

ABSTRACT

Natural disasters pose a major threat to coastal environments, often causing major loss of life and property damage. Coastal disasters such as erosion, floods, sea level rise and landslides often occur so it is very important to carry out disaster risk assessments in order to inform the formulation of coastal area management policies. The classification is divided into five hazard categories, namely very low, low, moderate, high and very high. Machine learning methods have been used as a method for classifying coastal vulnerability. One method that has been used to classify disaster vulnerability is the artificial neural network (ANN) method. However, there is a problem of imbalance in class distribution in the dataset, making it difficult for the model to recognize certain vulnerability classes.

Therefore, this research proposes the application of an artificial neural network method to classify the vulnerability of coastal areas using the dataset balancing method. Five dataset balancing methods are applied to the model, namely SMOTE, SMOTESVM, SMOTEENN, SMOTEKNN, and random over sampler. This research also uses SVM, logistic regression and decision tree methods to compare the performance of ANN models in classifying coastal vulnerability.

The research results show that the SMOTEENN balancing technique can increase recall for the minority (high) class from 56% when using the ANN model without balancing to 96% when the SMOTEENN method is applied. The ANN method with the SMOTEENN balancing technique achieved an accuracy of 98.3% after grid search optimization, and an accuracy of 97.7% with the initial model. The accuracy of the ANN method exceeds the classification results using the SVM, Logistic Region and Decision Tree methods. These results indicate that ANN with the SMOTEENN data balancing technique is more suitable for coastal vulnerability classification.

Keywords : ANN, SVM, Logistic Resgesion, SMOTEENN, CVI, Multi-Class Dataset, Imbalance Dataset, Grid Search.

INTISARI

Bencana alam menimbulkan ancaman besar terhadap lingkungan pesisir, sering kali menyebabkan kerugian besar pada nyawa dan kerusakan properti. Bencana pesisir seperti erosi, banjir, kenaikan permukaan air laut, dan longsor sering terjadi sehingga sangat penting untuk melakukan penilaian risiko bencana guna untuk dijadikan ajuan perumusan kebijakan pengolahan wilayah pesisir. Klasifikasi dibagi menjadi lima kategori bahaya yaitu sangat rendah, rendah, sedang, tinggi, dan sangat tinggi. Metode *machine learning* telah digunakan sebagai metode klasifikasi kerentanan pesisir. Namun, terdapat masalah ketidakseimbangan distribusi kelas pada *dataset* menyebabkan model sulit mengenali kelas kerentanan tertentu.

Oleh karena itu, penelitian ini mengusulkan penerapan metode *artificial neural network* untuk mengklasifikasikan kerentanan wilayah pesisir dengan menggunakan metode penyeimbang *dataset*. Lima metode penyeimbang *dataset* dan diterapkan pada model, yaitu SMOTE, SMOTESVM, SMOTEENN, SMOTEKNN, dan *random over sampler*. Penelitian ini juga menggunakan metode SVM, *logistic regresion*, dan *decision tree* untuk membandingkan kinerja model ANN dalam mengklasifikasi kerentanan pesisir.

Hasil penelitian menunjukkan teknik penyeimbangan SMOTEENN Selain itu terbukti dapat menaikkan *recall* untuk kelas minoritas (*high*) dari 56% saat menggunakan model ANN tanpa penyeimbang menjadi 96% ketika telah diterapkan metode SMOTEENN. Metode ANN dengan teknik penyeimbangan SMOTEENN, mencapai akurasi sebesar 98,3% setelah optimasi *grid search*, dan akurasi 97,7% dengan model awal. Akurasi Metode ANN melampaui hasil klasifikasi dengan metode SVM, *logistic regresion*, dan *decision tree*. Hasil ini menunjukkan bahwa ANN dengan teknik penyeimbangan data SMOTEENN lebih sesuai digunakan untuk klasifikasi kerentanan pesisir.

Kata kunci – ANN, SVM, *Logistic Regresion*, SMOTEENN, CVI, Klasifikasi Multi Kelas, Dataset Tidak Seimbang, *Grid Search*.