



# **Analisis Media Sosial Sebagai Upaya Dini Deteksi Potensi Konflik Masyarakat di Dunia Maya**

**Indra Budi**

Pidato Pengukuhan Guru Besar Tetap  
**Bidang Ilmu Komputer**  
Fakultas Ilmu Komputer, Universitas Indonesia  
Depok, 20 Desember 2023

**Analisis Media Sosial Sebagai Upaya Dini Deteksi Potensi Konflik  
Masyarakat di Dunia Maya**

ISBN : 978-623-333-645-1

E-ISBN : 978-623-333-646-8 (PDF)

©Hak Cipta Dilindungi Undang-Undang

Dilarang mengutip, memperbanyak dan menerjemahkan sebagian atau seluruh isi buku ini tanpa ijin tertulis dari penulis dan penerbit.

Cetakan 2023

Diterbitkan pertama kali oleh UI Publishing

Anggota IKAPI & APPTI

Jalan Salemba 4, Jakarta 10430

0818 436 500

E-mail: [uipublishing@ui.ac.id](mailto:uipublishing@ui.ac.id)

**Bismillahirrahmanirrahim**

**Assalamu'alaikum Warahmatullahi Wabarakatuh,**

Salam sejahtera bagi kita semua, semoga Allah SWT selalu melimpahkan hidayah-Nya kepada kita semua

**Yang saya hormati,**

1. Bapak Rektor dan para Wakil Rektor Universitas Indonesia
2. Bapak Ketua dan Anggota Majelis Wali Amanat Universitas Indonesia
3. Bapak Ketua dan Anggota Senat Akademik Universitas Indonesia
4. Bapak Ketua dan Anggota Dewan Guru Besar Universitas Indonesia
5. Para Dekan dan Wakil Dekan di lingkungan Universitas Indonesia
6. Para Direktur dan Wakil Direktur di lingkungan Universitas Indonesia
7. Ketua dan Anggota Senat Akademik Fakultas Ilmu Komputer Universitas Indonesia
8. Ketua dan Anggota Dewan Guru Besar Fakultas Ilmu Komputer Universitas Indonesia
9. Kolega Dosen dan para tenaga kependidikan Fakultas Ilmu Komputer, Universitas Indonesia;
10. Para Guru Besar Tamu, sahabat, kerabat dan seluruh tamu undangan.

Terima kasih kepada Bapak Rektor Universitas Indonesia yang memberikan kesempatan kepada saya untuk menyampaikan pidato pengukuhan Guru Besar yang berjudul **“Analisis Media Sosial Sebagai Upaya Dini Deteksi Potensi Konflik Masyarakat di Dunia Maya”**

## **Perkembangan Internet dan Penggunaan Media Sosial di Indonesia**

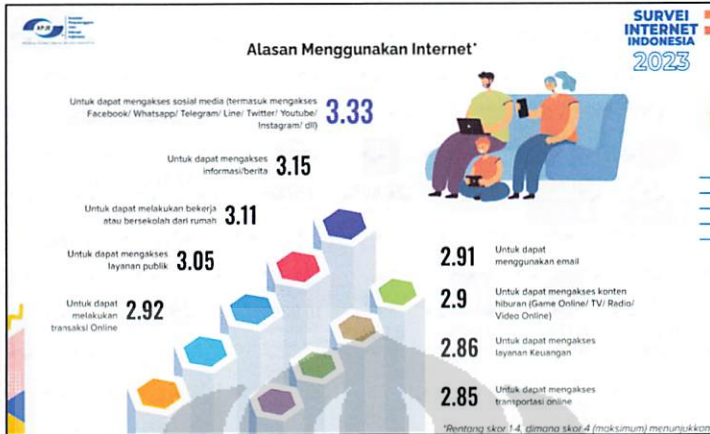
Saat ini, dalam kehidupan kita tidak terlepas dari penggunaan teknologi informasi, terutama teknologi Internet. Teknologi Internet telah mengubah pola dan kebiasaan hidup manusia dalam berbagai aspeknya. Berdasarkan data pada awal tahun 2023, jumlah pengguna Internet telah mencapai 5,16 miliar dari total populasi 8.01 miliar penduduk dunia (We are Social, 2023) yang berarti bahwa 64.4% penduduk dunia atau lebih dari separuhnya sudah menjadi pengguna Internet. Dari sumber yang sama, di Indonesia tercatat jumlah pengguna Internet adalah 212,9 juta dengan total penduduk Indonesia 276.4 juta yang berarti bahwa 77% penduduk Indonesia sudah menggunakan Internet. Sehingga secara persentase pengguna Internet di Indonesia lebih besar dibandingkan dengan yang ada di dunia. Hal ini juga dikonfirmasi oleh hasil survey Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2023 bahwa pengguna Internet di Indonesia mencapai 215 juta orang atau sebesar 78.19% dari total populasi penduduk Indonesia<sup>1</sup>.

Banyak hal yang dapat kita lakukan dengan menggunakan teknologi Internet seperti mencari informasi, berkomunikasi dengan orang lain, mencari konten hiburan, melakukan jual beli (*commerce*), dan lain sebagainya. Salah satu penggunaan Internet yang ternyata dominan adalah penggunaan media sosial. Berdasarkan data, terdapat 4,76 miliar dari 5,16 miliar (92%) pengguna Internet adalah pengguna media sosial, sedangkan untuk kasus di Indonesia, 167 juta dari 212,9 juta pengguna Internet (78%) adalah pengguna media sosial (We are Social, 2023).

Dengan semakin banyaknya pengguna media sosial, maka semakin banyak pula data atau informasi yang tercipta di Internet terutama pada media sosial. Terdapat beberapa alasan yang dikemukakan oleh pengguna ketika menggunakan Internet, khususnya media sosial, seperti terlihat pada Gambar 1.

---

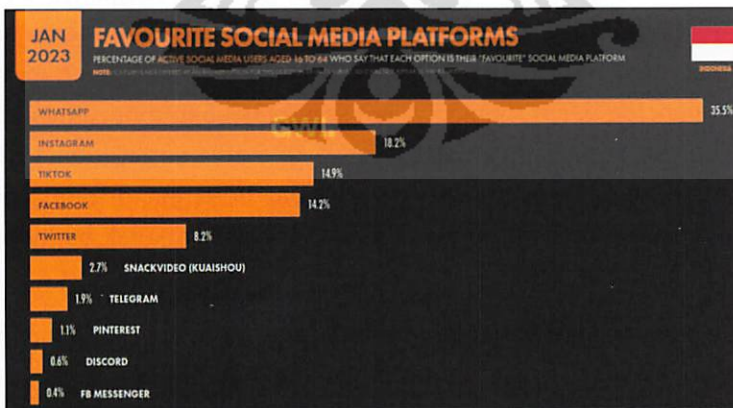
<sup>1</sup><https://apjii.or.id/berita/d/survei-apjii-pengguna-internet-di-indonesia-tembus-215-juta-orang>



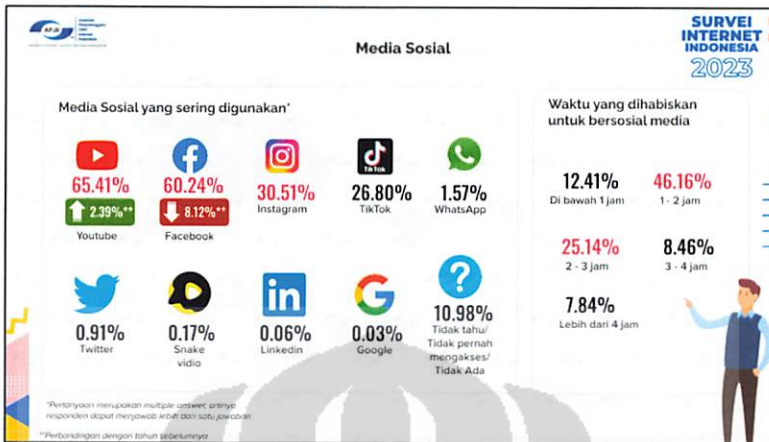
Gambar 1. Alasan menggunakan Internet (Sumber: Apjii, 2023)

Terlihat pada Gambar 1 bahwa untuk mengakses media sosial menjadi tujuan penggunaan Internet di Indonesia media sosial yang paling tinggi.

Sedangkan lima besar platform yang digunakan media sosial di Indonesia adalah Whatsapp, Instagram, Tiktok, Facebook dan Twitter, seperti terlihat pada Gambar 2 dan Gambar 3.



Gambar 2. Platform media sosial favorit di Indonesia (Sumber: We are Social 2023)



Gambar 3. Media sosial yang sering diakses di Indonesia (Sumber: Apjii, 2023)

Dengan aktifnya masyarakat menggunakan media sosial, menyebabkan data dan informasi semakin banyak tersebar di Internet, terutama di media sosial. Data dan informasi yang beredar dapat bersifat positif, misalnya konten-konten ajakan berbuat baik, informasi publik dari pemerintah, dsb. Namun ternyata juga banyak yang bersifat negatif, misalnya berita palsu (*hoax*) juga semakin banyak tersebar, penipuan melalui internet, ujaran kebencian, perundungan dan sebagainya. Dengan semakin banyaknya konten negatif yang bermunculan di media sosial, yang dapat mengakibatkan dampak negatif kepada masyarakat, misalnya ujaran kebencian dapat menjadi salah satu penyebab munculnya konflik di masyarakat. Walaupun awalnya hanya di media sosial (dunia maya), bukan tidak mungkin konflik tersebut dapat berlanjut di dunia nyata. Misalnya kasus baku hantam antar siswa SMA di Pontianak berawal dari saling ejek di media sosial<sup>2</sup>.

Agar potensi konflik tidak membesar dan menyebar ke dunia nyata maka perlu dilakukan upaya pendeteksian dini potensi konflik yang ada di

<sup>2</sup> <https://news.detik.com/berita/d-4506079/berawal-dari-bully-di-medsos-begini-kronologi-kasus-audrey>

dunia maya (media sosial). Salah satu upaya yang dapat dilakukan adalah dengan menggunakan analisis media sosial, misalnya dengan mendeteksi ujaran kebencian dan ujaran kasar yang ada di Twitter, mendeteksi potensi adanya perundungan (cyber bully) di Twitter, Instagram dan Facebook ataupun platform media sosial lainnya,

Mengingat bahwa konten yang di-post pengguna di media sosial menggunakan bahasa manusia yang dikenal dengan bahasa alami (*natural language*), dimana dalam ilmu komputer, bahasa alami dapat diproses dengan bidang ilmu pemrosesan bahasa alami (*natural language processing*). Pemrosesan Bahasa alami memiliki tantangan tersendiri, misalnya keambiguan (bahasa manusia dapat memiliki berbagai interpretasi), dan struktur gramatikal (sintaksis) dan makna kata serta kalimat (semantik) yang sangat bervariasi. Sejauhmana pendekatan atau metode yang ada dalam ilmu komputer dapat digunakan untuk melakukan analisis media sosial menjadi tema dalam pidato ini. Disamping itu juga dijelaskan penerapan analisis media sosial dalam berbagai domain, dan diberikan hasil studi kasus terkait dengan pendeteksian ujaran kebencian dan ujaran kasar pada Twitter di Indonesia.

**Hadirin yang saya hormati,**

### **Media Sosial dan Analisis Media Sosial**

Sebelum membahas bagaimana proses analisis dilakukan, terlebih dahulu kita perlu melihat definisi dari media sosial dan analisis media sosial.

Media sosial didefinisikan sebagai **platform berbasis Internet yang memungkinkan pengguna untuk membuat dan berbagi konten (seperti teks, video, audio dan grafis) dalam konteks *many-to-many*** (Khan, 2018). Disamping itu Khan menyebutkan terdapat lima karakteristik dari media sosial, yaitu sebagai media *peer-to-peer*, adanya partisipasi dan umpan balik dari pengguna, pemilik konten adalah penggunanya,

*conversational* (pengguna bebas berkomunikasi tanpa terbatas waktu dan lokasi), dan pengguna dapat menjalin hubungan sosial dengan lebih mudah dengan pengguna lainnya.

Telah banyak platform media sosial yang tersedia dan mempunyai banyak pengguna serta menjadi populer. Popularitas media sosial ini memberikan kesempatan untuk menciptakan tidak hanya hubungan pribadi, tetapi juga hubungan profesional bahkan hubungan bisnis. Beberapa *webometrics*, seperti statistik web dan situs analitik, seperti Alexa dan Google Analytics, menunjukkan bahwa situs web yang paling populer adalah halaman berbasis media sosial (Boyd & Ellison, 2007).

Media sosial juga menyediakan nilai sosial, yaitu data terkait dengan hubungan antar akun pengguna yang dapat diakses dan diolah oleh peneliti. Nilai sosial sangat penting, karena dapat digunakan dalam analisis graf sosial untuk mengekstrak informasi yang berguna, misalnya sejauhmana kaitan antar akun dalam suatu jaringan sosialnya (Pitas, 2016).

Dalam bukunya, Khan mendefinisikan bahwa analisis media sosial adalah **“Seni dan sains untuk mengekstrak wawasan (*insight*) tersembunyi yang berharga dari data media sosial yang besar baik yang semi-terstruktur dan tidak terstruktur yang dapat membantu dalam pengambilan keputusan”** (Khan, 2018).

Berikut penjelasan dari karakteristik analisis media sosial tersebut. Disebut sebagai seni (*art*), dikarenakan terkait dengan proses interpretasi dan menyelaraskannya dengan wawasan yang didapatkan dari hasil analisis yang dikaitkan dengan tujuan analisis tersebut dilakukan. Setiap orang bisa jadi mempunyai interpretasi yang berbeda untuk hasil proses analisis yang sama.

Proses analisis media sosial akan melibatkan proses identifikasi, ekstraksi, dan analisis data media sosial menggunakan alat, teknik dan metode ilmiah tertentu. Proses ini yang umumnya dipakai dalam penelitian ilmiah sehingga analisis media sosial disebut juga sebagai **sains**.



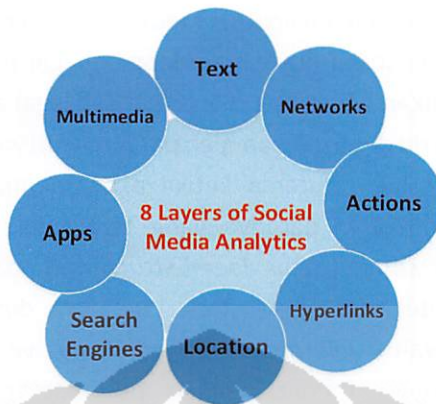
Data yang diproses dalam analisis media sosial merupakan data yang terdapat pada media sosial (yang tergolong sebagai basis data publik), sehingga memunculkan isu terkait dengan aspek legal atau etika. Hal ini menjadi hal yang perlu diperhatikan oleh para peneliti yang menggunakan data dari media sosial, terutama ketika akan mempublikasikan hasil penelitiannya. Selain jenis datanya yang tergolong tidak terstruktur (*unstructured*) dan semistruktur (*semi-structured*), data media sosial tersebut dapat dikategorikan juga sebagai *big data*, dimana mempunyai lima karakteristik yaitu *volume*, *velocity*, *variety*, *veracity* dan *value*. Sehingga pengetahuan terkait dengan pemrosesan big data juga diperlukan untuk melakukan proses analisis media sosial.

Kita tidak dapat langsung melihat wawasan/nilai bisnis atau informasi hanya dari koleksi data media sosial. Informasi atau wawasan/nilai bisnis tersebut perlu terlebih dahulu diekstraksi melalui proses analisis media sosial, dan inilah yang menjadi tujuan utama dari analisis media sosial, yaitu dapat membantu dalam pengambilan keputusan berdasarkan data atau informasi yang tersedia di media sosial.

**Hadirin yang saya hormati,**

### **Delapan Lapisan Analisis Media Sosial**

Dalam rangka membantu pengambilan keputusan berdasarkan data media sosial, maka analisis media sosial dapat dibagi kedalam delapan lapisan sesuai dengan sumber data dan peruntukannya seperti yang terlihat pada Gambar 4. Setiap lapisan mempunyai potensi informasi yang bernilai dan wawasan yang dapat diekstraksi untuk kebutuhan bisnis organisasi.



Gambar 4. Delapan lapisan analisis media sosial (Sumber: Khan, 2018)

### Lapisan 1: Text analytics

Analisis teks media sosial (*social media text analytics*) berkaitan dengan ekstraksi dan analisis wawasan bisnis dari elemen teks konten media sosial, seperti komentar, cuitan, postingan, atau pembaruan status. Hal ini dapat dilakukan dengan analisis teks ataupun pendekatan *text mining* (penambangan teks), diantaranya dapat memahami sentimen atau opini yang sedang *trending* ataupun sentimen pada suatu domain atau bidang tertentu.

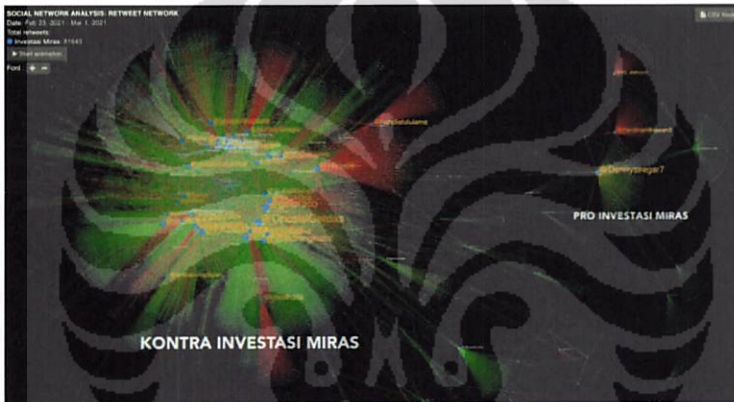


Gambar 5. Contoh twit di Twitter

Berdasarkan Gambar 5 yang merupakan salah satu contoh twit, dapat diekstrak informasi terkait dengan opini masyarakat terhadap suatu daerah wisata tertentu di Indonesia.

## **Lapisan 2. Network Analytics**

Analisis jaringan media sosial (*Social Media Network Analytics*) berkaitan dengan proses ekstraksi, analisis dan interpretasi jaringan sosial, misalnya jaringan pertemanan pada Facebook dan Twitter. Dengan melakukan ini, kita dapat antara lain menemukan *influencer* (mereka yang berpengaruh pada jaringan sosialnya). Disamping itu juga dapat dilihat struktur jaringan yang terbentuk jika dikaitkan dengan isu atau topik tertentu.



Gambar 6. Contoh jaringan sosial network terkait dengan isu investasi miras di Papua tahun 2021 (Sumber: DroneEmprit)

Pada Gambar 6, terlihat bahwa dapat dibuat jaringan dari akun-akun yang kontra atau pun pro pada media sosial Twitter terhadap investasi miras di Papua pada tahun 2021, dan terlihat bahwa mereka yang kontra lebih banyak.

## **Lapisan 3. Actions Analytics**

Analisis aksi media sosial (*actions analytics*) berkaitan dengan proses ekstraksi, analisis, dan interpretasi dari aksi yang dilakukan oleh pengguna pada media sosial, misalnya *likes*, *dislikes*, *shares*, *mentions*, atau

*endorsement*. Analisis ini antara lain dapat digunakan untuk memahami popularitas suatu *brand* tertentu (dari *likes*, *views* atau *clicks*), kinerja suatu strategi pemasaran di media sosial, pengukuran sejauh mana keterlibatan pengguna (*engagement*), dsb.

#### **Lapisan 4. Apps Analytics**

Analisis Aplikasi (*Apps Analytics*) berkaitan dengan pengukuran dan optimalisasi keterlibatan pengguna melalui aplikasi seluler (*mobile apps*). Salah satunya, *review* dan *rating* yang diberikan pengguna pada aplikasi seluler dapat menjadi objek untuk dianalisis, misalnya untuk melihat sejauh mana persepsi pengguna terhadap aplikasi tersebut.

#### **Lapisan 5. Hyperlink Analytics**

Analisis hyperlink (*Hyperlink analytics*) terkait dengan ekstraksi, analisis dan interpretasi hyperlinks pada media sosial, misalnya *in-links* dan *out-links*. Dengan melihat pola *in-links* dan *out-links* dari pengguna suatu media sosial, maka kita dapat melihat pola ketertarikan mereka terhadap suatu bidang/domain tertentu yang ada pada media sosial tersebut.

#### **Lapisan 6. Location Analytics**

Analisis Lokasi (*Location analytics*), sering juga dikenal dengan analisis spasial berkaitan dengan penambangan dan pemetaan lokasi dari pengguna media sosial, yang antara lain dapat ditentukan dari profilnya ataupun dari konten yang di-*post* oleh pengguna tersebut, karena seringkali profil dari pengguna di media sosial belum tentu akurat.

#### **Lapisan 7. Search Engine Analytics**

Analisis mesin pencari (*Search engines analytics*) berfokus pada analisis terhadap riwayat pencarian data pada suatu mesin pencari, sehingga dapat melakukan analisis tren, pengawasan kata kunci, riwayat hasil

pencarian dan iklan serta juga dapat dilakukan bagaimana statistik dari periklanan yang ada di mesin pencari tersebut.

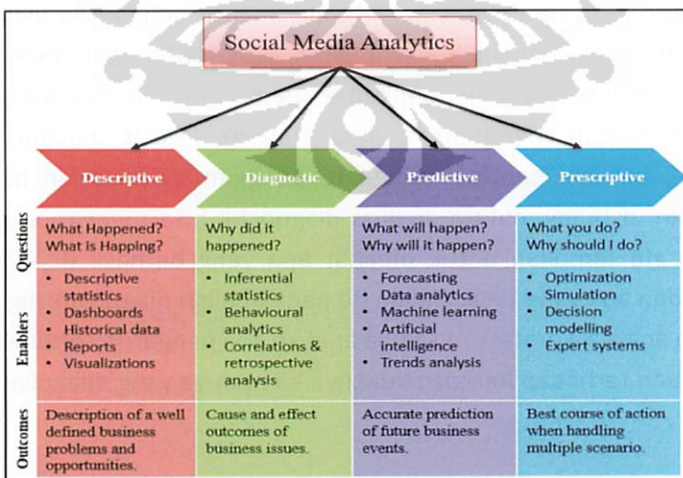
**Lapisan 8. Multimedia Analytics**

Analisis Multimedia (*Multimedia Analytics*) berkaitan dengan data yang menjadi objek analisis yang dapat bervariasi jenisnya, misalnya audio, gambar, teks dan video. Dimana jenis data tersebut memiliki karakteristik masing-masing dan memerlukan teknik pemrosesan sendiri.

Hadirin yang saya hormati,

**Tipe Analisis pada Media Sosial**

Analisis media sosial membantu mencapai tujuan bisnis melalui penggambaran data menjadi analisis tren, prediksi masa depan dan peluang dan optimasi proses bisnis dalam rangka meningkatkan kualitas dalam hal pengambilan keputusan organisasi (Delen, 2013). Analisis media sosial dapat dibagi menjadi empat tipe (Delen, 2013; Khan, 2018; Khan, 2019) yaitu analisis deskriptif, analisis diagnostik, analisis prediktif dan analisis preskriptif, seperti terlihat pada Gambar 7.



Gambar 7. Tipe Analisis Media Sosial (Sumber: Khan, 2018)

### **Analisis Deskriptif**

Analisis deskriptif bersifat reaktif dan menjawab pertanyaan “Apa yang sudah terjadi dan/atau apa yang sedang terjadi”. Jawaban dari pertanyaan tersebut didapatkan dengan berfokus pada pengumpulan dan mendeskripsikan data media sosial dalam bentuk laporan, visualisasi, dan mengelompokkannya untuk memahami masalahnya. Tindakan analisis (misalnya, menghitung jumlah *like*, *comment*, *twit* dan *views*) dan analisis teks adalah contoh analisis deskriptif. Analisis teks media sosial (misalnya terhadap komentar pengguna), dapat digunakan untuk memahami sentimen pengguna atau mengidentifikasi tren yang muncul dengan mengelompokkan tema dan topik tertentu. Salah satu ciri khas dari analisis deskriptif ini adalah, jika proses pengolahannya (statistik deskriptif) dilakukan dengan benar, maka hasilnya pasti benar.

### **Analisis Diagnostik**

Analisis diagnostik juga bersifat reaktif, akan menjawab pertanyaan “Mengapa sesuatu terjadi?” Pemrosesan pada analisis diagnostik mencakup statistik inferensial, analisis perilaku, korelasi dan analisis retrospektif serta hasilnya dapat berupa analisis sebab dan akibat dari masalah bisnis. Misalnya, pada analisis deskriptif dapat memberikan gambaran umum tentang kampanye pemasaran media sosial yang dilakukan dengan memberikan data statistik terkait dengan jumlah postingan, jumlah *mention*, pengikut, penggemar, tampilan halaman, ulasan, pin, dsb; sementara analisis diagnostik bisa menyaring data ini menjadi satu tampilan untuk melihat apa yang berhasil di kampanye tersebut dan apa yang tidak. Selain itu pada analisis diagnostik juga dapat dilakukan analisis korelasi, misalnya apakah ada pengaruh sentimen pada pemberitaan terhadap berhasil tidaknya kampanye yang dilakukan.

## **Analisis Prediktif**

Analisis prediktif melibatkan analisis sejumlah besar data media sosial yang digunakan untuk memprediksi sesuatu di masa depan. Analisis prediksi bersifat proaktif, berkaitan dengan pertanyaan “Apa yang akan terjadi dan/atau mengapa hal itu bisa terjadi?” Hasil utama dari pemodelan prediktif adalah proyeksi terkait dengan kejadian di masa depan dan alasan yang mendasari kejadian tersebut. Misalnya niat yang diungkapkan melalui media sosial (seperti membeli, menjual, merekomendasikan, berhenti, menginginkan, atau berharap) dapat digunakan sebagai basis untuk memprediksi suatu peristiwa di masa depan (seperti pembelian). Disamping itu, misalnya seorang manajer bisnis dapat memperkirakan angka penjualan berdasarkan kunjungan historis (atau tautan masuk) ke situs web perusahaan. Contoh lainnya, misalnya berdasarkan tweet yang ada di Twitter, kita dapat mengetahui opini masyarakat terhadap suatu isu atau topik tertentu. Ciri khas dari analisis prediktif adalah bahwa hasilnya tidak selalu akurat tergantung dari akurasi model yang digunakan.

## **Analisis Preskriptif**

Meskipun analisis prediktif membantu memprediksi masa depan, analisis preskriptif menyarankan tindakan terbaik yang dapat diambil ketika menangani suatu skenario tertentu. Misalnya, jika terdapat kelompok pengguna media sosial dengan perilaku pembelian tertentu, maka dengan analisis preskriptif, kita dapat memberikan penawaran yang berbeda untuk setiap kelompok. Analisis preskriptif dapat mencakup antara lain optimasi dan pemodelan simulasi, keputusan multi-kriteria pemodelan, ataupun sistem pakar.

**Hadirin yang saya hormati,**

### **Penggunaan Analisis Media Sosial pada Berbagai Domain di Indonesia**

Sebelum melihat bagaimana proses analisis media sosial dapat dilakukan, kita melihat dahulu contoh penggunaan atau penerapan analisis media sosial pada berbagai bidang di Indonesia.

Analisis media sosial dapat dilakukan pada berbagai bidang atau domain, baik untuk kepentingan bisnis (organisasi swasta) ataupun kepentingan pemerintahan (organisasi publik). Zachlod dalam artikelnya membagi penerapan analisis media sosial kedalam 14 area atau domain (Zachlod, 2022). Berikut merupakan contoh beberapa penggunaan analisis media sosial pada berbagai bidang atau area seperti pemerintahan, *commerce*, kota cerdas dan sosial kemasyarakatan di Indonesia.

#### **Bidang pemerintahan**

Pada bidang pemerintahan, analisis media sosial dapat dilakukan antara lain pada topik terkait dengan pemilihan umum, seperti kampanye, prediksi pemenang pemilu, deteksi buzzer, dsb. Disamping itu juga dapat dilakukan analisis opini atau persepsi masyarakat, misalnya sentimen analisis untuk melihat persepsi masyarakat terhadap topik tertentu. Tabel 1 berikut diberikan beberapa contohnya.

Pada Tabel 1 dapat terlihat beberapa contoh penerapan analisis media sosial pada bidang pemerintahan, misalnya untuk memprediksi pemenang pemilu presiden Indonesia tahun 2014, dimana berdasarkan data twitter dapat diprediksi pemenang pemilu dengan menggunakan tugas analisis sentimen, dan hasilnya dibandingkan dengan hasil survei dari beberapa lembaga survei. Ternyata hasil prediksi sistem berkorelasi positif dengan hasil dari berbagai lembaga survei (Ibrahim, 2015).



**Tabel 1. Contoh penerapan analisis media sosial pada bidang pemerintahan**

<b>No</b>	<b>Tugas</b>	<b>Tujuan</b>	<b>Sumber Data</b>	<b>Tipe</b>	<b>Sumber</b>
1	Analisis Sentimen	Memprediksi pemenang pemilihan presiden tahun 2014	Twitter	Analisis Deskriptif dan Analisis Prediktif	(Ibrahim, 2015)
2	Analisis Sentimen	Melihat kampanye pada pemilihan gubernur DKI tahun 2017	Twitter	Analisis Deskriptif dan Analisis Prediktif	(Mihardi, 2018)
3	Analisis Sentimen	Opini masyarakat terhadap bakal calon presiden 2024	Komentar Youtube	Analisis Deskriptif dan Analisis Prediktif	(Wisnubroto, 2022)
4	Analisis Sentimen dan Pemodelan Topik	Opini masyarakat terhadap kampanye kendaraan listrik	Twitter	Analisis Deskriptif dan Analisis Prediktif	(Ashari, 2023)

Selain pada aspek politik (terkait dengan pemilihan), juga dapat dilakukan analisis terkait dengan persepsi masyarakat terhadap kebijakan yang dikeluarkan pemerintah. Misalnya untuk melihat sentimen atau opini/persepsi masyarakat terhadap kebijakan kendaraan listrik menggunakan data Twitter (Ashari, 2023).

**Bidang Commerce atau Bisnis**

Pada *commerce* atau perdagangan atau bidang bisnis, analisis media sosial dapat dimanfaatkan pada berbagai area dan berbagai tujuan. Misalnya pada *e-commerce*, rekomendasi produk dapat menjadi salah satu aplikasi yang dikembangkan berdasarkan review pengguna.

Tabel 2. Contoh penerapan analisis media sosial pada bidang commerce

No	Tugas	Tujuan	Sumber Data	Tipe	Sumber
1	Rekomendasi	Memberikan rekomendasi item kepada pengguna pada e-commerce	data transaksi pada suatu e-commerce	analisis preskriptif	(Putra, 2017)
2	Analisis sentimen	mendeteksi kaitan intensi pembelian dgn sentimen pada social commerce	data komentar pada tiktok, facebook, dan instagram	analisis prediktif	(Virgananda, 2023)
3	Klasifikasi	deteksi konten mengandung promosi pada postingan di mobile apps	review pada aplikasi Female Daily	analisis prediktif	(Jonathan, 2021)
4	Klasifikasi	klasifikasi kategori komplain pada e-commerce	komplain pengguna Bukalapak	analisis prediktif	(Itsari, 2022)
5	Analisis sentimen	mendeteksi sentimen pengguna digital banking	Twitter	analisis prediktif	(Andrian, 2022)

Pada Tabel 2 terlihat bahwa berdasarkan data penjualan pada e-commerce tertentu, kita dapat memberikan rekomendasi terkait dengan suatu item kepada pengguna tertentu berdasarkan tingkat kesamaan item ataupun kesamaan profil penggunanya (Putra, 2017). Selain itu dapat pula dideteksi apakah ada kaitan antara intensi pembelian dengan sentimen berdasarkan postingan di Tiktok, Facebook dan Instagram (Virgananda, 2023). Berdasarkan review pada suatu aplikasi tertentu (dalam hal ini adalah aplikasi Female Daily), dapat dideteksi apakah review tersebut mengandung iklan atau tidak, sebab terdapat kebijakan di Female Daily bahwa review tidak boleh mengandung iklan, namun faktanya seringkali masih ditemukan review pengguna yang didalamnya masih mengandung iklan (Jonathan, 2021).

Klasifikasi juga dapat dilakukan pada komplain yang diberikan oleh pengguna pada suatu e-commerce, misalnya di Bukalapak, komplain dapat diklasifikasikan apakah terkait dengan aplikasi atau layanan yang diberikan (Itsari, 2022). Contoh lainnya yang juga termasuk klasifikasi adalah melakukan analisis sentimen untuk melihat persepsi atau opini pengguna digital banking apakah positif atau negatif berdasarkan tweet yang mereka *posting* di media sosial Twitter (Andrian, 2022).

### **Bidang Kota Cerdas (*Smart City*)**

Untuk mendukung penerapan kota cerdas (*smart city*), beberapa aplikasi dapat dikembangkan berbasis analisis media sosial, diantaranya untuk mendeteksi tingkat banjir, tingkat kemacetan, potensi wisata, dsb. Contoh penerapan analisis media sosial dapat dilihat pada Tabel 3.

Berdasarkan Tabel 3, dapat terlihat bahwa dengan menggunakan data Twitter, kita dapat melakukan deteksi terkait dengan ketinggian genangan air saat banjir (Felicia, 2021) ataupun dapat pula mendeteksi penyebab dan lokasi kemacetan (Putra, 2022).

Selain dapat menggunakan data dari media sosial dalam rangka mendeteksi ketinggian genangan air saat banjir, penyebab dan lokasi

kemacetan, dengan menggunakan data review terkait dengan restoran dapat ditentukan rekomendasi restoran berdasarkan kesamaan dari penggunaannya (Mustofa, 2018; Suciati, 2019). Rekomendasi restoran ini dapat dimanfaatkan untuk mendukung wisata kuliner pada suatu daerah. Aplikasi yang dapat memberikan rekomendasi restoran pada suatu kota tertentu tersebut dapat sebagai pendukung *e-tourism* sebagai bagian dari penerapan Kota Cerdas.

Tabel 3. Contoh penerapan analisis media sosial untuk mendukung penerapan Kota Cerdas

No	Tugas	Tujuan	Sumber Data	Tipe	Sumber
1	Klasifikasi	mendeteksi ketinggian genangan air saat banjir di Jakarta	Twitter	analisis prediktif	(Felicia, 2021)
2	Klasifikasi	mendeteksi penyebab dan lokasi kemacetan	Twitter	analisis prediktif	(Putra, 2022)
3	Rekomendasi	memberikan rekomendasi restoran	data review terkait restoran	analisis preskriptif	(Mustofa, 2018)
4	Analisis sentimen	analisis sentimen terhadap review restoran berdasarkan aspek makanan, suasana dan harga	review pengguna pada pergikuliner	analisis prediktif	(Suciati, 2019)

## Sosial Budaya dan Kemasyarakatan

Pada bidang sosial budaya dan kemasyarakatan, misalnya kita dapat melakukan deteksi ujaran kasar atau ujaran kebencian yang saat ini cukup marak muncul di media sosial. Beberapa contoh penerapan media sosial pada bidang sosial kemasyarakatan dapat dilihat pada Tabel 4.

Tabel 4. Contoh penerapan analisis media sosial pada bidang sosial budaya dan kemasyarakatan

No	Tugas	Tujuan	Sumber Data	Tipe	Sumber
1	Klasifikasi	mendeteksi ujaran kebencian terkait kekerasan seksual	Twitter	analisis prediktif	(Ramadhan, 2022)
2	Klasifikasi	Deteksi sebuah komentar apakah mengandung ujaran kebencian	Instagram	analisis prediktif	(Pratiwi, 2018)
3	Klasifikasi	deteksi apakah postingan mengandung unsur pornografi	Twitter	analisis prediktif	(Izzah, 2018)
4	Klasifikasi	deteksi ujaran kebencian pada lima bahasa daerah di Indonesia	Twitter	analisis prediktif	(Asti, 2021)
5	Klasifikasi	Deteksi ujaran kasar di Media Sosial	Twitter	analisis prediktif	(Ibrohim, 2018)

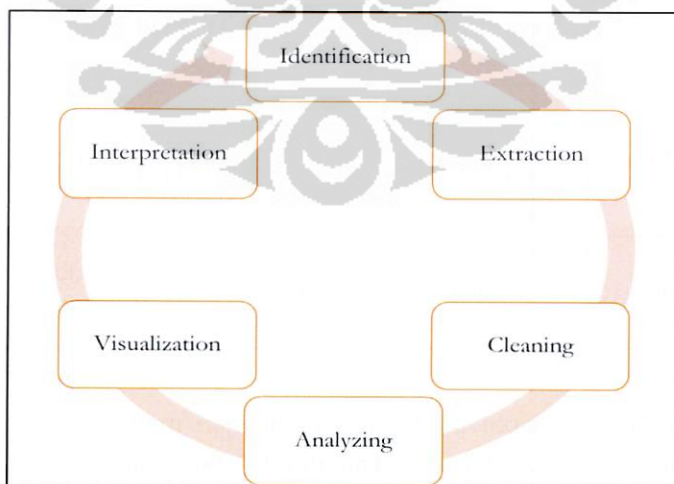
Pada Tabel 4 dapat dilihat bahwa dengan menggunakan analisis media sosial, kita dapat melakukan deteksi terhadap postingan yang muncul, apakah mengandung ujaran kebencian atau tidak, apakah juga mengandung ujaran kasar atau tidak (Ibrohim, 2018; Pratiwi, 2018; Asti,

2021). Disamping itu dapat pula membantu pendeteksian konten-konten negatif lain, misalnya mendeteksi apakah *postingan* yang ada mengandung unsur pornografi atau tidak (Izzah, 2018). Ujaran kasar dan ujaran kebencian yang muncul pada media sosial, menyebabkan keterbelahan pada masyarakat, terutama pada waktu menjelang pemilihan umum atau pemilihan kepala daerah. Secara tidak langsung dapat memicu terjadinya konflik di masyarakat. Untuk memberikan pemahaman lebih lanjut terkait dengan studi yang sudah dilakukan untuk pendeteksian ujaran kebencian dan ujaran kasar pada media sosial dalam konteks Indonesia dijelaskan pada bagian studi kasus dari pidato ini.

**Hadirin yang saya hormati.**

### **Tahapan Pengembangan Analisis Media Sosial**

Untuk mengembangkan aplikasi atau sistem berbasis analisis media sosial, perlu mengikuti alur atau tahapan tertentu. Salah satu tahapan yang dapat diikuti adalah seperti terlihat pada Gambar 8 yang juga terlihat sebagai sebuah siklus (Khan, 2018)



Gambar 8. Siklus pengembangan analisis media sosial (Sumber: Khan, 2018)

### **1. Identifikasi (*Identification*)**

Menemukan pertanyaan dan mengetahui data apa yang perlu dianalisis merupakan hal yang sangat penting dan krusial dalam mendapatkan wawasan bisnis yang berguna. Pencarian dan identifikasi sumber informasi yang benar untuk analisis menjadi hal yang penting untuk dilakukan pada awal proses pengembangan aplikasi. Selain itu, penting juga untuk merumuskan permasalahan dan potensial alternatif solusinya. Beberapa pertimbangan dalam menentukan sumber data atau informasi diantaranya: data bersifat publik atau bersumber dari sebuah organisasi; apakah data berkaitan dengan tujuan bisnis yang akan dicapai, media sosial apa yang menjadi sumber datanya.

### **2. Ekstraksi (*extraction*)**

Tahapan ekstraksi adalah tahapan untuk mendapatkan data atau koleksi data yang akan digunakan sebagai dasar dalam pengembangan model yang sesuai. Secara umum data dapat diperoleh dengan dua cara, yaitu *crawling* dan *scrapping* (Khder, 2021). *Crawling* adalah mengumpulkan data dari platform media sosial menggunakan API (*Application Programming Interface*), aplikasi khusus yang disediakan media sosial untuk pengaksesan data pada media sosial tersebut. Sedangkan *scrapping* adalah proses ekstraksi data dari halaman web atau situs web. Proses ini melibatkan penggunaan program komputer atau bot untuk mengakses dan mengambil informasi dari halaman-halaman web, termasuk teks, gambar, tautan, atau data lainnya.

### **3. Pembersihan Data (*Cleaning*)**

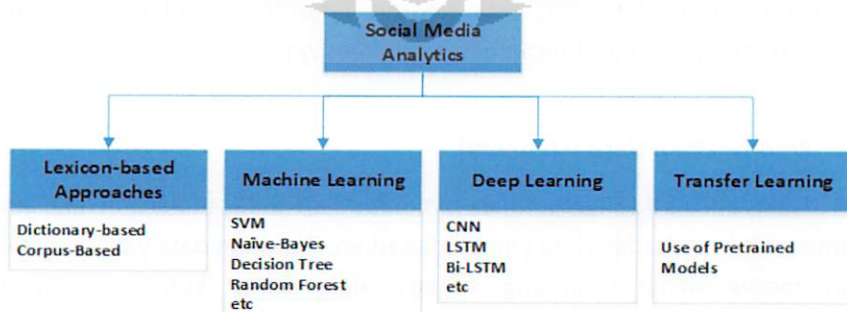
Tahapan berikutnya setelah mendapatkan data adalah melakukan pembersihan terhadap data yang didapatkan. Biasanya data yang berasal dari media sosial tergolong sebagai data yang 'kotor', misalnya mengandung iklan, emoji, dan karakter lainnya seperti tanda baca yang memang tidak diperlukan dalam rangka proses analisis. Pembersihan data

bertujuan untuk mengidentifikasi, memperbaiki, dan menghilangkan karakter-karakter yang tidak diperlukan, kesalahan eja (*error typos*), ketidakakuratan data, atau anomali dalam data. Data yang bersih dan berkualitas akan menghasilkan hasil analisis yang lebih tepat dan reliabel. Pada pembersihan data antara lain dilakukan penghilangan duplikat, penanganan data yang tidak konsisten, penanganan format data, normalisasi data, penghilangan karakter-karakter yang tidak diperlukan, misalnya tanda baca, angka, dsb.

#### 4. Analisis (*Analyzing*)

Setelah data dibersihkan, kemudian disesuaikan dengan format dari proses analisis yang akan dilakukan. Selanjutnya, dilakukan proses untuk penerapan metode tertentu dalam rangka penemuan pola sesuai dengan tugasnya. Penerapan metode ini akan menghasilkan model tertentu yang akan digunakan sesuai dengan tugasnya.

Secara umum pendekatan atau metode yang dapat digunakan dalam analisis media sosial dapat dibagi menjadi lima pendekatan seperti pendekatan yang dapat digunakan dalam salah satu tugas dalam analisis media sosial (analisis sentimen), yaitu *lexicon-based*, *machine learning*, *deep learning*, dan *transfer learning* serta *hybrid* diantara pendekatan-pendekatan tersebut, seperti yang terlihat pada Gambar 8 (Nandwani, 2021).



Gambar 8. Metode dalam analisis media sosial (modifikasi dari Nandwani, 2021)



***Lexicon-based approach:***

pada metode ini, digunakan kamus kata terkait dengan tugasnya, misalnya untuk analisis sentimen, metode ini menggunakan sebuah kamus kata di mana setiap kata positif dan negatif diberi nilai sentimen. Kemudian, jumlah atau rata-rata nilai sentimen digunakan untuk menghitung sentimen seluruh kalimat atau dokumen.

***Machine learning approach:***

Pada pendekatan pembelajaran mesin (*machine learning*), seluruh dataset dibagi menjadi dua bagian untuk tujuan pelatihan dan pengujian: dataset pelatihan dan dataset pengujian. Dataset pelatihan adalah informasi yang digunakan untuk melatih model dengan menyediakan karakteristik dari berbagai contoh. Dataset pengujian kemudian digunakan untuk melihat seberapa bagus model yang telah dilatih dari dataset pelatihan. Beberapa algoritma pembelajaran mesin antara lain: Naïve Bayes (NB), Support Vector Machine (SVM), Decision Trees (DT), Random Forest (RF), dsb.

***Deep learning approach:***

Walaupun pendekatan *deep learning* juga tergolong *machine learning*, namun pada pendekatan ini, kita tidak perlu mendefinisikan fitur-fitur terlebih dahulu seperti halnya pada pendekatan pembelajaran mesin secara tradisional. Fitur-fitur atau pola akan dibuat secara otomatis berdasarkan data teks yang diberikan. Beberapa contoh algoritma *deep learning*, diantaranya: CNN, LSTM, Bi-LSTM, dsb.

***Transfer Learning approach:***

Pendekatan ini juga merupakan bagian dari pembelajaran mesin. Sebuah model yang dilatih pada kumpulan data yang besar untuk menyelesaikan satu masalah dapat diterapkan pada masalah terkait

lainnya. Menggunakan kembali model yang telah dilatih sebelumnya pada suatu domain untuk dapat digunakan pada domain yang lain dapat menghemat waktu dan menghasilkan hasil yang lebih efisien.

## 5. Visualisasi (*Visualization*)

Setelah model diperoleh dan diterapkan pada dataset, maka langkah berikutnya yang dapat dilakukan adalah melakukan visualisasi dalam rangka memberikan kepada pengguna tampilan tertentu sesuai dengan tugas yang dapat digunakan oleh pengguna untuk mendapatkan wawasan tertentu. Visualisasi dapat menggambarkan misalnya jaringan data (siapa yang terlibat pada suatu konteks tertentu), topik data yang sedang dibahas, misalnya dengan wordcloud, gambaran tren data, atau gambaran spasial terhadap data. Sebagai contoh dapat dilihat pada Gambar 9 berikut.



Gambar 9. Contoh wordcloud sebagai contoh visualisasi<sup>3</sup>

---

<sup>3</sup> <https://pacmann.io/blog/ide-proyek-data-analysis-untuk-bangun-portofolio>

## **6. Interpretasi (*Interpretation*)**

Tahapan terakhir adalah proses interpretasi terhadap hasil analisis dan atau visualisasi yang dilakukan oleh pengguna. Proses ini tentunya dilakukan oleh pengguna yang akan menggunakan hasil dari proses analisis yang dilakukan berdasarkan visualisasi data. Interpretasi ini tentunya dapat berbeda antara satu pengguna dengan pengguna yang lainnya. Proses interpretasi yang akan dilakukan oleh pengguna, pastinya disesuaikan dengan tujuan dari pengembangan aplikasinya dalam rangka membantu dalam mendukung pengambilan keputusan.

### **Studi Kasus: Deteksi Ujaran Kebencian dan Ujaran Kasar pada Bahasa Daerah di Indonesia**

Pada bagian ini dijelaskan bagaimana analisis media sosial dapat dilakukan untuk mendeteksi ujaran kebencian dan ujaran kasar yang kerap kali muncul di media sosial, terutama di Twitter. Seperti disampaikan sebelumnya bahwa dalam konten di media sosial dapat bersifat positif ataupun negatif. Kedua konten tersebut dapat digunakan sebagai basis analisis dalam rangka mendukung pengambilan keputusan pada domain tertentu, seperti yang sudah dijelaskan pada contoh sebelumnya.

Pada bagian ini, dijelaskan lebih detail terkait dengan analisis terhadap konten negatif yaitu pendeteksian ujaran kebencian dan ujaran kasar pada media sosial di Indonesia. Intensitas kemunculan ujaran kebencian dan ujaran kasar di media sosial, terutama Twitter meningkat menjelang pemilihan umum atau pemilihan langsung pilkada seperti yang saat ini kita hadapi. Dengan melakukan deteksi hal ini pada media sosial, diharapkan sebagai upaya dini dalam rangka pencegahan potensi konflik di dunia maya.

Menurut Komnas HAM (2015), suatu ujaran kebencian mempunyai target, golongan, dan tingkatan tertentu. Ujaran kebencian bisa terhadap suatu golongan tertentu seperti suku, agama, ras, orientasi seks, dan lain-

lain yang ditargetkan kepada suatu individu atau kelompok tertentu. Selain mempunyai bentuk dan tingkatan tertentu, penyampaian suatu ujaran kebencian seringkali disertai dengan ujaran kasar (Davidson, dkk., 2017). Menurut Wijana dan Rohmadi (2011), ujaran kasar merupakan suatu ujaran yang disertai kata-kata kasar/kotor, baik dalam konteks bercanda maupun untuk menghina/memaki seseorang. Di Indonesia, kata-kata yang sering digunakan untuk menyampaikan ujaran kasar biasanya berasal dari suatu keadaan yang tidak menyenangkan atau sesuatu yang menjijikan dan diharamkan oleh agama. Namun tidak semua ujaran kasar adalah ujaran kebencian, begitu pula tidak semua ujaran kebencian mengandung ujaran kasar. Beberapa contoh ujaran kebencian dan ujaran kasar dapat dilihat pada Tabel 5.

Tabel 5. Contoh twit yang mengandung ujaran kebencian dan ujaran kasar  
(Sumber: Ibrohim dan Budi, 2023)

No	Twit	Ujaran Kebencian	Ujaran Kasar
1	Buddha Myanmar emang babi!	V	V
2	Bangsat! Dasar Cina kocak lu coy wkwk	-	V
3	Saya benci muslim, mereka suka membunuh orang!	V	-
4	Sesama manusia, muslim & buda harus hidup damai:)	-	-

Selain menggunakan Bahasa Indonesia dan Bahasa Inggris, ujaran kebencian yang ada di Indonesia juga dilakukan menggunakan bahasa daerah. Indonesia memiliki lebih dari 700 bahasa daerah dengan karakteristiknya masing-masing (Putri, 2018). Perbedaan kata yang memiliki arti yang sama merupakan salah satu kekhasan dari masing-masing bahasa daerah di Indonesia. Sebagai contoh, kata “*bodoh*” yang

sering digunakan dalam melontarkan ujaran kasar dan ujaran kebencian di Indonesia memiliki kata yang berbeda dalam masing-masing bahasa daerah di Indonesia, seperti “*bodho*” (Jawa), “*bodo/belet*” (Sunda), “*bhudhu*” (Madura), “*pakak/bengak*” (Minang), dan “*buyan*” (Musi).

Pada tahun 2017, ada sekitar 1.400 laporan ujaran kebencian yang masuk ke Bareskrim Polri, sedangkan pada tahun 2018 sampai kuartal pertama sudah ada sekitar 900 laporan yang masuk (Hernanto dan Jeihan, 2018). Disamping itu Kementerian Komunikasi dan Informatika (Kemenkominfo) menyatakan bahwa dari tahun 2018 hingga April 2021 terdapat sebanyak 3.640 ujaran kebencian pada ruang digital<sup>4</sup>, terutama media sosial. Banyaknya ujaran kebencian di media sosial dapat membuat pihak berwenang seperti Kemenkominfo dan Polri kesulitan dalam menentukan prioritas penyelesaian permasalahan ujaran kebencian. Selain itu pihak berwenang juga akan kesulitan serta membutuhkan waktu yang cukup lama dalam melakukan identifikasi ujaran kasar dan ujaran kebencian beserta target, kategori, dan level ujaran kebencian jika tidak memahami bahasa daerah yang digunakan oleh pengguna media sosial atau jika identifikasi dilakukan secara manual.

Tabel 6 memperlihatkan penelitian-penelitian terkait dengan ujaran kebencian dan ujaran kasar. Secara umum dapat terlihat bahwa sudah ada beberapa penelitian terkait dengan beberapa Bahasa antara lain India, Inggris dan juga Bahasa Indonesia dan juga Bahasa daerah di Indonesia. Selanjutnya penjelasan terkait dengan studi kasus ini, yaitu identifikasi ujaran kebencian dan ujaran kasar pada Bahasa daerah di Indonesia mengacu kepada penelitian Asti dan kawan-kawan (Asti, dkk, 2021). Pada naskah pidato ini tidak semua bagian dijelaskan, sehingga jika ingin mengetahui penjelasan lebih detail terkait dengan bagian tertentu dapat melihat di (Asti, 2022).

---

<sup>4</sup><https://diskominfotik.riau.go.id/2021/04/27/sejak-2018-kominfo-tangani-3-640-ujaran-kebencian-berbasis-sara-di-ruang-digital/>

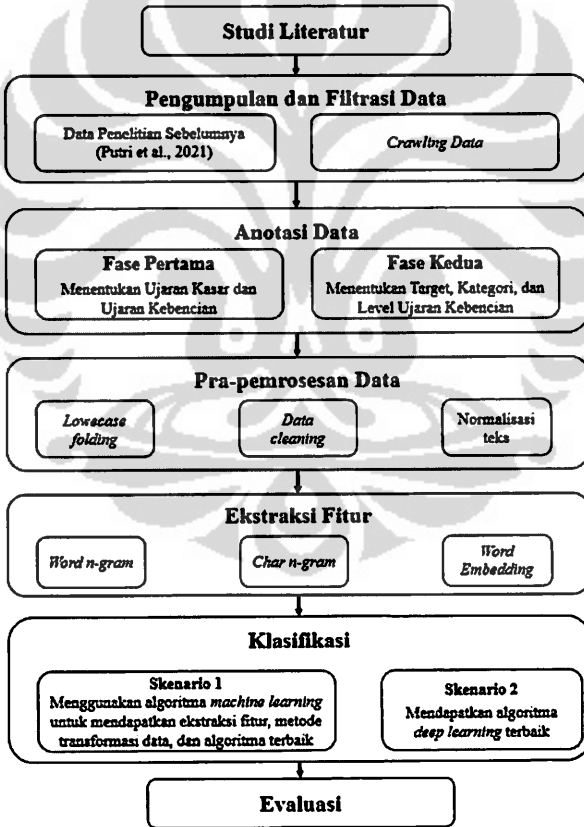
Tabel 6. Penelitian terkait ujaran kebencian dan ujaran kasar (Sumber: Asti, 2022)

Sumber	Bahasa	Metode	Fitur
(Sazany & Budi, 2018)	Indonesia	LSTM	Word2Vec, fastText
Sandaruwan et al. (2019)	Sinhala	SVM, MNB, RFDT	<i>Bag-of-Word, Word n-gram, Character n-gram, dan Word skip-gram</i>
Sutejo & Lestari (2019)	Indonesia	LSTM	fitur tekstual, fitur akustik, dan kombinasi keduanya
Ibrohim & Budi (2019)	Indonesia	SVM, NB, RFDT	Frekuensi term, ortografi, leksikon
Hana et al. (2020)	Indonesia, Inggris	SVM, CNN, DistilBERT	TF-IDF
Hendrawan et al. (2020)	Indonesia, Inggris	RFDT, BiLSTM, BiLSTM dengan pre-trained BERT	<i>Word unigram TF-IDF</i>
Kurniawan & Budi (2020)	Indonesia	SVM, Naïve Bayes, RFDT	<i>Bag-of-Words (BOW), TF-IDF</i>
Rahul et al. (2021)	Code-mixing Hindi dan English (Hinglish)	MNB, Logistic Regression, RFDT, SVM, CNN, BiLSTM, Machine Learning Ensemble, Deep Learning Ensemble	<i>Count Vectorizer, TF-IDF, fastText</i>
Malik et al. (2021)	Inggris	<i>Machine Learning (LR, SVM, DT, RF, XGBoost) dan Deep Learning (CNN, MLP, LSTM)</i>	<i>n-gram, POS tagging, TF-IDF, BERT, fastText</i>
Putri et al. (2021)	Bahasa daerah di Indonesia	SVM, Naïve Bayes, RFDT	TF-IDF

Sumber	Bahasa	Metode	Fitur
Asti, et al. (2023)	5 Bahasa daerah di Indonesia	SVM, MNB, dan RFDT	n-gram

**Tahapan Penelitian**

Dapat terlihat pada Gambar 10 tahapan penelitian yang dapat dilakukan terkait dengan pendeteksian ujaran kebencian dan ujaran kasar pada Bahasa daerah pada Twitter.



Gambar 10. Tahapan penelitian pendeteksian ujaran kebencian dan ujaran kasar pada Bahasa daerah di Twitter

## **Pengumpulan Data**

Data diambil dari Twitter dengan cara *crawling* menggunakan Twitter API (Tweepy Library)<sup>5</sup>. Crawling data dilakukan dalam rentang waktu 2 Desember 2020 hingga 21 Maret 2021 guna mendapatkan 5000 twit untuk data masing-masing bahasa daerah. Crawling data tersebut menggunakan kueri dari penelitian Ibrohim & Budi (2019) yang diterjemahkan ke masing-masing Bahasa daerah. Bahasa daerah yang digunakan adalah lima Bahasa daerah yaitu: Jawa, Sunda, Madura, Minang, dan Musi.

## **Anotasi Data**

Anotasi data dilakukan oleh dua hingga tiga orang untuk masing-masing bahasa daerah. Kriteria utama pemilihan anotator adalah orang yang fasih dalam masing-masing bahasa daerah. Anotasi dilakukan berdasarkan panduan anotasi yang telah dibuat. Masing-masing twit tersebut dianotasi untuk menentukan target, kategori, dan level ujaran kebencian.

## **Pra-pemrosesan Data**

Pra-pemrosesan data yang dilakukan antara lain *lowercase folding*, *data cleaning*, dan normalisasi teks. *Lowercase folding* yaitu mengubah semua kata pada twit menjadi huruf kecil. Selanjutnya dilakukan *data cleaning* yakni menghapus karakter-karakter yang tidak diperlukan saat proses klasifikasi seperti menghapus *Retweet* (RT), *link* atau URL, tanda baca, angka, dan lain sebagainya. Normalisasi teks dilakukan pada kata-kata yang tidak baku dalam penulisannya menggunakan kamus typo.

---

<sup>5</sup> <http://docs.tweepy.org/en/latest/>



### Ekstraksi Fitur

Fitur-fitur yang digunakan pada pendekatan *classical machine learning* adalah *word n-gram* dan *char n-gram* dengan pembobotan TF-IDF. Pada pendekatan *deep learning*, fitur yang digunakan adalah *word embedding*. Dua jenis *word-embedding* yang digunakan, yaitu yang berasal dari data latih dan *pre-trained model fastText* dalam Bahasa daerah. Tabel 7 merupakan contoh ilustrasi ekstraksi fitur *word n-gram* dan *char n-gram*.

Tabel 7. Contoh ekstraksi fitur

Fitur	Teks	Nilai n	Hasil ekstraksi fitur
<i>Word n-gram</i>	ini merupakan contoh ekstraksi fitur	3	'ini merupakan contoh', 'merupakan contoh ekstraksi', 'contoh ekstraksi fitur'
<i>Char n-gram</i>	ekstraksi fitur	4	'ekst', 'kstr', 'stra', 'trak', 'raks', 'aksi', 'ksi_', 'si_f', 'i_fi', '_fit', 'fitu', 'itur'

### Klasifikasi

Algoritma klasifikasi yang digunakan pada pendekatan tradisional machine learning adalah Support Vector Machine (SVM), Multinomial Naive Bayes (MNB), dan Random Forest Decision Tree (RFDT). Selain menggunakan algoritma tradisional machine learning, juga menggunakan algoritma deep learning Long Short-Term Memory (LSTM) dan Convolutional Neural Network (CNN).

## **Evaluasi**

Setelah proses klasifikasi selesai, maka dilakukan evaluasi dari hasil klasifikasi tersebut. Evaluasi pada penelitian ini menggunakan 5-fold cross validation dan menggunakan metrik evaluasi **F1-score**. Metrik ini menggabungkan dua metrik lainnya, yaitu presisi (*precision*) dan *recall*, ke dalam satu nilai tunggal, yaitu:

$$\text{F1-score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}$$

Sedangkan yang dimaksud dengan presisi adalah mengukur sejauh mana hasil positif yang diprediksi oleh model benar-benar positif dengan rumusan sebagai berikut:

$$\text{Presisi} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Dan yang dimaksud dengan *recall* adalah mengukur sejauh mana model mengenali semua instance positif yang sebenarnya, dengan rumusan sebagai:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

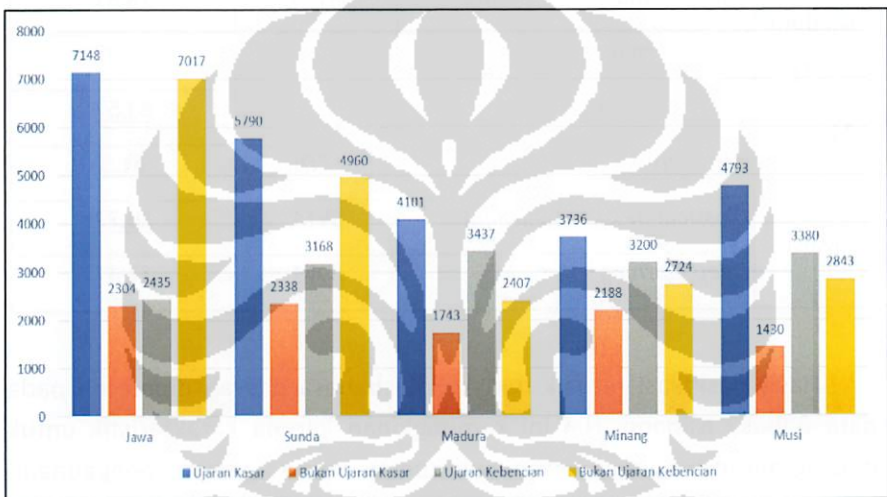
Dengan definisi dari *True Positive*, *False Positive* dan *False Negative* mengacu kepada *confusion matrix* yang ada pada Tabel 8.

Tabel 8. *Confusion Matrix*

		Aktual	
		Positive	Negative
Prediksi	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Dapat dilihat pada Tabel 8 bahwa True Positive adalah instance yang positive dan diprediksi positive oleh model, True negative adalah instance yang negative dan diprediksi sebagai negative oleh model. False Positive adalah instance negative yang diprediksi sebagai positive oleh model sedangkan False Negative adalah instance yang positive diprediksi sebagai negative oleh model.

### Hasil dan Analisis



Gambar 11. Distribusi tweet pada masing-masing bahasa daerah (Sumber: Asti, dkk., 2021)

Data yang digunakan dalam eksperimen adalah dari lima Bahasa daerah (Jawa, Sunda, Madura, Minang dan Musi) dengan proporsi dapat terlihat pada Gambar 11.

Tabel 8. Hasil eksperimen deteksi ujaran kebencian (Sumber: Asti, 2022)

Bahasa	Pendekatan	Algoritma	F1-score (%)
Jawa	<i>Classical machine learning</i>	SVM	<u>70,43</u>
	<i>Deep learning</i>	CNN	64,06
Sunda	<i>Classical machine learning</i>	SVM	<u>68,79</u>
	<i>Deep learning</i>	LSTM	63,83
Madura	<i>Classical machine learning</i>	SVM	<u>78,81</u>
	<i>Deep learning</i>	CNN	76,93
Minang	<i>Classical machine learning</i>	SVM	<u>83,57</u>
	<i>Deep learning</i>	CNN	81,5
Musi	<i>Classical machine learning</i>	SVM	<u>80,72</u>
	<i>Deep learning</i>	SVM	77,47

Hasil klasifikasi terbaik dari kelima bahasa daerah diperoleh pada data Bahasa Minang. Hal ini kemungkinan karena karakteristik untuk masing-masing bahasa daerah berbeda. Keberagaman penggunaan Bahasa Minang tidak seperti Bahasa Jawa dan Sunda. Pada Bahasa Jawa dan Sunda terdapat tingkatan bahasa yang membuat kosakata pada data lebih beragam. Namun hal tersebut memungkinkan algoritma salah dalam proses klasifikasi karena terdapat kata yang memiliki arti yang sama. Contoh kesalahan klasifikasi dapat terlihat pada Tabel 9.

Tabel 9. Contoh kesalahan klasifikasi pada Bahasa Sunda dan Bahasa Minang (Sumber: Asti, 2022)

Kesalahan klasifikasi	Contoh Twit Bahasa Sunda	Contoh Twit Bahasa Minang
<i>False Positive</i>	@USER kecap bagong	Jiah kareh kareh cirik anjiang, baun se yang kareh nyo, dipijak lunak
	@USER Neehh loe di laknat junjungan loe sendiri.... wakakaka	manga aden dulu kok pantek
<i>False Negative</i>	@USER eta si undergod banci mereun nya, ni rajin pisan nguruskeun hirup batur. tina seblak kana munafik, anjir jararauh.asa geuleuh nempo lalaki gibah julid ciga kitu	@USER Raisa bongak
	@user ya lo anjir kupu2 malam	Mario anjiang

Contoh kesalahan klasifikasi pada Bahasa Sunda dan Bahasa Minang dapat terlihat pada Tabel 9. Terlihat pada Tabel 9 bahwa twit pada data *false positive* umumnya mengandung kata kasar dimana kata tersebut sering digunakan sebagai ujaran kebencian seperti kata ‘bagong’ dan ‘laknat’ di Bahasa Sunda, ‘pantek’ dan ‘anjiang’ di Bahasa Minang. Sedangkan pada kesalahan klasifikasi false negative pada Bahasa Sunda kemungkinan terjadi karena twit tersebut mengandung kata yang dapat digunakan sebagai ujaran kasar dan bukan ujaran kasar seperti kata ‘geuleuh’. Sedangkan pada Bahasa Minang, *false negative* kemungkinan terjadi karena kesalahan klasifikasi karena algoritma yang digunakan belum dapat menangkap penggunaan nama orang dalam menyatakan ujaran kebencian dengan menggunakan kata kasar.

**Hadirin yang saya hormati,**

### **Tantangan dalam Analisis Media Sosial**

Berdasarkan uraian diatas, untuk mengembangkan aplikasi berbasis analisis media sosial, terutama terkait dengan analisis teks pada media sosial, paling tidak membutuhkan ketersediaan data dan pengetahuan tentang *natural language processing* (NLP). Disamping itu juga ditemukan beberapa tantangan dalam analisis media sosial terkait dengan data tersebut.

Berikut adalah tantangan dalam melakukan analisis media sosial, terutama terkait dengan karakteristik dari datanya.

#### ***Volume dan Velocity***

Media sosial menghasilkan data dalam jumlah yang sangat besar (*volume*) dan bertambah atau berubah dengan sangat cepat (*velocity*). Hal ini sesuai dengan karakteristik dari *big data*. Memproses, menyimpan, dan menganalisis volume data yang besar bisa menjadi tantangan teknis tersendiri, disamping itu perubahan data yang sangat cepat, diperlukan akurasi waktu dalam memprosesnya sehingga hasilnya dapat dipergunakan sebagaimana mestinya. Hal ini juga memerlukan kemampuan untuk mengikuti tren dan perubahan data tsb.

#### **Kualitas Data**

Data di media sosial seringkali berisi banyak *noise*, termasuk spam, akun palsu, dan informasi yang salah. Memisahkan data yang berguna dari yang tidak berguna adalah tantangan, hal ini dapat ditangani dalam tahapan *preprocessing* dalam analisis media sosial. Namun yang dapat ditangani adalah yang terkait dengan struktur atau sintaks dari kalimat atau bahasanya, sedangkan untuk menentukan apakah konten-nya valid atau tidak akan menjadi tantangan tersendiri untuk menentukannya.

## **Variabilitas Bahasa dan Kultur**

Media sosial adalah platform global, dan data dapat berisi berbagai bahasa, dialek, dan konteks budaya. Apalagi dalam konteks Indonesia, dalam media sosial, mengingat pengguna dari berbagai latar belakang daerah dan budaya yang berbeda, seringkali pengguna menggunakan bahasa daerah dalam postingan atau komentarnya di media sosial. Menganalisis data yang bervariasi dalam hal bahasa dan budaya memerlukan pemahaman yang mendalam tentang konteksnya. Hal ini tercermin dalam contoh kasus yang dijelaskan sebelumnya.

## **Privacy Concerns**

Mengumpulkan dan menganalisis data media sosial sering melibatkan masalah privasi. Penting untuk mematuhi regulasi privasi dan etika yang berlaku. Sehingga peneliti diharapkan juga memahami aturan yang ada terkait dengan sejauhmana data yang ada di media sosial dapat diambil dan dipublikasikan.

## **Keterbatasan Sumber Data**

Tidak semua data di media sosial dapat diakses atau dianalisis. Beberapa platform mungkin membatasi akses ke data, yang dapat menghambat analisis komprehensif, misalnya Twitter saat ini sudah mulai membatasi pengaksesan datanya.

Untuk mengatasi tantangan dalam analisis media sosial, analis perlu mengembangkan keterampilan teknis, memahami konteks data, mengikuti regulasi dan etika, dan menggunakan alat dan teknik analisis yang sesuai. Kesadaran akan tantangan-tantangan ini adalah langkah pertama dalam menghasilkan hasil analisis yang bermanfaat dan akurat.

## **Penutup**

Media sosial merupakan tempat berbagi data dan informasi yang saat ini banyak digunakan oleh masyarakat Indonesia. Pada media sosial tersebar konten yang tidak hanya bersifat positif, namun juga yang bersifat negatif. Analisis media sosial dengan menggunakan pendekatan ilmu komputer umumnya dan *text mining* secara khusus dapat digunakan dalam mengolah data media sosial terutama yang bersifat teks untuk mendapatkan wawasan (*insight*) yang ada pada media sosial. Pengetahuan terkait dengan *natural language processing* juga dibutuhkan mengingat bahwa konten yang di-*posting* pada media sosial adalah berupa teks Bahasa manusia (Bahasa alami).

Tujuan utama analisis media sosial adalah dalam rangka membantu dalam pengambilan keputusan dalam berbagai domain atau bidang. Selain itu analisis media sosial dapat pula digunakan untuk menganalisis konten negatif yang muncul di media sosial, misalnya deteksi ujaran kebencian. Hal ini dapat sebagai upaya dini yang dapat dilakukan dalam rangka mencegah terjadinya potensi konflik di masyarakat di dunia maya, apalagi memasuki tahun pemilu ini.



## **Daftar Pustaka**

- Asosiasi Penyelenggara Jasa Internet Indonesia.(2023). Laporan Survei Internet APJII 2022. Jakarta: Asosiasi Penyelenggara Jasa Internet Indonesia.
- Andrian, B., Simanungkalit, T., Budi, I., and Wicaksono, A.F. (2022). Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia. In *International Journal of Advanced Computer Science and Applications(IJACSA)*, 13(3), 2022. <http://dx.doi.org/10.14569/IJACSA.2022.0130356>
- Asti, A.D. (2022). Identifikasi target kategori dan level ujaran kebencian pada data twitter berbahasa daerah menggunakan classifical machine learning dan deep learning, Tesis Magister, Fakultas Ilmu Komputer, Universitas Indonesia.
- Asti, A. D., Budi, I., and Ibrohim, M. O. (2021). Multi-label Classification for Hate Speech and Abusive Language in Indonesian-Local Languages. In *2021 International Conference on Advanced Computer Science and Information Systems (ICACISIS)*, Depok, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICACISIS53237.2021.9631316.
- Ashari, N., Mifta Al Firdaus, M. Z., Budi, I., Santoso, A. B., and Putra, P.K. (2023). Analyzing Public Opinion on Electrical Vehicles in Indonesia Using Sentiment Analysis and Topic Modeling. In *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, Jakarta, Indonesia, 2023, pp. 461-465, doi: 10.1109/ICCoSITE57641.2023.10127834.
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230.

- Can, U., & Alatas, B. (2019). A new direction in social network analysis: Online social network analysis problems and applications. *Physica A: Statistical Mechanics and Its Applications*, 535, 122372. <https://doi.org/https://doi.org/10.1016/j.physa.2019.122372>.
- Davidson, T., Warmesley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017, Icwsm*, 512–515.
- Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support System*, 55(1), 359-363. <http://dx.doi.org/10.1016/j.dss.2012.05.044>
- Felicia, I.K., and Budi, I.(2021).Classification of Inundation Level using Tweets in Indonesian Language. In *ACM International Conference Proceeding Series*, pp. 137–143.
- Ibrahim, M., Abdillah, O., Wicaksono, A.F., dan Adriani, M. (2015).Buzzer Detection and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation, 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 2015, pp. 1348-1353, doi: 10.1109/ICDMW.2015.113
- Ibrohim, M.O., dan Budi, I.(2018).A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media, *Procedia Computer Science Volume 135*, 2018, Pages 222-229
- Ibrohim, M.O., Budi, I. (2023).Hate speech and abusive language detection in Indonesian social media: Progress and challenges, *Heliyon*, 2023, 9(8). <https://doi.org/10.1016/j.heliyon.2023.e18647>
- Itsari, M. Y. I., and Budi, I. (2022). Classification of Complaint Categories in E-Commerce: A Case Study of PT Bukalapak. In *2022 5th International Conference on Information and Communications Technology*

(ICOIACT), Yogyakarta, Indonesia, 2022, pp. 317-324, doi: 10.1109/ICOIACT55506.2022.9971933.

Izzah, N., Budi, I., and Louvan, S. (2018). Classification of pornographic content on Twitter using support vector machine and Naive Bayes. In 2018 4th International Conference on Computer and Technology Applications (ICCTA), Istanbul, Turkey, 2018, pp. 156-160, doi: 10.1109/CATA.2018.8398674.

Jonathan, B., and Budi, I. (2021). The Use of Text Mining for Classification of Product Selling Content in Social Media Female Daily. In 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICACSIS53237.2021.9631329.

Khan, Gohar, F. (2018). *Creating Value with Social Media Analytics: Managing, Aligning and Mining Social Media Text, Networks, Actions, Locations, Apps, Hyperlinks, Multimedia, & Search Engines Data*. CreateSpace, Seattle, WA 98108-0683, USA.

Khan, Gohar dan Dillon, Stuart. (2019). What Should Businesses Know About Social Media Analytics?. ACIS 2019 Proceedings. 14. <https://aisel.aisnet.org/acis2019/14>

Khder, M.(2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. In International Journal of Advances in Soft Computing and its Applications 13(3):145-168 DOI:10.15849/IJASCA.211128.11

Komnas HAM. (2015). *Buku Saku Penanganan Ujaran Kebencian*. Jakarta: Komisi Nasional Hak Asasi Manusia Republik Indonesia.

Mihardi, M. dan Budi, I. (2018). Public Sentiment on Political Campaign Using Twitter Data in 2017 Jakarta's Governor Election. In 2018 International Conference on Applied Information Technology and

- Innovation (ICAITI), Padang, Indonesia, 2018, pp. 67-72, doi: 10.1109/ICAITI.2018.8686740.
- Mustafa, A. A., and Budi, I. (2018). Recommendation System Based on Item and User Similarity on Restaurants Directory Online. In 2018 6th International Conference on Information and Communication Technology (ICoICT), Bandung, Indonesia, 2018, pp. 70-74, doi: 10.1109/ICoICT.2018.8528775.
- Nandwani, P., Verma, R. (2021). A review on sentiment analysis and emotion detection from text. In Soc. Netw. Anal. Min. 11, 81 (2021). <https://doi.org/10.1007/s13278-021-00776-6>
- Pitas, I. (Ed.). (2016). Graph-Based Social Media Analysis. Chapman and Hall/CRC. <https://doi.org/10.1201/b19513>
- Pratiwi, N. I., Budi, I., and Alfina, I. (2018). Hate Speech Detection on Indonesian Instagram Comments using FastText Approach. In 2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Yogyakarta, Indonesia, 2018, pp. 447-450, doi: 10.1109/ICACSIS.2018.8618182.
- Putra, A. A., Mahendra, R., Budi, I., and Munajat, Q. (2017). Two-steps graph-based collaborative filtering using user and item similarities: Case study of E-commerce recommender systems. In 2017 International Conference on Data and Software Engineering (ICoDSE), Palembang, Indonesia, 2017, pp. 1-6, doi: 10.1109/ICoDSE.2017.8285891.
- Putra, P.K., Mahendra, R. & Budi, I. (2022). Traffic and road conditions monitoring system using extracted information from Twitter. J Big Data 9, 65 (2022). <https://doi.org/10.1186/s40537-022-00621-3>
- Ramadhan, M. N., Budi, I., Santoso, A. B., and Suryono, R. R. (2022). Sexual Violence Classification as Hate Speech using Indonesian Tweet. In 2022 International Symposium on Information Technology and

Digital Innovation (ISITDI), Padang, Indonesia, 2022, pp. 114-120, doi: 10.1109/ISITDI55734.2022.9944482.

Suciati, A., and Budi, I. (2019). Aspect-based Opinion Mining for Code-Mixed Restaurant Reviews in Indonesia. In 2019 International Conference on Asian Language Processing (IALP), Shanghai, China, 2019, pp. 59-64, doi: 10.1109/IALP48816.2019.9037689.

Virgananda, M.A., and Budi, I., Kamrozi and Suryono, R.R. (2023). Purchase Intention and Sentiment Analysis on Twitter Related to Social Commerce. In International Journal of Advanced Computer Science and Applications (IJACSA), 14(7), 2023. <http://dx.doi.org/10.14569/IJACSA.2023.0140760>

Wankhade, M., Rao, A.C.S. & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artif Intell Rev* 55, 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>

We are Social, Digital 2023 Indonesia, <https://datareportal.com/reports/digital-2023-indonesia> diakses tanggal 16 Oktober 2023

Wijana, I.D.P. & Rohmadi., M. (2011). *Sosiolinguistik: Kajian, Teori, dan Analisis* (3rd ed.). Yogyakarta: Pustaka Pelajar.

Wisnubroto, A. S., Saifunas, A., Santoso, A. B., Putra, P. K. and Budi, I. (2022). Opinion-based sentiment analysis related to 2024 Indonesian Presidential Election on YouTube, 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2022, pp. 318-323, doi: 10.1109/ISRITI56927.2022.10052848.

Zachlod, C., Olga Samuel, Andrea Ochsner, Sarah Werthmüller. (2022). Analytics of social media data – State of characteristics and application, *Journal of Business Research* Volume 144, May 2022, Pages 1064-1076

## **UCAPAN TERIMA KASIH**

Hadirin yang saya hormati, perkenankan saya pada akhir pidato ini menyampaikan ucapan terima kasih kepada semua pihak yang telah memberikan kesempatan dan pembelajaran, sehingga saya bisa berada pada posisi saya sekarang ini.

Puji syukur Alhamdulillah kepada Allah Subhanahu Wa Ta'ala atas semua limpahan rahmat, karunia, rezeki, dan semua anugerah yang telah diberikan kepada saya, terutama nikmat iman dan kesehatan serta gelar Guru Besar ini. Semoga dengan gelar Guru Besar ini, dapat memberikan keberkahan bagi semua orang.

Sholawat dan salam, semoga senantiasa tercurah kepada junjungan kita Nabi Besar Rasulullah SAW dan juga keluarganya dan para sahabatnya serta para pengikutnya. Semoga kita semua bisa menjadikan Rasulullah sebagai suri tauladan dan istiqomah menjadi pengikutnya hingga akhir akhir hayat.

Pertama-tama, ucapan terima kasih saya haturkan kepada Pemerintah Republik Indonesia melalui Bapak Menteri Pendidikan, Kebudayaan, Riset dan Teknologi, Bapak Nadiem Anwar Makarim, B.A., M.B.A yang berkenan menandatangani SK Guru Besar saya. Terima kasih atas kepercayaannya yang diberikan kepada saya untuk menjadi Guru Besar di lingkungan kementerian Bapak. Terima kasih juga saya haturkan kepada Rektor Universitas Indonesia Prof. Ari Kuncoro, S.E., M.A., Ph.D., Terima kasih juga kepada seluruh jajaran Pimpinan Universitas Indonesia, Ketua dan Anggota Senat Akademik Universitas Indonesia, Ketua dan Anggota Dewan Guru Besar Universitas Indonesia.

Kedua, ucapan terima kasih saya ucapkan kepada Manajemen Fasilkom UI, Prof. Petrus Mursanto, Pak Ari Saptawijaya, Prof Achmad Nizar Hidayanto, Pak Alfian Farizki Wicaksono, Prof Putu Wuri Handayani, Ibu Betty Purwandari, Prof. Harry Budi Santoso, Pak Adila Alfa Krisnadhi, Bu Dina Chahyati, Pak Rizal Fathoni Aji, Pak Panca Hadi Putra, Ibu Laksmi Rahadiani, Pak Muhammad Hafizhuddin Hilman, Ibu Edina Ayuningtyas,

Pak Hafied Nur Siddiqi, dan Ibu Hennie Marianie. Ketua DGB Fasilkom UI, Prof. Eko K Budiardjo beserta anggota DGBF, Ketua Senat Akademik Fasilkom UI, Prof Heru Suhartanto beserta seluruh anggota SAF. Anggota UPMA Fasilkom, Pak Amril Syalim dan Pak Wahyu Catur Wibowo. Serta kepada Dekan Fasilkom UI sebelumnya, Pak Bagyo Y. Moeliodihardjo, M.Sc., Prof Toemin A. Masoem, Prof T. Basaruddin dan Ibu Dra. Mirna Adriani, PhD (alm). Terima kasih atas semua bantuan, dukungan, perhatian, dan masukannya selama saya menjalankan kewajiban saya dalam Tri Dharma Perguruan Tinggi.

Ketiga, terima kasih yang mendalam saya ucapkan kepada Prof. Ir. Zainal Arifin Hasibuan, MLS., Ph.D. yang telah membimbing saya dari sejak S1 sampai S3, dan yang mengajak saya untuk menjadi dosen di Fasilkom UI. Terima kasih atas semua ilmu yang diberikan dan kepercayaannya membantu berbagai tugas di Fasilkom. Kepada Prof Stéphane Bressan dari School of Computing, National University of Singapore, terima kasih telah memperkenalkan saya kedalam dunia penelitian. Semoga karya yang telah dihasilkan dapat memberikan keberkahan. Dan semua guru-guru saya, dari SD, SMP, SMA, S1, S2 dan S3 yang tidak mungkin saya sebutkan satu persatu, terima kasih untuk semua ilmu yang sudah kalian berikan untuk saya. Semoga Allah SWT memberikan balasan yang terbaik untuk semua pengorbanan kalian.

Keempat, terima kasih kepada rekan-rekan dosen Fasilkom UI yang telah banyak memberikan ilmu dan pengalaman terutama dalam kegiatan pengajaran dan penelitian: Prof. Aniati Murni Arymurthy, Prof. Belawati H. Widjaja, Prof. T. Basaruddin, Prof. Heru Suhartanto, Prof. Eko Kuswardono Budiardjo, Prof. Dana Indra Sensuse, Prof. Kasiyah, Prof. Yudho Giri Sucahyo, Prof. Wisnu Jatmiko, Ibu Erdefi Rakun, Pak Yugo Kartono Isal, Pak Bobby A. A. Nazief, Pak Denny, Pak Adhi Yuniarto, Pak L. Yohanes, Pak Ade Azurat, Pak Lim Yohanes Stefanus, Ibu Siti Aminah, Pak Amril Syalim, Pak Ari Wibisono, Pak Muhammad Rifki Shihab, Pak Heri Kurniawan, Pak Wahyu Catur Wibowo, Pak Bob Hardian Syahbuddin, Ibu Yova Ruldeviyani, Pak Dadan Hardianto, Ibu Ave Adriana Pinem, Ibu

Fatimah Azzahro, Ibu Widia Resti Fitriani, Ibu Nabila Clydea Harahap, Ibu Ika Chandra Hapsari, Ibu Imairi Eitiveni, Ibu Ika Alfina, Ibu Annisa Monicha Sari, Ibu Meganingrum Arista Jiwanggi, Ibu Maya Retno Ayu Setyautami, Ibu Iis Solichach, Ibu Iis Afriyanti, Pak Ichlasul Affan, Pak Fariz Darari, Ibu Evi Yulianti, Ibu Kurniawati Azizah, Bapak Made Harta Dwijaksana, Ibu Dinial Utami Nurul Qomariah, Ibu Puspa Indahati Sandhyaduhita, Pak Rahmad Mahendra, Ibu Arlisa Yulawati, Pak Rahmat Mustafa Samik-Ibrahim, Pak Muhammad Mishbah, Pak Pudy Prima, Pak Satrio Baskoro Yudhoatmojo, Ibu Arawinda Dinakaramani, Pak Suryana Setiawan, Pak Setiadi Yazid, Pak Widijanto Satyo Nugroho.

Kelima, terima kasih dari lubuk hati paling dalam kepada kawan-kawan dosen di Lab Information Manajemen dan Lab Information Retrieval dan NLP. Dan juga kepada mahasiswa dan alumni: Mas Oky, Mas Ryan, Bu Yani, Pak Yani, Pak Tosan, Bu Rizka, Pak Bambang, Bu Nurul, Pak Firdaus, dan yang lainnya yang tidak dapat saya sebutkan namanya satu persatu. Terima kasih atas diskusi dan sharing serta publikasinya serta bantuan dalam pelaksanaan hibah-hibah. Tetap semangat dalam melakukan penelitian dan publikasi.

Keenam, kepada Direktur Direktorat SDM UI dan segenap jajarannya, terutama tim SDM UI yaitu Pak Agus Anang dan tim Penilai AA dan Lektor UI, terima kasih atas semua dukungannya sehingga proses pengurusan Guru Besar saya dapat berjalan dengan baik. Terima kasih yang tak terhingga pula kepada tim SDM Fasilkom UI, terutama Mbak Hennie dan Mbak Juwita dan semua tim SDM Fasilkom UI (Mbak Aisyah dan Mbak Mala) yang sangat membantu dalam proses pengurusan Guru Besar ini.

Ketujuh, kepada seluruh tenaga kependidikan Fasilkom UI: Staf sekretariat akademik terutama Ibu Dewi, Ibu Rita, Ibu Nunik, Mbak Fiqoh, Mas Dede, Mas Handys, Mas Muslih, Mbak Afifah, Mbak Dira, dan yang di MTI: Mas Fajar, Mbak Ika W dan Mbak Nurul; staf perpustakaan (Ibu Lily dan Mas Wiryo); tim Riset & Kemahasiswaan (Mbak Ifa, Mbak Buana dan Mbak Ika P); staf TI (Pak Maman, Uda Baginda, Mas Dullah, Mas Galuh, Mas Edda, Mas Akmal, Mas Febi, Mas Sidik dan Mas Ganda); tim Humas



(Mbak Trisna, Mbak Sendi, dan Mas Rama); tim Keuangan (Mbak Tri, Mbak Selvia, Mbak Lail, Mbak Lidya); tim Fasum (Pak Harry, Pak Margo, Pak Acil, Pak Ali, Mbak Tita dan Mas Nur), tim Sekretariat Dekan (Mbak Inggit dan Mbak Anggitha), dan semua satpam dan janitor Fasilkom UI; terima kasih atas semua bantuan yang telah diberikan sehingga tugas saya menjadi lebih mudah dijalankan. Dan juga kepada tenaga kependidikan Fasilkom yang sudah pensiun: Bu Nining, Bu Sylvia, Bu Nunuk, Bu Puji, Bu Lala, Pak Udin, Pak Kresno (alm) dan Pak Sudirman (alm), terima kasih atas ilmu yang diberikan dan juga bantuannya sehingga memudahkan saya menjalankan tugas yang diberikan.

Kedelapan, terima kasih saya ucapkan kepada kawan-kawan dan mentor di Aptikom dan LAM-Infokom yang telah memberikan kesempatan bagi saya untuk berkontribusi dalam upaya sharing terkait dengan keilmuan dan penjaminan mutu pendidikan tinggi di Indonesia. Serta kepada Pemerintah Kota Depok, terutama Kepala Dinas dan Staf Dinas Komunikasi dan Informatika yang telah mempercayai saya menjadi anggota Dewan Kota Cerdas, mohon maaf jika belum maksimal dalam menjalankan tugas.

Kesembilan, kepada keluarga besar Rajo Alam Pagaruyung, keluarga besar Lubuk Jambi dan keluarga besar Rumah Baru pewaris Rajo Ibadat Sumpur Kudus, terima kasih atas dukungannya, semoga diberikan balasan setimpal oleh Allah SWT.

Kesepuluh, kepada kawan-kawan dan sahabat alumni SD Negeri Duren Tiga 04 Jakarta angkatan 89, SMPN 182 Jakarta angkatan 92, SMAN 28 Jakarta angkatan 95 serta kawan-kawan Fasilkom UI angkatan 95, terima kasih telah menjadi teman seperjuangan serta terima kasih atas dukungan dan doa'nya. Sehat dan sukses selalu bagi kawan-kawan semua.

Kepada Bapak Mertua Sutan Maradjo (alm) dan Ibu Mertua Nurlely, terima kasih atas bimbingan dan doa'nya serta telah merestui saya

sebagai menantu. Semoga Papa mendapatkan tempat terbaik di sisi Allah SWT.

Kakak dan adik ipar: Neci dan Mas Dedy, Niwed, Dadon dan Uni Phit, Nepha dan Mas Arpri, Anggi dan Lahandi, Ni Phit, serta Asmul Khairi, terima kasih atas doa dan dukungannya.

Kakak dan adik-adikku: Da Syaf, Andri, dan Ria, terima kasih atas doa dan dukungan kalian. Semoga senantiasa kita berada dalam lindungan Allah SWT.

Ponakan-ponakanku: Kakak Zaza, Kakak Karin, Abang Razes, Abang Alta, Abang Ifan, Abang Ifal, Abang Hafizh, Kakak Hani, Kakak Hanum, Adik Maryam, Abang Dimas, Abang Fikri, Adik Farhan, Kakak Hafsha, Adik Abyan. Terima kasih atas doanya, teruslah mengejar cita-cita kalian.

Putra-putra yang Abi sayangi dan cintai: Abang Izzudin Abdul Rasyid, Kakak Muhammad Azzam Mushlih, Uda Raihan Abdurrahman dan Adik Rizqi Muhammad Ramadhan. Terima kasih atas do'a kalian, kalianlah qurrota 'ayun bagi Abi dan Ummi. Abang dan Kakak tetaplah menjaga hafalan quran kalian, Insya Allah syurga merindukan kalian. Uda dan Rizqi semoga juga dapat menjadi hafizh yaa...Mohon maaf Abi tidak selalu ada menemani kalian.

Istriku tercinta Ayu Maturia, terima kasih telah menjadi pendamping selama 22 tahun ini, terima kasih atas doa dan dukungannya, dan mohon maaf jika belum bisa menjadi suami yang diimpikan.

Terakhir, kepada Bapak Maksus Zakaria dan Amak Nurlis yang paling saya cintai dan sayangi, terima kasih telah mendidik dan memberikan kasih sayang dan selalu mendoa'kan Ananda dari kecil hingga dapat mencapai gelar ini. Mohon maaf atas salah dan kurang Ananda, sampai kapanpun Ananda tidak akan dapat membalas segala kebaikan yang telah Bapak dan Amak berikan, hanya kepada Allah lah berharap memberikan balasan terbaik atas kebaikan Bapak dan Amak.

Hadirin yang saya hormati,

Akhir kata, dari lubuk hati yang paling dalam, saya mohon maaf yang sebesar-besarnya apabila terdapat hal yang kurang berkenan dari pidato ini. Terima kasih sekali lagi atas kehadiran para hadirin semua di acara pengukuhan Guru Besar saya ini. Semoga keberkahan dan kebahagiaan dari Allah SWT senantiasa dilimpahkan pada kita semua.

Aamiin,

Wassalamu'alaikum warohmatullahi wabarokatuh.



## **RIWAYAT HIDUP**

### **Data Pribadi**

Nama : Prof. Dr. Indra Budi, S.Kom, M.Kom.  
Tempat, Tanggal Lahir : Sumpur Kudus (Sijunjung), 1 Maret 1976  
Jenis Kelamin : Laki-laki  
Agama : Islam  
Kewarganegaraan : Indonesia  
E-mail : [indra@cs.ui.ac.id](mailto:indra@cs.ui.ac.id)  
Pendidikan Tertinggi : Doktor  
Jabatan Akademik : Guru Besar dalam Bidang Ilmu Komputer  
Pangkat/Gol : Pembina Utama Muda/IVc  
Institusi : Universitas Indonesia  
Fakultas : Fakultas Ilmu Komputer  
Alamat : Fakultas Ilmu Komputer, Kampus UI Depok  
Nomor Telepon/Faks : 021-7863419/021-7863415  
Google Scholar h-index : 24  
Scopus h-index : 17

### **Data Keluarga**

Ayah : Maksus Zakaria  
Ibu : Nurlis  
Ayah Mertua : Sutan Maradjo (alm)  
Ibu Mertua : Nurlely

Istri : Ayu Maturia  
Anak : 1. Izzudin Abdul Rasyid  
2. Muhammad Azzam Mushlih  
3. Raihan Abdurrahman  
4. Rizqi Muhammad Ramadhan  
Kakak : Safrizal  
Adik : 1. Andri 2. Fitria Sari

**Riwayat Pendidikan**

<b>Nama Sekolah</b>	<b>Jenjang</b>	<b>Program Studi/Jurusan</b>	<b>Lulus</b>
Universitas Indonesia	Doktor	Ilmu Komputer	2008
Universitas Indonesia	Magister	Ilmu Komputer	2003
Universitas Indonesia	Sarjana	Ilmu Komputer	2000
SMA Negeri 28 Jakarta	SMA	A1 (Fisika)	1995
SMP Negeri 182 Jakarta	SMP		1992
SDN Duren Tiga 04 Jakarta	SD		1989

**Riwayat Jabatan Fungsional**

No.	Jabatan	Mulai
1	Pengajar	1 Desember 2008
2	Asisten Ahli	1 Februari 2011
3	Lektor	1 Agustus 2012
4	Lektor Kepala	1 Agustus 2018
5	Guru Besar	1 Agustus 2023

**Pengalaman Profesional/Organisasi**

No	Deskripsi	Durasi
1	Asesor LAM-Infokom	2022 - sekarang
2	Asesor BAN-PT	2011 - sekarang
3	Anggota Pokja Tim Smart City Kota Depok	2020 - sekarang
4	Anggota Divisi 4 DE LAM-Infokom	2021 – 2024
5	Unit Penjaminan Mutu Fakultas Fasilkom UI	2021 – 2024
6	Ketua Prodi Sistem Informasi Fasilkom UI	2017 - 2021
7	Manajer Umum Fasilkom UI	2014-2017
8	Anggota Tim Ahli Standar Data BSNP	2015-2016
9	Dosen dan Peneliti Fasilkom UI	2001- sekarang

**Penghargaan**

No	Penghargaan	Lokasi	Tahun	Keterangan
1	Best Paper pada SNATI 2010	Yogyakarta	2010	
2	Best Paper pada KNSI 2015	Manado	2015	
3	Satyalancana Karya Satya 10 Tahun	Nasional	2019	Keppres No 77/TK/TAHUN 2019

**Pengalaman Penelitian dalam 5 Tahun Terakhir**

No.	Tahun	Skema Hibah	Judul Proposal	Sumber Pendanaan
1	2023	PUTI Pascasarjana	Automated Short Answer Scoring using Sentence Transformers and Deep Learning	UI
2	2023	Product Market Fit UI Incubate	Langing Academy: One Stop Service for AI, DS, NLP, and CV Education Needs	UI
3	2022	PUTI Q1	Exploring Sentiment and Focus Discussion about Green Related Topic in The Indonesian Social Media	UI
4	2022	PUTI Pascasarjana	A New Model of Fintech Supervision in Indonesia Based on Online News, Twitter and Google Playstore Reviews	UI

<b>No.</b>	<b>Tahun</b>	<b>Skema Hibah</b>	<b>Judul Proposal</b>	<b>Sumber Pendanaan</b>
5	2021	Publikasi Doktor (PD)	Penyusunan dan Evaluasi Skema Klasifikasi Fungsi Sitasi dalam Teks	UI
6	2020	Penelitian Disertasi Doktor	Ekstraksi Informasi Dokumen Keputusan Pengadilan yang Bersifat Adaptif Menggunakan Metode Berbasis Aturan dan Pembelajaran Mesin	Dikti
7	2020	Penelitian Disertasi Doktor	Challenges and Trends Financial Technology (Fintech) di Indonesia	Dikti
8	2020	PUTI Q2	Analisis Tingkat Equitas Merk Menggunakan Data Media Sosial Twitter	UI
9	2020	PUTI Doktor	Ekstraksi Informasi Dokumen Keputusan Pengadilan Yang Bersifat Adaptif Menggunakan Pendekatan Pembelajaran Mesin	UI
10	2020	PUTI Q2	Deteksi Buzzer, Thread Starter, dan Lokasi Penyebar Ujaran Kebencian pada Code-Mixed Twitter Bahasa Indonesia-Bahasa Daerah	UI
11	2020	PUTI Doktor	Analisis Sosial Media Berbasis Data Fusion untuk Mendeteksi	UI



<b>No.</b>	<b>Tahun</b>	<b>Skema Hibah</b>	<b>Judul Proposal</b>	<b>Sumber Pendanaan</b>
			Kasus Fintech P2P Lending di Indonesia	
12	2020	PUTI Doktor	Analisis Sitasi dalam Teks untuk Memperbaiki Metode Pemingkatan dan Pemetaan Sains	UI
13	2020	PUTI Saintekes	Analisis Sentimen Berbasis Aspek dan Deteksi Emosi pada Ulasan Berbahasa Campuran	UI
14	2020	PUTI Prosiding	Pendekatan Deep Learning untuk Deteksi Bot, Analisis Sentimen dan Question Generation	UI
15	2020	PUTI Prosiding	Perbandingan Machine Learning dan Deep Learning untuk Identifikasi Bahasa Kasar, Ujaran Berpropaganda dan Analisis Sentimen pada Twitter dan Artikel Berita	UI
16	2019	PIT-9	Penggunaan Teknik Data Mining untuk Deteksi Ujaran Kebencian (Hate Speech) dan Review Analysis	UI
17	2019	PITTA-A	Identifikasi Ujaran Kasar, Ujaran Kebencian, dan Konten Bermuatan Asusila pada Twitter Berbahasa Indonesia	UI

No.	Tahun	Skema Hibah	Judul Proposal	Sumber Pendanaan
18	2018	PITTA	Deteksi Ujaran Kebencian ( Hate Speech ) pada Konten Media Sosial Menggunakan Pendekatan Machine Learning	UI
19	2018	PITTA	Klasifikasi E-commerce, Emoji Twitter dan Legal Document menggunakan Pendekatan Machine Learning	UI
20	2018	PITTA	Analisis Konten Media Sosial untuk Melihat Partisipasi Masyarakat dan Interaksinya dengan Pemerintah	UI

**Publikasi Artikel Ilmiah dalam 5 Tahun Terakhir**

No	Artikel
1	Ashari, N., Mifta Al Firdaus, M. Z., Budi, I., Santoso, A. B., and Putra, P.K. (2023). Analyzing Public Opinion on Electrical Vehicles in Indonesia Using Sentiment Analysis and Topic Modeling. In 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), pp. 461-465, doi: 10.1109/ICCoSITE57641.2023.10127834.
2	Ibrohim, M.O., dan Budi, I. (2023). Hate speech and abusive language detection in Indonesian social media: Progress and challenges, <i>Heliyon</i> , 2023, 9(8). <a href="https://doi.org/10.1016/j.heliyon.2023.e18647">https://doi.org/10.1016/j.heliyon.2023.e18647</a>

<b>No</b>	<b>Artikel</b>
3	Virgananda, M.A., and Budi, I., Kamrozi and Suryono, R.R. (2023). Purchase Intention and Sentiment Analysis on Twitter Related to Social Commerce. In International Journal of Advanced Computer Science and Applications (IJACSA), 14(7), 2023. <a href="http://dx.doi.org/10.14569/IJACSA.2023.0140760">http://dx.doi.org/10.14569/IJACSA.2023.0140760</a>
4	Budi, I., dan Suryono, R.R. (2023). Application of named entity recognition method for Indonesian datasets: a review, Bulletin of Electrical Engineering and Informatics 12 (2), pp. 969-978
5	Al Hakim, S., Sensuse, D.I., Budi, I. et al. (2023). Expert retrieval based on local journals metadata to drive small-medium industries (SMI) collaboration for product innovation. Social Network Analysis and Mining 13(1),68 <a href="https://doi.org/10.1007/s13278-023-01044-5">https://doi.org/10.1007/s13278-023-01044-5</a>
6	Al Hakim, S., Sensuse, D.I., Budi, I., Mishbah, W.M., Khusni, U.(2022). An empirical study of knowledge mapping implementation in Indonesian organizational context, in VINE Journal of Information and Knowledge Management Systems 51(5), pp. 773-791.
7	Amrullah, M.S., Budi, I., Santoso, A.B., dan Putra, P.K. (2023). The effect of using Emoji and Hashtag in sentiment analysis on Twitter case study: Indonesian online travel agent, in 2023 AIP Conference Proceedings 2654 (1).
8	Andrian, B., Simanungkalit, T., Budi, I., and Wicaksono, A.F. (2022). Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia. In International Journal of Advanced Computer Science and Applications(IJACSA), 13(3), 2022. <a href="http://dx.doi.org/10.14569/IJACSA.2022.0130356">http://dx.doi.org/10.14569/IJACSA.2022.0130356</a>

<b>No</b>	<b>Artikel</b>
9	Budi, I., dan Yaniasih. (2022). Understanding the meanings of citations using sentiment, role, and citation function classifications, <i>Scientometrics</i> 128 (1), pp. 735-759
10	Itsari, M. Y. I., and Budi, I. (2022). Classification of Complaint Categories in E-Commerce: A Case Study of PT Bukalapak. In 2022 5th International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2022, pp. 317-324, doi: 10.1109/ICOIACT55506.2022.9971933.
11	Maulana, P., Budi, I., Budi Santoso, A. (2022). Sentiment Analysis of Indonesian Government's Effort to Overcome the Unemployment Problem during COVID-19 Pandemic, In ICOIACT 2022 - 5th International Conference on Information and Communications Technology, pp. 144-149
12	Mazidnianto, M.F., Damanik, A.P., Budi, I., Santoso, A.B., Putra, P.K. (2022). Abortion drug products classification using text mining: A case study of PT XYZ, in 2022 AIP Conference Proceedings 2658,090003
13	Putra, P.K., Mahendra, R. & Budi, I. (2022). Traffic and road conditions monitoring system using extracted information from Twitter. <i>J Big Data</i> 9, 65 (2022). <a href="https://doi.org/10.1186/s40537-022-00621-3">https://doi.org/10.1186/s40537-022-00621-3</a>
14	Prisdayanti, A., Budi, I., Santoso, A.B., Putra, P.K. (2022). Sentiment analysis of government regulations regarding the implementation of reverse-transcriptase polymerase chain reaction (RT-PCR) during the Covid-19 pandemic in Indonesia (Case study: Air transportation mode), in 2022 AIP Conference Proceedings 2658,090002

<b>No</b>	<b>Artikel</b>
15	Marzuqi, T.A., Kristiani, E., Budi, I., Santoso, A.B., Putra, P.K. (2022). Analysis of government regulations sentiment regarding the Eid al-Fitr homecoming during Covid-19 pandemic in Indonesia using Twitter and Instagram data, in 2022 AIP Conference Proceedings 2658,040001.
16	Pamungkas, H., Yudhianto, R.H., Sembiring, B.J., Budi, I. (2022). Finding the Right Influencers on Instagram for Endorsement Product Using Text Mining, in International Journal on Advanced Science, Engineering and Information Technology 12(4), pp. 1395-1402.
17	Putra, P.O.H., Dewi, R.A.W.W.C.K., Budi, I. (2022). Usability factors that drive continued intention to use and loyalty of mobile travel application, in 2022 Heliyon 8(9),e10620
18	Ramadhan, M. N., Budi, I., Santoso, A. B., and Suryono, R. R. (2022). Sexual Violence Classification as Hate Speech using Indonesian Tweet. In 2022 International Symposium on Information Technology and Digital Innovation (ISITDI), Padang, Indonesia, 2022, pp. 114-120, doi: 10.1109/ISITDI55734.2022.9944482.
19	Rizquillah, R.D., dan Budi, I. (2022). Text Normalization on Code-Mixed Twitter Text using Language Detection, In 2022 Seventh International Conference on Informatics and Computing (ICIC),
20	Sinaga, R.F.P., dan Budi, I. (2022). Influencer Detection Through Social Network Analysis on Twitter of the Indonesian Smartphone Industry, Lecture Notes in Networks and Systems 685 LNNS, pp. 97-107.

<b>No</b>	<b>Artikel</b>
21	Sinatrya, N.S., Budi, I., Budi Santoso, A. (2022). Classification of Stock Price Movement With Sentiment Analysis and Commodity Price: Case Study of Metals and Mining Sector, in Proceedings - ICAC SIS 2022: 14th International Conference on Advanced Computer Science and Information Systems pp. 59-64.
22	Wisnubroto, A. S., Saifunas, A., Santoso, A. B., Putra, P. K. and Budi, I. (2022). Opinion-based sentiment analysis related to 2024 Indonesian Presidential Election on YouTube, 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2022, pp. 318-323, doi: 10.1109/ISRITI56927.2022.10052848
23	Yani, M., Krisnadhi, A.A., Budi, I. (2022). A better entity detection of question for knowledge graph question answering through extracting position-based patterns, in Journal of Big Data 9(1),80
24	Yuliawati, A., Tohari, H., Mahendra, R., Budi, I. (2022). On the Long Tail Products Recommendation using Tripartite Graph, in International Journal of Advanced Computer Science and Applications 13(1), pp. 816-822.
25	Akbar, A.F., Santoso, A.B., Putra, P.K., Budi, I. (2021). A classification model to identify public opinion on the lockdown policy using Indonesian tweets, in Journal of Theoretical and Applied Information Technology 99(14), pp. 3394-3402.
26	Artari, V.K.P., Mahendra, R., Jiwanggi, M.A., Anggraito, A., Budi, I. (2021). A Multi-Pass Sieve Coreference Resolution for Indonesian, in 2021 International Conference Recent Advances in Natural Language Processing, RANLP pp. 79-85

<b>No</b>	<b>Artikel</b>
27	Asti, A. D., Budi, I., and Ibrohim, M. O. (2021). Multi-label Classification for Hate Speech and Abusive Language in Indonesian-Local Languages. In 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICACSIS53237.2021.9631316.
28	Fatimah, N., Budi, I., Santoso, A.B., Putra, P.K. (2021). Analysis of Mental Health During the Covid-19 Pandemic in Indonesia using Twitter Data, in 2021 Proceedings - 2021 8th International Conference on Advanced Informatics: Concepts, Theory, and Application, ICAICTA 2021.
29	Felicia, I.K., and Budi, I. (2021).Classification of Inundation Level using Tweets in Indonesian Language. In ACM International Conference Proceeding Series, pp. 137–143.
30	Fransisca, D., Sulistyowati, I., Budi, I., Santoso, A.B., Putra, P.K. (2021). Sentiment Analysis of Office Automation Application in One of Indonesian Ministries, in Proceedings - International Conference on Informatics and Computational Sciences 2021- November, pp. 181-186.
31	Jonathan, B., and Budi, I. (2021).The Use of Text Mining for Classification of Product Selling Content in Social Media Female Daily. In 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICACSIS53237.2021.9631329.
32	Nababan, A.H., Mahendra, R., Budi, I. (2021).Twitter stance detection towards Job Creation Bill, in 2021 Procedia Computer Science 197, pp. 76-81

<b>No</b>	<b>Artikel</b>
33	Nurdeni, D.A., Budi, I., Santoso, A.B. (2021). Sentiment Analysis on Covid19 Vaccines in Indonesia: From the Perspective of Sinovac and Pfizer, in 3rd 2021 East Indonesia Conference on Computer and Information Technology, EIconCIT 2021 9431852, pp. 122-127
34	Putri, S.D.A., Ibrohim, M.O., Budi, I. (2021). Abusive language and hate speech detection for Javanese and Sundanese languages in tweets: Dataset and preliminary study, in 2021 11th International Workshop on Computer Science and Engineering, WCSE 2021, pp. 461-465.
35	Putri, S.D.A., Ibrohim, M.O., Budi, I. (2021). Abusive Language and Hate Speech Detection for Indonesian-Local Language in Social Media Text, in 2021 Lecture Notes in Networks and Systems 251, pp. 88-98
36	Ramdhani, T.W., Budi, I., Purwandari, B. (2021). Optical Character Recognition Engines Performance Comparison in Information Extraction, in 2021 International Journal of Advanced Computer Science and Applications 12(8), pp. 120-127
37	Saputro, B., Ma'mun, S., Budi, I., Santoso, A.B., Putra, P.K. (2021). Customer churn factors detection in Indonesian postpaid telecommunication services as a supporting medium for preventing waste of IT component, in IIOF Conference Series: Earth and Environmental Science 700(1),012015.
38	Solihin, F., Budi, I., Aji, R.F., Makarim, E. (2021). Advancement of information extraction use in legal documents, in 2021 International Review of Law, Computers and Technology 35(3), pp. 322-351.



<b>No</b>	<b>Artikel</b>
39	Suryono, R.R., Budi, I., Purwandari, B. (2021). Detection of fintech P2P lending issues in Indonesia, <i>Heliyon</i> 7(4),e06782
40	Utomo, D., Damanik, N., Budi, I. (2021). Classification on Participants Renewal Process in Insurance Company: Case Study PT XYZ, in 9th International Conference on Information and Communication Technology, <i>ICoICT 2021</i> , pp. 576-581
41	Wiraguna, A., Komarodin, P., Budi, I (2021). Twitter Sentiment Analysis of Indonesia Internet Service Provider: A Case Study of Indihome and Firstmedia, in 2021 Proceedings - 2nd International Conference on Computer Science and Engineering: The Effects of the Digital World After Pandemic (EDWAP), <i>IC2SE 2021</i> .
42	Wisnu Prabowo, M., Budi, I., Budi Santoso, H. (2021). Developing Question Generation System for Bahasa Indonesia Using Indonesian Standard Language Regulation, in <i>ACM International Conference Proceeding Series</i> pp. 258-261.
43	Yaniasih, Y., Budi, I. (2021). Systematic design and evaluation of a citation function classification scheme in Indonesian journals, in <i>Publications</i> 9(3),27.
44	Yaniasih, Y., Budi, I. (2021). Analysis of in-text citation patterns in local journals for ranking scientific documents, in <i>DESIDOC Journal of Library and Information Technology</i> 41(2), pp. 94-101.
45	Al Hakim, S., Sensuse, D.I., Budi, I. (2020). Conceptual model smart knowledge mapping with process and activity combination quadrant: Finalization and implementation, in <i>Journal of High Technology Management Research</i> 31(2),100393

<b>No</b>	<b>Artikel</b>
46	Alfina, I., Zeman, D., Dinakaramani, A., Budi, I., Suhartanto, H. (2020). Selecting the UD v2 Morphological Features for Indonesian Dependency Treebank, in 2020 International Conference on Asian Language Processing, IALP 2020 9310513, pp. 104-109.
47	Arafat, T.A., Budi, I., Mahendra, R., Salehah, D.A. (2020). Demographic Analysis of Candidates Supporter in Twitter during Indonesian Presidential Election 2019, in 7th International Conference on ICT for Smart Society: A IoT for Smart Society, ICISS 2020 – Proceeding 9307598.
48	Dewantara, D.S., Budi, I. (2020). Combination of LSTM and CNN for Article-Level Propaganda Detection in News Articles, in 5th International Conference on Informatics and Computing, ICIC 2020 9288532
49	Dewi, T.B.T., Indrawan, N.A., Budi, I., Santoso, A.B., Putra, P.K (2020). Community Understanding of the Importance of Social Distancing Using Sentiment Analysis in Twitter, in 3rd International Conference on Computer and Informatics Engineering, IC2IE 2020.
50	Eksa Permana, M., Ramadhan, H., Budi, I., Budi Santoso, A., Kresna Putra, P. (2020). Sentiment analysis and topic detection of mobile banking application review, in 5th International Conference on Informatics and Computing, ICIC 2020 9288616
51	Gustisa Wisnu, G.R., Ahmadi, Muttaqi, A.R., Santoso, A.B., Putra, P.K., Budi, I. (2020). Sentiment analysis and topic modelling of 2018 central java gubernatorial election using twitter data, in International Workshop on Big Data and Information Security, IWBS 2020 9255583, pp. 35-40

<b>No</b>	<b>Artikel</b>
52	Kurniawan, S., Budi, I. (2020). Indonesian tweets hate speech target classification using machine learning, in 5th International Conference on Informatics and Computing, ICIC 2020 9288515.
53	Laily, I.L., Budi, I., Santoso, A.B., Putra, P.K. (2020). Mining Indonesia Tourism's Reviews to Evaluate the Services through Multilabel Classification and LDA, in Proceedings of the International Conference on Electrical Engineering and Informatics 2020-October,9315392
54	Maryati, I., Purwandari, B., Budi Santoso, H., Budi, I. (2020). Implementation Strategies for Adopting Digital Library Research Support Services in Academic Libraries in Indonesia, in Proceedings - 2nd International Conference on Informatics, Multimedia, Cyber, and Information System, ICIMCIS 2020 9354327, pp. 188-194.
55	Suroso, H., Budi, I., Santoso, A.B., Putra, P.K. (2020). Sentiment Analysis on 'Homecoming Tradition Restriction' Policy on Twitter, in 3rd International Conference on Computer and Informatics Engineering, IC2IE 2020 9274609, pp. 75-79
56	Suryono, R.R., Budi, I., Purwandari, B. (2020). Challenges and trends of financial technology (Fintech): A systematic literature review, in 2020 Information (Switzerland) 11(12),590, pp. 1-20
57	Trivonanda, R., Mahendra, R., Budi, I., Hidayat, R.A. (2020). Sequential pattern mining for e-commerce recommender system, in International Conference on Advanced Computer Science and Information Systems, ICACSIS 2020 9263192, pp. 393-398

<b>No</b>	<b>Artikel</b>
58	Nasiri, D. F. dan Budi, I. (2019). Aspect Category Detection on Indonesian E-commerce Mobile Application Review, in 2019 International Conference on Data and Software Engineering (ICoDSE), Pontianak, Indonesia, pp. 1-6, doi: 10.1109/ICoDSE48700.2019.9092619.
59	Pakaya, F. N., Ibrohim, M. O., dan Budi, I. (2019). Malicious Account Detection on Twitter Based on Tweet Account Features using Machine Learning, in Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985840.
60	Sazany, E., dan Budi, I. (2019). Hate Speech Identification in Text Written in Indonesian with Recurrent Neural Network, in International Conference on Advanced Computer Science and information Systems (ICAC SIS), Bali, Indonesia, 2019, pp. 211-216, doi: 10.1109/ICAC SIS47736.2019.8979959.
61	Suciati, A., and Budi, I. (2019). Aspect-based Opinion Mining for Code-Mixed Restaurant Reviews in Indonesia. In 2019 International Conference on Asian Language Processing (IALP), Shanghai, China, 2019, pp. 59-64, doi: 10.1109/IALP48816.2019.9037689.
62	Ibrohim, M.O., dan Budi, I.(2018).A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media, Procedia Computer Science Volume 135, 2018, Pages 222-229

<b>No</b>	<b>Artikel</b>
63	Izzah, N., Budi, I., and Louvan, S. (2018). Classification of pornographic content on Twitter using support vector machine and Naive Bayes. In 2018 4th International Conference on Computer and Technology Applications (ICCTA), Istanbul, Turkey, 2018, pp. 156-160, doi: 10.1109/CATA.2018.8398674.
64	Mihardi, M. dan Budi, I. (2018). Public Sentiment on Political Campaign Using Twitter Data in 2017 Jakarta's Governor Election. In 2018 International Conference on Applied Information Technology and Innovation (ICAITI), Padang, Indonesia, 2018, pp. 67-72, doi: 10.1109/ICAITI.2018.8686740.
65	Mustafa, A. A., and Budi, I. (2018). Recommendation System Based on Item and User Similarity on Restaurants Directory Online. In 2018 6th International Conference on Information and Communication Technology (ICoICT), Bandung, Indonesia, 2018, pp. 70-74, doi: 10.1109/ICoICT.2018.8528775.
66	Pratiwi, N. I., Budi, I., and Alfina, I. (2018). Hate Speech Detection on Indonesian Instagram Comments using FastText Approach. In 2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Yogyakarta, Indonesia, 2018, pp. 447-450, doi: 10.1109/ICACSIS.2018.8618182.