

Analisis Sentimen Layanan Pelanggan Provider Internet Dengan Algoritma Support Vector Machine Dan Naïve Bayes

Oleh:

Muhammad Arshiel Naufal Dzaki ,

Hindarto

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Mei, 2025



Pendahuluan

Di era digital saat ini, media sosial tidak hanya digunakan untuk hiburan, tetapi juga sebagai wadah menyuarakan opini, termasuk keluhan dan pujian terhadap layanan internet. Hal ini membuka peluang bagi perusahaan seperti Biznet untuk melakukan analisis sentimen guna memahami kebutuhan pelanggan dan meningkatkan kualitas layanan. Biznet sendiri menghadapi berbagai keluhan terkait gangguan koneksi dan layanan pelanggan yang kurang memuaskan. Dengan menganalisis data dari media sosial secara real-time menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM), perusahaan dapat mengidentifikasi masalah dan meresponsnya lebih cepat. Analisis sentimen ini mengklasifikasikan opini ke dalam tiga kategori, yaitu positif, negatif, dan netral, melalui proses text mining yang mencakup pengambilan, pemrosesan, dan analisis data. Penelitian ini bertujuan mengimplementasikan data mining untuk menganalisis sentimen masyarakat terhadap layanan Biznet berdasarkan data dari X.com, sehingga diharapkan dapat membantu meningkatkan kualitas layanan internet di Indonesia serta memberikan kontribusi di bidang akademik dan industri.

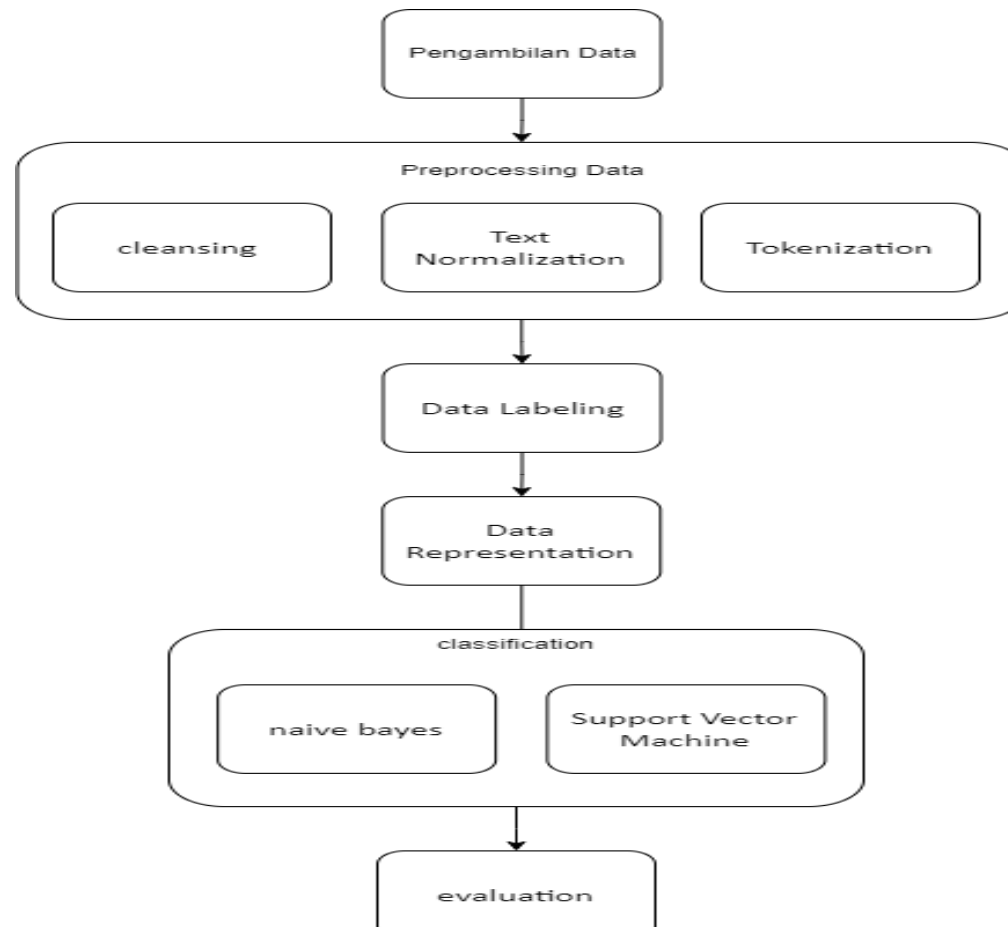
Pertanyaan Penelitian (Rumusan Masalah)

Rumusan Masalah

1. bagaimana cara merancang sistem analisis sentimen untuk kebutuhan pelayanan pelanggan?
2. mencari tahu perbandingan dari algoritma Naive Bayes dan Support Vector Machines (SVM)?

Metode

DIAGRAM ALUR



Metode

- **1. Pengumpulan Data**

Mengambil 7.272 data dari X.com (Twitter) menggunakan Python dan Tweepy, disimpan dalam format CSV.

- **2. Pemrosesan Data**

Meliputi:

Cleansing: Menghapus data tidak relevan (spam, iklan).

Normalisasi: Huruf kecil, hapus tanda baca, ubah kata tidak baku.

Tokenisasi: Memisah teks jadi kata-kata untuk analisis lanjutan.

- **3. Pelabelan Data**

Klasifikasi sentimen (positif, negatif, netral) menggunakan **VADER Sentiment Analysis** berbasis leksikon.

- **4. Representasi Data**

Menggunakan metode **TF-IDF** untuk mengubah teks menjadi angka berdasarkan bobot kata dalam dokumen.

Metode

- **5.klasifikasi**

- a) Naive Bayes adalah algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes. Algoritma ini mengasumsikan bahwa fitur-fitur (kata dalam teks) saling independen satu sama lain. Cocok untuk klasifikasi teks karena sederhana, cepat, dan efektif meskipun dengan data besar.
- b) Support Vector Machine adalah algoritma klasifikasi yang bekerja dengan mencari garis pemisah terbaik (hyperplane) untuk memisahkan data ke dalam kelas berbeda. Sangat efektif untuk data berdimensi tinggi seperti teks, dan dikenal memiliki akurasi tinggi serta tangguh terhadap data tidak seimbang

- **6.evaluasi**

Evaluasi model dilakukan untuk mengukur kinerja algoritma dalam mengklasifikasikan sentimen dengan metrik seperti akurasi, precision, recall, F1-score, dan confusion matrix.

Hasil

Evaluasi dilakukan menggunakan confusion matrix untuk mengukur akurasi, presisi, recall, dan F1-score pada algoritma SVM dan Naive Bayes dengan pembagian data menggunakan rasio 80:20, sehingga menghasilkan variasi dalam hasil evaluasi.

Algoritma	Akurasi	Precision	Recall	F1-Score
Multinomial Naïve Bayes	87,59%	93%	50%	65%
Bernoulli Naïve Bayes	88,21%	88%	69%	77%
Complement Naïve Bayes	82,38%	82%	85%	83%
SVM (Linear Kernel)	93,92%	99%	74%	84%
SVM (RBF Kernel)	91,32%	99%	66%	79%
SVM (Polynomial Kernel)	89,45%	98%	58%	73%
SVM (Sigmoid Kernel)	94,29%	96%	76%	85%

Pembahasan

SVM dengan kernel Sigmoid dan Linear menunjukkan performa terbaik secara keseluruhan dalam mengklasifikasikan sentimen pelanggan Biznet, dengan keunggulan pada akurasi dan precision yang tinggi serta keseimbangan recall yang baik. Hal ini mengindikasikan SVM lebih efektif dalam menangani data teks dengan dimensi tinggi dan klasifikasi sentimen. Meskipun demikian, Naive Bayes, khususnya varian Bernoulli dan Complement, tetap memberikan hasil yang kompetitif dan dapat menjadi alternatif yang efisien dalam skenario tertentu, meskipun performanya sedikit lebih rendah dibandingkan SVM.

Temuan Penting Penelitian

- **Akurasi tertinggi** dicapai oleh SVM kernel Sigmoid (94,29%) dan kernel Linear (93,92%).
- **Precision sangat tinggi** pada semua varian SVM (96%–99%), menandakan prediksi positif yang sangat akurat.
- **Recall terbaik** diperoleh oleh Complement Naive Bayes (85%), diikuti oleh SVM Sigmoid (76%).
- **F1-Score tertinggi** adalah SVM kernel Sigmoid (85%), menunjukkan keseimbangan optimal antara precision dan recall.
- Dari Naive Bayes, **Bernoulli Naive Bayes** memiliki performa terbaik dengan akurasi 88,21% dan F1-score 77%.

Manfaat Penelitian

- **Meningkatkan Pemahaman Sentimen Pelanggan**

Membantu perusahaan Biznet memahami opini dan perasaan pelanggan secara lebih akurat melalui analisis sentimen di media sosial.

- **Mendukung Pengambilan Keputusan**

Hasil analisis dapat digunakan sebagai dasar untuk perbaikan layanan dan strategi pemasaran Biznet agar lebih sesuai dengan kebutuhan pelanggan.

- **Mengoptimalkan Penggunaan Algoritma**

Memberikan insight tentang kelebihan dan kekurangan algoritma Naive Bayes dan SVM dalam analisis sentimen, sehingga memudahkan pemilihan metode yang tepat.

- **Kontribusi pada Pengembangan Teknologi Text Mining**

Menambah referensi dan penerapan teknik text mining khususnya untuk analisis sentimen di konteks layanan internet provider.

- **Efisiensi Analisis Data Media Sosial**

Membantu mengotomatisasi proses pengolahan data besar dari media sosial sehingga analisis lebih cepat dan sistematis

Referensi

- [1] F. D. Ananda And Y. Pristyanto, “Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine,” *Matrik J. Manajemen, Tek. Inform. Dan Rekayasa Komput.*, Vol. 20, No. 2, 2021, Doi: 10.30812/Matrik.V20i2.1130.
- [2] B. W. Sari And F. F. Haranto, “Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet,” *J. Pilar Nusa Mandiri*, Vol. 15, No. 2, 2019, Doi: 10.33480/Pilar.V15i2.699.
- [3] S. Rahayu And K. Handoko, “Implementasi Data Mining Untuk Memprediksi Pengaruh Media Sosial Terhadap Semangat Belajar Anak,” *J. Comasie*, Vol. 11, No. 02, 2024.
- [4] M. Bohlouli, J. Dalter, M. Dornhöfer, J. Zenkert, And M. Fathi, “Knowledge Discovery From Social Media Using Big Data-Provided Sentiment Analysis (Somabit),” 2015. Doi: 10.1177/0165551515602846.
- [5] S. B. Kotsiantis, I. D. Zaharakis, And P. E. Pintelas, “Machine Learning: A Review Of Classification And Combining Techniques,” *Artif. Intell. Rev.*, Vol. 26, No. 3, 2006, Doi: 10.1007/S10462-007-9052-3.
- [6] E. Rahm And H. Do, “Data Cleaning: Problems And Current Approaches,” *Ieee Data Eng. Bull.*, Vol. 23, No. 4, 2000.

Referensi

- [7] D. Jurafsky And J. H. Martin, “Speech And Language Processing An Introduction To Natural Language Processing, Computational Linguistics, And Speech Recognition With Language Models Third Edition Draft Summary Of Contents.”
- [8] C. D. Manning, P. Raghavan, And H. Schütze, *Introduction To Information Retrieval*. 2008. Doi: 10.1017/Cbo9780511809071.
- [9] C. J. Hutto And E. Gilbert, “Vader: A Parsimonious Rule-Based Model For Sentiment Analysis Of Social Media Text,” In *Proceedings Of The 8th International Conference On Weblogs And Social Media, Icwsn 2014*, 2014. Doi: 10.1609/Icwsn.V8i1.14550.
- [10] A. Rajaraman And J. D. Ullman, *Mining Of Massive Datasets*, Vol. 9781107015357. 2011. Doi: 10.1017/Cbo9781139058452.
- [11] A. McCallum And K. Nigam, “A Comparison Of Event Models For Naive Bayes Text Classification,” *Aaai/Icml-98 Work. Learn. Text Categ.*, 1998, Doi: 10.1.1.46.1529.
- [12] C. Cortes And V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, Vol. 20, No. 3, 1995, Doi: 10.1023/A:1022627411411.
- [13] J. W. Iskandar And Y. Nataliani, “Perbandingan Naïve Bayes, Svm, Dan K-Nn Untuk Analisis Sentimen Gadget Berbasis Aspek,” *J. Resti (Rekayasa Sist. Dan Teknol. Informasi)*, Vol. 5, No. 6, 2021, Doi: 10.29207/Resti.V5i6.3588.

