

# **Analisis Sentimen Media Sosial dengan Teknik Data Mining**

## **USULAN PENELITIAN**

**Disusun Oleh:**

**Tryanda Anggita Suwito  
2203010343**



**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNIK  
UNIVERSITAS PERJUANGAN TASIKMALAYA**

**2024**

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Di era digital ini, media sosial telah menjadi bagian integral dari kehidupan sehari-hari. Platform seperti Twitter, Facebook, dan Instagram tidak hanya digunakan untuk berkomunikasi, tetapi juga sebagai sumber informasi dan tempat untuk menyuarakan opini. Setiap hari, jutaan pengguna berbagi pikiran, perasaan, dan reaksi mereka terhadap berbagai isu, produk, dan layanan melalui media sosial. Data yang dihasilkan dari aktivitas ini merupakan tambang emas informasi yang dapat dimanfaatkan oleh berbagai pihak, mulai dari perusahaan hingga peneliti akademis.

Salah satu bidang penelitian yang muncul dari pertumbuhan media sosial adalah analisis sentimen. Analisis sentimen adalah proses mengidentifikasi dan mengkategorikan opini yang dinyatakan dalam teks, terutama untuk menentukan sikap penulis terhadap subjek tertentu, apakah positif, negatif, atau netral. Dalam konteks bisnis, analisis sentimen dapat membantu perusahaan memahami persepsi konsumen terhadap produk mereka, memungkinkan mereka untuk merespon dengan lebih cepat dan efektif terhadap kebutuhan dan keinginan pelanggan. Dalam konteks politik, analisis sentimen dapat digunakan untuk memantau opini publik terhadap kebijakan atau kandidat tertentu.

Namun, analisis sentimen tidaklah sederhana. Data dari media sosial seringkali tidak terstruktur, mengandung banyak noise, dan menggunakan bahasa informal serta slang yang dapat berubah dengan cepat. Selain itu, volume data yang sangat besar memerlukan teknik pengolahan data yang efisien dan efektif. Oleh karena itu, teknik data mining menjadi sangat penting dalam proses ini. Data mining, yang mencakup berbagai teknik untuk ekstraksi informasi dari dataset yang besar, dapat digunakan untuk mengidentifikasi pola dan tren dalam data sosial yang tidak langsung terlihat.

Snort, salah satu alat open-source yang populer digunakan dalam analisis keamanan jaringan, menunjukkan bagaimana alat open-source dapat dimanfaatkan untuk analisis data. Namun, dalam konteks analisis sentimen, berbagai teknik seperti klasifikasi teks, pengolahan bahasa alami (NLP), dan machine learning digunakan untuk mengembangkan model yang dapat mengkategorikan sentimen secara otomatis.

Penelitian ini bertujuan untuk mengeksplorasi dan mengimplementasikan teknik data mining untuk analisis sentimen media sosial. Dengan menggunakan data dari platform media sosial seperti Twitter, penelitian ini akan mengembangkan model yang mampu mengidentifikasi sentimen pengguna terhadap topik tertentu. Hasil dari penelitian ini diharapkan dapat memberikan wawasan yang berguna bagi perusahaan dan pembuat kebijakan dalam memahami opini publik dan mengambil keputusan yang lebih informasional.

Melalui penelitian ini, diharapkan dapat dikembangkan metode analisis sentimen yang tidak hanya akurat tetapi juga dapat diimplementasikan secara efisien pada skala besar. Dengan demikian, penelitian ini tidak hanya akan berkontribusi pada literatur akademis tetapi juga memiliki aplikasi praktis yang luas.

## **1.2 Rumusan Masalah**

1. Bagaimana cara mengumpulkan data sentimen yang relevan dari platform media sosial seperti Twitter secara efisien?
  - Apa saja sumber data yang akan digunakan?
  - Alat atau API apa yang akan digunakan untuk pengumpulan data?
2. Bagaimana cara melakukan pra-pemrosesan data teks dari media sosial agar siap untuk analisis sentimen?
  - Teknik apa yang digunakan untuk membersihkan data dari noise?
  - Bagaimana cara melakukan tokenisasi, stemming, dan penghapusan stopwords pada data teks?
3. Teknik data mining apa yang paling efektif untuk menganalisis sentimen dari data media sosial?
  - Algoritma klasifikasi apa yang akan digunakan (misalnya, Naive Bayes, SVM, neural networks)?
  - Bagaimana cara mengoptimalkan model untuk mencapai akurasi yang tinggi?

## **1.3 Batasan Masalah**

1. Penelitian ini akan fokus pada data yang diambil dari satu atau dua platform media sosial saja, misalnya Twitter dan Facebook. Hal ini dilakukan untuk memastikan kelayakan dalam pengumpulan dan analisis data.
2. Penelitian akan dibatasi pada teks berbahasa Indonesia. Data yang diperoleh dalam bahasa lain tidak akan dianalisis dalam penelitian ini.
3. Analisis sentimen akan dibatasi pada tiga kategori utama: positif, negatif, dan netral. Sentimen yang lebih spesifik atau kompleks (misalnya, sangat positif atau sangat negatif) tidak akan dibahas secara terpisah.

## **1.4 Tujuan Penelitian**

### **1. Mengembangkan Metode Pengumpulan Data:**

- Membangun metode yang efisien untuk mengumpulkan data sentimen dari platform media sosial seperti Twitter dan Facebook.
- Menggunakan alat atau API yang relevan untuk mengakses dan mengekstrak data secara sistematis.

### **2. Melakukan Pra-pemrosesan Data Teks:**

- Mengimplementasikan teknik pra-pemrosesan yang efektif untuk membersihkan data teks dari noise.
- Melakukan tokenisasi, stemming, dan penghapusan stopwords untuk mempersiapkan data teks bagi analisis lebih lanjut.

### **3. Mengembangkan Model Analisis Sentimen:**

- Menggunakan algoritma data mining seperti Naive Bayes, Support Vector Machine (SVM), dan neural networks untuk membangun model analisis sentimen.
- Mengoptimalkan model untuk mencapai tingkat akurasi, presisi, recall, dan F1-score yang tinggi.

## **1.5. Manfaat Penelitian**

- Pengembangan Ilmu Pengetahuan: Penelitian ini dapat memperkaya literatur dalam bidang analisis sentimen dan teknik data mining, terutama dalam konteks analisis data media sosial berbahasa Indonesia.
- Pengambilan Keputusan Bisnis: Hasil penelitian dapat digunakan oleh perusahaan untuk memahami opini konsumen terhadap produk atau layanan mereka. Informasi ini dapat membantu dalam merumuskan strategi pemasaran, pengembangan produk, dan peningkatan layanan pelanggan.
- Pengembangan Aplikasi: Penelitian ini dapat menghasilkan aplikasi atau sistem yang dapat digunakan untuk analisis sentimen secara otomatis. Aplikasi ini bisa bermanfaat bagi berbagai industri yang membutuhkan pemahaman tentang opini publik secara real-time.

## **1.6 Metodologi Penelitian**

Penelitian ini akan dilakukan dengan menggunakan pendekatan kuantitatif yang melibatkan pengumpulan, pra-pemrosesan, analisis, dan evaluasi data media sosial. Metodologi penelitian ini dibagi menjadi beberapa tahapan utama sebagai berikut:

1. Pengumpulan Data
2. Pra-pemrosesan Data
3. Pengembangan Model Analisis Sentimen
4. Evaluasi Model
5. Analisis dan Interpretasi Hasil
6. Implementasi dan Pengujian

## **1.7 Sistematika Penulisan**

Sistematika penulisan menggambarkan secara singkat organisasi penulisan laporan, serta isi dari setiap bagiannya.

### **BAB I        PENDAHULUAN**

Dalam bab ini di jelaskan latar belakang materi kerja praktik, tujuan pembahasan dan sistematika pembahasan.

### **BAB II        TINJAUAN UMUM**

Bab ini menjelaskan sejarah singkat dan struktur organisasi di instansi DPMPTSP Kota Tasikmalaya.

### **BAB III        METODOLOGI**

Dalam bab ini berisi tentang metode dan uraian landasan teori sebagai dasar untuk menganalisis pokok permasalahan dengan menjelaskan teori-teori yang berasal dari studi pustaka yang berkaitan dengan laporan yang dibahas.

### **BAB IV        HASIL DAN PEMBAHASAN**

Bab ini berisi tentang hasil dan pembahasan masalah yang meliputi analisis sistem yang ada dan analisi sistem usulan dengan alat bantu diagram aliran data, normalisasi table, dialog layar, rancangan sistem masukan dan keluaran serta pembahasan dan pengujian implementasi sistem dari hasil rancangan yang dilakukan.

## **BAB V            KESIMPULAN DAN SARAN**

Bagian kesimpulan dan saran memuat rangkuman yang merupakan hasil dari keseluruhan penelitian yang telah dilakukan serta menjadi jawaban dari tujuan penelitian yang telah ditetapkan. Bagian ini juga berisi saran-saran yang didapatkan berdasarkan kekurangan yang ditemukan pada penelitian ini.

## **DAFTAR PUSTAKA**

Pada bagian daftar Pustaka berisi sumber-sumber kepustakaan yang digunakan dalam penelitian, baik berupa jurnal penelitian, buku, prosiding, ataupun sumber pustakawan lainnya.

## **LAMPIRAN**

Bagian lampiran berisi keterangan tabel, gambar, dan hal-hal lain yang wajib dilampirkan untuk memperjelas uraian dalam laporan.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Analisis Sentimen Media Sosial**

Analisis sentimen adalah proses untuk mengidentifikasi dan mengevaluasi sentimen, opini, atau sikap yang terkandung dalam teks. Dalam konteks media sosial, analisis sentimen memungkinkan kita untuk memahami opini dan reaksi pengguna terhadap berbagai topik, seperti produk, layanan, atau peristiwa. Konsep ini menjadi penting karena membantu perusahaan dan pembuat kebijakan dalam memahami persepsi dan kebutuhan pelanggan serta memantau opini publik.

#### **2.2 Teknik Data Mining dalam Analisis Sentimen**

Teknik data mining merupakan kumpulan metode dan algoritma untuk mengekstrak pola atau pengetahuan yang bermanfaat dari dataset yang besar dan kompleks. Dalam analisis sentimen, teknik data mining digunakan untuk mengidentifikasi pola dan tren dalam data teks yang diperoleh dari media sosial. Beberapa teknik data mining yang umum digunakan dalam analisis sentimen meliputi klasifikasi teks, pengolahan bahasa alami (NLP), dan machine learning.

#### **2.3 Studi Terkait**

Studi-studi terdahulu telah dilakukan dalam bidang analisis sentimen menggunakan teknik data mining. Beberapa penelitian telah berhasil mengembangkan model yang efektif untuk mengklasifikasikan sentimen secara otomatis, sementara yang lain berfokus pada pengembangan metodologi baru untuk meningkatkan akurasi dan efisiensi analisis sentimen. Tinjauan terhadap literatur ini menjadi penting untuk memahami perkembangan terkini dalam bidang analisis sentimen dan menentukan pendekatan yang tepat untuk penelitian yang dilakukan.

**BAB III**  
**METODOLOGI PENELITIAN**