ANALISIS DATA SENTIMEN NEGATIF PADA OPINI PENGGUNA TWITTER TERHADAP BERITA SEPAK BOLA LIGA 1 TAHUN 2022 DENGAN PENERAPAN *SUPPORT VECTOR MECHINE*

Nurul Dalifah, Nana Suarna, Willy Prihartono
Teknik Informatika,Komputerisasi Akuntansi, STMIK IKMI Cirebon
Jln.Perjuangan No.10B Majasem Kec. Kesambi Kota Cirebon
nuruldalifah80@gmail.com

ABSTRAK

Perkembangan media sosial semakin pesat dikalangan masyarakat salah satunya twitter yang menggunakan bahasa alami untuk bisa membuka potensi pada interaksi manusia yang bisa mengetahui sentiment para penggunanya. Permasalahan yang didapat dari masalah tersebut yaitu menggali informasi,menganalisis pendapat sentiment, evaluasi sikap dan emosi orang dari bahasa tertulis. Tujuannya untuk memahami dan mengkategorikan sentiment dari opini pengguna *Twitter* terkait berita liga 1 indonesia tahun 2022. Bagaimana dataset opini pengguna *twitter* terhadap berita liga 1 tahun 2022 dikumpulkan dan diolah untuk analisis. Penelitian ini akan menggunakan data teks dari pengguna *Twitter* yang terkait dengan berita liga 1 tahun 2022. *SVM* akan digunakan untuk memprediksi sentimen opini dari tweet tersebut menjadi kategori positif, negatif, atau netral. Hasil penelitian ini menggunakan algoritma *support vector mechine* menunjukan bahwa akurasi sebesar 91% presisi 30%, recall 33%,F1 score 32% dengan Jumlah sentimen Negatif = 445, Positif = 50, dan Netral = 4. Secara keseluruhan Model algoritma *support vector mechine* menghasilkan akurasi yang lebih tinggi. Hasil model menunjukan bahwa sentimen opini pengguna twitter menanggapi berita tersebut yang bersifat kontrovesrsial

Kata kunci : *analisis sentiment, opini pengguna twitter, berita sepak bola indonesia liga 1, support vector mechine*

1. PENDAHULUAN

Dalam era kemajuan pesat dibidang informatika, teknologi informasi telah mengalami kemajuan yang signifikan dan memberikan dampak yang luas terhadap berbagai sector kehidupan. Transformasi digital telah menjangkau sector-sektor krusial seperti teknologi, bisnis,dan pendidikan.Salah satunya yang menarik perhatian yaitu berita dunia sepak bola, terutama berita liga 1 tahun 2022 yang merespon dan menanggapi opini, pandangan reaksi terhadap berita tersebut.

Permasalahan yang ditemukan yaitu mengidentifikasi mengapa berita liga 1 tahun 2022 perlu dilakukan analisis sentimen,karena tantangan pada berita tersebut mengacu pada peristiwa yang sudah terjadi, masalah yang harus dipecahkan termasuk relevan untuk merujuk pada tren terbaru melalui berita atau isu isu.

Penelitian yang dilakukan oleh Minardi dkk yang berjudul Sentiment Analysis Terhadap Perspektif Warganet Atas Tragedi Kanjuruhan Malang di *Twitter* Menggunakan *Naïve Bayes Classifier* hasil dari penelitian tersebut adalah Analisis sentimen dilakukan dengan mengindeks tweet pengguna *Twitter* dengan tagar Usut Tuntas Tragedi Kanjuruhan, mengambil (crawling) data 1.500 tweet yang ada sebagai kumpulan data (dataset), setelah itu data untuk diproses diidentifikasi (labeling) untuk langkah selanjutnya yaitu tahap *Preprocessing* data yang terdiri dari *Annotation Removal, Remove Hashtag, Transformation Remove Url, Regexp,* Indonesian Steaming, Indonesian *Stopword Removal* dipadukan dengan operator *Smote Upsampling*. Pembuatan *Confusion Matrix* yang menunjukkan hasil akhir analisis berjalan dengan baik yaitu nilai *accuracy* 77,67%, nilai *precision* sebesar 77,19%, nilai *recall s*ebesar 78,50%, dan nilai *AUC* 0.820 *(good classification*).[1]

Penelitian sebelumnya tentang analisis sentimen, yaitu Penelitian yang dilakukan oleh Hanafi dan Sholihin yang Berjudul Analisis Sentimen Terhadap Pssi Atas Tragedi Kanjuruhan Menggunakan Multinomial Naïve Bayes hasil dari penelitian tersebut adalah Hasil dari pengujian menggunakan metode Multinomial Naïve Bayes mampu menghasilkan analisa pengujian akurasi tertinggi sebelum Tragedi Kanjuruhan bernilai 73% dengan rasio terbaik ialah 80% data latih dan 20% data uji, sedangkan akurasi tertinggi sesudah Tragedi Kanjuruhan bernilai 68% dengan rasio terbaik ialah 90% data latih dan 10% data uji.[2]

Pendapat lain yang dilakukan oleh Nugraha dkk tentang analisis sentimen, yaitu yang berjudul "Analisis Sentimen Twitter Terhadap Menteri Indonesia Dengan Algoritma Support Vector Machine Dan Naive Bayes" Penelitian ini fokus pada tanggapan pengguna Twitter terhadap pelantikan kabinet dan bertujuan untuk mengelompokkan opini dalam kategori analisis sentimen. Dalam penelitian ini, digunakan model algoritma Support Vector Machine (SVM) dan Naïve Bayes.Hasil penelitian menunjukkan bahwa pada pengujian silang, algoritma

SVM mencapai akurasi 89,60%, recall 90,91%, dan presisi 97,64%. Sementara algoritma Naïve Bayes memiliki akurasi 85,74%, recall 85,74%, dan presisi 100,00%. Penelitian ini memberikan wawasan kepada pemerintah dan masyarakat mengenai pandangan terhadap kinerja Menteri Indonesia melalui opini di media sosial.[3]

Tujuan penelitian ini yaitu untuk menganalisis sentimen opini pengguna twitter terhadap berita sepak bola Indonesia Liga 1 tahun 2022 menggunakan metode *support vector machine* (SVM). Dengan penelitian ini, berupaya untuk memahami perasaan dan sikap pengguna twitter terhadap berita liga 1 2022, apakah positif, negatif, atau netral.

Metode penelitian ini terdiri dari beberapa langkah penting yaitu fokus utama adalah melakukan analisis sentimen terhadap opini pengguna *twitter* terkait berita sepak bola liga 1 tahun 2022 menggunakan metode *support vector machine (SVM).* Selanjutnya *SVM* telah terbukti menjadi pendekatan yang efektif dalam analisis sentimen dan klasifikasi teks di berbagai konteks sebelumnya[4]

Implikasi hasil penelitian ini memberikan kontribusi yang signifikan pada pemahaman tentang bagaimana masyarakat Indonesia merespons dan merasa terkait berita liga 1 tahun 2022. Ini dapat membantu memahami sentimen dan opini umum terkait acara olahraga populer ini. Kedua, temuan dari penelitian ini dapat sangat berguna bagi para praktisi di industri media, olahraga, dan pemasaran

2. TINJAUAN PUSTAKA

2.1. Twitter

Twitter adalah platform yang dioperasikan oleh Twitter Inc., menyediakan layanan jejaring sosial berbentuk mikroblogging. Di sini, pengguna dapat mengirim dan membaca pesan dalam bentuk 'tweet' (Twitter, 2013). Mikroblogging adalah bentuk komunikasi online yang memungkinkan pengguna memperbarui status mengenai pemikiran, kegiatan, pandangan tentang objek atau fenomena tertentu, dan sebagainya. Tweet merupakan pesan teks yang terbatas hingga 140 karakter yang ditampilkan di halaman profil pengguna. Meskipun tweet bisa dilihat secara publik, pengirim memiliki opsi untuk membatasi pengiriman pesan hanya kepada daftar teman mereka.Pengguna juga dapat melihat 4.444 tweet dari pengguna lain yang dikenal sebagai pengikut.[5]

2.2. Analisis Sentimen

Analisis Sentimen merupakan proses yang bertujuan untuk menafsirkan makna dari suatu pendapat, pandangan, atau ekspresi emosional yang terdapat dalam teks, percakapan, postingan (aktivitas internet), atau dalam basis data dengan menggunakan Natural Language Processing (NLP). Penelitian dalam analisis sentimen terdiri dari beberapa tugas, seperti ekstraksi sentimen, klasifikasi sentimen, dan identifikasi spam . Dalam konteks bisnis, analisis sentimen sangat bermanfaat karena memungkinkan pemahaman terhadap respons pelanggan terhadap produk yang ditawarkan. Metode dasar yang digunakan dalam analisis sentimen pada teks adalah melalui konstruksi kamus kata yang memilah kata-kata menjadi positif dan negatif berdasarkan sifat atau makna intrinsik dari setiap kata. Tingkat keakuratan dari pendekatan ini sangat bergantung pada kelengkapan kamus kata yang telah dibangun. [6]

2.3. Fenomena Sepak Bola.

Banyak pelatih yang menjadi korban yang diawali dari Robert Albert yang mengundurkan diri pada pekan ketiga liga 1 2023-2023 dan diakhri oleh Milomir Seslija yang juga ditendang Borneo Fc pada pekan ke-10 Lalu terjadilah tragedi Kanjuruhan tepat pada tanggal 1 oktober 2022 yang mengguncang liga 1 tahun 2022-2023 saat laga derbi Jawa Timur antara Arema Fc melawan persebaya diStadion Kanjuruhan, Kabupaten Malang, yang berakhir ricuh. Pada saat akhir score arema Fc menderita kekalahan 2-3 dari sang tamu yang lantas membuat suporter tuan rumah hingga turun kelapangan. Suporter yang pada awalnya ingin memberikan semangat terhadap fans pemainnya,tetapi karna jumlah kapasitas penonton melebihi akhirnya pihak keamanan mengambil tindakan represif dengan menembak gas air mata. Gas air mata ini kemudan menciptakan kepanikan yang luar biasa hebat, terlebih setelah proyektil ditembakkan ketribbun penonton. Jumlah total 135 korban meninggal dunia karena terinjak injak dan sesak nafas karena gas air mata. Tragedi ini membuat liga 1 tahun 2022-2023 dan seluruh kompetisi sepak bola lain dihentikan secara total selama kurang lebih 2 bulan.[7]

2.4. Text Mining

Text mining adalah proses penambangan data dalam bentuk teks. Biasanya, sumber data diambil dari dokumen.Tujuannya adalah untuk menemukan kata-kata yang menggambarkan isi dokumen sehingga dapat dilakukan analisis hubungan antar dokumen. Tujuan dari penambangan teks adalah untuk mengekstrak informasi yang berguna dari sumber data. Oleh karena itu, sumber data yang digunakan dalam text mining adalah kumpulan dokumen dengan format tidak terstruktur yang melaluinya pola-pola menarik dapat diidentifikasi dan dipelajari. Tugas khusus dalam penambangan teks meliputi klasifikasi teks dan pengelompokan teks. dalam penelitian menggunakan metode SVM dengan kelebihan SVM adalah dapat menganalisis pola klasifikasi data dengan cara yang sama dan akurasi yang sama. Ini memiliki fleksibilitas tinggi dan beragam skenario ilmu data. SVM telah menjadi metode klasifikasi yang umum digunakan.[8]

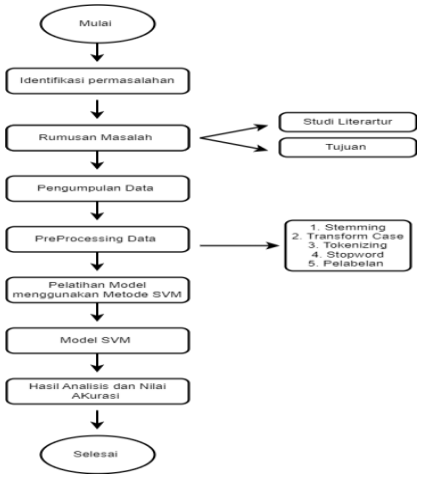
3. METODE PENELITIAN

Metode pada tahapan penelitian ini yang digunakan untuk mengklasifikasi opini pengguna *twitter* pada berita liga 1 tahun 2022 yaitu dengan menerapkan metode *Support Vector Mechine*(SVM), pada rancangan yang akan dbangun yang akan ditampilkan pada Gambar 1.

Pada Gambar 1, terdapat 5 (lima) tahapan dalam penelitian diantaranya yaitu pengumpulan data, preprocessing,pelatihan model menggunakan pelatihan svm, proses model svm, hasil analisis dan nilai akurasi sekaligus evaluasi didalamnya.

3.1. Pengumpulan Data

Tahapan pengumpulan data yang digunakan yaitu dari dataset twitter yang berisi 12 atribut dengan jumlah dataset 509 record. Hasil dari *crawling* data pada *google collab*. Setelah data yang dikumpulkan, melakukan tahap pembersihan data terlebih dahulu untuk diproses dtahap analisis diantaranya tahapan proses penelitian diatas proses *cleaning* data pada *twitter t*erhadap dataset yaitu menghapus url (https://) pada tiap komenan, menghapus *mention*(@),menghapus hastag(#), menghapus karakter selain huruf dan spasi maka setelah proses pembersihan data, maka dataset siap untuk digunakan dalam analisis ketahap selanjutnya.



Gambar 1. Tahapan Metode Penelitian

3.2. PraProcessing

Praprocessing adalah proses tahapan awal yaitu diantaranya Proses *case folding*, tokenisasi, stopword dan *stemming* untuk menjadi dataset yang digunakan pada penelitian.[9]

1) Proses Case Folding
   Case folding adalah proses mengubah semua karakter dalam sebuah dokumen menjadi huruf kecil atau huruf besar yang sama, untuk mempercepat perbandingan selama proses pemrosesan data.
2) Tokenisasi
   Tokenisasi adalah untuk memecah teks menjadi bagian-bagian yang lebih kecil, yang nantinya dapat diolah lebih lanjut dalam analisis bahasa alami, pemrosesan bahasa alami, atau pemodelan bahasa. Memisahkan teks menjadi bagian-bagian yang lebih kecil, seperti kata-kata, frasa, atau karakter, memungkinkan data lebih efisien memahami dan menganalisis teks yang diberikan.
3) Stopword
   Stopword adalah istilah yang mengacu pada kata-kata umum yang sering ditemukan dalam teks tetapi kurang memberikan nilai atau informasi yang signifikan dalam analisis bahasa alami atau saat memproses teks. Ini adalah kata-kata seperti "the", "is", "at", "on", "in", "and", yang sering muncul namun jarang memberikan kontribusi informasi khusus tentang teks yang sedang diproses.
4) Stemming
   Stemming adalah proses dalam pemrosesan bahasa alami yang bertujuan untuk mengubah kata-kata menjadi bentuk dasarnya atau "akar kata." Ini dilakukan dengan menghapus awalan, akhiran, atau infleksi dari kata sehingga hanya sisa akar kata yang mewakili makna dasarnya.
5) Pelabelan
   Pelabelan adalah memberikan kategori atau tanda pada data agar atribut tertentu dapat diidentifikasi. Dengan pelabelan ini memudahkan menentukan kalimat sentimen yang terdiri dari negative positif ataupun netral.

3.3. Pelatihan Model menggunakan SVM

Pelatihan model menggunakan svm terdapat beberapa tahapan diantaranya ektraksi fitur yaitu perhitungan tfi-idf dan pembagian data untuk pelatihan model.

1) Ekstraksi Fitur
   TF-IDF *(Term Frequency-Inverse Document Frequency)* memberi nilai penting pada kata dalam satu dokumen dengan mempertimbangkan seberapa sering kata tersebut muncul di dokumen itu dan seberapa jarang kata itu muncul di seluruh kumpulan dokumen. Ini membantu menentukan kepentingan kata dalam konteks spesifik dan banyak digunakan dalam pemrosesan teks, pengklasifikasian dokumen, dan sistem rekomendasi. Rumus pada TF-IDF seperti pada dibawah berikut ini :

$$TF_{(t,d)} = \frac{Number\ of\ times\ t\ appear\ in\ document\ d}{Total\ number\ of\ terms\ in\ document\ d}$$

$$IDF_{(t,d)} = \log\left(\frac{Total\ number\ of\ documents\ D}{Number\ of\ document\ the\ term\ in\ it}\right)$$

$$TFIDF(t, d\ D) = TF_{(t,d)} * IDF_{(t,d)}$$

dimana :

t_i = Dokumen ke – i,

df = Document frequency,
n = Banyaknya data
idf = Inverse document frequency.

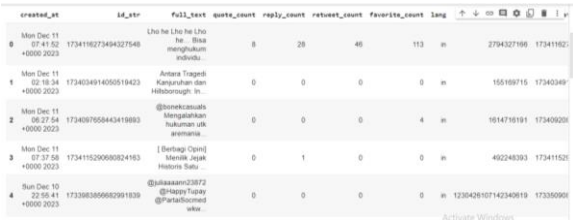
2) Pembagian Data
Tahapan selanjutnya adalah tahap partisi data yang membagi data menjadi dua bagian yaitu data latih dan data uji. Pembagian menjadi data pelatihan dan pengujian didasarkan pada beberapa penelitian yang dilakukan. Terdapat perbandingan rincian data yang digunakan dari beberapa penelitian: skenario 90: 10, 80: 20, 70: 30, dan 60: 40.

3.4. Pemodelan *Suport Vector Mechine*

Support Vector Mechine (SVM) adalah metode klasifikasi pola umum yang digunakan di banyak aplikasi berbeda. Pengaturan parameter kernel dan pemilihan fitur pada prosedur pelatihan *SVM* secara signifikan mempengaruhi akurasi klasifikasi[10]. Pada penelitian ini menggunakan *support vector mechine* Proses klasifikasi dibagi menjadi 3 sentimen kelas yaitu positif, negatif dan netral yang dimana data dibagi menjadi dua yaitu data training dan data testing menggunakan ratio 0,1 yang artinya 90 % data digunakan untuk proses training dan digunakan untuk 10 % untuk data testing.

4. HASIL DAN PEMBAHASAN

4.1. Data



Gambar 2. hasil carwling data

Hasil crawling dala bentuk soft file csv dari jumlah keseluruhan data yang kerecord berjumlah 509 data yang berjumlah 12 atribut diantaranya *created_id,idr_string,full_text, quote_count, rpreply_count, retweet_count, favorite_count, lang, user_id_str, conversation_id_str, username, tweet_url*, dari 12 atribut yang ditampilkan. setelah itu proses pembersihan data dan menyeleksi atribut teks.

Tabel 1. Pembersihan Data

Sebelum	Sesudah
Lho he Lho he Lho heâ€¦ Bisa menghukum individual SUporter ta? Apa kabar Tragedi Kanjuruhan? Ga ada satupun nama SUporter perorangan maupun organisasi yang dihukum, padahalâ€¦.. https://t.co/IJI0iacWGF	Lho he Lho he Lho he Bisa menghukum individual Suporter ta Apa kabar Tragedi Kanjuruhan Ga ada satupun nama Suporter perorangan maupun organisasi yang dihukum padahal

Pada tabel 1 menunjukan hasil pembersihan data pada atribut teks untuk menghilangkan atau menghapus bagian yang tidak diperlukan seperti url, hastag, mention dan karakter. Pada data yang sudah bersih maka akan melanutkan ketahap preprocessing.

4.2. Praprocessing

a) Case Folding

Tabel.2. Hasil case folding

Antara Tragedi Kanjuruhan dan Hillsborough Indonesia Susah Belajar	antara tragedi kanjuruhan dan hillsborough indonesia susah belajar

Pada tabel 1 menunjukan hasil pembersihan data pada atribut teks untuk menghilangkan atau menghapus bagian yang tidak diperlukan seperti url, hastag, mention dan karakter. Pada data yang sudah bersih maka akan melanutkan ketahap preprocessing.

b) Tokenisasi
Pada tahapan tokenisasi kalimat “antara tragedi kanjuruhan dan hillsborough indonesia susah belajar” yang jika ditokenisasikan menjadi “['antara', 'tragedi', 'kanjuruhan', 'hillsborough', 'indonesia', 'susah', 'belajar']. dari kalimat tersebut telah terbagi menjadi kata token yang terpisah.

4.3. Stopword

tahapan stoprword menghilangkan kata yang tidak berguna seperti dan,pada,sebagai dll. sperti hasil dari kalimat berikut “mengalahkan hukuman utk aremania yg rusuh saat tragedi kanjuruhan eh malah gak ada yang dihukum sih” yang diberi label kuning yaitu kata stopword yang harus dihilangkan menjadi “mengalahkan hukuman utk aremania yg rusuh tragedi kanjuruhan eh malah gak yang dihukum sih”

4.4. Stemming

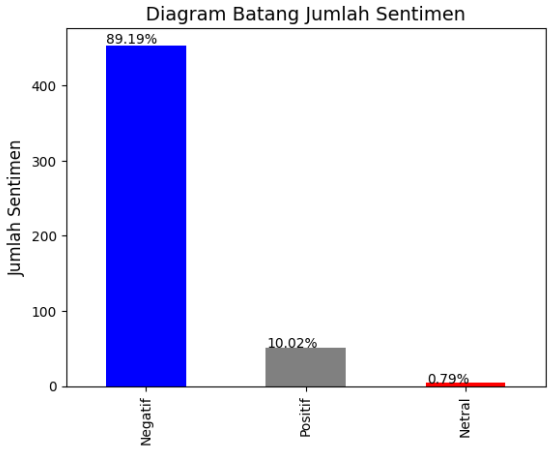
Merubah kalimat kata imbuhan menjadi kata dasar pada dataset.

Tabel 3. Proses Stemming

mengalahkan hukuman utk aremania yg rusuh saat tragedi kanjuruhan eh malah gak ada yg dihukum sih hehehe	['kalah', 'hukum', 'utk', 'aremania', 'yg', 'rusuh', 'tragedi', 'kanjuruhan', 'eh', 'gak', 'yg', 'hukum', 'sih', 'hehehe']

Pada tabel diatas menunjukan proses hasil atemming dengan kata imbuhan menjadi kata dasar seperti menghukum menjadi hukum.

4.5. Pelabelan



Gambar 3. Hasil jumlah sentimen

Pada gambar diatas menunjukan bahwa jumlah negatif lebih banyak dengan jumlah 89,19%, positif, 10,02% dan netral di 0,79% lebih dominan pada sentimen negatif.

4.6. Pelatihan Support Vector Mehine

a) Ekstraksi Fitur

```
(0, 1852)   0.141672218716548980   (507, 1000)   0.21259761966127710
  (0, 581)  0.122968722287433390     (507, 2068) 0.13959590892430810
 (0, 1442) 0.196457658739980750     (508, 1244) 0.31143423236488920
 (0, 1989) 0.103790596739896780     (508, 460)  0.29683237633615167
 (0, 1279) 0.196457658739980750     (508, 1659) 0.28550629688232176
 (0, 1276) 0.135062757405650600     (508, 657)  0.31143423236488920
```

Gambar 4. Hasil dari perhitungan tf-idf

Pada gambar diatas menunjukan hasil nilai yang diperoleh dari perhitugan tf-idf yang menggunakan tools google collab .

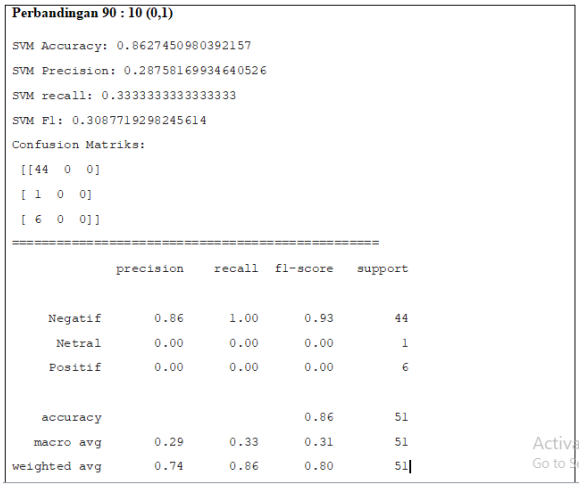
b) Pembagian Data

Tabel 4. Hasil pembagian data

Perbandingan	Data latih	Data uji	Total
Percobaan 1 (90 : 10)	458	51	509
Percobaan 2 (80 : 20)	407	102	509
Percobaan 3 (70 : 30)	356	153	509
Percobaan 4 (60 :40)	305	204	509

4.7. Pemodelan Suppor Vector Mechine

Dari perbandingan 0,1 yaitu 90 : 10 dengan pembagian masing masing data 458 untuk data data latih dan 51 untuk data uji menghasilkan Akurasi sebesar 0,86, presisi 0,29, recall 0,33 F1-score 0,30 Dari perhitungan data latih menghasilkan Negatif sebesar 44, positif dengan jumlah 1 dan netral berjumlah 6 maka jika ditotalkan ada 51 ulasan dan hasil akhir evaluasi model pada tabel dibawah ini



Perbandingan 90 : 10 (0,1)

```
SVM Accuracy: 0.8627450980392157
SVM Precision: 0.28758169934640526
SVM recall: 0.3333333333333333
SVM F1: 0.3087719298245614
Confusion Matriks:
 [[44  0  0]
 [ 1  0  0]
 [ 6  0  0]]
================================================
              precision    recall  f1-score   support

     Negatif       0.86      1.00      0.93        44
      Netral       0.00      0.00      0.00         1
     Positif       0.00      0.00      0.00         6

    accuracy                           0.86        51
   macro avg       0.29      0.33      0.31        51
weighted avg       0.74      0.86      0.80        51
```

Gambar 4. Hasil dari pemodelan algorirma svm

Tabel 5. Evaluasi model

Keterangan	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4
Data Trainig	90%	80%	70%	60%
Data Latih	10%	20%	30%	40%
Negatif	44	90	139	183
Positif	1	1	2	2
Netral	6	11	12	19
Akurasi	86%	88%	91%	90%
Presisi	29%	29%	30%	30%
Recall	33%	33%	33%	33%
F1-Score	31%	31%	32%	32%

Pada tabel 5 diatas mengindikasikan bahwa dari empat rangkaian percobaan, skenario yang memberikan nilai akurasi paling tinggi adalah pada percobaan ketiga yaitu dengan data trainning 70% dan data latih 30% dengan nilai akurasinya sebesar 91%

5. KESIMPULAN DAN SARAN

Penelitian menggunakan algoritma Support Vector Machine (SVM) berhasil mengklasifikasikan sentimen menjadi negatif, positif, dan netral melalui beberapa tahapan. Proses dimulai dengan pengambilan data melalui crawling menggunakan Google Colab, dilanjutkan dengan pembersihan, preprocessing, dan terjemahan data. Selanjutnya, dilakukan ekstraksi fitur atau pembobotan kata, pembagian data, implementasi algoritma SVM, dan evaluasi model.Ekstraksi data dilakukan melalui tiga tahap. Setelah tokenisasi, kalimat diubah menjadi unit kata kecil, dihitung Term-Frequency (TF), kemudian dilanjutkan dengan perhitungan Inverse Document Frequency (IDF). Perkalian TF dengan IDF menghasilkan nilai TF-IDF sebagai hasil akhir proses.Dari hasil pengujian, akurasi tertinggi tercapai pada percobaan ketiga, menggunakan 70% data latih dan 30% data testing, dengan akurasi mencapai 91%, lebih tinggi dibandingkan dengan dua percobaan

sebelumnya.Saran yang dapat diberikan adalah mengenai pengambilan sampel data. Dari 509 ulasan, sebanyak 445 memiliki sentimen negatif. Jika pengambilan data diperluas menjadi lebih dari 500, kemungkinan besar akan memperoleh jumlah ulasan yang lebih representatif.

DAFTAR PUSTAKA

[1] M. Minardi, R. Lasepa, S. Riyadi, S. Ramadhan, and D. D. Saputra, "Sentiment Analysis Terhadap Perspektif Warganet Atas Tragedi Kanjuruhan Malang di Twitter Menggunakan Naïve Bayes Classifier," *J. Inform.*, vol. 10, no. 1, pp. 45–53, 2023, doi: 10.31294/inf.v10i1.14546.

[2] M. A. Hanafi and A. Solichin, "Sentiment Analysis on Pssi Over Kanjuruhan Tragedy Using Multinomial Naïve Bayes," vol. 15, pp. 21–28, 2023.

[3] S. N. Nugraha, R. Pebrianto, A. Latif, and M. R. Firdaus, "Analisis Sentimen Twitter Terhadap Menteri Indonesia Dengan Algoritma Support Vector Machine Dan Naive Bayes," *E-Link J. Tek. Elektro dan Inform.*, vol. 17, no. 1, p. 1, 2022, doi: 10.30587/e-link.v17i1.3965.

[4] K. M. L. Muhammad Fadli Asshiddiqi, "Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI," vol. 39, no. 6, pp. 0–2, 2020.

[5] D. A. Wulandari, R. Rohmat Saedudin, and R. Andreswari, "Analisis Sentimen Media Sosial Twitter Terhadap Reaksi Masyarakat Pada Ruu Cipta Kerja Menggunakan Metode Klasifikasi Algoritma Naive Bayes," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 9007–9016, 2021.

[6] V. A. Flores, L. Jasa, and L. Linawati, "Analisis Sentimen untuk Mengetahui Kelemahan dan Kelebihan Pesaing Bisnis Rumah Makan Berdasarkan Komentar Positif dan Negatif di Instagram," *Maj. Ilm. Teknol. Elektro*, vol. 19, no. 1, p. 49, 2020, doi: 10.24843/mite.2020.v19i01.p07.

[7] M. F. Nasvian, M. A. Sugiharto, and M. F. Setiamukti, "Day 7 of Kanjuruhan Tragedy: Twitter Data Analysis," *J. Audiens*, vol. 2, no. 2, 2023, [Online]. Available: https://doi.org/

[8] Y. A. Wijaya, N. Suarna, Iin, R. Hamonangan, and R. Nining, "Comparison of machine learning algorithm for Santander dataset," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, p. 012032, 2021, doi: 10.1088/1757-899x/1088/1/012032.

[9] R. Umar, S. Sunardi, and M. N. Ardhiansyah, "Penerapan Algoritma Support Vector Machine Pada Analisis Sentimen Hashtag Twitter," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 5, p. 1607, 2022, doi: 10.30865/jurikom.v9i5.4877.

[10] A. Nurhadi, "Klasifikasi Konten Berita Digital Bahasa Indonesia Menggunakan Support Vector Machines (SVM) Berbasis Particle Swarm Optimization (PSO)," vol. 3, no. 2, pp. 1–9, 2015.