

Penerapan Metode Naïve Bayes Untuk Wacana Kenaikan Harga Tiket Candi Borobudur Pada Twitter

Dani Juhaeni^{1*}, Arief Wibowo²

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

Email: ^{1*}danijuhaeni22@gmail.com, ²arief.wibowo@budiluhur.ac.id
(*: corresponding author)

Abstrak-Penyebaran informasi mengenai wacana kenaikan harga tiket candi Borobudur sempat menimbulkan reaksi pro dan kontra dari masyarakat. Hasil dari keluarnya wacana tersebut menimbulkan perdebatan pada media sosial, terdapat berbagai macam opini yang diungkapkan oleh masyarakat. Dari opini-opini yang diungkapkan masyarakat bisa ditemukan suatu pola tertentu yang dapat digunakan untuk melakukan analisis sentimen terhadap opini-opini tersebut. Maka diperlukan sistem analisis sentimen untuk mengklasifikasikan opini atau pandangan masyarakat mengenai wacana kenaikan harga tiket candi Borobudur guna mendapatkan gambaran seberapa besar tingkat penerimaan atau penolakan dari masyarakat terkait wacana kenaikan harga tiket candi borobudur. Penelitian ini bertujuan untuk membuat sistem yang dapat melakukan klasifikasi sentimen opini positif dan opini negatif mengenai wacana kenaikan harga tiket candi Borobudur pada media sosial twitter. Metode pada penelitian ini menggunakan metode Naïve Bayes untuk melakukan klasifikasi sentimen terhadap wacana kenaikan harga tiket candi Borobudur yang direncanakan oleh pemerintah, beserta nilai akurasi dari implementasi algoritme Naïve Bayes. Dataset yang digunakan bersumber dari media sosial twitter berupa tweet yang didapatkan dengan fitur pencarian kata dengan kata kunci “borobudur naik harga” dan “tiket borobudur”. Untuk melakukan pengumpulan data tweet diperoleh dengan cara melakukan proses *crawling* data dengan menggunakan *library tweepy* yang memanfaatkan API twitter. Dari hasil pengujian dan evaluasi tertinggi yang didapat dari penelitian ini mendapatkan nilai akurasi sebesar 70.0%, nilai presisi sebesar 71.0% dan nilai recall sebesar 68.0%.

Kata Kunci: Naïve Bayes, Borobudur, Twitter

Application Naïve Bayes Method To Discourse Increase Of Borobudur Temple Ticket Prices On Twitter

Abstract-Dissemination of information regarding the discourse of increasing ticket prices for Borobudur temple had caused pro and contra reactions from the public. The results of the release of the discourse caused debate on social media, there were various kinds of opinions expressed by the public. From the opinions expressed by the public, a certain pattern can be found that can be used to conduct sentiment analysis on these opinions. So a sentiment analysis system is needed to classify public opinions or views regarding the discourse on increasing ticket prices for Borobudur temple to get an idea of how much acceptance or rejection from the public regarding the discourse on increasing ticket prices for Borobudur temple is needed. This study aims to create a system that can classify positive opinion sentiments and negative opinions regarding the discourse on increasing ticket prices for Borobudur temple on Twitter social media. The method in this study uses the Naïve Bayes method to classify sentiment on the discourse on the increase in ticket prices for Borobudur temple planned by the government, along with the accuracy value of the implementation of the Naïve Bayes algorithm. The dataset used is sourced from Twitter social media in the form of tweets obtained with the word search feature with the keywords "Borobudur rising prices" and "Borobudur tickets". To collect tweet data is obtained by crawling the data using the tweepy library that utilizes the Twitter API. From the results of the highest testing and evaluation obtained from this study, it obtained an accuracy value of 70.0%, a precision value of 71.0% and a recall value of 68.0%.

Keywords: Naïve Bayes, Borobudur, Twitter

1. PENDAHULUAN

Penyebaran informasi mengenai wacana kenaikan harga tiket candi Borobudur sempat menimbulkan pro-kontra masyarakat. Media sosial mempunyai peranan penting dalam penyebaran informasi, Twitter merupakan salah satu media sosial yang sedang populer untuk digunakan pada saat ini[1]. Twitter adalah situs microblogging yang awalnya mengizinkan pesan teks (*tweet*) hingga 140 karakter, yang kemudian diubah menjadi 280 karakter pada tahun 2017[2]. Text mining merupakan variasi dari data mining yang mencoba menemukan pola menarik dari dalam sejumlah besar data teks[3].

Berdasarkan opini masyarakat yang terdapat pada media sosial twitter dapat dibuat sebuah sentimen analisis, yang bertujuan untuk menganalisa pendapat, evaluasi, sikap, penilaian dan emosi seseorang terhadap suatu hal terhadap kejadian tertentu[4]. Dalam melakukan analisis sentimen pada penelitian ini menggunakan metode naïve bayes. Metode naïve bayes dipilih karena memiliki algoritma yang sederhana dengan nilai akurasi yang tinggi[5].

Untuk memperoleh dataset dari twitter berupa *tweet* dilakukan proses *crawling* data menggunakan *library tweepy*, dengan cara mengunduh data dari server twitter dengan bantuan Application Programming Interface (API) twitter[6]. Setelah data di dapatkan maka akan dilakukan proses *preprocessing* untuk mengeliminasi data yang tidak dibutuhkan agar memudahkan sistem untuk memproses data. Proses *preprocessing* dibutuhkan terutama untuk data yang berasal dari media sosial karena Sebagian besar data berisi kalimat yang tidak baku dan tidak terstruktur serta terdapat *noise* yang besar[7].

Pada penelitian sebelumnya [8] melakukan klasifikasi dengan persentase sentimen positif sebanyak 86%, sedangkan dengan sentimen negatif sebanyak 14%. menghasilkan nilai akurasi dari perhitungan menggunakan algoritma Naïve Bayes adalah 92,96%. Penelitian berikutnya [9] melakukan klasifikasi menggunakan metode Naïve Bayes dan *Lexicon Based Features* didapatkan nilai *accuracy*, *precision*, *recall*, dan *f-measure* sebesar 0,8, 0,8, 0,8 dan 0,8. Penelitian berikutnya [5] menunjukkan hasil evaluasi dari model Naive Bayes Classification yaitu diperoleh tingkat akurasi klasifikasi sebesar 87,34%, sensitivitas sebesar 93,43%, dan spesifisitas 71,76%.

Tujuan penelitian ini adalah untuk membuat sistem yang dapat melakukan klasifikasi sentimen opini positif dan opini negatif terhadap wacana kenaikan harga tiket candi Borobudur pada media sosial twitter. Penelitian ini diharapkan dapat membantu untuk memberikan gambaran mengenai sentimen masyarakat yang berkaitan dengan wacana kenaikan harga tiket candi Borobudur dengan cara melakukan melakukan proses klasifikasi *tweet* pada twitter.

2. METODE PENELITIAN

2.1 Data Penelitian

Dataset yang digunakan pada penelitian ini menggunakan data *tweet* dari sosial media twitter yang didapatkan dengan cara melakukan proses *crawling* data. Proses *crawling* data ini menggunakan tools *library tweepy* dengan bahasa pemrograman python. *Crawling* data dilakukan pada tanggal 5 Juni 2022 sampai 17 Juni 2022, dengan jumlah data *tweet* yang didapat sebanyak 413 data dengan kata kunci “borobudur naik harga” dan “tiket borobudur”. Data yang diambil yaitu *id*, *text (tweet)*, *username*, dan *date* (tanggal *tweet* di posting).

2.2 Penerapan Metode

Proses tahapan awal hingga akhir dari sistem yang berjalan dari rancangan sistem klasifikasi dengan menggunakan metode naïve bayes dapat dilihat pada Gambar 1 :



Gambar 1. Tahapan Metode

Untuk tahapan pengumpulan data akan dilakukan proses *crawling* data twitter menggunakan *library tweepy* dengan memanfaatkan API yang disediakan oleh twitter. *Crawling* data twitter meliputi *id*, *text (tweet)*, *username*, dan *date*. Kemudian akan dilakukan proses *preprocessing* data *tweet* untuk mengubah data menjadi data bersih. Setelah data *tweet* melalui proses preprocessing, maka tahapan selanjutnya yaitu proses *labeling* untuk memberikan *label* kelas kepada data *tweet* berupa kelas positif dan kelas negatif. Setelah data *tweet* memiliki *label* maka dilanjutkan dengan melakukan pembagian data. Pada proses pembagian data *tweet*, data *tweet* akan dibagi menjadi data *train* dan data *test* yang nanti akan digunakan dalam proses klasifikasi data. Agar pembagian data tetap seimbang untuk setiap kelas-nya maka digunakan teknik *stratified random sampling*. Lalu proses selanjutnya yaitu melakukan klasifikasi dengan melakukan pengujian terhadap data *test* berdasarkan data *train* yang telah di latih dengan menggunakan metode Naïve Bayes sehingga menghasilkan klasifikasi berupa data prediksi. Setelah data prediksi di dapatkan maka akan dibandingkan dengan data aktual untuk menghasilkan evaluasi *confusion matrix* berupa akurasi, presisi dan *recall*.

2.3 Pengumpulan Data

Pada tahapan ini dilakukan proses pengumpulan data *tweet* dari media sosial twitter dengan menggunakan *library tweepy*, dengan menggunakan kata kunci “borobudur naik harga” dan “tiket borobudur”. Untuk tahapan proses *crawling* data dari media sosial twitter proses awal yang perlu dilakukan adalah mendapatkan API twitter dengan menggunakan akun twitter *developer*. API yang dibutuhkan meliputi *api_key*, *api_key_secret*, *access_token*, dan *access_token_secret*. Setelah mendapatkan API twitter kemudian dilakukan proses *crawling* data dengan menggunakan *library tweepy*. data yang telah dikumpulkan dalam proses *crawling* akan disimpan kedalam bentuk format file excel.

2.4 Preprocessing

tahapan preprocessing ini akan dilakukan beberapa proses agar menghasilkan data tweet yang bersih, sehingga bisa dilakukan pengujian terhadap algoritma Naïve Bayes dengan lebih optimal. Proses preprocessing yang dilakukan pada tahap ini diantaranya yaitu, *case folding*, *cleansing*, *change slangword*, *remove stopword*, dan *stemming*.

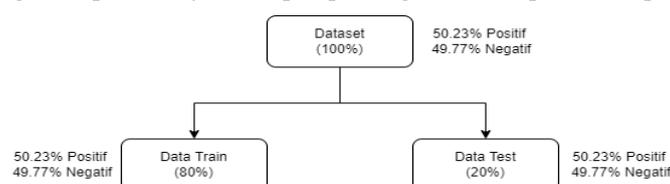
2.5 Labeling

Pada tahapan proses *labeling* akan dilakukan pemberian *label* berdasarkan 2 kategori kelas yaitu kelas positif dan kelas negatif, pengelompokan kelas ini berdasarkan isi opini suatu kalimat di dalam data *tweet*. Untuk kelas positif akan diberikan kepada data tweet yang mengandung opini sepakat dan menyetujui tentang wacana kenaikan harga tiket candi Borobudur, sedangkan kelas negatif akan diberikan kepada data tweet yang mengandung opini menolak dan mengkritik tentang wacana kenaikan harga tiket candi Borobudur. Kelas positif dan negatif tersebut dipilih berdasarkan hasil *tweet* yang dikumpulkan dalam kurun waktu tertentu dan untuk memudahkan proses klasifikasi. Proses tahapan *labeling* dilakukan dengan melakukan *labeling* manual, hasil *labeling* yang telah didapatkan kemudian di verifikasi oleh bapak Dr. Bayquni, M.IKom. selaku pakar komunikasi politik pemerintahan dan luar negeri. Pada tahapan ini diperoleh dataset yang telah memiliki *label* sebanyak 221 data.

2.6 Pembagian Data

Pada tahapan proses pembagian data akan dilakukan pembagian *dataset* yang telah memiliki *label* untuk mendapatkan data *train* dan data *test*. Pembagian *dataset* menggunakan perbandingan 80% data *train* dan 20% data *test*. Alasan digunakan perbandingan 80% data *train* dan 20% data *test* karena berdasarkan penelitian sebelumnya hasil perbandingan ini mendapatkan hasil yang baik[10].

Dalam proses pembagian dataset menggunakan teknik *stratified random sampling* sehingga proses pembagian *dataset* akan dilakukan berdasarkan strata/label dan untuk pengambilan data akan dilakukan secara acak sederhana[11]. Sehingga setiap data memiliki peluang yang sama besar untuk dapat terpilih[12]. Teknik *stratified random sampling* digunakan agar pembagian *dataset* menjadi data *train* dan data *test* menjadi proporsional untuk pembagian tiap kelas-nya. Tahapan pembagian data dapat dilihat pada Gambar 2 :



Gambar 2. Tahapan Pembagian Data

2.7 Klasifikasi Algoritma Naïve Bayes

Naïve Bayes merupakan salah satu teknik klasifikasi dalam machine learning yang berbasis probabilistik sederhana berdasarkan pada aturan bayes (teorema bayes) dengan asumsi ketergantungan yang kuat (naif)[13]. Naïve Bayes Classifier memiliki bentuk umum dapat dilihat pada persamaan 1 seperti berikut[9] :

$$P(c|w) = \frac{P(w|c) \cdot P(c)}{P(w)} \quad (1)$$

$P(c|w)$: adalah posterior atau peluang kelas c terhadap kata w .

$P(w|c)$: adalah likelihood atau peluang kata w terhadap kelas c .

$P(c)$: adalah prior atau peluang kemunculan pada kelas c .

$P(w)$: adalah evidence atau peluang kemunculan kata w .

2.8 Rancangan Pengujian

Pada tahapan proses ini akan dilakukan evaluasi terhadap hasil pengujian untuk mengetahui nilai dari akurasi, presisi dan recall berdasarkan train model yang akan dibuat menggunakan metode algoritme Naïve Bayes. Proses pengujian dilakukan dengan cara menghitung perbandingan dari data hasil prediksi dari proses klasifikasi dengan data labeling aktual yang didapatkan dari proses labeling manual. Berikut merupakan tabel 1 *confusion matrix*[1]:

Tabel 1. Confusion Matrix

| Cofusion Matrix | | Kelas Aktual | |
|-----------------|---------|--------------|---------|
| | | Positif | Negatif |
| Kelas Prediksi | Positif | TP | FP |
| | Negatif | FN | TN |

Berikut merupakan Confusion matrix yang terdapat pada sistem yang meliputi, akurasi, presisi dan recall.

a. Akurasi

Untuk mendapatkan akurasi dilakukan perhitungan dengan menggunakan persamaan 2 sebagai berikut :

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

b. Presisi

Untuk mendapatkan presisi dilakukan perhitungan dengan menggunakan persamaan 3 sebagai berikut :

$$Presisi = \frac{TP}{TP + FP} \quad (3)$$

c. Recall

Untuk mendapatkan recall dilakukan perhitungan dengan menggunakan persamaan 4 sebagai berikut :

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

3. HASIL DAN PEMBAHASAN

Berdasarkan hasil pembagian *dataset* yang telah dilakukan proses *preprocessing* dan *labeling* yang mendapatkan data sebanyak 221 data. Kemudian akan dilakukan proses pembagian data terhadap *dataset* menjadi data *train* dan data *test* dengan perbandingan 80% data *train* dan 20% data *test*, maka menghasilkan 177 data *train* dan 44 data *test*.

3.1 Pengujian metode

Pengujian merupakan salah satu proses penting dalam pembuatan sistem, guna untuk mengevaluasi dan menganalisis seberapa besar tingkat akurasi sistem yang telah dibuat. Pada penelitian ini menggunakan pengujian pada akurasi, presisi dan recall. Data yang digunakan untuk melakukan pengujian sebanyak 20% data sebesar 44 data dari total keseluruhan data dari 221 data. Pada tabel 2 merupakan hasil pengujian data dari sistem yang telah dibuat.

Tabel 2. Data hasil pengujian

| No | Teks | Label Aktual | Label Prediksi |
|----|--|--------------|----------------|
| 1 | hindar candi rusak lantai aus rangin unjun harga tiket candi tiket masuk area candi borobudur dagang aman turut | Positif | Positif |
| 2 | jaga lestari kaya sejarah budaya nusantara langkah naik harga tiket candi borobudur unesco | Positif | Positif |
| 3 | pro kontra kasi dagang borobudur pas nonton berita wawancara keluh harga wisatawan jarang harga kali lipat naik | Negatif | Negatif |
| 4 | harga masuk borobudur candi paket tour bayar masuk akal harga masuk segituan atas maintenance program batas kunjung atas timbang | Positif | Negatif |
| 5 | tuju harga biar harga masuk area borobudur bayar segitu bicara budha doa gratis niat beda wisata | Positif | Negatif |
| 6 | ingat mahal tiket atas stupa harga tiket ajar candi borobudur bangun abad situs budaya indonesia masuk dalam situs waris dunia | Positif | Positif |

| | | | |
|-----|--|---------|---------|
| 7 | tiket borobudur candi latar taman harga normal borobudur masuk ajaib dunia rusak sayang banget tambah material batu tumpuk rawan turun kikis | Positif | Positif |
| 8 | harga tiket stupa masuk kawasan borobudur lokal keliling kawasan hijau olahraga stupa waris dunia lihat bayar monggo kuota hari | Positif | Positif |
| 9 | konsep benar naikin harga tiket masuk borobudur minta tawar tawar zink jaga lestari harga tiket kaya disney | Positif | Positif |
| 10 | batas puncak candi borobudur naikin harga gila | Negatif | Negatif |
| ... | ... | ... | ... |
| 44 | tiket borobudur wisata lokal enggan kunjung borobudur menteri gratis | Negatif | Negatif |

Berdasarkan hasil pengujian diatas dapat dilihat pada Tabel 1 dengan *confusion matrix* sebagai berikut :

| Cofusion Matrix | | Kelas Aktual | |
|-----------------|---------|-------------------------|-------------------------|
| | | Positif | Negatif |
| Kelas Prediksi | Positif | 16 TP (True Positif) | 7 FP (False Positif) |
| | Negatif | 6 FN (False Negatif) | 15 TN (True Negatif) |

Berdasarkan tabel confusion matrix maka dapat diperoleh nilai akurasi, presisi dan recall sebagai berikut :

$$\begin{aligned}
 1. \text{ Akurasi} &= (TP + TN) / (TP + TN + FP + FN) \\
 &= (15 + 16) / (15 + 16 + 6 + 7) \\
 &= (31) / (44) \\
 &= 0.7 * 100\% \\
 &= 70.0\%
 \end{aligned}$$

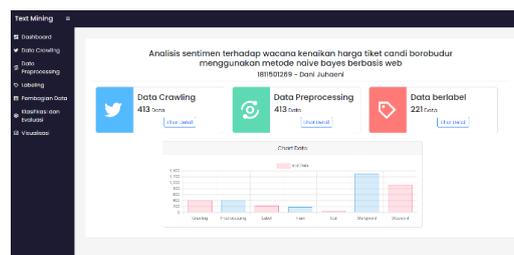
$$\begin{aligned}
 2. \text{ Presisi} &= TP / (TP + FP) \\
 &= 15 / (15 + 6) \\
 &= 15 / (21) \\
 &= 0.71 * 100\% \\
 &= 71.0\%
 \end{aligned}$$

$$\begin{aligned}
 3. \text{ Recall} &= TP / (TP + FN) \\
 &= 15 / (15 + 7) \\
 &= 15 / (22) \\
 &= 0.68 * 100\% \\
 &= 68.0\%
 \end{aligned}$$

Berdasarkan hasil pengujian dapat diketahui bahwa dengan menggunakan algoritma Naïve Bayes mampu mendapatkan nilai pengujian tertinggi dengan akurasi 70%, presisi 71.0% dan recall 68.0%. sedangkan sentimen yang diperoleh pada penelitian ini dengan topik bahasan wacana kenaikan harga tiket candi Borobudur sebesar 50,23% positif dan 49,77% negatif.

3.2 Dashboard

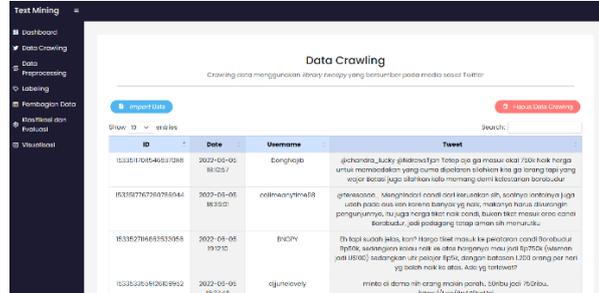
Tampilan layar *dashboard* merupakan tampilan awal dari sistem ini, pada menu *dashboard* memuat informasi yang meliputi jumlah data *crawling*, jumlah data *Preprocessing*, jumlah data berlabel dan *chart* data keseluruhan dari *dataset* yang ada. Untuk tampilan *dashboard* dapat dilihat pada Gambar 3.



Gambar 3. Dashboard

3.3 Data Crawling

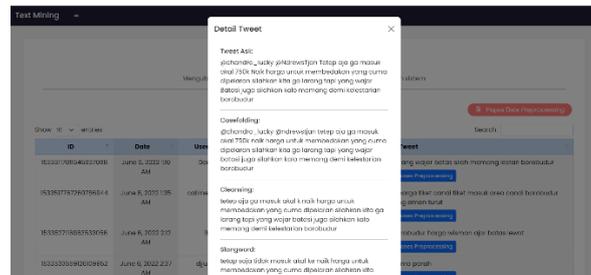
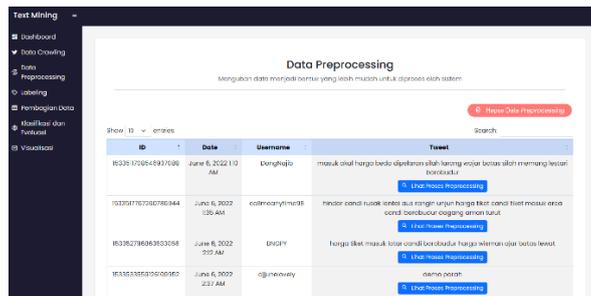
Tampilan layar data crawling berisikan tentang informasi mengenai tabel data *crawling* yang terdiri dari id, *date*, *username*, dan *tweet*. Pada menu ini terdapat proses *import* data *crawling* dengan format excel (.xlsx). Terdapat juga hapus data *crawling* yang berfungsi untuk menghapus data. Untuk tampilan *crawling* dapat dilihat pada Gambar 4.



Gambar 4. Data Crawling

3.4 Data Preprocessing

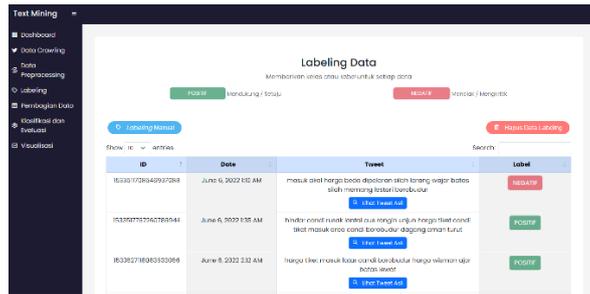
Tampilan layar preprocessing data berisikan tentang informasi mengenai tabel preprocessing data yang terdiri dari id, *date*, *username*, dan *tweet*. Pada menu ini dapat menampilkan langkah-langkah hasil dari data yang telah di preprocessing. Untuk tampilan *preprocessing* dan contoh tahapan preprocessing yang digunakan dapat dilihat pada Gambar 5



Gambar 5. Data Preprocessing

3.5 Labeling

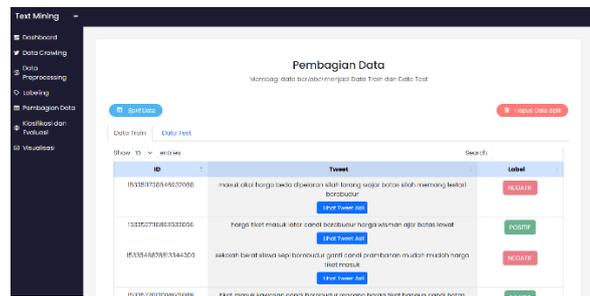
Tampilan layar labeling dari program aplikasi yang telah dibuat. Pada menu ini terdapat fungsi untuk menambahkan *label* pada data. Untuk tampilan *labeling* dapat dilihat pada Gambar 6.



Gambar 6. Labeling

3.6 Pembagian Data

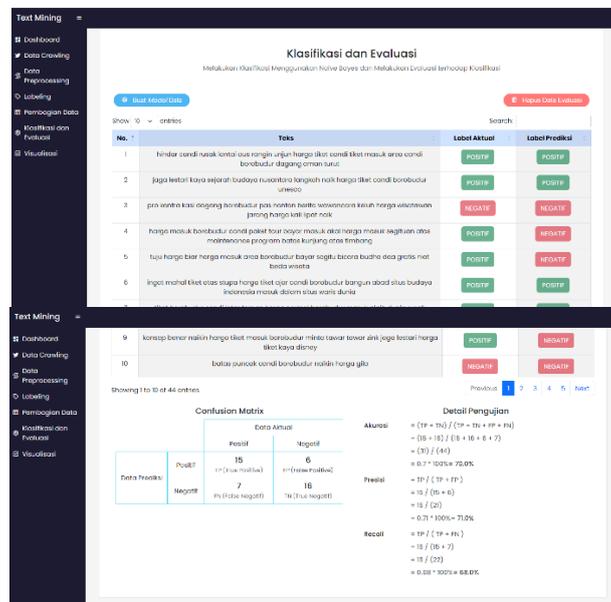
Tampilan layar pembagian data dari program aplikasi yang telah dibuat. Pada menu ini terdapat fungsi untuk melakukan pembagian *dataset* menjadi data *train* dan data *test*. Untuk tampilan pembagian data dapat dilihat pada Gambar 7.



Gambar 7. Pembagian Data

3.7 Klasifikasi dan Evaluasi

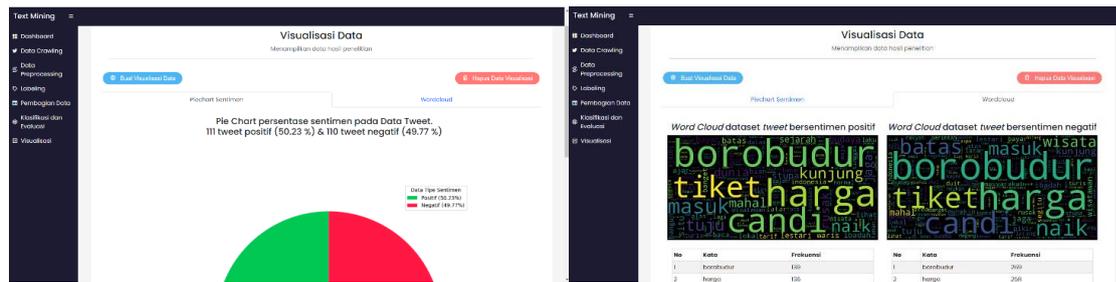
Tampilan layar klasifikasi dan evaluasi dari program aplikasi yang telah dibuat. Terdapat menu untuk membuat klasifikasi dengan menggunakan metode naïve bayes yang kemudian ditampilkan hasil evaluasi dari hasil klasifikasi. Untuk tampilan klasifikasi dan evaluasi dapat dilihat pada Gambar 8.



Gambar 8. Klasifikasi dan Evaluasi

3.8 Visualisasi

Tampilan layar visualisasi dari program aplikasi yang telah dibuat. Terdapat menu untuk membuat visualisasi dengan menggunakan *wordcloud* dan *piechart*. Untuk tampilan *visualisasi* dapat dilihat pada Gambar 9.



Gambar 9. Klasifikasi dan Evaluasi

4. KESIMPULAN

Berdasarkan hasil evaluasi dari aplikasi yang dibuat menggunakan dataset dan algoritma yang diusulkan, maka dapat disimpulkan bahwa :

- Berdasarkan dataset yang didapat dari proses crawling didapatkan sentimen masyarakat terhadap wacana kenaikan harga tiket candi Borobudur sebesar 50,23% positif dan 49,77% negatif.
- Penggunaan metode Naïve Bayes dalam melakukan analisis sentimen terhadap 20% data dari total keseluruhan data dapat berjalan dengan baik, dengan nilai evaluasi tertinggi yang didapat yaitu: nilai akurasi sebesar 70.0%, nilai presisi sebesar 71.0% dan nilai recall sebesar 68.0%.

DAFTAR PUSTAKA

- [1] Z. Firmansyah and N. F. Puspitasari, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Berdasarkan Opini Pada Twitter Menggunakan Algoritma Naive Bayes," *J. Tek. Inform.*, vol. 14, no. 2, pp. 171–178, 2021, [Online]. Available: <https://doi.org/10.15408/jti.v14i2.24024>
- [2] F. Fathonah and A. Herliana, "Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naïve Bayes," *J. Sains dan Inform.*, vol. 7, no. 2, pp. 155–164, 2021, doi: 10.34128/jsi.v7i2.331.
- [3] B. S. Tanjung, "Pendekatan Text Mining sebagai Sistem Pendeteksi Pemberitaan Palsu yang Tersebar dalam Twitter," *Semin. Nas. Ilmu Terap. 2018 – Univ. Widya Kartika*, pp. 1–6, 2018.
- [4] R. Mubarak, "Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Pembatasan Pembatasan Sosial Berskala Besar (Psbb) Dengan Metode ...," *J. Siliwangi Seri Sains dan Teknol.*, vol. 7, no. 1, pp. 19–24, 2021, [Online]. Available: <http://jurnal.unsil.ac.id/index.php/jssainstek/article/view/3726>
- [5] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, p. 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [6] B. M. Pintoko and K. M. L., "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier," *e-Proceeding Eng.*, vol. 5, no. 3, pp. 8121–8130, 2018.
- [7] Fachrul Rozy Saputra Rangkuti, "Analisis Sentimen Pengguna Twitter Terhadap Covid-19 Di Indonesia Menggunakan Metode Naïve Bayes Classifier(Nbc)," vol. 4, no. 2, pp. 43–50, 2018, [Online]. Available: <http://eprints.itn.ac.id/id/eprint/5044>
- [8] S. Lestari and S. Saepudin, "Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.)*, pp. 163–170, 2021.
- [9] Arifin Kurniawan, Indriati Indriati, and Sigit Adinugroho, "Analisis Sentimen Opini Film Menggunakan Metode Naïve Bayes dan Lexicon Based Features," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 9, pp. 8335–8342, 2019.
- [10] T. A. Tutupoly and I. Alfaro, "Komparasi Algoritma C4.5 dan Naive Bayes yang Dikembangkan menjadi Web Intelligence pada Perhitungan Bonus Tahunan Karyawan di PT. ABC," *J. Mitra Pendidik.*, vol. 3, no. 1, pp. 11–2292–103, 2019.
- [11] A. Putri, "Analisis Quick Count dengan Menggunakan Metode Stratified Random Sampling Studi Kasus Pemilu Gubernur Kalimantan Barat 2018," *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 10, no. 1, pp. 43–50, 2021.
- [12] S. F. Ulya, Y. Sukestiyarno, and P. Hendikawati, "Analisis Prediksi Quick Count Dengan Metode Stratified Random Sampling Dan Estimasi Confidence Interval Menggunakan Metode Maksimum Likelihood," *Unnes J. Math.*, vol. 7, no. 1, pp. 108–119, 2018.
- [13] R. Kararisma, S. Lestanti, and T. Chulkamdi, "Aplikasi Klasifikasi Sentimen Pada Ulasan Smartphone Di Situs Jual Beli Online Berbasis Web Menggunakan Naive Bayes Dengan Tf-Idf," *J. Mhs. Tek. Inform.*, vol. 6, no. 1, pp. 31–37, 2022.