

BAB 2 TINJAUAN PUSTAKA

2.1 Hoaks dalam Kaidah Bahasa

Hoaks merupakan konstruksi informasional digital yang kompleks, dirancang secara sistematis untuk menyesatkan dan memanipulasi persepsi publik. Dari perspektif komputasional, hoaks tidak sekadar dipahami sebagai informasi palsu, melainkan sebagai objek penelitian multidimensional yang membutuhkan pendekatan analisis canggih melalui teknologi pemrosesan bahasa alami (*Natural Language Processing*) dan kecerdasan buatan (*Artificial Intelligence*). Secara teknis, hoaks memiliki karakteristik linguistik dan struktural yang dapat diidentifikasi melalui algoritma machine learning. Penelitian Tajrian et al. [10] mengungkap bahwa hoaks umumnya memiliki pola bahasa spesifik, mencakup beberapa aspek penting sebagai berikut.

1. Penggunaan muatan emosional yang provokatif.
2. Pengaburan sumber informasi.
3. Struktur kalimat yang bertujuan memanipulasi persepsi. Serta,
4. Penggunaan retorika yang menimbulkan ketegangan atau keraguan.

Kompleksitas deteksi hoaks terletak pada kemampuan algoritma untuk melampaui sekadar identifikasi ketidakakuratan informasi. Model komputasional mutakhir dituntut untuk mampu memahami konteks semantik yang mendalam, mengidentifikasi intensi komunikasi, dan menganalisis nuansa bahasa yang tersembunyi dalam teks. Hal ini menjadikan hoaks sebagai objek penelitian yang sangat dinamis dalam ranah kecerdasan buatan.

Dalam konteks media sosial di Indonesia, karakteristik hoaks menunjukkan pola yang lebih spesifik. Berdasarkan pengamatan terhadap penyebaran berita hoaks di platform media sosial, ciri khas berita hoaks yang menjadi fokus dalam pelatihan model IndoBERT ini adalah sebagai berikut.

1. Penggunaan bahasa Indonesia informal yang dominan, mencerminkan karakteristik komunikasi digital sehari-hari.
2. Terdapat kecenderungan penggunaan substitusi huruf dengan angka (misalnya, 'A' menjadi '4', 'I' menjadi '1', dan 'O' menjadi '0') sebagai strategi untuk menghindari deteksi otomatis.

3. Penggunaan huruf besar yang berlebihan mencerminkan bentuk provokasi yang disampaikan oleh pembuat berita hoaks.
4. Terdapat penggunaan kata-kata substitusi dari bahasa Inggris yang signifikan, mencerminkan fenomena pencampuran bahasa yang umum di media sosial Indonesia.

Sebagai ilustrasi, salah satu contoh berita hoaks yang mencakup karakteristik-karakteristik dan menjadi bahan latih model adalah sebagai berikut.

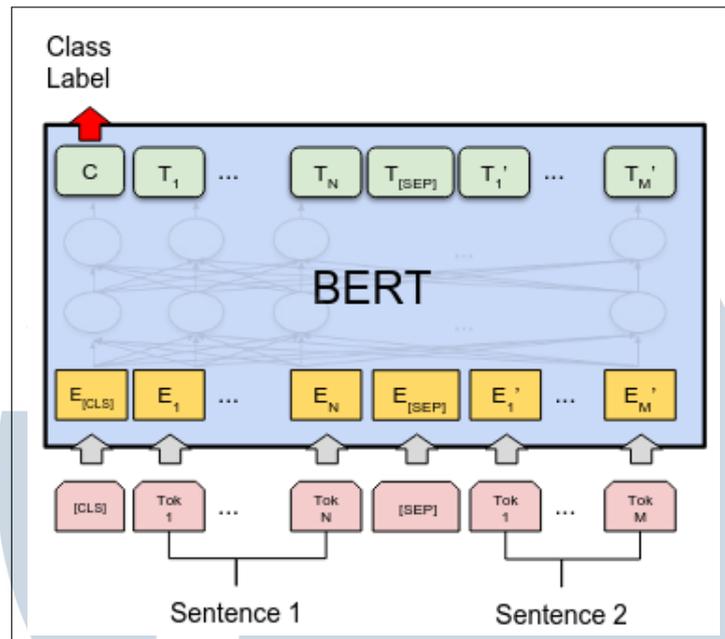
GEMP4RR!! KOALISII SUDAH PEC4H, PRABOWO DAN JOKOWI SALING SER4NG! ~
BREAKING NEWS
JOKOWI DAN PRABOWO RETAK
KINI SALING BONGKAR KASUS BESAR RI

2.2 BERT dan IndoBERT

2.2.1 BERT (*Bidirectional Encoder Representations from Transformers*)

Perkembangan model bahasa berbasis transformer telah merevolusi pendekatan *natural language processing*, dengan BERT (atau *Bidirectional Encoder Representations from Transformers*) sebagai salah satu tonggak penting dalam inovasi ini [13].

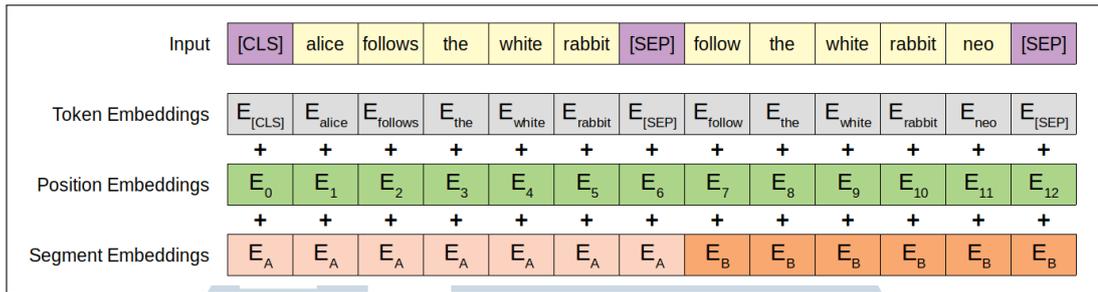
Model yang diperkenalkan oleh Google pada 2018 ini membawa paradigma baru dalam memahami konteks bahasa melalui mekanisme *bidirectional encoding* yang memungkinkan model memahami kata dalam keseluruhan konteks kalimat, seperti yang diilustrasikan pada Gambar 2.1. Arsitektur BERT dibangun atas prinsip transformer yang memungkinkan model menangkap hubungan kompleks antarkata melalui mekanisme *attention*. Tidak seperti model LLM lainnya yang memproses teks secara linier, BERT mampu memproses konteks kata dari kedua arah secara simultan. Hal ini memberikan kemampuan yang jauh lebih canggih dalam memahami nuansa semantik dan pragmatik bahasa.



Gambar 2.1. Ilustrasi BERT dalam proses *fine-tuning* untuk klasifikasi kalimat. Diambil di [13]

Model BERT menggunakan metode *pre-training* yang inovatif, yakni *masked language modeling* dan *next sentence prediction*. Dalam *masked language modeling*, beberapa kata dalam kalimat disembunyikan, dan model dilatih untuk memprediksi kata-kata tersebut berdasarkan konteks. Pendekatan ini memungkinkan model mengembangkan pemahaman mendalam tentang struktur dan makna bahasa. Dalam konteks deteksi hoaks, keunggulan BERT terletak pada kemampuannya menangkap pola linguistik yang kompleks. Penelitian yang dilakukan oleh Raza dan Ding menunjukkan bahwa model transformer seperti BERT secara signifikan mengungguli metode klasifikasi tradisional dalam mendeteksi karakteristik kebahasaan yang menjadi ciri khas informasi palsu [18].

Agar dapat diproses oleh BERT, maka data-data perlu ditransformasi menjadi format yang dipahami oleh BERT, yaitu dengan proses tokenisasi. Tokenisasi sendiri merupakan salah satu langkah *preprocessing* yang mengubah data tekstual mentah menjadi format yang dapat dipahami oleh model BERT seperti yang diilustrasikan pada Gambar 2.2. Proses tokenisasi melibatkan beberapa langkah sebagai berikut.



Gambar 2.2. Contoh proses tokenisasi pada BERT. Diambil di [19]

- **Subword Tokenization:** Berbeda dengan tokenisasi model lainnya, tokenizer pada model IndoBERT menggunakan *Subword Tokenization*. Pendekatan ini memecah kata-kata menjadi unit-unit yang lebih kecil dan bermakna yang dapat menangani kata-kata di luar kosakata dengan lebih efektif. Misalnya, kata-kata kompleks bahasa Indonesia dapat diuraikan menjadi komponen akar dan imbuhan, memungkinkan representasi linguistik yang lebih baik.
- **Special Token Handling:** Tokenizer BERT ini secara otomatis menambahkan token khusus yang penting untuk klasifikasi *sequence*:
 1. Token [CLS] (Klasifikasi): Disisipkan di awal setiap *sequence*
 2. Token [SEP] (Pemisah): Digunakan untuk memisahkan segmen teks input yang berbeda
- **Normalisasi *sequence*:** Untuk memastikan konsistensi di seluruh input, proses tokenisasi menerapkan dua teknik kunci:
 1. **Padding:** *Sequence* yang lebih pendek diisi hingga panjang yang seragam
 2. **Truncation:** *Sequence* yang lebih panjang melebihi panjang sekuens maksimum dipotong

Proses tokenisasi ini juga dapat direpresentasikan secara matematis seperti yang ada pada Rumus 2.1.

$$\text{Tokenize : Text} \rightarrow \text{Tokens} \subset \mathbb{Z}^n \quad (2.1)$$

Di mana:

- Text merepresentasikan artikel berita input
- Token adalah sekuens token yang dikodekan sebagai integer
- n merepresentasikan panjang sekuens maksimum

2.2.2 IndoBERT

IndoBERT merupakan varian khusus BERT untuk bahasa Indonesia yang dibangun oleh tim dari IndoNLP [14]. Melalui proses *pre-training* pada korpus berbahasa Indonesia yang luas, IndoBERT mampu menangkap nuansa spesifik bahasa, termasuk variasi dialek, struktur kalimat, dan konteks kultural yang unik. Studi Rahmawati et al. [20] membuktikan efektivitas IndoBERT dalam tugas klasifikasi, menunjukkan bahwa model ini dapat mencapai akurasi yang lebih tinggi dibandingkan model sebelumnya dalam mendeteksi hoaks berbahasa Indonesia. Keberhasilan ini tidak terlepas dari kemampuan model dalam memahami konteks dan struktur bahasa Indonesia secara mendalam. Penelitian mutakhir oleh Sibaroni et al. [21] semakin menegaskan potensi IndoBERT, menunjukkan bahwa model ini tidak hanya unggul dalam deteksi hoaks, tetapi juga mampu memberikan wawasan yang lebih komprehensif tentang karakteristik linguistik informasi palsu di media sosial Indonesia.

2.3 Evaluation Metrics

Evaluation metrics adalah alat ukur yang digunakan untuk menilai performa model dalam tugas klasifikasi tertentu. Dalam konteks deteksi berita hoaks, metrik ini digunakan untuk mengevaluasi seberapa baik model dalam mencapai tujuannya; dalam konteks penelitian ini yaitu dengan membedakan antara berita hoaks dan berita asli. Dengan menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *F₁-score*, peneliti dapat memahami berbagai aspek dari kinerja model, seperti tingkat prediksi yang benar, kemampuan model untuk mendeteksi semua hoaks, serta keseimbangan antara ketepatan dan sensitivitas deteksi.

Untuk memahami dasar perhitungan metrik-metrik tersebut, diperlukan pemahaman tentang *confusion matrix* sebagai komponen fundamental dalam evaluasi model klasifikasi. *Confusion matrix* adalah tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi terhadap nilai aktual. Dalam konteks deteksi berita hoaks, *confusion matrix*

terdiri dari empat komponen utama yang merepresentasikan kombinasi antara nilai prediksi dan nilai aktual, seperti yang ditunjukkan pada Tabel 2.1.

Nilai Aktual	Nilai Prediksi	
	Benar	Salah
Benar	TP	FN
Salah	FP	TN

Tabel 2.1. Struktur Confusion Matrix untuk Klasifikasi Berita Hoaks

di mana:

- *TP (True Positive)*: Jumlah data benar yang teridentifikasi dengan benar
- *TN (True Negative)*: Jumlah data salah yang teridentifikasi dengan benar
- *FP (False Positive)*: Jumlah data benar yang teridentifikasi dengan salah
- *FN (False Negative)*: Jumlah data salah yang teridentifikasi dengan salah

Berdasarkan *confusion matrix* tersebut, dalam konteks penelitian ini True Positive (TP) menunjukkan jumlah berita hoaks yang berhasil diidentifikasi dengan benar, True Negative (TN) merepresentasikan berita asli yang teridentifikasi dengan tepat, False Positive (FP) mengindikasikan berita asli yang salah diklasifikasikan sebagai hoaks, dan False Negative (FN) menunjukkan berita hoaks yang salah diklasifikasikan sebagai berita asli. Komponen-komponen ini menjadi dasar perhitungan berbagai metrik evaluasi yang digunakan dalam penelitian ini.

Proses evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap label kebenaran (*ground truth*) menggunakan data uji yang telah ditentukan sebelumnya. Metrik-metrik ini memberikan pandangan kuantitatif yang membantu dalam membandingkan berbagai model atau pendekatan yang digunakan dalam penelitian. Adapun perhitungan metrik-metrik ini adalah sebagai berikut.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

1. **Accuracy**

Metrik ini mengukur tingkat klasifikasi benar secara keseluruhan. Metrik ini dapat dihitung dengan menggunakan Rumus 2.2.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.2)$$

2. **Precision**

Metrik ini mengukur proporsi data yang teridentifikasi dengan benar. Metrik ini dapat dihitung dengan menggunakan Rumus 2.3.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.3)$$

3. **Recall**

Metrik ini menilai kemampuan model dalam mendeteksi semua data benar yang sebenarnya. Metrik ini dapat dihitung dengan menggunakan Rumus 2.4.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.4)$$

4. **F_1 -score**

Metrik F_1 adalah rata-rata harmonis dari presisi dan *recall*. Metrik ini memberikan gambaran tunggal yang seimbang antara presisi dan *recall* dan dapat digambarkan dengan menggunakan Rumus 2.5.

$$F_1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.5)$$

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A