



## Analisis Sentimen Game Genshin Impact pada Play Store Menggunakan Naïve Bayes Clasifier

**Primandani Arsi**

Prodi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

**Pungkas Subarkah**

Prodi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

**Bagus Adhi Kusuma**

Prodi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

Alamat: Universitas Amikom Purwokerto, Jalan Letjend Pol. Soemarto, Watumas,  
Purwanegara, Purwokerto Utara, Kabupaten Banyumas, Jawa Tengah  
Korespondensi penulis: [ukhti.prima@amikompurwokerto.ac.id](mailto:ukhti.prima@amikompurwokerto.ac.id)

**Abstract.** *Online games are an entertainment medium that cannot be separated from some groups of people, especially during the co-19 pandemic. The existence of a policy of violence reduces the interaction of people in the world, including in Indonesia, resulting in an increase in online activity, especially playing games. Game is an interesting online gaming platform, where players can still interact socially boldly. Genshin Impact is a mobile game that provides other platforms such as PC, PlayStation and Nintendo Switch. Although some people like this game, others are not satisfied with the game play. Sentiment analysis is also very useful when game developers want to know what users think about the game experience. This research aims to build a model of the classification of reviews on the Genshin Impact game available on the Google Play platform, so that the resulting model can provide recommendations for developers for improvement. The results obtained are reviews on the Google Play Store tend to be positive with an accuracy score of 87%, precision of 67%, memory of 98%, and f1 score of 67%. Evaluation is done by comparing the model that has been obtained in this study with the previous model with the same algorithm.*

**Keywords:** *genshin impact, naive bayes, online*

**Abstrak.** *Game Online merupakan salah satu media hiburan yang tidak dapat di pisahkan dari beberapa kalangan masyarakat terutama selama pandemi covid-19. Adanya kebijakan pembatasan sosial mengurangi interaksi masyarakat di dunia, tidak terkecuali di Indonesia berakibat pada peningkatan aktivitas online terutama bermain game. Game merupakan platform permainan online yang menarik, dimana para pemain masih dapat berinteraksi sosial secara daring. Genshin Impact merupakan game mobile yang menyediakan platform lain seperti PC, PlayStation, dan Nintendo Switch. Meskipun sebagian kalangan menyukai game ini, namun sebagian yang lain merasa tidak puas dengan permainan game tersebut. Analisis sentimen juga sangat berguna saat pihak pengembang game ingin mengetahui*

opini pengguna tentang pengalaman bermain *game* tersebut. Penelitian ini bertujuan untuk membangun sebuah model dari klasifikasi ulasan pada *game Genshin Impact* yang tersedia pada platform *Google Play*, sehingga dengan model yang dihasilkan dapat memberikan rekomendasi bagi pengembang untuk perbaikan. Hasil yang didapatkan adalah ulasan pada *Google Play Store* cenderung positif dengan nilai akurasi sebesar 87%, precision sebesar 67%, recall 98%, dan f1-score 67%. Evaluasi dilakukan dengan cara membandingkan model yang telah diperoleh pada penelitian ini dengan model sebelumnya dengan algoritma yang sama.

**Kata kunci:** *genshin impact, naive bayes, online*

## 1. PENDAHULUAN

*Game Online* merupakan salah satu media hiburan yang tidak dapat di pisahkan dari beberapa kalangan masyarakat selama bertahun-tahun, terutama selama pandemi covid-19 beberapa tahun yang lalu. Adanya kebijakan pembatasan sosial mengurangi interaksi masyarakat di dunia, tidak terkecuali di Indonesia. *Game* merupakan platform permainan online yang menarik, dimana para pemain masih dapat berinteraksi sosial secara daring melalui platform *game* itu sendiri. Riset menunjukkan bahwa terjadi peningkatan durasi bermain *game* sejak adanya regulasi pembatasan sosial[1]. Selain menjadi alternatif hiburan daring, *game* juga memiliki sisi negatif diantaranya menyebabkan kecanduan sehingga interaksi sosial didunia nyata menjadi terganggu. Bahkan efek negatif dari *game* dapat menyebabkan penggunanya menjadi antisosial[2].

Beberapa jenis *game online* yang banyak diminati masyarakat Indonesia yang berbasis mobile adalah *Mobile Legend, PUBG Mobile, Free Fire* dan *Genshin Impact*. *Genshin Impact* merupakan *game mobile* yang menyediakan platform lain seperti PC, *PlayStation*, dan *Nintendo Switch*. *Game* ini memadukan elemen *action RPG* dengan dunia terbuka yang indah dan menarik untuk dijelajahi[3]. *Genshin Impact* dirilis oleh miHoYo pada tahun 2020[4]. Sejak perilisannya, *Genshin Impact* telah menjadi salah satu *game mobile* yang sangat populer di seluruh dunia, termasuk di Indonesia [5]. Meskipun sebagian kalangan menyukai *game* ini, namun sebagian yang lain merasa tidak puas dengan permainan *game* tersebut. *Google Play* adalah salah satu platform bagi pengguna untuk dapat memberikan ulasan/komentar pada aplikasi yang telah didownload salah satunya *Genshin Impact*. Bagi pengguna lainnya ulasan pengguna sebelumnya merupakan informasi berharga yang dapat menjadi masukan terkait dengan kualitas suatu *game*. Analisis sentimen dapat digunakan untuk mengetahui pendapat pengguna tentang *game* tersebut tergolong positif, netral, atau

negatif [6]. Analisis sentimen juga sangat berguna saat pihak pengembang *game* ingin mengetahui opini pengguna tentang pengalaman bermain *game* tersebut. Sehingga dapat dimanfaatkan untuk memutuskan arah pengembangan dan peningkatan konten selanjutnya. Teknik analisa sentimen juga dapat mengetahui popularitas sebuah produk, dalam hal ini *game* [7]. Metode analisa sentimen yang umum digunakan adalah *machine learning*, salah satunya menggunakan algoritma Naïve Bayes.

Naive Bayes adalah algoritma yang dapat mengklasifikasikan dengan cara sederhana dan paling umum digunakan. Model klasifikasi Naive Bayes menghitung probabilitas posterior suatu kelas berdasarkan distribusi kata dalam dokumen. Hal itu bergantung pada representasi dokumen yang sangat sederhana sebagai *Bag of words*. Model ini bekerja dengan mengekstraksi fitur *bag of words* yang mengabaikan posisi kata dalam dokumen. Teknik ini menggunakan Teorema Bayes untuk memprediksi probabilitas bahwa set fitur yang diberikan milik label tertentu[8]. Penelitian ini bertujuan untuk membangun sebuah model dari klasifikasi ulasan pada *game Genshin Impact* yang tersedia pada *platform Google Play*, sehingga dengan model yang dihasilkan dapat memberikan rekomendasi bagi pengembang untuk perbaikan. Selain itu sebagai referensi bagi pengguna terkait penilaian performa dari *game Genshin Impact*.

## 2. TINJAUAN PUSTAKA

Naive Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dg menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naive Bayes Classifier ini adalah asumsi yg sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian. Keunggulan metode ini yaitu, hanya membutuhkan jumlah data pelatihan (training data) yg kecil untuk menentukan estimasi parameter yg diperlukan dalam proses pengklasifikasian.

Tahapan proses Naive Bayes

1. Menghitung jumlah kelas / label
2. Menghitung Jumlah Kasus Per Kelas

3. Kalikan Semua Variable Kelas

4. Bandingkan Hasil Per Kelas

Pada Teorema Bayes, bila terdapat dua kejadian yang terpisah

(misalkan X dan H), maka Teorema Bayes dirumuskan sebagai berikut [20]:

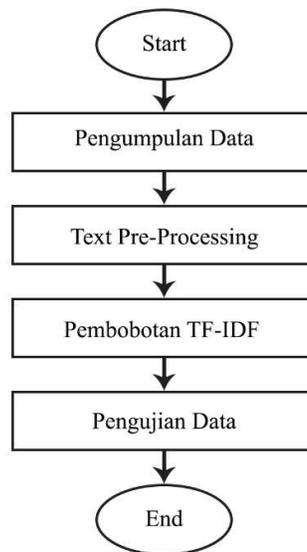
$$P(H|X) = \frac{P(X|H)}{P(X)} \cdot P(H)$$

Keterangan

- a. X : Data dengan class yang belum diketahui
- b. H : Hipotesis data merupakan suatu class spesifik
- c.  $P(H|X)$  : Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
- d.  $P(H)$  : Probabilitas hipotesis H (prior probabilitas)
- e.  $P(X|H)$  : Probabilitas X berdasarkan kondisi pada hipotesis H
- f.  $P(X)$  : Probabilitas X

### 3. METODOLOGI PENELITIAN

Tahapan dalam penelitian digambarkan seperti alur yang tampak pada gambar 1 dibawah ini.



Pengumpulan Data yang dilakukan pada penelitian ini didapatkan dari ulasan pengguna *game Genshin Impact* di *Google Play Store* dengan menggunakan teknik *web scraping*. Teknik ini melibatkan ekstraksi data dari elemen HTML, teks, gambar, tautan, dan informasi lainnya yang ada pada halaman web. Web scraping digunakan untuk

mendapatkan data secara terstruktur dari berbagai sumber *google Play Store* dan kemudian data tersebut dapat diolah atau dianalisis lebih lanjut [9]. *Text Pre-Processing* dalam penelitian ini meliputi Transform Case (mengubah semua huruf kapital yang ada pada dataset menjadil kecil), *Labeling* (mengkategorikan sentiment berdasarkan nilai rating ke positif dan negatif), *Tokenizing* (melakukan pemecahan sekumpulan kalimat menjadi potongan kata-kata atau yang sering disebut token), *Filter tokens by length* (menyortir data dengan batas minimal token tertentu), *Remove Punctuation* (melakukan penghapusan terhadap tanda baca ), Pembobotan *TF-IDF*.

Setelah dataset melalui proses *text pre-processing*, selanjutnya adalah melakukan pembobotan pada teks tersebut untuk mendapatkan nilai agar bisa diklasifikasikan [10]. Pembobotan kata dilakukan di dalam sebuah dokumen menggunakan salah satu metode pembobotan kata yang populer, yaitu metode *TF-IDF (Term Frequency-Inversed Document Frequency)*. *Term-Frequency* memiliki fokus pada seberapa seringnya suatu istilah yang muncul dalam suatu dokumen. Sementara itu, *Inversed Document Frequency* memiliki fokus pada memberikan bobot terendah terhadap istilah yang muncul dalam banyak dokumen[11].

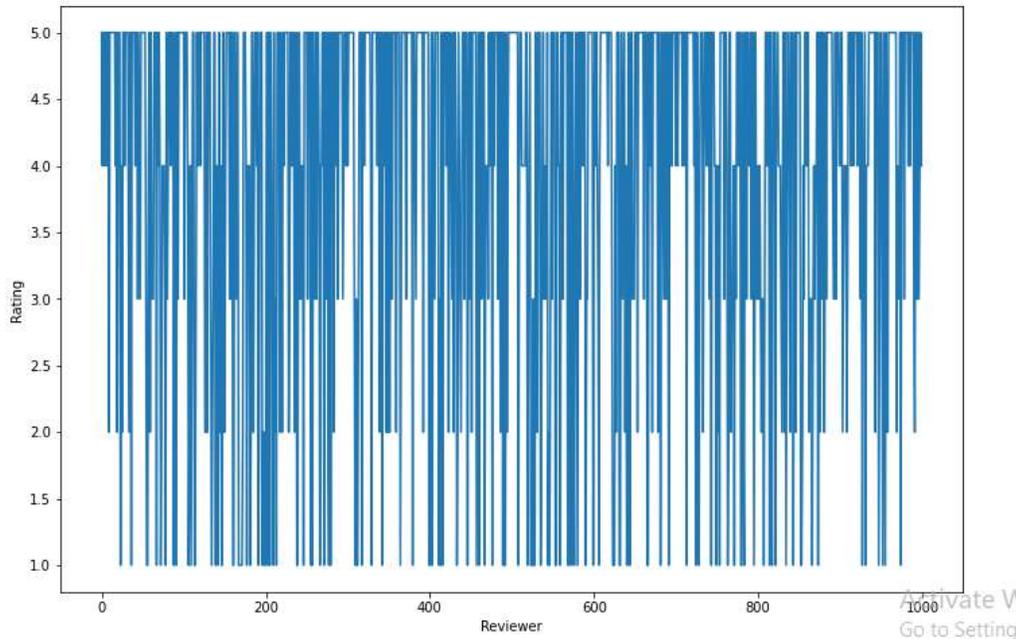
Pengujian atau evaluasi menggunakan algoritma Naïve Bayes yang memakai statistika sederhana dengan dasarnya yaitu teorema bayes dengan memperkirakan ada atau tidaknya suatu class [11]. Multinomial Naïve Bayes merupakan metode klasifikasi yang difokuskan untuk melakukan klasifikasi pada data tekstual. Multinomial Naïve Bayes memiliki fitur dimana hasil yang didapat dari masing-masing kelas bersifat independen karena dokumen yang satu dengan dokumen yang berikutnya tidak memiliki keterkaitan, sehingga hasil yang didapatkan adalah murni hanya diolah dari dokumen itu sendiri[12][13].

#### 4. HASIL DAN PEMBAHASAN

##### Data Scraping

Data Scraping dilakukan dengan menggunakan *tools Google Colaboratory* dengan bahasa pemrograman Python. Scraping data dilakukan pada *game Genshin Impact* yang ada di *play store* pada rentang waktu 29 Desember 2022 hingga 3 Januari 2022. Adapun ulasan yang diambil adalah pada ulasan yang dianggap paling relevan sejumlah 1000

ulasan. Berdasarkan analisa awal terhadap visualisasi ulasan data yang telah diperoleh ditunjukkan pada gambar 2 berikut ini.



Gambar 1. Grafik Rating

Visualisasi data ulasan diatas menunjukkan range rating 1 sampai dengan 5. Adapun jumlah ulasan terbanyak dapat didefinisikan dari gambar tersebut adalah pada rating 5. Selanjutnya pengelompokan *rating* ulasan ini akan menjadi masukan pada tahap *labeling*.

#### *Text-Preprocessing*

Tahapan selanjutnya dari penelitian ini yakni *Text-Preprocessing*, tahap ini penting dilakukan untuk mengubah data tidak terstruktur menjadi terstruktur agar dapat diimplementasikan pada algoritma. Adapun *text-preprocessing* yang dilakukan adalah:

1. *Transform Case*, atau biasa disebut juga *lowering text* adalah mengubah semua huruf kapital yang ada pada dataset menjadil huruf kecil. Gambar 4 berikut merupakan output tahap *transform case*.

	score	content
970	4	bermain sangat menyenangkan, dengan kualitas g...
137	5	enak, gamenya tidak kaku, game yang sangat fle...
82	5	game nya udah bagus tapi pas saya baru downloa...
188	5	i like this map nya luas bisa dijelajahin semu...
358	5	game ini sangat keren dan saya menyukai game b...

Gambar 2. Hasil tahap *Transform Case*

2. *Labeling* adalah proses mengkategorikan sentimen berdasarkan nilai *rating* ke positif dan negatif, jika rating 1 sampai 3 akan memberikan nilai sentiment 0 yang artinya negatif, sedangkan 4 dan 5 memberikan nilai sentiment 1 untuk positif.

	score	content	sentiment
970	4	bermain sangat menyenangkan, dengan kualitas g...	1.0
137	5	enak, gamenya tidak kaku, game yang sangat fle...	1.0
82	5	game nya udah bagus tapi pas saya baru downloa...	1.0
188	5	i like this map nya luas bisa dijelajahin semu...	1.0
358	5	game ini sangat keren dan saya menyukai game b...	1.0

Gambar 3. Proses *labeling*

3. *Tokenizing* yaitu melakukan pemecahan sekumpulan kalimat menjadi potongan kata-kata atau yang sering disebut token.

```
204  kalo kalian maen game keren mirip game playsta...
71   agak kecewa sama quest archon terbaru tapi tet...
594  gamenya bagus dalam segi grafis character styl...
672  game paling nyaman kalo buat dimainin waktu se...
14   untuk meningkatkan grafik tolong sediakan smaa
```

4. *Filter tokens by length* melakukan sortir data dengan batas minimal token tertentu, semisal jika kurang dari 3 token data tersebut akan dihilangkan.

```
204  kalo kalian maen game keren mirip game playsta...
71   agak kecewa sama quest archon terbaru tapi tet...
594  gamenya bagus dalam segi grafis character styl...
672  game paling nyaman kalo buat dimainin waktu se...
14   untuk meningkatkan grafik tolong sediakan smaa
...
647  hanya menyarankan untuk secepatnya bagi yang m...
797  game sangat seru bisa explore dapat character ...
605  malas leveling kenapa harus level tinggi untuk...
611  semua sudh bingung ngapain lagi overall gameny...
988  sangat bagus apalagi untuk update terbaru alan...
Name: content_clean, Length: 200, dtype: object
```

Gambar 5. Hasil *Filtering*

5. *Remove Punctuation* melakukan penghapusan terhadap tanda baca pada dataset seperti titik, koma, dan sebagainya.

	score	content	sentiment	content_clean
204	5	kalo kalian mau maen game keren mirip game pla...	1.0	kalo kalian maen game keren mirip game playsta...
71	4	agak kecewa sama quest archon terbaru 😞 tapi t...	1.0	agak kecewa sama quest archon terbaru tapi tet...
594	5	gamenya bagus dalam segi grafis, character sty...	1.0	gamenya bagus dalam segi grafis character styl...
672	5	game paling nyaman kalo buat dimainin pas wakt...	1.0	game paling nyaman kalo buat dimainin waktu se...
14	5	untuk meningkatkan grafik, tolong sediakan ant...	1.0	untuk meningkatkan grafik tolong sediakan sma...
64	5	gamenya bagus, tapi mengapa ya saat wish selal...	1.0	gamenya bagus tapi mengapa saat wish selalu da...
340	5	tolong yang ngasih review jelek mikirin lagi, ...	1.0	tolong yang ngasih review jelek mikirin lagi m...
135	2	ada masalah memo yang agak aneh padahal masih ...	0.0	masalah memo yang agak aneh padahal masih tapi...
350	5	game ini keren bangeett, bebas mau ngapain aja...	1.0	game keren bangeett bebas ngapain nggak ditunt...
976	5	overall gamenya bagus, dari story gamenya, mek...	1.0	overall gamenya bagus dari story gamenya mekan...

Tahapan selanjutnya adalah tahapan pembobotan yang dilakukan dengan algoritma TF-IDF. Berikut hasil pembobotan pada *term*. Output tf-idf ini yang nantinya digunakan pada inputan *naive bayes*.

```
(0, 2951) 0.21893356672879696
(0, 2949) 0.21893356672879696
(0, 2381) 0.28438813234074944
(0, 2196) 0.2678981874245861
(0, 1923) 0.2397084403966413
(0, 1577) 0.21151869336869653
(0, 1565) 0.18166874814478923
(0, 1345) 0.15110925075836718
(0, 1211) 0.1440798704100185
(0, 1210) 0.4104990145405799
(0, 904) 0.26654416362977235
(0, 507) 0.4464369909609559
(0, 477) 0.18332894634075175
(0, 429) 0.1300285921293466
(0, 242) 0.1196052465502595
(0, 226) 0.12031875973289921
(0, 200) 0.15546288012761547
(0, 128) 0.17425387478468882
```

### Pengujian Data

Pengujian data dilakukan dengan melakukan *splitting data* dengan rasio 80:20 untuk *training* dan *testing*. Implementasi menggunakan Algoritma Naive Bayes dengan pembobotan TF-IDF. Hasil yang didapatkan adalah ulasan pada Google Play Store cenderung positif dengan nilai akurasi sebesar 87%, precision sebesar 67%, recall 98%, dan f1-score 67%. Evaluasi dilakukan dengan cara membandingkan model yang telah diperoleh pada penelitian ini dengan model sebelumnya dengan algoritma yang sama. Penelitian terdahulu oleh George KL, 2022 menghasilkan akurasi (83.3%), precision (65%), recall (63%)[14]. Perbandingan ketiga pada penelitian kedua oleh Ryo K, 2021 menghasilkan nilai presisi tertinggi yaitu sebanyak 86%, 78% recall dan 82% di f1-score[15]. Sehingga jika dibandingkan dengan beberapa penelitian tersebut, dapat diambil

pernyataan jika menggunakan Multinomial Naïve Bayes dan TF-IDF menghasilkan akurasi yang baik yakni 87%.

## 5. KESIMPULAN DAN SARAN

Berdasarkan eksperimen yang dilakukan terhadap data ulasan *Game Genshim Impact* dan evaluasi yang dilakukan dengan cara membandingkan model dengan model yang telah ada sebelumnya dapat diambil kesimpulan bahwa algoritma Naive Bayes akan mengalami performa yang baik jika dikombinasikan dengan algoritma TF-IDF sebagai pembobotan kata. Namun demikian dalam penelitian ini belum dapat mengklasifikasikan jenis kalimat yang mengandung simbol emoji. Saran untuk perbaikan pada penelitian selanjutnya adalah dengan mempertimbangkan simbol emoji pada saat *labeling*.

## DAFTAR PUSTAKA

- [1] D. L. King, P. H. Delfabbro, J. Billieux, and M. N. Potenza, "Problematic online gaming and the COVID-19 pandemic," *J. Behav. Addict.*, vol. 9, no. 2, pp. 184–186, 2020.
- [2] R. J. Santi, D. Setiawan, and I. A. Pratiwi, "Perubahan Tingkah Laku Anak Sekolah Dasar Akibat Game Online," *J. Penelit. dan Pengemb. Pendidik.*, vol. 5, no. 3, p. 385, 2021.
- [3] A. Zahraputeri and L. Kusdiby, "Analisis Persepsi Pemain Terhadap Game Cross-Platform : Studi Kasus Permainan Genshin Impact," *Pros. 12th Ind. Res. Work. Natl. Semin.*, pp. 1273–1278, 2021.
- [4] C. Angelia, F. A. M. Hutabarat, N. Nugroho, Arwin, and Ivone, "Perilaku Konsumtif Gamers Genshin Impact terhadap Pembelian Gacha," *J. Bus. Econ. Res.*, vol. 2, no. 3, pp. 61–65, 2021.
- [5] R. I. Haikal, D. P. Agustino, and I. M. P. P. Wijaya, "Evaluasi User Experience pada Game Genshin Impact menggunakan Metode Cognitive Walkthrough dan Persona," *J. Sist. dan Inform.*, vol. 16, no. 1, pp. 17–25, 2021.
- [6] Y. Wang, K. T. Kim, B. J. Lee, and H. Y. Youn, "Word clustering based on POS feature for efficient twitter sentiment analysis," *Human-centric Comput. Inf. Sci.*, vol. 8, no. 1, 2018.
- [7] E. Indrayuni, "Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vectore Machine Berbasis Particle Swarm Optimization," *J. Evolusi*, vol. 4, 2016.
- [8] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *Jdmsi*, vol. 2, no. 2, pp. 1–9, 2021.

- [9] R. M. Amir Latif, M. Talha Abdullah, S. U. Aslam Shah, M. Farhan, F. Ijaz, and A. Karim, "Data scraping from google play store and visualization of its content for analytics," *2019 2nd Int. Conf. Comput. Math. Eng. Technol. iCoMET 2019*, pp. 1–8, 2019.
- [10] J. Homepage, N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "MALCOM: Indonesian Journal of Machine Learning and Computer Science The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Goo," vol. 2, no. April, pp. 47–54, 2022.
- [11] M. T. Anjasmosos, I. Istiadi, and F. Marisa, "Analisis Sentimen Aplikasi Go-Jek Menggunakan Metode SVM Dan NBC (Studi Kasus: Komentar Pada Play Store)," *Conf. Innov. Appl. Sci. Technol. (CIASTECH 2020)*, no. Ciastech, pp. 489–498, 2020.
- [12] D. P. Santoso and W. Wibowo, "Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Metode Naïve Bayes Classifier pada Situs Google Play Store," *J. Sains dan Seni ITS*, vol. 11, no. 2, 2022.
- [13] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier," *J. Minfo Polgan*, vol. 11, no. 2, pp. 30–35, 2022.
- [14] G. K. Locarso, "Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan NBC," *J. Tek. Inform. Kaputama*, vol. 6, no. 2, pp. 353–361, 2022.
- [15] R. Kusnadi, Y. Yusuf, A. Andriantony, R. Ardian Yaputra, and M. Caintan, "Analisis Sentimen Terhadap Game Genshin Impact Menggunakan Bert," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 6, no. 2, pp. 122–129, 2021.