

PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE DALAM MENGANALISIS SENTIMEN ULASAN PELANGGAN SHOPEEFOOD BERDASARKAN TWITTER

Siti Nur Rismanah¹, Rini Astuti², Fadhil M. Basysyar³

^{1,3} Teknik Informatika, STMIK IKMI Cirebon

² Sistem Informasi, STMIK LIKMI Bandung

Jalan Perjuangan No 10B Majasem, Kotra Cirebon, Indonesia

srismanah@gmail.com

ABSTRAK

Platform seperti *ShopeeFood* yang dapat digunakan untuk memesan makanan secara *online* telah menjadi fenomena di industri *e-commerce*. Fenomena ini muncul sebagai hasil dari perubahan perilaku konsumen yang semakin mengutamakan kenyamanan, efisiensi, dan berbagai pilihan pengalaman makanan. Pesatnya pertumbuhan pemesanan makanan secara online juga memunculkan tantangan dan peluang baru, seperti analisis sentimen terhadap ulasan pelanggan untuk mengetahui tingkat kepuasan, kekhawatiran, dan preferensi pelanggan. Penelitian ini bertujuan untuk menganalisis sentimen terhadap ulasan pelanggan *ShopeeFood* berdasarkan pengguna Twitter dengan mengambil data melalui teknik data *crawling*. Pada penelitian ini, data ulasan pelanggan *ShopeeFood* dari media sosial Twitter yang telah terkumpul dari tanggal 6 September hingga 16 Oktober 2023. Data ini mencakup ulasan pelanggan *ShopeeFood* sebanyak 2246 data ulasan pelanggan, dan menggunakan metode *Knowledge Discovery in Database (KDD)*. Setelah dilakukan nya pengambilan data, tahapan selanjutnya adalah *preprocessing* data. Tahapan *preprocessing* data ini mencakup *cleaning data*, *case folding*, *labeling*, *tokenize*, *filter tokens by length*, *stopword* dengan menggunakan aplikasi RapidMiner dan pengklasifikasian menggunakan algoritma *Support Vector Machine*. Hasil dari pengujian ini didapatkan nilai akurasi sebesar 80.31%, presisi 73.22%, recall 95.58% dan error rate 19.68%.

Kata kunci : *Support Vector Machine, Analisis Sentimen, ShopeeFood, Twitter*

1. PENDAHULUAN

Pada era digital saat ini yang terus berkembang, kemajuan pesat di bidang informatika telah memberikan dampak besar pada berbagai aspek kehidupan manusia. Bidang bisnis dan layanan pelanggan adalah contoh nyata dari pengaruh teknologi informatika. Platform seperti *ShopeeFood* yang dapat digunakan untuk memesan makanan secara *online* telah menjadi fenomena di industri *e-commerce*. Fenomena ini muncul sebagai hasil dari perubahan perilaku konsumen yang semakin mengutamakan kenyamanan, efisiensi, dan berbagai pilihan pengalaman makanan. Jika layanan diberikan dengan lebih baik dan lebih cepat, setiap pelanggan akan lebih puas dengan penyedia layanan yang ada, dan kepuasan pelanggan adalah nilai utama bagi perusahaan penyedia layanan [1].

Dalam beberapa tahun terakhir, media sosial sebagai platform untuk berbagi pengalaman dan ulasan telah berkembang pesat, menghasilkan data dalam jumlah besar yang berisi informasi penting. Meningkatnya penggunaan media sosial sebagai platform dimana pelanggan dapat mengekspresikan diri mereka sendiri mengenai barang dan jasa. Salah satunya adalah Twitter, Twitter ini sering digunakan untuk berbagi informasi, berita terkini, gagasan, serta interaksi sosial antar pengguna di seluruh dunia. Tren ini telah menyebabkan masalah utama yaitu volume data yang sangat besar. Sangat tidak mungkin untuk menganalisis jutaan tweet dan ulasan pelanggan yang ada di Twitter secara manual. Pesatnya pertumbuhan pemesanan makanan secara online juga memunculkan

tantangan dan peluang baru, seperti analisis sentimen terhadap ulasan pelanggan untuk mengetahui tingkat kepuasan, kekhawatiran, dan preferensi pelanggan.

Analisis sentimen juga dikenal sebagai penggalian opini, adalah cabang studi yang meneliti opini, sentimen, penilaian, sikap, dan emosi orang yang berkaitan dengan produk, layanan, organisasi, individu, isu, peristiwa, atau topik tertentu. *Opinion mining* digunakan untuk meneliti sikap terhadap produk atau layanan tertentu, isu politik dan sosial, serta kegiatan yang sedang berlangsung[2]. *Support Vector Machine (SVM)* adalah salah satu metode klasifikasi yang paling populer. Metode klasifikasi ini menggunakan pembelajaran terawasi yang memprediksi kelas berdasarkan pola dari proses pelatihan. Klasifikasi dilakukan dengan *hyperplane* yang memisahkan kelas positif dan negatif. *Hyperplane* yang baik adalah yang memiliki jarak terbesar ke titik data training terdekat dari setiap kelas, karena pada umumnya semakin besar margin, semakin rendah kesalahan generalisasi pemisahnya. *Hyperplane* pemisah terbaik antara dua kelas dapat dilakukan dengan mengukur margin *hyperplane* dan mencari titik maksimumnya.[3]. Pengujian algoritma *Support Vector Machine (SVM)* dilakukan oleh [4], untuk mengklasifikasikan sentimen dalam data Twitter. Proses tersebut melibatkan tahap *preprocessing* yang kompleks, termasuk pembersihan data, tokenisasi, normalisasi kata, penghilangan *stopword*, serta analisis jenis kata dan *stemming*. Hasilnya menunjukkan bahwa SVM mampu mencapai

akurasi tertinggi sebesar 93% saat diuji dengan data uji yang diperoleh selama pengumpulan data. Evaluasi dengan *K-Fold Cross Validation* juga menunjukkan konsistensi performa SVM dengan rata-rata akurasi sebesar 83.5%. Kesimpulannya, SVM merupakan pilihan yang efektif dalam mengklasifikasikan sentimen pada data Twitter setelah dilakukan preprocessing yang matang.

Tujuan dari penelitian ini adalah menerapkan algoritma *Support Vector Machine* (SVM) dalam menganalisis sentimen ulasan pelanggan ShopeeFood berdasarkan pengguna Twitter. Sehingga mendapatkan pemahaman tentang bagaimana *Support Vector Machine* (SVM) dapat digunakan untuk mengklasifikasikan ulasan pelanggan terkait layanan pesan-antar makanan dan bagaimana hasil analisis ini dapat memberikan informasi kepada perusahaan seperti ShopeeFood. Pemahaman yang baik tentang sentimen pelanggan dapat memberikan manfaat yang signifikan, membantu layanan mereka dan memenuhi harapan pelanggan dengan lebih baik. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan kontribusi positif bagi perkembangan industri layanan pesan-antar makanan.

Dalam penelitian ini dilakukan pengambilan data dari media sosial Twitter dengan cara *crawling* data menggunakan kata kunci “*shopeefood*”. Sehingga didapat data ulasan pelanggan *ShopeeFood* sebanyak 2246 data ulasan terhitung dari 6 September hingga 16 Oktober 2023. *Knowledge Discovery in Database* (KDD) digunakan untuk menganalisis data dan mengidentifikasi pola sentimen yang mungkin tersembunyi dalam data yang besar dan kompleks. Dengan menggunakan pendekatan metode data *mining* dengan menerapkan algoritma *Support Vector Machine* (SVM). Metode ini akan membantu mendapatkan wawasan yang lebih mendalam tentang layanan pesan-antar makanan dan mengklasifikasikan ulasan pelanggan.

Diharapkan bahwa penelitian ini memiliki potensi implikasi di bidang Informatika, khususnya dalam analisis sentimen pelanggan di era digital. Keberhasilan pencapaian tujuan penelitian akan memberikan kontribusi dalam memahami efektivitas algoritma *Support Vector Machine* (SVM) dalam menganalisis sentimen ulasan pelanggan pada industri pesan-antar makanan seperti ShopeeFood. Memberikan pemahaman tentang kemampuan *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen pelanggan secara akurat, serta dalam bahasa yang tidak terstruktur dan beragam yang sering ditemui di media sosial seperti Twitter.

2. TINJAUAN PUSTAKA

ShopeeFood adalah salah satu dari layanan transportasi daring yang sedang populer saat ini, yang mengkhususkan diri dalam pengantaran makanan kepada pelanggan[5]. Pentingnya kepuasan pelanggan untuk membantu pelanggan lain dalam memilih atau menentukan rekomendasi makanan enak dengan harga

murah. Ulasan dari pengguna seringkali menjadi sumber informasi yang efektif dan efisien dalam mencari informasi tentang suatu produk atau aplikasi [6]. Penelitian yang dilakukan Wijaya dkk membahas tentang kepuasan konsumen dari penyedia layanan pengantaran makanan dengan menggunakan data dari Twitter menggunakan tiga algoritma yang berbeda, yaitu *Stochastic Gradient Descent*, *Support Vector Machine*, dan *K-Nearest Neighbor*, untuk mengklasifikasikan sentimen pelanggan sebagai positif atau negatif. Data juga diproses dengan dan tanpa menggunakan TF-IDF untuk membandingkan hasilnya. Hasil evaluasi menunjukkan bahwa ketiga model tersebut memiliki akurasi dan nilai F di atas 78%, dengan peningkatan sekitar 1% setelah menggunakan TF-IDF [1].

Text Mining merupakan metode yang digunakan untuk mengeksplorasi dan mengekstrak informasi dari teks tertulis. Pendekatan ini memungkinkan penggalian informasi dari data teks dengan cara menganalisis, mengidentifikasi pola. Dengan memproses dan menganalisis sejumlah data besar, *text mining* juga merupakan proses untuk menemukan informasi atau tren baru. Twitter adalah salah satu platform sosial yang sangat terkenal di antara pengguna internet karena kemudahannya dalam penggunaan dan kesederhanaannya, memungkinkan pengguna untuk menyuarakan pendapat atau opini mereka secara bebas [7]. Dalam penelitiannya Alhaq dkk membahas mengenai penerapan metode *Support Vector Machine* untuk analisis sentimen pengguna twitter terhadap *marketplace* Bukalapak. Tujuan dari penggunaan metode *Support Vector Machines* (SVM) adalah untuk mengkategorikan ulasan tersebut ke dalam kategori yang bersifat negatif, netral, atau positif terhadap layanan Bukalapak. Hasil penelitian menunjukkan bahwa SVM berhasil mencapai tingkat akurasi sebesar 93%, menunjukkan kehandalan metode ini dalam mengidentifikasi pendapat pengguna dengan tepat. Selain itu, SVM dapat membantu bagaimana pengguna melihat layanan Bukalapak di platform media sosial seperti Twitter. Hasil penelitian ini menunjukkan bahwa analisis sentimen sangat penting untuk memahami bagaimana pengguna bertindak terhadap layanan e-commerce di era digital saat ini [4].

Teknik klasifikasi dalam *support vector machine* yang didasarkan pada metode machine learning yang dikembangkan oleh Vladimir Vapnik adalah algoritma *Support Vector Machine* (SVM). Metode ini memproyeksikan kelas berdasarkan pola yang telah dipelajari dari proses *training*. Dalam SVM, klasifikasi dilakukan dengan menggunakan *hyperplane* untuk memisahkan kelas opini positif dan negatif. Dalam konsepnya, sebuah *hyperplane* yang optimal dianggap sebagai yang memiliki jarak terbesar dari titik data pelatihan terdekat dari setiap kelasnya. Hal ini karena pada umumnya, semakin besar margin (jarak), semakin rendah kesalahan generalisasi dari

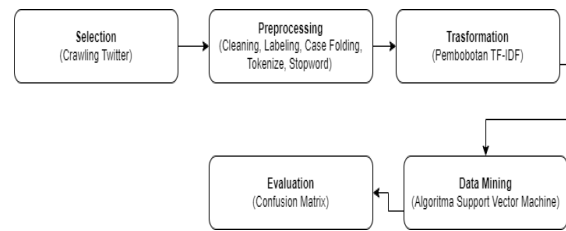
proses pemisahan kelas. Margin ini merupakan jarak dari titik vektor di suatu kelas terhadap hyperplane [8].

Penelitian yang dilakukan Tiniges dkk bertujuan untuk menganalisis sentimen pengguna layanan *Indihome* berdasarkan data tweet yang terdapat di Twitter. Metode yang digunakan dalam penelitian ini adalah *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen pengguna menjadi positif atau negatif. Hasil dari penelitian ini menunjukkan bahwa metode SVM mampu memberikan akurasi sebesar 87% dengan *precision* 86%, *recall* 95%, *error rate* 13%, dan *f1-score* 90%. Dari hasil analisis sentimen didapatkan bahwa tingkat kepuasan pengguna layanan *Indihome* cenderung rendah, dengan sekitar 81,6% sentimen negatif. Kesimpulan dari penelitian ini adalah metode SVM dapat digunakan secara efektif dalam menganalisis sentimen pengguna layanan berbahasa Indonesia di Twitter, dan hasilnya menunjukkan tingkat kepuasan yang kurang baik terhadap layanan *Indihome* [9].

Knowledge Discovery in Database (KDD) adalah suatu metode yang digunakan untuk mendapatkan informasi atau pengetahuan yang terdapat dalam suatu database yang sudah ada, yang terdiri dari tabel-tabel yang memiliki hubungan satu sama lain. Hasil dari proses pengetahuan ini dapat dijadikan sebagai basis pengetahuan untuk mendukung pengambilan keputusan. KDD dan data mining sering digunakan untuk menggali informasi yang tersembunyi dari database yang berukuran besar. Proses KDD terdiri dari 5 tahap, diantaranya adalah pemilihan data, pre-processing, transformasi, data mining, dan evaluasi [10]. Dengan menggunakan data yang diperoleh dari Twitter, Agustina dkk membahas analisis sentimen publik terhadap tiga marketplace besar di Indonesia yaitu *Shopee*, *Tokopedia*, dan *Bukalapak*. Mengklasifikasikan persepsi publik terhadap layanan dan masalah di ketiga pasar tersebut menjadi persepsi positif atau negatif adalah tujuan utama penelitian. Berbagai langkah pre-processing teks, seperti pembersihan data, pengolahan teks, dan pengelompokan sentimen, digunakan. Penelitian ini juga menjelaskan hasil analisis, menunjukkan bahwa sebagian besar tweet memiliki sentimen positif, dan menunjukkan hasil klasifikasi terbaik dari pengujian *cross-validation* k-fold. Hasil ini menunjukkan respons media sosial publik terhadap ketiga pasar tersebut[11]

3. METODE PENELITIAN

Dalam penelitian ini, menerapkan pendekatan *Knowledge Discovery in Databases* (KDD) sebagai metode utama untuk mengeksplorasi informasi dari dataset yang tersedia. Tahapan KDD yang diterapkan terdiri dari beberapa langkah yang mencakup pemilihan dan pemahaman data, pra-pemrosesan data, transformasi data, pemodelan, evaluasi model. Rangkaian langkah-langkah dalam penelitian ini mencakup beberapa tahap, yakni:



Gambar 1. Tahapan Penelitian KDD

Berdasarkan gambar 1 dapat dijabarkan langkah-langkah dari proses KDD yang dilakukan sebagai berikut:

3.1. Selection

Pada tahap ini, dilakukan pengumpulan data ulasan pelanggan *ShopeeFood* pada Twitter menggunakan teknik *crawling* dengan rentang waktu yang digunakan yaitu pada tanggal 6 September hingga 16 Oktober 2023 dan data yang berhasil terkumpul sebanyak 2246 data.

3.2. Preprocessing

Tahapan *preprocessing* ini mencakup beberapa langkah yang diantaranya adalah *cleaning*, *labeling*, *case folding*, *tokenize*, dan *stopword* untuk membersihkan data dari nilai yang hilang atau tidak valid, menangani *outlier*, menangani duplikasi, dan menyesuaikan data dengan format yang benar.

3.3. Transformation

Tahap selanjutnya adalah *transformation*, dimana dilakukan pembobotan kata menggunakan metode TF-IDF guna mengekstraksi fitur untuk mengubah teks ulasan menjadi numerik.

3.4. Data Mining

Pada tahap data mining, menggunakan algoritma *Support Vector Machine* untuk mengklasifikasikan kelas. Dilakukan pengujian data dan mengevaluasi hasil temuan yang dihasilkan oleh model algoritma SVM.

3.5. Evaluation

Tahapan terakhir adalah *evaluation*, pada tahap ini dilakukan evaluasi model menggunakan *confusion matrix* yang diantaranya adalah akurasi, presisi dan recall.

4. HASIL DAN PEMBAHASAN

4.1. Selection

Pada penelitian ini didapatkan data hasil *crawling* pada media sosial Twitter sebanyak 2246 data dengan 12 variabel yang diantaranya adalah *created_at*, *id_str*, *full_text*, *quote_count*, *reply_count*, *retweet_count*, *favorite_count*, *lang*, *user_id_str*, *conversation_id_str*, *username* dan *tweet_url*.

```

python --yes tweet-harvest@2.2.8 -o "(filename)" -s "[search_keyword]" -l (limit) --token (twitter_auth_token)
  
```

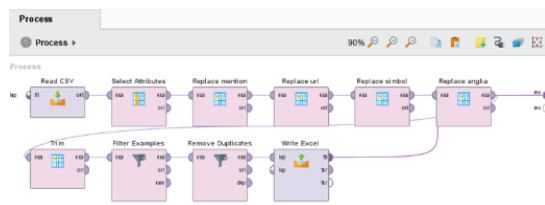
Gambar 2. Crawling Data

Gambar 2 menunjukkan script yang digunakan untuk melakukan crawling data dengan menggunakan bahasa pemrograman *python* pada google *colab*.

4.2. Preprocessing

4.2.1. Cleaning Data

Pada tahap ini dilakukan pembersihan teks dari username dan URL, simbol, angka, dan data duplikat.



Gambar 3. Cleaning data

Pada gambar 3 menunjukkan operator yang digunakan pada Rapidminer untuk tahap *cleaning*, operator *Replace* digunakan untuk menghapus simbol, url, username. Operator *Trim* digunakan untuk menghapus spasi yang berada diawal kalimat. Sedangkan operator *Filter Examples* digunakan untuk membatasi entri yang tidak valid, seperti nilai kosong, nilai yang hilang. Operator *Remove Duplicates* digunakan untuk menghapus baris-baris yang duplikat dalam dataset, dan memastikan bahwa baris tersebut hanya muncul sekali. Setelah tahapan *cleaning* dataset dilakukan selanjutnya adalah menyimpan dataset yang bersih dan siap digunakan kedalam file excel dengan menggunakan operator *Write Excel*.

Tabel 1. Hasil Cleaning Data

Sebelum	Sesudah
skrg shopeefood hiks, KNP SIHHHHHHH https://t.co/juvDvPKHwo	skrg shopeefood hiks, KNP SIHHHHHHH

Tabel 1 menampilkan hasil antara teks sebelum dan setelah dilakukan proses *cleaning* data terlihat sebuah kalimat yang semula terdapat url akan terhapus karena telah dilakukan tahap *cleaning* data.

4.2.2. Labeling

Setelah dilakukan pembersihan dataset maka tahap selanjutnya adalah *labeling* yaitu proses memberikan atau menandai teks atau data dengan label atau klasifikasi untuk menunjukkan pendapat atau perasaan yang terkandung di dalamnya.

Sentimen	Text
Positif	pls w tr tau shopeefood ternyata semurah itu
Negatif	mewah pisan nkr rek mesen dahar jeng awewe mah kuat ka kabeh aplikasi dibuka gofood shopeefood grabfood terus milih daharun ...
Negatif	hangis bgl idup lgi cpek cepetnya mesen shopeefood dapet driver tolol
Negatif	Jangan belanja di shopeefood pake kartu kredit gais Udah pas pembayaran bermasalah dicancel malah kagak balik duitnya
Positif	kesempatan makan gratis SHOPEEFood
Positif	Shopeefood diskonnya ampe rb anjay
Positif	Nasi Goreng Empal Mercon Seuhah Nasi Goreng Spesial dengan bumbu khas Sambel Seuhah ditambah Empal Mercon Seuhah yang S...
Negatif	gue sehatan scroll in shopeefood taunya beli ayam bang dawa lagi crsbar lagi
Positif	Diskon Shopeeefood cek jam nya disini ya
Positif	janjan lupa kalm pocer shopeefood diskon buat jajan hari ini scroll agak ke bawah

Gambar 4. Labeling

Gambar 4 menunjukkan label sentimen yang diberi warna hijau dengan ulasan disampingnya.

4.2.3. Case Folding

Pada tahap ini teks diubah dari huruf kapital menjadi huruf kecil secara keseluruhan (*lower case*) menggunakan operator *Transform Cases* pada RapidMiner.

Tabel 2. Hasil Case Folding

Sebelum	Sesudah
BELI DI SHOPEEFOOD ADA DISKON SUMPAAH SETENGAH HARGA TADI AKU BELI E	beli di shopeefood ada diskon sumpah setengah harga tadi aku beli e

Tabel 2 menunjukkan hasil antara teks sebelum dan setelah proses *Case Folding*, terlihat dari perubahan huruf yang semula berupa huruf kapital menjadi huruf kecil.

4.2.4. Tokenizing

Proses memisahkan teks menjadi token-token atau kata-kata tunggal. Pada saat yang sama, proses ini juga menghilangkan karakter lain, seperti tanda baca dengan menggunakan operator *Tokenize* pada Rapidminer.

Tabel 3. Hasil Tokenizing

Sebelum	Sesudah
napa shopeefood driver double order beda resto gini cancel pilih driver hadeuuh lapeeeer laper	[napa], [shopeefood], [driver], [double], [order], [beda], [resto], [gini], [cancel], [pilih], [driver], [hadeuuh], [lapeeeer], [laper]

Tabel 3 menunjukkan baris teks sebelum dan sesudah dilakukan *Tokenize* yang semula sebuah kalimat menjadi potongan kata.

4.2.5. Stopword

Digunakan operator *Stopword* untuk menghilangkan kata-kata umum yang tidak memberikan nilai tambah dalam analisis.

Tabel 4. Hasil Stopword

Sebelum	Sesudah
tetap lah aku pilih promo ayam geprek di shopeefood	tetap aku pilih promo ayam geprek shopeefood

Tabel 4 menunjukkan perbedaan teks antara sebelum dan sesudah digunakannya *stopword*.

4.2.6. Filter Token by Length

Pada tahapan ini digunakan parameter dari operator *Filter Token by Length* untuk mengatur jika ada teks yang kurang dari 4 huruf dan lebih dari 25 huruf maka akan terhapus.

4.3. Transformation

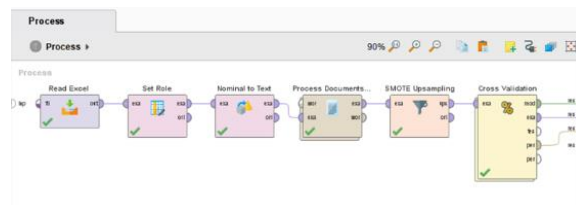
Tahapan selanjutnya adalah transformation, pada tahap ini akan dilakukan pembobotan kata menggunakan TF-IDF untuk menghitung bobot kata-kata kunci dari dataset terkait dengan ShopeeFood. Dalam pembobotan term, nilai atau bobot diberikan kepada setiap term yang ada pada setiap tweet yang telah melalui tahap *preprocessing*. Metode yang digunakan dalam pembobotan ini adalah TF-IDF, yang bertujuan untuk memberikan nilai kepada sebuah term untuk digunakan sebagai input dalam proses klasifikasi yang ditunjukkan pada gambar 5.

[illegible]

Gambar 5. Pembobotan TF-IDF

4.4. Data Mining

Pada tahap Data Mining melibatkan penggunaan beberapa operator di RapidMiner yang digambarkan pada gambar 6.



Gambar 6. Tahapan Klasifikasi Algoritma SVM

Pada gambar 6 menunjukkan proses klasifikasi yang diterapkan, dimulai dengan mengimpor data dari file Excel menggunakan operator "*Read Excel*". Kemudian, menggunakan operator "*Set Role*" untuk menetapkan peran atribut dalam dataset, "*Nominal to Text*" untuk mengubah atribut nominal menjadi teks, dan "*Process Documents from Data*", yang terdiri dari serangkaian operator, termasuk "*Tokenize*" untuk memecah teks menjadi kata-kata. "*Transform Cases*" menyesuaikan huruf, "*Stopword*" menghilangkan kata-kata stopwords, dan "*Filter Token by Length*" menghilangkan token yang tidak diinginkan. Operator "*SMOTE Upsampling*" digunakan untuk mengatasi ketidakseimbangan kelas dalam dataset.



Gambar 7. Proses didalam Cross Validation

Pada gambar 7 pemodelan algoritma SVM digunakan dalam tahap data *mining* untuk klasifikasi. Sebelum penerapan SVM, dilakukan pembagian data *training* dan *testing* menggunakan operator *cross validation* dengan 10-fold. SVM digunakan untuk mengklasifikasikan kelas berdasarkan data yang telah di proses sebelumnya, setelah dilakukan klasifikasi digunakan operator *performance* untuk mendapatkan hasil performa model SVM yang sudah diterapkan dan mengetahui nilai akurasi yang didapat oleh SVM.

4.5. Evaluation

Pada tahap *evaluation* digunakan operator *performance* yang nantinya menghasilkan *confusion matrix* diterapkan pada data testing yang kemudian akan diproses oleh model algoritma *Support Vector Machine* yang telah dibuat sebelumnya untuk melakukan uji performa *algoritma Support Vector Machine*.

Tabel 5. Confusion Matrix

	Actual Negatif	Actual Positif
Pred. Negatif	TN	FN
Pred. Positif	FP	TP

Pada tabel 5 *Confusion matrix* ada empat bagian utama yakni True Positif(TP), True Negatif(TN), False Positif(FP), dan False Negatif(FN). Adapun rumus perhitungan secara manual untuk mengetahui selisih nilai dari perhitungan menggunakan Rapidminer dan perhitungan secara manual.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$Precision = \frac{TP}{TP+FP} \times 100\%$$

$$Recall = \frac{TP}{TP+FN} \times 100\%$$

Berdasarkan hasil penelitian yang telah dilakukan menggunakan tools RapidMiner dengan algoritma SVM didapatkan hasil *accuracy* pada confusion matrix sebesar 80.31%.

accuracy: 80.31% +/- 3.49% (micro average: 80.31%)			
	true Negatf	true Positf	class precision
pred. Negatf	456	31	93.63%
pred. Positf	245	670	73.22%
class recall	65.05%	95.58%	

Gambar 8. Hasil RapidMiner

Pada tabel 8 dapat dilihat *True Negatif* (TN) sejumlah 456, menandakan jumlah ulasan yang benar-benar negatif dan diprediksi secara tepat sebagai negatif oleh model. Sementara *True Positif* (TP) mencapai 670, mewakili jumlah ulasan positif yang berhasil diprediksi dengan benar. Namun, ditemukan 31 *False Negatif* (FN) yang merupakan ulasan positif namun salah diprediksi sebagai negatif, serta 245 *False Positif* (FP) yang seharusnya negatif tapi salah diprediksi sebagai positif. *Precision* ulasan negatif sebesar 93.63% menunjukkan ketepatan dalam mengklasifikasikan ulasan negatif dan *precision*

ulasan positif sebesar 73.22% menunjukkan seberapa tepat model dalam mengklasifikasikan ulasan sebagai positif. Sedangkan *Recall* untuk ulasan negatif sebesar 65.05% dan tingkat pendeteksian untuk ulasan positif sebesar 95.58% menunjukkan seberapa baik model dalam mengidentifikasi ulasan positif atau negatif. Pada penelitian ini ditemukan *error rate* 19.68%. Proses perhitungan nilai *performance confusion matrix* pada metode *Support Vector Machine* (SVM) dapat dilakukan dengan perhitungan manual menggunakan rumus sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$Accuracy = \frac{670+456}{670+456+245+31} \times 100\%$$

$$Accuracy = \frac{1126}{1402} \times 100\% = 80.31\%$$

Dari hasil perhitungan diatas, menunjukan bahwa tingkat *accuracy* sebesar 80.31% dari uji klasifikasi yang dilakukan menggunakan algoritma *Support Vector Machine* (SVM).

$$Recall = \frac{TP}{TP+FN} \times 100\%$$

$$Recall = \frac{670}{670+31} \times 100\%$$

$$Recall = \frac{670}{701} \times 100\% = 95.58\%$$

Dari hasil perhitungan diatas, nilai *recall* sebesar 95.58% menunjukan seberapa baik model dalam mengidentifikasi ulasan dari uji klasifikasi yang dilakukan menggunakan algoritma *Support Vector Machine* (SVM).

$$Precision = \frac{TP}{TP+FP} \times 100\%$$

$$Precision = \frac{670}{670+245} \times 100\%$$

$$Precision = \frac{670}{915} \times 100\% = 73.22\%$$

Dari hasil perhitungan diatas, nilai *precision* sebesar 73.22% menunjukan ketepatan model dalam mengklasifikasikan ulasan.

Tabel 6. Perbandingan Perhitungan Menggunakan RapidMiner dan Manual

Perhitungan	Accuracy	Recall	Precision
RapidMiner	80.31%	95.58%	73.22%
Manual	80.31%	95.58%	73.22%

Pada Tabel 6, ditemukan bahwa hasil perhitungan yang dilakukan secara manual sama dengan hasil perhitungan menggunakan Rapidminer. Dengan menggunakan 1000 dataset ulasan

Shopeefood hasil *crawling* data dari Twitter 701 diantaranya adalah ulasan negatif dan 299 ulasan positif. Kata-kata yang sering muncul dalam kelas positif dan negatif akan divisualisasikan tampilan *wordcloud* dengan menggunakan sample 50 kata yang sering muncul.



Gambar 9 Wordcloud Sentimen Negatif

Pada gambar sentimen negatif, kata kunci yang sering muncul adalah “driver”. Ini dikarenakan banyak pelanggan yang mengeluh pada saat pengantaran makan driver lama karena driver mengambil *double order* atau mengambil lebih dari satu orderan. Bukan hanya itu, pelanggan juga mengeluh dikarenakan driver kurang begitu paham dalam membaca titik lokasi pengantaran sehingga sering terjadi salah alamat. Kata kunci “mahal” juga sering muncul pada sentimen negatif, dikarenakan biaya ongkos kirim yang mahal dan adanya biaya parkir yang ditanggung kepada pelanggan. Rekomendasi dari karakteritik sentimen negatif adalah peningkatan kualitas layanan, menonaktifkan fitur *double order* dan evaluasi keakuratan titik lokasi pengantaran.



Gambar 10 Wordcloud Sentimen Positif

Pada gambar sentimen positif, kata kunci yang sering muncul adalah “voucher”, pelanggan cukup puas dengan adanya voucher atau diskon yang diberikan kepada pelanggan. Kata kunci “enak” juga sering muncul dikarenakan makanan yang dipesan melalui ShopeeFood dinilai enak. Rekomendasi dari ulasan positif adalah penawaran promo dan diskon menarik bagi pengguna dipertahankan.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan yang telah dijabarkan, maka dapat ditarik kesimpulan terkait penerapan algoritma *Support Vector Machine* dalam

menganalisis ulasan pelanggan shopeefood berdasarkan twitter adalah hasil evaluasi model *confusion matrix* menunjukkan bagaimana SVM digunakan untuk memahami dan mengklasifikasikan sentimen ulasan pelanggan ShopeeFood. Ini menunjukkan kemampuan model untuk memprediksi baik sentimen positif maupun negatif dari ulasan yang diberikan, didapatkan nilai *accuracy* sebesar 80.31%, *precision* sebesar 73.22%, *recall* sebesar 95.58% dan *error rate* 19.68%.

ShopeeFood dapat meningkatkan layanan dengan langkah-langkah strategis. Prioritas utamanya adalah memperbaiki layanan pengiriman dengan pengawasan ketat terhadap aktivitas driver untuk mencegah kesalahan pengantaran dan meninjau ulang terkait biaya pengiriman yang dianggap tinggi serta menghapus biaya parkir untuk meningkatkan kepuasan pelanggan. Untuk penelitian berikutnya, analisis sentimen ulasan pelanggan dapat diperdalam dengan membandingkan metode *machine learning* seperti *Random Forest*, *Naïve Bayes*, *KNN* dan *Decision Tree*. Serta memperbanyak data untuk mendapatkan hasil yang bagus dan penerapan *Natural Language processing* (NLP) serta keterlibatan ahli bahasa akan meningkatkan pemahaman dan analisis terhadap ulasan pelanggan.

DAFTAR PUSTAKA

- [1] A. Wijaya, "Analisis Sentimen Kepuasan Pelanggan pada Penyedia Layanan Pengantaran Makanan," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 3, pp. 1699–1708, 2022, doi: 10.35957/jatisi.v9i3.2196.
- [2] Fadhillah Dwi, Ananda, and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 2, pp. 407–416, 2021, doi: 10.30812/matrik.v20i2.1130.
- [3] N. Fitriyah, B. Warsito, and D. A. I. Maruddani, "Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm)," *J. Gaussian*, vol. 9, no. 3, pp. 376–390, 2020, doi: 10.14710/j.gauss.v9i3.28932.
- [4] Z. Alhaq, A. Mustopa, S. Mulyatun, and J. D. Santoso, "Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter," *J. Inf. Syst. Manag.*, vol. 3, no. 2, pp. 44–49, 2021, doi: 10.24076/joism.2021v3i2.558.
- [5] M. Z. Farhan, "Analisis Sentimen Layanan ShopeeFood Pada Twitter Dengan Metode K-Nearest Neighbor, Support Vector Machine, dan Decision Tree," *J. Ilm. Inform.*, vol. 7, no. 2, pp. 95–106, 2023, doi: 10.35316/jimi.v7i2.95-106.
- [6] K. A. Rokhman, Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *J. Inf. Syst. Manag.*, vol. 2, no. 2, 2021.
- [7] D. L. Rianti, Y. Umaidah, and A. Voutama, "Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel Support Vector Machine," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 6, no. 1, p. 98, 2021, doi: 10.30998/string.v6i1.9993.
- [8] B. W. Sari and F. F. Haranto, "Implementasi Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter Terhadap Pelayanan Telkom Dan Biznet," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 171–176, 2019, doi: 10.33480/pilar.v15i2.699.
- [9] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [10] F. Romadoni, Y. Umaidah, and B. N. Sari, "Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 2, pp. 247–253, 2020, doi: 10.32736/sisfokom.v9i2.903.
- [11] D. A. Agustina, S. Subanti, and E. Zukhronah, "Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine," *Indones. J. Appl. Stat.*, vol. 3, no. 2, p. 109, 2021, doi: 10.13057/ijas.v3i2.44337.