

**EKSTRAKSI PENGETAHUAN DARI ULASAN APLIKASI CAPCUT  
MENGUNAKAN METODE ANALISIS SENTIMEN BERBASIS ASPEK  
DAN METODE KLASIFIKASI**

**SKRIPSI**

Program Studi Sistem Informasi

Jenjang Sarjana S1



Oleh :

**Ishlah Putri Ariyani**

**09031282126101**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA**

**2024**

**LEMBAR PENGESAHAN**

**SKRIPSI**

**EKSTRAKSI PENGETAHUAN DARI ULASAN APLIKASI CAPCUT  
MENGUNAKAN METODE ANALISIS SENTIMEN BERBASIS ASPEK  
DAN METODE KLASIFIKASI**

Sebagai salah satu syarat untuk penyelesaian Studi

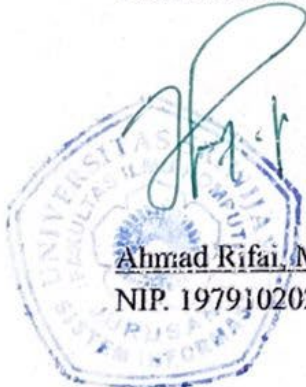
Di Program Studi Sistem Informasi

Oleh

**Ishlah Putri Ariyani**

**NIM 09031232126101**

Mengetahui,  
Ketua Jurusan Sistem Informasi,



**Ahmad Rifai, M.T**  
**NIP. 197910202010121003**

Palembang,

Pembimbing,

A green ink signature of Ken Ditha Tania, M.Kom., PhD. is written over the text.

**Ken Ditha Tania, M.Kom., PhD**  
**NIP. 198507182012122003**

## HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini :

Nama : Ishlah Putri Ariyani  
NIM : 09031282126101  
Program Studi : Sistem Informasi  
Judul Skripsi : Ekstraksi Pengetahuan dari Ulasan Aplikasi CapCut  
Menggunakan Metode Analisis Sentimen Berbasis Aspek  
dan Metode Klasifikasi

Hasil Pengecekan iThenticate/Turnitin: 5%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya saya sendiri dan bukan penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 27 Desember 2024



Ishlah Putri Ariyani  
NIM. 09031282126101

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Senin

Tanggal : 23 Desember 2024

Nama : Ishlah Putri Ariyani

NIM : 09031282126101

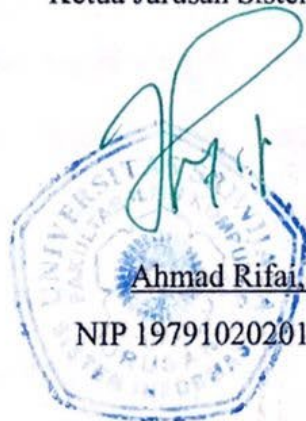
Judul : Ekstraksi Pengetahuan dari Ulasan Aplikasi CapCut Menggunakan  
Metode Analisis Sentimen Berbasis Aspek dan Metode Klasifikasi

Komisi Penguji :

1. Ketua : Allsela Meiriza, M.T
2. Sekretaris : Pacu Putra Suarli, M.CS
3. Pembimbing : Ken Ditha Tania, M.Kom., PhD
4. Penguji : Ari Wedhasmara, M.TI



Mengetahui,  
Ketua Jurusan Sistem Informasi



Ahmad Rifai, M.T

NIP 197910202010121003

## HALAMAN PERSEMBAHAN DAN MOTTO

*” Life can be heavy, especially if u try to carry it all at one, part of growing up and moving into new chapters of your life is about catch or release. What I mean by that is, knowing what things to keep and what things to release. You can’t carry all things, all grudges, decide what is yours to hold and let the rest go”*

– Taylor Alison Swift

*”Perang Telah Usai, Aku Bisa Pulang. Kubaringkan Panah dan Berteriak  
MENANG!”* – Nadin Amizah

*“Angan-Angan Yang Dulu Mimpi Belaka, Kita Gapai Segala Yang Tak Disangka”*

– Baskara Putra

**Skripsi ini saya persembahkan untuk :**

- 1. Allah SWT**
- 2. Diriku Sendiri, Ishlah Putri Ariyani**
- 3. Ayah, Ibu, Nyai, Yai, Intan, Bintang beserta Seluruh Keluarga Besar**
- 4. Dosen Pembimbing Skripsi dan Dosen Pembimbing Akademik**
- 5. Dosen Sistem Informasi Universitas Sriwijaya**
- 6. Seluruh teman – teman Sistem Informasi Angkatan 2021**
- 7. Fakultas Ilmu Komputer, Universitas Sriwijaya**
- 8. Almamater, Universitas Sriwijaya**

## KATA PENGANTAR

Puji dan syukur kepada Allah SWT atas ridho dan karunia-Nya, Penulis dapat menyelesaikan laporan tugas akhir yang berjudul "Ekstraksi Pengetahuan dari Ulasan Aplikasi CapCut Menggunakan Metode Analisis Sentimen Berdasarkan Aspek dan Metode Klasifikasi" sebagai salah satu syarat menyelesaikan Pendidikan Jenjang Sarjana di Program Studi Sistem Informasi pada Fakultas Ilmu Komputer Universitas Sriwijaya.

Selama penyusunan skripsi ini penulis banyak mendapatkan bantuan, bimbingan, dukungan serta doa dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis mengucapkan terima kasih kepada :

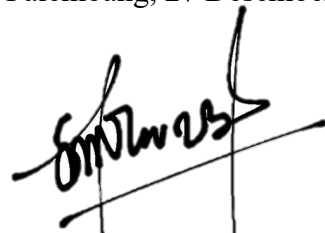
1. Allah SWT yang telah memberikan Rahmat dan anugerah berupa ilmu yang bermanfaat, kesempatan, kesabaran serta kesehatan jasmani dan rohani sehingga penulis mampu menyelesaikan skripsi ini dengan tepat waktu.
2. Kedua orang tua penulis, Bapak Abun Yamin, S.T dan Ibu Ria Sari, S.Pd, selaku dua orang yang sangat berjasa dalam hidup penulis. Terima kasih atas doa, kasih sayang, kepercayaan dan segala bentuk yang telah diberikan, sehingga penulis merasa terdukung di segala pilihan dan keputusan yang diambil oleh penulis, serta tanpa lelah mendengar keluh kesah penulis hingga di titik ini. Semoga Allah SWT memberikan keberkahan di dunia serta tempat terbaik di akhirat kelak, karena telah menjadi figur orangtua terbaik bagi penulis.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.

4. Bapak Ahmad Rifai, M.T selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Ken Ditha Tania, M.Kom., PhD, selaku Dosen Pembimbing Skripsi yang telah membantu penulis dalam pengerjaan tugas akhir ini, dengan bantuan dan bimbingan Beliau lah penulis dapat menjalankan dan menyelesaikan tugas akhir ini dengan baik dan lancar.
6. Ibu Putri Eka Sevtiyuni, S.Si., M.T, selaku Dosen Pembimbing Akademik yang telah membimbing penulis dari semester awal hingga semester akhir.
7. Muhamad Anfakhri yang berkontribusi banyak dalam tugas akhir ini, baik tenaga, waktu, maupun pikirannya. Terimakasih untuk segala hal yang menemani, mendukung ataupun menghibur dalam kesedihan, mendengar keluh kesah, memberi semangat dan banyak motivasi kepada penulis.
8. Siska Septiyanah selaku teman, sahabat penulis yang sudah banyak membantu penulis, memberikan semangat kepada penulis, selalu bersedia 24/7 untuk menolong kesulitan penulis, dan memberikan tebengan kepada penulis, mendengarkan curhatan penulis dari awal perkuliahan hingga akhir perkuliahan serta tidak menolak untuk direpotkan oleh penulis.
9. Anna Fadila Inayah dan Belia Citra selaku teman dari awal perkuliahan hingga sekarang yang telah banyak membantu tugas akhir penulis, selalu bersedia untuk menemani penulis dari awal perkuliahan hingga saat ini, memberikan dukungan dan semangat kepada penulis.
10. Teman – teman sesepuh selaku teman PP Palembang-Indralaya yang sudah memberikan celotehan yang lucu membuat penulis terhibur dan sudah membantu perkuliahan penulis kurang lebih 5 semester.

11. Teman – teman SI Reguler B Angkatan 2021 yang sudah membantu selama proses perkuliahan.
12. Teman – Teman HIMSI Tahun 2023 sampai dengan 2024 yang sudah menjadi tempat belajar, menjadi teman bercerita dan bertukar pikiran.
13. Semua teman – teman seperjuangan di Sistem Informasi Angkatan 2021 yang tidak bisa saya sebutkan satu persatu.
14. *last but not least*, kepada diri saya sendiri. Apresiasi sebesar-besarnya karena telah berjuang untuk menyelesaikan apa yang telah dimulai dan memilih untuk tidak menyerah dalam kondisi apapun, saya bangga pada diri saya sendiri bisa menyelesaikan laporan tugas akhir ini dengan penuh lika-liku kehidupan yang dijalani.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan dengan segala kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran terhadap skripsi ini.

Palembang, 27 Desember 2024



Ishlah Putri Ariyani  
NIM 09031282126101



**EKSTRAKSI PENGETAHUAN DARI ULASAN APLIKASI CAPCUT  
MENGUNAKAN METODE ANALISIS SENTIMEN BERBASIS ASPEK  
DAN METODE KLASIFIKASI**

**Oleh :**

**Ishlah Putri Ariyani**

**09031282126101**

**ABSTRAK**

Indonesia mengalami perkembangan teknologi yang pesat, khususnya dalam penggunaan internet dan platform *editing* seperti CapCut. Platform ini memungkinkan pengeditan video di berbagai perangkat, namun kepuasan pengguna tidak selalu terjamin karena perbedaan pengalaman individu. Untuk memahami opini dan persepsi pengguna terhadap CapCut, penelitian ini menggunakan *Aspect-Based Sentiment Analysis* (ABSA) dan membandingkan empat algoritma *machine learning*—*Naïve Bayes*, *Support Vector Machine (SVM)*, *Decision Tree*, dan *Random Forest* dengan Teknik SMOTE diterapkan untuk meningkatkan representasi data minoritas dan akurasi analisis sentimen. Hasil penelitian menunjukkan bahwa algoritma *Decision Tree* yang paling efektif dibandingkan dengan algoritma lainnya dengan nilai akurasi 0.97, *Cross-validation accuracy* untuk *Decision Tree* adalah  $0.97 \pm 0.00$ , yang menunjukkan konsistensi model yang sangat baik pada data yang berbeda serta algoritma *Decision Tree* juga tidak terpengaruh banyak oleh penggunaan Teknik SMOTE. Penelitian ini juga menghasilkan ekstraksi pengetahuan yang berguna berupa XML.

**Kata Kunci :** ABSA, CapCut, Ekstraksi Pengetahuan, SMOTE

**KNOWLEDGE EXTRACTION FROM CAPCUT APP REVIEWS USING  
ASPECT-BASED SENTIMENT ANALYSIS AND CLASSIFICATION  
METHODS**

**Oleh :**

**Ishlah Putri Ariyani**

**09031282126101**

***ABSTRACT***

*Indonesia is experiencing rapid technological development, especially in the use of the internet and editing platforms such as CapCut. These platforms allow video editing on various devices, but user satisfaction is not always guaranteed due to differences in individual experience. To understand users' opinions and perceptions of CapCut, this study uses Aspect-Based Sentiment Analysis (ABSA) and compares four machine learning algorithms-Naïve Bayes, Support Vector Machine (SVM), Decision Tree, and Random Forest with SMOTE technique applied to improve minority data representation and sentiment analysis accuracy. The results show that the Decision Tree algorithm is the most effective algorithm compared to the other algorithms with an accuracy value of 0.97, Cross-validation accuracy for Decision Tree is  $0.97 \pm 0.00$ , which indicates excellent model consistency on different data and Decision Tree algorithm is also not affected much by the use of SMOTE Technique. This research also resulted in the extraction of useful knowledge in the form of XML.*

**Keywords :** ABSA, CapCut, Knowledge Extraction, SMOTE

## DAFTAR ISI

LEMBAR PENGESAHAN .....	Error! Bookmark not defined.
HALAMAN PERNYATAAN BEBAS PLAGIAT.....	Error! Bookmark not defined.
HALAMAN PERSETUJUAN.....	iii
HALAMAN PERSEMBAHAN DAN MOTTO.....	v
KATA PENGANTAR.....	vi
ABSTRAK .....	ix
<i>ABSTRACT</i> .....	x
DAFTAR ISI .....	xi
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL .....	xiv
DAFTAR LAMPIRAN .....	xv
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah.....	7
<b>BAB II TINJAUAN PUSTAKA</b> .....	8
2.1 Landasan Teori .....	8
2.1.1 Ekstraksi Pengetahuan .....	8
2.1.3 <i>Aspect-Based Sentiment Analysis</i> (ABSA) .....	9
2.1.4 Aplikasi CapCut.....	10
2.1.5 <i>Text Mining</i> .....	11
2.1.6 <i>Synthetic Minority Oversampling Technique</i> (SMOTE).....	11
2.1.7 <i>Web Scraping Data</i> .....	12
2.1.8 <i>Natural language toolkit</i> (NLTK).....	12
2.1.9 Sastrawi.....	13
2.1.10 Pembobotan TF-IDF .....	13
2.1.11 Google Collab .....	14
2.2 Penelitian Sebelumnya .....	15
<b>BAB III METODOLOGI PENELITIAN</b> .....	18

3.1 Objek Penelitian .....	18
3.2 Alat dan Bahan .....	18
3.3 Tahapan Penelitian.....	19
3.4 Metode Pengumpulan Data .....	22
3.5 Tahapan Pre-Processing Data .....	22
3.5.1 <i>Cleaning</i> .....	22
3.5.2 <i>Case Folding</i> .....	23
3.5.3 <i>Tokenize</i> .....	24
3.5.4 <i>Filtering</i> .....	25
3.5.5 <i>Stemming</i> .....	25
3.6 <i>Aspect-Based Sentiment Analysis ( ABSA)</i> .....	26
<b>BAB IV PEMBAHASAN HASIL</b> .....	28
4.1 Pengumpulan Data.....	28
4.1.1 Proses <i>Crawling</i> Data .....	28
4.1.2 Hasil <i>Crawling</i> Data .....	28
4.2 Data <i>Preprocessing</i> .....	30
4.2.1 <i>Cleaning</i> .....	31
4.2.2 <i>Case Folding</i> .....	32
4.2.3 <i>Tokenizing</i> .....	32
4.2.4 <i>Filtering/stopword removal</i> .....	33
4.2.5 <i>Stemming</i> .....	33
4.3 Pelabelan Data .....	34
4.4 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	35
4.5 Split Data.....	35
4.6 <i>Synthetic Minority Oversampling Technique (SMOTE)</i> .....	36
4.7 Evaluasi Model.....	36
4.8 Analisis Kesalahan Model .....	42
<b>BAB V PENUTUP</b> .....	47
5.1 Kesimpulan.....	47
5.2 Saran.....	48
<b>DAFTAR PUSTAKA</b> .....	49
<b>LAMPIRAN</b> .....	51

## DAFTAR GAMBAR

Gambar 2. 1 Logo Aplikasi CapCut.....	10
Gambar 3. 1 Tahapan Penelitian .....	19
Gambar 3. 2 Tahapan Pre-Processing .....	20
Gambar 3. 3 <i>Library</i> Python .....	22
Gambar 4. 1 Proses <i>Crawling</i> .....	28
Gambar 4. 2 <i>Crawling</i> Data .....	29
Gambar 4. 3 Hasil <i>Crawling</i> .....	29
Gambar 4. 4 <i>Import</i> Dataset.....	30
Gambar 4. 5 Kode <i>clean</i> slang dan membuat <i>dictionary</i> .....	30
Gambar 4. 6 Kode mengganti slang.....	30
Gambar 4. 7 Kode Menghapus <i>missing values</i> .....	31
Gambar 4. 8 Menghapus Duplikasi Data .....	31
Gambar 4. 9.Kode Menghapus simbol dan karakter khusus .....	31
Gambar 4. 10. Menghapus Teks Kosong Setelah Pembersihan.....	31
Gambar 4. 11 Kode Proses <i>Case Folding</i> .....	32
Gambar 4. 12 Kode <i>Tokenizing</i> .....	32
Gambar 4. 13 Kode <i>Filtering/Stopword</i> .....	33
Gambar 4. 14. Kode proses <i>stemming</i> .....	34
Gambar 4. 15. Hasil Sentimen berdasarkan ABSA.....	34
Gambar 4. 16. Visual Aspek <i>User Experience</i> .....	35
Gambar 4. 17 Visualisasi Aspek Fitur.....	35
Gambar 4. 18. Proses Transformasi Teks ke Matriks TF-IDF Menggunakan TfidfVectorizer.....	35
Gambar 4. 19. Pembagian Dataset Menjadi <i>Training</i> dan <i>Testing</i> .....	36
Gambar 4. 20 Penyeimbangan Data Latih dengan Metode SMOTE.....	36
Gambar 4. 24. ROC-Curve NB + SMOTE .....	37
Gambar 4. 23 ROC-Curve NB Tanpa SMOTE.....	37
Gambar 4. 27 ROC-Curve SVM + SMOTE .....	39
Gambar 4. 28 ROC-Curve SVM Tanpa SMOTE.....	39
Gambar 4. 31 ROC - Curve Tanpa SMOTE DT .....	40
Gambar 4. 32 ROC - Curve + SMOTE DT .....	40
Gambar 4. 35 ROC-Curve + SMOTE RF.....	42
Gambar 4. 36 ROC-Curve Tanpa SMOTE RF .....	42
Gambar 4. 37 <i>Confusion Matrix</i> NB + SMOTE .....	43
Gambar 4. 38 <i>Confusion Matrix</i> NB Tanpa SMOTE.....	43
Gambar 4. 39 <i>Confusion Matrix</i> SVM + SMOTE .....	43
Gambar 4. 40 <i>Confusion Matrix</i> SVM Tanpa SMOTE.....	43
Gambar 4. 41 <i>Confusion Matrix</i> DT Tanpa SMOTE .....	44
Gambar 4. 42 <i>Confusion Matrix</i> DT + SMOTE .....	44
Gambar 4. 43 <i>Confusion Matrix</i> RF + SMOTE.....	45
Gambar 4. 44 <i>Confusion Matrix</i> RF + SMOTE.....	45
Gambar 4. 45 <i>Knowledge Extraction</i> Menggunakan SMOTE.....	46
Gambar 4. 46. <i>Knowledge Extraction</i> Tanpa SMOTE.....	46

## DAFTAR TABEL

Tabel 2.1 Penelitian Sebelumnya .....	15
Tabel 3. 1 Hasil <i>Cleaning</i> .....	22
Tabel 3. 2 Hasil <i>Case Folding</i> .....	23
Tabel 3. 3 Hasil <i>Tokenize</i> .....	24
Tabel 3. 4 Hasil <i>Stopword Removal</i> .....	25
Tabel 3. 5 Hasil <i>Stemming</i> .....	26
Tabel 3. 6 Aspek -Aspek ABSA .....	26
Tabel 4. 1 Hasil Akurasi, <i>Cross Validation</i> , AUC NB.....	36
Tabel 4. 2 Naive Bayes Tanpa SMOTE .....	37
Tabel 4. 3 Naive Bayes + SMOTE.....	37
Tabel 4. 4 Hasil Akurasi, <i>Cross-validation</i> , AUC SVM .....	38
Tabel 4. 5 SVM Tanpa SMOTE .....	38
Tabel 4. 6 SVM + SMOTE .....	38
Tabel 4. 7 Hasil Nilai <i>Accuracy</i> , <i>Cross-validation Accuracy</i> , AUC DT.....	39
Tabel 4. 8 DT Tanpa SMOTE .....	39
Tabel 4. 9 30. DT + SMOTE.....	40
Tabel 4. 10 Nilai Akurasi, <i>Cross-validation Accuracy</i> , AUC RF.....	41
Tabel 4. 11 Random Forest Tanpa SMOTE .....	41
Tabel 4. 12 Random Forest + SMOTE.....	41

## DAFTAR LAMPIRAN

<b>Lampiran 1. 1</b>	Surat Kesiediaan Membimbing .....	A-1
<b>Lampiran 1. 2</b>	Surat Keputusan Tugas Akhir ( SK TA) .....	B-1
<b>Lampiran 1. 3</b>	Form Desk Evaluasi SemPro .....	C-1
<b>Lampiran 1. 4</b>	Tabel Dataset Hasil Scraping .....	D-1
<b>Lampiran 1. 5</b>	Surat Keterangan Pengecekan Similarity .....	E-1
<b>Lampiran 1. 6</b>	Log Book Dosen Pembimbing Skripsi .....	F-1
<b>Lampiran 1. 7</b>	Kartu Konsultasi .....	G-1

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia mengalami perkembangan teknologi yang signifikan, terutama terhadap penggunaan internet yang meluas, terlebih dalam bidang teknologi kreatif seperti penggunaan *platform editing*. *Platform editing* menjadi salah satu kunci untuk memenuhi kebutuhan kreatif pengguna internet. *Platform editing* merupakan aplikasi atau perangkat lunak yang menyediakan *tools* dan fitur untuk mengedit konten digital seperti gambar, teks, audio maupun video. *Platform editing* yang semakin populer saat ini adalah CapCut. CapCut yang dikembangkan oleh *ByteDance*, telah menarik perhatian khalayak ramai berkat fitur-fitur inovatifnya.

CapCut merupakan platform digital kreatif lengkap yang memungkinkan pengeditan video melalui browser, iOS, Android dan perangkat internet lainnya. CapCut menawarkan berbagai fitur menarik yang dapat digunakan untuk mengedit video untuk berkreasi antara lain menambah musik, teks, stiker, efek dan filter untuk mempercantik, menghilangkan *background*, *green screen*, mengubah ukuran dan durasi video seperti halnya *text-to-speech* atau sebaliknya yang mampu mengenali berbagai bahasa (Noor Hasan, 2024a). Melonjaknya popularitas penggunaan aplikasi CapCut mempengaruhi preferensi pengguna terhadap perubahan dan peningkatan fitur yang ada. Aplikasi CapCut juga tidak dapat menjamin kepuasan pengguna, dikarenakan setiap pengguna memiliki pemikiran dan pengalaman yang berbeda-beda.

Selain aplikasi CapCut terdapat platform *editing* lainnya seperti VN sebuah aplikasi pengeditan video yang juga menawarkan fitur canggih seperti pengeditan



multi-layer, penyesuaian efek, dan dukungan berbagai format video. Menurut penelitian (Ardiansyah et al., 2023), aplikasi CapCut dinilai lebih unggul dibandingkan VN dalam beberapa aspek. Tampilan antarmuka CapCut yang intuitif dan dirancang untuk kemudahan penggunaan membuatnya lebih disukai oleh sebagian besar pengguna. Pengguna mengapresiasi CapCut karena kemudahan navigasi dan pengoperasiannya yang sederhana, yang tidak selalu ditemukan pada VN, yang memiliki kurva pembelajaran sedikit lebih tinggi. Fakta ini menunjukkan bahwa preferensi pengguna terhadap aplikasi editing tidak hanya bergantung pada fitur yang ditawarkan, tetapi juga pada kenyamanan dan kemudahan penggunaan aplikasi tersebut.

Analisis sentimen adalah metode untuk mengeksplorasi opini atau teks yang diperoleh dari berbagai platform media sosial, menggunakan pembelajaran mesin untuk analisisnya (Nandwani & Verma, 2021). Analisis sentimen adalah metode otomatis untuk mengkategorikan sejumlah besar teks menjadi sentimen positif atau negatif (Dashtipour et al., 2021; Thavareesan & Mahesan, 2020). Analisis sentimen cocok untuk mengamati bagaimana masyarakat, terutama pengguna aplikasi CapCut, merespons kebijakan tersebut (Czeranowska et al., 2023; Nandwani & Verma, 2021).

Kualitas layanan aplikasi dapat dilakukan dengan memanfaatkan data mining melalui analisis sentimen terhadap ulasan yang masuk secara massal di berbagai platform online, khususnya di Play Store. Analisis ini diperlukan untuk memperoleh informasi berharga terkait persepsi dan tingkat kepuasan pengguna terhadap aplikasi CapCut (Czeranowska et al., 2023). Meski Play Store memberikan gambaran awal melalui rating dan ulasan, informasi tersebut sering kali tidak

mencakup konteks rinci mengenai aspek-aspek penting yang relevan untuk mengevaluasi pengalaman pengguna. Untuk mengatasi kekurangan ini, analisis sentimen berbasis aspek (ABSA) dapat digunakan sebagai solusi untuk memahami opini dan persepsi pengguna terhadap aplikasi tersebut (Samsir et al., 2022) (Pandunata et al., 2023). Penting untuk menganalisis persepsi dan kepuasan pengguna terhadap CapCut guna mengidentifikasi aspek-aspek yang menjadi keunggulan utama serta kelemahan yang perlu diperbaiki. Analisis ini diharapkan mampu memberikan wawasan mendalam bagi pengembang untuk meningkatkan kualitas aplikasi dan mempertahankan loyalitas pengguna di tengah persaingan yang semakin ketat.

Penelitian ini bertujuan untuk menganalisis ulasan pengguna terkait aplikasi CapCut berdasarkan aspek dengan mengadopsi beberapa algoritma seperti *Naïve Bayes*, *Support Vector Machine (SVM)*, *Decision Tree* dan *Random Forest* yang akan menghasilkan ekstraksi pengetahuan dari hasil analisis sentimen ulasan pengguna aplikasi CapCut. Ekstraksi pengetahuan sangat penting dalam berbagai aplikasi, seperti pengolahan bahasa alami, analisis media sosial, dan pengolahan dokumen. Ini memainkan peran kunci dalam mengelola informasi besar yang tidak terstruktur, yang perlu diubah menjadi pengetahuan yang berguna dan dapat diproses secara komputer (Chasseray et al., 2023; Muppavarapu et al., 2021). Pentingnya ekstraksi pengetahuan pada ulasan aplikasi adalah untuk membantu pengembang dan pemangku kepentingan memahami dan memanfaatkan informasi yang terkandung dalam ulasan pengguna untuk meningkatkan kualitas dan kesuksesan aplikasi. Kemudian, membandingkan beberapa algoritma untuk menilai kinerja model dapat memberikan wawasan tentang kepuasan dan persepsi secara

efektif. Informasi ini sangat berguna bagi perusahaan dalam pengambilan keputusan strategis di masa depan (Tahyudin et al., 2023).

Sampai saat ini, banyak penelitian terkait analisis sentimen diantaranya oleh (Noor Hasan, 2024a), analisis sentimen pada penelitian ini menggunakan algoritma *naïve bayes* yang menghasilkan *confusion matrix*, akurasi sebesar 84.09%, presisi sebesar 91.91%, dan *recall* sebesar 73.53%. Penelitian terkait oleh (Zai & Rahman Isnain, n.d.), analisis sentimen aplikasi CapCut menggunakan algoritma *naïve bayes* dan *Support Vector Machine (SVM)* menggunakan SMOTE (*Synthetic Minority Oversampling Technique*) menunjukkan bahwa hasil optimasi menggunakan SMOTE untuk model *Support Vector Machine (SVM)* mengalami kenaikan dibandingkan model *naïve bayes* yang terjadi penurunan pada nilai presisi. Penelitian selanjutnya oleh (Al Habesyah et al., 2024), analisis sentimen terhadap penutupan tiktok shop yang menggunakan empat algoritma *SVM*, *Random Forest*, *Decision Tree*, dan *Deep Learning (H2O)* dengan menunjukkan hasil AUC 0.900 menggunakan model *Deep Learning (H2O)* dengan SMOTE memiliki kinerja yang paling baik untuk analisis sentimen pada penelitian ini. Penelitian berikutnya oleh (Pandunata et al., 2023), penelitian ini menggunakan metode *Aspect-Based Sentiment Analysis (ABSA)* dan menggunakan metode SMOTE. Hasil dari penelitian ini menunjukkan bahwa algoritma *Support Vector Machine* menggunakan SMOTE adalah algoritma yang paling efektif dari setiap aspek yang sudah ditentukan. Penelitian selanjutnya terkait dengan (Tesfa et al., 2024), penelitian ini menggunakan tiga algoritma yaitu *naïve bayes*, *Support Vector Machine (SVM)* dan *K-Nearest Neighbor (KNN)* yang digunakan untuk menganalisis sentimen berbasis ABSA dimana data yang dikumpulkan dari

halaman facebook resmi *Ethio-telecom* dan menghasilkan bahwa model SVM yang dilatih dengan teknik representasi kata TF-IDF dengan dataset *oversampled* SMOTE telah mencapai akurasi tertinggi sebesar 98% dibandingkan dengan algoritma *naïve bayes* dan *KNN*. Penelitian lainnya dilakukan oleh (Khabour et al., 2023), objek dalam penelitian ini adalah perusahaan layanan perbankan seluler dan menguji beberapa algoritma klasifikasi seperti *Decision Tree*, *support vector machine*, *logistic regression*, *naïve bayes*, *k-nearest neighbors* dan *Random Forest* yang menunjukkan nilai akurasi tertinggi didapatkan sebanyak 94% dan nilai presisi 94% menggunakan algoritma *logistic regression*.

Penelitian ini memberikan wawasan mengenai efektivitas masing-masing algoritma dalam memahami persepsi dan kepuasan pengguna dengan menerapkan dan membandingkan beberapa algoritma pembelajaran mesin dengan menggunakan metode *Aspect-Based Sentiment Analysis* (ABSA) pada ulasan aplikasi CapCut. Kemudian penelitian ini akan meningkatkan representasi data minoritas dan mengurangi bias pada hasil model menggunakan teknik oversampling SMOTE sehingga membantu meningkatkan akurasi analisis sentimen.

Berdasarkan uraian diatas, maka penulis bermaksud untuk melakukan penelitian judul ” **Ekstraksi Pengetahuan dari Ulasan Aplikasi CapCut Menggunakan Metode Analisis Sentimen Berbasis Aspek dan Metode Klasifikasi** ”.

## 1.2 Rumusan Masalah

Berikut rumusan masalah untuk dibahas berdasarkan latar belakang :

1. Bagaimana ABSA akan diterapkan dalam analisis sentimen ulasan pengguna aplikasi CapCut berdasarkan aspek Fitur dan User Experience ?
2. Bagaimana hasil dari ekstraksi pengetahuan dengan menggunakan ABSA memberikan pengetahuan baru terhadap aplikasi CapCut?
3. Menganalisis dengan membandingkan algoritma mana yang paling efektif untuk diterapkan dalam analisis sentimen ulasan aplikasi CapCut menggunakan ABSA.
4. Cara menggunakan teknik SMOTE untuk menganalisis sentimen di ulasan aplikasi CapCut.

## 1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Melakukan analisis sentimen terhadap ulasan aplikasi CapCut berdasarkan aspek.
2. Mendapatkan informasi dari ulasan pengguna aplikasi CapCut.
3. Mengetahui algoritma yang paling efektif untuk digunakan dalam analisis sentimen ulasan aplikasi CapCut.
4. Mendapatkan wawasan mengenai teknik SMOTE dalam analisis sentimen ulasan aplikasi CapCut.

#### 1.4 Manfaat Penelitian

Hal yang dapat diperoleh dari penelitian ini sebagai berikut.

1. Bagi Penulis, dapat memberikan pemahaman terkait ekstraksi pengetahuan dari ulasan aplikasi CapCut dengan menggunakan ABSA (*Aspect-Based Sentiment Analysis*), mengetahui manakah algoritma yang paling efektif dalam penelitian ini dan memberikan wawasan dalam memanfaatkan teknik SMOTE (*Synthetic Minority Over-sampling Technique*) untuk menangani ketidakseimbangan data.
2. Bagi Perusahaan, hasil dari analisis sentimen menggunakan ABSA (*Aspect-Based Sentiment Analysis*) diharapkan dapat membantu pengembang aplikasi dalam memahami persepsi pengguna dan meningkatkan kualitas layanan berdasarkan aspek khususnya pada fitur aplikasi dan *user experience*.

#### 1.5 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut.

1. Ulasan aplikasi CapCut dari *Google play store* digunakan dalam penelitian ini.
2. Menggunakan dua jenis klasifikasi data yaitu positif dan negatif.
3. Penelitian ini dilakukan hanya menggunakan aspek yang sudah ditentukan yaitu fitur dan *user experience*.
4. Menggunakan *tool Google Collab* untuk mengolah data mulai dari *scraping* hingga menganalisis menggunakan ABSA (*Aspect-Based Sentiment Analysis*).

## DAFTAR PUSTAKA

465620-none-715b89ab. (n.d.).

- Adli, N. B. Z., Ahmad, M., Ghani, N. A., Ravana, S. D., & Norman, A. A. (2024). An Ensemble Classification of Mental Health in Malaysia related to the Covid-19 Pandemic using Social Media Sentiment Analysis. *KSII Transactions on Internet and Information Systems (TIIS)*, 18(2), 370–396.
- Al Habesyah, N. Z., Herteno, R., Indriani, F., Budiman, I., & Kartini, D. (2024). Sentiment Analysis of TikTok Shop Closure in Indonesia on Twitter Using Supervised Machine Learning. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 6(2), 148–156. <https://doi.org/10.35882/jeeemi.v6i2.381>
- Alqaryouti, O., Siyam, N., Abdel Monem, A., & Shaalan, K. (2024). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*, 20(1–2), 142–161. <https://doi.org/10.1016/j.aci.2019.11.003>
- Amrie, S., Kurniawan, S., Windiatmaja, J. H., & Ruldeviyani, Y. (2022). Analysis of Google Play Store's Sentiment Review on Indonesia's P2P Fintech Platform. *2022 IEEE Delhi Section Conference (DELCON)*, 1–5.
- Ardiansyah, M., M, M., & Riswanto. (2023). Analisis Komparasi Ketertarikan Masyarakat Kota Batam Dalam Penggunaan Video Editor Capcut dan VN. *Jurnal Informasi Dan Teknologi*, 5(3), 91–102. <https://doi.org/10.60083/jidt.v5i3.398>
- Chamorro-Atalaya, O., Aldana-Trejo, F., Alvarado-Bravo, N., Nieves-Barreto, C., Aguilar-Loyaga, S., Farfán-Aguilar, J., Torres-Quiroz, A., Riveros-Cuellar, A., Pérez-Samanamud, M., & Pérez-Guevara, L. (2023). Student Satisfaction Classification Algorithm Using the Minority Synthetic Oversampling Technique. *International Journal of Information and Education Technology*, 13(7), 1094–1100. <https://doi.org/10.18178/ijiet.2023.13.7.1909>
- Chasseray, Y., Barthe-Delanoë, A.-M., Négny, S., & Le Lann, J.-M. (2023). Knowledge Extraction from textual data and performance evaluation in an unsupervised context. *Information Sciences*, 629, 324–343.
- Chia, Y. K., Chen, H., Han, W., Chen, G., Aljunied, S. M., Poria, S., & Bing, L. (2023). Domain-expanded aste: Rethinking generalization in aspect sentiment triplet extraction. *ArXiv Preprint ArXiv:2305.14434*.
- Czeranowska, O., Chlasta, K., Miłkowski, P., Grabowska, I., Kocoń, J., Hwaszcz, K., Wiczorek, J., & Jastrzębowska, A. (2023). Migrants vs. stayers in the pandemic—A sentiment analysis of Twitter content. *Telematics and Informatics Reports*, 10, 100059.

- Dashtipour, K., Gogate, M., Adeel, A., Larijani, H., & Hussain, A. (2021). Sentiment analysis of persian movie reviews using deep learning. *Entropy*, 23(5). <https://doi.org/10.3390/e23050596>
- Febby Wilyani, Qonaah Nuryan Arif, & Fitri Aslimar. (2024). Pengenalan Dasar Pemrograman Python Dengan Google Colaboratory. *Jurnal Pelayanan Dan Pengabdian Masyarakat Indonesia*, 3(1), 08–14. <https://doi.org/10.55606/jppmi.v3i1.1087>
- Gelar Guntara, R. (2023). Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(1), 55–60. <https://doi.org/10.47233/jteksis.v5i1.750>
- Gunawan, K. I., & Santoso, J. (2021). Multilabel Text Classification Menggunakan SVM dan Doc2Vec Classification Pada Dokumen Berita Bahasa Indonesia. *Journal of Information System, Graphics, Hospitality and Technology*, 3(01), 29–38. <https://doi.org/10.37823/insight.v3i01.126>
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1–34. <https://doi.org/10.3390/bdcc4010001>
- Khabour, S. M., Mustafa, D., & AlThebyan, Q. (2023). Arabic Sentiment Analysis of Mobile Banking Services Reviews. *2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 1–8. <https://doi.org/10.1109/SNAMS60348.2023.10375427>
- Muppavarapu, V., Ramesh, G., Gyrard, A., & Noura, M. (2021). Knowledge Extraction using semantic similarity of concepts from Web of Things knowledge bases. *Data & Knowledge Engineering*, 135, 101923.
- Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 81.
- Noor Hasan, F. (2024a). Analisis Sentimen Pengguna Aplikasi CapCut Pada Ulasan di Play Store Menggunakan Metode Naïve Bayes. *Media Online*, 4(4). <https://doi.org/10.30865/klik.v4i4.1555>
- Noor Hasan, F. (2024b). Analisis Sentimen Pengguna Aplikasi CapCut Pada Ulasan di Play Store Menggunakan Metode Naïve Bayes. *Media Online*, 4(4). <https://doi.org/10.30865/klik.v4i4.1555>
- Pandunata, P., Nurdiansyah, Y., & Alfina, F. D. (2023). Aspect-Based Sentiment Analysis of Avatar 2 Movie Reviews on IMDb Using Support Vector Machine. *E3S Web of Conferences*, 448. <https://doi.org/10.1051/e3sconf/202344802041>
- Pendidikan, B., Pelatihan Keuangan, D., Keuangan, K., Djufri, M., & Pajak, P. (n.d.). *JURNAL BPPK PENERAPAN TEKNIK WEB SCRAPING UNTUK*



*PENGGALIAN POTENSI PAJAK (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak)* (Vol. 13).

- Qahar, M. Y. Al, Ruldeviyani, Y., Mukharomah, U. N., Fidyawan, M. A., & Putra, R. (2024). Factor analysis influencing Mobile JKN user experience using sentiment analysis. *IAES International Journal of Artificial Intelligence*, 13(2), 1782–1793. <https://doi.org/10.11591/ijai.v13.i2.pp1782-1793>
- Rachman, F. F., & Pramana, S. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. In *Health Information Management Journal ISSN* (Vol. 8, Issue 2).
- Rofiqi, M. A., Fauzan, A. C., Agustin, A. P., & Saputra, A. A. (2019). Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 1(2), 58–64.
- Samsir, S., Kusmanto, K., Dalimunthe, A. H., Aditiya, R., & Watrianthos, R. (2022). Implementation naïve bayes classification for sentiment analysis on internet movie database. *Building of Informatics, Technology and Science (BITS)*, 4(1), 1–6.
- Sofyan, S., & Prasetyo, A. (2021). Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Data Tidak Seimbang Pada Tingkat Pendapatan Pekerja Informal Di Provinsi D.I. Yogyakarta Tahun 2019. In *Seminar Nasional Official Statistics*.
- Sukmana, F., & Rozi, F. (2019). Extraction keyterm in work order for decision support. *J. Theor. Appl. Inf. Technol*, 97(22), 3262–3272.
- Suryani, L., & Edy, K. (n.d.). PENGEMBANGAN APLIKASI “LOST & FOUND” BERBASIS ANDROID DENGAN MENGGUNAKAN METODE TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) DAN COSINE SIMILARITY APPLICATION DEVELOPMENT “LOST & FOUND” ANDROID-BASED USING TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) AND COSINE SIMILARITY METHOD. In *JEC* (Vol. 6, Issue 2).
- Tahyudin, I., Hananto, A. R., Rahayu, S. A., Anjani, R. M., & Nurhopipah, A. (2023). Sentiment Analysis Model Development on E-Money Service Complaints. *TEM Journal*, 12(4).
- Tesfa, T., Belete, B., Abera, S., Mohapatra, S. K., & Das, T. K. (2024). Aspect-Based Sentiment Analysis on Amharic Text for Evaluating Ethio-Telecom Services. *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*, 1–6. <https://doi.org/10.1109/ic-ETITE58242.2024.10493748>

- Thavareesan, S., & Mahesan, S. (2020). Sentiment Lexicon Expansion using Word2vec and fastText for Sentiment Prediction in Tamil texts. *2020 Moratuwa Engineering Research Conference (MERCon)*, 272–276. <https://doi.org/10.1109/MERCon50084.2020.9185369>
- Virgana, R. A. E., Sapanji, T., Hamdani, D., & Harahap, P. (2023). Sentiment Analysis of the Top 5 E-commerce Platforms in Indonesia using Text Mining and Natural Language Processing (NLP). In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 7, Issue 2). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Wang, M., & Hu, F. (2021a). The application of nltk *Library* for python natural language processing in corpus research. *Theory and Practice in Language Studies*, *11*(9), 1041–1049. <https://doi.org/10.17507/tpls.1109.09>
- Wang, M., & Hu, F. (2021b). The application of nltk *Library* for python natural language processing in corpus research. *Theory and Practice in Language Studies*, *11*(9), 1041–1049.
- Yang, X., Cai, B., & Xue, Y. (2022). Review on Optimization of Nuclear Power Development: A Cyber-Physical-Social System in Energy Perspective. In *Journal of Modern Power Systems and Clean Energy* (Vol. 10, Issue 3, pp. 547–561). State Grid Electric Power Research Institute Nanjing Branch. <https://doi.org/10.35833/MPCE.2021.000272>
- Zai, C., & Rahman Isnain, A. (n.d.). *Komparasi Algoritma Naïve Bayes dan Support Vector Machine (SVM) pada Analisis Sentimen Capcut*. *9*(1), 2024.