

## Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website

Retno Sari<sup>1</sup>, Ratih Yulia Hayuningtyas<sup>2</sup>

<sup>1,2</sup>STMIK Nusa Mandiri

<sup>1</sup>retno.rnr@nusamandiri.ac.id, <sup>2</sup>ratih.ryl@nusamandiri.ac.id

**Abstract:** The tourists should first know the condition of the attraction that you want to visit by looking at the opinion of previous tourists. Data in this study were obtained from Tripadvisor, then classified into negative classes and positive classes. From this data an application will be made using Naive Bayes. The Naive Bayes Algorithm was chosen because it gives the right result in the classification of opinions in paragraph. Data processing in this study includes tokenization, stopword removal, and stemming then data will be tested with Naive Bayes Algorithm. To make it easier for tourists to know the classification of public opinion, in this study sentiment analysis applications made using the PHP Programming language and HTML to make it easier for tourist to find conclusions from an opinion.

**Keywords:** Sentiment Analysis, Naive Bayes

**Abstrak:** Para wisatawan sebaiknya terlebih dahulu mengetahui kondisi dari objek wisata yang ingin dikunjungi dengan melihat opini wisatawan sebelumnya. Data dalam penelitian ini didapat dari Tripadvisor, yang kemudian diklasifikasikan ke dalam class negatif dan class positif. Dari data tersebut akan dibuatkan aplikasi dengan Naive Bayes. Algoritma Naive Bayes dipilih karena memberikan hasil yang tepat dalam pengklasifikasian opini dalam bentuk paragraf. Pengolahan awal dalam penelitian ini meliputi *tokenization*, *stopword removal*, dan *stemming* kemudian data akan diuji dengan algoritma Naive Bayes. Untuk memudahkan wisatawan mengetahui klasifikasi opini publik maka pada penelitian ini dibuat aplikasi analisis sentimen menggunakan pemrograman PHP dan HTML untuk memudahkan wisatawan mencari kesimpulan dari sebuah opini.

**Kata kunci:** Analisa Sentimen, Naive Bayes



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2019 by author and IJSE-Indonesian Journal on Software Engineering.

### A. PENDAHULUAN

Salah satu objek wisata yang mempunyai daya tarik anjungan rumah adat, museum dan wisata alam yaitu Taman Mini Indonesia Indah (TMII). Berkunjung ke Taman Mini Indonesia Indah meninggalkan kesan dan pesan yang berbeda-beda bagi setiap wisatawan yang dikenal dengan ulasan (Hayuningtyas & Sari, 2019). Ulasan atau *opini* dari sebuah objek wisata begitu cepat meningkat dan meluas dikarenakan adanya teknologi informasi yaitu internet.

Penggunaan jejaring sosial atau media sosial semakin bertambah, ini sebagai bukti dari perkembangan internet yang mempengaruhi media komunikasi (Attabi, Muflikhah, & Fauzi, 2018). Media sosial sudah menjadi bagian terpenting untuk percakapan diseluruh dunia, karena orang-orang dapat dengan bebas mengekspresikan sebuah pendapat tentang suatu objek yang ada (Al-khafaji & Habeeb, 2017).

Sebelum para wisatawan berkunjung ke suatu objek wisata, sebaiknya wisatawan mengetahui terlebih dahulu kondisi dari objek wisata tersebut. Hal ini dapat dilihat dari opini wisatawan yang pernah mengunjungi Taman Mini Indonesia Indah. Pada penelitian ini menggunakan data ulasan yang didapat dari Tripadvisor yang menyediakan berbagai informasi mengenai restoran, objek wisata, tempat penginapan dan sebagainya.

Dari data yang dikumpulkan akan dibuatkan analisis sentimen, dimana untuk membuat sebuah analisis sentimen banyak hal yang harus dipersiapkan terlebih dahulu, salah satunya dengan memilih *classifier* yang akan digunakan. *Classifier* merupakan metode yang dapat mengklasifikasikan sebuah data menjadi beberapa kelas yaitu positif dan (Fanissa, Fauzi, & Adinugroho, 2018). Metode *Classifier* yang digunakan yaitu Algoritma Naive Bayes karena memberikan hasil yang tepat dalam pengklasifikasian sebuah opini. Opini dapat berupa paragraph yang terdiri dari beberapa kalimat dengan kalimat negatif atau kalimat positif (Nurzahputra & Muslim, 2016).

## B. TINJAUAN PUSTAKA

### 1. Analisis Sentimen

Suatu proses untuk memahami, mengekstrak dan mengolah suatu data untuk mendapatkan informasi sentimen yang terkandung pada suatu kalimat opini disebut dengan sentimen analisis (Yusnitasari, Ikasari, Pratiwi, Syahri, & Ramdani, 2017). Sedangkan menurut (Fanissa et al., 2018) "Analisis sentimen adalah bidang interdisipliner, sebuah bidang dimana pendekatan pemecahan masalahnya dengan menggunakan tinjauan dari berbagai sudut pandang ilmu serumpun secara relevan dan terpadu". Analisis sentimen terdiri dari pemrosesan bahasa, analisis teks dan komputasi linguistik untuk mengidentifikasi suatu sentimen dari dokumen (Vinodhini & Chandrasekaran, 2016).

Analisis sentimen disebut juga dengan *opinion mining*, sedangkan *opinion mining* bisa dikatakan kombinasi antara *text mining* dan *natural language processing*. *Text mining* memiliki arti penambahan sebuah data yang berupa teks yang bersumber dari sebuah dokumen, tujuan dari *text mining* yaitu mencari kata yang mewakili dari sebuah dokumen sehingga dapat dilakukan sebuah analisis yang terhubung antar dokumen (Nurzahputra & Muslim, 2016).

### 2. Algoritma Naive Bayes

Algoritma Naive Bayes merupakan salah satu algoritma yang termasuk kedalam klasifikasi. Metode ini merupakan metode terbaik dalam pengklasifikasian yang dikemukakan oleh *Thomas Bayes* dengan memprediksi peluang di masa depan berdasarkan data di masa sebelumnya.

*Naive Bayes* menerapkan fungsi statistik sederhana berdasarkan teorema bayes dengan asumsi keberadaan dari suatu fitur tertentu terhadap suatu kelas yang tidak berhubungan dengan fitur lainnya (Wilianto, Pudjiantoro, & Umbara, 2017). Naive Bayes merupakan suatu metode klasifikasi yang menggunakan perhitungan probabilitas (Ling, Kencana, & Oka, 2014).

Persamaan dari probabilitas prior:

$$P(H) = \frac{N_j}{N}$$

$N_j$  : Jumlah data pada suatu class  
 $N$  : Jumlah total data

Persamaan dari teorema bayes:

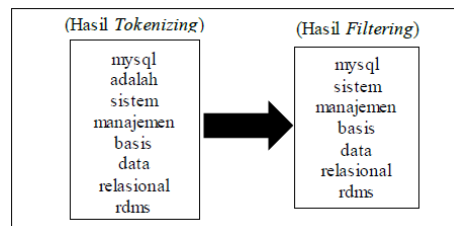
$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

$X$  : Data class belum diketahui  
 $H$  : Hipotesis data class spesifik  
 $P(H|X)$  : Probabilitas Hipotesis H terhadap kondisi X  
 $P(H)$  : Probabilitas Hipotesis H  
 $P(X|H)$  : Probabilitas X terhadap kondisi hipotesis H  
 $P(X)$  : Probabilitas X

### 3. Tokenization

*Tokenization* yaitu tahapan untuk memecahkan sekumpulan teks menjadi sebuah kata. *Tokenization* untuk menghapus tanda seperti titik (.), koma (,), spasi, dan angka yang ada pada teks tersebut (Nurzahputra & Muslim, 2016). Pada proses ini juga melakukan pembersihan karakter-karakter tertentu seperti tanda baca dengan melakukan fungsi penghilangan tanca

baca serta merubah huruf menjadi huruf kecil (Rizqiyani, Mulwinda, & Putri, 2017). Berikut adalah contoh dari *tokenization*.



Sumber: (Rizqiyani et al., 2017)

Gambar 1. Contoh *Tokenization*

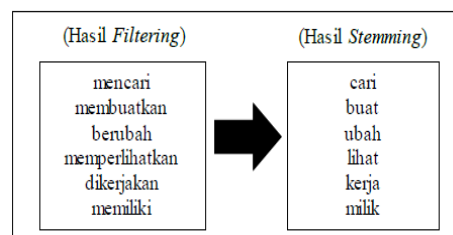
#### 4. Stopword Removal

*Filtering* merupakan tahapan untuk menghilangkan kata yang tidak memiliki makna atau *stopword*. Sedangkan *stopword* adalah kosakata yang bukan ciri/uni dari suatu dokumen (Nurzahputra & Muslim, 2016).

*Stopwords* umumnya digunakan dalam pengambilan informasi salah satu contohnya adalah mesin pencari *google*. Pengurangan indeks dalam teks dengan penghilangan beberapa kata kerja, kata sifat, dan kata keterangan lainnya dapat dimasukkan ke dalam daftar *stopwords* (Ling et al., 2014).

#### 5. Stemming

*Stemming* merupakan proses pemetaan atau pengurangan bentuk menjadi kata dasarnya. Tujuan dari proses *stemming* yaitu menghilangkan imbuhan baik itu berupa prefix, sufiks, maupun konfiks yang ada disetiap kata (Nurzahputra & Muslim, 2016). Tahap *stemming* merupakan tahap mencari akar (*root*) kata dari setiap kata hasil *filtering* (Rizqiyani et al., 2017). Berikut adalah contoh gambar dari *stemming*.



Sumber: (Rizqiyani et al., 2017)

Gambar 2. Contoh *Stemming*

#### 6. Bahasa Pemrograman

##### a. HTML

Suatu bahasa markup yang digunakan untuk membuat sebuah halaman website dan dapat memberikan berbagai informasi menggunakan browser disebut dengan *Hyper Text Markup Language* (HTML) (Saleh, 2015).

##### b. PHP

PHP merupakan singkatan dari *Hypertext Processor* yaitu suatu bahasa pemrograman yang berbasis kode-kode atau *script* yang dapat mengolah suatu data serta mengirimkannya kembali ke *browser* yang diubah menjadi kode HTML (Oktavian, 2010).

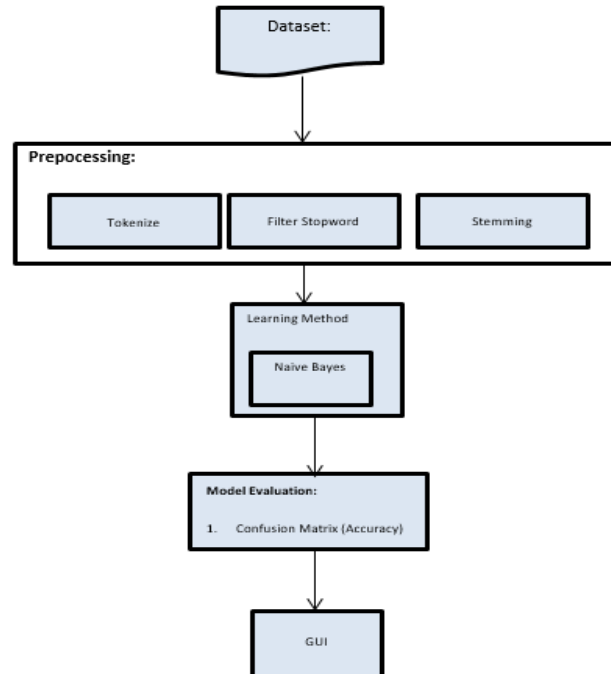
##### c. MySQL

MySQL adalah salah satu jenis *database server* yang sangat terkenal. Kepopuleran disebabkan karena MySQL menggunakan SQL sebagai bahasa dasar untuk mengakses basis datanya. MySQL merupakan server basis data yang menggunakan teknik relasional untuk menghubungkan antara tabel-tabel dalam basis datanya atau mendukung RDBMS (*Relational Database Management System*) (Rizqiyani et al., 2017).

### C. METODE PENELITIAN

Metode yang dipakai dalam penelitian ini yaitu *Naive Bayes*, pada metode ini terdapat dua proses yaitu proses *training* dan proses *testing*. Langkah awal dilakukan data *training* dan selanjutnya dilakukan proses *testing* dengan mengacu pada probabilitas dari dataset.

Berikut adalah model yang diusulkan pada penelitian ini sebagai berikut:



Sumber: Hasil Penelitian (2019)

Gambar 3. Model yang diusulkan

#### 1. Pengumpulan Data

Data untuk pengolahan data awal sampai dengan pengujian metode dari [https://www.tripadvisor.co.id/Attraction\\_Review-g294229-d379330-Reviews-Beautiful\\_Indonesia\\_in\\_Minature\\_Park-Jakarta\\_Java.html](https://www.tripadvisor.co.id/Attraction_Review-g294229-d379330-Reviews-Beautiful_Indonesia_in_Minature_Park-Jakarta_Java.html). Data yang digunakan sebanyak 100 ulasan terdiri dari 50 ulasan positif dan 50 ulasan negatif.

#### 2. Pengolahan Data Awal

Terdapat tiga proses pada tahap pengolahan data awal yaitu *tokenization*, *stopword removal*, *stemming*.

#### 3. Eksperimen dan Pengujian

Pada tahap ini data sudah dikelompokkan kedalam *Class* ulasan negatif dan positif untuk dilakukan eksperimen dan pengujian. Data training yang sudah dikelompokkan akan dilakukan pengujian menggunakan metode *Naive Bayes*.

#### 4. Model Evaluation

Model Evaluation menggunakan *confusion matrix* yaitu sebagai alat visualisasi yang biasa digunakan pada *supervised learning*.

Data class	Classified as <i>pos</i>	Classified as <i>neg</i>	$\begin{bmatrix} tp & fn \\ fp & tn \end{bmatrix}$
<i>pos</i>	true positive ( <i>tp</i> )	false negative ( <i>fn</i> )	
<i>neg</i>	false positive ( <i>fp</i> )	true negative ( <i>tn</i> )	

Sumber: (Sokolova & Lapalme, 2009)

Gambar 4. *Confusion Matrix*

Matriks evaluasi yang dapat digunakan dari gambar di atas yaitu *precision*, *recall* dan *Fscore*. Nilai *precision* diambil dari nilai positif yang benar dibagi dengan jumlah yang diberikan label positif. Sedangkan nilai *recall* diperoleh dari nilai positif yang benar dibagi dengan jumlah positif dalam data (Sokolova & Lapalme, 2009).

Measure	Formula	Evaluation focus
Accuracy	$\frac{tp+tn}{tp+fn+fp+tn}$	Overall effectiveness of a classifier
Precision	$\frac{tp}{tp+fp}$	Class agreement of the data labels with the positive labels given by the classifier
Recall (Sensitivity)	$\frac{tp}{tp+fn}$	Effectiveness of a classifier to identify positive labels

Sumber: (Sokolova & Lapalme, 2009)

Gambar 5. Klasifikasi Nilai *Confusion Matrix*

*Sensitivity (True Positive Rate/TPR)* dan *Specificity (True Negative Rate/TNR)* termasuk kedalam sebuah akurasi dimana yang pertama untuk sampel positif actual dan yang kedua untuk sampel actual sampel negatif. *Sensitivity* dan *Specificity* dapat digunakan untuk mengevaluasi suatu klasifikasi data (Tharwat, 2018).

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{P}, \quad TNR = \frac{TN}{FP + TN} = \frac{TN}{N}$$

Sumber: (Tharwat, 2018)

Gambar 6. Klasifikasi *Sensitivity* dan *Specificity*

*Positiv prediction value (PPV)* yaitu mewakili sampel positif yang diklasifikasikan dengan benar terhadap jumlah sampel yang diprediksi positif, sedangkan *Negative prediction value (NPV)* yaitu mengukur proporsi sampel negatif yang diklasifikasikan dengan benar ke total jumlah sampel yang diprediksi negatif (Tharwat, 2018). Persamaannya dapat dilihat dibawah ini

$$PPV = \text{Precision} = \frac{TP}{FP + TP} = 1 - FDR$$

$$NPV = \frac{TN}{FN + TN} = 1 - FOR$$

Sumber: (Tharwat, 2018)

Gambar 7. Klasifikasi *PPV* dan *NPV*

## 5. User Interface

Website dirancang menggunakan bahasa pemrograman MySQL, PHP, dan HTML dengan menggunakan server Xampp.

## D. HASIL DAN PEMBAHASAN

### 1. Proses Training

Dataset yang sudah diklasifikasikan kedalam *class* positif dan *class* negatif. Sebanyak 100 ulasan terdiri dari 50 *class* negatif dan 50 *class* positif.

Tabel 1. Data *Training*

Dokumen	Fitur	Class
Pos3	Sangat Menarik dan merupakan edukasi yang bagus untuk anak-anak agar mengenal secara langsung bentuk rumah adat provinsi di Indonesia sayangnya bangunan terlihat kusam	Positif
Pos15	Taman Mini Indonesia Indah sangat bagus buat tempat rekreasi dan edukasi. Disini banyak rumah adat dari seluruh indonesia	Positif
Pos18	Menurut Saya Tempat Reaksi Bersama Keluarga Adalah Yang Cocok Hanya Di Taman Mini Indonesia Indah. Berbagai Ragam Rumah Budaya Adat Indonesia Semua Ada Disini. Sekalian Mengetahui Luas, Menambah Wawasan Kita Kepada Putra Putri Kita.	Positif
Pos24	Tempatnya bagus, penuh sejarah, ada sungainya juga, pergi kesana hemat, cocok untuk anak sekolah maupun wisatawan yang ingin mengetahui budaya indonesia tetapi agak kotor	Positif
Neg2	Bagus, Sayangnya agak kotor dan bangunan terlihat kusam mungkin perlu diupgrade.	Negatif
Neg12	Merupakan ikon wisata nasional bagi bangsa ini sayangnya kurang terawat dan terlihat kotor. danau buatan yang kering sehingga bentang pulau2 di indonesia terlihat hanya menjadi gundukan tanah. beberapa anjungan propinsi yang tutup dan kurang informasi menjadi kekecewaan bagi pengunjung.	Negatif
Neg21	Taman Mini banyak menampilkan budaya dan saat ini pengelolaannya terlihat kurang sehingga banyak yang terlihat kumuh dan kusam	Negatif
Neg24	Taman Mini Indonesia Indah sebenarnya merupakan perwujudan miniatur dari daerah-daerah di Indonesia. Hanya sayangnya kurang di optimalkan, mungkin dapat mencontoh di Shenchen yang merupakan miniatur dunia, dengan adanya atraksi yang menarik. selain itu terkesan juga kotor dan kurang dijaga kebersihannya , terlalu banyak pedagang yang kurang teratur.	Negatif
Neg25	Perawatan tempat ini agak kurang sehingga ada beberapa spot baik anjungan atau museum, yang terlihat kotor dan kusam kurang terpelihara, baik bangunannya maupun kebersihannya. Tarif masuk untuk beberapa permainan pun terasa agak mahal dan yang paling membuat malas adalah macetnya.	Negatif

Sumber: Hasil Penelitian (2019)

## 2. Proses *Testing*

Berikut adalah data *testing* yang digunakan untuk uji data dengan Naive Bayes.

Tabel 2. Data *Testing*

Dokumen	Fitur	Class
x-52	Rumah adat masih terjaga, hanya di museum indonesianya kurang terpelihara, tempatnya kotor dan kurang banyak benda budaya yang bisa dimusiumkan mengingat sebenarnya tempatnya cukup besar	?

Sumber: Hasil Penelitian (2019)

Dari data *testing* tersebut dilakukan pengolahan data awal yaitu:

### a. *Tokenization*

Semua tanda baca (?.,:;',", dst), simbol (\*,\$,&, dst) dan angka (0..9) dan merubah huruf *Uppercase* menjadi *lowercase* yang ada didalam ulasan akan dihilangkan.

### b. *Stopword Removal*

Sebuah ulasan yang sudah melalui tahapan *tokenization* akan dilakukan *stopword removal*. Pada tahapan ini menghilangkan beberapa kata kerja, kata sifat, kata keterangan, dan kata penghubung seperti yang, atau, di, dengan dan sebagainya.

### c. *Stemming*

*Stemming* menghilangkan kata atau ulasan yang memiliki imbuhan menjadi kata dasar dari sebuah token.

Berikut adalah hasil dari pengolahan data dimulai dari *tokenization*, *stopword removal*, dan *stemming*.

Tabel 3. Hasil Pengolahan Data Awal

Ulasan	Tokenization	Stopword Removal	Stemming
Rumah adat masih terjaga, hanya di museum indonesianya kurang terpelihara, tempatnya kotor dan kurang banyak benda budaya yang bisa dimusiumkan mengingat sebenarnya tempatnya cukup besar	rumah adat masih terjaga hanya di museum indonesianya kurang terpelihara tempatnya kotor dan kurang banyak benda budaya yang bisa dimusiumkan mengingat sebenarnya tempatnya cukup besar	rumah adat masih terjaga hanya museum indonesianya kurang terpelihara tempatnya kotor kurang banyak benda budaya bisa dimusiumkan mengingat sebenarnya tempatnya cukup besar	rumah adat masih terjaga hanya museum indonesia kurang terpelihara tempat kotor kurang banyak benda budaya bisa museum ingat benar tempat cukup besar

Sumber: Hasil Penelitian (2019)

### 3. Eksperimen dan Pengujian Metode

Eksperimen dan pengujian menggunakan algoritma *Naive Bayes*, dengan menentukan klasifikasi sebuah ulasan termasuk ke dalam ulasan negatif atau positif. Tabel berikut adalah klasifikasi dokumen.

Tabel 4. Klasifikasi Dokumen

Dokumen	Bagus	Budaya	Kotor	Kusam	Class
Pos3	1	0	0	1	Positif
Pos15	1	0	0	0	Positif
Pos18	0	1	0	0	Positif
Pos24	1	1	1	0	Positif
Neg2	1	0	1	1	Negatif
Neg12	0	0	1	0	Negatif
Neg21	0	1	0	1	Negatif
Neg24	0	0	1	0	Negatif
Neg25	0	0	1	1	Negatif
x-51	0	1	1	0	?

Sumber: Hasil Penelitian (2019)

Berikut adalah perhitungan probabilitas dari sebuah dokumen yaitu:

#### a. Probabilitas Prior

$$P(Class="positif")=4/9=0,44$$

$$P(Class="negatif")=5/9=0,56$$

Untuk *Class* positif

$$P(x-51|Positif)$$

$$= P(bagus=3|positif)*P(budaya=2|positif)*P(kotor=1|positif)*P(kusam=1|positif)$$

$$= 0,75*0,5*0,25*0,25$$

$$= 0,0234$$

Untuk *Class* negatif

$$= P(bagus=1|negatif)*P(budaya=1|negatif)*P(kotor=4|negatif)*P(kusam=3|negatif)$$

$$= 0,2*0,2*0,44*0,33$$

$$= 0,0058$$



**b. Probabilitas Posterior**

Untuk *Class* positif  
 =  $P(X|positif) \cdot P(class="positif")$   
 =  $0,0234 \cdot 0,44$   
 = 0,010296

Untuk *Class* negatif  
 =  $P(X|negatif) \cdot P(class="negatif")$   
 =  $0,0058 \cdot 0,56$   
 = 0,0032248

Dokumen x-51 termasuk kedalam *class* positif

**4. Model Evaluation**

Pada model *evaluation* ini menggunakan perhitungan *confusion matrix* yang dapat dilihat pada tabel dibawah ini

Tabel 5. *Confusion Matrix*

Predication	Actual	
	Negatif	Positif
Pred Negatif	41 (TN)	21 (FP)
Pred Positif	9 (FN)	29 (TP)

Dari tabel di atas terdapat sebanyak 50 data *class* negatif yang terdiri dari 41 data diprediksi negatif dan 9 data yang diprediksi positif. Sebaliknya untuk data *class* positif sebanyak 50 data yang terdiri dari 21 data diprediksi negatif dan 29 data yang diprediksi positif. Dari data tersebut dapat dihitung nilai *precision*, *recall*, dan *accuracy*.

Tabel 6. Nilai *Precision*, *Recall*, *Accuracy*

<b>Precision</b>	<b>Recall</b>	<b>Accuracy</b>
0.58	1	0.7

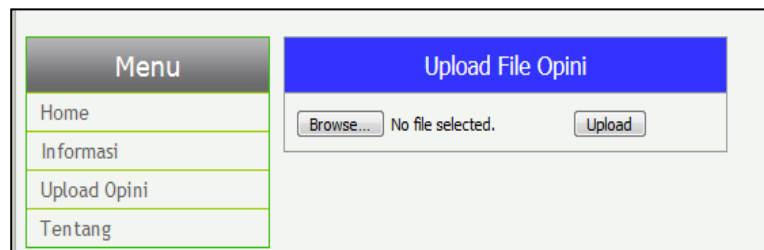
Dari tabel diatas dapat dilihat bahwa untuk nilai *precision* sebesar 0,58, *recall* sebesar 1, dan *accuracy* sebesar 0,7. Dari tabel *confusion matrix* dapat dihitung jg nilai *sensitivity*, *specificity*, *ppv* dan *npv*.

Tabel 7. Nilai *Sensitivity*, *Spesificity*, *PPV*, dan *NPV*

<b>Sensitivity</b>	<b>Spesificity</b>	<b>PPV</b>	<b>NPV</b>
0,414	0,661	0,58	0,82

**5. User Interface**

*User Interface* pada penelitian ini meliputi tampilan menu utama yang terdiri dari *home*, *informasi*, *upload opini*, dan *tentang*. Untuk menghasilkan suatu keputusan sebuah ulasan termasuk kedalam *class* negatif atau *class* positif maka pilih menu *upload opini*. Kemudian *browse file* yang akan di *upload* dan klik *upload* untuk mendapatkan hasilnya. Berikut adalah tampilan dari menu *upload opini* dan hasil.



Sumber: Hasil Penelitian (2019)

Gambar 8. Menu *Upload File Opini*





Sumber: Hasil Penelitian (2019)

Gambar 9. Hasil Klasifikasi Opini

## E. KESIMPULAN

Dari pembuatan website ini penerapan algoritma *Naive Bayes* untuk analisis sentimen bahasa Indonesia pada wisata Taman Mini Indonesia Indah berbasis website dapat disimpulkan bahwa website ini memberikan informasi mengenai opini yang ada dan mengklasifikasinya serta mendapatkan hasil apakah opini tersebut opini yang positif atau opini yang negatif. Algoritma *Naive Bayes* diterapkan dalam analisis sentimen ini dikarenakan dapat memberikan hasil yang terbaik dalam mengklasifikasikan suatu ulasan.

## REFERENSI

- Al-khafaji, H. K., & Habeeb, A. T. (2017). Efficient Algorithms for Preprocessing and Stemming of Tweets in a Sentiment Analysis System. *Journal of Computer Engineering*, 19(3), 44–50. <https://doi.org/10.9790/0661-1903024450>
- Attabi, A. W., Muflikhah, L., & Fauzi, M. A. (2018). Penerapan Analisis Sentimen untuk Menilai Suatu Produk pada Twitter Berbahasa Indonesia dengan Metode *Naive Bayes Classifier* dan Information Gain. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2, 4548–4554.
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018). Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode *Naive Bayes* dan Seleksi Fitur Query Expansion Ranking. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2, 2766–2770.
- Hayuningtyas, R. Y., & Sari, R. (2019). Analisis Sentimen Opini Publik Bahasa Indonesia Terhadap Wisata TMII Menggunakan *Naive Bayes* dan PSO. *Jurnal Techno Nusa Mandiri*, 16(1), 37–42.
- Ling, J., Kencana, I. P. E. N., & Oka, T. B. (2014). ANALISIS SENTIMEN MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER* DENGAN SELEKSI FITUR CHI SQUARE. *E-Jurnal Matematika*, 3(3), 92–99.
- Nurzahputra, A., & Muslim, M. A. (2016). Analisis Sentimen pada Opini Mahasiswa Menggunakan Natural Language Processing. In *Seminar Nasional Ilmu Komputer (SNIK 2016)* (pp. 114–118).
- Oktavian, D. P. (2010). *Menjadi Programmer Jempolan menggunakan PHP*. Yogyakarta: Mediakom.
- Rizqiyani, V., Mulwinda, A., & Putri, R. D. M. (2017). Klasifikasi Judul Buku dengan Algoritma *Naive Bayes* dan Pencarian Buku pada Perpustakaan Jurusan Teknik Elektro. *Jurnal Teknik Elektro*, 9, 60–65.

- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, 2.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45, 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*. <https://doi.org/10.1016/j.aci.2018.08.003>
- Vinodhini, G., & Chandrasekaran, R. . (2016). A Comparative Performance Evaluation of Neural Network Based Approach For Sentiment Classification of Online Reviews. *Journal Of King Saud University*, 2–12.
- Wilianto, L., Pudjiantoro, T. H., & Umbara, F. R. (2017). Analisis Sentimen Terhadap Tempat Wisata Dari Komentar Pengunjung Dengan Menggunakan Metode Naive Bayes Classifier Studi Kasus Jawa Barat. In *SNATIF*.
- Yusnitasari, T., Ikasari, D., Pratiwi, E. E. S., Syahri, N., & Ramdani. (2017). Analisis Sentimen Terhadap Review Restoran Fish Streat Pada Aplikasi Zomato Menggunakan Stemming Nazief Adriani dan Naive Bayes Classifier. In *SentriNov* (Vol. 3, pp. 2477–2097).