

SENTIMEN ANALISIS PADA MEDIA SOSIAL MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER (NBC)

Muhamad Imam Syafii
Informatika
syafiimam@gmail.com

Abstrak

Twitter merupakan salah satu situs microblogging yang memungkinkan penggunanya untuk menulis tentang berbagai topik. Menjelang pemilihan umum, para politisi atau tokoh publik sering memanfaatkan media sosial untuk berkampanye dan meningkatkan popularitas mereka. Hal tersebut dapat digunakan sebagai sumber data untuk menilai sentimen pada Twitter. Pengguna sering menggunakan singkatan kata dan ejaan kata yang salah, dimana dapat menyulitkan fitur yang diambil serta mengurangi ketepatan klasifikasi. Dalam penelitian ini penulis menerapkan proses *text mining* dan untuk mengklasifikasinya dengan menggunakan algoritma Naive Bayes Classifier. Penelitian ini menggunakan 450 data tweet tentang sentimen kepada akun twitter @Sandiuno. Data tersebut diklasifikasi secara manual dan dibagi kedalam masing-masing kategori, 150 untuk opini positif, 150 untuk opini negatif dan 150 untuk opini netral. Kemudian skenario pengujian dibagi menjadi 5 yaitu 50:50, 60:40, 70:30, 80:20 dan 90:10. Hasil penelitian ini menghasilkan tingkat akurasi yang baik pada rasio 70:30, dengan hasil pengujian mencapai 65,18% dengan menggunakan 135 data testing dan 315 data training.

Kata Kunci: *Twitter, Klasifikasi, Naive Bayes Classifier*

PENDAHULUAN

Media jejaring sosial seperti Twitter, Facebook dan Youtube merupakan beberapa media perangkat komunikasi terpopuler di masyarakat saat ini. Menjelang pemilihan umum, para politisi atau tokoh publik sering memanfaatkan media sosial untuk berkampanye dan meningkatkan popularitas mereka (A. D. Putri, 2021);(Ahmad et al., 2021);(Putra et al., 2009). Salah satu media jejaring sosial yang telah dimanfaatkan dalam pemilihan umum adalah Twitter (Sulistiani et al., 2021);(Dewi, 2021);(Very et al., 2021). Twitter telah dimanfaatkan dalam pemilihan umum di beberapa negara seperti Singapura, Jerman, dan Amerika (Puspaningrum et al., 2020);(Borman et al., 2018);(Permata & Rahmawati, 2018). Dalam implementasinya, media sosial lebih banyak digunakan untuk kegiatan jual beli, menyampaikan informasi, bahkan sebagai media untuk mengekspresikan diri (Hana et al., 2019);(Saputra & Permata, 2018);(Siwi & Puspaningtyas, 2020). Pertumbuhan media sosial sangat cepat, tidak hanya penggunaanya yang terus menerus meningkat, namun semakin banyaknya media sosial yang ditawarkan melalui aplikasi mobile ataupun website. Sehingga hal tersebut dimanfaatkan oleh para politisi atau tokoh publik untuk

menyampaikan pesan dan informasi kepada masyarakat (Wahyudi et al., 2021);(Ayu et al., 2021);(Samsugi et al., 2018). Kata yang terkandung dalam twitter adalah bahasa alami manusia yang merupakan bahasa dengan struktur kompleks. Informasi yang terdapat didalam twitter mayoritas adalah berupa data teks atau opini serta pendapat yang tertulis dalam kalimat-kalimat yang tidak terstruktur. Oleh karena itu diperlukan sebuah analisis yang tepat untuk mengelompokkan opini secara tepat sehingga memberikan informasi yang akurat dan memberikan suatu informasi berkualitas tinggi dari teks untuk pengklasifikasian atau kategorisasi. Analisis sentimen adalah studi komputasional dari opini-opini orang, penilaian dan emosi melalui entitas, event dan atribut yang dimiliki (Saputra & Febriyanto, 2019);(Fernando et al., 2021). Tugas dasar analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau fitur/tingkat aspek apakah pendapat yang dikemukakan dalam teks tersebut bersifat positif, negatif atau netral. Text mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen, dimana text mining merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data teks yang berjumlah besar. Salah satu metode klasifikasi yang dapat digunakan dalam melakukan text mining adalah metode Naive Bayes Classifier (NBC) (Febria Lina & Setiyanto, 2021);(Fakhrurozi et al., 2022);(Abidin et al., 2021). Kelebihan NBC yaitu merupakan metode yang sederhana tetapi memiliki tingkat akurasi tinggi. NBC terbukti dapat digunakan secara efektif untuk mengklasifikasikan teks secara otomatis. Algoritma NBC yang sederhana dan kecepatannya tinggi dalam proses pelatihan dan klasifikasi membuat algoritma ini menarik untuk digunakan sebagai salah satu metode klasifikasi.

KAJIAN PUSTAKA

Text Mining

Text mining (penambangan teks) adalah penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru, sesuatu yang tidak diketahui sebelumnya atau menemukan kembali informasi yang tersirat secara implisit, yang berasal dari informasi yang diekstrak secara otomatis dari sumber-sumber data teks yang berbeda-beda (Alita, 2021);(Hasri & Alita, 2022).

Sentiment Analysis

Sentiment analysis adalah mengekstraksi pendapat, sentimen, evaluasi dan emosi orang tentang suatu topik tertentu yang tertulis menggunakan teknik pemrosesan bahasa alami. Sejumlah karya-karya besar lainnya menyebutkan *sentiment analysis* fokus pada aplikasi spesifik yang mengklasifikasikan mengenai opini positif, negatif dan netral (Rahman Isnain et al., 2021);(Syah & Witanti, 2022);(Isnain et al., n.d.).

Preprocessing

Tahap *preprocessing* diperlukan untuk membersihkan data dari teks yang tidak diperlukan, dimana data teks yang tidak terstruktur akan diubah menjadi data teks yang terstruktur atau semi terstruktur (*Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis*, 2021);(Hendrastuty et al., 2021);(Styawati et al., 2021). Tahap dari *preprocessing* untuk mengolah data yaitu *case folding*, *cleansing*, *tokenizing*, *stopword removal* dan *stemming*.

1. Case Folding

Case folding merupakan tahapan merubah bentuk kata-kata menjadi sama bentuknya, baik semuanya menjadi *lower case* ataupun menjadi *upper*.

2. Cleansing

Cleansing merupakan proses membersihkan kata-kata yang tidak diperlukan untuk mengurangi noise (Sentiaji & Bachtiar, 2015). Kata yang dihilangkan adalah URL, *hashtag* (#), *username* (@username), email dan tanda baca (!@#\$%^&*()_+=.,<>?!|{}[]1234567890).

3. Tokenizing

Tokenizing berguna untuk memisahkan deretan kata menjadi beberapa urutan yang terpotong oleh spasi atau karakter spesial.

4. Stopword Removal

Stopword removal berperan untuk membuang kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansinya dengan teks.

5. Stemming

Stemming adalah tahapan untuk membuat kata yang berimbuhan kembali ke bentuk asalnya. Kata-kata yang muncul di dalam dokumen yang sering mempunyai banyak varian morfologik. Karena itu, setiap kata yang bukan *stopwords* direduksi ke *term* yang cocok untuk mendapatkan bentuk akarnya dengan menghilangkan awalan atau akhiran.

Twitter

Twitter merupakan suatu kumpulan kata yang berisikan maksimal 140 karakter. Didalam *text twitter*, masyarakat umum biasa menyebutkan dengan kata “*tweet*” atau kicauan. Direktur Jenderal Sumber Daya Perangkat Pos dan Informatika (SDPP) Kementerian Komunikasi dan Informatika menyatakan perkembangan dunia teknologi berkembang sangat pesat di dunia tak terkecuali Indonesia dengan mencapai peringkat ketiga di Asia untuk jumlah pengguna internet (Assuja & Saniati, 2016);(Setiawansyah et al., 2021);(Yuliana et al., 2021). Dari 245 juta penduduk Indonesia, pengguna internet di Indonesia mencapai 55 juta orang. Sejak data terakhir pada Desember 2011, tercatat jumlah pengguna internet di Indonesia mencapai 55 juta orang atau menguasai Asia sebesar 22,4% setelah Jepang. Tercatat Indonesia menjadi negara kelima terbesar pengguna *Twitter* di bawah Inggris dan negara besar lainnya, dengan jumlah pengguna sebanyak 19,5 juta pengguna.

Naive Bayes Classifier (NBC)

Algoritma *Naive Bayes Classifier* merupakan algoritma yang digunakan untuk mencari nilai probabilitas tertinggi untuk mengklasifikasi data uji pada kategori yang paling tepat (Alita et al., 2021);(Amalia et al., 2021);(S. eka Y. Putri & Surahman, 2019). Dalam penelitian ini yang menjadi data uji adalah dokumen opini Facebook. Ada dua tahap pada klasifikasi dokumen. Tahap pertama adalah pelatihan (*training*) terhadap dokumen yang sudah diketahui kategorinya. Sedangkan tahap kedua adalah pengujian (*testing*), yaitu proses klasifikasi dokumen yang belum diketahui kategorinya. Dalam algoritma *Naive Bayes Classifier* setiap opini direpresentasikan dengan pasangan atribut “ $x_1, x_2, x_3, \dots, x_n$ ” dimana x_1 adalah kata pertama, x_2 adalah kata kedua dan seterusnya. Sedangkan V adalah himpunan kategori opini (Gandhi et al., 2021);(Damuri et al., 2021);(Wibisono et al., 2020). Pada saat klasifikasi algoritma akan mencari probabilitas tertinggi dari semua kategori dokumen yang diujikan (V_{MAP}).

METODE

A. Pengumpulan Data

Pengumpulan data dilakukan untuk mendapatkan data yang sesuai mengenai informasi yang dibutuhkan peneliti (Budiman et al., 2021);(Sidiq & Manaf, 2020);(Eka Saputri, 2018). Data yang digunakan merupakan opini teks berbahasa Indonesia yang diambil

dari akun *Twitter* @sandiuno. Data terbagi atas opini positif, opini negatif dan opini netral. Sebagian data akan dijadikan data *training* dan sebagian sebagai data *testing*. Data *training* adalah data yang digunakan sebagai acuan untuk membangun model klasifikasi, sedangkan data *testing* adalah data yang digunakan untuk menguji performa dari model klasifikasi *Naive Bayes Classifier*. Jumlah data yang digunakan yaitu 450 dokumen opini, yang dibagi 150 opini positif, 150 opini negatif dan 150 opini netral. Dari 450 dokumen tersebut peneliti membagi skenario untuk proses pengujian antara data *training* dan data *testing* yaitu 50:50, 60:40, 70:30, 80:20 dan 90:10.

B. Analisis Data

Pada tahap ini dilakukan analisis data terhadap data opini yang diperoleh dari akun *Twitter* @sandiuno. Data opini nantinya akan melewati tahap *preprocessing* dan selanjutnya akan diklasifikasikan menurut kategorinya. Dalam penelitian ini, peneliti mengklasifikasikan opini ke dalam tiga kategori, yaitu opini positif, opini negatif dan opini netral. Setiap orang memiliki ciri khas dalam penulisan opini. Dari hasil pengamatan yang dilakukan pada akun *Twitter* @sandiuno, terdapat beberapa karakteristik dalam penulisan opini, seperti :

1. Jumlah rata-rata kata

Dari 450 data tersebut jumlah kata rata-rata yaitu 26 kata untuk kelas positif, 20 kata untuk kelas negatif dan 11 kata untuk kelas netral.

2. Data opini memiliki kata yang tidak baku

Penggunaan kata yang digunakan dalam memberikan opini tidak sesuai dengan kaidah bahasa Indonesia. Contoh kata yang tidak baku yang terdapat dalam opini tersebut diantaranya kata “gak” dan “ntar”.

3. Data opini terdapat singkatan

Terdapat beberapa tulisan opini yang penulisannya tidak lengkap atau disingkat. Contoh kata yang disingkat yaitu kata ‘kepada’, jakarta, ‘rakyat’, ‘cepat’, ‘mungkin’ menjadi kata ‘kpd’, ‘jkt’, ‘rky’, ‘cpt’, ‘mgkn’ dan lain sebagainya.

4. Makna kata opini

Dalam kelas negatif kata yang mengandung makna negatif diantaranya kata ‘munafik’, ‘sombong’, ‘kecewa’, ‘bohong’. Sedangkan untuk kelas positif yaitu ‘baik’, ‘keren’, ‘semangat’ dan untuk kelas netral seperti ‘sehat’, ‘lucu’, ‘sabar’ dan sebagainya.

5. Data opini terdapat kesalahan dalam pengetikan opini

Kesalahan yang terjadi dalam penulisan opini seperti penulisan kata bapak menjadi “bapka”.

C. Rancangan Arsitektur

a. Rancangan Arsitektur *Training*

Rancangan arsitektur *training* bertujuan untuk melakukan proses mencari nilai probabilitas yang nantinya dapat digunakan dalam proses *testing*. Berikut ini merupakan penjelasan dari rancangan arsitektur *training* :

1. Data *Training*

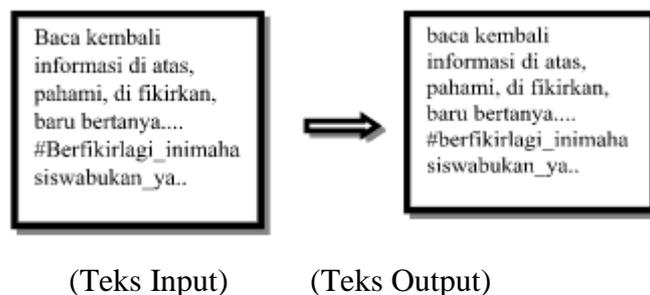
Dalam penelitian ini ada tiga kategori opini yang akan digunakan, yaitu : opini positif, negatif dan opini netral yang diambil dari akun *Twitter* @sandiuono. Dokumen opini yang digunakan ada 450 dokumen dan dibagi menjadi 150 dokumen untuk opini positif, 150 dokumen negatif dan 150 dokumen netral. Dari 450 dokumen tersebut peneliti membagi skenario untuk pengujian antara data *training* dan data *testing*, yaitu 50:50, 60:40, 70:30, 80:20 dan 90:10.

2. *Preprocessing*

Tahap *preprocessing* bertujuan untuk mempersiapkan data agar formatnya sesuai dengan kebutuhan sistem. Proses-proses yang berjalan dalam tahap ini adalah *case folding*, *cleansing*, *tokenizing*, *stopword removal* dan *stemming*.

a. *Case Folding*

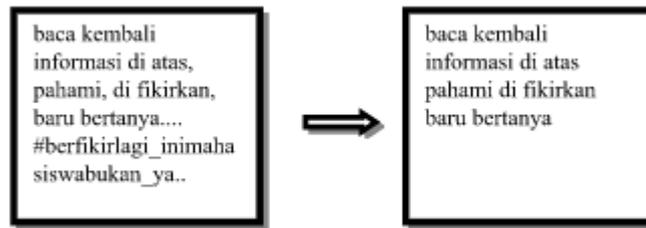
Proses *case folding* yaitu menyeragamkan bentuk huruf menjadi sama bentuknya, baik menjadi *lower case* ataupun menjadi *upper case*. Berikut ini adalah proses dari *case folding*, dapat dilihat pada gambar 2. :



Gambar 2. Proses *Case Folding*

b. *Cleansing*

Proses *cleansing* adalah proses untuk membersihkan tanda baca dan kata-kata yang tidak diperlukan untuk mengurangi noise (data yang tidak konsisten). *Cleansing* dalam penelitian ini menghapus hastag (#), URL dan tanda baca (!@#\$\$%^&*()_+=.,<>?/|{}[]1234567890). Berikut ini adalah proses dari *cleansing*, dapat dilihat pada gambar 3. :

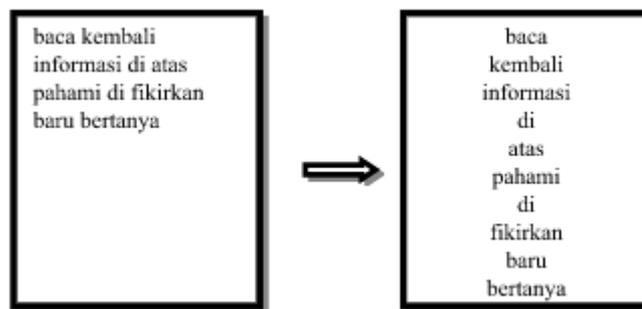


(Teks Input) (Teks Output)

Gambar 3. Proses *Cleansing*

c. *Tokenizing*

Proses *tokenizing* untuk memisahkan deretan kata didalam kalimat menjadi potongan kata tunggal. Kata dapat dipisahkan ketika ada “spasi”, tanda titik “.”, koma “,”, titik dua “:”, titik koma “;”, dan lain sebagainya. Berikut ini adalah proses dari *tokenizing*, dapat dilihat pada gambar 4. :



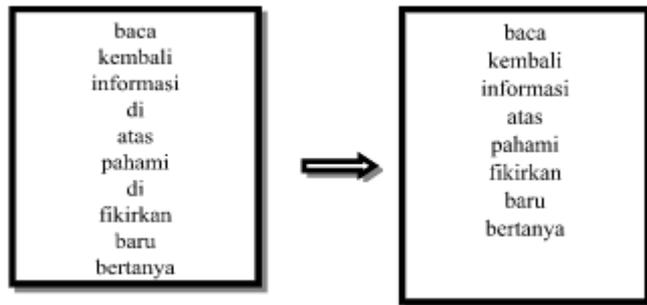
(Teks Input) (Teks Output)

Gambar 4. Proses *Tokenizing*

d. *Stopword Removal*

Proses *stopword removal* yaitu proses untuk membuang kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansinya dengan teks. Daftar *stopword* pada penelitian ini menggunakan *stopword* pada penelitian Talla (2003) dengan memilah kata tertentu. Contoh beberapa kata yang sering masuk ke dalam

stopword list adalah “sebuah”, “yang”, “itu”, “di” dan lain-lain. Berikut ini adalah proses dari *stopword removal*, dapat dilihat pada gambar 5. :

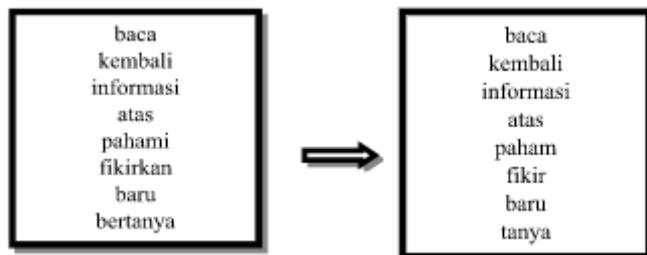


(Teks Input) (Teks Output)

Gambar 5. Proses *Stopword Removal*

e. *Stemming*

Proses *stemming* yaitu proses untuk mencari akar kata dari sebuah kata, kata yang dipecah yaitu kata yang berimbuhan seperti awalan (prefiks), sisipan (infiks), akhiran (sufiks), awalan akhiran (konfiks). Hasil dari proses *stemming* ini akan menghasilkan akar kata yaitu bagian dari kata yang tersisa setelah dihilangkan imbuhan. Pada penelitian ini proses *stemming* yang digunakan adalah dari penelitian (Febriyani, 2015). Berikut ini adalah proses dari *stemming*, dapat dilihat pada gambar 6. :



(Teks Input) (Teks Output)

Gambar 6. Proses *Stemming*

3. Menghitung Probabilitas Kata

Menghitung probabilitas kata dilakukan untuk mendapatkan *term* dengan nilai yang lebih penting dan dianggap relevan untuk dijadikan kata kunci. Proses pembobotan menggunakan algoritma *Naive Bayes Classifier*, dalam proses perhitungan menggunakan persamaan 1:

$$P(V_j) = \frac{n_k + 1}{n + |Kosakata|} \quad (1)$$

Keterangan :

X_i = Kata

V_j = Kategori Opini

Jadi ($P(W_k|V_j)$) sama dengan n_k (frekuensi munculnya W_k) ditambah 1 kemudian dibagi dengan n (banyaknya seluruh kata dalam kategori) dan ditambah dengan banyaknya kosakata $|kosakata|$.

4. Hasil Hitung

Hasil hitung persamaan nantinya akan digunakan untuk proses *testing*. Setelah mendapatkan hasil dari probabilitas kata, kemudian akan menghitung probabilitas dari dokumen sampel atau contoh, diasumsikan $P(V_j)$ (probabilitas kategori dokumen) sama dengan $docs_j$ (jumlah dokumen setiap kategori) dibagi dengan $|contoh|$ (jumlah dokumen yang digunakan sebagai data *training* dari seluruh kategori). Diperoleh persamaan 2 berikut :

$$P(V_j) = \frac{docs_j}{|contoh|} \quad (2)$$

HASIL DAN PEMBAHASAN

Implementasi Sistem

Berdasarkan dari hasil analisis yang telah dilakukan pada bab sebelumnya, maka tahap selanjutnya yang dilakukan adalah melakukan klasifikasi opini dengan metode *text mining* yang menggunakan algoritma *Naive Bayes Classifier* ke dalam bentuk program dengan menggunakan bahasa pemrograman JAVA dan menggunakan bantuan *tool* WEKA.

Skenario Pengujian

Dataset yang digunakan berjumlah 450 data opini berbahasa Indonesia yang dibagi kedalam kategorinya masing-masing yakni 150 untuk data positif, 150 untuk data negatif dan 150 untuk data netral. Data didapat dari akun *twitter* @sandiuono. Setiap dokumen yang diujikan hanya memiliki satu kategori dan *dataset* dibagi menjadi data *training* dan *testing*. Data *training* adalah data yang digunakan sebagai acuan untuk membangun model

klasifikasi, sedangkan data *testing* adalah data yang digunakan untuk menguji performa dari model klasifikasi tersebut. Skenario pengujian data dapat dilihat pada tabel 4.1.

Tabel 4. Skenario Pengujian Data *Testing* dan *Training*

No	Rasio	Training			Testing			Jumlah
		Pos	Neg	Net	Pos	Neg	Net	
1	50 : 50	75	75	75	75	75	75	450
2	60 : 40	90	90	90	60	60	60	450
3	70 : 30	105	105	105	45	45	45	450
4	80 : 20	120	120	120	30	30	30	450
5	90 : 10	135	135	135	15	15	15	450

Pada skenario pengujian, data *training* dan data *testing* sebelumnya sudah dipersiapkan dan disimpan ke dalam *dataset*. Hasil data *training* yang didapat akan digunakan untuk melakukan proses *testing* dengan data baru yang belum diketahui kategorinya. Pada penelitian ini terdapat proses *pre-processing* yaitu *case folding*, *cleansing*, *tokenizing*, *stopword removal* dan *stemming*. Proses ini bertujuan untuk mendapatkan kata kunci dan mengetahui frekuensi kemunculan kata pada dokumen.

Hasil Pengujian

Proses pengujian dilakukan untuk menghasilkan kata kunci yang mewakili setiap kategori dan mengetahui ketepatan akurasi dokumen. Untuk mengukur akurasi ditentukan dengan perbandingan banyaknya data pengujian tepat (benar) dengan banyaknya seluruh pengujian.

Analisis Hasil Pengujian

Berdasarkan hasil pengujian maka dapat dilakukan beberapa analisis dari data *training* dan *testing*. Seperti yang sudah dijelaskan pada skenario pengujian (tabel 4.1) terdapat 450 dokumen, dengan skenario pengujian yaitu :

Tabel 4.7 Rasio Hasil Akurasi

No	Rasio	Hasil Akurasi
----	-------	---------------

1	50 : 50	58,22%
2	60 : 40	58,33%
3	70 : 30	65,18%
4	80 : 20	57,77%
5	90 : 10	55,55%

Berdasarkan hasil pengujian maka dapat dilakukan beberapa analisis. Seperti yang sudah dijelaskan pada skenario pengujian, terdapat 3 kategori opini sebanyak 450 dokumen (150 positif, 150 negatif dan 150 netral). Setelah melalui tahap *preprocessing* dan perhitungan akurasi maka didapatkan akurasi pada tabel 4.7. Didalam tabel 4.7 terdapat akurasi terbesar pada skenario pengujian 70:30 dengan asumsi 315 data *training* dan 135 data *testing* dengan presentasi akurasi 65,18%, dimana persentase akurasi setiap masing-masing dari kategori adalah 77,77% untuk kategori positif, 68,88% untuk kategori negatif dan 48,88% untuk kategori netral.

Berdasarkan hasil akurasi pada tabel 4.7, terdapat beberapa faktor yang mengalami hasil akurasi yang berbeda-beda pada tiap kategori dan hasil akurasi. Untuk hasil akurasi mengalami persentase yang beragam seiring dengan skenario pengujian. Hal ini dikarenakan proses *stemming* mempengaruhi klasifikasi karena proses *stemming* terkadang tidak menemukan kata dasar yang sebenarnya, hal ini terjadi karena *stemming* mengubah kata yang berimbuhan kedalam kata dasar dengan melakukan perubahan aturan dengan pencarian akar kata, sehingga untuk kata yang mengalami proses *stemming* akan mengalami penghapusan imbuhan yang sesuai dengan aturan *stemming* tersebut. Hasil tersebut juga dapat merubah makna dari sentimen awal. Sebagai contoh ketika ada opini seperti ini misalnya “@forum_relawan @sandiuno @detikcom : *Pemikirannya bagus, tapi terlihat tdk mengikuti dinamika pembangunan saat ini, cenderung menafikan*”. Pada proses penentuan kelas opini tersebut masuk dalam kelas negatif, tetapi pada saat proses pembelajaran maka hasil ekstraksi kata yang didapatkan yaitu kata *pikir, bagus, lihat, ikut, bangun, cenderung*. Kata-kata yang dihasilkan tersebut cenderung memiliki arti kata yang bersifat positif seperti kata *bagus* dan *bangun* dan bersifat netral seperti kata *ikut, lihat* dan *cenderung*. Oleh karena itu setiap hasil yang didapatkan tidak sepenuhnya benar semua, tergantung dari ekstraksi fitur kata yang dihasilkan dari proses pembelajaran.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang dilakukan pada bab sebelumnya, maka kesimpulan yang dapat diambil adalah sebagai berikut :

1. Rancangan arsitektur digunakan untuk membangun arsitektur *training* dan arsitektur *testing*. Arsitektur *training* terdapat proses *preprocessing*, pembobotan dan pengelompokan data untuk menghasilkan kata kunci setiap kategori opini kemudian digunakan untuk *testing* terhadap dokumen yang belum diketahui kategorinya.
2. Evaluasi hasil pengujian dengan menggunakan algoritma *Naive Bayes Classifier* menghasilkan akurasi yang beragam dari masing-masing pengujian tergantung dari skenario pengujian telah ditentukan. Perbedaan rasio antara data *training* dan data *testing* didapatkan nilai persentase tertinggi 65,18%, yaitu pada rasio 70:30. Sedangkan untuk akurasi pada tiap kategori adalah 77,77% untuk kategori positif, 68,88% untuk kategori negatif dan 48,88% untuk kategori netral.

REFERENSI

- Abidin, Z., Wijaya, A., & Pasha, D. (2021). Aplikasi Stemming Kata Bahasa Lampung Dialek Api Menggunakan Pendekatan Brute-Force dan Pemograman C. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 1–8.
- Ahmad, I., Borman, R. I., Caksana, G. G., & Fakhrurozi, J. (2021). Penerapan Teknologi Augmented Reality Katalog Perumahan Sebagai Media Pemasaran Pada PT. San Esha Arthamas. *SINTECH (Science and Information Technology) Journal*, 4(1), 53–58.
- Alita, D. (2021). Multiclass SVM Algorithm for Sarcasm Text in Twitter. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(1), 118–128. <https://doi.org/10.35957/jatisi.v8i1.646>
- Alita, D., Sari, I., Isnain, A. R., & Styawati, S. (2021). Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), 17–23.
- Amalia, F. S., Setiawansyah, S., & ... (2021). Analisis Data Penjualan Handphone Dan Elektronik Menggunakan Algoritma Apriori (Studi Kasus: Cv Rey Gasendra). ... *Journal of Telematics and ...*, 2(1), 1–6. <https://ejurnal.teknokrat.ac.id/index.php/telefortech/article/view/1810>

- Assuja, M. A., & Saniati, S. (2016). Analisis Sentimen Tweet Menggunakan Backpropagation Neural Network. *Jurnal Teknoinfo*, 10(2), 48–53.
- Ayu, M., Sari, F. M., & Muhaqiqin, M. (2021). Pelatihan Guru dalam Penggunaan Website Grammar Sebagai Media Pembelajaran selama Pandemi. *Al-Mu'awanah: Jurnal Pengabdian Kepada Masyarakat*, 2(1), 49–55.
- Borman, R. I., Putra, Y. P., Fernando, Y., Kurniawan, D. E., Prasetyawan, P., & Ahmad, I. (2018). Designing an Android-based Space Travel Application Trough Virtual Reality for Teaching Media. *2018 International Conference on Applied Engineering (ICAE)*, 1–5.
- Budiman, A., Sunariyo, S., & Jupriyadi, J. (2021). Budiman, Arief, Sunariyo Sunariyo, and Jupriyadi Jupriyadi. 2021. “Sistem Informasi Monitoring Dan Pemeliharaan Penggunaan SCADA (Supervisory Control and Data Acquisition).” *Jurnal Tekno Kompak* 15(2): 168. Sistem Informasi Monitoring dan Pemeliharaan Pengg. *Jurnal Tekno Kompak*, 15(2), 168. <https://doi.org/10.33365/jtk.v15i2.1159>
- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *Jurnal Riset Komputer*, 8(6), 219–225. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Dewi, P. S. (2021). E-Learning : Penerapan Project Based Learning pada Mata Kuliah Media Pembelajaran. *Prisma*, 10(1), 97. <https://doi.org/10.35194/jp.v10i1.1012>
- Eka Saputri, R. (2018). Pengaruh Kecerdasan Emosional Terhadap Hasil Belajar Siswa. *Syntax Literate : Jurnal Ilmiah Indonesia*, 3(4), 93–102.
- Fakhrurozi, J., Adrian, Q. J., Mulyanto, A., Informasi, S. S., Teknokrat, U., & Online, M. (2022). Pelatihan Penulisan Jurnalistik dan Naskah Video Bagi Siswa SMK Widya Yahya Gading Rejo. 2(5), 503–509.
- Febria Lina, L., & Setiyanto, A. (2021). Privacy Concerns in Personalized Advertising Effectiveness on Social Media. *SIJDEB*, 5(2), 147–156. <https://doi.org/10.29259/sijdeb.v5i2.147-156>

- Fernando, Y., Ahmad, I., Azmi, A., & Borman, R. I. (2021). Penerapan Teknologi Augmented Reality Katalog Perumahan Sebagai Media Pemasaran Pada PT. San Esha Arthamas. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 5(1), 62–71.
- Gandhi, B. S., Megawaty, D. A., & Alita, D. (2021). Aplikasi Monitoring Dan Penentuan Peringkat Kelas Menggunakan Naïve Bayes Classifier. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(1), 54–63.
- Hana, P., Rusliyawati, R., & Damayanti, D. (2019). Pengaruh Media Richness Dan Frequently Update Terhadap Loyali Tas Civitas Akademika Perguruan Tinggi. *Jurnal Tekno Kompak*, 13(2), 7. <https://doi.org/10.33365/jtk.v13i2.328>
- Hasri, C. F., & Alita, D. (2022). Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, 3(2), 145–160. <http://jim.teknokrat.ac.id/index.php/informatika>
- Hendrastuty, N., Rahman Isnain, A., Yanti Rahmadhani, A., Styawati, S., Hendrastuty, N., Isnain, A. R., Rahman Isnain, A., Yanti Rahmadhani, A., Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT*, 6(3), 150–155. <http://situs.com>
- Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis, (2021).
- Isnain, A. R., Marga, N. S., & Alita, D. (n.d.). Sentiment Analysis Of Government Policy On Corona Case Using Naive Bayes Algorithm. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(1), 55–64.
- Permata, P., & Rahmawati, W. D. (2018). Pengembangan Media Pembelajaran Berbasis Multimedia pada Materi Kalkulus. *UNION: Jurnal Ilmiah Pendidikan Matematika*, 6(3), 277–286.
- Puspaningrum, A. S., Suaidah, S., & Laudhana, A. C. (2020). MEDIA PEMBELAJARAN TENSES UNTUK ANAK SEKOLAH MENENGAH PERTAMA BERBASIS

ANDROID MENGGUNAKAN CONSTRUCT 2. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 1(1), 25–35. <https://doi.org/10.33365/jatika.v1i1.150>

Putra, A. D., Suryono, R. R., & Darmini, D. (2009). Rancang bangun media pembelajaran TOEFL berbasis web. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*.

Putri, A. D. (2021). Maksimalisasi Media Sosial untuk Meningkatkan Pendapatan dan Pengembangan Diri Generasi Z di MAN 1 Pesawaran. *Journal of Social Sciences and Technology for Community Service (JSSTCS)*, 2(2), 37. <https://doi.org/10.33365/jsstcs.v2i2.1180>

Putri, S. eka Y., & Surahman, A. (2019). Penerapan Model Naive Bayes Untuk Memprediksi Potensi Pendaftaran Siswa Di Smk Taman Siswa Teluk Betung Berbasis Web. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 1(1), 93–99. <https://doi.org/10.33365/jatika.v1i1.228>

Rahman Isnain, A., Indra Sakti, A., Alita, D., Satya Marga, N., Isnain, A. R., Sakti, A. I., Alita, D., Marga, N. S., Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jdmsi*, 2(1), 31–37. <https://t.co/NfhnmJtXw>

Samsugi, S., Ardiansyah, A., & Kastutara, D. (2018). Arduino dan Modul Wifi ESP8266 sebagai Media Kendali Jarak Jauh dengan antarmuka Berbasis Android. *Jurnal Teknoinfo*, 12(1), 23–27.

Saputra, V. H., & Febriyanto, E. (2019). Media Pembelajaran Berbasis Multimedia Untuk Anak Tuna Grahita. *Mathema: Jurnal Pendidikan Matematika*, 1(1), 15–23.

Saputra, V. H., & Permata, P. (2018). Media Pembelajaran Interaktif Menggunakan Macromedia Flash Pada Materi Bangun Ruang. *WACANA AKADEMIKA: Majalah Ilmiah Kependidikan*, 2(2), 116–125.

Setiawansyah, S., Adrian, Q. J., & Devija, R. N. (2021). SISTEM PAKAR DIAGNOSA PENYAKIT SELAMA KEHAMILAN MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB. *Jurnal Manajemen Informatika (JAMIKA)*, 11(1), 24–36.

Sidiq, M., & Manaf, N. A. (2020). Karakteristik Tindak Tutur Direktif Tokoh Protagonis

Dalam Novel Cantik Itu Luka Karya Eka Kurniawan. *Lingua Franca: Jurnal Bahasa, Sastra, Dan Pengajarannya*, 4(1), 13–21.

Siwi, F., & Puspaningtyas, N. D. (2020). PENERAPAN MEDIA PEMBELAJARAN KOGNITIF DALAM MATERI PERSAMAAN GARIS LURUS MENGGUNAKAN VIDEO DI ERA 4.0. *Jurnal Ilmiah Matematika Realistik*, 1(1), 7–10.

Styawati, Andi Nurkholis, Zaenal Abidin, & Heni Sulistiani. (2021). Optimasi Parameter Support Vector Machine Berbasis Algoritma Firefly Pada Data Opini Film. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 904–910.
<https://doi.org/10.29207/resti.v5i5.3380>

Sulistiani, H., Putra, A. D., Rahmanto, Y., & ... (2021). Pendampingan Dan Pelatihan Pengembangan Media Pembelajaran Interaktif Dan Video Editing Di Smkn 7 Bandar Lampung. *Journal of Social ...*, 2(2), 160–166.
<https://ejournal.teknokrat.ac.id/index.php/JSSTCS/article/view/1375>

Syah, H., & Witanti, A. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm). *Jurnal Sistem Informasi Dan Informatika (Simika)*, 5(1), 59–67.
<https://doi.org/10.47080/simika.v5i1.1411>

Very, V. H. S., Pasha, D., Hendra Saputra, V., & Pasha, D. (2021). Komik Berbasis Scientific Sebagai Media Pembelajaran di Masa Pandemi Covid-19. *SJME (Supremum Journal of Mathematics Education)*, 5(1), 85–96.
<https://doi.org/10.35706/sjme.v5i1.4514>

Wahyudi, A. D., Surahman, A., & ... (2021). Penerapan Media Promosi Produk E-Marketplace Menggunakan Pendekatan AIDA Model dan 3D Objek. *Jurnal Informatika ...*, 6(1), 35–40.
<http://ejournal.poltektegal.ac.id/index.php/informatika/article/view/2304>

Wibisono, A. D., Rizkiono, S. D., & Wantoro, A. (2020). Filtering Spam Email Menggunakan Metode Naive Bayes. *Telefortech: Journal Of Telematics And Information Technology*, 1(1), 9–17.

Yuliana, Y., Paradise, P., & Kusrini, K. (2021). Sistem Pakar Diagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Classifier Berbasis Web. *CSRID (Computer Science Research and Its Development Journal)*, 10(3), 127. <https://doi.org/10.22303/csrid.10.3.2018.127-138>