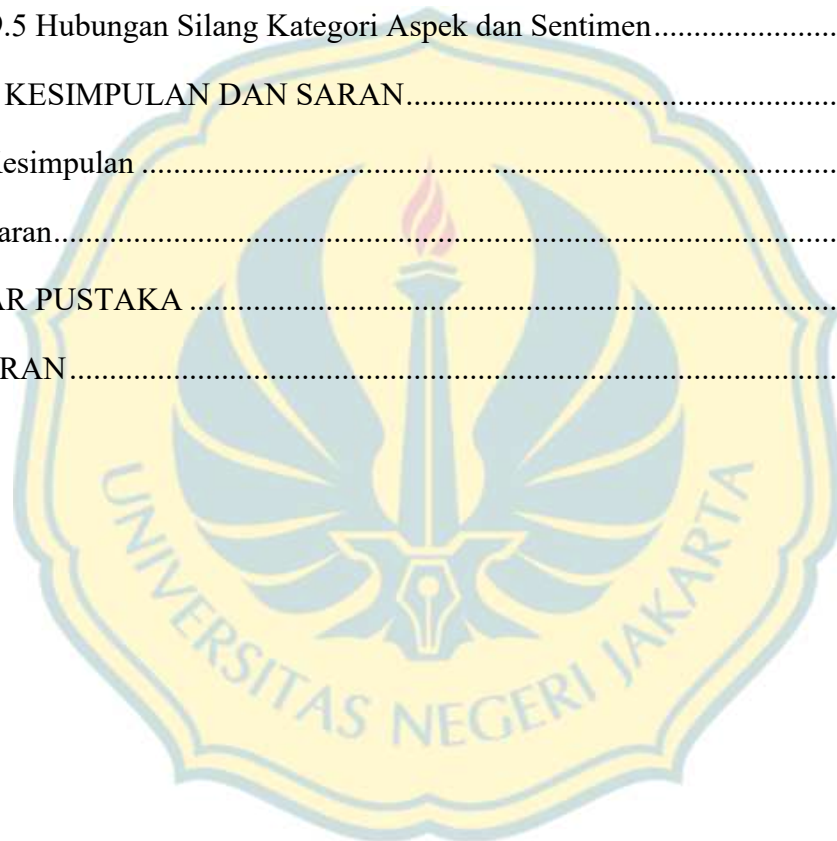
DAFTAR ISI

LEMBAR PENGESAHAN TUGAS AKHIR	i
LEMBAR PERNYATAAN	ii
KATA PENGANTAR	iii
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Pembatasan Masalah	6
1.4 Rumusan Masalah	6
1.5 Tujuan Penelitian	6
1.6 Manfaat Penelitian	7
1.6.1 Manfaat Teoritis	7
1.6.2 Manfaat Praktis	7
BAB II TINJAUAN PUSTAKA	6
2.1 Kerangka Teoritik	6
2.1.1 Pinjaman Online	6
2.1.2 Media Sosial X	9
2.1.3 Natural Language Processing (NLP)	9
2.1.4 Aspect Based Sentiment Analysis (ABSA)	9
2.1.5 Teks Preprocessing	10

2.1.6 Word Embedding FastText	12
2.1.7 Stratified K-Fold Cross Validation	13
2.1.8 Synthetic Minority Over-sampling Technique (SMOTE)	13
2.1.9 Bayesian Network	14
2.1.10 C4.5	16
2.1.11 Knowledge Discovery in Database (KDD)	19
2.1.12 Confusion Matrix	22
2.2 Penelitian Relevan	24
2.3 Kerangka Berpikir	32
BAB III METODOLOGI PENELITIAN	35
3.1 Tempat dan Waktu Penelitian	35
3.2 Alat dan Bahan Penelitian	35
3.2.1 Alat Penelitian	35
3.2.2 Bahan Penelitian	37
3.3 Langkah-Langkah Penelitian	37
3.4 Prosedur Penelitian	38
3.4.1 <i>Data Integration</i> (Integrasi Data)	39
3.4.2 <i>Data Cleaning</i> (Pembersihan Data)	39
3.4.3 <i>Data Selection</i> (Seleksi Data)	41
3.4.4 <i>Data Transformation</i> (Transformasi Data)	42
3.4.5 Data Mining	43
3.4.6 <i>Pattern Evaluation</i> (Evaluasi Pola dan Perbandingan)	44
3.4.7 Representasi Pengetahuan (Knowledge Presentation)	45
3.5 Teknik Pengumpulan Data	45
3.5.1 Sumber Data dan Alat Pengumpulan	45
3.5.2 Kata Kunci Pencarian	45

3.5.3 Parameter dan Volume Data	45
3.5.4 Instrumen Pelabelan Data	45
BAB IV HASIL DAN PEMBAHASAN	47
4.1 Hasil Pengumpulan dan Integrasi Data	47
4.2 Hasil Teks	49
4.2.1 Hasil Case Folding	49
4.2.2 Hasil Normalisasi	50
4.2.3 Hasil Cleaning	51
4.2.4 Hasil Stopword Removal	52
4.2.5 Hasil Stemming	54
4.3 Hasil Seleksi dan Pelabelan Data	55
4.3.1 Distribusi Label Aspek	55
4.3.2 Distribusi Label Sentimen	56
4.4 Transformasi Data	57
4.4.1 Representasi Teks dengan FastText	57
4.4.2 SMOTE	58
4.4.3 Pembagian Data	59
4.5 Data Mining	59
4.5.1 Pelatihan Model Bayesian Network	60
4.5.2 Pelatihan Model C4.5	61
4.6 Hasil Klasifikasi dan Evaluasi Aspek	62
4.6.1 Pengujian Aspek dengan Bayesian Network	63
4.6.2 Pengujian Aspek dengan C4.5	67
4.7 Hasil Klasifikasi dan Evaluasi Sentimen	70
4.7.1 Pengujian Sentimen dengan Bayesian Network	71
4.7.2 Pengujian Sentimen dengan C4.5	74

4.8 Analisis Perbandingan Model	78
4.9 Representasi Pengetahuan.....	82
4.9.1 Visualisasi Pohon Keputusan C4.5	82
4.9.2 Visualisasi <i>Directed Acyclic Graph</i> Bayesian Network	83
4.9.3 Distribusi Label Sentimen dan Aspek.....	85
4.9.4 Visualisasi Frekuensi Kata.....	87
4.9.5 Hubungan Silang Kategori Aspek dan Sentimen.....	88
BAB V KESIMPULAN DAN SARAN.....	90
5.1 Kesimpulan	90
5.2 Saran.....	91
DAFTAR PUSTAKA	93
LAMPIRAN.....	100



Intelligentia - Dignitas

DAFTAR TABEL

Tabel 2.1 Confusion Matrix	23
Tabel 2.2 Penelitian Relevan.....	27
Tabel 3.1 Spesifikasi Perangkat Keras.....	35
Tabel 3.2 Spesifikasi Perangkat Lunak.....	36
Tabel 3.3 Contoh Hasil Transformasi Data.....	40
Tabel 3.4 Contoh Hasil Pelabelan Data	42
Tabel 3.5 Skenario Pengujian	43
Tabel 4.1 Rincian Atribut Dataframe.....	47
Tabel 4.2 Ringkasan Dataframe Setelah Proses Seleksi	48
Tabel 4.3 Ringkasan Dataframe Setelah Proses Casefolding	49
Tabel 4.4 Ringkasan Dataframe Setelah Proses Normalisasi	50
Tabel 4.5 Ringkasan Dataframe Setelah Proses Cleaning	52
Tabel 4.6 Ringkasan Dataframe Setelah Proses Stopword Removal.....	53
Tabel 4.7 Ringkasan Dataframe Setelah Proses Stemming	54
Tabel 4.8 Distribusi Label Aspek.....	55
Tabel 4.9 Distribusi Label Sentimen.....	56
Tabel 4.10 Persebaran Data Awal dan Setelah SMOTE.....	58
Tabel 4.11 Hasil Grid Search Cross Validation pada Model Bayesian Network Aspek	60
Tabel 4.12 Hasil Grid Search Cross Validation pada Model Bayesian Network Sentimen.....	60
Tabel 4.13 Hasil Grid Search Cross Validation pada Model C4.5 Aspek	61
Tabel 4.14 Hasil Grid Search Cross Validation pada Model C4.5 Sentimen	62
Tabel 4.15 Analisis Perbandingan Model	79

DAFTAR GAMBAR

Gambar 2.1 Arsitektur FastText (Liebenlito, Yesinta, & Musti, 2024).....	12
Gambar 2.2 Struktur Bayesian Network (Regina, Saragih, & Kartini, 2023)	14
Gambar 2.3 Ilustrasi Proses Algoritma C4.5 (Junaedi dkk., 2019).....	17
Gambar 2.4 Tahapan KDD (Han, Kamber, & Pei, 2012)	20
Gambar 2.5 Kerangka Berpikir	34
Gambar 3.1 Diagram Alir Langkah-Langkah Penelitian	38
Gambar 4.1 Visualisasi FastText	57
Gambar 4.2 Confusion Matrix Bayesian Network Aspek Baseline.....	63
Gambar 4.3 Confusion Matrix Bayesian Network Aspek SMOTE	65
Gambar 4.4 Confusion Matrix C4.5 Aspek Baseline.....	67
Gambar 4.5 Confusion Matrix C4.5 Aspek SMOTE	69
Gambar 4.6 Confusion Matrix Bayesian Network Sentimen Baseline.....	71
Gambar 4.7 Confusion Matrix Bayesian Network Sentimen SMOTE	73
Gambar 4.8 Confusion Matrix C4.5 Sentimen Baseline.....	75
Gambar 4.9 Confusion Matrix C4.5 Sentimen SMOTE	77
Gambar 4.10 Pohon Keputusan Aspek	83
Gambar 4.11 Pohon Keputusan Sentimen	83
Gambar 4.12 Struktur DAG Aspek	84
Gambar 4.13 Struktur DAG Sentimen	84
Gambar 4.14 Visualisasi Distribusi Label Aspek	85
Gambar 4.15 Visualisasi Distribusi Label Sentimen	86
Gambar 4.16 Representasi Kata dengan WordCloud	87
Gambar 4.17 Diagram Frekuensi Kemunculan Kata	88
Gambar 4.18 Hubungan Silang Aspek dan Sentimen.....	88

DAFTAR LAMPIRAN

Lampiran 1 Lembar Pernyataan Kelayakan Judul	100
Lampiran 2 Surat Tugas Dosen Pembimbing	101
Lampiran 3 Lembar Konsultasi Skripsi	102
Lampiran 4 Kartu Kehadiran Seminar Proposal Skripsi.....	106
Lampiran 5 Biodata Peserta Ujian Skripsi	108
Lampiran 6 Surat Permohonan Ujian Skripsi	111
Lampiran 7 Surat Persetujuan Dosen Pembimbing Skripsi	112
Lampiran 8 Surat Persetujuan Dosen Pembimbing Akademik.....	114
Lampiran 9 Pra-Transkrip Akademik	115
Lampiran 10 Surat Pernyataan Menyelesaikan Mata Kuliah Wajib.....	118
Lampiran 11 Dataset Asli	119
Lampiran 12 Dataset Setelah Preprocessing	120
Lampiran 13 Dataset Setelah Distribusi Label.....	121
Lampiran 14 Dokumentasi Kode	122

