

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pada saat ini penggunaan skin care oleh kaum wanita menjadi suatu kewajiban di era modern zaman sekarang. Bahkan produk-produk perawatan kulit digunakan untuk kebutuhan kecantikan wanita maupun pria dari kalangan muda sampai tua. Berbagai macam produk mempunyai kegunaannya masing-masing, mulai dari untuk mengurangi jerawat, mempercerah kulit, dan menjaga kesehatan kulit. Namun sebelum melakukan perawatan kulit atau menggunakan skin care sebaiknya diketahui terlebih dahulu jenis kulit yang dimiliki. Sebab jika menggunakan produk yang salah dan tidak sesuai dengan jenis kulit maka hal itu dapat memperburuk masalah kulit yang dialami. Selain memperhatikan jenis kulit, perlu juga mengetahui fungsi produk dan urutan pemakaian secara umum, serta memperhatikan bahan aktif yang tidak boleh digunakan secara bersamaan.

Penelitian Billy Gunawan, Helen Sasty Pratiwi, dan Enda Esyudha Pratama Sabagian besar konsumen mengambil saran, ulasan ataupun pendapat dari opini orang lain sebelum mereka memutuskan untuk membeli barang tertentu. Dengan adanya user (aktif) dalam komunikasi online, sentiment membawa pengaruh besar terhadap perusahaan yang ingin mengetahui umpan balik (feedback) dari masyarakat terhadap merk produk yang mereka jual, sehingga munculnya data besar (big data) yang merupakan himpunan data dalam jumlah yang sangat besar, rumit, tidak terstruktur menjadi salah satu sumber daya yang dapat di olah menjadi data yang akurat. Review dan dan opini sangat banyak akan menyulitkan pembaca untuk mengelompokan review dan opini yang sesuai kelasnya, kelas sentiment di bagi menjadi kelas positif, negative, dan netral sehingga pengguna dapat membaca sesuai opini yang diinginkan [1].

Penelitian Wahdatun Novi Yanti membuat system analisis sentiment menggunakan metode Multinomial Naïve Bayes. Komentar dibagi menjadi

dua sentiment positif dan negative. Hasil penelitian yang dilakukan Multinomial Naïve Bayes berhasil sebagai salah satu klasifikasi yang ditandai dengan nilai yang cukup untuk semua kelas. Terutama mampu menyelesaikan masalah analisis sentiment [2].

Dengan demikian, dari ulasan – ulasan tersebut dapat dibangun analisis sentiment dengan menggunakan algoritma klasifikasi Multinomial Naïve Bayes. Metode Multinomial Naïve Bayes merupakan pengembangan dari Naïve Bayes, Algoritma Naïve Bayes adalah metode yang menggunakan prinsip probability untuk membuat model prediksi klasifikasi. Dengan memanfaatkan data tentang kejadian dimasa lalu. Keunggulan Naïve Bayes adalah sifatnya yang efektif dan cepat untuk mengolah data berjumlah besar. Berdasarkan uraian diatas maka peneliti diharapkan membuat sebuah system “Analisis Sentiment terhadap ulasan produk skin care menggunakan metode Multinomial Naïve Bayes” untuk klasifikasi review dari opini masyarakat terhadap penggunaan skin care kedalam dua sentiment yaitu positif dan negatif.

### **1.2. Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah diuraikan sebelumnya, maka dirumuskan permasalahan penelitian yaitu berapa tingkat akurasi yang didapatkan dari hasil klasifikasi Multinomial Naïve Bayes terhadap ulasan produk skin care.

### **1.3. Batasan Masalah**

Beberapa batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Mengolah data ulasan yang berhubungan dengan skin care.
2. Bahasa yang diambil dari Twitter adalah Bahasa Indonesia.
3. Analisis sentimen dalam penelitian ini menggunakan data tweet dari Twitter.
4. Analisis sentimen menggunakan metode MULTINOMIAL NAIVE BAYES.
5. Dari sentimen diklasifikasikan kedalam 2 kelas yaitu positif dan negatif.
6. Bahasa pemrograman yang digunakan adalah Python.

#### **1.4. Maksud dan Tujuan Penelitian**

Tujuan dari pembuatan Skripsi ini adalah:

1. Menerapkan metode Multinomial Naïve Bayes pada analisis sentimen terhadap data twitter ulasan produk skin care.
2. Membantu pembaca untuk menentukan opini yang sesuai dengan kelasnya.
3. Mencari nilai akurasi Algoritma Multinomial Naïve Bayes dalam mengklasifikasi sentimen terhadap tingkat kepuasan Produk Skin Care.

#### **1.5. Manfaat Penelitian**

Manfaat penelitian ini adalah sebagai berikut:

1. Memberikan aplikasi tentang kepuasan pelanggan dengan analisis sentiment.
2. Mengetahui cara memperoleh dan mengolah data untuk sentiment analisis.
3. Memperoleh hasil analisis terhadap ulasan produk skin care dengan multinomial Naïve Bayes.

#### **1.6. Metode Penelitian**

Dalam penelitian ini akan menjelaskan tahapan – tahapan sentimen analisis sebagai berikut:

##### **1.6.1 Metode Pengumpulan Data**

###### **1.6.1.1 Studi Pustaka**

Mencari literatur atau jurnal yang bersangkutan dengan Analisis Sentimen, Opinion Mining, Multinomial Naïve Bayes, TF-IDF.

###### **1.6.1.2 Eksperimen**

Menguji kode program dan mengimplementasi metode Multinomial naive bayes kedalam program. Membuat user interface yang mudah untuk dipahami pengguna serta dapat memberikan informasi kepada pengguna.

##### **1.6.2 Metode Analisis**

Dalam metode analisis ini terdapat beberapa tahapan seperti dibawah ini:

1. Collecting Data

Data yang dikumpulkan diperoleh dari crawling data twitter menggunakan Twitter API.

## 2. Data Preprocessing

Tahap preprocessing adalah tahap awal dalam analisis sentimen. Berikut tahap – tahapnya :

- a. Stopword Removal Menghilangkan term yang tidak memiliki arti atau tidak relevan.
- b. Cleaning Data merupakan tahapan penghapusan kata-kata yang tidak dipakai seperti hastag dan username.
- c. Case Folding atau Lower Case digunakan untuk merubah semua dokumen atau kalimat menjadi huruf kecil.
- d. Penyesuaian PUEBI merubah kata kekinian sesuai PUEBI (Pedoman Umum Ejaan Bahasa Indonesia).

## 3. Feature Weighting (Fitur Pembobotan)

- a. Term Frequency (TF) merupakan frekuensi kemunculan kata pada kalimat.
- b. Document Frequency (IDF) kemunculan berapa banyak kata dalam satu kalimat.

## 4. Klasifikasi dengan Naïve Bayes

Pada tahap ini data diklasifikasikan menggunakan Naïve Bayes dengan menentukan opini ulasan produk skin care pada twitter menjadi 2 kelas yaitu, opini positive dan negative.

### 1.6.3 Metode Perancangan

1. Perancangan Aplikasi menggunakan flowchart.
2. Perancangan Antarmuka Pengguna.

#### **1.6.4 Metode Implementasi**

1. Implementasi Sistem di shell python dengan menggunakan bahasa python.
2. Manual Program.
3. Hasil Implementasi.

#### **1.6.5 Metode Testing**

Confision Matrix Testing merupakan suatu metode yang umumnya digunakan untuk melakukan perhitungan tingkat akurasi pada data mining. Confusion matrix memuat informasi tentang klasifikasi yang di prediksi dengan benar oleh sebuah system klasifikasi.

#### **1.7. Sistematika Penulisan**

Sistematika dalam penulisan skripsi ini dibagi menjadi lima bab, antara lain sebagai berikut :

##### **BAB I PENDAHULUAN**

Pada bab ini akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan, manfaat dan sistematika penulisan penelitian.

##### **BAB II LANDASAN TEORI**

Pada bab ini menjelaskan mengenai teori yang digunakan sebagai dasa dalam analisis penelitian. Mencakup tinjauan pustaka meliputi Pengertian Data Mining, Sentiment Analysis, Klasifikasi, Pre-Processing, Term\_Frequency- Inverse Document Frequency (TF IDF), Flowchart, Bahasa Python, Naïve Bayes Classifier, Naïve Bayes multinomial, Waterfall Model, dan Confusion Matrix .

##### **BAB III ANALISIS DAN PERANCANGAN**

Pada bab ini akan menjelaskan mengenai perancangan dan model cara kerja yang akan dilakukan di dalam studi kasus sentimen analisis ulasan produk skin care.

##### **BAB IV IMPLEMENTASI DAN PEMBAHASAN**

Pada bab ini akan menjelaskan Analisis Sentimen dari Pre-Processing data, Tokenization, Steaming, TF-IDF, SVM dan Confusion Matrix.

## **BAB V PENUTUP**

Bab ini akan menjelaskan kesimpulan yang didapat dari hasil analisis penelitian ini, perancangan dan website serta saran untuk pengembangan yang lebih baik lagi.

