

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Dalam masa perkembangan digital yang tidak ada habisnya menjadikan internet sebagai bagian yang tidak terpisahkan dari kehidupan sehari-hari. Berbagai aktifitas bisa dilakukan dengan internet seperti, komunikasi, mencari informasi, mengakses konten hiburan, pendidikan, dan pekerjaan, semuanya terhubung dengan internet.

Berkembangnya internet dengan cepat membuat sosial media seperti *Facebook* dan *Twitter* sudah menjadi saluran utama bagi masyarakat untuk berbagi pendapat, mengungkapkan pengalaman, dan memberikan ulasan tentang berbagai topik. *Twitter*, sebagai salah satu sosial media yang lebih ringkas memungkinkan pengguna menyampaikan pandangan mereka tentang ulasan suatu produk, disisi lain *Facebook* sebagai platform media sosial memberikan ruang yang luas memungkinkan pengguna untuk memberikan pendapat atau ulasan mereka secara terperinci melalui fitur ulasan dan komentar. Besarnya pengguna sosial media merupakan peluang yang tepat untuk mendapatkan data sentimen suatu produk menjadi lebih efektif dan efisien. Salah satu produk yang sering diperbincangkan adalah *IndiHome*.

*IndiHome* merupakan salah satu penyedia layanan internet yang dimiliki oleh PT. Telkom Indonesia, layanan ini sangat populer di Indonesia. Namun, *IndiHome* juga menghadapi tantangan dalam menjaga kestabilan dan kecepatan koneksi internet mereka. Banyak pengguna yang merasa kesal dengan masalah-

masalah seperti koneksi terputus, kecepatan yang lambat, atau gangguan teknis lainnya. Permasalahan tersebut perlu diperhatikan agar segera melakukan peningkatan kualitas pelayanan. Langkah tersebut dapat dilakukan dengan memanfaatkan data dari pelanggan yang dapat diekstrak menjadi sebuah informasi. Saat ini PT. Telkom Indonesia hanya mengetahui *review* atau pendapat pelanggan melalui telepon *Customer Service* (CS). Dengan adanya sosial media yang luas data dapat dikumpulkan berupa kritikan atau keluhan pelanggan IndiHome yang dapat dilihat dari sentimen opini pengguna pada media sosial *Twitter* dan *Facebook*, salah satu teknik yang dapat digunakan adalah *Data Mining*.

Berdasarkan permasalahan tersebut menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes* klasifikasi terhadap pengaruh ulasan pelanggan *IndiHome* di *Twitter* dan *Facebook* dengan mengklasifikasikan berdasarkan sentimen positif dan negatif, penelitian ini bertujuan untuk memahami pendapat pelanggan *IndiHome* dengan lebih mudah. Hasil analisis sentimen pelanggan dapat memberikan manfaat yang berharga dalam mengevaluasi dan meningkatkan pelayanan di area tertentu, sehingga dapat memberikan kepuasan yang lebih baik kepada pelanggan.

Berdasarkan permasalahan diatas maka, skripsi ini akan membahas tentang pengaruh kepuasan pelanggan *IndiHome* di *Twitter* dan *Facebook* dengan klasifikasi berdasarkan sentimen positif dan negatif.

Dengan menggunakan metode tersebut diharapkan memberikan manfaat dan hasil dari skripsi ini dapat digunakan untuk bahan evaluasi untuk peningkatan

pelayanan agar dapat memberikan kepuasan terhadap pelanggan yang lebih baik lagi. Dengan berbagai macam pertimbangan yang dilakukan, peneliti memilih judul yaitu :

**“ANALISIS SENTIMEN ULASAN PELANGGAN INDIHOME DI FACEBOOK DAN TWITTER DENGAN METODE SUPPORT VECTOR MACHINE DAN NAIVE BAYES KLASIFIKASI”.**

### **1.2 Rumusan Masalah**

Berdasarkan latar belakang dari uraian di atas, berikut adalah rumusan masalah yang terdapat dalam Skripsi ini :

1. Bagaimana menerapkan metode web scraping dengan algoritma *Support Vector Machine* dan *Naïve Bayes* untuk klasifikasi data ulasan pelanggan *IndiHome* ?
2. Berapa perbandingan akurasi yang dihasilkan algoritma *Support Vector Machine* dan *Naïve Bayes* dalam melakukan proses klasifikasi ulasan pelanggan *IndiHome*?
3. Apakah algoritma *Support Vector Machine* dan *Naïve Bayes* yang sudah dilatih dan diuji cukup baik berdasarkan tingkat akurasinya?

### **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk menganalisis sentimen ulasan pelanggan *IndiHome* yang tersebar di media sosial *Twitter* dan *Facebook* menggunakan metode klasifikasi *Support Vector Machine* dan *Naïve Bayes* sebagai pembanding.

## 1.4 Manfaat Penelitian

Menerapkan ilmu yang sudah didapatkan dari perkuliahan khususnya *data mining*, serta mengetahui gambaran umum persepsi pengguna tentang layanan *IndiHome* berdasarkan ulasan di sosial media seperti *Facebook* dan *Twitter*. Hasil klasifikasi sentimen dapat dijadikan sebagai evaluasi untuk meningkatkan dan memperbaiki kualitas pelayanan.

## 1.5 Batasan Masalah

Dalam ilmu data terdapat cakupan yang sangat luas, sehingga peneliti membuat batasan masalah, yaitu :

1. Dalam *web scraping*, data yang digunakan adalah data hasil *scraping* pada ulasan di *Twitter* dan *Facebook*.
2. Data *scraping* yang digunakan berjumlah 5000 data pada sosial media *Facebook* dan *Twitter*
3. Menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes* untuk melakukan perbandingan klasifikasi keluhan pelanggan.
4. Menggunakan *Jupyter Lab* sebagai lingkungan pengembangan, *Python* sebagai bahasa pemrograman, dan *Streamlit* sebagai *framework web*.
5. Proses deployment website menggunakan *streamlit*

## 1.6 Metodologi Penelitian

### 1.6.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan pada “*Analisis Sentimen Ulasan Pelanggan Facebook Dan Twitter Dengan Metode Support Vector*”

*Machine Dan Naive Klasifikasi (Studi Kasus: PT. Telkom Indonesia)*”, yaitu sebagai berikut :

1. Studi Literatur

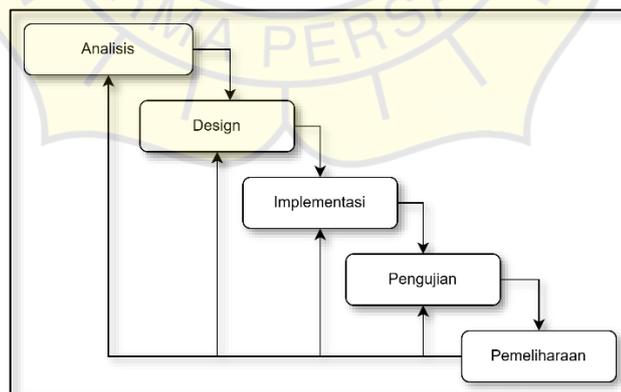
Peneliti mengumpulkan data dari jurnal atau karya tulis ilmiah yang relevan, sehingga dapat membantu peneliti dalam menambah referensi sesuai dengan topik yang diangkat.

2. Filter Data

Peneliti menggunakan *filtering data* dari *dataset* yang akan digunakan, untuk mendapatkan variabel dan data yang dapat digunakan.

### 1.6.2 Metode Pengembangan Sistem

Penelitian ini menggunakan System Development Life Cycle (SDLC) Waterfall yaitu proses untuk membangun sistem dari satu tahap menuju tahap lainnya layaknya air terjun. Dalam pengembangan sistem menggunakan model waterfall, memiliki beberapa tahapan dengan uraian sebagai berikut:



**Gambar 1. 1** SDLC Diagram Waterfall

Metode Waterfall adalah cara membuat sistem dengan langkah-langkah yang berurutan. Setiap langkah harus selesai sebelum langkah berikutnya dimulai.

Salah satu keuntungannya adalah semua kebutuhan harus jelas sebelum mulai membuat sistem. Terdapat 5 tahapan pada metode waterfall yaitu Analisis, Desain, Implementasi, Pengujian, dan Pemeliharaan

#### 1. Analisis

Tahap ini merupakan persyaratan dan kebutuhan sistem yang akan dikembangkan dikumpulkan dan dianalisis dengan seksama agar mendapatkan pemahaman yang jelas tentang kebutuhan sistem yang telah ditetapkan sebelum melanjutkan ke tahap desain dan implementasi.

#### 2. Desain

Tahapan ini menerjemahkan kebutuhan pengguna dari tahap analisis kedalam bentuk UML seperti usecase, dan flowchart dan mockup aplikasi.

#### 3. Implementasi

Tahap ini merupakan proses transformasi hasil perencanaan desain sebelumnya menjadi aplikasi fungsional melalui pelaksanaan pengkodean. Dalam pembangunan sistem ini, bahasa pemrograman *Python* digunakan sebagai alat pengembangan utama. Tahap ini mengadopsi kerangka kerja CRISP-DM, sebuah pendekatan yang telah terstandarisasi dalam proses eksplorasi dan *data mining*. Tahapan ini juga dikenal sebagai inti dari proses pembangunan sistem, yang menghasilkan program komputer yang sesuai dengan rancangan yang telah dirancang sebelumnya.

#### 4. Pengujian

Tahap ini melakukan pengujian terhadap sistem yang sudah selesai dibuat pada tahap sebelumnya, model yang dihasilkan CRISP-DM akan

diujikan pada tahap ini, serta fungsionalitas aplikasi, hal ini dilakukan untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan. Pengujian dilakukan dengan 2 tahap, pertama pengujian akurasi algoritma dengan menggunakan *classification report*.

## 5. Pemeliharaan

Tahap ini adalah tahap pemeliharaan terhadap sistem, yang memungkinkan sebuah sistem akan mengalami perubahan ketika sudah melalui tahap *deployment*. Perubahan bisa terjadi ketika adanya kesalahan yang muncul dan tidak terdeteksi saat pengujian atau perangkat lunak sedang beradaptasi dengan lingkungan pengembangan yang baru.

### 1.6.3 Metode Implementasi

Dalam melakukan implementasi model yang peneliti paparkan, peneliti menggunakan metode *CRISP-DM (Cross Industry Standard Process for Data Mining)* dengan melakukan beberapa tahapan, yaitu :

#### 1. *Business Understanding*

Analisis dilakukan untuk memahami pelanggan *IndiHome* pada sosial media, seperti *Twitter* dan *Facebook* apakah memiliki sentimen positif atau negatif terhadap *IndiHome*.

#### 2. *Data Understanding*

Setelah mengumpulkan data menggunakan metode *scraping* pada sosial media *Twitter* dan *Facebook*, data akan dianalisis apakah sudah cukup layak untuk dilakukan pengolahan data.



### 3. *Data Preparation*

Selanjutnya dilakukan tahap *Data Preparation*, data akan dibersihkan dan diolah sehingga menghasilkan data yang relevan agar dapat dilakukan tahap selanjutnya yaitu labeling atau pelabelan kata positif dan negatif.

### 4. *Modeling*

Tahapan ini dilakukan pembobotan kata positif dan negatif dengan menggunakan *TF-IDF*, dan juga menerapkan algoritma *Support Vector Machine* dan *Naïve Bayes*.

### 5. *Evaluation*

Setelah berhasil mendapatkan akurasi dengan menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes*, selanjutnya kedua algoritma tersebut akan dievaluasi menggunakan *Confussion Matrix*.

### 6. *Deployment*

Setelah berhasil mendapatkan hasil akurasi serta mengetahui hasil evaluasi. Web Sentimen Analysis akan di deploy menggunakan layanan hosting milik streamlit dan deta space sebagai *Database Cloud*.

## 1.7 **Sistematika Penulisan**

Adapun sistematika penulisan pada laporan penelitian skripsi ini adalah sebagai berikut:

## **BAB I PENDAHULUAN**

Pada bab ini, berisi gambaran umum hingga spesifik tentang permasalahan yang ditemukan di lapangan. Penjabaran masalah terdiri dari latar belakang

masalah, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta metodologi penelitian yang digunakan.

## **BAB II LANDASAN TEORI**

Pada bab ini, berisi tentang pemaparan teori yang mengutip dari penelitian sebelumnya serta beberapa referensi lain seperti buku dan website resmi sebagai pendukung penelitian ini.

## **BAB III ANALISIS DAN PERANCANGAN**

Pada bab ini, berisi tentang uraian secara rinci mengenai metode yang digunakan saat penelitian, Kemudian melakukan analisis kebutuhan, perancangan, dan implementasi sistem sesuai dengan metode penelitian yang digunakan.

## **BAB IV PEMBAHASAN**

Pada bab ini, berisi tentang Implementasi Algoritma *Support Vector Machine* dan *Naïve Bayes* untuk pengelompokan teks yang berguna untuk mengklasifikasikan ulasan bernada negatif atau positif, metode ini akan melakukan kalkulasi probabilitas yang digunakan untuk penjumlahan frekuensi serta kombinasi nilai dari dataset. Pada persamaan rumus *Support Vector Machine* dan *Naïve Bayes*.

## **BAB V PENUTUP**

Pada bab ini, berisi tentang kesimpulan yang didapatkan dari hasil pengujian dan evaluasi Implementasi yang dilakukan oleh peneliti, serta saran untuk pembaca dalam mengembangkan penelitian ini agar menjadi lebih baik



**TEKNOLOGI INFORMASI**

**UNIVERSITAS DARMA PERSADA**