# 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 juga orang tua. Menurut Hadion Wijoyo (2020), diketahui bahwa guru menyenangi kelas daring sebesar 67% sedangkan yang lainya lebih menyenangi kelas luring dikarenankan membutuhkan waktu lebih dalam mempersiapkan bahan ajar di kelas daring termasuk pemahaman perangkat IT yang digunakan. Sistem pembelajaran yang semula dianggap sebagai solusi mulai menuai beragam pendapat dari masyarakat.

Menurut Ronal Watrianthos (2020) melalui penelitian yang berjudul Analisis Pembelajaran Daring di Era Pandemic Covid-19, melalui kata kunci ‘pembelajaran daring’, ‘kuliah’, ‘belajar’ dan hashtag-hashtag seperti ‘#belajardirumah’ atau ‘#belajardarirumah‘. Hasil penelitian menunjukkan pendapat (sentimen) masyarakat terhadap pembelajaran daring cenderung mengarah pada hasil sentimen yang negatif sebesar 83% pada bulan Juli 2020. Dalam penelitian tersebut juga dilakukan analisis emosi, menunjukkan bahwa *‘trust’* atau kepercayaan sangat mendominasi yang menandakan kepercayaan terhadap pembelajaran daring telah jauh menurun.

Penelitian lain yang pernah dilakukan tentang analisis sentimen, diantaranya adalah melakukan analisis sentimen pada *tweet* bahasa Indonesia melalui media sosial Twitter terhadap persepakbolaan Indonesia (Andre, dkk, 2019), menyatakan bahwa algoritme *K-Nearest Neighbors* (KNN) mempu memperoleh nilai akurasi 79.99% dengan nilai K=23. Penelitian lain juga dilakukan oleh Nova, dkk (2019), menyatakan bahwa KNN dapat digunakan untuk analisis sentimen dengan nilai akurasi sebesar 80% terhadap 127 data dan mampu mengimbangi algoritme *Naive Bayes Classifier*. KNN juga digunakan oleh Novelty (2020) dalam penelitian yang berjudul *Sentiment Analysis on Movie Reviews Using Information Gain and K-Nearest Neighbor,* untuk melakukan analisis sentimen terhadap dataset *review* film dengan total 2000 data, memperoleh hasil yang baik pada K=3 dengan nilai akurasi sebesar 96.8%.

Penelitian ini bertujuan untuk melakukan analisis sentimen masyarakat terhadap pembelajaran daring pada periode waktu 1 Desember 2020 hingga 31 Januari 2021. Metode yang digunakan adalah dengan melakukan analisis sentimen melalui pendekatan *machine learning* disertai fitur kamus sentimen, dengan ekstraksi fitur menggunakan *CountVectorizer* dan algoritme klasifikasi *K-Nearest Neighbors*. *Dataset* yang digunakan berupa teks kicauan (*tweet*) yang bersumber pada media sosial Twitter dengan kata kunci X, Y, dan Z. Pengumpulan *dataset* dilakukan pada tanggal 1 Desember 2020 hingga 31 Januari 2021. Tercatat ada sekitar X.XXX *tweet* yang diperoleh dengan kata kunci dan rentang tanggal yang diusulkan.

## Perumusan Masalah

Berdasarkan uraian latar belakang yang telah disampaikan, maka dapat disimpulkan rumusan masalah sebagai berikut:

* 1. Bagaimana persentase pandangan (sentimen) masyarakat Indonesia terhadap pembelajaran daring pada periode waktu 1 Desember 2020 hingga 31 Januari 2021?
  2. Bagaimana cara menganalisis sentimen berdasarkan pendapat masyarakat Indonesia melalui media sosial Twitter menggunakan algoritme *K-Nearest Neighbors*?
  3. Berapa nilai akurasi yang diperoleh algoritme *K-Nearest Neighbors* dalam melakukan analisis sentimen?

## Batasan Masalah

Adapun batasan atau ruang lingkup masalah dalam 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. Aplikasi hanya mengklasifikasikan tweet menjadi dua buah kategori sentimen, yaitu: “positif” dan “negatif”.
6. Waktu pemprosesan meningkat seiring dengan jumlah data yang diproses.

## Tujuan

Adapun tujuan dari dilakukan penelitian ini adalah sebagai berikut:

1. Melakukan analisis sentimen masyarakat terhadap pembelajaran daring melalui media sosial Twitter.
2. Merancang sebuah model penelitian untuk menganalisis sentimen dengan topik terkait pembelajaran daring menggunakan algoritme *K-Nearest Neighbors*.
3. Menguji keakuratan algoritme *K-Nearest Neighbors* dalam melakukan analisis sentimen*.*

## Manfaat

Adapun manfaat dari penelitian adalah untuk menganalisis pandangan (sentimen) masyarakat Indonesia berdasarkan *tweet* yang dipublikasikan melalui media sosial Twitter. Sehingga dapat diperoleh gambaran sentimen masyarakat terkait topik pembelajaran daring di tengah pendemi Covid-19. Hasil penelitian ini juga diharapkan menjadi bahan evaluasi untuk sistem pembelajaran daring yang akan berlangsung. Penelitian ini juga dilakukan untuk menguji kinerja dan nilai akurasi algoritme *K-Nearest Neighbors* 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, perumusan 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, antara lain pengertian dan pemahaman materi terkait *text mining*, analisis sentimen, Twitter, *crawling*, *preprocessing*, *labeling*, *CountVectorizer, K-Nearest Neighbors, modeling*, dan pengujian serta studi literatur.

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

Bagian ini berisi tentang sumber data penelitian, penerapan dan tahapan metode yang digunakan, serta rancangan pengujian yang akan dilakukan.

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

Bagian ini berisi mengenai lingkungan percobaan dari sistem yang dibuat, implementasi metode, flowchart tahapan metode, dan uraian algoritme pada proses, serta analisis pengujian sistem yang telah dibangun pada sisi akurasi.

### **BAB V: PENUTUP**

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