

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Teknologi *internet* adalah jaringan data yang mampu menghubungkan seluruh dunia, menghilangkan batasan ruang dan waktu, serta memungkinkan berbagai aktivitas elektronik seperti *e-commerce* dan layanan data publik. Dengan beragam kemudahan dan fasilitas yang ditawarkan, internet menjadi elemen penting dalam komunikasi modern. Seiring dengan semakin luasnya penggunaan *internet* oleh masyarakat, trafik *backbone* menjadi padat dan kualitas koneksi menjadi tantangan. Penyelenggara akses *internet*, baik penyelenggara jaringan (*network operator*) maupun penyelenggara jasa *Internet Service Provider*, secara kompetitif menyediakan layanan beragam untuk pelanggan. (Ruth, 2013)

Menurut laporan *We Are Social*, pada Januari 2023, ada 167 juta orang di Indonesia yang aktif menggunakan media sosial, yang merupakan 60,4% dari total populasi. Jumlah pengguna aktif ini mengalami penurunan sebesar 12,57% dibandingkan tahun sebelumnya, di mana jumlahnya mencapai 191 juta orang. Penurunan ini merupakan yang pertama dalam satu dekade terakhir (Widi, 2023).

Berdasarkan laporan terbaru dari *We Are Social*, per Oktober 2023, jumlah pengguna *Twitter* (sekarang dikenal sebagai X) di Indonesia mencapai sekitar 27,5 juta orang. Jumlah ini menempatkan Indonesia di peringkat keempat dunia dalam hal jumlah pengguna *Twitter*, naik dari peringkat keenam pada Mei 2023. Pertumbuhan ini menunjukkan peningkatan yang signifikan dalam

adopsi platform di Indonesia, dengan peningkatan 71,2% secara kuartalan. Amerika Serikat masih memimpin dengan jumlah pengguna terbanyak, yaitu 108,55 juta, diikuti oleh Jepang dengan 74,1 juta dan India dengan 30,3 juta pengguna. Di bawah Indonesia, ada Inggris dengan 24,3 juta pengguna dan Brasil dengan 24,15 juta pengguna.

Lonjakan ini mencerminkan meningkatnya popularitas *Twitter* di Indonesia, meskipun platform tersebut menghadapi tantangan seperti penurunan valuasi setelah diakuisisi oleh Elon Musk. *Twitter* tetap menjadi salah satu media sosial utama yang digunakan untuk berbagai aktivitas, termasuk diskusi publik dan layanan pelanggan di Indonesia (Katadata, 2023).

Analisis sentimen di *Twitter* telah menjadi alat penting untuk memahami opini publik dan dapat digunakan untuk berbagai tujuan, termasuk meningkatkan layanan pelanggan dan membuat keputusan bisnis yang lebih baik (Wang et al., 2022). Analisis sentimen yang dilakukan pada data *Twitter* dapat membantu dalam memprediksi tren pasar, memahami kebutuhan pelanggan, dan mengidentifikasi area yang memerlukan perbaikan (Wang et al., 2022).

Sehingga berdasarkan latar belakang tersebut, penelitian ini dibuat untuk menganalisis sentimen pengguna dan calon pengguna layanan *Internet Service Providers (ISP)* di Indonesia melalui media sosial *Twitter*. Penelitian ini mengambil data *tweet* dari media sosial *Twitter* sejumlah 1000 *tweet* 6 perusahaan penyedia layanan *internet* dengan kata kunci yang relevan. Analisis sentimen dalam penelitian ini menggunakan pemodelan *NLP (natural*

*language processing*) dan *Clustering* dengan metode *K-Means* yang diharapkan bisa mengklasifikasikan opini pengguna maupun calon pengguna *ISP*. Dengan adanya penelitian ini, diharapkan dapat memberikan wawasan baru berupa hasil penelitian yang selanjutnya bisa digunakan oleh pihak-pihak terkait yang membutuhkan informasi mengenai opini pengguna dan calon pengguna *ISP* dalam media sosial *Twitter*, baik untuk membuat kebijakan maupun penelitian lanjutan.

Penelitian ini merupakan penelitian replikasi yang dimodifikasi dari penelitian yang dilakukan oleh Hashfi et al., (2022). Modifikasi atas penelitian ini adalah penggantian metode analisis sentimen dari sebelumnya menggunakan *SVM* dan *Naïve Bayes* menjadi menggunakan menggunakan metode *NLP (natural language processing)* dan penambahan evaluasi terhadap sentimen pengguna *ISP* di media sosial *Twitter* menggunakan metode *Clustering K-Means*. Dengan adanya penelitian ini, diharapkan dapat memberikan wawasan baru berupa hasil penelitian yang selanjutnya bisa digunakan oleh pihak-pihak terkait yang membutuhkan informasi mengenai opini pengguna dan calon pengguna *ISP* dalam media sosial *Twitter*, baik untuk membuat kebijakan maupun penelitian lanjutan

## **1.2 Perumusan Masalah**

Pada hasil penelitian terdahulu yang meneliti sentimen pelanggan terhadap layanan penyedia layanan *internet* menggunakan algoritma klasifikasi *Naïve Bayes* dan *Support Vector Machine (SVM)*, ditemukan bahwa algoritma *SVM* memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan *Naïve*

*Bayes* (Siregar et al., 2021). Penelitian tersebut berfokus pada pengklasifikasian sentimen positif dan negatif dari *tweet* pengguna layanan *internet*. Namun, penelitian tersebut memiliki beberapa keterbatasan, seperti proses pelabelan sentimen yang dilakukan secara manual.

Penelitian ini berbeda dengan penelitian terdahulu karena menggunakan metode *K-Means Clustering* untuk mengidentifikasi tema-tema utama yang dibahas dalam *tweet* pengguna layanan *ISP* di Indonesia. Penelitian ini juga memanfaatkan analisis sentimen dengan menggunakan model *NLPTown/bert-base-multilingual-uncased-sentiment* untuk mengklasifikasikan sentimen *tweet* menjadi positif, negatif, dan netral. Dengan menggunakan 6000 *tweet*, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai pola dan masalah utama yang dihadapi oleh pengguna layanan *ISP*. Berdasarkan uraian di atas maka pertanyaan penelitian yang akan dijawab di penelitian ini adalah sebagai berikut:

- 1) Bagaimana sentimen pengguna terhadap layanan *ISP* di Indonesia yang tercermin dalam *tweet* di platform media sosial X?
- 2) Apa saja tema utama yang sering dibahas oleh pengguna terkait layanan *ISP* di Indonesia?
- 3) Bagaimana *clusterisasi tweet* dapat membantu mengidentifikasi pola dan masalah utama yang dihadapi oleh pengguna layanan *ISP*?
- 4) Apa Saja Rekomendasi yang Bisa Diberikan untuk Peningkatan Layanan *ISP*?

- 5) Bagaimana Penelitian Ini Dapat Mengembangkan Wawasan Baru untuk Penelitian Lanjutan?

### 1.3 Tujuan Penelitian

Berdasarkan perumusan masalah yang telah dijelaskan, tujuan dari penelitian ini adalah sebagai berikut:

**1) Menganalisis Sentimen Pengguna *ISP* di Indonesia:**

Mengetahui sentimen pengguna terhadap layanan *ISP* di Indonesia yang tercermin dalam *tweet* di platform media sosial X, apakah positif, negatif, atau netral.

**2) Mengidentifikasi Tema Utama yang Dibahas oleh Pengguna:**

Mengidentifikasi tema-tema utama yang sering dibahas oleh pengguna terkait layanan *ISP* di Indonesia melalui analisis *clustering tweet*. Tujuan ini membantu dalam mengungkapkan masalah dan keluhan yang paling sering dibicarakan oleh pengguna.

**3) Menggunakan *Clustering* untuk Menemukan Pola dan Masalah Utama:**

Menggunakan metode *K-Means Clustering* untuk mengelompokkan *tweet* dan menemukan pola serta masalah utama yang dihadapi oleh pengguna layanan *ISP*. Dengan clusterisasi ini, diharapkan dapat ditemukan *cluster-cluster* tertentu yang mewakili berbagai isu seperti kualitas jaringan, layanan pelanggan, harga, dan sebagainya.

**4) Memberikan Rekomendasi untuk Peningkatan Layanan *ISP*:**

Berdasarkan hasil analisis sentimen dan *clustering*, penelitian ini bertujuan untuk memberikan rekomendasi yang dapat digunakan oleh perusahaan *ISP* untuk meningkatkan kualitas layanan dan kepuasan pelanggan. Informasi yang diperoleh dari penelitian ini diharapkan dapat digunakan untuk membuat kebijakan yang lebih efektif dan strategi pemasaran yang lebih baik.

#### **5) Mengembangkan Wawasan Baru untuk Penelitian Lanjutan:**

Penelitian ini diharapkan dapat memberikan kontribusi akademis dengan menyediakan data dan analisis yang dapat digunakan sebagai dasar bagi penelitian lanjutan di bidang analisis sentimen dan clusterisasi data. Hasil penelitian ini juga diharapkan dapat digunakan oleh pihak-pihak terkait yang membutuhkan informasi mengenai opini pengguna layanan *ISP* di media sosial *Twitter*.

#### **1.4 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis sebagai berikut:

1. Pengembangan Ilmu Pengetahuan: Penelitian ini berkontribusi pada pengembangan ilmu pengetahuan khususnya dalam bidang analisis sentimen dan clusterisasi data. Dengan memanfaatkan metode *K-Means Clustering* dalam analisis sentimen pengguna layanan *ISP* di media sosial *Twitter*, penelitian ini memberikan wawasan baru tentang bagaimana teknik ini dapat digunakan untuk mengevaluasi opini publik dan persepsi pengguna.

2. Landasan untuk Penelitian Lanjutan: Hasil dari penelitian ini dapat dijadikan landasan bagi penelitian-penelitian lanjutan yang berkaitan dengan analisis sentimen dan clusterisasi data. Penelitian ini dapat memberikan dasar metodologis dan analitis yang kuat bagi akademisi dan peneliti yang ingin mengeksplorasi lebih jauh tentang penggunaan metode *machine learning* dalam analisis sentimen media sosial.
3. Pengembangan Metodologi Analisis Sentimen: Penelitian ini menunjukkan bagaimana metode *unsupervised learning* seperti *K-Means Clustering* dapat diterapkan dalam analisis sentimen, yang biasanya menggunakan metode *supervised learning*. Hal ini dapat memotivasi pengembangan dan penerapan metode-metode baru dalam analisis sentimen di masa depan.
4. Bagi Perusahaan *ISP*: Penelitian ini memberikan informasi yang berharga bagi perusahaan *ISP* mengenai sentimen dan opini pengguna layanan mereka. Dengan mengetahui tema-tema utama yang sering dibahas oleh pengguna, perusahaan dapat mengidentifikasi masalah-masalah utama yang dihadapi pelanggan dan mengembangkan strategi untuk meningkatkan kualitas layanan dan kepuasan pelanggan.
5. Bagi Pemasar dan Manajer Media Sosial: Hasil analisis sentimen dan *clustering* dapat digunakan oleh tim pemasaran dan manajer media sosial untuk memahami persepsi publik terhadap merek mereka. Informasi ini dapat membantu dalam merancang kampanye pemasaran yang lebih efektif dan responsif terhadap kebutuhan dan keluhan pelanggan.
6. Bagi Pembuat Kebijakan Informasi yang diperoleh dari penelitian ini dapat digunakan oleh pembuat kebijakan dalam merumuskan kebijakan yang

lebih baik untuk meningkatkan kualitas layanan internet di Indonesia. Kebijakan yang berbasis data ini diharapkan dapat menjawab kebutuhan masyarakat dan meningkatkan penetrasi serta kualitas layanan *internet*.

7. Bagi Peneliti dan Akademisi: Penelitian ini menyediakan data dan temuan yang dapat digunakan oleh peneliti dan akademisi dalam studi lebih lanjut. Selain itu, metodologi yang digunakan dalam penelitian ini dapat diaplikasikan dalam konteks yang berbeda untuk mengeksplorasi aspek-aspek lain dari analisis sentimen dan clusterisasi data.
8. Bagi Konsumen: Dengan adanya penelitian ini, konsumen dapat memperoleh gambaran yang lebih jelas tentang kualitas layanan *ISP* yang ada di Indonesia. Informasi ini dapat membantu konsumen dalam mengambil keputusan yang lebih baik dalam memilih layanan *ISP* yang sesuai dengan kebutuhan mereka.

### **1.5 Sistematika Penulisan**

Penulisan penelitian ini disusun secara sistematis dengan mengikuti beberapa bab utama yang diuraikan sebagai berikut:

- 1) BAB I.

Bab ini memberikan gambaran umum mengenai latar belakang, perumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

- 2) BAB II.

Bab ini berisi tinjauan teori dan penelitian sebelumnya yang berkaitan dengan topik penelitian. Dalam bab ini dijelaskan teori-teori dasar mengenai analisis sentimen, clusterisasi, dan penggunaan media

sosial dalam konteks layanan *ISP*. Selain itu, juga dibahas penelitian-penelitian sebelumnya yang relevan sebagai dasar dan pembanding penelitian ini.

### 3) BAB III.

Bab ini menjelaskan metode yang digunakan dalam penelitian ini, termasuk jenis penelitian, sumber data, teknik pengumpulan data, dan teknik analisis data. Dijelaskan pula bagaimana proses pengolahan data dilakukan, mulai dari *preprocessing*, vektorisasi data teks menggunakan *TF-IDF*, hingga penerapan metode *K-Means Clustering* dan analisis sentimen menggunakan model *nlptown/bert-base-multilingual-uncased-sentiment*.

### 4) BAB IV.

Bab ini memaparkan hasil dari penelitian yang telah dilakukan, termasuk hasil analisis sentimen, tema utama yang ditemukan dalam *tweet*, dan hasil *clustering tweet*. Pembahasan juga dilakukan untuk menginterpretasikan hasil temuan, mengaitkan dengan teori dan penelitian sebelumnya, serta menjelaskan implikasi praktis dari hasil penelitian.

### 5) BAB V.

Bab ini menyajikan kesimpulan dari penelitian yang telah dilakukan berdasarkan hasil dan pembahasan pada bab sebelumnya. Selain itu, diberikan pula saran-saran yang diharapkan dapat berguna bagi penelitian selanjutnya dan bagi pihak-pihak yang berkepentingan.