BAB I
PENDAHULUAN

1.1 Latar Belakang

Media sosial telah berkembang pesat dalam beberapa tahun terakhir sebagai platform yang tidak hanya memungkinkan komunikasi antar individu, tetapi juga sebagai ruang untuk berbagi pendapat, informasi, dan kritik terhadap berbagai isu, termasuk kebijakan pemerintah. Salah satu platform media sosial yang banyak digunakan adalah X, yang memungkinkan penggunanya untuk menyampaikan pandangan atau pendapat dalam bentuk "tweets" yang bersifat publik dan dapat dibaca oleh siapa saja [1].

Dengan jumlah pengguna X yang sangat besar, analisis terhadap sejumlah besar tweet yang berkaitan dengan CPNS dapat memberikan wawasan yang berharga mengenai persepsi dan reaksi publik terhadap kebijakan tersebut. Analisis sentimen digunakan untuk memahami apakah pandangan masyarakat terhadap CPNS cenderung positif, negatif, atau netral. Dalam konteks ini, pendekatan analisis sentimen berfungsi untuk mengategorikan teks dalam tweet berdasarkan emosi yang terkandung di dalamnya [2].

Salah satu isu yang sering menjadi perhatian publik di Indonesia adalah Calon Pegawai Negeri Sipil (CPNS), yang menarik perhatian ribuan bahkan jutaan orang yang berharap untuk bergabung dengan instansi pemerintah. Setiap tahunnya, CPNS menjadi topik perbincangan hangat di media sosial, dengan berbagai reaksi yang muncul dari peserta ujian, calon pelamar, dan masyarakat umum. Reaksi ini dapat berupa pendapat positif, negatif, atau netral terkait dengan proses, persyaratan, hasil ujian, dan berbagai aspek lain yang memengaruhi kepuasan atau kekecewaan para pelamar [3].

Perbandingan kinerja antara Random Forest dan SVM dalam menganalisis sentimen pada X terkait CPNS. Tujuan dari perbandingan ini adalah untuk mengevaluasi akurasi, presisi, dan recall kedua metode dalam mengklasifikasikan sentimen yang terdapat dalam tweets mengenai CPNS. Dengan demikian,

penelitian ini diharapkan dapat memberikan kontribusi dalam pemilihan metode yang lebih tepat dan efektif untuk analisis sentimen di platform X, khususnya dalam konteks isu-isu sosial yang besar seperti CPNS[4].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan diatas maka rumusan masalah dalam penelitian ini adalah "Bagaimana perbandingan kinerja antara metode Random Forest dan Support Vector Machine (SVM) dalam analisis sentimen terhadap tweet mengenai Calon Pegawai Negeri Sipil (CPNS) di X berdasakan matrik evaluasi seperti akurasi, recall, precision, f1-score?".

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan yang perlu diperhatikan untuk memastikan fokus dan keberhasilan penelitian. Batasan-batasan tersebut adalah sebagai berikut:.

1. Penelitian ini hanya akan menggunakan data tweet yang diambil dari platform X dengan memanfaatkan auth_token pada X untuk melakukan crawling. Hanya tweet yang relevan dengan topik Calon Pegawai Negeri Sipil (CPNS).
2. Data tweet yang diambil dibatasi dari 01 Agustus sampai 30 November.
3. Fokus penelitian ini hanya melibatkan analisis sentimen terhadap tweet yang dikategorikan ke dalam tiga kategori: positif, negatif, dan netral.
4. Penelitian ini akan membandingkan kinerja dua metode analisis sentimen, yaitu Random Forest dan Support Vector Machine (SVM), dan tidak akan mencakup metode lain dalam pembandingannya.
5. Kata kunci yang digunakan untuk crawling data hanya #CPNS
6. Penelitian ini tidak akan membahas lebih lanjut mengenai pengolahan data, analisis sentimen hanya menggunakan confusion metrik standar seperti akurasi, presisi, recall, dan f1-score.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk Menganalisis dan membandingkan kinerja metode Random Forest dan Support Vector Machine (SVM) dalam melakukan analisis sentimen terhadap tweet mengenai CPNS di X. Perbandingan ini dilakukan dengan mengukur metrik evaluasi seperti akurasi, recall, precision, dan f1-score, sehingga dapat diketahui metode mana yang lebih efektif dalam mengklasifikasikan sentimen publik terhadap topik CPNS.

1.5 Manfaat Penelitian

Tuliskan Manfaat yang dihasilkan dari penelitian ini diuraikan sebagai berikut:

1. Bagi Peneliti

   Penelitian ini memberikan wawasan tentang efektivitas dua metode Random Forest dan Support Vector Machine (SVM) dalam menganalisis sentimen data teks dari tweeter. Temuan dari penelitian ini dapat menjadi referensi untuk penelitian lebih lanjut di bidang analisis sentimen, dan pengolahan bahasa alami (NLP), khususnya dalam media sosial X.

2. Bagi pemerintah dan penyelenggara CPNS

   Hasil penelitian ini dapat digunakan untuk memahami sentimen masyarakat terkait dengan Calon Pegawai Negeri Sipil (CPNS), yang dapat memberikan gambaran mengenai persepsi publik, baik yang positif, negatif, atau netral. Pemerintah atau instansi terkait dapat menggunakan informasi ini untuk memperbaiki atau menyesuaikan kebijakan dan prosedur CPNS agar lebih transparan dan diterima oleh masyarakat..

1.6 Sistematika Penulisan

Adapun sistematika penulisan skripsi ini sebagai berikut:

BAB I PENDAHULUAN memuat tentang Latar belakang masalah, rumusan masalah, batasan masalah, Tujuan penelitian, dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA, berisi studi literatur, dasar-dasar teori yang digunakan dalam penelitian sebelumnya.

BAB III METODE PENELITIAN, memuat tentang objek penelitian, alur penelitian, aplikasi apa saja yang digunakan, dan bahan penelitian

BAB IV HASIL DAN PEMBAHASAN, berisi tentang hasil dan pembahasan penelitian

BAB V PENUTUP, berisi tentang kesimpulan dan saran dalam penelitian…