

IMPLEMENTASI SUPPORT VECTOR MACHINE (SVM) DENGAN QUERY EXPANSION RANKING PADA REVIEW PENGUNAAN JAMU MADURA

IMPLEMENTATION OF SUPPORT VECTOR MACHINE (SVM) WITH QUERY EXPANSION RANKING IN THE REVIEW OF MADURA TRADITIONAL HERBAL MEDICINE

Rika Yunitarini¹⁾, Hambali Fitrianto²⁾, Fifin Ayu Mufarroha³⁾

^{1), 2), 3)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo
Jl Raya Telang, PO. BOX 2, Kamal, Bangkalan

Email : rika.yunitarini@trunojoyo.ac.id¹⁾, hambali.fitrianto01@gmail.com²⁾,
fifin.mufarroha@trunojoyo.ac.id³⁾

Abstrak

Jamu Madura adalah minuman herbal tradisional yang terbuat dari bahan-bahan alami dan sudah terkenal khasiatnya. Popularitas jamu Madura tidak hanya didasarkan pada keragaman produk jamu dan manfaatnya bagi kesehatan, tetapi juga pada nilai-nilai tradisional yang telah diwariskan secara turun-temurun. Salah satu jamu Madura yang cukup banyak peminatnya adalah jamu peluntur. Jamu peluntur adalah rangkaian produk obat atau jamu yang dirancang khusus sebagai solusi untuk mengatasi keterlambatan menstruasi atau ketidaklancaran haid, yang seringkali menjadi sumber kekhawatiran bagi ibu dan remaja putri. Dengan latar belakang meningkatnya permintaan terhadap produk jamu Madura ini, maka dilakukan analisis sentimen terhadap ulasan produk jamu Madura di aplikasi Shopee, Lazada, dan Tokopedia. Penelitian ini menerapkan Support Vector Machine dan Query Expansion Ranking untuk mencapai akurasi tertinggi dalam melakukan review terhadap penggunaan jamu Madura. Hasil yang diperoleh untuk penggunaan algoritma Support Vector Machine yaitu memiliki akurasi 93%, sementara untuk penggunaan algoritma Support Vector Machine dan Query Expansion Ranking pada rasio seleksi fitur 50% dan 100% meningkat akurasinya menjadi 94%.

Kata kunci: Analisis Sentimen, Jamu Madura, Jamu Peluntur, Query Expansion Ranking, Support Vector Machine

Abstract

Madura traditional herbal medicine is a traditional herbal medicine made from natural ingredients and is well-known for its efficacy. The popularity of Madura traditional herbal medicine is not only based on the diversity of traditional herbal medicine products and their health benefits, but also on traditional values that have been passed down from generation to generation. One of the most popular Madura traditional herbal medicine is Peluntur traditional herbal medicine. Peluntur traditional herbal medicine is a series of medicinal or herbal products specifically designed as a solution to overcome late menstruation or irregular menstruation, which is often a source of concern for mothers and young women. With the background of the increasing demand for Madura traditional herbal medicine products, a sentiment analysis was conducted on Madura traditional herbal medicine product reviews on the Shopee, Lazada, and Tokopedia applications. This study applies Support Vector Machine and Query Expansion Ranking to achieve the highest accuracy in reviewing the use of Madura traditional herbal medicine. The results obtained for the use of the Support Vector Machine algorithm have an accuracy of 93%, while for the use of the Support Vector Machine and Query Expansion Ranking algorithms at feature selection ratios of 50% and 100% the accuracy increases to 94%.

Keywords: Madura traditional herbal medicine, Peluntur traditional herbal medicine, Query Expansion Ranking, Sentiment Analysis, Support Vector Machine.

1. PENDAHULUAN

Jamu merupakan minuman herbal tradisional yang terbuat dari bahan-bahan alami sebagai alternatif pengobatan. Jamu dikonsumsi untuk merawat kesehatan, seperti mengurangi kelelahan, mengurangi nyeri, serta menjaga stamina. Selain itu, jamu juga sering digunakan untuk pencegahan, pengobatan, dan pemulihan dari penyakit, serta memperlancar menstruasi atau haid bagi para wanita [1], jamu Madura telah menjadi terkenal sejak zaman nenek moyang kita dan

terus diwariskan dari satu generasi ke generasi berikutnya secara turun-temurun [2]. Menurut hasil Riset Kesehatan Dasar, diketahui bahwa sebanyak 59.12% dari penduduk Indonesia yang berusia di atas 15 tahun pernah menggunakan obat tradisional, terutama jamu. Penggunaan obat tradisional ini melibatkan berbagai wilayah, termasuk baik di pedesaan maupun perkotaan. Pada kelompok usia 55-64 tahun, prevalensi penggunaan obat tradisional mencapai 67.69%, dengan perempuan memiliki presentase yang lebih tinggi (61.87%) dibandingkan dengan laki-laki (56.33%) [3]. Prevalensi penggunaan dan konsumsi jamu di Indonesia dipengaruhi oleh sejumlah faktor, di antaranya tingkat pengetahuan individu. terbatasnya pengetahuan mengenai manfaat jamu, disertai dengan penyebaran informasi yang kurang akurat, menjadi tantangan utama yang menyebabkan variasi dalam penerimaan jamu sebagai obat tradisional.

Diantara beberapa jamu Madura terdapat jamu yang sering dikonsumsi yaitu jamu peluntur. Dibandingkan jamu Madura yang lain, jamu peluntur memiliki frekuensi yang lebih sering dalam pengkonsumsian. Jamu peluntur adalah rangkaian produk obat yang dirancang untuk merangsang menstruasi yang terlambat atau tidak teratur. terbuat dari bahan herbal alami dan diproses dengan resep tradisional Madura, produk ini aman digunakan oleh wanita dari berbagai usia, mulai dari remaja hingga ibu dewasa. jamu ini digunakan secara terus menerus dan dianggap sebagai kebutuhan bagi wanita yang ingin memperlancar menstruasi dan menjaga siklusnya tetap teratur. Dengan bahan-bahan alami yang terkandung dalam jamu ini, diharapkan dapat memberikan solusi tanpa efek samping yang berlebihan [4]. Berdasarkan analisis di atas maka peneliti tertarik untuk melakukan penelitian lebih lanjut terkait dengan jamu peluntur. Keberlanjutan konsumsi jamu peluntur menjadi fokus penelitian, karena jamu ini diminum secara terus menerus (rutin) sebagai suatu kebutuhan, terutama bagi wanita yang ingin menjaga kelancaran menstruasi.

Tingginya permintaan terhadap produk jamu Madura [5], merupakan peluang bagi para pemangku kepentingan industri jamu Madura untuk mengembangkan usahanya. Selain itu, bahan baku jamu Madura yang relatif aman mendorong produsen jamu Madura untuk terus berinovasi memberikan alternatif pengobatan kepada masyarakat. Penjualan jamu Madura semakin meningkat [1]. Dalam penelitian ini, dilakukan analisis sentimen terhadap 1000 ulasan produk di aplikasi Shopee, Lazada dan Tokopedia. Analisis sentimen adalah tugas pengelompokan atau klasifikasi teks yang tidak terstruktur yang sering berisi pendapat orang tentang topik tertentu dengan tujuan untuk lebih memahami pendapat orang tentang topik tersebut. Misalnya, sebuah organisasi melakukan analisis sentimen pada halaman media sosial mereka (mis., Twitter, Facebook) untuk memahami bagaimana masyarakat umum memandang organisasi dan produk/layannya [6]. Bagian ulasan produk terdiri dari konten komentar dalam format teks bebas dan peringkat bintang dari 1 hingga 5. Konten komentar ulasan digunakan untuk mengetahui informasi apa yang menjadi fokus pembeli saat memberikan ulasan. Konten komentar ulasan dapat mencakup beberapa peringkat atribut produk, tetapi karena pada setiap ulasan hanya memiliki satu peringkat bintang, jumlah bintang tidak dapat mewakili semua fitur produk yang dinilai oleh pembeli [7]. Keadaan ini membutuhkan analisis sentimen dari data ulasan konsumen dan nantinya analisis sentimen ini akan ditujukan pada produsen jamu Madura [1]. Untuk ulasan positif berarti produk tersebut bagus dan dapat dipertahankan, sedangkan ulasan negatif dapat menjadi bahan untuk mengevaluasi produk dan memungkinkan pengembangan produk yang lebih efektif [8]. Berdasarkan penelitian tersebut, penelitian ini mengusulkan untuk menganalisis sentimen terhadap jamu Madura dengan menggunakan *Support Vector Machine* dan *Query Expansion Ranking* untuk mendapatkan hasil yang paling akurat [9].

2. DASAR TEORI

2.1. Jamu Peluntur

Jamu peluntur adalah rangkaian produk obat yang dirancang untuk merangsang menstruasi yang terlambat atau tidak teratur. terbuat dari bahan herbal alami dan diproses dengan resep tradisional Madura, produk ini aman digunakan oleh wanita dari berbagai usia, mulai dari remaja hingga ibu dewasa. jamu ini digunakan secara terus menerus dan dianggap sebagai kebutuhan bagi wanita yang ingin memperlancar menstruasi dan menjaga siklusnya tetap teratur.

dengan bahan-bahan alami yang terkandung dalam jamu ini, diharapkan dapat memberikan solusi tanpa efek samping yang berlebihan [4].

2.2. Analisis Sentimen

Analisis sentimen, di sisi lain, merupakan bidang studi yang fokus pada analisis opini, sentimen, evaluasi, sikap, dan emosi yang diekspresikan oleh masyarakat. Ini merupakan salah satu bidang penelitian yang paling aktif dalam pemrosesan bahasa alami, serta menjadi fokus dalam penambahan data, penambahan web, dan penambahan teks [10].

2.3. Scraping

Penelitian ini, teknik *scraping* digunakan sebagai alat yang efektif untuk mengumpulkan data ulasan jamu Madura yang nantinya akan menjadi dasar analisis sentimen dan pemahaman lebih lanjut mengenai pandangan masyarakat terhadap jamu madura. Teknik *scraping* adalah sebuah pendekatan yang memungkinkan kita untuk mengubah data web yang awalnya tidak terstruktur menjadi data terstruktur yang dapat disimpan dalam basis data pusat atau *spreadsheet* pusat. Proses *scraping* web terdiri dari tiga tahap penting: pengambilan data melalui protokol HTTP dengan pengiriman dan penerimaan permintaan, ekstraksi data dari halaman HTML, dan analisis data yang terkumpul. Setelah data terkumpul, transformasi data juga menjadi langkah penting dalam mempersiapkan data untuk analisis lebih lanjut [11].

2.4. Preprocessing

Preprocessing data dilakukan agar data yang akan dimasukkan ke dalam model analisis sesuai dengan karakteristik model yang akan dihasilkan [12]. Terdapat lima tahap dalam *preprocessing* ini, dimulai dari “*case folding*” yang mengubah semua huruf dalam teks menjadi huruf kecil, “*Punctuation Removal*” yang menghilangkan elemen seperti *mention*, *link*, *hashtag*, URL, *emoticon*, angka, dan tanda baca yang tidak relevan untuk analisis sentimen. Tahap selanjutnya adalah “*tokenizing*,” yaitu memotong teks menjadi bagian-bagian lebih kecil yang disebut token. Kemudian ada “*stopword removal*” yang menghilangkan kata-kata yang tidak berperan penting dalam analisis sentimen, seperti ‘dan’ atau ‘atau.’ Terakhir, tahap “*stemming*” menghilangkan awalan dan akhiran kata, mengembalikan kata-kata tersebut ke bentuk dasar [13].

2.5. Ekstraksi Fitur TF-IDF

Setelah melakukan proses *preprocessing* data tahapan berikutnya dalam proses analisis sentimen adalah tahapan pembobotan, di mana kata-kata dasar tersebut diubah menjadi sebuah vektor. Hal ini berkaitan dengan metode dalam penelitian ini. Pada penelitian ini, metode pembobotan, *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) [14], *Term Frequency – Inverse Document Frequency* (TF-IDF) merupakan sebuah algoritma untuk menunjukkan frekuensi kata yang keluar dalam dokumen. *Term Frequency* (TF) adalah frekuensi munculnya kata dalam suatu dokumen, sedangkan *Document Frequency* (DF) adalah banyaknya dokumen yang mengandung kata tertentu. Nilai bobot didapat dari TF dan *Inverse* dari DF [15].

$$IDF(w) = \log\left(\frac{N}{DF(w)}\right) \quad \dots (1)$$

$$TF - IDF(w, d) = TF(w, d) \times IDF(w) \quad \dots (2)$$

2.6. Seleksi Fitur Query Expansion Ranking (QER)

Metode ini mengurangi kompleksitas komputasi dan untuk mempercepat proses komputasi serta tidak menurunkan kualitas analisis sentimen. Query dimasukkan oleh pengguna dan digabungkan dengan model pembobotan probabilitas [16]. Semakin tinggi skor fitur, semakin besar perbedaan probabilitas antara kelas positif dan negatif. Hal ini mengindikasikan bahwa fitur tersebut lebih spesifik untuk setiap kelas dan lebih berarti untuk proses klasifikasi [17].

2.7. Support Vector Machine

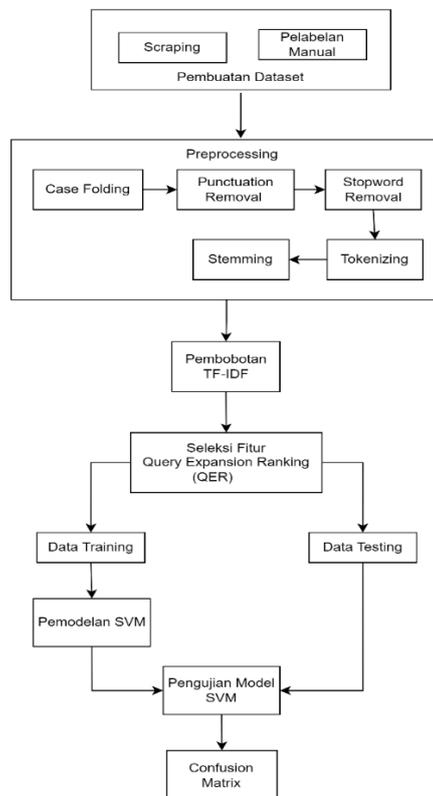
SVM merupakan metode klasifikasi yang termasuk dalam kategori pembelajaran terawasi atau *supervised learning* [18]. karena memiliki label atau target pembelajaran tertentu, Svm memahami hubungan antara input (fitur) dan output (label atau target) dengan menggunakan data pelatihan yang sudah diberi label atau target [19]. Tujuan dari SVM adalah untuk membuat keputusan antara dua kelas yang memprediksi label. Biasanya ada batas keputusan, yang disebut *hyperplane*, yang sejauh mungkin dari titik data terdekat dari setiap kelas [20].

2.8. Confusion Matrix

Confusion Matrix memiliki peran penting dalam memberikan informasi tentang sejauh mana metode ini berhasil dalam mengklasifikasikan data dengan benar. Parameter yang digunakan adalah akurasi [21], Akurasi digunakan untuk mengetahui seberapa besar keakuratan untuk mendeteksi nilai suatu angka hasil pengukuran terhadap angka sebenarnya [22].

3. METODOLOGI PENELITIAN

Metode ulasan yang digunakan dalam penelitian ini adalah analisis sentimen berdasarkan ulasan konsumen tentang jamu Madura. Ulasan tentang jamu Madura, khususnya jamu peluntur, dikumpulkan melalui beberapa situs *market place* (*Shopee, Lazada, dan Tokopedia*) berjumlah 1000 data. Pengumpulan data ini digunakan untuk mengetahui bagaimana tanggapan atau sentimen pengguna jamu peluntur (konsumen) terhadap jamu yang mereka konsumsi. Sehingga analisis ini dapat membantu proses perhitungan terkait banyaknya sentiment positif atau negatif, yang nantinya bermanfaat bagi para produsen jamu Madura. Serta dapat menjadi bahan untuk mengevaluasi produk dan memungkinkan pengembangan produk jamu yang lebih efektif lagi bagi penggunaannya. Data ulasan yang diambil adalah data teks berbahasa Indonesia menggunakan *instant data scraper*.



Gambar 1. Metodologi Penelitian

Gambar 1 merupakan metode penelitian dengan Algoritma *Support Vector Machine* (SVM) digunakan sebagai metode analisis sentimen utama, sedangkan *Query Expansion Ranking* digunakan untuk mengetahui pengaruh nilai akurasi dalam analisis sentimen. Langkah pertama dalam arsitektur di atas, adalah pembuatan dataset yang meliputi proses *scraping* dan pelabelan

manual. Selanjutnya dataset yang terbentuk akan dilakukan *preprocessing*. Tahap *preprocessing* yang dilakukan adalah dengan melakukan ekstraksi data dengan tujuan agar data dokumen ulasan Jamu Madura lebih terstruktur dan memudahkan proses-proses selanjutnya. Setelah *preprocessing* adalah melakukan pembobotan TF-IDF, dimana kata diubah menjadi bentuk numerik yang menghasilkan matriks vektor. Setelah itu masuk ke tahapan seleksi fitur menggunakan *query expansion ranking*, yang kemudian dilanjutkan dengan pembagian dataset, pemodelan klasifikasi SVM, pengujian dan diakhiri evaluasi untuk mengetahui akurasi dari model yang sudah dibangun.

Penelitian sebelumnya yang dilakukan oleh Muhammad Rangga Aziz Nasution dkk tahun 2019 yang berjudul “Perbandingan Akurasi dan Waktu Pemrosesan Algoritma K-NN dan SVM pada Analisis Sentimen Twitter”. Penelitian tersebut menyimpulkan bahwa *Support Vector Machine* lebih akurat daripada pengklasifikasi *K-Nearest Neighbour*, yaitu 89,70% tanpa validasi *K-Fold cross-validation* dan 88,76% tanpa validasi *K-Fold cross-validation* [14]. Selanjutnya, penelitian yang dilakukan pada tahun 2021 oleh Bintang Sifa Amalia dkk. yang berjudul “Analisis Sentimen Ulasan Pelanggan Restoran Menggunakan Algoritma *Support Vector Machine* dan *K-Nearest Neighbour*”. Dapat disimpulkan bahwa hasil SVM pada penelitian ini lebih unggul dengan menghasilkan nilai akurasi sebesar 81.92% dan AUC sebesar 0.918 [12]. Penelitian yang dilakukan pada tahun 2022 oleh Hasbi Atsqualani dkk. yang berjudul “Analisis Sentimen Data Twitter Indonesia Menggunakan *Support Vector Machine* dan *Query Expansion Ranking*”. Dalam penelitian ini, dapat disimpulkan bahwa tujuan dari penelitian ini adalah untuk mengklasifikasikan data Twitter Indonesia ke dalam polaritas sentimen positif dan negatif dengan menggunakan *Support Vector Machine* dan *Query Expansion Ranking*. Dengan menggunakan *support vector machine* dan *query augmented ranking* untuk mengklasifikasikan data twitter Indonesia ke dalam polaritas sentimen positif dan negatif, maka informasi yang terkandung di dalamnya dapat dieksplorasi dan informasi yang berguna dapat diberikan kepada pihak-pihak yang membutuhkan dari data yang diamati. Tahapan penelitian meliputi data *crawling*, *pre-processing* data, term *frequency-inverse document frequency* (TF-IDF), pemilihan fitur *query expansion ranking* dan klasifikasi data menggunakan metode *support vector machine* (SVM). Untuk mengetahui performa dari proses klasifikasi ini, maka dilakukan pengujian dengan menggunakan matriks konfigurasi. Hasil penelitian menunjukkan bahwa klasifikasi yang diusulkan mencapai akurasi sebesar 77% dan nilai *F-measure* sebesar 68% [13].

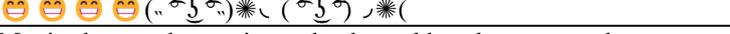
Selanjutnya, penelitian yang dilakukan oleh Adi Ariyo Munandar dkk pada tahun 2022 yang berjudul ‘Analisis Sentimen Aplikasi Pembelajaran Online Menggunakan Klasifikasi SVM’. Penelitian ini menyimpulkan bahwa rata-rata akurasi data klasifikasi sentimen SVM adalah 99% untuk Ruang Guru, 96% untuk Zenius dan 82% untuk Quipper. Pada penelitian ini, akurasi SVM berhasil ditingkatkan hingga mencapai akurasi 99%. Hasil ini didapatkan dengan membandingkan data dari Ruang Guru, Zenius dan Quipper [11].

4. PENGUJIAN DAN PEMBAHASAN

4.1. Pengumpulan Dataset

Data diambil dengan teknik Scraping dengan total data yang akan digunakan sebanyak 1000 ulasan pada jamu Madura di aplikasi *market place* shopee, Lazada, dan Tokopedia. Berikut merupakan contoh data ulasan jamu madura :

Tabel 1. Contoh Data Ulasan Jamu Madura

D	Ulasan	Tanggal	Sumber
1	mantap banget Jamunya 😄 bakalan coba lagi, mantap pokoknya	24/02/2021	shopee
2	Beneran ampuh banget jamunya,  	02/08/2022	shopee
3	Manjur banget, baru minum dua kapsul besoknya mens dong	01/10/2022	sazada
4	Keren pokoknya, jamunya manjur 🤗🤗🤗 haid langsung lancer	10/03/2022	tokopedia
5	Lancar banget haid setelah minum jamu ini 🐱🐱	18/07/2022	shopee

6	Gak ada efeknya, sudah habis saya minum tapi gak haid juga 😞😞	01/02/2022	shopee
7	Dari hari Rabu Sampek sekarang belum mempan juga 😞😞😞😞	04/09/2021	tokopedia
8	Sudah habis dua kotak tapi gak ada reaksi 😞😞	17/11/2019	lazada

4.2. Preprocessing

Tahap *preprocessing* yang dilakukan adalah dengan melakukan ekstraksi data dengan tujuan agar data dokumen ulasan jamu Madura lebih terstruktur dan memudahkan proses-proses selanjutnya. Dalam rangka mempermudah proses *preprocessing*, digunakan sampel data sebanyak 8 dokumen yang ada pada Tabel 2. Hasil tahapan *preprocessing* adalah sebagai berikut:

Tabel 2. Contoh Data Sesudah di *Preprocessing*

D	Hasil <i>Preprocessing</i>
1	mantap banget jamu coba mantap pokok
2	bener ampuh banget jamu
3	manjur banget minum kapsul besok mens
4	keren pokok jamu manjur haid langsung lancer
5	lancar banget haid minum jamu
6	efek habis minum haid
7	rabu sampek mempan
8	habis kotak aksi

Di bawah ini terdapat perhitungan dengan metode TD-IDF. Tabel 3 merupakan Contoh perhitungan dilakukan dengan menggunakan delapan data ulasan yang telah melewati tahap pra-pemrosesan sebagaimana yang telah dijelaskan dalam Tabel 3.8 sebelumnya. Dalam proses ekstraksi fitur ini, bobot diberikan kepada setiap kata yang terdapat dalam dokumen, dengan nilai TF-IDF ditentukan berdasarkan frekuensi kata tersebut dalam dokumen

Tabel 3. Hasil Pembobotan TF-IDF

Bobot(W) tf*idf							
Wd1	Wd2	Wd3	Wd4	Wd5	Wd6	Wd7	Wd8
1,806179 974	0	0	0	0	0	0	0
0,301029 996	0,301029 996	0,301029 996	0	0,301029 996	0	0	0
0,301029 996	0,301029 996	0	0,301029 996	0,301029 996	0	0	0
0,903089 987	0	0	0	0	0	0	0
0,602059 991	0	0	0,602059 991	0	0	0	0
0	0,903089 987	0	0	0	0	0	0
0	0,903089 987	0	0	0	0	0	0

Bobot(W) tf*idf							
Wd1	Wd2	Wd3	Wd4	Wd5	Wd6	Wd7	Wd8
0	0	0,602059 991	0,602059 991	0	0	0	0
0	0	0,425968 732	0	0,425968 732	0,425968 732	0	0
0	0	0,903089 987	0	0	0	0	0
0	0	0,903089 987	0	0	0	0	0
0	0	0,903089 987	0	0	0	0	0
0	0	0	0,903089 987	0	0	0	0
0	0	0	0,425968 732	0,425968 732	0,425968 732	0	0
0	0	0	0,903089 987	0	0	0	0
0	0	0	0,602059 991	0,602059 991	0	0	0
0	0	0	0	0	0,903089 987	0	0
0	0	0	0	0	0,602059 991	0	0,602059 991
0	0	0	0	0	0	0,903089 987	0
0	0	0	0	0	0	0,903089 987	0
0	0	0	0	0	0	0,903089 987	0
0	0	0	0	0	0	0	0,903089 987
0	0	0	0	0	0	0	0,903089 987

4.3. Query Expansion Ranking (QER)

Langkah selanjutnya adalah menghitung nilai skor yang akan digunakan dalam *Query Expansion Ranking* pada tabel 4. Perhitungan nilai skor atau *score function* dapat dilakukan dengan menggunakan Persamaan 3. Contoh perhitungan score untuk kata “haid” dapat dilakukan sebagai berikut.

$$scoref = \frac{|pf+qf|}{|pf-qf|} \dots (3)$$

$$scoref = \frac{|0,490196078 + 0,428571429|}{|0,490196078 - 0,428571429|} = 14,90909091$$

Tabel 4. Query Expansion Ranking (QER)

TERM	Dff+	Pf	Dff-	Qf	Score	Ranking
Minum	2	0,490196078	1	0,428571429	14,90909091	1
Haid	2	0,490196078	1	0,428571429	14,90909091	2
Coba	1	0,294117647	0	0,142857143	2,888888889	3
Bener	1	0,294117647	0	0,142857143	2,888888889	4
Ampuh	1	0,294117647	0	0,142857143	2,888888889	5
kapsul	1	0,294117647	0	0,142857143	2,888888889	6
besok	1	0,294117647	0	0,142857143	2,888888889	7
mens	1	0,294117647	0	0,142857143	2,888888889	8
keren	1	0,294117647	0	0,142857143	2,888888889	9
langsung	1	0,294117647	0	0,142857143	2,888888889	10
mantap	2	0,490196078	0	0,142857143	1,822580645	11
pokok	2	0,490196078	0	0,142857143	1,822580645	12
manjur	2	0,490196078	0	0,142857143	1,822580645	13
lancar	2	0,490196078	0	0,142857143	1,822580645	14
banget	4	0,882352941	0	0,142857143	1,386363636	15
jamu	4	0,882352941	0	0,142857143	1,386363636	16
habis	0	0,098039216	2	0,714285714	-1,318181818	17
efek	0	0,098039216	1	0,428571429	-1,593220339	18
rabu	0	0,098039216	1	0,428571429	-1,593220339	19
sampek	0	0,098039216	1	0,428571429	-1,593220339	20
mempan	0	0,098039216	1	0,428571429	-1,593220339	21
kotak	0	0,098039216	1	0,428571429	-1,593220339	22
aksi	0	0,098039216	1	0,428571429	-1,593220339	23

4.4. Klasifikasi Support Vector Machine (SVM)

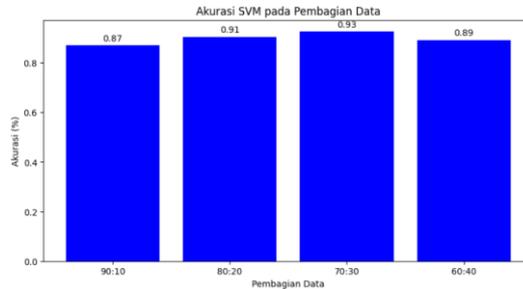
Selanjutnya masuk ke tahapan klasifikasi SVM dengan memasukkan hasil Seleksi Fitur, jika hasil klasifikasi lebih dari 0 maka hasil dinyatakan sebagai positif dan jika hasil kurang dari 0 atau sama dengan 0 maka hasil dinyatakan negatif, dan Langkah selanjutnya pengukuran performa dengan confusion matrix.

4.5. Analisis dan Uji Coba

Tahap ini dilakukan pembelajaran model untuk proses klasifikasi yakni mengelompokkan sentimen dalam sebuah kalimat kedalam kategori kelas positif dan negatif. Model yang digunakan untuk klasifikasi sentimen adalah SVM tanpa SVM dan SVM + QER dengan rasio seleksi fitur untuk QER 25%, 50%, 75%, 100% dengan pembagian data latih dan uji berdasarkan persentase 90% data latih 10% data uji, 80% data latih 20% data uji, 70% data latih 30% data uji, 60% data latih 40% data uji untuk setiap model. Proses analisis dan uji coba ini

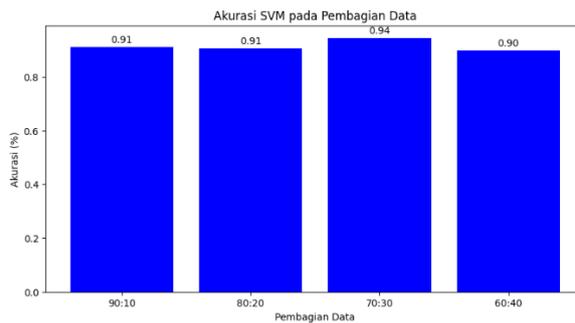
dilakukan menggunakan google collaboratory yaitu editor bahasa pemrograman Python secara online. Dalam artikel ini akan ditampilkan dua hasil uji coba, yang pertama tanpa QER dan yang kedua dengan QER 75%.

Berdasarkan implementasi proses klasifikasi SVM seperti pada gambar 4.32 akan menghasilkan akurasi yang ditampilkan dalam bentuk grafik sebagai berikut.



Gambar 2. Hasil perbandingan akurasi SVM tanpa QER

Gambar 2 mendapatkan hasil dari perbandingan dataset sebagai berikut, 90:10 sebesar 87%, 80:20 sebesar 91%, 70:30 sebesar 93%, dan 60:40 sebesar 89%, jadi untuk akurasi terbaik pada pengujian klasifikasi SVM ada pada pembagian dataset 70:30 yaitu sebesar 93%. Berdasarkan implementasi proses klasifikasi SVM dan QER dengan rasio seleksi fitur 100% seperti pada gambar 7 akan menghasilkan akurasi yang ditampilkan dalam bentuk grafik sebagai berikut.



Gambar 3. Hasil perbandingan akurasi SVM dan QER dengan rasio seleksi fitur 100%.

Gambar 3 mendapatkan hasil dari perbandingan dataset sebagai berikut, 90:10 sebesar 91%, 80:20 sebesar 91%, 70:30 sebesar 94%, dan 60:40 sebesar 90%, jadi untuk akurasi terbaik pada pengujian klasifikasi SVM dan QER dengan rasio seleksi fitur 50% ada pada pembagian dataset 70:30 yaitu sebesar 94%.

Tabel 5 menunjukkan perbandingan nilai akurasi dengan menggunakan SVM tanpa QER serta menggunakan SVM dengan QUR, didapatkan hasil bahwa penggunaan *Query Expansion Ranking* dapat meningkatkan kemampuan model dalam mengenali sentimen ulasan Jamu Madura.

Tabel 5. Perbandingan Hasil Akurasi

	Perbandingan Dataset	Rasio Seleksi Fitur	Akurasi
<i>Support Vector Machine (SVM)</i>	90:10	-	87%
	80:20	-	91%
	70:30	-	93%
	60:40	-	89%
	90:10	25%	76%
	80:20		84%
	70:30		85%
	60:40		83%
	90:10		90%

	Perbandingan Dataset	Rasio Seleksi Fitur	Akurasi
<i>Support Vector Machine (SVM) dan Query Expansion Ranking (QER)</i>	80:20	50%	92%
	70:30		94%
	60:40		87%
	90:10	75%	93%
	80:20		92%
	70:30		90%
	60:40		89%
	90:10	100%	91%
	80:20		91%
	70:30		94%
	60:40		90%

5. KESIMPULAN

Berdasarkan hasil penelitian, pengujian, dan analisis yang telah dilakukan, penelitian ini menghasilkan kesimpulan dari skenario pengujian pertama, yang melibatkan *support vector machine*, menunjukkan akurasi tertinggi pada pembagian dataset 70:30 sebesar 93%. Pengujian ini membuktikan efektivitas *support vector machine* dalam mengklasifikasikan sentimen ulasan jamu Madura. Sementara itu pada skenario pengujian kedua yang melibatkan *support vector machine* dan *query expansion ranking*, diperoleh bahwa penggunaan *query expansion ranking* memberikan kontribusi positif terhadap peningkatan akurasi klasifikasi sentimen. Akurasi tertinggi tercapai pada pembagian dataset 70:30, yaitu 94%, dengan rasio seleksi fitur 50% dan pada pembagian dataset 70:30, yaitu 94%, dengan rasio seleksi fitur 100%. Ini membuktikan bahwa penggunaan *query expansion ranking* dapat meningkatkan kemampuan model dalam mengenali sentimen ulasan Jamu Madura. Untuk pengembangan lebih lanjut dari penelitian ini, perlu melakukan optimasi pada parameter *support vector machine* seperti kernel dan parameter C. Selain itu perlunya menambah jumlah data ulasan Jamu Madura dapat berkontribusi pada peningkatan keakuratan model. Dan dapat menambahkan kelas netral dalam analisis sentimen dapat membantu mengatasi situasi di mana ulasan tidak jelas atau tidak menunjukkan sentimen positif atau negatif secara tegas.

DAFTAR PUSTAKA

[1] R. Yunitarini, E. Widiawanti, P. Adi, and P. Nugroho, “Menggunakan Metode Waterfall Information System of Madura Herb Stock Using Waterfall Method,” J. SimanteC, vol. 11, no. 1, pp. 65–72, 2022.

[2] S. Kristianto, J. Batoro, S. Widyarti, and S. B. Sumitro, “Exploration and economic value of medicinal plants as traditional herbal ingredients in bangselok, madura, indonesia,” Proc. Int. Conf. Ind. Eng. Oper. Manag., no. August, 2020.

[3] M. R. Adiyasa and M. Meiyanti, “Pemanfaatan obat tradisional di Indonesia: distribusi dan faktor demografis yang berpengaruh,” J. Biomedika dan Kesehat., vol. 4, no. 3, pp. 130–138, 2021, doi: 10.18051/jbiomedkes.2021.v4.130-138.

[4] Eva, “Jamu Pluntur /Tepat Bulan (herbs to avoid late menstrual period),” Maduherbal, 2021. <https://maduraherbal.com/jamu-pluntur-tepat-bulan-herbs-to-avoid-late-menstrual-period-herbal-code-jt-011/> (Diakses 3 September 2024).

[5] I. F. Pramasari and N. Q. Wijaya, “Strategi Pengembangan Jamu Ramuan Madura Di Kabupaten Sumenep,” J. Pertan. Cemara, vol. 18, no. 1, pp. 50–63, 2021, doi: 10.24929/fp.v18i1.1365.

[6] Prasetyo, Vincentius Riandaru, Ihza Akbar Ryanda, and Delta Ardy Prima. “Analisis Sentimen Dan Kategorisasi Review Pelanggan Pada Cafe Kopi Paste Dengan Metode Naive Bayes Dan K-Nearest Neighbor.” NERO (Networking Engineering Research Operation) 8.1 (2023): 1-8.

- [7] E. H. Muktafin, K. Kusriani, and E. T. Luthfi, "Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing," *J. Eksplora Inform.*, vol. 10, no. 1, pp. 32–42, 2020, doi: 10.30864/eksplora.v10i1.390.
- [8] R. Ireland and A. Liu, "Application of data analytics for product design: Sentiment analysis of online product reviews," *CIRP J. Manuf. Sci. Technol.*, vol. 23, pp. 128–144, 2018, doi: 10.1016/j.cirpj.2018.06.003.
- [9] A. P. Natasuwarna, "Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring," *Techno.Com*, vol. 19, no. 4, pp. 437–448, 2020, doi: 10.33633/tc.v19i4.4044.
- [10] S. S. Istia and H. D. Purnomo, "Sentiment analysis of law enforcement performance using support vector machine and K-nearest neighbor," *Proc. - 2018 3rd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2018*, pp. 84–89, 2018, doi: 10.1109/ICITISEE.2018.8720969.
- [11] A. A. Munandar, F. Farikhin, and C. E. Widodo, "Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 8, no. 2, p. 77, 2023, doi: 10.31328/jointecs.v8i2.4747.
- [12] B. Sifa Amalia, Y. Umaidah, R. Mayasari, S. Karawang Jl HSRonggo Waluyo, K. Telukjambe Timur, and K. Karawang, "Analisis Sentimen Review Pelanggan Restoran Menggunakan Algoritma Support Vector Machine Dan K-Nearest Neighbor," *SITEKIN J. Sains, Teknol. dan Ind.*, vol. 19, no. 1, pp. 28–34, 2021.
- [13] H. Atsqalani, N. Hayatin, and C. S. K. Aditya, "Sentiment Analysis from Indonesian Twitter Data Using Support Vector Machine And Query Expansion Ranking," *J. Online Inform.*, vol. 7, no. 1, p. 116, 2022, doi: 10.15575/join.v7i1.669.
- [14] M. R. A. Nasution and M. Hayaty, "Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter," *J. Inform.*, vol. 6, no. 2, pp. 226–235, 2019, doi: 10.31311/ji.v6i2.5129.
- [15] S. Rizkia, E. Budi Setiawan, and D. Puspendari, "Analisis Sentimen Kepuasan Pelanggan Terhadap Internet Provider Indihome di Twitter Menggunakan Metode Decision Tree dan Pembobotan TF-IDF," *e-Proceeding Eng.*, vol. 6, no. Agustus, pp. 9683–9693, 2019.
- [16] S. Fanissa, M. A. Fauzi, and S. Adinugroho, "Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, pp. 2766–2770, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] Y. A. V. Gunawan, N. A. S. ER, I. B. M. Mahendra, I. M. Widiartha, I. G. N. A. C. Putra, and I. G. A. G. A. Kadyanan, "Analisis Sentimen Ulasan Aplikasi Transportasi Online Menggunakan Multinomial Naïve Bayes dan Query Expansion Ranking," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 11, no. 1, p. 121, 2022, doi: 10.24843/jlk.2022.v11.i01.p13.
- [18] F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan, "Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine," *J. Inform. Polinema*, vol. 4, no. 2, p. 93, 2018, doi: 10.33795/jip.v4i2.152.
- [19] S. Huang, C. A. I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, and X. U. Wayne, "Applications of support vector machine (SVM) learning in cancer genomics," *Cancer Genomics and Proteomics*, vol. 15, no. 1, pp. 41–51, 2018, doi: 10.21873/cgp.20063.
- [20] F. Prasetiawan¹, S. Widiyanesti², and T. Widarmanti³, "Analisis Sentimen Mengenai Kualitas Layanan Jasa Ekspedisi Barang Sicepat Di Media Sosial Twitter Sentiment Analysis Regarding Quality of Sicepat Expedition Services On Twitter Social Media," *e-Proceeding Manag.*, vol. 9, no. 2, pp. 147–160, 2022.

- [21] F. R. Irawan, A. Jazuli, and T. Khotimah, “Analisis Sentimen Terhadap Pengguna Gojek Menggunakan Metode K-Nearset Neighbors Sentiment Analysis of Gojek Users Using K-Nearest Neighbor,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 62–68, 2022, doi: 10.33387/jiko.
- [22] M. Birjali, M. Kasri, and A. Beni-Hssane, “A comprehensive survey on sentiment analysis: Approaches, challenges and trends,” *Knowledge-Based Syst.*, vol. 226, p. 107134, 2021, doi: 10.1016/j.knosys.2021.107134.