

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Saat ini Indonesia berada di bawah kepemimpinan Presiden Prabowo Subianto yang dilantik pada 20 Oktober 2024 (BPMI Setpres, 2024), situasi kondisi sosial-politik Indonesia dalam tahun pertama pemerintahannya diwarnai dinamika kebijakan yang memicu respons publik luas Gelombang unjuk rasa “Indonesia Gelap” pada 21 Februari 2025 menyoroti 17 tuntutan terkait agenda legislasi dan kebijakan pemerintah (Tirto.id, 2025). Pada Maret 2025, aksi penolakan terhadap revisi UU TNI dan pembahasan RUU Polri turut mencuat, alasan penolakan yang kerap dikemukakan di ruang publik antara lain kekhawatiran perluasan kewenangan dan implikasinya terhadap kebebasan sipil serta tata kelola keamanan yang akuntabel (Tirto.id, 2025)(Tempo, 2025). Di ranah perdagangan internasional, Amerika Serikat mengumumkan tarif impor 19% untuk sejumlah produk Indonesia mulai 7 Agustus 2025, yang memantik diskusi mengenai daya saing ekspor dan konsekuensi lanjutan bagi pelaku usaha (Kompas, 2025). Pada Agustus 2025, polemik tunjangan DPR memicu protes di sejumlah kota, keberatan publik umumnya berkisar pada ketimpangan prioritas anggaran, kepekaan sosial di tengah tekanan ekonomi, serta transparansi dan akuntabilitas penggunaan fasilitas wakil rakyat (Detik.com, 2025). Rangkaian isu dari tuntutan kebijakan, perdebatan revisi sektor pertahanan-keamanan, hingga tekanan eksternal seperti tarif impor ini membentuk konteks ketidakpastian yang relevan dengan kecemasan (*anxiety*) warga negara Indonesia.

Kecemasan dimaknai sebagai kumpulan respons afektif–fisik yang mencakup rasa khawatir, tegang, takut, dan gejala somatik (seperti palpitasi, sesak, pusing), yang berakar pada mekanisme *fight-or-flight*. Faktor pemicunya mencakup pengalaman hidup yang menekan (konflik keluarga atau pekerjaan, beban akademik, tekanan finansial), pola hidup, kondisi kesehatan fisik/jiwa komorbid,

serta obat/zat (mind.org.uk, 2021)(Bragdon, 2024). Indikasi kecemasan dapat dibagi tiga, yaitu Indikasi kecemasan kognitif, fisik, dan perilaku. Dalam konteks sosial dan politik modern, terutama pada situasi penuh ketidakpastian, ekspresi kecemasan tidak hanya tampak dalam bentuk keluhan fisik atau emosional, tetapi juga terekam melalui perilaku dan ekspresi linguistik masyarakat di media sosial (Attia et al., 2022).

Dalam konteks kebijakan publik, kekhawatiran terhadap arah kebijakan negara, dampak ekonomi, dan isu keamanan/lingkungan dapat memperbesar rasa tidak pasti, sehingga memicu ekspresi kecemasan di ruang publik, termasuk pada media sosial. Pada tataran komunikasi digital, platform X (sebelumnya Twitter) tetap menjadi arena percakapan *real time* mengenai isu kebijakan. X merupakan hasil rebrand Twitter sejak Juli 2023 dan mempertahankan karakter *microblogging* dengan linimasa vertikal serta interaksi beruntun (*threads*) (Kompas, 2022)(TIME, 2023). Di Indonesia, platform ini sering menjadi kanal opini kebijakan dan respons publik, sehingga wajar dijadikan sumber data untuk mengekstrak indikator emosi, termasuk sinyal kecemasan.

Pemantauan kecenderungan kecemasan secara dini penting karena dapat membantu pengambil kebijakan memahami dampak psiko-sosial kebijakan, sekaligus memberi sinyal intervensi yang diperlukan. Literatur mutakhir menunjukkan bahwa tanda-tanda kondisi kesehatan mental dapat direfleksikan melalui bahasa alami di media sosial dan dapat dipelajari secara komputasional (Putri & Setiawan, 2025). Selain itu, studi klinis dan populasi memperlihatkan bahwa skrining berbasis gejala serta faktor gaya hidup dapat dipadukan dengan pembelajaran mesin (*machine learning*) untuk memperkirakan status kesejahteraan mental (Putri & Setiawan, 2025)(Ku & Min, 2024).

Secara teknis, beberapa algoritma pembelajaran mesin telah banyak dipakai untuk deteksi/indikasi masalah kesehatan mental dari teks, seperti Naïve Bayes (NB), Random Forest (RF), dan Long Short-Term Memory (LSTM). Naïve Bayes banyak dipakai pada data teks berfitur tinggi karena efisien, sederhana, dan skalabel

cocok untuk volume besar unggahan media sosial (Sathvika et al., 2023). Di sisi lain, sejumlah studi menemukan keunggulan Random Forest pada tugas klasifikasi kesehatan mental. Pada klasifikasi depresi atau kecemasan berbasis cuitan, (Rahayu et al., 2023) melaporkan Random Forest mencapai akurasi hingga 96% pada skema bagi data 80:20 (Rahayu et al., 2023). Penelitian lain menekankan pentingnya mengevaluasi ketahanan model (*robustness*) terhadap noise atau bias jawaban subyektif. Dalam skenario demikian, Random Forest termasuk algoritma yang stabil saat dibandingkan bersama CNN, XGBoost, dan regresi logistik (Ku & Min, 2024). Pada pendekatan *deep learning*, Long Short-Term Memory efektif menangkap dependensi jangka panjang dalam teks, sehingga sering unggul pada deteksi kondisi mental berbasis bahasa alami. Pada penelitian (Nurochman & Ashiilah, 2024), menunjukkan Long Short-Term Memory lebih unggul daripada SVM pada tiga metrik utama (*precision, recall, F1-score*), meskipun SVM unggul tipis pada akurasi. Sejalan dengan itu, penelitian lain berbasis data bahasa Indonesia juga memposisikan arsitektur LSTM sebagai kandidat kuat untuk deteksi depresi dari teks di platform X (Putri & Setiawan, 2025).

Beberapa penelitian sebelumnya telah berupaya mendeteksi atau memprediksi kecemasan (*anxiety*) menggunakan pendekatan *machine learning* dan *deep learning*. Namun, penelitian oleh (Nurochman & Ashiilah, 2024) masih berfokus pada prediksi kecemasan secara umum, tanpa mengklasifikasikan lebih lanjut ke dalam dimensi *anxiety* seperti kognitif, fisik, dan perilaku. Selain itu, penelitian oleh (Putri & Setiawan, 2025), (Torres Cruz, 2023), dan (Abdul Rahman et al., 2023) belum melakukan perbandingan komprehensif antara tiga algoritma yaitu Random Forest, Naïve Bayes, dan Long Short-Term Memory secara bersamaan, sehingga belum dapat diketahui algoritma mana yang memiliki kinerja paling optimal. Di sisi lain, penelitian yang dilakukan oleh (Rizki Rinaldi & Basysyar, 2025) hanya menggunakan evaluasi terbatas melalui *confusion matrix* dan pembagian data *train-test split*, tanpa penerapan metode validasi yang lebih kuat seperti *cross validation* untuk menguji konsistensi performa model.

Berdasarkan kekurangan tersebut, penelitian ini bertujuan untuk mengisi celah dengan membandingkan kinerja tiga algoritma, yaitu Random Forest, Naïve Bayes, dan Long Short-Term Memory (LSTM) secara bersamaan dalam indikasi kecemasan (*anxiety*) masyarakat Indonesia di platform X berdasarkan dimensi kognitif, fisik, dan perilaku, dengan menerapkan evaluasi yang lebih menyeluruh menggunakan *confusion matrix*, *train-test split*, dan *cross validation* untuk memperoleh hasil yang lebih akurat dan stabil.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dijelaskan, identifikasi masalah dalam penelitian ini adalah :

1. Bagaimana proses pengumpulan dan pra-pemrosesan data teks berbahasa Indonesia dari platform X agar dapat digunakan untuk mendeteksi indikasi *anxiety*?
2. Bagaimana kinerja algoritma Naïve Bayes, Random Forest, dan LSTM dalam mengindikasikan *anxiety* pada unggahan berbahasa Indonesia di platform X?
3. Bagaimana perbandingan performa algoritma Naive Bayes, Random Forest, dan LSTM dalam mendeteksi indikasi *anxiety* pada teks berbahasa Indonesia di platform X berdasarkan hasil evaluasi menggunakan *confusion matrix*, *train-test split*, dan *cross validation*?

## 1.3 Tujuan

Tujuan dari penelitian ini adalah :

1. Penelitian ini mengumpulkan dan memproses data teks berbahasa Indonesia dari platform X yang berkaitan dengan isu atau kebijakan publik sehingga dapat digunakan sebagai dataset untuk mendeteksi indikasi *anxiety*.
2. Penelitian ini menerapkan dan menguji kinerja tiga algoritma, yaitu Naïve Bayes, Random Forest, dan Long Short-Term Memory (LSTM) dalam

mengindikasikan *anxiety* pada teks media sosial berbahasa Indonesia di platform X.

3. Penelitian ini membandingkan performa ketiga algoritma berdasarkan *confusion matrix*, *train-test split* dan *K-fold cross validation* sehingga dapat diketahui algoritma yang paling optimal.

#### 1.4 Batasan Masalah

Untuk menjaga fokus dan ruang lingkup penelitian agar lebih terarah, maka batasan masalah dalam penelitian ini adalah :

1. Sumber data terbatas pada unggahan publik berbahasa Indonesia di platform X yang berkaitan dengan isu/kebijakan nasional pada rentang waktu yang ditetapkan peneliti.
2. Objek analisis berada pada level unggahan (*post-level*) sebagai indikasi kecemasan, penelitian ini tidak dimaksudkan untuk diagnosis klinis.
3. Metode yang digunakan hanya mencakup tiga algoritma, yaitu Naïve Bayes, Random Forest, dan Long Short-Term Memory (LSTM).
4. Evaluasi menggunakan *confusion matrix*, *train-test split* dan *cross validation*.

#### 1.5 Kontribusi

Kontribusi dari penelitian ini adalah :

1. Memperkuat dasar teoritis dalam penerapan *machine learning* dan *deep learning* untuk menganalisis teks, khususnya pada indikasi *anxiety* (kecemasan).
2. Menyediakan evaluasi komparatif tiga algoritma utama Naïve Bayes, Random Forest, dan Long Short Term Memory pada tugas klasifikasi teks. Analisis performa yang melibatkan *confusion matrix*, *train-test split* dan *cross validation* memberikan pemahaman mendalam mengenai kelebihan

dan kekurangan masing-masing algoritma, serta menjadi referensi dalam pemilihan model untuk penelitian selanjutnya.

3. Menghubungkan bidang ilmu komputer dengan psikologi komputasional, di mana algoritma klasifikasi digunakan untuk mendeteksi indikasi *anxiety* (kecemasan) dalam konteks sosial-politik masyarakat Indonesia.

