

Abstrak

Twitter merupakan salah satu media sosial yang masih banyak digunakan hingga saat ini. Pesan di Twitter dibatasi hanya 280 karakter, sehingga memberikan batasan pada pengguna untuk bertukar informasi dan memaksa pengguna menggunakan berbagai strategi seperti bahasa gaul dan singkatan untuk memadatkan lebih banyak informasi. Hal tersebut membuat pengguna kesulitan untuk memahami konteks atau topik dari informasi sehingga dibutuhkan solusi untuk permasalahan ini. Telah banyak penelitian yang melakukan klasifikasi topik pada twitter tetapi memberikan hasil akurasi rendah dikarenakan batasan-batasan tersebut dan belum menerapkan metode *hybrid deep learning* pada klasifikasi topik. Oleh karena itu, penelitian ini melibatkan fitur ekspansi untuk mengatasi keterbatasan tersebut dengan mencari kata-kata yang mirip menggunakan Glove dan menerapkan metode hybrid deep learning pada model klasifikasi untuk mengungguli model tunggal. Model yang digunakan dalam penelitian ini adalah Bidirectional Long Short-Term Memory (Bi-LSTM), Convolutional Neural Network (CNN), *hybrid* BILSTM-CNN, dan *hybrid* CNN-BILSTM. Hasil penelitian menunjukkan bahwa model Bi-LSTM dengan fitur expansion menghasilkan akurasi sebesar 93,82%, diikuti oleh CNN dengan fitur expansion menghasilkan akurasi sebesar 93,75%. Model terbaik untuk klasifikasi topik pada penelitian ini dengan rasio terbaik (80:20), variasi N-gram terbaik (unigram + bigram + trigram) , dan fitur ekspansi glove dengan top 5 similarity yaitu hybrid CNN-BILSTM dengan akurasi sebesar 94,34% menghasilkan peningkatan akurasi sebesar 1,71% terhadap Bi-LSTM dan 1,47% terhadap CNN.

Kata kunci : klasifikasi topik, Glove, Bi-LSTM, CNN, hybrid deep learning
