

ANALISIS SENTIMEN MASYARAKAT TERHADAP KINERJA BUPATI KABUPATEN LANGKAT PERIODE 2019-2024 MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE(SVM)

Chairul Sani Harahap¹, Hotler Manurung², Muamar Khadapi³
Sistem Informasi, STMIK Kaputama, Binjai
E-mail: harahapsani088@gmail.com¹

ABSTRAK

Kinerja kepala daerah sering menjadi sorotan masyarakat dan memunculkan beragam opini, khususnya di media sosial seperti *Twitter*. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kinerja Bupati Kabupaten Langkat periode 2019–2024 dengan menggunakan algoritma *Support Vector Machine* (SVM). Data diperoleh melalui teknik *crawling* dengan *library Python*, kemudian dilakukan tahap *preprocessing* yang mencakup *case folding*, *cleansing*, tokenisasi, *stopword removal*, *stemming*, serta representasi data menggunakan metode TF-IDF. Data yang telah diproses diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Hasil penelitian menunjukkan bahwa mayoritas sentimen masyarakat cenderung negatif, diikuti sentimen positif, dan sebagian kecil netral. Model klasifikasi dengan algoritma SVM memberikan akurasi sebesar 74%, yang menunjukkan bahwa metode ini cukup efektif dalam mengelompokkan opini masyarakat berbasis teks. Penelitian ini diharapkan dapat menjadi masukan bagi pemerintah daerah dalam mengevaluasi kinerja kepemimpinan, serta memberikan kontribusi terhadap pengembangan analisis sentimen di bidang pemerintahan.

Kata kunci

Analisis Sentimen, Support Vector Machine, TF-IDF, Kinerja Bupati, Kabupaten Langkat.

ABSTRACT

The performance of regional leaders often becomes a public concern and generates various opinions, particularly on social media platforms such as Twitter. This study aims to analyze public sentiment toward the performance of the Regent of Langkat Regency for the 2019–2024 period using the Support Vector Machine (SVM) algorithm. The data were collected through a crawling technique with Python libraries and then processed through several stages, including case folding, cleansing, tokenization, stopword removal, stemming, and data representation using the TF-IDF method. The processed data were classified into three sentiment categories: positive, negative, and neutral. The results show that the majority of public sentiment tends to be negative, followed by positive sentiment, with a smaller proportion being neutral. The classification model using the SVM algorithm achieved an accuracy of 74%, indicating that this method is fairly effective in categorizing text-based public opinions. This research is expected to provide insights for local government in evaluating leadership performance and contribute to the development of sentiment analysis in the field of public administration.

Keywords

Sentiment Analysis, Support Vector Machine, TF-IDF, Regent's Performance, Langkat Regency.

1. PENDAHULUAN

Pembangunan daerah merupakan faktor fundamental dalam upaya meningkatkan kesejahteraan masyarakat di seluruh wilayah Indonesia. Kabupaten Langkat, Sumatera Utara, selama masa kepemimpinan periode 2019–2024 telah mengimplementasikan berbagai program pembangunan yang memberikan dampak signifikan terhadap berbagai indikator sosial dan ekonomi, baik dalam skala perkotaan maupun pedesaan. Berbagai

inisiatif ini mencakup peningkatan infrastruktur, pemberdayaan ekonomi masyarakat, serta penguatan pelayanan publik yang diharapkan mampu mendorong pertumbuhan dan pemerataan hasil pembangunan. Meskipun demikian, sejumlah indikator sosial-ekonomi seperti tingkat kemiskinan, pengangguran, dan ketimpangan pendapatan masih menunjukkan angka yang cukup mengkhawatirkan. Ketidakseimbangan ini mencerminkan bahwa manfaat pembangunan belum sepenuhnya dirasakan oleh seluruh lapisan masyarakat.

Lemahnya implementasi program peningkatan taraf hidup dan pendidikan di beberapa wilayah bahkan berdampak pada munculnya berbagai masalah sosial, seperti meningkatnya tingkat pengangguran, kriminalitas, dan premanisme. Kondisi tersebut mengindikasikan perlunya evaluasi mendalam terhadap efektivitas kebijakan yang telah diterapkan serta perbaikan pada sektor-sektor strategis agar pembangunan dapat berjalan lebih merata dan inklusif. Seiring dengan berkembangnya teknologi informasi, media sosial seperti Twitter kini telah berkembang menjadi wadah utama bagi masyarakat untuk menyuarakan pendapat, memberikan masukan yang membangun, serta mengungkapkan pandangan mereka terhadap berbagai kebijakan yang dikeluarkan pemerintah. Platform ini memberikan kesempatan yang demokratis bagi setiap orang untuk turut serta dalam pembahasan publik mengenai berbagai persoalan pembangunan di daerah mereka. Dalam hal ini, analisis sentimen terhadap teks di media sosial dengan memanfaatkan pendekatan Natural Language Processing (NLP) dan algoritma Support Vector Machine (SVM) memungkinkan para peneliti untuk mengenali dan memahami kecenderungan pendapat masyarakat dengan cara yang lebih objektif dan menyeluruh (Syahputra, Rahayudi, dan Muflikhah, 2022).

Penerapan metode analisis sentimen ini memiliki kelebihan dalam mengelompokkan pendapat publik ke dalam tiga kategori pokok: positif, negatif, dan netral, dengan tingkat ketepatan yang cukup tinggi (Fauzan et al., 2023). Pada tahap pengambilan fitur, penelitian ini menggunakan pendekatan TF-IDF (Term Frequency-Inverse Document Frequency) yang dapat mengubah setiap tweet menjadi bentuk vektor angka berdasarkan seberapa penting setiap kata dalam keseluruhan kumpulan data (Runimeirati, Abdul Muis, dan Muhammad, 2023). Analisis sentimen adalah analisis pendapat seseorang yang menggambarkan bagaimana mereka menyampaikan pendapat mereka tentang situasi atau topik yang ditulis. Ini biasanya dibagi menjadi ekspresi senang, tidak senang, dan biasa saja, atau biasanya dibagi menjadi kategori positif, negatif, dan netral (Hendrawan dan Sela, 2024).

Metode ini terbukti ampuh untuk menentukan bobot kata dan menyaring kata-kata yang kurang memberikan informasi berarti dalam proses pengelompokan (Emarapenta, Sinulingga, Sitorus, 2024). Dengan mengkaji data Twitter dari masyarakat Kabupaten Langkat secara menyeluruh, penelitian ini bertujuan untuk mendapatkan gambaran yang nyata dan asli mengenai tanggapan masyarakat terhadap program-program pembangunan daerah yang telah dijalankan selama kurun waktu tersebut.

Penelitian terdahulu mengenai analisis sentimen dengan SVM telah banyak dilakukan. Hendrastuty et al. (2021) menunjukkan klasifikasi sentimen terhadap Program Kartu Prakerja dengan akurasi 98.34%. Sementara itu, penelitian Putra, Haeirudin, dan Khairunnisa (2021) tentang kebijakan PPKM mendapatkan akurasi 64%. Di sisi lain, Seran et al. (2024) mencatat akurasi 66% pada analisis sentimen kinerja Presiden Joko Widodo. Studi lain oleh Caroline et al. (2024) pada kasus korupsi PT. Timah memperoleh akurasi sebesar 88%. Syam et al. (2024) berhasil mencapai akurasi 97% pada analisis sentimen konflik Israel-Palestina dengan skema preprocessing yang tepat. Hasil-hasil ini

menunjukkan bahwa performa SVM sangat dipengaruhi oleh kualitas data dan tahapan pra-pemrosesan yang diterapkan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *text mining* dengan kerangka kerja CRISP-DM yang terdiri dari enam tahap utama. CRISP-DM dipilih karena bersifat sistematis, fleksibel, dan dapat diterapkan pada berbagai permasalahan analisis teks.

2.1 Pengumpulan Data

Data dikumpulkan dari media sosial Twitter menggunakan metode crawling dengan kata kunci "Kinerja Bupati Langkat" dan "Bupati Langkat" dari tahun 2019 hingga 2024. Data diperoleh menggunakan bahasa pemrograman Python dengan bantuan berbagai alat pendukung melalui platform Google Colab (Oleh dan Fawwazrin, 2024).

2.2 Tahapan Preprocessing

Data yang telah dikumpulkan kemudian diolah melalui serangkaian tahapan:

- Case Folding*: Mengubah seluruh teks menjadi huruf kecil.
- Cleaning*: Menghilangkan tautan, tag pengguna, tagar, angka, dan tanda baca.
- Tokenisasi: Memisahkan teks menjadi kata-kata terpisah.
- Stopword Removal*: Menghapus kata-kata umum yang tidak relevan.
- Stemming*: Mengubah kata berimbuhan menjadi kata dasar.

2.3 Ekstraksi Fitur

Setelah pra-pemrosesan, data teks diubah menjadi vektor numerik menggunakan metode TF-IDF agar dapat diolah oleh algoritma klasifikasi.

2.4 Pemodelan dan Evaluasi

Pada tahap ini, data yang telah diproses dibagi menjadi data latih (70%) dan data uji (30%). Model klasifikasi dibangun menggunakan algoritma SVM dengan kernel linear. Kinerja model dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score, yang ditampilkan dalam bentuk confusion matrix.

Rumus Akurasi:

$$accuracy = \frac{TP + TN}{P + N}$$

3. HASIL DAN PEMBAHASAN

3.1 Akurasi Model

Pengujian model SVM menunjukkan akurasi sebesar 74.47%. Hasil ini menunjukkan bahwa model cukup efektif dalam mengklasifikasikan sentimen, terutama pada kelas negatif (F1-score 0.84) dan positif (F1-score 0.68), meskipun performanya kurang optimal pada kelas netral (F1-score 0.33).

Akurasi: 0.74

Akurasi: 74.47%

Laporan Klasifikasi:

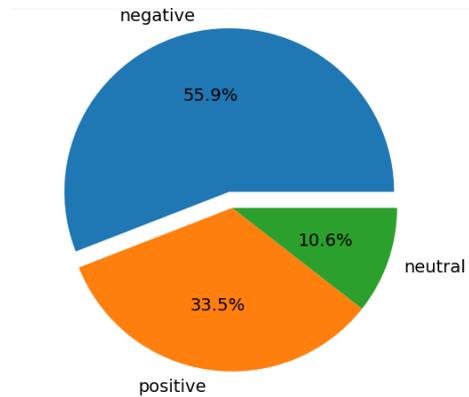
	precision	recall	f1-score	support
negative	0.82	0.87	0.84	109
neutral	0.35	0.32	0.33	19
positive	0.71	0.65	0.68	60
accuracy			0.74	188
macro avg	0.63	0.61	0.62	188
weighted avg	0.74	0.74	0.74	188

Gambar 1. Akurasi Model

3.2 Distribusi Sentimen

Analisis terhadap 2.209 tweet menunjukkan distribusi sentimen sebagai berikut:

- Negatif: 55.9%
- Positif: 33.5%
- Netral: 10.6%

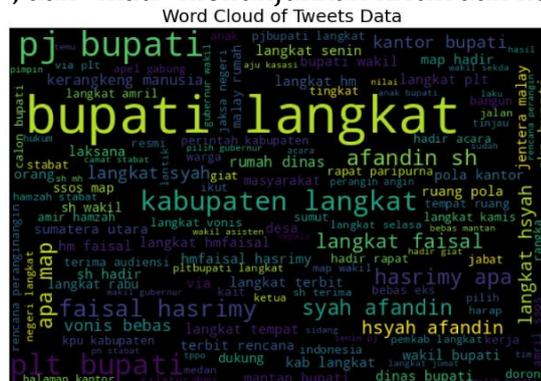


Gambar 2. Distribusi Sentimen(Pie Chart)

Hasil ini mengindikasikan bahwa opini masyarakat di Twitter terhadap kinerja Bupati Kabupaten Langkat cenderung lebih banyak yang bersifat kritis atau negatif.

3.3 Analisis Kata Kunci (Word Cloud)

Word cloud dari data tweet menunjukkan bahwa kata yang paling banyak muncul antara lain “bupati”, “langkat”, “kabupaten”, “faisal”, dan “syah afandin”. Kata-kata dominan pada sentimen positif mencakup “buka”, “kunjung”, “doa”, dan “layanan”, yang merefleksikan apresiasi masyarakat. Sebaliknya, kata-kata pada sentimen negatif seperti “hadir”, “tempat”, “manusia”, dan “mati” menunjukkan kritik dan keluhan.



Gambar 3. Word Cloud Seluruh Data



Gambar 4. Word Cloud Positif Dan Negatif

4. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, dapat ditarik beberapa kesimpulan utama:

- a. Penelitian ini berhasil mengimplementasikan tahapan *preprocessing* data teks dan ekstraksi fitur TF-IDF untuk mempersiapkan data *tweet* yang relevan.
- b. Penerapan algoritma SVM menghasilkan model klasifikasi dengan akurasi 74.47%, yang cukup efektif dalam menganalisis sentimen terhadap kinerja bupati.
- c. Analisis sentimen menunjukkan bahwa opini masyarakat di Twitter didominasi oleh sentimen negatif, diikuti oleh positif dan netral, mencerminkan adanya kritik yang signifikan terhadap kepemimpinan daerah.
- d. Pendekatan berbasis text mining dengan kerangka kerja CRISP-DM terbukti dapat menjadi alat yang berguna untuk mendukung evaluasi kebijakan pemerintah daerah.

5. DAFTAR PUSTAKA

- Ade Dwi Dayani, Yuhandri, dan Gunadi Widi Nurcahyo. (2024). Analisis Sentimen Terhadap Opini Publik Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine. *Jurnal Komtekinfo*, 11, 1–10.
- Alexander, Nyongki, Radja Bria, dan Arita Witanti. (2024). Analisis Sentimen Masyarakat Indonesia Menggunakan Algoritma Support Vector Machine Tentang Pilpres 2024. 7(6).
- Azrul, Ahmad, Ade Irma Purnamasari, dan Irfan Ali. (2024). Analisis Sentimen Pengguna Twitter Terhadap Perkembangan Artificial Intelligence Dengan Penerapan Algoritma Long Short-Term Memory (LSTM). 8(1), 413–21.
- Caroline, Fionna, Raden George, Samuel Budi, Muhammad Ezar, dan Al Rivian. (2024). Analisis Sentimen Masyarakat Terhadap Kasus Korupsi PT . Timah Menggunakan Metode Support Vector Machine. 4(1), 43–50.
- Emarapenta, Jesica, Br Sinulingga, Hizkya Cesar, dan Kayika Sitorus. (2024). Analisis Sentimen Masyarakat Terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF Sentiment Analysis Of Public Towards Indonesian Horror Films Using SVM And TF- IDF Methods. 14, 42–53.
- Fauzan, Muazim Rahman, Harma Oktafia, Lingga Wijaya, dan Joni Karman. (2023). Kenaikan Harga BBM Di Media Sosial Twitter. 1(1), 82–89.
- Hendrawan, Ari, dan Enny Itje Sela. (2024). Analisis Sentimen Komentar Youtube Tentang Resesi Global 2023 Menggunakan LSTM. *Jurnal Indonesia: Manajemen Informatika Dan Komunikasi*, 5(1), 587–93.
- Hendrastuty, Nirwana, Auliya Rahman Isnain, Ari Yanti Rahmadhani. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. 6(3), 150–55.
- Pratama, Raistiwar. (2024). Bupati Or Regent? Martanagara Of Bandung 1893–1918. 20(1).
- Putra, Aldiansyah, Dede Haeirudin, dan Hasna Khairunnisa. (2021). Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm.
- Runimeirati, Abdul Muis, dan Figur Muhammad. (2023). Pelatihan Text Mining Menggunakan Bahasa Pemrograman Python. *Abdimas Langkanae*, 3(1), 36–46.
- Seran, Yulius Bambang, dkk. (2024). Analisis Sentimen Masyarakat Terhadap Kinerja Kerja Presiden Joko Widodo Enggunakan Algoritma Support Vector Machine. 8(4), 7190–95.

- Syahputra, Dhimas Wida, Bayu Rahayudi, dan Lailil Muflikhah. (2022). Analisis Sentimen Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Metode Support Vector Machine. 6(3), 1067–72.
- Syam, Abd Azis, Galang Hardy M, Agus Salim, Dewi Fatmarani Suriyanto, dan Muhammad Fajar B. (2024). Analisis Teknik Preprocessing Pada Sentimen Masyarakat Terkait Konflik Israel-Palestina Menggunakan Support Vector Machine. 9(3), 1464–72.
- Wati, Risa, dan Siti Ernawati. (2021). Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python. Jurnal Teknik Informatika UNIKA Santo Thomas, 06, 240–47.
- Yanti, Fajri, Betha Nurina Sari, dan Sofi Defiyanti. (2024). Implementasi Algoritma LSTM Pada Peramalan Stok Obat (Studi Kasus: Puskesmas Beber). 8(4), 6082–89.
- Zufria, Ilka, Aidil Halim Lubis, dan Siti Septia Febiyaula. (2024). Kepolisian Republik Indonesia Menggunakan. 4307, 1266–72.