

# BAB 1

## PENDAHULUAN

### 1.1 LATAR BELAKANG

Menurut Kamus Besar Bahasa Indonesia, *hoax* merupakan berita dan reaksi negatif terhadap informasi yang tidak benar dan tersebar luas secara daring, khususnya di media sosial. *Hoax* merupakan usaha bisnis atau sebagai sarana mengumpulkan dan menganalisis informasi untuk menangani situasi yang jelas salah dengan benar. Dengan maksud mencegah penerima informasi menerima informasi yang tidak akurat atau menyesatkan tentang situasi tertentu, informasi tersebut diteruskan dengan jelas oleh orang yang memberikannya.

Informasi *hoax* tersebar luas karena masyarakat umum dapat dengan mudah mengungkapkan pendapat mereka tentang informasi apa pun yang mereka dapatkan, dan karena masyarakat umum memiliki ketergantungan pada sumber mana pun yang secara konsisten menyebarkan informasi yang akurat. Informasi *hoax* juga dapat mempengaruhi kesehatan mental, dalam sebuah studi di *University of California San Francisco*, para psikolog sepakat bahwa *hoax* bisa memberikan dampak buruk bagi kesehatan mental, seperti *Post-Traumatic Stress Syndrome* (PTSD), menimbulkan kecemasan sampai kekerasan.

Tujuan dari informasi *hoax* ini adalah untuk mempengaruhi persepsi, membentuk opini, dan menciptakan sudut pandang yang menantang kepercayaan orang-orang yang menggunakan media sosial dan internet untuk mengakses informasi. Dari segi epistemologi tradisional, *hoax* lahir dengan reputasi dapat dipercaya dan memberikan efek positif bagi masyarakat. Akibatnya, setiap informasi tentang *hoax* tidak dapat dimasukkan dalam kategori informasi yang dapat dipercaya [1].

Saat ini, perkembangan internet telah membuat setiap orang dapat mengaksesnya dengan mudah. Selain itu, menjamurnya *smartphone* telah memudahkan pengguna untuk menggunakan internet. Penggunaan internet saat ini sudah dikenal mulai dari orang tua hingga anak kecil, media internet dikenal tidak terbatas waktu lagi. Baik wilayah maupun batas lain yang terpengaruh olehnya. Sehubungan dengan hal tersebut, pengguna media

sosial tentunya mengalami dampak positif dan negatif. Menyadari bahwa terdapat perbedaan dalam setiap peraturan yang berkaitan dengan kekayaan intelektual yang diterapkan di setiap negara. Oleh karena itu, ketika sesuatu dapat digunakan di suatu negara tanpa batasan, saat itulah situasinya menjadi pelanggaran hukum bagi negara tersebut.

Di era globalisasi ini, penyebaran *hoax* sebagian besar disebabkan oleh konsumsi media massa dan penyebaran informasi teknologi. Minimnya informasi dan data menjadi salah satu penyebab utama berita *hoax*. Ini karena fanatisme berdasarkan tokoh atau golongan menyebabkan banyak orang disesatkan oleh penyebab lain dari berita *hoax*. Semakin berkembangnya informasi di tengah masyarakat, maka berita *hoax* juga sangat cepat menyebar, maka dari permasalahan tersebut kami menawarkan sebuah inovasi Sistem Pendeteksi *Hoax* Pada Berita Seputar Pemerintahan dengan mengedepankan berita yang tersebar di banyak *website*. Dengan inovasi ini dapat membantu *user* dalam melihat kebenaran informasi yang diinginkan

*Hoax* sangat berbahaya dan berpotensi memecah belah masyarakat. Maka dari itu, sangat esensial untuk mendalami ciri-ciri *hoax* seperti yang telah dijelaskan sebelumnya. Berita tidak benar ini dapat diketahui dari sumber yang problematis, menghasut kebencian, dan tidak netral. *Hoax* biasanya mengeksploitasi fanatisme agama, memiliki tajuk utama yang provokatif, berusaha menjadi viral, dan sering kali memanipulasi foto dan keterangan. *Hoax* berdampak signifikan terhadap kesehatan mental masyarakat, menyebabkan kecemasan, kepanikan berlebihan, ketakutan ekstrem, dan mudah tersinggung [2].

Ada kemungkinan informasi akan dirilis kemudian atau diproduksi nanti dengan memasukkan pendapat pribadi, yang mungkin juga mengubahnya dari informasi yang awalnya faktual menjadi informasi yang *hoax*. Saat membuat produk informasi, seseorang harus mencari data yang faktual dan berkorelasi secara konsisten sehingga orang yang membaca informasi tersebut dapat memahaminya dengan jelas. Ini adalah kesalahan umum di antara oknum yang berbagi dan menciptakan produk informasi.

Perkembangan teknologi terus melejit seiring bertambahnya waktu, kebutuhan manusia pun semakin hari semakin kompleks tak terbantahkan juga kebutuhan tentang berita yang ramai dicari, namun sayangnya tidak semua berita yang disajikan itu benar. Seringkali ditemukan berita *hoax* yang beredar di tengah-tengah masyarakat. Berita *hoax*

ini umumnya tersebar di *website* maupun media sosial yang sulit terlepas dari kehidupan masyarakat saat ini.

Ada kategori berita yang bisa ditemukan di portal agar orang lebih cepat menemukan berita yang mereka cari. Namun klasifikasi jenis berita ini sangat terbatas, jadi jika seseorang ingin menggunakan klasifikasi yang lebih spesifik, mereka perlu melakukannya berdasarkan petunjuk dengan menentukan berita dari klasifikasi tersebut dan mengidentifikasi sub kategori yang lebih spesifik. Ini sulit karena fakta bahwa satu berita akan bertahan setidaknya selama satu hari, dan masing-masing kabar akan mempunyai karakteristik mirip. Maka dari itu, metode analisis data adalah langkah penting yang harus diambil.

Proses klasifikasi merupakan bagian dari sekian banyaknya bidang dalam data *mining*, khususnya *text mining*. Proses klasifikasi menghasilkan kegiatan mempelajari dan menganalisis tulisan yang telah direklasifikasi untuk membuat model yang berfungsi untuk mengklasifikasikan tulisan baru yang belum mempunyai kategori. Klasifikasi adalah metode mengkategorikan atau mengklasifikasikan sebuah tulisan berdasarkan ciri-cirinya [3].

Biasanya pelabelan pada metode pengklasifikasian tengah dilaksanakan dengan tangan tanpa bantuan aplikasi tambahan yang dilakukan oleh ahli dan disusun berdasarkan dataset yang berukuran besar. Akibatnya, dalam penelitian ini akan dikembangkan proses klasifikasi yang akan dipakai pada metode/proses pelabelan secara spontan dengan jumlah yang besar untuk mengambil alih proses pelabelan manual. *Naïve Bayes Classifier* (NBC) merupakan teknik klasifikasi yang paling umum dipakai pada klasifikasi data. Algoritma *Naïve Bayes* digunakan untuk menentukan nilai kemungkinan terbesar untuk mengkategorikan data uji. Dalam penelitian ini, menggunakan data dari situs *web* yang terkait dengan pemerintahan Indonesia [4].

Klasifikasi *Bayesian* merupakan pengelompokan terstruktur yang bisa dipakai dalam memprediksi peluang kedudukan dalam kelas. Teorema *Bayes* mendasari klasifikasi *Bayesian*, yang yang mampu mengklasifikasikan mirip rumitnya jaringan saraf. Menurut Kusriani, klasifikasi *Bayesian* 2009 sudah pasti mempunyai tingkat akurasi dan kecepatan yang besar bila diterapkan pada database data yang besar [5].

Berdasarkan latar belakang, maka diberikan solusi sebuah *website* yang dirancang untuk mendeteksi berita *hoax* yang tersebar di tengah-tengah masyarakat. *Website* ini dirancang dengan memanfaatkan algoritma *naïve bayes*, yang mana algoritma *naïve bayes* ini dimanfaatkan dalam memperkirakan kemungkinan dari beberapa tingkat variabel target. Keuntungan dari menggunakan algoritma *naïve bayes* yaitu bahwa algoritma *naïve bayes* hanya membutuhkan sedikit data uji guna menetapkan perkiraan batasan yang diperlukan didalam proses pengklasifikasian [6].

## 1.2 RUMUSAN MASALAH

Berikut adalah rumusan masalah dari penelitian ini:

- 1) Bagaimana tingkat keakuratan dan kinerja *naïve bayes* dalam penelitian ini?
- 2) Bagaimana algoritma *naïve bayes* dapat mempermudah *user*?

## 1.3 BATASAN MASALAH

Berikut adalah batasan masalah dari penelitian ini:

- 1) Berita pemerintahan Indonesia yang bersumber dari *website* pemberitaan Indonesia yaitu *kompas.com*, *detik.com*, *liputan6.com*, dan *TribunNews.com* untuk berita benar, dan *turnbackhoax.id* untuk berita *hoax*.
- 2) Pemberitaan Indonesia yang mengangkat tentang pemerintahan, korupsi, partai, ekonomi, Jokowi, dan pemilu.

## 1.4 TUJUAN

Berikut tujuan dari penelitian ini:

- 1) Untuk mengetahui keakuratan serta kinerja sebagai klasifikasi menggunakan metode *naïve bayes*.
- 2) Mempermudah *user* dalam mengetahui berita, sehingga tidak terjadi simpang siur pada berita yang belum jelas kebenarannya, dengan cara merancang sebuah aplikasi berbasis *website*.

## **1.5 MANFAAT**

Penulis mengharapkan penelitian yang dilakukan dapat mempermudah pengguna ketika mengakses berita yang meragukan sehingga dapat memilahnya dengan bantuan aplikasi ini. Dengan demikian dapat juga guna mendapatkan ketepatan akurasi dan berupa klasifikasi berita yang bermanfaat untuk memperoleh hasil klasifikasi berita berdasarkan kategori yang dikehendaki.

## **1.6 SISTEMATIKA PENULISAN**

Penelitian ini dibagi menjadi beberapa bab. Latar belakang, rumusan masalah, manfaat dan tujuan penelitian, definisi masalah, dan sistematika penulisan semuanya tercakup dalam Bab 1. Bab kedua membahas konsep berita, penambahan teks, metode penelitian seperti dataset yang diperlukan, jalur penelitian seperti pengumpulan dataset, pengujian berita, parameter kinerja sistem, dan prosedur pembuatan situs web, yang dibahas di bab ketiga. Bab 4 membahas simulasi sistem dan hasil analisis berbasis simulasi. Bab 5 menjelaskan kesimpulan dan saran untuk pengembangan tesis di masa mendatang.