

## ABSTRACT

### ENHANCING STOCK MARKET INSIGHT THROUGH SENTIMENT ANALYSIS: A COMPARATIVE STUDY OF TRADITIONAL AND TRANSFORMER MODEL

Faris Rineksa Palagan

20/454532/PA/19563

This study presents a comparative analysis of transformer-based language models (BERT and FinBERT) and conventional machine learning algorithms (Logistic Regression, Naïve Bayes, Random Forest, and K-Nearest Neighbours) for financial sentiment analysis, utilizing datasets of formal financial news headlines and informal financial tweets, with the aim of potential application in stock market forecasting. The experimental results demonstrate a significant disparity in performance, where conventional models exhibited limited capability in capturing contextual meaning, resulting in suboptimal classification accuracy (generally below 60%). In contrast, transformer-based models consistently outperformed these traditional approaches, with the BERT model achieving a notably high accuracy of **96.9%** on the tweet-based dataset and **88.3%** on the more formal PhraseBank dataset, confirming the superiority of contextual language models for extracting nuanced sentiment from financial text. However, a key limitation identified is the substantial computational demand and inference latency of the high-performing BERT model, which makes it less suitable for immediate real-time deployment in time-sensitive financial applications. Therefore, future research should focus on improving deployment efficiency by investigating lightweight transformer variants and optimization techniques like quantization, and by integrating sentiment predictions with downstream financial indicators, such as stock price movements or trading volume, to develop hybrid pipelines for short-term price forecasting and causal analysis.

**Keyword:** Stock market forecasting, Transformer-based language models, Financial sentiment analysis

## INTISARI

### PENINGKATAN WAWASAN PASAR SAHAM MELALUI ANALISIS SENTIMEN: STUDI KOMPAFATIF ANTARA MODEL TRADISIONAL DAN MODEL TRANSFORMER

Faris Rineksa Palagan

20/454532/PA/19563

Penelitian ini menyajikan analisis komparatif antara model bahasa berbasis transformer (*BERT* dan *FinBERT*) dan algoritma *machine learning* konvensional (Regresi Logistik, *Naïve Bayes*, *Random Forest*, dan *K-Nearest Neighbours*) untuk analisis sentimen keuangan, menggunakan kumpulan data yang terdiri dari berita utama keuangan formal dan *tweet* keuangan informal, dengan tujuan aplikasi potensial dalam peramalan pasar saham. Hasil eksperimen menunjukkan perbedaan kinerja yang signifikan, di mana model konvensional menunjukkan kemampuan terbatas dalam menangkap makna kontekstual, yang menghasilkan akurasi klasifikasi yang suboptimal (umumnya di bawah 60%). Sebaliknya, model berbasis *transformer* secara konsisten mengungguli pendekatan tradisional ini, dengan model *BERT* mencapai akurasi yang sangat tinggi sebesar **96,9%** pada kumpulan data *tweet* dan **88,3%** pada kumpulan data *PhraseBank* yang lebih formal, menegaskan superioritas model bahasa kontekstual untuk mengekstraksi sentimen yang bernuansa dari teks keuangan. Namun, keterbatasan utama yang diidentifikasi adalah tuntutan komputasi yang substansial dan latensi inferensi dari model *BERT* dengan kinerja tinggi, yang membuatnya kurang sesuai untuk penyebaran waktu nyata (*real-time*) dalam aplikasi keuangan yang sensitif terhadap waktu. Oleh karena itu, penelitian di masa depan harus berfokus pada peningkatan efisiensi penyebaran dengan menyelidiki varian *transformer* yang lebih ringan dan teknik optimasi seperti kuantisasi, serta dengan mengintegrasikan prediksi sentimen dengan indikator keuangan hilir, seperti pergerakan harga saham atau volume perdagangan, untuk mengembangkan alur kerja hibrida untuk peramalan harga jangka pendek dan analisis kausal.

**Kata Kunci:** Prediksi pasar saham, Model bahasa berbasis Transformer, Analisis sentimen keuangan