

ABSTRAK

Nama : Muhammad Farhan
NIM : 41518110084
Program Studi : Teknik Informatika
Judul : Implementasi Sentimen Analisis Pada Twitter
Menggunakan BERT Terhadap Komentar Kenaikan Harga
Beberapa komoditas Dalam Dua Bahasa.
Pembimbing : Anis Cherid, SE, MTI.

Twitter merupakan platform media sosial terkemuka dengan lebih dari 1 miliar pengguna di Google PlayStore. Ini merupakan tempat buat berbagi komentar, data, serta berhubungan. Peningkatan harga komoditas semacam bahan bakar, bahan pokok, serta yang lain kerap dibahas di sini. Ketersediaan pangan sangat berarti, serta riset ini memanfaatkan pendekatan baru ialah model BERT guna menganalisis sentimen di Twitter. Kami menguji model pre- training BERT, khususnya mBERT, pada komentar Tweet tentang peningkatan harga. Kami mendeteksi kalau model fine- tuned mBERT bahasa Indonesia mempunyai akurasi besar(95%) dengan hyperparameter tertentu. Tata cara pelabelan data mempengaruhi hasil, dengan pelabelan berbasis TextBlob bahasa Indonesia memberikan akurasi terbaik. Model mBERT dapat memprediksi sentimen dengan baik, serta hasil terbaik diperoleh dengan akurasi 95% pada data bahasa Indonesia memakai hyperparameter yaitu Learning Rate 0.00002, Batch Size 32, jumlah Epoch 5 dan waktu pelatihan 10 menit.

Kata kunci: Analisis sentimen, BERT, Twitter, NLP

UNIVERSITAS
MERCU BUANA

ABSTRACT

Name : Muhammad Farhan
Student Number : 41518110084
Study Program : Engineering Informatics
Title : Implementation of Sentiment Analysis on Twitter Using BERT to Comment on Price Increases for Several Commodities in Two Languages
Counsellor : Anis Cherid, SE, MTI.

Twitter is a leading social media platform with over 1 billion users on Google PlayStore. This is a place to share feedback, data and connect. Here we often discuss increases in the prices of raw materials such as fuel, basic materials and others. Food availability is very important and this research uses a new approach, namely the BERT model, to analyze sentiment on Twitter. We tested BERT pre-training models, specifically mBERT, on Tweet comments about price increases. We found that the optimized Indonesian mBERT model had great accuracy (95%) with some hyperparameters. The data labeling procedure influences the results, with TextBlob-based Indonesian labeling providing the highest accuracy. The mBERT model can predict sentiment well and the best results are achieved with 95% accuracy on Indonesian language data using hyperparameters, i.e. learning rate 0.00002, batch size 32, number of epochs 5 and training time 10 minutes.

Keywords: Sentiment Analysis, BERT, Twitter, NLP

UNIVERSITAS
MERCU BUANA