

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dinamika penyampaian pendapat dan tanggapan publik terhadap berbagai kebijakan pemerintah kini semakin terfasilitasi melalui kehadiran media sosial, dengan TikTok sebagai salah satu platform yang paling diminati. Menurut temuan laporan Digital 2025: Indonesia oleh DataReportal, jumlah akun aktif TikTok dari kalangan dewasa (18 tahun ke atas) di tanah air menyentuh angka kurang lebih 108 juta pada awal 2025. Angka tersebut mengindikasikan pergeseran fungsi platform ini menjadi arena interaksi digital yang aktif, di mana audiens tidak lagi sebatas menjadi pembaca pasif, melainkan turut aktif memberikan respons, sokongan, ataupun penolakan terhadap isu aktual melalui fitur komentar. Potensi teks digital yang dihasilkan pengguna ini, seperti dijelaskan oleh (Apriani dkk., 2024), , sangat memadai untuk dijadikan basis data penelitian. Dengan demikian, TikTok menawarkan peluang yang strategis dalam mengkaji fluktuasi opini masyarakat terhadap berbagai isu sosial yang sedang mengemuka.

Perbincangan hangat di kalangan warganet TikTok turut menyoroti aturan baru mengenai pelaksanaan tugas kedinasan dari rumah atau Work From Home setiap hari Jumat bagi Aparatur Sipil Negara. Regulasi ini didasarkan pada Surat Edaran Menteri Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 3 Tahun 2026 yang mengatur mekanisme kerja pegawai pemerintah dalam kerangka akselerasi transformasi tata kelola birokrasi. Ditetapkan pada 30 Maret 2026, aturan tersebut resmi berlaku efektif mulai 1 April 2026 (Kementerian PANRB, 2026). Melalui kebijakan ini, pola kerja pegawai dibagi menjadi empat hari masuk kantor dari Senin sampai Kamis, serta satu hari bekerja secara fleksibel dari rumah setiap hari Jumat. Berbagai pro dan kontra yang bermunculan dari masyarakat maupun kalangan pegawai negeri sendiri menjadikan wacana ini sangat relevan dan menarik untuk dikaji secara mendalam, terutama melihat bagaimana diskursus tersebut berkembang di ruang digital.

Pengumpulan data dalam studi ini dilakukan melalui teknik *web scraping* yang menyorot bagian komentar dari video TikTok terkait aturan WFH Jumat bagi pegawai negeri. Berbagai data mentah yang terkumpul kemudian harus melalui serangkaian pembersihan teks guna membersihkan *noise* sekaligus mempersiapkan korpus sebelum masuk ke tahap pemodelan. Pentingnya langkah persiapan ini diperkuat oleh pandangan (Rifaldi dkk., 2023) yang menilai bahwa tahap tersebut memegang peran esensial sebelum analisis lanjutan dilakukan. Tahap berikutnya melibatkan penggunaan IndoBERT untuk pelabelan sentimen secara otomatis. (Koto dkk., 2020) mengembangkan IndoBERT sebagai model bahasa yang sudah dilatih sebelumnya untuk memproses bahasa Indonesia. Hasil pelabelan tersebut kemudian dikonversi ke dalam representasi TF-IDF guna menimbang tingkat kepentingan suatu kata. Pendekatan TF-IDF yang dikombinasikan dengan metode SMOTE pernah dimanfaatkan oleh (Wahid dkk., 2025a) dalam analisis sentimen berbahasa Indonesia; SMOTE berfungsi mengatasi ketidakseimbangan kelas data, sementara metode Chi-Square dipakai untuk menyaring fitur-fitur yang paling relevan. Penerapan Chi-Square dalam seleksi fitur analisis sentimen berbasis data TikTok juga pernah dilakukan oleh (Purnamawati dkk., 2023). Guna melakukan klasifikasi sentimen, penelitian ini mengaplikasikan algoritma *Multinomial Naive Bayes*. Penggunaan *Multinomial Naive Bayes* bersama pembobotan TF-IDF untuk klasifikasi teks berbahasa Indonesia turut didukung oleh (Wibawa & Pramarta, 2023). Terakhir, kinerja model dievaluasi memakai sejumlah metrik pengujian, yakni *accuracy*, *precision*, *recall*, dan *F1-score*.

Fleksibilitas algoritma Naive Bayes dalam ranah analisis sentimen telah terbukti melalui pemanfaatannya pada berbagai macam objek riset. Sebagai contoh, (Anugrah dkk., 2025) menerapkan metode ini untuk menelaah sentimen warganet di platform media sosial X. Penelitian (Muhtadi dkk., 2025) menggunakan algoritma serupa guna mengkaji ulasan konsumen pada lokapasar Tokopedia. Ragam aplikasi lainnya juga terlihat dalam penelitian (Musaad dkk., 2024) yang meneliti respons publik terhadap Program Rekrutmen Bersama BUMN, serta penelitian (Andrian & Isnain, 2024) yang mengklasifikasikan sentimen terkait TikTok Shop di Twitter dengan tingkat akurasi mencapai 78,22%. Deretan temuan tersebut menegaskan bahwa Naive Bayes memiliki

keandalan yang baik dalam mengelompokkan sentimen dari bermacam-macam karakteristik data yang berbeda.

Meskipun penelitian sebelumnya sudah menggunakan Naive Bayes untuk berbagai objek, belum ditemukan penelitian yang khusus menganalisis komentar TikTok tentang kebijakan kerja dari rumah di hari Jumat bagi pegawai negeri sipil. Penelitian sebelumnya belum menemukan penggunaan metode pelabelan otomatis berbasis IndoBERT yang digabungkan dengan TF-IDF, SMOTE, Chi-Square, dan Multinomial Naive Bayes pada objek tersebut. Sebab itu, penelitian ini dilakukan untuk menganalisis kecenderungan sentimen komentar pengguna TikTok terhadap kebijakan WFH hari Jumat bagi ASN serta mengevaluasi kemampuan model Multinomial Naive Bayes dalam mengklasifikasikan komentar ke dalam kategori positif, negatif, atau netral. Hasil yang diharapkan adalah mendapatkan model klasifikasi yang memiliki performa yang baik, serta gambaran mengenai respons pengguna TikTok terhadap kebijakan WFH hari Jumat bagi ASN.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana prosedur pengunduhan dan penghimpunan data teks dari kolom komentar TikTok yang menyoroti aturan *work from home* (WFH) hari Jumat bagi pegawai negeri?
2. Bagaimana bentuk penerapan kombinasi pembobotan kata TF-IDF serta algoritma dalam melakukan analisis sentimen pada komentar TikTok tersebut?
3. Apa saja hasil klasifikasi akhir yang diperoleh ketika opini komentar TikTok terkait kebijakan WFH Jumat bagi ASN dikelompokkan ke dalam kategori positif, negatif, dan netral??
4. Bagaimana tingkat efektivitas dan unjuk kerja dari model klasifikasi yang diuji melalui parameter *Accuracy*, *Precision*, *Recall*, serta *F1-Score*?

1.3 Tujuan

Tujuan dari penelitian ini adalah:

1. Mengetahui proses pengumpulan data komentar TikTok terhadap Kebijakan WFH hari Jumat bagi ASN.
2. Menerapkan metode TF-IDF dan Multinomial Naïve Bayes dalam analisis sentimen komentar TikTok.
3. Mengklasifikasikan komentar TikTok terhadap kebijakan WFH hari Jumat bagi ASN ke dalam kategori positif, negatif, dan netral.
4. Mengukur performa model TF-IDF dan Multinomial Naïve Bayes berdasarkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat baik secara akademis, praktis, maupun teknis.

1.4.1 Manfaat Akademis

Diharapkan penelitian ini akan menjadi referensi ilmiah untuk pengembangan penelitian di bidang *Text Mining*, *Natural Language Processing* (NLP), dan Analisis Sentimen, khususnya yang menggunakan metode Naive Bayes untuk memanfaatkan data dari media sosial TikTok..

1.4.2 Manfaat Praktis

Penelitian ini diharapkan dapat memberikan informasi mengenai pengguna TikTok terhadap kebijakan *Work From Home* hari Jumat bagi Aparatur Sipil Negara berdasarkan analisis komentar pada media sosial TikTok. Hasil penelitian dapat menjadi salah satu masukan bagi pemerintah dalam melakukan evaluasi terhadap implementasi kebijakan tersebut.

1.4.3 Manfaat Teknis

Penelitian ini diharapkan dapat menjadi contoh penggunaan teknik web scraping menggunakan Bright Data untuk mengumpulkan data media sosial serta penggunaan metode Naive Bayes untuk melakukan klasifikasi sentimen pada data teks berbahasa Indonesia.

1.5 Ruang Lingkup Masalah

Agar penelitian lebih terarah, maka ruang lingkup penelitian ini dibatasi sebagai berikut:

1. Penelitian ini membahas analisis sentimen komentar TikTok terkait kebijakan WFH hari Jumat bagi ASN.
2. Penelitian ini hanya menggunakan satu video TikTok milik akun @detik.com sebagai sumber data karena video tersebut dipilih secara purposif berdasarkan kriteria tertentu. Video yang digunakan merupakan salah satu video dengan jumlah komentar terbanyak dan paling representatif dalam membahas kebijakan WFH hari Jumat bagi ASN pada periode penelitian.
3. Periode pengambilan data ditetapkan sejak video dipublikasikan pada 31 Maret 2026 hingga proses pengambilan data pada 7 Juli 2026
4. Pelabelan sentimen dilakukan menggunakan IndoBERT kemudian data disaring berdasarkan *confidence score* sebesar 0,90.
5. Analisis dilakukan melalui tahapan *text preprocessing*, pelabelan otomatis menggunakan IndoBERT, pembobotan TF-IDF, penyeimbangan data menggunakan SMOTE, seleksi fitur Chi-Square, dan klasifikasi menggunakan algoritma Multinomial Naive Bayes.
6. Sentimen diklasifikasikan menjadi 3 kelas, yaitu positif, negatif, dan netral.
7. Evaluasi model menggunakan *Accuracy, Precision, Recall, dan F1-Score*.

1.6 Kebaharuan

Penelitian ini memiliki beberapa aspek kebaruan dibandingkan penelitian terdahulu, yaitu sebagai berikut.

1. Penelitian ini menggunakan komentar pada media sosial TikTok sebagai sumber data untuk menganalisis sentimen masyarakat terhadap kebijakan publik, sementara mayoritas penelitian sebelumnya memakai data dari platform Twitter atau X.

2. Dataset penelitian diperoleh melalui proses *web scraping* dengan Bright Data, sehingga memungkinkan pengumpulan data komentar secara otomatis dari video TikTok yang menjadi objek penelitian.
3. Penelitian ini menitikberatkan pada kebijakan Work From Home (WFH) Hari Jumat bagi Aparatur Sipil Negara, yakni isu kebijakan publik yang tengah menarik perhatian masyarakat; oleh karena itu, hasil penelitian diharapkan dapat memberikan gambaran mengenai kondisi opini public terutama pengguna TikTok pada saat kebijakan tersebut diterapkan.
4. Penelitian ini menggunakan metode Naive Bayes dalam proses klasifikasi sentimen komentar TikTok untuk mengukur tingkat performa algoritma pada data teks berbahasa Indonesia yang berkaitan dengan kebijakan publik.

1.7 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian, batasan masalah atau ruang lingkup, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini memuat teori-teori yang mendukung penelitian, seperti analisis sentimen, text mining, Natural Language Processing, TikTok, TF-IDF, Multinomial Naïve Bayes, serta penelitian terdahulu yang relevan.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang digunakan, meliputi pengumpulan data, preprocessing, ekstraksi fitur, klasifikasi sentimen, dan evaluasi model.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil pengolahan data, hasil klasifikasi sentimen, serta evaluasi performa model yang digunakan dalam penelitian.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian dan saran untuk penelitian selanjutnya.