

BAB I

PENDAHULUAN

1.1 Latar Belakang

Corona Virus (Covid19) merupakan kelompok virus yang dapat menginfeksi sistem pernapasan manusia, virus ini menyebar dari partikel yang keluar dari mulut atau hidung saat batuk atau bersin karena penyebaran virus ini sangat cepat peningkatan orang terinfeksi pun dari berbagai negara meningkat seperti yang terjadi pada 6 oktober 2020 total kasus terinfeksi Covid-19 sebesar 4.223.094 kasus dan meningkat di setiap harinya (Kemkes, 2020). Pada 30 Januari 2020, WHO menyatakan wabah itu sebagai darurat kesehatan masyarakat yang membuat resah dunia (Astini, Sari 2020). Hal ini tentu saja menyita perhatian seluruh masyarakat dan pemerintahan indonesia, berbagai upaya pemerintah pun melakukan PSBB (Pembatasan Sosial Berskala Besar) dan PPKM (Pemberlakuan pembatasan kegiatan masyarakat), kegiatan belajar dan mengajar pun terkena dampak dari kebijakan ini pada bulan Maret 2020 pemerintah memberlakukan pembelajaran daring pada seluruh pelajar di Indonesia (Kominfo, 2021).

Pembelajaran daring ini melahirkan banyaknya opini karena timbul dari perubahan sistem yang ada sebelumnya dari banyaknya opini tersebut perlu dilakukan penggalan informasi di media sosial salah satunya di *Twitter*. Batas 140 karakter per *Tweet* memungkinkan pengguna memposting lebih ringkas dan bermakna daripada jejaring sosial dan blog lainnya. Hasil dari analisis sentimen ini juga dapat digunakan dan dijadikan bahan evaluasi bagi Kementerian Pendidikan dan Kebudayaan, dan seluruh peran di dunia pendidikan supaya dapat meningkatkan kualitas pengajaran yang efektif.

Naive Bayes Classifier menggunakan metode probabilistik dan statistik yang diusulkan oleh ilmuwan Inggris Thomas *Bayes*, fitur utama dari pengklasifikasi *Naive Bayes* ini adalah sangat kuat (naif). *Naive Bayes* adalah teknik untuk membangun model pengklasifikasi yang tidak rumit tetapi memiliki nilai akurasi yang tinggi dalam mengklasifikasi (Putra et al., 2020)

Penelitian dengan menggunakan algoritma NBC pernah dilakukan oleh Dwiraswati (2019) mencapai akurasi 80% dengan mengklasifikasikan 32.847 *tweet* terkait kolonoskopi, mammografi, atau *Pap smear* untuk memahami opini publik tentang alat skrining kanker.

Penelitian yang dilakukan oleh Iryana dan Adikara (2021) melakukan penelitian tentang analisis sentimen masyarakat Jakarta tentang MRT (*Mass Rapid Transit*) menggunakan data media sosial atau *Twitter* yang mengandung kata kunci "MRTJakarta" selama periode pengujian MRI publik dari 5 hingga 23 Maret 2019. *Tweet* yang diekstraksi adalah 1000 *tweet*, 70% untuk data latih dan 30% untuk pengujian. data. Dalam penelitian ini, *Naive Bayes* dapat memprediksi sentimen dari *tweet* yang dikumpulkan tentang minat publik terhadap MRT Jakarta dengan akurasi 75%.

Fikri, Sabrila, dan Azhar (2020) membandingkan algoritma *Naive Bayes* dan *Support Vector Machine* dengan menggunakan 2654 data *tweets*, menunjukkan bahwa algoritma *Naive Bayes* mencapai hasil keakuratan 73,45%, *presisi* 0,75, dan *recall* 0,75 yang lebih unggul sebesar dan sedangkan algoritma *Support Vector Machine* mendapatkan hasil keakuratan 70,20%, *presisi* 0,73, dan *recall* 0,71. Pada penelitian selanjutnya juga dibandingkan algoritma *Naive Bayes* dengan *K-Nearest Neighbor* oleh Devita, Herwanto, and Wibawa (2018) penelitian ini menggunakan 40 data jurnal dan menunjukkan bahwa algoritma *Naive Bayes* menghasilkan keakuratan yang cukup baik yaitu mempunyai *Accuracy* 70%, *Recall* 70%, *Precision* 70,9%, dan *Error* 30%, sedangkan *K-Nearest Neighbor* menghasilkan keakuratan 40%, *Recall* 40%, *Precision* 64,1%, dan *Error* 60%, karena metode tersebut tidak efektif dengan jumlah data latih yang sedikit. Mohammad Hamid Asnawi, Isal Firmansyah, Rafly Novian, Resa Septiani Pontoh (2021) melakukan penelitian mengklasifikasikan aplikasi media sosial dari Google Playstore yang terpopuler di dunia dan membandingkan dengan Algoritma *Naive Bayes*, *K-Nearest Neighbor* dan *Support Vector Machine* data yang digunakan dalam penelitian tersebut yaitu 2000 data didapatkan akurasi sebesar 78.96% pada algoritma *Naive Bayes*, dengan menggunakan algoritma *K-Nearest Neighbor* didapat nilai akurasi 50.23%, dan dengan menggunakan algoritma *Support Vector Machine* didapat nilai akurasi 75.29%.

Data yang diperoleh dari proses *Crawling* dengan kata kunci dalam pencarian *tweet* yaitu “Pembelajaran *Online*”, “Sekolah *Online*”, “Kuliah *Online*”, “Sekolah Daring” dan “Kuliah Daring” selama pandemi Covid-19 dan diklasifikasikan ke dalam tiga kategori sentimen yaitu positif, negatif, dan netral pada penelitian ini menghitung *accuracy*, *precision*, *recall*, dan *f1-score*. Penelitian ini menggunakan metode *Naive Bayes* yang diharapkan menghasilkan analisis data dengan tingkat keakuratan yang tinggi.

1.2 Rumusan Masalah

Rumusan masalah ini berlandaskan dari latar belakang dan permasalahan yang disampaikan, dan dapat disimpulkan di antaranya sebagai berikut:

1. Bagaimana pengaplikasian algoritma *Naive Bayes* pada analisis sentimen pembelajaran daring pada masa Covid-19?
2. Bagaimana keakuratan yang diperoleh dari hasil klasifikasi sentimen pembelajaran daring pada masa Covid-19 menggunakan metode *Naive Bayes*?

1.3 Tujuan Penelitian

Tujuan penelitian dapat di ambil dari rumusan masalah yang disampaikan, maka dapat disimpulkan beberapa tujuan penelitian di antaranya sebagai berikut :

1. Mengetahui pengaplikasian algoritma *Naive Bayes* pada analisis sentimen mengenai pembelajaran daring dimasa Covid-19.
2. Mengetahui tingkat keakuratan yang di capai dari analisis sentimen dengan metode *Naive Bayes*.

1.4 Manfaat

Harapan untuk penelitian ini dapat memberikan manfaat positif bagi kedua belah pihak, antara lain :

1. Membagikan ilmu tentang klasifikasi opini kuliah *online* selama masa Covid-19 dengan metode *Naive Bayes*.
2. Mengklasifikasikan sebagian besar opini/tanggapan ke dalam kelas positif, negatif dan netra

