

**PENGEMBANGAN MODEL PENGUKURAN *ECHO*  
*CHAMBER* DENGAN ANALISIS SENTIMEN BERBASIS  
ASPEK BEBAS DOMAIN DAN KONSENSUS *GROUP*  
*DECISION MAKING***

**Laporan Tugas Akhir**

Disusun sebagai syarat kelulusan tingkat sarjana

Oleh

**PUTI NABILLA AIDIRA**

**NIM : 13521088**



**PROGRAM STUDI TEKNIK INFORMATIKA  
SEKOLAH TEKNIK ELEKTRO & INFORMATIKA  
INSTITUT TEKNOLOGI BANDUNG**

**September 2025**

**PENGEMBANGAN MODEL PENGUKURAN *ECHO*  
*CHAMBER* DENGAN ANALISIS SENTIMEN BERBASIS  
ASPEK BEBAS DOMAIN DAN KONSENSUS *GROUP*  
*DECISION MAKING***

**Laporan Tugas Akhir**

Oleh  
**PUTI NABILLA AIDIRA**  
**NIM : 13521088**

**Program Studi Teknik Informatika**  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung

Telah disetujui dan disahkan sebagai Laporan Tugas Akhir  
di Bandung, pada tanggal 18 September 2025

Pembimbing,



Dr. Masayu Leylia Khodra, S.T., M.T.  
NIP. 19760429 200812 2 000

## **LEMBAR PERNYATAAN**

Dengan ini saya menyatakan bahwa:

1. Pengerajan dan penulisan Laporan Tugas Akhir ini dilakukan tanpa menggunakan bantuan yang tidak dibenarkan.
2. Segala bentuk kutipan dan acuan terhadap tulisan orang lain yang digunakan di dalam penyusunan laporan tugas akhir ini telah dituliskan dengan baik dan benar.
3. Laporan Tugas Akhir ini belum pernah diajukan pada program pendidikan di perguruan tinggi mana pun.

Jika terbukti melanggar hal-hal di atas, saya bersedia dikenakan sanksi sesuai dengan Peraturan Akademik dan Kemahasiswaan Institut Teknologi Bandung bagian Penegakan Norma Akademik dan Kemahasiswaan khususnya Pasal 2.1 dan Pasal 2.2.

Bandung, 18 September 2025



Puti Nabilla Aidira

NIM 13521088

## ABSTRAK

# PENGEMBANGAN MODEL PENGUKURAN *ECHO CHAMBER* DENGAN ANALISIS SENTIMEN BERBASIS ASPEK BEBAS DOMAIN DAN KONSENSUS *GROUP DECISION MAKING*

Oleh

PUTI NABILLA AIDIRA

NIM: 13521088

Media sosial menjadi ruang penting bagi masyarakat Indonesia untuk berinteraksi dan membentuk opini, namun pola komunikasi di dalamnya kerap melahirkan *echo chamber* yang memperkuat polarisasi. Meskipun begitu, studi *echo chamber* di Indonesia masih terbatas pada analisis manual atau analisis jaringan tanpa mempertimbangkan sentimen konten.

Penelitian ini mengembangkan model pengukuran *echo chamber* pada percakapan media sosial berbahasa Indonesia dengan Analisis Sentimen Berbasis Aspek (ASBA) bebas domain dan metrik konsensus dalam kerangka *Group Decision Making* (GDM). Rancangan mencakup lima tahap: praproses teks, ASBA bebas domain, pembentukan graf interaksi, deteksi komunitas, dan perhitungan konsensus GDM. Pada ASBA, berbagai pendekatan dievaluasi untuk menemukan model bebas domain terbaik. Hasil menunjukkan bahwa metode *zero-shot LLM* generatif dengan model Gemini-2.5-flash memberikan performa terbaik, dengan F1 makro 0,4178 pada dataset Pilkada DKI dan 0,4990 pada dataset *Indo Vaccination*, serta *Penalized Brier Score* lebih baik dibanding *baseline*.

Pengujian model pengukuran *echo chamber* dilakukan pada tiga dataset percakapan X berbahasa Indonesia: *Indo Vaccination*, *Budget Efficiency*, dan Pilpres 2024. Hasil menunjukkan nilai *Echo Chamber Ratio* (ECR) yang konsisten dengan analisis pakar terdahulu mengenai *echo chamber* di Indonesia. Nilai ECR tersebut antara lain, 1,13 pada *Indo Vaccination*, 4,55 pada *Budget Efficiency*, dan 2,92 pada Pilpres 2024. Uji hipotesis dengan *null* atribut (pengacakan sentimen) dan *null* topologi (*rewiring* jaringan) memperkuat temuan ini. Nilai z-score antara 3,37–303,99 dan *empirical p-value*  $\leq 0,01$  menunjukkan hasil signifikan secara statistik dan jauh dari distribusi acak. Temuan ini menunjukkan bahwa pendekatan berbasis ASBA dan konsensus GDM dapat memberikan ukuran ringkas melalui nilai ECR yang bermakna dan konsisten dengan analisis pakar. Pendekatan ini juga memungkinkan analisis kualitatif terhadap konsensus hingga tingkat komunitas.

**Kata kunci:** *echo chamber*, media sosial, Analisis Sentimen Berbasis Aspek (ASBA), konsensus *Group Decision Making*

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah Swt., atas rahmat, karunia, dan petunjuk-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini yang berjudul “Pengembangan Model Pengukuran *Echo Chamber* dengan Analisis Sentimen Berbasis Aspek Bebas Domain dan Konsensus *Group Decision Making*” sebagai salah satu syarat untuk meraih gelar sarjana di Institut Teknologi Bandung.

Penulis menyadari bahwa penyelesaian Tugas Akhir ini tidak terlepas dari dukungan dan bantuan berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Masayu Leylia Khodra, S.T., M.T., selaku dosen pembimbing Tugas Akhir, yang telah memberikan bimbingan, arahan, serta motivasi sejak awal hingga selesaiya Tugas Akhir ini.
2. Ibu Dr. Eng. Ayu Purwarianti, S.T., M.T. dan Ibu Dr. Fariska Zakhralativa Ruskanda, S.T., M.T., selaku dosen penguji yang telah memberikan masukan, kritik, serta saran berharga untuk penyempurnaan Tugas Akhir ini.
3. Ibu Dessi Puji Lestari, S.T., M.Eng., Dr. Nur Ulfa Maulidevi, S.T., M.Sc., Ph.D., dan Dr. Eng. Ayu Purwarianti, S.T., M.T., selaku dosen wali, atas bimbingan dan dukungan akademik yang diberikan selama penulis menempuh studi.
4. Ibu Robithoh Annur, S.T., M.Eng., Ph.D., dan Ibu Tricya Esterina Widagdo, S.T., M.Sc., selaku dosen koordinator Tugas Akhir, yang telah menyediakan materi, arahan, dan masukan penting terkait penyusunan Tugas Akhir ini.
5. Seluruh dosen Program Studi Sarjana Teknik Informatika dan Tahap Persiapan Bersama (TPB), atas ilmu dan pengalaman berharga yang telah diberikan selama perkuliahan di Institut Teknologi Bandung.
6. Staf Tata Usaha Program Studi Sarjana Teknik Informatika, atas bantuan dalam pemenuhan kebutuhan administrasi perkuliahan dan Tugas Akhir.

7. Beasiswa Kaltim Tuntas dari Pemerintah Provinsi Kalimantan Timur, atas dukungan finansial yang memungkinkan penulis menempuh studi dan menyelesaikan Tugas Akhir ini dengan baik.
8. Orang tua, adik, dan seluruh keluarga, atas dukungan moral, doa, dan bantuan finansial.
9. Teman-teman Teknik Informatika 2021, atas dukungan moral, semangat, serta kerja sama selama menempuh perkuliahan.
10. Seluruh pihak yang telah memberikan kontribusi, baik langsung maupun tidak langsung, yang tidak dapat disebutkan satu per satu.

Penulis berharap Tugas Akhir ini dapat memberikan kontribusi yang berarti bagi pengembangan penelitian terkait pengukuran *echo chamber* dan analisis sentimen berbasis aspek, khususnya dalam konteks percakapan media sosial berbahasa Indonesia.

Bandung, 18 September 2025

Penulis

## DAFTAR ISI

<b>LEMBAR PERNYATAAN .....</b>	<b>iii</b>
<b>ABSTRAK .....</b>	<b>iv</b>
<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>DAFTAR ISI.....</b>	<b>vii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>x</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>DAFTAR SINGKATAN.....</b>	<b>xiv</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
I.1 Latar Belakang .....	1
I.2 Rumusan Masalah.....	3
I.3 Tujuan dan Ukuran Keberhasilan Pencapaian .....	4
I.4 Batasan Masalah .....	4
I.5 Metodologi .....	4
I.6 Sistematika Pembahasan .....	5
<b>BAB II KAJIAN PUSTAKA .....</b>	<b>7</b>
II.1 <i>Echo Chamber</i> .....	7
II.2 Kategori Pendekatan Studi <i>Echo Chamber</i> .....	8
II.3 Analisis Sentimen Berbasis Aspek (ASBA) .....	9
II.3.1 Metode ASBA Bebas Domain .....	11
II.3.1.1 <i>Double Propagation</i> .....	12
II.3.1.2 <i>Zero-Shot Large Language Model (LLM)</i> generatif .....	17
II.3.2 Metrik Evaluasi ASBA .....	20

II.4	Representasi Vektor Teks .....	23
II.5	Metrik Konsensus dalam Kerangka <i>Group Decision Making</i> .....	25
II.6	Metode Deteksi Komunitas.....	29
II.7	Interaksi Pada X.....	33
II.8	<i>Null Model</i> dan Uji Hipotesis pada Analisis Jaringan Sosial .....	35
II.9	Penelitian Terkait .....	38
II.9.1	<i>Towards Echo Chamber Assessment by Employing Aspect-Based Sentiment Analysis and GDM Consensus Metrics</i> (Amendola dkk., 2024)....	38
II.9.2	<i>Aspect-based sentiment analysis for Indonesian restaurant reviews</i> (Ekawati dan Khodra, 2017) .....	42
<b>BAB III ANALISIS persoalan DAN RANCANGAN SOLUSI .....</b>		<b>45</b>
III.1	Analisis Persoalan.....	45
III.2	Analisis Solusi .....	48
III.2.1	Ketersediaan <i>Dataset</i> .....	48
III.2.2	Kualitas Teks pada <i>Dataset</i> .....	53
III.2.3	Metode ASBA Bebas Domain .....	55
III.2.4	Pembangunan Graf Interaksi dan Metode Deteksi Komunitas.....	58
III.2.5	Perhitungan Metrik Konsensus GDM.....	59
III.3	Rancangan Solusi.....	60
III.3.1	Praproses Teks .....	62
III.3.2	ASBA Bebas Domain .....	62
III.3.3	Pembangunan Graf Interaksi dan Deteksi Komunitas .....	63
III.3.4	Perhitungan Metrik Konsensus GDM.....	64
III.4	Implementasi Solusi.....	67
III.4.1	Praproses Teks .....	67
III.4.2	Metode ASBA Bebas Domain: <i>Double Propagation</i> .....	67

III.4.3	Metode ASBA Bebas Domain: <i>Zero-Shot LLM</i> Generatif .....	70
III.4.4	Pembangunan Graf Interaksi dan Deteksi Komunitas .....	71
III.4.5	Perhitungan Metrik Konsensus GDM.....	72
<b>BAB IV</b>	<b>EKSPERIMEN DAN PENGUJIAN.....</b>	<b>74</b>
IV.1	Eksperimen Praproses .....	74
IV.2	Eksperimen Konfigurasi ASBA.....	76
IV.3	Pengujian ASBA.....	79
IV.4	Pengujian Pengukuran <i>Echo Chamber</i> .....	81
IV.5	Uji Hipotesis Pengukuran <i>Echo Chamber</i> .....	85
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN.....</b>	<b>88</b>
V.1	Kesimpulan .....	88
V.2	Saran .....	89
<b>DAFTAR REFERENSI.....</b>		<b>91</b>

## **DAFTAR LAMPIRAN**

<b>Lampiran A. Deskripsi <i>Dataset</i> Mentah .....</b>	<b>95</b>
A.1    Deskripsi <i>Dataset</i> Mentah untuk ASBA .....	95
A.2    Deskripsi <i>Dataset</i> Mentah untuk Pengukuran <i>Echo Chamber</i> .....	96
<b>Lampiran B. Deskripsi <i>Dataset</i> Final .....</b>	<b>98</b>
B.1    Deskripsi <i>Dataset</i> Final untuk Eksperimen Konfigurasi ASBA.....	98
B.2    Deskripsi <i>Dataset</i> Final untuk Pengujian ASBA.....	100
B.3    Deskripsi <i>Dataset</i> Final untuk Pengujian Pengukuran <i>Echo Chamber</i>	
102	
<b>Lampiran C. Isu Kualitas Teks pada <i>Dataset</i> .....</b>	<b>104</b>
<b>Lampiran D. Keluaran Detail Konsensus dan Graf Interaksi dengan Komunitas.....</b>	<b>106</b>
D.1    Keluaran Detail Konsensus Dalam Komunitas.....	106
D.2    Keluaran Detail Konsensus Antarkomunitas .....	107
D.3    Keluaran Graf Interaksi dengan Komunitas .....	109
<b>Lampiran E. Graf Histogram Distribusi Null pada Uji Hipotesis .....</b>	<b>110</b>

## DAFTAR GAMBAR

Gambar II.1 Visualisasi Contoh Elemen Sentimen .....	10
Gambar II.2 Metode <i>Double Propagation</i> .....	14
Gambar II.3 Aturan Propagasi (Qiu dkk., 2011) .....	14
Gambar II.4 Algoritma Propagasi (Qiu dkk., 2011) .....	16
Gambar II.5 Hasil F-score Eksperimen yang Membandingkan <i>Double Propagation</i> (Prop-dep) dengan Pendekatan Lain (Qiu dkk., 2011) .....	17
Gambar II.6 Contoh <i>Prompt</i> yang Digunakan (Mughal dkk., 2024).....	18
Gambar II.7 Hasil Penelitian Komparasi <i>Zero-Shot LLM Generatif</i> dengan Model Lain pada Tugas ACSA (Mughal dkk., 2024) .....	19
Gambar II.8 Visualisasi Proses <i>Multilevel Graph Partitioning</i> (Karypis dan Kumar, 1995) .....	30
Gambar II.9 Ilustrasi Penerapan Metode Louvain Pada Graf Tak Berbobot Sederhana.....	33
Gambar II.10 Pipeline Sistem (Amendola dkk., 2024).....	39
Gambar II.11 Perbandingan Metrik Konsensus dengan Metrik SOTA Berbasis Topologi (Amendola dkk., 2024).....	41
Gambar II.12 Perbandingan Hasil Konsensus Berbasis ASBA dengan Hasil Konsensus Berbasis VADER pada <i>Dataset ROT</i> (Amendola dkk., 2024) .....	42
Gambar III.1 Partisi yang Dihasilkan METIS <i>Multilevel K-way Graph Partitioning</i> dengan k = 2 pada <i>Dataset Echo Chamber Pilpres 2024</i> (kiri) dan <i>Budget Efficiency</i> (kanan). .....	47
Gambar III.2 Diagram Alur Pembangunan <i>Dataset</i> Eksperimen dan Pengujian ASBA.....	50
Gambar III.3 Hasil Analisis Sentimen Warganet Media Sosial Terkait Isu Efisiensi Anggaran (Fahmi, 2025).....	53

Gambar III.4 Diagram Alur Pembangunan <i>Dataset</i> Pengujian Pengukuran <i>Echo Chamber</i> .....	53
Gambar III.5 Perbandingan Hasil Deteksi Komunitas dengan METIS ( $k = 2$ ) dan Louvain pada <i>Dataset Pilpres 2024</i> .....	59
Gambar III.6 Rancangan Solusi .....	61
Gambar III.7 Contoh Keluaran Graf Interaksi untuk <i>Dataset Pilpres 2024</i> .....	64
Gambar III.8 <i>Pseudocode</i> Praproses Teks .....	67
Gambar III.9 <i>Pseudocode</i> Iterasi <i>Double Propagation</i> ..	68
Gambar III.10 <i>Pseudocode</i> Perhitungan Bobot Sentimen <i>Double Propagation</i> ..	69
Gambar III.11 <i>Pseudocode</i> Parsing Respons <i>Zero-Shot LLM Generatif</i> .....	70
Gambar III.12 <i>Pseudocode</i> Pembangunan Graf Interaksi .....	71
Gambar III.13 <i>Pseudocode</i> Deteksi Komunitas dan <i>Merging User</i> .....	72
Gambar III.14 <i>Pseudocode</i> Perhitungan Metrik Konsensus GDM .....	72
Gambar IV.1 <i>Confusion Matrix</i> Hasil Pengujian ASBA untuk <i>Dataset Pilkada DKI</i> dengan Aspek Paslon .....	80
Gambar IV.2 <i>Confusion Matrix</i> Hasil Pengujian ASBA untuk <i>Dataset Indo Vaccination</i> .....	80

## DAFTAR TABEL

Tabel II.1 Hasil Konsensus <i>Dataset “Twitter Vaccination”</i> dan ROT (Amendola dkk., 2024) .....	39
Tabel III.1 Kasus dan Aturan Resolusi Konflik Manual untuk Anotasi Klasifikasi Aspek .....	49
Tabel III.2 Deskripsi Ringkas <i>Dataset Final</i> untuk Eksperimen dan Pengujian ASBA .....	50
Tabel III.3 Deskripsi Ringkas <i>Dataset Final</i> untuk Pengujian Pengukuran <i>Echo Chamber</i> .....	52
Tabel III.4 Contoh Hasil Praproses untuk Setiap Tingkat .....	54
Tabel III.5 Desain <i>system prompt</i> dan <i>user prompt</i> .....	56
Tabel III.6 Contoh Baris Keluaran ASBA untuk <i>Dataset Pilpres 2024</i> .....	62
Tabel III.7 Contoh Baris Keluaran Deteksi Komunitas untuk <i>Dataset Pilpres 2024</i> .....	63
Tabel III.8 Jumlah Komunitas yang Dihasilkan .....	64
Tabel IV.1 Hasil Eksperimen Praproses .....	75
Tabel IV.2 Hasil Eksperimen Konfigurasi ASBA .....	78
Tabel IV.3 Hasil Pengujian ASBA .....	79
Tabel IV.4 Hasil Pengujian Pengukuran <i>Echo Chamber</i> .....	82
Tabel IV.5 Hasil <i>Baseline</i> Pengukuran <i>Echo Chamber</i> Amendola dkk. (2024) ..	84
Tabel IV.6 Hasil Uji Hipotesis Pengukuran <i>Echo Chamber</i> dengan <i>Null</i> Atribut	86
Tabel IV.7 Hasil Uji Hipotesis Pengukuran <i>Echo Chamber</i> dengan <i>Null</i> Topologi .....	86

## **DAFTAR SINGKATAN**

ACSA	: <i>Aspect Category Sentiment Analysis</i>
ASBA	: Analisis Sentimen Berbasis Aspek
ATSA	: <i>Aspect Term Sentiment Analysis</i>
BERT	: <i>Bidirectional Encoder Representations from Transformers</i>
CRP	: <i>Consensus Reaching Process</i>
GDM	: <i>Group Decision Making</i>
LLM	: <i>Large Language Model</i>
NLP	: <i>Natural Language Processing</i>
POS	: <i>Part-of-Speech</i>
PBS	: <i>Penalized Brier Score</i>
SOTA	: <i>State-of-the-art</i>

# BAB I

## PENDAHULUAN

Bab ini menjelaskan latar belakang yang menjadi dasar pemilihan topik serta urgensi dari penelitian tugas akhir mengenai pengembangan model pengukuran *echo chamber* dengan analisis sentimen berbasis aspek dan konsensus *Group Decision Making*. Selanjutnya, bab ini merumuskan permasalahan yang akan diselesaikan, tujuan yang ingin dicapai, serta ukuran keberhasilan dalam pelaksanaan penelitian. Batasan masalah turut dijabarkan untuk memperjelas ruang lingkup agar penelitian tetap fokus dan terarah. Terakhir, disampaikan gambaran umum metodologi yang digunakan sebagai dasar penyusunan dan pelaksanaan tugas akhir ini.

### I.1 Latar Belakang

Di era digital saat ini, media sosial menjadi wadah utama bagi masyarakat untuk menyampaikan opini, berinteraksi, dan membentuk narasi atas isu-isu yang sedang berkembang. Sebuah survei oleh Katadata Insight Center (KIC) dan Kementerian Komunikasi dan Informatika menunjukkan bahwa 73% masyarakat Indonesia memperoleh informasi melalui media sosial (Karmila dan Yuningsih, 2025). Selain itu, penelitian oleh Juleha dkk. (2024) menyoroti bahwa di Indonesia, media sosial turut membentuk pola komunikasi, partisipasi politik, dan opini publik, baik dalam dimensi positif maupun negatif. Namun, di balik peran penting tersebut, interaksi di media sosial tidak lepas dari permasalahan. Salah satu isu yang kerap muncul adalah terbentuknya *echo chamber*. *Echo chamber* merujuk pada ruang media yang tertutup dan terbatas, yang memperkuat pesan-pesan tertentu sekaligus mengisolasi dari pesan yang berlawanan (Jamieson dan Cappella, 2008). Fenomena seperti ini dapat mendorong polarisasi sosial ekstrem yang pada akhirnya berpotensi mempermudah penyebaran misinformasi (Cinelli dkk., 2021).

*Echo chamber* telah menjadi objek kajian di berbagai studi. Morini dkk. (2021) mengidentifikasi tiga pendekatan utama dalam studi terkait isu ini, di antaranya

pendekatan berbasis konten, berbasis jaringan, dan *hybrid*. Sebagian besar penelitian terdahulu berfokus pada pendekatan berbasis jaringan melalui deteksi komunitas untuk memetakan segregasi sosial. Meskipun efektif dalam mengidentifikasi struktur komunitas, pendekatan ini memiliki kelemahan karena mengabaikan informasi semantik dan sentimen dari isi percakapan. Padahal, informasi ini penting untuk memahami pandangan seseorang terhadap suatu isu. Beberapa penelitian lain yang menggunakan pendekatan berbasis konten atau *hybrid* juga masih berfokus pada sentimen keseluruhan terhadap suatu isu, tanpa mempertimbangkan aspek-aspek spesifik di dalamnya. Padahal, setiap isu umumnya memiliki dimensi yang kompleks, yang mencakup beragam aspek yang dapat memicu respons berbeda. Ketika aspek-aspek ini tidak dipertimbangkan, pemahaman terhadap respons individu menjadi kurang representatif.

Studi Amendola dkk. (2024) menawarkan pendekatan baru untuk mengukur *echo chamber* melalui pendekatan *hybrid* yang berfokus pada pemanfaatan teknik Analisis Sentimen Berbasis Aspek (ASBA) dan metrik konsensus dalam kerangka *Group Decision Making* (GDM). Analisis Sentimen Berbasis Aspek (ASBA) merupakan teknik analisis teks yang mengidentifikasi sentimen secara spesifik terhadap berbagai aspek dalam suatu isu. Setelah mendapatkan keluaran ASBA, dilakukan pengukuran tingkat kesepakatan antarpengguna menggunakan metrik konsensus, yang biasa diterapkan dalam studi GDM, terutama dalam fase pencapaian konsensus (*Consensus Reaching Process* atau CRP).

Amendola dkk. (2024) mengawali pendekatannya dengan mendefinisikan komunitas melalui tahap deteksi komunitas menggunakan algoritma *multi-level graph partitioning*. Selanjutnya, digunakan model ASBA hasil *fine-tuning* BERT (*Bidirectional Encoder Representations from Transformers*) untuk mengukur sentimen masing-masing pengguna untuk tiap aspek. Setelah itu, pengukuran konsensus dilakukan dengan menghitung *cosine similarity* antarpasangan vektor sentimen pengguna terhadap aspek tertentu. Nilai konsensus antarpengguna untuk setiap aspek ini kemudian diagregasikan menjadi nilai konsensus aspek. Terakhir, nilai konsensus aspek tersebut diagregasikan lagi untuk mendapatkan nilai konsensus dalam komunitas dan antarkomunitas.

Pendekatan Amendola dkk. (2024) dinilai unggul karena mampu menghasilkan metrik ringkas untuk mengukur *echo chamber* dengan mempertimbangkan berbagai aspek, polarisasi internal komunitas, dan disparitas antarkomunitas. Di balik keunggulannya, keterbatasan utama penelitian ini adalah tidak adanya pengkajian mendalam mengenai bagaimana model ASBA yang digunakan dapat diterapkan pada berbagai domain. Model ASBA *state-of-the-art* (SOTA) umumnya dikembangkan khusus untuk domain tertentu. Akibatnya, model perlu di-*fine-tune* untuk setiap domain baru. Hal ini menjadi tantangan karena data berlabel untuk *fine-tuning* model pada setiap domain baru tidak selalu tersedia dan membutuhkan anotasi yang *resource-intensive*. Di sisi lain, fenomena *echo chamber* pada media sosial dapat terjadi di berbagai domain, sehingga membutuhkan model ASBA yang lebih fleksibel. Terlebih lagi, penelitian mengenai pengukuran *echo chamber* otomatis berbasis *hybrid* dalam konteks bahasa Indonesia saat ini belum ditemukan.

## I.2 Rumusan Masalah

Pendekatan pengukuran *echo chamber* yang sudah ada, seperti metode berbasis jaringan, masih memiliki keterbatasan dalam menangkap dinamika opini secara semantik dan kontekstual. Di sisi lain, Analisis Sentimen Berbasis Aspek (ASBA) dan metrik konsensus dalam kerangka *Group Decision Making* (GDM) telah terbukti efektif mengevaluasi keselarasan opini dalam komunitas media sosial. Meski demikian, fenomena *echo chamber* pada media sosial dapat terjadi di berbagai domain. Sedangkan, model ASBA SOTA berbasis *supervised deep learning* yang umum digunakan sangat bergantung pada domain spesifik. Padahal, *fine-tuning* untuk setiap domain baru sulit dilakukan karena memerlukan data berlabel yang tidak selalu tersedia. Anotasi untuk setiap domain baru juga *resource-intensive* dan sering kali tidak mungkin dilakukan. Selain itu, penelitian pengukuran *echo chamber* otomatis berbasis *hybrid* dalam konteks bahasa Indonesia saat ini belum ditemukan. Berdasarkan beberapa pertimbangan tersebut, masalah yang ingin diselesaikan dalam tugas akhir ini adalah bagaimana mengembangkan model pengukuran *echo chamber* pada percakapan media sosial berbahasa Indonesia dengan memanfaatkan ASBA bebas domain dan metrik konsensus GDM?

### **I.3 Tujuan dan Ukuran Keberhasilan Pencapaian**

Tujuan dari tugas akhir ini adalah menghasilkan model pengukuran *echo chamber* pada percakapan media sosial berbahasa Indonesia dengan memanfaatkan ASBA bebas domain dan metrik konsensus GDM. Ukuran keberhasilan ditentukan melalui kinerja model yang lebih baik dibandingkan pendekatan terdahulu dalam mengukur *echo chamber* pada percakapan media sosial berbahasa Indonesia.

### **I.4 Batasan Masalah**

Batasan masalah dalam tugas akhir antara lain:

1. Pada komponen jaringan, deteksi komunitas menggunakan metode yang telah terbukti efektif berdasarkan studi literatur tanpa eksperimen lebih lanjut.
2. Subtugas ASBA yang digunakan pada tugas akhir ini adalah ACSA (*Aspect Category Sentiment Analysis*) dengan pendefinisian kategori aspek dilakukan secara manual tanpa menggunakan metode otomatis.
3. *Dataset* yang digunakan berupa percakapan berbahasa Indonesia dari media sosial X.

### **I.5 Metodologi**

Tahapan yang dilalui selama pelaksanaan tugas akhir ini antara lain:

1. Analisis Persoalan

Dalam tahap ini, dilakukan analisis komprehensif terhadap berbagai tantangan dan keterbatasan yang relevan dalam studi *echo chamber* di media sosial. Berbagai studi sebelumnya, termasuk penelitian Amendola dkk. (2024), ditelaah untuk mengenali potensi masalah dalam mengadaptasi atau mengembangkan model pengukuran *echo chamber*. Fokus utama adalah menguraikan secara sistematis kendala-kendala yang perlu diatasi. Hasil dari analisis ini akan menjadi dasar bagi perancangan solusi pada tahap selanjutnya.

## 2. Analisis dan Perancangan Solusi

Tahap ini dimulai dengan menganalisis berbagai alternatif solusi yang berpotensi mengatasi tantangan dan keterbatasan yang telah diidentifikasi pada tahap sebelumnya. Analisis ini didasarkan pada studi literatur terhadap persoalan-persoalan serupa atau yang memiliki potensi untuk diadaptasi. Selanjutnya, dilakukan perancangan arsitektur solusi yang akan menghasilkan representasi visual dan deskriptif dari sistem yang diusulkan.

## 3. Pengumpulan dan Pengolahan Data

Data dikumpulkan dari *dataset* percakapan media sosial X berbahasa Indonesia yang tersedia, dengan fokus pada percakapan yang membahas topik-topik nyata yang kontroversial. Setelah itu, data diproses melalui tahap praproses dan anotasi manual sesuai kebutuhan.

## 4. Implementasi Solusi

Pada tahap ini, solusi diimplementasikan berdasarkan hasil dari tahap perancangan solusi. Hal ini mencakup eksperimen pemilihan model ASBA bebas domain, serta implementasi dan integrasi seluruh *pipeline* pengukuran *echo chamber* mulai dari ASBA, deteksi komunitas, hingga perhitungan konsensus.

## 5. Evaluasi

Untuk mengukur kinerja model, evaluasi akan dilakukan berdasarkan metrik yang telah ditetapkan dalam tahap perancangan solusi. Proses evaluasi ini bertujuan untuk mengukur kemampuan model dalam menyelesaikan masalah, menemukan potensi kesalahan dan area perbaikan, serta memberikan gambaran menyeluruh tentang efektivitas solusi dan peluang pengembangannya.

## I.6 Sistematika Pembahasan

Sistematika pembahasan tugas akhir ini disusun untuk memberikan alur yang jelas dan terstruktur mengenai pengembangan model pengukuran *echo chamber* berbasis

Analisis Sentimen Berbasis Aspek (ASBA) dan konsensus *Group Decision Making* (GDM). Setiap bab memiliki fokus dan tujuan tertentu sebagai berikut:

1. Bab I Pendahuluan

Menjelaskan latar belakang yang menjadi dasar pemilihan topik serta urgensi tugas akhir. Bab ini merumuskan permasalahan, tujuan, dan ukuran keberhasilan tugas akhir, serta membatasi ruang lingkup tugas akhir. Selain itu, disampaikan gambaran umum metodologi yang digunakan.

2. Bab II Kajian Pustaka

Memuat hasil studi literatur terkait pengukuran *echo chamber* dengan ASBA dan konsensus GDM. Bab ini membahas konsep teoretis, temuan penelitian sebelumnya, serta kerangka pemikiran yang menjadi dasar penelitian.

3. Bab III Analisis Persoalan dan Rancangan Solusi

Memaparkan analisis persoalan untuk mengidentifikasi tantangan dalam pengukuran *echo chamber*. Analisis ini kemudian dilanjutkan dengan evaluasi berbagai metode relevan yang diakhiri dengan perancangan dan implementasi solusi.

4. Bab IV Eksperimen dan Pengujian

Menyajikan proses eksperimen dan pengujian yang dilakukan, termasuk tujuan eksperimen, skenario yang digunakan, hasil yang diperoleh, serta analisis terhadap hasil untuk mengevaluasi efektivitas solusi yang diusulkan.

5. Bab V: Kesimpulan dan Saran

Menyimpulkan temuan penelitian, menilai pencapaian tujuan penelitian, serta memberikan saran untuk penelitian atau pengembangan lebih lanjut di masa mendatang.

## **BAB II**

### **KAJIAN PUSTAKA**

Bab ini memuat hasil studi literatur yang berkaitan dengan pengembangan model pengukuran *echo chamber* dengan ASBA Bebas Domain dan konsensus GDM. Dalam bab ini, dibahas berbagai konsep teoretis, temuan dari penelitian sebelumnya, serta kerangka pemikiran yang menjadi dasar studi.

#### **II.1 *Echo Chamber***

Jamieson dan Capella dalam *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment* (2008) mendefinisikan *echo chamber* sebagai ruang media yang tertutup dan terbatas, yang memperkuat pesan-pesan tertentu sekaligus mengisolasi pesan yang berlawanan. Penguatannya terjadi melalui interaksi berulang antarindividu dengan keyakinan atau pandangan yang sama (Morini dkk., 2021). Sebagai media interaksi yang masif dan terbuka, media sosial kerap menjadi ruang terbentuknya *echo chamber*.

Meskipun kerap disamakan, *echo chamber* memiliki perbedaan mendasar dengan *filter bubble*. *Filter bubble*, sebagaimana dikemukakan oleh Eli Pariser dalam *The Filter Bubble* (2011), terjadi akibat algoritma yang menyaring informasi sesuai dengan preferensi pengguna tanpa keterlibatan aktif pengguna tersebut. Sebaliknya, *echo chamber* terbentuk tidak hanya karena algoritma, tetapi juga karena pilihan aktif individu (Colleoni dkk., 2014).

Fenomena *echo chamber* memiliki serangkaian dampak negatif yang signifikan. Pertama, fenomena ini akan memperkuat bias kognitif individu, menyebabkan seseorang makin yakin akan kebenaran perspektifnya sendiri. Hal ini berpotensi mempertajam polarisasi pada masyarakat dan menghambat tercapainya pemahaman bersama antara kelompok-kelompok dengan pandangan yang berbeda (Djusr dkk., 2025). Lebih lanjut, terbentuknya polarisasi sosial ekstrem ini dapat memfasilitasi penyebaran misinformasi dengan lebih mudah dan cepat (Cinelli dkk., 2021).

## **II.2 Kategori Pendekatan Studi *Echo Chamber***

Berdasarkan penelitian-penelitian terdahulu, pendekatan studi *echo chamber* pada media sosial dapat dikelompokkan menjadi tiga kategori: (Morini dkk., 2021)

### **1. Pendekatan Berbasis Konten**

Pendekatan berbasis konten mengevaluasi *echo chamber* dengan asumsi bahwa lingkungan yang terpolarisasi dapat teridentifikasi dengan menganalisis kecenderungan konten yang dibagikan atau dikonsumsi oleh pengguna serta mengukur sentimennya terhadap isu kontroversial. Pendekatan ini berfokus pada kecenderungan ideologis pengguna tanpa memperhitungkan pola interaksi dengan pengguna lain (Morini dkk., 2021).

Salah satu metrik yang sering digunakan dalam pendekatan ini adalah analisis sentimen yang mengukur polaritas emosi dalam konten yang dibagikan pengguna. Tingkat sentimen yang seragam dalam suatu komunitas dapat mengindikasikan adanya *echo chamber*. Selain itu, terdapat teknik analisis topik yang digunakan untuk mengidentifikasi tema dominan dalam percakapan. Konsistensi topik dalam komunitas tertentu dapat menunjukkan polarisasi atau isolasi informasi (Mahmoudi dkk., 2024).

### **2. Pendekatan Berbasis Jaringan**

Pendekatan berbasis jaringan mengevaluasi *echo chamber* dengan menganalisis pola interaksi antarpengguna dalam suatu jaringan. Pendekatan ini berfokus pada struktur keterhubungan dalam jaringan sosial. *Echo chamber* diidentifikasi melalui keberadaan komunitas-komunitas yang terkelompok secara topologis, tanpa memperhatikan konten interaksi (Morini dkk., 2021).

Salah satu metrik yang sering digunakan dalam pendekatan ini adalah modularitas, yang membandingkan kepadatan sisi dalam komunitas dengan kepadatan sisi yang diharapkan dalam model *null*. Nilai modularitas yang tinggi menunjukkan bahwa komunitas lebih terisolasi satu sama lain, yang dapat mengindikasikan keberadaan *echo chamber*. Selain itu, *coverage* mengukur rasio jumlah sisi dalam komunitas terhadap total sisi dalam jaringan. Makin

mendekati satu, makin besar kemungkinan komunitas tersebut terisolasi dari jaringan lain. Metrik *Random Walk Controversy* (RWC) mengevaluasi probabilitas dua *random walk* tetap berada dalam komunitas yang sama dibandingkan dengan probabilitas berpindah ke komunitas lain. Jika probabilitasnya tinggi, komunitas tersebut cenderung mengalami polarisasi. Variasi dari metrik ini meliputi *Authoritative RWC* (ARWC), yang memulai *random walk* dari simpul otoritatif, serta *Displacement RWC* (DRWC), yang mengukur perbandingan jumlah langkah antara dua komunitas terhadap total panjang *walk*. Terakhir, *Boundary Connectivity* (BC) mengukur hubungan antara simpul internal dan simpul batas komunitas. Jika hubungan antaranggota dalam komunitas lebih kuat dibandingkan dengan koneksi ke luar komunitas, maka tingkat isolasi dan potensi *echo chamber* makin tinggi (Amendola dkk., 2024).

### 3. Pendekatan *Hybrid*

Pendekatan *hybrid* menggabungkan pendekatan berbasis konten dan berbasis jaringan untuk studi *echo chamber* yang lebih komprehensif. Pendekatan ini tidak hanya mengukur kecenderungan ideologis pengguna berdasarkan konten yang dikonsumsi dan dibagikan, tetapi juga memetakan interaksi sosialnya dalam jaringan. Pada praktiknya, pendekatan *hybrid* mengintegrasikan metrik-metrik yang digunakan pada pendekatan berbasis konten dan jaringan.

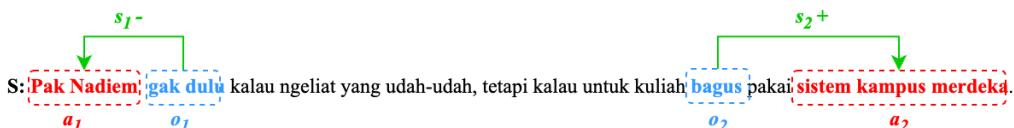
### II.3 Analisis Sentimen Berbasis Aspek (ASBA)

Analisis Sentimen Berbasis Aspek (ASBA) adalah teknik analisis teks yang mengelompokkan data berdasarkan aspek tertentu dan menentukan polaritas sentimen yang terkait dengan masing-masing aspek tersebut (Amendola dkk., 2024). Berbeda dengan analisis sentimen tradisional yang mengukur sentimen secara keseluruhan dalam sebuah teks, ASBA memungkinkan identifikasi sentimen yang lebih terperinci terhadap berbagai aspek dalam teks.

Beberapa elemen utama ASBA di antaranya: (Pontiki dkk., 2016)

1. Kategori Aspek (*Aspect Category, c*): aspek khusus dari sebuah entitas yang menjadi objek opini dalam sebuah kalimat. Kategori aspek bersifat *predefined* untuk bidang atau domain tertentu.
2. Ekspresi Aspek (*Aspect Term, a*): ekspresi linguistik yang menjadi target ekspresi sentimen. Ekspresi linguistik ini dapat disebutkan secara eksplisit maupun implisit dalam teks.
3. Ekspresi Sentimen (*Opinion Term, o*): kata atau frasa yang mengungkapkan sentimen dari pemegang opini terhadap ekspresi aspek tertentu.
4. Kategori Sentimen (*Sentiment Polarity, s*): klasifikasi sentimen yang diberikan terhadap kategori atau ekspresi aspek. Dapat berupa positif, negatif, atau netral.

Sebagai contoh, pada kalimat *tweet* tentang kampus merdeka: "Pak Nadiem gak dulu kalau *ngeliat* yang *udah-udah*, tetapi kalau untuk kuliah bagus pakai sistem kampus merdeka". Pada kalimat tersebut, kategori aspek yang muncul adalah "kebijakan menteri" (karena mengacu langsung pada kebijakan menteri) dan *umum* (karena mengacu pada program kampus merdeka secara keseluruhan). Ekspresi aspek berupa "Pak Nadiem" dan "sistem kampus merdeka", sedangkan ekspresi sentimen adalah "gak dulu" dan "bagus". Kategori sentimennya yaitu "negatif" untuk kebijakan menteri dan "positif" untuk umum. Visualisasi contoh elemen sentimen dapat dilihat pada Gambar II.1.



Gambar II.1 Visualisasi Contoh Elemen Sentimen

Berdasarkan elemen-elemen sentimen di atas, ASBA dapat didefinisikan sebagai proses mengidentifikasi ekspresi sentimen yang terkait dengan ekspresi aspek tertentu dalam berbagai kategori aspek, lalu mengklasifikasikan kategori sentimen

dari aspek yang terdeteksi (Amendola dkk., 2024). Dalam implementasinya, terdapat beberapa subtugas utama dalam ASBA. Salah satunya adalah *Aspect Category Sentiment Analysis* (ACSA), yang mengklasifikasikan sentimen berdasarkan kategori aspek yang telah ditentukan tanpa perlu mengidentifikasi ekspresi aspek secara eksplisit. Sementara itu, *Aspect Term Sentiment Analysis* (ATSA) berfokus pada klasifikasi sentimen terhadap ekspresi aspek spesifik dalam kalimat.

### **II.3.1 Metode ASBA Bebas Domain**

Penelitian-penelitian ASBA terkini menunjukkan bahwa sebagian besar metode yang ada sangat bergantung pada domain tertentu (*domain-dependent*). Ketergantungan ini muncul karena baik ekstraksi aspek maupun klasifikasi sentimen sangat dipengaruhi oleh konteks dan topik domain dari data pelatihan. Misalnya, kata "dingin" bisa merujuk pada ruangan yang sejuk atau pelayanan yang buruk tergantung konteksnya. Konteks seperti ini tidak hanya bersifat lokal (sekitar kata target), tetapi juga global (keseluruhan dokumen), sehingga model yang tidak mampu menangkap konteks relevan akan kesulitan melakukan klasifikasi secara akurat (Hua dkk., 2024).

Oleh karena itu, pendekatan ASBA SOTA umumnya dikembangkan khusus untuk domain tertentu dengan memanfaatkan sumber daya seperti *domain-specific dataset*, tag POS, relasi dependensi sintaksis, dan basis pengetahuan berbasis ontologi (Hua dkk., 2024). Hal ini membuat performa model cenderung baik hanya pada domain yang digunakan saat pelatihan, dan tidak *robust* saat diterapkan pada domain baru (*unseen domain*). Pendekatan *cross-domain transfer* yang bertujuan untuk mengurangi ketergantungan terhadap domain juga masih memiliki keterbatasan karena perbedaan distribusi probabilitas dan ruang fitur antardomain sering kali menyebabkan inkonsistensi performa. Oleh sebab itu, pemilihan domain sumber yang tepat menjadi krusial, sebab kualitas data berlabel pada domain sumber sangat menentukan kesesuaian transfer ke domain target. Hal ini seringkali tidak memungkinkan terutama pada ruang domain yang sangat luas. Lebih jauh lagi, beberapa metode *cross-domain* tetap membutuhkan data berlabel dari domain

target untuk proses *fine-tuning*, yang dalam praktiknya mahal dan sering kali tidak tersedia (Nazir dkk., 2020).

Permasalahan muncul ketika ASBA diterapkan pada skenario bebas domain (*domain-independent*), yaitu kondisi ketika sistem harus mampu menganalisis opini secara efektif tanpa dibatasi oleh domain tertentu. Dalam konteks ini, beberapa metode ASBA dapat digunakan karena kemampuannya untuk tetap *robust* pada domain baru dengan kebutuhan konfigurasi minimal dan tanpa kebutuhan data berlabel *domain-specific*. Metode-metode tersebut antara lain *double propagation* dan *zero-shot Large Language Model* (LLM) generatif.

#### **II.3.1.1 Double Propagation**

*Double propagation* adalah sebuah metode berbasis *bootstrapping* yang secara simultan melakukan ekspansi leksikon opini (ekspresi sentimen) dan target opini (ekspresi aspek) menggunakan keterkaitan sintaktis dalam kalimat. Metode yang diperkenalkan oleh Qiu dkk. (2011) ini unggul karena dapat bekerja hanya dengan *seed* leksikon awal yang terbatas, tanpa memerlukan data berlabel tambahan.

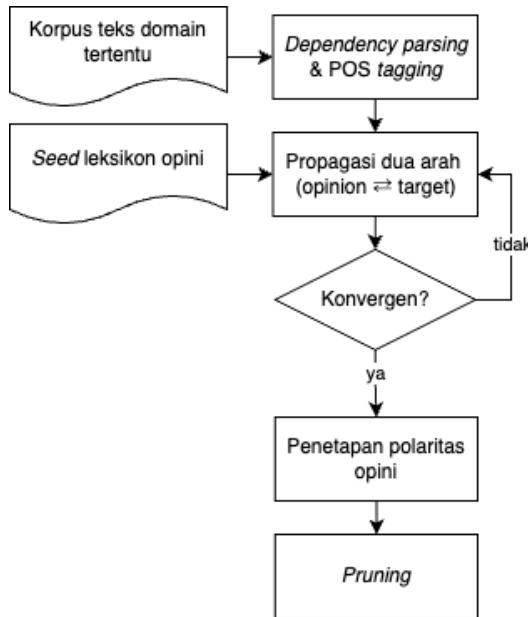
Proses yang dilakukan dalam metode *double propagation* dapat dilihat pada Gambar II.2. Proses dimulai dengan tahap inisialisasi, sistem memerlukan dua masukan utama, yaitu sebuah *seed* leksikon opini bebas domain dan korpus teks tanpa label dari domain tertentu. *Seed* leksikon opini terdiri dari sejumlah kata sifat yang sudah diketahui memiliki polaritas sentimen, seperti “*good*”, “*bad*”, dan “*excellent*”. Leksikon ini berfungsi sebagai titik awal proses ekstraksi opini dan target opini secara *semi-supervised* (Qiu dkk., 2011).

Langkah selanjutnya adalah analisis sintaktis dan pelabelan kelas kata terhadap setiap kalimat dalam korpus dengan *Part-of-Speech (POS) Tagging*. Dengan menggunakan alat bantu seperti *POS Tagger* dan *dependency parser*, struktur gramatis antarkata dianalisis untuk mengidentifikasi hubungan antarkata. Struktur yang dianalisis khususnya antara kata sifat (JJ, JJR, JJS) dan kata benda (NN, NNS). Hasil *parsing* ini menjadi dasar untuk proses berikutnya (Qiu dkk., 2011).

Proses inti dari metode ini adalah proses propagasi yang dibagi menjadi empat subtugas utama. Setiap subtugas dirancang untuk secara iteratif memperluas

kumpulan kata opini atau target yang diketahui. Keempat subtugas tersebut adalah: (1) ekstraksi target menggunakan kata opini; (2) ekstraksi target menggunakan target yang telah diekstraksi; (3) ekstraksi kata opini menggunakan target yang telah diekstraksi; dan (4) ekstraksi kata opini menggunakan kata opini yang diberikan maupun yang telah diekstraksi. Untuk menyelesaikan subtugas-subtugas ini, berbagai jenis relasi dependensi sintaksis dimanfaatkan, di antaranya, OT-Rel (relasi antara opini dan target) untuk subtugas (1) dan (3), TT-Rel (relasi antartarget) untuk subtugas (2), dan OO-Rel (relasi antaropini) untuk subtugas (4). Didefinisikan empat jenis aturan propagasi untuk masing-masing subtugas, dapat dilihat pada Gambar II.3. Penting untuk dicatat bahwa pendekatan ini hanya menggunakan dependensi langsung (*direct dependency*). Hal ini didasari pertimbangan bahwa *parsing* dengan ekspresi informal yang sering ditemukan di lingkungan *web* cenderung lebih sulit dan rawan kesalahan dibanding teks formal. Untuk teks formal lebih cocok menggunakan dependensi tidak langsung (*indirect dependency*). Semua aturan yang didefinisikan bersifat *domain-independent*, artinya dapat diterapkan pada berbagai domain teks tanpa memerlukan penyesuaian khusus (Qiu dkk., 2011).

Dalam definisi aturan, o (atau t) merepresentasikan kata opini (atau target) yang merupakan keluaran atau yang diekstraksi. Simbol {O} (atau {T}) adalah himpunan kata opini (atau target) yang sudah diketahui, baik dari masukan awal maupun hasil ekstraksi sebelumnya. Simbol H menunjukkan sembarang kata. POS(O(atau T)) mengacu pada informasi dari kata opini (O) atau target (T), sedangkan O(atau T)-Dep merujuk pada relasi dependensi sintaksis dari kata opini (O) atau target (T) (Qiu dkk., 2011).



Gambar II.2 Metode *Double Propagation*

RuleID	Observations	output	Examples
R1 <sub>1</sub>	$O \rightarrow O\text{-}Dep \rightarrow T$ s.t. $O \in \{O\}$ , $O\text{-}Dep \in \{MR\}$ , $POS(T) \in \{NN\}$	$t = T$	The phone has a good “screen”. (good → mod → screen)
R1 <sub>2</sub>	$O \rightarrow O\text{-}Dep \rightarrow H \leftarrow T\text{-}Dep \leftarrow T$ s.t. $O \in \{O\}$ , $O/T\text{-}Dep \in \{MR\}$ , $POS(T) \in \{NN\}$	$t = T$	“iPod” is the <u>best</u> mp3 player. (best → mod → player ← subj ← iPod)
R2 <sub>1</sub>	$O \rightarrow O\text{-}Dep \rightarrow T$ s.t. $T \in \{T\}$ , $O\text{-}Dep \in \{MR\}$ , $POS(O) \in \{JJ\}$	$o = O$	same as R1 <sub>1</sub> with screen as the known word and good as the extracted word
R2 <sub>2</sub>	$O \rightarrow O\text{-}Dep \rightarrow H \leftarrow T\text{-}Dep \leftarrow T$ s.t. $T \in \{T\}$ , $O\text{-}Dep \in \{MR\}$ , $POS(O) \in \{JJ\}$	$o = O$	same as R1 <sub>2</sub> with iPod as the known word and best as the extract word
R3 <sub>1</sub>	$T_{i(j)} \rightarrow T_{i(j)}\text{-}Dep \rightarrow T_{j(i)}$ s.t. $T_{i(j)} \in \{T\}$ , $T_{j(i)} \in \{T\}$ , $T_{i(j)}\text{-}Dep \in \{CONJ\}$ , $POS(T_{i(j)}) \in \{NN\}$	$t = T_{i(j)}$	Does the player play dvd with <u>audio</u> and “video”? (video → conj → audio)
R3 <sub>2</sub>	$T_i \rightarrow T_i\text{-}Dep \rightarrow H \leftarrow T_j\text{-}Dep \leftarrow T_j$ s.t. $T_i \in \{T\}$ , $T_i\text{-}Dep == T_j\text{-}Dep$ , $POS(T_j) \in \{NN\}$	$t = T_j$	Canon “G3” has a great <u>lens</u> . (lens → obj → has ← subj ← G3)
R4 <sub>1</sub>	$O_{i(j)} \rightarrow O_{i(j)}\text{-}Dep \rightarrow O_{j(i)}$ s.t. $O_{j(i)} \in \{O\}$ , $O_{i(j)}\text{-}Dep \in \{CONJ\}$ , $POS(O_{i(j)}) \in \{JJ\}$	$o = O_{i(j)}$	The camera is <u>amazing</u> and “easy” to use. (easy → conj → amazing)
R4 <sub>2</sub>	$O_i \rightarrow O_i\text{-}Dep \rightarrow H \leftarrow O_j\text{-}Dep \leftarrow O_j$ s.t. $O_i \in \{O\}$ , $O_i\text{-}Dep == O_j\text{-}Dep$ , $POS(O_j) \in \{JJ\}$	$o = O_j$	If you want to buy a <u>sexy</u> , “cool”, accessory-available mp3 player, you can choose iPod. (sexy → mod → player ← mod ← cool)

Gambar II.3 Aturan Propagasi (Qiu dkk., 2011)

Sementara itu, himpunan *tag* POS yang digunakan untuk identifikasi di antaranya {JJ} yaitu himpunan *tag* POS untuk kata opini potensial, yang mencakup JJ (kata sifat), JJR (kata sifat komparatif), dan JJS (kata sifat superlatif). Selanjutnya, {NN} adalah himpunan *tag* POS untuk target potensial, yang terdiri dari NN (kata benda tunggal) dan NNS (kata benda jamak). Himpunan {MR} berisi relasi dependensi yang menggambarkan hubungan antara kata opini dan target (misalnya, "mod", "pnmod", "subj", "s", "obj", "obj2", dan "desc"). Himpunan {CONJ} hanya berisi relasi "conj" (konjungsi). Panah dalam deskripsi aturan ( $\rightarrow$ ) menunjukkan arah dependensi, misalnya, " $O \rightarrow O\text{-Dep} \rightarrow T$ " berarti  $O$  bergantung pada  $T$  melalui relasi sintaksis  $O\text{-Dep}$ . Operator "==" menunjukkan kesamaan atau ekuivalensi, ekuivalensi berarti "mod" sama dengan "pnmod", dan "s" atau "subj" sama dengan "obj". Contohnya, " $T_i\text{-Dep} == T_j\text{-Dep}$ " berarti " $T_i\text{-Dep}$ " sama atau ekuivalen dengan " $T_j\text{-Dep}$ " (misalnya, "subj" dan "obj" pada aturan R3<sub>2</sub>) (Qiu dkk., 2011).

Secara spesifik, aturan R1<sub>i</sub> digunakan untuk mengekstrak target (t) menggunakan kata opini (O), aturan R2<sub>i</sub> untuk mengekstrak kata opini (o) menggunakan target (T), aturan R3<sub>i</sub> untuk mengekstrak target (t) menggunakan target yang telah diekstraksi (T), dan aturan R4<sub>i</sub> untuk mengekstrak kata opini (o) menggunakan kata opini yang telah diketahui (O<sub>i</sub>) (Qiu dkk., 2011).

Sebagai contoh, pada aturan R1<sub>i</sub> yang berbunyi " $O \rightarrow O\text{-Dep} \rightarrow T$  s.t.  $O \in \{O\}$ ,  $O\text{-Dep} \in \{MR\}$ ,  $POS(T) \in \{NN\}$ " dengan keluaran  $t = T$ , dijelaskan bahwa jika diberikan kata opini O yang diketahui, dan kata tersebut bergantung pada kata lain T melalui relasi O-Dep yang termasuk dalam himpunan {MR}, serta kata T memiliki *tag* POS sebagai kata benda (NN), maka kata T diekstraksi sebagai target. Sebagai ilustrasi, pertimbangkan kalimat "*The phone has a good screen*". Jika "good" dikenal sebagai kata opini, dan "good" bergantung pada "screen" melalui relasi "mod" (yang ada di {MR}), dan "screen" di-tag sebagai NN, maka R11 dapat diterapkan untuk mengekstrak "screen" sebagai target (Qiu dkk., 2011).

Secara umum, tahap propagasi ini dilakukan secara iteratif, kata opini dan target yang baru diekstraksi akan digunakan kembali untuk menemukan entitas lainnya. Proses ini terus berjalan hingga tidak ditemukan lagi pasangan baru yang valid, menandakan proses telah konvergen. Dengan demikian, meskipun hanya berangkat

dari leksikon kecil, metode ini mampu memperluas cakupan opini dan target secara bertahap dengan cakupan yang tinggi (Qiu dkk., 2011). Algoritma propagasi dapat dilihat pada Gambar II.4.

```
Input: Opinion Word Dictionary { O} , Review Data R
Output: All Possible Features { F} , The Expanded Opinion Lexicon { O-Expanded}
Function:
1. { O-Expanded}={ O}
2. { F} = Ø , { O} = Ø
3. for each parsed sentence in R
4.     if(Extracted features not in { F})
5.         Extract features { Fi} using R11 and R12 based on opinion words in { O-Expanded}
6.     endif
7.     if(Extracted opinion words not in { O-Expanded})
8.         Extract new opinion words { Oj} using R41 and R42 based on opinion words in { O-Expanded}
9.     endif
10.    endfor
11. Set { F} ={ F} +{ Fi} , { O-Expanded} ={ O-Expanded} +{ Oj}
12. for each parsed sentence in R
13.     if(Extracted features not in { F})
14.         Extract features { F'} using R31 and R32 based on features in { Fi}
15.     endif
16.     if(Extracted opinion words not in { O-Expanded})
17.         Extract opinion words { O'} using R21 and R22 based on features in { F'}
18.     endif
19. end for
20. Set { F} ={ F} +{ F'}, { O} ={ O} +{ O'}
21. Set { F} ={ F} +{ F'}, { O-Expanded} ={ O-Expanded} +{ O'}
22. Repeat 2 till size({ F}) =0, size({ O}) =0
```

Gambar II.4 Algoritma Propagasi (Qiu dkk., 2011)

Setelah entitas opini baru ditemukan, sistem melanjutkan ke tahap penetapan polaritas opini. Penetapan ini didasarkan pada sejumlah aturan. Pertama, opini yang diekstrak dari target atau opini lain serta target yang diekstraksi dari target lain akan mewarisi polaritas sumbernya dengan memperhatikan kata-kata negasi (*heterogeneous* dan *homogeneous rule*). Sementara itu, opini yang tidak memiliki referensi langsung akan mengambil polaritas keseluruhan dokumen (*intra-review rule*). Sistem juga memperhitungkan kata-kata negasi seperti “*not*”, “*but*”, atau “*however*” untuk menyesuaikan polaritas secara kontekstual (Qiu dkk., 2011).

Terakhir, dilakukan proses *pruning* target opini untuk meningkatkan presisi hasil ekstraksi dengan tiga teknik utama. Pertama, *clause-based pruning* untuk menghindari duplikasi target dalam satu klausa. Kedua, *heuristic pruning* untuk

membuang kata-kata yang bukan merupakan aspek opini, seperti nama penjual dan pembeli pada domain ulasan. Ketiga, *phrase merging* untuk menggabungkan kata benda menjadi frasa aspek yang utuh, misalnya “*battery*” dan *life* menjadi “*battery life*”. Selain itu, target yang hanya muncul sekali dalam korpus juga akan disaring melalui *frequency-based pruning* untuk mengurangi *noise* yang tidak relevan (Qiu dkk., 2011).

Metode ini menunjukkan hasil yang signifikan dalam perluasan leksikon opini dan ekstraksi target. Dalam serangkaian eksperimen yang membandingkan metode ini dengan pendekatan lainnya seperti CRF, CRF-D, KN06, Hu, Popescu, dan PLSA, *double propagation* secara konsisten meraih F-score tertinggi. Hasil eksperimen dapat dilihat pada Gambar II.5. Keunggulan utama *double propagation* terletak pada kemampuannya untuk beroperasi secara *semi-supervised*, hanya membutuhkan *seed* leksikon opini awal berukuran kecil, serta secara efektif memanfaatkan hubungan dependensi sintaktis antara kata opini dan target (Qiu dkk., 2011).

	F-score						
	Hu	Popescu	PLSA	CRF	CRF-D	noProp-dep	Prop-dep
D1	0.78	0.84	0.58	0.46	0.55	0.80	0.84
D2	0.75	0.80	0.53	0.50	0.58	0.73	0.85
D3	0.74	0.81	0.56	0.40	0.50	0.79	0.88
D4	0.75	0.83	0.49	0.42	0.52	0.79	0.82
D5	0.77	0.84	0.54	0.38	0.52	0.75	0.89
Avg	0.76	0.82	0.54	0.43	0.54	0.77	0.86

Gambar II.5 Hasil F-score Eksperimen yang Membandingkan *Double Propagation* (Prop-dep) dengan Pendekatan Lain (Qiu dkk., 2011)

### II.3.1.2 *Zero-Shot Large Language Model (LLM) generatif*

Di era perkembangan *Large Language Models* (LLM) generatif, teknologi ini makin banyak digunakan dalam berbagai tugas pemrosesan bahasa alami, termasuk ASBA. Model generatif adalah model yang belajar untuk menghasilkan data baru,

seperti teks atau gambar, berdasarkan pola yang ditemukan dalam data yang sudah ada. Model generatif mencoba untuk memahami bagaimana suatu data terbentuk lalu menciptakan data baru yang menyerupai (Mitchell, 2015). Hal ini berbeda dengan model diskriminatif yang fokus pada membedakan atau mengklasifikasikan data ke dalam kategori tertentu (Mitchell, 2015). Sementara itu, *Large Language Models* merupakan subset dari pembelajaran mendalam yang memanfaatkan arsitektur model besar untuk memahami dan menganalisis teks dengan tingkat akurasi yang tinggi. Oleh karena itu, LLM generatif adalah LLM yang mampu menghasilkan teks baru berdasarkan konteks yang telah dipelajari.

<b>Prompt Example</b>
<p>As an aspect-based sentiment analyzer, your task is to extract sentiments (Positive, Negative, or Neutral) from people's review for a given list of aspects.</p> <p>Each review consists of a text and may contain single or multiple aspects and their corresponding sentiments. You will be provided with the review and the list of aspects.</p> <p>Your goal is to process the reviews and given list of aspects. Then, predict sentiment of each given aspect.</p> <p>Please present the result in specified format.</p> <p>Review: {text}</p> <p>Aspects: {aspects}</p> <p>Required output Format:</p> <p style="color: blue;">['sentiment1', 'sentiment2', 'sentiment3', ..]</p>

Gambar II.6 Contoh *Prompt* yang Digunakan (Mughal dkk., 2024)

Mughal dkk. (2024) melakukan analisis komparatif terhadap LLM generatif, yaitu GPT-3.5-turbo dan PaLM bison-001, dengan model SOTA berbasis pembelajaran mendalam lainnya, seperti DeBERTa dan ATAE-LSTM. Studi ini mengevaluasi performa model di berbagai domain dalam dua subtugas utama ASBA, yaitu *Aspect*

*Category Sentiment Analysis* (ACSA) dan *Aspect Term Sentiment Analysis* (ATSA). Analisis dilakukan menggunakan *zero-shot prompt engineering*, yang berarti model diuji kemampuannya untuk melakukan tugas-tugas tanpa memerlukan contoh atau pelatihan tambahan untuk setiap tugas atau domain. *Prompt* tersusun dari teks masukan, kumpulan kategori aspek dan ekspresi aspek, tiga kelas sentimen, serta format keluaran yang diharapkan. Contoh *prompt* dapat dilihat pada Gambar II.6.

Hasil penelitian pada tugas ACSA (Gambar II.7) menunjukkan bahwa PaLM secara konsisten memiliki performa yang baik dalam berbagai domain, mengungguli DeBERTa dalam hal akurasi dan F1-score. Namun, performanya menurun pada *dataset* kompleks yang tiap ulasannya terdiri dari banyak aspek dan sentimen. Sebaliknya, GPT-3.5 menunjukkan potensi besar dalam menangani *dataset* kompleks, dengan hasil yang bersaing dengan model khusus seperti ATAE-LSTM. Hasil ini dicapai tanpa pelatihan tambahan maupun *dataset* berlabel. Hal ini menunjukkan kapabilitas generalisasi yang kuat pada domain baru (Mughal dkk., 2024).

Akurasi						
Dataset	Domain	Baseline	ATAE-LSTM	DeBERTa	PaLM	GPT-3.5
Sentihood[18]	Neighbourhood	93.3[61]	84.0	<b>87.6</b>	<b>90.7</b>	67.8
SemEval16[17]	Laptop	88.4 [81]	68.7	<b>81.9</b>	<b>91.0</b>	67.8
	Restaurant	84.0[13]	80.3	<b>83.5</b>	<b>92.7</b>	81.1
MAMS[13]	Restaurant	74.0 [13]	63.9	<b>77.4</b>	55.8	<b>64.5</b>

F1-Score						
Dataset	Domain	Baseline	ATAE-LSTM	DeBERTa	PaLM	GPT-3.5
Sentihood	Neighbourhood	87.9	83.6	87.4	<b>90.9</b>	65.4
SemEval16	Laptop	-	68.8	79.0	<b>88.7</b>	75.4
	Restaurant	-	79.0	80.7	<b>90.9</b>	85.7
MAMS	Restaurant	-	63.1	77.2	48.4	<b>64.4</b>

Gambar II.7 Hasil Penelitian Komparasi *Zero-Shot* LLM Generatif dengan Model Lain pada Tugas ACSA (Mughal dkk., 2024)

Keunggulan utama LLM generatif dalam ASBA terletak pada beberapa faktor utama. Pertama, kemampuan *transfer learning* yang diperoleh dari *pre-training* pada tugas bahasa umum memberikan dasar yang kuat untuk penyesuaian pada tugas ASBA, sekaligus mendukung generalisasi yang lebih baik terhadap domain baru. Selain itu, pemahaman bahasa yang mendalam memungkinkan LLM generatif menangkap struktur gramatis kompleks serta hubungan semantik dalam teks. Terakhir, *reasoning* kontekstual yang dimiliki LLM generatif memungkinkan analisis sentimen yang lebih akurat, terutama dalam kalimat yang mengandung opini yang kontras (Mughal dkk., 2024).

### II.3.2 Metrik Evaluasi ASBA

Evaluasi untuk ASBA, khususnya pada subtugas ACSA (*Aspect Category Sentiment Analysis*) penting untuk menilai sejauh mana model mampu mengklasifikasikan kategori aspek dan sentimen. Klasifikasi kategori aspek umumnya bersifat multikelas, karena terdapat lebih dari dua kelas, dan multilabel, karena setiap teks dapat dikaitkan dengan lebih dari satu kategori aspek. Metrik utama yang umum digunakan dalam klasifikasi multikelas adalah *precision*, *recall*, serta *F1-score* baik dalam bentuk makro maupun mikro.

*Precision* mengukur sejauh mana prediksi positif yang dibuat oleh model benar-benar relevan. Secara matematis *precision* untuk dirumuskan pada persamaan (2.1) (Opitz, 2024).

$$Precision = \frac{TP}{TP+FP} \quad (2.1)$$

Dengan *TP* adalah jumlah *true positive* dan *FP* adalah jumlah *false positive*. *Precision* menjadi penting ketika kesalahan prediksi aspek tertentu (*false positive*) dapat menurunkan reliabilitas analisis, misalnya dalam sistem rekomendasi produk yang harus menghindari mengaitkan ulasan dengan aspek yang salah. Berbeda dengan *precision*, *recall* menekankan pada kemampuan model dalam menemukan semua sampel yang relevan. Secara matematis *recall* dirumuskan pada persamaan (2.2) (Opitz, 2024).

$$Recall = \frac{TP}{TP+FN} \quad (2.2)$$

Dengan  $FN$  adalah jumlah *false negative*. Untuk menyeimbangkan *trade-off* antara *precision* dan *recall*, digunakan *F1-score*, yang didefinisikan sebagai rata-rata harmonik keduanya. Secara matematis *F1-score* dirumuskan pada persamaan (2.3) (Opitz, 2024).

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.3)$$

*F1-score* dianggap lebih representatif dibanding hanya menggunakan *precision* atau *recall* secara terpisah, khususnya pada *dataset* dengan distribusi kelas tidak seimbang. Pada klasifikasi multikelas dan multilabel seperti dalam kategori aspek, terdapat dua pendekatan utama untuk menghitung *F1*, yaitu *F1 makro* dan *F1 mikro*. *F1makro* menghitung nilai *F1* untuk setiap kelas secara terpisah kemudian merata-ratakannya, sehingga setiap kategori aspek diberi bobot yang sama terlepas dari ukuran kelasnya. Sebaliknya, *F1 mikro* menghitung total *TP*, *FP*, dan *FN* dari semua kelas terlebih dahulu, kemudian menghasilkan satu skor global. Dengan demikian, *F1 makro* lebih adil dalam menilai performa pada kelas minoritas, terutama jika seluruh kelas sama pentingnya (Opitz, 2024).

Pada klasifikasi sentimen *soft-label*, dibutuhkan evaluasi yang tidak hanya menilai prediksi kelas, tetapi juga kualitas probabilitas yang dihasilkan model. Oleh karena itu, metrik yang banyak digunakan adalah *Brier Score*, yang mengukur akurasi probabilistik dengan menghitung rata-rata kuadrat selisih antara probabilitas prediksi dan label aktual. Semakin rendah *Brier Score*, semakin baik kualitas probabilitas yang dihasilkan model. Karena memiliki tiga kelas (positif, negatif, dan netral), klasifikasi sentimen juga termasuk klasifikasi multikelas. Secara matematis, *Brier Score* untuk klasifikasi probabilitas multikelas dirumuskan pada persamaan (2.4) (Glenn, 1950).

$$BS = \frac{1}{N} \sum_{t=1}^N \sum_{k=1}^K (f_{t,k} - o_{t,k})^2 \quad (2.4)$$

Keterangan:

$N$  = jumlah sampel,

$K$  = jumlah kelas,

$f_{t,k}$  = probabilitas prediksi bahwa sampel ke- $t$  termasuk ke dalam kelas  $k$ ,

$o_{t,k}$  = label aktual sampel ke- $t$  untuk kelas  $k$ , bernilai 1 jika kelas  $k$  adalah label sebenarnya, dan 0 sebaliknya.

Namun, kelemahan *Brier Score* adalah bahwa prediksi yang yakin (probabilitas tinggi) tetapi salah kadang mendapat skor lebih baik daripada prediksi yang ragu (probabilitas rendah) namun benar. Untuk mengatasi hal tersebut, Ahmadian dkk. (2025) mengembangkan *Penalized Brier Score* (PBS), yang menambahkan penalti sebesar  $(K - 1)/K$  apabila kelas prediksi tertinggi tidak sama dengan label aktual.

Dalam evaluasi probabilistik klasifikasi umumnya diperlukan *baseline* agar skor model dapat dibandingkan dengan prediksi naif. Untuk *Brier Score*, *baseline* umumnya didefinisikan sebagai model yang selalu memprediksi probabilitas *prior* kelas, yaitu distribusi frekuensi kelas dalam *dataset* (Ferro dkk., 2012). Misalnya, jika dalam *dataset* 70% sampel bernilai positif, 20% netral, dan 10% negatif, maka *baseline* model akan selalu mengeluarkan prediksi [0,7, 0,2, 0,1] untuk semua sampel. Secara matematis, probabilitas *prior* dituliskan pada persamaan (2.5). *Baseline Brier Score* kemudian dihitung dengan memasukkan nilai probabilitas *prior* ini ke persamaan (2.4). *Baseline* ini berfungsi sebagai tolak ukur minimal, karena setiap model yang lebih buruk dari *baseline* probabilistik *prior* dianggap tidak lebih baik dari tebakan distribusi kelas.

$$p_k = \frac{n_k}{N} \quad (2.5)$$

Keterangan:

$N$  = jumlah sampel,

$n_k$  = jumlah sampel pada kelas  $k$ ,

$p_k$  = probabilitas *prior* kelas  $k$ .

Lebih lanjut, *Brier Skill Score* (BSS) digunakan untuk menilai seberapa baik model dibandingkan *baseline*. BSS dihitung menggunakan persamaan (2.6).

$$BSS = 1 - \frac{BS}{BS_{ref}} \quad (2.6)$$

Dengan  $BS$  adalah *Brier Score* model yang diuji dan  $BS_{ref}$  adalah *Brier Score* dari model *baseline*. Nilai  $BSS > 0$  menunjukkan bahwa model lebih baik daripada *baseline*, sementara  $BSS < 0$  berarti model lebih buruk dari *baseline*. Dengan demikian, *Brier Skill Score* memungkinkan interpretasi yang lebih intuitif dalam membandingkan performa model relatif terhadap prediksi naif berbasis distribusi *prior* kelas (Ferro dkk., 2012).

#### II.4 Representasi Vektor Teks

Representasi vektor teks merupakan pendekatan penting dalam pemrosesan bahasa alami (NLP) untuk mengubah teks menjadi bentuk numerik sehingga dapat diolah oleh komputer. Setiap kata, kalimat, atau dokumen dipetakan ke dalam ruang vektor berdimensi banyak, di mana kedekatan antarvektor dapat digunakan untuk mengukur kesamaan semantik atau kesamaan dokumen (Salton dkk., 1975). Secara umum, metode representasi vektor dapat dibedakan menjadi pendekatan *sparse* seperti TF-IDF dan pendekatan *dense* seperti Word2Vec, FastText, serta Sentence-

BERT yang lebih padat dan mampu menangkap informasi semantik dengan lebih baik (Jha, 2025).

Salah satu metode klasik yang banyak digunakan adalah *Term Frequency–Inverse Document Frequency* (TF-IDF). Metode ini menghitung bobot sebuah kata dalam dokumen berdasarkan frekuensi kemunculannya, dengan mengurangi pengaruh kata yang sering muncul di banyak dokumen dalam korpus (Arroyo-Fernández dkk., 2017). Dengan demikian, kata-kata yang jarang tetapi relevan akan memiliki bobot lebih tinggi. TF-IDF sederhana, efisien, dan efektif untuk tugas yang mengandalkan pencocokan kata kunci, seperti pencarian informasi. Namun, kelemahannya adalah tidak mampu menangkap konteks maupun urutan kata dalam kalimat serta menghasilkan vektor yang sangat *sparse* dengan dimensi besar seiring bertambahnya kosakata (Jha, 2025).

Selanjutnya, Mikolov dkk. (2013) memperkenalkan Word2Vec. Word2Vec menggunakan jaringan saraf (*neural network*) sederhana dengan dua arsitektur utama, yakni *Continuous Bag-of-Words* (CBOW) yang memprediksi kata target dari konteks sekitarnya, dan Skip-gram yang memprediksi konteks dari kata target. Melalui jaringan ini, Word2Vec menghasilkan *embedding* kata berupa vektor *dense* yang mampu menangkap hubungan semantik dan sintaksis. Keunggulan Word2Vec terletak pada efisiensi dan kemampuannya menghasilkan representasi kata yang bermakna. Meski demikian, Word2Vec masih memiliki keterbatasan karena setiap kata direpresentasikan sebagai satu vektor tetap, sehingga tidak dapat membedakan makna kata polisemi (bermakna ganda) yang bergantung pada konteks kalimat (Mikolov dkk., 2013).

Untuk mengatasi keterbatasan tersebut, FastText dikembangkan oleh Facebook AI Research dengan menambahkan representasi berbasis sub-kata (n-gram karakter) dalam pembentukan *embedding*. Dengan pendekatan ini, *embedding* suatu kata diperoleh tidak hanya dari kata utuh, tetapi juga dari potongan-potongan karakter penyusunnya. Hal ini membuat FastText lebih unggul dalam menangani bahasa dengan morfologi kompleks serta lebih *robust* terhadap kata baru (*out-of-vocabulary*) maupun kesalahan pengetikan. FastText terbukti lebih fleksibel dibandingkan Word2Vec, namun tetap memiliki keterbatasan karena belum mampu

menangkap konteks penuh dalam kalimat. Dengan demikian, meskipun lebih baik daripada TF-IDF dan Word2Vec dalam hal generalisasi kosakata, FastText masih kurang optimal dalam memahami makna kalimat yang bergantung pada konteks (Bojanowski dkk., 2017).

Lebih lanjut, *Sentence-BERT* (SBERT) diperkenalkan oleh Reimers dan Gurevych (2019) sebagai modifikasi dari BERT yang dirancang untuk menghasilkan representasi kalimat secara lebih efisien. SBERT memanfaatkan arsitektur *siamese* dan triplet *network*, sehingga memungkinkan pembelajaran *embedding* kalimat yang dapat langsung dibandingkan menggunakan ukuran kesamaan vektor seperti *cosine similarity*. Pendekatan ini secara signifikan mengurangi biaya komputasi yang biasanya sangat tinggi pada BERT standar ketika menghitung kemiripan antar pasangan kalimat dalam jumlah besar. SBERT menunjukkan kinerja unggul pada berbagai *benchmark Semantic Textual Similarity* (STS) sekaligus mempercepat proses inferensi hingga ratusan kali lipat dibandingkan BERT konvensional. Meskipun demikian, SBERT tetap memiliki keterbatasan, misalnya kebutuhan komputasi relatif tinggi pada tahap pelatihan, serta ketergantungan pada data berlabel untuk *fine-tuning* jika digunakan pada domain yang spesifik (Reimers & Gurevych, 2019).

## **II.5 Metrik Konsensus dalam Kerangka *Group Decision Making***

Masalah pengambilan keputusan kelompok (*Group Decision Making* atau GDM) dapat didefinisikan sebagai situasi pengambilan keputusan yang melibatkan: (1) dua atau lebih ahli dengan gagasan, sikap, motivasi, dan pengetahuan masing-masing; (2) suatu masalah yang perlu diselesaikan; (3) upaya untuk mencapai solusi bersama. Untuk menyelesaikan masalah GDM, terdapat dua tahap solusi, yaitu proses pencapaian konsensus (*Consensus Reaching Process* atau CRP) dan proses seleksi (Herrera-Viedma dkk., 2005).

Metrik konsensus digunakan pada tahap pertama, CRP, yang berkaitan dengan upaya mencapai tingkat kesepakatan atau persetujuan maksimal di antara para ahli terhadap himpunan alternatif solusi. *Consensus Reaching Process* bekerja dengan mengumpulkan preferensi individu, mengagregasikannya menjadi preferensi kelompok, dan mengevaluasi tingkat kesepakatan. Tingkat kesepakatan inilah yang

diukur dengan metrik konsensus. Metrik konsensus biasanya terbagi menjadi dua tingkat abstraksi, yaitu metrik untuk menghitung nilai konsensus pasangan dan metrik untuk menghitung nilai konsensus global atau agregat (Herrera-Viedma dkk., 2005).

Terdapat berbagai metrik konsensus, salah satunya adalah pada penelitian Morente-Molinera dkk. (2019) yang menghitung nilai konsensus pasangan berdasarkan selisih jumlah kata positif dan kata negatif dibandingkan dengan jumlah total kedua nilai tersebut, kemudian dikonversi secara linier ke rentang [0,1]. Nilai konsensus global kemudian dihitung dari rata-rata (*mean*) nilai konsensus pasangan. Pada penelitian ini digunakan ambang batas (*threshold*) 0.5, nilai konsensus di atas 0.5 menandakan kesepakatan, sedangkan nilai konsensus di bawah 0.5 menandakan ketidaksepakatan.

Selain itu, pada penelitian Amendola dkk., (2024) yang menerapkan metrik konsensus pada hasil ASBA, nilai konsensus pasangan dihitung berdasarkan *cosine similarity* antarpasangan vektor sentimen. Karena diterapkan pada hasil ASBA, sebelum menghitung konsensus global terdapat tingkat abstraksi tambahan yaitu *Aspect Consensus*. Nilai *Aspect Consensus* dihitung dengan *power average mean* dari setiap nilai konsensus pasangan. Nilai konsensus global, atau yang dalam penelitian ini disebut *Community Consensus*, kemudian dihitung dengan *power average mean* dari setiap nilai *Aspect Consensus*. Secara lebih detail berikut adalah rumus yang digunakan dalam penelitian Amendola dkk., (2024).

1. Konsensus pasangan, dihitung dengan *cosine similarity* antarvektor sentimen aspek dari setiap pasangan pengguna, seperti pada persamaan (2.7).

$$sv_k^{i,j} = \frac{s_k^i s_k^j}{\|s_k^i\| \|s_k^j\|} \quad (2.7)$$

Keterangan:

$S_k^i$  = kolom dari matriks sentimen aspek  $s^i$  yang menunjukkan sentimen vektor tiga nilai dari pengguna ke- $i$  pada aspek ke- $k$ .

Hasilnya adalah vektor kesamaan yang mencerminkan kemiripan sentimen pengguna terhadap masing-masing aspek yang dianalisis.

2. *Aspect consensus* untuk setiap aspek topik, dihitung dengan mengagregasikan konsensus pasangan menggunakan *power average mean*, seperti pada persamaan (2.8).

$$ac_z = \left( \frac{1}{t} \sum_{i=1}^n \sum_{j=1}^n |sv_z^{i,j}|^p \right)^{\frac{1}{p}} \quad (2.8)$$

Keterangan:

$sv_z^{i,j}$  = konsensus pasangan antara pengguna ke- $i$  dan ke- $j$  pada aspek ke- $z$ ,

$n$  = jumlah pengguna,

$$t = \frac{n(n-1)}{2},$$

$p$  = nilai  $p$ -norm.

Hasilnya adalah vektor *aspect consensus* yang menunjukkan tingkat kesepakatan dalam komunitas untuk masing-masing aspek.

3. *Community consensus*, dihitung dengan mengagregasikan *aspect consensus* menggunakan *power average mean*, seperti pada persamaan (2.9).

$$cc = \left( \frac{1}{m} \sum_{z=1}^m |ac_z|^p \right)^{\frac{1}{p}} \quad (2.9)$$

Keterangan:

$ac_z = aspect \ consenus$  untuk aspek ke- $z$ ,

$z = 1, 2, \dots, m$ ,

$m =$  jumlah aspek,

$p =$  nilai  $p$ -norm.

Hasilnya adalah satu nilai tunggal yang mencerminkan kesepakatan dalam komunitas terhadap suatu topik secara keseluruhan.

Berikut adalah contoh perhitungan konsensus dalam komunitas untuk tiga pengguna dan dua kategori aspek. Sentimen vektor tiga nilai untuk setiap pengguna pada setiap aspek (positif, netral, negatif) sebagai berikut.

1. Pengguna 1 memiliki sentimen  $[0,8, 0,1, 0,1]$  untuk Aspek 1 dan  $[0,2, 0,6, 0,2]$  untuk Aspek 2
2. Pengguna 2 memiliki sentimen  $[0,7, 0,2, 0,1]$  untuk Aspek 1 dan  $[0,1, 0,8, 0,1]$  untuk Aspek 2
3. Pengguna 3 memiliki sentimen  $[0,3, 0,3, 0,4]$  untuk Aspek 1 dan  $[0,3, 0,2, 0,5]$  untuk Aspek 2.

Dengan persamaan (2.7), didapatkan untuk Aspek 1, konsensus pasangan adalah  $sv_1^{1,2} \approx 0,988$  (Pengguna 1 dan Pengguna 2),  $sv_1^{1,3} \approx 0,655$  (Pengguna 1 dan Pengguna 3), dan  $sv_1^{2,3} \approx 0,722$  (Pengguna 2 dan Pengguna 3). Sementara itu, untuk Aspek 2, hasilnya adalah  $sv_2^{1,2} \approx 0,966$ ,  $sv_2^{1,3} \approx 0,686$ , dan  $sv_2^{2,3} \approx 0,480$ . Selanjutnya, *aspect consensus* didapat melalui persamaan (2.8). Dengan jumlah pasangan unik ( $t$ ) sebanyak 3 dan nilai  $p = 1$ , konsensus untuk Aspek 1 ( $ac_1$ ) adalah sekitar 0,788. Untuk Aspek 2 ( $ac_2$ ), nilai konsensus yang diperoleh adalah sekitar 0,711. Hal ini menunjukkan bahwa Aspek 1 memiliki tingkat kesepakatan yang lebih tinggi di antara pengguna dibandingkan Aspek 2. Terakhir, *community consensus* dihitung melalui persamaan (2.9). Dengan jumlah aspek ( $m$ ) sebanyak 2 dan nilai  $p = 1$ , konsensus komunitas ( $cc$ ) mencapai nilai sekitar 0,750. Nilai ini merupakan satu nilai tunggal yang mencerminkan tingkat kesepakatan

keseluruhan dalam komunitas terhadap topik yang dianalisis, menunjukkan tingkat konsensus yang cukup tinggi secara keseluruhan.

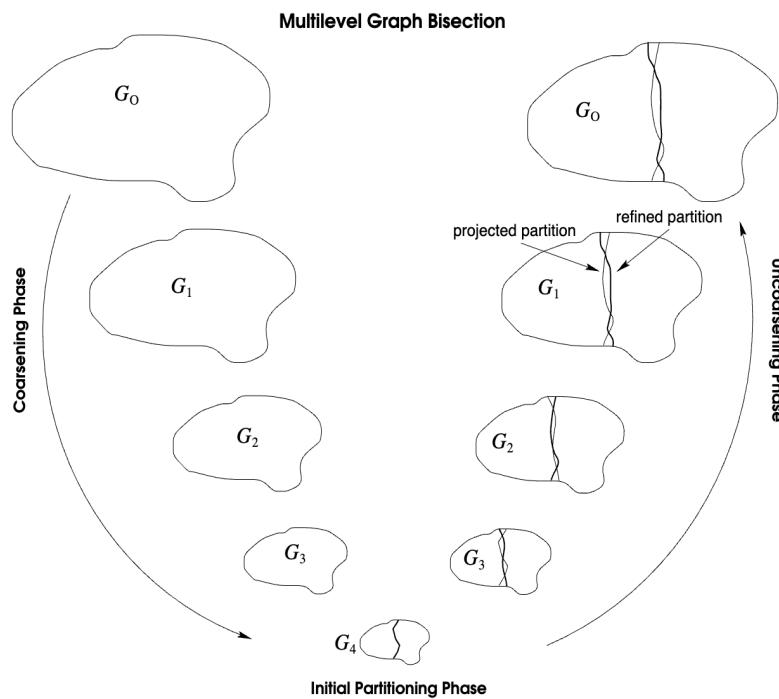
## II.6 Metode Deteksi Komunitas

Komunitas, atau yang disebut juga *cluster* atau modul, didefinisikan sebagai kelompok simpul (*vertices*) dalam sebuah graf yang memiliki kesamaan sifat dan/atau peran dalam graf (Fortunato, 2010). Dalam kehidupan nyata, masyarakat secara alami membentuk kelompok semacam ini dalam berbagai bentuk, seperti keluarga, lingkaran pertemanan, kota, dan lain-lain. Dengan hadirnya internet, bentuk-bentuk komunitas virtual juga bermunculan dan berkembang pesat. Dalam ilmu jaringan (*network science*), tujuan utama dari deteksi komunitas adalah mengidentifikasi modul-modul, lalu, bila memungkinkan, menemukan struktur hierarkisnya dengan hanya memanfaatkan informasi yang terkandung dalam topologi graf (Fortunato, 2010).

Salah satu pendekatan klasik yang digunakan untuk menyelesaikan masalah deteksi komunitas adalah *graph partitioning*. Pendekatan ini bertujuan untuk membagi simpul-simpul di dalam graf menjadi beberapa kelompok dengan ukuran yang telah ditentukan, sedemikian rupa sehingga jumlah sisi (*edges*) yang menghubungkan antarkelompok tersebut bernilai minimal. Jumlah sisi yang berada di antara kelompok-kelompok ini disebut *cut size* (Fortunato, 2010).

Salah satu metode *graph partitioning* klasik adalah *multilevel graph partitioning*. Secara umum, metode ini memperkecil ukuran graf melalui proses *coarsening*, yaitu menggabungkan simpul dan sisi sehingga terbentuk representasi graf yang lebih kecil. Pemilihan pasangan simpul untuk penggabungan didasarkan pada bobot sisi pada graf berbobot, atau menggunakan strategi acak (*random*) pada graf tak berbobot. Partisi kemudian dihitung pada graf hasil penyusutan ini, lalu diproyeksikan kembali (*uncoarsening*) ke graf semula. Pada tahap ini dilakukan langkah *refinement* untuk menyesuaikan penempatan simpul agar kualitas pemotongan (*cut quality*) semakin baik tanpa mengorbankan keseimbangan ukuran partisi, misalnya dengan algoritma Kernighan–Lin (KL) (Karypis dan Kumar, 1995). Visualisasi proses ini dapat dilihat pada Gambar II.8. *Multilevel graph partitioning* telah diimplementasikan dalam perangkat lunak METIS. Perangkat

lunak METIS versi terbaru (versi 5.1.0) dapat membagi graf tak berstruktur menjadi  $k$  bagian dengan dua metode, *multilevel recursive bisection* dan *multilevel k-way partitioning* (Karypis, 2013). *Multilevel recursive bisection* membagi graf menjadi dua secara rekursif hingga tercapai jumlah partisi yang diinginkan (Karypis dan Kumar, 1998a), sedangkan *multilevel k-way partitioning* langsung membagi graf menjadi  $k$  bagian pada tahap partisi awal. Hasil dari *multilevel k-way partitioning* cenderung lebih seimbang dengan *cut size* minimum, tetapi algoritmanya lebih kompleks dibandingkan *multilevel recursive bisection* (Karypis dan Kumar, 1998b).



Gambar II.8 Visualisasi Proses *Multilevel Graph Partitioning* (Karypis dan Kumar, 1995)

Pendekatan *graph partitioning* klasik yang berfokus pada *minimum cut size* memiliki keterbatasan karena memerlukan informasi awal berupa jumlah partisi atau ukuran komunitas yang diinginkan. Selain itu, meminimalkan *cut size* saja tidak selalu menghasilkan partisi komunitas yang baik. Sebab, kualitas sebuah komunitas tidak hanya ditentukan oleh minimnya sisi antarkomunitas, terutama dalam jaringan dunia nyata seperti jaringan sosial. Hal ini mendorong

pengembangan metode deteksi struktur komunitas yang berangkat dari asumsi bahwa jaringan secara alami terbagi menjadi subkelompok, sehingga jumlah dan ukuran komunitas ditentukan oleh struktur jaringan itu sendiri (Newman, 2006). Newman (2006) memperkenalkan metode deteksi struktur komunitas dengan optimalisasi modularitas. Modularitas didefinisikan sebagai selisih antara jumlah sisi yang berada di dalam sebuah kelompok dengan nilai ekspektasi jumlah sisi yang sama dalam jaringan acak (model *null*) yang ekuivalen. Secara matematis, modularitas diformulasikan sebagai persamaan (2.10) (Newman, 2006).

$$Q = \frac{1}{2m} \sum_{i=1}^N \sum_{j=1}^N \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (2.10)$$

Keterangan:

$A_{ij}$  = elemen matriks ketetanggaan graf, bernilai 1 (atau bobot sisi) jika terdapat sisi antara simpul  $i$  dan  $j$ , 0 jika tidak,

$k_i$  = derajat (jumlah sisi yang terhubung) simpul  $i$ ,

$m$  = jumlah total sisi dalam graf,

$\delta(c_i, c_j)$  = fungsi delta Kronecker, bernilai 1 jika simpul  $i$  dan  $j$  berada pada komunitas yang sama, 0 jika tidak.

Berdasarkan persamaan (2.10), modularitas suatu graf dengan  $c$  komunitas dapat dihitung dengan persamaan (2.11) (Fortunato, 2010).

$$Q = \sum_{c=1}^{n_c} \left[ \frac{l_c}{m} - \left( \frac{d_c}{2m} \right)^2 \right] \quad (2.11)$$

Keterangan:

$n_c$  = jumlah komunitas,

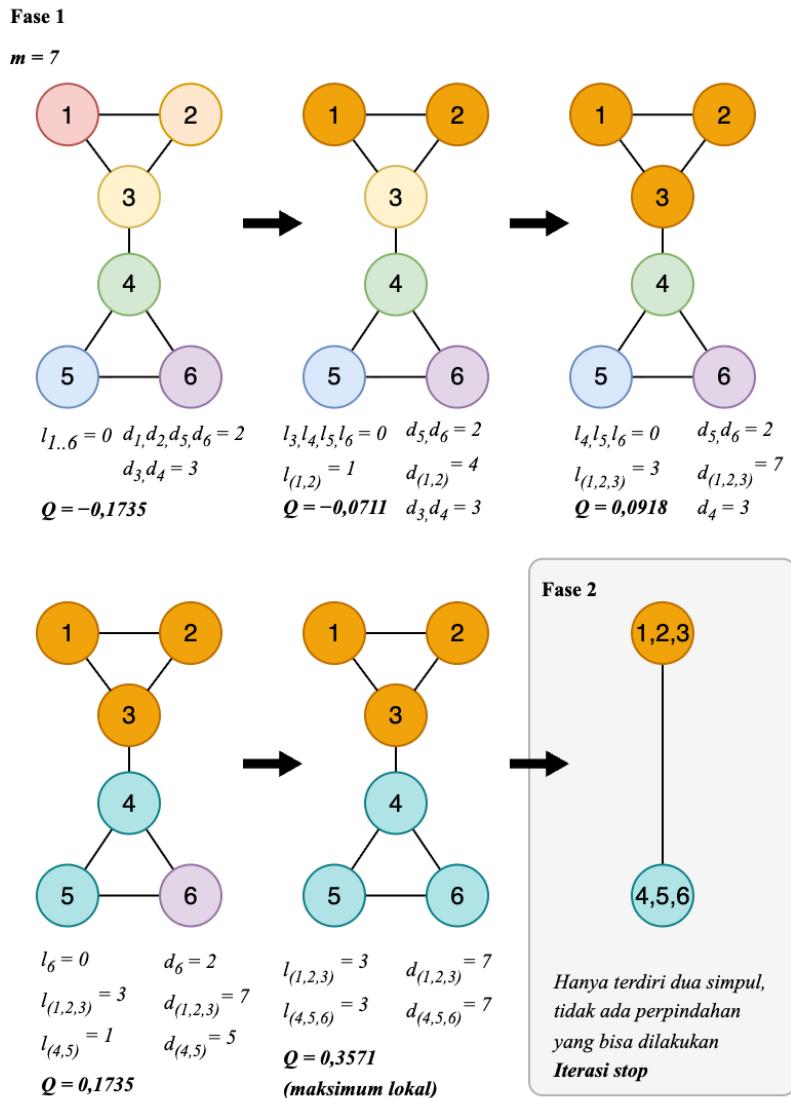
$l_c$  = jumlah sisi di dalam komunitas  $c$ ,

$d_c$  = total derajat semua simpul dalam komunitas  $c$ ,

$m$  = jumlah total sisi dalam graf.

Berbagai metode deteksi komunitas telah dikembangkan dengan prinsip optimalisasi modularitas, salah satu yang umum digunakan adalah metode Louvain yang diperkenalkan oleh Blondel dkk. (2008). Metode Louvain terdiri dari dua fase iteratif. Fase pertama adalah pengelompokan simpul ke dalam komunitas berdasarkan modularitas. Pada awal fase ini, setiap simpul dianggap sebagai komunitasnya sendiri. Lalu, untuk setiap simpul, dicek apakah memindahkan simpul tersebut ke komunitas tetangganya bisa meningkatkan modularitas. Jika terdapat peningkatan modularitas, simpul tersebut dipindahkan ke komunitas yang memberi peningkatan maksimum. Proses ini diulang hingga tidak ada lagi perpindahan yang bisa meningkatkan modularitas (mencapai maksimum lokal). Hasil akhir fase pertama adalah simpul-simpul yang sudah terkelompok. Fase kedua adalah fase agregasi simpul. Pada fase ini, komunitas yang terbentuk dari fase pertama diperlakukan sebagai simpul baru dengan bobot sisi ditentukan dari jumlah bobot sisi antarkomunitas asal. Hasil dari fase kedua adalah graf baru dengan jumlah simpul yang lebih sedikit. Setelah fase kedua selesai, iterasi berlanjut dengan mengulang proses pada fase pertama. Iterasi dilakukan sampai tercapai modularitas maksimum.

Gambar II.9 mengilustrasikan penerapan metode Louvain pada graf tak berbobot sederhana dengan enam simpul dan tujuh sisi. Ilustrasi setiap graf pada fase pertama menunjukkan pemindahan simpul ke komunitas yang memberi peningkatan maksimum untuk setiap pengecekan simpul. Pada ilustrasi ini, pengecekan simpul dilakukan berdasarkan urutan angka simpul. Nilai modularitas ( $Q$ ) dihitung berdasarkan persamaan (2.11). Fase pertama berhenti dengan nilai modularitas 0,3571. Karena agregasi simpul pada fase kedua hanya menyisakan dua simpul, tidak ada lagi perpindahan yang dapat dilakukan, sehingga iterasi cukup berjalan sekali.



Gambar II.9 Ilustrasi Penerapan Metode Louvain Pada Graf Tak Berbobot Sederhana

## II.7 Interaksi Pada X

Platform X (sebelumnya dikenal sebagai Twitter) menyediakan berbagai bentuk interaksi antarpengguna, beberapa di antaranya adalah *reply*, *mention*, dan *retweet*. Ketiga fitur ini menjadi instrumen penting dalam membangun percakapan, memperluas jangkauan pesan, serta memperkuat jejaring sosial digital. Secara

teknis, interaksi ini juga tercermin dalam metadata yang tersedia melalui X API, sehingga dapat ditelusuri secara komputasional.

*Reply* merupakan bentuk interaksi eksplisit di mana pengguna merespons unggahan (*post*) tertentu. Fitur ini ditandai dengan metadata "in\_reply\_to\_status\_id" (ID dari *post* yang dibalas), "in\_reply\_to\_user\_id" (ID *user* yang dibalas), dan "in\_reply\_to\_screen\_name" (*username* yang dibalas) (X, 2025). Penelitian oleh Zade dkk. (2024) menemukan bahwa *reply* umumnya diarahkan kepada pengunggah utama (*root tweet author*) dengan proporsi 73,5%, jauh lebih tinggi dibanding *quote tweet* yang hanya 43,3%. Hal ini menunjukkan bahwa *reply* digunakan sebagai medium komunikasi dua arah langsung, baik untuk menyetujui, menolak, maupun membingkai ulang (*reframe*) suatu diskusi (Zade dkk., 2024).

*Mention* berbeda dengan *reply* karena dapat muncul di *tweet* baru yang tidak harus terikat pada suatu unggahan tertentu. *Mention* biasanya digunakan untuk menarik perhatian pengguna lain dengan menyertakan *username* pengguna tersebut ("@*username*"). Dalam metadata *tweet* dari API X, *mention* tercatat dalam bagian "entities.user\_mentions", yang memuat informasi tentang akun yang disebut (X, 2025). *Mention* juga dapat diidentifikasi langsung dari teks *tweet* dengan mendeteksi penggunaan simbol "@*username*". Menurut Jensen dkk. (2023), *mention* berfungsi memperluas percakapan melampaui konteks awal sebuah *post*, sehingga lebih cenderung menciptakan jaringan interaksi lateral (antarpengguna) dibanding vertikal (pengguna dengan konten).

*Retweet*, di sisi lain, berfungsi sebagai mekanisme amplifikasi konten. *Retweet* memungkinkan pengguna menyebarkan ulang unggahan orang lain ke linimasa pengikutnya, baik secara langsung maupun dengan tambahan komentar (*quote tweet*). Metadata *retweet* dapat diidentifikasi melalui atribut "retweeted\_status", yang menyimpan informasi tentang *tweet* asli yang disebarluaskan (X, 2025). Selain itu, *retweet* juga dapat diidentifikasi dari teks dengan mendeteksi penggunaan pola "RT *username*". Studi sebelumnya menunjukkan bahwa *retweet* sering kali digunakan untuk memperkuat posisi atau opini tertentu dengan memperbesar jangkauan pesan tanpa harus membangun diskusi baru (Jensen dkk., 2023).

Dalam dokumentasi resmi X, interaksi ini diproses secara berbeda di linimasa. *Reply* ditautkan langsung ke *post* yang dibalas dan muncul dalam notifikasi pengunggah. *Mention* dapat menjangkau lebih luas karena tidak terbatas pada konteks *reply*. Sementara itu, *retweet* meningkatkan visibilitas pesan tanpa harus menambah percakapan. Perbedaan fungsi ini menegaskan bahwa *reply*, *mention*, dan *retweet* sama-sama menjadi indikator keterlibatan publik, namun dengan pola interaksi dan implikasi sosial yang berbeda (X, 2025).

## II.8 Null Model dan Uji Hipotesis pada Analisis Jaringan Sosial

Salah satu tantangan yang dihadapi dalam studi analisis jaringan sosial terdahulu, termasuk studi *echo chamber* di media sosial, adalah kesulitan dalam memperoleh *dataset* kontrol negatif yang valid. *Dataset* kontrol negatif diperlukan untuk membandingkan dan menguji hipotesis secara objektif. Untuk mengatasi masalah ini, diperlukan penggunaan *null model* yang dapat memberikan dasar pembanding yang valid dalam analisis jaringan sosial (Hobson dkk., 2021).

Menurut Hobson dkk. (2021), *null model* adalah model acuan yang dibuat dengan mengacak komponen data observasi. Tujuannya adalah untuk menghasilkan distribusi acak yang hanya mencerminkan struktur dasar jaringan, tanpa adanya pola spesifik yang diteliti. Model ini kemudian digunakan sebagai pembanding untuk menguji apakah pola yang diamati pada data riil bersifat signifikan secara statistik atau hanya muncul karena kebetulan.

Berbagai metode telah dikembangkan untuk membangun *null model* dalam analisis jaringan sosial. Hobson dkk. (2021) mengidentifikasi empat pendekatan utama:

1. Permutasi: Pendekatan ini melibatkan pengacakkan fitur tertentu dalam data sambil mempertahankan fitur yang lain. Misalnya, dalam konteks jaringan sosial, permutasi dapat dilakukan dengan mengacak koneksi antarsimpul tanpa mengubah derajat simpul tersebut. Pendekatan ini berguna untuk menguji hipotesis mengenai fitur yang diamati.
2. *Resampling*: Pendekatan ini membangkitkan distribusi referensi dengan cara mengambil sampel secara berulang kali dari data asli. Teknik yang populer digunakan salah satunya adalah *bootstrap* yang memungkinkan

perkiraan ketidakpastian statistik tanpa harus membuat asumsi mengenai bentuk distribusi data.

3. *Sampling* dari distribusi teoretis: Pendekatan ini menghasilkan data sintetis dengan mengambil sampel dari suatu distribusi teoretis yang diasumsikan relevan. Misalnya, sebuah jaringan acak dapat dibangun dengan mengambil sampel dari distribusi derajat yang polanya sengaja dibuat mirip dengan jaringan data asli.
4. Model generatif: Pendekatan ini menggunakan model matematis atau algoritma untuk membangun sebuah jaringan baru dari awal. Jaringan ini dibentuk berdasarkan aturan atau prinsip yang telah ditetapkan, seperti pada *configuration model* atau model berbasis agen (*agent-based model*).

Selanjutnya, *null model* akan digunakan sebagai pembanding dalam uji hipotesis. Uji hipotesis adalah prosedur statistik yang digunakan untuk menguji klaim atau dugaan tentang parameter populasi berdasarkan data sampel. Proses ini melibatkan beberapa langkah penting (González dkk., 1997):

1. Menyusun Hipotesis: Terdiri dari dua hipotesis yang saling bertentangan:
  - a. Hipotesis Nol ( $H_0$ ): Menyatakan bahwa tidak terdapat efek atau perbedaan.
  - b. Hipotesis Alternatif ( $H_1$ ): Menyatakan bahwa terdapat efek atau perbedaan.
2. Menentukan Tingkat Signifikansi ( $\alpha$ ): Biasanya ditetapkan pada 0,05, yang berarti terdapat 5% kemungkinan menolak hipotesis nol yang benar.
3. Mengumpulkan dan Menganalisis Data: Mengumpulkan data yang relevan dan menghitung statistik uji yang sesuai.
4. Menghitung Nilai  $p$ : Menentukan probabilitas mendapatkan hasil yang sama ekstremnya dengan hasil yang diamati, dengan asumsi hipotesis nol benar.
5. Membuat Keputusan: Jika nilai  $p$  lebih kecil dari  $\alpha$ , hipotesis nol ditolak; jika tidak, hipotesis nol gagal ditolak.

Sebagai contoh, dalam uji dua sisi, hipotesis nol menyatakan bahwa rata-rata populasi sama dengan nilai tertentu, sedangkan hipotesis alternatif menyatakan bahwa rata-rata populasi berbeda dari nilai tersebut.

Dalam konteks uji hipotesis dengan *null model*, terdapat dua ukuran yang umum digunakan untuk menilai signifikansi, yaitu *z-score* dan *empirical p-value*. *Z-score* menyatakan jarak suatu nilai observasi dari rata-rata distribusi dalam satuan simpangan baku. Rumus *z-score* dapat dilihat pada persamaan (2.12) (Kreyszig, 1979).

$$z = \frac{x - \mu}{\sigma} \quad (2.12)$$

dengan  $x$  sebagai nilai observasi,  $\mu$  sebagai rata-rata distribusi, dan  $\sigma$  sebagai simpangan baku. Nilai *z-score* yang besar (positif maupun negatif) menunjukkan bahwa hasil observasi semakin jauh dari kondisi yang diharapkan jika hipotesis nol benar (Kreyszig, 1979).

Sementara itu, *empirical p-value* diperoleh melalui pendekatan simulasi atau permutasi, yaitu dengan menghitung proporsi hasil acak yang sama ekstrem atau lebih ekstrem dibandingkan nilai observasi. Rumus *empirical p-value* dapat dilihat pada persamaan (2.13) (Phipson dkk., 2010).

$$\text{emp\_p} = \frac{\text{jumlah nilai simulasi } \geq x}{N} \quad (2.13)$$

dengan  $x$  sebagai nilai observasi dan  $N$  jumlah total simulasi atau permutasi. Nilai *empirical p* yang kecil menunjukkan bahwa kejadian yang sama ekstremnya dengan observasi jarang terjadi secara kebetulan, sehingga memberikan dasar untuk menolak hipotesis nol (Phipson dkk., 2010).

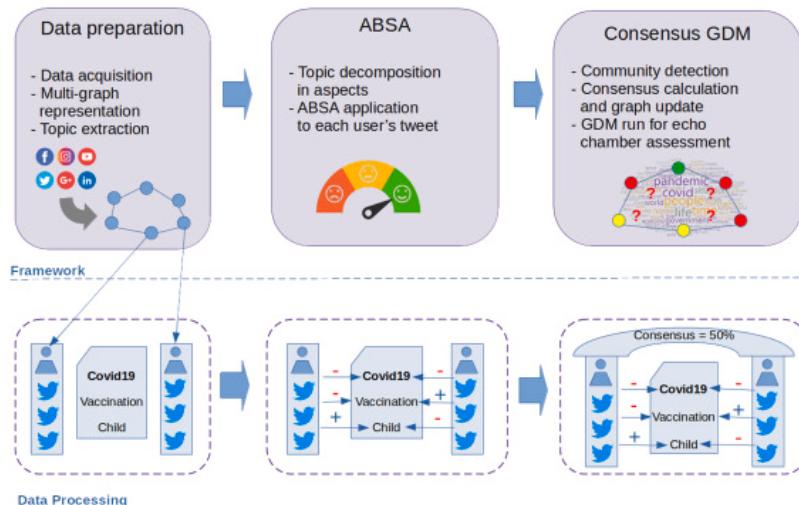
## **II.9 Penelitian Terkait**

### **II.9.1 Towards Echo Chamber Assessment by Employing Aspect-Based Sentiment Analysis and GDM Consensus Metrics (Amendola dkk., 2024)**

Penelitian ini memperkenalkan metode baru untuk mengukur efek *echo chamber* dalam media sosial dengan menggunakan ASBA dan metrik konsensus GDM. Pendekatan ini bertujuan untuk mengatasi keterbatasan metode berbasis topologi dengan memasukkan analisis semantik dari diskusi dan opini pengguna terhadap berbagai aspek untuk suatu topik.

*Pipeline* yang digunakan untuk sistem dalam penelitian ini dapat dilihat pada Gambar II.10. Pemrosesan dimulai dengan tahap persiapan data yang dikumpulkan dari media sosial Twitter (sekarang X). Grafik jaringan sosial kemudian dibentuk berdasarkan hubungan pengguna, seperti *follower/followee*, *mention*, dan *retweet* pada topik tertentu. Topik diekstraksi menggunakan *tf/idf metric*. Setelah dibentuk, grafik jaringan sosial dibagi menjadi dua komunitas (komunitas-0 dan komunitas-1) menggunakan algoritma *multilevel graph partitioning* dengan *software package* METIS.

Selanjutnya, ASBA dengan BERT diterapkan untuk menganalisis sentimen yang diekspresikan oleh pengguna terhadap berbagai aspek dari suatu topik, yang telah diekstraksi sebelumnya. Analisis ini mengidentifikasi aspek-aspek spesifik dari topik dan menentukan polaritas sentimen (positif, netral, negatif) yang terkait dengan masing-masing aspek untuk setiap pengguna. Sebagai contoh, dalam diskusi tentang "Covid," ASBA dapat membedakan antara sentimen terhadap "vaksin" dan "pembatasan." Setiap opini pengguna terhadap suatu aspek dalam topik direpresentasikan dalam bentuk vektor polaritas sentimen aspek, yang berisi tiga nilai probabilitas sentimen (positif, netral, dan negatif).



Gambar II.10 Pipeline Sistem (Amendola dkk., 2024)

Tabel II.1 Hasil Konsensus Dataset “Twitter Vaccination” dan ROT (Amendola dkk., 2024)

Dataset “Twitter Vaccination”						
Komunitas	Aspect consensus					Community Consensus
	Child	Health	Measles	Vaccination	Vaccine	
0	0,80	0,69	0,79	0,58	0,69	0,71
1	0,81	0,70	0,79	0,61	0,71	0,72
antarkomunitas	0,80	0,69	0,79	0,59	0,69	0,71
Dataset ROT						
Komunitas	Aspect consensus					Community Consensus
	Biden	Trump	President	Government	Boris	
0	0,93	0,95	0,95	0,92	0,91	0,93
1	0,85	0,85	0,89	0,90	0,84	0,86
antarkomunitas	0,76	0,78	0,93	0,85	0,82	0,83

Dalam tahap berikutnya, pengukuran *echo chamber* direformulasikan sebagai masalah GDM. Pada tahap ini, pengguna dalam suatu komunitas dianggap sebagai ahli yang mengungkapkan preferensinya (melalui sentimen) terhadap berbagai aspek suatu topik. Metrik konsensus kemudian diterapkan pada vektor polaritas sentimen aspek untuk mengukur tingkat kesepakatan di antara pengguna pada tiga tingkat, yaitu konsensus pasangan, *Aspect Consensus*, dan *Community Consensus*.

Metode ini juga mengevaluasi konsensus antarkomunitas. Konsensus antarkomunitas dihitung dengan proses yang sama dengan konsensus dalam komunitas, tetapi melibatkan pengguna dari komunitas yang berbeda. *Echo chamber* kemudian diidentifikasi dengan membandingkan konsensus dalam komunitas dan konsensus antarkomunitas. Jika suatu komunitas memiliki tingkat konsensus dalam komunitas yang tinggi namun konsensus antarkomunitas yang rendah, maka dapat disimpulkan bahwa efek *echo chamber* dalam komunitas tersebut kuat.

Metode ini divalidasi menggunakan dua *dataset* Twitter (sekarang X) yang terkait dengan diskusi tentang vaksinasi dan politik sayap kanan radikal. *Dataset* pertama, "*Twitter Vaccination*," berisi sekitar 90.000 *tweet* dari 52.510 pengguna antara Juni hingga Oktober 2019. *Dataset* kedua diperoleh dari "*Radical Right On Twitter*" (ROT), terdiri dari 82.000 *tweet* ID yang diposting oleh sekitar 41.400 pengguna antara Juli 2020 hingga Januari 2021, terkait dengan aktivitas 35 tokoh sayap kanan radikal. Hasil konsensus yang didapat dari kedua *dataset* dapat dilihat pada Tabel II.1.

Secara umum, penelitian menunjukkan bahwa metrik konsensus GDM dengan ASBA dapat memberikan wawasan yang lebih mendalam tentang *echo chamber* dalam media sosial. Dalam *dataset* vaksinasi, tingkat konsensus dalam dan antarkomunitas mencerminkan opini negatif yang seragam terhadap aspek tertentu, seperti "*measles*" dan "*vaccine*." Sementara itu, dalam *dataset* ROT, tingkat konsensus dalam komunitas yang lebih tinggi dibandingkan dengan konsensus antarkomunitas menunjukkan adanya komunitas yang lebih terpolarisasi, terutama terkait perbedaan opini terhadap aspek seperti "*Biden*."

Metric	Twitter	ROT
<b>Vaccination</b>		
Modularity (Mod) [24]	0.29	0.44
Coverage (Cov) [25]	0.98	0.96
Random Walk	0.95	0.93
Controversy (RWC) [36]		
Authoritative Random	0.48	0.71
Walk Controversy (ARWC) [5]		
Displacement Random	0.99	0.96
Walk Controversy (DRWC) [5]		
Boundary Connectivity (BC) [37]	0.23	0.11
<b>ASBA-based consensus</b>	0.71	0.83

Gambar II.11 Perbandingan Metrik Konsensus dengan Metrik SOTA Berbasis Topologi  
 (Amendola dkk., 2024)

Selain itu, dibandingkan dengan metrik SOTA berbasis topologi (Gambar II.11), metrik konsensus menawarkan perspektif yang lebih berfokus pada isi diskusi pengguna daripada hanya struktur jaringan. Metrik topologi mengukur polarisasi dari sisi hubungan antarpengguna, sedangkan metrik konsensus lebih menangkap perubahan opini yang dipengaruhi sentimen. Dengan demikian, kedua pendekatan ini dapat digunakan secara komplementer untuk memahami fenomena *echo chamber* secara lebih menyeluruh.

Penelitian ini juga menyoroti pentingnya ASBA dalam analisis sentimen, yang terbukti lebih efektif dibandingkan metode analisis sentimen tradisional seperti VADER. Dengan ASBA, analisis sentimen dapat difokuskan pada aspek spesifik dari suatu topik, sehingga memberikan wawasan yang lebih mendalam tentang reaksi pengguna. Dalam *dataset* ROT, hasil konsensus berbasis VADER berbeda secara signifikan dari hasil berbasis ASBA (Gambar II.10), menunjukkan bahwa analisis sentimen secara keseluruhan tidak cukup untuk menangkap perbedaan kecil pada opini dalam diskusi sosial.

Community	VADER-based	ASBA-based
	consensus <i>cc</i>	consensus <i>cc</i>
0	0.72	0.93
1	0.71	0.86
In between	0.43	0.83

Gambar II.12 Perbandingan Hasil Konsensus Berbasis ASBA dengan Hasil Konsensus Berbasis VADER pada *Dataset ROT* (Amendola dkk., 2024)

Secara keseluruhan, pendekatan yang dikembangkan dalam penelitian ini menawarkan metrik ringkas yang komprehensif untuk mengukur *echo chamber*, dengan mempertimbangkan baik tingkat polarisasi dalam suatu komunitas maupun perbedaan opini antarkomunitas pada berbagai aspek tertentu dari suatu topik.

### II.9.2 *Aspect-based sentiment analysis for Indonesian restaurant reviews* (Ekawati dan Khodra, 2017)

Penelitian ini membahas analisis sentimen berbasis aspek pada ulasan restoran berbahasa Indonesia dengan mengadaptasi metode terbaik dari SemEval 2016 untuk meningkatkan kinerja analisis sentimen berbasis aspek pada teks berbahasa Indonesia. Penelitian ini menggunakan lima tahap analisis sentimen berbasis aspek, di antaranya, praproses kalimat, ekstraksi ekspresi aspek, kategorisasi ekspresi aspek, klasifikasi polaritas sentimen, dan pembentukan struktur opini. Pembelajaran mesin digunakan untuk pembangunan model klasifikasi token pada tahap ekstraksi ekspresi aspek, serta klasifikasi kalimat untuk tahap kategorisasi ekspresi aspek dan klasifikasi polaritas sentimen (Ekawati dan Khodra, 2017).

Pada tahap ekstraksi ekspresi aspek, model klasifikasi token menggunakan algoritma *Conditional Random Field* (CRF) dengan 992 kalimat data latih. Fitur yang digunakan meliputi leksikal, sintaksis, dan model semantik terdistribusi seperti *Continuous Bag-of-Word* (CBOW), *Global Vectors for Word Representation* (GloVe), dan Latent Dirichlet Allocation (LDA). Probabilitas bersyarat CRF dihitung dengan persamaan (2.14).

$$p(y|x) = \frac{1}{z(x)} \exp\left(\sum_i \sum_j \lambda_j f_j(y_i, y_{i-1}, x, i)\right) \quad (2.14)$$

$f_j(y_i, y_{i-1}, x, i)$  adalah fungsi fitur transisi dan  $\lambda_j$  adalah parameter yang diestimasi. F1-measure terbaik untuk tahap ini adalah 0.793 (Ekawati dan Khodra, 2017).

Untuk tahap kategorisasi ekspresi aspek dan klasifikasi polaritas sentimen, model menggunakan algoritma *Maximum Entropy* (MaxEnt) *Classifier* dengan data latih masing-masing 1100 dan 1585 kalimat. Probabilitas kelas MaxEnt *classifier* dihitung dengan persamaan (2.15).

$$p(y|x) = \frac{1}{z(x)} \exp\left(\sum_{i=0}^n \lambda_i f_i(y, x)\right) \quad (2.15)$$

$f_i(y, x)$  adalah fungsi fitur,  $\lambda_i$  adalah bobot fitur, dan  $z(x)$  adalah konstanta normalisasi. F1-measure untuk kategorisasi ekspresi aspek adalah 0.828 dan untuk klasifikasi polaritas sentimen adalah 0,812 (Ekawati dan Khodra, 2017).

Proses implementasi dilakukan menggunakan Python di Linux, memanfaatkan *library* InaNLP untuk praproses kalimat, CRFSuite untuk model ekstraksi aspek, Scikit-learn untuk kategorisasi dan klasifikasi sentimen, serta Gensim dan StanfordNLP (GloVe) untuk model semantik terdistribusi. Eksperimen dilakukan dengan 10-fold cross validation untuk menentukan model terbaik. Model CBOW, GloVe, dan LDA dibangun dari 11470 ulasan restoran TripAdvisor yang terdiri dari 36766 kalimat. Jumlah kluster untuk CBOW dan GloVe divariasikan (100, 500, 1000, 5000), sementara LDA dengan jumlah topik (50, 100, 200, 300, 400, 500) dan jumlah kluster yang sama (Ekawati dan Khodra, 2017).

Secara keseluruhan, F1-*measure* sistem adalah 0,642. Kelebihan metode ini meliputi penggunaan berbagai fitur, termasuk model semantik terdistribusi, yang memungkinkan ekstraksi aspek yang tidak ada dalam data latih sebelumnya. Selain itu, tidak diperlukan terjemahan ekspresi aspek ke bahasa Inggris untuk kategorisasi, yang merupakan perbaikan dari metode sebelumnya. Namun, kelemahannya adalah kesulitan dalam pembentukan struktur opini, terutama untuk kalimat dengan beberapa ekspresi aspek dengan kategori yang sama namun sentimen berbeda, serta kesalahan klasifikasi akibat kalimat dengan banyak kategori (Ekawati dan Khodra, 2017).

## **BAB III**

### **ANALISIS PERSOALAN DAN RANCANGAN SOLUSI**

Bab ini memaparkan analisis persoalan yang mengidentifikasi tantangan dan hambatan dalam pengembangan model pengukuran *echo chamber* berbasis ASBA dan metrik konsensus GDM, dilanjutkan dengan analisis solusi yang mengevaluasi metode-metode yang relevan, serta rancangan solusi yang merinci rancangan implementasi sistem.

#### **III.1 Analisis Persoalan**

Tugas akhir ini mengadaptasi pendekatan Analisis Sentimen Berbasis Aspek (ASBA) dan metrik konsensus dalam kerangka *Group Decision Making* (GDM) untuk mengukur *echo chamber* pada percakapan berbahasa Indonesia. Oleh karena itu, dibutuhkan metode ASBA bebas domain, pada subtugas ACSA, yang dapat menghasilkan keluaran berupa distribusi probabilitas sentimen terhadap kategori-kategori aspek tertentu dalam teks berbahasa Indonesia. Untuk mendapatkan metode ASBA bebas domain terbaik, akan dilakukan eksperimen dan pengujian untuk metode *double propagation* dan *zero-shot LLM* generatif. Setelah itu, akan dilakukan pengujian pada *pipeline* penuh pengukuran *echo chamber* yang terdiri dari ASBA, deteksi komunitas, dan perhitungan konsensus. Dalam proses ini terdapat beberapa persoalan yang perlu diselesaikan.

Persoalan pertama dalam melakukan eksperimen dan pengujian ASBA serta pengujian pengukuran *echo chamber* adalah belum adanya *dataset* yang dapat digunakan. Untuk eksperimen dan pengujian ASBA, dibutuhkan *dataset* dengan label ASBA yang dapat menggambarkan performa metode ketika digunakan untuk mengukur *echo chamber* pada percakapan berbahasa Indonesia di *platform X (tweet)* dalam berbagai domain. Sementara itu, pengujian pengukuran *echo chamber* memerlukan *dataset tweet* berbahasa Indonesia dengan label interaksi yang teridentifikasi sebagai *echo chamber*. Selain itu, *negative control dataset* diperlukan untuk memastikan bahwa model tidak mendekripsi *echo chamber* pada data yang seharusnya bukan *echo chamber*. Namun, hingga saat ini belum ada himpunan data yang secara eksplisit dilabeli sebagai “*echo chamber*” maupun

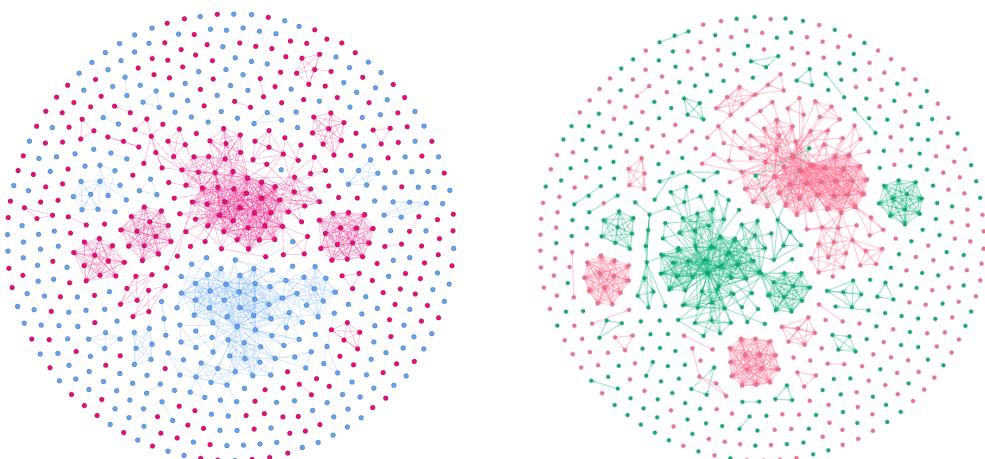
“bukan *echo chamber*,” sehingga evaluasi objektif terhadap model menjadi sulit dilakukan. Anotasi manual untuk melabeli *dataset* sebagai “*echo chamber*” maupun “bukan *echo chamber*” juga sulit dilakukan karena keterbatasan akses terhadap pakar. Persoalan ini juga menjadi persoalan evaluasi model pengukuran *echo chamber* yang umum ditemukan pada studi-studi terdahulu.

Persoalan kedua adalah penyesuaian metode ASBA berbasis *double propagation* dan *zero-shot LLM* generatif. Pada *double propagation* (Qiu dkk., 2011), tantangan utama terletak pada ketergantungan terhadap *dependency parser* dan *POS tagger*, yang kinerjanya rentan menurun pada bahasa Indonesia informal dengan struktur tidak baku, singkatan, *code-mixing*, dan salah ketik. Selain itu, metode ini hanya menghasilkan daftar ekspresi aspek dengan label sentimen biner, padahal subtugas ACSA pada tugas akhir ini membutuhkan probabilitas tiga kelas (positif, netral, negatif) per kategori aspek. Dengan demikian, diperlukan penyesuaian berupa pemetaan ekspresi aspek ke kategori dan modifikasi kalkulasi sentimen.

Sementara itu, pada *zero-shot LLM* generatif, model yang dipilih harus berperforma baik pada subtugas ACSA. Model juga harus dapat diakses secara gratis, mengingat keterbatasan sumber daya pada tugas akhir ini. Lebih lanjut, persoalan utama *zero-shot LLM* generatif adalah inkonsistensi hasil. Model yang sama dapat memberi keluaran berbeda meski diberi *prompt* identik, baik dalam penentuan kategori maupun probabilitas sentimen. Format keluaran pun sering tidak sesuai instruksi, misalnya muncul *bullet point* atau parafrasa yang tidak diperlukan. Hal ini menyulitkan replikasi dan evaluasi, karena penelitian membutuhkan hasil stabil untuk perbandingan metode.

Persoalan ketiga berkaitan dengan pemilihan jenis interaksi dalam pembangunan graf serta penyesuaian metode deteksi komunitas agar mampu mengidentifikasi komunitas alami. Setiap *dataset* memiliki karakteristik berbeda, sebagian *dataset* sudah menyertakan label komunitas, sebagian lainnya tidak. Terdapat pula variasi dalam ketersediaan label interaksi. Selain itu, pada tahap deteksi komunitas, Amendola dkk. (2024) menggunakan *multilevel k-way graph partitioning* dengan perangkat lunak METIS dengan  $k = 2$ . Pendekatan ini secara eksplisit memaksa graf interaksi terbagi menjadi dua kelompok, namun hasilnya sering kali gagal

merefleksikan komunitas alami, terutama ketika jumlah interaksi relatif sedikit. Gambar III.1 menampilkan contoh partisi yang dihasilkan METIS dengan  $k = 2$  pada dataset *echo chamber* Pilpres 2024 dan *Budget Efficiency*. Visualisasi dihasilkan menggunakan perangkat lunak Gephi dengan algoritma *layout* Fruchterman Reingold. Dari gambar tersebut dapat dilihat bahwa hasil partisi kurang dapat merefleksikan struktur komunitas yang sebenarnya dalam jaringan interaksi. Banyak simpul dengan derajat nol ditempatkan secara arbitrer. Beberapa simpul yang sama sekali tidak terhubung juga dikelompokkan menjadi satu komunitas. Amendola dkk. (2024) menyoroti keterbatasan ini dan menyarankan agar metrik konsensus dieksplorasi lebih lanjut dengan cara mengintegrasikannya dengan metode deteksi komunitas yang dapat menemukan lebih dari dua komunitas.



*Warna yang berbeda menunjukkan komunitas yang berbeda*

Gambar III.1 Partisi yang Dihasilkan METIS *Multilevel K-way Graph Partitioning* dengan  $k = 2$  pada *Dataset Echo Chamber* Pilpres 2024 (kiri) dan *Budget Efficiency* (kanan).

Persoalan keempat adalah penyesuaian perhitungan metrik konsensus GDM. Amendola dkk. (2024) menggunakan pendekatan tiga tingkat mulai dari konsensus pasangan, *aspect consensus*, dan *community consensus*. Perhitungan konsensus secara rinci dijelaskan pada Subbab II.4. Meskipun pendekatan ini memberikan kerangka kerja yang sistematis, terdapat beberapa permasalahan yang perlu

diperhatikan dalam penerapannya. Pertama, asumsi pembagian paksa menjadi dua komunitas tidak praktis untuk jaringan dengan banyak komunitas karena menghasilkan terlalu banyak nilai konsensus dalam komunitas sehingga hasil sulit disimpulkan. Kedua, perhitungan konsensus antarkomunitas untuk lebih dari dua komunitas dengan menganggap seluruh simpul berada pada satu komunitas yang sama menimbulkan dominasi komunitas besar. Misalnya, dua komunitas beranggotakan 100 simpul akan menghasilkan 10.000 pasangan silang, sehingga nilai konsensus antarkomunitas global hanya akan mencerminkan perilaku komunitas terbesar. Ketiga, pendekatan ini tidak menangani perbedaan jumlah aspek yang dibahas tiap komunitas, jika aspek yang tidak dibahas diberi nilai [0,0,0], semakin banyak aspek tidak dibahas akan menurunkan konsensus secara berlebihan. Padahal tidak semua komunitas perlu membahas setiap aspek. Selain itu, variasi jumlah aspek antardomain topik juga menghambat kemampuan generalisasi jika aspek ditangani dengan cara ini.

### **III.2 Analisis Solusi**

#### **III.2.1 Ketersediaan *Dataset***

Untuk menjawab persoalan ketersediaan *dataset* eksperimen dan pengujian ASBA, dikumpulkan beberapa *dataset* yang memenuhi kriteria: kumpulan *tweet* berbahasa Indonesia dalam berbagai domain dengan topik kontroversial (terdapat sentimen kuat yang bervariasi). Terdapat tiga *dataset* yang berhasil didapatkan, di antaranya *dataset Indo Vaccination*, Pilkada DKI, dan Kampus Merdeka. Ketiga *dataset* diperoleh dari penelitian sebelumnya, yaitu Pratiwi dkk. (2024) terkait topik Kampus Merdeka yang dipublikasikan di Zenodo (<https://zenodo.org/records/14234851>), Rofiqoh dkk. (2017) terkait Pilkada DKI yang tersedia melalui GitHub (<https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia>), serta Purwitasari dkk. (2023) terkait vaksinasi COVID-19 yang dipublikasikan di Mendeley Data (<https://data.mendeley.com/datasets/7ky2bjwtn/3>). Deskripsi lengkap *dataset* mentah dapat dilihat pada Lampiran A. Selanjutnya, dilakukan pembagian (*split*) *dataset* berdasarkan tujuan: *dataset* Kampus Merdeka dan Pilkada DKI dipakai untuk eksperimen konfigurasi (*dataset* validasi), sementara *dataset* *Indo*

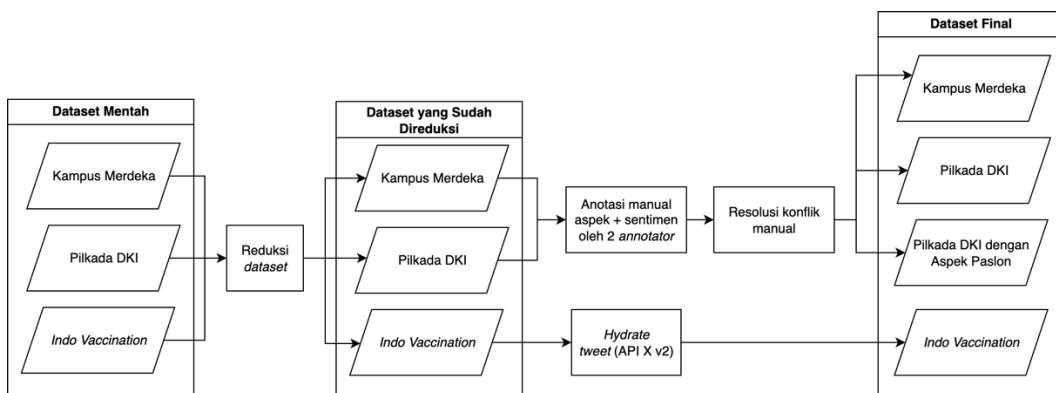
*Vaccination* dan Pilkada DKI (dengan aspek pasangan calon) digunakan sebagai data pengujian akhir (*dataset pengujian*).

Terdapat permasalahan pada *dataset* Kampus Merdeka dan Pilkada DKI yaitu tidak tersedianya label ASBA berupa kategori aspek serta sentimen pada level aspek. Oleh karena itu, dilakukan anotasi manual untuk kedua *dataset* tersebut. Anotasi diawali dengan eksplorasi potensi aspek menggunakan bantuan LLM (Gemini-2.5-Flash) terhadap sampel *tweet*, yang kemudian diseleksi secara manual menjadi daftar aspek dengan deskripsi rinci. Dua orang *annotator* dilibatkan untuk memberi label aspek sekaligus sentimen. Proses anotasi juga mencakup mekanisme resolusi konflik manual dengan beberapa kasus, kasus-kasus dan resolusinya diuraikan pada Tabel III.1. Untuk klasifikasi sentimen, resolusi konflik dilakukan dengan pendekatan serupa, namun tanpa kasus kesalahan parsial. Awalnya, anotasi sentimen dilakukan dengan lima kelas: “Sangat Positif”, “Positif”, “Netral”, “Negatif”, “Sangat Negatif”. Namun, pada tahap evaluasi, label sentimen dikonversi menjadi *hard label* “[1, 0, 0]” untuk positif, “[0, 1, 0]” untuk netral, dan “[0,0,1]” untuk negatif, agar hasil lebih sederhana dan konsisten. Karena keterbatasan sumber daya, anotasi hanya dilakukan pada 300 *tweet*, yang kemudian mengalami penyaringan pada tahap resolusi konflik serta pembuangan teks kosong, sehingga jumlah data akhir yang digunakan menjadi kurang dari 300.

Tabel III.1 Kasus dan Aturan Resolusi Konflik Manual untuk Anotasi Klasifikasi Aspek

No.	Kasus	Resolusi
1.	Salah satu atau kedua <i>annotator</i> melakukan kesalahan parsial pada penentuan kategori aspek. Ditandai dengan kategori aspek yang tidak selaras dengan ekspresi aspek yang ditulis, atau ekspresi aspek tidak ditemukan pada teks.	<i>Annotator</i> diminta melakukan anotasi ulang.
2.	Salah satu <i>annotator</i> keliru sepenuhnya.	Label final diambil dari anotasi <i>annotator</i> lain.
3.	Kedua <i>annotator</i> benar dengan menemukan aspek berbeda.	Kedua aspek dijadikan label final.
4.	Aspek rancu atau tidak jelas.	<i>Tweet</i> dihapus

Selain itu, terdapat permasalahan pada *dataset Indo Vaccination* yang hanya menyimpan *tweet ID*. Untuk *dataset* ini dilakukan *hydrate tweet* dengan X API v2. Secara keseluruhan, diagram alur proses pembangunan *dataset* untuk eksperimen dan pengujian ASBA dapat dilihat pada Gambar III.2. Deskripsi ringkas *dataset* final yang digunakan untuk eksperimen konfigurasi dan pengujian ASBA dapat dilihat pada Tabel III.2. Deskripsi lengkap *dataset* final dapat dilihat pada Lampiran B.



Gambar III.2 Diagram Alur Pembangunan *Dataset* Eksperimen dan Pengujian ASBA

Tabel III.2 Deskripsi Ringkas *Dataset* Final untuk Eksperimen dan Pengujian ASBA

Jenis Dataset	Nama Dataset	Jumlah Tweet	Jumlah Kategori Aspek	Kategori Aspek
Dataset Eksperimen	Kampus Merdeka	297	3	a. Anggaran b. Kebijakan Menteri c. Umum
	Pilkada DKI	264	5	a. Politik Uang b. Pendukung/Pemilih c. SARA d. Kampanye e. Tokoh dan Partai
Dataset Pengujian	Indo Vaccination	437	7	a. Aplikasi c. Implementasi d. Peserta e. Layanan f. Produk Vaksin g. Umum
	Pilkada DKI dengan Aspek Paslon	900	3	a. Agus-Sylvi b. Ahok-Djarot c. Anies-Sandi

Selanjutnya, untuk menjawab persoalan ketersediaan *dataset* pengujian pengukuran *echo chamber*, dikumpulkan beberapa *dataset* yang memenuhi kriteria: kumpulan *tweet* dengan label interaksi pada domain yang telah diteliti dan terdokumentasi terkait fenomena *echo chamber* dalam percakapan berbahasa Indonesia di *platform X*.

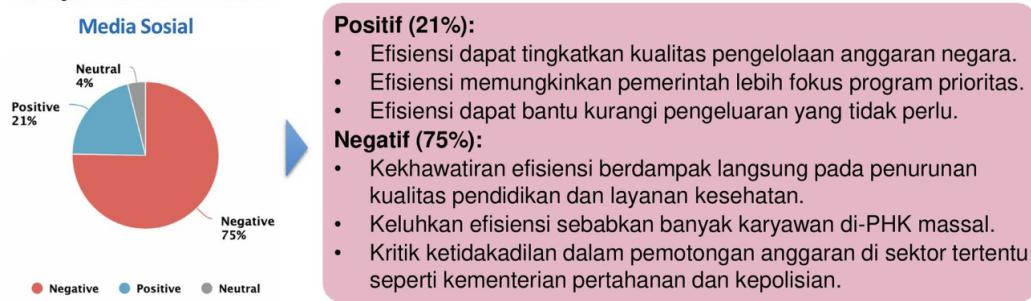
Terdapat tiga *dataset* yang berhasil didapatkan, di antaranya *dataset Indo Vaccination* (sama dengan yang digunakan pada pengujian *echo chamber*), *Budget Efficiency*, dan Pilpres 2024. *Dataset Indo Vaccination* dipilih karena isu vaksin terbukti melahirkan polarisasi yang kuat akibat penguatan kelompok antivaksin di media sosial dalam perspektif *echo chamber* (Sugiono, 2021). *Dataset Budget Efficiency* didapatkan dari sumber terbuka melalui Mendeley Data (<https://data.mendeley.com/datasets/25dhk4dn3r/1>). *Dataset* ini dipilih berdasarkan laporan Fahmi (2025) yang meskipun tidak secara eksplisit menyatakan keberadaan *echo chamber*, hasil analisis SNA (*Social Network Analysis*) dan isi konten menunjukkan karakteristik *echo chamber* berupa tingginya homofili opini dan penguatan internal dalam komunitas (Cinelli dkk., 2021). *Dataset Pilpres 2024* didapatkan dari sumber terbuka yang disediakan oleh Drone Emprit melalui program SATRIA DATA 2024 (<https://pers.droneemprit.id/twitter-x-data-set-for-researchers/>). *Dataset* ini digunakan karena topik ini telah terbukti memiliki bias politik dan polarisasi yang mengindikasikan *echo chamber* dalam percakapan di media sosial selama Pemilihan Presiden 2024 (Manuputty dkk., 2025). Deskripsi lengkap *dataset* mentah dapat dilihat pada Lampiran A.

Terdapat beberapa *tweet* duplikat dan teks kosong pada ketiga *dataset*, sehingga dilakukan penghapusan baris yang bermasalah. Selanjutnya, guna mengurangi ukuran *dataset* karena keterbatasan sumber daya dalam pemrosesan ASBA, pada *dataset Indo Vaccination* hanya dipilih dua komunitas terbesar (tidak termasuk komunitas nol yang berisi sisa pengguna) serta tiga komunitas berukuran sedang yang dipilih secara acak. Sementara itu, pada *dataset Budget Efficiency*, dilakukan deteksi komunitas terlebih dahulu, kemudian dipilih sepuluh komunitas terbesar.

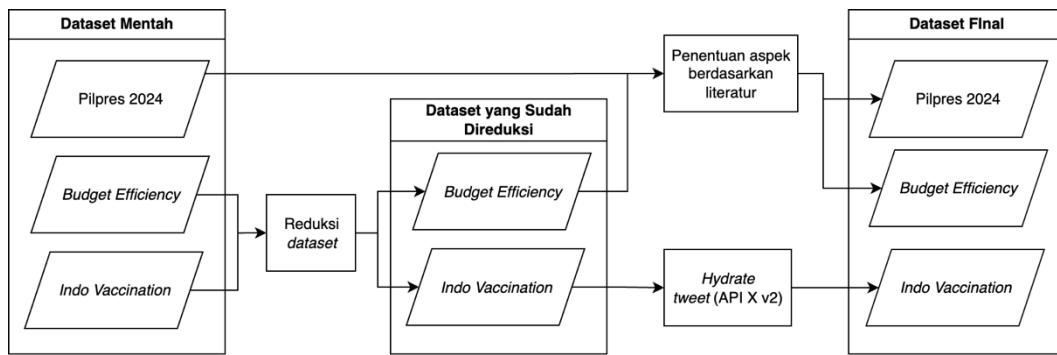
Tabel III.3 Deskripsi Ringkas *Dataset* Final untuk Pengujian Pengukuran *Echo Chamber*

Nama Dataset	Jumlah Tweet	Jumlah Kategori Aspek	Kategori Aspek	Jenis Interaksi
<i>Indo Vaccination</i>	437	7	a. Aplikasi b. Biaya c. Implementasi d. Peserta e. Layanan f. Produk Vaksin g. Umum	<i>Mention</i> dan <i>retweet</i>
<i>Budget Efficiency</i>	745	4	a. Kualitas Pengelolaan b. Dampak c. Prioritas Alokasi d. Umum	<i>Reply</i>
Pilpres 2024	970	3	a. Anies b. Prabowo c. Ganjar	<i>Reply</i>

Lebih lanjut, karena *dataset Budget Efficiency* dan Pilpres 2024 belum memiliki label aspek, dilakukan pemilihan kategori aspek. Untuk *dataset Budget Efficiency*, pemilihan dilakukan dengan merujuk pada analisis Fahmi (2025) yang membahas sentimen warganet di media sosial terkait isu efisiensi anggaran. Hasil analisis ini dapat dilihat pada Gambar III.3. Berdasarkan analisis tersebut, aspek yang digunakan di antaranya, Kualitas Pengelolaan, Dampak, Prioritas Alokasi, serta Umum. Selain itu, untuk *dataset Pilpres 2024*, kategori aspek yang dipilih adalah ketiga pasangan calon (paslon), yaitu Anies, Prabowo, dan Ganjar, merujuk pada analisis Manuputty dkk. (2025) yang menunjukkan bahwa percakapan warganet di media sosial cenderung membentuk *echo chamber* seputar pasangan calon yang didukung. Secara keseluruhan, diagram alur proses pembangunan *dataset* untuk pengujian pengukuran *echo chamber* dapat dilihat pada Gambar III.4. Secara ringkas, deskripsi *dataset* final yang digunakan untuk pengujian pengukuran *echo chamber* dapat dilihat pada Tabel III.3. Deskripsi *dataset* final secara lengkap dapat dilihat pada Lampiran B.



Gambar III.3 Hasil Analisis Sentimen Warganet Media Sosial Terkait Isu Efisiensi Anggaran  
(Fahmi, 2025)



Gambar III.4 Diagram Alur Pembangunan *Dataset* Pengujian Pengukuran *Echo Chamber*

### III.2.2 Kualitas Teks pada *Dataset*

Lebih lanjut, terdapat berbagai isu kualitas teks pada *tweet* dari *dataset*, baik *dataset* untuk eksperimen dan pengujian ASBA maupun *dataset* untuk pengujian pengukuran *echo chamber*. Sebagaimana teks informal pada umumnya, ditemukan masalah berupa penggunaan huruf besar dan kecil yang tidak sesuai kaidah, penggunaan bahasa gaul (*slang*), singkatan serta ejaan yang salah, hingga campuran bahasa asing (*code-mixing*). Di sisi lain, dari fitur khusus *platform X*, banyak *tweet* memuat *mention* terhadap akun pengguna lain, *hashtag*, tautan eksternal, serta penggunaan emoji. Contoh *tweet* dengan isu-isu ini dapat dilihat pada Lampiran C. Isu-isu ini menunjukkan bahwa teks dalam *dataset* cenderung *noisy* dan tidak terstandarisasi, sehingga perlu dilakukan tahapan praproses.

Untuk mengatasi persoalan ini dirancang tahapan praproses dalam tiga tingkat. Pertama, tingkat pembersihan (*cleaning*) minimal yang mencakup penghapusan

*whitespace* berlebih dan koma. Tahap ini bertujuan untuk mempertahankan teks mentah (*raw text*), penghapusan koma dan *whitespace* berlebih hanya untuk menghindari kesalahan format, karena data disimpan dalam format CSV, dan meningkatkan keterbacaan. Kedua, tingkat pembersihan penuh yang terdiri dari *case folding* untuk mengurangi ukuran kamus dan meningkatkan konsistensi. Dilakukan juga penghapusan elemen spesifik X seperti *link* (tautan), *hashtag*, *mention*, "RT" (untuk *retweet*), serta karakter non-alfanumerik (termasuk emoji dan tanda baca). Penghapusan elemen-elemen tersebut dilakukan untuk mengurangi *noise* yang dapat mempengaruhi kinerja analisis sentimen secara negatif. Pengurangan *noise* ini juga terbukti dapat meningkatkan performa *classifier* (Haddi dkk., 2013).

Tabel III.4 Contoh Hasil Praproses untuk Setiap Tingkat

Tingkat Praproses	Teks
Sebelum Praproses	"Jadi yah tweps.. udah tahu kan sekarang kalau semua vaksin baik, so.. jangan pilih-pilih, segera vaksin. Menunda vaksin Covid-19 karena memilih merek tertentu memperlambat percepatan target vaksinasi  Yuk bangun antibodi #VaksinKokPilihPilih"
Pembersihan minimal	"Jadi yah tweps.. udah tahu kan sekarang kalau semua vaksin baik so.. jangan pilih-pilih segera vaksin. Menunda vaksin Covid-19 karena memilih merek tertentu memperlambat percepatan target vaksinasi Yuk bangun antibodi #VaksinKokPilihPilih"
Pembersihan penuh	"jadi yah tweps udah tahu kan sekarang kalau semua vaksin baik so jangan pilih pilih segera vaksin menunda vaksin covid19 karena memilih merek tertentu memperlambat percepatan target vaksinasi yuk bangun antibodi"
Pembersihan penuh + normalisasi	"tweps sudah kan vaksin baik sok jangan pilih pilih vaksin menunda vaksin covid19 memilih merek memperlambat percepatan target vaksinasi yuk bangun antibodi"

Terakhir adalah tingkat normalisasi yang meliputi penggantian kata *slang* dengan bentuk baku menggunakan leksikon dari TheOnlyDo (2023) dan penelitian Ibrohim dan Budi (2018), serta penghapusan *stopwords* menggunakan daftar dari Tala (2003). Khusus untuk *stopwords*, dilakukan penyaringan terhadap kata-kata yang

penting dalam analisis sentimen, yaitu: negasi, konjungsi, kata penguat, dan kata kerja modal. Kata-kata tersebut dipertahankan karena memiliki peran penting dalam menentukan polaritas sentimen, terutama kata negasi yang dapat membalik makna kalimat dan kata penguat yang memperkuat atau memperlemah sentimen. Selain itu, konjungsi dan kata kerja modal juga dipertahankan agar *dependency parser* pada metode *double propagation* tidak terpengaruhi oleh penghapusan kata-kata yang secara sintaktis berperan dalam hubungan antarkata. Tahapan normalisasi ini bertujuan untuk meningkatkan kualitas representasi teks (Singh dan Kumari, 2016).

Eksperimen akan dilakukan pada ketiga tingkatan praproses untuk menentukan tingkat praproses yang menghasilkan performa terbaik pada kedua metode ASBA. Sebagai catatan, penanganan *code-mixing* tidak dilakukan karena membutuhkan model yang lebih kompleks, terlebih bahasa yang muncul tidak hanya bahasa Inggris dan bahasa Indonesia, tetapi juga mencakup bahasa Arab dan bahasa daerah. praproses dalam tugas akhir ini dibatasi pada praproses berbasis aturan saja. Contoh hasil praproses pada setiap tingkat dapat dilihat pada Tabel III.4.

### III.2.3 Metode ASBA Bebas Domain

Untuk mengatasi persoalan dalam penerapan metode ASBA *double propagation* pada bahasa Indonesia informal, dilakukan beberapa penyesuaian. Pertama, untuk mengurangi potensi kesalahan pada tahap *dependency parsing* dan POS *tagging*, digunakan dua *processor* berbeda, yaitu UDPipe Indonesian-CSUI-UD-2.12-230717 dan Stanza-ID. Pemilihan ini didasarkan pada pertimbangan bahwa UDPipe dilatih dengan *treebank* Indonesian-CSUI yang cukup representatif untuk struktur bahasa Indonesia, sedangkan Stanza-ID menawarkan keunggulan melalui *neural pipeline* modern yang relatif lebih adaptif terhadap variasi bahasa alami. Dengan membandingkan keluaran keduanya dalam eksperimen, diharapkan dapat diperoleh hasil *parsing* yang lebih stabil pada teks media sosial. Selain itu, untuk leksikon kata opini (ekspresi sentimen) digunakan InSet Lexicon (Koto dkk., 2017) yang dibangun dengan kata-kata dari *tweet*.

Tabel III.5 Desain *system prompt* dan *user prompt*

Jenis <i>Prompt</i>	Isi <i>Prompt</i>
<i>System Prompt</i>	<p>Anda adalah model ABSA (Aspect-Based Sentiment Analysis) yang ahli dalam menganalisis sentimen berbasis aspek pada teks berbahasa Indonesia.</p> <p><b>ASPEK YANG TERSEDIA:</b></p> <ul style="list-style-type: none"> <li>- "category": description (sesuai daftar aspects)</li> <li>- ...</li> </ul> <p><b>TUGAS ANDA:</b></p> <ol style="list-style-type: none"> <li>1. Untuk setiap tweet yang diberikan, identifikasi aspek mana yang disebutkan atau tersirat</li> <li>2. Hanya sertakan aspek yang benar-benar relevan dengan tweet</li> <li>3. Untuk setiap aspek yang relevan, analisis sentimennya (positif, netral, negatif)</li> </ol> <p><b>FORMAT OUTPUT YANG WAJIB:</b></p> <p><b>Tweet:</b> &lt;teks_tweet_persis_sama&gt;</p> <p><b>Aspect:</b> &lt;nama_aspek_dari_daftar&gt;</p> <p><b>Sentiment:</b> [p_pos, p_neu, p_neg]</p> <p><b>ATURAN PENTING:</b></p> <ul style="list-style-type: none"> <li>- Probabilitas sentimen harus berjumlah 1.0 (contoh: [0.70, 0.20, 0.10])</li> <li>- Gunakan nama aspek yang PERSIS SAMA dengan daftar yang diberikan</li> <li>- Setiap tweet memiliki setidaknya satu aspek</li> <li>- PASTIKAN format output sama persis seperti yang ditentukan</li> </ul>
<i>User Prompt</i>	<p>Analisis tweet-tweet berikut untuk aspek dan sentimen yang relevan:</p> <p>"tweet1"</p> <p>"tweet2"</p> <p>"..."</p> <p>Berikan analisis ABSA untuk setiap tweet yang relevan dengan format yang telah ditentukan.</p>

Selanjutnya, untuk menyesuaikan keluaran ekspresi aspek dari *double propagation* dengan kategori aspek, dilakukan proses pemetaan berbasis kesamaan semantik menggunakan *cosine similarity* dari representasi vektor. Beberapa pendekatan vektorisasi yang diuji meliputi Sentence-BERT (model *pretrained* firqaaa/indo-sentence-bert-base), FastText (model *pretrained* cc.id.300.vec.gz), dan TF-IDF. Sentence-BERT dipilih karena kemampuannya menangkap makna kontekstual kalimat secara lebih baik melalui arsitektur transformer yang telah *fine-tuned* untuk bahasa Indonesia. FastText bahasa Indonesia dengan dimensi vektor 300 dipertimbangkan karena keunggulannya dalam menangani morfologi kompleks dan kata tidak baku yang sering muncul pada teks informal melalui representasi *subword*, serta telah dilatih pada korpus *Common Crawl* bahasa Indonesia yang mencakup ragam teks web. Sedangkan TF-IDF digunakan sebagai *baseline* berbasis frekuensi sederhana.

Sementara itu, untuk *zero-shot* LLM generatif, dipilih model Gemini-2.0-Flash dan Gemini-2.5-Flash. Pemilihan Gemini didasarkan pada temuan Mughal dkk. (2024) yang menunjukkan keunggulan PaLM pada tugas ASBA. Namun, mengingat PaLM kini sudah *deprecated* dan digantikan oleh Gemini, maka penggunaan Gemini menjadi pilihan yang logis. Gemini-2.5-Flash merupakan model *cost-efficient* terkini dari Gemini API dengan kemampuan *adaptive thinking*, sedangkan Gemini-2.0-Flash adalah model *non-thinking* yang dioptimalkan untuk kecepatan. Keduanya mendukung masukan multimodal (teks, audio, gambar, video) dengan keluaran berupa teks. Model Gemini-2.5-Pro tidak dipilih karena keterbatasan sumber daya pada *tier* gratis, meskipun model ini memiliki performa tertinggi dalam akurasi dan memiliki kemampuan *enhanced thinking*. Alternatif lain, yaitu GPT yang juga dilaporkan Mughal dkk. (2024), tidak dipertimbangkan karena keterbatasan akses berbayar.

Selanjutnya, untuk mengatasi persoalan inkonsistensi keluaran pada penerapan metode ASBA *zero-shot* dengan model Gemini, diterapkan beberapa solusi. Pertama, digunakan *system prompt* dengan struktur format yang jelas dan eksplisit, termasuk penekanan agar model tidak mengubah struktur keluaran. Desain *prompt* ini diadaptasi dari rancangan yang digunakan Mughal dkk. (2024). Desain *system prompt* dan *user prompt* dapat dilihat pada Tabel III.5. Kedua, parameter

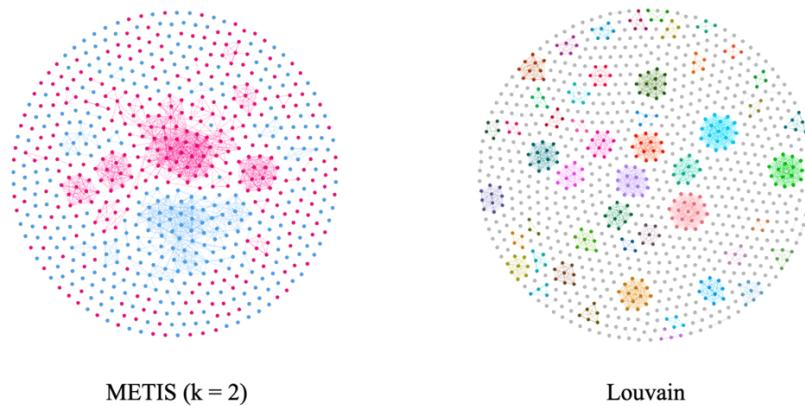
*temperature* model diatur ke nilai 0 untuk membuat keluarannya lebih konsisten dan tidak acak. Ketiga, dalam tahap *parsing* keluaran, diterapkan teknik *fuzzy matching* dengan SequenceMatcher menggunakan ambang 0,8, sehingga keluaran yang memiliki sedikit variasi penulisan tetap dapat dikenali dengan benar. Terakhir, parameter *temperature* model diatur ke nilai 0 untuk membuat keluarannya lebih konsisten dan tidak acak.

### **III.2.4 Pembangunan Graf Interaksi dan Metode Deteksi Komunitas**

Untuk mengatasi masalah pembangunan graf interaksi, untuk *dataset Pilpres 2024* dan *Budget Efficiency*, diterapkan empat aturan koneksi interaksi antar-tweet: (1) menghubungkan *tweet* yang membalas *status\_id* (*tweet id*) yang sama, (2) menghubungkan *tweet* yang membalas *screen\_name* (*username*) yang sama, (3) menghubungkan balasan terhadap *screen\_name* tertentu dengan *tweet* dari pengguna bersangkutan, dan (4) menghubungkan balasan terhadap *status\_id* dengan *tweet* pemilik *id* tersebut. Sementara itu, pada *dataset Indo Vaccination* yang sudah memiliki label komunitas, interaksi yang digunakan adalah *retweet* dan *mention*, sesuai dengan proses deteksi komunitas yang dilakukan pada *dataset* tersebut.

Selanjutnya, untuk mengatasi keterbatasan pendekatan *multilevel k-way graph partitioning* dengan METIS yang memaksa partisi ke dalam dua kelompok, penelitian ini menggunakan metode Louvain (Blondel dkk., 2008). Berbeda dengan METIS, Louvain tidak menetapkan jumlah komunitas sejak awal, melainkan berangkat dari asumsi bahwa jaringan secara alami terbagi ke dalam beberapa subkelompok, dan struktur komunitas ditentukan oleh optimisasi modularitas (Newman, 2006). Dengan demikian, jumlah dan ukuran komunitas dapat muncul secara organik sesuai dengan pola interaksi yang ada. Pendekatan ini lebih sesuai untuk kasus percakapan media sosial, di mana distribusi komunitas tidak selalu biner dan sering kali bersifat multikluster. Implementasi Louvain pada *dataset Pilpres 2024* dengan parameter *default* (resolusi = 1) menghasilkan pembagian komunitas yang lebih alami, sebagaimana ditunjukkan pada Gambar III.5. Makna dari komunitas yang terbentuk dari algoritma Louvain adalah bahwa anggota dalam satu komunitas saling berinteraksi secara intens, sementara interaksi dengan

anggota komunitas lain relatif sedikit. Hasil ini memperlihatkan bahwa Louvain mampu merefleksikan struktur interaksi yang lebih realistik dalam jaringan sosial dibandingkan METIS.



*Warna yang berbeda menunjukkan komunitas yang berbeda*

Gambar III.5 Perbandingan Hasil Deteksi Komunitas dengan METIS ( $k = 2$ ) dan Louvain pada Dataset Pilpres 2024

### III.2.5 Perhitungan Metrik Konsensus GDM

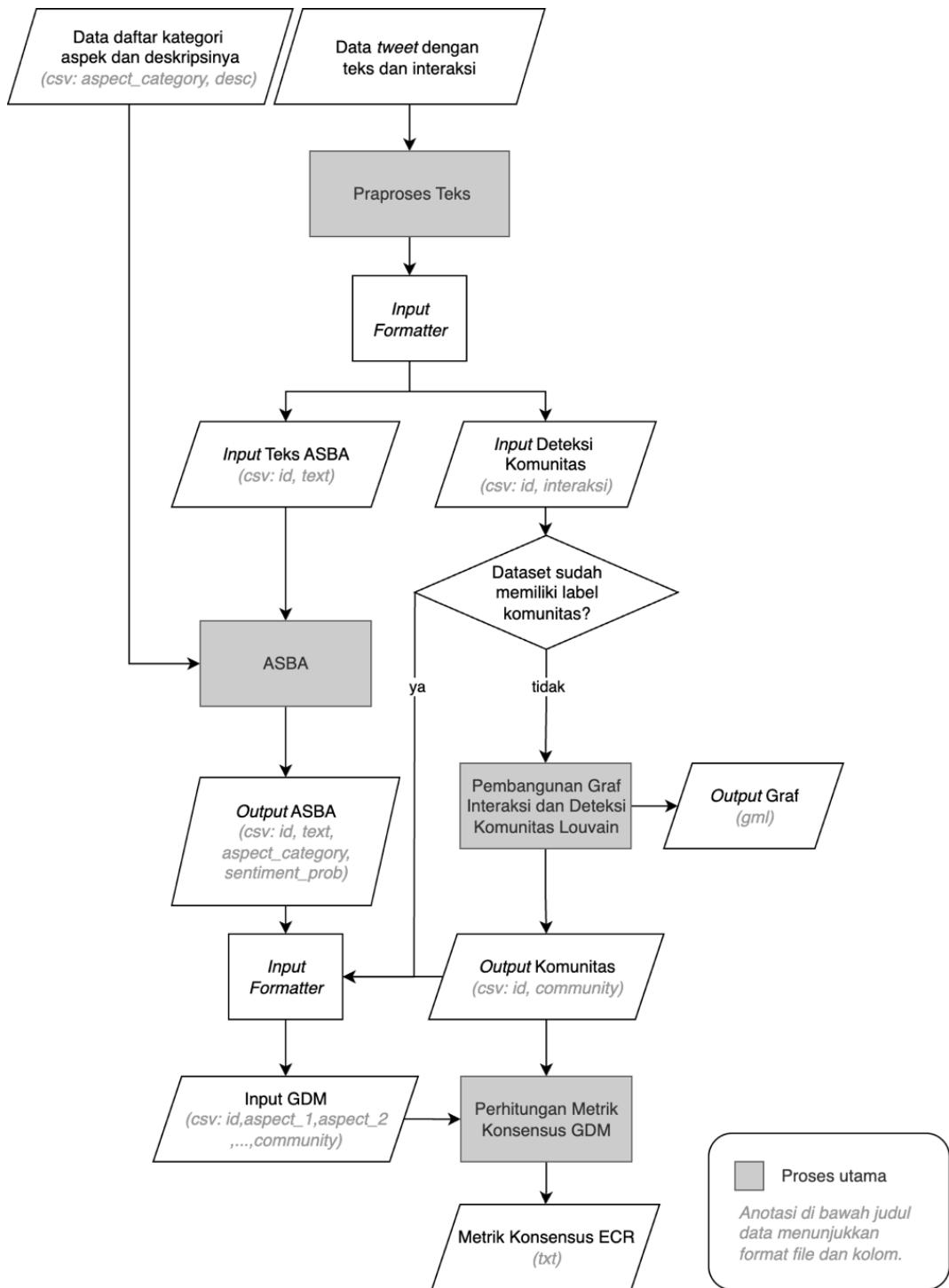
Untuk mengatasi permasalahan perhitungan konsensus pada penelitian Amendola dkk. (2024), diusulkan beberapa solusi. Pertama, konsensus dalam komunitas global dihitung menggunakan rata-rata (*arithmetic mean*) dari seluruh komunitas. Perlu dicatat bahwa pendekatan ini memberikan bobot yang sama antara komunitas kecil dan komunitas besar. Walaupun secara intuitif komunitas kecil cenderung lebih mudah mencapai konsensus, hal ini tidak menjadi masalah karena dalam konteks *echo chamber* yang menjadi fokus adalah tingkat polarisasi suatu komunitas terlepas dari ukurannya. Namun demikian, perlu diperhatikan bahwa komunitas dengan jumlah anggota yang terlalu sedikit ( $<5$ ) umumnya dianggap tidak stabil dan kurang merepresentasikan struktur sosial yang bermakna. Beberapa studi menunjukkan bahwa komunitas kecil sering kali bersifat *under-developed* atau artifisial (Wagenseller dan Wang, 2017; Zhang dkk., 2016). Oleh karena itu, pada tugas akhir ini, akan diterapkan penyaringan untuk komunitas dengan ukuran

lebih kecil dari lima. Dengan demikian, perbedaan ukuran komunitas tidak menjadi persoalan selama ukuran komunitas lebih dari lima anggota, karena di titik ini komunitas sudah dapat dianggap cukup stabil untuk dianalisis tingkat konsensus dan polarisasinya.

Kedua, konsensus antarkomunitas global dihitung secara berpasangan antarpasangan komunitas, kemudian hasilnya dirata-ratakan. Dengan cara ini, perhitungan tetap selaras dengan kerangka penelitian Amendola dkk. (2024) yang menghitung konsensus antara dua komunitas, sekaligus menghindari dominasi komunitas berukuran besar dalam hasil perhitungan konsensus global. Ketiga, dalam perhitungan konsensus aspek, pasangan pengguna dengan nilai sentimen [0, 0, 0] (tidak membahas aspek tertentu) dikeluarkan dari perhitungan. Dengan demikian, aspek yang sama sekali tidak pernah dibahas tidak akan menurunkan nilai konsensus, sehingga konsensus suatu komunitas hanya mencerminkan aspek-aspek yang benar-benar relevan dengan diskusi di komunitas tersebut. Selanjutnya, metrik akhir pengukuran *echo chamber* diturunkan dalam bentuk *Echo Chamber Ratio* (ECR), yaitu rasio antara konsensus dalam komunitas global dengan konsensus antarkomunitas global. Interpretasi dari nilai ini adalah: jika  $ECR < 1$ , maka tidak terjadi *echo chamber* karena konsensus antarkomunitas lebih tinggi daripada konsensus dalam komunitas; jika  $ECR > 1$ , maka konsensus dalam komunitas lebih dominan dibandingkan konsensus antarkomunitas, yang menunjukkan adanya polarisasi internal yang lebih kuat daripada kesepakatan lintas komunitas. Semakin tinggi nilai ECR, semakin kuat pula tingkat *echo chamber* dalam jaringan yang diamati.

### **III.3 Rancangan Solusi**

Gambar III.6 menunjukkan rancangan solusi untuk tugas akhir ini. Proses ini terdiri atas lima tahapan utama: praproses teks, ASBA, pembangunan graf interaksi, deteksi komunitas, dan konsensus GDM.



Gambar III.6 Rancangan Solusi

### III.3.1 Praproses Teks

Masukan untuk proses ini berupa *dataset tweet* mentah dari X. Teknik praproses yang diterapkan berdasarkan konfigurasi terbaik dari eksperimen yang dijelaskan pada Bab IV. Tingkatannya meliputi:

1. Pembersihan minimal: penghapusan *whitespace* berlebih dan koma.
2. Pembersihan penuh: *case folding*, penghapusan tautan, hashtag, mention, simbol “RT”, dan karakter non-alfanumerik.
3. Normalisasi: penggantian kata *slang* dengan bentuk baku menggunakan leksikon dari TheOnlyDo (2023) dan penelitian Ibrohim dan Budi (2018) serta penghapusan *stopwords* daftar dari Tala (2003).

Keluaran dari proses ini berupa teks yang siap dijadikan masukan ASBA, dengan format bersih dan konsisten.

### III.3.2 ASBA Bebas Domain

Masukan untuk proses ini adalah teks yang telah dipraproses dan daftar kategori aspek dengan deskripsi. ASBA dilakukan menggunakan model dengan konfigurasi terbaik dari eksperimen yang dijelaskan pada Bab IV, yang mencakup dua pendekatan:

1. Metode *Double Propagation*:
  - a. *Processor (Dependency Parser* dan *Pos Tagging*): UDPipe Indonesian-CSUI-UD-2.12-230717 dan Stanza-ID
  - b. Representasi vektor: Sentence-BERT, FastText, TF-IDF
2. Metode *zero-shot LLM* generatif:
  - a. Model: Gemini-2.0-flash dan Gemini-2.5-flash

Tabel III.6 Contoh Baris Keluaran ASBA untuk *Dataset Pilpres 2024*

id	text	aspect_category	sentiment_prob
1742969			
1508705	“pidato anies baswedan jarang diviralkan benar pidato manusia waras sosok negarawan berfihak rakyat mozza”	anies	“[1.0, 0.0, 0.0]”
40296			

Model terbaik digunakan untuk mengekstraksi pasangan aspek-sentimen dari setiap teks. Keluaran dari proses ini berupa daftar pasangan aspek-sentimen per *tweet*. Contoh keluaran dapat dilihat pada Tabel III.6.

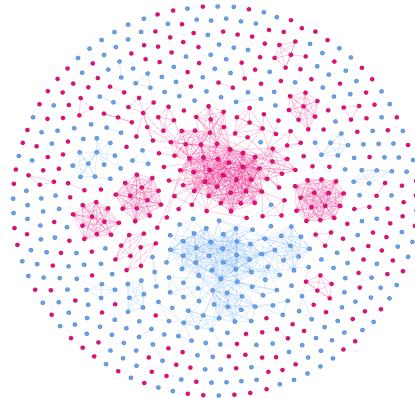
### III.3.3 Pembangunan Graf Interaksi dan Deteksi Komunitas

Masukan dari proses ini adalah *dataset tweet* dengan teks dan interaksi. Sebelum memulai proses akan dilakukan pengecekan apakah *dataset* sudah memiliki label komunitas. Jika sudah, maka proses ini dilewatkan. Proses pembangunan graf dilakukan berdasarkan aturan:

1. Menghubungkan *tweet* yang membahas “status\_id” yang sama.
2. Menghubungkan *tweet* yang membahas “screen\_name” yang sama.
3. Menghubungkan *tweet* yang membahas “screen\_name” tertentu dengan *tweet* dari pengguna dengan “name” tersebut.
4. Menghubungkan *tweet* yang membahas “status\_id” dengan *tweet* yang memiliki “id” sesuai “status\_id”.

Tabel III.7 Contoh Baris Keluaran Deteksi Komunitas untuk *Dataset Pilpres 2024*

id	text	in_reply_to_screen_name	in_reply_to_status_id	Community (nama komunitas)
1742969150870540296	pidato anies baswedan jarang diviralkan benar pidato manusia waras sosok negarawan berfihak rakyat mozza	21_mozza	1742722427296395520	66



*Warna yang berbeda menunjukkan komunitas yang berbeda.*

Gambar III.7 Contoh Keluaran Graf Interaksi untuk *Dataset Pilpres 2024*

Simpul merepresentasikan tiap *tweet*. Dengan catatan, setelah deteksi komunitas, *tweet* dari pengguna yang berada di komunitas yang sama akan di-*merge* menjadi satu simpul. Sisi menunjukkan adanya interaksi antar-*tweet*. Setelah membangun graf interaksi proses dilanjutkan ke deteksi komunitas menggunakan metode Louvain dengan parameter default (resolusi = 1). Keluaran dari proses ini adalah graf interaksi (*file .gml*) dan label komunitas tiap simpul/*tweet*. Contoh keluaran graf dapat dilihat pada Gambar III.7. Contoh keluaran label komunitas dapat dilihat pada Tabel III.7. Jumlah komunitas yang dihasilkan pada *dataset Budget Efficiency* dan Pilpres 2024 dapat dilihat pada Tabel III.8.

Tabel III.8 Jumlah Komunitas yang Dihasilkan

<i>Dataset</i>	Jumlah Komunitas
<i>Budget Efficiency</i>	10
Pilpres 2024	564

### III.3.4 Perhitungan Metrik Konsensus GDM

Masukan untuk proses ini berupa hasil gabungan keluaran ASBA dan deteksi komunitas. Tahap perhitungan konsensus dilakukan dengan pendekatan GDM.

Komunitas dengan ukuran (jumlah anggota)  $< 5$  tidak diperhitungkan dalam konsensus. Konsensus dihitung melalui beberapa tingkatan berikut:

1. Konsensus pasangan:

Konsensus dihitung untuk setiap pasangan pengguna berdasarkan kesamaan vektor sentimen aspek menggunakan *cosine similarity*. Pasangan pengguna dengan vektor sentimen “[0, 0, 0]” (tidak membahas aspek tertentu) dikeluarkan dari perhitungan.

2. *Aspect consensus*:

Nilai konsensus pasangan kemudian diagregasikan pada level aspek menggunakan *power average mean* dengan  $p = 1$  (*arithmetic mean*). Hasilnya menggambarkan tingkat kesepakatan dalam komunitas terhadap masing-masing aspek.

3. *Community consensus*:

Nilai *aspect consensus* untuk seluruh aspek dalam satu komunitas diagregasikan kembali menggunakan *power average mean* dengan  $p = 1$  (*arithmetic mean*), menghasilkan satu nilai tunggal yang mencerminkan kesepakatan komunitas terhadap suatu topik secara keseluruhan.

4. Konsensus global dan antarkomunitas:

- a. Konsensus dalam komunitas global dihitung dengan rata-rata (*arithmetic mean*) dari seluruh komunitas, tanpa membedakan ukuran komunitas.
- b. Konsensus antarkomunitas global dihitung secara *pairwise* untuk setiap pasangan komunitas, lalu dirata-ratakan (*arithmetic mean*). Perhitungan konsensus hanya dilakukan jika jumlah komunitas minimal tiga.

Keluaran dari proses ini adalah metrik ECR (*Echo Chamber Ratio*), yaitu rasio konsensus dalam komunitas global terhadap konsensus antarkomunitas global. Interpretasi hasil ECR adalah jika  $ECR < 1$ , tidak terdeteksi *echo chamber*.

Sedangkan, jika ECR > 1, terdeteksi *echo chamber*, semakin tinggi ECR semakin kuat *echo chamber*.

Berikut adalah ilustrasi perhitungan untuk *dataset Budget Efficiency*. Pada *dataset* ini untuk komunitas 0 dengan ukuran 114 anggota, pertama dihitung konsensus pasangan menggunakan *cosine similarity* antar vektor sentimen. Sebagai contoh, pasangan pengguna dengan *id tweet* 1889965433048141891 dengan vektor sentimen [0,9, 0,05, 0,05] dan pengguna dengan *id tweet* 1889966195480404154 dengan vektor [0,9, 0,1, 0,0] pada kategori aspek “umum” dihitung dengan rumus *cosine similarity* seperti pada persamaan (3.1).

$$\begin{aligned} \text{similarity} &= \frac{(0,9 \times 0,9) + (0,05 \times 0,1) + (0,05 \times 0)}{\sqrt{0,9^2 + 0,05^2 + 0,05^2} \times \sqrt{0,9^2 + 0,1^2 + 0^2}} = \\ &= \frac{0,81 + 0,005 + 0}{\sqrt{0,815} \times \sqrt{0,82}} = \frac{0,815}{0,9028 \times 0,9055} \approx 0,9976 \end{aligned} \quad (3.1)$$

Hasil ini menunjukkan tingkat kesepakatan yang sangat tinggi pada kategori aspek “umum” untuk pasangan tersebut. Perhitungan serupa dilakukan untuk seluruh pasangan dalam komunitas (dengan jumlah anggota  $\geq 5$ ), kemudian hasilnya diagregasikan menggunakan *arithmetic mean* untuk setiap kategori aspek. Misalnya, untuk kategori aspek “umum” diperoleh nilai konsensus 0,1241. Selanjutnya, seluruh nilai konsensus aspek dirata-ratakan kembali sehingga diperoleh konsensus komunitas sebesar 0,1616 untuk komunitas 0.

Langkah yang sama dilakukan untuk seluruh komunitas (0–9), kemudian nilai konsensus komunitas dirata-ratakan sehingga diperoleh konsensus dalam komunitas global sebesar 0,2200. Setelah itu, konsensus antarkomunitas global dihitung dengan cara serupa, yaitu pertama-tama membentuk pasangan komunitas. Sebagai contoh, komunitas 0 dan 1 digabung, lalu konsensus antarpengguna, konsensus aspek, dan konsensus komunitas dihitung kembali menggunakan rumus yang sama, menghasilkan konsensus komunitas sebesar 0,0087. Proses ini dilakukan untuk seluruh pasangan komunitas, lalu hasilnya dirata-ratakan sehingga diperoleh konsensus antarkomunitas global sebesar 0,0484.

Terakhir, nilai ECR dihitung dengan membandingkan konsensus dalam komunitas global terhadap konsensus antarkomunitas global pada persamaan (3.2).

$$ECR = \frac{0,2200}{0,0484} = 4,5497 \quad (3.2)$$

### III.4 Implementasi Solusi

#### III.4.1 Praproses Teks

Bagian ini menjelaskan implementasi praproses teks. Tingkatan praproses teks dapat dilihat pada Subbab III.3.1. Implementasi tiap tingkat praproses dapat dilihat pada Gambar III.8.

```
# Pembersihan Minimal
for each row in dataset:
    text ← row[text_column]
    if text is null: continue
    text ← replace ',' with ' ' → normalize whitespaces → join
    splitlines
    row[text_column] ← text
# Pembersihan Penuh
for each row in dataset:
    text ← row[text_column]
    if text is null: continue
    text ← lowercase(text)
    text ← remove RT at start, urls, mentions, hashtags, non-
    alphanumeric chars
    text ← normalize whitespaces → join splitlines
    row[text_column] ← text
# Normalisasi
load slang maps + stopwords
combined_map ← merge(slang maps)
for each row in dataset:
    tokens ← split(row[text_column] or "")
    tokens ← replace using combined_map → remove stopwords
    row[text_column] ← join(tokens, " ")
```

Gambar III.8 *Pseudocode* Praproses Teks

#### III.4.2 Metode ASBA Bebas Domain: *Double Propagation*

Bagian ini menjelaskan implementasi *double propagation*. Digunakan dua pilihan *dependency parser*, yaitu Stanza (lokal) atau UDPipe API (*server*). Untuk UDPipe, teks dikirim ke API lalu hasil *dependency parser* dikembalikan dalam format CoNLL-U yang kemudian di-parsing menjadi struktur dependensi. Bagian utama

sistem adalah iterasi *double propagation*. Aturan yang digunakan adalah (1) jika sebuah *opinion word* diketahui, kata benda (NOUN/PROPN) yang menjadi dependennya ditandai sebagai target; (2) jika sebuah *target word* diketahui, kata sifat/keterangan (ADJ/ADV) yang terkait menjadi *opinion*; (3) Jika ada hubungan “conj” (konjungsi), kata baru dapat diwariskan sebagai target/*opinion* tambahan. Implementasi iterasi *double propagation* dapat dilihat dalam bentuk *pseudocode* pada Gambar III.9.

```

known_opinions ← seed lexicon (kata & bobot awal)
known_targets ← Ø
found_pairs ← Ø

for iteration in 1..MAX_ITER:
    new_opinions ← Ø
    new_targets ← Ø

    for each tweet in dataset:
        for each sentence in tweet.doc:
            for each word in sentence.words:
                # Opinion → Target
                if word ∈ known_opinions:
                    for child in dependents(word):
                        if child.pos ∈ {NOUN, PROPN}:
                            new_targets.add(child)
                            found_pairs[tweet].add((child,word))
                        if child.deprel == conj and child.pos ∈ {ADJ, ADV}:
                            new_opinions.add(child)

                # Target → Opinion
                if word ∈ known_targets:
                    for child in dependents(word):
                        if child.pos ∈ {ADJ, ADV}:
                            new_opinions.add(child)
                        if child.deprel == conj and child.pos ∈ {NOUN, PROPN}:
                            new_targets.add(child)

    if new_opinions ∪ new_targets = Ø:
        break    # sudah konvergen

    known_opinions ∪= new_opinions
    known_targets ∪= new_targets

```

Gambar III.9 *Pseudocode Iterasi Double Propagation*

Selanjutnya, dilakukan proses *phrase expansion* dan *pruning*. Target diperluas menjadi frasa aspek dengan menambahkan kata sifat (ADJ) maupun kata benda (NOUN) di sekitarnya. Setelah itu, dilakukan *global pruning* dengan menyaring hanya target yang memiliki frekuensi di atas satu. Setelah itu, bobot sentimen dihitung dari *seed lexicon*. Sistem juga menangani negasi, jika kata opini berada dalam *window* ( $\pm 5$  kata) dengan kata negasi, polaritas dibalik. Implementasi perhitungan bobot sentimen dapat dilihat pada Gambar III.10. Dari bobot akhir, sistem kemudian menghasilkan probabilitas sentimen (positif, netral, negatif) dengan fungsi normalisasi.

```
function calculate_contextual_sentiment(target, opinions, doc):
    total_weight ← 0
    opinion_count ← 0

    for op in opinions:
        weight ← lexicon[op]
        if op berada dekat kata negasi dalam doc:
            weight ← -weight
        total_weight += weight
        opinion_count += 1

    if opinion_count == 0:
        total_weight ← doc_level_sentiment * 0.3

    return total_weight, opinion_count
```

Gambar III.10 *Pseudocode* Perhitungan Bobot Sentimen *Double Propagation*

Terakhir, untuk memetakan target dengan kategori aspek digunakan teks aspek yang merupakan gabungan kategori aspek dan deskripsinya. Untuk representasi vektor digunakan tiga opsi mode, di antaranya:

1. TF-IDF: menghitung bobot kata pada teks aspek.
2. FastText (*model pretrained cc.id.300.vec.gz*): representasi kata berbasis *pre-trained embeddings*.
3. Sentence-BERT (*model pretrained firqaaa/indo-sentence-bert-base*): representasi kalimat berbasis model *transformer*.

### III.4.3 Metode ASBA Bebas Domain: *Zero-Shot LLM* Generatif

Bagian ini menjelaskan implementasi *zero-shot LLM* generatif dengan model Gemini. Opsi model yang digunakan adalah Gemini-2.0-Flash dan Gemini-2.5-Flash dari API GoogleGenerativeAI. Sistem akan mengirimkan *system prompt* berisi definisi tugas, daftar aspek, dan aturan format keluaran. Kemudian, *user prompt* berisi *batch tweet* yang akan dianalisis. Setiap *batch* terdiri dari 25 *tweet*, jumlah ini dipilih untuk memaksimalkan kapasitas token. Detail *prompt* dapat dilihat pada Subbab III.2.2.

```
for response in model_outputs:
    candidates ← regex_extract(response)
    if candidates = ∅:
        candidates ← block_extract(response)

    for (tweet_text, aspect_raw, probs_raw) in candidates:
        probs ← parse_probabilities(probs_raw)

        # Tweet matching
        tweet_id ← exact_match(tweet_text, tweets_batch)
        if tweet_id = ∅:
            tweet_id ← fuzzy_match(tweet_text, tweets_batch)
        if tweet_id = ∅:
            tweet_id ← positional_fallback(tweets_batch)

        # Aspect matching
        aspect ← exact_match(aspect_raw, aspect_categories)
        if aspect = ∅:
            aspect ← fuzzy_match(aspect_raw, aspect_categories)

        probs ← normalize(probs)
        parsed_results.add({id, aspect, probs})
```

Gambar III.11 *Pseudocode Parsing Respons Zero-Shot LLM Generatif*

*Parsing* respons dilakukan dengan terlebih dahulu menerapkan pola *regex* ketat untuk mendeteksi struktur *Tweet-Aspect-Sentiment*. Apabila pola tersebut tidak sesuai, sistem menggunakan strategi *parsing* fleksibel dengan memecah teks ke dalam blok-blok terpisah dan mengekstrak komponennya secara individual. Selanjutnya, probabilitas sentimen diproses dengan cara menghapus simbol atau label tambahan, mengonversinya ke dalam bentuk numerik, dan melakukan normalisasi agar total bernilai 1,0. Untuk *tweet*, teks hasil *parsing* dicocokkan

melalui pencocokan langsung. Jika tidak berhasil, sistem menerapkan *fuzzy matching* dengan SequenceMatcher (*threshold* = 0,8). Apabila kedua metode gagal, tersedia opsi *positional matching*, yaitu memasangkan hasil dengan *tweet* input berdasarkan urutan kemunculannya dalam *batch*. Sementara itu, nama aspek yang dihasilkan model dinormalisasi terhadap daftar aspek resmi dengan pendekatan serupa, yaitu kombinasi pencocokan langsung dan *fuzzy matching*. Implementasi *parsing* respons dapat dilihat pada Gambar III.11.

#### III.4.4 Pembangunan Graf Interaksi dan Deteksi Komunitas

```
# Membangun Graf
G ← Ø
status_map ← {}
screen_map ← {}
name_map ← {}

for t in tweets:
    G.add_node(t.id, {text, name, reply_info})
    status_map[t.status_id] ← t.id
    screen_map[t.screen_name] ← t.id
    name_map[t.name] ← t.id

# Menambahkan Simpul dengan Aturan
for t in tweets:
    if t.reply_status_id ∈ status_map:
        G.add_edge(t.id, status_map[t.reply_status_id])
    if t.reply_screen_name ∈ screen_map:
        G.add_edge(t.id, screen_map[t.reply_screen_name])
    if t.reply_screen_name ∈ name_map:
        G.add_edge(t.id, name_map[t.reply_screen_name])
    if t.reply_status_id ∈ tweets.id:
        G.add_edge(t.id, t.reply_status_id)
```

Gambar III.12 *Pseudocode* Pembangunan Graf Interaksi

Bagian ini menjelaskan implementasi proses pembangunan graf interaksi dan deteksi komunitas dengan Louvain. Proses pembangunan graf menggunakan aturan yang dijelaskan pada Subbab III.3.3. Implementasi pembangunan graf interaksi menggunakan kakas NetworkX. Implementasi dalam bentuk *pseudocode* dapat dilihat pada Gambar III.12.

Selanjutnya, deteksi komunitas diimplementasikan menggunakan kakas python-louvain dengan parameter *default*. Setelah komunitas dideteksi, jika beberapa *tweet* dari satu pengguna berada dalam komunitas yang sama, simpul-simpul tersebut di-

*merge* menjadi satu, teks digabung, dan sisi dibangun ulang. Implementasi deteksi komunitas dan *merging user* dapat dilihat pada Gambar III.13.

```
# Deteksi Komunitas
partition ← louvain(G)

for node in G:
    node.community ← partition[node]

# Merge Pengguna
if "name" ∈ tweets.columns:
    for user in unique(tweets.name):
        nodes ← {n ∈ G | n.name = user}
        if |nodes| > 1 and all(nodes.community = c):
            merged ← merge_nodes(nodes)
            G.replace(nodes, merged)
```

Gambar III.13 *Pseudocode* Deteksi Komunitas dan *Merging User*

### III.4.5 Perhitungan Metrik Konsensus GDM

```
# Masukan
aspects_data, communities ← load_aspects_csv(file)
filtered ← {c: members ∈ communities | |members| ≥ MIN_MEMBERS}
if filtered = Ø:# Filter komunitas
    return "No valid communities"
within_ccs ← Ø
between_ccs ← Ø
# Konsensus Dalam Komunitas
for comm, members in filtered:
    comm_pairs ← combinations(members, 2)
    cc, ac_values ← calculate_consensus(comm_pairs, members,
                                            p_inner)
    if cc ≠ Ø:
        within_ccs.add(cc)
        log("Community", comm, "size", |members|, "CC(within)=", cc)
keys ← list(filtered.keys())# Konsensus Antarkomunitas
for (a, b) in combinations(keys, 2):
    cross_pairs ← {(u,v) | u ∈ filtered[a], v ∈ filtered[b]}
    cc, ac_values ← calculate_consensus(cross_pairs,
                                            aspects_data, p_inner)
    if cc ≠ Ø:
        between_ccs.add(cc)
        log("Between", a, b, "CC=", cc)
within_mean ← mean(within_ccs) # Rata-Rata Agregat
between_mean ← mean(between_ccs)
ECR ← within_mean / between_mean
save_summary(within_mean, between_mean, ECR, file_summary)
save_details(within_ccs, between_ccs, file_details)
```

Gambar III.14 *Pseudocode* Perhitungan Metrik Konsensus GDM

Bagian ini menjelaskan implementasi proses perhitungan metrik konsensus GDM. Proses perhitungan sesuai dengan yang dijelaskan pada Subbab III.3.4. Masukan berupa hasil gabungan keluaran ASBA dan deteksi komunitas. Perhitungan konsensus dilakukan dalam beberapa tingkat: (1) konsensus pasangan, dihitung dengan *cosine similarity* pada vektor sentimen aspek; (2) konsensus aspek, agregasi konsensus pasangan menggunakan *arithmetic mean*; (3) konsensus komunitas, agregasi konsensus aspek dengan *arithmetic mean*; (4) konsensus dalam komunitas global, rata-rata seluruh konsensus komunitas; dan (5) konsensus antarkomunitas global, rata-rata konsensus antara setiap pasang komunitas. Komunitas dengan anggota <5 tidak diperhitungkan. Implementasi perhitungan metrik konsensus GDM dapat dilihat pada Gambar III.14.

## **BAB IV**

### **EKSPERIMENT DAN PENGUJIAN**

Bab ini menyajikan proses eksperimen dan pengujian yang dilakukan dalam tugas akhir ini. Bab ini menjelaskan tujuan eksperimen, skenario yang digunakan, hasil yang diperoleh, serta analisis terhadap hasil tersebut untuk mengevaluasi efektivitas metode yang diusulkan.

#### **IV.1 Eksperimen Praproses**

Tujuan dari eksperimen ini adalah menentukan konfigurasi praproses teks terbaik yang dapat meningkatkan kualitas input untuk ASBA *double propagation* dan *zero-shot* LLM generatif. Eksperimen dilakukan pada *dataset* Pilkada DKI dan Kampus Merdeka (*dataset* eksperimen konfigurasi ASBA) menggunakan metode ASBA *double propagation* dengan konfigurasi model *default*, yaitu *processor Stanza* dan representasi vektor TF-IDF, serta pada metode *zero-shot* LLM generatif menggunakan model gemini-2.0-flash. Tiga tingkat praproses yang diuji adalah sebagai berikut:

1. Pembersihan minimal: penghapusan *whitespace* berlebih dan tanda koma.
2. Pembersihan penuh: *case folding*, penghapusan tautan, *hashtag*, *mention*, kata “RT”, serta karakter non-alfanumerik.
3. Pembersihan penuh + normalisasi: penggantian kata *slang* menjadi bentuk baku menggunakan leksikon TheOnlyDo (2023) dan Ibrohim dan Budi (2018), serta penghapusan *stopwords* menggunakan daftar Tala (2003).

Pengukuran performa klasifikasi sentimen dilakukan menggunakan *Penalized Brier Score* (PBS). Karena PBS terdiri dari *Standard Brier Score* (0-2) dan *penalty* (0,67 untuk 3 kelas), rentang PBS untuk klasifikasi 3 kelas adalah 0 (sempurna) hingga 2,67 (terburuk). PBS dipilih karena merupakan metrik yang baik dan standar untuk pengukuran *soft label*, sedangkan representasi sentimen dalam bentuk *soft label* dibutuhkan untuk menghitung kesamaan sentimen pada pengukuran *echo*

*chamber*. Selain itu, dihitung pula *Baseline Brier Score* berdasarkan probabilitas prior kelas sebagai *baseline*. Sementara itu, pengukuran performa deteksi kategori aspek dilakukan menggunakan F1 makro.

Tabel IV.1 Hasil Eksperimen Praproses

<i>Double Propagation</i>				
<i>Dataset</i>	Tingkat Praproses	F1 Makro (klasifikasi kategori aspek)	<i>Penalized Brier Score</i> (klasifikasi sentimen)	<i>Baseline Brier Score</i>
Pilkada DKI	Pembersihan Minimal	0,1460	<b>1,4719</b>	1,1268
	Pembersihan Penuh	<b>0,2274</b>	1,5720	
	Pembersihan Penuh + Normalisasi	0,1818	1,6046	
Kampus Merdeka	Pembersihan Minimal	0,3804	1,5418	1,1417
	Pembersihan Penuh	<b>0,3970</b>	<b>1,4219</b>	
	Pembersihan Penuh + Normalisasi	0,3557	1,4766	
<i>Zero-shot LLM generatif</i>				
<i>Dataset</i>	Tingkat Praproses	F1 Makro (klasifikasi kategori aspek)	<i>Penalized Brier Score</i> (klasifikasi sentimen)	<i>Baseline Brier Score</i>
Pilkada DKI	Pembersihan Minimal	0,4980	0,7313	1,1268
	Pembersihan Penuh	<b>0,6562</b>	<b>0,7118</b>	
	Pembersihan Penuh + Normalisasi	0,5894	0,7344	
Kampus Merdeka	Pembersihan Minimal	0,6322	0,6706	1,1417
	Pembersihan Penuh	0,6200	<b>0,6310</b>	
	Pembersihan Penuh + Normalisasi	<b>0,7335</b>	0,7402	

Tabel IV.1 menampilkan hasil eksperimen praproses. Hasil eksperimen praproses menunjukkan bahwa pembersihan penuh merupakan pendekatan yang paling konsisten untuk kedua metode. Pada metode *double propagation*, skor F1 dan PBS meningkat dari pembersihan minimal ke pembersihan penuh. Namun, penambahan normalisasi justru menurunkan performa, kemungkinan akibat hilangnya variasi kata yang penting. Meskipun begitu, pada Pilkada DKI, PBS untuk klasifikasi sentimen sedikit lebih baik pada pembersihan minimal dibanding pembersihan penuh. Hal ini menunjukkan bahwa dengan metode *double propagation* dalam beberapa kasus pembersihan minimal bisa lebih menguntungkan untuk klasifikasi sentimen. Namun, secara keseluruhan pembersihan menyeluruh tanpa normalisasi masih menjadi konfigurasi paling optimal. Di sisi lain, metode *zero-shot* LLM generatif memperoleh peningkatan performa signifikan dari pembersihan teks, terutama pada *dataset* dengan *noise* tinggi seperti Pilkada DKI, yang ditandai oleh peningkatan F1 secara substansial. Disisi lain, efek normalisasi cenderung bervariasi, pada *dataset* Kampus Merdeka normalisasi memberikan dampak positif. Namun pada *dataset* Pilkada DKI dan tugas klasifikasi sentimen pada kedua *dataset*, normalisasi menurunkan performa. Oleh karena itu, dapat disimpulkan bahwa pembersihan penuh adalah langkah praproses paling optimal untuk kedua metode ASBA.

## IV.2 Eksperimen Konfigurasi ASBA

Tujuan dari eksperimen ini adalah menentukan model dan konfigurasi ASBA terbaik untuk mengklasifikasikan kategori aspek dan sentimen dari teks dalam konteks *tweet* bahasa Indonesia informal dalam berbagai domain topik untuk *double propagation* dan *zero-shot* LLM generatif.

Eksperimen dilakukan menggunakan *dataset* Pilkada DKI dan Kampus Merdeka (*dataset* validasi) yang telah melalui tahap praproses penuh, sesuai dengan konfigurasi yang memberikan hasil terbaik pada eksperimen praproses sebelumnya. Konfigurasi eksperimen ini adalah sebagai berikut:

1. Metode *Double Propagation*:
  - a. *Processor (Dependency Parser* dan *POS Tagging)*: UDPipe Indonesian-CSUI-UD-2.12-230717 dan Stanza-ID.

- b. Representasi vektor: Sentence-BERT (*model pretrained* firqaaa/indo-sentence-bert-base), FastText (*model pretrained* cc.id.300.vec.gz), TF-IDF.
2. Metode *zero-shot* LLM generatif:
- a. Model: Gemini-2.0-flash dan Gemini-2.5-flash

Pengukuran performa dilakukan menggunakan F1 makro untuk evaluasi deteksi kategori aspek dan *Penalized Brier Score* (PBS) untuk evaluasi klasifikasi sentimen.

Tabel IV.2 menampilkan eksperimen konfigurasi ASBA. Berdasarkan hasil eksperimen, performa metode *double propagation* dan *zero-shot* LLM generatif menunjukkan perbedaan yang signifikan. Untuk metode *double propagation*, performa sangat bergantung pada kombinasi *processor* dan representasi vektor. Pada *dataset* Pilkada DKI, Stanza dengan Fasttext memberikan F1 makro tertinggi (0,3287) untuk klasifikasi aspek, sementara UDpipe dengan Fasttext sedikit lebih unggul dalam klasifikasi sentimen dengan PBS 1,5082. Sedangkan pada *dataset* Kampus Merdeka, kombinasi UDpipe dengan TF-IDF menunjukkan performa paling konsisten baik untuk klasifikasi aspek (F1 makro 0,4371) maupun klasifikasi sentimen (PBS 1,3450), menunjukkan bahwa TF-IDF lebih unggul untuk *dataset* yang lebih besar atau beragam. Meskipun begitu, seluruh konfigurasi *double propagation* memiliki PBS yang lebih buruk dari *baseline*, sehingga tidak dapat disebut baik.

Sementara itu, metode *zero-shot* LLM generatif, khususnya Gemini-2.5-flash, secara signifikan mengungguli semua konfigurasi *double propagation*. Pada Pilkada DKI, Gemini-2.5-flash mencapai F1 makro 0,7256 dan PBS 0,6134, sedangkan pada Kampus Merdeka, meski Gemini-2.0-flash sedikit lebih baik untuk PBS (0,6258), Gemini-2.5-flash tetap menawarkan keseimbangan terbaik antara klasifikasi aspek dan klasifikasi sentimen dengan F1 makro 0,6479. Secara keseluruhan, LLM generatif, khususnya model Gemini-2.5-flash, terbukti jauh lebih efektif dan seimbang dibanding metode *double propagation*.

Tabel IV.2 Hasil Eksperimen Konfigurasi ASBA

<i>Double Propagation</i>					
<i>Dataset</i>	<i>Processor</i>	Representasi Vektor	F1 Makro (klasifikasi kategori aspek)	<i>Penalized Brier Score</i> (klasifikasi sentimen)	<i>Baseline Brier Score</i>
Pilkada DKI	Stanza	TF-IDF	0,2274	1,6493	1,1268
		Fasttext	<b>0,3287</b>	<b>1,5235</b>	
		Sentence-BERT	0,2866	1,5550	
	UDpipe	TF-IDF	0,2622	1,6826	
		Fasttext	<b>0,3158</b>	<b>1,5082</b>	
		Sentence-BERT	0,2911	1,5156	
Kampus Merdeka	Stanza	TF-IDF	0,3970	<b>1,3994</b>	1,1417
		Fasttext	<b>0,4162</b>	1,4534	
		Sentence-BERT	0,3968	1,4808	
	UDpipe	TF-IDF	<b>0,4371</b>	<b>1,3450</b>	
		Fasttext	0,4130	1,3863	
		Sentence-BERT	0,4307	1,4031	
<i>Zero-shot LLM generatif</i>					
<i>Dataset</i>	Model		F1 Makro (klasifikasi kategori aspek)	<i>Penalized Brier Score</i> (klasifikasi sentimen)	<i>Baseline Brier Score</i>
Pilkada DKI	Gemini-2.0-flash		0,6768	0,7184	1,1268
	<b>Gemini-2.5-flash</b>		<b>0,7256</b>	<b>0,6134</b>	
Kampus Merdeka	Gemini-2.0-flash		0,6200	<b>0,6258</b>	1,1417
	Gemini-2.5-flash		<b>0,6479</b>	0,6879	

### IV.3 Pengujian ASBA

Tujuan dari pengujian ini adalah menampilkan hasil kinerja model ASBA terpilih dengan konfigurasi terbaik dari eksperimen konfigurasi sebelumnya, yaitu model Gemini-2.5-flash. Pengujian dilakukan menggunakan *dataset* Pilkada DKI dengan aspek paslon dan *dataset Indo Vaccination* (*dataset* pengujian ASBA).

Tabel IV.3 Hasil Pengujian ASBA

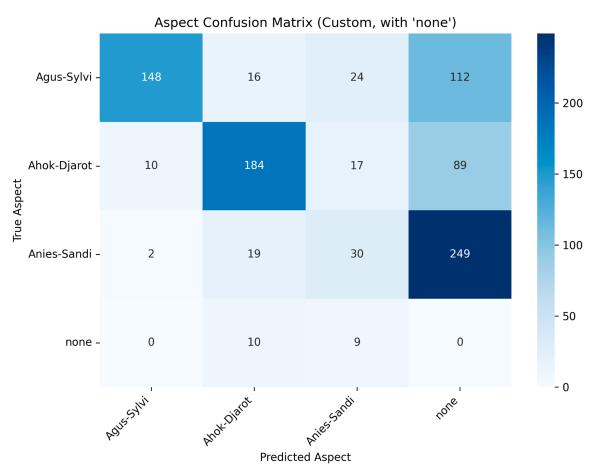
<i>Dataset</i>	Model	F1 Makro (klasifikasi kategori aspek)	<i>Penalized Brier Score</i> (klasifikasi sentimen)	<i>Baseline Brier Score</i>
Pilkada DKI dengan Aspek Paslon	Gemini-2.5-flash	0,4178	0,4834	0,9026
<i>Indo Vaccination</i>	Gemini-2.5-flash	0,4990	0,6935	0,8333

Tabel IV.3 menampilkan hasil pengujian ASBA, dengan *confusion matrix* dapat dilihat pada Gambar IV.1 (untuk *dataset* Pilkada DKI dengan aspek paslon) dan Gambar IV.2 (untuk *dataset Indo Vaccination*). Pada *dataset* Pilkada DKI dengan aspek paslon, model mampu memisahkan kelas secara relatif baik. Misalnya, aspek Anies-Sandi (249 benar) dan Ahok-Djarot (184 benar) dikenali dengan baik dengan tingkat kesalahan antarkelas yang kecil. Hal ini menunjukkan bahwa ketika aspek diekspresikan secara eksplisit model lebih mudah menangkap pola.

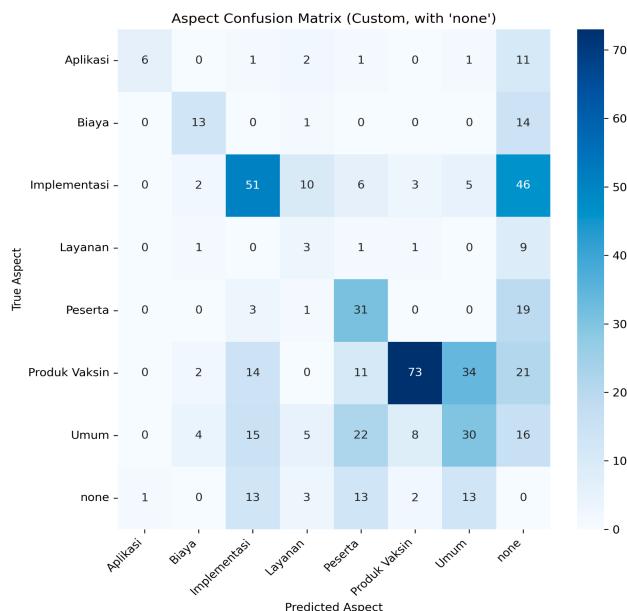
Sebaliknya, pada *dataset Indo Vaccination*, performa model lebih rendah. Banyak aspek konseptual seperti implementasi, layanan, atau peserta tercampur dengan kategori umum atau tidak terdeteksi, sehingga menurunkan akurasi klasifikasi. Hal ini mengindikasikan bahwa aspek yang tidak diekspresikan secara eksplisit lebih sulit diidentifikasi.

Selain itu, terdapat temuan menarik pada kasus *dataset* Pilkada DKI dengan aspek paslon. Karena banyak *hashtag* yang langsung merepresentasikan nama paslon sebagai target opini, ketika praproses dilakukan hanya sampai tahap pembersihan minimal, performa F1 makro meningkat hingga 0,6982, meskipun di sisi lain PBS

memburuk menjadi 0,6438. Hal ini mengindikasikan bahwa menjaga informasi berupa *hashtag* tertentu dapat membantu model mengenali aspek dengan lebih baik, walaupun membuat klasifikasi sentimen menjadi lebih tidak stabil. Sebaliknya, pola ini tidak terlihat pada *dataset Indo Vaccination*. Fakta ini menegaskan bahwa efektivitas praproses sangat kontekstual dan harus disesuaikan dengan karakteristik domain.



Gambar IV.1 *Confusion Matrix* Hasil Pengujian ASBA untuk *Dataset* Pilkada DKI dengan Aspek Paslon



Gambar IV.2 *Confusion Matrix* Hasil Pengujian ASBA untuk *Dataset* Indo Vaccination

Secara umum, hasil evaluasi menunjukkan bahwa model cenderung konservatif. Ketika model memprediksi sebuah aspek, hasilnya umumnya tepat dengan *precision* tinggi, misalnya Agus-Sylvi 92,5%, Aplikasi 85,7%, dan Produk Vaksin 83,9%. Namun, banyak aspek relevan yang luput terdeteksi sehingga *recall* rendah, seperti Anies-Sandi yang hanya 10,0% atau Layanan 33,3%. Dengan kata lain, model lebih sering gagal mengenali aspek daripada salah mengklasifikasikan. Misalnya, pada *tweet Indo Vaccination* “*gw asn teh blm dpt giliran pdhl hari2 ktemu bnyk orng apa lg di lapangan kyk taun kmrn kerja campur harap2 cemas plg krmh g brani dkt2 anak istri tdr dkmr terpisah ee bca brita koruptor di vaksin dluan,*” *gold label* menunjukkan aspek Peserta dan Implementasi, tetapi model hanya berhasil mendeteksi Peserta dan melewatkannya Implementasi.

Dari sisi pengukuran *echo chamber*, karakteristik ini masih dapat diterima. Hal ini karena, label yang berhasil diprediksi relatif bersih dari *false positives*, sehingga distribusi sentimen dalam komunitas tidak banyak terdistorsi. Sementara itu, rendahnya *recall* hanya mengurangi jumlah opini yang terbaca, bukan mengubah arah polarisasi. Selama opini yang terlewat tidak terkonsentrasi pada komunitas tertentu, pola polarisasi tetap tercermin dengan baik. Selain itu, kinerja klasifikasi sentimen cukup baik, tercermin dari nilai *Penalized Brier Score* yang lebih rendah dibanding *baseline*, sehingga prediksi sentimen dapat dianggap stabil. Dengan pertimbangan tersebut, model dengan konfigurasi terbaik ini tetap digunakan dalam *pipeline* pengukuran *echo chamber*.

#### **IV.4 Pengujian Pengukuran *Echo Chamber***

Tujuan dari pengujian ini adalah menampilkan hasil pengukuran *echo chamber* untuk *dataset Indo Vaccination* dengan label ASBA dari *gold label*, *Indo Vaccination* dengan label ASBA dari hasil inferensi, *Budget Efficiency*, dan Pilpres 2024 (*dataset* pengujian pengukuran *echo chamber*) menggunakan keseluruhan *pipeline* dengan konfigurasi terbaik berdasarkan eksperimen sebelumnya. Sebagai pembanding, dilakukan pula pengukuran *echo chamber* dengan model *baseline* dari Amendola dkk. (2024) dengan deteksi komunitas METIS  $k = 2$ . Sebagai catatan, walaupun pada model *baseline* tidak terdapat keluaran metrik ECR, ECR tetap dikalkulasi untuk kemudahan perbandingan.

Tabel IV.4 Hasil Pengujian Pengukuran *Echo Chamber*

<i>Dataset</i>	Konsensus Dalam Komunitas Global	Konsensus Antarkomunitas Global	ECR
<i>Indo Vaccination dengan gold label</i>	0,0464	0,0359	1,2918
<i>Indo Vaccination</i>	0,0554	0,0489	1,1315
<i>Budget Efficiency</i>	0,2200	0,0484	4,5497
Pilpres 2024	0,3891	0,1330	2,9259

Tabel IV.4 menampilkan hasil pengujian pengukuran *echo chamber*. Hasil keluaran detail konsensus serta keluaran graf interaksi dan komunitas dapat dilihat pada Lampiran D. Sebagai informasi, komunitas yang dievaluasi untuk *dataset Pilpres 2024* berjumlah 29 komunitas, untuk *dataset Budget Efficiency* berjumlah 10 komunitas, dan untuk *dataset Indo Vaccination* berjumlah 5 komunitas.

Untuk *dataset Pilpres 2024*, terlihat bahwa konsensus dalam komunitas global cukup tinggi (0,3891), sedangkan konsensus antarkomunitas global sangat rendah (0,1330). Kondisi ini menghasilkan nilai ECR sebesar 2,9259 yang menandakan terbentuknya *echo chamber* yang cukup kuat yang terjadi saat interaksi internal komunitas cukup jauh lebih kohesif dibanding interaksi lintas komunitas. Hal ini semakin jelas jika melihat komunitas dengan konsensus tertinggi, seperti komunitas 148 yang mencapai 0.998 dan secara eksklusif membicarakan Ganjar–Mahfud dalam sentimen positif, serta komunitas 46 dengan konsensus 0.9954 yang fokus pada narasi positif mengenai Anies. Menariknya, konsensus antarkomunitas antara keduanya bernilai 0.0000, menandakan tidak ada titik temu sama sekali dalam percakapan kedua kelompok ini. Hasil ini sejalan dengan temuan Drone Emprit pada Pilpres 2024 yang menunjukkan bahwa klaster pendukung Anies, Ganjar, dan Prabowo cenderung hanya berinteraksi di dalam kelompoknya masing-masing tanpa dialog lintas kubu (Manuputty dkk., 2025). Minimnya interaksi lintas komunitas dan tingginya keseragaman opini di dalam kelompok masing-masing ini

menegaskan adanya polarisasi tajam di media sosial dengan karakteristik *echo chamber* yang kuat.

Sementara itu untuk *dataset Budget Efficiency*, hasil perhitungan menunjukkan bahwa konsensus dalam komunitas global relatif rendah (0,2200), sedangkan konsensus antarkomunitas global jauh lebih rendah lagi (0,0484). Perbedaan tajam ini menghasilkan ECR sebesar 4,5497, yang menandakan terbentuknya *echo chamber* yang sangat kuat, meskipun keseragaman dalam komunitas tidak tinggi secara absolut, perbandingan dengan rendahnya konsensus antarkomunitas tetap memperlihatkan adanya polarisasi. Hal ini karena opini lebih seragam di dalam kelompok tetapi hampir tidak ada titik temu antarkelompok. Temuan ini juga selaras dengan analisis Fahmi (2025) yang menunjukkan adanya klaster publik yang terpisah antara positif dan negatif, sementara klaster media hanya positif saja. Hal ini tercermin dari konsensus dalam komunitas 4 (0,9872) yang merupakan media dengan sentimen positif terhadap aspek dampak, namun memiliki konsensus antarkomunitas yang sangat rendah dengan kelompok lain sehingga ikut menurunkan nilai konsensus antarkomunitas global. Sebaliknya, komunitas publik menunjukkan konsensus internal yang lebih rendah, hanya sekitar 0,1, sehingga mencerminkan adanya heterogenitas di dalam kelompok publik. Namun, bila dibandingkan dengan kelompok lain, tetap terlihat adanya disparitas yang tajam sehingga memperkuat polarisasi. Pola ini menegaskan terjadinya *echo chamber* yang kuat, sejalan dengan nilai ECR yang sangat tinggi.

Untuk *dataset Indo Vaccination*, hasil perhitungan menunjukkan konsensus dalam komunitas global sebesar 0,0554 dan konsensus antarkomunitas global sebesar 0,0489, dengan nilai ECR 1,1315. Nilai ini masih menunjukkan adanya *echo chamber*, namun tidak terlalu kuat. Temuan ini sedikit kontradiktif dengan hasil analisis Sugiono (2021) yang menyebutkan bahwa *echo chamber* pada isu vaksinasi di Indonesia cukup kuat, terutama karena agresivitas kelompok antivaksin dalam menyebarkan narasi negatif. Perbedaan ini dapat dipahami karena *dataset* yang digunakan di sini hanya mencakup lima komunitas, sehingga tidak mewakili keseluruhan interaksi yang lebih luas. Selain itu, dalam lima komunitas tersebut tidak terdapat kelompok antivaksin, sehingga *echo chamber* yang kuat tidak terdeteksi. Hal ini justru menunjukkan keunggulan model, karena mampu

mengidentifikasi secara lebih detail sampai ke tingkat kelompok-kelompok yang terlibat dalam suatu *echo chamber*, sehingga absennya kelompok antivaksin dalam sampel data langsung tercermin pada rendahnya nilai ECR.

Selain itu, dilakukan juga pengujian terhadap *dataset Indo Vaccination* menggunakan label ASBA *gold label*. Hasilnya tidak jauh berbeda dengan hasil inferensi, dengan perbedaan nilai ECR hanya sekitar 0,16 serta perbedaan konsensus dalam dan antarkomunitas global sekitar 0,01. Hal ini menunjukkan bahwa hasil inferensi ASBA cukup dapat diandalkan, karena meskipun performa akurasinya pada pengujian ASBA terlalu tinggi, perbedaan tersebut tidak berpengaruh signifikan terhadap hasil perhitungan ECR.

Tabel IV.5 Hasil *Baseline* Pengukuran *Echo Chamber* Amendola dkk. (2024)

<i>Baseline</i>				
<i>Dataset</i>	Konsensus Dalam Komunitas 0	Konsensus Dalam Komunitas 1	Konsensus Antarkomunitas 0-1	ECR
<i>Indo Vaccination</i> dengan <i>gold label</i>	0,0606	0,0401	0,0394	1,2790
<i>Indo Vaccination</i>	0,0568	0,0567	0,0512	1,1087
<i>Budget Efficiency</i>	0,1174	0,1131	0,0514	2,2422
Pilpres 2024	0,1201	0,1040	0,1115	1,0049

Tabel IV.5 menampilkan hasil pengukuran *echo chamber* dengan model *baseline*. Pada hasil, keempat *dataset* menunjukkan tingkat *echo chamber* yang lebih rendah dibandingkan dengan hasil model yang dikembangkan dalam tugas akhir ini. Bahkan pada *dataset* Pilpres 2024, nilai ECR mendekati 1, menandakan *echo chamber* yang relatif lemah. Hasil ini menunjukkan bahwa model *baseline* gagal mengidentifikasi *echo chamber* yang cukup kuat. Salah satu faktor penyebabnya adalah pemartisian graf menjadi hanya dua komunitas, yang menghasilkan pembagian kurang alami dengan ukuran komunitas yang besar sehingga konsensus dalam komunitas menurun. Sementara itu, pada *dataset Indo Vaccination*

perbedaan nilai ECR relatif kecil. Hal ini terjadi karena jumlah komunitas yang dihasilkan oleh deteksi komunitas Louvain (sebanyak 4 komunitas) tidak terlalu berbeda jauh dengan hasil partisi METIS  $k = 2$ , sehingga struktur komunitas yang terbentuk masih memiliki kemiripan.

Meskipun demikian, hasil model *baseline* tetap menunjukkan pola yang sejalan dengan model yang dikembangkan, dengan urutan nilai ECR: *Budget Efficiency* > Pilpres 2024 > *Indo Vaccination*. Selain itu, nilai konsensus antarkomunitas antara hasil model *baseline* dengan model tugas akhir hampir sama, dengan selisih kurang dari 0,1. Hal ini menegaskan bahwa model baru lebih sensitif dibandingkan *baseline* dalam mendeteksi kekuatan *echo chamber*, namun tetap mempertahankan konsistensi pola perbedaan relatif antar-*dataset* dan interaksi antarkomunitas yang didapat dari model *baseline*.

#### IV.5 Uji Hipotesis Pengukuran *Echo Chamber*

Tujuan dari uji hipotesis ini adalah untuk membuktikan bahwa hasil model pengukuran *echo chamber* pada *dataset echo chamber Indo Vaccination, Budget Efficiency*, dan Pilpres 2024 bersifat signifikan secara statistik dan bukan sekadar kebetulan jika dibandingkan dengan kondisi acak. Dengan demikian, nilai pengukuran yang diperoleh dapat dipastikan mencerminkan fenomena *echo chamber* yang nyata, bukan sekadar hasil dari distribusi acak.

Secara formal, hipotesis yang diuji adalah sebagai berikut:

1.  $H_0$  (hipotesis nol): nilai ECR observasi tidak berbeda secara signifikan dari distribusi acak (tidak terdapat pola *echo chamber* nyata).
2.  $H_1$  (hipotesis alternatif): nilai ECR observasi berbeda secara signifikan dari distribusi acak (terdapat pola *echo chamber* nyata).

Uji hipotesis dirancang dengan memanfaatkan *null model* sebagai pembanding. *Null model* dibuat melalui dua pendekatan permutasi:

1. *Null Atribut*: distribusi sentimen antarpengguna diacak (*shuffle*), sementara struktur komunitas (topologi jaringan) dipertahankan.

2. *Null Topologi*: struktur jaringan diacak melalui algoritma *double-edge swap* (Fosdick dkk., 2018), sementara distribusi sentimen tetap dipertahankan.

Hasil uji hipotesis dapat dilihat pada Tabel IV.6 (*null* atribut) dan Tabel IV.7 (*null* topologi). Hasil distribusi histogram untuk model null dapat dilihat pada Lampiran E. Karena keterbatasan komputasi, uji hipotesis dijalankan dengan 100 permutasi baik pada model *null* atribut maupun topologi. Walaupun jumlah 100 permutasi hanya memberikan estimasi *p-value* dengan presisi hingga dua desimal, hasil yang diperoleh sudah cukup ekstrem (ECR observasi berada jauh di atas rata-rata distribusi *null*). Oleh karena itu, penambahan jumlah permutasi diperkirakan tidak akan mengubah kesimpulan signifikan secara substantif, melainkan hanya meningkatkan presisi estimasi *p-value*.

Tabel IV.6 Hasil Uji Hipotesis Pengukuran *Echo Chamber* dengan *Null* Atribut

Dataset	ECR Observasi	<i>Null</i> atribut			
		<i>mean</i>	<i>std</i>	<i>z-score</i>	<i>empirical p-value</i>
<i>Indo Vaccination</i>	1,1315	0,9919	0,0413	3,3777	0,0
<i>Budget Efficiency</i>	4,5497	0,9788	0,0136	262,16	0,0
Pilpres 2024	2,9259	0,9922	0,1220	15,849	0,0

Tabel IV.7 Hasil Uji Hipotesis Pengukuran *Echo Chamber* dengan *Null* Topologi

Dataset	ECR Observasi	<i>Null</i> topologi			
		<i>mean</i>	<i>std</i>	<i>z-score</i>	<i>empirical p-value</i>
<i>Indo Vaccination</i>	1,1315	0,9854	0,0411	3,5514	0,01
<i>Budget Efficiency</i>	4,5497	0,9789	0,0117	303,99	0,0
Pilpres 2024	2,9259	1,0159	0,0499	38,214	0,0

Hasil uji hipotesis menunjukkan bahwa seluruh *dataset* yang diuji (Indo-Vaccination, *Budget Efficiency*, dan Pilpres 2024) memiliki nilai ECR observasi yang secara konsisten lebih tinggi dibandingkan dengan nilai rata-rata distribusi *null*, baik pada model *null* atribut maupun *null* topologi. Pada uji *null* atribut, nilai *z-score* yang diperoleh berkisar antara 3,3777 hingga 262,16, sedangkan pada uji *null* topologi berkisar antara 3,5514 hingga 303,99. Nilai *z-score* yang sangat tinggi ini menunjukkan bahwa hasil pengukuran berada jauh dari distribusi acak, sehingga kecil kemungkinan diperoleh secara kebetulan. Hal ini juga diperkuat dengan nilai *empirical p-value* yang seluruhnya berada pada tingkat 0,01 atau lebih kecil, yang berarti hasil pengukuran signifikan pada taraf kepercayaan 95% maupun 99%.

Dengan demikian, hipotesis nol ( $H_0$ ) ditolak dan hipotesis alternatif ( $H_1$ ) diterima pada seluruh *dataset*. Artinya, fenomena *echo chamber* yang terdeteksi pada ketiga *dataset* bukan merupakan artefak dari distribusi acak, melainkan mencerminkan pola konsensus yang signifikan secara statistik.

## BAB V

### KESIMPULAN DAN SARAN

#### V.1 Kesimpulan

Kesimpulan yang dapat diambil dari tugas akhir ini adalah sebagai berikut.

1. Model pengukuran *echo chamber* pada percakapan media sosial berbahasa Indonesia dapat dikembangkan dengan mengintegrasikan Analisis Sentimen Berbasis Aspek (ASBA) bebas domain yang sesuai untuk konteks bahasa Indonesia di media sosial yaitu dengan pendekatan *zero-shot LLM* generatif (Gemini 2.5 Flash). Data interaksi antar pengguna diolah menjadi graf berbasis *reply*, *retweet*, dan/atau *mention*, kemudian dilakukan deteksi komunitas menggunakan algoritma Louvain. Selanjutnya, tingkat konsensus dihitung dengan kerangka *Group Decision Making* (GDM), yaitu dengan memperoleh rata-rata konsensus dalam komunitas (konsensus dalam komunitas global) dan rata-rata konsensus antarpasangan komunitas (konsensus antarkomunitas global). Nilai *Echo Chamber Ratio* (ECR) diperoleh dari perbandingan kedua metrik konsensus tersebut, yang menjadi indikator kuantitatif utama untuk mengukur keberadaan *echo chamber*.
2. Hasil pengukuran *echo chamber* dengan model berbasis ASBA dan metrik konsensus GDM menunjukkan nilai ECR yang konsisten dengan analisis pakar pada studi *echo chamber* terdahulu. Pada dataset *Indo Vaccination*, konsensus dalam komunitas global sebesar 0,0554 dan konsensus antarkomunitas global 0,0489, menghasilkan ECR 1,1315. Dataset *Budget Efficiency* memperlihatkan tingkat *echo chamber* tertinggi, dengan konsensus dalam komunitas global 0,2200, konsensus antarkomunitas global 0,0484, dan ECR 4,5497. Adapun dataset Pilpres 2024 menunjukkan konsensus dalam komunitas global 0,3891, konsensus antarkomunitas global 0,1330, serta ECR 2,9259.
3. Hasil uji hipotesis menunjukkan bahwa nilai ECR observasi pada seluruh dataset (*Indo Vaccination*, *Budget Efficiency*, dan Pilpres 2024) konsisten lebih tinggi daripada rata-rata distribusi *null*, baik *null* atribut (hasil

pengacakan sentimen) maupun *null* topologi (hasil *rewiring* sisi). Nilai *z-score* yang sangat tinggi (3,37–303,99) serta *empirical p-value*  $\leq 0,01$  menunjukkan bahwa temuan *echo chamber* signifikan secara statistik dan bukan sekadar artefak acak.

4. Hasil pengukuran *echo chamber* pada *dataset Indo Vaccination* dengan *gold label* ( $ECR = 1,2918$ ) tidak jauh berbeda dibandingkan *dataset* dengan label hasil inferensi ( $ECR = 1,1315$ ), dengan selisih 0,1603. Hasil ini menunjukkan bahwa inferensi ASBA cukup *reliable* dan mampu menghasilkan estimasi *echo chamber* yang mendekati hasil dengan *gold label*.
5. Hasil eksperimen ASBA menunjukkan metode ASBA bebas domain terbaik untuk *tweet* berbahasa Indonesia adalah *zero-shot* LLM generatif dengan model Gemini-2.5-flash. Pada pengujian dengan *dataset* Pilkada DKI dengan Aspek Paslon, model ini mencapai F1 makro untuk klasifikasi aspek sebesar 0,4178 dan *Penalized Brier Score* untuk klasifikasi sentimen 0,4834 dengan *Baseline Brier Score* 0,9026. Untuk *dataset Indo Vaccination*, model yang sama meraih F1 makro 0,4990 dan *Penalized Brier Score* 0,6935 dengan *Baseline Brier Score* 0,8333.
6. Penelitian tugas akhir ini berhasil menunjukkan bahwa pengukuran *echo chamber* dengan ASBA dan metrik konsensus GDM tidak hanya memberikan hasil kuantitatif ringkas melalui nilai ECR, tetapi juga dapat dipadukan dengan analisis kualitatif untuk menjelaskan karakteristik komunitas melalui nilai konsensus dalam komunitas dan antarkomunitas.

## V.2 Saran

Saran yang dapat dipertimbangkan untuk penelitian lebih lanjut terkait topik tugas akhir ini adalah sebagai berikut.

1. Penelitian selanjutnya perlu menggunakan *dataset* dengan cakupan komunitas yang lebih luas untuk menangkap dinamika interaksi yang lebih lengkap dan representatif. Hal ini akan meningkatkan validitas temuan serta memperkuat generalisasi hasil analisis.

2. Perlu dilakukan eksplorasi metode ASBA dengan performa yang lebih baik, termasuk integrasi pemodelan topik otomatis untuk penentuan aspek. Penentuan aspek secara manual pada penelitian ini dapat mempengaruhi objektivitas dan skalabilitas hasil, sehingga penggunaan pendekatan otomatis dapat meningkatkan konsistensi analisis.
3. Penelitian selanjutnya dapat mengintegrasikan metrik-metrik lain, seperti kesamaan kata, kesamaan topik, atau metrik topologi lain. Pendekatan ini memungkinkan pemahaman yang lebih holistik tentang struktur dan interaksi komunitas.
4. Penelitian dapat dikembangkan untuk menetapkan *threshold echo chamber* yang masuk tahap bahaya, sehingga dapat menjadi dasar pengembangan strategi preventif untuk memitigasi efek negatif *echo chamber*.
5. Perlu dilakukan kerja sama dengan pakar di bidang komunikasi atau sosiologi untuk memverifikasi temuan penelitian dan memberikan masukan dari perspektif ilmu sosial yang relevan. Hal ini penting agar hasil penelitian dapat bermakna dalam konteks sosial.

## DAFTAR REFERENSI

- Ahmadian, R., Ghatee, M., & Wahlström, J. (2025). Superior scoring rules for probabilistic evaluation of single-label multi-class classification tasks. *International Journal of Approximate Reasoning*, 182, 109421. <https://doi.org/10.1016/j.ijar.2025.109421>
- Amendola, M., Cavaliere, D., De Maio, C., Fenza, G., & Loia, V. (2024). Towards echo chamber assessment by employing aspect-based sentiment analysis and GDM consensus metrics. *Online Social Networks and Media*, 39–40, 100276. <https://doi.org/10.1016/j.osnem.2024.100276>
- Arroyo-Fernández, I., Méndez-Cruz, C.-F., Sierra, G., Torres-Moreno, J.-M., & Sidorov, G. (2019). Unsupervised sentence representations as word information series: Revisiting TF-IDF. *Computer Speech & Language*, 56, 104–129. <https://doi.org/10.1016/j.csl.2019.01.005>
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542. <https://doi.org/10.1177/0956797615594620>
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. [https://doi.org/10.1162/tacl\\_a\\_00051](https://doi.org/10.1162/tacl_a_00051)
- Chumakov, S., Kovantsev, A., & Surikov, A. (2022). Generative approach to aspect based sentiment analysis with GPT language models. *Procedia Computer Science*, 229, 284–293. <https://doi.org/10.1016/j.procs.2023.12.030>
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, 64(2), 317–332. <https://doi.org/10.1111/jcom.12084>
- Ekawati, D., & Khodra, M. L. (2017, August). Aspect-based sentiment analysis for Indonesian restaurant reviews. In *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICAICTA.2017.8090973>
- Fahmi, I. (2025). Drone Emprit: Analisis percakapan publik: Efisiensi anggaran. *Drone Emprit*. <https://pers.droneemprit.id/analisis-percakapan-publik-efisiensi-anggaran/>
- Ferro, C. A., & Fricker, T. E. (2012). A bias-corrected decomposition of the Brier score. *Quarterly Journal of the Royal Meteorological Society*, 138(668), 1954–1960.
- Firdaus, A. A. (2025). Budget efficiency in Indonesia [Dataset]. *Mendeley Data*, VI. <https://doi.org/10.17632/25dhk4dn3r.1>
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3–5), 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>
- Fosdick, B. K., Larremore, D. B., Nishimura, J., & Ugander, J. (2018). Configuring random graph models with fixed degree sequences. *SIAM Review*, 60(2), 315–355.
- Glenn, W. B. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1), 1–3.
- González, W., & Wilson, G. (1997, September 10). Hypothesis testing. *Environmental Sampling & Monitoring Primer*. Virginia Tech. Diarsipkan dari asli pada 11 Desember 2011. Diakses pada 4 September 2025, dari

<https://web.archive.org/web/20111211083531/http://www.cee.vt.edu/ewr/environmental/tech/smprimer/hypotest/ht.html>

- Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science*, 17, 26–32. <https://doi.org/10.1016/j.procs.2013.05.005>
- Herrera-Viedma, E., Martínez, L., Mata, F., & Chiclana, F. (2005). A consensus support system model for group decision-making problems with multigranular linguistic preference relations. *IEEE Transactions on Fuzzy Systems*, 13(5), 644–658. <https://doi.org/10.1109/TFUZZ.2005.856561>
- Hobson, E. A., Silk, M. J., Fefferman, N. H., Larremore, D. B., Rombach, P., Shai, S., & Pinter-Wollman, N. (2021). A guide to choosing and implementing reference models for social network analysis. *Biological Reviews*, 96(6), 2716–2734.
- Hua, Y. C., Denny, P., Wicker, J., & Taskova, K. (2024). A systematic review of aspect-based sentiment analysis: Domains, methods, and trends. *Artificial Intelligence Review*, 57(11), 296. <https://doi.org/10.1007/s10462-024-10761-9>
- Ibrohim, M. O., & Budi, I. (2018). A dataset and preliminaries study for abusive language detection in Indonesian social media. *Procedia Computer Science*, 135, 222–229. <https://doi.org/10.1016/j.procs.2018.08.169>
- Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press.
- Jensen, K. S., Taylor, M. L., & Williamson, S. H. (2023). A typology of Twitter interactions to effectively analyse engagement and evidence research impact. *Research for All*, 7(1). <https://doi.org/10.14324/RFA.07.1.09>
- Jha, A. (2025). Vectorization techniques in NLP [Guide]. *Neptune.ai*. <https://neptune.ai/blog/vectorization-techniques-nlp>
- Juleha, J., Yuniar, J., & Marsuki, N. R. (2024). Peran media sosial dalam dinamika opini publik dan partisipasi politik era digital. *Concept: Journal of Social Humanities and Education*, 3(1), 38–45. <https://doi.org/10.35905/concept.v3i1.5582>
- Karmila, L., & Yuningsih, A. (2025). The dynamics of political identity in the digital era: The role of social media, polarization, and challenges to democracy. *TOPLAMA*, 2(2), 28–36.
- Karypis, G., & Kumar, V. (1998a). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1), 359–392. <https://doi.org/10.1137/S1064827595287997>
- Karypis, G., & Kumar, V. (1998b). Multilevel k-way partitioning scheme for irregular graphs. *Journal of Parallel and Distributed Computing*, 48(1), 96–129. <https://doi.org/10.1006/jpdc.1997.1404>
- Khairullah, P., & Karmagatri, M. (2025). Dataset of Fore Coffee Indonesia sentiment analysis [Dataset]. *Mendeley Data*, V1. <https://doi.org/10.17632/yp4t48y62.1>
- Koto, F., & Rahmaningtyas, G. Y. (2017, December). Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. In *2017 International Conference on Asian Language Processing (IALP)* (pp. 391–394). IEEE.
- Kreyszig, E. (1979). *Advanced engineering mathematics* (4th ed., p. 880, eq. 5). Wiley. ISBN 0-471-02140-7.
- Mahmoudi, A., Jemielniak, D., & Ciechanowski, L. (2024). Echo chambers in online social networks: A systematic literature review. *IEEE Access*, 12, 9594–9620. <https://doi.org/10.1109/ACCESS.2024.3350846>
- Manuputty, C. R., Hanita, M., & Ritonga, R. (2025). Fenomena bias politik media sosial dalam polarisasi selama pemilihan umum presiden 2024. *Warta Ikatan Sarjana Komunikasi Indonesia*, 8(1), 206–221. <https://doi.org/10.25008/wartaiski.v8i1.543>

- Mikolov, T., Corrado, G., Chen, K., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint*. <https://arxiv.org/abs/1301.3781>
- Mitchell, T. M. (2017). Generative and discriminative classifiers: Naive Bayes and logistic regression. In *Machine learning* (Chapter 3). McGraw-Hill Education.
- Morente-Molinera, J. A., Kou, G., Samuylov, K., Ureña, R., & Herrera-Viedma, E. (2019). Carrying out consensual group decision-making processes under social networks using sentiment analysis over comparative expressions. *Knowledge-Based Systems*, 165, 335–345. <https://doi.org/10.1016/j.knosys.2018.12.017>
- Morini, V., Pollacci, L., & Rossetti, G. (2021). Toward a standard approach for echo chamber detection: Reddit case study. *Applied Sciences*, 11(12), 5390. <https://doi.org/10.3390/app11125390>
- Mughal, N., Mujtaba, G., Shaikh, S., Kumar, A., & Daudpota, S. M. (2024). Comparative analysis of deep natural networks and large language models for aspect-based sentiment analysis. *IEEE Access*, 12, 60943–60959. <https://doi.org/10.1109/ACCESS.2024.3381141>
- Nasoha, A. M. M., Atqiyah, A. N., Mujahidin, M., & Ahnaf, R. A. (2025). Memahami Pancasila dalam algoritma media sosial: Analisis wacana digital tentang ideologi bangsa. *Jembatan Hukum: Kajian Ilmu Hukum, Sosial dan Administrasi Negara*, 2(2), 55–72. <https://doi.org/10.31000/jh.v2i2.15276>
- Nazir, A., Rao, Y., Wu, L., & Sun, L. (2020). Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. *IEEE Transactions on Affective Computing*, 13(2), 845–863. <https://doi.org/10.1109/TAFFC.2020.2973156>
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577–8582. <https://doi.org/10.1073/pnas.0601602103>
- Opitz, J. (2024). A closer look at classification evaluation metrics and a critical reflection of common evaluation practice. *Transactions of the Association for Computational Linguistics*, 12, 820–836.
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.
- Phipson, B., & Smyth, G. K. (2010). Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology*, 9(1), 39. <https://doi.org/10.2202/1544-6115.1585>
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., & Manandhar, S. (2016). SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* (pp. 19–30). Association for Computational Linguistics. <https://doi.org/10.18653/v1/S16-1002>
- Pratiwi, M. C., Puspito, S. D. A., Razani, M. F., & Rakhmawati, N. A. (2023). Perbandingan algoritma analisis sentimen pada isu kampus merdeka di X (Twitter). *Prosiding Seminar Nasional Informatika*.
- Purwitasari, D., Putra, C. B. P., & Raharjo, A. B. (2023). A stance dataset with aspect-based sentiment information from Indonesian COVID-19 vaccination-related tweets. *Data in Brief*, 47, 108951. <https://doi.org/10.1016/j.dib.2023.108951>
- Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion word expansion and target extraction through double propagation. *Computational Linguistics*, 37(1), 9–27. [https://doi.org/10.1162/coli\\_a\\_00034](https://doi.org/10.1162/coli_a_00034)
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3980–3990). Association for Computational Linguistics. <https://aclanthology.org/D19-1410/>

- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). Analisis sentimen tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler Indonesia pada Twitter dengan metode support vector machine dan lexicon based features. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(12), 1725–1732. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620. <https://doi.org/10.1145/361219.361220>
- Singh, T., & Kumari, M. (2016). Role of text pre-processing in Twitter sentiment analysis. *Procedia Computer Science*, 89, 549–554. <https://doi.org/10.1016/j.procs.2016.06.095>
- Sugiono, S. (2021). Polarization as the impact of strengthening of anti-vaccine groups in social media (echo chamber perspective). *Jurnal Penelitian Komunikasi dan Opini Publik*, 25(2), 149–164.
- Tala, F. Z. (2003). A study of stemming effects on information retrieval in Bahasa Indonesia (Master's thesis). Institute for Logic, Language and Computation, Universiteit van Amsterdam, The Netherlands.
- TheOnlyDo. (2023). Indonesia slang dictionary [Dataset]. *Hugging Face*. <https://huggingface.co/datasets/theonlydo/indonesia-slang>
- Wagenseller, P., & Wang, S. (2017). Size matters: A comparative analysis of community detection algorithms. *arXiv*. <https://arxiv.org/abs/1712.01690>
- Wardianto, W., Farikhin, F., & Kusumo Nugraheni, D. M. (2023). Analisis sentimen berbasis aspek ulasan pelanggan restoran menggunakan LSTM dengan Adam optimizer. *JOINTECS (Journal of Information Technology and Computer Science)*, 8(2), 679–686. <https://doi.org/10.31328/jointecs.v8i2.3580>
- X. (2025). Tweet object. *Dokumentasi X API v1*. <https://developer.x.com/en/docs/x-api/v1/data-dictionary/object-model/tweet>
- X Help Center. (2025). Mentions and replies. *Dokumentasi daring*. <https://help.x.com/en/using-x/mentions-and-replies>
- Zade, H., Williams, S., Tran, T. T., Smith, C., Venkatagiri, S., Hsieh, G., & Starbird, K. (2024). To reply or to quote: Comparing conversational framing strategies on Twitter. *ACM Journal on Computing and Sustainable Societies*, 2(1), 1–27. <https://doi.org/10.1145/3610105>
- Zhang, Z.-K., Liu, C., Zhang, Y.-C., & Zhou, T. (2016). Centralized nucleation in online social networks leads to high social clustering. *Scientific Reports*, 6, 35663. <https://doi.org/10.1038/srep35663>

## Lampiran A. Deskripsi Dataset Mentah

### A.1 Deskripsi Dataset Mentah untuk ASBA

Tabel A.1 Deskripsi Dataset Mentah untuk ASBA

Dataset Mentah untuk ASBA			
Nama Dataset	Kampus Merdeka	Pilkada DKI	Indo Vaccination
Sumber	<a href="https://zenodo.org/records/14234851">https://zenodo.org/records/14234851</a>	<a href="https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/blob/master/dataset_tweet_sentiment_pilkada_DKI_2017.csv">https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/blob/master/dataset_tweet_sentiment_pilkada_DKI_2017.csv</a>	<a href="https://data.mendeley.com/datasets/7ky2bjwtn/3">https://data.mendeley.com/datasets/7ky2bjwtn/3</a>
Deskripsi	<i>Dataset</i> diambil dari penelitian Pratiwi dkk. (2024), berupa <i>tweet X</i> dengan kata kunci “kampus merdeka” yang dikumpulkan melalui Tweet-Harvest pada periode 30 April–30 Oktober 2024 yang telah disaring untuk menghapus data yang tidak relevan.	<i>Dataset</i> diambil dari penelitian Rofiqoh dkk. (2017), berupa <i>tweet X</i> terkait dengan pelaksanaan Pemilihan Gubernur DKI Jakarta Tahun 2017 dari 3 pasangan calon yaitu Agus-Sylvi, Ahok-Djarot, dan Anies-Sandi. Telah dilakukan praproses berupa penggantian emoji dengan penanda spesial serta penghapusan <i>stopwords</i> .	<i>Dataset</i> diambil dari penelitian Purwitasari dkk. (2023), berupa <i>tweet X</i> terkait vaksinasi COVID-19 yang dikumpulkan dengan kata kunci tertentu pada periode Januari–Oktober 2021. <i>Dataset</i> telah disaring untuk menghapus data non-Bahasa Indonesia, non-topik vaksinasi, spam, dan duplikasi, lalu dilabeli secara manual dengan informasi <i>stance</i> dan sentimen berbasis aspek.
Kolom	a. full_text: teks <i>raw tweet</i> b. label: sentimen level dokumen, positif (1) dan negatif (0).	a. Id: <i>identifier</i> baris b. Sentiment: sentimen level dokumen, positif (“positive”) dan negatif (“negative”). c. Pasangan Calon: “Agus-Sylvi”, “Ahok-Djarot”, “Anies-Sandi”. d. Text Tweet: teks <i>tweet</i> .	a. id: ID <i>tweet</i> b. user_id: ID pengguna <i>author tweet</i> c. community: komunitas pengguna berdasarkan jaringan interaksi <i>retweet</i> dan <i>mention</i> dengan metode deteksi komunitas Louvain d. aspect_category: kategori aspek yang dibahas dalam <i>tweet</i> (satu <i>tweet</i> dapat memiliki lebih dari satu aspek yang dipisahkan spasi, jika anotator tidak sepakat dengan aspek maka diberi label “Invalid”) e. aspect_sentiment: sentimen terhadap aspek (“Negative”/“Positive”):

Dataset Mentah untuk ASBA			
Nama Dataset	Kampus Merdeka	Pilkada DKI	Indo Vaccination
			negatif/positif terhadap seluruh aspek; “Conflict”: aspek berbeda memiliki sentimen yang berbeda; “Invalid”: anotator tidak sepakat) f. stance: sikap pengguna terhadap isu (“favor”/“against”/“neutral”)
Jumlah Tweet	600	900	9030
Jumlah Kategori Aspek	-	-	7
Kategori Aspek	-	-	a. Implementation b. Services c. Costs d. Participants e. Apps f. Vaccine products g. General

## A.2 Deskripsi Dataset Mentah untuk Pengukuran Echo Chamber

Tabel A.2 Deskripsi Dataset Mentah untuk Pengukuran Echo Chamber

Dataset Mentah untuk Pengukuran Echo Chamber		
Nama Dataset	Pilpres 2024	Budget Efficiency
Sumber	<a href="https://pers.droneemprit.id/twitter-x-data-set-for-researchers/">https://pers.droneemprit.id/twitter-x-data-set-for-researchers/</a>	<a href="https://data.mendeley.com/datasets/25dhk4dn3r/1">https://data.mendeley.com/datasets/25dhk4dn3r/1</a>
Deskripsi	<i>Dataset</i> ini disediakan oleh Drone Emprit melalui program SATRIA DATA 2024, berupa <i>tweet X</i> terkait percakapan pemilihan presiden 2024 dengan kata kunci “Anies”, “Prabowo”, dan “Ganjar”. Data dikumpulkan pada periode Januari 2024 dan dibagikan untuk kepentingan riset dan kompetisi akademik.	<i>Dataset</i> diambil dari Mendeley Data (Firdaus, 2025), berupa <i>tweet X</i> berbahasa Indonesia terkait isu efisiensi anggaran yang diakses pada Februari 2025 melalui API X. <i>Dataset</i> adalah data primer yang belum diproses, namun telah disaring dengan memperhatikan kebijakan privasi pengguna serta etika penelitian.
Kolom	a. id: ID <i>tweet</i> b. name: <i>username author tweet</i> c. author: metadata akun <i>author tweet</i>	a. conversation_id_str : ID percakapan tempat <i>tweet</i> berada b. created_at : waktu pembuatan <i>tweet</i>

Dataset Mentah untuk Pengukuran Echo Chamber		
Nama Dataset	Pilpres 2024	Budget Efficiency
	d. date_created: waktu saat tweet dibuat e. date_inserted_dt: waktu saat <i>tweet</i> masuk ke basis data f. type: jenis konten g. tcode: tipe interaksi <i>tweet</i> (misal: “ <i>reply</i> ”, “ <i>retweet</i> ”, “ <i>original</i> ”) h. content: teks isi <i>tweet</i> i. contentJson: metadata <i>tweet</i> (jumlah <i>reply</i> , <i>retweet</i> , <i>like</i> , <i>view</i> , URL, dsb.) j. num_retweets: jumlah <i>retweet</i> k. lang: bahasa <i>tweet</i> l. in_reply_to_screen_name: <i>username</i> akun yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> ) m.in_reply_to_status_id: ID <i>tweet</i> yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> )	c. favorite_count : jumlah <i>likes</i> d. full_text : isi teks <i>tweet</i> e. id_str : ID <i>tweet</i> f. image_url : URL gambar (jika ada) dalam <i>tweet</i> g. in_reply_to_screen_name : <i>username</i> akun yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> ) h. lang : bahasa yang digunakan i. location : lokasi pengguna (jika tersedia) j. quote_count : jumlah kutipan ( <i>quote tweet</i> ) terhadap <i>tweet</i> k. reply_count : jumlah <i>reply</i> terhadap <i>tweet</i> l. retweet_count : jumlah <i>retweet</i> terhadap <i>tweet</i>
Jumlah Tweet	1000 ( <i>sample</i> yang tersedia, <i>dataset</i> asli memiliki >9 juta <i>tweet</i> )	>10000
Jumlah Kategori Aspek	-	-
Kategori Aspek	-	-

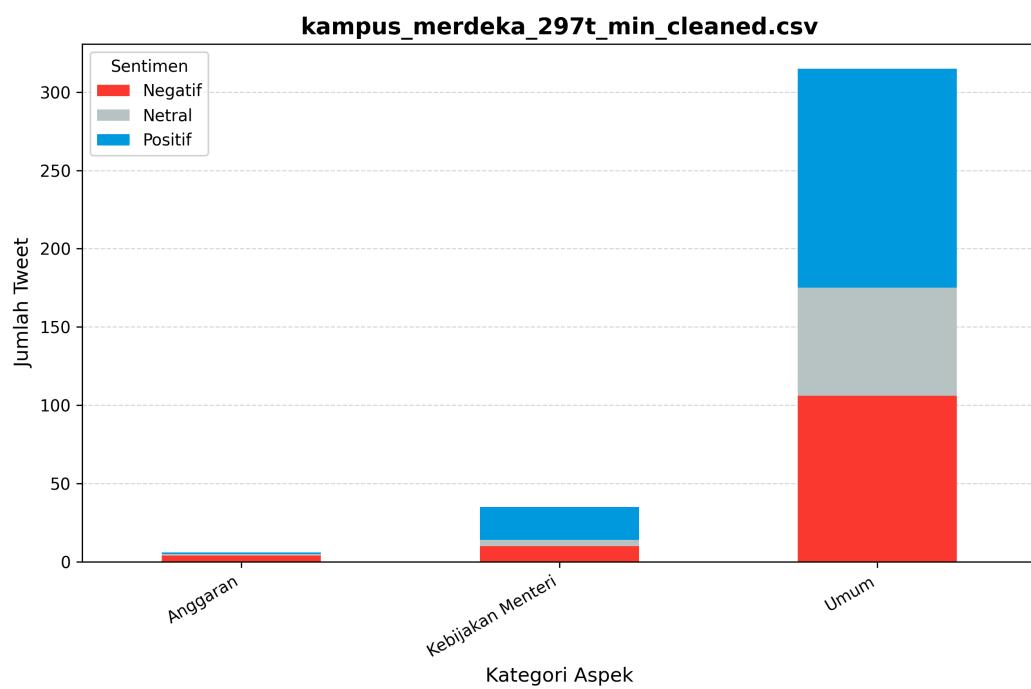
## Lampiran B. Deskripsi Dataset Final

### B.1 Deskripsi Dataset Final untuk Eksperimen Konfigurasi ASBA

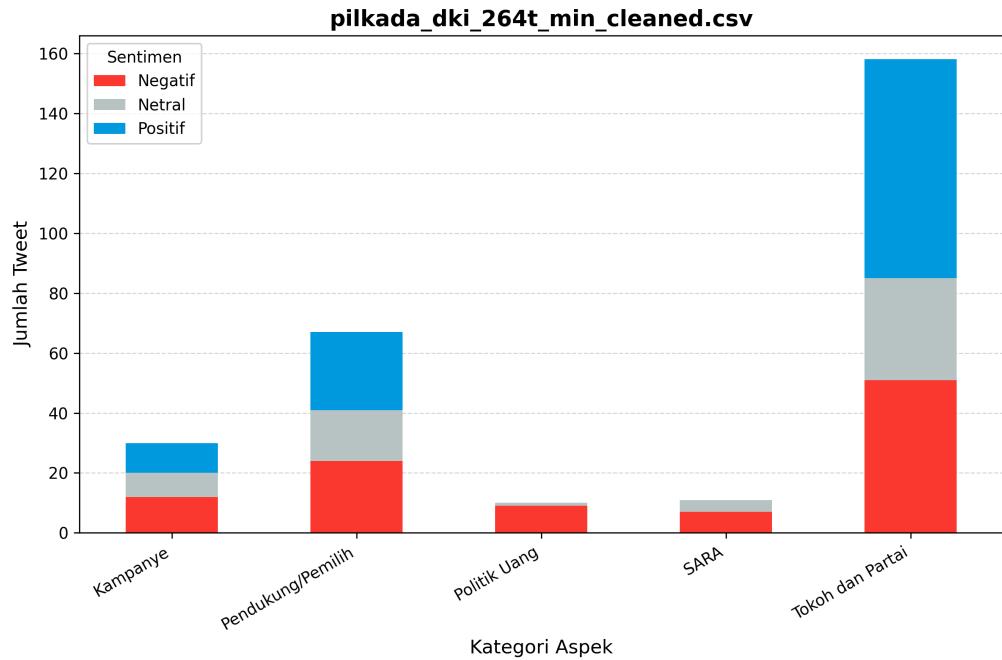
Tabel B.1 Deskripsi Dataset Final untuk Eksperimen Konfigurasi ASBA

Dataset Final untuk Eksperimen Konfigurasi ASBA		
Nama Dataset	Kampus Merdeka	Pilkada DKI
Kolom	a. id: <i>identifier tweet</i> b. text: teks <i>tweet</i> c. aspect_term_0: ekspresi aspek hasil anotasi <i>annotator 1</i> . d. aspect_category_0: kategori aspek hasil anotasi <i>annotator 1</i> . e. sentiment_0: kelas sentimen (contoh: “Positif”) hasil anotasi <i>annotator 1</i> . f. soft_label_0: <i>soft label</i> (contoh: “[0, 0.3, 0.7]”) sentimen hasil anotasi <i>annotator 1</i> . g. hard_label_0: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen hasil anotasi <i>annotator 1</i> . h. aspect_term_1: ekspresi aspek hasil anotasi <i>annotator 2</i> . i. aspect_category_1: kategori aspek hasil anotasi <i>annotator 2</i> . j. sentiment_1: kelas sentimen (contoh: “Positif”) hasil anotasi <i>annotator 2</i> . k. soft_label_1: <i>soft label</i> (contoh: “[0, 0.3, 0.7]”) sentimen hasil anotasi <i>annotator 2</i> . l. hard_label_1: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen hasil anotasi <i>annotator 2</i> . m. aspect_category_final: kategori aspek final hasil resolusi konflik. n. soft_label_final: <i>soft label</i> sentimen final hasil resolusi konflik. o. hard_label_final: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen final hasil resolusi konflik	a. id: <i>identifier tweet</i> b. text: teks <i>tweet</i> c. aspect_term_0: ekspresi aspek hasil anotasi <i>annotator 1</i> . d. aspect_category_0: kategori aspek hasil anotasi <i>annotator 1</i> . e. sentiment_0: kelas sentimen (contoh: “Positif”) hasil anotasi <i>annotator 1</i> . f. soft_label_0: <i>soft label</i> (contoh: “[0, 0.3, 0.7]”) sentimen hasil anotasi <i>annotator 1</i> . g. hard_label_0: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen hasil anotasi <i>annotator 1</i> . h. aspect_term_1: ekspresi aspek hasil anotasi <i>annotator 2</i> . i. aspect_category_1: kategori aspek hasil anotasi <i>annotator 2</i> . j. sentiment_1: kelas sentimen (contoh: “Positif”) hasil anotasi <i>annotator 2</i> . k. soft_label_1: <i>soft label</i> (contoh: “[0, 0.3, 0.7]”) sentimen hasil anotasi <i>annotator 2</i> . l. hard_label_1: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen hasil anotasi <i>annotator 2</i> . m. aspect_category_final: kategori aspek final hasil resolusi konflik. n. soft_label_final: <i>soft label</i> sentimen final hasil resolusi konflik. o. hard_label_final: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen final hasil resolusi konflik
Jumlah Tweet	297	264
Jumlah Kategori Aspek	3	5
Kategori Aspek	a. Anggaran: berkaitan dengan alokasi dan penggunaan, anggaran dana, atau keuangan kampus merdeka. b. Kebijakan Menteri: berkaitan dengan peran atau kebijakan Menteri	a. Politik Uang: membahas politik uang pada pemilihan umum DKI Jakarta tahun 2017. b. Pendukung/Pemilih: berkaitan dengan pendukung atau pemilih

Dataset Final untuk Eksperimen Konfigurasi ASBA		
Nama Dataset	Kampus Merdeka	Pilkada DKI
	<p>Pendidikan dan Kebudayaan Nadiem Makarim dalam komteks kampus merdeka.</p> <p>c. Umum: opini terhadap Kampus Mardeka secara umum yang mencakup berbagai program atau inisiatif seperti pertukaran pelajar, PMM, MSIB, Magang Merdeka, Kurikulum Merdeka, atau IISMA.</p>	<p>termasuk kelompok pendukung atau basis pemilih pada pemilihan umum DKI Jakarta tahun 2017.</p> <p>c. SARA: membahas isu SARA (Suku, Agama, Ras, dan Antargolongan) pada pemilihan umum DKI Jakarta tahun 2017.</p> <p>d. Kampanye: mengenai kampanye yang digunakan termasuk metode dan pendekatan yang diambil oleh calon pada pemilihan umum DKI Jakarta tahun 2017.</p> <p>e. Tokoh dan Partai: membahas nama tokoh politik (contohnya: Ahok, Anies, SBY, AHY, Hary Tanoe) atau nama partai politik (contohnya: Gerindra, PDIP, Golkar).</p>



Gambar B.1 Distribusi Aspek untuk *Dataset* Kampus Merdeka



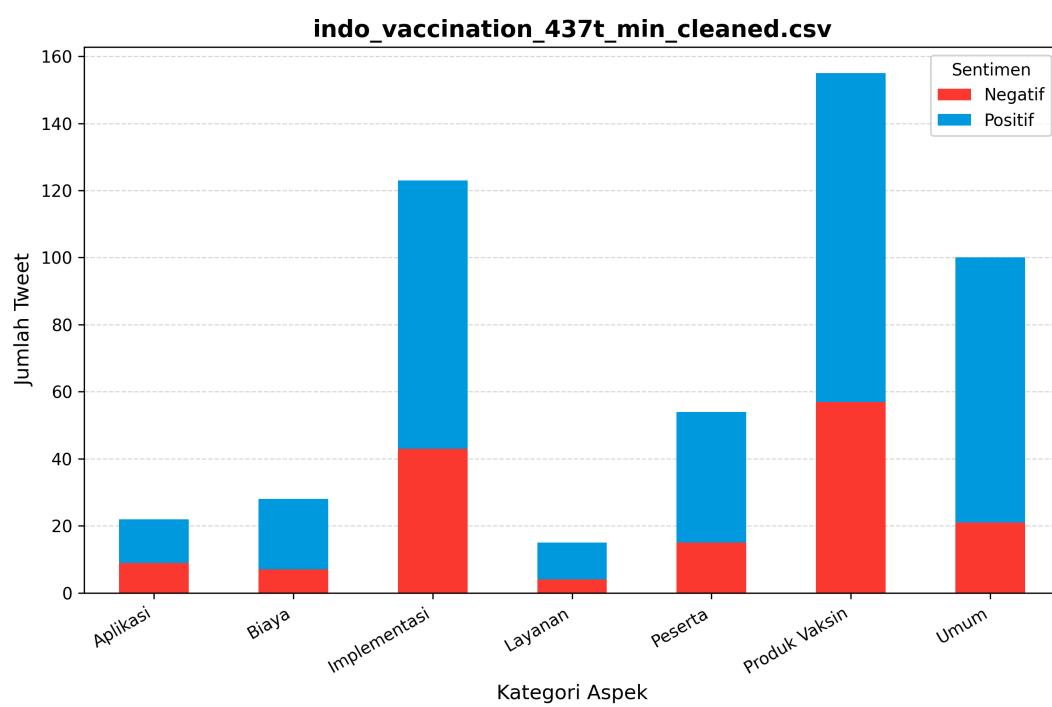
Gambar B.2 Distribusi Aspek untuk *Dataset Pilkada DKI*

## B.2 Deskripsi *Dataset Final* untuk Pengujian ASBA

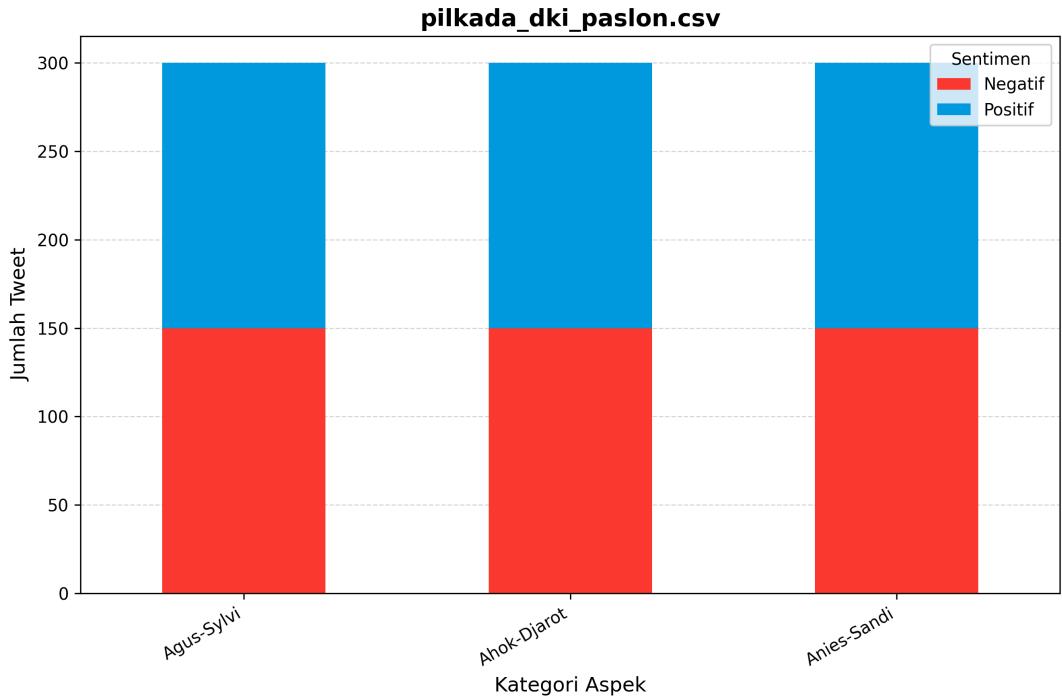
Tabel B.2 Deskripsi *Dataset Final* untuk Pengujian ASBA

Dataset Final untuk Pengujian ASBA		
Nama Dataset	Indo Vaccination	Pilkada DKI dengan Aspek Paslon
Kolom	a. id: <i>identifier tweet</i> b. text: teks <i>tweet</i> c. aspect_category_final: kategori aspek final. d. aspect_sentiment: kelas sentimen (contoh: “Positive”). e. hard_label_final: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen final.	a. id: <i>identifier tweet</i> b. text: teks <i>tweet</i> c. sentiment: kelas sentimen (contoh: “positive”). d. hard_label_final: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen final.
Jumlah Tweet	437	900
Jumlah Kategori Aspek	7	3

Dataset Final untuk Pengujian ASBA		
Nama Dataset	Indo Vaccination	Pilkada DKI dengan Aspek Paslon
Kategori Aspek	<ul style="list-style-type: none"> <li>a. Aplikasi: terkait aplikasi atau platform digital untuk pendaftaran verifikasi atau pemantauan vaksinasi.</li> <li>b. Biaya: terkait biaya harga pembayaran subsidi atau beban ekonomi vaksinasi.</li> <li>c. Implementasi: tentang penerapan vaksinasi di lapangan, distribusi, atau pelaksanaan.</li> <li>d. Peserta: tentang individu atau kelompok yang menjadi target atau penerima vaksinasi.</li> <li>e. Layanan: tentang kualitas layanan, antrian, jadwal, atau penanganan vaksinasi di lokasi.</li> <li>f. Produk Vaksin: membahas merek vaksin seperti pfizer, astrazeneca, dan lain lain, jenis vaksin ketersediaan produk vaksin atau perbandingan antarproduk.</li> <li>g. Umum: relevan dengan topik besar vaksinasi tetapi tidak cocok dengan aspek lain.</li> </ul>	<ul style="list-style-type: none"> <li>a. Agus-Sylvi: mengenai pasangan calon gubernur dan wakil gubernur nomor urut satu Agus Harimurti Yudhoyono dan Sylviana Murni.</li> <li>b. Ahok-Djarot: mengenai pasangan calon gubernur dan wakil gubernur nomor urut dua Basuki Tjahaja Purnama (Ahok) dan Djarot Saiful Hidayat.</li> <li>c. Anies-Sandi: mengenai pasangan calon gubernur dan wakil gubernur nomor urut tiga Anies Baswedan dan Sandiaga Uno.</li> </ul>



Gambar B.3 Distribusi Aspek untuk *Dataset Indo Vaccination*



Gambar B.4 Distribusi Aspek untuk *Dataset* Pilkada DKI dengan Aspek Paslon

### B.3 Deskripsi *Dataset* Final untuk Pengujian Pengukuran *Echo Chamber*

Tabel B.3 Deskripsi *Dataset* Final untuk Pengukuran *Echo Chamber*

Dataset Final untuk Pengukuran <i>Echo Chamber</i>			
Nama Dataset	Indo Vaccination	Budget Efficiency	Pilpres 2024
Kolom	a. id: <i>identifier tweet</i> b. text: teks <i>tweet</i> c. aspect_category_final: kategori aspek final. d. aspect_sentiment: kelas sentimen (contoh: “Positive”). e. hard_label_final: <i>hard label</i> (contoh: “[0, 0, 1]”) sentimen final. f. community: komunitas hasil deteksi komunitas Louvain dengan interaksi <i>mention</i> dan <i>retweet</i>	a. id : <i>identifier tweet</i> b. text : teks <i>tweet</i> c. in_reply_to_screen_name : <i>username</i> akun yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> )	a. id : <i>identifier tweet</i> b. text : teks <i>tweet</i> c. in_reply_to_screen_name : <i>username</i> akun yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> ) d. in_reply_to_status_id: ID <i>tweet</i> yang dibalas (jika <i>tweet</i> merupakan <i>reply</i> ) e. name: <i>username author tweet</i>
Jumlah Tweet	437	745	970
Jumlah	7	4	3

Dataset Final untuk Pengukuran Echo Chamber			
Nama Dataset	<i>Indo Vaccination</i>	<i>Budget Efficiency</i>	Pilpres 2024
Kategori Aspek			
Kategori Aspek	<p>a. Aplikasi: terkait aplikasi atau platform digital untuk pendaftaran verifikasi atau pemantauan vaksinasi.</p> <p>b. Biaya: terkait biaya harga pembayaran subsidi atau beban ekonomi vaksinasi.</p> <p>c. Implementasi: tentang penerapan vaksinasi di lapangan, distribusi, atau pelaksanaan.</p> <p>d. Peserta: tentang individu atau kelompok yang menjadi target atau penerima vaksinasi.</p> <p>e. Layanan: tentang kualitas layanan, antrian, jadwal, atau penanganan vaksinasi di lokasi.</p> <p>f. Produk Vaksin: membahas merek vaksin seperti pfizer, astrazeneca, dan lain lain, jenis vaksin ketersediaan produk vaksin atau perbandingan antarproduk.</p> <p>g. Umum: relevan dengan topik besar vaksinasi tetapi tidak cocok dengan aspek lain.</p>	<p>a. Kualitas Pengelolaan: mengenai transparansi dan akuntabilitas dalam pengelolaan anggaran.</p> <p>b. Dampak: mengenai dampak dari efisiensi anggaran terhadap kualitas pendidikan, layanan kesehatan, dan PHK (pemutusan hubungan kerja) massal.</p> <p>c. Prioritas Alokasi: mengenai prioritas alokasi anggaran seperti program prioritas dan antara, MBG (makan bergizi gratis), dan sektor-sektor strategis lain.</p> <p>d. Umum: opini umum tentang efisiensi anggaran yang tidak cocok dengan aspek lain.</p>	<p>a. Anies: mengenai Anies Baswedan dan Muhammin Iskandar (Cak Imin) sebagai pasangan calon presiden dan wakil presiden nomor urut satu yang didukung partai PKB, mungkin mengandung <i>tagline</i> AMuhaimin, IskandarN, atau AMIN</p> <p>b. Prabowo: mengenai Prabowo Subianto dan Gibran Rakabuming sebagai pasangan calon presiden dan wakil presiden nomor urut dua yang didukung partai Gerindra, mungkin mengandung <i>tagline</i> PrabowoGibran</p> <p>c. Ganjar: mengenai Ganjar Pranowo dan Mahfud MD sebagai pasangan calon presiden dan wakil presiden nomor urut tiga yang didukung partai PDIP, mungkin mengandung <i>tagline</i> Ganjar-Mahfud</p>
Jenis Interaksi	<i>Mention</i> dan <i>retweet</i>	<i>Reply</i>	<i>Reply</i>

## Lampiran C. Isu Kualitas Teks pada Dataset

Tabel C.1 Isu Kualitas Teks pada Dataset

Permasalahan	Teks Tweet	Dataset
<i>Tweet</i> dengan huruf besar dan kecil tidak sesuai aturan	“kampus merdeka super <b>LELET</b> ”	Kampus Merdeka
	“Saya bilang <b>PASTI MENANG</b> karena bentuk keyakinan akan keberhasilan goal goal <b>LAIN</b> dari <b>IMPROVISASI</b> politik Pak SBY & Pak AHY,di Pilkada DKI.”	Pilkada DKI
<i>Tweet</i> dengan mention <i>username</i>	“@gitaputrid Kampus merdeka oke sih”	Kampus Merdeka
	“bagi2 sembako gratis tim @basuki_btp itu melanggar Pasal 73 Ayat 1 UU No.10 Thn 2016”	Pilkada DKI
	“@jokowi Alhamdulillah saya sudah vaksin pak tinggal nyari kerjaan buat biaya kehidupan sehari hari”	<i>Indo Vaccination (hasil hydrate)</i>
<i>Tweet</i> dengan <i>slang</i>	“@mfs_ub kalo udah kuliah mah kampus merdeka kali ya. kalo kata adekku mah <b>b</b> aja haha	Kampus Merdeka
	“wah si AHY 70% sentimennya negatif. <b>Wkwkwk</b> ”	Pilkada DKI
<i>Tweet</i> dengan singkatan dan ejaan yang salah	“Kampus merdeka malah <b>jd</b> kampus kapitalis <b>yha</b> ? Dilepas gitu aja bikin kebijakan biar neraca <b>ga</b> merah”	Kampus Merdeka
	“AHY siap siap jadi vocalis BAND rock <b>yaaaa</b> , lompat lompat mulu bang kerjaannya #PilkadaDKI #QuickCount”	Pilkada DKI
<i>Tweet</i> dengan bahasa asing	“korban kampus merdeka pasti <b>relate</b> ”	Kampus Merdeka
	"Salam "" <b>Wallahulmuwafiq IlaAqwamithoriq</b> @sandiuno mencerminkan beliau NU sejak itu saya jatuh hati meski waktu itu saya #AHY @AgusYudhoyono"	Pilkada DKI
<i>Tweet</i> dengan <i>hashtag</i> berupa ungkapan	“Niatnya jadi kampus merdeka eh dia malah jadi kampus romusha <b>#hadeh</b> ”	Kampus Merdeka
<i>Tweet</i> dengan <i>hashtag</i> berupa kata kerja	“Cakeepppp....#bersatunya pendukung #AniesSandi dan #agussilvy”	Pilkada DKI
<i>Tweet</i> dengan <i>hashtag</i> berupa tokoh target opini	“#ahy terbaik”	Pilkada DKI
<i>Tweet</i> dengan <i>hashtag</i> berupa topik	“Setelah negara yang transaksional dengan rakyat kampus ikut-ikutan transaksional dengan	Kampus Merdeka

Permasalahan	Teks <i>Tweet</i>	<i>Dataset</i>
	mahasiswa. Hidup #KampusMerdeka”	
	“Dilarang Pilih Cagub Nonmuslim? Habib: Pilih Djarot! #PilkadaDKI2017”	Pilkada DKI
<i>Tweet</i> dengan tautan	“Semoga dgn menteri yg baru program Kampus Merdeka dihapuskan... Amiinn... <a href="https://t.co/SWUxv71ir5">https://t.co/SWUxv71ir5</a> ”	Kampus Merdeka
<i>Tweet</i> dengan penanda emoji	“Pingin lihat mas #AHY debat <FACE WITH STUCK-OUT TOUNGES>”	Pilkada DKI
<i>Tweet</i> dengan emoji	“@republikaonline 6 juta vaksin sinovac datang lagi 😱😱😱 apakah Rumor jema'ah haji indonesia akan ditolak lagi karna vaksin trsbt 😱😱”	<i>Indo Vaccination (hasil hydrate)</i>

## Lampiran D. Keluaran Detail Konsensus dan Graf Interaksi dengan Komunitas

### D.1 Keluaran Detail Konsensus Dalam Komunitas

Tabel D.1 Keluaran Detail Konsensus Dalam Komunitas untuk *Dataset Budget Efficiency*

Komunitas	Ukuran	Konsensus Dalam Komunitas	Konsensus Aspek			
			Dampak	Kualitas Pengelolaan	Prioritas Alokasi	Umum
0	114	0,1616	0,1299	0,3606	0,0319	0,1241
1	52	0,1012	0,1039	0,2055	0,0725	0,0230
2	50	0,1963	0,0426	0,0000	0,4199	0,1264
3	60	0,1419	0,2040	0,2421	0,1131	0,0085
4	187	0,9872	0,0000	0,0000	0,9872	0,0000

Tabel D.2 Keluaran Detail Konsensus Dalam Komunitas untuk *Dataset Indo Vaccination*

Komunitas	Ukuran	Konsensus Dalam Komunitas	Konsensus Aspek						
			Applikasi	Biaya	Implementasi	Layanan	Peserta	Produk Vaksin	Umum
156	216	0,0574	0,0071	0,0193	0,0286	0,0084	0,0658	0,0903	0,1822
31	101	0,0606	0,0000	0,0273	0,1342	0,0204	0,0213	0,0695	0,1516
38	46	0,0460	0,0000	0,0007	0,0557	0,0422	0,0756	0,0227	0,1250
11	44	0,0637	0,0054	0,0187	0,0375	0,0000	0,0622	0,0389	0,2195
49	46	0,0491	0,0340	0,0285	0,0894	0,0203	0,0863	0,0062	0,0790

Tabel D.3 Keluaran Detail Konsensus Dalam Komunitas untuk *Dataset Indo Vaccination* dengan *Gold Label*

Komunitas	Ukuran	Konsensus Dalam Komunitas	Konsensus Aspek						
			Applikasi	Biaya	Implementasi	Layanan	Peserta	Produk Vaksin	Umum
156	205	0,0635	0,0099	0,0277	0,0748	0,0000	0,0386	0,2057	0,0880
31	96	0,0394	0,0080	0,0126	0,0843	0,0000	0,0080	0,0620	0,1006
38	46	0,0380	0,0076	0,0076	0,0275	0,0374	0,0306	0,0392	0,1160
11	44	0,0416	0,0000	0,0000	0,0293	0,0118	0,0247	0,1739	0,0513
49	46	0,0494	0,0279	0,0076	0,0841	0,0000	0,0392	0,0275	0,1105

Tabel D.4 Keluaran Detail Konsensus Antarkomunitas untuk *Dataset Pilpres 2024*

Komunitas	Ukuran	Konsensus Dalam Komunitas	Konsensus Aspek		
			Anies	Ganjar	Prabowo
0	5	0,9990	0,9990	0,0000	0,0000
2	8	0,2504	0,5008	0,0000	0,0000
5	15	0,6325	0,6325	0,0000	0,0000
7	7	0,1332	0,2664	0,0000	0,0000
19	9	0,5808	0,5808	0,0000	0,0000
20	19	0,2760	0,4955	0,0000	0,0566
22	14	0,1518	0,2401	0,0000	0,0636
31	10	0,0627	0,0585	0,0000	0,1297
36	5	0,3330	0,3330	0,0000	0,0000
39	7	0,1142	0,0317	0,0000	0,1968
42	6	0,4273	0,4273	0,0000	0,0000
46	5	0,9954	0,9954	0,0000	0,0000
48	9	0,1701	0,2239	0,0000	0,1163
49	13	0,5948	0,5948	0,0000	0,0000
53	9	0,2553	0,5105	0,0000	0,0000
57	5	0,3202	0,3202	0,0000	0,0000
63	13	0,1671	0,2870	0,0000	0,0472
65	8	0,2360	0,0736	0,0000	0,3984
69	6	0,6658	0,6658	0,0000	0,0000
71	5	0,5693	0,1424	0,0000	0,9962
105	6	0,0556	0,0000	0,0000	0,1111
126	11	0,0000	0,0000	0,0000	0,0000
136	6	0,4281	0,0000	0,4281	0,0000
148	15	0,9998	0,0000	0,9998	0,0000
452	6	0,4279	0,0000	0,0000	0,4279
476	5	0,2984	0,0000	0,0000	0,5969
482	5	0,3626	0,0000	0,0000	0,3626
499	5	0,6000	0,0000	0,0000	0,6000
510	11	0,1762	0,0000	0,0000	0,1762
0	5	0,9990	0,9990	0,0000	0,0000
2	8	0,2504	0,5008	0,0000	0,0000

## D.2 Keluaran Detail Konsensus Antarkomunitas

Tabel D.5 Keluaran Detail Konsensus Antarkomunitas untuk *Dataset Budget Efficiency*

Komunitas	0	1	2	3	4	5	6	7	8	9
0	<b>0,162</b>	0,009	0,069	0,136	0,018	0,112	0,080	0,022	0,008	0,013
1		<b>0,101</b>	0,006	0,012	0,002	0,007	0,005	0,102	0,102	0,097
2			<b>0,196</b>	0,076	0,199	0,038	0,109	0,004	0,010	0,008
3				<b>0,142</b>	0,054	0,126	0,118	0,022	0,012	0,016
4					<b>0,987</b>	0,005	0,095	0,001	0,007	0,005
5						<b>0,163</b>	0,127	0,012	0,005	0,008
6							<b>0,153</b>	0,003	0,006	0,006
7								<b>0,103</b>	0,099	0,100
8									<b>0,098</b>	0,103
9										<b>0,095</b>

Teks **tebal** menunjukkan konsensus dalam komunitas.

Tabel D.6 Keluaran Detail Konsensus Antarkomunitas untuk *Dataset Indo Vaccination*

Komunitas	11	31	38	49	156
11	<b>0,0637</b>	0,0516	0,0505	0,0464	0,0515
31		<b>0,0606</b>	0,0483	0,0503	0,0505
38			<b>0,0460</b>	0,0498	0,0473
49				<b>0,0491</b>	0,0431
156					<b>0,0574</b>

Teks **tebal** menunjukkan konsensus dalam komunitas.

Tabel D.7 Keluaran Detail Konsensus Antarkomunitas untuk *Dataset Indo Vaccination* dengan *Gold Label*

Komunitas	11	31	38	49	156
11	<b>0,0416</b>	0,0359	0,0314	0,0274	0,0275
31		<b>0,0394</b>	0,0353	0,0386	0,0354
38			<b>0,0380</b>	0,0414	0,0395
49				<b>0,0494</b>	0,0466
156					<b>0,0635</b>

Teks **tebal** menunjukkan konsensus dalam komunitas.

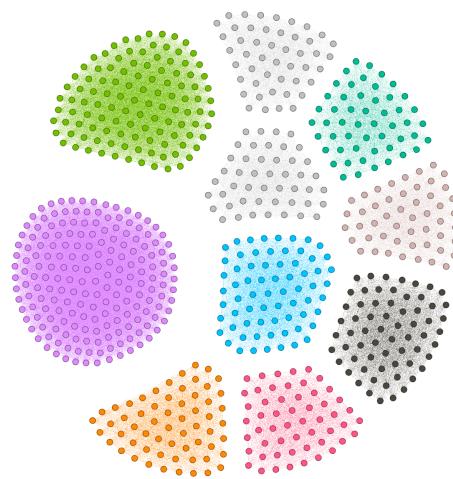
Tabel D.8 Keluaran Detail Konsensus Antarkomunitas untuk *Dataset Pilpres 2024*

Komunitas	0	2	5	7	19	20	22	31	36	39
0	<b>0,999</b>	0,354	0,756	0,027	0,773	0,343	0,184	0,067	0,599	0,084
2		<b>0,250</b>	0,296	0,118	0,291	0,266	0,192	0,077	0,221	0,080
5			<b>0,633</b>	0,094	0,610	0,274	0,182	0,050	0,446	0,086
7				<b>0,133</b>	0,057	0,059	0,120	0,034	0,013	0,056
19					<b>0,581</b>	0,277	0,158	0,056	0,484	0,074
20						<b>0,276</b>	0,178	0,078	0,229	0,146
22							<b>0,152</b>	0,077	0,129	0,094
31								<b>0,063</b>	0,059	0,055
36									<b>0,333</b>	0,067
39										<b>0,114</b>

Teks **tebal** menunjukkan konsensus dalam komunitas. Tabel hanya menampilkan 10 dari 29 komunitas.

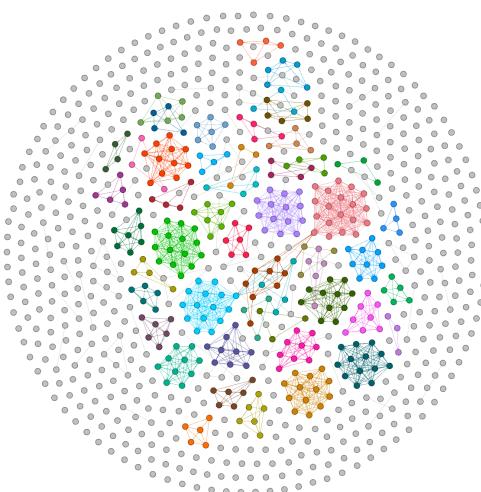
karena keterbatasan tempat.

### D.3 Keluaran Graf Interaksi dengan Komunitas



*Warna berbeda menunjukkan komunitas yang berbeda.*

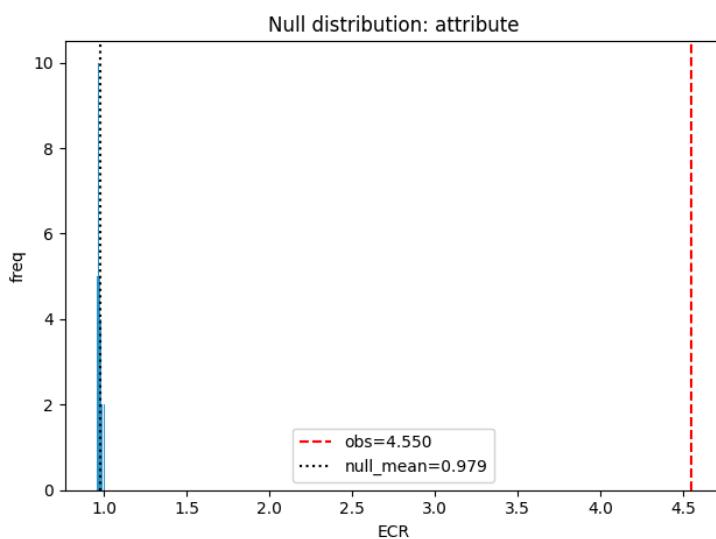
Gambar D.1 Keluaran Graf Interaksi dengan Komunitas untuk *Dataset Budget Efficiency*



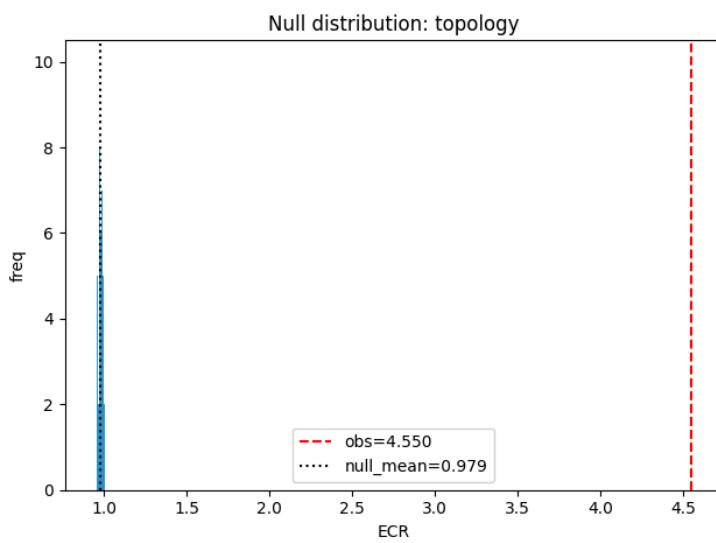
*Warna berbeda menunjukkan komunitas yang berbeda. Beberapa komunitas yang terlalu kecil diwaranai abu-abu.*

Gambar D.2 Keluaran Graf Interaksi dengan Komunitas untuk *Dataset Pilpres 2024*

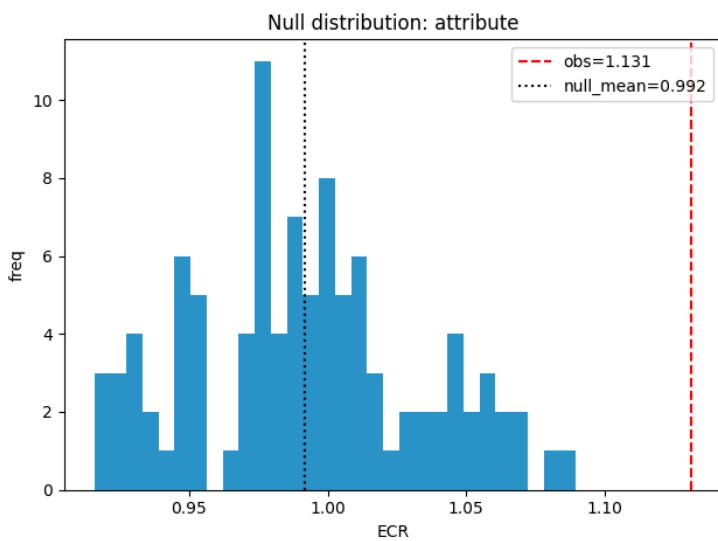
## Lampiran E. Graf Histogram Distribusi Null pada Uji Hipotesis



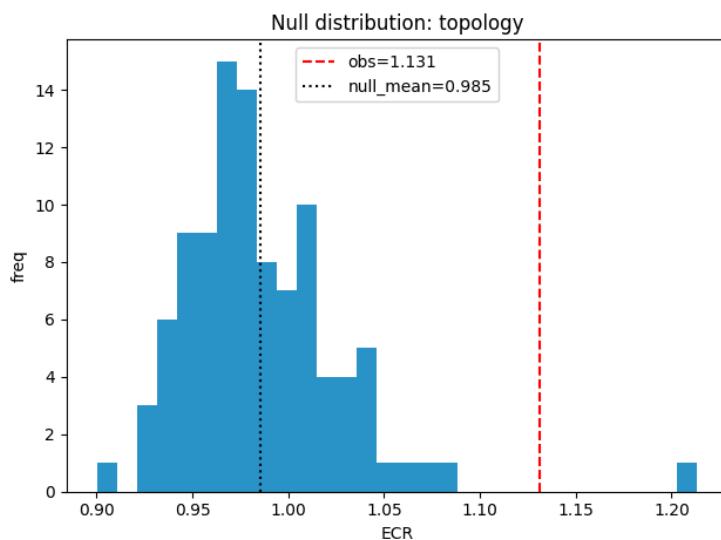
Gambar E.1 Graf Histogram Distribusi Null Atribut pada Uji Hipotesis untuk *Dataset Budget Efficiency*



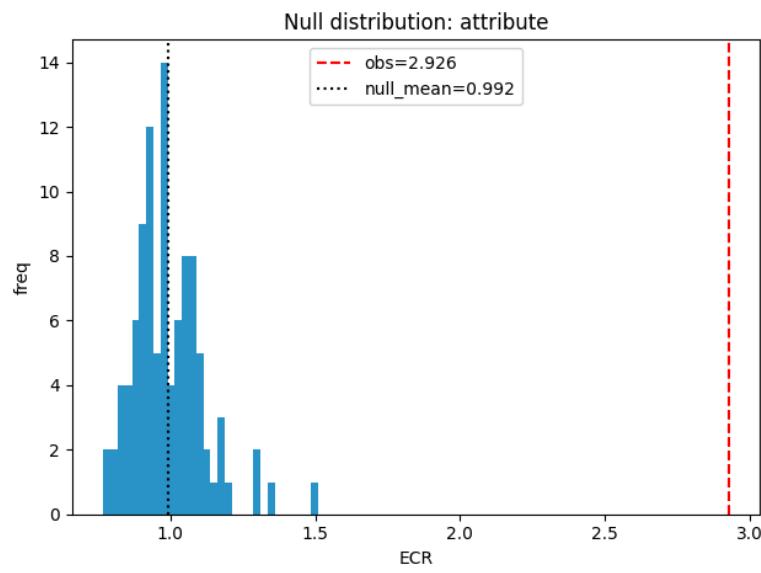
Gambar E.2 Graf Histogram Distribusi Null Topoloi pada Uji Hipotesis untuk *Dataset Budget Efficiency*



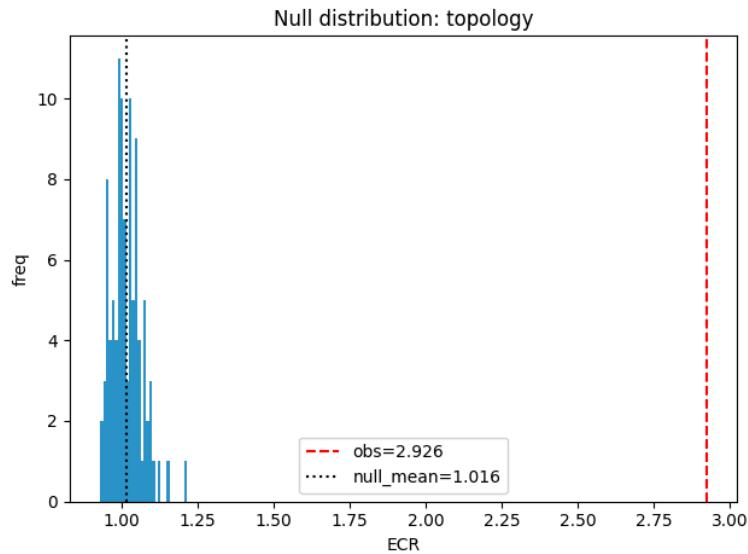
Gambar E.3 Graf Histogram Distribusi Null Atribut pada Uji Hipotesis untuk *Dataset Indo Vaccination*



Gambar E.4 Graf Histogram Distribusi Null Topologi pada Uji Hipotesis untuk *Dataset Indo Vaccination*



Gambar E.5 Graf Histogram Distribusi Null Atribut pada Uji Hipotesis untuk *Dataset Pilpres 2024*



Gambar E.6 Graf Histogram Distribusi Null Topologi pada Uji Hipotesis untuk *Dataset Pilpres 2024*