

**NASKAH PUBLIKASI**

**ANALISIS SENTIMEN PRODUK IPHONE MENGGUNAKAN  
NAIVE BAYES CLASSIFIER**



Disusun oleh:

**BAYU DAMAR JATI**

**5150411214**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI DAN ELEKTRO  
UNIVERSITAS TEKNOLOGI YOGYAKARTA  
2020**

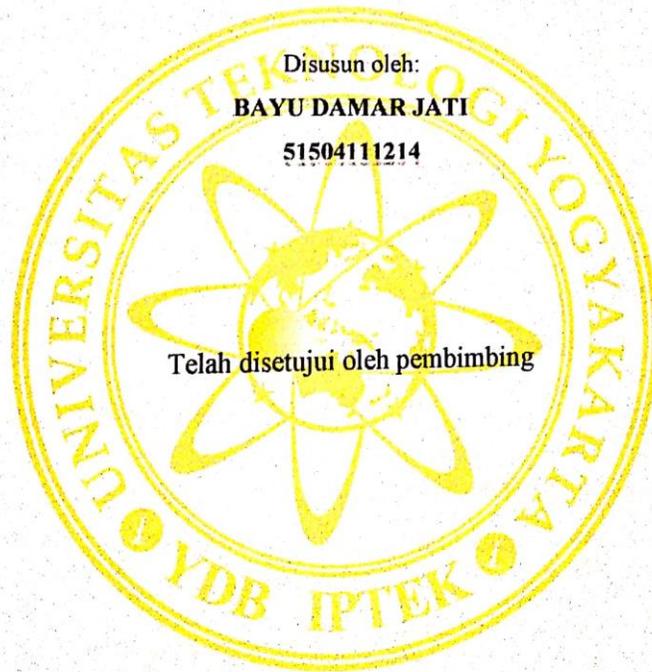
NASKAH PUBLIKASI

**ANALISIS SENTIMEN PRODUK IPHONE MENGGUNAKAN  
NAIVE BAYES CLASSIFIER**

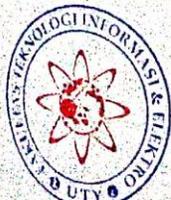
Disusun oleh:  
**BAYU DAMAR JATI**

**51504111214**

Telah disetujui oleh pembimbing



Pembimbing



**Dr. Enny Itje Sela, S. Si., M. Kom.**

Tanggal: 5 Desember 2019

# ANALISIS SENTIMEN PRODUK IPHONE MENGGUNAKAN NAIVE BAYES CLASSIFIER

**BAYU DAMAR JATI**

*Program Studi Informatika, Fakultas Teknologi Informasi dan Elektro  
Universitas Teknologi Yogyakarta*

*Jl. Ringroad Utara Jombor Sleman Yogyakarta*

*E-mail : [bayudamar123@gmail.com](mailto:bayudamar123@gmail.com)*

## ABSTRAK

*Ponsel Cerdas atau yang juga populer dengan sebutan Smartphone merupakan suatu perangkat yang banyak digunakan dan membantu manusia dalam kehidupan sehari-hari mulai dari kebutuhan komunikasi hingga hiburan. Iphone merupakan salah satu raksasa dalam industri smartphone. Analisis sentimen produk Iphone adalah proses mengekstraksi komentar dan sentimen tentang produk Iphone yang tertulis. Twitter adalah salah satu media sosial yang kerap digunakan untuk menulis berbagai komentar salah satunya komentar mengenai Iphone. Komentar yang tertulis pada media sosial Twitter dapat digunakan sebagai objek penelitian mengenai sentimen analisis dan metode Naive Bayes Classifier dapat digunakan untuk membantu proses sentimen analisis secara otomatis. Proses sentimen analisis yang dilakukan oleh sistem dapat membantu manusia dalam menganalisis komentar dalam jumlah yang banyak. Penelitian dilakukan dengan menggunakan 200 data latih yang terdiri dari masing-masing 100 data bersentimen positif dan 100 data bersentimen negatif serta 100 data uji menghasilkan akurasi sebesar 94%.*

**Kata kunci :** iphone, analisis sentimen, naive bayes classifier

## 1. PENDAHULUAN

Ponsel Cerdas atau juga sering disebut dengan *Smartphone* adalah suatu perangkat yang banyak digunakan manusia dalam kehidupan sehari-hari di era modern ini. *Smartphone* dapat membantu manusia mulai dari fungsi dasarnya sebagai sarana untuk berkomunikasi hingga fungsi sebagai sarana hiburan seperti berselancar di internet, mendengarkan musik hingga untuk bermain *game online*. Salah satu raksasa dalam industri *smartphone* adalah Apple melalui produk Iphone-nya. Iphone banyak digemari dan dibicarakan oleh banyak orang di berbagai media.

Media sosial khususnya Twitter sekarang ini menjadi perangkat komunikasi yang sangat populer di kalangan pengguna internet. Pada konferensi resmi pengembang Twitter Chirp 2010, perusahaan tersebut menyampaikan statistik mengenai situs dan pengguna Twitter. Statistik tersebut menyebutkan bahwa pada bulan April 2010, Twitter memiliki 106 juta akun dan sebanyak 180 juta pengunjung unik setiap bulannya [2].

*Tweet* adalah teks status pengguna yang digunakan untuk memberikan informasi di Twitter [7]. Pengguna Twitter dapat memberitahu tentang apa yang sedang dilakukan atau dirasakan, percakapan, berbagi informasi, dan pelaporan berita melalui *tweet*. Isi *tweet* dapat mengekspresikan opini dari penggunanya, contohnya “aku sangat suka dengan iphone”. Opini melalui *tweet* inilah yang digunakan untuk melihat sentimen yang dimunculkan salah satunya opini mengenai produk iphone.

Penentuan sentimen opini atau komentar dapat dilakukan dan dipahami secara manual akan tetapi apabila terdapat data yang banyak maka akan memakan waktu yang lebih lama. Hal ini bisa diatasi dengan membangun sistem analisis sentimen yang dapat melakukan klasifikasi sentimen positif atau negatif. Metode Naive Bayes Classifier dapat digunakan untuk membantu klasifikasi sentimen, oleh karena itu penelitian ini dilakukan dengan membangun sistem analisis sentimen dengan metode Naive Bayes Classifier untuk mengetahui performa Naive Bayes Classifier dan untuk membantu melakukan sentimen analisis secara otomatis oleh sistem.

## 2. LANDASAN TEORI

### 2.1. Analisis Sentimen

Menurut Liu [6], analisis sentimen juga disebut penambangan opini, adalah bidang studi yang menganalisis pendapat, sentimen, evaluasi, penilaian, sikap dan emosi orang-orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik dan atributnya. Terdapat banyak nama dan tugas yang sedikit berbeda, misalnya, analisis sentimen, penambangan opini, ekstraksi pendapat, penambangan sentimen, analisis subjektivitas, analisis pengaruh, analisis emosi, penambangan peninjauan, dan lain-lain namun sekarang semuanya berada di bawah payung analisis sentimen atau penambangan opini.

### 2.2. Naive Bayes Classifier

Menurut Supriyanto, [12] Naive Bayes Classifier (NBC) adalah metode klasifikasi berdasarkan probabilitas dan Teorema Bayes dengan asumsi bahwa setiap variabel X bersifat bebas (*independence*), dengan mengasumsikan bahwa keberadaan sebuah atribut (variabel) tidak ada kaitannya dengan atribut (variabel) yang lain, sedangkan menurut Rosi, F. dkk. [10], Naive Bayes adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk machine learning dan *data mining*. Performa Naive Bayes yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi keindependenan atribut (tidak ada kaitan antar atribut).

Naive Bayes Classifier dilakukan dengan melakukan perhitungan *prior probability*, perhitungan *conditional probability*, dan perhitungan *posterior probability*. Rumus perhitungan *prior probability* menurut Rahman, A. dkk. [9] dapat dilihat pada persamaan 1.

$$P(c) = \left(\frac{N_c}{N}\right) \quad (1)$$

Keterangan:

$P(c)$  = *Prior Probability* dari kelas c.

$N$  = Jumlah keseluruhan dokumen.

$\frac{N_c}{N}$  = Jumlah keseluruhan dokumen kelas c.

Perhitungan *conditional probability* dilakukan dengan model multinomial. Sementara menurut Rahman, A. dkk. [9], perhitungan *posterior probability* dilakukan dengan rumus yang dapat dilihat pada persamaan 2.

$$P(c|term \text{ dokumen } d) = P(c) \times P(T_1|c) \times P(T_2|c) \times \dots \times P(T_n|c) \quad (2)$$

Keterangan:

$P(c | \text{term dokumen } d)$  = *Posterior Probability* suatu dokumen termasuk kelas c.

$P(c)$  = *Prior Probability* dari kelas c.

$P(T_n|c)$  = *Conditional Probability* kata ke-n dengan diketahui kelas c.

### 2.3. Multinomial Naive Bayes

Multinomial Naive Bayes digunakan untuk menghitung *conditional probability* setiap kata yang ada. Menurut Rahman, A. dkk. [9], model multinomial memperhitungkan frekuensi setiap kata yang muncul pada dokumen. Rumus perhitungan model multinomial dapat dilihat pada persamaan 3.

$$P(T_n|c) = \frac{\text{count}(tn,c)+1}{\text{count}(c)+|V|} \quad (3)$$

Keterangan:

$P(T_n|c)$  = *Conditional Probability* kata ke-n dengan diketahui kelas c.

$\text{count}(tn, c)$  = Jumlah *term* tn yang ditemukan di seluruh data pelatihan dengan kategori c.

$\text{count}(c)$  = Jumlah *term* di seluruh data pelatihan dengan kategori c.

$|V|$  = Jumlah seluruh *term* pada data pelatihan.

Model multinomial juga dapat dipadukan dengan pembobotan kata TF-IDF. Rumus perhitungan dari model multinomial yang dikombinasikan dengan pembobotan kata TF-IDF menurut Rahman, A. dkk. [9] dapat dilihat pada persamaan 4.

$$P(T_n|c) = \frac{W_{ct} + 1}{\sum W' \in V W'_{ct} + B'} \quad (4)$$

Keterangan:

$P(T_n|c)$  = *Conditional Probability* kata ke-n dengan diketahui kelas c.

$W_{ct}$  = Nilai pembobotan TF-IDF atau W dari *term* t di kategori c.

$\sum W' \in V W'_{ct}$  = Jumlah total W dari keseluruhan *term* yang berada di kategori c.

$B'$  = Jumlah W kata unik (nilai IDF tidak dikali dengan TF) pada seluruh dokumen.

### 2.4. Naive Bayes Classifier

TF-IDF adalah salah satu metode yang bisa digunakan untuk pembobotan kata. Menurut Rahman dkk. [9], *Term Weighting* atau pembobotan kata bertujuan untuk memberikan bobot nilai pada setiap kata. Perhitungan bobot ini memerlukan dua hal, yaitu *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). *Term Frequency* merupakan banyaknya jumlah kata atau term tertentu yang ada dalam suatu dokumen. Sementara *Inverse Document Frequency* adalah frekuensi kemunculan kata atau term pada seluruh dokumen. Rumus perhitungan TF-IDF menurut Rahman, A. dkk. [9] dapat dilihat pada persamaan 5.

$$W_{dt} = tf_{dt} \times idf_t = tf_{dt} \times \log\left(\frac{N}{df_t}\right) \quad (5)$$

Keterangan:

- $W_{dt}$  = Bobot *term* ke *t* dokumen *d*.
- $tf_{dt}$  = Jumlah kemunculan *term* *t* pada dokumen *t*.
- $N$  = Jumlah dokumen secara keseluruhan.
- $idf_t$  = Jumlah dokumen yang mengandung *term* *t*.

## 2.5. Sastrawi

Sastrawi adalah *library* yang dikembangkan untuk melakukan proses pencarian kata dasar. Menurut Antonio Rachmat dan Yuan Lukito [3] kemampuan Sastrawi *stemming* tergolong baik dalam mengambil kata dasar dari setiap kata-kata yang menjadi data masukan. Instalasi Sastrawi menggunakan PHPComposer sehingga *library* Sastrawi dapat terunduh dan digunakan. Sastrawi juga dapat digunakan untuk proses *stopword removal*, yaitu menghilangkan kata *stopword*. Sastrawi sudah dilengkapi dengan daftar kata dalam bahasa Indonesia dan daftar kata *stopword* dalam bahasa Indonesia

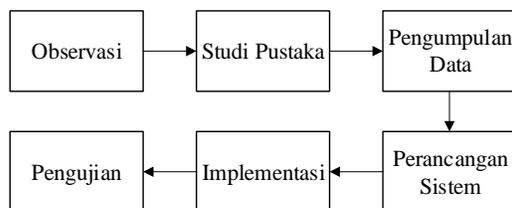
## 3. METODOLOGI PENELITIAN

### 3.1 Data Penelitian

Data yang digunakan pada penelitian ini adalah data berupa komentar tekstual. Komentar tekstual ini dikumpulkan dari media sosial Twitter. Pengumpulan data dilakukan dengan memanfaatkan Twitter API. Data yang dikumpulkan adalah *tweet* yang mengandung kata kunci 'iphone' dalam rentang tanggal 1 Juni 2019 hingga 21 Juli 2019. Data yang telah dikumpulkan akan dibagi menjadi dua jenis data yaitu data latih dan data uji yang akan digunakan untuk pengujian sistem.

### 3.2 Metode Penelitian

Metode penelitian adalah langkah-langkah yang dilakukan oleh peneliti dalam rangka untuk mengumpulkan informasi atau data serta melakukan investigasi pada data yang telah didapatkan tersebut. Metode penelitian yang dilakukan penulis dalam penelitian yang dilakukan untuk membangun sistem analisis dapat dilihat pada gambar 1.



Gambar 1 Diagram Proses Penelitian

### 3.2.1 Observasi

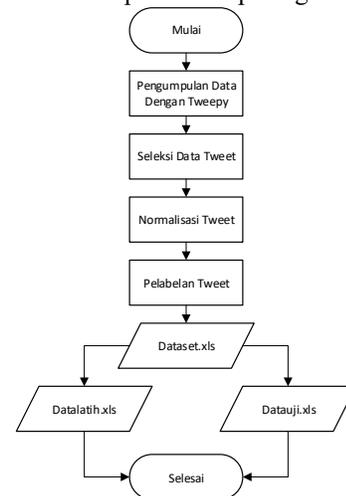
Observasi merupakan proses yang dilakukan untuk mengamati komentar tertulis pada twitter. Observasi dengan melakukan pengamatan langsung dengan mengakses sosial media twitter untuk mengetahui *tweet* yang ditulis oleh pengguna twitter. Observasi dilakukan untuk memberikan gambaran mengenai data yang akan digunakan untuk proses analisis sentimen.

### 3.2.2 Studi Pustaka

Studi pustaka adalah proses pengumpulan informasi yang dibutuhkan untuk melakukan penelitian. Studi pustaka dilakukan dengan melakukan pencarian bahan-bahan dan pengambilan informasi yang relevan. Sumber studi pustaka mengambil informasi dari jurnal, skripsi dan sumber lain dengan kasus mengenai analisis sentimen.

### 3.2.3 Pengumpulan Data

Proses pengumpulan data adalah proses yang dilakukan untuk mengumpulkan data yang akan digunakan sebagai data latih dan data uji. Proses pengumpulan data dilakukan dengan mengambil data *tweet* dari twitter dengan Twitter API. Langkah pengumpulan data dapat dilihat pada gambar 2.



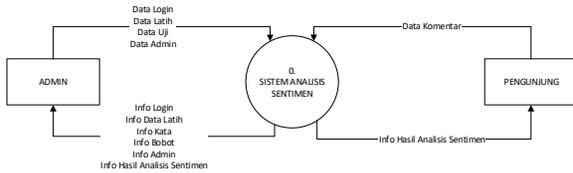
Gambar 2 Langkah Pengumpulan Data

Data yang diambil menggunakan *tweepy* kemudian diseleksi dan merapikan teks yang didapat. Setelah diseleksi dan dirapikan kemudian dilabeli apakah positif atau negatif. Dataset yang telah dikumpulkan dibagi menjadi data latih dan data uji yang disimpan dalam bentuk *.xls*.

### 3.2.4 Perancangan Sistem

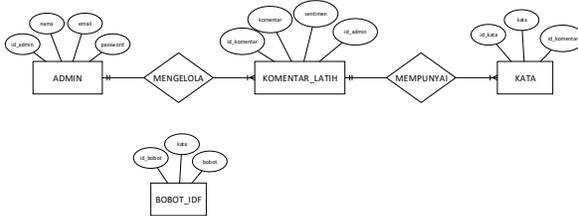
Perancangan dilakukan untuk mengetahui gambaran bagaimana sistem yang akan dibangun. Perancangan dilakukan dengan menggunakan Diagram Aliran Data (DAD), *Entity Relationship*

Diagram (ERD). Diagram Kontek dapat dilihat pada gambar 3 dan ERD bisa dilihat pada gambar 4.



Gambar 3 Diagram Konteks

Pada sistem yang dibangun, admin akan mampu melakukan login, memanipulasi data latih, melakukan pengujian terhadap data latih dan melakukan pelabelan secara otomatis sedangkan pengunjung hanya bisa melakukan analisis sentimen terhadap komentar. Hasil pengujian dan pelabelan akan disajikan dalam bentuk diagram pie.



Gambar 4 ERD

Admin bisa mengelola data latih yang akan disimpan pada tabel komentar\_latih. Setiap data latih yang berbentuk kalimat akan dipecah menjadi kata per kata yang kemudian disimpan pada tabel kata. Terdapat juga tabel bobot\_idf yang akan digunakan untuk menyimpan hasil perhitungan nilai bobot idf

### 3.2.5 Implementasi

Implementasi dilakukan untuk mengimplementasikan hasil perancangan ke dalam kode atau bahasa yang dimengerti oleh mesin komputer dengan Bahasa pemrograman tertentu. Dalam sistem ujian yang akan dibangun, sistem dibangun menggunakan menggunakan bahasa pemrograman PHP. Selain menggunakan bahasa pemrograman PHP, penyimpanan data untuk sistem ujian disimpan dalam basis data mySQL.

### 3.2.6 Pengujian

Sistem yang telah selesai dibangun kemudian dilakukan pengujian untuk mengetahui apakah sistem telah berjalan seperti yang diharapkan dan untuk mengetahui kekurangan sistem. Pengujian juga dilakukan untuk mengetahui akurasi yang dihasilkan oleh sistem. Pengujian dilakukan dengan mengolah data yang telah dikumpulkan menggunakan sistem yang telah dibuat. Data terdiri dari 200 data latih dan 100 data uji. Data latih terdiri dari 100 data berlabel positif dan 100 data berlabel negatif.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Hasil

Hasil sistem yang telah dibuat diuji dengan skenario untuk pembuktian kebenaran perhitungan yang dilakukan oleh sistem. Pembuktian dilakukan dengan membandingkan hasil perhitungan manual dengan hasil yang dihasilkan oleh perhitungan sistem. Pembuktian dilakukan dengan mengerjakan perhitungan dengan data latih yang bisa dilihat pada tabel 1 dan data uji yang bisa dilihat pada tabel 2.

Tabel 1 Tabel Data Latih

No	Komentar	Sentimen
1	@Bayu18 Iphone kameranya bagus	Positif
2	@zain_xx iphone mantap di kameranya	Positif
3	@Nanda_loween Boros Iphone buat Gaming	Negatif
4	@Todd_AAP Baterai iphone Boros	Negatif

Tabel 2 Tabel Data Uji

No	Komentar	Sentimen
1	Mantap iphone ini kameranya bagus	Positif

#### 4.1.1 Perhitungan Manual

Perhitungan manual dimulai dengan melakukan *preprocessing* terhadap data latih. *Preprocessing* dilakukan untuk mendapat kata dasar dan menghilangkan username, url, dan simbol. Hasil *Preprocessing* dapat dilihat pada tabel 3.

Tabel 3 Tabel Hasil Preprocessing Data Latih

No	Komentar	Hasil Preprocessing
1	@Bayu18 Iphone kameranya bagus	iphone kamera bagus
2	@zain_xx iphone mantap di kameranya	iphone mantap kamera
3	@Nanda_loween Boros Iphone buat Gaming	boros iphone buat gaming
4	@Todd_AAP Baterai iphone Boros	baterai iphone boros

Setelah proses *preprocessing* kemudian menghitung kata unik dan bobot IDF menggunakan rumus IDF. Hasil kata unik dan bobot IDF nya dapat dilihat pada tabel 4.

Tabel 4 Tabel Kata Unik dan Nilai IDF

No	Kata	D	df	IDF
1	iphone	4	4	$\text{Log}(4/4) = 0$
2	kamera	4	2	$\text{Log}(4/2) = 0,30103$
3	bagus	4	1	$\text{Log}(4/1) = 0,60206$
4	mantap	4	1	$\text{Log}(4/1) = 0,60206$
5	boros	4	2	$\text{Log}(4/2) = 0,30103$
6	buat	4	1	$\text{Log}(4/1) = 0,60206$
7	gaming	4	1	$\text{Log}(4/1) = 0,60206$
8	baterai	4	1	$\text{Log}(4/1) = 0,60206$
Total				3,612

Setelah mendapat kata unik dan bobotnya kemudian menentukan kata unik yang terdapat pada masing-masing kelas dan menghitung bobot tiap kata unik di masing-masing kelas (*bag of words*). *Bag of words* kelas sentimen positif dapat dilihat pada tabel 5 dan *bag of words* kelas sentimen negatif dapat dilihat pada tabel 6.

Tabel 5 Tabel Bag Of Words Kelas Sentimen Positif

No	Kata	Term	IDF	TF-IDF
1	iphone	2	0	$2 * 0 = 0$
2	kamera	2	0,30103	$2 * 30103 = 0,60206$
3	bagus	1	0,60206	$1 * 60206 = 0,60206$
4	mantap	1	0,60206	$1 * 60206 = 0,60206$
Total				1,806

Tabel 6 Tabel Bag Of Words Kelas Sentimen Negatif

No	Kata	Term	IDF	TF-IDF
1	iphone	2	0	$2 * 0 = 0$
2	boros	2	0,30103	$2 * 30103 = 0,60206$
3	buat	1	0,60206	$1 * 60206 = 0,60206$
4	gaming	1	0,60206	$1 * 60206 = 0,60206$
5	baterai	1	0,60206	$1 * 60206 = 0,60206$
Total				2,408

Setelah didapatkan daftar kata dan bobot TF-IDFnya kemudian dilakukan perhitungan naive bayes

pada data uji. Sebelum diuji data uji juga harus di *preprocessing* terlebih dahulu. Hasil dari *preprocessing* data uji dapat dilihat pada tabel 7.

Tabel 7 Tabel Hasil Preprocessing Data Uji

No	Komentar	Hasil Preprocessing
1	Mantap iphone ini kameranya bagus	mantap iphone kamera bagus

Setelah *dipreprocessing* data uji kemudian dihitung probabilitasnya terhadap kelas positif dan negatif dengan rumus naive bayes yang telah dijelaskan pada landasan teori.

Menghitung probabilitas komentar terhadap kelas positif dimulai dengan *menghitung prior probability* dengan cara sebagai berikut:  
 $P(\text{Positif}) = (2/4) = 0,5$

Setelah menghitung *prior probability* kemudian menghitung masing-masing kata yakni mantap, iphone, kamera, bagus dengan perhitungan conditional probability sebagai berikut:

$$P(\text{mantap} | \text{Positif}) = (0,60206 + 1) / (1,806 + 3,612) = 0,295692$$

$$P(\text{iphone} | \text{Positif}) = (0 + 1) / (1,806 + 3,612) = 0,18457$$

$$P(\text{kamera} | \text{Positif}) = (0,60206 + 1) / (1,806 + 3,612) = 0,295692$$

$$P(\text{bagus} | \text{Positif}) = (0,60206 + 1) / (1,806 + 3,612) = 0,295692$$

Langkah selanjutnya adalah menghitung *posterior probability* dengan perhitungan sebagai berikut:

$$P(\text{Positif} | \text{mantap iphone kamera bagus}) = 0,5 \times 0,295692 \times 0,18457 \times 0,295692 \times 0,295692 = 0,002385887$$

Menghitung probabilitas komentar terhadap kelas negatif dimulai dengan menghitung *prior probability* dengan cara sebagai berikut:  
 $P(\text{Negatif}) = (2/4) = 0,5$

Setelah menghitung *prior probability* kemudian menghitung masing-masing kata yakni mantap, iphone, kamera, bagus dengan perhitungan conditional probability sebagai berikut:

$$P(\text{mantap} | \text{Negatif}) = (0 + 1) / (2,408 + 3,612) = 0,166113$$

$$P(\text{iphone} | \text{Negatif}) = (0 + 1) / (2,408 + 3,612) = 0,166113$$

$$P(\text{kamera} | \text{Negatif}) = (0 + 1) / (2,408 + 3,612) = 0,166113$$

$$P(\text{bagus} | \text{Negatif}) = (0 + 1) / (2,408 + 3,612) = 0,166113$$

Langkah selanjutnya adalah menghitung *posterior probability* dengan perhitungan sebagai berikut:

$$P(\text{Negatif} \mid \text{mantap iphone kamera bagus}) = 0,5 \times 0,166113 \times 0,1660964 \times 0,1660964 \times 0,1660964 = 0,0003807014$$

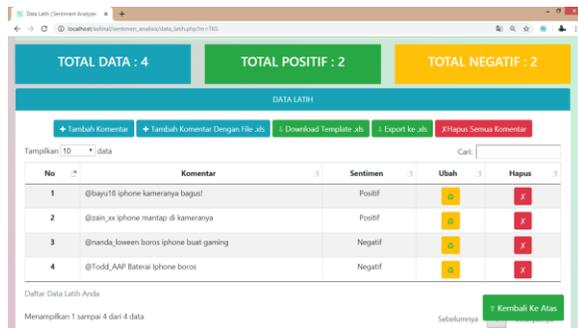
Probabilitas data uji terhadap masing-masing sentimen sudah diketahui, selanjutnya adalah mencari nilai tertinggi. Karena nilai posterior probability kelas positif (0,002385887) lebih besar daripada posterior probability kelas negatif (0,0003807014) maka data uji termasuk dalam kelas positif.

Perhitungan manual telah selesai dilakukan. Hasil dari perhitungan manual menunjukkan bahwa data uji bersentimen Positif.

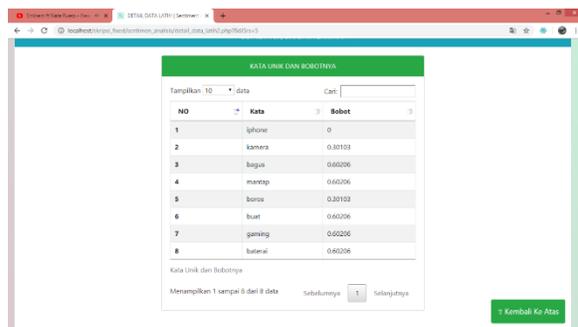
#### 4.1.2 Perhitungan Oleh Sistem

Perhitungan oleh sistem dilakukan dengan mengunggah data latih dan data uji yang sama dengan data latih dan data uji dalam perhitungan manual. Data latih dan data uji yang diunggah adalah berupa file yang berekstensi .xls.

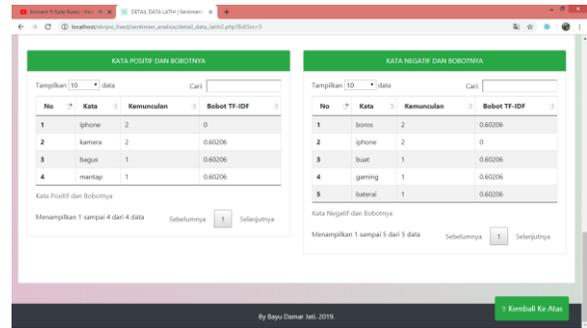
Pertama mengunggah data latih ke dalam sistem. Pengunggahan data latih dilakukan pada menu data latih. Setelah data latih terunggah maka akan ditampilkan data latih yang sukses diunggah seperti pada gambar 5 dan juga fitur dari data latih seperti pada gambar 6 dan gambar 7.



Gambar 5 Data Latih yang Diunggah

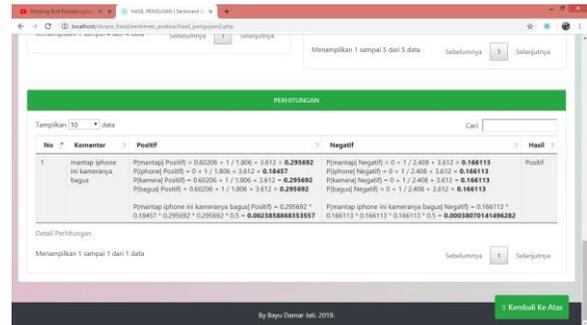


Gambar 6 Nilai IDF



Gambar 7 Kata dan Bobot pada Tiap Kelas

Setelah data latih sukses diunggah maka langkah selanjutnya adalah mengunggah data uji melalui form pengujian dan memilih data latih yang digunakan. Setelah data uji diproses maka akan muncul hasil perhitungan. Hasil perhitungan oleh sistem dapat dilihat pada gambar 8.



Gambar 7 Hasil Perhitungan oleh Sistem

Sistem yang telah melakukan perhitungan akan menampilkan detail perhitungan dan hasilnya. Perhitungan oleh sistem menghasilkan bahwa data uji bersentimen positif.

#### 4.1.3 Kesimpulan Perhitungan

Pembuktian perhitungan dilakukan dengan data latih dan data uji yang sama. Perhitungan yang dilakukan secara manual dan perhitungan yang dilakukan oleh sistem menghasilkan hasil yang sama. Hal ini membuktikan bahwa sistem telah berjalan sesuai dengan apa yang diharapkan yaitu menghasilkan hasil yang sama dengan perhitungan manual.

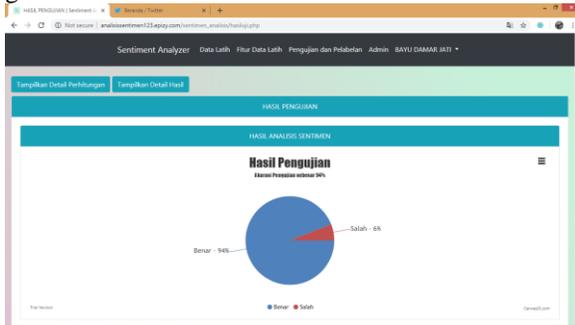
### 4.2 Pembahasan

#### 4.2.1 Pengujian Dengan Dataset Uji

Sistem analisis sentimen menggunakan naive bayes ini diuji dengan menggunakan 200 data latih dan 100 data uji. Data latih terdiri dari 100 data bersentimen positif dan 100 data bersentimen negatif. Kata unik yang dihasilkan dari 200 komentar adalah 739 kata unik, sedangkan komentar bersentimen

positif memiliki 411 kata unik dan komentar bersentimen negatif memiliki 479 kata unik.

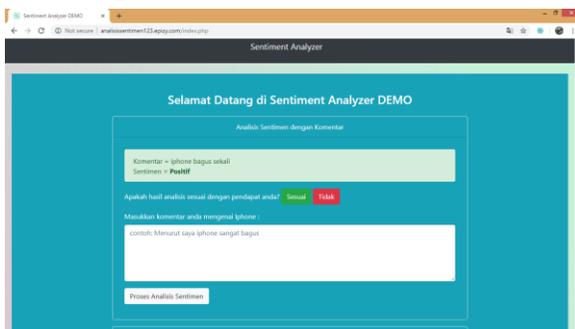
Akurasi yang dihasilkan menggunakan 200 data latih dan 100 data uji adalah 94%. Artinya sistem analisis sentimen mampu menganalisis 94 komentar dengan benar dari 100 komentar yang digunakan. Hasil pengujian dengan dataset uji dapat dilihat pada gambar 8.



Gambar 8 Hasil Pengujian Dengan Dataset Uji

#### 4.2.2 Pengujian Dengan Komentar Pengguna

Pengujian ini dilakukan dengan meminta beberapa pengguna untuk memasukkan komentar mengenai Iphone. Setelah hasil keluar, pengguna diminta pendapatnya apakah hasil analisis sentimen sesuai dengan dengan pendapatnya. Contoh pengujian oleh pengguna bisa dilihat pada gambar 9.



Gambar 8 Pengujian oleh pengguna

Pengujian ini mendapatkan beberapa komentar yang diuji oleh pengguna. Terdapat 12 komentar yang diuji. Hasil dari pengujian beragam, terdapat beberapa yang sesuai dan beberapa yang tidak sesuai. Komentar yang diuji dan hasilnya dapat dilihat pada tabel 7.

Tabel 8 Tabel Komentar yang Diuji oleh Pengguna

No	Komentar	Hasil	Kesesuaian
1	harga yang ditawarkan oleh iphone tidak seimbang dengan spesifikasi yang diberikan	Positif	Tidak Sesuai

Tabel 8 Lanjutan

No	Komentar	Hasil	Kesesuaian
2	Iphone overprice	Positif	Tidak Sesuai
3	Iphone tidak Bagus	Positif	Tidak Sesuai
4	Harganya sesuai dengan fitur yang ditawarkan	Positif	Sesuai
5	Menurut saya laptop Acer bagus sekali	Positif	Sesuai
6	Baterai cukup impresif	Negatif	Tidak Sesuai
7	Jelek karena tidak ada penyimpanan eksternal	Negatif	Sesuai
8	hardware wifinya iphone bagus ya	Positif	Sesuai
9	Tukang Review Ini Bilang Desain iPhone 8 Plus Sampah	Negatif	Sesuai
10	Kejadian Lagi, Baterai iPhone 8 Plus Bengkok Merusak Layar	Negatif	Sesuai
11	iPhone ko perlu di cas buat maen game, itu hp apa komputer rumah	Positif	Tidak Sesuai
12	Iphone hp kentang	Positif	Tidak Sesuai

Hasil yang didapatkan dari pengujian terhadap 12 komentar oleh pengguna adalah 6 hasil analisis sesuai dengan pendapat pengguna sedangkan 6 hasil analisis tidak sesuai dengan pendapat pengguna. Pada hasil yang tidak sesuai terdapat kata yang tidak terdapat pada fitur data latih seperti contohnya kata 'overprice' pada komentar 'iphone overprice', kata 'kentang' pada komentar 'iphone hp kentang' dan juga kata 'impresif' pada komentar 'baterai cukup impresif'. Pada komentar 'iphone tidak bagus' terdapat kata negasi yaitu 'tidak' yang diikuti oleh kata 'bagus' sehingga mengubah arti kalimat menjadi negatif, namun hasil dari sistem adalah positif. Kedua kesalahan yang terjadi karena tidak ditemukannya kata dalam fitur yang digunakan untuk perhitungan dan juga kata negasi yang mampu merubah arti ini kemungkinan bisa lebih diminimalisir lagi dengan cara menambahkan data latih yang lebih banyak dan bervariasi dan juga menambahkan algoritma deteksi kata negasi

## 5. PENUTUP

### 5.1. Kesimpulan

Berdasarkan keseluruhan proses analisis, perancangan, dan implementasi, pada pembuatan sistem analisis sentimen produk iphone menggunakan metode naive bayes, diperoleh kesimpulan sebagai berikut:

1. Sistem analisis sentimen produk Iphone menggunakan metode Naive Bayes mampu melakukan analisis sentimen terhadap data yang berupa komentar tekstual.
2. Tingkat keakuratan sistem analisis sentimen produk iphone menggunakan metode naive bayes yang diuji menggunakan data uji mendapatkan nilai 94%.
3. Masih terdapat beberapa kesalahan ketika diuji oleh beberapa pengguna karena kalimat yang lebih bervariasi.

### 5.2. Saran

Berdasarkan kesimpulan diatas, penulis mencantumkan beberapa saran antara lain:

1. Disarankan menambahkan algoritma yang mampu menangani kesalahan penulisan atau algoritma perbaikan kata tidak baku.
2. Disarankan untuk menambahkan data latih yang lebih banyak lagi dan bervariasi sehingga fitur kata yang dimiliki lebih banyak.
3. Peneliti selanjutnya disarankan menambahkan fitur deteksi negasi (convert negation).

## DAFTAR PUSTAKA

- [1] Agusvianto, H. (2017), *Sistem Informasi Inventori Gudang Untuk Mengontrol Persediaan Barang Pada Gudang Studi Kasus: PT. Alaisys Sidoarjo*, Journal Information Engineering and Educational Technology, 01(01), 40–46.
- [2] Buntoro, G.A. (2017), *Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter*, Integer Journal, 2(1), 32–41.
- [3] C, A.R. and Lukito, Y. (2016), *Klasifikasi Sentimen Komentar Politik dari Facebook Page Menggunakan Naive Bayes*, JUISI, 02(02), 26–34.
- [4] Fanissa, S., Fauzi, M.A. and 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(8), 2766–2770.
- [5] Ling, J., Kencana, I.P.E.N. and Oka, T.B. (2014), *Analisis Sentimen Menggunakan Metode Naive Bayes Classifier Dengan Seleksi Fitur Chi Square*, E-Jurnal Matematika, 3(3), 92–99.
- [6] Liu, B. (2012), *Sentiment Analysis and Opinion Mining*, Chicago: Morgan & Claypool Publisher.
- [7] Nurhuda, F., Sihwi, S.W. and Doewes, A. (2013), *Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier*, Jurnal ITSMART, 2(2), 35–42.
- [8] Purnama, C. (2016), *Sistem Informasi Manajemen*, Mojokerto: Insan Global.
- [9] Rahman, A., Wiranto and Doewes, A. (2017), *Online News Classification Using Multinomial Naive Bayes*, ITSMART: Jurnal Ilmiah Teknologi Dan Informasi, 6(1), 32–38.
- [10] Rosi, F., Fauzi, M.A. and Perdana, R.S. (2017), *Prediksi Rating Pada Review Produk Kecantikan Menggunakan Metode Naive Bayes dan Categorical Proportional Difference (CPD)*, Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(5), 1991–1997.
- [11] Sela, E.I. (2015), *Penentuan Indikator Yang Berperan Pada Identifikasi Kemiskinan Menggunakan Data Mining*, Jurnal Riset Daerah Edisi Khusus, 16–32.
- [12] Supriyanto, A., (2018). *Analisis Sentimen Media Sosial Tokoh Politik Menggunakan Metode Naive Bayes*, Skripsi, S. Kom., Teknik Informatika, Universitas Teknologi Yogyakarta, Yogyakarta