

Nomor Daftar:034/S/BD/R/16/VIII/2023

**Penerapan Algoritma *Support Vector Machine* untuk Analisis Sentimen
Review Pengguna pada Twitter Ruang Guru**

SKRIPSI

**Diajukan untuk Persyaratan Penelitian dan Penulisan Skripsi sebagai Bagian
dari Syarat Gelar Sarjana Program Studi Bisnis Digital**



Oleh

Bagus Haryanto

NIM 190441

**PROGRAM STUDI BISNIS DIGITAL
KAMPUS TASIKMALAYA
UNIVERSITAS PENDIDIKAN INDONESIA
2023**

**Penerapan Algoritma *Support Vector Machine* untuk Analisis Sentimen
Review Pengguna pada Twitter Ruang Guru**

**Oleh
Bagus Haryanto**

**Diajukan untuk memenuhi sebagian dari Persyaratan Memperoleh Gelar
Sarjana Binis Program Studi Bisnis Digital**

**© Bagus Haryanto
Universitas Pendidikan Indonesia
Agustus 2023**

Hak Cipta dilindungi undang-undang.

**Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
Dengan dicetak ulang, difoto copy, atau cara lainya tanpa izin dari penulis**

BAGUS HARYANTO

LEMBAR PENGESAHAN SKRIPSI

BAGUSHARYANTO

Penerapan Algoritma *Support Vector Machine* untuk Analisis Sentimen *Review*
Pengguna pada Twitter Ruang Guru

Disetujui dan disahkan oleh dosen pembimbing:

Pembimbing I



Odiog Dardiana, S.Kom., M.Kom

NIP. 920200419860711101

Pembimbing II



Rangga Celar Guntarn, S.Kom, M.Kom

NIP. 920200819880616101

Mengetahui,

Ketua Program Studi S1 Bisnis Digital UPI

Kampus Tasihmalaya



Syti Sarah Maesaroh, S.P., M.M

NIP. 920190219900625201

SURAT PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan Algoritma *Support Vector Machine* Untuk Analisis Sentimen *Review* Pengguna Pada *Twitter* Ruang Guru” ini beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung resiko atau sanksi apabila dikemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Tasikmalaya, 9 Agustus 2023

Yang membuat pernyataan,



Bagus Haryanto

NIM 1904441

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian ini. Penulis juga mengucapkan rasa hormat dan terima kasih yang setinggi-tingginya kepada semua pihak yang telah memberikan dukungan, bimbingan, serta inspirasi dalam proses penyusunan skripsi ini. Penelitian ini merupakan hasil dari upaya kolaboratif dan kerja keras yang melibatkan berbagai pihak. Penulis ingin mengucapkan terima kasih kepada Bapak Oding Herdiana, S.Kom., M.Kom dan Rangga Gelar Guntara, S.Kom., M.Kom selaku dosen pembimbing penulis yang telah memberikan arahan, pengarahan, serta dorongan yang luar biasa. Bimbingan dan masukan yang diberikan telah sangat berharga dalam merumuskan fokus penelitian dan mengatasi berbagai kendala yang penulis hadapi. Tidak lupa pula penulis ucapkan terima kasih kepada seluruh dosen di Program Studi Bisnis Digital yang telah memberikan ilmu dan pengetahuan yang sangat berarti selama masa studi penulis di universitas ini. Penulis juga ingin berterima kasih kepada seluruh rekan-rekan di Program Studi Bisnis Digital yang telah memberikan dukungan moril dan saling berbagi pengalaman dalam perjalanan penelitian ini. Penulis menghargai setiap kritik, saran, dan masukan yang dapat berkontribusi untuk kemajuan penelitian ini. Akhir kata, semoga hasil dari penelitian ini dapat memberikan manfaat dan kontribusi positif bagi pengembangan ilmu pengetahuan, teknologi, dan sumber inspirasi bagi penelitian selanjutnya.

Tasikmalaya, 9 Agustus 2023

Penulis

UCAPAN TERIMA KASIH

Penulis ingin mengucapkan rasa terima kasih yang tulus dan mendalam kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian skripsi ini, berjudul "Penerapan Algoritma *Support Vector Machine* untuk Analisis Sentimen *Review* Pengguna pada Twitter Ruang Guru". Keberhasilan penelitian ini tidak lepas dari kerja sama dan bantuan dari berbagai individu dan institusi, yang dengan rendah hati kami sampaikan dalam ucapan terima kasih ini:

1. Dengan limpahan rahmat dan karunia dari Allah SWT, penulis dapat menyelesaikan skripsi ini dengan penuh rasa syukur.
2. Kedua orang tua tercinta, Bapak Kaerudin dan Ibu Sunarsih, terima kasih atas seluruh dukungan, cinta dan doa yang tak henti-hentinya mengalir selama perjalanan studi penulis. Kalian adalah sumber inspirasi utama penulis.
3. Bapak Oding Herdiana, S.Kom., M.Kom, terima kasih atas arahan, bimbingan, serta kesabaran Bapak selama proses penyusunan skripsi ini. Ilmu dan pengalaman yang Bapak berikan telah membantu penulis mengembangkan pengetahuan dan keterampilan dalam penelitian ini.
4. Bapak Rangga Gelar Guntara, S.Kom., M.Kom, terima kasih atas bimbingan Bapak yang sangat berharga dan dukungan yang luar biasa kepada penulis selama melakukan penelitian ini.
5. Ibu Syti Sarah, S.P., M.M, sebagai pembimbing akademik penulis. Terima kasih atas segala ilmu, bimbingan, kritik, saran dan motivasi selama masa perkuliahan.
6. Seluruh tenaga pengajar dan staff Program Studi Bisnis Digital terimakasih atas dedikasi dan kontribusi yang luar biasa yang telah diberikan.
7. Rekan-rekan terbaik, Howard Lauson, Kris Adam Gunanta, Novian Rachman Iskandar, Moch Gangan Sidiq yang senantiasa memberikan dukungan yang luar biasa.

8. Terima kasih kepada semua teman seperjuangan di Program Studi Bisnis Digital UPI dalam menyusun skripsi. Bersama-sama, kita melewati tantangan dan berbagi dukungan. Sukses untuk kita semua.
9. Semua pihak yang turut serta, penulis mengucapkan terima kasih dan mengapresiasi semua pihak yang terlibat, baik secara langsung maupun tidak langsung, dalam dukungan dan dorongan positif untuk penyelesaian skripsi ini.

Penerapan Algoritma *Support Vector Machine* untuk Analisis Sentimen Review Pengguna pada Twitter Ruang Guru

ABSTRAK

Oleh

Bagus Haryanto

Perkembangan teknologi dan informasi yang pesat dalam dunia pendidikan telah menciptakan transformasi baru yang dikenal sebagai *e-learning*. Ruang Guru adalah perusahaan teknologi yang fokus pada layanan pendidikan berbasis teknologi informasi, pada tahun 2023 jumlah pengguna Ruang Guru mencapai 38 juta pengguna aktif. Dengan meningkatnya jumlah pengguna secara signifikan, *review* mengenai layanan Ruang Guru juga mengalami peningkatan, mencakup *review* yang menyatakan kepuasan dan ketidakpuasan. Analisis sentimen menjadi instrumen penting untuk memahami secara umum apakah *review* tersebut memiliki sentimen positif atau negatif. Twitter menjadi *platform* media sosial yang banyak digunakan oleh pengguna untuk berbagi opini tentang layanan Ruang Guru. Penelitian ini bertujuan untuk melakukan klasifikasi terhadap *review* pengguna Ruang Guru di Twitter menggunakan algoritma *Support Vector Machine* (SVM). Selama jangka waktu satu bulan (1-31 Maret 2023), data *review* pengguna mengenai Ruang Guru yang berasal dari platform Twitter telah dikumpulkan. Penjaringan data ini dilakukan sejalan dengan pelaksanaan Seleksi Nasional Bersama Tes (SNBT). Dari hasil *scrapping*, total terhimpun 406 entitas data. Data-data tersebut selanjutnya mengalami proses pelabelan, yang mengategorikan mereka ke dalam dua kelas berbeda: sentimen positif dan sentimen negatif. Hasil pengujian melalui perhitungan *TF-IDF* dan *confusion matrix* menunjukkan hasil berikut: *accuracy* mencapai 78,5%, *precision* sebesar 60%, *recall* sebesar 30%, dan nilai *f1_score* sebesar 40%. Hasil dari penelitian ini menyimpulkan bahwa algoritma *Support Vector Machine* (SVM) berhasil diimplementasikan dan dapat menjadi bahan evaluasi Ruang Guru di masa SNBT yang akan datang.

Kata kunci: analisis sentiment, *support vector machine*, twitter, ruang guru

Implementation of Support Vector Machine Algorithm for Sentiment Analysis of User *Reviews* on Twitter Ruang Guru

ABSTRACT

by

Bagus Haryanto

The rapid development of technology and information in education has created a new transformation known as e-learning. Ruang Guru is a technology company that focuses on information technology-based education services, by 2023 the number of Ruang Guru users will reach 38 million active users. With the significant increase in the number of users, *reviews* about Ruang Guru's services have also increased, including *reviews* expressing satisfaction and dissatisfaction. Sentiment analysis becomes an important instrument to generally understand whether the *review* has a positive or negative sentiment. Twitter is a social media platform that is widely used by users to share opinions about Ruang Guru services. This research aims to classify Ruang Guru user *reviews* on Twitter using the Support Vector Machine (SVM) algorithm. Over a period of one month (March 1-31, 2023), user *review* data about Ruang Guru from the Twitter platform was collected. This data scraping was conducted in line with the implementation of the National Joint Test Selection (SNBT). From the scrapping results, a total of 406 data entities were collected. The data were then subjected to a labeling process, which categorized them into two different classes: positive sentiment and negative sentiment. Test results through TF-IDF and confusion matrix calculations show the following results: 78.5% accuracy, 60% precision, 30% recall, and 40% f1_score. The results of this study conclude that the Support Vector Machine (SVM) algorithm was successfully implemented and can be used as an evaluation material for the Ruang Guru in the future SNBT period..

Keywords: sentiment analysis, support vector machine, twitter, ruang guru

DAFTAR ISI

SURAT PERNYATAAN KEASLIAN SKRIPSI.....	i
KATA PENGANTAR	ii
UCAPAN TERIMA KASIH.....	iii
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang Penelitian.....	1
1.2. Rumusan Masalah Penelitian	3
1.3. Batasan Masalah Penelitian.....	4
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
1.6. Struktur Organisasi Skripsi.....	5
BAB II KAJIAN PUSTAKA	7
2.1. Kajian Pustaka	7
2.1.1. <i>Data Mining</i>	7
2.1.2. Klasifikasi	7
2.1.3. Ruang Guru	8
2.1.4. <i>Web Scrapper</i>	9
2.1.5. Analisis Sentimen.....	9
2.1.6. <i>Machine Learning</i>	10
2.1.7. <i>Cross Industry Standard Process for Data Mining</i>	11
2.1.8. <i>Text PreProcessing</i>	13
2.1.9. <i>TF-IDF</i>	14

2.1.10.	Algoritma <i>Support Vector Machine</i> (SVM).....	15
2.1.11.	<i>Confusion matrix</i>	16
2.1.12.	<i>Python</i>	17
2.1.13.	Google Colab	19
2.1.14.	<i>Word cloud</i>	20
2.1.15.	Penelitian Terdahulu	21
BAB III METODE PENELITIAN.....		21
3.1.	Desain Penelitian	25
3.2.	Pengumpulan Data.....	26
3.2.1.	Sumber Data	26
3.2.2.	Teknik Pengumpulan Data.....	27
3.3.	Penerapan Metode CRISP-DM	27
3.3.1.	<i>Business Understanding</i>	28
3.3.2.	<i>Data Understanding</i>	28
3.3.3.	<i>Data Preparation</i>	28
3.3.4.	<i>Modelling</i>	29
3.3.6.	<i>Deployment</i>	31
BAB IV HASIL DAN PEMBAHASAN		32
4.1.	Gambaran Umum Perusahaan	32
4.1.1.	Profil Perusahaan	32
4.1.2.	Visi dan Misi.....	32
4.2.	Deskripsi Data <i>Scrapping</i>	33
4.3.	Analisis Sentimen Menggunakan Metode CRISP-DM	35
4.3.1.	<i>Business Understanding</i>	35
4.3.2.	<i>Data Understanding</i>	35
4.3.3.	<i>Data Preparation</i>	37
4.3.4.	Modeling.....	39
4.3.5.	<i>Evaluation</i>	42
4.3.6.	<i>Deployment</i>	44
4.4.	Pembahasan	46
4.4.1.	Hasil Penerapan Algoritma <i>Support Vector Machine</i>	46
4.4.2.	Perbandingan Evaluasi SVM dengan Algoritma lain	47
4.4.3.	Perbandingan Hasil Penelitian dengan Hasil Penelitian Terdahulu.....	47
4.4.4.	terpretasi Visualisasi <i>Word Cloud</i> Sentimen Pengguna Ruang Guru	48

BAB V SIMPULAN, IMPLIKASI DAN REKOMENDASI	50
5.1. Simpulan.....	50
5.2. Implikasi.....	51
5.3. Rekomendasi	52
DAFTAR PUSTAKA	53
LAMPIRAN	56
RIWAYAT HIDUP	65

DAFTAR TABEL

Tabel 2. 1 Struktur <i>confusion matrix</i>	16
Tabel 2. 2 Penelitian terdahulu.....	21
Tabel 4. 1 Keterangan variabel data <i>scrapping</i>	34
Tabel 4. 2 Variabel <i>dataset</i>	35
Tabel 4. 3 Contoh hasil <i>labeling data</i>	36
Tabel 4. 4 Hasil <i>case folding</i>	37
Tabel 4. 5 Hasil <i>tokenizing</i>	38
Tabel 4. 6 Hasil <i>filtering</i>	38
Tabel 4. 7 Hasil <i>stemming</i>	39
Tabel 4. 8 Hasil akhir tahap <i>data preparation</i>	39
Tabel 4. 9 Hasil TF-IDF.....	41
Tabel 4. 10 <i>Splitting</i> data 80:20	41
Tabel 4. 11 <i>Confusion matrix SVM</i>	42
Tabel 4. 12 Hasil evaluasi SVM.....	43
Tabel 4. 13 <i>Confusion matrix Naive Bayes</i>	43
Tabel 4. 14 Hasil evaluasi <i>Naive Bayes</i>	43
Tabel 4. 15 <i>Confusion matrix Decision Tree</i>	44
Tabel 4. 16 Hasil evaluasi <i>Decision Tree</i>	44
Tabel 4. 17 Perbandingan evaluasi SVM dengan model lain.....	47
Tabel 4. 18 Perbandingan hasil penelitian dengan penelitian terdahulu	48

DAFTAR GAMBAR

Gambar 2. 1 <i>Cycle</i> CRISP-DM.....	12
Gambar 3. 1 Desain penelitian	25
Gambar 3. 2 Teknik pengumpulan data	27
Gambar 3. 3 Alur kerja <i>data preparation</i>	29
Gambar 4. 1 Hasil <i>scrapping data</i>	34
Gambar 4. 2 Perbandingan data sentimen positif dan negatif.....	36
Gambar 4. 3 <i>Syntax count</i> TF-IDF	40
Gambar 4. 4 <i>Output count</i> TF-IDF	40
Gambar 4. 5 <i>Input syntax splitting data</i>	41
Gambar 4. 6 Sentimen positif <i>wordcloud</i>	45
Gambar 4. 7 Sentimen negatif <i>wordcloud</i>	46

DAFTAR LAMPIRAN

Lampiran 1 <i>Source code scapping data twitter</i>	56
Lampiran 2 <i>Source code EDA sentiment snalysis</i>	56
Lampiran 3 Hasil <i>scrapping data</i>	60
Lampiran 4 Hasil <i>labeling data</i>	61
Lampiran 5 Hasil <i>data preprocessing</i>	62
Lampiran 6 Hasil <i>modeling dan evaluation</i>	63