

BAB I

PENDAHULUAN

1.1 Latar belakang

Pada zaman modern ini, internet telah menjadi bagian integral dari kehidupan kita sehari-hari. Penggunaan media sosial telah meluas di seluruh dunia, dengan banyak orang memanfaatkannya untuk berbagi pendapat, perasaan, pengalaman, dan berbagai hal yang menarik perhatian mereka (Troussas et al., 2013). Twitter, salah satu platform media sosial yang masih populer, memungkinkan pengguna untuk berbagi pemikiran mereka dalam bentuk "tweet" yang terbatas hingga 140 karakter. Pengguna sering menggunakan tweet untuk menyampaikan pesan singkat dengan berbagai simbol, emotikon, dan singkatan untuk memberi makna lebih pada pesan yang disampaikan. Tweet di Twitter sering kali menyentuh topik-topik yang sedang hangat dan beberapa di antaranya bahkan menjadi trending topic. Sebagai salah satu platform yang memungkinkan pengguna untuk berbagi pandangan, berita, dan pemikiran dengan cepat, Twitter memiliki lebih dari 330 juta pengguna aktif bulanan, menjadikannya sumber data yang sangat kaya dan dinamis. Platform ini sering digunakan untuk menganalisis opini publik tentang berbagai isu, mulai dari produk hingga kebijakan politik (Statista, 2021).

Analisis sentimen adalah teknik untuk mengidentifikasi opini pengguna di media sosial, yang berupaya memahami perasaan atau emosi yang melekat pada teks yang diunggah. Algoritma seperti Support Vector Machine (SVM) dan Naive Bayes banyak digunakan dalam analisis sentimen karena terbukti berguna dalam mengklasifikasikan teks dengan akurasi tinggi (Ramneet et al., 2020). Namun, masalah yang dihadapi dalam analisis data Twitter adalah banyaknya data yang harus diproses secara cepat dan akurat. Oleh karena itu, penelitian ini berfokus pada penerapan algoritma SVM dan Naive Bayes dalam klasifikasi sentimen di Twitter, untuk mengevaluasi kinerja kedua algoritma dalam konteks analisis sentimen.

Support Vector Machine (SVM) adalah algoritma yang efektif dan efisien untuk klasifikasi, terutama dalam menangani masalah pemisahan linear dan non-linier. Dalam analisis sentimen, SVM digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau sikap dalam teks (Septianingrum & Irawan, 2021). Keunggulan SVM terletak pada kemampuannya untuk menangani data dengan dimensi tinggi dan menghasilkan akurasi

yang lebih tinggi dibandingkan metode lainnya.

Seiring dengan meningkatnya kepedulian terhadap opini publik, banyak peneliti mulai menggali teknik text mining, khususnya analisis sentimen, untuk mengklasifikasikan sentimen menjadi tiga kategori: positif, negatif, dan netral. Salah satu algoritma yang sering digunakan dalam hal ini adalah Naive Bayes, yang berdasarkan asumsi bahwa atribut dalam data bersifat independen, meskipun pada kenyataannya hal ini tidak selalu berlaku. Keuntungan utama dari Naive Bayes adalah kemampuannya untuk melakukan klasifikasi dengan data pelatihan yang relatif sedikit (Rahayu et al., 2019). Salah satu isu yang sedang ramai dibicarakan adalah penilaian terhadap Pemimpin, yang dianggap kurang berhasil memimpin sektor pendidikan. Sejak menjabat sebagai Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi, Pemimpin telah menggagas berbagai kebijakan inovatif dalam bidang pendidikan. Salah satu inisiatif yang paling signifikan adalah program "Merdeka Belajar", yang dirancang untuk memberikan keleluasaan lebih kepada peserta didik dan tenaga pendidik dalam menjalani proses pembelajaran. Program ini mencakup sejumlah perubahan, termasuk revisi kurikulum, penghapusan sistem Ujian Nasional, integrasi teknologi dalam pendidikan, serta peningkatan peran guru dalam menentukan metode pengajaran yang lebih efektif. Kendati demikian, kebijakan ini mendapat tanggapan beragam dari masyarakat, dengan sebagian mendukung karena dinilai progresif, sementara yang lain mengkritik pelaksanaannya yang dianggap kurang optimal (Prasetyo et al., 2021). DPR pun menyarankan agar sektor pendidikan ditinjau kembali oleh pemerintahan yang akan datang. Topik ini menimbulkan berbagai pendapat dari masyarakat, baik yang mendukung maupun mengkritik. Pendapat-pendapat ini bisa dianalisis untuk membantu pengambilan keputusan atau kebijakan terkait sektor pendidikan. Salah satu metode yang dapat digunakan adalah analisis sentimen.

Banyak penelitian terkait analisis sentimen yang telah dilakukan sebelumnya. Salah satunya oleh Ruhyana (2019), yang menganalisis sentimen terhadap kebijakan sistem plat nomor ganjil/genap menggunakan algoritma Naive Bayes, dengan hasil akurasi 86,67%, precision 71,43%, dan recall 80,00% (Ruhyana, 2019). Penelitian lainnya oleh Rusdian dan Rosiyadi (2019) juga menganalisis sentimen terhadap tokoh publik menggunakan Naive Bayes dan SVM. Hasilnya menunjukkan bahwa SVM memberikan kinerja yang lebih baik dengan akurasi 73,96%, sedangkan Naive Bayes mencatatkan akurasi 71,94% (Rusdian & Rosiyadi, 2019). Selain itu, beberapa penelitian telah dilakukan untuk meneliti opini publik terhadap kebijakan pendidikan di Indonesia.

Analisis sentimen terhadap kebijakan yang diterapkan oleh Pemimpin telah dilakukan dengan menerapkan algoritma Naïve Bayes dan SVM. Hasil penelitian tersebut mengungkapkan bahwa Naïve Bayes memiliki tingkat akurasi lebih tinggi dibandingkan SVM dalam mengelompokkan sentimen publik di Twitter. Namun, penelitian lain yang menyoroiti opini masyarakat terkait pemindahan ibu kota Indonesia justru menunjukkan hasil yang berbeda, di mana algoritma SVM memiliki akurasi lebih unggul dibandingkan Naïve Bayes dalam analisis sentimen berbasis Twitter. Perbedaan hasil ini mengindikasikan bahwa efektivitas suatu algoritma tidak bersifat mutlak, melainkan bergantung pada berbagai faktor, seperti karakteristik dataset, metode preprocessing, serta fitur yang digunakan dalam pemodelan. Selain itu, struktur teks yang bervariasi dalam setiap topik juga dapat memengaruhi performa algoritma yang digunakan. Misalnya, dalam analisis kebijakan publik, opini masyarakat cenderung lebih kompleks dan terdiri atas berbagai aspek yang harus diperhatikan, sehingga efektivitas algoritma tertentu dapat berbeda-beda. Oleh karena itu, penelitian lebih lanjut diperlukan untuk memahami bagaimana variasi karakteristik dataset dapat memengaruhi hasil analisis sentimen menggunakan Naïve Bayes dan SVM (Widowati & Sadikin, 2021). Penelitian lain mengungkap bahwa perbedaan akurasi antara Naïve Bayes dan SVM dalam analisis sentimen sangat bergantung pada jenis data yang digunakan, terutama dalam konteks opini publik mengenai pemindahan ibu kota Indonesia. SVM cenderung lebih efektif dalam mengelola data dengan banyak fitur dan mampu memisahkan sentimen dengan lebih baik, sedangkan Naïve Bayes lebih cocok untuk teks pendek seperti tweet di Twitter. Selain itu, langkah-langkah preprocessing, seperti menghapus stopwords dan menerapkan stemming, juga berperan penting dalam meningkatkan kinerja algoritma. Jika dilakukan secara optimal, SVM dapat mengenali pola hubungan antar kata dengan lebih akurat, sementara Naïve Bayes lebih fleksibel dalam menghadapi variasi bahasa. Oleh karena itu, pemilihan algoritma harus disesuaikan dengan jenis data dan tujuan analisis agar hasilnya lebih optimal (Supian et al., 2024).

Meskipun berbagai penelitian telah dilakukan, masih terdapat beberapa gap yang harus diperhatikan dalam menganalisis sentimen terhadap kinerja Pemimpin di sektor pendidikan. Pertama, masih sedikit penelitian yang secara khusus membandingkan efektivitas algoritma Naïve Bayes dan SVM dalam mengevaluasi kebijakan pendidikan di Indonesia. Kedua, sebagian besar penelitian yang telah dilakukan menggunakan dataset yang berbeda-beda, sehingga hasil yang diperoleh sulit untuk dibandingkan secara langsung. Ketiga, diperlukan penelitian yang lebih komprehensif guna memahami

faktor-faktor yang mempengaruhi performa masing-masing algoritma, seperti penyesuaian parameter (parameter tuning), pemilihan fitur yang optimal, serta metode preprocessing yang digunakan dalam pengolahan teks dari media sosial.

Dengan mempertimbangkan gap yang telah diidentifikasi, studi ini bertujuan untuk mengisi kesenjangan tersebut melalui analisis sentimen terhadap kinerja Pemimpin di sektor pendidikan menggunakan dataset yang lebih luas dan terstruktur. Selain itu, penelitian ini juga berupaya membandingkan performa algoritma Naïve Bayes dan SVM secara langsung guna mengevaluasi efektivitas masing-masing dalam mengklasifikasikan opini masyarakat di Twitter. Diharapkan penelitian ini dapat memperdalam pemahaman tentang bagaimana pandangan publik terhadap sektor pendidikan, sehingga memberikan informasi yang bermanfaat bagi pengambil keputusan atau pembuat kebijakan. Dengan begitu, hasil penelitian ini dapat dijadikan dasar untuk merumuskan kebijakan pendidikan yang lebih sesuai dan responsif di masa mendatang.

1.2 Rumusan Masalah

1. Bagaimana performa algoritma SVM dan Naive Bayes dalam mengklasifikasikan sentimen pada data Twitter yang membahas evaluasi kinerja Pemimpin di sektor pendidikan?
2. Algoritma mana yang mampu memberikan hasil akurasi tertinggi dalam klasifikasi sentimen positif, negatif dan netral terkait evaluasi kinerja Pemimpin?
3. Apa saja tantangan yang dihadapi dalam penerapan kedua algoritma tersebut pada data Twitter yang beragam dan besar terkait evaluasi kinerja Pemimpin.

1.3 Tujuan Penelitian

1. Mengenalisis sentimen pengguna Twitter terhadap evaluasi kinerja Pemimpin di sektor pendidikan dengan memanfaatkan algoritma Support Vector Machine (SVM) dan Naive Bayes.
2. Membandingkan efektivitas algoritma SVM dan Naive Bayes dalam mengelompokkan sentimen positif, negatif, dan netral pada percakapan Twitter yang menyoroti kinerja Pemimpin di bidang pendidikan.
3. Mengidentifikasi algoritma yang paling akurat dan efisien untuk menganalisis sentimen di media sosial Twitter terkait evaluasi kinerja Pemimpin dalam ranah

pendidikan.

1.4 Batasan Masalah

1. Data yang digunakan merupakan tweet terkait evaluasi kinerja Pemimpin di sektor pendidikan yang diambil secara acak dari Twitter selama periode satu tahun terakhir (2023-2024).
2. Penelitian ini hanya akan fokus pada tiga kategori sentimen, yaitu positif, negatif, dan netral.
3. Algoritma yang diterapkan hanya terbatas pada SVM dan Naive Bayes, tanpa mempertimbangkan algoritma lain seperti Random Forest atau K-Nearest Neighbors (KNN).

1.5 Manfaat Penelitian

Penelitian ini kemungkinan akan memberi manfaat besar dalam analisis sentimen. Manfaat penelitian ini terdiri dari manfaat teoritis dan manfaat praktis yang signifikan bagi kemajuan ilmu pengetahuan dan aplikasi yang lebih efektif dalam memprediksi analisis sentimen. Manfaatnya dijelaskan sebagai berikut :

1.5.1 Manfaat Teoritis

1. Penelitian ini membawa wawasan baru dalam domain Text Mining, Machine Learning, dan Natural Language Processing (NLP), khususnya dalam penerapan metode Support Vector Machine (SVM) dan Naive Bayes untuk menganalisis sentimen.
2. Hasil dari penelitian ini bisa menjadi rujukan bagi para peneliti lain yang tertarik untuk mengembangkan metode analisis sentimen di berbagai bidang, khususnya dalam menilai kebijakan publik berdasarkan opini masyarakat di media sosial.
3. Penelitian ini berkontribusi dalam membandingkan efektivitas algoritma SVM dan Naive Bayes dalam analisis sentimen, sehingga dapat membantu peneliti lain dalam menentukan algoritma yang paling cocok untuk kebutuhan analisis teks yang mereka lakukan.
4. Penelitian ini semakin menguatkan teori bahwa media sosial dapat digunakan

sebagai indikator opini publik, terutama dalam mengevaluasi kebijakan yang diterapkan pemerintah di sektor pendidikan.

1.5.2 Manfaat Praktis

1. Analisis sentimen yang dilakukan dalam penelitian ini dapat membantu Kementerian Pendidikan memahami bagaimana masyarakat menanggapi kebijakan yang telah diterapkan oleh Pemimpin, sehingga bisa menjadi bahan evaluasi untuk perbaikan di masa depan.
2. Dengan adanya penelitian ini, masyarakat juga bisa melihat bagaimana opini mereka di media sosial dapat berperan dalam mengevaluasi kinerja pemerintah, sehingga dapat mendorong keterbukaan dan meningkatkan partisipasi publik dalam kebijakan pendidikan.
3. Studi ini dapat menjadi pijakan awal dalam pengembangan sistem otomatis yang mampu memantau opini publik secara real-time di media sosial, sehingga pemerintah dan pemangku kebijakan bisa merespons lebih cepat terhadap berbagai isu yang berkembang.
4. Para akademisi dan peneliti juga dapat memanfaatkan hasil penelitian ini sebagai referensi untuk riset lebih lanjut di bidang analisis sentimen, termasuk mengeksplorasi algoritma lain atau menggunakan data yang lebih luas untuk mendapatkan hasil yang lebih akurat.

