

# Classification Of Sentiment Analysis In The Vn Video Editor Maker Vlognow Application Using The Method Support Vector Machine (SVM)

(Case Study: Play Store Comments/Reviews)

## Klasifikasi Analisis Sentimen Pada Aplikasi Vn Video Editor Maker Vlognow Dengan Menggunakan Metode Support Vector Machine (SVM)

(Studi Kasus: Komentar/Ulasan Play Store)

Hazmi Ali Wardhana<sup>1)</sup>, Mochammad Alfian Rosid<sup>2)</sup>, Ade eviyanti<sup>3)</sup>

<sup>1)</sup>Program Studi Teknik Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

<sup>2)</sup> Program Studi Teknik Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

Email Penulis Korespondensi: alfanrosid@umsida.ac.id

**Abstract.** VlogNow video editor maker is an editing platform application with a rating of 4.6 found on the Google Playstore. On the Google Playstore, the application is assessed using a rating and followed by reviews from users. Reviews can contain good opinions or even bad opinions based on the level of user satisfaction with the application. In this study, sentiment analysis will be carried out using the Support Vector Machine algorithm to determine user sentiment towards the VN application. The kernel used is the RBF (Radial Basis Function) kernel with a total of 1,000 data. The best accuracy results are obtained through the scenario of dividing the dataset with a ratio of 70:30 resulting in an accuracy of 88.40%.

**Keywords -** VN video editor, Playstore, Support Vector Machine, RBF

**Abstrak.** VlogNow video editor maker merupakan aplikasi platform editing dengan rating 4,6 yang terdapat di google playstore. Pada google playstore penilaian aplikasi dengan menggunakan rating dan diikuti review dari pengguna. Review dapat mengandung opini yang bersifat baik atau bahkan opini buruk berdasarkan tingkat kepuasan pengguna terhadap aplikasi. Dalam penelitian ini akan dilakukan analisis sentimen dengan menggunakan algoritma Support Vector Machine untuk mengetahui sentimen pengguna terhadap aplikasi VN. Kernel yang digunakan adalah kernel RBF (Radial Basis Function) dengan total data sebanyak 1.000 data. Hasil akurasi terbaik didapatkan melalui skenario pembagian dataset dengan perbandingan 70:30 dihasilkan akurasi sebesar 88,40%..

**Kata Kunci -** VN video editor, Playstore, Support Vector Machine, RBF

## I. Pendahuluan

Proses editing menghasilkan sebuah karya yang menampilkan berbagai kumpulan video yang telah menjadi sebuah cerita runtut dan juga dapat dengan layak menjadi konsumsi publik, karena semua proses itu editing video di perlukan *software* atau aplikasi untuk para editor video seperti aplikasi VlogNow video editor maker [1] di era serba digital ini untuk pemanfaatan pembelajaran bahasa Indonesia di sekolah. Peneliti menggunakan media editor video VN dalam menambah variasi peserta didik menulis teks prosedur. Dengan adanya penggunaan aplikasi edit video VN untuk media pembelajaran teks prosedur. VN atau yang memiliki nama lengkap VlogNow ini memiliki berbagai fitur menarik dan juga terbilang mudah digunakan. Belakangan ini, aplikasi tersebut menjadi buah bibir di kalangan pengguna smartphone karena aplikasinya yang sederhana namun tergolong powerful. Keistimewaan dari aplikasi VlogNow lebih mudah digunakan untuk orang-orang pemula. Karena tidak menampilkan banyak fitur sehingga tidak membuat peserta didik bingung. Selain itu, kapasitas memori di handphone tidak temakan terlalu banyak oleh aplikasi VN ini jadi tidak terlalu berat.

Google playstore merupakan sebuah laman yang memiliki fungsi untuk mendownload aplikasi pada android, penggunanya dapat memilih berbagai aplikasi yang diperlukan untuk menunjang kebutuhannya. Untuk mengetahui bagaimana kualitas sebuah aplikasi, google play store menyediakan fitur penilaian terhadap bagaimana performa aplikasi tersebut ketika digunakan. Hal tersebut dikarenakan pengguna terlebih dahulu ingin mengetahui aplikasi yang akan di *install*, sehingga fitur rating dan penilaian sangat sering dimanfaatkan terhadap aplikasi [9].

Namun terdapat berbagai rating yang berbeda dengan komentar yang ada, hal tersebut pada akhirnya menyebabkan calon pengguna kesulitan untuk mengetahui terkait penilaian pada aplikasi.

*Text mining* merupakan sebuah proses dalam eksplorasi data teks yang susunannya tidak terstruktur dengan didukung perangkat lunak yang mampu melakukan identifikasi terkait bagaimana konsep, pola, serta kata kunci dan atribut yang lainnya, ini biasa disebut dengan analisis teks. Salah satu tujuan *text mining* ini adalah melakukan analisis sentiment. Selanjutnya analisis sentiment merupakan sebuah aplikasi yang juga termasuk dalam text mining serta banyak dipakai untuk melacak sentiment pelanggan mengenai perusahaan. Ini juga dikenal dengan penambahan opini, analisis sentiment menambang teks yang berasal dari tinjauan online, hasil jejaring sosial, email, serta berbagai interaksi panggilan, dan sumber data yang lain dengan tujuan mengidentifikasi ulasan umum yang menuju kepada perasaan positif ataupun negative yang dirasakan oleh pelanggan aplikasi. [2].

Salah satu metode dalam *machine learning* untuk melakukan analisis sentiment adalah *Support Vector Machine* (SVM), *Support vector machine* bisa juga dijelaskan dengan lebih mudah sebagai upaya dalam mencari *hyperplane* dengan kondisi terbaik serta memiliki kegunaan untuk memisah dua buah *class* yang terdapat pada input *space*. Pada gambar 1a menunjukkan beberapa pattern yang menunjukkan sebagai anggota dari dua buah *class*: dengan keterangan positif (dinotasikan dengan +1) dan negative (dinotasikan dengan -1). Selanjutnya *pattern* pada *class* negatif ditunjukkan dengan symbol kotak, sebaliknya *pattern* pada *class* positif ditunjukkan dengan symbol lingkaran. Pada kegiatan pembelajaran problemklasifikasi diartikan bahwa usaha untuk menemukan garis (*hyperplane*) yang memisahkan dua kelompok yang telah dibahas tersebut.

Beberapa penelitian sebelumnya yang berkaitan dengan *penggunaan text mining* untuk mengklasifikasikan suatu objek adalah diantaranya penelitian yang dilakukan oleh Lutfi Budi Ilmawan, Muhammad Aliyazid Mude pada tahun 2020, Penelitian mengenai Perbandingan Metode Klasifikasi *Support Vector Machine* dan *Naïve Bayes* untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. Dalam penelitian itu dataset yang digunakan adalah komentar atau ulasan pada *google play store* dengan menggunakan metode *support vector machine* dan mendapatkan hasil yakni SVM *classifier* mendapatkan akurasi sebesar 81,46% dan *Naïve Bayes classifier* sebesar 75,41%, hal tersebut menunjukkan jika metode SVM lebih sesuai digunakan sebagai metode klasifikasi dengan tujuan melakukan proses analisis sentiment ulasan tekstual dengan Bahasa Indonesia pada *Google Play Store*. Selanjutnya penelitian yang dilakukan oleh Mochmad Tri Anjasmoros, Istiadi, Fitri Marisa pada tahun 2020, Penelitian mengenai Analisis Sentimen Aplikasi GO-JEK Menggunakan Metode *Support Vector Machine* dan *Naïve Bayes*, penelitian ini memakai data yang didapatkan pada ulasan aplikasi gojek pada play store dan mendapatkan hasil. Berdasarkan data yang diperoleh dari metode SVM dan NBC yang memiliki data training dengan perbedaan 10% hingga 90%. Selanjutnya grafik 1 *score accuracy* maksimal yakni 0.8612 dengan memakai metode SVM kernel linear dengan komponen data testing 30% data training 70%. *Score accuracy* yang paling rendah yakni 0.5 dengan memakai metode SVM kernel polynomial, RBF serta sigmoid.

Setelah didapatkan latar belakang tersebut sehingga dilaksanakan penelitian dengan tujuan mengklasifikasikan ulasan atau komentar terhadap aplikasi Vn video editor pada play store, dengan menerapkan metode *Support Vector Machine*, maka dilakukan penelitian proposal skripsi yang berjudul “Klasifikasi Analisis Sentiment Pada Aplikasi VN Video Editor Maker VlogNow Dengan Menggunakan Metode *Support Vector Machine* (SVM) (Studi Kasus : Komentar/Ulasan Play Store)” dari judul tersebut akan memperoleh tingkat akurasi dalam orientasi sentiment dan klasifikasi pada dua kelas, yakni ulasan dengan isi positif dan negative yang mempergunakan metode algoritma SVM.

## II. METODE

### Sumber Data

Dataset pada penelitian ini didapatkan melalui metode *crawling* pada situs Google Playstore, dengan menggunakan google collab untuk mengambil dataset yang berupa ulasan dengan disertai rating oleh pengguna aplikasi VN editor pada rentang waktu 10 Februari 2022 sampai dengan 20 Februari 2023. Dataset berjumlah 1.000 data yang akan diklasifikasikan kedalam sentimen positif, netral, dan negatif. Berikut dataset yang akan digunakan pada penelitian kali ini :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	userName	score	at	content															
2	Bayu April	1	2/20/2023 3:33	Perbaiki kualitas teks yang tidak rapi setelah selesai disimpan lalu dibagikan ke sosial media.															
3	aso papeh	1	2/20/2023 3:05	Mantap															
4	The Gimbi	1	2/19/2023 11:45	Good															
5	Cattleya P	1	2/19/2023 10:54	Great															
6	Mujahid. J	-1	2/19/2023 6:24	Kalo di sev gak pernah masuk ke galeri Kessel dengan apk yg satu ini susahÃ² ngedit malah ga bisa di sev															
7	FERY offici	1	2/18/2023 15:03	Sangat puas dengan apk ini, mohon untuk tambahkan fitur fitur lain yang lebih keren lagi, buat temanÃ² follow akun tiktok aku ya BG FERY															
8	Eman Cpk	1	2/18/2023 13:51	Knpa gak bisa download															
9	Anwar Mu	-1	2/18/2023 12:18	Di update malah makin error															
10	SUASOLI K	1	2/18/2023 6:14	Ok deh															
11	Aisyah ica	0	2/17/2023 10:58	Afaan tuh															
12	Damanhui	0	2/17/2023 10:41	Aplikasi bagus sangat membantu...tapi kenapa begitu udah beres editing.dan download vedeo nya ko delay trus yah gak sinkron. Ada yg bisa bantu jawab?															
13	milo lyrick	-1	2/17/2023 6:45	Awalnya suka banget sama ni apk, lebih simple ngeditnya dari apk yg lain.. tapi akhirÃ² ini suka tekeluar sendiri pdhl sdh di update... tolong dong perbaikannya															
14	Pangeran	-1	2/17/2023 4:26	Tak bisa mengirim video															
15	Nazib Mul	-1	2/17/2023 3:15	tiap ngedit lirik atau subtitle pasti selalu ngga pas kalau udah create, mabok															
16	duNo kesi	1	2/16/2023 16:44	Best after Cc															
17	Agung Pra	1	2/16/2023 13:52	Kangen gw sama aplikasi ini Alhamdulillah ketemu akhirnya ini pembuat auditor musik gw fyp nih makasih															
18	Iqbal Ptk	1	2/16/2023 6:40	GG gaming															
19	akunback	1	2/16/2023 6:00	Ok															
20	Wahyu Nu	1	2/15/2023 18:29	ðŸ™															
21	Tamimah	1	2/15/2023 17:21	Selama ini tidak ada masalah,aplikasi berjalan lancar															
22	Nazwa Pu	-1	2/15/2023 17:19	video gue udh cape ngedit ilangggg ajaggg lu gatau rasanya nangis malem malem ajggggfffd fucjshsgsvagahhggggggg trs gmnnnnnn capekkbbgtttcguee nged															
23	Niko eyl 0	-1	2/15/2023 15:57	VN:NOðŸrc CC:YESðŸ™ðŸ™															
24	Mr. One	0	2/15/2023 15:26	Setelah diupgrade, rendernya 99 persen malah ga bisa2 jadi....															
25	ORGENES	1	2/15/2023 13:39	Sangat super															
26	Nayaka At	1	2/15/2023 13:13	bagus untuk mnghd kn video															
27	Nila Sri W	0	2/15/2023 10:40	Dulu pas maka apk ini, seneng banget nggak ngelag kek gitu. Tapi setelah mengalami beberapa update, suka ngelag, ditambah pas ngimpor lama di awal, dan keti															

Gambar 1.Contoh Data

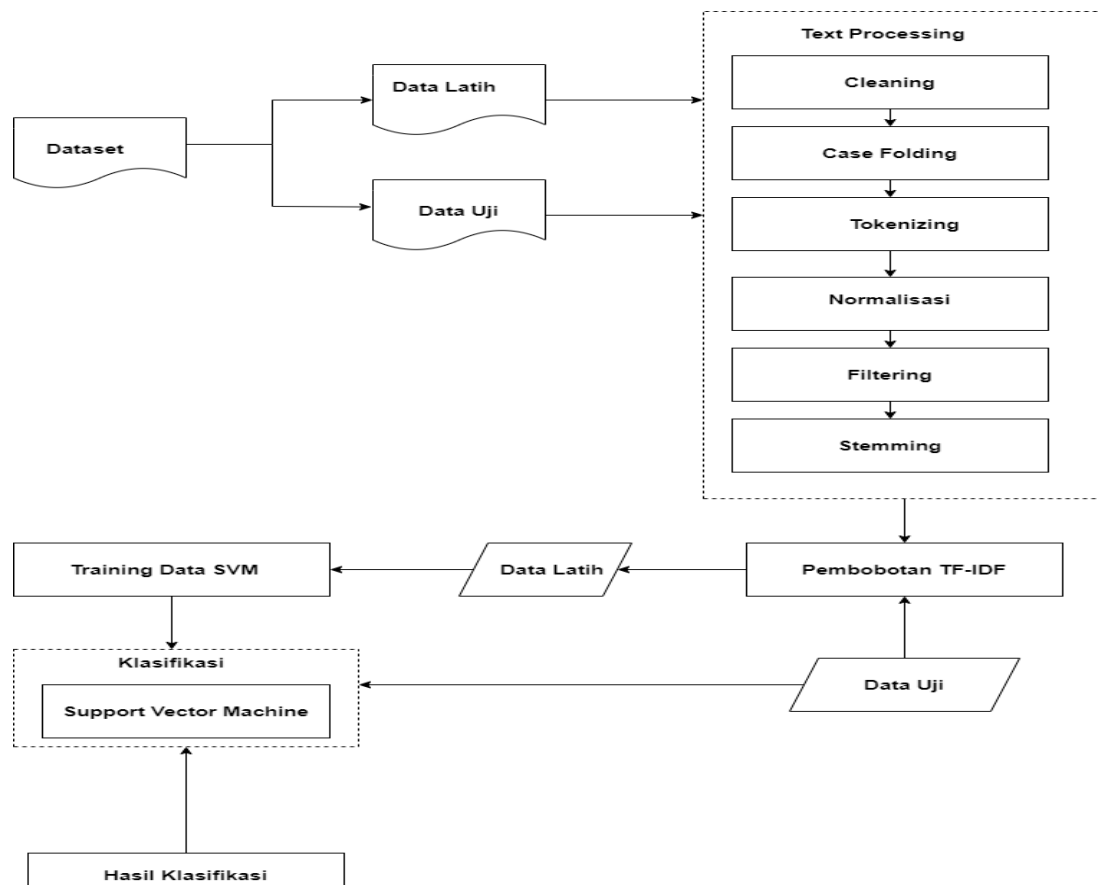
Selanjutnya dataset akan dikelompokkan menjadi 3 kelas sentiment yaitu positif, netral dan negatif. Metode dalam pelabelan dataset didasarkan pada rating yang diberikan oleh pengguna (Daryfayi Edyt & Asror, 2020). Ulasan dengan rating 4-5 yang berisi kata yang menunjukkan kepuasan pengguna aplikasi VN editor akan didefinisikan sebagai ulasan positif sedangkan ulasan dengan rating 1-2 yang berisikan kata-kata ketidakpuasan pengguna terhadap aplikasi VN editor akan didefinisikan sebagai ulasan negatif. Ulasan dengan rating 3 akan otomatis dikategorikan sebagai ulasan netral. Ulasan positif akan diberikan nilai 1 sedangkan untuk ulasan negatif akan diberikan nilai -1 sedangkan ulasan netral akan bernilai 0. Berikut sentimen beserta total data pada penelitian ini :

Tabel 1. Jumlah Data

No.	Sentimen	Total
1.	Positif	555
2.	Netral	79
3.	Negatif	366
Total		1.000

## Flowchart

Flowchart merupakan gambaran jalannya sebuah program dari satu proses ke proses lainnya. Sehingga, alur dari system menjadi simple dan mudah dipahami oleh semua orang. Selain itu, fungsi lain flowchart adalah menyederhanakan rangkaian prosedur pemahaman terhadap informasi tersebut. Flowchart arsitektur umum penelitian dapat dilihat pada Gambar 1. dibawah ini.



Gambar 2. Flowchart Arsitektur Umum Penelitian

## Analisis Sentimen

*Sentiment analysis* atau analisis sentimen jika diamati berdasarkan Bahasa Indonesia memiliki definisi sebuah Teknik ataupun cara yang memiliki kegunaan untuk bisa mengidentifikasi terkait bagaimana sentiment dapat diekspresikan dengan teks dan juga bagaimana dapat dikategorikan dalam sentiment yang positif dan negatif. Pada hasil *system prototype* telah mendapatkan presisi yang tinggi yakni 75%-95% kembali menyesuaikan dengan data dalam prosesnya untuk menemukan sentiment pada halaman web serta artikel berita [3].

*Sentiment analysis* atau *opinion mining* menganut pada bidang yang luas berdasarkan pengolahan Bahasa yang sesungguhnya, komputasi *linguistic* serta *text mining* yang bertujuan untuk melakukan Analisa terkait opini, sentiment, sikap, evaluasi, serta bagaimana emosi seseorang terkait kesediaan pembicara ataupun penulis terhadap topik, layanan, produk, organisasi, individu dan kegiatan lainnya[4].

Penulis dalam melakukan penelitian kali ini terdiri dari beberapa proses dalam mengelompokkan ulasan pengguna aplikasi VN. Berikut rincian tahapan penulis dalam penelitian ini :

- Tahapan proses diawali dengan splitting dataset menjadi data latih dan data uji yang dibagi secara acak, selanjutnya data akan diteruskan untuk selanjutnya dilakukan preprocessing. Penulis akan membagi kumpulan data menjadi data latih dan data uji untuk berbagai perbandingan, karena banyaknya data latih akan mempengaruhi akurasi. Semakin banyak data pelatihan, semakin banyak model yang dipelajari dan karenanya semakin baik akurasi [8].
- Setelah melalui rangkaian preprocessing, data selanjutnya akan dilakukan pembobotan dengan metode TF-IDF. Data latih akan diproses terlebih dahulu dalam metode SVM untuk melatih algoritma dalam mengenali sentiment.
- Selanjutnya Algoritma Support Vector Machine diterapkan pada data uji untuk mengklasifikasikan dataset menjadi ulasan positif, netral dan negatif serta menghitung akurasi yang didapatkan.

d. Pengujian menggunakan confusion matrix dilakukan untuk memastikan akurasi yang didapat.

## Text Preprocessing

Dalam kegiatan permrosesan teks (*Text Preprocessing*) memiliki definisi yakni sebuah proses dalam merubah suatu tatanan bentuk yang masih belum tertata hingga berubah menjadi lebih tertata dan terstruktur menyesuaikan kebutuhan, dan keterkaitan dengan proses mining selanjutnya (*sentiment analysis*, peringkasan, *clustering* dokumen, etc.). menurut Feldmant *text processing* merupakan sebuah tahap pada proses yang pertama pada teks dengan bertujuan untuk menyiapkan sebuah teks untuk menjadi data yang selanjutnya akan dilakukan pengolahan lagi. Selanutnya teks yang sudah ada terlebih dahulu harus dipersiapkan dan dapat dijalankan dalam tingkatan yang berbeda. Dalam sebuah dokumen bisa dilakukan pemecahan sehingga menjadi bagian-bagian bab, sub-bab, paragraph, kalimat dan pada akhirnya menjadi sebuah potongan kata/token. Berikut ini merupakan tahapan text processing[5]. Berikut tahapan dalam preprocessing :

### a. Case Folding

*Case Folding*, yakni sebuah proses yang bertujuan untuk merubah seluruh karakter pada teks menjadi huruf kecil dan juga menghapus tanda baca dan juga angka. *Case folding* memiliki cara kerja yakni dengan cara melakukan proses pada huruf alphabet dari mulai “a” sampai “z” saja, hal tersebut memunculkan karakter selain yang disebutkan akan dihapus [6].

### b. Cleaning

*Cleaning* merupakan proses dimana dataset yang sebelumnya telah disatukan dari beberapa artikel website yang ada, kedalam satu file excel dengan bertujuan melakukan perbaikan tata Bahasa sehingga dapat sesuai dengan inputan yang bisa diterima oleh sistem yang ada, kemudian selanjutnya akan diubah menjadi format file text yang didalamnya terdapat beberapa atribut penting contohnya yakni opini, topik, serta kelas dalam proses melakukan klasifikasi dan juga peretrievean dokumen selanjutnya.[4].

### c. Tokenizing

*Tokenizing* merupakan sebuah proses dalam memisahkan berbagai kata yang terdapat pada kalimat, paragraph ataupun juga halaman sehingga dapat menjadi token ataupun potongan kata tunggal atau termed word. Pada waktu yang sama, tokenisasi turut serta membuang beberapa karakter yang sebenarnya dianggap memiliki fungsi sebagai tanda baca [6].

### d. Normalisasi

Normalisasi merupakan sebuah proses mengembalikan kata-kata non baku dan diubah menjadi Bahasa baku pada kamus Bahasa Indonesia.

### e. Filtering

*Filtering* yakni sebuah tahapan/proses yang bertujuan mengambil kata-kata kunci dari hasil pada token [7].

### f. Stemming

*Stemming*, yakni proses yang bertujuan memperoleh kata dasar dengan cara menghapus awalan, sisipan, serta akhiran dan juga kombinasinya. Dalam penelitian ini yang dipilih untuk digunakan adalah algoritma stemming yakni *confix striping* yang termasuk dalam hasil pengembangan algoritma nazief and Andriani’s Stemmer [8].

## Term Frequency Inverse Document Frequency ( TF.IDF)

TFIDF merupakan metode yang berjenis integrasi antar *term frequency* (TF), dan *inverse document frequency* (IDF). TF ditentukan dengan cara persamaan dengan term frequency ke-I yakni frekuanesi terkait munculnya term ke-i pada dokumen ke-j.

*Inverse Document Frequency* (IDF) merupakan logaritma yang berasal dari rasio jumlah semua dokumen pada korpus dengan jumlah dokumen yang mempunyai term sesuai dengan yang disebutkan secara matematis pada persamaan. Nilai tersebut dihasilkan dari melakukan perkalian keduanya yang kemudian diformulasikan pada persamaan.

$$tf_{ji} = \frac{freq_i(dj)}{\sum_{i=1}^k freq_i(dj)} \pi r^2 \quad (1)$$

$$idf = \log \frac{|D|}{|\{d:ti \in d\}|} \quad (2)$$

$$(tf - idf)_{ij} = tf_{ji}(dj) * idf_{ji} \quad (3)$$

Fungsi metode TFIDF adalah yakni dalam kaitannya menemukan representasi nilai pada setiap dokumen yang berasal dari kumpulan data training, kemudian selanjutnya dibentuk sebuah vector antara dokumen dan kata dan dilanjutkan dengan menentukan kesamaan antar keduanya oleh protorype vector yang juga disebut dengan cluster centroid [8].

## Support Vector Machine

“*Support Vector Machine*” (SVM) merupakan algoritma dalam kegiatan pembelajaran mesin laser yang dapat dimanfaatkan terhadap kegiatan klasifikasi ataupun permintaan regresi. Namun, sebagian besar digunakan dalam masalah klasifikasi. Algoritma SVM memplot setiap elemen data sebagai titik ruang penyimpanan NDIMENSI (di mana n adalah nilai dari setiap fungsi dari nilai koordinat tertentu) dengan nilai masing-masing fungsi. Kemudian kami diklasifikasikan dengan menemukan pesawat super, yang membedakan dua kelas dengan sangat baik.

Dalam hal ini penulis menggunakan Metode One-Against-One untuk menentukan banyaknya model klasifikasi biner. Tahapan pengerjaan SVM Multiclass dengan menggunakan metode One-Against-One diantaranya :

- Membangun “k” buah model SVM biner dengan k adalah jumlah kelas. Dibangun  $k(k-1)/2$  buah model klasifikasi biner. Tiap-tiap model klasifikasi akan dilatih pada data dari dua kelas (klasifikasi biner). Pada penelitian ini dilakukan klasifikasi untuk 3 kelas sehingga didapat total jumlah model klasifikasi biner yang dilakukan adalah:  $3(3-1)/2 = 3$  kali
- Menentukan Model SVM biner yang akan digunakan untuk pelatihan. Berikut disajikan tabel yang menunjukkan penentuan kelas sentimen dengan menggunakan algoritma SVM multiclass dengan metode One Against One.

Tabel 2. Penentuan Class

$Y_i = 1$	$Y_i = -1$	Hipotesis
Kelas 1	Kelas 2	$f_{12}(x) = (w_{12})x + b_{12}$
Kelas 1	Kelas 3	$f_{13}(x) = (w_{13})x + b_{13}$
Kelas 2	Kelas 3	$f_{23}(x) = (w_{23})x + b_{23}$

- Penelitian yang akan dilakukan adalah mengimplementasikan metode Support Vector Machine menggunakan Kernel Radial Basis Function, Linear, Sigmoid, Polinomial. Untuk mengetahui hasil

keakuratan algoritma, akan dilakukan pengujian dengan menggunakan *Confusion Matrix* dengan menggunakan nilai *precision*, *recall*, dan *f1-score*.

## Metode Pengujian

Jika sistem menjalankan *classifier*, tentu diharapkan dan diprediksi dengan tepat. Namun hal tersebut tetap tidak menjamin jika penerapan sistem disarankan tidak boleh 100 nar. Ini membutuhkan pengukuran kinerja sistem untuk mengeksplorasi berbagai aspek pengujian.

### a. Accuracy

*Accuracy* adalah sebuah metode dalam pengujian dengan berdasar pada tingkat kedekatan diantara nilai prediksi dengan nilai actual. Untuk mengetahui akurasi hasil dari prediksi, maka perlu untuk mengetahui terlebih dahulu mengenai jumlah data set yang akan dilakukan klasifikasi secara tepat. Berikut merupakan persamaan akurasi :

$$Accuracy = \frac{N \text{ benar}}{N} \times 100\% \quad (4)$$

### b. Precision

*Precision* adalah metode dengan bertujuan melakukan pengujian yang melakukan perbandingan terkait jumlah terhadap data informasi yang relevan serta diterima dari sistem dan turut memperhatikan jumlah total informasi yang diperoleh dari sistem, apakah relevan atau tidak. Rumus presisi terlihat seperti ini:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

### c. Recall

*Recall* adalah metode yang bertujuan melakukan pengujian dengan cara melaksanakan perbandingan terhadap jumlah pada informasi relevan yang diterima dari sistem dengan jumlah total informasi relevan pada kumpulan sebuah informasi (termasuk yang dipertahankan maupun yang disimpan sistem. Berikut merupakan rumus penarikan :

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

## III. Hasil dan Pembahasan

Proses dimulai dengan langkah pra-pemrosesan, karena data tidak menggunakan kata-kata standar yang sempurna. Jalankan fase prapemrosesan Bantuan dengan penggunaan perpustakaan dalam bahasa pemrograman Python3. Kembali Praproses data dan simpan hasil praproses sebagai File baru, yang akan digunakan sebagai kumpulan data nanti dalam proses klasifikasi. Hasil prapemrosesan ditunjukkan pada gambar 4.

	userName	score	at	content	Case_folded	Tokenized	Stemmed	No_Stop
0	Putra Subandi	5	12/11/2021 19:13	APLIKASINYA MUDAH DAN BUAGUS BANGETTTTT 🥰🥰	aplikasinya mudah dan buagus bangetttt	[aplikasinya, mudah, dan, buagus, bangetttt]	[aplikasi, mudah, dan, buagus, bangetttt]	[aplikasi, mudah, buagus, bangetttt]
1	Eli Eli tarkim	5	12/11/2021 17:37	Saya sngat suka	saya sngat suka	[saya, sngat, suka]	[saya, sngat, suka]	[sngat, suka]
2	Sinta Aisyah	5	12/11/2021 17:12	Ini aplikasi keren banget, tp kenapa ya skrg v...	ini aplikasi keren banget tp kenapa ya skrg vi...	[ini, aplikasi, keren, banget, tp, kenapa, ya, skrg, vi...]	[ini, aplikasi, keren, banget, tp, kenapa, ya, skrg, video, ...]	[aplikasi, keren, banget, tp, ya, skrg, video, ...]
3	Deni Achmad	5	12/11/2021 16:14	Apk bagus banget aku suka aku suka	apk bagus banget aku suka aku suka	[apk, bagus, banget, aku, suka, aku, suka]	[apk, bagus, banget, aku, suka, aku, suka]	[apk, bagus, banget, suka, suka]
4	intans yoo	5	12/11/2021 16:05	Baguuusss bangettt aplikasinya... Semoga kede...	baguuusss bangettt aplikasinya... semoga kedepan...	[baguuusss, bangettt, aplikasinya, semoga, ke...]	[baguuusss, bangettt, aplikasinya, moga, depa...]	[baguuusss, bangettt, aplikasinya, moga, grat...]

Gambar 3. Hasil Preprocessing

## Penerapan Algoritma Support Vector Machine

Pada metode SVM, terdapat beberapa kernel yang digunakan, seperti kernel Linear, Polynomial, Radial Basis Function (RBF), dan Sigmoid. Dari setiap kernel akan dicari akurasi terbaik yang nantinya akan digunakan. Pada penelitian kali ini akan dilakukan pembagian dataset sebanyak 3 kali yang selanjutnya akan diterapkan metode SVM pada tiap kernelnya. Perbandingan dataset dapat dilihat pada tabel berikut :

- a. Perbandingan data *training* dan data *testing* sebesar 70% : 30% :

Tabel 3. Perbandingan 70% : 30% pada Kernel SVM

Kernel	Akurasi
Linear	83,80%
Polynomial	77,40%
Radial Basis Function (RBF)	88,40%
Sigmoid	82,00%

- b. Perbandingan data *training* dan data *testing* sebesar 80% : 20% :

Tabel 4. Perbandingan 80% : 20% pada Kernel SVM

Kernel	Akurasi
Linear	81,68%
Polynomial	75,97%
Radial Basis Function (RBF)	87,08%
Sigmoid	79,27%

- c. Perbandingan data *training* dan data *testing* sebesar 90% : 10% :

Tabel 5. Perbandingan 90% : 10% pada Kernel SVM

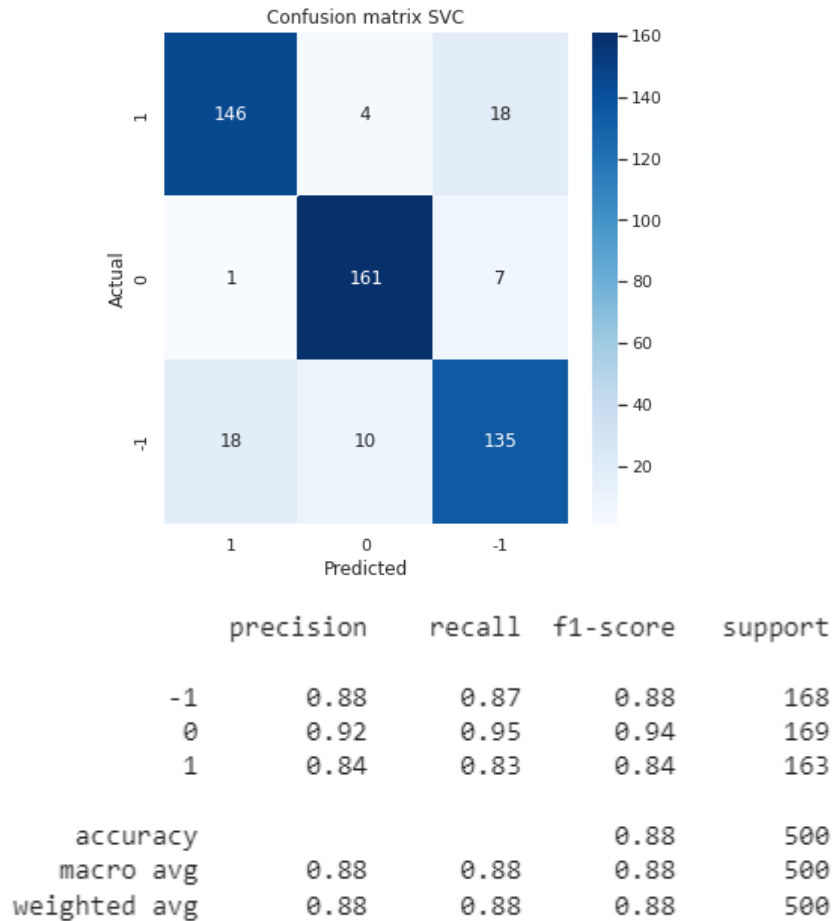
Kernel	Akurasi
Linear	80,23%
Polynomial	73,05%
Radial Basis Function (RBF)	87,42%
Sigmoid	78,44%

Dapat kita lihat dari Tabel 2 terlihat bahwa kernel Radial Basis Function (RBF) pada perbandingan data *training* dan data *testing* 70% : 30% memiliki akurasi paling tinggi dibandingkan dengan kernel yang lainnya yakni mampu menghasilkan 88,40%.

## Tahapan Pengujian

Peringkat digunakan untuk memvisualisasikan tolak ukur tentang kinerja algoritma Support Vector Machine (SVM) saat runtime analisis sentiment metode evaluasi dilakukan dengan confusion matrix, yang akan membuat kesimpulan tentang nilai presisi, f-score, recall terhadap tiap kategori sentiment dan nilai akurasi keseluruhan. Hasil dari confusion Matrix pada perbandingan 70%:30% ditunjukkan pada gambar 5.



Gambar 4. Hasil *Confusion Matrix*

Berdasarkan evaluasi dengan menggunakan metode *Confusion Matrix*, didapatkan nilai presisi untuk sentimen negatif(-1), netral (0) dan positif(1) masing-masing 88%, 92%, dan 84%. Hasil recall sentimen negatif(-1), netral(0) dan positif(1) sebesar 87%, 95%, dan 83%. Sedangkan untuk nilai dari *f1-score* untuk sentimen negatif(-1), netral(0), dan positif(1) berturut-turut sebesar 88%, 94%, dan 84%. Melalui evaluasi dengan menggunakan metode *Confusion Matrix* diatas, algoritma SVM mampu memperoleh akurasi sebesar 88%.

#### IV. KESIMPULAN

Berdasarkan hasil evaluasi, penggunaan kernel Radial Basis Function (RBF) di penelitian ini memberikan akurasi yang paling tinggi dibandingkan dengan kernel lainnya. Dengan demikian, penerapan algoritma support vector machine pada klasifikasi review video editor dengan total 1.000 data dapat memberikan hasil yang cukup akurat yaitu mencapai total 88,40% pada perbandingan dataset 70% : 30%.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih yang sebesar besarnya kepada pihak-pihak yang berkontribusi dalam penelitian ini. Terimakasih kami sampaikan kepada :

1. Laboratorium Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Sidoarjo.
2. Bapak/Ibu dosen Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Sidoarjo

## REFERENSI

- [1] Purmadi, A. (2021). Pemanfaatan Aplikasi VN Untuk Pembuatan Video Pembelajaran Menggunakan Telepon Pintar Kepada Guru PAUD. *Jurnal Pengabdian Masyarakat*, Vol1 No1.
- [2] Anjasmoros, M. T. (2020). ANALISIS SENTIMEN APLIKASI GO-JEK MENGGUNAKAN METODE SVM DAN NBC (STUDI KASUS: KOMENTAR PADA PLAY STORE). *CIASTECH 2020*.
- [3] Maulana, Y. (2018). Implementasi Text Mining Pada Penilaian Otomatis Ujian Jawaban Esai Dengan Metode Hirarki Clustering. *Jurnal Universitas Dian Nuswantoro Semarang*.
- [4] Syarifuddin, M. (2020). ANALISIS SENTIMEN OPINI PUBLIK TERHADAP EFEK PSBB PADA TWITTER DENGAN ALGORITMA DECISION TREE-KNN-NAÏVE BAYES. Malang: PPPM Nusa Mandiri.
- [5] Latius Hermawan, M. B. (2020). Pembelajaran Text Preprocessing berbasis Simulator. *journals.usm.ac.id/index.php/transformatika*, Vol.17, No.2,.
- [6] Muhammad, S. (2020). ANALISIS SENTIMEN OPINI PUBLIK MENGENAI COVID-19 PADA TWITTER MENGGUNAKAN METODE NAÏVE BAYES DAN KNN. *INTI NUSA MANDIRI*, VOL. 15. NO. 1.
- [7] Dyah Auliya Agustina, S. S. (2020). Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine . *Indonesian Journal of Applied Statistics* , Volume 3 No. 2.
- [8] Nelly Indriani, E. R. (2017). Peringkasan dan Support Vector Machine pada Klasifikasi. *JURNAL INFOTEL*, Vol.9 No.4 .
- [9] Rijallahudin, R. K. (2016). Pengaruh Penggunaan Layanan Aplikasi Digital Google Play. *Jurnal Komunikatio ISSN 2442-3882*, Volume 2 Nomor 2.
- [10] Fonda. (2020) IMPLEMENTASI TEXT MINING PADA TWITTER DENGAN ALGORITMA K-MEANS CLUSTERING SEBAGAI DASAR KEBIJAKAN MARKETING BIRO PERJALANAN WISATA. *JIK*. 2020;9 (2): 138 - 147
- [11] Budi Ilmawan (2020) Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store Vol. 12 No. 2, Agustus 2020, pp.154-161
- [12] Meisya Permata, (2021) PERBANDINGAN KERNEL SUPPORT VECTOR MACHINE (SVM) DALAM PENERAPAN ANALISIS SENTIMEN VAKSINISASI COVID-19, ISSN 2598-7305 | E-ISSN 2598-9642 Vol. 4 No 2
- [13] Siti Masripah, L. D. (2020). Algoritma Klasifikasi Naïve Bayes untuk Analisa Sentimen Aplikasi Shopee . *JURNAL SWABUMI*, Vol.8 No.2.
- [14] Lubis, A. A. (2020). Prediksi Akurasi Perusahaan Saham Menggunakan SVM dan. *Jurnal SIFO Mikroskil*, VOL 21, NO 1.
- [15] Rahmawati Hakim, (2020, Analisis Sentimen Pengguna Instagram Terhadap Kebijakan Kemdikbud Mengenai Bantuan Kuota Internet dengan Metode Support Vector Machine (SVM), Vol. 8 No. 2 Ed. Juli - Des. 2020.

### **Conflict of Interest Statement:**

*The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.*