

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Seperti yang ditunjukkan oleh Organisasi Kesehatan Dunia (WHO) di situs otoritasnya, Indonesia adalah negara dengan 4.232.099 kasus positif dan 142.828 kasus meninggal angka dari informasi tersebut adalah angka yang besar. Angka ini secara umum akan terus bertambah jika tidak dilakukan pengawasan oleh pemerintah. Lockdown PSBB (Pembatasan Sosial Bersekala Besar), PPKM level 1-4, serta imunisasi di berbagai daerah telah dilakukan di Indonesia untuk mengurangi peningkatan COVID-19, dengan penanganan yang telah berjalan. Kasus COVID-19 di bulan Oktober, serta untuk mengendalikan COVID-19, Pemerintah Indonesia juga telah membuat aplikasi bernama Peduli Lindungi yang dibuat untuk memantau aktifitas serta diharapkan untuk mengurangi lonjakan penyebaran Covid - 19 di Indonesia [1]

Dengan adanya Aplikasi Peduli Lindungi ini menimbulkan berbagai tanggapan masyarakat terhadap aplikasi Peduli Lindungi sehingga mengundang banyak kalangan untuk memberikan pendapat opini di berbagai kalangan melalui media sosial, media sosial yang sering digunakan masyarakat untuk menyuarakan pendapat dan opini adalah Twitter.

Twitter telah menjadi media berbasis web dan smartphone yang paling populer dengan lebih dari 200 juta pengguna dan lebih dari 10,6 miliar tweet yang telah dihasilkan. Melalui posting Twitter, dapat berbagi dan mendapatkan data tentang masalah apapun dengan menggunakan informasi dari Twitter, dapat digunakan untuk menganalisis pendapat dan opini masyarakat terhadap Aplikasi

Peduli Lindungi melalui analisis sentimen dengan menggunakan klasifikasi pendapat ke dalam 2 kelas yaitu "Positif" dan "Negatif". [2]

Analisis Sentimen merupakan suatu peristiwa dari penelitian yang bergantung pada sikap individu terhadap suatu hal. Analisis Sentimen biasanya dilakukan untuk mengumpulkan dan menemukan opini public dalam postingan Blog, Twitter, Facebook dan lainnya. Analisis sentimen diperlukan untuk mengetahui penilaian umum pada suatu objek. Dengan menggunakan informasi dari media online Twitter, dapat dilakukan Analisis Sentimen terhadap Aplikasi Peduli Lindungi. [2]

Support Vector Machine merupakan salah satu algoritma klasifikasi yang baik. Kemampuan Algoritma Support Vector Machine mendefinisikan hyperplane yang baik menjadikan algoritma SVM memiliki tingkat akurasi yang baik dibandingkan algoritma lainnya. Sedangkan algoritma Naïve Bayes merupakan algoritma klasifikasi machine learning dengan penalaran probabilitas yang tak kalah baik dengan algoritma lainnya. Algoritma Naïve Bayes juga memiliki kelebihan yaitu dengan menggunakan data latih yang tidak terlalu banyak dan memiliki tingkat akurasi yang baik. [3]

Terdapat beberapa penelitian analisis sentimen menggunakan algoritma klasifikasi yang pernah dilakukan oleh Henni Tuhuteru [4] dengan penelitian yang berjudul Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier pada penelitian ini menggunakan data set sebanyak 2378 dengan menggunakan algoritma Naïve Bayes dan SVM dengan pengujian validasi menggunakan K Fold Cross Validation dengan hasil SVM lebih baik dari pada Naïve Bayes

dengan tingkat akurasi sebesar 81,61% sedangkan Naïve Bayes memiliki tingkat akurasi 67.29%

Selanjutnya penelitian yang pernah dilakukan oleh Friska Fitriana [3] mengenai Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada media sosial media twitter menggunakan Support Vector Machine serta Naïve Bayes dengan menggunakan data set yang diambil dari twitter sebanyak 1000 dengan keyword "Vaksin Covid-19" berdasarkan penelitian tersebut metode SVM lebih baik sedikit dibandingkan metode Naive Bayes dengan tingkat akurasi 90,47% untuk algoritma SVM sedangkan Naïve Bayes 88,64%

Kemudian penelitian dilakukan oleh Arifin Oki mengenai Analisa Perbandingan Tingkat Performansi Metode Support Vector Machine dan Naïve Bayes untuk Klasifikasi Jalur Minat Bakat SMA. Data set diambil dari dua buah sma dengan data sejumlah 288 siswa, dalam penelitian ini digunakan pengujian menggunakan K Fold Cross Validation, serta menggunakan algoritma Naïve Bayes dan serta SVM dari penelitian ini didapatkan hasil Algoritma SVM lebih baik dari pada Algoritma Naïve Bayes. [5]

Berdasarkan perbandingan ketiga jurnal diatas algoritma SVM dan Naïve Bayes memiliki beberapa kelebihan serta kekurangan masing masing. Pada penelitian ini akan membahas proses analisis sentimen dengan melakukan perbandingan dua metode yaitu Naïve Bayes Classifier dan Support Vector Machine untuk mengetahui perbandingan tingkat akurasi serta waktu pemrosesan yang diimplementasikan pada analisis sentimen twitter terhadap aplikasi Peduli Lindungi.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu dirumuskan suatu masalah penelitian yaitu perbandingan Akurasi serta Waktu Pemrosesan algoritma Naïve Bayes dengan algoritma Support Vector Machine dalam analisis sentimen Aplikasi Peduli Lindungi.

## 1.3 Batasan Masalah

Adapun batasan masalah pada penelitian adalah:

- Penelitian ini mengambil objek berupa opini masyarakat terhadap aplikasi Peduli Lindungi.
- Algoritma yang digunakan untuk mengolah data ialah metode Naïve Bayes serta Support Vector Machine.
- Dataset yang digunakan diambil dari twitter dengan menggunakan tools twint dengan keyword "PLindungi", "Peduli Lindungi"
- Tweet yang digunakan berupa text dan tidak mengandung gambar.

## 1.4 Maksud dan Tujuan Penelitian

Berdasarkan rumusan masalah yang terdapat diatas maksud dan tujuan yang dikedepankan dalam penelitian ini adalah membandingkan akurasi serta waktu proses algoritma Naïve Bayes dengan algoritma Support Vector Machine dalam opini pengguna Twitter terhadap Aplikasi Peduli Lindungi.

## 1.5 Manfaat Penelitian

Manfaat yang diharapkan dari adanya penelitian ini yaitu:

1. Mengetahui seberapa besar tingkat akurasi dan performansi algoritma Support Vector Machine dengan Naïve Bayes
2. Mengetahui algoritma yang sesuai untuk klasifikasi analisis sentimen Aplikasi Peduli Lindungi
3. Sebagai referensi untuk penelitian selanjutnya dan mengembangkan penelitian ini.

## 1.6 Metode Penelitian

1. Metode Pengumpulan Data



Dalam metode ini penulis mencoba mendapatkan data dari twitter dengan cara crawling data menggunakan tools Twint dari akun akun acak berdasarkan keyword “Peduli Lindungi”, “Plindungi”, “Aplikasi Peduli Lindungi”

## 2. Metode Pengolahan data

Pada pengolahan data ini terdiri dari beberapa tahapan yang memiliki masing masing metode untuk Analysis Sentiment. Hal yang pertama kali dilakukan adalah mengcrawling data twitter menggunakan tools Twint selanjutnya data tersebut disimpan dalam bentuk file document, lalu data yang telah diproses dibagi menjadi dua tipe data yaitu data training dan data testing, berikut tahapan selanjutnya dalam pengolahan data training dan data testing.

1. Pelabelan Manual
2. Cleaning
3. Tweet tokenize
4. Stopwords
5. Normalization
6. Text Stemming

## 3. Metode Analisis

Penelitian ini menganalisis setiap data yang di peroleh melalui cara kerja algoritma Naïve Bayes dan Support Vector Machine dalam mengklasifikasikan setiap data tweet yang diperoleh dari crawling data itu sendiri.

### 1.7 Sistematika Penulisan

Adapun sistematika penulisan Skripsi ini adalah sebagai berikut:

#### 1. BAB 1 Pendahuluan

Pada Bab 1 diuraikan isi dan rencana pengerjaan Skripsi secara keseluruhan yang meliputi latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat, metode penelitian yang digunakan.

#### 2. BAB 2 Landasan Teori

Bab 2 memaparkan dasar dasar teori yang berkaitan dengan Analisis Sentiment, Metode Naïve Bayes, Metode Support Vector Machine.

3. BAB 3 Perancangan dan Implementasi  
Perancangan sistem dan contoh implementasi dari sistem yang dibangun dan dipaparkan pada bab ini. Selanjutnya dilakukan proses implementasi.
4. BAB 4 Pengujian dan Analisis  
Pada bab ini dibahas skenario dan hasil dari pengujian yang dilakukan pada hasil implementasi system.
5. BAB 5 Kesimpulan dan Saran  
Bab ini berisi kesimpulan dan saran yang didapatkan dari hasil implementasi sistem secara keseluruhan.

