

Analisis dan Visualisasi Berbasis Web Sentimen Pengguna Jenius Menggunakan Naïve Bayes Classifier

Jimmy Pratama Setiadi¹, Sugiyamta²

Sistem Informasi, Universitas STIKUBANK Semarang, Jl. Tri Lomba Juang, Mugassari, Semarang Utara, 50241

e-mail: ¹jimmypratama.std@gmail.com, ²sugiyamtagik@edu.unisbank.ac.id

Submitted Date: January 18th, 2024

Reviewed Date: January 22nd, 2024

Revised Date: January 26th, 2024

Accepted Date: January 29th, 2024

Abstract

Along with technology development, many varied digital banking or Mobile Banking applications have emerged. Mobile Banking is an innovative result of technological advances in the banking sector and Jenius is one of them. The high level of competition urges Jenius to maintain its competitive advantage and continue to innovate so that it can continue to survive. To overcome these problems, researchers propose a sentiment analysis process to understand the wants and needs of user feedback. This research uses the Naïve Bayes Classifier algorithm as a sentiment analysis method with visualization results presented via the web. The dataset is obtained by scraping method with three sentiment categories: positive, neutral, and negative. The data processing process uses the Preprocessing method with Naïve Bayes testing performed on three configurations of training data: testing data to determine the highest accuracy. The results of this study show that the Naïve Bayes algorithm gets the highest accuracy rate of 90% with the results of positive sentiment word clouds that describe the ease of use of the application as a reference for Jenius to continue to maintain these advantages, while neutral sentiment word clouds that describe the difficulty of the verification process and negative sentiment word clouds that describe application performance that is not optimal as an improvement material for Jenius to increase satisfaction and increase customer satisfaction.

Keywords: Sentiment analysis; Web; Naïve Bayes Classifier; Preprocessing; Word Cloud

Abstrak

Seiring berkembangnya teknologi, banyak muncul aplikasi perbankan digital atau *Mobile Banking* yang bervariasi. *Mobile Banking* adalah sebuah hasil inovasi kemajuan teknologi dari bidang perbankan dan Jenius merupakan salah satunya. Tingginya tingkat persaingan mendesak Jenius untuk mempertahankan keunggulan kompetitif yang dimiliki dan terus berinovasi sehingga dapat terus bertahan. Untuk mengatasi permasalahan tersebut, peneliti mengusulkan proses analisis sentimen dengan tujuan untuk memahami keinginan dan kebutuhan dari umpan balik pengguna. Penelitian ini menggunakan algoritma *Naïve Bayes Classifier* sebagai metode analisis sentimen dengan hasil visualisasi yang dipresentasikan melalui web. *Dataset* didapatkan dengan metode *Scraping* dengan tiga kategori sentimen yaitu positif, netral, dan negatif. Proses pengolahan data menggunakan metode *Preprocessing* dengan pengujian *Naïve Bayes* yang dilakukan pada tiga konfigurasi *data training: data testing* untuk menentukan akurasi tertinggi. Hasil penelitian ini menunjukkan algoritma *Naïve Bayes* mendapatkan tingkat akurasi tertinggi sebesar 90% dengan hasil *word cloud* sentimen positif yang menggambarkan kemudahan penggunaan aplikasi sebagai acuan Jenius untuk terus mempertahankan keunggulan tersebut, sedangkan *word cloud* sentimen netral yang menggambarkan sulitnya proses verifikasi dan *word cloud* sentimen negatif yang menggambarkan kinerja aplikasi yang tidak optimal sebagai bahan *improvement* bagi Jenius untuk meningkatkan kepuasan dan loyalitas pengguna.

Kata Kunci: Analisis Sentimen; Web; *Naïve Bayes Classifier*; *Preprocessing*; *Word Cloud*

1 Pendahuluan

Kemajuan teknologi telah memberikan dorongan yang signifikan terhadap penggunaan layanan perbankan melalui perangkat seluler, yang sering dikenal sebagai *mobile banking*. *Mobile banking* merupakan layanan yang memberikan kenyamanan dan aksesibilitas tinggi untuk mendapatkan informasi terbaru serta melakukan transaksi keuangan secara instan (Maulana et al., 2018).

Jenius merupakan sebuah inovasi Fintech atau Financial Technology dari Bank Tabungan Pensiun Nasional (BTPN), berupa aplikasi perbankan berplatform digital yang diciptakan pada tahun 2016 dan telah memiliki 4,6 juta pengguna per Maret 2023. Selain transaksi perbankan dasar seperti transfer dan pembayaran tagihan, Jenius juga menawarkan produk-produk inovatif seperti kartu debit fisik dan virtual yang dapat dikontrol melalui aplikasi, tabungan berjangka dengan bunga kompetitif, pembayaran dengan QRIS, top up e-Wallet dan berbagai promo discount menarik bagi pemilik m-card atau kartu debit jenius yang membuatnya menonjol dalam pasar *mobile banking*.

Ketersediaan aplikasi *mobile banking* yang bervariasi menyebabkan persaingan semakin ketat. Tingginya tingkat persaingan mendesak agar Jenius untuk dapat mempertahankan keunggulan kompetitif yang dimiliki serta melakukan inovasi agar dapat terus bertahan dan mempertahankan loyalitas pengguna. Pengukuran sentimen pengguna merupakan solusi untuk dapat memahami keinginan dari umpan balik pengguna.

Analisis sentimen mampu menyelesaikan masalah ini dengan menganalisis emosi dan konteks dari umpan balik yang diberikan (Fiarni et al., 2016). Analisis sentimen dapat mengkonversi data yang awalnya tidak terstruktur menjadi data yang terstruktur (Natasuwarna, 2020). Analisis sentimen merupakan usaha untuk mengevaluasi pendapat atau opini seseorang terhadap suatu objek atau masalah dengan hasil akhir menentukan apakah penilaian tersebut bersifat positif atau negatif. (Somantri & Apriliani, 2018).

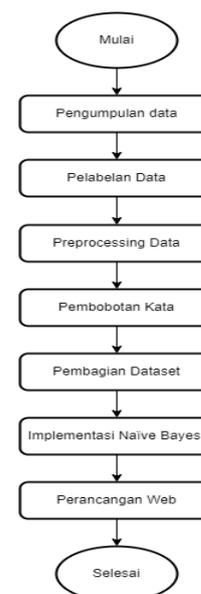
Berdasarkan penelitian analisis sentimen sebelumnya oleh (Nugroho et al., 2016), algoritma *Naïve Bayes* menghasilkan tingkat akurasi sebesar 80%. Pada penelitian (Apriani & Gustian, 2019), *Naïve Bayes* menghasilkan akurasi sebesar

97,13%. Sedangkan pada penelitian (Sanjaya et al., 2023), *Naïve Bayes* menghasilkan akurasi sebesar 85% sedangkan *Support Vector Machine* sebesar 81%. *Naïve Bayes Classifier* adalah metode *machine learning* yang menggunakan konsep perhitungan probabilitas dan statistik dari konsep yang diusulkan oleh ilmuwan Inggris, Thomas Bayes. Metode ini bertujuan untuk memprediksi probabilitas di masa depan berdasarkan pengalaman dari kejadian sebelumnya. (Ginting & Trinanda, 2013). Oleh karena itu, pada penelitian ini *Naïve Bayes Classifier* digunakan untuk menganalisis sentimen pengguna Jenius.

Bahasa pemrograman python digunakan untuk melakukan pemrosesan data serta *tools* WordPress dan XAMPP untuk keperluan perancangan web pada penelitian ini. Sentimen akan dikategorikan ke dalam tiga jenis yaitu positif, netral, dan negatif. Kata-kata dominan dari setiap jenis sentimen divisualisasikan ke dalam bentuk *word cloud*. *Word cloud* menghasilkan citra visual dari teks tertulis yang digunakan sebagai faktor pendukung proses analisis teks dengan menampilkan kata-kata yang ada dengan besar kecilnya huruf bergantung pada frekuensi munculnya kata tersebut di dalam teks.

2 Metode Penelitian

Berikut ini merupakan gambar *flowchart* alur penelitian analisis sentimen pengguna Jenius.



Gambar 1. Flowchart Penelitian

Berdasarkan *flowchart* di atas, penelitian dilakukan dengan 7 tahapan proses. Proses pertama adalah pengumpulan data yang dilanjutkan proses pelabelan data, *preprocessing* data, pembobotan kata, pembagian *dataset*, implementasi *Naïve Bayes*, dan perancangan web sebagai tahap terakhir.

2.1 Metode Pengumpulan Data

Data ulasan Jenius dikumpulkan dengan menggunakan metode *Scraping* melalui Google Colab dengan bahasa pemrograman Python. Target pengambilan data berjumlah 1000 data ulasan yang akan tersimpan dalam bentuk CSV.

2.2 Pelabelan Data

Pelabelan data adalah proses menetapkan kategori tertentu pada setiap *record* di dalam *dataset*. Label mewakili informasi atau klasifikasi yang ingin diprediksi atau dipelajari oleh model klasifikasi *Naïve Bayes*. Pelabelan data dilakukan secara manual berdasarkan kriteria *score*. Kriteria pelabelan data dimuat dalam Tabel 1.

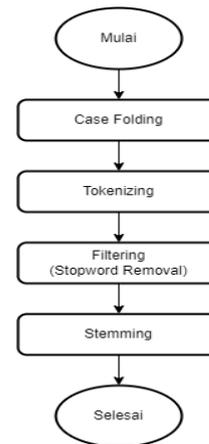
Tabel 1. Kriteria Pelabelan Data

Score	Label
5 – 4	Positif
3	Netral
2 – 1	Negatif

Berdasarkan Tabel 1, data ulasan dibagi ke dalam tiga jenis kategori sentimen. Sentimen positif dilabelkan pada ulasan dengan nilai *score* 5 atau 4. Sentimen netral dilabelkan pada ulasan dengan nilai *score* 3. Sentimen negatif dilabelkan pada ulasan dengan nilai *score* 2 atau 1.

2.3 Preprocessing data

Preprocessing adalah tahap persiapan data yang bertujuan untuk mempermudah proses pengolahan data (Lukmana et al., 2019). Peneliti melakukan proses *Preprocessing* untuk mempersiapkan data mentah agar siap untuk diolah. *Preprocessing Data* berfungsi untuk membersihkan dan meningkatkan kualitas data agar hasil analisis menjadi menjadi lebih akurat. Urutan tahapan dalam *Preprocessing* digambarkan melalui *flowchart* pada Gambar 2.



Gambar 2. Flowchart Preprocessing Data

Tahapan *Preprocessing Data* adalah sebagai berikut:

1. **Case Folding**
Proses perubahan karakter kata menjadi *lower case* atau huruf kecil agar setiap teks memiliki format yang seragam.
2. **Tokenizing**
Tokenizing adalah proses pemecahan data menjadi potongan – potongan kata berdasarkan penggunaan spasi. Tujuan dilakukannya tahap ini untuk mempersiapkan data ulasan menjadi token – token individual agar dapat diolah lebih lanjut.
3. **Filtering**
Berfokus pada penghapusan kata tidak bermakna maupun tidak relevan atau sering disebut dengan proses *Stopword Removal*. *Filtering* dilakukan untuk meningkatkan kualitas data yang akan diolah. Kata-kata yang tidak relevan tadi biasa disebut dengan *stop words* yang akan di eliminasi, contoh *stop words* dalam Bahasa Indonesia adalah di, dari, ke, atau.
4. **Stemming**
Proses perubahan kata ke bentuk dasarnya. *Stemming* berguna untuk mengurangi variasi kata di dalam dataset pada saat akan diolah.

2.4 Pembobotan Kata

Pemberian nilai untuk setiap kata dalam suatu dokumen dikenal sebagai proses pembobotan kata. (Gunawan et al., 2018). Pada penelitian ini, pemberian bobot pada kata-kata dilakukan dengan menerapkan metode TF-IDF untuk mengukur

frekuensi atau jumlah kemunculan kata tertentu dalam *dataset* Jenius.

2.4.1 TF-IDF

Term frequency-inverse document frequency (TF-IDF) adalah suatu metode vektorisasi teks yang mengubah representasi teks menjadi vektor yang dapat digunakan secara efektif. TF-IDF mengintegrasikan dua konsep kunci, yaitu *Term Frequency* (TF) untuk mengukur seberapa sering sebuah kata muncul dalam suatu dokumen dan *Document Frequency* (DF) untuk mengukur seberapa umum kata tersebut muncul di seluruh koleksi dokumen. Untuk mengurangi bobot suatu istilah jika kemunculannya terlalu umum, digunakan konsep *Inverse Document Frequency* (IDF). IDF memberikan bobot yang lebih tinggi kepada istilah yang muncul jarang, sehingga istilah tersebut dianggap lebih informatif. Pembobotan

$$idf_i = \log\left(\frac{n}{df_i}\right) \quad (1)$$

$$w_{i,j} = tf_{i,j} \times idf_i \quad (2)$$

TF-IDF dapat dilakukan dengan rumus berikut

Dengan keterangan sebagai berikut:

- idf_i = nilai IDF untuk istilah i
- df_i = total dokumen yang berisikan istilah i
- n = jumlah total dokumen
- $w_{i,j}$ = bobot TF-IDF untuk istilah i dalam dokumen j
- $tf_{i,j}$ = jumlah kemunculan istilah i dalam dokumen j

2.5 Pembagian Dataset

Proses pembagian dataset melibatkan pemisahan dataset menjadi dua bagian, yaitu *dataset* latihan (*data training*) dan *dataset* pengujian (*data testing*). *Dataset* latihan digunakan sebagai metode pelatihan model atau algoritma *machine learning*, sementara *dataset* pengujian digunakan sebagai evaluasi performa dari model yang telah di-*train*. Pada pengujian kali ini, *dataset* Jenius dibagi menjadi tiga konfigurasi *data training* dan *data testing* yang berbeda yaitu 90:10, 80:20, dan 70:30. Fungsi *train_test_split* dari pustaka bahasa pemrograman Python *scikit-learn* digunakan untuk melakukan pembagian dataset.

2.6 Implementasi Naïve Bayes

Proses implementasi *Naïve Bayes Classifier* dimulai dengan menyiapkan dua kategori klasifikasi data mencakup *dataset* latihan dan *dataset* uji yang dihasilkan dari tahapan pembagian dataset sebelumnya. *Dataset* latihan digunakan untuk melatih model *Naïve Bayes Classifier* dengan memahami pola dan relasi antara fitur (atribut) dan label kelas. *Dataset* pengujian merupakan sebagian dari *dataset* yang tidak dimanfaatkan selama fase pelatihan atau *training* yang digunakan sebagai variabel pengujian kinerja model yang dilatih pada *data training*. Dalam implementasinya, klasifikasi dengan Naïve Bayes digunakan untuk mengukur nilai atau bobot dari setiap kata pada subset data dari hasil proses pembagian dataset.

Pada penelitian ini, proses klasifikasi dilakukan dengan membuat objek *Naïve Bayes* menggunakan *MultinomialNB*. Pelatihan model *Naïve Bayes* dilakukan dengan menerapkan fungsi *fit* pada data training. Hasil pelatihan model *Naïve Bayes* digunakan untuk memprediksi kelas pada data training dengan menggunakan fungsi *predict*. Hasil akurasi model Naïve Bayes dapat diperoleh dengan menggunakan fungsi *accuracy_score*. Hasil dari pengimplementasian algoritma *Naïve Bayes Classifier* akan menghasilkan *confusion matrix* yang memuat nilai *accuracy*, *precision*, *recall*, dan *f1*.

2.7 Perancangan Web

Web visualisasi sentimen pengguna Jenius dirancang dengan menggunakan WordPress *local* yang terkoneksi dengan Localhost XAMPP. Perancangan web dilakukan dengan Langkah-langkah berikut:

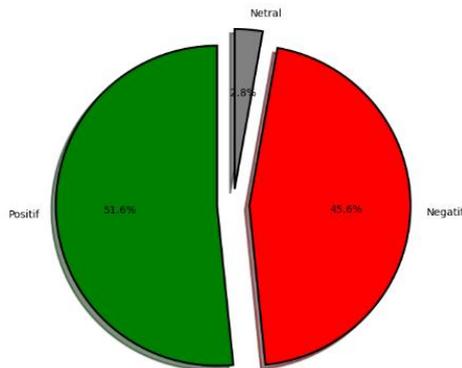
1. Membuat database baru yang akan dikoneksikan dengan WordPress.
2. Melakukan setup WordPress melalui Localhost XAMPP dengan database yang telah dibuat.
3. Melakukan instalasi WordPress melalui Localhost XAMPP.
4. Merancang tampilan web setelah berhasil masuk ke halaman dashboard WordPress.

3 Hasil dan Pembahasan

3.1 Pengumpulan Data

Proses pengambilan data dari ulasan aplikasi Jenius di Google Play menggunakan metode

scraping sukses mengumpulkan sejumlah total 1000 data ulasan dengan distribusi sentimen yang divisualisasikan pada Gambar 3.



Gambar 3. Distribusi Sentimen

Berdasarkan Gambar 3, didapati bahwa dari 1000 data ulasan di dominasi oleh sentimen positif dan negatif. Presentase distribusi sentimen terbagi menjadi 51,6% (516 data) sentimen positif, 45,6% (456 data) sentimen negatif, dan 2,8% (28 data) sentimen netral.

3.2 Preprocessing Data

3.2.1 Case Folding

Pada tahap ini, semua karakter dalam teks diubah menjadi huruf kecil agar teks menjadi seragam. Perbandingan sebelum dan setelah penerapan proses *Case Folding* dapat ditemukan dalam Tabel 2.

Tabel 2. Proses Case Folding

Sebelum Case Folding	Setelah Case Folding
Aplikasi keuangan yang terbaik si ini Penggunaannya mudah dan simpel bisa top up dengan mudah mempermudah juga untuk membayar berlangganan netflix e walet yang mantablah	aplikasi keuangan yang terbaik si ini penggunaannya mudah dan simpel bisa top up dengan mudah mempermudah juga untuk membayar berlangganan netflix e walet yang mantablah

3.2.2 Tokenizing

Pada tahap ini setiap teks dibagi menjadi token-token individual berdasarkan penggunaan spasi. Tabel 3 di bawah ini memuat hasil sebelum dan sesudah dilakukannya proses *Tokenizing*.

Tabel 3. Proses Tokenizing

Sebelum Tokenizing	Setelah Tokenizing
Aplikasi keuangan yang terbaik si ini Penggunaannya mudah dan simpel bisa top up dengan mudah mempermudah juga untuk membayar berlangganan netflix e walet yang mantablah	aplikasi keuangan yang terbaik si ini penggunaannya mudah dan simpel bisa top up dengan mudah mempermudah juga untuk membayar berlangganan netflix e walet yang mantablah

aplikasi keuangan yang terbaik si ini penggunaannya mudah dan simpel bisa top up dengan mudah mempermudah juga untuk membayar berlangganan netflix e walet yang mantablah	['aplikasi', 'keuangan', 'yang', 'terbaik', 'si', 'ini', 'penggunaannya', 'mudah', 'dan', 'simpel', 'bisa', 'top', 'up', 'dengan', 'mudah', 'mempermudah', 'juga', 'untuk', 'membayar', 'berlangganan', 'netflix', 'e', 'walet', 'yang', 'mantablah']
--	---

3.2.3 Filtering

Pada proses *Filtering* atau *Stopword Removal*, dilakukan penghapusan karakter tertentu dari teks seperti tanda baca, karakter khusus, atau angka yang tidak relevan seperti kata “yang”, “dan”, “di”. Hasil sebelum dan sesudah proses *Filtering* berjalan dapat dilihat pada Tabel 4.

Tabel 4. Proses Filtering

Sebelum Filtering	Setelah Filtering
['aplikasi', 'keuangan', 'yang', 'terbaik', 'si', 'ini', 'penggunaannya', 'mudah', 'dan', 'simpel', 'bisa', 'top', 'up', 'dengan', 'mudah', 'mempermudah', 'juga', 'untuk', 'membayar', 'berlangganan', 'netflix', 'e', 'walet', 'yang', 'mantablah']	['aplikasi', 'keuangan', 'terbaik', 'si', 'penggunaannya', 'mudah', 'simpel', 'top', 'up', 'mudah', 'mempermudah', 'membayar', 'berlangganan', 'netflix', 'e', 'walet', 'mantablah']

3.2.4 Stemming

Pada tahap *stemming* dilakukan penghapusan akhiran kata atau imbuhan untuk mengembalikan kata ke bentuk dasarnya. Perbandingan hasil sebelum dan setelah penerapan *Stemming* dapat ditemukan dalam Tabel 5.

Tabel 5. Proses Stemming

Sebelum Stemming	Setelah Stemming
['aplikasi', 'keuangan', 'terbaik', 'si', 'penggunaannya', 'mudah', 'simpel', 'top', 'up', 'mudah', 'up', 'mudah', 'mempermudah', 'membayar', 'berlangganan', 'netflix', 'e', 'walet', 'mantablah']	['aplikasi', 'uang', 'baik', 'si', 'guna', 'mudah', 'simpel', 'top', 'up', 'mudah', 'mudah', 'bayar', 'langgan', 'netflix', 'e', 'walet', 'mantablah']

1.1. Pembobotan TF-IDF

Proses pembobotan TF-IDF terhadap dataset Jenius dilakukan dengan menghitung frekuensi kata yang muncul di dalam *dataset*. Gambar 4 di bawah ini menampilkan hasil dari pembobotan TF-IDF.

```
print(X_final)
(0, 827)    0.2165270302330912
(0, 1234)   0.2165270302330912
(0, 1245)   0.2165270302330912
(0, 777)    0.22941870699938124
(0, 2359)   0.22941870699938124
(0, 770)    0.22941870699938124
(0, 424)    0.22941870699938124
(0, 425)    0.12644561530707765
(0, 767)    0.32660661202020613
(0, 1000)   0.1546572314360704
(0, 994)    0.2165270302330912
(0, 235)    0.07565015368050211
(0, 550)    0.159558419473949
(0, 1776)   0.3564939426242633
(0, 999)    0.16330330601010307
(0, 507)    0.18958737519878652
(0, 1366)   0.22941870699938124
(0, 1105)   0.1546572314360704
(0, 987)    0.25543515280353146
(0, 240)    0.18958737519878652
(0, 1727)   0.15176480312095628
(0, 790)    0.11173659894494278
(0, 1177)   0.1109610778871533
(0, 183)    0.0943387055186848
(1, 131)    0.03207634738957438
:
:
```

Gambar 4. Hasil TF-IDF

Berdasarkan Gambar 4, rendah atau tingginya nilai bobot suatu kata dapat ditafsirkan berdasarkan hasil pembobotan TF-IDF yang mendekati 0 atau 1. Jika nilai pembobotan mendekati 0 menandakan bahwa kata tersebut tidak begitu penting dengan frekuensi kemunculan yang tinggi di dalam dokumen. Sedangkan nilai pembobotan yang mendekati 1 menandakan bahwa kata tersebut memiliki nilai kepentingan tinggi dengan frekuensi kemunculan sedikit.

3.3 Klasifikasi Naïve Bayes

Proses klasifikasi diterapkan terhadap hasil pembagian *dataset* sebelumnya. Klasifikasi dilakukan pada tiga konfigurasi *data training: data testing* untuk mencari konfigurasi dengan tingkat akurasi tertinggi.

3.3.1 Konfigurasi 90:10

Pada konfigurasi 90:10, klasifikasi *Naïve Bayes* diterapkan pada 90% *data training* dan 10% *data testing*. Hasil klasifikasi *Naïve Bayes* terhadap dataset Jenius dengan konfigurasi 90:10 dapat dilihat pada Gambar 5.

```
MultinomialNB Accuracy: 0.8838709677419355
MultinomialNB Precision: 0.8945251880735753
MultinomialNB Recall: 0.8838709677419355
MultinomialNB F1: 0.883355402454968
confusion matrix:
[[45  9  0]
 [ 0 52  0]
 [ 7  2 40]]
-----
              precision    recall  f1-score   support

 Negatif      0.87         0.83         0.85         54
  Netral      0.83         1.00         0.90         52
  Positif      1.00         0.82         0.90         49

 accuracy          0.88         0.88         0.88        155
 macro avg         0.90         0.88         0.88        155
 weighted avg      0.89         0.88         0.88        155
```

Gambar 5. Confusion Matrix Naïve Bayes konfigurasi 90:10

Pada konfigurasi 90:10, klasifikasi *Naïve Bayes* menghasilkan tingkat *accuracy* sebesar 0.8838 atau 88%, tingkat *precision* sebesar 0.8945 atau 89%, tingkat *recall* sebesar 0.8838 atau 88%, dan tingkat *f1* sebesar 0.8833 atau 88%.

3.3.2 Konfigurasi 80:20

Pada konfigurasi 80:20, klasifikasi *Naïve Bayes* diterapkan pada 80% *data training* dan 20% *data testing*. Hasil klasifikasi *Naïve Bayes* terhadap dataset Jenius dengan konfigurasi 80:20 dapat dilihat pada Gambar 6.

```
MultinomialNB Accuracy: 0.8935483870967742
MultinomialNB Precision: 0.9045756776353343
MultinomialNB Recall: 0.8935483870967742
MultinomialNB F1: 0.8936800448474785
confusion matrix:
[[ 92 19  0]
 [  0 100  0]
 [ 10  4 85]]
-----
              precision    recall  f1-score   support

 Negatif      0.90         0.83         0.86        111
  Netral      0.81         1.00         0.90        100
  Positif      1.00         0.86         0.92         99

 accuracy          0.89         0.89         0.89        310
 macro avg         0.90         0.90         0.89        310
 weighted avg      0.90         0.89         0.89        310
```

Gambar 6. Confusion Matrix Naïve Bayes konfigurasi 80:20

Pada konfigurasi 80:20, klasifikasi *Naïve Bayes* menghasilkan tingkat akurasi sebesar 0.8935 atau 89%, tingkat *precision* sebesar 0.9045 atau 90%, tingkat *recall* sebesar 0.8935 atau 89%, dan tingkat *f1* sebesar 0.8936 atau 89%.

3.3.3 Konfigurasi 70:30

Pada konfigurasi 70:30, klasifikasi *Naïve Bayes* diterapkan pada 70% *data training* dan 30% *data testing*. Gambar 7 dibawah ini menunjukkan hasil klasifikasi *Naïve Bayes* pada *dataset* Jenius dengan konfigurasi 70:30.

MultinomialNB Accuracy:	0.9032258064516129			
MultinomialNB Precision:	0.914615330597147			
MultinomialNB Recall:	0.9032258064516129			
MultinomialNB F1:	0.9038019550302426			
confusion matrix:	[[131 26 0] [0 145 0] [11 8 144]]			
	precision	recall	f1-score	support
Negatif	0.92	0.83	0.88	157
Netral	0.81	1.00	0.90	145
Positif	1.00	0.88	0.94	163
accuracy			0.90	465
macro avg	0.91	0.91	0.90	465
weighted avg	0.91	0.90	0.90	465

Gambar 7. Confusion Matrix *Naïve Bayes* konfigurasi 70:30

Pada konfigurasi 70:30, klasifikasi *Naïve Bayes* menghasilkan tingkat *accuracy* sebesar 0.9032 atau 90%, tingkat *precision* sebesar 0.9146 atau 91%, tingkat *recall* sebesar 0.9032 atau 90%, dan tingkat *f1* sebesar 0.9038 atau 90%.

Untuk memudahkan perbandingan akurasi, hasil pengujian *Naïve Bayes* pada ketiga konfigurasi dituliskan ke dalam tabel berikut

Tabel 6. Hasil Pengujian *Naïve Bayes*

Pengujian <i>Naïve Bayes</i>	
Konfigurasi	Akurasi
90:10	88%
80:20	89%
70:30	90%

Berdasarkan Tabel 6, didapati bahwa konfigurasi 70:30 menempati posisi tertinggi dengan hasil akurasi sebesar 90% diikuti konfigurasi 80:20 dengan hasil akurasi sebesar 89% dan konfigurasi 90:10 menempati posisi terakhir dengan hasil akurasi sebesar 88%.

3.4 Word cloud

3.4.1 Worcloud Sentimen Positif

Kata dominan dari kategori sentimen positif dipresentasikan ke dalam *word cloud* yang dapat dilihat pada Gambar 8.



Gambar 8. Word Cloud Sentimen Positif

Berdasarkan Gambar 8, *word cloud* dari sentimen positif menampilkan beberapa kata dominan. Kata dominan yang diperlihatkan *word cloud* sentimen positif seperti “mudah”, “simpl”, “bagus”, dan “cepat”.

3.4.2 Worcloud Sentimen Netral

Kata dominan dari kategori sentimen Netral dipresentasikan ke dalam *word cloud* yang dapat dilihat pada Gambar.



Gambar 9. Word Cloud Sentimen Netral

Berdasarkan Gambar 9, *word cloud* dari sentimen netral menampilkan beberapa kata dominan. Kata dominan yang diperlihatkan *word cloud* sentimen netral seperti “tolong”, “loadingnya”, “lambat”, dan “ngelag”.

3.4.3 Wordcloud Sentimen Negatif

Kata dominan dari kategori sentimen Negatif dipresentasikan ke dalam word cloud yang dapat dilihat pada Gambar 10.



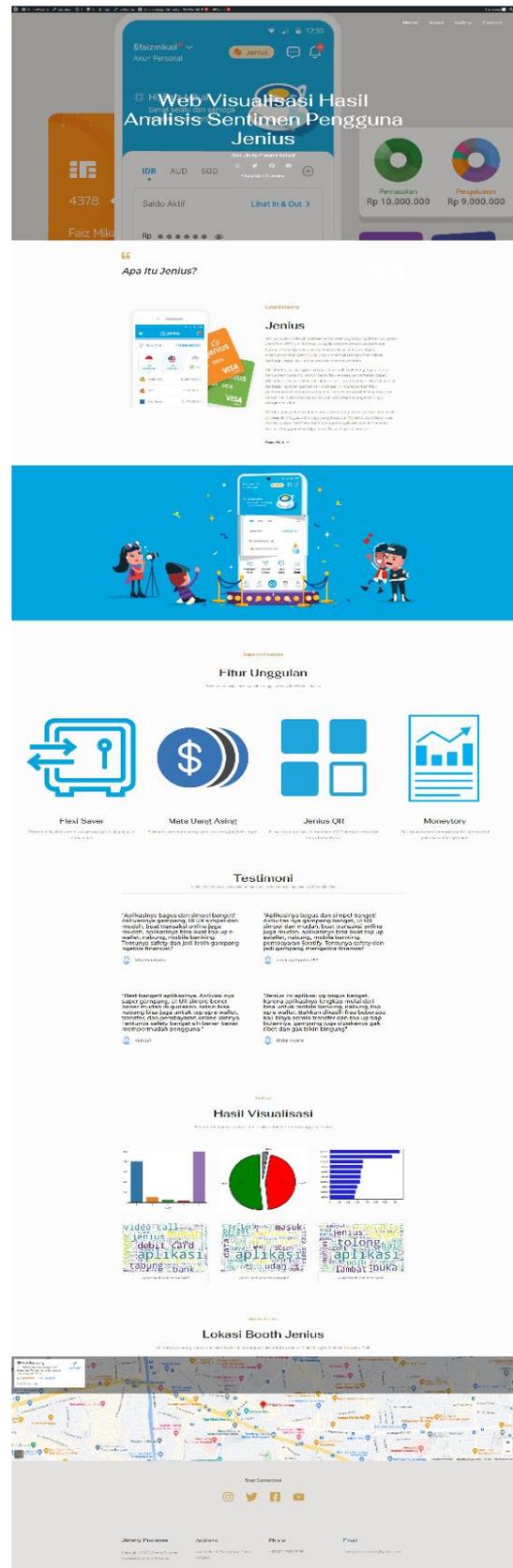
Gambar 10 Word Cloud Sentimen Negatif

Berdasarkan Gambar 10, *word cloud* dari sentimen negatif menampilkan beberapa kata dominan. Kata dominan yang diperlihatkan *word cloud* sentimen negatif seperti “verifikasi”, “sulit”, “ribet”, dan “gagal”.

3.5 Tampilan Web

3.5.1 Halaman Home

Pada halaman Home, menampilkan beberapa *section* yaitu pengenalan Jenius, fitur unggulan yang dimiliki Jenius, testimoni aplikasi Jenius yang diambil dari ulasan Google Play, sekilas hasil visualisasi, dan *maps* dari beberapa lokasi stan Jenius di Kota Semarang. Gambar 11 di bawah ini menampilkan halaman Home.



Gambar 11. Halaman Home



3.5.2 Halaman About

Pada halaman ini, ditampilkan beberapa *section* yaitu latar belakang Jenius, penghargaan yang diraih Jenius, dan latar belakang peneliti yang memuat alasan dibuatnya web ini. Gambar 12 berikut menunjukkan antarmuka halaman About.



Gambar 12. Halaman About

3.5.3 Halaman Gallery

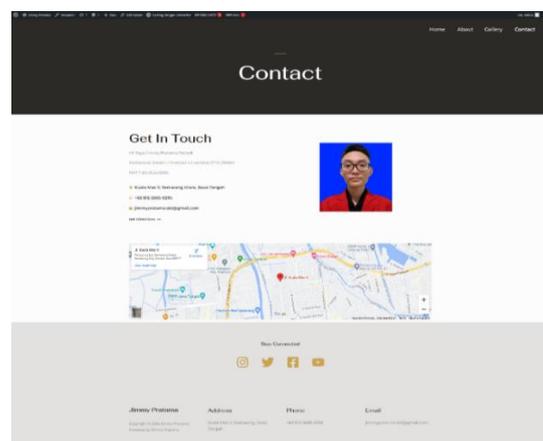
Pada halaman ini, informasi visualisasi dari hasil analisis sentimen Jenius ditampilkan ke dalam beberapa *section* yaitu visualisasi *dataset* Jenius, visualisasi data sentimen yang berisi word cloud dan diagram batang 10 kata paling populer, dan hasil pengujian *Naive Bayes* beserta kesimpulannya. Gambar 13 di bawah ini menampilkan antarmuka halaman Gallery.



Gambar 13. Halaman Gallery

3.5.4 Halaman Contact

Pada halaman Contact, menampilkan informasi data diri peneliti yang memuat foto, alamat, no telp, email, dan media social peneliti. Gambar 14 berikut menunjukkan tampilan halaman Contact.



Gambar 14. Halaman Contact

4 Kesimpulan

Dari hasil penelitian ini, peneliti dapat menyimpulkan bahwa:

1. Penelitian ini sukses mengategorikan data ulasan Jenius ke dalam tiga jenis kategori sentimen. Hasil kategori sentimen memuat sentimen positif, sentimen netral, dan sentimen negatif.
2. Algoritma *Naïve Bayes Classifier* terbukti sangat efektif dalam melakukan analisis sentimen, tingkat akurasi rata-rata dari ketiga pengujian sebesar 89% dengan akurasi tertinggi sebesar 90% pada konfigurasi pengujian 70% *data training* dan 30% *data testing*.
3. Hasil visualisasi *word cloud* menunjukkan kata – kata dominan yang membentuk kategori sentimen tersebut, diantaranya sentimen positif terbentuk dari respon pengguna atas mudahnya penggunaan aplikasi yang digunakan sebagai acuan bagi Jenius untuk mempertahankan keunggulan tersebut. Sentimen negatif terbentuk dari sulitnya proses verifikasi dan sentimen netral terbentuk dari keluhan pengguna atas kinerja aplikasi Jenius yang kurang optimal karena waktu *loading* yang lama dan sering terjadi *lag* yang dapat digunakan oleh Jenius sebagai acuan untuk melakukan *improvement* guna meningkatkan kepuasan dan loyalitas pengguna.

References

- Apriani, R., & Gustian, D. (2019). Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia. *Jurnal Rekayasa Teknologi Nusa Putra*, 6(1), 54-62.
- Fiarni, C., Maharani, H., & Pratama, R. (2016). *Sentiment Analysis System for Indonesia Online Retail Shop Review Using Hierarchy Naive Bayes Technique. Conference: 2016 4th International Conference on Information and Communication Technology (ICoICT)*, 212-217. <https://doi.org/10.1109/ICoICT.2016.7571912>
- Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2016). Analisis Sentimen pada Jasa Ojek Online Menggunakan Metode Naïve Bayes. *Prosiding Seminar Sains Nasional dan Teknologi*, 1(1), 156-161.
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode *Naive Bayes*. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 4(2), 17-29. www.femaledaily.com
- Ginting, S. L. B., & Trinanda, R. P. (2013). Teknik Data Mining Menggunakan Metode Bayes Classifier untuk Optimalisasi Pencarian pada Aplikasi Perpustakaan (Studi Kasus: Perpustakaan Universitas Pasundan – Bandung). *Jurnal Teknologi dan Informasi (JATI)*, 3(2), 37-50. <https://doi.org/https://doi.org/10.34010/jati.v3i2.794>
- Maulana, R., Iskandar, & Mailany, M. (2018). Pengaruh Penggunaan Mobile Banking terhadap Minat Nasabah dalam Bertransaksi Menggunakan Technology Acceptance Model. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 2(2), 146-155.
- Natasuwarna, A. P. (2020). Seleksi Fitur Support Vector Machine pada Analisis Sentimen Keberlanjutan Pembelajaran Daring. *Techno.Com: Jurnal Teknologi Informasi*, 19(4), 437-448.
- Sanjaya, T. P. R., Ahmad Fauzi, & Anis Fitri Nur Masruriyah. (2023). Analisis Sentimen Ulasan pada E-Commerce Shopee Menggunakan Algoritma *Naive Bayes* dan *Support Vector Machine*. *INFOTECH: Jurnal Informatika & Teknologi*, 4(1), 16-26. <https://doi.org/10.37373/infotech.v4i1.422>
- Somantri, O., & Apriliani, D. (2018). Support Vector Machine Berbasis Feature Selection untuk Sentiment Analysis Kepuasan Pelanggan terhadap Pelayanan Warung dan Restoran Kuliner Kota Tegal. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 5(5), 537-548. <https://doi.org/10.25126/jtiik20185867>
- Lukmana, D. T., Subanti, S., & Susanti, Y. (2019). Analisis Sentimen terhadap Calon Presiden 2019 dengan Support Vector Machine di Twitter. *Seminar & Conference Proceedings of UMT*, 154-160.

