

BAB I PENDAHULUAN

1.1 Latar Belakang

Kinerja Presiden Joko Widodo selama sepuluh tahun terakhir menjadi perhatian utama masyarakat Indonesia. Media sosial menjadi platform utama untuk menyampaikan opini, kritik, dan saran, menggantikan media konvensional yang kurang aksesibel. Opini publik yang meluas ini menciptakan kebutuhan untuk menganalisis sentimen masyarakat secara terstruktur [1].

Analisis sentimen berbasis media sosial efektif dalam memahami opini publik, dengan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) sering digunakan karena akurasi yang tinggi. Sebagai contoh, penelitian pendahulu menunjukkan bahwa SVM memiliki performa lebih baik dibandingkan *Naive Bayes* dalam klasifikasi sentimen. Penelitian Julio menunjukkan bahwa SVM akurasi 90,78%, lebih tinggi dibanding *Naive Bayes* dengan akurasi 81,37% [2].

Optimasi algoritma seperti *Grid Search* terbukti dapat meningkatkan akurasi *Naive Bayes* dan SVM secara signifikan, seperti pada analisis sentimen. Namun, penerapannya pada analisis opini terkait kinerja pemerintahan masih jarang dilakukan, membuka peluang penelitian di bidang ini [3].

Penelitian ini bertujuan untuk membandingkan kinerja *Naive Bayes* dan SVM yang dioptimasi dengan *Grid Search* dalam analisis sentimen terhadap opini masyarakat tentang kinerja Presiden Jokowi. Dengan mengidentifikasi algoritma terbaik, penelitian ini diharapkan memberikan pandangan lebih akurat untuk evaluasi kebijakan yang lebih baik [4].

1.2 Rumusan Masalah

1. Bagaimana efektivitas algoritma *Naive Bayes* dan SVM dalam klasifikasi analisis sentimen publik terkait kinerja Presiden Jokowi?

2. Apa perbedaan klasifikasi sentimen (positif, negatif) antara kedua algoritma tersebut untuk menganalisis opini publik terhadap kebijakan dan tindakan kinerja pemerintah Jokowi?
3. Bagaimana metode *tuning hyperparameter* melalui *Grid Search* dapat mempengaruhi kinerja terhadap kedua algoritma dalam analisis sentimen?

1.3 Batasan Masalah

Dalam batasan masalah dalam penelitian ini diperlukan menjaga fokus serta ruang lingkup analisis agar tetap relevan dengan tujuan penelitian. Oleh karena itu, batasan masalah yang ditetapkan sebagai berikut:

1. **Sumber Data:** Penelitian ini akan mengambil data pada sosial media, khususnya YouTube. Data dikumpulkan dalam bentuk komentar dan opini publik yang mencerminkan sentimen masyarakat mengenai kinerja Presiden Jokowi. Alasan mengambil komentar dari youtube karena proses *scraping* data yang mudah dibanding media sosial lainnya.
2. **Metode Algoritma:** Untuk analisis sentimen dalam penelitian ini hanya terbatas hanya pada *Naive Bayes* dan SVM.
3. **Teknik Optimasi:** Untuk Penelitian ini fokus pada metode *Grid Search* sebagai teknik *tuning hyperparameter* meningkatkan akurasi setiap algoritma yang digunakan.
4. **Jenis Analisis Sentimen:** Analisis sentimen yang dilakukan penelitian ini hanya mencakup sentimen positif, negatif.
5. **Parameter Evaluasi:** Kinerja model dievaluasi berdasarkan metrik tertentu, seperti *akurasi*, *precision*, *recall*, dan *f1-score*.

1.4 Tujuan Penelitian

1. Dapat melakukan analisis sentimen dengan algoritma *Naive Bayes* dan SVM terkait pada kinerja Presiden Jokowi.
2. Mengetahui hasil setelah dua algoritma dioptimasi oleh *Grid Search* dalam analisis sentimen terkait kinerja Presiden Jokowi.
3. Dapat mengetahui mendalam terkait analisis sentimen publik (positif, negatif) terkait kinerja Presiden Jokowi.

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi untuk pengembangan teori analisis sentimen terkait membandingkan kinerja dua algoritma machine learning, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM), dalam konteks analisis sentimen terhadap opini publik mengenai kinerja Presiden Jokowi. Penelitian ini dapat memperkaya literatur dalam bidang *text mining* dan *Natural Language Processing* (NLP), serta mengidentifikasi kelebihan dan kekurangan masing-masing algoritma dalam menangani data teks yang bersifat subyektif dan tidak terstruktur. Selain itu, penerapan *Grid Search* untuk optimasi parameter dalam SVM memberi wawasan tentang bagaimana pencarian parameter optimal dapat mempengaruhi performa model, sehingga dapat dijadikan acuan untuk penelitian selanjutnya dalam pengembangan teknik klasifikasi teks.

2. Manfaat Praktis

Secara praktis, hasil penelitian sangat bermanfaat bagi berbagai pihak, seperti pemerintah, partai politik, atau lembaga riset yang ingin memahami sentimen publik terhadap kinerja Jokowi. Dengan mengetahui algoritma yang lebih efektif dalam menganalisis sentimen, pihak-pihak dapat merumuskan kebijakan atau strategi komunikasi yang lebih sesuai dengan opini masyarakat. Selain itu, penelitian ini juga dapat memberikan panduan bagi pengembang sistem analisis sentimen otomatis yang dapat digunakan untuk memonitor opini publik di media sosial, survei, atau forum diskusi. Bagi perusahaan atau organisasi, penelitian ini bisa menjadi dasar dalam memanfaatkan analisis sentimen untuk strategi pemasaran atau pengambilan keputusan berbasis data.

Melalui manfaat tersebut penelitian ini diharapkan dapat memberikan nilai tambah secara teoritis serta praktis untuk pengembangan model analisis sentimen yang lebih optimal dan relevan bagi kebutuhan berbagai pihak.

1.6 Sistematika Penulisan

Sistematika penulisan diuraikan dalam beberapa bab sebagai berikut :

BAB 1 PENDAHULUAN:

Pada bab ini berisi latar belakang masalah yang mendasari dilakukannya penelitian, rumusan masalah yang ingin dijawab, batasan masalah untuk menjaga fokus penelitian, tujuan penelitian yang ingin dicapai, manfaat penelitian baik secara teoritis maupun praktis, serta sistematika penulisan skripsi secara keseluruhan.

BAB 2 TINJAUAN PUSTAKA

Pada bab ini menjelaskan studi literatur yang relevan dan teori-teori dasar yang digunakan dalam penelitian, seperti konsep dasar analisis sentimen, algoritma Naïve Bayes dan Support Vector Machine (SVM), metode optimasi Grid Search, teknik penyeimbangan data menggunakan SMOTE, dan evaluasi performa model dengan confusion matrix. Bab ini juga mencakup hasil penelitian sebelumnya yang mendukung penelitian ini serta tinjauan umum terkait pengumpulan dan pengolahan data dari media sosial.

BAB 3 METODE PENELITIAN

Pada bab ini menjelaskan alur penelitian yang dilakukan, mulai dari pengumpulan data dengan teknik scraping menggunakan YouTube Data API, tahap praproses data seperti case folding, cleaning, slang word normalization, stop word removal, tokenizing, dan stemming. Selanjutnya, dilakukan labeling data dengan Sentiment Intensity Analyzer yang dimodifikasi, ekstraksi fitur menggunakan TF-IDF, penyeimbangan data dengan SMOTE, dan pembagian data menjadi data latih dan data uji. Bab ini juga mencakup proses pelatihan model menggunakan algoritma Naive Bayes dan

SVM, optimasi parameter dengan Grid Search, serta evaluasi performa menggunakan confusion matrix.

BAB 4 HASIL DAN PEMBAHASAN

Pada bab ini memuat hasil penelitian yang diperoleh, mulai dari hasil scraping data hingga evaluasi performa model. Hasil penelitian mencakup distribusi sentimen pada dataset, performa masing-masing algoritma sebelum dan setelah optimasi Grid Search, serta perbandingan antara algoritma Naive Bayes dan SVM berdasarkan metrik akurasi, precision, recall, dan F1-score. Bab ini juga membahas hasil confusion matrix dari setiap model untuk memahami keunggulan dan kelemahan masing-masing algoritma dalam klasifikasi sentimen terkait kinerja Presiden Jokowi.

BAB 5 PENUTUP

Pada Bab ini berisi kesimpulan yang merupakan intisari dari hasil penelitian yang telah dilakukan, menjawab rumusan masalah, dan menyimpulkan algoritma mana yang memiliki performa terbaik dalam analisis sentimen. Selain itu, bab ini juga memberikan saran untuk pengembangan lebih lanjut, seperti penggunaan dataset yang lebih besar, eksplorasi algoritma lain, dan penerapan teknik representasi teks yang lebih kompleks untuk meningkatkan hasil analisis.