

ABSTRAK

Peningkatan volume berita pariwisata di Indonesia menghadirkan tantangan dalam pengelompokan berita ke dalam sub kategori seperti wisata alam, wisata budaya, wisata buatan, dan bukan pariwisata. Untuk mengatasi masalah ini, penelitian ini mengembangkan model klasifikasi teks berbasis *Artificial Neural Network* (ANN). Sebanyak 5.261 dokumen berita yang dikumpulkan dari situs Detik.com digunakan sebagai dataset. Tahapan penelitian meliputi *preprocessing* data, seperti *cleaning*, tokenisasi, penghapusan *stopword* menggunakan *nlTK*, dilakukan pembobotan *TF-IDF* untuk merepresentasikan teks dalam bentuk numerik serta penggunaan *K-fold Cross Validation* untuk pembagian data dan *SMOTE* untuk *balancing* data. Model ANN dirancang dengan fungsi aktivasi *ReLU* pada *hidden layer* dan *softmax* pada *output layer* untuk mendukung klasifikasi *multiclass*. Selain itu, penelitian ini mengevaluasi pengaruh variasi *epoch* dan penggunaan *SMOTE* terhadap kinerja model. Hasil eksperimen menunjukkan bahwa metode ANN mampu mencapai tingkat akurasi yang memuaskan, dengan nilai akurasi 97.33% dalam klasifikasi teks berita pariwisata. Penelitian ini memberikan kontribusi signifikan dalam penerapan metode ANN untuk pengolahan teks berbahasa Indonesia, khususnya dalam lingkup berita pariwisata.

Kata kunci: *Artificial Neural Network*, Klasifikasi teks, pariwisata

\

ABSTRACT

The increasing volume of tourism news in Indonesia presents challenges in grouping news into sub-categories such natural tourism, cultural tourism, artificial tourism, and non-tourism. To overcome this problem, this research develops a text classification model based on Artificial Neural Network (ANN). A total of 5,261 news documents collected from the Detik.com site were used a dataset. The research stages include data preprocessing, such cleaning, tokenization, removing stop words using nltk, TF-IDF weighting to represent text in numerical form and using K-fold Cross Validation for data division and SMOTE for data balancing. The ANN model is designed with a ReLU activation function in the hidden layer and softmax in the output layer to support multiclass classification. In addition, this study evaluates the effect of epoch variations and the use of SMOTE on model performance. The experimental results show that the ANN method is able to achieve a satisfactory level of accuracy, with an accuracy value of 97.33% in the classification of tourism news texts. This research makes a significant contribution to the application of the ANN method for processing Indonesian language texts, especially in the scope of tourism news.

Keywords: *Artificial Neural Network, Text Classification, Tourism*