

BAB I PENDAHULUAN

1.1 Latar Belakang

Hoax adalah informasi atau berita yang belum terverifikasi atau bukan fakta yang sebenarnya.[1] Penyebaran berita *hoax* menjadi tantangan terbesar dalam ekosistem informasi. Berita dapat dengan mudah diakses. Tetapi dengan banyaknya berita yang tersedia, sulit untuk memvalidasi berita fakta dan berita *hoax*. Berita *hoax* menjadi masalah besar atau menjadi masalah serius karena berita *hoax* dapat menyebabkan kebingungan, dan kerugian.

Banyaknya berita sering menarik minat untuk membacanya, tetapi kadang terlalu banyaknya berita tersebut membuat pembaca susah mendapatkan informasi yang terpercaya. Berita *hoax* merupakan kumpulan kata atau kalimat yang mengandung informasi yang tidak benar yang berupaya untuk membohongi atau mengarahkan pembaca atau pendengarnya agar mendukung atau percaya dengan isi beritanya. Penyebar berita *hoax* umumnya mengetahui bahwa berita yang disebarkan tidak benar.[2]

Menurut *overview* yang telah dilakukan oleh Lembaga Asosiasi Penyelenggara Jasa Web Indonesia (APJII) pada tahun 2017 dijelaskan bahwa penetrasi pengguna *web* Indonesia mencapai 143,26 juta jiwa atau 54,56% *add up to* populasi penduduk Indonesia 262 juta orang.[3] Di lansir situs CNN Indonesia, menurut Kementerian Komunikasi dan Informatika di Indonesia. Dilaporkan ada hingga 800.000 situs yang menyebarkan berita bohong dan ujar kebencian.[4] Setidaknya 30% hingga 60% Masyarakat Indonesia terpapar *scam* saat mengakses dan berkomunikasi melalui dunia maya. Sementara hanya 21% hingga 36% saja yang dapat mengenali atau mendeteksi adanya *deception*. Sebagian besar *scam* yang ditemukan terkait dengan isu politik, kesehatan, dan agama.[5]

Hoax yaitu informasi yang dibuat-buat untuk menutupi berita yang sebenarnya. *Hoax* diartikan sebagai upaya untuk memutar balikkan fakta menggunakan informasi yang seolah-olah meyakinkan, akan tetapi tidak dapat diverifikasi kebenarannya. Berita *hoax* dibuat menyerupai kejadian yang sebenarnya tetapi point-point penting akan di rubah oleh oknum yang membuat berita *hoax* tersebut. Berita *hoax* masih tersebar luas dan memiliki resiko yang besar untuk memanipulasi pola pikir. Sistem deteksi berita, terutama di situs publik atau pemerintah, sangat penting. Namun, dalam penerapannya, penting juga untuk menemukan metode terbaik dalam menentukan apakah suatu berita *hoax* atau fakta.[6]

Deception seringkali sengaja di buat dan dibagikan sehingga dapat menyebar lebih cepat. Informasi yang didapatkan dari berita *hoax* tentunya dapat memperngaruhi pola pikir karena menimbulkan keraguan dan kebingungan terhadap informasi yang di terima, serta mampu merusak citra individu dan kelompok yang berkaitan. Judul yang digunakan oleh berita *hoax* sering kali bersifat sensasional dan provokatif. Hal tersebut sengaja di buat untuk menarik minat dan keingintahuan pembaca.[5]

Dalam penelitian ini akan membangun sistem deteksi berita *hoax*, penelitian ini berfokus membangun aplikasi yang mampu memverifikasi berita fakta dan *hoax* dengan menggunakan metode *Machine Learning* algoritma *Support Vector Machine*. Sumber data yang digunakan dalam penelitian ini berasal dari berbagai portal berita seperti Detik.com, Liputan6.com, CNNIndonesia.com, Riaupos.jawapos.com, Riauterkini.com, prokopim.bengkaliskab.go.id, Pikiran-rakyat.com, idntimes.com dan situs verifikasi fakta seperti *Trunbackhoax.id*, *Cekfakta.com*. Data dikumpulkan melalui teknik web *scrapping*, kemudian diklasifikasikan menggunakan algoritma *Support Vector Machine* untuk membedakan antara berita fakta dan berita *hoax*. Hasil dari penelitian ini berupa sebuah aplikasi deteksi berita dan menyediakan berita dari berbagai sumber di dalam satu aplikasi

1.2 Permasalahan

Dari latar belakang yang telah dirumuskan diatas menjelaskan bahwa dengan banyaknya informasi yang tersebar, sulit untuk membedakan antara berita *hoax* dan fakta. Berita *hoax* seringkali memanfaatkan judul sensasional untuk menarik perhatian, meskipun informasi yang disampaikan tidak benar. Oleh karena itu, permasalahan yang ingin dipecahkan adalah bagaimana membangun aplikasi yang dapat mengidentifikasi berita *hoax* dan fakta dengan menggunakan algoritma *Support Vector Machine* (SVM). Aplikasi ini menggunakan data yang dikumpulkan dari beberapa portal berita terpercaya dan situs verifikasi fakta. Sistem yang dikembangkan akan memproses teks berita untuk melakukan klasifikasi apakah berita tersebut termasuk *hoax* atau fakta.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut:

1. Sumber data yang digunakan ada 10 sumber pada portal berita dan situs verifikasi fakta. Untuk sumber pada portal berita ada beberapa situs sumber seperti Detik.com, Liputan6.com, CNNIndonesia.com, Riaupos.jawapos.com, Riauterkini.com, Prokopim.bengkaliskab.go.id, Pikiran-rakyat.com, idntimes.com dan situs verifikasi fakta seperti *Trunbackhoax.id*, *Cekfakta.com*.
2. Sistem deteksi berita *hoax* menggunakan algoritma *Support Vector Machine* (SVM) untuk klasifikasi berita berdasarkan teks. Model berkerja dalam teks berita.

1.4 Tujuan

Penelitian ini bertujuan untuk membuat sistem deteksi berita berbasis *Machine Learning* menggunakan algoritma *Support Vector Machine* (SVM) untuk mengidentifikasi berita *hoax* dan fakta berdasarkan teks berita yang dikumpulkan dari beberapa sumber seperti Detik.com, Liputan6.com, CNNIndonesia.com, Riaupos.jawapos.com, Riauterkini.com, Prokopim.bengkaliskab.go.id, Pikiran-rakyat.com, idntimes.com dan situs

verifikasi fakta seperti *Trunbackhoax.id*, *Cekfakta.com*. Membangun aplikasi *mobile* untuk mendapatkan berita dari beberapa situs berita dan beragam di dalam satu aplikasi.

1.5 Manfaat

Manfaat dari penelitian ini adalah sebagai berikut:

1. Dapat membantu dalam verifikasi berita, aplikasi ini diharapkan dapat mempercepat proses verifikasi berita
2. Dapat memudahkan pengguna dalam mengakses berita dari beberapa sumber di dalam satu aplikasi.