

ABSTRAK

Meningkatnya konsumsi berita daring di Indonesia, terutama pada konten pariwisata seperti wisata alam, buatan, budaya, dan religi, mendorong perlunya sistem klasifikasi otomatis yang mampu mengelompokkan berita sesuai isi kontennya. Penelitian ini memfokuskan pada pengembangan sistem klasifikasi berita pariwisata yang diperoleh dari situs Kompas.com menggunakan teknik web scraping. Permasalahan utama yang diangkat adalah belum tersedianya sistem klasifikasi otomatis yang akurat dan terstruktur untuk berita pariwisata, yang menyebabkan kesulitan dalam pencarian informasi spesifik dan berpotensi meningkatkan penyebaran misinformasi akibat literasi digital yang rendah. Berbagai algoritma pembelajaran mesin telah digunakan dalam klasifikasi teks, namun memiliki keterbatasan dalam menangani relasi non-linier dalam data teks. Untuk itu, penelitian ini menggunakan algoritma *Multilayer Perceptron* (MLP), yang mampu memodelkan hubungan non-linier secara efektif dan melakukan ekstraksi fitur otomatis. Tujuan dari penelitian ini adalah mengembangkan sistem klasifikasi berita pariwisata berbasis empat kategori utama, yakni wisata alam, wisata buatan, budaya, dan wisata religi. Hasil penelitian terbaik ditunjukkan oleh Skenario 3 yang menggunakan fungsi aktivasi *tanh* dan solver *Adam*, menghasilkan akurasi 86.83%, *precision* 84.61%, *recall* 84.12%, dan *F1 – score* 84.34%. Model *Multilayer Perceptron* (MLP) yang dihasilkan terbukti cukup akurat dalam mengklasifikasikan berita pariwisata ke dalam empat kategori, yaitu wisata alam, buatan, budaya, dan religi.

Kata kunci: Klasifikasi Berita, Kompas.com, *Multilayer Perceptron*, Pariwisata, Teks Digital.

ABSTRACT

The increasing consumption of online news in Indonesia, particularly in tourism-related content such as natural, artificial, and religious destinations, highlights the need for an automatic classification system capable of grouping news articles based on their content. This study focuses on the development of a tourism news classification system using data collected from Kompas.com through web scraping techniques. The main issue addressed in this research is the absence of an accurate and structured automatic classification system for tourism news, which leads to difficulties in information retrieval and increases the potential spread of misinformation due to low digital literacy. Various machine learning algorithms have been used in text classification, but they have limitations in handling non-linear relationships in text data. Therefore, this research adopts the Multilayer Perceptron (MLP) algorithm, which is capable of effectively modeling complex non-linear relationships and performing automatic feature extraction. Previous studies have shown that MLP achieves higher accuracy compared to other algorithms in similar classification tasks. The purpose of this study is to develop a tourism news classification system based on four main categories, namely nature tourism, artificial tourism, culture, and religious tourism. The best research results were shown by Scenario 3, which used the tanh activation function and Adam solver, resulting in an accuracy of 86.83%, precision of 84.61%, recall of 84.12%, and F1 score of 84.34%. The Multilayer Perceptron (MLP) model produced proved to be sufficiently accurate in classifying tourism news into four categories: natural tourism, man-made tourism, cultural tourism, and religious tourism.

Keywords: *news classification, Kompas.com, Multilayer Perceptron, tourism, digital text.*