

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada akhir tahun 2019, WHO menyatakan adanya virus baru, yaitu virus corona (COVID-19) yang ditemukan pertama kalinya di Wuhan, China. Virus ini dapat menyebar dengan cepat melalui tetesan kecil (droplet) dari hidung atau mulut ketika batuk atau bersin. Pertama kali Indonesia melaporkan adanya kasus Covid-19 pada 2 Maret 2020 dengan 2 kasus, dan kasus positif covid-19 terus meningkat [1]. Semakin banyaknya kasus setiap harinya yang terus bertambah, pemerintah telah melakukan berbagai upaya untuk mencegah banyaknya penyebaran covid-19 yang terjadi.

Dengan perkembangan teknologi yang sangat pesat, berbagai informasi bisa kita dapatkan dengan cepat dan mudah. Salah satu media sosial yang dapat memberikan berbagai penyebaran suatu berita dengan cepat yaitu twitter. Twitter adalah situs microblogging di mana pengguna dapat memposting pesan waktu nyata yang dikenal sebagai tweet [2]. Semua pengguna dapat memposting hal apa saja dan kapan saja, sehingga membuat orang lain tau lebih cepat mengetahui berita berita terbaru. Twitter juga dilengkapi dengan fitur *trending*, dimana orang orang dapat dengan mudah mengetahui hal apa yang sedang dibicarakan oleh banyak orang. Berbagai upaya kebijakan pemerintah terkait dampak ekonomi akibat covid-19 pernah menjadi trending topik di twitter yang memicu berbagai tanggapan dari masyarakat mulai dari tanggapan positif, tanggapan netral, hingga tanggapan negatif.

Opini atau sentimen masyarakat di twitter tentu dapat digunakan sebagai kritik dan saran bagi pemerintah dalam melakukan kebijakan dan pemulihan dampak dampak yang telah terjadi akibat covid-19. Opini masyarakat di twitter tentunya masih acak dan belum terklasifikasi. Untuk mendapatkan data yang lebih tertata dan rapi tentunya perlu dilakukan beberapa proses. Proses untuk analisis sentimen, yang pertama yaitu proses pengambilan data dari suatu sumber seperti media sosial, berita online, ulasan pengguna,dll. Setelah dilakukan proses pengambilan data, kemudian dilakukan proses *preprocessing* data, kemudian akan dilakukan proses klasifikasi data mengenai dampak ekonomi akibat covid-19 dengan menggunakan algoritma k-nearest neighbor.

Algoritma K-Nearest Neighbor (KNN) merupakan algoritma yang sederhana untuk pengklasifikasian tanpa harus melakukan perhitungan secara kompleks [3]. Algoritma ini bekerja berdasarkan jarak terpendek dari sampel data training. Algoritma KNN memiliki performa yang cukup baik dalam melakukan klasifikasi analisis sentimen dibandingkan dengan algoritma yang lain [6].

Berdasarkan latar belakang diatas akan dilakukan penelitian dengan topik “Penerapan Algoritma KNN pada Analisis Sentimen Dampak Ekonomi Akibat Covid di Media Sosial Twitter.”

1.2 Rumusan Masalah

Perumusan masalah dalam Skripsi ini adalah:

1. Bagaimana mengklasifikasi sentiment masyarakat berdasarkan cuitan pengguna twitter, menggunakan algoritma K-Nearest Neighbor?

2. Bagaimana nilai akurasi dan hasil analisis sentimen dengan menggunakan algoritma K-Nearest Neighbor dalam mengklasifikasi dampak ekonomi akibat covid-19.

1.3 Batasan Masalah

Batasan masalah dalam Skripsi ini adalah:

1. Sistem ini menggunakan bahasa pemrograman Python.
2. Metode untuk mengklasifikasi dampak ekonomi akibat covid-19 menggunakan algoritma K-Nearest Neighbor.
3. Dataset diambil menggunakan teknik *crawling* data dari python dan API twitter dengan library tweepy.
4. Pengumpulan data pada twitter menggunakan *keyword* : PSBB, PHK, Ekonomi, KerugianCovid, ImpacOfCovid, DampakCovid, EkonomiPascaCovid, PemulihanEkonomi, KerugianEkonomiAkibatCovid". dan *hashtag* : #PSBB, #PHK, #Ekonomi, #KerugianCovid, #DampakCovid19, #PemulihanEkonomi.
5. Tweet yang digunakan hanya tweet berupa text, tidak mengandung gambar maupun emoji.
6. Data yang digunakan merupakan Tweet berbahasa Indonesia.
7. Hasil dari klasifikasi adalah sentimen positif, sentimen netral dan sentimen negatif dari dampak ekonomi akibat covid-19.
8. Data yang diambil mengenai dampak ekonomi akibat covid-19 adalah 7 hari, mulai tanggal 15 November 2021 sampai dengan 21 November 2021.

9. Jumlah dataset yang digunakan kurang lebih 3500 data, kemudian akan dibagi menjadi data training dan data testing.

1.4 Maksud dan Tujuan

Tujuan dari pembuatan Skripsi ini adalah:

1. Melakukan klasifikasi analisis sentimen masyarakat dengan menjadikan 3 kategori yaitu positif, negatif, dan netral.
2. Menghasilkan nilai akurasi dan hasil analisis sentimen dengan menggunakan algoritma K-Nearest Neighbor dalam mengklasifikasi dampak ekonomi akibat covid-19.

1.5 Manfaat Penulisan

a. Bagi Masyarakat

1. Diharapkan dapat bermanfaat sebagai tambahan referensi untuk penelitian berikutnya.
2. Diharapkan dapat digunakan sebagai evaluasi untuk mengetahui respon masyarakat mengenai dampak yang terjadi pada perekonomian di Indonesia akibat pandemic covid-19.
3. Dapat mengetahui dampak positif maupun negatif yang terjadi akibat covid-19.

b. Bagi Ilmu Pengetahuan

1. Mengetahui bagaimana implementasi algoritma K-Nearest Neighbor dalam pemrograman Python.
2. Mengetahui cara crawling data dari twitter.
3. Mengetahui cara menganalisis sentimen masyarakat dengan algoritma K-Nearest Neighbor.

4. Mengetahui cara kerja algoritma K-Nearest Neighbor dalam mengklasifikasi suatu kondisi.
5. Mengetahui nilai akurasi algoritma K-Nearest Neighbor dalam mengklasifikasi suatu kondisi.

1.6 Metode Penelitian

Jenis penelitian yang digunakan dalam kasus ini termasuk dalam penelitian eksperimen dan *case study* (studi kasus). Penelitian eksperimen adalah penelitian yang melibatkan beberapa faktor dan dilakukan untuk membuktikan suatu konsep. Dalam penelitian eksperimen terdapat hubungan sebab akibat. hubungan sebab akibat ini terjadi jika dampaknya merupakan efek dari korelasi, dampaknya menimbulkan efek dan juga kita bisa mencari penjelasan dari hubungan sebab akibat. Dalam penelitian eksperimen ini saya melakukan proses *crawling* data twitter kemudian melakukan klasifikasi menggunakan algoritma k-nearest neighbor. Dalam penelitian Studi kasus dilakukan dengan menggunakan *keyword* : PSBB, PHK, Ekonomi, KerugianCovid, ImpacOfCovid, DampakCovid, EkonomiPascaCovid, PemulihanEkonomi, KerugianEkonomiAkibatCovid". dan *hashtag* : #PSBB, #PHK, #Ekonomi, #KerugianCovid, #DampakCovid19, #PemulihanEkonomi. Kemudian dilakukan proses pengolahan data, setelah proses pengolahan data dilakukan analisis dengan pendekatan kuantitatif yang dijelaskan dengan hasil perhitungan tabel.

1.6.1 Metode Pengumpulan Data

Metode pengumpulan data yang akan dilakukan untuk penelitian ini, yaitu:

1. Mencari literatur yang bersangkutan dengan algoritma K-Nearest Neighbor, Dampak Ekonomi akibat covid-19 dan Analisis Sentimen di media social Twitter.
2. Mendapatkan dataset masyarakat tentang dampak ekonomi akibat covid—19 dengan melakukan *crawling* data menggunakan API twitter.

1.6.2 Metode Analisis

1. Metode K-Nearest Neighbor

K-Nearest Neighbor adalah metode yang digunakan dalam melakukan klasifikasi data dalam analisis dampak ekonomi akibat covid-19. Metode ini digunakan untuk mengklasifikasi data sebagai Dampak Positif, Dampak Netral, dan Dampak Negatif dari sentimen masyarakat. Metode ini melakukan klasifikasi terhadap objek yang jaraknya paling dekat dengan objek tersebut.

2. Mengklasifikasi Sentimen Masyarakat.

Melakukan klasifikasi sentimen masyarakat, dengan membagi dalam 3 kelompok, yaitu : sentimen positif, sentimen netral, dan sentimen negatif.

3. Menghitung akurasi dari pengujian data

Untuk menghitung ketepatan dalam akurasi klasifikasi menggunakan Confusion matrix.

1.6.3 Metode Implementasi

Implementasi algoritma K-Nearest Neighbor menggunakan Bahasa Pemrograman Python. Dengan menggunakan beberapa tools, seperti : NLTK, tweepy, Pandas, NumPy, Scikit-Learn, Seabon, dll.

1.6.4 Metode Pengujian

Pengujian dalam penelitian ini yaitu untuk mengukur ketepatan dalam akurasi klasifikasi menggunakan Confusion matrix.

1.7 Sistematika Penulisan

Sistematika dalam penulisan Skripsi ini dibagi menjadi lima bab, antara lain sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penulisan, metode penelitian dan sistematika penulisan penelitian.

BAB II LANDASAN TEORI

Pada bab ini menjelaskan mengenai teori yang digunakan sebagai dasar dalam analisis penelitian. Mencakup tinjauan pustaka meliputi Dampak ekonomi akibat covid-19, Twitter, Text Mining, Analisis Sentimen, Python, Library Python, Flowchart, Google Colaboratory, Preprocessing data, TF-IDF, SMOTE, K-Fold Cross Validation, Algoritma K-Nearest Neighbor, Confusion matrix.

BAB III ANALISIS DAN PERANCANGAN

Pada bab ini menjelaskan mengenai perancangan dan model cara kerja yang akan dilakukan di dalam studi kasus Penerapan Algoritma K-Nearest Neighbor pada analisis sentimen dampak ekonomi akibat covid di media social twitter.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan menjelaskan implementasi algoritma K-Nearest Neighbor dalam analisis sentimen masyarakat, mengklasifikasi analisis sentimen masyarakat serta menghitung nilai akurasi.

BAB V PENUTUP

Pada bab ini menjelaskan kesimpulan yang diperoleh dari hasil analisis penelitian ini, serta saran untuk pengembangan yang lebih baik lagi.

