

1. Pendahuluan

Latar Belakang

Pada masa digital ini, informasi dapat menyebar secara luas dan cepat melalui media sosial. Saat ini, Twitter merupakan salah satu situs jejaring sosial yang paling banyak digunakan. Salah satu mekanisme difusi informasi pada Twitter adalah *retweet*[1]. *Retweet* adalah kegiatan dimana pengguna memposting ulang *tweet* pengguna lainnya. Dengan memprediksi *tweet* yang akan mendapatkan banyak *retweet* dapat membantu menyelesaikan masalah seperti analisis bisnis dan pemasaran yang membutuhkan *engagement* yang besar[2].

Pada penelitian sebelumnya, Thi Bich Ngoc Hoang dkk. Pada tahun 2017 melakukan penelitian dengan tujuan untuk melakukan prediksi *retweet*. Menggunakan metode klasifikasi Random Forest dengan fitur *user-based*, *time-based*, dan *content-based*[3]. Penelitian selanjutnya, Rakes dkk. Membuat sistem prediksi *retweet* berdasarkan fitur pengguna dengan metode klasifikasi SVM pada tahun 2021[4]. Penelitian berikutnya, Hamidan dkk. Melakukan studi mengenai prediksi *retweet* menggunakan metode klasifikasi ANN pada tahun 2021, dengan fitur berbasis pengguna dan konten[5].

Tujuan penelitian ini, yaitu untuk membangun model prediksi *retweet* dengan metode klasifikasi ANN yang dioptimasi algoritma Cat Swarm Optimization (CSO) menggunakan fitur pengguna, konten, dan waktu. CSO berpotensi untuk meningkatkan performansi model prediksi *retweet* dengan mencari kombinasi parameter terbaik sehingga memperoleh struktur jaringan ANN yang optimal. CSO merupakan algoritma pengoptimalan yang dikembangkan berdasarkan perilaku umum kucing saat berburu mangsanya. Peneliti memilih metode ini karena pada penelitian sebelumnya ANN-CSO dapat memperoleh hasil performansi yang baik dibandingkan algoritma optimasi lainnya[6]–[8].

Berdasarkan hal tersebut, topik pada penelitian ini membahas tentang pembuatan model prediksi *retweet* menggunakan fitur berbasis pengguna, konten, dan waktu dengan memanfaatkan metode klasifikasi ANN yang dioptimasi algoritma CSO. Batasan dari penelitian ini adalah *dataset* yang diambil dengan metode *crawling data* menggunakan kata kunci “capres”. Data yang diambil terdiri dari *tweet* berbahasa Indonesia dalam rentang waktu bulan Juni – Juli 2023.

Tujuan

Penelitian ini bertujuan untuk meningkatkan performansi model prediksi *retweet* dengan metode klasifikasi ANN yang dioptimasi CSO menggunakan fitur pengguna, konten, dan waktu. Serta mengetahui dampak penggunaan metode *resampling* terhadap performansi model klasifikasi dalam mengatasi ketidakseimbangan kelas pada data *tweet*.

Organisasi Tulisan

Penelitian ini akan membahas tentang metodologi dan studi terkait serupa dari penelitian sebelumnya yang berkaitan dengan topik *retweet*. Bagian berikutnya, akan menjelaskan urutan langkah yang akan diambil untuk membangun model prediksi *retweet* menggunakan informasi data pengguna, konten, dan waktu. Setelah itu, berbagai skenario pengujian akan dilakukan untuk mengevaluasi dan menganalisis hasil penelitian. Kesimpulan dan saran untuk penelitian lebih lanjut disajikan pada bagian terakhir.