

## ABSTRAK

Bencana alam adalah hal yang alami dan tidak bisa diperkirakan. Dampak bencana alam tergantung dari besar intensitasnya, yang dapat berupa tanah longsor, banjir, gempa bumi, bahkan korban jiwa sekalipun. Banyak hal yang bisa diminimalisir sebelum bencana tersebut meluas. Untuk itulah pentingnya informasi akan terjadinya bencana.

Sosial media adalah tempat yang dapat menghubungkan satu informasi ke informasi yang lain, sehingga dapat tersebar luas. Perkembangan media sosial saat ini sangatlah cepat terlebih dengan adanya berita yang berkaitan dengan bencana di sekitar. Salah satu media sosial yang sekarang sangat banyak digunakan adalah Twitter. Dengan menggunakan Twitter, masyarakat dapat dengan cepat memberi penyebaran informasi bencana melalui tweet agar dampak penanggulangan dapat dipercepat. Maka dari itu, inilah pentingnya untuk mengetahui informasi secara real-time tentang jumlah kejadian bencana alam agar antisipasi dapat dilakukan sejak dini.

Maka dari penjelasan yang disebutkan diatas dibutuhkan sistem yang dapat memilah dengan sendiri data bencana alam pada Tweet. Hasil pengujian dibuat untuk menampilkan mapping bencana yang terjadi di wilayah Indonesia yang diklasifikasikan berdasarkan area yang memiliki data tweet dalam bentuk visualisasi peta wilayah bencana mana yang lebih tinggi frekuensi serta jenis bencananya. Klasifikasi wilayah dilakukan menggunakan metode algoritma K-Nearest Neighbor. Dalam penelitian ini hasil nilai uji Confusion matrix memiliki akurasi yang terbaik dengan menggunakan metric Jaccard sebesar 86% dan untuk performansi data sharing menggunakan k-fold cross validation 10 Fold, hasil akurasi terbaiknya adalah metric Jaccard yaitu sebesar 83% pada Fold 8.

**Kata Kunci:** *twitter, bencana alam, klasifikasi, confusion matrix, k-fold cross validation*