

Tesis

**Analisis Sentimen Layanan *E-commerce*
pada Sosial Media X Menggunakan
Metode *SVM Linier* dan *Non Linier***

**Fiki Andrianto
2008048021**



**Program Studi S2 Informatika
Fakultas Teknologi Industri
Universitas Ahmad Dahlan
Yogyakarta
2024**

Tesis

Analisis Sentimen Layanan *E-commerce* pada Sosial Media X Menggunakan Metode *SVM Linier dan Non Linier*

Fiki Andrianto
2008048021

Dipertahankan di depan Dewan Penguji
Tanggal 22 Maret 2024

Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D.
Ketua Penguji



Prof. Dr. Ir. Imam Riadi, M.Kom.
Penguji 1



Rusydi Umar, S.T., M.T., Ph.D.
Penguji 2



Dr. Murinto, S.Si., M.Kom.
Penguji 3



Mengetahui,

Prof. Dr. Ir. Siti Jamilatun, M.T.
Dekan Fakultas Teknologi Industri

Pernyataan Tidak Plagiat

Saya yang bertanda tangan di bawah ini:

Nama : Fiki Andrianto
NIM : 2008048021
Email : Fiki2008048021@webmail.uad.ac.id
Program Studi : S2 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Analisis Sentimen Layanan *E-commerce* pada Sosial media X
Menggunakan Metode *SVM Linier dan Non Linier*

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Ahmad Dahlan maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Ahmad Dahlan.



Surabaya, 22 Maret 2024

Menyatakan

Fiki Andrianto

Pernyataan Persetujuan Akses

Saya yang bertanda tangan di bawah ini:

Nama : Fiki Andrianto
NIM : 2008048021
Email : Fiki2008048021@webmail.uad.ac.id
Program Studi : S2 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Analisis Sentimen Layanan E-commerce pada Sosial media X
Menggunakan Metode SVM Linier dan Non Linier

Dengan ini saya menyerahkan hak Sepenuhnya kepada Pusat Sumber Belajar Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tesis elektronik sebagai berikut (beri tanda pada kotak):

Saya mengizinkan karya tersebut diunggah ke dalam aplikasi Repository Perpustakaan Universitas Ahmad Dahlan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Yogyakarta, 22 Maret 2024

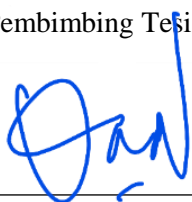
Yang Menyatakan



Fiki Andrianto

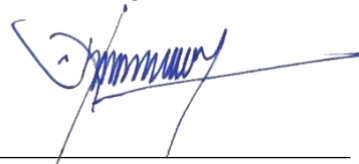
Mengetahui,

Pembimbing Tesis 1



Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D.

Pembimbing Tesis 2



Prof. Ir. Dr. Imam Riadi, M.Kom.

Motto dan Persembahan

“Peluh adalah cerita, kesuksesan adalah karyanya.” -Fiki Andrianto

“Jika kita berani bermimpi, maka kita harus berani bekerja keras untuk mewujudkannya.” - Fiki Andrianto

Tesis ini kupersembahkan untuk:

1. Alm. Ayah dan Ibu, Terimakasih telah menjadi orang tua yang selalu memberikan pengajaran yang baik dimasa hidup.
2. Suami/istri/anak yang memberikan dukungan dan semangat selama proses kuliah dan penulisan tesis
3. Keluarga, Sahabat, MTI10, MTI11, MTI12, MTI13, MTI14 yang selalu menjadi teman diskusi yang baik penghibur selama proses penulisan Tesis.
4. Para pembaca naskah tesis ini semoga menjadi inspirasi dalam pengembangan ilmu pengetahuan.

Kata Pengantar

Bismillahirrahmaanirrahiim,

Assalaamu'alaikum warahmatullaahi wabarakaatuh

Segala puji bagi Allah swt yang telah memberikan rahmat, hidayah, dan inayah kepada hamba-Nya sehingga penulis dapat menyelesaikan Tesis ini.

Penulis menyadari bahwa keberhasilan dalam menyelesaikan laporan ini berkat dorongan dan bimbingan berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis menghaturkan terimakasih dan penghargaan kepada Bapak Prof. Dr. Muchlas, M.T. Rektor Universitas Ahmad Dahlan, Ibu Prof. Dr. Ir. Siti Jamilatun, M.T. Dekan Fakultas Teknologi Industri, Bapak Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D. selaku Kaprodi S2 Informatika, Bapak Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D. selaku Pembimbing Tesis 1 dan Bapak Prof. Dr. Ir. Imam Riadi, M.Kom. selaku Pembimbing Tesis 2 yang rela diganggu setiap saat selama proses penulisan tesis. Terimakasih juga penulis haturkan kepada seluruh dosen S2 Informatika yang telah memberikan banyak ilmu dan wawasan.

Selanjutnya penulis juga menyampaikan terima kasih kepada rekan-rekan angkatan 10 di S2IF UAD dan berbagai pihak yang tidak dapat disebutkan satu persatu.

Terakhir, penulis sangat mengharapkan saran dan kritik yang membangun, karena Tesis ini sangat jauh dari sempurna. Semoga Allah meridhoi langkah kita. Amin.

Wassalaamu'alaikum warahmatullaahi wabarakaatuh

Yogyakarta, 22 Maret 2024

Fiki Andrianto

Daftar Isi

Pernyataan Tidak Plagiat	iii
Pernyataan Persetujuan Akses	iv
Motto dan Persembahan	v
Kata Pengantar	vi
Daftar Isi	vii
Daftar Gambar	ix
Daftar Tabel	x
Abstrak	xi
Abstract	xii
Bab 1 Pendahuluan	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	2
1.3 Batasan Masalah	2
1.4 Rumusan Masalah.....	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian.....	4
Bab 2 Tinjauan Pustaka	5
2.1 Kajian Penelitian Terdahulu	5
2.2 Landasan Teori.....	16
2.2.1 Sosial media X	16
2.2.2 <i>Data Collection</i>	17
2.2.3 Analisis Sentimen.....	17
2.2.4 <i>Preprocessing</i>	18
2.2.5 <i>Term Weighting (TF-IDF)</i>	19
2.2.6 <i>Support Vector Machine</i>	19
2.2.6.1 SVM Linier	20
2.2.6.2 SVM Non Linier	21
2.2.7 <i>Spliting Data</i>	22
2.2.8 <i>Evaluation</i>	24
Bab 3 Metodologi	26
3.1 Objek Penelitian	26
3.2 Alat dan Bahan	26

3.3 Perancangan Sistem.....	27
3.3.1 <i>Data Collection</i>	28
3.3.2 Analisis Sentimen.....	29
3.3.3 <i>Preprosesing</i>	30
3.3.4 <i>TF-IDF</i>	31
3.3.5 Algoritma <i>Support Vector Machine</i>	31
3.3.6 Pengujian	32
3.4 <i>Evaluation</i>	34
Bab 4 Hasil dan Pembahasan.....	35
4.1 Pengambilan Data.....	35
4.2 <i>Labelling</i>	36
4.3 <i>Preprosesing</i>	37
4.3.1 <i>Case Folding</i>	37
4.3.2 <i>Tokenizing</i>	38
4.3.3 <i>Stopword Removal</i>	40
4.3.4 <i>Stemming</i>	41
4.4 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	43
4.5 <i>Split Data</i>	45
4.6 <i>Support Vector Machine</i>	46
4.6.1 SVM Kernel Linier.....	46
4.6.2 <i>SVM Kernel Non Linier</i>	48
Bab 5 Penutup	56
5.1 Kesimpulan.....	56
5.2 Saran	57
Daftar Acuan	58
Lampiran 1 Daftar Riwayat Hidup.....	62
Lampiran 2 <i>Source Code</i>	63

Daftar Gambar

Gambar 2.1 <i>Support Vector Machine</i>	20
Gambar 3.1 Kerangka Penelitian	27
Gambar 3. 2 Alur <i>Scraping Sosial media X</i>	29
Gambar 3. 3 Analisis Sentimen.....	30
Gambar 3. 4 <i>Preprosesing</i>	30
Gambar 4. 1 Dataset Penelitian Tweet.....	35
Gambar 4. 2 Total <i>Labelling Dataset</i>	36
Gambar 4. 3 Proses <i>Case Folding</i>	37
Gambar 4. 4 <i>Coding Case Folding</i>	38
Gambar 4. 5 Proses <i>Tokenization</i>	39
Gambar 4. 6 <i>Coding Tokenization</i>	39
Gambar 4. 7 Proses <i>Stopword Removal</i>	40
Gambar 4. 8 <i>Coding Lists Stopword Removal</i>	41
Gambar 4. 9 Proses Pengolahan <i>Stemmed</i>	42
Gambar 4. 10 <i>Coding Stemmed</i>	42
Gambar 4. 11 Alur Split penelitian	45
Gambar 4. 12 Hasil Evaluasi Matrix SVM Kernel Linier	48
Gambar 4. 13 Hasil Evaluasi Matrix SVM Kernel Polynomial.....	51
Gambar 4. 14 Hasil Evaluasi Matrix SVM Kernel Rbf	52
Gambar 4. 15 Hasil Evaluasi Matrix SVM Kernel Sigmoid.....	53

Daftar Tabel

Tabel 2. 1 Rangkuman Penelitian Terdahulu	11
Tabel 2. 2 Perbandingan Penelitian Terdahulu	15
Tabel 3. 1 Alat dan Bahan Penelitian	26
Tabel 4. 1 Perhitungan TF-IDF	44
Tabel 4. 2 Penggunaan Dataset	46
Tabel 4. 3 Hasil classification <i>SVM Kernel Linier</i>	46
Tabel 4. 4 Hasil classification <i>SVM Kernel Non Linier</i>	49
Tabel 4. 5 Hasil Komparasi <i>SVM Kernel Linier</i> dan <i>Non Linier</i>	53
Tabel 4. 6 Perbandingan <i>Splitting</i> dan <i>Accuration</i>	54

Andrianto, F. (2023). **Analisis Sentimen Layanan E-commerce pada Sosial Media X Menggunakan Metode SVM Linier dan Non Linier**. Tesis, Magister Informatika, Universitas Ahmad Dahlan, Yogyakarta.

Abstrak

Dalam era perkembangan teknologi digital, sektor perdagangan *e-commerce* telah menjadi pusat bisnis dan belanja *online*. Sosial media X, telah menjadi platform bagi pelanggan untuk menyampaikan pengalaman mereka terkait layanan *e-commerce*. Sentimen yang terkandung dalam *tweet* pelanggan dapat memberikan wawasan penting tentang tingkat kepuasan atau ketidakpuasan mereka terhadap sebuah layanan. Penggunaan Algoritma *Machine Learning* dalam analisis sentimen digunakan untuk mendeteksi kepuasan pelanggan apakah termasuk dalam komentar positif atau negatif. Tujuan penelitian ini adalah melakukan analisis sentimen terhadap kepuasan pelanggan melalui pengguna sosial media X menggunakan algoritma *Machine Learning*, guna identifikasi algoritma yang tepat dalam mengenali *tweet* pengunjung dalam kategori sentimen positif atau negatif terhadap tiga *e-commerce online* besar Indonesia yaitu Tokopedia, Shopee, dan Lazada.

Besarnya data yang dihasilkan oleh platform media sosial membuat analisis klasifikasi yang tepat diperlukan. Tahapan klasifikasi meliputi *preprocessing*, *feature extraction and selection*, *split data klasifikasi* dan evaluasi. Model SVM linier dan non-linier dipilih untuk penelitian ini karena SVM linier merupakan salah satu pilihan yang populer dalam dunia machine learning, terutama untuk menangani dataset yang besar dan fitur. Model SVM linier dapat memberikan hasil yang baik dalam kasus-kasus dengan hubungan linier antara fitur dan label kelas. Sedangkan SVM non-linier memiliki fleksibilitas untuk menangani hubungan yang kompleks dan tidak linier antara fitur dan label kelas. Oleh karena itu, penggunaan kedua model ini dapat memberikan pemahaman yang holistik tentang pola data yang kompleks dalam analisis sentimen pelanggan terhadap layanan *e-commerce*.

Hasil evaluasi menunjukkan bahwa model SVM dengan berbagai jenis kernel memberikan performa yang baik dalam memprediksi data. Kernel Linear, dengan akurasi tertinggi 94% pada rasio split 80%, menonjol dalam keseimbangan antara data pelatihan dan pengujian. Kernel Polynomial, meskipun sedikit lebih rendah dengan akurasi 93% pada rasio split 75%, masih menunjukkan kinerja yang solid. Kernel RBF memberikan hasil terbaik pada rasio split 90%, menunjukkan ketangguhan dalam menangani lebih banyak data pelatihan daripada pengujian. Terakhir, kernel Sigmoid juga menunjukkan konsistensi dengan akurasi tertinggi 94% pada rasio split 85%. Kesimpulannya, pemilihan kernel yang tepat bergantung pada karakteristik data tertentu, namun secara umum, SVM menunjukkan fleksibilitas dan kemampuan yang baik dalam berbagai situasi menangkap tanggapan pelanggan terhadap *e-commerce* pada platform sosial media X.

Kata kunci: SVM, *Machine Learning*, *Marketplace Online*, *Social media X*

Andrianto, F. (2023). **Sentiment Analysis of E-commerce Services on Sosial Media X Using Linear and Non-Linear SVM Methods**. Thesis, Master Program of Informatics, Universitas Ahmad Dahlan, Yogyakarta.

Abstract

In the era of digital technology development, the e-commerce trade sector has become the center of online business and shopping. Social media X, has become a platform for customers to convey their experiences with e-commerce services. The sentiments contained in customer tweets can provide important insights into their level of satisfaction or dissatisfaction with a service. The use of Machine Learning Algorithms in sentiment analysis is used to detect whether customer satisfaction is included in positive or negative comments. The purpose of this research is to conduct sentiment analysis of customer satisfaction through X social media users using Machine Learning algorithms, in order to identify the right algorithm in recognizing visitor tweets in the positive or negative sentiment category towards three large Indonesian online e-commerce namely tokopedia, shopee, and lazada.

The amount of data generated by social media platforms makes proper classification analysis necessary. Classification stages include preprocessing, feature extraction and selection, split data classification and evaluation. Linear and non-linear svm models are chosen for this research because linear svm is one of the popular choices in the world of machine learning, especially for handling large datasets and features. The linear svm model can provide good results in cases with a linear relationship between features and class labels. Whereas non-linear svm has the flexibility to handle complex and non-linear relationships between features and class labels. Therefore, the use of these two models can provide a holistic understanding of complex data patterns in the analysis of customer sentiment towards e-commerce services.

The evaluation results show that SVM models with different types of kernels provide good performance in predicting data. The Linear kernel, with the highest accuracy of 94% at 80% split ratio, stands out in the balance between training and testing data. The Polynomial kernel, although slightly lower with 93% accuracy at 75% split ratio, still showed solid performance. The RBF kernel gave the best results at 90% split ratio, showing resilience in handling more training data than testing. Finally, the Sigmoid kernel also showed consistency with the highest accuracy of 94% at a split ratio of 85%. In conclusion, the selection of the right kernel depends on the specific data characteristics, but in general, SVM shows good flexibility and ability in various situations to capture customer responses to e-commerce on social media platform X.

Keywords: SVM, Machine Learning, Marketplace Online, Social media X,