

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis sentimen adalah proses memahami dan mengolah data tekstual secara otomatis untuk mendapatkan informasi. Analisis sentimen dilakukan untuk mendeteksi opini terhadap suatu subjek dan objek (misalnya individu, organisasi ataupun produk) dalam sebuah kumpulan data. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian ataupun aplikasi mengenai analisis sentimen berkembang pesat. (Alita & Isnain, 2020). Juga untuk mengetahui siapa pemberi opini yang banyak mendapatkan respon dari pengguna twitter. Analisis sentiment dapat digunakan untuk mengungkap opini publik terhadap suatu isu (Kurniawan & Apriliani, 2020). Tidak hanya itu, analisis sentimen merupakan cara mengumpulkan pendapat khalayak umum menggunakan jejaring sosial yang didalamnya terdapat mengandung pelayanan umum, serta isu terkini (Syarifuddin, 2020). Contohnya seperti yang sedang ramai saat ini yaitu virus corona.

Virus corona atau dikenal sebagai pandemi Covid-19 pertama kali di deteksi di kota wuhan Desember 2019 (Wikipedia, 2019). Penyebaran virus secara cepat melalui kontak fisik seperti droplets yang dikeluarkan selama berbicara, batuk, atau bersin (Dave dkk, 2020). Pada awal tahun 2020 virus ini menyebar ke berbagai negara, termasuk Indonesia. Tidak hanya itu, pengaruh yang ditimbulkan akibat virus corona berdampak hampir disemua aspek, seperti kehidupan masyarakat (Burhanuddin & Abdi, 2020).

Kehidupan masyarakat berubah semenjak adanya wabah pandemi corona terlebih sejak dilakukannya *social distancing* yang mana masyarakat diharuskan untuk menjaga jarak kontak fisik antara satu orang dengan yang lainnya, berdiam diri dirumah, menghindari kerumunan, selalu mencuci tangan hingga diharuskan untuk selalu menggunakan masker (Yanuarita dan Haryati, 2020). Banyak masyarakat yang mengutarakan pendapatnya melalui media sosial twitter dalam menanggapi dampak yang ditimbulkan akibat wabah virus corona (Kaur & Sharma, 2020).

Media sosial Twitter adalah salah satu sebagai media penghubung yang diminati oleh seluruh masyarakat di dunia. Hal ini dapat dilihat dari jumlah peningkatan pengguna Twitter yang tercatat di seluruh dunia salah satunya Indonesia (Darwis dkk, 2021). Twitter digunakan tidak hanya untuk bersosialisasi dan berkomunikasi namun juga digunakan untuk menyampaikan aspirasi dan mempresentasikan hal-hal yang sedang terjadi pada masyarakat (Astari dkk, 2020). Hal ini dapat dijadikan sebagai sebuah acuan untuk dilakukan analisis sentimen masyarakat terhadap dampak virus Corona. Analisis sentimen merupakan cara untuk melakukan proses klasifikasi terhadap dokumen-dokumen teks yang dilandasi dari opini dan pandangan dari masyarakat tertentu bisa dilakukan dengan proses memahami, mengolah hingga mengekstrak data bersifat dokumen teks secara otomatis (Hendriani & Sianturi, 2021). Hal ini dilakukan untuk mengetahui apakah opini maupun tanggapan tersebut bersifat positif, negatif atau netral. Salah satu metode yang digunakan dalam analisis sentimen yaitu metode Naïve Bayes Classifier (Ratino dkk, 2019).

Algoritma *Naive Bayes* merupakan salah satu algoritma klasifikasi yang banyak digunakan pada Data Mining ataupun Text Mining (Wijaya dkk, 2016). Salah satu metode klasifikasi untuk menentukan gambaran persepsi masyarakat di dalam Text Mining adalah *Naive Bayes Classifier* (Isnain dkk, 2020). Metode ini dipilih dikarenakan sangat cocok untuk *short* data teks. Kelebihan dari NBC adalah metode sederhana tetapi memiliki akurasi serta performansi yang tinggi dalam pengklasifikasian teks (Cahyono, 2017).

Penelitian berjudul perbandingan algoritma klasifikasi Naive Bayes dan SVM pada studi kasus pemberian penerima beasiswa ppa oleh Linawati, Safitri, Alfian, Pangesti, dan Winarto pada tahun 2020. Hasil dari penelitian ini yaitu komparasi antara algoritma Naive Bayes dan SVM (Support Vector Machine) yang dilakukan untuk mengklasifikasikan nilai akurasi tertinggi dengan 5 variabel dan jumlah data sebesar 122 dapat disimpulkan bahwa algoritma Naive Bayes memiliki tingkat akurasi lebih tinggi yaitu 90.90% dibandingkan dengan metode SVM yaitu 89.25%. Selanjutnya penelitian yang berjudul perbandingan algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam klasifikasi sms spam berbahasa Indonesia oleh Widyawati dan Sutanto, 2019. Dalam makalah ini, membandingkan

kinerja klasifikasi oleh Naïve Bayes dan Mendukung Mesin Vektor (SVM). Prosesnya meliputi *preprocessing* (tokenizing, stopwords, penghapusan, dan stemming) dari setiap set data pelatihan dan pengujian sebelum memproses informasi melalui Naïve Bayes dan SVM. Hasil dari pengolahan data sebanyak 1143 record (765 training dan 378 pengujian) menunjukkan bahwa kinerja Naïve Bayes melampaui SVM dalam hal penarikan (94%) dan Presisi (95%). Penelitian berjudul perbandingan metode Naïve Bayes dan support vector machine pada analisis sentimen twitter oleh Mujaddid Izzul Fikri, Trifebi Shina Sabrila, Yufis Azhar, 2020. Dalam penelitian ini dilakukan analisis sentimen terhadap tweets yang berhubungan dengan Universitas Muhammadiyah Malang (UMM) untuk mengetahui opini masyarakat mengenai kampus ini. Analisis dilakukan dengan mengklasifikasikan tweets yang berisi sentimen masyarakat mengenai UMM. Metode klasifikasi yang digunakan dalam penelitian ini adalah Naïve Bayes dan Support Vector Machine (SVM) dengan pembobotan menggunakan TF-IDF. Hasil komparasi kedua metode menunjukkan bahwa Naïve Bayes mendapatkan hasil akurasi yang lebih baik dari SVM dengan akurasi sebesar 73,65%. Penelitian berjudul Analisa Sentiment Untuk Opini Alumni Pada Perguruan Tinggi oleh Dharmendra, Saputra dan Pramaita, 2019. Penelitian ini melakukan pengelompokan opini dari alumni Perguruan Tinggi dengan menggunakan algoritma SVM dan NBC dimana dalam penelitian ini dibagi menjadi 3 komponen utama yaitu komponen input, sistem pengelompokan opini, serta komponen output. Opini yang akan diolah merupakan data berupa file opini format *.csv, yang selanjutnya melakukan proses pengelompokan terhadap opini. Kemudian sistem menghasilkan output berupa 3 jenis opini yaitu, opini positif, opini netral dan opini negative. Dari hasil pengujian secara umum menunjukkan akurasi dengan nilai akurasi tertinggi pada algoritma NBC mencapai 94,45%, sedangkan tingkat akurasi tertinggi pada algoritma SVM mencapai 75,76%.

Berdasarkan uraian diatas maka diperlukannya analisis sentimen untuk memisahkan mana yang merupakan opini berisi informasi yang bersifat positif, netral dan negatif terutama dalam aspek sosial. Penelitian ini bertujuan menganalisis opini publik untuk menemukan klasifikasi positif, negatif dan netral terhadap kehidupan masyarakat baik itu sebelum dan sesudah wabah pandemi

Corona. Adapun metode klasifikasi data yang digunakan oleh penulis yaitu *Naive Bayes Classifier* dikarenakan metode ini memiliki tingkat akurasi, penarikan dan presisi yang lebih baik daripada metode SVM.

1.2 Rumusan Masalah

Berdasarkan masalah yang telah diuraikan sebelumnya, maka didapat rumusan masalah dalam penelitian ini, yaitu

1. Bagaimana hasil penelitian kelas sentimen pendapat masyarakat terhadap dampak virus corona?
2. Bagaimana cara menghitung opini masyarakat pada media sosial twitter secara akurat?

1.3 Batasan Masalah

Adapun batasan pada penelitian ini adalah sebagai berikut:

1. Metode klasifikasi data yang digunakan yaitu *Naive Bayes Classifier*.
2. Media sosial yang digunakan yaitu *twitter*.
3. Data yang diambil berupa tweet berbahasa Indonesia.
4. Aspek yang dianalisis berupa aspek ekonomi dan sosial.
5. Klasifikasi dibagi menjadi tiga kelas, yaitu positif, negatif dan netral.
6. Dataset yang diambil yaitu tahun 2020 – 2021.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah Menganalisis sentimen dampak dari virus corona terhadap perekonomian pada media sosial twitter dengan menggunakan metode *Naive Bayes Classifier* dengan harapan agar memudahkan masyarakat menilai suatu opini dilihat dari berbagai persepektif atau sudut pandang.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut

1. Dapat membantu menyajikan informasi tentang opini yang bersifat positif, netral atau negatif.

2. Dapat mengetahui akurasi metode *Naïve Bayes Classifier* dalam menghasilkan perhitungan data positif, negatif dan netral terhadap opini masyarakat pada media sosial.
3. Sebagai acuan untuk dapat mengetahui seberapa besar pengaruh media sosial terhadap penilaian suatu informasi.