

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Di era globalisasi ini, pertukaran informasi adalah hal yang sering terjadi dalam sehari-hari. Informasi tersebut diperlukan untuk mengambil strategi juga keputusan dari pendapat orang lain. Opini yang orang lain miliki tentunya memberikan dampak dan kontribusi akan cerminan topik yang orang tersebut bicarakan [1].

Media sosial memudahkan manusia untuk berinteraksi satu sama lain, beropini, juga bertukar informasi. Akses ke jejaring sosial yang cukup mudah dan hanya perlu menggunakan ponsel pintar, kita bisa mencari hal yang sedang terjadi di dunia, dengan mengakses berita melalui jejaring sosial media. Akibat adanya sosial media ini memunculkan adanya perbedaan pendapat masyarakat yang tentunya dapat berefek fakta maupun *hoax* tergantung akan topik yang dibahas [2].

Saat ini banyaknya pertukaran akan informasi dan *data* melalui online media tidak hanya digunakan oleh situs-situs berita yang sudah dikenali oleh kebanyakan orang-orang, namun oleh siapapun pemakai media sosial dapat ikut berperan dalam penyebaran suatu *data* maupun informasi. Sayangnya banyak sekali *data* dan informasi yang disebarkan secara individu maupun secara berkelompok lebih banyak yang kemungkinan besar bertolak belakang fakta atau teindikasi *hoax*. *Hoax* merupakan

berita atau informasi yang berisi *data* yang belum diketahui kebenarannya atau yang sebenarnya tidak terjadi [3].

Dampak yang disebabkan oleh berita-berita *hoax* merupakan dampak yang tidak bisa disadari disaat itu juga, karena informasi *hoax* akan mendistorsi pembaca *hoax* tersebut. Berita yang bersifat *hoax* akan diserap dan bila tidak berjaga-jaga maka akan mempengaruhi cara berpikir pembacanya. Dan ini sangat berbahaya bila persebaran *hoax* terjadi dalam kehidupan sehari-hari[4].

Berdasarkan latar belakang maraknya penyebaran berita *hoax* pada *Social Media twitter*, maka akan dibuat penelitian dengan judul “Analisa Deteksi *Hoax* Menggunakan *Algoritma Support Vector Machine, Naive Bayes, Random Forest dan K-Nearest Neighbor* Mengenai Berita Vaksin Covid-19 pada Twitter”.

1.2 Batasan Masalah

Batasan-batasan masalah dalam pengerjaan penelitian ini yaitu :

- a. *Data* sentimen mengenai berita vaksin covid-19 pada *twitter*.
- b. *Data-data* sentimen yang akan diklasifikasikan ke dalam dua kelas yaitu *hoax* dan *non hoax*.
- c. Penelitian ini melakukan analisa terhadap *Tweet* dengan Bahasa Indonesia.
- d. Penelitian ini menggunakan *Algoritma Support Vector Machine, Naive Bayes, Random Forest dan K-Nearest Neighbor* untuk analisa deteksi *hoax*.

1.3 Rumusan Masalah

Berdasarkan latar belakang masalah maka dapat dirumuskan masalah yang akan dibahas yaitu, Bagaimana tingkat keefektifan algoritma *Support Vector Machine*, *Naïve Beyes*, *K-NN*, *Random Forest* dalam Sistem Deteksi *Hoax* menggunakan bahasa Indonesia?

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan yang ingin dicapai yaitu membangun model deteksi berita *hoax* dari media sosial Twitter dengan metode *Support Vector Machine*, *Naïve Beyes*, *K-Nearest Neighbor*, *Random Forest*.

1.4.2 Manfaat Penelitian

Manfaat penelitian ini yaitu memberikan pihak-pihak pengambil kebijakan terkait dengan efek penyebaran berita *hoax* di masyarakat. Hasil penelitian ini diharapkan bisa mengkonter isu-isu negatif terkait dengan dampak penyebaran *hoax* di masyarakat, terutama di media sosial. Juga untuk mengecek keakuratan algoritma-algoritma dalam penerapan *hoax detection*.