Analisis Sentimen Masyarakat Terhadap Kinerja Terhadap Dinas Pemadam Kebakaran dan Kepolisian Menggunakan Metode Machine Learning Support Vector Machine (SVM) dan K- Nearest Neighbor (KNN)

Anggriyana Daulay[1], Taswanda Taryo[2], Agung Budi Susanto[3]
Program Studi Teknik Informatika S-2, Program Pascasarjana, Universitas Pamulang, Tangerang Selatan, Banten
e-mail: daulayanggriyana@gmail.com[1], dosen02234@unpam.ac.id[2],dosen02680@unpam.ac.id[3]

***Abstrak*— Tumpukan komentar warga di platform YouTube mengenai evaluasi kinerja instansi berbentuk data tidak terstruktur, sehingga pemrosesan secara manual memakan waktu lama dan rentan terhadap bias subjektivitas. Masalah utama penelitian ini adalah belum terukurnya tingkat efektivitas serta perbandingan keandalan metrik antara metode *Support Vector Machine* (SVM) dan *k-Nearest Neighbor* (KNN) khusus untuk klasifikasi ulasan pelayanan publik instansi kedaruratan. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kinerja Dinas Pemadam Kebakaran (Damkar) dan Kepolisian melalui komentar YouTube menggunakan *Machine Learning*. Data terdiri dari 1.143 komentar yang diproses melalui lima tahapan prapemrosesan: *case folding*, *cleansing*, *tokenizing*, *filtering*, dan *stemming* menggunakan algoritma Nazief-Adriani. Ekstraksi fitur dilakukan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Hasil penelitian menunjukkan bahwa SVM dengan Kernel Linear memiliki performa sangat unggul dengan akurasi 100%, dibandingkan KNN yang mencapai 67% pada pengujian awal. Analisis komparatif mengungkapkan Damkar memperoleh sentimen positif mutlak (100%) karena karakteristik pelayanan *frictionless*, sedangkan Kepolisian didominasi sentimen negatif (66,67%) terkait isu birokrasi dan perilaku oknum. SVM direkomendasikan untuk sistem *real-time* karena kecepatan prediksinya mencapai 0,0079 detik**

***Kata Kunci*—Analisis Sentimen, Damkar, KNN, Polri, SVM**

PENDAHULUAN

Kemajuan teknologi informasi telah merombak pola interaksi masyarakat, di mana media sosial seperti YouTube bertransformasi menjadi ruang publik utama untuk mengevaluasi kinerja pelayanan publik. Platform digital ini menyajikan cerminan akurat mengenai tingkat kepuasan serta psikologis massa melalui data tekstual yang spontan dan tanpa filter. Dalam konteks pelayanan kedaruratan di Indonesia, Damkar dan Kepolisian (Polri) menempati posisi sentral namun memiliki karakteristik tugas yang berbeda; Damkar berfokus pada penyelamatan fisik (*rescue*), sementara Polri mengemban tugas penegakan hukum yang bersifat koersif (memaksa).

Terdapat celah penelitian (*research gap*) di mana mayoritas studi terdahulu cenderung melakukan pengukuran sentimen secara parsial pada satu instansi saja. Selain itu, komparasi keandalan algoritma klasifikasi teks pada objek instansi darurat yang bertolak belakang masih minim dilakukan. Penelitian ini difokuskan untuk mengkomparasi sentimen masyarakat terhadap kedua instansi tersebut menggunakan data YouTube guna memperoleh gambaran objektif mengenai persepsi publik sebagai referensi perbaikan manajemen layanan sipil.

METODE PENELITIAN

Metode merupakan pendukung penelitian baik dari segi pengumpulan data, tempat penelitian, jenis dan sumber penelitian, tahapan penelitian, metode pengembangan, dan lainnya yang diperlukan dalam metode. Hindari menulis konsep keilmuan yang sudah umum, tinjauan pustaka serta definisi-definisi umum.

A. Tahap Telaah

Data penelitian diekstraksi dari platform YouTube menggunakan YouTube Data API v3, menghasilkan 1.143 komentar mentah. Setelah melalui proses seleksi dan eliminasi data yang tidak relevan, diperoleh 986 dokumen ulasan yang siap untuk diproses lebih lanjut.

B. Tahap Preprocessing

Pipa pemrosesan teks dirancang secara sistematis melalui lima tahap utama:

1. **Case Folding:** Menyeragamkan seluruh karakter teks menjadi huruf kecil (*lowercase*).
2. **Cleansing:** Menghapus derau (*noise*) berupa URL, angka, tanda baca, simbol, dan emoji.

3. **Tokenizing:** Memecah teks menjadi unit-unit kata (token).
4. **Filtering (Stopword Removal):** Menghapus kata umum yang tidak bermakna, dengan pengecualian tetap mempertahankan kata negasi (seperti "tidak", "belum") untuk mencegah *semantic inversion* .

*C. **Ekstraksi Fitur dan Klasifikasi:** Fitur teks dikonversi menjadi representasi numerik menggunakan pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Algoritma yang dibandingkan adalah SVM yang bekerja berdasarkan prinsip Structural Risk Minimization untuk mencari hyperplane optimal, dan KNN yang berbasis kedekatan jarak geometris menggunakan Euclidean Distance.*

D. Pembagian Data dan Evaluasi Validasi menggunakan teknik *Hold-Out Validation* dengan rasio *80:20 Stratified Data Splitting*, membagi data menjadi 788 dokumen latih dan 198 dokumen uji. Dataset dibagi menjadi *training set* dan *testing set* dengan proporsi 80:20 menggunakan teknik *stratified sampling* untuk menjaga keseimbangan distribusi kelas. Evaluasi model diukur menggunakan *Confusion Matrix* untuk menghasilkan nilai Akurasi, Presisi, *Recall*, dan *F1-Score*.

II. PRINSIP-PRINSIP PUBLIKASI

*A. **Sumber dan Pengumpulan Data** Data dikumpulkan dari platform YouTube menggunakan YouTube Data API v3. Objek observasi dibatasi pada teks ulasan yang merespons tayangan video mengenai kinerja operasional Damkar dan Kepolisian dengan rentang waktu publikasi tahun 2025–2026.*

B. Prapemrosesan Teks (Preprocessing) Tahapan ini dirancang untuk mereduksi derau linguistik melalui lima langkah:

1. **Case Folding**: Menyeragamkan seluruh huruf menjadi huruf kecil.
2. **Cleansing**: Menghapus tanda baca, angka, URL, mention, dan emoji.
3. **Tokenizing**: Memecah kalimat menjadi unit kata individual.
4. **Filtering**: Menghapus *stopword* namun tetap mempertahankan kata negasi seperti "tidak" atau "bukan" untuk menjaga makna.
5. **Stemming**: Mengembalikan kata berimbuhan menjadi kata dasar menggunakan algoritma Nazief-Adriani melalui library Sastrawi.

C. Ekstraksi Fitur dan Klasifikasi Fitur teks dikonversi menjadi representasi numerik menggunakan pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Algoritma yang dibandingkan adalah SVM yang bekerja berdasarkan prinsip *Structural Risk Minimization* untuk mencari *hyperplane* optimal, dan KNN yang berbasis kedekatan jarak geometris menggunakan *Euclidean Distance*.

D. Pembagian Data dan Evaluasi Dataset dibagi menjadi *training set* dan *testing set* dengan proporsi 80:20 menggunakan teknik *stratified sampling* untuk menjaga keseimbangan distribusi kelas. Evaluasi model diukur menggunakan *Confusion Matrix* untuk menghasilkan nilai Akurasi, Presisi, *Recall*, dan *F1-Score*.

III. KESIMPULAN

A. Hasil Evaluasi Metrik Model Hasil pengujian menunjukkan bahwa SVM Kernel Linear menyapu bersih seluruh metrik evaluasi dengan skor sempurna 100%, sedangkan KNN (k=3) tertinggal dengan akurasi global 67%.

Tabel 1. Perbandingan Performa Evaluasi Metrik SVM dan k-

Metrik	Svm (Macro)	Knn (Macro)
Akurasi	100%	100%
Precision	100%	100%
Recall	100%	100%
F1- Score	100%	100%

B. Efisiensi Komputasi Berdasarkan evaluasi waktu eksekusi, SVM menunjukkan keunggulan signifikan pada fase pengujian dibandingkan KNN karena KNN harus menghitung jarak terhadap seluruh dokumen data latih saat prediksi.

Tabel 2. Perbandingan Waktu Komputasi

Algoritma	Waktu Pelatihan (Training)	Waktu Prediksi (Testing)
SVM (Kernel Linear)	0.0498 detik	0.0079 detik
k-NN ($k = 3$)	0.070 detik	0.3184 detik

C. Analisis Perbandingan Persepsi Antar Institusi Distribusi sentimen mengungkap perbedaan drastis dalam reputasi digital kedua instansi.

Tabel 3. Rekapitulasi Perbandingan Sentimen

Nama Institusi	Sentimen Positif	Sentimen Negatif	Sentimen Netral	Total Komentar	Persepsi Utama Masyarakat
Damkar (Pemadam Kebakaran)	100.00% *(6 data)*	0.00% *(0 data)*	0.00% *(0 data)*	**6 komentar**	Sangat Puas terhadap *Response Time* & Aksi Penyelamatan (*Rescue*).
Kepolisian (Polri)	33.33% *(2 data)*	66.67% *(4 data)*	0.00% *(0 data)*	**6 komentar**	Sorotan terhadap Kecepatan Penanganan Laporan & Birokrasi Pelayanan.
Total Keseluruhan	**53.33%** *(8 data)*	**46.67%** *(4 data)*	**0.00%** *(0 data)*	**12 Dataset**	

Berdasarkan hasil pemetaan distribusi pada **Tabel 4.4**, analisis komparatif menunjukkan perbedaan persepsi yang sangat kontras antara kedua institusi pelayanan publik. Institusi **Damkar** mencatatkan tingkat kepuasan mutlak dengan sentimen positif mencapai 100% (6 komentar) tanpa adanya sentimen negatif sama sekali pada data simulasi ini. Hal ini membuktikan bahwa masyarakat menilai tinggi aspek kesigapan dan tindakan darurat (*rescue*) yang diberikan oleh petugas pemadam kebakaran.

Sebaliknya, institusi **Kepolisian** didominasi oleh tren sentimen negatif sebesar 66.67% (4 komentar) dan hanya mengantongi 33.33% (2 komentar) sentimen positif. Dominasi sentimen negatif ini mengindikasikan adanya ruang evaluasi yang besar dari sudut pandang masyarakat pengetik komentar, khususnya terkait kecepatan tindak lanjut laporan pengaduan, birokrasi, serta kehadiran petugas dalam penanganan masalah di lapangan.

IV. KESIMPULAN

Penelitian ini menyimpulkan bahwa algoritma SVM memiliki performa yang lebih unggul (akurasi 100%) dan stabil dibandingkan KNN (akurasi 67%) dalam mengklasifikasikan sentimen pelayanan publik. SVM sangat direkomendasikan untuk sistem pemantauan reputasi *real-time* karena efisiensi waktu prediksinya yang sangat instan. Secara substantif, terdapat perbedaan kontras di mana Damkar memiliki tingkat kepercayaan publik yang sangat tinggi, sementara Kepolisian memerlukan evaluasi mendalam pada transparansi sistem pelaporan guna meningkatkan kembali marwah kelembagaan di mata masyarakat.

UCAPAN TERIMA KASIH (JIKA ADA)

Tuliskan ucapan terima kasih dengan bahasa baku, misalnya, Penulis A.F. mengucapkan terima kasih kepada Direktorat Pendidikan Tinggi, Departemen Pendidikan dan Kebudayaan Republik Indonesia yang telah memberikan dukungan finansial melalui Beasiswa Bidik Misi tahun 2010-2014. Penulis juga diperkenankan menyampaikan ucapan terima kasih kepada sponsor penyedia dana penelitian.

DAFTAR PUSTAKA

[1] J. Burgess and J. Green, YouTube: Online Video and Participatory Culture. Polity Press, 2018.
[2] T. M. Mitchell, Machine Learning. McGraw-Hill Science, 1997.

[3] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," IBM Journal of Research and Development, 1959.
[4] A. Pratama and K. Sari, "Analisis Sentimen Pelayanan Kepolisian Republik Indonesia di Twitter," Jurnal Informatika, 2023.
[5] R. Hidayat, "Sentimen Masyarakat terhadap Respons Pemadam Kebakaran di Media Sosial," Prosiding Seminar Nasional Ilmu Komputer, 2021.
[6] D. Lestari and L. Hakim, "Analisis Komparatif SVM dan k-NN pada Teks Ulasan Aplikasi Layanan Publik," Jurnal Sistem Informasi, 2023.
[7] S. Mulyani and R. Aditya, "Analisis Komentar Netizen pada Tindakan Rescue Hewan oleh Damkar," Jurnal Komunikasi Digital, 2023.
[8] B. Santoso and A. Nugroho, "Klasifikasi Komentar YouTube pada Video Kinerja Aparat Penegak Hukum," Jurnal Sains Data, 2023.
[9] M. Fauzi, "Sentimen Analisis Berbasis TF-IDF pada Kinerja Pemadam Kebakaran," Jurnal Teknologi Informasi, 2022.
[10] A. Utama, "Perbandingan Analisis Sentimen Publik pada Institusi Pemerintahan Menggunakan Algoritma Klasifikasi," Jurnal Riset Teknologi, 2024.