# BAB I

# PENDAHULUAN

## Latar Belakang

Organisasi Kesehatan Dunia (WHO) telah menyatakan bahwa *Coronavirus disease* 2019 atau Covid-19 dikategorikan sebagai pandemi global pada 11 Maret 2020 (Ducharme, 2020). Pandemi tersebut menyebar dengan sangat cepat dan telah melanda 215 negara di dunia (Sadikin & Sadikin, 2020). Penyebaran virus melalui kontak fisik memaksa semua negara untuk menerapkan *social distancing* dan *physical distancing* guna mengurangi interaksi antara orang-orang. Pemerintah Indonesia melalui Presiden Jokowi telah mengeluarkan pernyataan terkait *social distancing* dan *physical distancing* ini dengan dikeluarkannya kebijakan Pembatasan Sosial Berskala Besar (PSBB) untuk mencegah penyebaran virus (Ristyawati, 2020).

Pendidikan merupakan salah satu bidang yang terkena dampak pandemi Covid-19. Kementrian Pendidikan dan Kebudayaan Republik Indonesia melalui menterinya Nadiem Makarim telah mengeluarkan kebijakan tentang pelaksanaan pendidikan dalam masa darurat Covid-19. Kebijakan tersebut menjelaskan tentang pelaksanaan proses Belajar Dari Rumah (BRD) secara daring atau *online*. Karenanya, seluruh institusi pendidikan diminta untuk menghentikan proses belajar mengajar baik di sekolah maupun di kampus dan menggantinya dengan sistem belajar jarak jauh. Hal ini mengakibatkan semua model pembelajaran saat ini harus berlangsung secara daring atau *online* dengan bantuan alat perantara seperti *hand phone*, komputer, atau laptop (Watrianthos, Ronal, 2020). Guru, murid, dan orang tua harus menyesuaikan diri dengan model pembelajaran tersebut. Bagi sebagian Guru yang tidak mahir dalam penggunaan teknologi akan merasa terkejut dan harus segera beradaptasi, demikian juga murid dan orang tua (Wijoyo, Hadion, 2020). Sistem pembelajaran yang semula dianggap sebagai solusi mulai menuai beragam pendapat dari masyarakat.

Penelitian terdahulu yang pernah dilakukan diantaranya adalah analisis pembelajaran daring pada bulan Juli 2020 terhadap kata kunci ‘pembelajaran daring’, ‘kuliah’, ‘belajar’ dan hashtag-hashtag seperti ‘#belajardirumah’ atau ‘#belajardarirumah‘(Watrianthos, Ronal, 2020), memperoleh hasil sentimen negatif sebesar 83%. Studi lain pernah dilakukan perbandingan algoritme melalui analisis sentimen (Nova , dkk, 2019), menyatakan bahwa *K-Nearest Neighbor* (KNN) dapat digunakan untuk klasifikasi sentimen dengan nilai akurasi sebesar 80% dan mampu mengimbangi *Naive Bayes Classifier*. KNN juga digunakan untuk melakukan analisis sentimen terhadap dataset *review* film dengan total 2000 data (Novelty, 2020), menghasilkan hasil yang baik pada K=3 dengan akurasi sebesar 96.8%.

Penelitian ini bertujuan untuk melakukan analisis sentimen masyarakat terhadap sistem pembelajaran daring, menggunakan metode analisis sentimen dengan algoritme klasifikasi *K-Nearest Neighbors* dan pembobotan kata *CountVectorizer*, disertai fitur kamus sentimen yang digunakan untuk menentukan sentimen pada saat pelabelan data. *Dataset* yang digunakan berupa teks kicauan (*tweet*) yang bersumber pada media sosial Twitter dengan kata kunci X, Y, dan Z. Pengumpulan *dataset* dilakukan dari rentang tanggal 1 Desember 2020 sampai dengan 31 Januari 2021 dengan menggunakan fitur *standard* *search* *API* Twitter dan libary *tweepy*. Tercatat ada sekitar X.XXX *tweet* yang diperoleh dengan kata kunci dan rentang tanggal yang diusulkan.

## Perumusan Masalah

Berdasarkan uraian latar belakang di atas, maka dapat dirumuskan masalah sebagai berikut:

* 1. Bagaimana pandangan(sentimen) masyarakat Indonesia terhadap sistem pembelajaran jarak jauh?
  2. Bagaimana cara menganalisis sentimen berdasarkan pendapat masyarakat Indonesia melalui media sosial Twitter menggunakan algoritme *K-Nearest Neighbors*?
  3. Apakah sentimen yang dihasilkan akurat?

## Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

1. Aplikasi menggunakan bahasa pemrograman *Python.*
2. *Platform* yang digunakan hanya berbasis *web.*
3. *Dataset* bersumber pada Twitter, terbatas pada *tweet* berbahasa Indonesia kata kunci X,Y dan Z pada rentang tanggal 1 Desember 2020 sampai dengan 31 Januari 2021.
4. Fitur *import* hanya dapat mengenali file masukan berupa *excel* dengan ekstensi .xls atau .xlsx.
5. Pelabelan sentimen membutuhkan data dari kamus kata positif dan data kata negatif.
6. Aplikasi hanya mengklasifikasikan tweet menjadi dua buah kategori sentimen, yaitu: “positif” dan “negatif”.
7. Waktu pemprosesan meningkat seiring dengan jumlah data yang diproses.

## Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis pandangan(sentimen) masyarakat Indonesia melalui media sosial Twitter terhadap sistem pembelajaran jarak jauh.
2. Menguji keakuratan algoritme *K-Nearest Neighbors* dengan pembobotan kata *CountVectorizer*.

## Manfaat

Manfaat dari penelitian adalah untuk menganalisis pandangan(sentimen) masyarakat Indonesia berdasarkan *tweet* yang dipublikasikan melalui media sosial Twitter. Hasil analisis digunakan untuk untuk mendapatkan informasi terhadap pandangan masyarakat Indonesia terkait sistem pembelajaran jarak jauh. Penelitian ini juga dilakukan guna menguji keakuratan algoritme *K-Nearest Neighbors* dengan pembobotan kata *CountVectorizer* untuk analisis sentimen.

## Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut:

### BAB I: PENDAHULUAN

Bagian ini berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, dan juga membahas mengenai sistematika penulisan.

### **BAB II: LANDASAN TEORI**

Bagian ini berisi tentang algoritme dan metode yang akan dibahas, serta teori-teori yang berkaitan dengan penelitian ini, yaitu pengertian dan pemahaman materi terkait *text mining*, *preprocessing*, *labeling*, *modeling*, pengujian dan evaluasi serta studi literatur.

### **BAB III: METODOLOGI PENELITIAN**

Bagian ini berisi tentang sumber data penelitian, penerapan atau tahapan metode yang digunakan. Bab ini juga berisi tentang rancangan pengujian dari ekstraksi informasi yang didapat.

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

Bagian ini berisi mengenai lingkungan percobaan sistem yang dibuat, implementasi metode, flowchart tahapan metode, dan uraian algoritme pada proses, serta analisa pengujian sistem yang telah dibangun apakah data hasil pengelompokan yang didapat sudah sesuai dan relevan.

### **BAB V: PENUTUP**

Bagian ini berisi tentang kesimpulan yang dapat ditarik dari penelitian dan saran untuk pengembangan lebih lanjut mengenai topik terkait dalam penelitian berikutnya.