

ABSTRAK

MUHAMAD RIVAI. Analisis Sentimen Judul Berita pada Media Massa Menggunakan RoBERTa. Dibimbing oleh TEKAD MATULATAN dan NOLAN EFRANDA.

Perkembangan internet di Indonesia yang sangat pesat mendorong konsumsi berita daring yang sangat tinggi, dimana judul berita sering kali dijadikan dasar penilaian oleh pembaca tanpa membaca isi keseluruhan berita. Berita tentang program makan bergizi gratis (MBG) menuai pro dan kontra di khalayak umum. Analisis sentimen digunakan untuk menganalisis teks untuk mengidentifikasi apakah maknanya memiliki sentimen positif, negatif, atau netral. Penelitian ini bertujuan untuk mengembangkan model untuk klasifikasi sentimen judul berita menggunakan algoritma RoBERTa. Data judul berita diperoleh dari tiga media massa daring terpopuler yaitu Detik, Kompas, dan YouTube dari Januari sampai Maret 2025 melalui proses *scraping*. Pelabelan data menggunakan algoritma IndoBERT. Proses pelatihan model menggunakan metode *k-fold cross validation* dengan dataset sebanyak 2093 data. Hasil dari penelitian ini didapatkan akurasi model RoBERTa pada data *testing* sebesar 79% , presisi 74%, *recall* 77%, dan *f1-score* 75%. Hasil menunjukkan pelatihan RoBERTa memakan waktu 14 menit 29 detik dengan konsumsi RAM sebesar 18,3% dan GPU sebesar 12,79%.

Kata kunci: Analisis Sentimen, *Deep Learning*, Judul Berita, Media Massa, RoBERTa

ABSTRACT

MUHAMAD RIVAI. Sentiment Analysis of News Headlines in Mass Media Using RoBERTa. Supervised by TEKAD MATULATAN dan NOLAN EFRANDA.

The rapid development of the internet in Indonesia has led to a very high consumption of online news, where news headlines are often used as a basis for judgment by readers without reading the entire content of the news. News about the free nutritious meal program (MBG) has sparked pros and cons among the general public. Sentiment analysis is used to analyze text to identify whether its meaning has positive, negative, or neutral sentiment. This study aims to develop a model for classifying the sentiment of news headlines using the RoBERTa algorithm. The headline data was obtained from three of the most popular online media outlets, namely Detik, Kompas, and YouTube, from January to March 2025 through a scraping process. Data labeling was performed using the IndoBERT algorithm. The model training process used the k-fold cross validation method with a dataset of 2093 data. The results of this study show that the accuracy of the RoBERTa model on the testing data is 79%, precision 74%, recall 77%, and f1-score 75%. The results show that RoBERTa training took 14 minutes 29 seconds with RAM consumption of 18,3% and GPU consumption of 12,79%.

Keywords: Deep Learning, Mass Media, News Headlines, RoBERTa, Sentiment Analysis

