

Implementasi Metode Multinomial Naïve Bayes Untuk Sentiment Analysis Terhadap Data Ulasan Produk Colearn Pada Google Play Store

Fadhlan Hadaina^{1*}, Utomo Budiyo²

^{1,2*}Fakultas Teknologi Informasi, Teknik Informatika, Universitas Budi Luhur, Jakarta, Indonesia

Email: ^{1*}1811500824@student.budiluhur.ac.id, ²utomo.budiyo@budiluhur.ac.id
(* : corresponding author)

Abstrak-Startup CoLearn merupakan sebuah perusahaan rintisan yang memiliki tujuan untuk menciptakan sebuah aplikasi bimbel untuk memudahkan siswa belajar di masa pandemi. Dalam pengembangan aplikasi tersebut tak lepas dari ulasan-ulasan pengguna khususnya para siswa sekolah untuk bahan evaluasi pengembangan aplikasi menjadi lebih baik. Terlepas dari pemanfaatan ulasan untuk bahan evaluasi pengembangan aplikasi, dapat dilakukan analisis sentimen terhadap data ulasan dengan menghasilkan 2 kelas sentimen yaitu kelas positif dan negatif. Analisis sentimen dilakukan dengan penerapan metode *multinomial naive bayes*. *Multinomial naive bayes* merupakan salah satu metode *text mining* pada proses pengklasifikasian teks dengan menggunakan nilai probabilitas suatu kelas dalam suatu dokumen. Hasil pengujian metode *Multinomial naive bayes* mampu menghasilkan analisa pengujian akurasi bernilai 88,89% menggunakan 536 dataset dengan rincian 439 data positif dan 97 data negatif dengan dukungan tahapan *preprocessing* diantaranya *case folding*, *cleaning*, *tokenizing*, penggantian *slang word*, penghapusan *stop word* dan *stemming*. Dari hasil pengujian tersebut dapat diambil kesimpulan bahwa metode *Multinomial naive bayes* dapat digunakan untuk melakukan analisis sentimen terhadap data ulasan dengan hasil yang cukup baik. Dengan adanya sistem analisis sentimen terhadap data ulasan ini dapat bermanfaat bagi *startup* CoLearn dalam mengevaluasi ulasan untuk meningkatkan kualitas aplikasi.

Kata Kunci: *sentiment analysis, multinomial naive bayes, ulasan produk, colearn, google play store*

Implementation Of Multinomial Naive Bayes Method For Sentiment Analysis Of Colearn Product Review Data On Google Play Store

Abstract- Startup CoLearn is a start-up company that aims to create a tutoring application to make it easier for students to learn during a pandemic. In developing the application, it cannot be separated from user reviews, especially students, to evaluate application development for the better. Apart from the use of reviews for application development evaluation materials, sentiment analysis can be carried out on data reviews by generating 2 positive and negative class sentiments. Sentiment analysis is done by applying *Multinomial Naive Bayes*. *Multinomial Naive Bayes* is one of the *text mining* methods in the *text classification* process by using the probability value of a class in a document. The results of the *Multinomial Naive Bayes* test were able to produce a feasibility test analysis of 88.89% using 536 data sets with details of 439 positive data and 97 negative data with the support of *pre-processing* stages including *case folding*, *cleaning*, *tokenizing*, *slang words*, *deletion of stop words* and *stemming*. From the test results, it can be concluded that the *Multinomial Naive Bayes* method can be used to conduct *sentiment analysis* on data reviews with fairly good results. The existence of a *sentiment analysis* system for data reviews can be useful for *CoLearn Startup* in reviews to improve the quality of the application.

Keywords: *sentiment analysis, multinomial naive bayes, product reviews, colearn, google play store*

1. PENDAHULUAN

Pandemi Covid-19 memaksa banyak orang untuk beraktivitas di dalam rumah salah satunya proses pembelajaran. Dalam hal proses pembelajaran daring ini tidak lepas kaitannya dengan sebuah aplikasi penyedia jasa edukasi atau biasa disebut *startup* edukasi. Tujuan utama dari aplikasi ini adalah untuk memudahkan proses pembelajaran siswa yang bisa dilakukan dimana saja dengan perangkat apa saja serta waktu yang fleksibel pula. Karena hal ini, banyak *startup* edukasi di Indonesia yang mulai menerapkan proses pembelajaran khususnya bimbel dengan media daring salah satunya *startup* asal Indonesia CoLearn.

CoLearn merupakan *startup* edukasi yang memiliki tujuan untuk membantu siswa menghadapi kesulitan dalam mengerjakan soal-soal dengan banyak fitur yang ditawarkan seperti fitur tanya pengajar dan fitur *live class* dimana siswa dapat bertatap muka langsung dengan pengajar melalui *video conference*. Selain itu aplikasi colearn dapat diakses dimana saja mulai dari *web* hingga *smartphone* yang bisa di unduh gratis melalui *google play store* maupun *app store*.

Pengguna yang telah menggunakan aplikasi CoLearn dapat memberikan *rating* (skor berbintang 1–5) disertai ulasan atau komentar mengenai pengalaman penggunaan melalui fitur *ratings & reviews*. Hingga pertengahan Juni 2022, tercatat aplikasi CoLearn pada *google play store* memiliki nilai *rating* 4.5. Nilai *rating* yang mendominasi adalah *rating* berbintang dengan skor 5, kemudian disusul *rating* berbintang dengan skor 4. Semakin kecil nilai skor *rating* berbintang menunjukkan kinerja aplikasi yang tidak maksimal ketika digunakan pengguna. Referensi

[1] berpendapat ulasan dari 2 pengguna pada *google play store*, terdiri dari ulasan positif dan negatif seperti keluhan, kritik, atau saran.

Beberapa penelitian sebelumnya telah membuat *sentiment analysis* menggunakan beberapa algoritma yang mempunyai fungsi sama sebagai klasifikasi sentimen dan mendapat akurasi sebesar 92,47 % untuk *naive bayes* dan 94,14 % untuk *support vector machine* [2]. Kemudian pada penelitian lain telah dilakukan uji coba model klasifikasi *multinomial naive bayes* menggunakan 978 *data training* dan 245 *data testing*. Hasil terbaik didapatkan terhadap data *testing* pada jumlah fitur 1481 dengan frekuensi kemunculan fitur atau kata minimal 1 kali pada data *training* atau parameter ke-1. Hasil akurasi sebesar 80 % [3]. Kemudian analisis sentimen terhadap ulasan produk *online* berbahasa Indonesia menggunakan metode *naive bayes* dengan hasil nilai akurasi tertinggi yaitu sebesar 77,78% [4]. Dan penelitian terhadap pemindahan Ibu Kota Negara menggunakan metode *bernoulli naive bayes* dan *multinomial naive bayes* dengan hasil sensitivitas metode *bernoulli naive bayes* sebesar 93,45% dan *multinomial naive bayes* sebesar 90,19% [5].

Sehingga Penelitian ini bertujuan mengimplementasi metode *multinomial naive bayes* untuk analisis sentimen pengguna aplikasi CoLearn melalui ulasan pada *google play store*. Hasil analisis tersebut menghasilkan 2 polaritas yaitu ulasan positif dan negatif. Dari penelitian ini dapat dihasilkan model klasifikasi yang terbaik, sehingga diharapkan akurasi yang didapatkan cukup baik.

2. METODE PENELITIAN

2.1 Sentiment Analysis

Sentiment analysis atau analisis sentimen adalah bidang studi yang menganalisis pendapat, penilaian, sikap, dan perasaan terhadap entitas seperti layanan, produk, masalah, atau peristiwa tertentu. Analisis sentimen memiliki aplikasi dalam berbagai bidang bisnis dan sosial, karena opini dapat mempengaruhi perilaku individu dalam proses pengambilan keputusan organisasi [2]. Analisis sentimen juga digunakan untuk menemukan informasi berharga yang dibutuhkan dari data yang tidak terstruktur [6].

Pentingnya analisis sentimen berkaitan dengan pesatnya perkembangan opini di media sosial, seperti ulasan suatu produk atau layanan dan berbagai diskusi di forum blog, *Twitter*, dan jejaring sosial. Dikarenakan data yang diproses pada text mining merupakan sebuah teks yang tidak terstruktur, maka diperlukan pemilihan teks sebelum dilakukan proses selanjutnya, pada tahap ini dikenal dengan prapemrosesan (*preprocessing*).

2.2 Text Preprocessing

Tahapan *preprocessing* atau praproses merupakan bagian yang sangat penting dalam menyiapkan data, hal ini dikarenakan struktur data yang dihasilkan pada tahap pengumpulan tidak beraturan, sehingga menyebabkan proses menjadi tidak berjalan dengan baik.

Merujuk pada Referensi [7] maka pada penelitian ini akan dibahas beberapa tahapan *preprocessing* antara lain, *case folding*, *cleaning*, *tokenizing*, mengganti *slang word*, menghapus *stop word* dan *stemming*.

2.3 Multinomial Naïve Bayes

Multinomial naive bayes merupakan salah satu metode khusus dari *naive bayes* sebagai metode text mining dalam proses klasifikasi teks menggunakan probabilitas kelas dalam dokumen [8]. Proses dimulai dengan memasukkan data latih yang digunakan untuk pembelajaran kemudian dilakukan perhitungan peluang kemunculan suatu kelas pada data latih yang dilakukan menggunakan Persamaan 1 [3].

$$P(c) = \frac{N_c}{N_{doc}} \quad (1)$$

Keterangan:

- c : Kategori atau kelas
- doc : Dokumen
- N_c : Banyaknya kategori c pada dokumen latih
- N_{doc} : Banyaknya keseluruhan dokumen latih yang digunakan

Perhitungan selanjutnya dari probabilitas bahwa kata i termasuk dalam kategori atau kelas tertentu dapat dilakukan dengan menggunakan Persamaan 2 [3].

$$P(w_i, c) = \frac{count(w_i, c) + 1}{\sum_w count(w, c) + |V|} \quad (2)$$

Keterangan:

- w_i : Kata ke- i dalam seluruh dokumen yang berkategori c
- $count(w_i, c)$: Jumlah kata tertentu yang muncul dalam suatu kategori atau kelas
- $\sum_w count(w, c)$: Jumlah seluruh kata pada kelas
- $|V|$: merupakan jumlah seluruh kata unik pada kelas

Penambahan nilai satu berfungsi untuk mencegah hasil probabilitas bernilai 0, manipulasi ini disebut dengan *laplace smoothing*. Setelah tahap pembelajaran selesai dilakukan maka selanjutnya dilakukan tahap klasifikasi data baru berdasarkan hasil pembelajaran. Untuk klasifikasi data baru dilakukan perhitungan menggunakan Persamaan 3 [9].

$$P(c, d) = P(c) \prod_{1 \leq i \leq \text{count}(v, d)} P(w_i, c) \quad (3)$$

Keterangan:

$\text{count}(v, d)$: Jumlah kata unik pada dokumen

2.4 Confusion Matrix

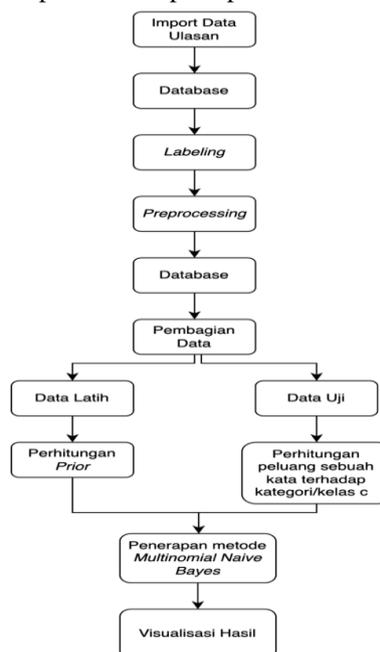
Confusion Matrix adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining [10]. Untuk menghasilkan akurasi dibutuhkan *confusion matrix* dengan data yang dihasilkan dari model klasifikasi berupa data *True Positif* (TP), *True Negatif* (TN), *False Positif* (FP), dan *False Negatif* (FN) dengan membandingkan data aktual dan data hasil klasifikasi. Untuk hasil akurasi dilakukan perhitungan menggunakan Persamaan 4.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (4)$$

3. HASIL DAN PEMBAHASAN

3.1 Penerapan Metode

Pada Gambar 1 dijelaskan alur penerapan metode pada penelitian ini.



Gambar 1. Tahapan Metode

Pada Gambar 1 tahapan pertama merupakan proses *import* data ulasan produk CoLearn di *Google Play Store*, kemudian disimpan ke dalam *database* sebagai *dirty dataset*. Tahapan selanjutnya dilakukan *labeling* terhadap data ulasan yang tersimpan di dalam *database*. Setelah *dataset* dilabelkan, maka dilakukan *preprocessing* terhadap *dataset* dan menghasilkan *clean dataset* yang digunakan untuk proses klasifikasi. Setelah *clean dataset* dihasilkan maka dilakukan pembagian menjadi dua yaitu data latih dan data uji dengan rasio 9:1. Pada masing-masing data dilakukan perhitungan, yaitu perhitungan *prior* terhadap data latih dan perhitungan peluang sebuah kata dari setiap dokumen pada data uji. Setelah itu, dilakukan klasifikasi terhadap dokumen untuk menghasilkan sentimen positif atau negatif terhadap data ulasan yang kemudian divisualisasikan hasilnya.

3.2 Import Data Ulasan

Pada tahap ini data ulasan di-*insert* ke *database* sebagai *dirty dataset* yang digunakan untuk proses pelabelan terhadap data ulasan. Data ulasan bersumber dari ulasan produk CoLearn yang terdapat di *Google Play Store*.

3.3 Labeling

Pada tahap ini setiap dirty dataset ditandai dengan label positif atau negatif untuk kebutuhan data latih dan sebagai acuan perhitungan akurasi terhadap hasil klasifikasi data uji.

3.4 Preprocessing

Pada tahapan *preprocessing* ini, beberapa proses dilakukan untuk mendapatkan kumpulan data yang bersih, sehingga membuat proses pengujian hasil klasifikasi sentimen lebih akurat. Proses *preprocessing* yang digunakan dalam penelitian ini meliputi:

a. *Case Folding*

Case Folding merupakan proses penyetaraan kata yang mengandung huruf besar untuk diubah menjadi huruf kecil, misalnya ‘Metode’ menjadi ‘metode’, ‘Penelitian’ menjadi ‘penelitian’, dan seterusnya.

b. *Cleaning*

Cleaning merupakan proses untuk menghapus karakter selain a sampai z atau karakter selain huruf, karakter tersebut dihilangkan, seperti contoh 8266, / (garis miring), dan seterusnya sehingga hanya menyisakan karakter huruf saja.

c. *Tokenizing*

Setelah melalui tahap *cleaning* maka dilakukan tahap *tokenizing* atau tokenisasi, dimana tahapan ini memenggal kalimat menjadi potongan kata agar memudahkan proses selanjutnya.

d. Mengganti *slang word*

Pada tahapan ini dilakukan perbandingan kata terhadap kata yang terdapat pada *slang word list*. Penulis menggunakan data *json* untuk menyimpan *slang word list* untuk dibandingkan dengan kata yang sudah di tokenisasi.

e. Menghapus *stop word*

Setelah penggantian *slang word* dilakukan penghapusan *stop word*. Prosesnya sama dengan tahapan mengganti *slang word* namun menggunakan data *stop word list* yang diperoleh dari file *.csv*.

f. *Stemming*

Stemming merupakan suatu proses mengubah suatu kata menjadi kata dasar sehingga semua kata imbuhan baik itu awalan, sisipan, akhiran, dan kombinasi awalan dan akhiran akan dihilangkan. Seperti: kata ‘ke atas’ diubah menjadi kata ‘atas’ dan kata ‘melihat’ diubah menjadi kata ‘lihat’. Proses pengubahan tersebut dilakukan dengan melibatkan pustaka JSastrawi.

3.5 Klasifikasi Multinomial Naïve Bayes

Pada tahap ini dilakukan perhitungan menggunakan sampel data yang sudah dibagi menjadi data latih dan data uji yang dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Data Latih Sampel Ulasan Positif Hasil *Preprocessing*

No	Ulasan	Label
1	susah kerjain tugas aplikasi rekomendasi sih cuman dapet jelas biar pas ulang bingung	<i>true</i>
2	video bahas nya mudah paham nyesel download aplikasi keren pokok nya	<i>true</i>
3	bantu ragu	<i>true</i>
4	maaf sebener nya colern bagus kadang nyari nu sesuai pas ketemu bikin paham pas jelasin mohon baik ya kakak terimakasih	<i>false</i>
5	aplikasi nya an cocok	<i>false</i>

Tabel 2. Data Uji Sampel Ulasan Hasil *Preprocessing*

No	Ulasan	Label
1	bagus bantu ajar	<i>true</i>

Setelah data dibagi menjadi data latih dan data uji, kemudian dilakukan perhitungan peluang kelas menggunakan Persamaan 1.

$$P(\text{positif}) = \frac{3}{5}$$

$$P(\text{negatif}) = \frac{2}{5}$$

Kemudian dilanjutkan menghitung peluang terhadap kata-kata yang bersumber dari data uji menggunakan Persamaan 2.

Data ulasan : bagus bantu ajar

$$\begin{aligned}
 \text{Perhitungan kelas positif} : P(\text{bagus, positif}) &= \frac{0 + 1}{26 + 42} = \frac{1}{68} \\
 P(\text{bantu, positif}) &= \frac{1 + 1}{26 + 42} = \frac{2}{68} \\
 P(\text{ajar, positif}) &= \frac{0 + 1}{26 + 42} = \frac{1}{68}
 \end{aligned}$$

$$\begin{aligned}
 \text{Perhitungan kelas negatif} : P(\text{bagus, negatif}) &= \frac{1 + 1}{24 + 42} = \frac{2}{66} \\
 P(\text{bantu, negatif}) &= \frac{0 + 1}{24 + 42} = \frac{1}{66} \\
 P(\text{ajar, negatif}) &= \frac{0 + 1}{24 + 42} = \frac{1}{66}
 \end{aligned}$$

Setelah perhitungan peluang terhadap kata-kata data uji, kemudian dilakukan perhitungan peluang dokumen data uji terhadap kelas sentimen menggunakan Persamaan 3.

$$\begin{aligned}
 P(\text{positif}, d1) &= \frac{3}{5} \times \left(\frac{1}{68}\right)^1 \times \left(\frac{2}{68}\right)^1 \times \left(\frac{1}{68}\right)^1 \\
 &= 0.6 \times 0.014705^1 \times 0.029411^1 \times 0.014705^1 \\
 &= 0.00000381584
 \end{aligned}$$

$$\begin{aligned}
 P(\text{negatif}, d1) &= \frac{2}{5} \times \left(\frac{2}{66}\right)^1 \times \left(\frac{1}{66}\right)^1 \times \left(\frac{1}{66}\right)^1 \\
 &= 0.4 \times 0.030303^1 \times 0.015151^1 \times 0.015151^1 \\
 &= 0.00000278245
 \end{aligned}$$

Dari hasil perhitungan peluang dokumen terhadap kelas diatas, diketahui bahwa peluang dokumen terhadap kelas positif lebih besar dari peluang dokumen terhadap kelas negatif dengan nilai 0.00000381584 untuk peluang positif dan 0.00000278245 untuk peluang negatif. Maka dapat diambil kesimpulan bahwa sampel data uji diklasifikasikan sebagai sentimen positif.

3.6 Pengujian Data Ulasan

Pengujian data ulasan dilakukan dengan 536 *dataset* yang diperoleh dari hasil *crawling google play store review* yang dapat dilihat pada Tabel 3.

Tabel 3. Data Pengujian

No	Ulasan	Label
1	Buat kalian yang sering kesusahan saat kerjain tugas, Aplikasi ini Rekomendasi sih.Karena di sini kita bukan cuman Dapat jawaban Tapi kita dapet Penjelasan sekaligus,biar pas ulangan ga Bingung 😊	<i>true</i>
2	Aplikasi ini sangat membantu saya, kakak kakak nya juga baik, tlaten banget, enak pokoknya disini bisa ikut zoom nya juga	<i>true</i>
3	Aplikasinya sangat bagussss..... Ga ribett ribett lagi tinggal foto dapet deh dan ada pembahasannya juga... Pokoknya bagus banget 😊😁	<i>true</i>
.
536	kok video nya beda banget dengan foto yang dipotong??	<i>false</i>

Dari hasil pengujian menggunakan dengan implementasi metode *multinomial naïve bayes* pada sistem *desktop* yang telah dibangun diperoleh *confusion matrix* yang dapat dilihat pada Tabel 4.

Tabel 4. *Confusion Matrix*

		Predicted Label	
		Positif	Negatif
Actual Label	Positif	45 (true positif)	4 (false positif)
	Negatif	2 (false negatif)	3 (true negatif)

Dari hasil perhitungan *confusion matrix* pada Tabel 4 kemudian dilakukan perhitungan akurasi menggunakan Persamaan 4.

$$\begin{aligned}
 \text{Accuracy} &= \frac{45 + 3}{45 + 3 + 4 + 2} \times 100\% \\
 &= \frac{48}{54} \times 100\% \\
 &= 0.888888889 \times 100\% \\
 &= 88.888888889 \% \approx 88.89 \%
 \end{aligned}$$

3.7 Analisa Hasil Pengujian

Dari hasil pengujian diperoleh hasil akurasi yang tinggi yaitu 88,89%. Dari hasil pengujian tersebut ada beberapa faktor pendukung didapatkannya akurasi yang tinggi diantaranya sebagai berikut:

a. *Labeling*

Tahapan *labeling* dijadikan faktor pendukung karena pada tahapan ini dilakukan seleksi data yang baik agar meningkatkan kualitas kamus data pada kelas tertentu.

b. *Preprocessing*

Tahapan *preprocessing* ini juga menjadi salah satu faktor pendukung karena tahapan ini dengan menghilangkan anomali-anomali pada data sehingga mampu menguatkan sebuah data dengan mengurangi redundansi kata atau bahkan kata yang tidak berguna. Sebagai contoh ditemukan 2 kata yang artinya sama yaitu “jlk ☹” dan “jelek”, pada tahap *preprocessing* kata “jlk ☹” diproses dengan melakukan *cleaning*, penggantian *slang word* dan kata berubah menjadi “jelek” sehingga kata “jelek” memiliki 2 frekuensi. Kemudian contoh lain diterapkan proses *stemming* terhadap kata “bantu” dan “membantu” maka kata “membantu” diubah menjadi “bantu” sehingga kata “bantu” memiliki 2 frekuensi.

c. Variasi kata yang minim

Bila dibandingkan dengan penelitian yang disebutkan pada sub-bab 3.2. Penggunaan data ulasan tidak lebih kompleks dibanding penggunaan data *tweet* dari *twitter* dimana masyarakat dapat beropini dengan ruang lingkup yang cukup luas tentang suatu hal karena tidak hanya berisi sebuah ulasan.

4. KESIMPULAN

Berdasarkan hasil evaluasi dari sistem analisis sentimen ulasan produk CoLearn pada *google play store*, maka dapat disimpulkan bahwa:

- Sistem ini dapat menganalisis sentimen terhadap ulasan produk CoLearn pada *google play store*.

- b. Penggunaan metode *multinomial naïve bayes* untuk analisis sentimen cukup baik dengan menghasilkan persentase akurasi dengan nilai 88,89% menggunakan 536 *dataset* yang terbagi menjadi data latih dan data uji dengan rasio 9:1.
- c. Pemilihan data pada tahapan labeling sangat menentukan kualitas kamus kata terhadap kelas-kelas sentimen.
- d. Penerapan proses *cleaning*, penggantian *slang word* dan *stemming* pada tahapan *preprocessing* sangat berpengaruh dalam meningkatkan frekuensi kata sehingga perhitungan peluang pun dapat meningkat.

Sebagai pengembangan lebih lanjut dari sistem ini, berikut adalah beberapa saran yang dapat peneliti berikan untuk membuatnya bekerja dengan sempurna dengan fitur-fitur yang lebih baik:

- a. Pengembangan sistem dapat dikombinasikan dengan metode lain untuk hasil yang lebih baik.
- b. Penambahan fitur *scraping* data ulasan pada sistem.
- c. Dapat melakukan pendeteksian *slang word* yang lebih baik.
- d. Penambahan *bot spam checker* pada tahap *preprocessing*.
- e. Penambahan kelas sentimen netral.

DAFTAR PUSTAKA

- [1] R. Watrianthos, S. Suryadi, D. Irmayani, M. Nasution, and E. F. S. Simanjanrang, "Sentiment Analysis Of Traveloka App Using Naïve Bayes Classifier Method," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 8, no. 7, pp. 786–788, 2019, [Online]. Available: www.ijstr.org
- [2] R. I. Alhaqq, I. Made, K. Putra, and Y. Ruldeviyani, "Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 11, no. 2, pp. 105–113, 2022.
- [3] L. Lesmana, Mukrodin, and F. Nabyla, "Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Sistem Zonasi PPDB Menggunakan Algoritma Multinomial Naive Bayes," *Jurnal Sistem Informasi dan Teknologi Peradaban (JSITP)*, vol. 1, no. 1, pp. 24–28, 2020, [Online]. Available: <http://journal.peradaban.ac.id/index.php/jsitp>
- [4] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *Jurnal Edukasi dan Penelitian Informatika*, vol. 4, no. 2, pp. 17–29, 2018, [Online]. Available: www.femaledaily.com
- [5] N. S. Wardani, A. Prahutama, and P. Kartikasari, "Analisis Sentimen Pemindahan Ibu Kota Negara Dengan Klasifikasi Naive Bayes Untuk Model Bernoulli Dan Multinomial," *Jurnal Gaussian*, vol. 9, no. 3, pp. 237–246, 2020, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [6] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentiment Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [7] A. Filcha and M. Hayaty, "Implementasi Algoritma Rabin-Karp untuk Pendeteksi Plagiarisme pada Dokumen Tugas Mahasiswa," *JUITA*, vol. 7, no. 1, pp. 25–32, May 2019.
- [8] J. Wisudawati, Adiwijaya, and S. al Faraby, "Klasifikasi Sentimen pada Movie Review dengan Metode Multinomial Naïve Bayes," *e-Proceeding of engineering*, vol. 4, no. 2, pp. 2978–2988, Aug. 2017.
- [9] S. A. Karunia, R. Saptono, and R. Anggrainingsih, "Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection," *Jurnal Ilmiah Teknologi dan Informasi*, vol. 6, no. 1, pp. 10–15, 2017, [Online]. Available: <https://www.researchgate.net/publication/319963217%0AOnline>
- [10] B. P. Pratiwi, A. S. Handayani, and Sarjana, "Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi WSN Menggunakan Confusion Matrix," *Jurnal Informatika UPGRIS*, vol. 6, no. 2, pp. 66–75, 2020.