

**ANALISIS SENTIMEN TERHADAP APLIKASI SIREKAP
MOBILE BERBASIS ASPEK PADA PEMILIHAN
UMUM 2024 DENGAN METODE
NAÏVE BAYES**

SKRIPSI

**ANANDA MUHARRIZ SINAGA
171402052**



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA**

2024

**ANALISIS SENTIMEN TERHADAP APLIKASI SIREKAP
MOBILE BERBASIS ASPEK PADA PEMILIHAN
UMUM 2024 DENGAN METODE
NAÏVE BAYES**

SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah Sarjana
Teknologi Informasi

ANANDA MUHARRIZ SINAGA

171402052



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
2024**

PERSETUJUAN

Judul : Analisis Sentimen Terhadap Aplikasi Sirekap Mobile
Berbasis Aspek Pada Pemilihan Umum 2024 Dengan
Metode Naïve Bayes

Kategori : Skripsi

Nama Mahasiswa : Ananda Muharriz Sinaga

Nomor Induk Mahasiswa : 171402052

Program Studi : Sarjana (S-1) Teknologi Informasi

Fakultas : Ilmu Komputer dan Teknologi Informasi
Universitas Sumatera Utara

Medan, 12 Juli 2024

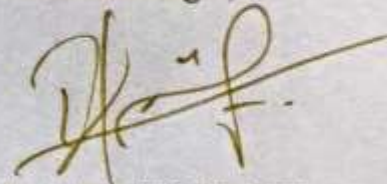
Komisi Pembimbing:

Pembimbing 2,



Sarah Purnamawati S.T., M.Sc.
NIP. 198302262010122003

Pembimbing 1,



Dedy Arisandi S.T., M.Kom.
NIP. 197908312009121002

Diketahui/disetujui oleh
Program Studi S-1 Teknologi Informasi
Ketua,




Dedy Arisandi S.T., M.Kom.
NIP. 197908312009121002

PERNYATAAN

**ANALISIS SENTIMEN TERHADAP APLIKASI SIREKAP MOBILE
BERBASIS ASPEK PADA PEMILIHAN
UMUM 2024 DENGAN METODE
NAÏVE BAYES**

SKRIPSI

Saya Mengakui bahwa skripsi ini adalah karya saya sendiri, kecuali beberapa kutipan dari ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, Juli 2024

Ananda Muharriz Sinaga

171402052

UCAPAN TERIMA KASIH

Puji dan syukur penulis hadirkan kepada Allah SWT yang mana telah memberikan penulis kemampuan dan kesehatan sehingga dapat menyelesaikan skripsi ini sebagai syarat untuk mendapatkan gelar Sarjana Komputer Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Penulis juga ingin mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang senantiasa memberikan doa, bimbingan, dan bantuan kepada penulis mulai dari masa perkuliahan sampai selesainya penyusunan skripsi ini. Rasa terima kasih, dari hari yang paling dalam penulis ucapkan kepada :

1. Kedua orang tua penulis, Ayahanda Alm. Sariaman Sinaga dan Ibunda Nurawati Sitompul karena sudah mendidik, membesarkan dan memberikan dukungan penuh kepada penulis.
2. Saudara kandung penulis, Rakha Adwitya Sinaga yang sudah menemani penulis dalam mengerjakan penulisan skripsi.
3. Keluarga Besar Ibunda Nurawati Sitompul, yang sudah sangat banyak membantu penulis dan keluarga sehingga bisa menyelesaikan masa perkuliahan.
4. Ibu Dr. Maya Silvi Lydia B.Sc., M.Sc. selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi.
5. Bapak Dedy Arisandi S.T., M.Sc. selaku Ketua Program Studi S1 Teknologi Informasi.
6. Bapak Dedy Arisandi S.T., M.Sc. dan Ibu Sarah Purnamawati S.T., M.Sc., selaku pembimbing I dan pembimbing II yang sudah meluangkan waktunya untuk membimbing, mengarahkan dan memberikan masukan serta saran kepada penulis sehingga penulis bisa menyelesaikan penulisan skripsi ini.
7. Seluruh Dosen Program Studi S1 Teknologi Informasi yang sudah memberikan penulis ilmu semasa perkuliahan.
8. Seluruh staff dan pegawai akademik Fakultas Ilmu Komputer dan Teknologi Universitas Sumatera Utara yang sudah membantu penulis dalam urusan administrasi pada masa perkuliahan hingga selesai.
9. Teman-teman seperjuangan penulis, Jackie Chandra, Dinul Iman, Fakhri Rizha Ananda, Ibnu Maulana, M. Rafif Rasyidi, Rafid Miftah, M. Taufik Baskoro,

Varrel Preston Herman Jr, M. Rizki Fatihah, M. Farras Siraj Polem, Ade Rizky, Haqi, Deo Pranata Silitonga, Ulwan Azmi, Prima Julawal, dan M. Fajar Harahap.

10. Teman-teman angkatan 2017 yang sudah membantu penulis saat masa perkuliahan.
11. Dan orang-orang yang sudah membantu penulis mulai dari awal masa perkuliahan sampai selesainya penulisan skripsi ini yang tidak mungkin bisa penulis sebutkan satu persatu.

Semoga seluruh pihak yang terkait senantiasa diberikan Rahmat dan Karunia-Nya dan dimudahkan dalam urusan di dunia. Penulis sadar bahwa banyak kekurangan pada penulisan skripsi ini, oleh karena itu penulis sangat terbuka untuk menerima saran dan kritik yang membangun demi menyempurnakan skripsi ini.

Medan, Juli 2024

Penulis

ABSTRAK

Sirekap Mobile merupakan aplikasi yang digunakan sebagai fasilitas rekapitulasi suara di Tempat Pemungutan Suara (TPS) pada Pemilihan Umum 2024 yang dibuat oleh Komisi Pemilihan Umum (KPU) untuk memudahkan proses rekapitulasi suara. Tetapi, implementasi aplikasi ini memiliki tantangan teknis yang dapat mempengaruhi persepsi publik terhadap KPU. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Sirekap Mobile berdasarkan ulasan pengguna di Google Play Store. Metode yang digunakan adalah *Naïve Bayes* yang nantinya akan mengklasifikasikan sentimen menjadi sentimen positif, netral dan negatif yang dikelompokkan berdasarkan aspek *usability*, *credibility*, dan *availability*. Dataset terdiri dari sejumlah ulasan yang dikumpulkan dari platform Google Play Store yang nantinya akan melalui tahap *preprocessing* sebelum masuk ke proses analisis data. Hasil dari penelitian ini dijelaskan menggunakan evaluasi *confusion matrix* dan perhitungan *accuracy*, *precision*, *recall* dan *f1-score* dengan akurasi model sebesar 85%. Hasil sentimen dari keseluruhan ulasan Sirekap Mobile pada ulasan Google Play Store menunjukkan bahwa aspek *usability* mendapatkan sentimen positif sebanyak 9,77%, netral sebanyak 63,69% dan negatif sebanyak 26,54%. Aspek *credibility* mendapat sentiment positif sebanyak 1,99%, netral sebanyak 92,32% dan negatif sebanyak 5,69%. Aspek *availability* mendapat sentiment positif sebanyak 2,98%, netral sebanyak 55,38% dan negatif sebanyak 41,64%. Penelitian ini akan memberikan wawasan tentang persepsi masyarakat terhadap implementasi teknologi *mobile app* dalam proses berdemokrasi di Indonesia dan menjadi masukan bagi KPU dalam meningkatkan aplikasi Sirekap Mobile untuk Pemilu mendatang, serta bagi peneliti dan praktisi dalam bidang analisis sentimen dan teknologi informasi.

Kata Kunci: Analisis Sentimen, Aplikasi Sirekap Mobile, Pemilu 2024, *Naïve Bayes*, *Aspect Based Sentiment Analysis*

SENTIMENT ANALYSIS OF SIREKAP MOBILE APPLICATION BASED ON ASPECTS IN THE 2024 GENERAL ELECTION USING THE NAÏVE BAYES

ABSTRACT

Sirekap Mobile is an application used for vote tallying at Polling Stations (TPS) in the 2024 General Election, developed by the General Election Commission (KPU) to streamline the vote tallying process. However, its implementation faces technical challenges that may affect public perception of the KPU. This study aims to analyze user sentiment towards the Sirekap Mobile application based on user reviews from the Google Play Store. The Naïve Bayes method is employed to classify sentiments into positive, neutral, and negative categories, focusing on usability, credibility, and availability aspects. The dataset consists of reviews collected from the Google Play Store, which undergo preprocessing before analysis. Results are evaluated using a confusion matrix and metrics such as accuracy, precision, recall, and f1-score, achieving a model accuracy of 85%. The sentiment analysis of overall reviews for Sirekap Mobile on the Google Play Store shows that the usability aspect received positive sentiment from 9.77%, neutral from 63.69%, and negative from 26.54% of reviewers. The credibility aspect received positive sentiment from 1.99%, neutral from 92.32%, and negative from 5.69% of reviewers. The availability aspect received positive sentiment from 2.98%, neutral from 55.38%, and negative from 41.64% of reviewers. This research provides insights into public perception of mobile technology implementation in democratic processes in Indonesia and offers recommendations for KPU to enhance the Sirekap application for future elections, as well as for sentiment analysis researchers and IT practitioners.

Keywords: Sentiment Analysis, Sirekap Mobile Application, 2024 Elections, Naïve Bayes, Aspect-Based Sentiment Analysis

DAFTAR ISI

PERSETUJUAN	iii
PERSETUJUAN PENGGANDAAN SKRIPSI	iii
PERNYATAAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	4
1.6 Metodologi Penelitian	4
1.6.1 Studi Literatur	4
1.6.2 Analisis Permasalahan	4
1.6.3 Perancangan Sistem	4
1.6.4 Dokumentasi dan Penyusunan Laporan	4
1.7 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	6
2.1 Sirekap Mobile	6
2.2 Natural Language Processing	6
2.3 Sentimen Analisis Berbasis Aspek	7
2.4 Term Frequency Inverse Document Frequency	7
2.5 Metode Klasifikasi Naïve Bayes	8
2.6 Stratified K-Fold Validation	8
2.7 Penelitian Terdahulu	9
2.8 Perbedaan Penelitian Terdahulu	14
BAB 3 ANALISIS PERANCANGAN	15
3.1 Dataset	15

3.2 Penentuan Aspek	16
3.3 Arsitektur Umum	17
3.3.1 Data Input	18
3.3.2 Data Pre-processing	20
3.3.3 Transform Case	21
3.3.4 Punctuation Removal	22
3.3.5 Word Stemming	23
3.3.6 <i>Negation Handling</i>	24
3.3.7 Tokenization	25
3.3.8 Stopwords Removal	26
3.3.9 Term Frequency – Inverse Document Frequency	27
3.3.10 Complement Naïve Bayes	29
3.4 Perancangan Sistem	32
3.4.1 Rancangan program	32
3.4.2 Rancangan Antarmuka	33
3.5 Metode Evaluasi	35
BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM	38
4.1 Implementasi Sistem	38
4.1.1 Spesifikasi Perangkat Keras dan Perangkat Lunak	38
4.1.2 Implementasi Perancangan Tampilan Antarmuka	38
4.2 Pelatihan dan Pengujian Model	40
4.2.1 Proses Data Training	41
4.2.2 Uji Model	44
4.3 Evaluasi Model	45
4.3.1 Perhitungan pada aspek <i>usability</i>	49
4.3.2 Perhitungan pada aspek <i>credibility</i>	50
4.3.3 Perhitungan pada aspek <i>availability</i>	50
4.3.4 Perhitungan rata-rata akurasi seluruh aspek	51
4.4 Visualisasi Hasil Analisis Sentimen Aplikasi Sirekap Mobile	51
BAB 5 KESIMPULAN DAN SARAN	54
5.1 Kesimpulan	54
5.2 Saran	54
DAFTAR PUSTAKA	55

DAFTAR TABEL

Tabel 2.1 Tabel penelitian terdahulu	11
Tabel 3.1 Hasil scraping ulasan pengguna Sirekap Mobile	15
Tabel 3.2 Kata kunci tiap aspek	19
Tabel 3.3 Kata kunci tiap sentimen	19
Tabel 3.4 <i>Polarity</i> dataset tiap kelas dalam tiap aspek	19
Tabel 3.5 Contoh data yang sudah dilabeli	20
Tabel 3.6 Contoh proses <i>transform case</i>	21
Tabel 3.7 Contoh proses <i>punctuation removal</i>	22
Tabel 3.8 Contoh proses <i>word stemming</i>	24
Tabel 3.9 Contoh proses <i>negation handling</i>	24
Tabel 3.10 Contoh proses <i>tokenization</i>	25
Tabel 3.11 Contoh proses <i>stopwords removal</i>	27
Tabel 3.12 Contoh pembobotan dengan TF-IDF	28
Tabel 3.13 Contoh dokumen yang akan diklasifikasi	30
Tabel 3.14 Perhitungan komplemen kelas target	31
Tabel 3.15 Confusion Matrix	36
Tabel 4.1 Sampel data latih pada training	41
Tabel 4.2 Sample data latih yang setelah proses preprocessing	42
Tabel 4.3 Rata-rata akurasi dan <i>f1-score</i> aspek <i>usability</i>	43
Tabel 4.4 Rata-rata akurasi dan <i>f1-score</i> aspek <i>availability</i>	43
Tabel 4.5 Rata-rata akurasi dan <i>f1-score</i> aspek <i>credibility</i>	43
Tabel 4.6 Hasil proses uji model	44
Tabel 4.7 <i>Confusion Matrix</i> sentimen negatif <i>usability</i>	46
Tabel 4.8 <i>Confusion Matrix</i> sentimen netral <i>usability</i>	46
Tabel 4.9 <i>Confusion Matrix</i> sentimen positif <i>usability</i>	47
Tabel 4.10 <i>Confusion Matrix</i> sentimen negatif <i>credibility</i>	47
Tabel 4.11 <i>Confusion Matrix</i> sentimen netral <i>credibility</i>	48
Tabel 4.12 <i>Confusion Matrix</i> sentimen positif <i>credibility</i>	48
Tabel 4.13 <i>Confusion Matrix</i> sentimen negatif <i>availability</i>	49
Tabel 4.14 <i>Confusion Matrix</i> sentimen netral <i>availability</i>	49
Tabel 4.15 <i>Confusion Matrix</i> sentimen positif <i>availability</i>	49

DAFTAR GAMBAR

Gambar 3.1 Hasil plot kelompok 1	16
Gambar 3.2 Hasil plot kelompok 2	16
Gambar 3.3 Hasil plot kelompok 3	17
Gambar 3.4 Arsitektur Umum	18
Gambar 3.5 Alur Kerja <i>Pre-processing</i>	20
Gambar 3.6 Alur kerja <i>transform case</i>	21
Gambar 3.7 Alur kerja proses <i>punctuation removal</i>	22
Gambar 3.8 Alur kerja proses <i>word stemming</i>	23
Gambar 3.9 Alur kerja <i>tokenization</i>	25
Gambar 3.10 Alur kerja proses <i>stopwords removal</i>	26
Gambar 3.11 Alur kerja proses <i>TF-IDF</i>	28
Gambar 3.12 Alur kerja Complement Naïve Bayes	30
Gambar 3.13 Pseudocode proses training	32
Gambar 3.14 Alur program	33
Gambar 3.15 Tampilan halaman utama	33
Gambar 3.16 Tampilan halaman untuk <i>upload csv</i>	34
Gambar 3.17 Tampilan halaman untuk <i>manual input</i>	35
Gambar 4.1 Tampilan halaman utama	39
Gambar 4.2 Halaman Upload CSV	39
Gambar 4.3 Tampilan Hasil Upload CSV	40
Gambar 4.4 Tampilan Manual Input	40
Gambar 4.5 <i>Plot</i> hasil uji model aspek <i>usability</i>	46
Gambar 4.6 <i>Plot</i> hasil uji model aspek <i>credibility</i>	47
Gambar 4.7 <i>Plot</i> hasil uji model aspek <i>availability</i>	48
Gambar 4.8 Chart hasil analisis pada aspek <i>usability</i>	52
Gambar 4.9 Chart hasil analisis pada aspek <i>credibility</i>	52
Gambar 4.10 Chart hasil analisis pada aspek <i>availability</i>	53

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indonesia adalah negara demokrasi yang dipimpin oleh presiden dan wakil presiden sebagai kepala negara sekaligus kepala pemerintahan. Presiden dan wakil presiden dipilih oleh rakyat dan memiliki masa jabatan selama lima tahun, lalu dilakukan lagi Pemilihan Umum(Pemilu) untuk menentukan presiden berikutnya. Presiden yang sudah menjabat dapat mencalonkan diri lagi untuk pemilu berikutnya maksimal sebanyak satu kali.

Pemilu diselenggarakan oleh sebuah lembaga negara bernama Komisi Pemilihan Umum (KPU). KPU didirikan pada tahun 1999 dan bersifat independen. Dalam menyelesaikan tugasnya, KPU membentuk badan Ad Hoc yang berguna untuk membantu KPU dalam menyelenggarakan pemilihan. Badan-badan tersebut adalah Panitia Pemilihan Kecamatan (PPK), Panitia Pemilihan Suara (PPS), dan Kelompok Penyelenggara Pemungutan Suara (KPPS). PPK adalah badan yang bertugas untuk melaksanakan Pemilu di tingkat Kecamatan yang berisikan lima orang tokoh masyarakat. Panitia Pemungutan Suara (PPS) merupakan badan yang bertanggung jawab untuk melaksanakan Pemilu di tingkat Kelurahan atau Desa dan berisikan 3 orang dari tokoh masyarakat. Kemudian badan yang terakhir adalah Kelompok Penyelenggara Pemungutan Suara (KPPS). Kelompok ini bertanggung jawab melakukan pemungutan dan perhitungan suara di Tempat Pemungutan Suara (TPS). KPPS berisi 7 orang di setiap TPS, dimana satu orang sebagai ketua merangkap anggota, dan enam lainnya sebagai anggota.

Pada pemilu 2024, terdapat perubahan sistem dalam proses rekapitulasi suara, dimana KPU membuat aplikasi bernama “Sirekap Mobile“, yang akan menjadi sarana yang digunakan oleh KPPS untuk membagikan formulir C-Hasil kepada pengawas dan saksi di Tempat Pemungutan Suara (TPS) yang datanya diverifikasi oleh KPU pusat. Sayangnya pada praktik di lapangan, terjadi beberapa masalah seperti lambatnya proses rekapitulasi, inkonsisten data, dan informasi yang terbatas mengenai proses rekapitulasi suara yang diberikan oleh KPU. Hal-hal tersebut ditakutkan akan menimbulkan keraguan dan kepercayaan publik terhadap aplikasi Sirekap dan KPU (Silvianto, 2024). Meskipun sudah disiapkan secara matang, aplikasi Sirekap masih memiliki kendala

seperti yang terjadi di Kabupaten Tuban. Aplikasi Sirekap dikabarkan tidak bisa digunakan dikarenakan masalah *server* yang dinilai lemot sehingga sulit untuk *login* ke akun Sirekap (Ilahi, 2024). Selain dua kendala yang disebutkan, Sirekap Mobile juga mengalami kendala salah baca angka sehingga suara menggelembung di salah satu pasangan calon dan tidak sesuai dengan yang dituliskan di formulir C-Plano yang berada di TPS (Hidayat, 2024).

Penelitian sebelumnya sudah pernah dilakukan oleh (Amirul, Victor, & Arini, 2019). Amirul *et al* melakukan analisis sentimen terhadap kinerja KPU selama Pemilu 2019 dengan menggunakan komentar-komentar dari akun Facebook KPU. Penelitian ini menggunakan algoritma *K-Means* dan *Confix Stripping Stemmer* untuk memproses data. Hasil penelitian ini menunjukkan tingkat akurasi sebesar 84% untuk metode *K-Means*.

Penelitian selanjutnya dilakukan (Imam, Windu, & Budi, 2019). Penelitian ini fokus pada analisis sentimen terhadap KPU dengan menggunakan data cuitan dari Twitter. Mereka menggunakan algoritma *Support Vector Machine* (SVM) dan melakukan seleksi fitur dengan menggunakan metode *Weight by Correlation*. Hasilnya, akurasi SVM meningkat dari 66.49% menjadi 81.18% setelah penerapan seleksi fitur. Lalu penelitian berikutnya oleh (Nababan, Lumenta, Rindengan, Pontoh, & Akay, 2020). Nababan *et al* menganalisis sentimen masyarakat terhadap hasil Pilpres 2019 berdasarkan cuitan-cuitan Twitter dengan menggunakan metode analisis *lexicon*. Mereka menemukan bahwa sentimen positif terkait hashtag #JokowiAminSudahMenang mencapai 54%, sedangkan sentimen negatif terkait hashtag #PrabowoBukanPemimpin mencapai 51%. Mayoritas cuitan dengan hashtag #JagaDemokrasi dianggap netral dengan presentase 71%.

Penelitian terbaru oleh (Amirullah, Alam, & Sulisty, 2023). Amirullah *et al* melakukan analisis sentimen terhadap kinerja KPU menjelang Pemilu 2024 berdasarkan opini yang terdapat dalam cuitan-cuitan Twitter. Penelitian ini menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan sentimen dari 2475 data yang diambil melalui crawling Twitter. Hasilnya, mereka mencatat 321 data positif, 409 data netral, dan 1745 data negatif, dengan tingkat akurasi model mencapai 74% setelah evaluasi menggunakan *confusion matrix*.

Berdasarkan latar belakang yang sudah dijelaskan, penulis ingin mengajukan penelitian dengan judul “Analisis Sentimen terhadap Aplikasi Sirekap Berbasis Aspek

pada Pemilihan Umum 2024 dengan metode *Naïve Bayes*". Dataset yang akan diteliti akan menggunakan ulasan dari *Google Play Store*.

1.2 Rumusan Masalah

Pada Pemilu 2024, KPU membuat aplikasi bernama "Sirekap Mobile" yang digunakan untuk membantu proses rekapitulasi suara seperti memfoto C-Hasil dan C-Plano, membagikan hasil foto C-Hasil dan C-Plano kepada saksi dan pengawas yang sudah terdaftar dan sebagai alat bantu untuk mencetak formulir sertifikat hasil perolehan suara. Meskipun sudah dikembangkan selama 4 tahun, pada implementasinya Sirekap Mobile masih mengalami kendala seperti sulitnya *login* ke dalam akun, tidak bisa upload foto C-Hasil dan C-Plano, adanya kesalahan baca angka numerik pada dokumen C-Hasil dan C-Plano yang dilakukan oleh sistem sehingga suara membengkok di salah satu pasangan calon presiden. Kendala lainnya juga dapat dianalisa dari ulasan yang berasal dari *Google Play Store*. Untuk itu, dibutuhkan adanya penelitian yang dapat menganalisa ulasan-ulasan *Google Play Store* dari beberapa aspek yang telah ditentukan.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mendapatkan analisis sentimen dari aplikasi Sirekap Mobile terhadap aspek yang telah ditentukan menggunakan ulasan dari *Google Play Store* sebagai dataset penelitian dan *Naïve Bayes* sebagai metode klasifikasi sentimen.

1.4 Batasan Masalah

Batasan masalah yang ditetapkan pada penelitian ini adalah sebagai berikut:

1. Menggunakan data yang akan diambil dengan cara *crawling* pada ulasan *Google Play Store*,
2. Data akan dilabeli berdasarkan aturan *Aspect Based Sentiment Analysis* (ABSA) dan mengabaikan makna kalimat yang bersifat sarkasme dan satire dan menggunakan Bahasa Indonesia yang baku,
3. Ulasan *Google Play Store* yang akan digunakan sebagai dataset penelitian adalah ulasan yang menggunakan bahasa Indonesia,

4. Hasil dari klasifikasi sentimen akan dibagi menjadi beberapa aspek yang telah ditentukan, yaitu aspek *availability*, *usability* dan *credibility*.

1.5 Manfaat Penelitian

Manfaat dari penelitian yang akan dilakukan adalah sebagai berikut:

1. Memberikan informasi kepada Komisi Pemilihan Umum terkait *feedback* dalam implementasi Sirekap Mobile pada Pemilu 2024.
2. Menjadi referensi bagi penelitian selanjutnya yang memiliki masalah penelitian mengenai sentimen analisis maupun kinerja KPU dalam pemilihan umum.

1.6 Metodologi Penelitian

Adapun tahapan-tahapan yang dilakukan dalam membangun sistem dalam penelitian ini adalah :

1.6.1 Studi Literatur

Studi literatur mengenai *Machine Learning* dan *Natural Language Processing* dari buku, jurnal penelitian, dan sumber informasi lain.

1.6.2 Analisis Permasalahan

Analisis Permasalahan terhadap informasi yang sudah didapatkan pada studi literatur yang berhubungan dengan metode *Naïve Bayes* yang akan digunakan pada penelitian ini.

1.6.3 Perancangan Sistem

Melakukan perancangan sistem untuk menyelesaikan masalah yang sudah dianalisis sebelumnya.

1.6.4 Dokumentasi dan Penyusunan Laporan

Penyusunan Laporan dan Dokumentasi dari hasil evaluasi yang sudah dilakukan pada tahap sebelumnya.

1.7 Sistematika Penulisan

Sistematika penulisan pada penelitian ini diuraikan sebagai berikut:

Bab 1: Pendahuluan

Bab 1 berisikan latar belakang diangkatnya judul “ANALISIS SENTIMEN TERHADAP APLIKASI SIREKAP MOBILE BERBASIS ASPEK PADA PEMILIHAN UMUM 2024 DENGAN METODE NAÏVE BAYES”, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

Bab 2: Landasan Teori

Bab 2 memaparkan teori-teori dasar dan konsep-konsep pemecahan masalah yang akan digunakan dalam mengkaji permasalahan pada penelitian ini, terutama terkait dengan analisis sentimen berbasis aspek, ekstraksi fitur dan metode klasifikasi menggunakan *Naïve Bayes*.

Bab 3: Analisis dan Perancangan Sistem

Bab 3 menyajikan penjelasan mengenai arsitektur umum sistem yang dibangun, penerapan ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan metode klasifikasi *Naïve Bayes* dalam implementasi analisis sentimen berbasis aspek pada aplikasi Sirekap Mobile.

Bab 4: Implementasi dan Pembahasan

Bab 4 memaparkan implementasi analisis dan desain sistem. Pada bab ini juga akan dilakukan percobaan berdasarkan hasil dari sistem yang dikembangkan.

Bab 5: Kesimpulan dan Saran

Bab 5 memuat kesimpulan dari keseluruhan hasil penelitian yang telah dilakukan dan saran untuk pengembangan penelitian ini untuk yang akan datang.

BAB 2

LANDASAN TEORI

2.1 Sirekap Mobile

Sirekap Mobile adalah aplikasi yang dikembangkan oleh Komisi Pemilihan Umum (KPU) untuk memfasilitasi pengawasan dan pelaporan hasil penghitungan suara pada Pemilu atau Pilkada secara *real-time*. Aplikasi ini memungkinkan petugas KPPS (Kelompok Penyelenggara Pemungutan Suara) di TPS (Tempat Pemungutan Suara) untuk menginput hasil penghitungan suara secara langsung melalui perangkat mobile (biasanya smartphone atau tablet). Selain untuk melakukan rekapitulasi suara, Sirekap Mobile juga berguna untuk mengelola daftar hadir pengawas TPS dan saksi partai yang berada di dalam TPS. Dengan adanya Sirekap mobile, Pemilihan Umum 2024 diharapkan akan menjadi pemilihan yang transparan dan jauh lebih baik daripada pemilu sebelumnya (KPU, 2021).

2.2 Natural Language Processing

Natural Language Processing atau Pemrosesan bahasa alami adalah proses untuk membuat program memahami maksud dan tujuan dari sebuah teks (Kao & Poteet, 2006). NLP sendiri udah cukup banyak diimplementasikan dalam berbagai bidang seperti pengarsipan dokumen, *chatbot*, klasifikasi dan ekstraksi teks, dan analisis sentimen. Saat ini, penerapan NLP sangat erat kaitannya dengan Pembelajaran Mesin/*Machine Learning*. Meskipun begitu, penerapan NLP juga bisa digunakan menggunakan metode *Lexicon-based* yang tidak menerapkan *machine learning* di dalamnya.

Maraknya sosial media yang memungkinkan penggunanya menuangkan isi pikirannya di internet, dapat kita manfaatkan sebagai sumber acuan dan tolak ukur terhadap suatu kejadian, kebijakan ataupun fenomena alam. Dengan memanfaatkan NLP, kita dapat menganalisa bagaimana masyarakat memberikan respon terhadap kejadian tersebut. Proses analisa ini disebut dengan Sentimen Analisis.

Dalam kasus penelitian ini, NLP akan digunakan untuk menganalisa bagaimana respon masyarakat terhadap aplikasi Sirekap yang baru diterapkan pada Pemilihan Umum 2024. Respon-respon yang nantinya akan menghasilkan kesimpulan, dapat dijadikan bahan pertimbangan dalam perbaikan dan pembenahan aplikasi Sirekap untuk bisa diimplementasikan secara maksimal di Pemilihan Umum selanjutnya, yaitu Pemilu 2029.

2.3 Sentimen Analisis Berbasis Aspek

Sentimen analisis merupakan salah satu bagian dari *Natural Language Processing* (NLP) dimana program atau sistem berusaha mengenali opini dalam bentuk teks. Tujuan dari sentimen analisis adalah memahami pendapat dan emosi seseorang yang dituangkan dalam bentuk teks (Farhadloo & Rolland, 2016). Sentimen Analisis Berbasis Aspek merupakan salah satu pengembangan dari sentimen analisis. Sentimen Analisis Berbasis Aspek sangat baik dalam mengklasifikasi sentimen yang didalamnya terdapat beberapa aspek yang dibahas seperti teks ulasan pada sebuah *website* film, ulasan aplikasi, ulasan tempat makan pada *Google Maps*, dan lain sebagainya. Pada penelitian ini, Sentimen Analisis Berbasis Aspek adalah pilihan yang tepat dikarenakan objek yang akan diteliti adalah sentimen pengguna aplikasi terhadap aplikasi Sirekap Mobile yang mana akan memuat komentar yang berisikan opini pengguna dalam berbagai aspek dari aplikasi.

2.4 Term Frequency Inverse Document Frequency

Term Frequency-Inverse Document Frequency adalah sebuah metode statistik untuk mengukur seberapa penting sebuah kata di dalam dokumen relatif terhadap korpus. Korpus sendiri adalah koleksi teks dalam bahasa sehari-hari yang dapat dibaca oleh komputer. Karena TF-IDF termasuk ke dalam *Vector Space Model* (VSM), maka TF-IDF akan mengubah data teks dokumen menjadi nilai vektor dimana masing-masing nilai vektor mewakili bobot tiap kata di dalam dokumen. Adapun proses dalam pembobotan pada TF-IDF adalah dengan mengalikan *Term Frequency*(TF) dengan *Inverse Document Frequency*(IDF) sebagaimana ditulis dalam persamaan 2.1 berikut:

$$Weight(t, d) = tf(t, d) \times idf(t, D) \quad (2.1)$$

Dimana t adalah kata, d adalah dokumen, D adalah korpus, $tf(t,d)$ adalah frekuensi kata t di dalam dokumen d , dan $idf(t,D)$ adalah pengurangan bobot kata t jika kata memiliki frekuensi kemunculan yang banyak di dalam korpus D .

2.5 Metode Klasifikasi Naïve Bayes

Naïve bayes merupakan kumpulan metode klasifikasi yang diambil dari Teorema Bayes. Metode *Naïve Bayes* merupakan salah satu metode klasifikasi yang mudah untuk diimplementasikan. Selain mudah dalam implementasinya, metode ini juga salah satu metode yang lazim untuk digunakan dalam klasifikasi teks. *Naïve bayes* bekerja berdasarkan probabilitas bersyarat dalam menentukan kelas dari sebuah objek. Contoh dari probabilitas bersyarat adalah pelemparan dua koin. Peluang koin kedua koin untuk sama-sama menunjukkan bagian ekor adalah $\frac{1}{4}$. Peluang koin untuk menunjukkan satu ekor dan satu kepala adalah $\frac{1}{2}$. Probabilitas bersyarat dapat dihitung dengan persamaan 2.2 sebagai berikut:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (2.2)$$

Dimana $P(A/B)$ probabilitas bersyarat A yang diberikan oleh B, $P(B/A)$ adalah probabilitas bersyarat yang diberikan oleh A, $P(A)$ adalah probabilitas kejadian A dan $P(B)$ adalah probabilitas kejadian B.

2.6 Stratified K-Fold Validation

Stratified K-Fold Validation adalah metode validasi data dimana dataset dibagi menjadi k lipatan dan tiap lipatan dipastikan tiap kelas memiliki proporsi yang seimbang. Metode ini memastikan bahwa distribusi kelas pada data latih dan data validasi di setiap iterasi validasi serupa. Adapun langkah dalam proses *stratified k-fold validation* adalah sebagai berikut:

1. Pembagian dataset, dataset akan dibagi secara acak menjadi k lipatan yang sama besar. Misalkan nilai $k = 5$, maka dari 100 dataset yang ada, akan ada 5 kelompok dataset dimana tiap kelompok berisikan 20 data.
2. Stratifikasi, adalah proses dimana tiap lipatan dipastikan proporsi tiap kelas sama dengan proporsi kelas pada dataset asli. Misalkan pada dataset asli terdapat proporsi

kelas A sebanyak 45% dan kelas B sebanyak 55%, maka pada tiap lipatan akan memiliki proporsi yang serupa.

3. Iterasi Validasi, Proses *cross-validation* akan dilakukan sebanyak nilai k dimana tiap perulangan satu lipatan akan dijadikan sebagai data validasi dan data lainnya akan dijadikan data latih. Lipatan yang dijadikan data validasi akan ditukar tiap iterasi sampai seluruh lipatan sudah pernah menjadi data validasi sebanyak satu kali. Setelah proses latih, nilai rata-rata dari metrik evaluasi akan menjadi evaluasi untuk menilai performa model yang dihasilkan

2.7 Penelitian Terdahulu

Penelitian terkait Pemilihan Umum sudah pernah dilakukan, yang meneliti tentang analisis sentimen terhadap kinerja KPU pada Pemilu 2019 dengan judul “Analisis Sentimen Kinerja KPU Pemilu 2019 Menggunakan Algoritma *K-Means* dengan Algoritma *Confix Stripping Stemmer*” oleh (Amirul, Victor, & Arini, 2019). Dalam penelitian yang dilakukan oleh Amirul *et al*, Dataset yang digunakan diambil melalui komentar Facebook KPU. Data yang digunakan pada penelitian ini sebanyak 200 data dan dipisah 150 data sebagai data latih dan 50 data uji. Metode klasifikasi yang digunakan adalah K-Means dengan nilai $k=2$. Tingkat akurasi dari model yang dihasilkan oleh penelitian ini adalah 84% untuk metode *K-Means* dan 86% untuk metode *K-Means* dan dibantu oleh algoritma *Levenshtein Distance* dan *Confix Stripping Stemmer* pada bagian *preprocessing data*.

Penelitian selanjutnya dilakukan oleh (Imam, Windu, & Budi, 2019) dengan judul “Penggunaan *Feature Selection* di Algoritma *Support Vector Machine* untuk Sentimen Analisis Komisi Pemilihan Umum”. Dalam penelitian yang dilakukan Imam *et al*, digunakan data cuitan *Twitter(X)*. Tingkat akurasi dari algoritma SVM sendiri cukup rendah pada penelitian ini, namun setelah dibantu oleh feature selection *Weight by Correlation* akurasi meningkat cukup signifikan dari 66.49% menjadi 81.18%.

Setelah Pemilu 2019 selesai, dilakukan penelitian tentang sentimen masyarakat pengguna twitter terhadap hasil Pemilu yang diumumkan. Penelitian ini dilakukan oleh (Nababan, Lumenta, Rindengan, Pontoh, & Akay, 2020) dengan judul “Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode *Lexicon Analysis*”. Dibandingkan dengan dua penelitian sebelumnya, penelitian ini menggunakan metode *lexicon* sehingga tidak membutuhkan data latih. Hasil dari

penelitian ini adalah sentimen positif yang merupakan cuitan dengan *hashtag* #JokowiAminSudahMenang sebesar 54%, sedangkan sentimen negatif terbanyak terdapat pada cuitan dengan *hashtag* #PrabowoBukanPemimpin sebesar 51%. Sentimen netral berada pada cuitan dengan *hashtag* #JagaDemokrasi sebesar 71%.

Penelitian terkait pemilu 2024 dilakukan oleh (Amirullah, Alam, & Sulisty, 2023) dengan judul “Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini *Twitter* Menggunakan *Naïve Bayes*”. Berdasarkan penelitian yang dilakukan dari 2475 data yang didapat melalui *crawling* pada sosial media *Twitter*(X) sebanyak 321 data merupakan sentimen positif, 409 data merupakan sentimen netral dan 1745 data sentimen negatif. Hasil tersebut didapat melalui tahap klasifikasi menggunakan algoritma *naïve bayes* dan evaluasi data menggunakan *confusion matrix* dengan tingkat akurasi sebesar 74%.

Penelitian yang menggunakan metode Aspect-Based Sentiment Analysis (ABSA) pernah dilakukan oleh (Astuti, Liebenlito, & Fauziah, 2020) dengan judul “Analisis Sentimen Berbasis Aspek Pada Aplikasi Tokopedia Menggunakan LDA dan *Naïve Bayes*”. Penelitian ini menggunakan metode klasifikasi *Naïve Bayes* dan metode *clustering Latent Dirichlett Allocation* (LDA) untuk menentukan aspek yang akan diteliti. Penelitian ini bertujuan untuk membandingkan beberapa metode *resampling* pada dataset dan mendapatkan hasil akurasi 92,5% untuk teknik *RandomOverSampler*.

Berdasarkan referensi yang sudah disebutkan, penulis melakukan penelitian dengan judul “Analisis Sentimen Terhadap Aplikasi Sirekap Mobile Berbasis Aspek Pada Pemilihan Umum 2024 Dengan Metode *Naïve Bayes*”. Penulis akan menggunakan metode klasifikasi *Complement Naïve Bayes* (CNB) sebagai metode klasifikasi dengan pertimbangan penelitian yang dilakukan oleh (GOLDWIN & Priyanta, 2020) dengan judul “Penggunaan *Complement Naïve Bayes* Untuk Memecahkan Masalah Imbalance Dataset Pada *Twitter* Dengan Studi Kasus Jasa Transportasi Online”. Dimana pada penelitian tersebut didapatkan kesimpulan metode CNB memiliki nilai recall dan f1-score yang lebih unggul dibandingkan metode *Multinomial Naïve Bayes* jika diterapkan pada dataset yang tidak seimbang.

Tabel 2.1 Tabel penelitian terdahulu

No	Peneliti	Judul	Tahun	Keterangan
1	Amirul Haj, Victor Amrizal & Arini.	Analisis Sentimen Kinerja KPU Pemilu 2019 Menggunakan Algoritma <i>K-Means</i> dengan Algoritma <i>Confix Stripping Stemmer</i> .	2019	Penelitian ini berfokus pada perbandingan tingkat akurasi dari metode <i>K-Means</i> dalam mengklasifikasikan teks menggunakan dan tanpa normalisasi kata dan <i>stemming</i> . Jika menggunakan normalisasi kata dan <i>stemming</i> , tingkat akurasi meningkat sebesar 2% dari 84% menjadi 86%.
2	Imam Santoso, Windu Gata & Atik Budi Paryanti.	Penggunaan <i>feature selection</i> di Algoritma <i>Support Vector Machine</i> untuk Sentimen Analisis Komisi Pemilihan Umum.	2019	Penelitian yang dilakukan oleh Santoso <i>et al</i> juga berfokus pada perbandingan tingkat akurasi yang didapatkan oleh menggunakan metode <i>Support Vector Machine</i> apabila didukung oleh <i>feature selection</i> . Dengan adanya <i>feature selection</i> , tingkat

					akurasi meningkat dari 66,49% menjadi 81,18%.
3	Adi Pandu Rahmat Nababan, Arie S. M. Lumenta, Yaulie D. Y. Rindengan, Fransisca J. Pontoh & Yuri V. Akay.	Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode <i>Lexicon Analysis</i> .	2019	Penelitian dilakukan pasca Pemilihan Umum 2019 mengenai Hasil Pengumuman Pemilu dengan menggunakan metode <i>lexicon-based</i> dimana program akan menggunakan kamus yang sudah ditentukan skor kata atau kalimat dan dijadikan acuan dalam menentukan skor sentimen pada cuitan <i>twitter</i> .	
4	Fuad Amirullah, Syariful Alam & M. Imam Sulistyo.	Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini <i>Twitter</i> Menggunakan <i>Naïve Bayes</i> .	2023	Penelitian dilakukan sebelum dimulainya Pemilihan Umum 2024 dengan menggunakan data yang berasal dari sosial media <i>twitter</i> dengan cara <i>crawling</i> . Dari 2475 data, 321 data merupakan sentimen positif, 409 data	

					sentimen netral dan 1745 data sentimen negatif.
5	Shinta Prima Astuti, Muhaza Liebenlito & Irma Fauziah	Analisis Sentimen Berbasis Aspek pada Aplikasi Tokopedia Menggunakan LDA dan Naïve Bayes	2020		Topik penelitian ini adalah implementasi Aspect Based Sentiment Analysis (ABSA) dengan menggunakan metode klasifikasi Naïve Bayes dan metode <i>clustering Latent Dirichlett Allocation</i> (LDA) untuk menentukan aspek yang akan diteliti. Penelitian ini bertujuan untuk membandingkan beberapa metode resampling dataset dan mendapatkan hasil akurasi 92,5% untuk teknik <i>RandomOverSampler</i>
6	KEVIN GOLDWIN & Sigit Priyanta	Penggunaan Complement Naïve Bayes Untuk Memecahkan Masalah Imbalance Dataset Pada Twitter	2020		Penelitian ini membahas perbandingan performa dari metode <i>Multinomial Naïve Bayes</i> (MNB) dan <i>Complement Naïve Bayes</i> (CNB) jika

Dengan	Studi	digunakan	pada
Kasus	Jasa	dataset yang	tidak
Transportasi		seimbang.	Hasilnya
Online		CNB lebih	unggul
		pada nilai	<i>recall</i>
		sebesar 72%	dan <i>f1-</i>
		<i>score</i>	74%

2.8 Perbedaan Penelitian Terdahulu

Hal yang membedakan penelitian yang akan dilakukan dibanding penelitian yang sudah dilakukan sebelumnya adalah penelitian yang dilakukan adalah penelitian yang menggunakan aplikasi Sirekap Mobile sebagai objek yang diteliti dan ulasan pengguna aplikasi Sirekap Mobile pada Google Play Store sebagai dataset yang didapat menggunakan teknik *crawling data*. Ulasan pengguna akan dianalisa sentimennya berdasarkan aspek *usability*, *credibility* dan *availability* dan diklasifikasikan menjadi sentimen negatif, positif dan netral terhadap tiap aspek. Penelitian ini akan menggunakan metode *Term Frequency-Inverse Document Frequency* sebagai ekstraksi ciri dan *Complement Naïve Bayes* sebagai metode klasifikasi.

BAB 3

ANALISIS PERANCANGAN

3.1 Dataset

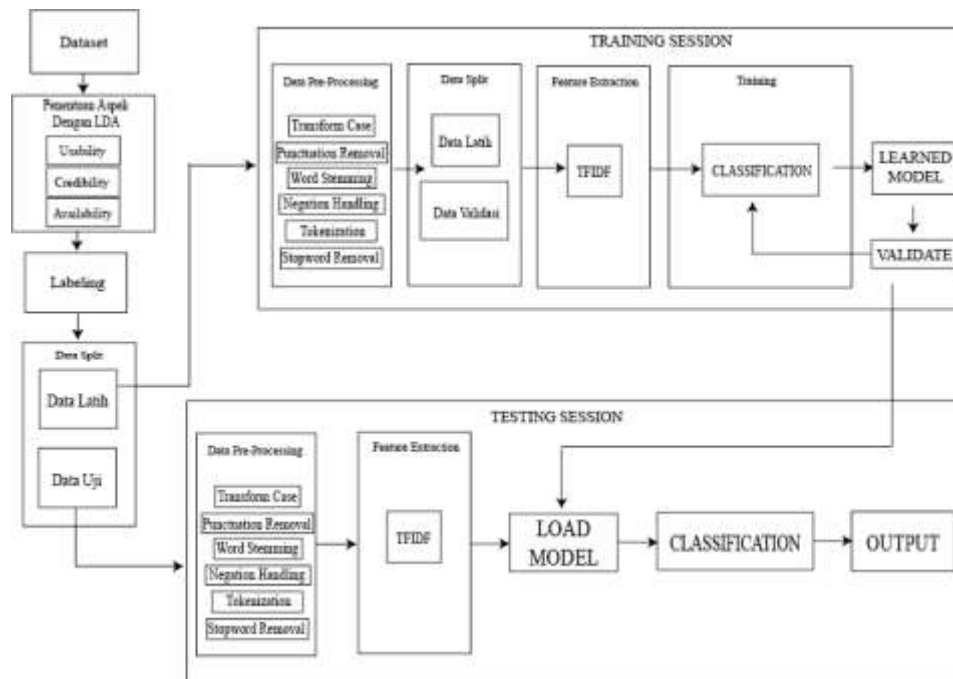
Dataset yang akan digunakan dalam penelitian ini adalah data yang diperoleh dari ulasan pengguna Sirekap Mobile pada situs Google Play Store yang diperoleh menggunakan metode *Web Scraping*. Dataset mulai diambil pada rentang waktu 2 Februari 2024 sampai dengan 16 Februari 2024. Dalam rentang waktu tersebut, data yang berhasil terkumpul sebanyak 2026 data yang disimpan dalam bentuk *Comma Separated Value(.csv)*. Data yang didapat diurutkan berdasarkan rating penilaian Google Play Store dengan skala satu sampai lima. Data tersebut nantinya akan dipisah dengan komposisi 80% sebagai data training dan 20% sebagai data uji. Data training akan digunakan sebanyak 1621 buah dan data uji sebanyak 405 buah. Contoh data dapat dilihat pada Tabel 3.1.

Tabel 3.1 Hasil scraping ulasan pengguna Sirekap Mobile

No.	content
1	MOHON DIBENAH KEMBALI SEBELUM KEMBALI DIOPERASIONALKAN!!! Dan jangan sesekali menyalahkan pengguna yg akan menggunakan untuk menginput data karena masalah dari awal kan harusnya ada uji coba dan kelayakan untuk pemakaian yg merata masa belum siap tapi udh di edarkan untuk menggunakan ke seluruh panitia. Pikir dulu evaluasi dulu dong...anggaran yg digunakan untuk membuat aplikasi hanya sia sia
2	aplikasi sangat membantu tapi blm BISA login
3	Menyenangkan makai aplikasinya
4	Informasinya kacau....banyak kesalahan input dll,,padahal biayanya muahaalll... Pakai uang pajak ,,waduuhh piye iki,,

Berdasarkan kumpulan kata hasil clustering menggunakan LDA, penulis menyimpulkan bahwa kelompok 1 akan menjadi aspek ketersediaan/*availability*, kelompok 2 adalah aspek kegunaan/*usability* dan kelompok 3 adalah aspek kredibilitas/*credibility*. Hasil sentimen yang didapatkan dari ulasan Google Play Store pada aplikasi Sirekap Mobile akan dikelompokkan berdasarkan tiga aspek yang sudah ditentukan. Aspek *usability* akan membahas bagaimana Sirekap Mobile dapat berguna dalam menyelesaikan tugas rekapitulasi suara di TPS dan mengelola daftar hadir saksi dan pengawas TPS. Aspek *credibility* akan membahas mengenai bagaimana Sirekap Mobile dapat membaca angka tulisan tangan dari formulir C-Hasil di TPS secara akurat. Aspek *availability* akan membahas apakah Aplikasi Sirekap dapat diakses dan digunakan pada saat proses rekapitulasi di TPS dimulai pada hari Pemilihan Umum 2024 berlangsung.

Tahap awal pada penelitian ini adalah pengumpulan data teks yang akan disimpan ke dalam format *Comma Separated Value*(.csv) yang nantinya akan diberikan label pada masing-masing datanya. Data yang sudah diberi label akan dipisahkan menjadi 2 bagian yaitu data latih, data uji dengan perbandingan 8:2. Tiap kelompok data akan memasuki tahap *pre-processing* lalu diubah menjadi data numerik menggunakan ekstraksi ciri dengan metode *Term Frequency – Inverse Document Frequency*. Selanjutnya data numerik tersebut akan dimasukkan ke dalam tahap klasifikasi. Gambar 3.4 merupakan gambaran dari arsitektur umum penelitian ini.



Gambar 3.4 Arsitektur Umum

3.3.1 Data Input

Data yang digunakan pada input proses latih dan uji diambil dari ulasan pengguna pada *Google Play Store* yang diambil menggunakan teknik *scraping*. Tiap ulasan akan diberi label pada tiap-tiap aspek. Aspek yang akan diambil adalah aspek *usability*, aspek *credibility* dan aspek *availability*. Aspek *usability* akan berisikan ulasan yang menyinggung tentang apakah aplikasi membantu pengguna dalam menyelesaikan pekerjaan mereka atau malah sebaliknya. Ulasan terkait *UI/UX* dan fitur aplikasi juga termasuk ke dalam aspek *usability*. Aspek *credibility* berisikan ulasan yang membahas tentang keamanan aplikasi dan sentimen masyarakat terhadap netralitas aplikasi di dalam pemilu. Akurasi aplikasi dalam membaca data atau hal-hal yang terkait dengan data perhitungan suara juga termasuk ke dalam aspek *credibility*. Aspek *availability* akan berisikan ulasan yang membahas tentang aspek *server*, *user account*, dan proses *upload* foto. Untuk lebih jelasnya, Tabel 3.2 akan menunjukkan kata kunci/*keyword* yang akan menjadi penentu apakah ulasan tersebut membahas aspek tertentu.

Tabel 3.2 Kata kunci tiap aspek

Aspek	Kata Kunci
Usability	aplikasi, tampilan, fitur, kamera, UI, gambar, penggunaan, sirekap, implementasi
Credibility	data, angka, huruf, baca, hasil, tulisan, masyarakat, input, keamanan, kunci, bukti, scan
Availability	login, upload, server, unggah, username, password, masuk, koneksi, jaringan, inisialisasi, token, respon

Setelah menentukan kategori berdasarkan kata kunci yang sudah ditentukan pada Tabel 3.2, selanjutnya adalah menentukan apakah sentimen terkait aspek tersebut bernilai negatif (-1), netral (0) atau positif (1). Cara menentukan nilai sentimen positif dan negatif adalah dengan menggunakan kata kunci sentimen yang ada pada Tabel 3.3, sedangkan sentimen netral adalah ulasan yang tidak menyinggung aspek apapun di dalam ulasannya. Setelah diberikan label sesuai dengan Tabel 3.3, didapat *polarity* dari dataset penelitian sebagaimana yang dapat dilihat pada Tabel 3.4.

Tabel 3.3 Kata kunci tiap sentimen

Sentimen	Kata Kunci
Positif (1)	efisien, membantu, berguna, mudah, sesuai, cepat, berhasil, tepat, baik, cukup, mantap, lancar, akurat, menyenangkan, keren, oke
Negatif (-1)	invalid, curang, penggelembungan, sulit, lambat, pusing, ribet, down, eror, hilang, menyusahkan, offline, macet, menghambat, gagal

Tabel 3.4 *Polarity* dataset tiap kelas dalam tiap aspek

Kelas	Usability	Credibility	Availability
Negatif (-1)	508	109	797
Netral (0)	1275	1879	1116
Positif (1)	243	38	113

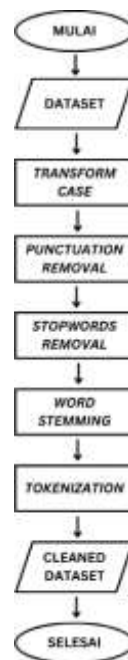
Tabel 3.5 Contoh data yang sudah dilabeli

Ulasan	Aspek		
	Usability	Credibility	Availability
Menyenangkan memakai aplikasinya	1	0	0

Tabel 3.5 adalah contoh data yang sudah dilabeli berdasarkan kata kunci aspek dan sentimen pada Tabel 3.2 dan Tabel 3.3. Dari Tabel 3.5 dapat dilihat ulasan tersebut bernilai “1” pada aspek *usability* dikarenakan di dalam ulasan tersebut terdapat kata “aplikasi” yang menandakan ulasan tersebut masuk ke dalam aspek *usability* dan kata “menyenangkan” yang merupakan kata kunci sentimen positif.

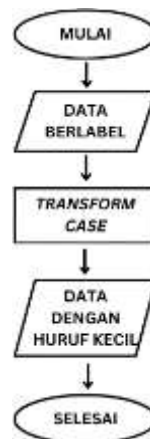
3.3.2 Data Pre-processing

Sebelum data masuk ke dalam proses latih, data akan terlebih dahulu diproses untuk menghilangkan detail yang tidak perlu dan meningkatkan kualitas dari model yang akan dihasilkan. Gambar 3.5 adalah alur kerja/*flowchart* dari proses *pre-processing* dari penelitian ini.

**Gambar 3.5 Alur Kerja *Pre-processing***

3.3.3 Transform Case

Proses *Transform Case* berguna untuk menyeragamkan setiap huruf yang ada menjadi huruf besar atau kecil. Pada penelitian ini setiap data akan diubah menjadi huruf kecil. Alur kerja proses *Transform Case* dapat dilihat pada Gambar 3.6.



Gambar 3.6 Alur kerja *transform case*

Tabel 3.6 adalah contoh data yang melalui proses *transform case*. Hasil dari proses ini adalah data yang seluruh hurufnya berubah menjadi huruf kecil. Hal ini berguna agar pembobotan pada kata “Aplikasi” akan menjadi satu dengan kata “aplikasi”.

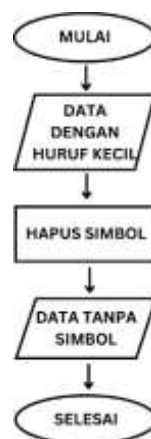
Tabel 3.6 Contoh proses *transform case*

Ulasan sebelum transform case	Ulasan sesudah <i>transform case</i>
Aplikasi ini masih banyak masalah, terutama bagian interface yang masih error, meloginan data yang tiba tiba ter-log out, tidak sinkron antara huruf dan angka, saya saran kan untuk bisa menginput data melalui ketik keyboard tetapi sebagai bukti kalau data tetap	aplikasi ini masih banyak masalah, terutama bagian interface yang masih error, meloginan data yang tiba tiba ter-log out, tidak sinkron antara huruf dan angka, saya saran kan untuk bisa menginput data melalui ketik keyboard tetapi sebagai bukti kalau data tetap

valid menggunakan scan foto, dan bisa di optimalkan kelancaran dan performa aplikasi untuk di berbagai device terutama OS Android 7 dan keatas Terimakasih. . . 😊😊	valid menggunakan scan foto, dan bisa di optimalkan kelancaran dan performa aplikasi untuk di berbagai device terutama os android 7 dan keatas terimakasih. . . 😊😊
---	---

3.3.4 Punctuation Removal

Data hasil proses *transform case* akan memasuki proses *punctuation removal* untuk menghilangkan simbol-simbol yang tidak relevan. Alur kerja pada proses *punctuation removal* dapat dilihat pada Gambar 3.7.



Gambar 3.7 Alur kerja proses *punctuation removal*

Data yang sudah diproses pada tahap *punctuation removal* dapat dilihat pada Tabel 3.7. Simbol-simbol seperti *emoticon*, dan tanda titik akan dihilangkan karena tidak relevan dengan sentimen pada ulasan.

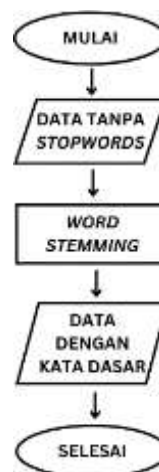
Tabel 3.7 Contoh proses *punctuation removal*

Ulasan sebelum punctuation removal	Ulasan sesudah punctuation removal
Aplikasi ini masih banyak masalah, terutama bagian interface yang masih	aplikasi ini masih banyak masalah terutama bagian interface yang masih

<p>error, meloginan data yang tiba tiba terlog out, tidak sinkron antara huruf dan angka, saya saran kan untuk bisa menginput data melalui ketik keyboard tetapi sebagai bukti kalau data tetap valid menggunakan scan foto, dan bisa di optimalkan kelancaran dan performa aplikasi untuk di berbagai device terutama OS Android 7 dan keatas Terimakasih. . . 😊😊</p>	<p>error meloginan data yang tiba tiba terlog out tidak sinkron antara huruf dan angka saya saran kan untuk bisa menginput data melalui ketik keyboard tetapi sebagai bukti kalau data tetap valid menggunakan scan foto dan bisa di optimalkan kelancaran dan performa aplikasi untuk di berbagai device terutama os android 7 dan keatas terimakasih</p>
--	--

3.3.5 Word Stemming

Proses selanjutnya adalah *word stemming*, yaitu merubah semua kata menjadi kata dasar. Alur proses dari *word stemming* dapat dilihat pada Gambar 3.8 dan contoh data yang memasuki proses *word stemming* dapat dilihat pada Tabel 3.8



Gambar 3.8 Alur kerja proses *word stemming*

Tabel 3.8 Contoh proses *word stemming*

Ulasan sebelum <i>word stemming</i>	Ulasan sesudah <i>word stemming</i>
aplikasi ini masih banyak masalah terutama bagian interface yang masih error meloginan data yang tiba tiba terlog out tidak sinkron antara huruf dan angka saya saran kan untuk bisa menginput data melalui ketik keyboard tetapi sebagai bukti kalau data tetap valid menggunakan scan foto dan bisa di optimalkan kelancaran dan performa aplikasi untuk di berbagai device terutama os android 7 dan keatas terimakasih	aplikasi ini masih banyak masalah utama bagi interface yang masih error meloginan data yang tiba tiba log out tidak sinkron antara huruf dan angka saya saran kan untuk bisa menginput data lalu ketik keyboard tetapi bagi bukti kalau data tetap valid guna scan foto dan bisa di optimal lancar dan performa aplikasi untuk di bagi device utama os android 7 dan atas terimakasih

3.3.6 *Negation Handling*

Negation handling pada penelitian ini akan menggunakan metode penggabungan kata negasi seperti “tidak”, “belum”, “jangan”, dan “bukan” dengan kata setelahnya. Penggabungan ini bertujuan untuk memisahkan kata sentimen negatif dan positif agar tidak dihitung sebagai satu kata pada proses pembobotan. Proses *negation handling* dapat dilihat pada Tabel 3.9

Tabel 3.9 Contoh proses *negation handling*

Ulasan sebelum <i>negation handling</i>	Ulasan sesudah <i>negation handling</i>
aplikasi ini masih banyak masalah utama bagi interface yang masih error meloginan data yang tiba tiba log out tidak sinkron antara huruf dan angka saya saran kan untuk bisa menginput data lalu ketik keyboard tetapi bagi bukti kalau data tetap valid guna scan foto dan bisa di optimal lancar dan	aplikasi ini masih banyak masalah utama bagi interface yang masih error meloginan data yang tiba tiba log out tidak_sinkron antara huruf dan angka saya saran kan untuk bisa menginput data lalu ketik keyboard tetapi bagi bukti kalau data tetap valid guna scan foto dan bisa di optimal lancar dan

performa aplikasi untuk di bagai device utama os android 7 dan atas terimakasih	performa aplikasi untuk di bagai device utama os android 7 dan atas terimakasih
---	---

3.3.7 Tokenization

Data yang sudah bersih, akan masuk ke dalam proses *tokenization* untuk dipenggal menjadi kata per kata. Gambar 3.9 adalah alur kerja dari proses *tokenization*.



Gambar 3.9 Alur kerja *tokenization*

Penggalan kata nantinya akan lebih mudah diproses pada tahap selanjutnya. Sebagai contoh, Tabel 3.10 akan menggambarkan proses *tokenization* yang terjadi pada program.

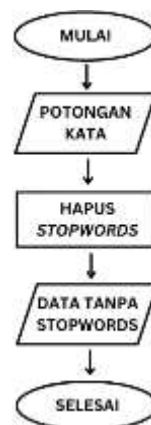
Tabel 3.10 Contoh proses *tokenization*

Ulasan sebelum <i>tokenization</i>	Ulasan sesudah <i>tokenization</i>
aplikasi ini masih banyak masalah utama bagi interface yang masih error meloginan data yang tiba tiba log out tidak_sinkron antara huruf dan angka saya saran kan untuk bisa menginput	'aplikasi', 'ini', 'masih', 'banyak', 'masalah', 'utama', 'bagi', 'interface', 'yang', 'masih', 'error', 'meloginan', 'data', 'yang', 'tiba', 'tiba', 'log', 'out', 'tidak_sinkron', 'antara', 'huruf', 'dan',

data lalu ketik keyboard tetapi bagi bukti kalau data tetap valid guna scan foto dan bisa di optimal lancar dan performa aplikasi untuk di bagi device utama os android 7 dan atas terimakasih	'angka', 'saya', 'saran', 'kan', 'untuk', 'bisa', 'menginput', 'data', 'lalu', 'ketik', 'keyboard', 'tetapi', 'bagai', 'bukti', 'kalau', 'data', 'tetap', 'valid', 'guna', 'scan', 'foto', 'dan', 'bisa', 'di', 'optimal', 'lancar', 'dan', 'performa', 'aplikasi', 'untuk', 'di', 'bagai', 'device', 'utama', 'os', 'android', '7', 'dan', 'atas', 'terimakasih'
--	--

3.3.8 Stopwords Removal

Setelah selesai diproses pada tahap *tokenization*, data yang sudah dipenggal per kata akan masuk tahap *stopwords removal* dimana kata yang memiliki frekuensi tinggi namun tidak atau sedikit memiliki makna dalam ulasan akan dihilangkan seperti kata “yang”, “dan”, “tetapi” dan seterusnya. Gambar 3.10 adalah alur kerja dari proses *stopwords removal* dan tabel 3.11 adalah contoh data yang memasuki proses *stopwords removal*.



Gambar 3.10 Alur kerja proses *stopwords removal*

Tabel 3.11 Contoh proses *stopwords removal*

Ulasan sebelum stopwords removal	Ulasan sesudah stopwords removal
'aplikasi', 'ini', 'masih', 'banyak', 'masalah', 'utama', 'bagi', 'interface', 'yang', 'masih', 'error', 'meloginan', 'data', 'yang', 'tiba', 'tiba', 'log', 'out', 'tidak_sinkron', 'antara', 'huruf', 'dan', 'angka', 'saya', 'saran', 'kan', 'untuk', 'bisa', 'menginput', 'data', 'lalu', 'ketik', 'keyboard', 'tetapi', 'bagai', 'bukti', 'kalau', 'data', 'tetap', 'valid', 'guna', 'scan', 'foto', 'dan', 'bisa', 'di', 'optimal', 'lancar', 'dan', 'performa', 'aplikasi', 'untuk', 'di', 'bagai', 'device', 'utama', 'os', 'android', '7', 'dan', 'atas', 'terimakasih'	'aplikasi', 'banyak', 'masalah', 'terutama', 'bagian', 'interface', 'error', 'meloginan', 'data', 'tiba', 'tiba', 'terlog', 'out', 'tiba', 'tidak_sinkron', 'huruf', 'angka', 'saran', 'menginput', 'data', 'melalui', 'ketik', 'keyboard', 'bukti', 'data', 'tetap', 'valid', 'menggunakan', 'scan', 'foto', 'optimalkan', 'kelancaran', 'performa', 'aplikasi', 'berbagai', 'device', 'terutama', 'os', 'android', '7', 'keatas', 'terimakasih'

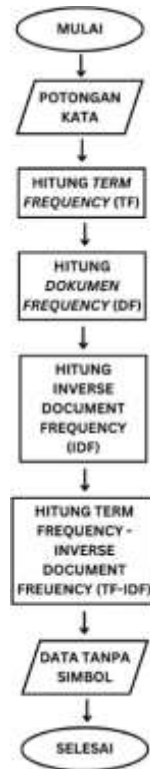
3.3.9 Term Frequency – Inverse Document Frequency

Term Frequency-Inverse Document Frequency adalah sebuah metode statistik untuk mengukur seberapa penting sebuah kata di dalam dokumen relatif terhadap korpus. Korpus sendiri adalah koleksi teks dalam bahasa sehari-hari yang dapat dibaca oleh komputer. Karena TF-IDF termasuk ke dalam *Vector Space Model* (VSM), maka TF-IDF akan mengubah data teks dokumen menjadi nilai vektor dimana masing-masing nilai vektor mewakili bobot tiap kata di dalam dokumen. Pembobotan dilakukan dengan menggunakan *unigram*, *bigram* dan *trigram*. Adapun proses dalam pembobotan pada TF-IDF adalah dengan mengalikan *Term Frequency* (TF) dengan *Inverse Document Frequency* (IDF) sebagaimana ditulis dalam rumus berikut:

$$Weight(t, d) = tf(t, d) \times idf(t, D)$$

Dimana t adalah kata, d adalah dokumen, D adalah korpus, $tf(t, d)$ adalah frekuensi kata t di dalam dokumen d , dan $idf(t, D)$ adalah pengurangan bobot kata t jika kata memiliki

frekuensi kemunculan yang banyak di dalam korpus D . Gambar 3.11 akan menunjukkan tentang alur kerja dari proses TF - IDF dalam penelitian ini.



Gambar 3.11 Alur kerja proses TF - IDF

Perhitungan bobot kata menggunakan TF , DF dan IDF dapat dilihat pada Tabel 3.12 apabila diketahui beberapa dokumen sebagai berikut:

- d1 : aplikasi bagus, tidak_bisa login
- d2 : aplikasi membantu
- d3 : data hasil scan tidak_sesuai, sulit login

Tabel 3.12 Contoh pembobotan dengan TF - IDF

Term(t)	TF			Df_i	IDF	TF-IDF		
	d1	d2	d3			d1	d2	d3
aplikasi	1	1	0	2	$\log \frac{5}{2} = 0.23$	0.23	0.23	0
bagus	1	0	0	1	$\log \frac{3}{1} = 0.47$	0.47	0	0
tidak_bisa	1	0	0	1	$\log \frac{3}{1} = 0.47$	0.47	0	0

login	1	0	1	2	$\log \frac{3}{2} = 0.23$	0.23	0	0.23
membantu	0	1	0	1	$\log \frac{3}{1} = 0.47$	0	0.47	0
data	0	0	1	1	$\log \frac{3}{1} = 0.47$	0	0	0.47
hasil	0	0	1	1	$\log \frac{3}{1} = 0.47$	0	0	0.47
scan	0	0	1	1	$\log \frac{3}{1} = 0.47$	0	0	0.47
tidak_sesuai	0	0	1	1	$\log \frac{3}{1} = 0.47$	0	0	0.47
sulit	0	0	1	1	$\log \frac{3}{1} = 0.47$	0	0	0.47

3.3.10 Complement Naïve Bayes

Setelah melalui pembobotan menggunakan metode TF-IDF, selanjutnya data akan masuk ke dalam proses latih data. Dalam klasifikasinya, penelitian ini akan menggunakan metode *Complement Naïve Bayes* dengan pertimbangan kelas dari tiap data yang kurang seimbang. Cara kerja dari algoritma *Complement Naïve Bayes* adalah dengan menghitung kemungkinan sebuah item berada di dalam setiap kelas yang ada selain target (komplemen kelas). Secara umum, teorema Bayes memiliki rumus sebagai berikut:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

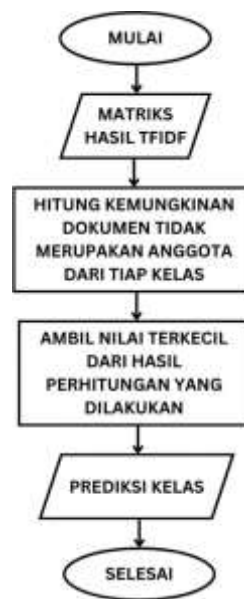
Dimana $P(A|B)$ adalah peluang terjadinya A jika kejadian B terjadi, $P(A)$ adalah peluang terjadinya kejadian A, $P(B)$ adalah peluang terjadinya kejadian B, $P(B|A)$ adalah peluang terjadinya B jika kejadian A sudah terjadi. Pada algoritma klasifikasi Naïve Bayes, digunakan formula berikut untuk menghitung kemungkinan sebuah kelas merupakan bagian dari kelas tertentu:

$$\operatorname{argmax} p(y). \prod p(w|y)^{f_i}$$

Dimana f_i merupakan banyaknya suatu kata yang muncul dalam suatu kalimat, kemudian dilakukan prediksi terhadap sebuah kelas tertentu menggunakan $p(w|y)$. Sedangkan pada *complement naïve bayes*, sebuah kejadian atau item akan dihitung kemungkinannya berada dalam kelas selain kelas target dengan rumus:

$$\operatorname{argmin} p(y) \cdot \prod \frac{1}{p(w|y')^{f_i}}$$

Dimana y' adalah komplemen kelas yang diprediksi dan f_i adalah frekuensi kemunculan kata dalam sebuah dokumen. Alur kerja dari proses prediksi kelas dari metode *Complement Naïve Bayes* dapat dilihat pada Gambar 3.12



Gambar 3.12 Alur kerja *Complement Naïve Bayes*

Untuk melihat bagaimana perhitungan CNB bekerja, kita dapat lihat Tabel 3.13 sebagai data yang akan diklasifikasi sebagai sentimen negatif, netral atau positif dengan dengan aspek *usability*.

Tabel 3.13 Contoh dokumen yang akan diklasifikasi

No	Kata	Sentimen
Dokumen 1	aplikasi, membantu, kamera, jernih	Positif
Dokumen 2	aplikasi, menyusahkan, sulit, login	Negatif
Dokumen 3	aplikasi, lumayan, membantu	Positif
Dokumen 4	aplikasi, membantu, mantap	?

Dari tabel 3.12, kita dapat menghitung *prior probabilities* dengan cara sebagai berikut.

$$P(Ps) = \frac{2}{3} \quad P(N) = \frac{1}{3}$$

Dimana $P(P)$ merupakan *prior probabilities* dari kelas sentimen positif dan $P(N)$ merupakan *prior probabilities* dari kelas sentimen negatif. Kemudian dilakukan perhitungan untuk setiap kelas yang bukan merupakan kelas target yang dapat dilihat pada Tabel 3.14

Tabel 3.14 Perhitungan komplemen kelas target

Sentimen Positif	Sentimen Negatif
$P(\text{aplikasi} Ps^{\wedge}) = \frac{1}{4}$	$P(\text{aplikasi} N^{\wedge}) = \frac{2}{7}$
$P(\text{membantu} Ps^{\wedge}) = \frac{0}{4}$	$P(\text{membantu} N^{\wedge}) = \frac{0}{7}$
$P(\text{mantap} Ps^{\wedge}) = \frac{0}{4}$	$P(\text{mantap} N^{\wedge}) = \frac{0}{7}$

Untuk menghindari pembagian nilai nol, akan dilakukan smoothing pada kata yang memiliki frekuensi nol dengan menggunakan *laplace smoothing* dengan nilai $\alpha = 1$.

Smoothing pada sentimen positif:

- $P(\text{aplikasi}|Ps^{\wedge}) = \frac{1+1}{4+2(1)} = \frac{2}{6}$
- $P(\text{membantu}|Ps^{\wedge}) = \frac{0+1}{4+2(1)} = \frac{1}{6}$
- $P(\text{mantap}|Ps^{\wedge}) = \frac{0+1}{4+2(1)} = \frac{1}{6}$

Smoothing pada sentimen negatif:

- $P(\text{aplikasi}|N^{\wedge}) = \frac{2+1}{7+2(1)} = \frac{2}{9}$
- $P(\text{membantu}|N^{\wedge}) = \frac{0+1}{7+2(1)} = \frac{1}{9}$
- $P(\text{mantap}|N^{\wedge}) = \frac{0+1}{7+2(1)} = \frac{1}{9}$

Kemudian dilakukan perhitungan kemungkinan dokumen 4 untuk setiap kelas menggunakan persamaan sebagai berikut:

$$\argmin p(y) \cdot \prod \frac{1}{p(w|y')^{f_i}}$$

Perhitungan untuk sentimen positif:

$$P(Ps|D4) = \frac{P(Ps)}{P(\text{aplikasi}|Ps^{\wedge}) \times P(\text{membantu}|Ps^{\wedge}) \times P(\text{mantap}|Ps^{\wedge})}$$

$$\begin{aligned}
&= \frac{\frac{2}{3}}{\frac{2}{6} \times \frac{1}{6} \times \frac{1}{6}} \\
&= \frac{2}{3} \times \frac{36}{2} \\
&= 12
\end{aligned}$$

Perhitungan untuk sentimen negatif:

$$\begin{aligned}
P(Ps|D4) &= \frac{P(Ps)}{P(\text{aplikasi}|Ps^{\wedge}) \times P(\text{membantu}|Ps^{\wedge}) \times P(\text{mantap}|Ps^{\wedge})} \\
&= \frac{\frac{1}{3}}{\frac{2}{9} \times \frac{1}{9} \times \frac{1}{9}} \\
&= \frac{1}{3} \times \frac{792}{2} \\
&= 132
\end{aligned}$$

Dari perhitungan yang sudah dilakukan didapat hasil perhitungan untuk sentimen positif adalah 12 sedangkan hasil perhitungan untuk sentimen negatif adalah 132. Secara otomatis, CNB akan mengambil nilai terkecil yaitu sentimen positif sebagai kelas prediksi. Gambar 3.13 adalah pseudo code proses training data jika menggunakan Stratified K-Fold Validation.

```

Begin

    k = 50

    fold = StratifiedKFold(dataset_split(k)

    For x in fold

        CNB.train(fold_data_train)

        CNB.predict(fold_data_val)

        train_val_switch(fold)

    End

```

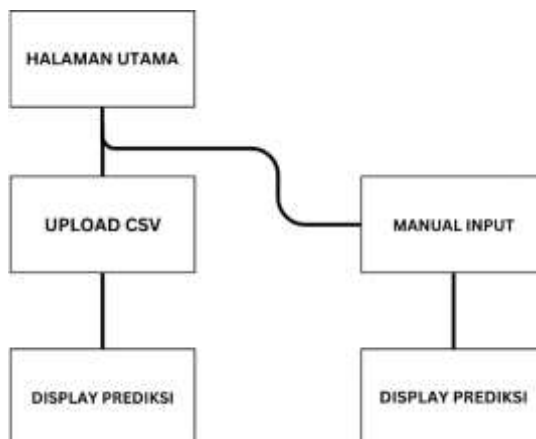
Gambar 3.13 Pseudocode proses training

3.4 Perancangan Sistem

Perancangan antarmuka sistem dan alur kerja sistem dibutuhkan untuk mempermudah pengguna menggunakan aplikasi yang dihasilkan dari penelitian ini.

3.4.1 Rancangan program

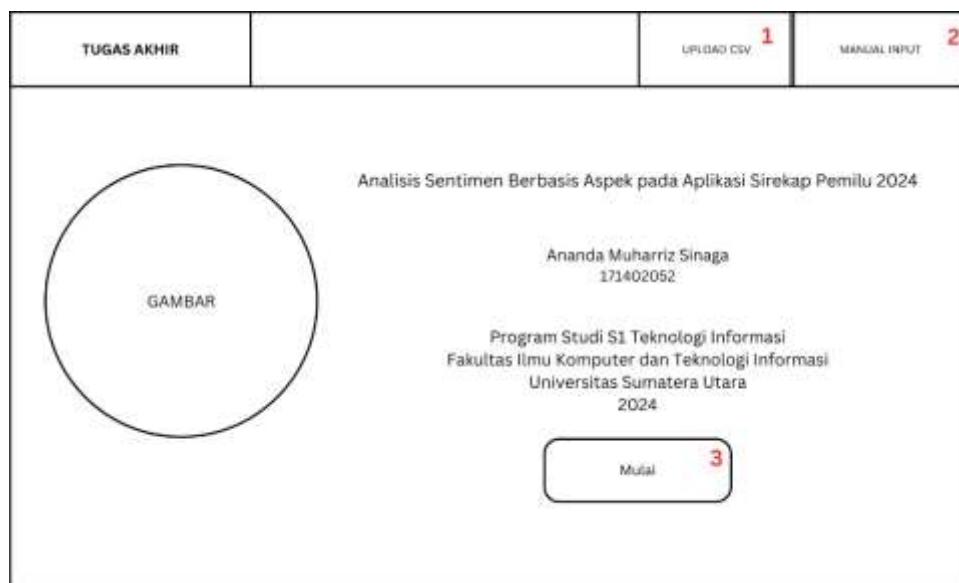
Alur program dibutuhkan untuk mengetahui bagaimana urutan sebuah program berjalan. Gambar 3.14 merupakan alur program dari aplikasi pada penelitian ini.



Gambar 3.14 Alur program

3.4.2 Rancangan Antarmuka

Rancangan antarmuka merupakan penggambaran tentang bagaimana antarmuka aplikasi akan terlihat setelah dikembangkan. Gambar 3.15 adalah penggambaran halaman utama dari aplikasi yang akan dibuat pada penelitian ini.



Gambar 3.15 Tampilan halaman utama

Keterangan :

1. Tombol (1) berfungsi untuk masuk ke halaman *upload csv*
2. Tombol (2) berfungsi untuk masuk ke halaman *manual input*

3. Tombol (3) berfungsi untuk masuk ke halaman *upload csv*

Gambar 3.16 adalah tampilan rancangan untuk halaman pengujian model yang berfungsi untuk melakukan uji coba pada model yang telah dihasilkan pada proses latih. Pada halaman ini, pengujian dilakukan dalam data majemuk yang dimuat dalam file dengan format *Comma Separated Value (csv)*.

The wireframe shows a web interface for uploading CSV files. At the top, there is a navigation bar with three sections. The first section on the left contains the text 'TUGAS AKHIR' and a red number '1'. The middle section is empty. The third section on the right contains a button labeled 'UPLOAD CSV' and the text 'MANUAL INPUT' with a red number '2'. Below the navigation bar, the main content area is titled 'Upload CSV'. It contains a large rectangular box with a red number '3' inside, representing a file upload area. Below this box is a button labeled 'Start' with a red number '4'. To the right of the 'Start' button, the text 'Hasil Klasifikasi Sentimen' is displayed with a red number '5'. Below this text is a table with three columns: 'Utama', 'Aktual', and 'Prediksi'. The table has four rows, including a header row.

Utama	Aktual	Prediksi

Gambar 3.16 Tampilan halaman untuk *upload csv*

Keterangan :

1. Tombol (1) berfungsi untuk kembali ke halaman utama
2. Tombol (2) berfungsi untuk masuk ke halaman *manual input*
3. Tombol (3) berfungsi untuk *upload file* dengan format csv
4. Tombol (4) berfungsi untuk memulai prediksi
5. Nomor (5) berfungsi untuk menampilkan hasil prediksi

Gambar 3.17 adalah tampilan rancangan untuk menguji model terhadap data tunggal yang diinput langsung pada aplikasi.

Gambar 3.17 Tampilan halaman untuk *manual input*

Keterangan :

1. Tombol (1) berfungsi untuk kembali ke halaman utama
2. Tombol (2) berfungsi untuk masuk ke halaman *upload csv*
3. Tombol (3) berfungsi untuk menginput sebuah *string*/kalimat ulasan
4. Tombol (4) berfungsi untuk memulai prediksi
5. Nomor (5) berfungsi untuk menampilkan hasil prediksi

3.5 Metode Evaluasi

Evaluasi model sangat diperlukan untuk mengetahui apakah suatu model sudah bisa dikatakan optimal atau belum. Pada penelitian ini metode evaluasi *Confusion Matrix* akan digunakan dalam menilai kinerja model yang telah dihasilkan. Tabel 3.15 merupakan matriks yang berisi elemen penting untuk mendapatkan nilai akurasi, presisi, *recall*, dan *F1-score*.

Tabel 3.15 Confusion Matrix

		Nilai Sebenarnya/Aktual	
		Positif	Negatif
Nilai Prediksi	Positif	True Positive (TP)	False Positive (FP)
	Negatif	False Negative (FN)	True Negative (TN)

Keterangan :

True Positive (TP) : kondisi dimana model memprediksi positif terhadap data dengan label positif

False Positive (FN) : kondisi dimana model memprediksi negatif terhadap data dengan label positif

True Negative (TN) : kondisi dimana model memprediksi negatif terhadap data dengan label negatif

False Negative (FN) : kondisi dimana model memprediksi positif terhadap data dengan label negatif

Dari matriks pada Tabel 3.12, kita dapat menghitung nilai akurasi, presisi, *recall* dan *F1-Score* dari sebuah model. Akurasi adalah nilai yang menunjukkan seberapa baik sebuah model dalam memprediksi nilai dengan benar. Persamaan 3.1 dan 3.2 adalah perhitungan untuk mendapatkan nilai akurasi dari sebuah model.

$$Accuracy = \frac{\text{Jumlah prediksi benar}}{\text{Jumlah prediksi yang dilakukan}} \times 100\% \quad (3.1)$$

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \times 100\% \quad (3.2)$$

Presisi adalah tingkat kualitas dari prediksi positif yang dilakukan oleh model. Persamaan 3.3 adalah perhitungan untuk mendapatkan nilai presisi.

$$Precision = \frac{TP}{(TP+FP)} \times 100\% \quad (3.3)$$

Recall adalah seberapa sering suatu model menemukan nilai *True Positive* (TP) dari seluruh kelas positif. Persamaan 3.4 adalah perhitungan untuk mendapatkan nilai *recall*.

$$Recall = \frac{TP}{(TP+FN)} \times 100\% \quad (3.3)$$

F1-score adalah nilai yang dapat dijadikan acuan apakah model yang dihasilkan sudah optimal atau belum. Model yang sempurna akan memiliki *F1-score* bernilai ‘1’. *F1-score* adalah perhitungan rata-rata dari presisi dan *recall*. Persamaan 3.4 adalah perhitungan untuk mendapatkan nilai *F1-score*.

$$F1 \text{ score} = 2 \times \frac{recall \times precision}{recall + precision} \times 100\% \quad (3.4)$$

BAB 4

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Implementasi Sistem

Perancangan sistem yang sudah dijabarkan pada Bab 3 akan diimplementasikan pada aplikasi Analisis Sentimen Berbasis Aspek terhadap aplikasi Sirekap Mobile pada Pemilu 2024. Pembuatan aplikasi dibuat menggunakan *hardware* dan *Software* sebagai berikut:

4.1.1 Spesifikasi Perangkat Keras dan Perangkat Lunak

Spesifikasi perangkat keras/*hardware* yang digunakan dalam pembuatan aplikasi ini adalah sebagai berikut:

1. Processor Intel Core i5-2410M CPU @ 2.30 GHz
2. Memory RAM 8 GB
3. HDD 500GB

Perangkat lunak/*software* yang digunakan dalam pembuatan aplikasi penelitian ini adalah sebagai berikut:

1. Windows 8.1 Pro 64 *bit*
2. Python 3.7.4 dengan menggunakan library *scikit-learn*, *pandas*, dan *numpy*.
3. Jupyter Notebook
4. Sublime Text 3

4.1.2 Implementasi Perancangan Tampilan Antarmuka

Berikut adalah desain tampilan program penelitian yang sudah dimuat rancangannya pada Bab 3:

1. Tampilan Halaman Utama

Tampilan halaman utama pada sistem dapat dilihat pada Gambar 4.1. Halaman ini adalah halaman yang pertama kali dilihat saat mengakses sistem.



Gambar 4.1 Tampilan halaman utama

2. Tampilan Upload CSV

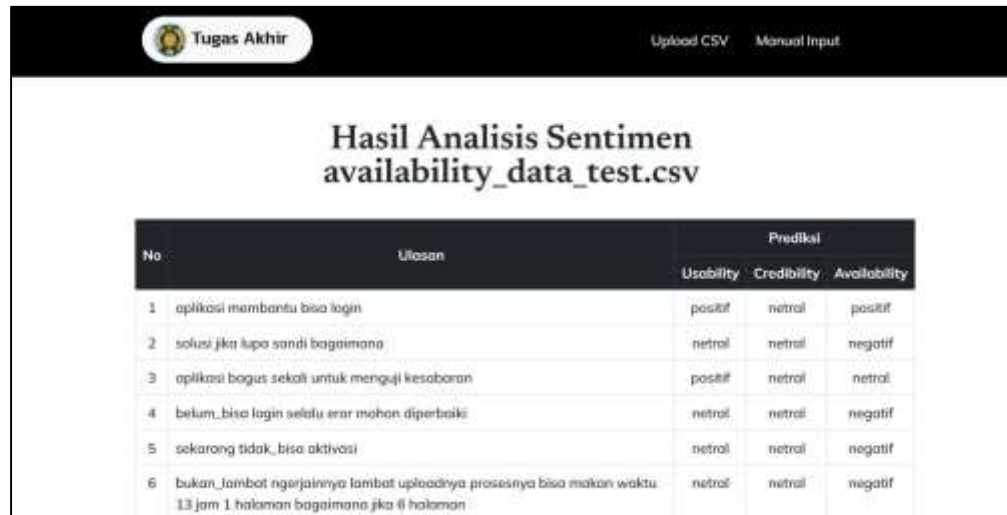
Tampilan halaman upload csv adalah tampilan yang digunakan apabila pengguna akan menggunakan data ulasan yang disimpan file dengan format .csv dan akan melakukan prediksi sentimen. Gambar 4.2 adalah tampilan halaman upload csv.



Gambar 4.2 Halaman Upload CSV

3. Tampilan Hasil Sentimen Upload CSV

Tampilan halaman hasil sentimen upload csv berisi tabel hasil klasifikasi sentimen ulasan yang sebelumnya sudah diupload pada halaman upload csv. Gambar 4.3 merupakan tampilan hasil sentimen upload csv.



No	Ulasan	Prediksi		
		Usability	Credibility	Availability
1	aplikasi membantu bisa login	positif	netral	positif
2	solusi jika lupa sandi bagaimana	netral	netral	negatif
3	aplikasi bagus sekali untuk menguji kesabaran	positif	netral	netral
4	belum bisa login selalu eror mohon diperbaiki	netral	netral	negatif
5	sekarang tidak bisa aktivasi	netral	netral	negatif
6	bukan jambat ngerjainnya lambat uploadnya prosesnya bisa makan waktu 13 jam 1 halaman bagaimana jika 6 halaman	netral	netral	negatif

Gambar 4.3 Tampilan Hasil Upload CSV

4. Tampilan Manual Input

Tampilan manual input berisi halaman berisi *input field* yang dapat diisi dengan ulasan secara manual dan langsung diklasifikasikan sentimennya berdasarkan aspek *usability*, *credibility*, dan *availability*. Gambar 4.4 adalah tampilan untuk halaman manual input.



Tugas Akhir Upload CSV Manual Input

Silahkan Input Ulasan

Aplikasi membantu

Submit

Usability:  positif

Credibility:  netral

Availability:  netral

Gambar 4.4 Tampilan Manual Input

4.2 Pelatihan dan Pengujian Model

Pada bagian ini akan dibahas proses latih model hingga pengujian model menggunakan data uji. Lalu akan dilakukan proses evaluasi model menggunakan *Confusion Matrix* dan perhitungan *accuracy*, *precision*, *recall*, dan *F1-Score*.

4.2.1 Proses Data Training

Pada proses ini dilakukan latih data sekaligus validasi dengan metode *Stratified K-Fold Validation* dengan nilai $k = 50$ pada latih aspek *usability* dan *availability* dan nilai $k = 10$ untuk latih aspek *credibility*. Tabel 4.1 adalah sampel data pada proses latih model.

Tabel 4.1 Sampel data latih pada training

No	Ulasan	Usability
1	tidak NIAT BANGET BIKIN APLIKASI , KAMERA 108 MEGAPIXEL JADI buram GARA-GARA FITUR KAMERA APLIKASI BURIK , KASIH AJA FITUR UPLOAD BIAR GAMBARNYA JELAS, TERUS APAAN COBA CROP MANUAL , MINIMAL SEPERTI CAMSCANNER BISA NGEPASIN LANGSUNG JADI tidak lambat-lambat sudah WAKTUNYA NGITUNG MASIH FOTONYA MEMPERLAMBAT PROSES PERHITUNGAN , MINIMAL ANGGARAN BUAT BIKIN SERVER JANGAN DISUNAT, DARI SIMULASI TERAKHIR BANYAK DATA yang SALAH BACA DARI PADA YANG benar, APALAGI UPLOADNYA lambat , PERSIAPAN 5 TAHUN NGAPAIN AJA, NOL BESAR	-1
2	Ya, apa yang bisa diharapkan dari KPU, developer dan pemerintah berwenang yang dari awalnya saja sudah mendukung kecurangan, melanggar etika, diam saat tahu ada yang salah. Get well soon, Indonesia! 1. Aplikasi tidak siap menampung "crowd system" di hari-H. 2. Server lemah dan bermasalah. Tidak ada limit atas jumlah suara sesuai jml maksimal DPT (300)	0
3	Server untuk semua hp tidak bisa	0
4	aplikasi Bagus dan membantu	1
5	kamera sudah jernih	1
6	Kalau bisa dipermudah cara login nya	0
7	Aplikasi Lambat	-1

8	Sangat membantu sekali aplikasinya	1
9	Kenapa tidak dapat mendapatkan kunci digital di aplikasi ini	0
10	Kurang bagus	0

Data latih akan melalui proses *preprocessing* untuk bisa dimasukkan ke dalam proses latih model. Tabel 4.2 merupakan sample data latih yang sudah melewati proses *preprocessing*.

Tabel 4.2 Sample data latih yang setelah proses preprocessing

No	Ulasan	Credibility
1	tidak niat banget bikin aplikasi kamera 108 megapixel jadi buram garagara fitur kamera aplikasi buruk kasih aja fitur upload biar gambar jelas terus apa coba crop manual minimal seperti camscanner bisa ngepasin langsung jadi tidak lambatlambat waktu ngitung foto lambat proses hitung minimal anggar buat bikin server jangan sunat dari simulasi akhir banyak data yang salah baca dari pada yang benar uploadnya lambat siap 5 tahun ngapain aja nol besar	-1
2	apa yang bisa harap dari kpu developer dan perintah wenang yang dari awal dukung curang langgar etika diam saat tahu yang salah get well soon indonesia 1 aplikasi tidak siap tampung crowd system di harih 2 server lemah dan masalah tidak limit atas jumlah suara sesuai jml maksimal dpt 300	0
3	server untuk semua hp tidak bisa	0
4	aplikasi bagus dan bantu	1
5	kamera jernih	1
6	bisa mudah cara login nya	0
7	aplikasi lambat	-1
8	sangat bantu sekali aplikasi	1
9	tidak dapat dapat kunci digital di aplikasi ini	0
10	kurang bagus	0

Setelah data latih bersih, data akan diacak dan dibagi menjadi k kelompok. Setiap kelompok akan dipastikan memiliki komposisi kelas yang sama seperti dataset seluruhnya. Penentuan nilai k untuk latih aspek usability terdapat pada Tabel 4.3 dan untuk aspek *availability* terdapat pada Tabel 4.4 menggunakan variabel rata-rata akurasi dan rata-rata *F1-Score*.

Tabel 4.3 Rata-rata akurasi dan *f1-score* aspek usability

No	Nilai k	Rata-Rata Akurasi	Rata-Rata <i>F1-Score</i>
1	20	0.8150491484	0.8024902890
2	30	0.8150537634	0.8029712031
3	40	0.8156914893	0.8031168267
4	50	0.8161861861	0.8031174868
5	60	0.8156093189	0.8028918732

Tabel 4.4 Rata-rata akurasi dan *f1-score* aspek availability

No	Nilai k	Rata-Rata Akurasi	Rata-Rata <i>F1-Score</i>
1	20	0.8275596079	0.8260545218
2	30	0.8254010223	0.8229092878
3	40	0.8259481996	0.8220107699
4	50	0.8291441441	0.8247508692
5	60	0.8249820788	0.8198341951

Tabel 4.3 dan Tabel 4.4 menunjukkan nilai optimal pada perhitungan rata-rata akurasi dan rata-rata *f1-score* pada nilai $k = 50$. Pada aspek credibility, uji coba nilai k dapat dilihat pada Tabel 4.5.

Tabel 4.5 Rata-rata akurasi dan *f1-score* aspek credibility

No	Nilai k	Rata-Rata Akurasi	Rata-Rata <i>F1-Score</i>
1	10	0.9371092831	0.9171646628
2	20	0.9376753155	0.9187034226
3	30	0.9371496562	0.9168810607

Tabel 4.5 menunjukkan nilai optimal dari rata-rata akurasi dan rata-rata *f1-score* terletak pada nilai $k = 20$. Setelah mendapatkan nilai k yang optimal pada masing-masing aspek, latih model akan dilakukan menggunakan nilai k yang sudah ditentukan. Dataset latih akan dikelompokkan menjadi k kelompok dan pada proses latih, satu kelompok akan menjadi data validasi dan kelompok yang lain akan menjadi data latih. Proses ini dilakukan secara bergantian dan dilakukan sebanyak k kali.

4.2.2 Uji Model

Setelah proses latih model, dilakukan uji model terhadap dataset uji yang sudah dipisahkan dan belum pernah digunakan pada proses latih model. Pada Tabel 4.6 merupakan hasil proses uji model menggunakan dataset uji.

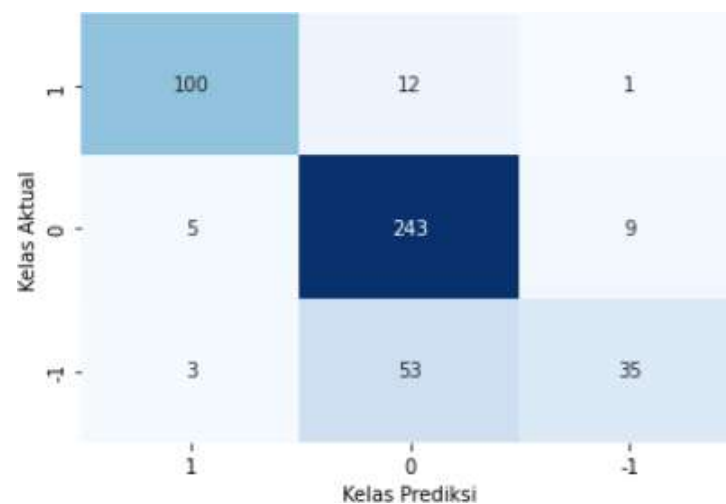
Tabel 4.6 Hasil proses uji model

No	Ulasan	Usability	Credibility	Availability	P. Usability	P. Credibility	P. Availability
1	Saya sarankan pemilu mendatang lebih baik tidak pakai Aplikasi. aplikasi ribet dan rawan kecurangan data alias penggelembungan suara	-1	-1	0	-1	-1	0
2	dibantu dong sis bro bisa submit dan upload data	0	0	-1	0	0	-1
3	tidak mau terbuka	0	0	0	0	0	0
4	Aplikasi tidak bagus, masa mau foto aja langsung logout sendiri. tidak lagi kamera bawaan aplikasinya jelek malah nyalahin kpps	-1	0	-1	-1	0	-1
5	bikin super sabar bukan karena tugas mikir 1000	0	0	-1	0	0	-1

	kali mentok di sertifikat digital coba berbagai macam cara biar tembus belum bisa semoga diperbagus dan cepat						
6	masak iya aplikasi tidak fitur opsi lupa kata sandi dah	0	0	0	0	0	-1
7	aplikasi sulit digunakan	-1	0	0	-1	0	0
8	Sebentar-sebentar harus update, Pembacaan tidak akurat, hingga bikin kisruh,server sering down dan banyak lagi. Aplikasi tidak layak pakai	-1	-1	-1	-1	-1	-1
9	berhasil sertifikat digital	0	0	1	0	0	1
10	banyak yang harus di update	0	0	0	0	0	0

4.3 Evaluasi Model

Untuk mengukur kinerja model yang sudah dibuat, diperlukan evaluasi model sehingga kita bisa mengetahui apakah model yang dihasilkan terlalu umum (*underfitting*) atau terlalu sempit (*overfitting*). Proses Evaluasi akan menggunakan metode *confusion matrix* pada setiap aspek, kemudian dihitung nilai akurasi, presisi, recall dan *f1-score*. Pada Gambar 4.5 kita dapat melihat plot perbandingan hasil prediksi dan aktual kelas dari proses uji model.



Gambar 4.5 *Plot hasil uji model aspek usability*

Melihat dari plot pada Gambar 4.5 kita dapat membuat *confusion matrix* tiap kelas pada aspek *usability*. Tabel 4.7 adalah *confusion matrix* untuk kelas sentimen negatif, Tabel 4.8 adalah *confusion matrix* untuk kelas sentimen netral dan Tabel 4.9 adalah *confusion matrix* untuk kelas sentimen positif.

Tabel 4.7 *Confusion Matrix sentimen negatif usability*

No	Confusion Matrix	Total
1	True Positive (TP)	35
2	True Negative (TN)	360
3	False Positive (FP)	10
4	False Negative (FN)	56

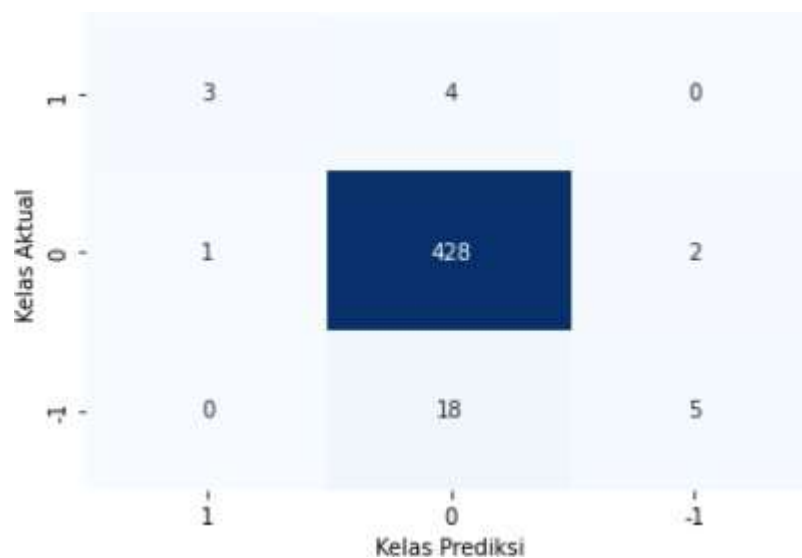
Tabel 4.8 *Confusion Matrix sentimen netral usability*

No	Confusion Matrix	Total
1	True Positive (TP)	243
2	True Negative (TN)	139
3	False Positive (FP)	55
4	False Negative (FN)	14

Tabel 4.9 *Confusion Matrix* sentimen positif *usability*

No	Confusion Matrix	Total
1	True Positive (TP)	100
2	True Negative (TN)	340
3	False Positive (FP)	8
4	False Negative (FN)	13

Kemudian Gambar 4.6 adalah hasil plot dari uji model pada aspek *credibility*.

**Gambar 4.6 Plot hasil uji model aspek *credibility***

Melihat dari plot pada Gambar 4.6 kita dapat membuat *confusion matrix* tiap kelas pada aspek *credibility*. Tabel 4.10 adalah *confusion matrix* untuk kelas sentimen negatif, Tabel 4.11 adalah *confusion matrix* untuk kelas sentimen netral dan Tabel 4.12 adalah *confusion matrix* untuk kelas sentimen positif.

Tabel 4.10 *Confusion Matrix* sentimen negatif *credibility*

No	Confusion Matrix	Total
1	True Positive (TP)	5
2	True Negative (TN)	436
3	False Positive (FP)	2
4	False Negative (FN)	18

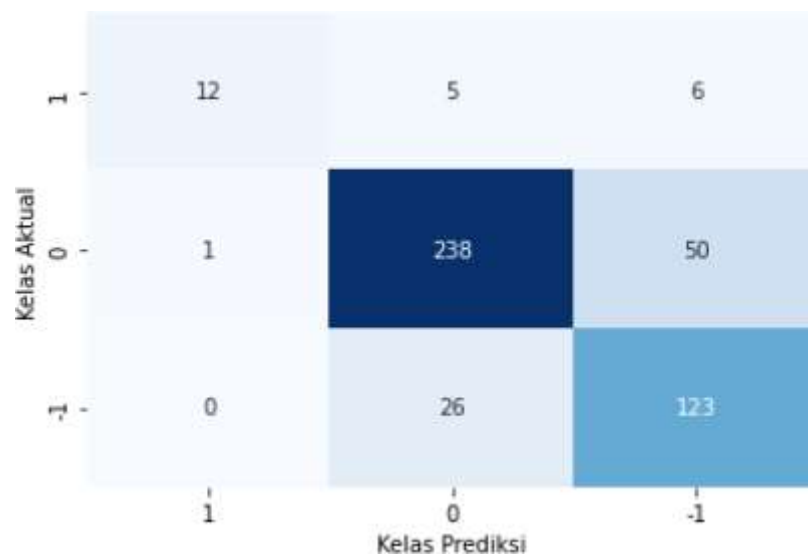
Tabel 4.11 *Confusion Matrix* sentimen netral *credibility*

No	Confusion Matrix	Total
1	True Positive (TP)	428
2	True Negative (TN)	8
3	False Positive (FP)	12
4	False Negative (FN)	3

Tabel 4.12 *Confusion Matrix* sentimen positif *credibility*

No	Confusion Matrix	Total
1	True Positive (TP)	3
2	True Negative (TN)	453
3	False Positive (FP)	1
4	False Negative (FN)	4

Kemudian Gambar 4.7 adalah hasil plot dari uji model pada aspek *availability*.

**Gambar 4.7 *Plot* hasil uji model aspek *availability***

Melihat dari plot pada Gambar 4.7 kita dapat membuat *confusion matrix* tiap kelas pada aspek *availability*. Tabel 4.13 adalah *confusion matrix* untuk kelas sentimen negatif, Tabel 4.14 adalah *confusion matrix* untuk kelas sentimen netral dan Tabel 4.15 adalah *confusion matrix* untuk kelas sentimen positif.

Tabel 4.13 Confusion Matrix sentimen negatif availability

No	Confusion Matrix	Total
1	True Positive (TP)	123
2	True Negative (TN)	256
3	False Positive (FP)	56
4	False Negative (FN)	26

Tabel 4.14 Confusion Matrix sentimen netral availability

No	Confusion Matrix	Total
1	True Positive (TP)	238
2	True Negative (TN)	282
3	False Positive (FP)	31
4	False Negative (FN)	51

Tabel 4.15 Confusion Matrix sentimen positif availability

No	Confusion Matrix	Total
1	True Positive (TP)	12
2	True Negative (TN)	437
3	False Positive (FP)	1
4	False Negative (FN)	11

Setelah mendapatkan nilai dari confusion matrix tiap kelas pada setiap aspek, nilai dari akurasi, presisi, *recall* dan *f1-score* dapat diperoleh menggunakan perhitungan sebagai berikut.

4.3.1 Perhitungan pada aspek *usability*

1. Presisi

$$\text{Perhitungan sentimen negatif} = \frac{TP}{TP+FP} = \frac{35}{45} = 0,77$$

$$\text{Perhitungan sentimen netral} = \frac{TP}{TP+FP} = \frac{243}{298} = 0,81$$

$$\text{Perhitungan sentimen positif} = \frac{TP}{TP+FP} = \frac{100}{108} = 0,92$$

2. Recall

$$\text{Pehitungan sentimen negatif} = \frac{TP}{TP+FN} = \frac{35}{91} = 0,38$$

$$\text{Pehitungan sentimen netral} = \frac{TP}{TP+FN} = \frac{243}{257} = 0,94$$

$$\text{Pehitungan sentimen positif} = \frac{TP}{TP+FN} = \frac{100}{113} = 0,88$$

3. F1-Score

$$\text{Pehitungan sentimen negatif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.29}{1.15} = 0,50$$

$$\text{Pehitungan sentimen netral} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.76}{1.75} = 0,86$$

$$\text{Pehitungan sentimen positif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.80}{1.8} = 0,88$$

4.3.2 Perhitungan pada aspek *credibility*

1. Presisi

$$\text{Pehitungan sentimen negatif} = \frac{TP}{TP+FP} = \frac{5}{7} = 0,71$$

$$\text{Pehitungan sentimen netral} = \frac{TP}{TP+FP} = \frac{428}{440} = 0,97$$

$$\text{Pehitungan sentimen positif} = \frac{TP}{TP+FP} = \frac{3}{4} = 0,75$$

2. Recall

$$\text{Pehitungan sentimen negatif} = \frac{TP}{TP+FN} = \frac{5}{23} = 0,21$$

$$\text{Pehitungan sentimen netral} = \frac{TP}{TP+FN} = \frac{428}{431} = 0,99$$

$$\text{Pehitungan sentimen positif} = \frac{TP}{TP+FN} = \frac{3}{7} = 0,42$$

3. F1-Score

$$\text{Pehitungan sentimen negatif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.14}{0.92} = 0,30$$

$$\text{Pehitungan sentimen netral} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.96}{1.96} = 0,97$$

$$\text{Pehitungan sentimen positif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0.31}{1.17} = 0,53$$

4.3.3 Perhitungan pada aspek *availability*

1. Presisi

$$\text{Pehitungan sentimen negatif} = \frac{TP}{TP+FP} = \frac{123}{179} = 0,68$$

$$\text{Pehitungan sentimen netral} = \frac{TP}{TP+FP} = \frac{238}{269} = 0,88$$

$$\text{Pehitungan sentimen positif} = \frac{TP}{TP+FP} = \frac{12}{13} = 0,92$$

2. Recall

$$\text{Pehitungan sentimen negatif} = \frac{TP}{TP+FN} = \frac{123}{149} = 0,82$$

$$\text{Pehitungan sentimen netral} = \frac{TP}{TP+FN} = \frac{238}{289} = 0,82$$

$$\text{Pehitungan sentimen positif} = \frac{TP}{TP+FN} = \frac{12}{23} = 0,52$$

3. F1-Score

$$\text{Pehitungan sentimen negatif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0,55}{1,5} = 0,73$$

$$\text{Pehitungan sentimen netral} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0,72}{1,7} = 0,84$$

$$\text{Pehitungan sentimen positif} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} = 2 \times \frac{0,47}{1,44} = 0,65$$

4.3.4 Perhitungan rata-rata akurasi seluruh aspek

$$\text{Akurasi aspek } usability = \frac{TP+TN}{TP+TN+FP+FN} = \frac{378}{461} = 0.81$$

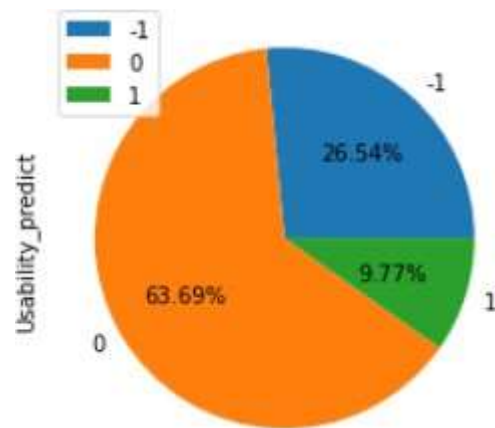
$$\text{Akurasi aspek } credibility = \frac{TP+TN}{TP+TN+FP+FN} = \frac{436}{461} = 0.94$$

$$\text{Akurasi aspek } availability = \frac{TP+TN}{TP+TN+FP+FN} = \frac{373}{461} = 0.80$$

$$\text{Rata-rata akurasi seluruh aspek} = \frac{0.81+0.94+0.80}{3} \times 100 = 0.85 \times 100 = 85\%$$

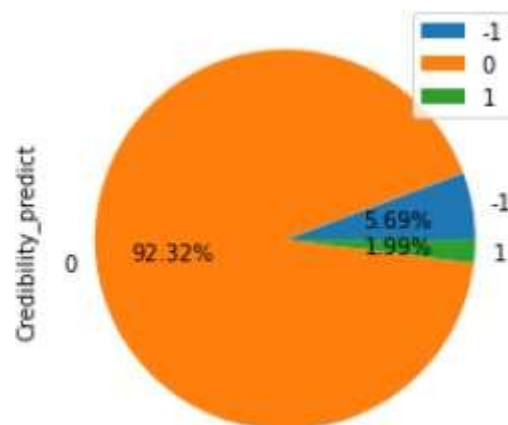
4.4 Visualisasi Hasil Analisis Sentimen Aplikasi Sirekap Mobile

Setelah dilakukan analisis sentimen dari keseluruhan ulasan Aplikasi Sirekap Mobile yang diambil dari *Google Play Store*, Data sentimen dari tiap aspek dapat dilakukan visualisasi sehingga memberikan informasi tentang bagaimana sentimen pengguna aplikasi terhadap aspek-aspek yang sudah ditentukan. Gambar 4.8 merupakan visualisasi data hasil analisis Sirekap mobile terkait aspek usability. Dapat dilihat bahwa sentimen positif sebanyak 9,77%, netral sebanyak 63,69% dan sentimen negatif sebanyak 26,54%. Berarti sebanyak 63,69% ulasan tidak membahas aspek usability. Sedangkan pengguna yang membahas aspek usability sebanyak 26,54% memiliki sentimen negatif terhadap aspek usability dari aplikasi Sirekap Mobile.



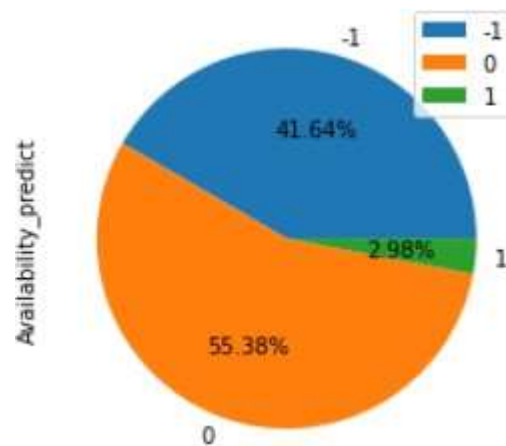
Gambar 4.8 Chart hasil analisis pada aspek *usability*

Kemudian aspek *credibility* dapat dilihat pada Gambar 4.9 dimana pada aspek tersebut, sentimen negatif memiliki persentase sebanyak 5,69%, sentimen positif sebanyak 1,99% dan sentimen netral sebanyak 92,32%. Artinya aspek *credibility* tidak begitu banyak mendapatkan *feedback* dari pengguna Sirekap Mobile.



Gambar 4.9 Chart hasil analisis pada aspek *credibility*

Aspek yang terakhir adalah aspek *availability* dimana pada Gambar 4.10 dapat diambil kesimpulan bahwa sebanyak 55,38% pengguna memiliki sentimen netral yang artinya tidak membahas aspek *availability*, sebanyak 41,64% pengguna memiliki sentimen negatif terkait aspek *availability* dan 2,98% pengguna memiliki sentimen positif terhadap aspek *availability* aplikasi Sirekap Mobile.



Gambar 4.10 Chart hasil analisis pada aspek *availability*

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil yang sudah didapatkan dari penelitian ini dapat disimpulkan bahwa metode *Complement Naïve Bayes* dapat mencapai akurasi sebesar 85% dalam mengklasifikasi sentimen menjadi tiga kelas, yaitu sentimen negatif, sentimen netral dan sentimen positif. Hasil sentimen dari keseluruhan ulasan Sirekap Mobile pada ulasan Google Play Store menunjukkan bahwa aspek *usability* mendapatkan sentimen positif sebanyak 9,77%, netral sebanyak 63,69% dan negatif sebanyak 26,54%. Aspek *credibility* mendapat sentiment positif sebanyak 1,99%, netral sebanyak 92,32% dan negatif sebanyak 5,69%. Aspek *availability* mendapat sentiment positif sebanyak 2,98%, netral sebanyak 55,38% dan negatif sebanyak 41.64%.

5.2 Saran

Adapun saran yang perlu menjadi perhatian pada penelitian selanjutnya adalah sebagai berikut:

1. Menggunakan dataset yang lebih besar dalam kurun waktu yang lebih lama agar dapat diimplementasikan pada metode *deep learning*.
2. Komposisi tiap kelas lebih seimbang dan variasi data lebih beragam.
3. Menggunakan metode normalisasi dan typo correction pada proses *data cleaning*

DAFTAR PUSTAKA

- Amirul, H., Victor, A., & Arini. (2019). Analisis Sentimen Kinerja KPU Pemilu 2019 Menggunakan Algoritma K-Means dengan Algoritma Confix Stripping Stemmer. *Journal of Innovation Information Technology and Application*.
- Amirullah, F., Alam, S., & Sulisty, M. I. (2023). Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naive Bayes. *Jurnal Ilmiah Teknik dan Ilmu Komputer*.
- Astuti, S. P., Liebenlito, M., & Fauziah, I. (2020). Analisis sentimen berbasis aspek pada aplikasi tokopedia menggunakan LDA dan naïve bayes. *Analisis sentimen berbasis aspek pada aplikasi tokopedia menggunakan LDA dan naïve bayes*. Institutional Repository UIN Syarif Hidayatullah Jakarta.
- Farhadloo, M., & Rolland, E. (2016). *Fundamentals of Sentiment Analysis and Its Application*. Switzerland: Springer International Publishing.
- GOLDWIN, K., & Priyanta, S. (2020). PENGGUNAAN COMPLEMENT NAIVE BAYES UNTUK MEMECAHKAN MASALAH IMBALANCE DATASET PADA TWITTER DENGAN STUDI KASUS JASA TRANSPORTASI ONLINE. *PENGGUNAAN COMPLEMENT NAIVE BAYES UNTUK MEMECAHKAN MASALAH IMBALANCE DATASET PADA TWITTER DENGAN STUDI KASUS JASA TRANSPORTASI ONLINE*. Indonesia: Universitas Gajah Mada.
- Hidayat, D. (2024, Februari 20). *Ketidaksesuaian Data Sirekap Akibat Kesalahan Sistem Baca Angka*. Retrieved from rri.co.id: <https://www.rri.co.id/pemilu/563200/ketidaksesuaian-data-sirekap-akibat-kesalahan-sistem-baca-angka>
- Ilahi, N. P. (2024, 02 06). *Keluhkan Aplikasi Sirekap Lemot dan Sulit Login, Begini Hasil Simulasi Ulang Penghitungan Suara di KPU Tuban*. Retrieved from <https://www.jawapos.com/>: <https://www.jawapos.com/berita-sekitar-anda/014124424/keluhkan-aplikasi-sirekap-lemot-dan-sulit-login-begini-hasil-simulasi-ulang-penghitungan-suara-di-kpu-tuban>
- Imam, S., Windu, G., & Budi, P. A. (2019). Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum. *Jurnal Rekayasa Sistem dan Teknologi Informasi*.

- Irfani, F. F., Triyanto, M., Hartanto, A. D., & Kusnawi. (2019). Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine. *Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine*. Jurnal Bisnis, Manajemen & Informatika.
- Julianto, Y., Setiabudi, D. H., & Silvia, R. (2019). Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine. *Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine*. Jurnal Teknologi Informasi dan Ilmu Komputer.
- Kao, A., & Poteet, S. R. (2006). *Natural Language Processing and Text Mining*. Bellevue: Springer.
- KPU. (2021, November 17). *Manfaatkan Sirekap, Transparan dan Kemudahan untuk Masyarakat*. Retrieved from Komisi Pemilihan Umum: <https://www.kpu.go.id/berita/baca/10143/manfaatkan-sirekap-transparan-dan-kemudahan-untuk-masyarakat>
- Nababan, A. P., Lumenta, A. S., Rindengan, Y. D., Pontoh, F. J., & Akay, Y. V. (2020). Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode Lexicon Analisis. *Jurnal Teknik Informatika*.
- Pitria, P. (2016). Analisis Sentimen Pengguna Twitter Pada Akun Resmi Samsung Indonesia Dengan Menggunakan Naive Bayes. *Analisis Sentimen Pengguna Twitter Pada Akun Resmi Samsung Indonesia Dengan Menggunakan Naive Bayes*. Unikom Repository.
- Ranjan, S., & Mishra, S. (2020). Comparative Sentiment Analysis of App Reviews. *Comparative Sentiment Analysis of App Reviews*. IEEE Xplore.
- Silvianto, H. (2024, 02 19). *Sirekap KPU Bermasalah, Kepercayaan Publik Dipertaruhkan*. Retrieved from Kompas.com: <https://nasional.kompas.com/read/2024/02/19/10052121/sirekap-kpu-bermasalah-kepercayaan-publik-dipertaruhkan?page=all>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET, DAN TEKNOLOGI
UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

Jalan Universitas No. 9A Gedung A, Kampus USU Medan 20155, Telepon: (061) 821007
Laman: <http://Fasilkomti.usu.ac.id>

KEPUTUSAN
DEKAN FAKULTAS ILMU KOMPUTER
DAN TEKNOLOGI INFORMASI
NOMOR : 2739/UN5.2.14.D/SK/SPB/2024
DEKAN FAKULTAS ILMU KOMPUTER
DAN TEKNOLOGI INFORMASI UNIVERSITAS SUMATERA UTARA

- Membaca : Surat Permohonan Mahasiswa Fasilkom-TI USU tanggal 11 Juli 2024 perihal permohonan ujian skripsi:
Nama : ANANDA MUHARRIZ SINAGA
NIM : 171402052
Program Studi : Sarjana (S-1) Teknologi Informasi
Judul Skripsi : Analisis Sentimen Terhadap Aplikasi Sirekap Mobile Berbasis Aspek Pada Pemilihan Umum 2024 Dengan Metode Naive Bayes
- Memperhatikan : Bahwa Mahasiswa tersebut telah memenuhi kewajiban untuk ikut dalam pelaksanaan Meja Hijau Skripsi Mahasiswa pada Program Studi Sarjana (S-1) Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara TA 2023/2024.
- Menimbang : Bahwa permohonan tersebut diatas dapat disetujui dan perlu ditetapkan dengan surat keputusan
- Mengingat : 1. Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional.
2. Peraturan Pemerintah Nomor 17 tahun 2010 tentang pengelolaan dan penyelenggara pendidikan.
3. Keputusan Rektor USU Nomor 03/UN5.1.R/SK/SPB/2021 tentang Peraturan Akademik Program Sarjana Universitas Sumatera Utara.
4. Surat Keputusan Rektor USU Nomor 1876/UN5.1.R/SK/SDM/2021 tentang pengangkatan Dekan Fasilkom-TI USU Periode 2021-2026

MEMUTUSKAN

- Menetapkan :
Pertama : Membentuk dan mengangkat Tim Penguji Skripsi mahasiswa sebagai berikut:
Ketua : Dr. Romi Fadillah Rahmat, B.Comp.Sc., M.Sc.
NIP: 198603032010121004
Sekretaris : Ade Sarah Huzaifah S.Kom., M.Kom
NIP: 198506302018032001
Anggota Penguji : Dedy Arisandi ST., M.Kom.
NIP: 197908312009121002
Anggota Penguji : Sarah Purnamawati ST., MSc.
NIP: 198302262010122003
Moderator : -
Panitera : -
- Kedua : Segala biaya yang diperlukan untuk pelaksanaan kegiatan ini dibebankan pada Dana Penerimaan Bukan Pajak (PNPB) Fasilkom-TI USU Tahun 2024.
- Ketiga : Keputusan ini berlaku sejak tanggal ditetapkan dengan ketentuan bahwa segala sesuatunya akan diperbaiki sebagaimana mestinya apabila dikemudian hari terdapat kekeliruan dalam surat keputusan ini.

- Tembusan :
- 1. Ketua Program Studi Sarjana (S-1) Teknologi Informasi
 - 2. Yang bersangkutan
 - 3. Arsip

Medan
Ditandatangani secara elektronik oleh:
Dekan



Maya Silvi Lydia
NIP 197401272002122001