ABSTRAK

Perkembangan media digital meningkatkan penyebaran berita tidak kredibel (*Fake News*) yang umumnya disajikan dalam bentuk kombinasi teks dan gambar. Pada berita bencana alam di Indonesia, penggunaan gambar yang tidak sesuai dengan isi berita dapat menimbulkan disinformasi dan memengaruhi persepsi masyarakat. Namun, penelitian klasifikasi berita di Indonesia masih didominasi pendekatan unimodal berbasis teks dan belum mempertimbangkan hubungan semantik antara teks dan gambar secara terukur. Oleh karena itu, penelitian ini bertujuan membangun sistem klasifikasi kredibilitas berita online berbasis analisis *Multimodal* dengan mengintegrasikan IndoBERT sebagai *text encoder* dan CLIP sebagai *image encoder*. Penelitian menggunakan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dengan dataset berita bencana alam periode 2024–2025 dari berbagai media online. Pada penelitian ini, modalitas teks yang digunakan berupa judul berita, sedangkan modalitas visual berupa gambar pendukung berita. Representasi judul berita dan gambar diekstraksi menggunakan IndoBERT dan CLIP, kemudian hubungan semantik antar-modalitas diukur menggunakan *Cosine Similarity*. Nilai keselarasan tersebut digabungkan melalui proses *Multimodal Fusion* untuk melakukan klasifikasi kredibilitas berita. Hasil evaluasi menunjukkan bahwa sistem memperoleh nilai *Accuracy*, *Precision*, *Recall,* dan *F1-Score* sebesar 93%. Hasil penelitian menunjukkan bahwa pendekatan *Multimodal* menggunakan IndoBERT dan CLIP mampu menghasilkan klasifikasi kredibilitas berita yang lebih objektif, konsisten, dan terukur berdasarkan keselarasan antara teks dan gambar.

Kata Kunci: *Multimodal Analysis*, *IndoBERT*, *CLIP*, *Cosine Similarity*, Kredibilitas Berita.

ABSTRACT

The development of digital media has increased the dissemination of non-credible news (Fake News), which is commonly presented through a combination of text and images. In Indonesian natural disaster news, the use of images that do not align with the content can lead to misinformation and influence public perception. However, news classification research in Indonesia is still dominated by unimodal text-based approaches and has not yet considered the semantic relationship between text and images in a measurable way. Therefore, this study aims to develop an online news credibility classification system based on Multimodal analysis by integrating IndoBERT as a text encoder and CLIP as an image encoder. The research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology using a natural disaster news dataset from the 2024–2025 period collected from various online media platforms. In this study, the textual modality consists of news headlines, while the visual modality consists of accompanying news images. Textual and visual representations are extracted using IndoBERT and CLIP, and the semantic relationship between modalities is measured using Cosine Similarity. These alignment values are then integrated through a Multimodal Fusion process to perform news credibility classification. Evaluation results indicate that the system achieved an Accuracy, Precision, Recall, and F1-Score of 93%. The findings demonstrate that a Multimodal approach using IndoBERT and CLIP is capable of producing a news credibility classification that is more objective, consistent, and measurable based on the alignment between text and images.

Keywords: *Multimodal* Analysis, *IndoBERT*, *CLIP*, *Cosine Similarity*, News Credibility.