

Prediksi Retweet Berdasarkan Fitur Pengguna, Konten, dan Waktu Menggunakan Metode Evolutionary UnderSampling Boosting

Ghina Khoerunnisa¹, Jondri², Widi Astuti³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹ghinak@students.telkomuniversity.ac.id, ²jondri@telkomuniversity.ac.id, ³widiwdu@telkomuniversity.ac.id

Abstrak

Twitter merupakan salah satu microblogging populer yang memungkinkan pengguna untuk menulis postingan. Retweet merupakan salah satu mekanisme penyebaran informasi di Twitter. Salah satu cara untuk memahami penyebaran informasi adalah dengan mempelajari mengenai prediksi retweet. Penelitian ini berfokus pada prediksi retweet dengan menggunakan Evolutionary Undersampling Boosting (EUSBoost) berdasarkan fitur berbasis pengguna, konten, dan waktu. Kami juga mempertimbangkan vektor teks sebagai fitur prediktif. Model yang dibangun dengan EUSBoost mampu mengungguli model dengan metode AdaBoost. Hasil evaluasi menunjukkan bahwa model terbaik dapat mencapai skor kinerja AUC sebesar 77,21% dan skor GM sebesar 77,18%. Sementara model berbasis Adaboost mencapai skor AUC mulai dari 68% hingga 69% dan skor GM mulai dari 62% hingga 63%. Selain itu, berdasarkan penelitian yang dilakukan ditemukan bahwa tidak ada perbedaan yang signifikan antara hanya menggunakan fitur numerik dan menggabungkan fitur numerik dan teks.

Kata kunci: difusi informasi, prediksi retweet, EUSBoost

