

PENERAPAN ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN BOY BAND BTS PADA MEDIA SOSIAL TWITTER

Rina Noviana^a, Isram Rasal^b

^aJurusan Informatika, rina_n@staff.gunadarma.ac.id, Universitas Gunadarma

^bJurusan Informatika, isramrasal@staff.gunadarma.ac.id, Universitas Gunadarma

ABSTRACT

BTS or Bangtan Sonyeondan is one of the vocal groups originating from South Korea, which is currently popular among Indonesian teenagers, resulting in many fans providing positive and negative comments through Twitter social media. The method used to determine whether these comments are positive or negative is by conducting sentiment analysis. The stages to perform data analysis are Preprocessing to clean the data, word weighting, labeling data into positive and negative classes, classification, and data visualization using pie charts. In this study, Naive Bayes and Support Vector Machine, were used for comparison, result of an accuracy score is 79% for Naive Bayes and 81% for Support Vector Machine. Among these two methods, Support Vector Machine achieved a higher accuracy score, and the sentiment analysis revealed that the comments obtained from Twitter users are predominantly positive.

Keywords: analysis, BTS, sentiment

ABSTRAK

BTS atau Bangtan Sonyeondan merupakan salah satu grup vokal yang berasal dari Korea Selatan, yang populer di kalangan remaja Indonesia saat ini, sehingga para penggemar atau fans banyak memberikan komentar positif dan negatif melalui media Twitter. Metode yang digunakan untuk mengetahui apakah komentar tersebut positif atau negatif adalah dengan melakukan analisis sentimen. Tahapan untuk melakukan analisis data yaitu Preprocessing untuk membersihkan data, pembobotan kata, pelabelan data menjadi kelas positif dan negatif, pengklasifikasian dan visualisasi data dengan diagram lingkaran. Dalam penelitian ini menggunakan 2 metode yaitu Naive Bayes dan Support Vector Machine untuk melakukan perbandingan dengan menghasilkan nilai akurasi pada Naive Bayes yaitu 79% dan Support Vector Machine yaitu 81%. Dari kedua metode tersebut, metode Support Vector Machine mendapatkan nilai akurasi yang lebih tinggi dan hasil dari analisis sentimen tersebut didapatkan informasi bahwa komentar yang didapatkan dari pengguna Twitter bersentimen positif.

Kata Kunci: analisis, BTS, sentimen

1. PENDAHULUAN

Berkembangnya budaya korea yang pesat dan meluas secara global, sehingga keberadaannya diterima oleh publik dari berbagai kalangan salah satunya Korean Pop Music atau yang biasa disebut K-Pop. Istilah K-Pop merujuk pada genre musik yang dibuat oleh masyarakat Korea dan dinikmati oleh masyarakat Korea maupun masyarakat Internasional [18]. BTS atau Bangtan Sonyeondan merupakan salah satu grup vokal yang berasal dari Korea Selatan, yang debut atau memulai karir sebagai idol pada tanggal 13 juni 2013 yang memiliki 7 orang member atau anggota. BTS memiliki fanbase yang disebut dengan ARMY. BTS merupakan salah satu Boyband dengan pengikut terbanyak, salah satunya di media sosial Twitter dengan pengikut sebanyak 48.143.096 juta Followers [3].

Twitter merupakan layanan jejaring sosial dan mikroblog daring yang mengizinkan penggunaannya untuk mengirim pesan secara real-time, dimana memungkinkan penggunaannya dapat saling berkomunikasi satu sama lain [29]. Media Twitter mempunyai fasilitas mengirim teks info, gambar atau video yang memungkinkan para pengguna dapat mengirim dan membaca pesan teks dengan panjang hingga 140 karakter, atau yang disebut dengan (tweet) [16].

Banyaknya para pengguna twitter sering memberikan opini, baik itu opini negatif atau positif. Opini dapat diungkapkan dengan kata-kata secara terbuka yang ditafsirkan dengan jelas verbal (lisan), maupun dapat diungkapkan melalui kata-kata melalui teks secara tidak langsung. Sehingga dibutuhkan text mining untuk menangani data dari Twitter yang memiliki karakteristik tidak terstruktur. Penggunaan teknik text mining mampu menganalisis data teks semi-terstruktur maupun tidak terstruktur, sehingga dapat menangani suatu masalah seperti analisis sentimen [6]. Metode yang digunakan untuk analisis sentimen antara lain Algoritma Naive Bayes dan Support Vector Machine. Dengan adanya analisis sentimen tanggapan (tweet) dari media Twitter bisa menjadikan bahan evaluasi bagi company (label perusahaan) dan boy grup untuk memperbaiki kualitas kinerja mereka di kemudian hari.

1.1. BTS

BTS atau *Bangtan Sonyeondan* merupakan salah satu grup vokal yang berasal dari Korea Selatan, yang debut atau memulai karir sebagai idol pada tanggal 13 juni 2013 dari salah satu agensi kecil yaitu *Bighit Entertainment*. BTS memiliki 7 orang member atau anggota. BTS memiliki fanbase yang disebut dengan ARMY atau Adorable Representative M. C For Youth terbentuk pada tanggal 9 Juli 2013 terpaut sekitar satu bulan setelah debutnya boyband BTS [3].

1.2. Twitter dan Twitter API (Application Programming Interface)

Twitter adalah layanan jejaring sosial dan mikroblog daring, yang mengizinkan penggunanya untuk mengirim pesan secara real-time, dimana memungkinkan penggunanya dapat saling berkomunikasi satu sama lain. Twitter dibuat pada bulan Maret 2006 dan diluncurkan pada bulan Juli di tahun yang sama oleh Jack Dorsey, Noah Glass, Biz Stone, dan Evan Williams [29]. Pada twitter terdapat fitur following dimana pengguna akun bebas melihat halaman twitter satu sama lain tanpa harus mengikuti. Selain itu media Twitter mempunyai fasilitas mengirim teks info, gambar, atau video, yang memungkinkan para pengguna dapat mengirim dan membaca pesan tersebut. Bagi pengguna Twitter yang ingin mengakses dan menggunakan informasi pada jejaring sosial tersebut, Twitter menyediakan Application Programming Interface (API) [16].

1.3. Text Mining

Text mining merupakan suatu proses menggali informasi dimana seorang user berinteraksi dengan sekumpulan dokumen menggunakan tools analisis yang merupakan komponen-komponen dalam data mining seperti kategorisasi informasi atau teks, mengelompokkan teks. Sumber data yang digunakan pada text mining diambil dari sekumpulan pola bahasa alami yang mampu menganalisis data teks semi-terstruktur dan tidak terstruktur [6]. Ada beberapa kategori yang termasuk ke dalam text mining salah satunya analisis sentimen yang mempelajari komputasi opini orang-orang, emosi, dan sentimen dan bertujuan untuk mengklasifikasikan sentimen pada suatu kalimat yang berupa opini atau pendapat, baik itu kalimat negatif ataupun positif [9].

1.4. Analisis Sentimen

Sentimen analysis atau opinion mining merupakan suatu proses pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistic dan text mining yang memiliki tujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian, dan emosi seseorang [9]. Kategori analisis sentimen yaitu mengelompokkan polaritas teks ke dalam suatu kalimat atau dokumen yaitu pengkategorian teks menjadi satu polaritas sentimen tertentu yang dianalisis apakah bersifat (positif, atau negatif, atau netral). Berdasarkan ruang lingkup teks tersebut [8].

1.5. Text Pre-Processing

Text preprocessing merupakan tahapan proses awal terhadap teks untuk mempersiapkan teks menjadi data yang akan diolah selanjutnya. Sekumpulan karakter yang bersambungan (teks) harus dipecah-pecah menjadi unsur yang lebih berarti, yang dapat dilakukan dalam tingkatan yang berbeda. Ada beberapa tahapan text preprocessing antara lain [13]:

- **Case Folding**

Proses mengubah semua huruf besar (upper case) dalam dokumen menjadi huruf kecil (lower case).

- **Cleansing**

Proses untuk membersihkan kata-kata yang tidak diperlukan untuk mengurangi noise pada proses klasifikasi.

- **Tokenizing**

Proses pemotongan kalimat menjadi sebuah kata, dengan memisahkan kata dan menentukan struktur dari setiap kata tersebut.

- **Normalization**

Proses perbaikan kata-kata yang masih salah dalam pengejaan atau merupakan singkatan berbentuk tertentu.

● **Stopwords**

Merupakan kumpulan kata-kata yang sering muncul dalam suatu dokumen, sehingga *stopword* digunakan untuk menghilangkan kata penghubung yang tidak begitu penting, tahapan *stopwords* adalah hasil dari *normalization*.

1.6. Machine Learning

Machine Learning (ML) atau pembelajaran mesin merupakan pendekatan AI yang digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan atau melakukan otomatisasi. Sesuai namanya, ML mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan menggeneralisasi [1]. Klasifikasi adalah metode dalam ML yang digunakan oleh mesin untuk memilah atau mengklasifikasi objek berdasarkan ciri tertentu sebagaimana manusia mencoba membedakan benda satu dengan yang lain. Sedangkan prediksi atau regresi digunakan oleh mesin untuk menerka keluaran dari suatu data masukan berdasarkan data yang sudah dipelajari dalam training [10].

1.7. Naive Bayes

Naive Bayes Classifier merupakan salah satu metode klasifikasi dengan probabilitas sederhana. Pada metode Naive Bayes Classifier menempuh dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan dilakukan proses terhadap sampel data yang sedapat mungkin dapat menjadi representasi data tersebut. Pada probabilitas prior bagi tiap kategori berdasarkan sampel data. Pada tahap klasifikasi ditentukan nilai kategori dari suatu data berdasarkan term yang muncul dalam data yang diklasifikasi. Persamaan teorema bayes dapat dinyatakan sebagai berikut [7]:

$$P(H|X) = \frac{P(X|H) * P(H) A}{P(H)}$$

Keterangan:

- X = Data dengan class yang belum diketahui
- H = Hipotesis data menggunakan suatu class yang spesifik
- P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (Parteori Probabilitas)
- P(H) = Probabilitas hipotesis H (Prior Probabilitas)
- P(X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) = Probabilitas H

1.8. Support Vector Machine

Algoritma Support Vector Machine (SVM) merupakan salah satu metode klasifikasi. Konsep dasar pada metode Support Vector Machine adalah memaksimalkan batas hyperplane yang memisahkan suatu set data. Support Vector Machine adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi linear dalam sebuah feature space (ruang fitur) berdimensi tinggi.

2. METODOLOGI PENELITIAN

2.1. Batasan Masalah Penelitian

Proses pengambilan data dilakukan melalui media Twitter, data yang diambil dari Twitter melalui akun publik bukan akun privat. Data yang diambil sebanyak 1700 data yang diambil dari sekumpulan tweet para pengguna Twitter yang mengandung kata BTS, Bangtan, Bighit dan Army. Yang akan diimplementasikan dengan algoritma Naive Bayes dan Support Vector Machine.

2.2. Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah membuat model machine learning menggunakan dataset dari Twitter dengan menggunakan metode algoritma Naive Bayes dan Support Vector Machine agar dapat mengklasifikasikan data teks seperti, komentar sentimen yang dikirimkan oleh penggemar/fans Korea Pop di indonesia dan masyarakat umum dengan melalui Twitter terhadap boy grup BTS. Dan menghitung akurasi tiap model untuk membandingkan performa dari tiap-tiap model.

2.3. Tinjauan Literatur

Pada tahun 2021, Peneliti dari Universitas Amikom Yogyakarta melakukan penelitian terhadap Analisis Sentimen Opini Terhadap Vaksin Covid-19 Pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes. Penulis mendapatkan hasil dari membandingkan dua algoritma tersebut

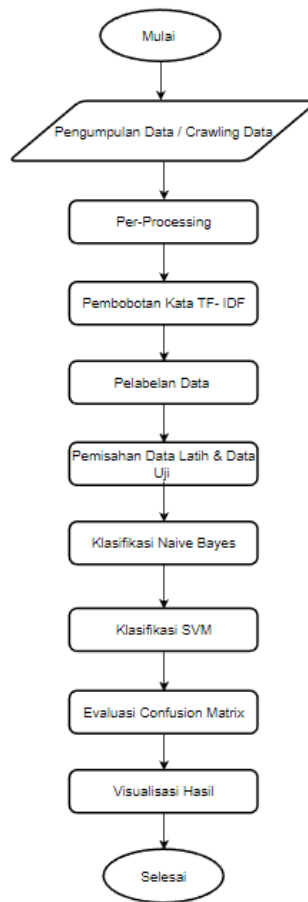
pada Algoritma SVM lebih unggul dalam hal akurasi, presisi dan recall dimana hasil yang didapatkan dari akurasi 90,47%, presisi 90,23% dan recall 90,78% karena proses pembobotan pada algoritma SVM lebih detail dan melakukan perulangan secara terus menerus untuk menentukan nilai bobot. Sedangkan Algoritma Naive Bayes lebih unggul dalam hal eksekusi waktu training dimana hasil akurasi 88,64%, presisi 87,32%, recall 88,13%, karena Algoritma Naive Bayes hanya melakukan satu kali perulangan untuk menentukan bobot yang tepat. Hal ini menyebabkan Naive Bayes lebih unggul dari segi waktu training dibandingkan SVM. Nilai masing-masing opini dengan menggunakan algoritma Naive Bayes dan SVM. Hasil dari Naive Bayes Netral 8,76%, Negatif 42,92%, dan Positif 48,32%, sedangkan hasil dari SVM Netral 10,56%, Negatif 41,28%, dan Positif 48,16%. Berdasarkan hasil yang diperoleh maka dapat dinyatakan bahwa opini positif tentang vaksin Covid-19 di Twitter lebih banyak dibandingkan opini negatif maupun yang netral [11].

Pada tahun 2022, Peneliti dari STMIK Borneo Internasional Balikpapan melakukan penelitian Analisis Sentimen Ibu Kota Negara Baru Menggunakan Metode Naive Bayes Classifier. Pada penelitian ini penulis mendapatkan Dataset sebanyak 1065 data, dari hasil tersebut mendapatkan nilai akurasi 94,18%, presisi 0,87%, dan recall 0,96%. Harmonized Recall and Precision atau yang juga dikenal dengan F1 - Score memiliki nilai 0.92% yang artinya sistem ini bagus dalam mendeteksi sentimen. Hasil dari analisis sentimen tersebut mendapatkan nilai Positif 289 dan nilai Negatif 776 sehingga metode Naive Bayes Mendapatkan nilai akurasi 94,18% [2].

Pada tahun 2022, Peneliti dari Universitas Majalengka melakukan penelitian Analisis Sentimen Pengguna Media Sosial Twitter Terhadap Wabah Covid-19 Dengan Metode Naive Bayes Classifier dan Support Vector Machine. Pada penelitian ini penulis mendapatkan dataset sebanyak 1652 data tweets. Pada metode Naive Bayes Classifier menghasilkan nilai akurasi dari pengklasifikasian sebesar 78,3%, sedangkan nilai akurasi yang diperoleh dari hasil klasifikasi yang didapatkan dari metode SVM sebesar 81,6%. Sementara hasil pengujian akurasi menggunakan metode Cross Validation dengan 10 K-Fold CV menghasilkan nilai rata-rata akurasi pada metode Naive Bayes Classifier sebesar 69,8% dan nilai rata-rata akurasi pada metode SVM sebesar 74,4%. Dari hasil perbandingan kedua nilai akurasi tersebut dapat disimpulkan bahwa metode SVM terbukti memiliki nilai akurasi yang lebih tinggi dari pada metode Naive Bayes Classifier [14].

2.4. Gambaran Umum Penelitian

Proses analisis sentimen yang dilakukan melalui jejaring media sosial Twitter, dilakukan untuk melihat tanggapan atau komentar Fans atau Penggemar Kpop yang berada di Indonesia terhadap Boy Grup BTS. Analisis sentimen ini menggunakan metode Naive Bayes dan Support Vector Machine untuk pengklasifikasian data tweets. Pada tahap analisis sentimen ini dibagi menjadi dua kelas yaitu kalimat positif atau negatif. Berikut ini tahapan penelitian analisis sentimen yang dapat dilihat pada gambar.



Gambar 1. Alur analisis

2.5. Spesifikasi Perangkat Keras dan Lunak

Perangkat keras atau hardware merupakan komponen penting yang digunakan untuk menyusun komputer sehingga dapat menjalankan tugasnya. Sedangkan perangkat lunak atau software merupakan program komputer yang digunakan untuk menjalankan tugas. Adapun spesifikasi perangkat keras dan perangkat lunak yang digunakan dalam penulisan ini dapat dilihat pada Tabel 1.

Tabel 1. Spesifikasi Perangkat Keras dan Lunak

Prosesor		AMD Ryzen 5 5500U with Radeon Graphics 2.10 GHZ
Memori RAM		16.0 GB
Sistem Operasi		Microsoft Windows 10 Home x64
Harddisk		516GB
Perangkat Penelitian	Lunak	Google Collabs, Python
Perangkat Dokumentasi	Lunak	Draw.io, OverLeaf Latex Editor

2.6. Dataset

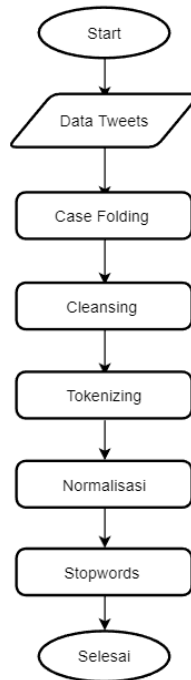
Pengambilan dataset merupakan komentar atau tweet para pengguna twitter melalui API Search Twitter yang menggunakan bahasa pemrograman Python. Setelah membuat akun pengembangan Twitter Developer, website akan memunculkan API Key, API Key Secret, Bearer Token, Access Token, dan Access Token Secret. Token ini yang akan digunakan untuk menyambungkan ke Twitter untuk melakukan pengambilan data tweet. Pengambilan dataset yang menggunakan Twitter ini, memakai kata kuncinya yaitu "BTS", "Bangtan", "Bighit", dan "Army". Dataset yang dikumpulkan sebanyak 1700 data yang diambil pada tanggal 2-3 September 2022 dan simpan kedalam file dengan format csv.

	id	date	username	tweet
0	1565878292208463873	2022-09-03 01:44:12+00:00	Emeraldie2	@etaerealte Soalnya kebanyakan army juga kek d...
1	1565878061186174977	2022-09-03 01:43:17+00:00	Sumantomumet	@btsfanbaseina @samsungID Misal kalau punya HP...
2	1565878010552930304	2022-09-03 01:43:04+00:00	Wint3rB3ar95	@yzooogi Klo ga hate Taehyung @BTS_twt ngapain ...
3	1565877207234654208	2022-09-03 01:39:53+00:00	Demianchael	Dan mereka masih bilang army b0ts 🤔👉 https://t...
4	1565877044420182017	2022-09-03 01:39:14+00:00	a_cherry_piece	jgn grgr rumor jd lupa streaming pink venom sm...

Gambar 2. Contoh Dataset

2.7. Pre-Processing

Tahap pre-processing merupakan tahap pembersihan data. Data tweet yang diambil dari media sosial Twitter merupakan data yang bentuknya tidak terstruktur. Beberapa tahapan preprocessing pada penelitian ini seperti, case folding, cleansing, tokenizing, normalization, dan stopwords.



Gambar 3. Alur Proses Pre-Processing

2.8. Pembobotan TF-IDF

Pada tahap ini merupakan proses pemberian bobot term pada dokumen. Term merupakan kata atau frasa yang digunakan untuk mengetahui konteks dari dokumen. Pembobotan ini yang nantinya akan digunakan oleh algoritma Machine Learning untuk klasifikasi dokumen.

2.9. Pelabelan

Konsep pelabelan akan dibagi menjadi dua kelas sentimen yaitu kelas sentimen positif dan kelas sentimen negatif. Nilai yang dihasilkan berupa nilai sentimen, jika nilai sentimen dibawah 0 (sentimen < 0) dinyatakan sebagai kelas negatif, sedangkan nilai sentimen diatas 0 (sentimen ≥) dinyatakan sebagai kelas positif.

	username	stopwords	label
0	Emeraldie2		1
1	Sumantomumet	ungu pertunjukan pamer keren banget	0
2	Wint3rB3ar95	benci tweet ngapain tweet buta dapat dukungan	1
3	Demianchael	bilang bts	1
4	a_cherry_piece	gara rumor lupa nonton sama mengeluarkan	0

Gambar 4. Hasil Pelabelan

2.10. Pembagian Data Latih dan Data Uji

Sebelum melakukan proses pelatihan model machine learning, dataset akan dibagi menjadi dua bagian terlebih dahulu. Dataset akan dibagi menjadi data latih dan data uji. Data uji ini digunakan untuk melihat persentase algoritma klasifikasi yang dilakukan dengan benar. Dalam tahap analisis ini banyaknya data latih 80% dan data uji 20% untuk masing-masing data.

2.11. Klasifikasi Naive Bayes

Proses klasifikasi Naive Bayes menggunakan fungsi MultinomialNB. MultinomialNB ini digunakan karena jenis dari klasifikasi Naive Bayes ini mampu mengolah data berupa teks. Setelah melakukan pengujian akurasi, berdasarkan perhitungan akurasi yang dilakukan, didapatkan hasil bahwa model dapat memprediksi dengan akurasi MultinomialNB yaitu mencapai 79%. Output dapat dilihat pada gambar 5.

MNB Accuracy Score --> 79.0909090909091

Gambar 5. Hasil Akurasi Naive Bayes

2.12. Klasifikasi Support Vector Machine

Proses klasifikasi SVM ini menggunakan metode kernel linear untuk melihat nilai akurasi yang di dapatkan. Setelah melakukan pengujian akurasi, berdasarkan perhitungan akurasi yang dilakukan, didapatkan hasil bahwa model dapat memprediksi dengan akurasi kernel yaitu mencapai 81%. Output dapat dilihat pada gambar 6.

SVM Accuracy Score --> 80.60606060606061

Gambar 6. Hasil Akurasi SVM

2.13. Confusion Matrix

Confusion matrix ini digunakan untuk mengevaluasi hasil dari prediksi dari metode klasifikasi, terdapat dua perintah yang digunakan untuk confusion matrix karena melakukan perbandingan dua algoritma. Hasil perbandingan yang di dapat dari kedua algoritma tersebut dapat dilihat pada tabel dibawah ini.

Tabel 2. Perbandingan Hasil Confusion Matrix

	Naive Bayes		SVM	
	Actual Positif	Actual Negatif	Actual Positif	Actual Negatif
Predicted Positif	257	4	252	9
Predicted Negatif	65	4	55	14
Accuracy	79%		81%	

Tabel 3. Hasil Perbandingan Nilai Evaluasi Model Naives Bayes dan SVM

	Accuracy	Precision	Recall	F1-Score
Naive Bayes	79%	80%	98%	88%
SVM	81%	82%	97%	89%

3. HASIL DAN PEMBAHASAN

3.1. Sentimen Positif

Data sentimen positif adalah hasil pelabelan yang masuk ke dalam kelas positif menggunakan analisis sentimen. Data sentimen positif tersebut diidentifikasi berdasarkan banyaknya frekuensi kata dalam dataset

tweet. Berikut adalah hasil visualisasi sentimen positif yang didapatkan dari komentar/tanggapan yang didapatkan melalui media Twitter.



Gambar 7. Hasil Sentimen Positif

3.2. Sentimen Negatif

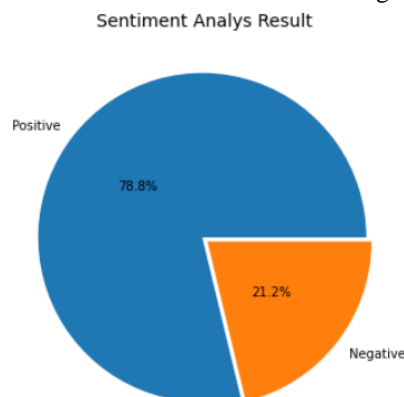
Data sentimen negatif adalah hasil pelabelan yang masuk ke dalam kelas negatif menggunakan analisis sentimen. Data sentimen negatif tersebut diidentifikasi berdasarkan banyaknya frekuensi kata dalam dataset tweet. Berikut adalah hasil visualisasi sentimen negatif yang didapatkan dari komentar/tanggapan yang didapatkan melalui media Twitter.



Gambar 8. Hasil Sentimen Negatif

3.3. Diagram Lingkaran (Pie Chart)

Diagram lingkaran atau Pie Chart digunakan untuk membandingkan tanggapan/komentar kalimat sentimen positif dan negatif. Hasil dari perbandingan tanggapan/komentar yang di dapatkan dari Twitter pada kalimat sentimen positif mendapatkan 78.8% dan kalimat sentimen negatif mendapatkan 21.2%.



Gambar 8. Diagram Lingkaran

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Berdasarkan Penelitian yang telah dilakukan penulis mengambil kesimpulan bahwa hasil dari analisis sentimen dengan menggunakan Machine Learning yang membandingkan 2 algoritma yaitu Naive Bayes dan Support Vector Machine telah berhasil dilakukan. Pengumpulan data untuk analisis sentimen dengan menggunakan media Twitter didapatkan sebanyak 1700 data yang diambil pada tanggal 2-3 September 2022. Setelah melalui tahap Preprocessing jumlah data yang digunakan untuk analisis sentimen sebanyak 1648 data. Pada analisis sentimen ini didapatkan hasil analisis sentimen positif sebesar 78.8% dan analisis sentimen negatif sebesar 21.2%. Kemudian algoritma yang mendapatkan hasil akurasi tertinggi pada analisis yang dilakukan adalah algoritma Support Vector Machine dengan nilai akurasi sebesar 81%. Sedangkan untuk algoritma Naive Bayes menghasilkan nilai akurasi sebesar 79%.

4.2. Saran

Penelitian yang dilakukan penulis masih banyak kekurangan yang mungkin dapat diperbaiki pada penulisan selanjutnya. Untuk mendapatkan hasil yang lebih baik, diharapkan pada penelitian selanjutnya, dapat menambah jumlah dataset yang lebih banyak, agar lebih banyak informasi yang akan didapatkan untuk hasil analisis sentimennya. Selain itu pada penelitian ini dapat menggunakan algoritma yang berbeda untuk melakukan perbandingan dan juga dapat menambahkan algoritma yang lain untuk melakukan perbandingan dengan algoritma Naive Bayes dan Support Vector Machine sehingga dapat menghasilkan nilai akurasi yang lebih baik.

DAFTAR PUSTAKA

- [1]. Ahmad, A. Mengenal artificial intelligence, machine learning, neural network, dan deep learning. <https://docplayer.info/62490785-Mengenal-artificialintelligence-machine-learning-neural-network-dan-deep-learning.html> , 2017.
- [2]. Aryanti, D. Analisis sentimen ibukota negara baru menggunakan metode naive bayes classifier. <https://ejurnal.seminarid.com/index.php/josh/article/view/1944/1189> , 2022.
- [3]. Azmi, H. T. Bts sebagai agensi besar industri k-pop dan relasinya dengan eksternal relations army indonesia. <https://www.researchgate.net/publication/361340303> , 2022.
- [4]. Cindie S. Feroza, D. M. Penggunaan media sosial instagram pada akun @yhoophii_official sebagai media komunikasi dengan pelanggan. <https://journal.binadarma.ac.id/index.php/jurnalinovasi/article/view/1397> , 2020.
- [5]. Delima Ayu Wulandari, Rd. Rohmat Saedudin, R. A. Analisis sentimen media sosial twitter terhadap reaksi masyarakat pada ruu cipta kerja menggunakan metode klasifikasi algoritma naive bayes. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15883/15596> , 2021.
- [6]. Deni Rusdian, D. R. Analisa sentimen terhadap tokoh publik menggunakan metode naive bayes classifier dan support vector machine. <https://jurnal.unimed.ac.id/2012/index.php/cess/article/view/13796> , 2019.
- [7]. Deni Rusdian, D. R. Analisa sentimen terhadap tokoh publik menggunakan metode naive bayes classifier dan support vector machine. <https://jurnal.unimed.ac.id/2012/index.php/cess/article/view/13796> , 2019.
- [8]. Dwi Widiastuti, Isram Rasal, D. W. A. P. Sentiment analysis of product reviews data on tokopedia by comparing the performance of classification algorithms. <https://garuda.kemdikbud.go.id/documents/detail/2903832> , 2022.
- [9]. Dyah Auliya .A, Sri Subanti, E. Z. Implementasi text mining pada analisis sentimen pengguna twitter terhadap marketplace di indonesia menggunakan algoritma support vector machine. <https://jurnal.uns.ac.id/ijas/article/view/44337> , 2020.
- [10]. Effendi, Rina Noviana. Perancangan Web Aplikasi Analisis Sentimen Media Sosial Twitter Dengan Metode Valence Aware Dictionary and Sentiment Reasoner (VADER) Menggunakan PHP dan MySQL pada Pemerintah Kota Bekasi. Jurnal Ilmiah KOMPUTASI, Volume 20 No : 1, Maret 2021, p-ISSN 1412-9434/e-ISSN 2549-7227. <https://ejournal.jakstik.ac.id/index.php/komputasi/article/view/369> , 2021
- [11]. Fermansah, D. Penggunaan metode traditional transformations data augmentation untuk peningkatan hasil akurasi pada model algoritma convolutional neural network (cnn) di klasifikasi gambar. <http://repositori.unsil.ac.id/233> , 2019.

- [12]. Frizka Fitriana, Ema Utami, H. A. F. Analisis sentimen opini terhadap vaksin covid-19 pada media sosial twitter menggunakan support vector machine dan naive bayes. <https://journal.unimma.ac.id/index.php/komtika/article/view/5185/2479>, 2021.
- [13]. Gare, O. B. Makalah pemrograman komputer sejarah perkembangan python. <https://pdfcoffee.com/download/makalah-pemrograman-doc-pdf-free.html>, 2018.
- [14]. Hakim, A. Klasifikasi sentimen terhadap bukalapak dengan menggunakan metode naïve bayes classifier. <https://repository.uin-suska.ac.id/14251/> , 2018.
- [15]. Harun Sujadi, Sandi Fajar, C. R. Analisis sentimen pengguna media sosial twitter terhadap wabah covid-19 dengan metode naive bayes classifier dan support vector machine. <https://ejournal.unma.ac.id/index.php/infotech/article/download/1883/1379> , 2022.
- [16]. Hutama, D. Perbandingan analisis sentimen pendidikan di indonesia selama covid-19 pada media sosial twitter dengan metode naïve bayes, knn, dan svm. <https://kc.umh.ac.id/15880/> , 2021.
- [17]. Maulizon, O. S. Klasifikasi sentimen masyarakat terhadap rokok pada twitter menggunakan naïve bayes classifier. <https://repository.uin-suska.ac.id/16689/> , 2018.
- [18]. Muhammad N. Muttaqin, I. K. Analisis sentimen aplikasi gojek menggunakan support vector machine dan k nearest neighbor. <http://lib.unnes.ac.id/50502/> , 2021.
- [19]. Nugroho, D. A. Promosi k-pop dalam media baru di amerika serikat. <https://eprints.umm.ac.id/74317/> , 2021.
- [20]. Purnomo, S. Opini wartawan terhadap faktor-faktor penghambat dalam peliputan berita studi pada wartawan sriwijaya post. <http://repository.radenfatah.ac.id/4753/> , 2019.