

## Abstrak

SMS (*Short Message Service*) merupakan salah satu layanan komunikasi yang masih menjadi pilihan utama meskipun sekarang ponsel semakin berkembang dengan berbagai aplikasi. Seiring berkembangnya berbagai media komunikasi lain, beberapa negara menurunkan tarif SMS agar tetap menarik minat pengguna ponsel. Hal itu mengakibatkan meningkatnya SMS *spam* yang dimanfaatkan oleh beberapa pihak salah satunya untuk iklan. Mengingat berkembangnya macam dokumen dengan multi-bahasa pada suatu pesan, Web, dan lain-lain, kebutuhan teknik pemrosesan multibahasa atau *cross-lingual* yang efektif menjadi semakin penting. Oleh sebab itu, pada tugas akhir ini dibuatlah SMS *spam filtering* untuk menyaring pesan menggunakan algoritma *Graph-based K-Nearest Neighbor*(GKNN). Tujuan dari tugas akhir ini adalah menganalisis penerapan metode *Graph-based KNN Algorithm* terhadap akurasi dan performansi dalam klasifikasi SMS spam dengan *cross-lingual*. Secara garis besar tahapan yang dilakukan dalam penelitian ini adalah data/pesan terlebih dahulu *dipreprocessing* kemudian direpresentasikan ke dalam suatu model graph. Selanjutnya dilakukan perhitungan dengan menggunakan metode GKNN. Dari penelitian ini didapatkan rata-rata nilai akurasi sebesar 96.27% dan rata-rata nilai F1 yaitu 87.28% untuk pengujian menggunakan dua bahasa (data training bahasa Indonesia + Bahasa Inggris (terjemah ke Indonesia)) begitu juga dengan data testing dengan nilai K untuk metode G-KNN = 10 dan *threshold* 0.001.

**Kata Kunci** : SMS, *spam*, *spam filtering*, klasifikasi, *graph-based k-nearest neighbor*, *supervised*.