

METODE BERNOULLI NAIVE BAYES UNTUK ANALISIS SENTIMEN TERHADAP LESBIAN, GAY, BISEKSUAL DAN TRANSGENDER (LGBT) PADA MEDIA SOSIAL TWITTER

Anggun Safira Kusmindasari, Wahyu Syaifullah Jauharis Saputra, Hendra Maulana

Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur
Jl. Rungkut Madya No. 1, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur
anggunsafira6@gmail.com

ABSTRAK

Saat ini teknologi berkembang dengan sangat pesat dan mempermudah pengguna untuk mengakses berbagai macam informasi. Salah satu media yang sering digunakan untuk mencari informasi selain google adalah twitter. Twitter merupakan salah satu media sosial yang banyak digunakan untuk menyampaikan pendapat dan mencari informasi berita terbaru. Salah satu pembahasan topik yang masih dibahas hingga saat ini adalah mengenai Lesbian, Gay, Biseksual dan Transgender (LGBT). Ulasan atau tweet pada pembahasan topik LGBT memuat banyak sentimen dari pengguna twitter. Untuk mendeteksi ulasan atau tweet yang dikirimkan pengguna pada twitter penting untuk dilakukan suatu teknik pengolahan data yang disebut dengan teknik analisis sentimen. Analisis sentimen menggunakan algoritma *Bernoulli Naive Bayes* yang akan di gunakan pada tweet atau ulasan pada twitter dalam penelitian ini. Tujuan penelitian ini untuk mengukur performa dari algoritma *Bernoulli Naive Bayes* dalam pengklasifikasian sentimen menjadi dua kelas berupa sentimen positif dan sentimen negatif. Sebanyak 2.000 data tweet pada twitter digunakan pada penelitian ini. Hasil pengujian menunjukkan *Bernoulli Naive Bayes* menghasilkan nilai akurasi 90,52%.

Kata kunci : Analisis sentimen, Bernoulli Naive Bayes, Twitter, LGBT

1. PENDAHULUAN

Teknologi informasi kini berkembang sangat pesat mempermudah mengakses berbagai macam informasi melalui berbagai platform salah satu platform yaitu sosial media yang terdiri berbagai macam dan salah satunya Twitter. Twitter merupakan sebuah situs web yang dapat dioperasikan oleh Twitter Inc [1]. Pengguna aktif harian twitter di Indonesia tergolong cukup banyak. Pengguna dapat menulis pesan berdasarkan topik dengan menggunakan tanda # (tagar). Topik yang banyak dibahas akan masuk dalam trending, banyak tanggapan yang disampaikan pada topik yang sedang trending sehingga dari sanalah dapat dilihat berbagai macam persepsi setiap orang berbeda-beda. Oleh karena itu penelitian ini dilakukan untuk memahami tweet atau mendeteksi sentimen yang di kirim pengguna twitter pada aplikasi, penting untuk dilakukan suatu pengolahan data yang melihat bagaimana sentimen masyarakat terhadap suatu topik. Analisis sentimen merupakan proses memahami, mengekstrak dan mengelola data tekstual secara otomatis untuk mendapatkan informasi sentimen. Tugas dasar analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen kalimat pendapat yang dikemukakan bersifat positif, negatif atau netral.

Penelitian terdahulu dilakukan oleh Dea dkk tentang Penerapan Bernoulli Naive Bayes untuk Analisis Sentimen Pengguna Twitter Terhadap Layanan Online Food Delivery di Indonesia memperoleh hasil akurasi sebesar 89% [2]. Lalu ada penelitian lainnya yang dilakukan oleh Enggar dan Hartatik tentang Analisis Sentimen Pengguna Twitter

Terhadap Layanan Provider IndiHome Menggunakan Metode Naive Bayes memperoleh nilai akurasi 81,33% [3].

Pada beberapa penelitian sebelumnya makan dilakukan penelitian ini untuk melihat keakurasi menggunakan Bernoulli Naive Bayes apa bisa diatas 90% dengan menggunakan data sebanyak 2.000 data tweet.

2. TINJAUAN PUSTAKA

2.1. Twitter

Twitter merupakan sebuah media sosial yang dapat digambarkan sebagai graph berarah yang artinya pengguna dapat mengikuti pengguna lain namun pengguna kedua tidak perlu untuk mengikutinya kembali. Indonesia diklaim menjadi salah satu megar yang pengguna aktif harian twitter paling besar berdasarkan finansial twitter kuartal ke-3 tahun 2019 [4]. Fokus utama twitter adalah pengamatan atau pengalaman orang terhadap peristiwa terkini yang dapat dengan mudah dibagikan pada pengguna lain [5].

2.2. Analisis Sentimen

Analisis sentimen adalah proses memahami, mengekstrak, dan mengelola data tekstual secara otomatis. Analisis sentimen dapat juga digunakan untuk menentukan kecenderungan suatu sikap terhadap sebuah sentimen yaitu sentimen positif atau sentimen negatif.

2.3. Preprocessing

Preprocessing merupakan suatu proses tahapan pada *text mining* yang bertujuan untuk membersihkan data agar siap untuk digunakan. *Preprocessing* memiliki tahapan yaitu *case folding*, *cleaning*, *tokenizing*, *normalization*, *stopword removal*, dan *stemming*. Data yang sudah melalui tahap *preprocessing* akan menjadi lebih bersih dan terstruktur sehingga meningkatkan kualitas analisis dan pemodelan pada data tersebut [6].

2.4. Term Frequency Inverse Document Frequency (TF IDF)

TF IDF adalah pembobotan kata yang akan digunakan dalam analisis data. TF adalah jumlah kata atau *term* tertentu yang ada dalam sebuah dokumen. IDF adalah frekuensi kemunculan kata pada seluruh dokumen [7]. Berikut rumus pembobotan kata dengan TF IDF sebagai berikut.

$$W_{j,i} = \frac{n_{j,i}}{\sum_k n_{k,i}} \cdot \log_2 \frac{D}{d_j} \quad (1)$$

Keterangan :

$W_{j,i}$: Pembobotan TF-IDF untuk kata (*term*) ke- j pada dokumen ke- i .

$n_{j,i}$: Jumlah kemunculan kata (*term*) ke- j pada dokumen ke- i .

$\sum_k n_{k,i}$: Jumlah kemunculan seluruh kata (*term*) pada dokumen ke- i .

D : Total dokumen.

d_j : Banyaknya dokumen yang mengandung kata (*term*) ke- j .

2.5. Bernoulli Naïve Bayes

Bernoulli naïve bayes merupakan salah satu metode dari *Naive bayes*. Model *Bernoulli naïve bayes* diaplikasikan pada data biner dan model ini memperhatikan kemunculan kata dalam dokumen dan mengabaikan frekuensi dari data yang muncul [8]. Berikut rumus *Bernoulli Naive Bayes*.

$$P(d|c) = \prod_{1 \leq i \leq M} (B_{it} P(e_i|c) + ((1 - B_{it})(1 - P(e_i|c))) \quad (2)$$

Keterangan :

$P(d|c)$: *Conditional probability*

$P(c)$: *Prior probability*

$P(e_i|c)$: Probabilitas dalam dokumen kategori c

2.6. Confusion Matrix

Confusion Matrix merupakan suatu metode yang digunakan untuk mengukur kinerja suatu metode klasifikasi. Performa dari sistem tersebut biasanya dievaluasi menggunakan data dalam sebuah *matrix*. Matriks dibagi menjadi empat bagian yang masing-masing mewakili kombinasi dua kemungkinan yaitu kelas yang di prediksi benar dan salah [9].

Tabel 1. Confussion Matrix

Kelas Sebenarnya	Kelas Prediksi	
	1 (+)	2 (-)
1 (+)	<i>True positif</i>	<i>False negative</i>
2 (-)	<i>False positif</i>	<i>True negative</i>

Keterangan :

- True Positive* merupakan jumlah data dengan kelas positif yang diklasifikasikan positif.
- True Negative* merupakan jumlah data dengan kelas negative yang diklasifikasikan negative.
- False Positive* merupakan jumlah data dengan kelas positif yang diklasifikasikan negative.
- False Negative* merupakan jumlah data dengan kelas negative yang diklasifikasikan positif.

Untuk pengukuran performa klasifikasi cara yang digunakan menggunakan rumus sebagai berikut :

$$Akurasi : \frac{TP+TN}{TP+TN+FN+FP} \times 100\% \quad (3)$$

$$Precision : \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$Recall : \frac{TP}{TP+FN} \times 100\% \quad (5)$$

$$F1 - score : 2x \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (6)$$

3. METODE PENELITIAN

Pada penelitian ini dilakukan dengan mengikuti langkah-langkah yang terstruktur agar dapat dilakukan secara baik dan terarah.

3.1. Identifikasi Masalah

Sebelum melakukan penelitian terlebih dulu untuk mengidentifikasi permasalahan yang terjadi mengenai topik yang diangkat dalam penelitian. Tujuan dari indentifikasi masalah adalah menemukan area yang dibutuhkan pemecah atau pemahaman lebih dalam pada penelitian [10].

3.2. Studi Literatur

Studi literatur merupakan salah satu langkah awal penelitian, tujuan utama studi literatur adaah untuk memahami topik penelitian dan membangun dasar teori dan metodologi yang tepat [11]. Sumber litelatur yang digunakan antara dari jurnal nasional ataupun internasional. Sumber literatur juga berfungsi untuk memperjelas teori atau istilah yang digunakan pada penelitian.

3.3. Pengumpulan Data

Pengumpulan data diambil dari media sosial twitter menggunakan kata kunci lgbt dan transgender. Data yang diperoleh sekitar 2.000 data tweet, data dibagi menjadi data latih dan data uji.

3.4. Analisis Proses

Analisis proses meliputi tahapan *preprocessing*, pembobotan TF IDF dan mengkalsifikasian dengan bernoulli naïve bayes. Tahapan *preprocessing* meliputi sebagai berikut.

- Case folding* : pengubahan pada kata yang menggunakan huruf kapital diubah menjadi huruf kecil.

- b. *Cleaning* : pembersihan kalimat dari kata yang tidak diperlukan untuk mengurangi *noise*.
- c. *Normalization* : mengubah kata menjadi bentuk yang baku.
- d. *Tokenizing* : pemotongan atau pemisahan deretan kalimat menjadi kata yang tersusun menjadi potongan kata tunggal.
- e. *Stopword removal* : menghilangkan kata yang tidak relevan.
- f. *Stemming* : penghilangan imbuhan dari sebuah kata sehingga menghasilkan kata dasar.

Selanjutnya data diubah menjadi data numerik menggunakan metode TF IDF agar data dapat dianalisis oleh algoritma klasifikasi. Setelah melalui proses TF IDF tahap berikutnya proses klasifikasi menggunakan algoritma *Bernoulli Naive Bayes*.

3.5. Evaluasi

Langkah selanjutnya setelah model analisis sentimen berhasil dibuat adalah evaluasi model menggunakan *confussion matix*. Hasil proses ini meliputi akurasi, *presicion*, *recall* dan *f1-score*.

4. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan bahasa pemrograman *pyhton* dengan mengimplementasikan

metode *Bernoulli Naïve Bayes* untuk analisis sentimen LGBT pada twitter.

4.1. Pengumpulan Data



Gambar 1. Pengambilan Data

Data yang digunakan merupakan data yang diambil atau *crawling* pada twitter. Data diambil dari tanggal 6 Agustus 2022 – 27 September 2023. Jumlah data tweet yang digunakan 2.000. Proses labeling data dilakukan secara manual menjadi sentimen positif atau pro dan sentimen negatif atau kontra. Notasi labeling untuk sentimen positif/pro adalah 1 dan sentimen negatif/kontra adalah 0.

4.2. Preprocessing

Tahapan *preprocessing* dilakukan agar data menjadi terstruktur. Proses *preprocessing* melalui beberapa tahapan beberapa diantaranya *case folding* atau mengubah menjadi huruf kecil dan *stemming* atau mengubah kata dasar [12], untuk contoh hasil proses *preprocessing* dapat dilihat pada table dibawah :

Tabel 1. Contoh Preprocessing

Tahapan	Hasil
Tweet	Kami wajib melindungi anak-anak dari promosi LGBT, termasuk di Citayam Fashion Week. Kami juga wajib menyediakan ruang untuk anak-anak kita berekspresi. Bersama DPRD DKI Jakarta, ruang terbuka yang keren akan terus kami perbanyak dan perjuangkan
Case folding	kami wajib melindungi anak-anak dari promosi lgbt, termasuk di citayam fashion week. kami juga wajib menyediakan ruang untuk anak-anak kita berekspresi. bersama dprd dki jakarta, ruang terbuka yang keren akan terus kami perbanyak dan perjuangkan
Cleaning	kami wajib melindungi anakanak dari promosi lgbt termasuk di citayam fashion week kami juga wajib menyediakan ruang untuk anakanak kita berekspresi bersama dprd dki jakarta, ruang terbuka yang keren akan terus kami perbanyak dan perjuangkan
Normalization	kami wajib melindungi anakanak dari promosi lgbt termasuk di citayam fashion week kami juga wajib menyediakan ruang untuk anakanak kita berekspresi bersama dprd dki jakarta, ruang terbuka yang keren akan terus kami perbanyak dan perjuangkan
Tokenizing	[kami, wajib, melindungi, anakanak, dari, promosi, lgbt, termasuk, di, citayam, fashion, week, kami, juga, wajib, menyediakan, ruang, untuk, anakanak, kita, berekspresi, bersama, dprd, dki, jakarta, ruang, terbuka, yang keren, akan, terus, kami, perbanyak, dan, perjuangkan]
Stopword removal	[wajib, melindungi, anakanak, promosi,lgbt, citayam, fashion, week, wajib, menyediakan, ruang, anakanak, berekspresi, bersama, dprd, dki, jakarta, ruang, terbuka, keren, perbanyak, perjuangan]
Stemming	[wajib, lindungi, anakanak, promosi, lgbt, citayam, fashion, week, wajib, sedia, ruang, anakanak, ekspresi, sama, dprd, dki, jakarta, ruang, buka, keren, banyak, uang]

4.3. TF IDF

Setelah data selesai di proses pada tahapan *stemming* dilanjutkan dengan proses perubahan kedalam bentuk numerik dengan TF IDF. Hasil TF IDF dapat dilihat pada gambar berikut.

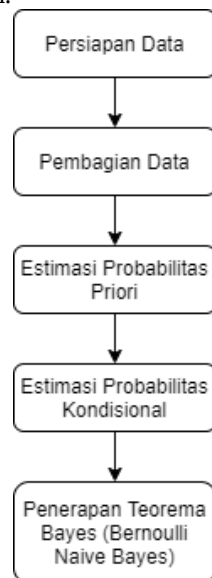
[illegible]

Gambar 2. Hasil TF IDF

4.4. Klasifikasi Bernoulli Naïve Bayes

Pada tahap klasifikasi *Bernoulli naive bayes* dilakukan pengumpulan data lebih dulu untuk bahan penelitian, dilanjutkan dengan pelabelan data dengan menggunakan variabel biner 0 dan 1. Lalu data dibagi menjadi dua, yaitu sebagai data latih dan data uji pada penelitian ini data dibagi menjadi 80% data latih dan 20% data uji. Setelah pembagian data, perhitungan probabilitas priori dan probabilitas kondisional (*likelihood*) dilakukan untuk menunjukan seberapa besar kemungkinan kelas muncul dalam dataset.

Tahap berikutnya penerapan teorema *bayes* untuk pengklasifikasian. _____



Gambar 3. Proses Klasifikasi

4.5. Word Cloud

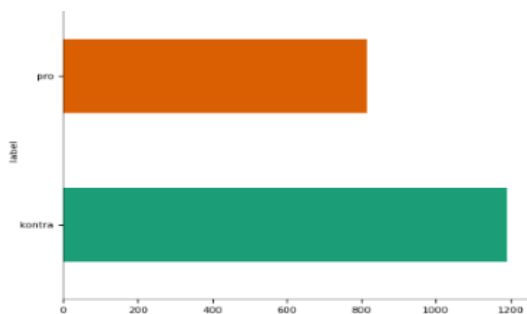
Word cloud merupakan sebuah representasi visual dari kata-kata yang sering muncul dalam sebuah kumpulan data. *Word cloud* sering digunakan untuk menunjukkan kata kunci dalam suatu teks.



Gambar 4. *Word cloud*

4.6. Evaluasi

Hasil dataset setelah melewati tahapan *preprocessing* menunjukkan bahwa jumlah sentimen positif atau pro mendukung sebanyak 800 data dan sentimen negative atau kontra menolak sebanyak 1.200 data, yang mana pembagian data tersebut dari daset yang berjumlah 2.000 data tweet. Untuk hasil dapat dilihat pada gambar berikut.



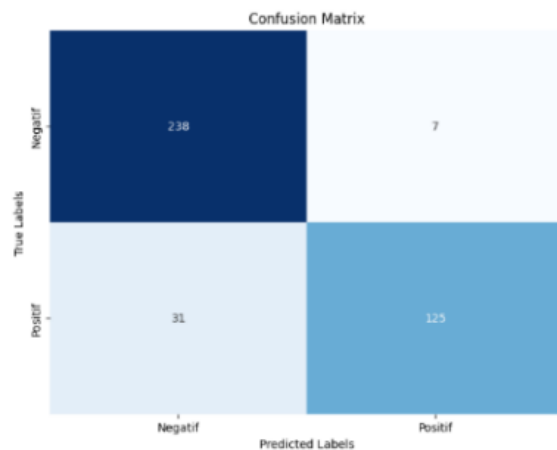
Gambar 5. Jumlah sentimen positif/pro dan negative/kontra

Tahap metode evaluasi yang digunakan menggunakan *confussion matrix* untuk mengetahui nilai akurasi, precision, *recall*, dan f1-score. Untuk hasil prediksi dan label asli diambil secara acak pada data uji untuk hasil dapat dilihat pada gambar berikut.

```
Prediksi: [1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0
0 0 1 1 1 0 0 0 0 1 0 0 0 1 1 0 0 1 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0
1 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 1 0 1 1 1 0
0 1 0 1 0 1 1 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 1 1 0 0 0 0 0 0 0
1 1 1 0 1 1 0 1 1 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 1 0 1 1 0 0 0 0 0 0 0 0 1
0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 1 1 0 0 0 1
0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0
0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 1 1 0 0 1 1 0 0 1 0 0 0 0 0 1 0 1 1 1 0 1 0 0 0 0 0 0 1 1 0 0 0
0 1 1 0 1 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 1 1 0 1 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0]
Label asli: [1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 0 1 1 0 0 0 0 0 0 0 1 1
0 0 1 1 1 0 0 0 0 0 1 0 0 0 1 1 1 0 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0
1 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 1 0 1 1 1 0
0 1 0 1 0 1 1 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 1 1 0 0 0 0 0 0 0 0 0
1 1 1 0 1 1 0 1 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 1 1 0 0 0 1 1 0 0 0 1 1 0 0 0 1
0 0 0 0 0 1 1 0 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 1 1 0 0 0 1
0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 1 0 0 1 1 0 1 1 0 0 0 0 0 0 0
0 0 0 0 0 1 1 0 0 1 1 0 0 0 0 0 1 0 0 1 1 1 1 0 0 0 0 1 1 1 0 1 1 0 0 0
0 0 1 0 1 1 0 1 0 1 1 0 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 1 1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 1 1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0]
```

Gambar 6. Hasil Prediksi Data Uji

Setelah model dilatih menggunakan data latih untuk evaluasi kinerja menggunakan data uji. Pada data uji yang diimplementasikan menghasilkan *confussion matix* seperti gambar berikut.



Gambar 7. Hasil *Confussion Matrix*

Dengan menggunakan *confusion matrix* diatas kita dapat menghitung akurasi, presisi, *recall* dan *f1-score* sebagai berikut.

$$Akurasi : \frac{238 + 125}{238 + 31 + 125 + 7} \times 100\% = 90,52\%$$

$$Precision : \frac{125}{125 + 7} \times 100\% = 94,7\%$$

$$Recall : \frac{125}{125 + 31} \times 100\% = 80,13\%$$

$$F1 - score : 2x \frac{0,947 \times 0,8013}{0,947 + 0,8013} \times 100\%$$

$$= 2x \frac{0,7588311}{1,7483} \times 100\% = 86,81\%$$

Accuracy = 90.52 %
 Recall = 80.13 %
 Precision = 94.7 %
 F1-Score = 86.81 %

Gambar 8. Hasil Perhitungan *Confussion Matrix*

Berdasarkan perhitungan diatas pada data uji didapatkan kinerja *Bernoulli Naïve Bayes* cukup baik dalam mengklasifikasikan sentimen. Perhitungan manual yang dilakukan dan perhitungan menggunakan sistem didapatkan nilai yang sama dengan akurasi 90,52%, precision 94,7%, recall 80,13%, dan *f1-score* 86,81%.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil yang didapat pada penelitian analisis sentimen pengguna twitter terhadap LGBT menggunakan metode *Bernoulli naive bayes* menunjukan tingkat akurasi 90,52% pada data uji. Dengan banyak data sebanyak 2.000 data tweet maka presentase akurasi meningkat dibanding dengan hanya menggunakan data 1.000 data yang hanya mendapat akurasi sekitar 70%. Untuk jumlah data sentimen positif atau pro mendukung berkisar 800 data, sedangkan untuk sentimen negatif atau kontra menolak berkisar 1.200 data. Dapat dilihat jika sentimen positif dan negative jumlah ulasan tidak begitu jauh berbeda. Maka dapat disimpulkan pada saat ini beberapa pengguna twitter mulai berani untuk menunjukan jati diri mereka dalam sosial media. Terlepas dari itu untuk analisis sentimen pada penelitian ini menggunakan metode *Bernoulli naive bayes* jumlah data sangat berpengaruh dengan tingkat akurasi. Saran untuk penelitian selanjutnya yang dapat diterapkan adalah memperluas cakupan variasi data, pelabelan data menggunakan metode lain serta penambahan algoritma pembandingan seperti *Logistic Regression*.

DAFTAR PUSTAKA

- [1] Kholilullah, M., & Hayati, U. "Analisis Sentimen Pengguna Twitter(X) Tentang Piala Dunia Usia 17 Menggunakan Metode Naive Bayes". *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no 1, Februari 2024.
- [2] Putri, D. F. A., Masjkur, M., & Indahwati. (2023). «Penerapan Bernoulli Naïve Bayes untuk Analisis Sentimen Pengguna Twitter Terhadap Layanan Online Food Delivery di Indonesia». *Journal of Statistics*, vol 12, no 1, pp 50–62, Januari 2023.
- [3] Hibbannuari, M. E. A., Hartatik. "Analisis Sentimen Pengguna Twitter Terhadap Layanan Provider IndiHome Menggunakan Algoritma". *Journal Automation Computer Information System* vol. 3, no 1, pp 19-26, Mei 2023.
- [4] Twitter Inc. (2019). Annual Report. https://www.annualreports.com/HostedData/AnnualReportArchive/t/NYSE_TWTR_2019.pdf.
- [5] Wang, J. Guo, C. Yuan, and B. Li, "Sentiment Analysis of Twitter Data," *Appl. Sci.*, vol. 12, no. 22, pp. 1–14, 2022, doi: 10.3390/app122211775.
- [6] M. D. Hendriyanto, A. A. Ridha dan U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *Journal Of Information Technology and Computer Science (IntecomS)*, pp. 1-7, 2022.
- [7] B. Herwijayanti, D. E. Ratnawati dan L. Muflikhah, "Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 1, pp. 306-312, 2017.
- [8] Dhuhiha, W. M. P., & Zone, F. "Perbandingan Kinerja Algoritma Multinomial dan Bernoulli Naïve Bayes dalam Mengklasifikasikan Komentar Cyberbullying". *Komputika : Jurnal Sistem Komputer*, vol 12, no 2, pp 109–117. 2023. <https://doi.org/10.34010/komputika.v12i2.9767>.
- [9] Sholihah. E. R. M., Diyas. I. G. S. M., Puspaningrum. E. Y. Perbandingan Kinerja Kernel Linear dan RBF Support Vector Machine Untuk Analisis Sentimen Ulasan Pengguna KAI Access Pada Google Play Store. *Jurnal Mahasiswa Teknik Informatika*. Vol 8, no 1. Februari 2024.
- [10] N. Chintalapudi, G. Battineni, and F. Amenta, "Sentimental Analysis of COVID-19 Tweets Using Deep Learning Models," *Infect. Dis. Rep.*, vol. 13, no. 2, pp. 329–339, 2021, doi: 10.3390/idr13020032.
- [11] R. Sistem et al., "Sentimen dan Pemodelan Topik Pariwisata Lombok," *J. RESTI (Rekayasa Sistem. dan Teknologi. Informasi)*, vol. 1, no. 10, pp. 123–131, 2021.
- [12] Vynska Amalia Permadi, "Analisis Sentimen Menggunakan Algoritma Naive bayes Terhadap Review Restoran di Singapura," *J. Buana Inform.*, vol. 11, pp. 141–151, 2020.