

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertama kali masalah Novel Corona Virus atau yang lebih dikenal dengan COVID-19 ditemukan yaitu pada tanggal 31 Desember 2019 bertempat di Kota Wuhan, China. Pelaporan kasus tersebut tentunya membuat gempar seluruh dunia. Data dari seluruh dunia, yang bersumber dari WHO menyebutkan bahwa terdapat 14.348.858 kasus COVID-19 pada Juli 2020, dapat dipastikan jumlah tersebut akan terus banyak setiap harinya. Di Indonesia sendiri, menurut data dari Satgas COVID-19 tercatat sudah 132.816 kasus pada pertengahan Agustus 2020 [1]. Pada awalnya, penyebaran virus COVID-19 ini hanya berdampak besar pada sektor ekonomi yang semakin lesu, namun kini dampaknya juga menjalar ke sektor pendidikan juga. Langkah awal negara di dunia termasuk Indonesia dalam menghadapi masalah tersebut adalah dengan meliburkan sementara seluruh kegiatan pendidikan diseluruh jenjang. Selanjutnya pemerintah dan lembaga lembaga terkait, memiliki tugas penting yaitu mencari alternatif proses kegiatan pendidikan bagi seluruh siswa maupun mahasiswa yang tidak dapat melangsungkan kegiatan pendidikan secara normal. Alternatif kegiatan pembelajaran yang ditetapkan pemerintah pada saat COVID 19 adalah *Pembelajaran Jarak Jauh* pada seluruh jenjang [2].

Namun pada saat ini, dengan semakin menurunnya kasus COVID-19 di Indonesia, Pemerintah yang awalnya menerapkan alternatif *Pembelajaran Jarak Jauh* mulai kembali menjadi penerapan *Pembelajaran Tatap Muka* [3]. Pembelajaran tersebut bisa terlaksana karena sudah sesuai dengan Instruksi Menteri Dalam Negeri Nomor 41 Tahun 2021 perihal pemberlakuan pembatasan kegiatan masyarakat level 3, level 2 dan level 1, juga mengoptimalkan posko penanganan corona virus disease 2019 pada tingkat desa maupun kelurahan dalam proses pengendalian penyebaran corona virus disease 2019. Agar aktivitas pendidikan bisa kembali normal seperti sedia kala, maka usaha yang dilakukan Pemerintah yaitu meminta seluruh instansi pendidikan baik sekolah maupun perguruan tinggi untuk selalu memprioritaskan protokol kesehatan yang sudah ada [1].

Dengan kembali normalnya kegiatan pembelajaran di Indonesia, maka kegiatan perkuliahan pada seluruh perguruan tinggi juga akan kembali dilaksanakan dengan cara *Kuliah Offline*, yang sebelumnya dilaksanakan dengan cara *Kuliah Online*. Hal tersebut tentunya menuai berbagai komentar dari Masyarakat, khususnya para *Mahasiswa, Dosen, maupun Civitas Akademika* yang terlibat langsung dalam proses *Kuliah Offline*. Terdapat berbagai macam sumber yang bisa dijadikan untuk memahami sikap serta tolak ukur pemahaman masyarakat terhadap kebijakan ini. Salah satu sumber yang terpercaya yang bisa digunakan yaitu media sosial *twitter* [1]. Pada media sosial *twitter* terdapat berbagai macam komentar atau *tweet* yang dituliskan masyarakat, mulai dari komentar positif maupun negatif dalam menanggapi hal tersebut. *Tweet* tersebut dapat digunakan sebagai sumber opini dari masalah yang muncul ditengah tengah masyarakat. Namun terdapat berbagai masalah dalam penyerapan *tweet* tersebut yaitu, beragam kata, bahasa tidak baku, singkatan, maupun simbol simbol, sehingga menjadikan hal tersebut menjadi sulit jika harus menghasilkan mana kata yang positif maupun negatif. Dilakukan Analisis Sentimen untuk bisa mengklasifikasi berbagai komentar tersebut [3].

Text mining merupakan sebuah teknik yang biasanya digunakan untuk proses klasifikasi dokumen, klustering, analisis sentimen, ekstrak informasi dan pencarian informasi yang mana Text Mining adalah bagian dari Data Mining yang digunakan untuk mencari pola yang terbaik dari data teks yang jumlahnya besar [4]. Analisis sentimen merupakan suatu pembelajaran komputasi dengan cara mencari pendapat, komentar bahkan emosi seseorang dalam berkomentar terhadap sebuah kasus pada media sosial. Kemampuan analisis sentimen yaitu bisa merubah data yang tidak tersusun menjadi tersusun. Manfaat analisis sentimen begitu besar pada seluruh sektor, karena digunakan sebagai bahan evaluasi serta ide dalam menanggapi sebuah peristiwa, komentar maupun pernyataan yang sedang diperbincangkan berbagai pihak maupun individu [3]. *Term Frequency-Inverse Document Frequency (TF-IDF)* merupakan proses pembobotan pada nilai masing masing kata yang ada pada sebuah dokumen. Fungsi dari proses ini adalah memperhitungkan seberapa pengaruh sebuah *term* dalam suatu kasus. Pada penelitian ini digunakan algoritma *Naïve Bayes Classifier (NBC)* dan *Social Network Analysis. Naïve Bayes*

Classifier digunakan untuk mengklasifikasikan sebuah *tweet* kedalam kategori sentimen positif atau negatif. *Social Network Analysis* digunakan untuk seleksi fitur. Seleksi fitur merupakan proses penghapusan atribut yang kurang relevan pada topik penelitian, sehingga nantinya bisa meningkatkan akurasi pada klasifikasi.

Pada penelitian yang berjudul “*Perbandingan Algoritma Naïve Bayes, K-NN, dan SVM dalam Pengklasifikasian Sentimen Media Sosial*”. Dari penelitian ini diketahui bahwa algoritma Naïve Bayes Classifier memiliki nilai akurasi lebih tinggi sebesar 79.8% dibandingkan algoritma K-NN sebesar 50.23% dan SVM sebesar 75.29%, sehingga algoritma Naïve Bayes Classifier lebih tepat digunakan untuk mengklasifikasi ulasan aplikasi media sosial [5].

Selanjutnya pada penelitian yang berjudul “*Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn*”. Dari hasil pengujian diketahui metode Naïve Bayes memiliki nilai lebih tinggi sebesar 63.21%, sedangkan metode KNN sebesar 58.10%, dan didapatkan pula kecenderungan opini masyarakat di twitter condong positif, hal tersebut dapat dilihat dari jumlah opini positif sebesar 610 sedangkan negatif 488 [6].

Pada penelitian ini algoritma *Naïve Bayes Classifier (NBC)* digunakan karena kecepatan dalam pengerjaannya serta akurasi yang dihasilkan sangat tinggi apabila diperuntukkan pada data dengan jumlah besar serta beragam [7]. Algoritma *Naïve Bayes Classifier (NBC)* juga pada studi literatur yang saya dapatkan memiliki tingkat akurasi yang paling tinggi. Pada penelitian ini, *Social Network Analysis* digunakan sebagai model visual sebuah informasi perihal *Kuliah Offline*, dengan cara mengumpulkan kata (node) pada jaringan kemudian melakukan identifikasi pada term tersebut mana yang memiliki pengaruh paling penting sehingga bisa diketahui kelompok kelompok jaringan yang terbentuk berdasarkan informasi *Kuliah Offline* di twitter.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Berdasarkan latar belakang yang dijelaskan, dapat dirumuskan dalam permasalahan dalam penelitian ini yaitu :

Kebijakan Pemerintah *Pasca Covid 19* mengenai perpindahan *Kuliah Online* menjadi *Kuliah Offline* menuai banyak sekali respon atau sentimen dari masyarakat salah satunya melalui media sosial *twitter*. Respon yang disampaikan juga beragam, baik berupa respon positif maupun negatif. Sehingga diperlukan analisis sentimen untuk mengetahui apakah lebih banyak sentimen positif atau sentimen negatif dari masyarakat.

1.2.2 Metode Usulan

Penelitian analisis sentimen ini mengusulkan metode *Naïve Bayes Classifier (NBC)* sebagai algoritma pengklasifikasian sentimen, *Social Network Analysis* sebagai metode seleksi fitur dan *Confusion Matrix* sebagai metode untuk evaluasi modelnya.

1.2.3 Pertanyaan Penelitian

1. Mengetahui kelas sentimen *Kuliah Offline pasca Covid 19* berdasarkan data tweet menggunakan metode *Naïve Bayes classifier*?
2. Berapa nilai akurasi algoritma *Naïve Bayes* seleksi fitur *Social Network Analysis* ?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

1. Mengetahui kemampuan dari algoritma *Naïve Bayes* dalam menganalisis sentimen pada tweet *Kuliah Offline pasca Covid-19* pada media sosial *twitter*.
2. Mengetahui tingkat akurasi yang didapatkan dari analisis sentimen menggunakan metode klasifikasi *Naïve Bayes* dan seleksi fitur *Social Network Analysis*.

1.3.2 Manfaat

1. Memperoleh informasi tentang bagaimana opini masyarakat pengguna *twitter* terhadap kebijakan *Kuliah Offline Pasca Covid 19*.
2. Sebagai model pelaksanaan *Kuliah Offline Pasca Covid 19* kedepannya.

1.4 Batasan Masalah

Penentuan batasan masalah dilakukan agar tidak terjadi pelebaran maupun penyimpangan pembahasan pokok masalah penelitian sehingga penelitian menjadi lebih terfokus, lebih mudah dalam pembahasan, dan tujuan penelitian akan berhasil. Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Data yang dipakai merupakan *tweet* yang berbahasa Indonesia dengan kata kunci "*kuliah offline*" Pasca Covid 19.
2. Informasi yang disajikan yaitu berupa komentar positif dan negatif pada twitter terhadap kebijakan pelaksanaan *Kuliah Offline Pasca Covid 19*.
3. Data yang digunakan merupakan *tweet* yang diperoleh dari *twitter* berjumlah 1029 *tweet*. *Tweet* tersebut diambil mulai 1 September 2022 sampai 1 Oktober 2022 yang berkaitan dengan pelaksanaan *Kuliah Offline Pasca Covid 19*.
4. Menggunakan kamus *slangword* pada python yang didalamnya terdapat kata baku dan tidak baku.
5. Menggunakan library Sastrawi untuk proses *Stemming*
6. Sistem dapat melakukan proses klasifikasi data yang sudah di proses dari awal, apabila ada data baru, sistem tidak bisa melakukan klasifikasi pada data tersebut.

1.5 Sistematika Pembahasan

BAB I. PENDAHULUAN

Pada bab ini memuat tentang latar belakang masalah, perumusan masalah, tujuan dan manfaat, batasan masalah, sistematika pembahasan.

BAB II. TEORI PENUNJANG

Pada bab ini berisi uraian mengenai tinjauan pustaka dan kerangka teori yang relevan dan terkait dengan tema penelitian.

BAB III. METODE USULAN

Pada bab ini membahas mengenai rancangan dari sistem yang akan dibuat setelah melakukan pengumpulan data dan analisa masalah mulai dari kebutuhan software, hardware, rancangan atau pola arsitektur serta alur dari sistem yang akan digunakan.

BAB IV. HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil dan implementasi dari analisis sentimen tentang pengguna indihome dengan metode naïve bayes dan social network analysis.

BAB IV. PENUTUP

Bab ini merumuskan kesimpulan dari hasil penelitian yang dilakukan serta memberikan saran terhadap para pembaca.

DAFTAR PUSTAKA

Memuat daftar jurnal atau referensi yang digunakan sebagai acuan dalam mengerjakan penelitian.