

BAB I

PENDAHULUAN

1.1 Latar Belakang

Natural Language Processing (NLP) adalah salah satu cabang dari *Artificial Intelligence* (AI) yang berfokus pada pemrosesan bahasa alami. Bahasa alami sendiri merupakan bahasa yang digunakan manusia dalam kehidupan sehari-hari untuk berkomunikasi. Sementara itu, komputer memerlukan proses khusus agar dapat memahami maksud atau tujuan dari bahasa yang digunakan oleh manusia Rahmadani & Tasrif, (2023). Kemampuan NLP dalam menganalisis bahasa alami menjadi semakin relevan ketika diterapkan pada percakapan digital yang berlangsung di media sosial. Dalam konteks komunikasi modern, NLP memungkinkan sistem komputer memahami opini, emosi, hingga maksud tersembunyi yang disampaikan melalui *text* oleh pengguna. Teknologi ini sangat penting dalam mengolah data percakapan yang kompleks, singkat, dan sering kali tidak terstruktur, seperti yang umum ditemukan di platform digital Rivaldi dkk., (2024). Dengan kemampuannya dalam menangkap pola bahasa, mengenali konteks, serta menginterpretasikan bahasa informal, NLP berperan strategis dalam memahami dinamika percakapan publik di ruang digital.

Peran tersebut semakin relevan seiring dengan berkembangnya media sosial sebagai wadah utama komunikasi daring, salah satunya adalah platform *X*. Platform ini menjadi ruang ekspresi masyarakat yang dinamis, memungkinkan warganet menyampaikan opini, mengkritisi kebijakan, atau mendiskusikan isu-isu yang sedang berkembang. Dengan karakteristiknya yang terbuka, cepat, dan singkat, *X* menjadi wadah utama untuk menyuarakan aspirasi dan emosi masyarakat. Dalam konteks ini, media sosial tidak hanya berperan sebagai sarana hiburan, tetapi juga sebagai instrumen penting dalam mencerminkan fenomena sosial-politik di masyarakat (Winahyu & Suharjo, 2021).

Fenomena yang menarik perhatian publik belakangan ini adalah munculnya tagar *#KaburAjaDulu* yang viral di media sosial. Tagar ini digunakan oleh banyak pengguna

untuk mengekspresikan keresahan mereka terhadap situasi ketenagakerjaan, tekanan sosial, hingga impian untuk mencari peluang yang lebih baik di luar negeri. Viralnya tagar tersebut menunjukkan adanya keresahan kolektif dan potensi krisis sosial yang perlu dipahami secara lebih dalam menurut Rachman & Pramana (2020). Hal ini menunjukkan bahwa media sosial tidak hanya merefleksikan tren sesaat, tetapi juga dapat berfungsi sebagai indikator awal terhadap isu-isu struktural yang dihadapi oleh masyarakat.

Melimpahnya opini publik yang tersebar di media sosial, khususnya dalam bentuk *cuitan* dan penggunaan tagar, menjadikan platform tersebut sebagai sumber data yang sangat penting dan relevan untuk dianalisis. Analisis terhadap opini-opini tersebut dapat memberikan gambaran mengenai persepsi, sikap, serta harapan masyarakat terhadap suatu isu tertentu. Pemanfaatan data media sosial dalam penelitian dapat membantu akademisi, pembuat kebijakan, maupun praktisi dalam memahami dinamika sosial secara lebih cepat, aktual, dan kontekstual. Oleh karena itu, diperlukan kajian yang sistematis untuk menelaah bagaimana respons masyarakat terhadap berbagai isu sosial diekspresikan melalui media sosial.

Mengingat besarnya volume data teks yang dihasilkan, seperti *cuitan*, diperlukan pendekatan analisis yang efektif dan efisien untuk mengolah serta mengekstraksi informasi yang terkandung di dalamnya. Salah satu metode yang telah terbukti dalam klasifikasi *text* pendek adalah algoritma *Multinomial Naïve Bayes*. Algoritma ini bekerja dengan prinsip probabilistik dan mampu menangani fitur-fitur dalam *text* seperti kata kunci atau frekuensi kemunculan kata. Karena *tweet* bersifat singkat dan padat, metode ini sangat cocok digunakan untuk mengelompokkan opini menjadi kategori tertentu, seperti sentimen positif atau negatif serta tema diskusi yang dominan (Ghimire, 2022).

Algoritma *Multinomial Naïve Bayes* didasarkan pada penerapan *Teorema Bayes* dengan anggapan bahwa setiap fitur dalam data saling bebas atau tidak bergantung satu sama lain Puspita Sari dkk., (2025). Menurut Hasri & Alita, (2022) walaupun metode ini tergolong sederhana, *Multinomial Naïve Bayes* telah terbukti efektif dalam

menangani tugas klasifikasi *text* dalam skala besar, seperti analisis sentimen, deteksi spam, hingga sistem rekomendasi.

Sejalan dengan permasalahan yang diuraikan sebelumnya, penelitian ini bertujuan untuk mengungkap kecenderungan opini masyarakat terhadap isu sosial, khususnya keresahan ketenagakerjaan yang tercermin dalam tagar #KaburAjaDulu, dengan menggunakan pendekatan klasifikasi *text* berbasis *Multinomial Naïve Bayes*. Dengan memahami pola-pola opini yang tersebar di media sosial, penelitian ini diharapkan dapat memberikan kontribusi dalam memetakan persepsi publik secara lebih akurat, serta menjadi bahan pertimbangan dalam merumuskan kebijakan yang lebih responsif terhadap kebutuhan dan aspirasi masyarakat.

1.2 Rumusan Masalah

Berdasarkan latar belakang rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana persepsi masyarakat terhadap hastag #KaburAjaDulu di media sosial *X* berdasarkan analisis sentimen?
2. Berapa nilai *accuracy*, *precision*, *recall*, dan *F1-Score* dari model klasifikasi sentimen yang dibangun menggunakan algoritma *Multinomial Naïve Bayes*?

1.3 Batasan Penelitian

Penelitian ini dibatasi agar fokus dan terarah, dengan ruang lingkup sebagai berikut:

1. Objek penelitian adalah komentar masyarakat mengenai #KaburAjaDulu.
2. Penelitian ini dibatasi pada komentar pengguna di media sosial *X* yang menggunakan tagar #KaburAjaDulu, dengan fokus pada data *text* berbahasa Indonesia yang dikumpulkan pada rentang waktu 10 Januari 2025 hingga 03 Maret 2025.
3. Menggunakan algoritma *Multinomial Naïve Bayes*.
4. Analisis sentiment hanya menggunakan dua kategori yaitu positif dan negatif dengan melakukan validasi secara manual.
5. Teknik klasifikasi yang digunakan algoritma *Multinomial Naïve Bayes*, dengan representasi fitur menggunakan *TF-IDF* dan *hiperparameter random search*.

6. Validasi model dilakukan dengan metode *K-Fold Cross Validation*, dan evaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-Score*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disampaikan, maka tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui persepsi masyarakat terhadap hastag #KaburAjaDulu di media sosial *X* melalui pendekatan analisis sentimen.
2. Untuk mengetahui performa model klasifikasi yang dibangun dengan mengukur tingkat *accuracy*, *precision*, *recall*, dan *F1-Score*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi sebagai berikut:

- a. Manfaat Teoritis:
 - Memberikan kontribusi pada literatur bidang analisis sentimen dan pemrosesan bahasa alami (NLP), khususnya dalam konteks penggunaan algoritma *Multinomial Naïve Bayes* untuk klasifikasi *text* berbahasa Indonesia.
 - Menjadi referensi bagi penelitian lanjutan yang membahas isu sosial melalui pendekatan komputasi berbasis *text*.
- b. Manfaat Praktis
 - Menyediakan data dan analisis yang dapat dimanfaatkan oleh pengambil kebijakan untuk memahami opini publik secara cepat terhadap isu ketenagakerjaan.
 - Memberikan dasar pengembangan sistem monitoring media sosial berbasis klasifikasi sentimen, yang berguna bagi lembaga pemerintah, media, atau pengembang aplikasi NLP.