

ABSTRAK

Penelitian ini bertujuan untuk meningkatkan kinerja klasifikasi sentimen pada dimensi spiritual dalam analisis sentimen multidimensi terkait twit mengenai vaksinasi COVID-19 di Indonesia. Tiga metode ekspansi kata kunci yang diuji meliputi *Mass Diffusion of Co-occurrence Approach* (MDCA), *Word2Vec*, dan TF-IDF. MDCA diuji dengan konfigurasi Top-10, Top-20, Top-30, dan Top-50. Seluruh kata kunci hasil ekspansi dievaluasi secara manual oleh tiga ahli spiritual. Kata kunci yang telah divalidasi digunakan untuk mengumpulkan dataset twit baru, yang selanjutnya diproses, diberi label sentimen menggunakan model IndoBERT yang telah di-*fine-tune*, lalu digabungkan dengan dataset awal. Setiap dataset yang dihasilkan diuji menggunakan *Naïve Bayes*, SVM, IndoBERT Base, dan IndoBERT Large P2. Hasil penelitian menunjukkan MDCA Top-30 memberikan kinerja terbaik pada *traditional classifier*, dengan peningkatan akurasi hingga +2,6% dan kenaikan *F1-score* sebesar +8,0% pada SVM. Sementara itu, MDCA Top-50 mencapai performa tertinggi pada model *deep learning*, dengan peningkatan akurasi hingga +7,8% dan *F1-score* sebesar +10,0% pada IndoBERT Base. Analisis kesalahan klasifikasi menunjukkan bahwa kesalahan sering terjadi pada istilah yang berkaitan dengan geografi dan pendidikan, sehingga menunjukkan perlunya penambahan dimensi analisis. Secara keseluruhan, temuan ini membuktikan MDCA, khususnya Top-30 dan Top-50, lebih efektif dibandingkan *Word2Vec* dan TF-IDF dalam memperluas kata kunci terkait dimensi spiritual serta meningkatkan performansi klasifikasi sentimen dalam analisis multidimensi wacana vaksin COVID-19 di Indonesia.

Kata kunci: ekspansi kata kunci, analisis sentimen multidimensi, dimensi spiritual, twit vaksinasi COVID-19, mass diffusion of co-occurrence approach (MDCA)