

ABSTRAK

Kurator berita adalah pengguna Twitter yang mengikuti, berbagi, membagikan kembali, atau memberikan tanggapan terhadap topik berita tertentu melalui tweet atau retweet di Twitter. Pengguna ini memiliki peran penting untuk membantu proses pengembangan berita dan sebagai sumber berita alternatif untuk masyarakat umum. Contoh kurator berita di Indonesia adalah @kurawa, akun Twitter yang membantu menyelidiki kasus pidana di Jakarta International School pada tahun 2016. Meskipun pengguna ini penting, masih kurangnya kepercayaan publik dikarenakan berbagai alasan. Seperti tweet berisi informasi yang menyesatkan, perkataan yang mendorong kebencian (*hate speech*), atau *hoax*. Serta, ada banyak pengguna jahat dengan *follower* palsu / bot yang dapat dengan mudah menaikkan tweet tersebut, terlebih lagi pengguna yang terverifikasi (memiliki tanda centang biru) tidak dapat menjamin untuk membagikan tweet yang bernilai. Sehingga, tantangannya adalah bagaimana cara untuk menilai dan mengklasifikasikan kredibilitas kurator berita di Twitter. Dalam studi ini, kredibilitas ditaksir menjadi 3 kelas yang berbeda: kredibel, tampak kredibel dan tidak kredibel. Dan kelas tersebut dianalisis oleh publik menggunakan survei pengguna. Kemudian, *classifier* dibangun menggunakan metode Naive Bayes Classifier dengan 159 fitur yang dikelompokkan dalam 7 grup fitur. Seperti *user-based*, *content-based*, *retweet-based*, *sentiment learning-based*, *topic learning-based*, *spam lexicon-based* dan *sentiment lexicon-based*. Hasil performansinya adalah 66,5% skor F1. Dalam pengoptimalan, *classifier* diuji secara berurutan menggunakan ratio data training dan testing yang berbeda, jumlah tweet yang berbeda, metode fitur seleksi yang berbeda, dan jumlah kurator berita yang tidak berlabel yang berbeda. Dan hasilnya menghasilkan skor F1 yang meningkat sebesar 7,3%, 1,8%, 3,8% dan 5,1% secara berurutan juga. Hasil performansi terbaik adalah 77,2% skor F1, dan diperoleh menggunakan rasio 90:10 data training dan testing, tweet sebanyak 75.852 tweet, fitur sebanyak 65 fitur yang dipilih oleh metode Information Gain, dan 200 kurator berita yang tidak berlabel yang dioptimalkan menggunakan metode Expectation-Maximization. Lalu, ada 3 grup fitur yang paling relevan dengan kelas kredibilitas. Fitur tersebut adalah *retweet-based*, *sentiment learning-based* dan *topic learning-based*.

Kata kunci: kurator berita, twitter, kredibilitas, klasifikasi, fitur