

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi akan terus berkembang bersamaan dengan kemajuan zaman. Begitu pula dengan data yang tersimpan di internet, maka akan semakin banyak. Data tersebut akan dilindungi oleh pihak-pihak terkait, agar tidak diambil dan disalahgunakan oleh pihak yang tidak bertanggung jawab. Kasus kebocoran data pernah terjadi beberapa kali di Indonesia. Contohnya pada kasus kebocoran data Tokopedia, Lazada, BPJS Kesehatan, dan yang baru-baru ini terjadi yaitu kebocoran data pada aplikasi PeduliLindungi.

PeduliLindungi merupakan sebuah aplikasi yang dibuat oleh instansi pemerintah guna membantu dalam melakukan pelacakan terhadap masyarakat yang terjangkit *Coronavirus Disease (COVID-19)* [1], untuk menghentikan penyebarannya. Aplikasi ini menyimpan data masyarakat yang sudah melakukan vaksinasi COVID-19 seperti Nama lengkap, Nomor induk kependudukan (NIK), Tanggal lahir, Kewarganegaraan, Email, dan Nomor telepon genggam. Selain itu aplikasi ini juga merupakan salah satu syarat untuk mengakses transportasi seperti pesawat, kereta, bus, dan kapal. Aplikasi ini bekerja dengan cara sistem meminta izin terlebih dahulu kepada pengguna untuk mengaktifkan data lokasi. Setelah data lokasi di Smartphone pengguna telah aktif, maka secara berkala aplikasi akan mengidentifikasi lokasi pengguna saat ini serta memberikan informasi terkait dengan zonasi penyebaran COVID-19 di lokasi tersebut [1].

Kemajuan teknologi informasi juga memberikan dampak positif kepada kehidupan manusia pada bidang komputer dan internet [2]. Banyaknya manfaat yang terdapat pada internet bawa transformasi yang sungguh besar terhadap bermacam aspek kehidupan manusia, seperti masyarakat mudah dalam mendapatkan informasi di internet dari berbagai sumber dengan waktu yang singkat. Serta dapat berkomunikasi menggunakan aplikasi media sosial semacam Facebook, WhatsApps, Twitter dan sebagainya [3]. Meskipun begitu, kemajuan teknologi informasi dapat disebut sebagai “pedang bermata dua” sebab selain

memberikan kemudahan terhadap kehidupan manusia, terdapat juga berbagai tindak kejahatan siber atau yang sering disebut “*cybercrime*” [4].

Untuk saat ini belum ada undang-undang khusus yang berkaitan dengan perlindungan data pribadi. Namun di Indonesia terdapat beberapa regulasi yang bersifat parsial dalam pengaturannya mengenai perlindungan data pribadi konsumen, contohnya pada Peraturan Menteri Komunikasi dan Informatika Republik Indonesia Nomor 20 Tahun 2016 tentang Perlindungan Data Pribadi Dalam Sistem Elektronik; Undang-Undang Nomor 19 Tahun 2016 tentang Informasi dan Transaksi Elektronik; dan Surat Edaran Otoritas Jasa Keuangan Nomor 14/SEOJK.07/2014 tentang Kerahasiaan dan keamanan Data dan/atau Informasi Pribadi Konsumen [5].

Pada umumnya, kebocoran data sering terjadi karena adanya celah kecil atau bug yang terdapat pada sistem keamanan di aplikasi tersebut. Atau juga dikarenakan pihak pengembang yang kurang memberikan keamanan ekstra pada aplikasi yang mereka kembangkan. Sehingga para pelaku kejahatan siber atau yang sering disebut *Hacker*, dengan mudah mengambil data-data tersebut untuk kepentingan mereka sendiri. Biasanya data-data yang mereka ambil akan dijual di situs terlarang seperti *Deep Web*, hal ini akan merugikan orang-orang yang datanya telah diperjual belikan. Karena bisa saja data mereka disalahgunakan untuk melakukan kejahatan yang mereka sendiri tidak lakukan.

Kasus seperti ini sangat meresahkan masyarakat, khususnya mereka yang sering melakukan kegiatan di internet. Menurut laporan *We Are Social* pada Januari 2021, jumlah pengguna internet di Indonesia mencapai 202,6 juta pengguna. Dari angka tersebut, terdapat 61,8 persen dari populasi di Indonesia yang memakai internet untuk mengakses media sosial [6]. Salah satu media sosial yang kerap diakses masyarakat yakni Twitter.

Twitter ialah layanan jejaring sosial serta microblogging yang membolehkan konsumennya untuk mengirim dan membaca teks pesan sampai 140 karakter [7]. Di twitter para pengguna dapat menulis mengenai bermacam topik hingga mangulas isu- isu yang tengah terjadi. Twitter sendiri dibuat pada bulan maret 2006 oleh Jack Dorsey, serta dirilis pada bulan Juli [7]. Sampai saat ini twitter

sudah jadi salah satu dari 10 web yang sangat sering dikunjungi di internet [7]. Dengan adanya layanan tersebut, membuat masyarakat menjadi lebih memilih meluangkan berbagai opininya melalui media sosial dibandingkan dengan menyampaikannya secara langsung. Opini-opini tersebutlah yang akan dimanfaatkan untuk menjadi sebuah informasi. Namun sebelum dimanfaatkan sebagai informasi, harus terlebih dahulu melalui proses analisis agar menjadi suatu informasi yang dapat membantu banyak pihak untuk mengambil suatu keputusan atau pilihan.

Pada penelitian sebelumnya [8] dalam menganalisis sentimen terhadap pemerintahan Joko Widodo menggunakan Algoritma *Naïve Bayes Classifier*, dapat menghasilkan nilai akurasi yang tinggi pada setiap sentimennya. Oleh karena itu, pada penelitian ini peneliti menggunakan metode klasifikasi *Naïve Bayes Classifier* untuk menganalisa sentimen data twitter terhadap kebocoran data pada aplikasi PeduliLindungi, agar bisa menjadi bahan masukan dan membantu para pengembang aplikasi untuk meningkatkan lagi keamanan dan perlindungan terhadap data pengguna pada aplikasi yang mereka kembangkan.

1.2 Perumusan masalah

Berdasar latar belakang masalah yang telah dikemukakan, maka permasalahan yang dapat dirumuskan yaitu sebagai berikut:

1. Berapa tingkat akurasi dari hasil analisis sentimen data twitter terhadap kebocoran data pada aplikasi PeduliLindungi-menggunakan algoritma *Naïve Bayes Classifier*?
2. Bagaimana hasil dari analisis data twitter terhadap kebocoran data pada aplikasi PeduliLindungi menggunakan algoritma *Naïve Bayes Classifier*?
3. Bagaimana respons masyarakat terhadap kebocoran data pada aplikasi PeduliLindungi dari media sosial twitter?

1.3 Batasan Penelitian

Berikut merupakan batasan masalah dari penelitian ini:

1. Sumber opini yang digunakan berasal dari media sosial twitter dan berbahasa Indonesia.

2. Media yang digunakan untuk mengambil data *tweet* pada twitter yaitu menggunakan *Twitter API*.
3. Analisis sentimen yang dilakukan hanya mencakup kelas positif, netral, dan negatif.
4. Sentimen twitter hanya sebatas respons masyarakat mengenai kebocoran data pada aplikasi PeduliLindungi.
5. Analisis sentimen yang dilakukan menggunakan algoritma *Naïve Bayes Classifier* dengan menggunakan bahasa pemrograman *Python* dan Platform *Google Collaboratory*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini yaitu sebagai berikut:

1. Mengetahui tingkat akurasi yang didapatkan dari analisis sentimen data twitter terhadap kebocoran data pada aplikasi PeduliLindungi menggunakan algoritma *Naïve Bayes Classifier*.
2. Mengetahui hasil dari analisis sentimen data twitter terhadap kebocoran data pada aplikasi PeduliLindungi menggunakan algoritma *Naïve Bayes Classifier*.
3. Mengetahui respons masyarakat terhadap kebocoran data pada aplikasi PeduliLindungi dari media sosial twitter.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat bagi para pengembang aplikasi terhadap keamanan dan perlindungan pada data para pengguna aplikasi yang mereka kembangkan. Sehingga tidak ada lagi kasus kebocoran data yang dapat merugikan para pengguna aplikasi yang datanya bocor. Serta membantu memberikan *feedback* mengenai keamanan dan perlindungan data pengguna aplikasi PeduliLindungi kepada pihak pengembang aplikasi.

1.6 Metode Penelitian

Berikut ini adalah beberapa tahapan yang dilakukan dalam memperoleh data-data yang akan digunakan.

1.6.1 Metode Pengumpulan Data

Penelitian ini menggunakan *tweet* yang berasal dari media sosial Twitter sebagai data yang akan digunakan. Dengan menggunakan *Twitter Application Programming Interface (Twitter API)* sebagai perantara dalam mengambil data, sehingga data tersebut dapat diolah menjadi informasi yang dibutuhkan.

1.6.2 Metode Study Literature

Merupakan teknik pengumpulan data dari berbagai sumber seperti jurnal ilmiah, paper, buku, maupun situs internet terpercaya yang berkaitan langsung dengan penelitian yang sedang dilakukan. Sehingga dapat membantu peneliti dalam menyelesaikan masalah penelitian.

1.6.3 Metode Analisis

Metode analisis yang digunakan pada penelitian ini adalah algoritma *Naïve Bayes Classifier*, dengan melakukan crawling tweet menggunakan *Twitter API*. Pada proses pengolahan datanya akan dibantu dengan menggunakan bahasa pemrograman *Python* dan *Platform Google Collaboratory*.

1.6.4 Metode Testing

Metode testing yang digunakan pada penelitian ini yaitu menggunakan *Confusion Matrix* dalam pengujiannya serta untuk mencari nilai akurasi pada algoritma *Naïve Bayes Classifier*.

1.7 Sistematika Penulisan

Sistematika penulisan ini dibuat bertujuan untuk mempermudah dalam pembuatan laporan penelitian yang dirangkum dalam beberapa sub-bab sebagai berikut:

BAB I: PENDAHULUAN

Bab ini menjelaskan mengenai latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Bab ini menjelaskan mengenai landasan teori yang digunakan dalam menganalisis sentimen pada masalah penelitian yang sedang dilakukan. Selain itu juga untuk menguatkan teori-teori yang digunakan dari berbagai sumber ahli yang dibahas secara detail.

BAB III: METODE PENELITIAN

Bab ini menjelaskan mengenai alat dan bahan yang digunakan dalam melakukan penelitian. Metode yang digunakan oleh peneliti yaitu menggunakan algoritma *Naïve Bayes Classifier*, serta tahap preprocessing data yang dilakukan.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil analisis yang telah dilakukan dalam penelitian mengenai analisis sentimen data twitter terhadap kebocoran data pada aplikasi PeduliLindungi menggunakan algoritma *Naïve Bayes Classifier*.

BAB V: KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang telah didapatkan dalam penelitian serta saran yang diharapkan dapat diterapkan untuk pengembangan lebih lanjut pada nantinya.