

ABSTRACT

Landslides are natural disasters which often occur in several regions in Indonesia that result in major losses in aspects of human life. To reduce and prevent major impacts resulting from landslides, it is important to create landslide susceptibility map as a way to mitigate and design land use in areas prone to landslides. In the process of creating a landslide susceptibility map, one way is to use a machine learning algorithm. One of the machine learning algorithms that has been widely used in this case is K-Nearest Neighbor (KNN) to carry out the data classification process of factors causing landslides and landslide data.

The exploration of KNN algorithm and creating a landslide susceptibility map is the aim of this research. In this research the soft voting classifier algorithm was used to combine the probability prediction results from calculating euclidean, manhattan, chebyshev, and cosine distance metrics with different k values. This research also used mutual information for the feature selection process and used distance weighting calculations. Testing was carried out by comparing the performance of the KNN ensemble algorithm with individual KNN algorithms and other machine learning algorithms.

Based on the research results, it shows that the proposed algorithm, the K-Nearest Neighbor ensemble soft voting has better performance than the individual KNN algorithm and other machine learning algorithms with a value of weighted precision 0.98923, weighted recall 0.98784, weighted f1-score 0.98825, and balanced accuracy 0.97906. The proposed algorithm can be used to create a landslide disaster vulnerability map for Kejajar District.

Keywords – K-Nearest Neighbor, distance metric, soft voting classifier, machine learning, landslide susceptibility map

INTISARI

Bencana tanah longsor merupakan bencana alam yang sering terjadi di beberapa wilayah di Indonesia yang mengakibatkan kerugian besar pada aspek kehidupan manusia. Untuk mengurangi dan mencegah dampak yang besar akibat bencana tanah longsor maka penting untuk membuat peta kerawanan bencana tanah longsor sebagai cara untuk mitigasi dan perancangan penggunaan lahan pada daerah yang rawan bencana tanah longsor. Dalam proses pembuatan peta kerawanan bencana tanah longsor salah satu cara adalah dengan menggunakan algoritma *machine learning*. Salah satu algoritma *machine learning* yang sudah banyak digunakan dalam kasus ini adalah *K-Nearest Neighbor* (KNN) untuk melakukan proses klasifikasi data faktor-faktor penyebab terjadinya bencana tanah longsor dan data bencana tanah longsor.

Melakukan eksplorasi terhadap algoritma KNN dan pembuatan peta kerawanan bencana tanah longsor menjadi tujuan dalam penelitian ini. Dalam penelitian ini menggunakan algoritma *soft voting classifier* untuk menggabungkan hasil probabilitas prediksi dari perhitungan metrik jarak *euclidean*, *manhattan*, *chebyshev*, dan *cosine* dengan nilai k yang berbeda. Dalam penelitian ini juga menggunakan *mutual information* untuk proses *feature selection* dan menggunakan perhitungan pembobotan jarak atau *weight distance*. Pengujian dilakukan dengan membandingkan kinerja dari algoritma *ensemble* KNN dengan algoritma KNN individu dan algoritma *machine learning* lain.

Berdasar hasil penelitian menunjukkan bahwa algoritma yang diusulkan yaitu *K-Nearest Neighbor ensemble soft voting* memiliki kinerja yang lebih baik dari algoritma KNN individu dan algoritma *machine learning* lain dengan nilai *weighted precision* 0,98923, *weighted recall* 0,98784, *weighted f1-score* 0,98825, dan *balanced accuracy* 0,97906. Algoritma yang diusulkan dapat digunakan untuk membuat peta kerawanan bencana tanah longsor untuk Kecamatan Kejajar.

Kata kunci – *K-Nearest Neighbor*, perhitungan metrik jarak, *soft voting classifier*, *machine learning*, peta kerawanan bencana tanah longsor