

**ANALISIS SENTIMEN PADA DATA KUESIONER EVALUASI DOSEN OLEH MAHASISWA
(EDOM) PRODI SISTEM INFORMASI TELKOM UNIVERSITY MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE
ANALYSIS OF SENTIMENT IN THE DATA EVALUATION OF DOSAGE STUDENTS (EDOM)
INFORMATION SYSTEMS TELKOM UNIVERSITY USING THE SUPPORT VECTOR
MACHINE ALGORITHM**

Indah Nur Fitri Astuti¹, Irfan Darmawan², Dita Pramesti³

Prodi S1 Sistem Informasi, Fakultas Rekayasa Industri, Universitas Telkom

¹ indahnf@students.telkomuniversity.ac.id, ² irfandarmawan@telkomuniversity.ac.id,

³ ditapramesti@telkomuniversity.ac.id

Abstrak

Universitas Telkom khususnya Fakultas Rekayasa Industri program studi Sistem Informasi menerapkan adanya kuesioner online dalam rangka penilaian kinerja terhadap tenaga pengajar (dosen) dengan menggunakan kuesioner yang diisi oleh mahasiswa pada setiap pergantian semester sebelum dilakukannya Ujian Akhir Semester (UAS). Dengan cara menuliskan kritik dan saran terhadap kegiatan perkuliahan. Evaluasi Dosen Oleh Mahasiswa atau yang biasa disebut dengan EDOM memiliki peranan yang sangat penting. Selain itu, data EDOM juga mampu untuk meningkatkan kualitas dalam pembelajaran dan standarisasi akademik yang perlu dievaluasi sehingga menghasilkan mahasiswa-mahasiswa yang berkualitas. Penelitian ini dilakukan dengan analisis sentimen menggunakan metode atau algoritma klasifikasi support vector machine (SVM), metode ini bisa digunakan untuk mengklasifikasikan opini ke dalam kelas positif, negatif maupun netral. Data yang digunakan adalah data komentar 1 semester ganjil tahun 2019/2020, data komentar yang terkumpul terdapat 2.465 dari data dosen pengajar aktif yang ada di Universitas Telkom Prodi Sistem Informasi. Dari data komentar ini bisa dilihat apakah yang ditulis oleh mahasiswa merupakan penilaian yang bernilai positif, negatif maupun netral. Kemudian dilakukannya analisis sentimen terhadap komentar, beberapa tahapan untuk analisis sentimen ini adalah pre-processing, ekstraksi fitur, klasifikasi dan evaluasi. Hasil dari penelitian ini untuk menentukan klasifikasi dari data EDOM sehingga hasilnya mudah untuk dibaca. Klasifikasi dalam menggunakan metode ini memperoleh tingkat akurasi 75%, precision 75%, recall 75%, dan f-1 score 74%.

Kata kunci: Telkom University, Analisis Sentimen, Support Vector Machine, Klasifikasi, Evaluasi Dosen Oleh Mahasiswa

Abstract

Telkom University in particular the Faculty of Industrial Engineering Information Systems study program applies an online questionnaire in the context of evaluating the performance of teaching staff (lecturers) using a questionnaire filled out by students at each turn of the semester before the end of the Semester Examination (UAS). By writing criticism and suggestions for lecturing activities. Lecturer Evaluation by Students or commonly referred to as EDOM has a very important role. In addition, EDOM data is also able to improve the quality of learning and academic standardization that needs to be evaluated so as to produce quality students. This research was conducted by sentiment analysis using a support vector machine (SVM) classification method or algorithm, this method can be used to classify opinions into positive, negative or neutral classes. The data used are commentary data 1 odd semester of 2019/2020, the comment data collected there are 2,465 from the data of active lecturers at Telkom University Information System Study Program. From this commentary data it can be seen whether what was written by the students is a rating that is positive, negative or neutral. Then do a sentiment analysis of comments, some stages for sentiment analysis are pre-processing, feature extraction, classification and evaluation. The results of this study determine the classification of EDOM data so that the results are easy to read. Classification in using this method obtained 75% accuracy, 75% precision, 75% recall, and f-1 score 74%.

Keywords: Telkom University, sentiment analysis, Support Vector Machine, classification, evaluation of lecturers by students

1. Pendahuluan

Kinerja adalah sebuah penilaian yang sistematis atas seseorang mengenai prestasi dalam pekerjaannya yang menggambarkan secara sistematis tentang kekuatan dan kelemahan yang terkait dengan pekerjaannya^[1]. Dosen merupakan salah satu elemen yang berperan penting dalam kegiatan belajar mengajar dan berkaitan dengan mahasiswa. Seseorang dosen adalah merupakan sosok yang berkewajiban untuk menyebarluaskan, mengembangkan ilmu dan mentransformasikan pengetahuan. Dosen khususnya di perguruan tinggi senantiasa harus meningkatkan pengetahuan dan cara mengajar yang efektif dan efisien agar dapat memberikan ilmu yang dapat diterima dengan baik oleh mahasiswa.

Universitas Telkom Bandung, menerapkan adanya pengisian kuesioner online untuk penilaian kinerja dosen pada akhir semester sebelum UAS (Ujian Akhir Semester). Fakultas Rekayasa Industri (FRI) merupakan salah satu fakultas yang menerapkan pengisian kuesioner online. Kuesioner online diisi oleh mahasiswa yang terdiri dari kuesioner tingkat kepuasan dan kuesioner komentar terhadap matakuliah dan dosen pengampu matakuliah yang bersangkutan. Pada penelitian ini penulis hanya menggunakan data komentar mahasiswa saja. Komentar yang sudah dikumpulkan kemudian dikelompokkan oleh penulis menjadi tiga kategori, yaitu komentar positif, negatif, dan netral. Saat ini, untuk pelabelan data EDOM dilakukan secara manual oleh penulis. Dengan adanya komentar dari data EDOM ini penulis melakukan penelitian untuk mengetahui tingkat akurasi analisis sentimen dari komentar mahasiswa terhadap dosen dengan menggunakan algoritma SVM.

Untuk mengatasi masalah diatas, maka dibutuhkan analisis yang bisa mengklasifikasikan opini ke dalam kelas positif, negatif atau netral. Klasifikasi ke tiga kelas sentimen ini disebut dengan analisis sentimen. Penelitian ini menggunakan metode *Support Vector Machine* (SVM) karena pada penelitian-penelitian sebelumnya yang pernah dilakukan, metode ini memiliki nilai akurasi yang cukup tinggi. Maka dari itu penulis mencari tingkat akurasi dan bagaimana mengimplementasikan algoritma SVM dalam analisis. Dengan menggunakan metode SVM ini maka peneliti dapat bisa melakukan analisis sentimen dengan gampang dalam melakukan klasifikasi, dengan tingkat akurasi dan efektivitas yang tinggi dan cepat. Sehingga hasil pengolahan dan analisis kemudian divisualisasikan agar membantu proses evaluasi kinerja dosen dan pengambilan langkah strategis dalam mengembangkan teknik pengajaran yang lebih baik.

2. Dasar Teori

2.1 Igracias

Integrated Academic Information System atau yang biasa disebut dengan *igracias* yaitu merupakan portal akademik yang dimiliki oleh Universitas Telkom untuk mengintegrasikan sistem informasi perguruan tinggi yang dirancang sebagai kepentingan yang komprehensif seperti layanan pendidikan, maupun pendukung yang dijalankan dengan berbasis *igracias*. *Igracias* juga memiliki sistem dengan berbagai komponen seperti akademis dan non-akademis, maka *igracias* pun dapat memudahkan pengguna dalam mengakses disatu tempat. Pengguna yang dapat mengakses *igracias* adalah tenaga pendidik (dosen), para pengelola, mahasiswa maupun staf^[2].

2.2 Text Mining

Text mining ialah salah satu proses yang digunakan untuk mengklasifikasi dokumen. Data yang didapatkan oleh text mining yaitu berupa pola yang bernilai unik dari beberapa data yang berjumlah tinggi^[3]. Text mining merupakan bagian dari data mining, yang mana text mining dapat mengolah data teks yang sulit ditangani menggunakan algoritma, tidak memiliki karakter, dan tidak terstruktur^[4]. Text mining juga dapat mengatasi pada kasus klustering, information retrieval, information extraction dan klasifikasi^[5].

2.3 Text Preprocessing

Preprocessing ialah tahap pertama sebelum melakukan proses klasifikasi. Data yang belum terstruktur atau yang masih mentah akan dibersihkan terlebih dahulu dengan proses cleansing, tahap ini dilakukan agar memudahkan dalam proses klasifikasi^[6]. *Preprocessing* merupakan proses yang sangat penting ketika melakukan analisis sentimen, salah satunya adalah komentar EDOM yang sebagian besar berisi kalimat yang memiliki *noise* (emot dan tag) yang tinggi dan tidak terstruktur. *Preprocessing* juga mempunyai tiga model untuk teks atau kalimat dengan *noise* yang tinggi^[7].

2.4 Analisis Sentimen

Analisis sentimen atau yang biasa dikenal dengan *opinion mining* ialah suatu metode dari data mining yang mempunyai fungsi untuk, mengubah data, mengolah data, dan mengkaji dari sebuah entitas. Analisis sentimen juga salah satu eksplorasi baru dari NLP atau yang biasa disebut dengan *Natural Language Processing*. Analisis sentimen juga mempunyai fungsi, yaitu menemukan data dari beberapa kelompok data^[8].

2.5 Natural Language Processing

Natural Language Processing (NLP) merupakan riset ilmiah yang menekuni bagaimana komputer bisa digunakan buat menekuni serta memanipulasi bahasa alami. NLP bisa digunakan dalam membuat bahasa lebih ter standarisasi

mengganti bahasa alami jadi bahasa yang lebih terstruktur, standar, serta terintegrasi sehingga mempermudah buat memastikan penaksiran yang tepat^[3].

2.6 Algoritma Support Vector Machine

Support Vector Machine atau yang biasa dikenal dengan sebutan SVM merupakan salah satu klasifikasi dengan cara *machine learning* yang menggunakan prinsip *Structural Risk Minimization* atau yang biasa disebut dengan SRM yang mempunyai tujuan untuk menemukan *hyperplane* yang terbaik yang dapat membagi dua kelas dalam *input space*^[8].

2.7 Pembobotan TF-IDF

Dalam pembobotan Term Frequency-Inverse Document Frequency atau yang sering disebut dengan TF-IDF merupakan bobot yang terdapat pada tiap kata dan dokumen. Dalam mendapatkan nilai dari metode TF-IDF ini yakni dengan memakai persamaan dua. Ketika ingin memperoleh nilai dari "idf", ini maka bisa dicari dengan menggunakan persamaan satu yang ada dibawah ini^[8]

$$idf_i = "log"_{10} (N/"df" _i)$$

Dapat dilihat dari rumus diatas bahwa N adalah jumlah dari data atau dokumen yang sudah diproses dan "df"_i merupakan total data atau dokumen yang berisi *term*.

$$TF - IDF = tf \times idf_i$$

Pada rumus diatas, "Idf" merupakan total data atau dokumen yang isinya beberapa sebutan atau istilah. Mengenai ini tiap data ataupun dokumen dinyatakan vektor yang memiliki satu bagian pada masing- masing *term* yang sesuai serta diperoleh pada kamus bersama bobot pada masing- masing bagian. Serta *term* yang tidak kerap timbul maka bernilai 0.

$$tf - idf_{(normalisasi)} = \frac{tf - idf(asli)}{\sqrt{\sum_{k=1}^t (tf - idf(k))^2}}$$

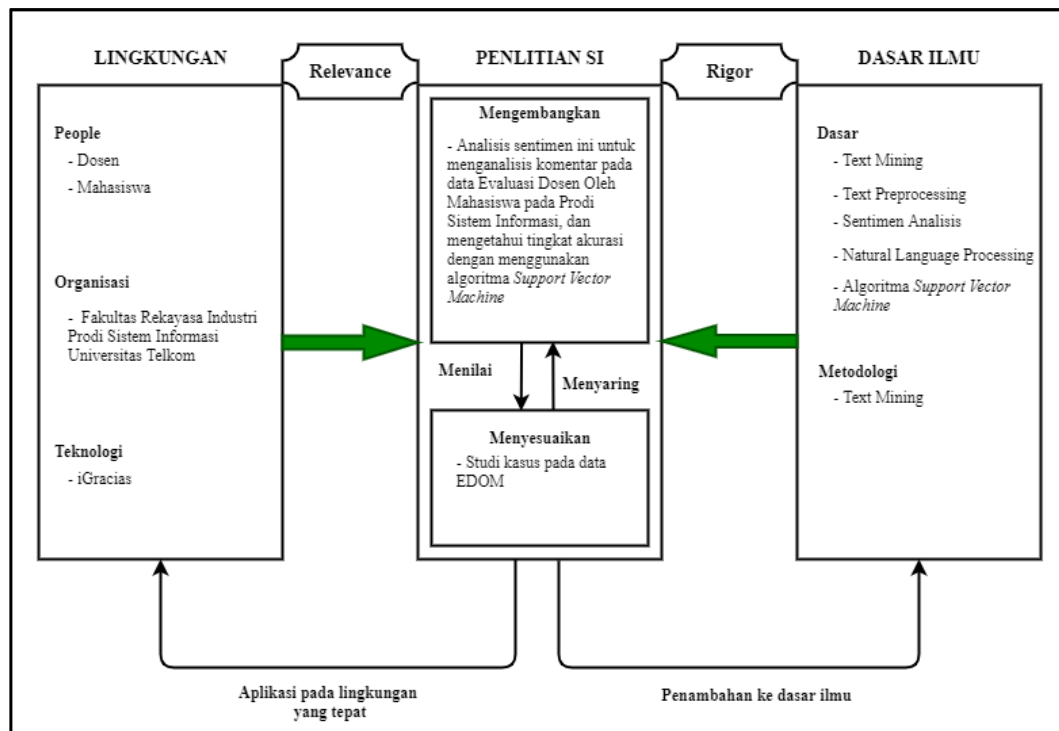
Ketika bobot sudah berhasil ditemukan maka langsung dilakukannya normalisasi dengan menggunakan bobot TF-IDF yang sudah didapatkan. Selain itu, dalam metode TF-IDF ini juga terdapat proses menghitung yang bisa mengacu ke dalam persamaan tiga^[8].

2.8 Klasifikasi Teks

Klasifikasi teks atau yang biasa disebut dengan pengkategorisasi teks ialah proses yang bekerja sendiri (otomatis) dalam memasukkan dokumen teks pada suatu bagian yang di dalamnya sudah berisi teks^[9]. Klasifikasi teks juga mempunyai tujuan, yaitu untuk mengklasifikasikan dokumen yang sudah dipilah menurut kelas nya masing-masing. Yang mana dalam satu dokumen mempunyai beberapa kelas.

3. Metodologi Penelitian

Konsep dan kerangka konseptual pada konseptual level bertujuan untuk mengidentifikasikan esensi dari tujuan riset dan ke ter hubungannya. Seperti pada gambar 1:



Gambar 1 Metodologi Konseptual [10]

Berdasarkan pada Gambar 1 merupakan penggambaran metodologi konseptual, yang dimana pada ruang lingkup lingkungan terdapat aspek *people*, organisasi dan teknologi. Dalam aspek *people* yaitu pengguna yang memakai sistem portal akademik Universitas Telkom (*igracias*), di dalamnya terdapat dosen dan mahasiswa yang menjadi pengguna *igracias*. Dalam aspek organisasi terdapat Fakultas Rekayasa Industri Program Studi Sistem Informasi Universitas Telkom. Dalam teknologi juga terdapat *igracias* yang dimana *igracias* itu sendiri adalah portal akademik yang biasa digunakan untuk kebutuhan akademik para mahasiswa dan dosen. Penelitian ini membahas tentang penilaian kinerja dosen yang dilihat dari sisi komentar pada data EDOM dan mengetahui tingkat akurasi dengan menggunakan algoritma *Support Vector Machine*. Menggunakan dasar ilmu *text mining*, *text preprocessing*, analisis sentimen, *natural language processing*, dan algoritma *support vector machine*, dengan menggunakan metodologi *text mining*.

4. Dataset

Data yang digunakan pada penelitian ini adalah data komentar yang ada pada komentar yang tersedia dalam data EDOM Telkom University. Data ini diambil dengan secara manual, kemudian dilakukan proses pre-processing untuk mengambil data komentar dan caption kiriman, kemudian disimpan pada file dengan format excel. Data yang telah berhasil dikumpulkan kemudian dilakukan penyesuaian dan pembersihan dengan menghilangkan simbol-simbol dan tanda baca pada dataset tersebut, hasil dapat dilihat dalam Tabel 1.

Tabel 1 Tabel Dataset

	Sebelum <i>Pre-processing</i>	Setelah <i>Pre-processing</i>
Komentar	3.090	2.465

Kemudian setelah data diurutkan, para peneliti memberi label pada sentimen secara manual. Label yang diberikan pada data terdiri dari positif, negatif, dan netral.

Tabel 2 Komentar yang diberi Label

No	Teks	Label	ID
1	semangat	1	ALB
2	puas	1	ALB
3	telat buk labil jadwal kelas ganti	2	ALV
4	dosen lambat menit jam kuliah jam	2	BMZ

	dosen jam dekat depan moga waktu		
5	moga	3	BMZ
6	terimakasih	3	BSR

Langkah selanjutnya adalah melakukan pra-pemrosesan teks dimulai dengan pelipatan huruf, yang mengubah semua huruf menjadi huruf kecil^[8].

Tabel 3 Data Komentar yang sudah di *Case Folding*

Sebelum <i>Case Folding</i>	Setelah <i>Case Folding</i>
Sudah sangat baik dalam menyampaikan materi dan sangat sesuai dengan kurikulum yang ada, tingkatkan terus dan sukses selalu pak	sudah sangat baik dalam menyampaikan materi dan sangat sesuai dengan kurikulum yang ada tingkatkan terus dan sukses selalu pak
Semangat	semangat
Mata kuliah ini terlalu rumit, ditambah dengan UTS nya yang membuat trauma membekas. Sampai saat ini, masih ada meme "Biar balance" di kelas saya.	mata kuliah ini terlalu rumit ditambah dengan uts nya yang membuat trauma membekas sampai saat ini masih ada meme biar balance di kelas saya

Tahap tokenization adalah merupakan proses pemotongan pada kalimat menjadi beberapa kata, yang biasa disebut dengan *token*^[11].

Tabel 4 Data Komentar yang sudah di *Tokenization*

Sebelum <i>Tokenization</i>	Setelah <i>Tokenization</i>
sudah sangat baik dalam menyampaikan materi dan sangat sesuai dengan kurikulum yang ada tingkatkan terus dan sukses selalu pak	['sudah', 'sangat', 'baik', 'dalam', 'menyampaikan', 'materi', 'dan', 'sangat', 'sesuai', 'dengan', 'kurikulum', 'yang', 'ada', 'tingkatkan', 'terus', 'dan', 'sukses', 'selalu', 'pak'],

Hentikan penghapusan kata-kata, pada tahap ini peneliti menghilangkan kata-kata yang tidak memiliki makna untuk digunakan ketika mengklasifikasikan sentimen, misalnya kata-kata: yang, untuk, pada, ke, para, ya, dan lain-lain^[12].

Tabel 5 Data Komentar yang sudah di *Stopword Removal*

Sebelum <i>Stopword</i>	Sesudah <i>Stopword</i>
sudah	materi
sangat	
baik	
dalam	
menyampaikan	
materi	
dan	
sangat	
sesuai	sesuai
dengan	
kurikulum	kurikulum
yang	
ada	
tingkatkan	tingkatkan
terus	
dan	
sukses	sukses
selalu	
pak	

Proses stemming adalah tahap menghilangkan afiks untuk setiap kata yang tercantum dalam setiap kalimat. Agar lebih mudah dipahami, stemming adalah proses mengubah kata-kata imbuhan menjadi kata-kata dasar. Tujuannya adalah untuk menghilangkan makna berlebih di setiap kata, sehingga ada standar untuk setiap kata^[13].

Tabel 6 Data Komentar yang sudah di *Stemming*

Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
materi	materi
sesuai	sesuai
kurikulum	kurikulum
tingkatkan	tingkat
sukses	sukses

5. Proses Klasifikasi

Proses klasifikasi menggunakan algoritma SVM dimulai dengan distribusi data pelatihan dan pengujian data untuk pengujian. Data yang digunakan adalah komentar dari data EDOM Program Studi Sistem Informasi Universitas Telkom. Dari total tiga ribu sembilan puluh komentar yang telah diberi label sentimen secara manual dan otomatis menggunakan python. Pada saat proses pelabelan, peneliti melakukan secara manual dan dibantu oleh beberapa orang mahasiswa dengan menandakan komentar ke dalam kategori positif, negatif, netral. Sehingga hasil yang didapatkan dari masing-masing label dapat dilihat dalam Tabel 7.

Tabel 7 Tabel Hasil Komentar per Kategori

	Positif	Negatif	Netral	Jumlah
Komentar	1.440	84	941	2.465

0.7484787018255578					
	precision	recall	f1-score	support	
1	0.75	0.88	0.81	288	
2	1.00	0.21	0.35	14	
3	0.74	0.59	0.65	191	
accuracy			0.75	493	
macro avg	0.83	0.56	0.61	493	
weighted avg	0.75	0.75	0.74	493	
[[254 0 34]					
[5 3 6]					
[79 0 112]]					

Gambar 2 Hasil Akurasi

Berdasarkan Gambar 2 menjelaskan hasil akurasi yang dilakukan dengan menggunakan tools python dengan menggunakan rasio 80:20. Hasil akurasi yang didapatkan dengan menggunakan rasio sebesar 75%, Precision 75%, Recall 75%, dan menghasilkan nilai F-1 Score 74%.

6. Kesimpulan

Data yang digunakan dalam analisis sentimen ini adalah data komentar EDOM Prodi Sistem Informasi Universitas Telkom, dengan menggunakan data komentar semester ganjil 2019/2020. Terdapat data komentar positif 1.440 komentar, komentar negatif 84 dan komentar netral 395 dengan total data 2.465. Dalam penelitian ini, setelah dilakukannya klasifikasi dengan menggunakan algoritma SVM dengan menggunakan metode TF-IDF dan algoritma SVM, maka dilakukan pengujian terhadap hasil klasifikasi dengan menggunakan perbandingan 80% data training dan 20% data testing, hasil yang didapat adalah akurasi sebesar 75%, precision sebesar 75%, recall sebesar 75% dan f-1 score sebesar 74%.

Daftar Pustaka:

- [1] A. S. SUTEDJO and A. P. MANGKUNEGARA, "Pengaruh Kompetensi dan Motivasi Kerja terhadap Kinerja Karyawan di PT. Inti Kebun Sejahtera," *BISMA (Bisnis dan Manajemen)*, vol. 5, no. 2, p. 120, 2018.
- [2] Y. R. P. Trisno Nugroho, "No Title," *Kepuasan Mhs. Telkom Univ. dalam Pencarian Inf. pada situs*, 2018.
- [3] and J. S. Feldman, Ronen, *No Title*. 2007.
- [4] I. H. Witten, "Text mining," *Pract. Handb. Internet Comput.*, pp. 14-1-14-22, 2004.
- [5] J. Berry, M.W. & Kogan, "Text Mining: Applications and Theory," *Text Min. Appl. Theory*.
- [6] G. Ngurah, M. Nata, and P. P. Yudiastra, "Preprocessing Text Mining Pada Email Box Berbahasa Indonesia," *Konf. Nas. Sist. Inform. 2017*, pp. 479-483, 2017.
- [7] A. Clark, "Pre-processing very noisy text ISSCO / TIM University of Geneva Switzerland," *Proc. Work. Shallow Process. Large Corpora*, pp. 1-10, 2003.
- [8] V. I. Santoso, G. Virginia, and Y. Lukito, "Penerapan Sentiment Analysis Pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine," *J. Transform.*, vol. 14, no. 2, p. 72, 2017.
- [9] X. F. Zhang, H. Y. Huang, and K. L. Zhang, "KNN text categorization algorithm based on semantic centre," *Proc. - 2009 Int. Conf. Inf. Technol. Comput. Sci. ITCS 2009*, vol. 1, pp. 249-252, 2009.
- [10] S. Hevner, Alan; Chatterjee, *Design Research in Information System : Theory and Practice*. Springer, 2010.
- [11] I. N. Fakhri and R. F. Umbara, "Analisis Sentimen pada Kuisisioner Kepuasan Terhadap Layanan dan Fasilitas Kampus Universitas Dengan Menggunakan Klasifikasi Support Vector Machine (SVM)," vol. 6, no. 2, pp. 8682-8691, 2019.
- [12] A. F. Sulaeman, A. A. Supianto, and F. A. Bachtiar, "Analisis Sentimen Opini Mahasiswa Terhadap Saran Evaluasi Kinerja Dosen Menggunakan TF-IDF dan Support Vector Machine .," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 6, pp. 5647-5655, 2019.

- [13] D. Wahyudi, T. Susyanto, and D. Nugroho, "Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia," *J. Ilm. SINUS*, vol. 15, no. 2, pp. 49–56, 2017.

