

1. Pendahuluan

Latar Belakang

Sosial media telah menjadi bagian dalam kehidupan manusia, seiring dengan perkembangan zaman sosial media terus berevolusi. Keberadaannya membuat banyak informasi mudah untuk didapat bagi penggunanya. Sudah hampir dua tahun sejak secara resmi diumumkan kasusnya pertama COVID19 dengan 2 orang terkonfirmasi positif di Indonesia pada tanggal 2 Maret 2020 oleh Presiden Joko Widodo, membuat banyaknya informasi dan opini masyarakat berupa pro dan kontra pada media sosial salah satunya twitter. Berbagai upaya telah dilakukan oleh pemerintah untuk menangani situasi pandemi yang saat ini sedang terjadi, mulai dari *lockdown*, pemberian vaksin, dan lain sebagainya.

Difusi informasi atau penyebaran informasi adalah salah satu proses dinamis yang dipelajari secara luas pada jaringan[1]. Penyebaran informasi yang sangat cepat di tengah masyarakat yang mudah untuk menelan informasi secara cepat tanpa mengetahui kebenarannya dapat mempengaruhi kebijakan dalam bermedia sosial. Di dalam twitter terdapat informasi yang dapat dilihat penggunanya baik dalam bentuk tweet atau tulisan yang bisa disisipkan dengan gambar, URL, hastag, dan lain sebagainya. Proses penyebaran informasi dalam twitter sangat bergantung pada jumlah followers, jumlah tweet yang disukai, jumlah retweet, dan umur akun user[2]. Terlebih di dalam twitter terdapat fitur retweet yang dapat dimanfaatkan pengguna yang menyebabkan informasi akan tersebar lebih luas. Difusi informasi pada twitter juga dipengaruhi beberapa *feature* yaitu user-based, time based, dan content-based[2]. Pada penelitian yang telah dilakukan oleh Adrian sebelumnya dengan menggunakan klasifikasi *multi-class* mendapatkan nilai akurasi dan *f1-score* kurang dari 80%, namun dengan penggunaan algoritma *deep learning* mampu membuat pelatihan model yang dibangun dengan lebih baik lagi[3].

Dengan adanya penelitian yang telah dilakukan Adrian sebelumnya, maka pada tugas akhir ini akan menggunakan metode Multi Level Stacking untuk memprediksi retweet yang berkaitan dengan tweet vaksinasi COVID19 di Indonesia pada media sosial *twitter* menggunakan *feature* user-based dan time-based. Dengan mengkombinasikan algoritma Decision Tree, KNN, dan *Naive Bayes* dalam metode Multi Level Stacking diharapkan menghasilkan performa yang baik.

Topik dan Batasan

Berdasarkan latar belakang yang telah disampaikan sebelumnya, maka rumusan masalah pada tugas akhir ini yaitu memprediksi tweet yang berkaitan dengan vaksinasi COVID19 akan di retweet berdasarkan *feature user-based* dan *time-based*. Serta, menganalisa metode Multi Level Stacking dalam melakukan prediksi. Dengan adanya batasan masalah yang bertujuan agar penelitian tugas akhir dilaksanakan secara spesifik, terdapat beberapa batasan yaitu data yang diambil dari Twitter mengenai tweet vaksinasi COVID19 dan data tweet dari Twitter yang digunakan menggunakan bahasa Indonesia.

Tujuan

Tujuan dari penelitian tugas akhir ini membuat sistem prediksi retweet yang berkaitan dengan tweet vaksinasi COVID19 berdasarkan *feature user-based* dan *time-based*. Serta mengimplementasi dan menganalisa metode Multi Level Stacking dalam melakukan prediksi.