

1. Pendahuluan

Latar Belakang

Bidang informasi dan komunikasi mendapat manfaat dari kemajuan teknologi yang terus berkembang. Penggunaan media sosial yang terus meningkat di masyarakat menjadi bukti akan hal ini. Salah satu platform digital yang saat ini telah merasuk ke dalam kehidupan sehari-hari adalah media sosial. Menurut hasil survei We Are Social [1], per Januari 2022, terdapat 191 juta pengguna media sosial aktif di Indonesia. Jumlah ini meningkat 12,35% dibandingkan dengan tahun lalu yang hanya berjumlah 171 juta orang. Bagi sebagian besar orang, media sosial dijadikan sebagai wadah untuk mencari informasi maupun menyampaikan pendapat. Kegiatan bermedia sosial dapat memberikan dampak positif maupun negatif bagi penggunanya. Dampak negatif akan terasa, ketika pengguna terlalu sering menggunakan media sosial, dimana pengguna akan merasa malas dan lebih mementingkan dirinya sendiri, sehingga akan berkurangnya rasa peduli terhadap sesama dan lingkungan sekitar [2]. Salah satu platform media sosial yang paling disukai di Indonesia adalah Twitter. Indonesia berada di peringkat kelima secara global dalam hal penggunaan Twitter pada tahun 2013, menurut situs resmi Kominfo [3], di bawah Amerika Serikat, Brasil, Jepang, dan Inggris. Penyampaian pendapat pada *twitter* tidak ditulis secara terstruktur sehingga menimbulkan kesulitan saat menganalisa informasi yang disampaikan ketika mengolah data dalam jumlah banyak. *Twitter* juga dapat dijadikan sebagai acuan dalam menganalisa sentimen masyarakat terhadap berbagai topik seperti tempat, situasi, tokoh publik dan lainnya.

Dengan 79 akun, pengguna Twitter yang menerima peringatan adalah yang terbanyak. Jumlah ini dikumpulkan antara tanggal 23 Februari dan 11 Maret 2021. UU ITE dibuat oleh pemerintah sebagai salah satu cara untuk mengatasi masalah ujaran kebencian. Sesuai dengan Ayat 2 Pasal 28 UU ITE, pengguna internet yang menyebarkan konten yang mengandung kebencian merupakan tindakan melanggar hukum. Surat Edaran (SE) Kapolri No. SE/06/X/2015 memperjelas definisi kasus ujaran kebencian serta jenis-jenis ujaran kebencian yang dapat diekspresikan di media sosial [4].

Gurinder Singh, PS Mishra, Muhammad Biki Hamzah, Najirah Umar, dan Cut Fiarni telah melakukan penelitian terkait dengan implementasi model *naïve bayes*. Dalam penelitiannya, model *naïve bayes* digunakan untuk melakukan klasifikasi analisis sentimen terhadap beberapa kasus yang ada, tidak hanya itu penggunaan model *naïve bayes* juga diterapkan dalam pengembangan sistem penambangan opini. Beberapa model *naïve bayes* yang digunakan dalam penelitian tersebut adalah *bernoulli naïve bayes*, *multinomial naïve bayes*, *gaussian naïve bayes*, dan *complement naïve bayes*. Dalam penerapannya, beberapa model *naïve bayes* tidak memberikan hasil yang maksimal dan tidak mendapatkan nilai akurasi yang cukup baik, sehingga beberapa peneliti mencoba untuk mengintergrasikan beberapa model *naïve bayes* dengan metode *machine learning* lainnya, dengan tujuan untuk dapat memaksimalkan nilai akurasi dari model *naïve bayes*.

Pada penelitian [5], menerapkan model dari *naïve bayes* untuk klasifikasi teks, tujuan dari penelitian ini adalah untuk memprediksi apakah sentimen artikel berita memiliki dampak positif atau negatif. Penelitian ini menggunakan dua pendekatan populer dari kategorisasi teks *naïve bayes* yaitu klasifikasi *bernoulli naïve bayes* dan klasifikasi *multinomial naïve bayes*. Penelitian ini juga bertujuan untuk mengidentifikasi nilai akurasi yang dihasilkan dari kedua model tersebut. Kemudian, nilai akurasi yang dihasilkan dari dua model *naïve bayes* tersebut dibandingkan untuk dapat disimpulkan model yang terbaik.

Tujuan dari penelitian ini, yang menggunakan [5] untuk menerapkan model *naïve bayes* untuk kategorisasi teks, adalah untuk menganalisis apakah sentimen cuitan berita akan memiliki efek yang baik atau buruk. Dalam penelitian ini, dua metode yang paling populer untuk kategorisasi teks *naïve bayes* - klasifikasi *naïve bayes multinomial* dan klasifikasi *naïve bayes bernoulli* - digunakan. Nilai akurasi yang dihasilkan oleh kedua model tersebut merupakan tujuan lain dari penelitian ini. Nilai akurasi dari kedua model *naïve bayes* tersebut kemudian dibandingkan untuk menentukan mana yang terbaik.

Pada [6], melakukan penelitian terkait analisis sentimen pada *twitter* dengan menggunakan teknik *machine learning* yaitu *logistic regression*, *naïve bayes*, dan *multinomial naïve bayes*. Teknik tersebut digunakan untuk mengklasifikasikan teks ke dalam beberapa kategori seperti positif, negatif, dan netral. Tujuan dari penelitian ini adalah untuk menguji ketiga teknik *machine learning* tersebut pada kumpulan data yang berbeda, yang didapatkan pada aplikasi *twitter* dan kemudian membandingkan hasil yang diperoleh untuk dapat disimpulkan teknik yang memiliki keakuratan lebih baik diantara ketiga teknik tersebut.

Pada [7], melakukan penelitian terkait dengan klasifikasi analisis sentimen pada ulasan film. Teknik yang diusulkan dalam penelitian ini adalah *multinomial naïve bayes* yang digunakan dalam klasifikasi analisis sentimen, yang kemudian teknik ini dikombinasikan dengan metode *adaboost* dengan tujuan untuk meningkatkan akurasi dari penggunaan *multinomial naïve bayes*. Pemilihan fitur *chi square* juga diterapkan untuk mengurangi jumlah fitur yang tidak relevan. Kesimpulan yang didapatkan dari penelitian ini adalah nilai akurasi pada klasifikasi analisis sentimen menggunakan algoritma *multinomial naïve bayes* mendapatkan nilai sebesar 81,39%, algoritma *multinomial naïve bayes* dengan penerapan *chi square* mendapatkan nilai sebesar 85,37%, dan hasil akhir akurasi

dari algoritma *multinomial naïve bayes* dengan penerapan *adaboost* dan seleksi fitur *chi square* adalah sebesar 87,74%.

Algoritma naïve bayes digunakan dalam [8] untuk melakukan analisis sentimen pada data yang diambil dari situs media sosial Twitter. Tujuan dari penelitian ini adalah untuk mengetahui pendekatan model naïve bayes yang ketika parameter frekuensi diterapkan dengan menggunakan TF-IDF, memiliki nilai akurasi yang maksimal. Model naïve bayes yang digunakan dalam penelitian ini antara lain model naïve bayes bernoulli, gaussian, komplemen, dan multinomial. Data yang dipartisi menjadi training dan test set diuji dengan menggunakan k-fold cross validation, dan tingkat akurasi kemudian dinilai dengan menggunakan confusion matrix untuk memilih model yang optimal.

Pada [9], menerapkan algoritma *multinomial naïve bayes* dalam membuat sistem penambahan opini untuk *non profit organization* (NPO) di Indonesia. Model yang digunakan dirancang untuk membagi opini *audiens* kedalam beberapa kategori seperti rasa terima kasih, penghargaan, keinginan mendorong, empati, menyarankan, berbagi, membutuhkan informasi lebih lanjut, niat perilaku, keterlibatan, persetujuan, dan ketidaksetujuan. Terdapat dua model yang disulkan dalam penelitian ini yaitu *binarized multinomial naïve bayes* yang digunakan dalam pengkategorian *caption* dan normal *multinomial naïve bayes* untuk pengkategorian komentar. Dari hasil pengujian, model yang diusulkan terbukti dalam pengkategorian *caption* dengan nilai *precision* sebesar 96,67% dan nilai *recall* sebesar 96,4%, sedangkan pengkategorian komentar menghasilkan nilai *precision* sebesar 83,3% dan nilai *recall* sebesar 80,3%.

Karena maraknya permasalahan mengenai ujaran kebencian pada media sosial Khususnya di Twitter, Pada penelitian penulis akan mengklasifikasi ujaran kebencian dengan mengimplementasikan model pada *naïve naves* untuk mendapatkan performansi yang paling optimal [10]. Model yang akan diukur meliputi *multinomial*, *gaussian* dan *bernoulli*. *Naive Bayes* sendiri dipilih karena memiliki performansi yang cukup baik untuk menyelesaikan banyak kasus dengan data yang besar. Oleh karena itu dengan mengimplementasikan model pada *naïve bayes* di harapkan dapat menentukan model pada *naïve bayes* untuk mengidentifikasi kalimat yang mengandung ujaran kebencian pada media sosial *twitter* dengan cepat dan efisien. Penelitian mengenai model *naïve bayes* masih terbilang sedikit oleh karena itu penulis memutuskan untuk mencoba meneliti mengenai model *naïve bayes* untuk klasifikasi ujaran kebencian, apakah model *naïve bayes* dapat mendeteksi ujaran kebencian dengan akurasi yang baik sehingga dapat di gunakan dalam deteksi ujaran kebencian dengan cepat dan efisien..

Topik dan Batasannya

Topik pada penelitian ini adalah melakukan analisis perbandingan model dalam classifier naïve bayes untuk mengidentifikasi ujaran kebencian dengan menggunakan model *Gaussian*, *Bernoulli* dan *Multinomial*. Keterbatasan dalam penelitian ini adalah masih sedikitnya referensi yang ada mengenai perbandingan model.

Tujuan

Penelitian ini memiliki tujuan untuk menganalisis ujaran kebencian dengan model *Naïve Bayes* sehingga dapat Mengetahui performansi optimal pada model *Naïve Bayes* terbaik

Organisasi Tulisan

Pada penelitian ini setelah pendahuluan, bagian yang dibahas adalah studi terkait dan hasil penelitian. Pada bagian selanjutnya akan membahas mengenai urutan sistem yang di bangun untuk penelitian. Pada bagian evaluasi menjelaskan mengenai hasil dari penelitian dan analisis hasil dengan beberapa skenario yang berbeda. Pada bagian terakhir berisi kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.