

## BAB 1

### PENDAHULUAN

#### 1.1 Latar Belakang

Teknologi telekomunikasi merupakan sarana yang penting dalam membantu dan memudahkan aktivitas masyarakat saat ini, terutama teknologi telepon seluler. Terutama pada saat masa pandemi COVID-19 seperti sekarang, dimana jaringan internet merupakan hal yang penting dalam menunjang aktivitas sehari-hari seperti bekerja dari rumah (*WFH*) dan belajar dari rumah (*Online Studying*). Pada saat ini sudah banyak perusahaan penyedia layanan telekomunikasi seluler yang muncul di Indonesia dan saling bersaing satu sama lain untuk mendapatkan konsumen dengan cara meningkatkan inovasi dan mutu pada layanan produk mereka. Salah satu perusahaan telekomunikasi tersebut adalah Indosat Ooredoo. PT Indosat Ooredoo Tbk adalah suatu perusahaan penyedia jasa telekomunikasi dan jaringan yang menawarkan saluran komunikasi untuk pengguna telepon genggam dan telepon pintar. Beberapa produk yang dihasilkan yaitu IM3 Ooredoo sebagai saluran komunikasi via suara untuk telepon. Indosat Ooredoo juga menyediakan layanan multimedia, internet dan komunikasi data.

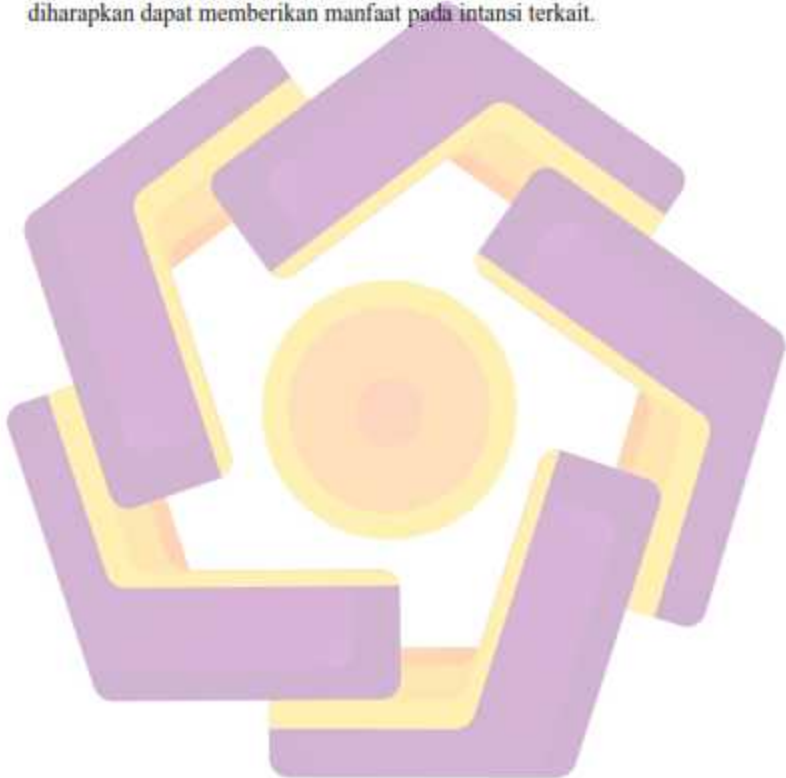
Salah satu tempat untuk pengguna layanan Indosat Ooredoo menyampaikan pendapat mengenai produk Indosat Ooredoo adalah melalui media social *twitter*. Pendapat itu dapat berupa komentar, keluhan, dan masukan. Semakin banyak pengguna yang menyampaikan pendapatnya maka hal tersebut memiliki manfaat yang baik untuk Indosat Ooredoo itu sendiri, mulai dari penilaian positif dan negatif

yang nantinya dapat mempengaruhi pandangan masyarakat mengenai produk Indosat Ooredoo.

Komentar-Komentar tersebut memiliki nilai manfaat yang dapat diimplementasikan untuk proses analisis sentimen sebagai sarana untuk pengambilan keputusan untuk kedepannya atau dapat menjadi evaluasi untuk Indosat Ooredoo sehingga dapat meningkatkan kualitas pemasaran, produk, dan layanan. Untuk menghasilkan informasi dari data opini atau komentar yang ada, maka perlu dilakukan pengolahan data dengan metode analisis sentimen untuk mendapatkan solusi dari permasalahan yang ada dengan memisahkan opini-opini ke dalam kelas positif dan negatif.

Analisis sentimen merupakan proses pengolahan data berbentuk teks yang bertujuan untuk mengklasifikasikan sentimen menjadi sentimen positif dan negatif. Dalam proses pengolahan, data teks akan melalui beberapa tahapan yaitu Pengumpulan Data, *Cleaning Data*, *Tokenization*, *Stopword Removal*, *Stemming*, *Feature List*, dan proses Pembobotan kemudian setelah itu diklasifikasikan menggunakan algoritma *Support Vector Machine* (SVM). SVM merupakan salah satu metode klasifikasi dengan menggunakan metode machine learning yang memprediksi kelas berdasarkan pola dari hasil proses *training* yang diciptakan oleh Vladimir Vapnik. Klasifikasi dilakukan dengan garis pembatas (*hyperplane*) yang memisahkan antara kelas opini positif dan opini negatif [1]. Dalam penelitian yang berjudul "Analisis Sentimen Hatespeech pada Twitter Dengan Metode Naïve Bayes Classifier Dan Support Vector Machine" yang dilakukan oleh Ghulam Asrofi Buntoro [2] dapat disimpulkan bahwa metode SVM merupakan metode terbaik

dibanding *Naïve Bayes Classifier* karena dengan metode SVM tingkat akurasi prediksi yang dihasilkan lebih tinggi. Oleh karena itu maka dalam penelitian ini akan menggunakan metode Support Vector Machine untuk mengukur tingkat kepuasan pelayanan dari Indosat Ooredoo. Dengan dilakukan penelitian ini diharapkan dapat memberikan manfaat pada instansi terkait.



## 1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, maka masalah yang menjadi pusat perhatian dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Menemukan sentimen yang diberikan oleh pengguna layanan telekomunikasi Indosat Ooredoo dari komentar twitter menggunakan Algoritma *Support Vector Machine*.
2. Mencari nilai akurasi algoritma *Support Vector Machine* pada klasifikasi sentimen terhadap pelayanan jasa telekomunikasi Indosat Ooredoo.

## 1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. Penelitian hanya dilakukan terhadap data *twitter* dengan *hashtag* yang berhubungan dengan Indosat Ooredoo yaitu *@indosatoredocare*.
2. Bahasa yang diambil dari Twitter adalah Bahasa Indonesia.
3. Analisis sentimen dengan metode SVM menggunakan alat bantu yang sudah tersedia di *library python*.
4. Sentimen diklasifikasikan kedalam 2 kelas yaitu positif dan negatif.
5. Bahasa pemrograman yang digunakan adalah Python.
6. *Training* data dilakukan secara terpisah dengan test predik suatu data.
7. Analisis sentimen untuk data pengujian data testing dalam penelitian ini menggunakan data tweet dari Twitter yang diunduh secara *realtime*.
8. Jumlah data yang digunakan pada saat proses prediksi adalah 100 data *twitter*.
9. Jumlah data *training* adalah 800.

#### 1.4 Tujuan Penelitian

Tujuan yang diharapkan dari penelitian ini adalah :

1. Menerapkan metode SVM pada analisis sentimen terhadap data twitter Indosat Ooredoo.
2. Menemukan sentimen yang berhubungan dengan Indosat Ooredoo dari komentar twitter menggunakan algoritma Support Vector Machine.
3. Mencari nilai akurasi Algoritma Support Vector Machine dalam mengklasifikasi sentimen terhadap tingkat kepuasan Pelayanan Indosat Ooredoo.

#### 1.5 Manfaat Penelitian

1. Bagi Indosat Ooredoo  
Mendapatkan nilai sentimen publik dari data *twitter* Indosat Ooredoo terhadap pelayanan yang diberikan guna meningkatkan pelayanan dan mutu dari Indosat Ooredoo.
2. Bagi Ilmu Pengetahuan
  - Diharapkan Mengetahui bagaimana implementasi *opinion mining* dalam pemrograman python.
  - Mengetahui cara memperoleh, mengolah data untuk sentimen analisis.

#### 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 Analisa Sentimen/ Opinion Mining, Algoritma Support Vector Machine, K-Fold Cross Validaton.

#### 1.6.1.2 Eksperimen

Menguji kode program agar mudah untuk dikembangkan dan mencari metode penulisan program yang tepat.

### 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. *Cleaning Data* merupakan tahapan penghapusan kata – kata yang tidak dipakai seperti *hashtag* dan *username*.
- b. *Case Folding* atau *Lower Case* digunakan untuk membuat keseluruhan data tweet menjadi *lowercase* (huruf kecil)

- c. *Tokenization* merupakan tahap untuk memisahkan kalimat menjadi kata dan menghilangkan data yang tidak bermakna.
  - d. *Stemming* merupakan tahapan untuk menghilangkan imbuhan awal atau akhir pada kata tersebut seperti berlari menjadi lari.
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 *feature Cosine Similarity*  
*Feature Cosine Similarity* adalah fitur untuk mencari tingkat kesamaan data untuk meningkatkan akurasi data.
5. Klasifikasi Support Vector Machine  
Pada tahap ini data diklasifikasikan menggunakan Algoritma Support Vector Machine dengan menentukan suatu kelas berdasarkan pola yang terbentuk dari hasil proses *training*. Disini data klasifikasi menjadi 2 tahap yaitu dengan tahap *training* yang kemudian akan diberikan label secara manual dan untuk data pengujian yang diunduh secara realtime.
6. Menghitung akurasi dari pengujian data  
Untuk menguji seberapa akurat penelitian ini maka akan dilakukan pengujian dengan 10-fold sebagai parameternya untuk melakukan *Cross Validation* yang diuji menggunakan tools pada python

### 1.6.3 Metode Perancangan

1. Perancangan Aplikasi menggunakan flowchart

#### **1.6.4 Metode Implementasi**

Implementasi Sistem di shell python dengan menggunakan bahasa python.

#### **1.6.5 Metode Testing**

K-fold merupakan metode yang digunakan untuk mengetahui rata-rata keberhasilan suatu sistem.

### **1.7 Sistematika Penulisan**

Pembahasan dalam skripsi ini akan disajikan dalam lima bab yang berurutan sebagai berikut:

#### **1. BAB I PENDAHULUAN**

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

#### **2. BAB II TINJAUAN PUSTAKA**

Pada bab ini menjelaskan mengenai teori yang digunakan sebagai dasa dalam analisis penelitian. Mencakup tinjauan pustaka meliputi Pengertian Data Mining, Support Vector Machine (SVM), Sentiment Analysis, Pre-Processing, Term Frequency-Inverse Document Frequency, Cosine Similarity, K-Fold Cross Validation dan Confusion Matrik.

#### **3. BAB III METODE PENELITIAN**

Pada bab ini akan menjelaskan mengenai perancangan dan model cara kerja yang akan dilakukan di dalam studi kasus sentimen analysis tweet Indosat Ooredoo Care.



#### 4. BAB IV HASIL DAN PEMBAHASAN

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

#### 5. BAB V KESIMPULAN DAN SARAN

Bab ini akan menjelaskan kesimpulan yang didapat dari hasil analisis penelitian ini.

