

# Sentimen Analisis Komentar Mahasiswa EDOM Dengan Metode Support Vector Machine (SVM)

**Ardan Nur Chaiya Putra, Sam Farisa Caherul Haviana, Badie'ah**

\* Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Sultan Agung Semarang

Correspondence Author: [ardanncp@std.unissula.ac.id](mailto:ardanncp@std.unissula.ac.id)

## Abstract

Universitas Sultan agung Semarang menerapkan sebuah kuisisioner online mahasiswa untuk mengevaluasi Dosen apakah sebuah dosen sudah melakukan tugas yang diberikan dan mahasiswa menilai apakah dosen sudah baik apa belum yang bisa di isi dengan melalui pilihan ganda dan sebuah komentar. Dalam hal ini yang akan digunakan adalah kuisisioner dalam bentuk komentar yang dibuat oleh mahasiswa dapat menghasilkan sebuah sentiment yang terbagi menjadi jadi tiga subjek berupa netral, positif, dan negatif. Namun terkadang dalam sebuah tulisan yang di buat oleh mahasiswa tidak bisa di identifikasi untuk menentukan apakah kalimat tersebut dapat dibagi menjadi tiga subjek yang berupa positif, negatif dan netral. Data yang berhasil didapatkan dari pengisian komentar dari kuisisioner online Bernama EDOM dengan mendapatkan hasil dengan jumlah data sekitar 1791 komentar, dan dengan penelitian ini akan menggunakan sentiment analisis dan algoritma Bernama Support Vector Machine yang menghasilkan hasil yang baik dengan tingkat Akurasi 82%, Presisi 72%, recall 61%, f-measure 65%..

Keyword: Klasifikasi, Komentar, Sentimen, *support vector machine*

## 1. PENDAHULUAN (10 PT)

### 1.1 Latar Belakang

Universitas Sultan agung Semarang menerapkan kuisisioner online dalam rangka mengevaluasi dosen. Kuisisioner ini di isi oleh mahasiswa sebagai bentuk penilaian kepada dosen yang terbagi menjadi pilihan dan sebuah komentar. Dan evaluasi ini memperoleh sebuah nilai subjek yang dapat dibagi menjadi 3 yaitu netral, positif, dan negatif akan tetapi untuk menemukan nilai sebuah objek dalam jumlah yang banyak terkadang sulit oleh sebab itu dalam makalah ini di buatlah sebuah sistem yang mempermudah mencari nilai subjek dari sebuah komentar yang Bernama *sentiment analysis*.

Sedangkan *sentiment analysis* ialah riset komputer dari sentimen, emosi, dan opini yang dikeluarkan secara teks lalu di pisah menjadi kelompok sentimen positif dan negatif, dan dari beberapa metode yang telah di teliti menghasilkan nilai yang berbeda, yang membandingkan *naïve bayes* dengan *support vector machine* yang menunjukkan bahwa SVM menunjukkan hasil dengan akurasi sebesar 81.10% dengan nilai AUC 0.904. dari hasil yang di bandingkan pada algoritma *feature selection* dengan *informatic gain*, *chi square*, *forward selection*, dan *backward elimination* mendapatkan sebuah parameter dengan nilai sebesar 200 yang paling bagus, yang mendapatkan nilai rata-rata akurasi sebesar 84,57% dan nilai AUC sebesar 0.899 (1).

Yang berdasar pada latar belakang permasalahan pada paper ini ialah Bagaimana Hasil dari sebuah algoritma *support vector machine* dalam melakukan analisis sentimen terhadap komentar mahasiswa di EDOM (evaluasi dosen oleh mahasiswa).

Bagaimana membuat system dengan metode *support vector machine* untuk analisis sentimen pada mahasiswa.

Yang menghasilkan tujuan penelitian berupa Menganalisa sentiment analisis komentar mahasiswa di EDOM (evaluasi dosen oleh mahasiswa) menggunakan *support vector machine*.

Merancang dan membangun aplikasi untuk klasifikasi sentiment dengan metode *support vector machine*.

### 1.2 Landasan Teori

Tinjauan Pustaka yang berasal dari penelitian sebelumnya yang dimulai dari penelitian menggunakan *support vector system* yang digunakan untuk penelitian sentimen analisis terhadap hasil evaluasi dokumen terpecah menjadi 3 bagian yaitu 103 positif, 163 negatif, dan 41 netral sentiment dengan akurasi sebesar 49.54% (1).

Kemudian penelitian yang melakukan analisis sentiment evaluasi dosen menggunakan TF-IDF dengan menghasilkan rata-rata 82% (2).

Dan penelitian berikutnya menggunakan *naïve boys* menggunakan 100 data tes yang memiliki sentimen manual mendapatkan akurasi 75% sehingga system mempunyai akurasi cukup tinggi (3).

Dan pada penelitian berikutnya mendapatkan hasil penelitiandan hasil pengujian algoritma SVM untuk kebijakan *lockdown* di Jakarta menggunakan twitter Bahasa Indonesia dengan fitur *tf-idf* yang menghasilkan hasil akurasi = 92%, presisi = 75%, *recall* = 92 % dan *f-score* = 83% (4).

Dan tinjauan terakhir menghasilkan bahwa analisis sentimen dengan metode SVM berhasil klasifikasi sentimen masyarakat ambon berdasarkan komentar di twitter dan facebook. Yang menghasilkan hasil sentiment netral lebih tinggi dari sentiment positif dan negative, yang terdiri netral 45%, positif 28%, dan negative 27% yang menunjukkan tidak puas terhadap program yang diterapkan oleh pemerintahan ambon yang perlu adanya sosialisasi (5).

### 1. Sentiment Analysis

Ialah suatu tulisan apakah apakah bersubjek positif, negatif, dan netral dari mendapatkan opini atau sikap dari seorang pembicara (6). Atau yang bisa diartikan sentimen analisis atau *mining opinion* adalah riset sentiment, anggapan dan rasa yang diciptakan secara harfiah yang diklasifikasikan dengan menjadi kelompok positif, negatif, dan netral. Sistem dapat mendapatkan akurasi yang tinggi yang melakukan perubahan terlalu berat pada dokumen dan sistem menggunakan jarak lebih dari 0 pada kelas positif sedangkan pada kelas negatif kurang dari 0 serta 0 untuk netral (1)

### 2. Machine Learning

Teknik untuk melakukan inferensi kepada data dengan melakukan pendekatan sistematis (7). Dengan pusat *machine learning* untuk menciptakan model (matematis) yang kemudian mereflesikan pola-pola data. Sedangkan *machine learning* ialah ilmu untuk mempelajari peningkatan metode computer yang menukar data menjadi polah yang pintar atau secara sederhana dapat disamakan sebagai cara mengganti data menjadi informasi (8)

### 3. Text Preprocessing

Ialah serangkaian Teknik yang digunakan sebelum penerapan metode penambangan data (9). Dan bagian terpenting dalam *language preprocessing*, dari karakter, kata dan kalimat untuk diidentifikasi terbagi menjadi 3 yaitu *Case folding*, *stemming*, *tokenisasi*.

### 4. Delta Term Frequency-Inverse Document Frequency (Delta TF-IDF)

Ialah dalam kantong kata setiap kata ngram pasangan kata diasosiasikan dengan sebuah nilai. Nilai-nilai ini biasanya jumlah kata mereka dalam dokumen. Terkadang nilai mengukur seberapa jarang istilah ini ada di dokumen lain. Sebagai gantinya, kamu menimbang Nilai-nilai ini dengan seberapa bias mereka terhadap satu corpus (10)

### 5. Support vector machine

Ialah algoritma *machine learning* yang memprediksi suatu kelas menurut dari pola pelatihan Langkah hasil yang dibuat oleh Vladimir vapnik. Klasifikasi dijalankan pada sebuah garis pembatas yang menafsirkan pada suatu objek hipotesis *positive* dan *negative* secara perseptis pada sebuah garis pembatas, *hyperplane* yang baik ialah mempunyai titik pada jarak yang besar ke data pelatihan pada setiap kelas, yang pada lazimnya besar suatu *margin* akan akan semakin kecil *error* abtaksi pemilah (1).

### 6. Pengujian metode klasifikasi

Untuk menguji seberapa besar efektivitas performa metode klasifikasi, empat parameter statistic digunakan untuk mengujinya, yang disebut *acuracy*, *precision*, *recal*, dan *f-measure*.

## 2. METODE PENELITIAN

### 2.1. Metode Pengambilan Data

Data yang dipakai dalam penelitian adalah data EDOM yang di dapatkan dari <http://edom.unissula.ac.id/> data yang diambil hanya berupa komentar yang disusun dalam file CSV.

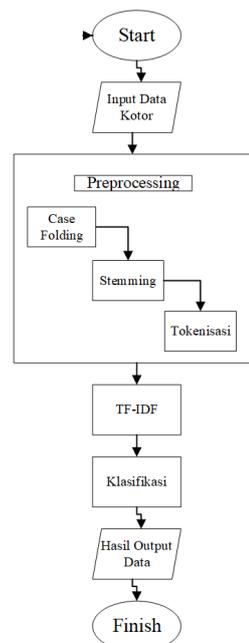
### 2.2. Metode Pengembangan Sistem

*Prototyping* adalah memudahkan pengguna untuk melakukan evaluasi terhadap desain yang diajukan oleh pengembang aplikasi dengan menggunakan langsung yang lebih efektif untuk menghindari terjaninya suatu *bug* atau *error* yang tidak diinginkan dengan alur sebagai berikut :

1. Pengumpul Kebutuhan
2. Desain *prototype*
3. Membangun *prototype*
4. Evaluasi
5. Perbaiki *prototype*
6. Membangun aplikasi secara keseluruhan

### 2.3. Perancangan Arsitektur Sistem

Dalam perencanaan arsitektur sistem diperlukan *flowchart* yang menunjukkan Langkah-langkah bagaimana sistem akan berjalan .



Gambar 1. Flowchart Arsitektur Sistem

## 3. HASIL DAN ANALISA

### 3.1 Hasil Implementasi *support vector machine*

Dalam melakukan pelaksanaan algoritma *support vector machine*, penelitian ini melakukan 3 tahap di awal yaitu pengumpulan data, melakukan pemrosesan data agar siap dilakukan *training* dan *training model support vector machine* yang dipakai nanti oleh program.

Pada tahap pengumpulan data, peneliti meminta data dari sistem EDOM unissula. Dan mendapatkan data komentar yang berjumlah 1791 komentar beserta sentimennya yang akan digunakan berupa data training 70% dan sisnya akan digunakan di data pengujian dan validasi sejumlah 537 data komentar, kemudian data yang sudah diproses dengan melakukan *preprocessing* yang terdiri dari *case folding*, *stemming*, tokenisasi.

Proses pertama dalam *preprocessing* akan digunakan ialah *case folding* yang merubah semua komentar yang memiliki huruf besar dan kecil agar di rubah ke huruf kecil semua setelah itu data komentar akan di proses oleh *stemming* yaitu kata-kata yang memiliki kata dasar sama akan di hitung sebagai kata yang sama. Sehingga algoritma mendapatkan hasil yang baik.

Proses yang akan dilangsungkan setelah *stemming* ialah tokenisasi yang membuat kata menjadi token, selanjutnya setelah melewati proses *preprocessing* Langkah berikutnya akan menggunakan TF-IDF agar data komentar memiliki bobot pada setiap kata-kata penting yang menjadi fitur data proses *training* algoritma. Selanjutnya setelah selesai pembobotan kata selesai, selanjutnya adalah menggunakan metode *Support Vector Machine* (SVM) yang memakai kernel RBF dan setelah selesai proses selesai, selanjutnya dilakukan pengujian.

Untuk memenuhi tujuan penelitian, maka diperlukan pengujian . penelitian ini akan menggunakan pengujian berupa akurasi, presisi, *recall*, *F-measure* untuk melihat performa metode *Support Vector Machine* dengan melakukan klasifikasi.

Tabel 1. Tabel hasil pengujian klasifikasi

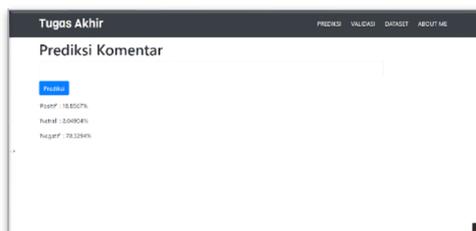
Pengujian	Hasil
Akurasi	0.82
Presisi	0.72

Recall	0.62
F-measure	0.65

Presisi 82%, Akurasi 72%, *Recall* 61%, *F-measure* 65%. Dalam hasil metode klasifikasi menampilkan hasil total data Mahasiswa EDOM berjumlah 1791. Yang hasil validasinya sudah baik, maka selanjutnya mode yang selesai training akan disimpan agar bisa digunakan dalam program.

### 3.2 Hasil Implementasi Sistem

Setelah hasil training sudah sesuai dengan keinginan, selanjutnya model *Support Vector Machine* di implementasi di sistem. Ini adalah hasil tangkapan layar pada sistem yang dibuat:



Gambar 2. Halaman Prediksi



Gambar 3. Halaman Validasi

## 4. KESIMPULAN

Setelah melakukan penelitian dan implementasi sistem mendapatkan kesimpulan sebagai berikut :

1. Algoritma *Support Vector Machine* dapat dipakai untuk melakukan klasifikasi komentar mahasiswa EDOM (evaluasi dosen oleh mahasiswa). Dengan performa klasifikasi yang dihasilkan oleh algoritma *Support Vector Machine* juga bisa disimpulkan baik, karena hasil pengujian model setelah training mencapai skor rata-rata Akurasi 82%, presisi 72%, recall 63%, F-measure 65%.
2. Implementasi sistem juga disimpulkan baik karena metode *Support Vector Machine* ini banyak digunakan sebagai salah satu algoritma dan mempunyai dokumentasi yang lengkap.
3. Pada pengujian sistem ada beberapa masalah seperti sistem merespon sangat lama dan jarang berubah

## ACKNOWLEDGEMENTS

Terima kasih atas fakultas Teknologi Industri dan Teknik Informatika atas paper tugas akhir mahasiswa ini.

## DAFTAR PUSTAKA

1. Santoso VI, Virginia G, Lukito Y. Penerapan Sentiment Analysis Pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine. *Jurnal Transformatika*. 2017;14(2):72.
2. Sulaeman AF, Supianto AA, Bachtiar FA. Analisis Sentimen Opini Mahasiswa Terhadap Saran Evaluasi Kinerja Dosen Menggunakan TF-IDF dan Support Vector Machine . 2019;3(6):5647–55.
3. Setiadi B S. Fakultas Teknik – Universitas Muria Kudus 153. *Prosiding SNATIF ke-4 Tahun 2017*. 2017;153–60.
4. Rahman Isnain A, Indra Sakti A, Alita D, Satya Marga N. Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jdmsi [Internet]*. 2021;2(1):31–7. Tersedia pada: <https://t.co/NfhnmJtXw>
5. Tuhuteru H. Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine. *Information System Development (ISD)*. 2020;5(2):7–13.

6. Lexalytics. Sentiment Analysis Explained [Internet]. Lexalytics.Com. 2019. Tersedia pada: <https://www.lexalytics.com/technology/sentiment-analysis>
7. Putra JWG. Pengenalan konsep pembelajaran mesin dan deep learning. Computational Linguistics and Natural Language Processing Laboratory [Internet]. 2019;4(July):1–235. Tersedia pada: <https://www.researchgate.net/publication/323700644>
8. Habibi M, Cahyo PW. Journal Classification Based on Abstract Using Cosine Similarity and Support Vector Machine. JISKA (Jurnal Informatika Sunan Kalijaga). 2020;4(3):48.
9. García S, Ramírez-Gallego S, Luengo J, Benítez JM, Herrera F. Big data preprocessing: methods and prospects. Big Data Anal. 2016;1(1).
10. Martineau J, Finin T, Fink C, Piatko C, Mayfield J, Syed Z. Delta TFIDF: An Improved Feature Space for Sentiment Analysis. Proceedings of the Second International Conference on Weblogs and Social Media (ICWSM [Internet]. 2008;29(May):490–497. Tersedia pada: <http://ebiquity.umbc.edu/papers/select/person/Tim/Finin/>