

BAB I PENDAHULUAN

1.1. Latar Belakang

Wacana mengenai kebijakan kenaikan fasilitas dan tunjangan bagi pejabat publik, khususnya Dewan Perwakilan Rakyat (DPR), kembali mencuat dan memicu polemik di tengah masyarakat. Di saat kondisi ekonomi nasional yang sedang menghadapi tantangan inflasi dan daya beli yang fluktuatif, rencana kenaikan tunjangan ini dinilai kontradiktif dengan semangat efisiensi anggaran negara. Kesenjangan antara fasilitas yang diterima wakil rakyat dengan kinerja legislasi yang dihasilkan sering kali menimbulkan persepsi ketidakadilan sosial[1]. Terkait produk kebijakan DPR *Omnibus Law* menunjukkan bahwa transparansi dan keberpihakan pada rakyat adalah faktor utama yang memengaruhi sentimen publik yang cenderung negatif. Hal ini diperkuat oleh temuan *Chrisinta & Simarmata*[2] yang menyatakan bahwa resistensi publik meningkat ketika isu sensitif pejabat terekspos secara luas di media massa.

Perkembangan media digital telah mengubah pola interaksi masyarakat dalam merespons isu politik. Saat ini, masyarakat tidak hanya mengonsumsi berita melalui televisi, tetapi beralih ke platform berbagi video seperti YouTube. Berbeda dengan Twitter yang terbatas karakter, kolom komentar YouTube memungkinkan pengguna menuliskan opini yang lebih panjang, deskriptif, dan argumentatif. Namun, besarnya volume data komentar ini menjadi tantangan tersendiri untuk dianalisis secara manual. Oleh karena itu, *Text Mining* dan *Natural Language Processing* (NLP) diperlukan untuk mengubah data komentar yang tidak terstruktur tersebut menjadi informasi berharga mengenai kecenderungan sentimen publik[3],[4]. Mengingat algoritma klasifikasi pada data mining ini telah terbukti sangat efektif untuk diterapkan di Indonesia[5].

Untuk mengatasi tantangan pengolahan data komentar YouTube yang memiliki dimensi fitur tinggi dan variasi bahasa tidak baku (*noise*), penelitian ini mengajukan penggunaan algoritma *Support Vector Machine* (SVM). SVM dipilih karena prinsip kerjanya yang berbasis pada pencarian *hyperplane* (garis pemisah) terbaik dengan margin maksimal untuk memisahkan kelas sentimen. Pendekatan ini memungkinkan model memiliki kemampuan generalisasi yang baik dan akurasi yang stabil, meskipun dihadapkan pada data teks yang kompleks. Keunggulan ini didukung oleh penelitian komparasi yang membuktikan bahwa SVM memiliki performa klasifikasi

yang lebih unggul dibandingkan *Naive Bayes*[6]. Terutama pada dataset dengan ketidakseimbangan kelas. Lebih lanjut, efektivitas SVM pada platform video juga dikonfirmasi oleh *Ramadhan & Mubarak*[7], di mana algoritma ini terbukti tangguh dalam menangani karakteristik teks YouTube yang penuh dengan bahasa *slang*.

1.2.Rumusan Masalah

1. Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah: Bagaimana penerapan tahap preprocessing dan pembobotan TF-IDF untuk menangani karakteristik data komentar YouTube yang tidak terstruktur agar siap diklasifikasikan?
2. Bagaimana perbandingan kinerja Kernel Linear dan Kernel RBF pada algoritma SVM dalam menentukan model klasifikasi sentimen yang paling optimal?
3. Berapa tingkat Akurasi, Presisi, Recall, dan F1-Score terbaik yang dihasilkan oleh algoritma SVM dalam mengklasifikasikan sentimen masyarakat terhadap isu kenaikan tunjangan DPR?

1.3.Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menerapkan teknik preprocessing dan metode pembobotan TF-IDF untuk mengubah data komentar YouTube yang penuh noise menjadi data terstruktur yang siap diproses.
2. Menganalisis dan membandingkan performa Kernel Linear dan Kernel RBF guna mengetahui konfigurasi kernel yang paling efektif untuk karakteristik data tersebut.
3. Mengukur kinerja model terbaik yang dihasilkan berdasarkan parameter evaluasi Akurasi, Presisi, Recall dan F1-Score.

1.4.Batasan Masalah

Agar penelitian lebih terarah, diberikan batasan masalah sebagai berikut:

1. Data yang digunakan adalah komentar berbahasa Indonesia pada video YouTube bertopik *Kenaikan Tunjangan/Gaji DPR* dari kanal berita nasional. Dengan rentang waktu pengambilan data pada periode Agustus 2025 hingga Desember 2025.
2. Algoritma klasifikasi yang digunakan adalah *Support Vector Machine (SVM)* dengan pembobotan TF-IDF.
3. Kelas sentimen dibagi menjadi dua kategori yaitu *Positif* dan *Negatif*, yang dilabeli secara otomatis menggunakan metode *Lexicon-Based*.

4. Penelitian ini tidak membahas aspek psikologis pengguna, melainkan fokus pada klasifikasi teks opini.

1.5.Keterbaharuan

Berikut penjelasan mengenai kebaruan penelitian ini berdasarkan kajian pustaka penelitian terdahulu:

1. Putri, Nama, & Sulistiono [8] dalam penelitiannya menggunakan dataset tweet dengan performa klasifikasi akurasi sebesar 80%. Berbeda dengan penelitian ini, kami tidak menggunakan Twitter melainkan memanfaatkan kolom komentar YouTube yang lebih deskriptif, serta menggunakan algoritma SVM yang terbukti lebih handal untuk teks kompleks.
2. Chrisinta & Simarmata [2] menghasilkan model dengan tingkat akurasi 64,55% menggunakan Naïve Bayes pada opini umum pejabat publik. Selain itu, pendekatan leksikon *lexicon-based* dan *Support Vector Machine SVM* juga telah terbukti andal dalam menganalisis sentimen terkait politik atau pejabat tinggi negara[9],[10]. Berbeda dengan penelitian-penelitian tersebut, penelitian ini memfokuskan objek secara spesifik pada isu *kenaikan tunjangan*.
3. Pane, Owen, & Prianto [1] dalam penelitiannya mengenai sentimen kebijakan DPR *Omnibus Law* menemukan bahwa masyarakat memberikan sentimen positif sebesar 47,06% dan sentimen negatif sebesar 40,52%. Berbeda dengan penelitian ini, kami memfokuskan objek secara spesifik pada isu kenaikan Tunjangan/Anggaran, yang diyakini memiliki sentimen penolakan yang lebih tajam.
4. Raharjo et al. [6] melakukan komparasi algoritma dan membuktikan bahwa *SVM* unggul dengan akurasi 87%, lebih tinggi dibandingkan *Naive Bayes* 81%. Berbeda dengan penelitian ini, penelitian kami tidak membandingkan algoritma, tetapi memperdalam optimasi SVM dengan membandingkan *Kernel Linear vs RBF*.
5. Ramadhan & Mubarak [7] mencapai akurasi 89% pada analisis sentimen BBM di YouTube. Berbeda dengan penelitian ini, penelitian tersebut membahas isu Kenaikan BBM. Penelitian ini mengisi celah dengan mengangkat isu Kenaikan Tunjangan Pejabat, serta melakukan optimalisasi internal SVM.