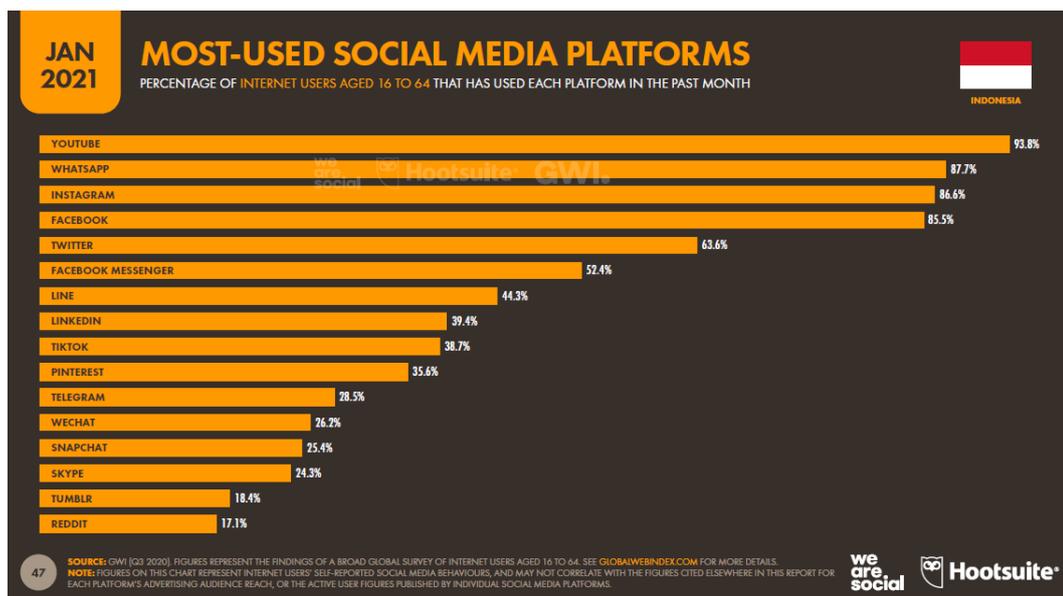


# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Jejaring media sosial pada masa sekarang mempunyai kedudukan yang besar terutama dalam perkembangan peristiwa yang terjadi saat ini. Media sosial merupakan suatu kelompok aplikasi berbasis internet yang menggunakan ideologi dan teknologi Web 2.0, yang mana penggunaanya dapat membuat atau saling bertukar informasi pada aplikasi tersebut (Kaplan and Haenlein, 2010). Banyak sekali media sosial yang berkembang dan digunakan oleh sekian banyak penduduk di Indonesia. Dari data yang diperoleh dari *We Are Social* dan *Hootsuite* disajikan beberapa platform media sosial yang paling sering digunakan di Indonesia, salah satunya media sosial *Twitter* (We Are Social, 2021).



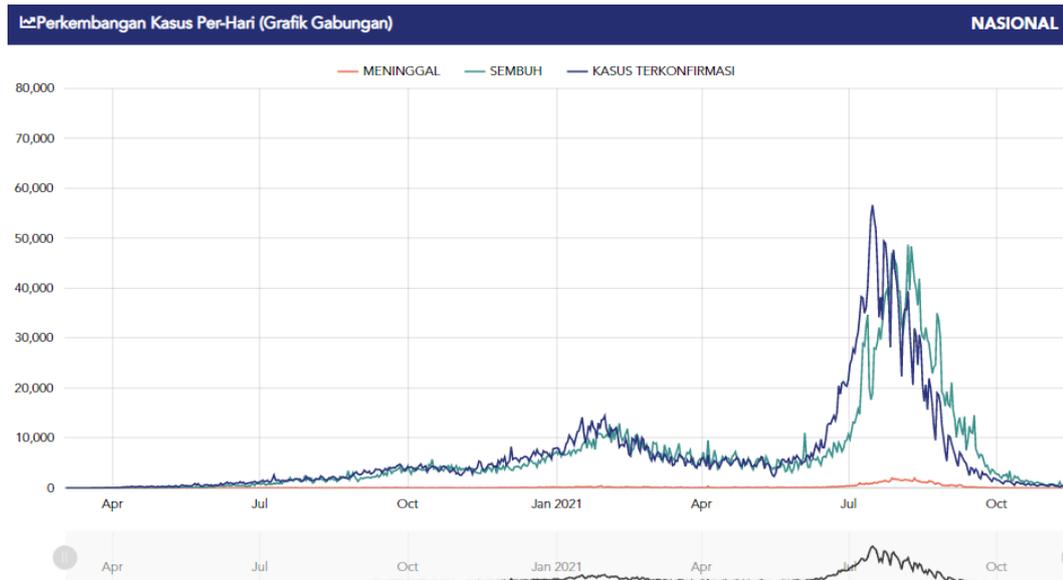
Gambar 1.1 Media sosial yang sering digunakan di Indonesia

*Twitter* merupakan layanan *microblogging* populer di mana pengguna dapat membuat status yang biasa disebut “*tweets*” (Go, Bhayani and Huang, 2009).

Aktivitas yang bisa dilakukan di media sosial *Twitter* yaitu membuat postingan atau “*tweets*”, memberi komentar, sampai dapat mengirimkan pesan pribadi kepada sesama pengguna *Twitter*. Kebiasaan masyarakat Indonesia mem-posting *tweet* ketika terjadi suatu peristiwa disekitar mereka menjadi salah satu acuan untuk mengetahui sentimen masyarakat terhadap lingkungan atau tempat tinggal mereka berupa kritik atau saran (Anjani and , Arif Djunaidy, 2014).

Penelitian ini menggunakan media sosial *Twitter* sebagai sumber data untuk menganalisis sentimen karena *Twitter* merupakan salah satu media sosial yang sering digunakan oleh masyarakat Indonesia. Kelebihan *Twitter* dari media sosial yang lain yaitu biasanya *Twitter* sebagai media untuk menyampaikan keluhan kesah secara langsung dari penggunanya, *Twitter* juga cepat dalam memperbarui sesuatu yang sedang hangat dibicarakan atau viral, dan *Twitter* menyediakan API yang membuat para pengguna dapat mengambil dan memodifikasi data dari *Twitter*. Sehingga *Twitter* merupakan media yang cukup baik untuk diambil datanya dalam penelitian ini.

Pandemi Covid-19 sudah berlangsung selama 1 tahun lebih di Indonesia. Kasus pertama Covid-19 terjadi di kota Wuhan pada awal Desember 2019. Virus Corona atau biasa disebut Covid-19 adalah kelompok virus terbesar dalam ordo *Nidovirales* termasuk dalam familia *Coronaviridae*, sub familia *Coronavirinae*, genus *Betacoronavirus*, subgenus *Sarbecovirus* (Parwanto, 2020). Saat ini di Indonesia pada awal bulan November 2021 kasus aktif penderita Covid-19 sudah mulai menurun menyentuh angka total 8.500 jiwa, angka sembuh sudah mencapai total 4 juta jiwa, dan yang meninggal mencapai total 140.000 jiwa (Satuan Tugas Penanganan covid-19, 2021).



Gambar 1.2 Perkembangan kasus Covid tahun 2021

Dengan semakin menurunnya angka kasus penderita Covid-19 ini maka pemerintah Indonesia mulai melonggarkan peraturan-peraturan terkait pandemi Covid-19 ini seperti peraturan yang pada awal pandemi kegiatan belajar mengajar harus dari rumah menggunakan sistem online sekarang sudah ada beberapa sekolah yang diperbolehkan tatap muka dengan syarat-syarat yang harus dipenuhi. Pada bulan Maret 2021 pemerintah telah melakukan beberapa kali penyesuaian peraturan mengikuti dinamika pandemi. Seperti pada keputusan penyesuaian terbaru, Surat Keputusan Bersama (SKB) Menteri Pendidikan dan Kebudayaan, Menteri Agama, Menteri Kesehatan, dan Menteri Dalam Negeri Republik Indonesia Nomor 03/KB/2021, Nomor 384 tahun 2021, Nomor HK.01.08/MENKES/4242/2021 dan Nomor 440-717 tahun 2021 Tentang Panduan penyelenggaraan Pembelajaran Di Masa Pandemi Corona Virus Disease 2019 (COVID-19). Dalam peraturan tersebut pemerintah sudah memberi izin pembelajaran tatap muka terbatas boleh dilaksanakan dengan syarat tetap menjalankan protokol kesehatan secara ketat.

Berdasarkan peraturan pemerintah tersebut, masyarakat Indonesia mulai membicarakannya di sosial media terutama di *Twitter*. Dari banyaknya opini yang diposting pada media sosial terutama Twitter, seperti perbedaan pendapat dari pernyataan setuju dan tidak setuju terkait pembelajaran tatap muka pada masa pandemi yang diterapkan pemerintah, maka dari itu dapat dilakukan analisis sentimen mulai dari yang positif, negatif sampai netral.

Penelitian ini melakukan analisis sentimen pengguna twitter terhadap penerapan pembelajaran tatap muka pada masa pandemi Covid-19. Untuk mengetahui sentimen dari masyarakat. Salah satu proses yang membedakan dari penelitian ini yaitu terdapat proses analisis bot/spam. Analisis bot digunakan untuk membedakan antara pengguna Twitter yang asli dengan bot/spam (Perdana, Muliawati and Alexandro, 2015). Sehingga ketika data yang dibuat oleh bot telah dibedakan dengan data yang dibuat oleh pengguna asli maka dapat dikatakan bahwa data yang dimiliki adalah data asli. Kemudian pada proses *text preprocessing* terdapat tahap *emoticon and emoji handling* yang berfungsi untuk mengatasi adanya emoticon dan emoji yang ada di dalam text pada *tweet* sehingga data *tweet* hanya berisi kalimat inti yang dibutuhkan saja.

Untuk melakukan analisis sentimen diperlukan teknik klasifikasi dokumen yang memiliki tingkat akurasi yang tinggi. Ada beragam teknik klasifikasi dokumen, beberapa diantaranya yaitu *Decision Trees*, *Naïve Bayes Classifier* (NBC), *K-Nearest Neighbour*, dan *Support Vector Machines* (SVM). Metode *Naïve Bayes Classifier* (NBC) banyak digunakan dalam penelitian mengenai *text mining*, kelebihan dari metode NBC ini salah satunya yaitu algoritma klasifikasi sederhana yang mempunyai kecepatan dan akurasi yang tinggi (Rish, 2001). Sedangkan

metode SVM dipilih karena memiliki kemampuan generalisasi dalam mengklasifikasikan suatu pola dimensi dan menghindari kesulitan dari permasalahan dimensionalitas (Tan, Steinbach and Kumar, 1981). Alasan menggunakan metode *Naïve Bayes* dan *Support Vector Machine* yaitu pada kedua metode tersebut meskipun berbeda tetapi masih memiliki kesamaan yaitu pada jenis model *Support Vector Machine* dan *Naïve Bayes* dapat diartikan sebagai model *Linier Clasifier*. Metode *Support Vector Machine* memiliki prinsip dasar linear classifier dan selanjutnya dikembangkan agar dapat bekerja pada problem non-linear (Nugroho, Witarto and Handoko, 2003). *Naïve Bayes* dikatakan memiliki model linear classifier karena jika memiliki varians ( $\sigma$ ) konstan di seluruh kelas (Gormley, 2018) dan dinyatakan dalam log-space (Hockenmaier, 2015). Selain itu tujuan dari pengambilan dua metode ini yaitu untuk memilih yang terbaik antara dua sekenario model yang berbeda dalam sentimen masyarakat pengguna *Twitter* terhadap pembelajaran tatap muka. Penelitian ini dilakukan dengan mengambil data *tweet* dari *Twitter* dimulai dari tanggal 30 Maret 2021 karena pada tanggal tersebut merupakan tanggal disahkannya peraturan pemerintah tentang panduan penyelenggaraan Pembelajaran Di Masa Pandemi Corona Virus Disease 2019 (COVID-19) sampai tanggal 23 Desember 2021 yang merupakan berakhirnya tahun ajaran akademik semester gasal tahun 2021. Hasil penelitian ini yaitu hasil analisis sentimen, perbandingan antara metode *Support Vector Machine* dan *Naïve Bayes*, dan juga visualisasi website yang dapat dilihat dan dimengerti oleh masyarakat. Dari hasil tersebut diharapkan dapat mengetahui sentimen masarakat terhadap kebijakan sekolah tatap muka di Indonesia ditengah pandemi *Covid-19* sehingga dapat membantu pihak-pihak yang ingin mengetahui kata kunci yang sama dan

sebagai bahan pertimbangan bagi pemerintah terhadap kebijakan atau program yang tepat dalam rangka pemberlakuan pembelajaran tatap muka ditengah pandemi *Covid-19*.

## **1.2 Perumusan Masalah**

Berdasarkan dari latar belakang yang sudah di uraikan maka dapat dirumuskan permasalahan yaitu:

1. Bagaimana melakukan analisis sentimen pada pengguna *Twitter* terhadap pembelajaran tatap muka pada masa pandemi *Covid-19* dari sebuah *tweet*?
2. Bagaimana mengetahui yang terbaik menggunakan filter bot atau tidak antara metode klasifikasi yang diuji?
3. Bagaimana mengetahui metode yang lebih optimal antara metode *Naïve Bayes* dan *Support Vector Machine (SVM)* dalam mengklasifikasikan *tweet*?
4. Bagaimana memvisualisasikan hasil dari klasifikasi kedalam bentuk website?

## **1.3 Batasan Masalah**

Dalam penyusunan penelitian ini terdapat beberapa batasan masalah, diantaranya sebagai berikut:

1. Data yang digunakan dalam penelitian ini merupakan dari media sosial *Twitter* dengan akun publik.
2. Penelitian ini hanya melakukan proses analisis sentimen terhadap *tweet* yang menggunakan bahasa Indonesia.
3. Bahasa pemrograman yang digunakan adalah bahasa pemrograman Python.
4. Data yang diambil yaitu *tweet* dari media sosial Twitter dalam rentang waktu 30 Maret – 23 Desember 2021.

5. Data yang digunakan untuk analisis sentimen berasal dari kata kunci “pembelajaran tatap muka” dan “ptm” yang nantinya akan digabung menjadi satu dataset.
6. Output yang dihasilkan dari penelitian ini yaitu berupa hasil klasifikasi sentimen positif, negatif dan netral dari masyarakat terhadap pembelajaran tatap muka di era pandemi Covid-19.

#### **1.4 Tujuan Penelitian**

Berdasarkan dari rumusan masalah yang sudah dijelaskan, tujuan dari penelitian ini adalah :

1. Mengetahui tanggapan masyarakat bersifat positif, negatif, atau netral yang berasal dari media sosial *Twitter* terhadap pembelajaran tatap muka di masa pandemi *Covid-19*.
2. Melakukan pengolahan data sampai klasifikasi model menggunakan filter bot dan tidak menggunakan filter bot.
3. Melakukan klasifikasi dari metode *Naïve Bayes* dan *Support Vector Machine (SVM)* kemudian memilih yang terbaik.
4. Membuat visualisasi dari klasifikasi data dari yang terbaik antara kedua metode ke dalam bentuk grafik, juga ditampilkan kata yang menjadi pemicu sentiment positif, negatif, dan netral dalam beropini dan dapat memprediksi label dari tweet yang dimasukkan ke dalam inputan yang akan diimplementasikan dalam suatu aplikasi berbasis web menggunakan Bahasa PHP.

## **1.5 Manfaat Penelitian**

Berdasarkan hasil penelitian analisis sentimen yang dilakukan, diharapkan dapat mengetahui sentimen masyarakat terhadap kebijakan sekolah tatap muka di Indonesia ditengah pandemi *Covid-19*. Sehingga hasil sentimen positif dan negatif maupun netral masyarakat Indonesia pengguna *Twitter* yang berupa opini atau tanggapan dapat membantu pihak-pihak yang ingin mengetahui kata kunci yang sama dan sebagai bahan pertimbangan bagi pemerintah terhadap kebijakan atau program yang tepat dalam rangka pemberlakuan pembelajaran tatap muka ditengah pandemi *Covid-19*.

## **1.6 Sistematika Penulisan**

Sistematika penulisan skripsi ini akan membantu mengarahkan penyusunan laporan supaya tidak menyimpang dan sebagai acuan dalam mencapai tujuan penulisan laporan skripsi sesuai dengan apa yang diharapkan. Langkah-langkah dalam proses penyusunan skripsi ini adalah sebagai berikut:

### **BAB I PENDAHULUAN**

Bab ini berisi tentang gambaran umum isi penelitian diantaranya latar belakang, rumusan masalah, batasan masalah, tujuan dan sistematika penulisan.

### **BAB II TINJAUAN PUSTAKA**

Bab ini berisi tentang penelitian terdahulu untuk membandingkan penelitian terdahulu dengan penelitian saat ini, dasar teori yang berhubungan dengan masalah yang akan dibahas seperti mengenai *text mining*, *preprocessing text*, analisis sentiment, klasifikasi teks,

*Twitter, naïve bayes, support vector machine, TF-IDF, dan confusion matrix.*

### **BAB III METODOLOGI PENELITIAN**

Bab ini berisi tentang metode penelitian yang akan dibuat meliputi studi literatur, analisis kebutuhan, dan perancangan model dan visualisasi

### **BAB IV HASIL DAN PEMBAHASAN**

Bab hasil dan pembahasan berisikan penjelasan penelitian mengenai pengumpulan data, pembangunan model, dan pengujian sistem, serta membahas hasil pengujian terhadap implementasi model pada sistem.

### **BAB V PENUTUP**

Bab penutup berisikan kesimpulan mengenai seluruh hasil penelitian dan berisikan saran terhadap penelitian selanjutnya.

### **DAFTAR PUSTAKA**

Berisi tentang literatur yang digunakan sebagai pedoman yang membantu pengerjaan skripsi.

### **LAMPIRAN**

Pada bagian ini berisi tentang data atau pelengkap yang menunjang dalam pembuatan skripsi.