

Analisis Sentimen Prosesor AMD Ryzen menggunakan Metode *Support Vector Machine*

Erlin¹, Josef Sianturi², Alyauma Hajjah³, Agustin⁴

¹ Institut Bisnis dan Teknologi Pelita Indonesia, erlin@lecturer.pelitaindonesia.ac.id, Pekanbaru, Indonesia

² STMik Amik Riau, josef4sianturi@gmail.com, Pekanbaru, Indonesia

³ Institut Bisnis dan Teknologi Pelita Indonesia, alyauma.hajjah@lecturer.pelitaindonesia.ac.id, Pekanbaru, Indonesia

⁴ STMik Amik Riau, agustin@sar.ac.id, Pekanbaru, Indonesia

Informasi Makalah

Submit : Desember 26, 2021

Revisi : Desember 27, 2021

Diterima : Desember 28, 2021

Kata Kunci :

Analisis Sentimen

Support Vector Machine

AMD Ryzen

Akurasi

Twitter

Confusion Matrix

Abstrak

Prosesor AMD menjadi salah satu pesaing prosesor Intel semenjak dikeluarkannya prosesor Ryzen generasi 3. Berbagai pendapat dan opini masyarakat mengenai prosesor ini sangat mudah ditemui pada media sosial *Twitter*. Opini ini dapat digunakan sebagai sistem untuk mendukung keputusan berkaitan dengan produk AMD Ryzen. Tujuan penelitian ini adalah untuk mengimplementasikan Analisis Sentimen dalam pendekatan *Data Mining* untuk menganalisa tekstual data yang terdapat pada *Twitter* menggunakan metode *Support Vector Machine*, mengeksplorasi dan memahami tren opini publik mengenai prosesor AMD Ryzen dan mengklasifikasikannya kedalam polaritas biner. Penelitian ini menggunakan *Tweet* dari *Library Tweepscraper*. Pelabelan dilakukan oleh *expert* untuk diklasifikasikan menjadi sentimen positif dan negatif. Selanjutnya melakukan pra-pemrosesan data untuk menghilangkan *noise*, mendeteksi nilai data yang hilang, data duplikat dan tidak relevan. Selanjutnya, algoritma *machine learning* digunakan untuk memprediksi data baru. Model yang dihasilkan dievaluasi menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa kinerja metode SVM sangat baik dalam hal akurasi, presisi, *recall* dan *F1 Score* dengan nilai masing-masing 96,67%, 96,43%, 100% dan 98,18%. Berdasarkan hasil yang diperoleh, sebagian besar publik memiliki opini yang positif terhadap prosesor AMD Ryzen. Penelitian ini juga membuktikan bahwa metode *Support Vector Machine* dapat digunakan sebagai algoritma cerdas untuk memprediksi sentimen di *Twitter* untuk data baru dengan cepat dan akurat.

Abstract

AMD processors have become one of the competitors for Intel processors since the company released the 3rd generation Ryzen processors. Various public opinions about this processor are easy to find on Twitter social media. This opinion can be used as a system to support decisions regarding AMD Ryzen

Erlin,

Email:erlin@lecturer.pelitaindonesia.ac.id

products. This study aims to employ Sentiment Analysis in the Data Mining approach to analyze textual data contained on Twitter using the Support Vector Machine method, explore and understand public opinion trends regarding AMD Ryzen processors and classify them into binary polarity. This research used Tweets from the Tweetscrapper Library. Labeling using the expert to be classified into positive and negative sentiments. Furthermore, the data preprocessing was conducted to remove noise, detect missing data values, duplicate and irrelevant data. Machine learning algorithms are applied to predict new data. The created model is evaluated using a confusion matrix. The results showed that the performance of the SVM method was highest in terms of accuracy, precision, recall, and F1 Score the with values of 87%, 86%, 87%, and 90%, respectively. Based on the results obtained, the most public has positive sentiments of AMD Ryzen processors. This research also proves that the Support Vector Machine method can be used as an intelligent algorithm to predict sentiment on Twitter for new data quickly and accurately.

1. Pendahuluan

Prosesor sering juga disebut sebagai unit pemrosesan merupakan komponen vital dalam sebuah perangkat komputer yang mengontrol semua bagian lain dari sistem dan menjalankan operasi di beberapa pangkalan data eksternal, normalnya pada memori dan beberapa aliran data lainnya (Dictionary, 2021). Pemakaian prosesor umumnya digunakan pada perangkat elektronik seperti smartphone, PC, tablet, dan piranti modern lainnya. Fungsinya untuk menampung masukan dalam format intruksi program serta menjalankan milyaran bahkan triliunan perhitungan untuk menghasilkan keluaran yang akan dipakai oleh pengguna (Towkeersah40, 2021). Kemajuan perkembangan prosesor sangat cepat sekali mengikuti hukum Moore, dimana jumlah transistor dalam prosesor akan berlipat ganda setiap dua tahun.

Intel dan AMD merupakan 2 (dua) merek prosesor yang menguasai pasar saat ini. Intel sudah menguasai pasar untuk waktu yang cukup lama, tetapi telah terjadi persaingan ketat sejak AMD mengeluarkan prosesor Ryzen generasi ke-3, sehingga sering menjadi perdebatan kelebihan dan kekurangan antara prosesor Intel Core i3 dan AMD Ryzen 3 sampai Intel Core i9 dan AMD Ryzen 9. Sebuah hasil review dari media online teknologi terkemuka mengklaim bahwa AMD Ryzen memiliki performa yang sangat baik

dengan harganya yang sangat kompetitif. AMD Ryzen terbukti sangat bagus untuk bermain game, mengedit video dan semua kebutuhan komputasi lainnya (Uy & Thomas, 2021).

Seiring dengan meningkatnya jumlah pengguna produk AMD Ryzen, maka semakin banyak pula ditemukan konsumen yang memberikan ulasan, testimoni dan pengalaman secara online. Kehadiran media sosial salah satunya *Twitter* sudah menjadi alat bantu yang banyak digunakan oleh masyarakat dari segala lapisan untuk mencari informasi dan sumber pengetahuan lainnya. Sebagian besar user menggunakan media ini untuk mencari hasil review terhadap produk tertentu maupun ulasan dalam bentuk forum diskusi yang akan digunakan dalam membuat sebuah keputusan pembelian. Review atau penilaian yang dilakukan oleh user biasanya dalam bentuk teks berupa ulasan atau komentar yang berisi pujian dan kepuasan yang terkategori penilaian positif atau sebaliknya keluhan, cacian dan ketidakpuasan yang terkategori penilaian negatif.

Pada sosial media *Twitter*, publik sudah terbiasa berbagi pendapat, pandangan, pikiran atau opini terhadap suatu produk yang sedang hangat dibicarakan. Kebiasaan konsumen memuat postingan atau *retweet* untuk menilai sebuah produk adalah salah satu cara untuk mengekspresikan tanggapan konsumen terhadap produk tersebut. Melalui *tweet* yang disampaikan oleh pengguna, publik dapat

melakukan penggalian informasi mengenai gambaran sentimen atau opini masyarakat terhadap kualitas produk yang ada. Salah satu teknik penggalian informasi pada *Twitter* adalah analisis sentimen.

Analisis sentimen adalah salah satu topik penelitian paling aktif dalam *data mining* dan pemrosesan bahasa alami. Analisis sentimen disebut juga penggalian opini (*opinion mining*), yang menganalisis sentimen, penilaian, perspektif, dan emosi orang terhadap objek dan atributnya yang dinyatakan dalam teks tertulis (B. Liu, 2015). Ketersediaan data opini dan sentimen dalam jumlah yang besar pada sosial media telah menyebabkan penerapan analisis sentimen menjadi lebih luas. Hal ini tidaklah mengherankan karena saat ini, setiap orang menyampaikan pendapat yang diungkapkan secara bebas di sosial media yang tersedia.

Banyak literatur tentang analisis sentimen memberikan perhatian yang signifikan untuk mengklasifikasikan ulasan produk dan penjualan (Stephenie et al., 2020; Yang et al., 2020; Singh & Sarraf, 2020) dengan hasil yang baik pada domain aplikasi ini (Hariguna et al., 2019; Mohd Nafis & Awang, 2021). Beberapa studi fokus pada analisis sentimen ulasan film menggunakan *machine learning* (Chakraborty et al., 2019; Nurdiansyah et al., 2018) dengan hasil yang sangat baik (Mamtesh & Mehla, 2019; Daeli & Adiwijaya, 2020). Selain itu, peneliti lain fokus pada penggunaan analisis sentimen untuk berbagai keperluan seperti penelitian yang dilakukan oleh (Fitri, 2020) tentang analisis sentimen terhadap aplikasi Ruangguru, (Fitriyah et al., 2020) dan (Wahyudi & Kusumawardana, 2021) membahas mengenai analisis sentimen gojek dan grab pada media sosial twitter, (Arsi & Waluyo, 2021) menggunakan metode SVM untuk menganalisis wacana pemindahan ibu kota Indonesia dan (Putra et al., 2020) yang meneliti mengenai implementasi metode Naïve Bayes dalam melakukan analisis sentimen dompet elektronik pada *Twitter*

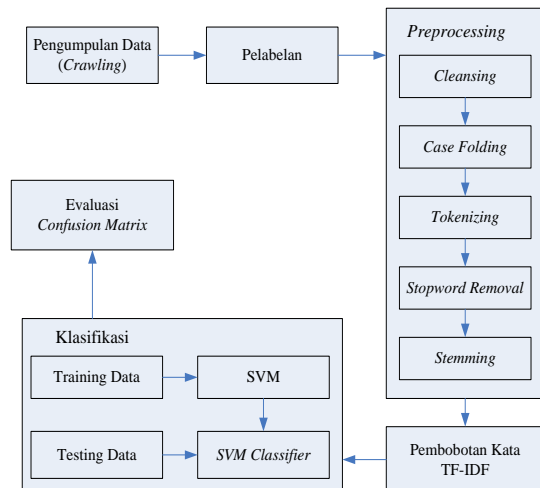
dengan rata-rata kinerja dari model pengklasifikasi yang sangat baik.

Support Vector Machine (SVM) adalah metode untuk masalah klasifikasi yang cocok untuk berbagai aplikasi dan dapat menangani masalah linier dan nonlinier. Metode SVM memberikan tingkat kinerja yang cukup tinggi dan sering digunakan dibandingkan dengan pohon keputusan (Suresh & Bharathi, 2016), Naïve Bayes (Isnain et al., 2021), dan pengklasifikasi model lainnya (N. Liu & Shen, 2020). Selain itu, SVM memiliki tingkat generalisasi yang tinggi dan juga terkenal karena ketahanannya terhadap data berdimensi tinggi (Prastyo et al., 2020).

Berdasarkan berbagai kelebihan dari SVM, penelitian ini menggunakan metode ini untuk mengklasifikasikan opini dan sentimen publik dalam teks tertulis menjadi klasifikasi sentimen binary "positif" dan "negatif", serta menunjukkan kemampuan metode yang diusulkan dalam hal kinerja, baik dari sisi akurasi, presisi, recall, dan F1 Score. Sejauh pengetahuan penulis, penelitian yang berfokus pada analisis sentimen terhadap review produk AMD Ryzen belum pernah dilakukan sebelumnya terlepas dari berbagai publikasi yang berhubungan dengan data multimodal dan analisis sentimen berbasis teks. Hasil penelitian ini memperlihatkan nilai akurasi, presisi, recall dan F1 score masing-masing 96,67%, 96,43%, 100%, dan 98,18%.

2. Metode Penelitian

Penelitian ini menggunakan metode *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen/opini publik mengenai produk prosesor AMD Ryzen. Proses klasifikasi *polarity* dari teks memerlukan serangkaian proses seperti diperlihatkan pada Gambar 1. Tahapan pertama adalah pengumpulan data melalui *web crawler* menggunakan *tool RapidMiner* dan tahapan terakhir adalah testing dan evaluasi model SVM menggunakan *Confusion Matrix*.



Gambar 1. Rancangan penelitian analisis sentimen untuk klasifikasi teks

2.1. Pengumpulan data dan Pelabelan

Data yang digunakan pada penelitian ini diambil dari kumpulan *tweets* bahasa Indonesia yaitu *tweets mention* yang berhubungan dengan AMD Ryzen melalui twitter search API menggunakan operator “Search Twitter” dengan jumlah data yang dibatasi dalam bahasa Indonesia dengan *query* “AMD Ryzen”. Jumlah data sebanyak 100 dibagi kedalam dua kelompok yaitu data training dan data testing dengan perbandingan 70:30. Tahap selanjutnya adalah memberikan label secara manual. Pelabelan dilakukan untuk mengelompokkan data ke dalam dua kelas sentimen yang akan digunakan yaitu positif dan negatif. Tujuan dari proses pelabelan ini untuk melatih *machine learning* mengelompokkan data berdasarkan kelas positif dan negatif.

2.2. Preprocessing

Prapemrosesan (*preprocessing*) data adalah langkah awal dan terpenting dalam membangun model *machine learning* untuk menyiapkan data mentah (*raw data*). Sebelum melakukan prosedur terkait data, maka penting untuk membersihkan dan memformat data karena umumnya *raw data* mengandung *noise*, nilai yang hilang, dan format yang tidak sesuai yang tidak dapat digunakan untuk model *machine learning* secara langsung. Oleh karena itu, prapemrosesan data

diperlukan untuk membersihkan dan membuat data sesuai untuk model pembelajaran mesin (*machine learning*), yang juga meningkatkan akurasi klasifikasi dan kualitas data. Data yang sesuai atau terstruktur membuat proses klasifikasi analisis sentimen menjadi lebih mudah dan sederhana. Tahapan ini melibatkan beberapa langkah termasuk *cleansing*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*.

2.3. Ekstraksi Fitur menggunakan TF-IDF

Algoritma *Machine Learning* tidak dapat bekerja langsung pada teks mentah oleh karena itu penelitian ini menggunakan *term frequency-inverse document frequency* (TF-IDF) untuk mengubah teks menjadi fitur matriks (atau vektor). Nilai TF-IDF akan naik sebanding dengan berapa kali sebuah kata muncul dalam dokumen dan akan turun sebanding dengan ukuran dataset. TF-IDF adalah hasil perkalian TF dan IDF dan diformulasikan dalam persamaan (1), $tf(t, d)$ dalam persamaan (2), dan $idf(t, D)$ pada persamaan (3).

$$tf - idf(t, d, D) = tf(t, d).idf(t, D) \quad (1)$$

dimana:

$$tf(t, d) = \log(1 + freq(t, d)) \quad (2)$$

$$idf(t, D) = \log\left(\frac{N}{count(d \in D) : t \in d}\right) \quad (3)$$

$tf(t, d)$ adalah frekuensi kata (*term*) yang menentukan seberapa sering suatu kata (t) muncul di seluruh dokumen (d) dan, $idf(t, D)$ adalah *inverse document frequency* dari kata (t). N adalah jumlah total dokumen dalam dataset, df adalah jumlah dokumen yang mengandung term/kata t .

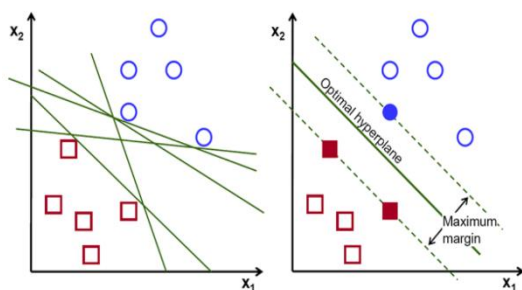
2.4. Model Support Vector Machine

Setelah tahapan *preprocessing*, langkah selanjutnya adalah pemodelan menggunakan SVM. Pengklasifikasi teks seperti SVM perlu dilatih sebelum digunakan untuk kategorisasi teks (Erlin et al., 2013; Erlin et al., 2014). Pelatihan data (*training data*) akan menghasilkan sebuah nilai atau pola yang akan digunakan pada proses pengujian (*testing data*). Penilaian kemudian dibuat

dengan menilai *score* yang mempresentasikan disisi mana dokumen tersebut diklasifikasikan.

Support Vector Machine (SVM) termasuk kedalam kelompok *supervised machine learning* (algoritma pembelajaran mesin terawasi) yang dapat digunakan untuk klasifikasi atau regresi. Namun, metode SVM lebih banyak diimplementasikan untuk masalah klasifikasi. Pada algoritma SVM, setiap item data perlu diplot sebagai titik dalam ruang n-dimensi (di mana n adalah sejumlah fitur yang dimiliki) dengan nilai setiap fitur menjadi nilai koordinat tertentu. Selanjutnya, menjalankan proses klasifikasi dengan menentukan *hyperplane* yang akan membedakan kedua kelas dengan sangat baik. SVM sangat cocok dalam ruang berdimensi tinggi, termasuk dalam kasus di mana jumlah dimensi lebih besar dari jumlah sampel. SVM juga bisa menghemat memori dengan menjadikan subset titik pelatihan dalam fungsi keputusan (scikit-learn, 2021).

Terdapat banyak kemungkinan *hyperplane* yang dapat dipilih untuk memisahkan kedua kelas titik data tersebut seperti diperlihatkan pada Gambar 2. Tujuannya adalah menemukan bidang yang memiliki margin maksimum, yaitu jarak maksimum antara titik data dari kedua kelas. Memaksimalkan jarak margin memberikan beberapa penguatan sehingga titik data berikutnya dapat diklasifikasikan dengan lebih baik (Gandhi, 2018).



Gambar 2. Kemungkinan *hyperplane* (Gandhi, 2018)

2.5. Evaluasi Model SVM

Matrik evaluasi model yang digunakan dalam penelitian ini adalah akurasi, presisi, recall, dan F1 score, yang sejalan dengan pengukuran evaluasi yang digunakan dalam penelitian lainnya. Keakuratan pemodelan pengklasifikasi dihitung dengan menggunakan *confusion matrix* yang proses kerjanya diperoleh dengan membandingkan label dari data uji yang telah ditentukan dengan prediksi label dari data uji yang dikelompokkan. Hasil dari penerapan teknik evaluasi ini adalah untuk mengetahui kesalahan nilai yang terjadi selama proses klasifikasi.

Tabel 1 adalah *confusion matrix* untuk klasifikasi Biner dengan ukuran 2 x 2 yang digunakan untuk mempresentasikan kelas aktual dan kelas prediksi.

Tabel 1. Confusion matrix untuk klasifikasi biner

Kelas Aktual	Kelas Prediksi	
	Positif	Negatif
Positif	tp	fp
Negatif	fn	tn

Beberapa pengukuran dalam bidang *machine learning dan information retrieval* sudah digunakan berdasarkan klasifikasi dari *confusion matrix* sebagaimana terdapat dalam persamaan (3-6).

$$Accuracy = \frac{tp+tn}{tp+fp+fn+tn} \quad (3)$$

$$Precision = \frac{tp}{tp+fp} \quad (4)$$

$$Recall/Sensitivity = \frac{tp}{tp+fn} \quad (5)$$

$$F1\ Score = \frac{2.(recall.precision)}{(recall+precision)} \quad (6)$$

Dalam persamaan (3-6), tp (*true positive*) adalah jumlah opini yang diprediksi positif dan aktualnya memang positif, tn (*true negative*) adalah jumlah opini yang diprediksi negatif dan aktualnya memang negatif, fp (*false positive*) adalah jumlah opini yang diprediksi positif sedangkan opini tersebut adalah positif, dan fn (*false negative*) adalah jumlah opini yang diprediksi positif sementara opini tersebut adalah negatif.

Tabel 2. *Output Cleansing*

Kode	Input Tweet Asli	Output Cleansing
D1	gunakan Processor Ryzen, bagus dan harganya Murah di @Tokokomputer #RyzenOnFire	gunakan Processor Ryzen bagus dan harganya Murah di
D2	CPU “berkelas dengan harga terjangkau” #RyzenOnFire.	CPU berkelas dengan harga terjangkau
D3	Ryzen CPU yang tidak terbilang murah, dengan performa buruk.	Ryzen CPU yang tidak terbilang murah dengan performa butuk
D4	Ryzen jelek lebih baik Intel http://.asnaskjf.avjavja	Ryzen jelek lebih baik Intel
D5	Processor AMD Ryzen lumayanlah untuk digunakan @jasgkasj	Processor AMD Ryzen lumayanlah untuk digunakan

Selanjutnya, fase *case folding* diperlihatkan pada Gambar 7. Proses mengubah dokumen dari huruf besar menjadi huruf kecil, menggunakan operator *transform case*.



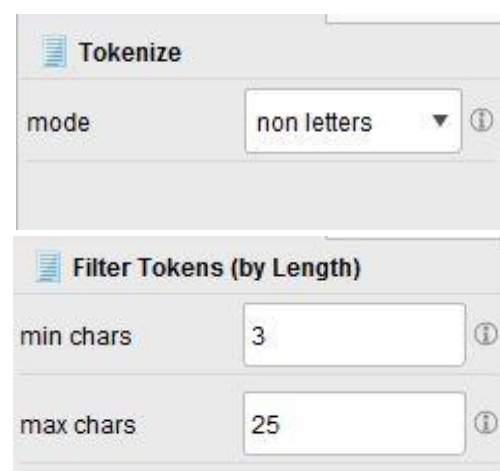
Gambar 7. Proses *case folding* dilakukan dengan operator *transform cases*

Pada penyeragaman bentuk huruf dalam dokumen, huruf kapital diubah menjadi huruf kecil, dan huruf seragam dari A sampai Z, dan juga menghilangkan tanda baca dan simbol yang tidak perlu seperti diperlihatkan pada Tabel 3 yang merupakan output dari proses *case folding*.

Tabel 3. *Output proses Case Folding*

Kode	Input Cleansing	Output Case Folding
D1	gunakan Processor Ryzen bagus dan harganya Murah di	gunakan processor ryzen bagus dan harganya Murah di
D2	CPU berkelas dengan harga terjangkau	cpu berkelas dengan harga terjangkau
D3	Ryzen CPU yang tidak terbilang murah dengan performa butuk	ryzen cpu yang tidak terbilang murah dengan performa butuk
D4	Ryzen jelek lebih baik Intel	ryzen jelek lebih baik intel
D5	Processor AMD Ryzen lumayanlah untuk digunakan	processor amd ryzen lumayanlah untuk digunakan

Tokenizing diasosiasikan dengan pemecahan kalimat menjadi beberapa kata atau pemecahan urutan *string* menjadi potongan-potongan yang disebut *token*. Dalam proses *tokenizing*, spasi berfungsi sebagai pembatas. Gambar 8 memperlihatkan proses *tokenizing* sedangkan Tabel 4 merupakan hasil proses *tokenizing*.



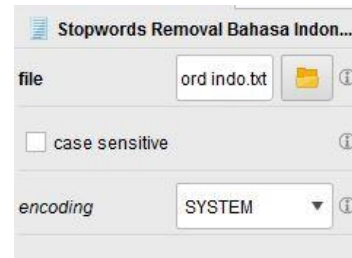
Gambar 8. Proses *tokenizing*

Tabel 4. Hasil Proses *Tokenizing*

Kode	Input <i>Case Folding</i>	Output <i>Tokenizing</i>
D1	gunakan prosessor ryzen bagus dan harganya Murah di	gunakan processor ryzen bagus dan harganya murah di
D2	cpu berkelas dengan harga terjangkau	cpu berkelas dengan harga terjangkau
D3	ryzen cpu yang tidak terbilang murah dengan performa buruk	ryzen cpu yang tidak terbilang murah dengan performa buruk
D4	ryzen jelek lebih baik intel	ryzen jelek lebih baik intel
D5	processor amd ryzen lumayanlah untuk digunakan	processor amd ryzen lumayanlah untuk digunakan

Stopword removal adalah proses menghilangkan kata-kata dari kumpulan data yang tidak menambahkan banyak arti pada sebuah kalimat atau tidak menyampaikan pesan yang signifikan. Proses *stopword removal* diperlihatkan pada Gambar 9. Sebuah *stoplist* dibuat untuk melacak *stopword* yang akan digunakan. Dalam dataset, setiap kata atau *token* akan dibandingkan dengan *stoplist*,

dan setiap kata yang cocok dengan *stopword* di *stoplist* akan dihilangkan.



Gambar 9. Proses *stopword removal*

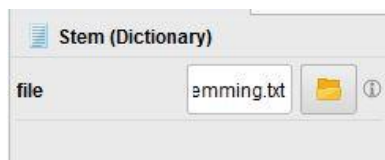
Tabel 5 memperlihatkan hasil *stopword removal* dengan menghilangkan kata-kata yang terdapat dalam *stoplist*.

Tabel 5. Hasil *Stopword Removal*

Kode	Input <i>Tokenizing</i>	Output <i>Stopword Removal</i>
D1	gunakan processor ryzen bagus dan harganya murah di	[gunakan, processor, ryzen, bagus, harga, murah]
D2	cpu berkelas dengan harga terjangkau	[cpu, berkelas, harga, terjangkau]
D3	ryzen cpu yang tidak terbilang murah dengan performa buruk	[ryzen, cpu, tidak, terbilang, murah, performa, buruk]
D4	ryzen jelek lebih baik intel	[ryzen, jelek, lebih, baik, intel]
D5	processor amd ryzen lumayanlah	[processor, amd, ryzen, lumayanlah]

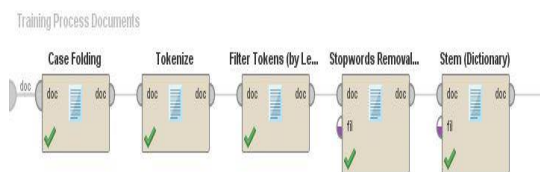
untuk digunakan

Stemming adalah langkah terakhir dari *preprocessing*, yang pada dasarnya menghilangkan prefiks, sufiks, sisipan, dan kombinasi prefiks dan sufiks dari sebuah kata dan mereduksinya menjadi kata dasar atau akar kata. Gambar 10 memperlihatkan proses stemming dengan operator stem dictionary yang akan mengembalikan kata-kata berimbuhan kedalam bentuk dasar atau akar kata tersebut.



Gambar 10. Proses *stemming* dengan operator *stem dictionary*

Gambar 11 merupakan *screenshot* dari *Process Document* dari tahap *preprocessing case folding, tokenizing, stopword removal* dan *stemming*.



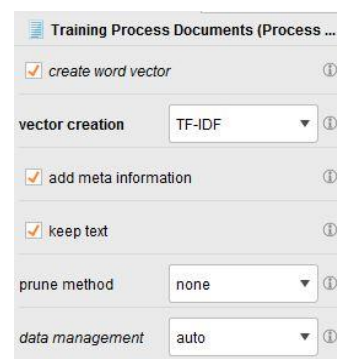
Gambar 11. Proses *pre-processing* keseluruhan

Tabel 6 merupakan hasil proses *labeling* sesudah dilakukan seluruh tahapan *preprocessing*.

Tabel 6. Hasil proses *Labeling*

Sebelum Proses <i>Pre-Processing</i>	Setelah Proses <i>Pre-Processing</i>	Kelas
gunakan Processor Ryzen, bagus dan harga Murah di @Tokokomputer #RyzenOnFire	[gunakan, prosessor, ryzen, bagus, harga, murah]	Positif

CPU “berkelas dengan harga terjangkau” #RyzenOnFire.	[cpu, berkelas, harga, terjangkau]	Positif
Ryzen CPU yang tidak terbilang murah, dengan performa buruk.	[ryzen, cpu, tidak, terbilang, murah, performa, buruk]	Negatif
Ryzen jelek lebih baik Intel http://.asnaskjf.avjavja	[ryzen, jelek, lebih, baik, intel	Negatif
Processor AMD Ryzen lumayanlah untuk digunakan @jasgkasj	[prosessor, amd, ryzen, lumayanlah,]	Positif
Ryzen, Posessor yang sangat layak di pilih.	[ryzen, prosessor, sangat, layak, pilih]	Positif



Gambar 12. Pengaturan TF-IDF

Gambar 12 merupakan proses pengaturan TF-IDF untuk melakukan seleksi fitur yang akan dipilih untuk proses selanjutnya.

Tabel 7 memperlihatkan cuplikan hasil kata yang sudah melewati tahapan *preprocessing* dan seleksi fitur melalui fase TF-IDF.

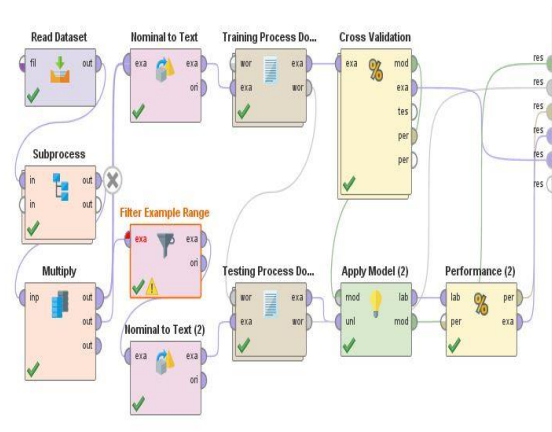
Tabel 7. Hasil proses TF-IDF

Word	Attribute ...	Tota...	Docum...	Positive	Negative
acer	acer	1	1	1	0
aga	aga	1	1	1	0
agam	agam	1	1	1	0
agresif	agresif	1	1	1	0
ahong	ahong	1	1	1	0
aja	aja	4	3	4	0
ajak	ajak	1	1	0	1
aktivitas	aktivitas	2	2	2	0
akun	akun	1	1	1	0
alexandravidia	alexandra...	1	1	1	0
amd	amd	49	42	47	2
amders	amders	1	1	1	0
amdingdo	amdingdo	1	1	1	0
amdingnesia	amdingnon...	5	5	5	0

3.3. Model Support Vector Machine

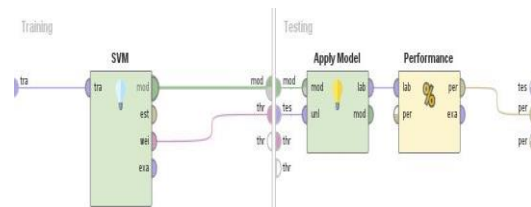
Klasifikasi pemodelan dilakukan dengan memakai perangkat lunak *open source RapidMiner*. Metode yang dipakai untuk mengklasifikasikan data penelitian ini menggunakan metode *Support Vector Machine*.

Pada pengujian ini, data yang dipakai merupakan data *tweet* hasil *crawling* yang tersimpan dalam file excel, kemudian dilakukan *preprocessing* oleh sistem analisis sentimen yang dibangun. Proses selanjutnya merupakan konversi data menjadi dokumen dengan operator *process document* melalui proses pembobotan kata. Tahap akhir merupakan tahap pengujian dengan memakai operator *cross validation*. Keseluruhan proses analisis sentimen yang dilakukan dengan perangkat lunak *RapidMiner* dapat dilihat pada Gambar 13.



Gambar 13. Keseluruhan rangkaian operator

Proses klasifikasi *Support Vector Machine* yang dilakukan didalam operator *cross validation*, diperlihatkan pada Gambar 14.



Gambar 14. Proses klasifikasi SVM

3.4. Evaluasi Model SVM

Tabel 8 memperlihatkan hasil evaluasi dari model dalam bentuk *confusion matrix*. Terdapat 29 data berada dalam klasifikasi yang tepat (*true positive* dan *true negative*) yang terdiri dari 27 data diprediksi positif dan aktualnya memang positif, 2 data diprediksi negatif dan aktualnya memang negatif, 1 diprediksi negatif, namun aktualnya adalah positif.

Tabel 8. Confusion Matrix hasil klasifikasi

accuracy: 96.67%			
	true Positive	true Negative	class precision
pred. Positive	27	1	96.43%
pred. Negative	0	2	100.00%
class recall	100.00%	66.67%	

Tabel 9 memperlihatkan pengukuran kinerja metode SVM. Hasil pengukuran berdasarkan tabel *confusion matrix* menunjukkan bahwa kinerja metode SVM sangat baik dalam hal akurasi, presisi, *recall*

dan *F1 Score* dengan nilai masing-masing 96,67%, 96,43%, 100% dan 98,18%.

Tabel 9. Pengukuran Kinerja SVM

Akurasi	Presisi	Recall	F1 Score
96,67%	96,43%	100%	98,18%

4. Simpulan

Penelitian ini sukses menggunakan metode SVM dalam mengklasifikasikan opini publik mengenai produk AMD Ryzen yang beredar di tengah masyarakat. Hasil penelitian memperlihatkan bahwa sebagian besar publik memberikan tanggapan positif terhadap prosesor AMD Ryzen. Hasil analisis sentimen ini bisa digunakan oleh masyarakat dalam membuat keputusan berkaitan dengan pembelian prosesor tersebut. Metode SVM memberikan kinerja terbaik pada hasil pengukuran akurasi, presisi, recall dan F1 Score dengan nilai rata-rata diatas 96%. Berdasarkan hasil penelitian ini, SVM dapat digunakan sebagai algoritma cerdas dalam bidang analisis sentimen untuk data baru. Untuk penelitian selanjutnya, menggabungkan beberapa algoritma machine learning dengan teknik ensemble dan menganalisis dampak dari penggabungan tersebut.

5. Referensi

arsy, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (Svm). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147. <https://doi.org/10.25126/Jtiik.0813944>

Chakraborty, K., Bhattacharyya, S., Bag, R., & Hassanien, A. A. (2019). Sentiment Analysis On A Set Of Movie Reviews Using Deep Learning Techniques. In *Social Network Analytics*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-815458-8.00007-4>

Daeli, N. O. F., & Adiwijaya. (2020). Sentiment Analysis On Movie Reviews Using Information Gain And K-Nearest

Neighbor. *Journal Of Data Science And Its Applications*, 3(1), 1–7. <https://doi.org/10.34818/Jdsa.2020.3.22>

Dictionary, O. (2021). Processor. In *Oxford Dictionary*. <https://www.oxfordlearnersdictionaries.com/definition/english/processor?q=processor>

Erlin, Rio, U., & Rahmiati. (2014). Two Text Classifiers In Online Discussion : Support Vector Machine Vs Back-Propagation Neural Network. *Telkomnika*, 12(1), 189–200. <https://doi.org/10.12928/Telkomnika.V12i1.1798>

Erlin, Rio, U., & Rahmiati. (2013). Text Message Categorization Of Collaborative Learning Skills In Online Discussion Using Support Vector Machine. *Proceeding - 2013 International Conference On Computer, Control, Informatics And Its Applications: "Recent Challenges In Computer, Control And Informatics"*, *Ic3ina 2013*, 295–300. <https://doi.org/10.1109/Ic3ina.2013.6819190>

Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71. <https://doi.org/10.26623/Transformatika.V18i1.2317>

Fitriyah, N., Warsito, B., & Maruddani, D. A. I. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (Svm). *Jurnal Gaussian*, 9(3), 376–390. <https://doi.org/10.14710/J.Gauss.V9i3.28932>

Gandhi, R. (2018). *Support Vector Machine — Introduction To Machine Learning Algorithms Svm Model From Scratch*. Towards Data Science. <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms->

- 934a444fca47
- Hariguna, T., Baihaqi, W. M., & Nurwanti, A. (2019). Sentiment Analysis of Product Reviews As A Customer Recommendation Using The Naive Bayes Classifier Algorithm. *Ijiis: International Journal Of Informatics And Information Systems*, 2(2), 48–55. <https://doi.org/10.47738/Ijiis.V2i2.13>
- Isnain, A. R., Marga, N. S., & Alita, D. (2021). Sentiment Analysis of Government Policy On Corona Case Using Naive Bayes Algorithm. *Ijccs (Indonesian Journal of Computing and Cybernetics Systems)*, 15(1), 55. <https://doi.org/10.22146/Ijccs.60718>
- Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, And Emotions. In *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press. <https://doi.org/10.1017/Cbo9781139084789>
- Liu, N., & Shen, B. (2020). Aspect-Based Sentiment Analysis with Gated Alternate Neural Network. *Knowledge-Based Systems*, 188(Xxxx), 105010. <https://doi.org/10.1016/J.Knosys.2019.105010>
- Mamtesh, & Mehla, S. (2019). Sentiment Analysis of Movie Reviews Using Machine Learning Classifiers. *International Journal of Computer Applications*, 182(50), 25–28. <https://doi.org/10.5120/Ijca2019918756>
- Mohd Nafis, N. S., & Awang, S. (2021). An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination For Sentiment Classification. *Ieee Access*, 9(Ml), 52177–52192. <https://doi.org/10.1109/Access.2021.3069001>
- Nurdiansyah, Y., Bukhori, S., & Hidayat, R. (2018). Sentiment Analysis System For Movie Review In Bahasa Indonesia Using Naive Bayes Classifier Method. *Journal of Physics: Conference Series*, 1008(1). <https://doi.org/10.1088/1742-6596/1008/1/012011>
- Prastyo, P. H., Sumi, A. S., Dian, A. W., & Permanasari, A. E. (2020). Tweets Responding To The Indonesian Government’s Handling of Covid-19: Sentiment Analysis Using Svm with Normalized Poly Kernel. *Journal of Information Systems Engineering And Business Intelligence*, 6(2), 112. <https://doi.org/10.20473/Jisebi.6.2.112-122>
- Putra, M. W. A., Susanti, Erlin, & Herwin. (2020). Analisis Sentimen Dompok Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier. *It Journal Research and Development*, 5(1), 72–86. [https://doi.org/10.25299/Itjrd.2020.Vol5\(1\).5159](https://doi.org/10.25299/Itjrd.2020.Vol5(1).5159)
- Scikit-Learn. (2021). *Svm*. Scikit-Learn.Org. <https://scikit-learn.org/stable/modules/svm.html>
- Singh, S. N., & Sarraf, T. (2020). Sentiment Analysis of A Product Based on User Reviews Using Random Forests Algorithm. *Proceedings Of The Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering*, 112–116. <https://doi.org/10.1109/Confluence47617.2020.9058128>
- Stephenie, Warsito, B., & Prahutama, A. (2020). Sentiment Analysis on Tokopedia Product Online Reviews Using Random Forest Method. *E3s Web Of Conferences*, 202, 1–10. <https://doi.org/10.1051/E3sconf/202020216006>
- Suresh, A., & Bharathi, C. R. (2016). Sentiment Classification Using Decision Tree Based Feature Selection. *International Journal Of Control Theory and Applications*, 9(36), 419–425.
- Towkeersah40. (2021). *Role Of Processor In A Pc*. Geeksforgeeks.

<https://www.geeksforgeeks.org/role-of-processor-in-a-pc/>

Uy, M. R., & Thomas, J. (2021). *Amd Processors: The Best Amd Cpus In 2021*. Techradar.

<https://www.techradar.com/news/amd-processors-the-best-amd-cpus-in-2021>

Wahyudi, R., & Kusumawardana, G. (2021). Analisis Sentimen Pada Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine. *Jurnal Informatika*, 8(2), 200–207.

<https://doi.org/10.31294/ji.v8i2.9681>

Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment Analysis For E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *Ieee Access*, 8, 23522–23530. <https://doi.org/10.1109/Access.2020.2969854>