



# **Klasifikasi Text Mining Untuk Analisa Keluhan Masyarakat Terhadap Pembelian Produk Kosmetik Online Dengan Menggunakan Metode Naïve Bayes**

**Lovina Pandwinata**

Gram Studi Teknik Informatika, ProFakultas Ilmu Komputer Dan Teknologi Informasi, Universitas Budi Darma,  
Jalan Sisingamangaraja No.338, Medan, Sumatera Utara, Indonesia  
Email: [lovinapandwinata@gmail.com](mailto:lovinapandwinata@gmail.com)

**Abstrak**—Keluhan masyarakat dalam pembelian produk kosmetik online dapat ditemukan pada forum internet ataupun sosial media yang sekarang tabu untuk di lihat dan di dengar pada kalangan masyarakat sekitar. Kalimat-kalimat yang sering di dengar dan dilihat seperti keluhan positif maupun negative, akan tetapi pengguna harus membaca satu persatu kalimat tersebut. Banyak nya keluhan pada pembelian suatu produk online saat ini membutuhkan pengklasifikasian sesuai pada keluhan yang dimiliki agar mempermudah mendapatkan keluhan positif maupun negative dalam pembelian produk kosmetik online, maka dari itu untuk mempermudah penulis membuat sistem analisa keluhan masyarakat sangat berguna yang akan sangat berguna pada kalangan pembeli produk kosmetik online. Keluhan masyarakat dalam pembelian produk online ini dalam penyelesaiannya menggunakan metode naïve bayes yang berdasarkan hasil dari pengklasifikasian dari keluhan negative maupun positif dalam proses seperti . text preprocessing, Case Folding, Cleaning Tokenizing, Filtering dan Pembobotan TF-IDF, data uji merupakan kata sifat yang sudah diberi label manual.

**Kata Kunci:** Klasifikasi; Keluhan Masyarakat; Produk Online; Algoritma naïve bayes

**Abstract**—Public complaints in purchasing cosmetic products online can be found on internet or social media forums which are now taboo to be seen and heard by the surrounding community. Sentences that are often heard and seen are positive or negative complaints, but the user must read the sentences one by one. The many complaints when purchasing an online product currently require classification according to the complaint they have in order to make it easier to get positive and negative complaints in purchasing cosmetic products online, therefore to make it easier for the author to make a very useful community complaint analysis system that will be very useful for buyers online cosmetic products. Public complaints in purchasing online products are resolved using the naïve Bayes method which is based on the results of the classification of negative and positive complaints in processes such as . text preprocessing, Case Folding, Cleaning Tokenizing, Filtering and TF-IDF Weighting, test data are adjectives that have been labeled manually.

**Keywords:** Classification; Community Complaints; Online Products; Naive Bayes Algorithm

## **1. PENDAHULUAN**

Di zaman sekarang yang serba teknologi, informasi menjadi kebutuhan yang sangat penting. Akibat adanya perkembangan informasi dituntut adanya media penyedia informasi yang dapat diakses kapanpun dengan mudah. Saat ini juga produk kosmetik sudah menjadi kebutuhan utama dari kaum wanita yang merupakan target utama dari industri kosmetik dan semakin banyak produk kosmetik. Semakin banyak merek kosmetik yang bermunculan baik dari dalam negeri maupun luar negeri membuat konsumen memiliki banyak pilihan. Program jual kosmetik online merupakan salah satu langkah yang sudah marak diaplikasikan sehingga masyarakat dapat memperoleh barang dengan instan. dalam bentuk tampilan yang dibuat semenarik mungkin yang terkadang dapat menimbulkan banyak tanggapan yang positif maupun negatif. Banyak website yang menyediakan informasi tentang produk kosmetik dengan memberikan banyak informasi berupa gambar dan review pengguna. Membaca semua review yang ada pada website tentu akan banyak memakan waktu, karna terlalu banyak opini yang ada dari berbagai sumber situs web yang berbeda. Disini juga text mining merupakan penerapan konsep dan teknik untuk mencari pola dalam teks yaitu proses penganalisaan teks berdasarkan informasi yang bermanfaat untuk dapat mengelompokkan keluhan tentang tanggapan yang positif maupun negatif. Oleh sebab itu analisa keluhan ini merupakan salah satu solusi penyelesaian masalah bertujuan untuk dapat mengelompokkan opini atau ulasan menjadi opini positif atau negatif secara otomatis dengan melakukan klasifikasi menggunakan naïve bayes yang secara cepat, sederhana serta memiliki akurasi tinggi dalam mengetahui selisih klasifikasi kelas positif dan negatif secara otomatis naïve bayes memiliki kelebihan nya yaitu sederhana[1].

## **2. METODOLOGI PENELITIAN**

### **2.1 Tahapan Penelitian**

Untuk mendukung kelancaran penelitian ini, maka dilakukan tahapan penelitian sebagai berikut:

- a. Studi Literatur  
Pada tahap ini dilakukan pengumpulan referensi yang diperlukan dalam penelitian. Hal ini dilakukan untuk memperoleh informasi dan data yang diperlukan untuk penelitian ini. Referensi yang digunakan dapat berupa buku, jurnal, artikel, paper, makalah baik berupa media cetak maupun media internet mengenai topik penelitian .
- b. Analisis Sistem  
Pada tahap ini akan dianalisis sistem yang akan dibuat, batasan sistem, kinerja sistem dan cara kerja sistem. Sehingga sistem dapat menerapkan Metode Naïve Bayes.



- c. Perancangan Sistem  
Merancang *input*, *output*, struktur file, program, prosedur, perangkat keras dan perangkat lunak yang diperlukan untuk mendukung sistem informasi
- d. Implementasi Sistem  
Sistem diimplementasikan dengan menggunakan Metode Naïve Bayes.
- e. Pengujian Sistem  
Pada tahap ini dilakukan pengujian kinerja sistem dan kebenaran hasil keamanan yang dilakukan dengan Metode Naïve Bayes
- f. Dokumentasi Sistem  
Merupakan langkah yang terakhir dilakukan, melakukan validasi dan hasil akhir (*output*) yang diperoleh dari sistem yang dirancang setelah melakukan evaluasi ketepatan maupun kecepatan terhadap kinerja sistem untuk membuat kesimpulan dari topik yang dikaji.

## 2.2 Text Mining

Teks mining adalah salah satu bidang khusus dalam data mining yang memiliki defenisi menambang data berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Teks mining dapat menganalisa dokumen, mengelompokkan dokumen berdasarkan kata-kata yang terkandung di dalamnya, serta menentukan kesamaan di antara dokumen untuk dapat mengetahui teks berhubungan dengan variabel lainnya, penerapan yang paling umum dilakukan teks mining misalnya seperti, analisa sentiment, mengukur preferensi pelanggan, meringkas dokumen, pengelompokan topik penelitian, dan banyak lainnya[2]

Teks mining juga memiliki bebera tipe diantaranya antara lain :

- a. *Search And Information Retrieval*  
Menyimpan dan menemukan kembali dokumen teks, termasuk pencarian dan kata kunci pencarian.
- b. *Document Clustering*  
Pengelompokan dan pengkategorian istilah, potongan, paragraph, atau dokumen menggunakan metode mining.
- c. *Web Mining*  
Data dan teks mining pada internet yang focus pada skala dan antar hubungan pada website.
- d. *Information Extraction*  
Identifikasi dan ekstraksi fakta yang relevan.
- e. *Natural Language Processing*  
Pemrosesan bahasa tingkat rendah yang biasanya digunakan untuk bahasa komputasi.

## 2.3 Metode Naïve Bayes

Naïve Bayes Classifier merupakan pengklasifikasi porbalistik sederhana yang didasarkan pada teorema Bayes, yang menyatakan bahwa kemungkinan terjadinya suatu peristiwa sama dengan porbaliats intristik terhitung dari data data yang tersedia, pengklasifikasian naïve bayes memiliki asumsi bahwa efek dari suatu nilai atribut tertentu tidak bergantung (independen) terhadap nilai atribut lainnya, pengklasifikasian ini dapat sangat efisien dan akurat, terutama ketika jumlah variabel yang tinggi, Metode pengkalisifikasian dengan menggunakan metode probabilitas dan statistik yang di kemukakkan oleh ilmuwan inggris Thomas bayes, yaitu memperediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga di kenal sebagai teorema Bayes. Ciri utama dari naïve Bayes classifier ini adalah asumsin yang sangat kuat (naïf) akan independensi dari masing masing kondisi / kejadian. Naïve Bayes Classifier dinilai berkerja sangat baik dibanding dengan model Classifier lainnya,yaitu Naïve Bayes Classifier memiliki tingkat akurat yang lebih baik di dibandingkan model Classifier lainnya. Naïve Bayes didasarkan bahwa nilai atribut seera kondisional saling bebas jika diberikan nilai output. Dengan kata lain diberikan nilai output porbalitas mengamati secara bersama adalah produk porbalitas individu sebelum menjelaskan Naive Bayes Classifier ini, akan dijelaskan terlebih dahulu Teorema Bayes yang menjadi dasar dari metode tersebut sebagai berikut. Perhitungan nilai probabilitas tersebut menggunakan persamaan [3]

$$(vj) = \frac{|docs\ j|}{|contoh|} \tag{1}$$

Keterangan:

(V j) : Peluang kemunculan dokumen yang memiliki kategori j.

|docsj| : Jumlah dokumen setiap kategori j.

|contoh| : Jumlah dokumen dari semua kategori.

Dan terakhir melakukan perhitungan probabilitas setiap kata pada data uji terhadap data uji pada setiap kelas j, dengan cara:

$$(xi|vj) = \frac{nK+1}{n+|kosakats|} \tag{2}$$

Keterangan:

P(Xi|V j) : Peluang kemunculan Xi pada kategori V j.

nk : Jumlah frekuensi kemunculan setiap kata.



$n$  : Jumlah frekuensi kemunculan kata dari setiap kategori.

$|kosakata|$  : Jumlah semua kata dari semua kategori.

Persamaan diatas dapat disederhanakan yaitu:

$$vmap = \operatorname{argmax}_{Vj \in v} p(Vj) \prod ip(ai|vj) \tag{3}$$

Keterangan:

$vj$  : Kategori komentar,  $j = 1, 2, 3, \dots, n$ .

$(ai|vj)$  : Peluang kemunculan  $ai$  pada kategori  $vj$

$(Vj)$  : Peluang kemunculan dokumen yang memiliki kategori  $j$ .

Berikut algoritma Naïve Bayes Classifier

a. Learning (pembelajaran)

Algoritma Naïve Bayes termasuk kedalam supervised learning, maka akan dibutuhkan pengetahuan awal untuk mengambil keputusan. Berikut langkah-lahkah Naive Bayes yaitu:

1. Langkah pertama betuk vocabulary pada setiap dokumen data training
2. Langkah dua hitung probabilitas pada setiap kategori ( $vj$ )
3. Langkah tiga tentukan frekuensi setiap kata  $wk$  pada setiap kategori  $P(wk|vj)$

b. Pengklasifikasian

1. Langkah pertama hitung  $(Vj) \prod P(ai \vee vj)$  untuk setiap kategori
2. Langkah dua tentukan kategori dengan nilai  $(Vj) \prod P(ai \vee vj)$  maksimal.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pembahasan

Pembelian produk online di zaman sekarang semakin meningkat menjuru keseluruhan dunia seperti halnya dengan produk kosmetik atau pun barang dan produk lainnya dengan adanya penjualan online semua akan menjadi serba praktis, situasi tersebut dimanfaatkan oleh penyedia layanan untuk bias mengembangkan bisnisnya, berbagai inovasi dilakukan oleh penyedia barang maupun jasa, namun dibalik kemudahan dan kecepatan belanja online itu semua tetap saja para masyarakat atau pun pembeli produk online masih terdapat keluhan-keluhan seperti dampak positif maupun negatif oleh para kalangan pembeli produk online.

Pada umumnya terkandung susah mengetahui keluhan-keluhan yang mana saja yang berdampak positif maupun negatif yang dikeluhkan oleh kalangan masyarakat karena terlaunya banyak tanggapan-tanggapan tersebut, untuk dapat mengetahui dari tanggapan-tanggapan tersebut perlu dilakukan pengelompokan-pengelompokan tentang tanggapan-tanggapan positif atau pun negatif tersebut.

Maka dari itu untuk menghindari dari kesulitan dari tanggapan masyarakat tersebut perlu adanya pengklasifikasian yang mempermudah untuk mengetahui tanggapan positif dan negatif yang dapat membantu mengetahui banyak informasi yang dikeluhkan pada kalangan masyarakat atau pun para pengguna jasa produk online.

Cotah kasus pada penei ini, menggunakan seperti akun media instagram untuk menentukan klasifikasi dan melihat komenan positif dan negative pada para pembeli secara online, Instagram ialah media sosial yang digunakan pengguna untuk mengambil foto atau video, menerapkan filter efek pada foto dan membagikannya ke berbagai media sosial termasuk Instagram itu sendiri.

Adapun prose *text processing* merupakan sebuah tahapan dalam membersihkan data yang digunakan dalam penelitian. Adapun tahap yang terdapat pada processing yaitu, sebagai berikut:

#### 1. Case Folding

*Case folding* adalah proses mengubah semua huruf dalam bentuk dokumen semua huruf menjadi huruf kecil. Berikut contoh data komentar yang akan dilakukan proses *case folding*.

**Tabel 2.** Komentar sebelum dan sesudah di *Case Folding*

| Instagram         | Komentar Sebelum Case Folding   | Komentar Sesudah Case Folding   | Kelas   | Komentar      |
|-------------------|---|---|---------|---------------|
| @auliyaskin_care  | Dr kalo pengen langsung konsul langung di klinik bisa?  | dr kalo pengen langsung konsul langung ke klinik bisa   | positif | @yannifebrian |
| @aprilia_kosmetik | Keadaan produk kosmetik yang di Instagram, membuat saya tidak percaya keaslian karena Foto produk tersebut sangat mencurigaka | keadaan produk kosmetik yang di instagram, membuat saya tidak percaya keaslian karena Foto produk tersebut sangat mencurigaka | negatif | @joegusm      |
| @beautygirls.bjb  | Peyampaian Produk kosmetik di Instagram, membuat saya yakin untuk melakukan pembelian.  | peyampaian Produk kosmetik di instagram, membuat saya yakin untuk melakukan pembelian.  | positif | @arsyilla_12  |



2. Cleaning

Adapun kata atau karakter yang akan dihilangkan pada data komentar adalah karakter atau simbol (@#%\$%^&\*()\_+";:{}<>.,?!~/[]), angka, link url (http://link.com), hashtag (#), dan mention (@username . contoh data komentar sebelum dan sesudah di Cleaning.

Tabel 3. Komentar sebelum dan sesudah di Cleaning

Table with 5 columns: Instagram, Komentar Sebelum Case Folding, Komentar Sesudah Case Folding, Kelas, and Komentar. It lists three examples of comments before and after cleaning.

3. Tokenizing

Proses memisahkan teks menjadi kata, dikenal sebagai token. Tokenisasi dilakukan berdasarkan pemisah yaitu karakter angka dan huruf dengan tanda baca dan spasi.

Tabel 4. Data komentar yang sesudah di tokenizing

Table with 3 columns: Data1, Data2, and Data3. It shows the individual tokens extracted from the comments in the previous table.

4. Filtering

Proses menghilangkan kata-kata yang dianggap tidak memiliki makna (stopwords).

Tabel 5. Data komentar sebelum dan sesudah di Filtering

Table with 6 columns: Data yang sebelum di filtering (Data1, Data2, Data3) and Data yang sudah di filtering (Data1, Data2, Data3). It shows the removal of stopwords from the tokenized data.



karena  
foto  
produk  
tersebut  
sangat  
mencurigakan

**5. Pembobotan TF-IDF**

Pembobotan adalah proses mengubah kata menjadi bentuk angka atau vector. TF-IDF merupakan untuk menentukan nilai frekuensi sebuah kata di dalam sebuah dokumen dan pemberian bobot pada setiap kata pada setiap dokumen untuk mencari dan menghitung berapa kali kata itu muncul. Berikut langkah-langkah cara kerja pembobotan TF-IDF.

**Tabel 6.** Pembobotan dengan TF-IDF

| Kata         | Data1 (TF) | Data2 (TF) | Data3 (TF) | DF | IDF                         | TF-IDF      |             |             |
|--------------|------------|------------|------------|----|-----------------------------|-------------|-------------|-------------|
|              |            |            |            |    | $Idf = \log(\frac{Td}{Df})$ | Data1       | Data2       | Data3       |
| Dr           | 1          | 0          | 0          | 1  | $\log(3/1) = 0,477121255$   | 0,477121255 | 0           | 0           |
| Kalo         | 1          | 0          | 0          | 1  | 0,477121255                 | 0,477121255 | 0           | 0           |
| Pengen       | 1          | 0          | 0          | 1  | 0,477121255                 | 0,477121255 | 0           | 0           |
| Langsung     | 1          | 0          | 0          | 1  | 0,477121255                 | 0,477121255 | 0           | 0           |
| Konsul       | 1          | 0          | 0          | 1  | 0,477121255                 | 0,477121255 | 0           | 0           |
| di           | 1          | 0          | 1          | 1  | 0,477121255                 | 0,477121255 | 0           | 0,477121255 |
| Klinik       | 1          | 0          | 0          | 1  | 0,477121255                 | 0,477121255 | 0           | 0           |
| Keadaan      | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Produk       | 0          | 1          | 1          | 1  | 0,477121255                 | 0           | 0,477121255 | 0,477121255 |
| Kosmetik     | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Instagram    | 0          | 1          | 1          | 1  | 0,477121255                 | 0           | 0,477121255 | 0,477121255 |
| Tidak        | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Percaya      | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Keaslian     | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Foto         | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| Produk       | 0          | 1          | 1          | 1  | 0,477121255                 | 0           | 0,477121255 | 0,477121255 |
| Mencurigakan | 0          | 1          | 0          | 1  | 0,477121255                 | 0           | 0,477121255 | 0           |
| penyampaian  | 0          | 0          | 1          | 1  | 0,477121255                 | 0           | 0           | 0,477121255 |
| Produk       | 0          | 1          | 1          | 1  | 0,477121255                 | 0           | 0,477121255 | 0,477121255 |
| Di           | 0          | 0          | 1          | 1  | 0,477121255                 | 0           | 0           | 0,477121255 |
| Instagram    | 0          | 1          | 1          | 1  | 0,477121255                 | 0           | 0,477121255 | 0,477121255 |
| Pembelian    | 0          | 0          | 1          | 1  | 0,477121255                 | 0           | 0           | 0,477121255 |

**3.2 Proses Klasifikasi Naïve Bayes Classifier**

Pada proses ini menjelaskan proses-proses yang akan dilakun dalam klasifikasi dan berikut merupakan penjelasan langkah-langkah perhitungn pembobotan TF-IDF dan proses klasifikasi *Naïve Bayes Classifier*.

a. Tahap *training*

Pada tahap ini data yang sudah didapatkan bobotnya seperti dijadikan pada data latih untuk menjadi acuan untuk membentuk model klasifikasi,selanjutnya akan mencari nilai porbalitas kategori dan porbalitas masing-masing kata dari kata latih. Menghitung kata menggunakan persamaan *trem* untuk setiap kelas dari data latih, perhitunganya yaitu

1. Pertama hitung porbalitas setiap kategori (*prior*) dengan menggunakan dua kelas yaitu kelas positif dan kelas negative

$$P(pos/neg) = \frac{d(pos/neg)}{|c|}$$

$$P(pos/neg) = \frac{d(pos/neg)}{|c|} = \frac{1}{3}$$

$P(pos/neg)$  = jumlah seluruh dokumen pada setiap kelas  $|c|$  = jumlah keseluruhan data yang dilatih.

2. Selanjutnya hitung porbalitas setiap *trem* dari dari semua dokumen jumlah seluruh pada data latih yang digunakan dalam perhitungan adalah sebanyak 13,7 dari kelas positif 4, proses processing dan stemming mempengaruhi banyak term. Berikut perhitungan porbalitas yang berdasarkan dari

$$P(w_k|pos/neg) = \frac{nk, pos/neg + 1}{n, pos/neg + |kosakata|}$$



$p(wk|pos/neg/net)$  = Peluang kemunculan kata pada kategori  $wk$  = Kata yang muncul pada sebuah kategori.  $(nk, pos/neg/net) + 1$  = Jumlah frekuensi kemunculan kata pada kategori.  $nk$  = Kemunculan setiap kata pada kategori.  $|kosakata|$  = Jumlah semua kata dari semua kategori. Diketahui :

Total nilai TF  $|kosakata| = 13$

Total nilai TF Positif = 7

Total nilai TF Negatif = 4

**Porbalitas kata "dr" berdasarkan table 6.**

$$P("dr"|pos) = \frac{"dr"|pos + 1}{(pos) + |kosakata|} = \frac{0,477121255 + 1}{7 + 13} = 0,0738560628$$

$$P("dr"|neg) = \frac{"dr"|neg + 1}{(neg) + |kosakata|} = \frac{0 + 1}{4 + 13} = 0,588235294$$

- b. Tahap Testing Tahap ini merupakan tahapan setelah proses training yang telah dilakukan diatas. Pada tahapan ini dilakukan pengujian dengan cara menggunakan data uji kedalam model yang telah dibentuk pada tahapan training diatas.

**Tabel 7.** Data setelah proses *Processing*

| No | Komentar   | Kelas   |
|----|--|---------|
| 1  | Keadan produk kosmetik yang di instagram membuat saya tidak percaya keasliannya karna foto produk tersebut sangat mencurigakan | Negatif |

Tahapan selanjutnya, akan melakukan probabilitas masing-masing kata didata uji dengan menggunakan hasil probabilitas term yang didapat pada data latih.

**Tabel 8.** Data latih dan data uji telah melalui tahapan *Preprocessing*

| Data latih | Komentar   | Kelas   |
|------------|--|---------|
| 1          | Dr kalo pengen langsung kosnul langsung di klinik bisa   | positif |
| 2          | Keadan produk kosmetik yang di instagram membuat saya tidak percaya keasliannya karna foto produk tersebut sangat mencurigakan | negatif |
| 3          | Penyampaian produk kosmetik di instagram membuat saya yakin untuk melakukan pembelian  | negatif |
| Data uji   |  |         |
| 4          | Keadan produk kosmetik yang di instagram membuat saya tidak percaya keasliannya karna foto produk tersebut sangat mencurigakan | ?       |

Selanjutnya dilakukan perhitungan probabilitas masing-masing kata diproses dengan menggunakan persamaan

$$P(w_k|pos/neg) = \frac{nk, pos/neg + 1}{n, pos/neg + |kosakata|}$$

**Tabel 9.** Pembobotan data uji

| No | Kata      | TF | DF | IDF         | TF-IDF      |
|----|-----------|----|----|-------------|-------------|
| 1  | produk    | 1  | 1  | 0,477121255 | 0,477121255 |
| 2  | kosmetik  | 1  | 1  | 0,477121255 | 0,477121255 |
| 3  | instagram | 1  | 1  | 0,477121255 | 0,477121255 |

Diketahui :

Total nilai TF  $|kosakata| = 13$

Total nilai TF Positif = 7

Total nilai TF Negatif = 4

**Probabilitas Data uji (Positif) berdasarkan tabel 5**

$$P("produk"|pos) = (0 + 1)/(7 + 13) = 0,05$$

$$P("kosmetik"|pos) = (0 + 1)/(7 + 13) = 0,05$$

$$P("instagram"|pos) = (0 + 1)/(7 + 13) = 0,05$$

**Probabilitas Data uji (Negatif) berdasarkan tabel 3.5**

$$P("produk"|neg) = (0,477121255 + 1)/(4 + 13) = 0,086889468$$

$$P("kosmetik"|neg) = (0,477121255 + 1)/(4 + 13) = 0,086889468$$

$$P("instagram"|neg) = (0,477121255 + 1)/(4 + 13) = 0,086889468$$

Tahapan selanjutnya menghitung probabilitas kategori dengan menggunakan persamaan

$$(pos/neg/d) = p(pos/neg/net) * \prod P(a_i|pos/neg/net) i$$

Keterangan :

$(pos/neg/net)$  = Peluang kemunculan kata pada kategori atau kelas dengan dokumen data uji.  $d$  = Dokumen data baru

$\prod (a_i|pos/neg/) i$  = Peluang kemunculan  $a_i$  pada kategori atau kelas  $a_i$  = Kata baru yang akan diuji.

$$P(uji|pos) = p(pos) \times p(produk|pos) \times p(kosmetik|pos) \times p(instagram|pos) (uji|pos)$$

$$= 0,333 \times 0,05 \times 0,05 \times 0,05 = \mathbf{0,00004162}$$





$$P(\text{uji}|\text{neg}) = p(\text{neg}) \times p(\text{produk}|\text{neg}) \times p(\text{kosmetik}|\text{neg}) \times p(\text{instagram}|\text{neg}) \times p(\text{uji}|\text{neg}) \\ = 0,333 \times 0,086889468 \times 0,086889468 \times 0,086889468 = \mathbf{0,0002184468}$$

Nilai probabilitas tertinggi adalah pada kategori negatif dengan nilai sebesar **0**., sehingga dapat diklasifikasikan kedalam kelas Negatif.

#### **4. KESIMPULAN**

Kesimpulan yang dapat diambil setelah melakukan hasil dan pembahasan dimana klasifikasi teks mining untuk analisa keluhan masyarakat dalam pembelian produk kosmetik online ini untuk bias mempermudah dan mengetahui tanggapan tanggapan positif maupun negative tentang pembelian produk kosmetik online dengan cara otomatis dengan melakukan klasifikasi menggunakan metode naïve bayes dimana hasil klasifikasi keluhan masyarakat masuk ke kategori positif .

#### **REFERENCES**

- [1] R. & Feldman, The Text Mining Handbook: Advanced Approaches in. Analyzing. 2007.
- [2] S. Natulis, Metode Naive Bayes Classifier Dan Penggunaanya Pada Klasifikasi Dokumen. Bandung: Institut Teknologi Bandung, 2010.
- [3] N. Afriani, Analisa Penanganan Kelusan Pelanggan Oleh Instalasi Pemasaran Dan Humas Rumah Sakit Umum Pusat fatmawati. UI Depok, 2012.
- [4] M. G. Ross, Community Organization: Theory, Principlesmand Praction. Second Edition,. NewYork: Harper & Row, 1967.
- [5] G. Kolter, Philip dan Armstrong, Prinsip-prinsip pemasaran. Jakarta: Erlangga, 2008.
- [6] BPOM., Peraturan Kepala Badan Pengawas Obat Dan Makanan Republik Indonesia Tentang Metode Analisa Kosmetik. Jakarta: Sekerta Negara, 2011.