

ANALISIS AMBIGUITAS SENTIMEN DAN KESEPAKATAN ANOTATOR PADA ISU GEOPOLITIK DI MEDIA SOSIAL

PUSPITA DEWI SURYANINGTYAS



PROGRAM STUDI MATEMATIKA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Ambiguitas Sentimen dan Kesepakatan Anotator pada Isu Geopolitik di Media Sosial” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Puspita Dewi Suryaningtyas
G5401221061

ABSTRAK

PUSPITA DEWI SURYANINGTYAS. Analisis Ambiguitas Sentimen dan Kesepakatan Anotator pada Isu Geopolitik di Media Sosial. Dibimbing oleh ELIS KHATIZAH dan SRI NURDIATI.

Komentar konflik geopolitik memuat opini, emosi, humor, ironi, dan interpretasi kontekstual yang tidak selalu dapat direpresentasikan melalui satu label sentimen. Penelitian ini menganalisis ambiguitas sentimen pada 2.606 komentar TikTok terkait konflik Iran dengan Amerika Serikat/Israel, dianotasi oleh sepuluh anotator Gen Z dan Non-Gen Z dengan label positif, netral, dan negatif. Shannon entropi digunakan untuk mengukur ambiguitas, Jensen-Shannon *divergence* digunakan untuk membandingkan distribusi entropi Gen Z dan Non-Gen Z, dan *Cohen's Kappa* dengan graf berbobot divisualisasikan untuk menganalisis *agreement* antaranotator. Distribusi sentimen hasil anotasi oleh 10 anotator memiliki *mean* entropi 0,768 dengan 56,60% komentar terkategori *highly ambiguous*. Hasil Jensen-Shannon *divergence* dan visualisasi graf menunjukkan bahwa perbedaan persepsi antargenerasi bersifat spesifik pada komentar tertentu, dengan struktur kesepakatan yang lebih dipengaruhi oleh karakteristik individual anotator, sedangkan *random forest* menunjukkan keterbatasan prediksi ambiguitas dari fitur teks. Hal ini mengonfirmasi relatif tingginya ambiguitas sentimen pada data yang digunakan sehingga menjadi bahan pertimbangan untuk analisis sentimen selanjutnya.

Kata kunci: ambiguitas, entropi, Jensen-Shannon *divergence*, sentimen geopolitik, tiktok.

ABSTRACT

PUSPITA DEWI SURYANINGTYAS. Analysis of Sentiment Ambiguity and Annotator Agreement on Geopolitical Issues on Social Media. Supervised by ELIS KHATIZAH and SRI NURDIATI.

Geopolitical conflict comments contain opinions, emotions, humor, irony, and contextual interpretations that cannot always be represented by a single sentiment label. This study analyzes sentiment ambiguity in 2,606 TikTok comments related to the Iran and United States/Israel conflict, annotated by ten Gen Z and Non-Gen Z annotators with positive, neutral, and negative labels. Shannon entropy is used to measure ambiguity, Jensen-Shannon divergence is used to compare the entropy distributions of Gen Z and Non-Gen Z, and Cohen's Kappa with a weighted graph is visualized to analyze inter-annotator agreement. The sentiment distribution annotated by the 10 annotators has a mean entropy of 0.768, with 56.60% of the comments categorized as highly ambiguous. The results of the Jensen-Shannon divergence and graph visualization show that intergenerational perception differences are specific to certain comments, with the agreement structure being more influenced by the individual characteristics of the annotators, whereas random forest shows the limitation of ambiguity prediction from text

features. This confirms the relatively high sentiment ambiguity in the data used, thus serving as a consideration for further sentiment analysis.

Keywords: ambiguity, entropy, geopolitical sentiment, Jensen-Shannon divergence, tiktok

Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

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

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

ANALISIS AMBIGUITAS SENTIMEN DAN KESEPAKATAN ANOTATOR PADA ISU GEOPOLITIK DI MEDIA SOSIAL

PUSPITA DEWI SURYANINGTYAS

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana Matematika pada
Program Studi Matematika

**PROGAM STUDI MATEMATIKA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:

1 Mohamad Khoirun Najib, S.Si., M.Mat.



Judul Skripsi : Analisis Ambiguitas Sentimen dan Kesepakatan Anotator pada
Isu Geopolitik di Media Sosial

Nama : Puspita Dewi Suryaningtyas
NIM : G5401221061

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Elis Khatizah, S.Si., M.Si., Ph.D

Pembimbing 2:
Prof. Dr. Ir. Sri Nurdianti, M.Sc.

Diketahui oleh

Ketua Program Studi:
Dr. Donny Citra Lesmana, S.Si., M.Fin.Math.
NIP.197902272005011001

Tanggal Ujian:
(18 Juni 2026)

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Maret 2026 sampai bulan Juli 2026 ini ialah berhasil diselesaikan, dengan judul “Analisis Ambiguitas Sentimen dan Kesepakatan Anotator pada Isu Geopolitik di Media Sosial”. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Ibu Yuni Wahyuni dan Alm Bapak Taryono selaku orang tua penulis, Melati selaku adik penulis yang selalu memberikan dukungan, doa, dan cinta kasih sayangnya selama penulis hidup sampai saat ini.
2. Ibu Elis Khatizah, S.Si. M.Si. Ph.D selaku dosen pembimbing 1 dan Ibu Prof. Dr. Ir. Sri Nurdianti, M.Sc. selaku dosen pembimbing 2 yang telah selalu sabar memberikan bimbingan, motivasi serta arahan dan saran selama proses penyusunan skripsi.
3. Penyelenggara dan panitia Seminar Nasional SSMI 2026 yang menerima skripsi penulis dan mendukung penulis untuk lulus tepat waktu. Serta seluruh dosen, staf dan tenaga kependidikan Program Studi Matematika yang telah memberikan ilmu dan wawasan, serta memfasilitasi selama masa perkuliahan
4. Muhammad Salman Alfarizi selaku *partner* dan sahabat penulis yang telah menemani, mengarahkan tenaga dan waktu, serta memberi dukungan kepada penulis selama proses penyusunan.
5. Teman-teman penulis sejak awal masuk kuliah, Oce, Feng, Rana, Wawa, Annas, dan Firlan yang selalu ada menemani hari-hari penulis dan selalu ada disaat penulis membutuhkan.
6. Teman terdekat di Program Studi Matematika yang senantiasa menemani hari-hari penulis selama menjalani perkuliahan, Jiel, Aufa, Melisa, Hana, Risma.
7. Teman-teman sepembimbingan, Alin, Zaidan, Abel, dan Ghifari yang selalu saling *support* dan memberikan motivasi untuk terus lanjut menyelesaikan tugas akhir ini.
8. Teman-teman Matematika 59 yang telah berjuang bersama penulis.
9. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan bantuan secara langsung maupun tidak langsung selama proses penyusunan skripsi ini berlangsung.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2026

Puspita Dewi Suryaningtyas



DAFTAR ISI

DAFTAR TABEL	ix
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	ix
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
TINJAUAN PUSTAKA	5
2.1 Analisis Sentimen dan Representasi Probabilistik Label	5
2.2 Shannon Entropi dan Ukuran Ambiguitas	6
2.3 Ukuran Divergensi dan Jensen-Shannon <i>Divergence</i>	7
2.4 <i>Agreement</i> dan <i>Cohen's Kappa</i>	8
2.5 Teori Graf dan <i>Network Analysis</i>	10
2.6 Korelasi Statistik dan Uji Nonparametrik	11
2.7 <i>Random Forest</i> sebagai Analisis Pendukung	12
III METODE	16
3.1 Sumber dan Jenis Data	16
3.2 Tahapan Penelitian	17
IV HASIL DAN PEMBAHASAN	19
4.1 Statistik Entropi dan Kategori Ambiguitas	22
4.2 Uji Signifikansi Entropi Gen Z dan Non-Gen Z	24
4.3 Analisis Jensen-Shannon <i>Divergence</i>	25
4.4 <i>Agreement</i> Graph Annotator	26
4.5 <i>Random Forest</i> sebagai Analisis Pendukung	29
V SIMPULAN DAN SARAN	31
5.1 Simpulan	31
5.2 Saran	31
DAFTAR PUSTAKA	33
LAMPIRAN	36
RIWAYAT HIDUP	43

DAFTAR TABEL

1	Contoh data komentar TikTok terkait konflik Iran dengan Amerika Serikat/Israel	19
2	Contoh hasil anotasi komentar	19
3	Contoh hasil perhitungan distribusi probabilitas	20
4	Hasil uji stabilitas anotasi	21
5	Contoh nilai entropi dan ambiguitas komentar kelompok anotator gabungan	22
6	Statistik dasar nilai entropi berdasarkan kelompok anotator	23
7	Distribusi kategori ambiguitas berdasarkan kelompok	23
8	Hasil uji Mann-Whitney entropi Gen Z dan Non-Gen Z	24
9	Statistik Jensen-Shannon <i>divergence</i>	25
10	<i>Degree centrality agreement graph</i> Gen Z	27
11	<i>Degree centrality agreement graph</i> Non-Gen Z	28
12	<i>Degree centrality agreement graph</i> gabungan	28
13	Contoh hasil tokenisasi komentar	29
14	Pembobotan TF-IDF pada komentar	30
15	Ringkasan metrik <i>random forest</i>	30

DAFTAR GAMBAR

1	Ilustrasi struktur <i>agreement</i> antaranotator dalam: (A) Graf berbobot berdasarkan (B) Matriks ketetanggaannya (Hapsari 2015).	10
2	Ilustrasi cara kerja model <i>random forest</i> (Agustina <i>et al.</i> 2025).	13
3	Tampilan apify TikTok comments scraper	16
4	Diagram alir tahapan penelitian: (A) Tahapan utama penelitian dan (B) Tahapan <i>random forest</i> sebagai analisis pendukung	18
5	Distribusi kategori berdasarkan level ambiguitas	24
6	Distribusi nilai JSD Gen Z vs Non-Gen Z	25
7	<i>Heatmap Cohen's Kappa</i> gabungan	26
8	<i>Agreement graph</i> : (A) Anotator Gen Z dan (B) Anotator Non-Gen Z	27
9	<i>Anotator agreement graph</i> gabungan	28

DAFTAR LAMPIRAN

1	Data komentar pada konten TikTok dengan isu geopolitik	36
2	Hasil anotasi label data dan perhitungan entropi dan <i>agreement</i>	36
3	Distribusi nilai entropi Gen Z	36
4	Distribusi nilai entropi Non-Gen Z	36
5	Distribusi nilai entropi gabungan	37



6	<i>Syntax</i> uji perbedaan Gen z dan Non-Gen Z dengan uji Mann-Whitney	37
7	<i>Syntax</i> untuk perhitungan Jensen Shanon <i>divergence</i>	38
8	<i>Syntax</i> untuk pra-pemrosesan data	39
9	<i>Syntax</i> untuk <i>split data</i>	40
10	<i>Syntax</i> untuk TF-IDF	40
11	<i>Syntax</i> untuk implementasi model <i>random forest</i>	41
12	<i>Classification report random forest</i> Gen Z	41
13	<i>Classification report random forest</i> Non-Gen Z	41
14	<i>Classification report random forest</i> gabungan	42