

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan pesat media sosial telah mengubah cara manusia berkomunikasi, salah satunya melalui platform diskusi online seperti Reddit. Reddit menjadi wadah bagi pengguna untuk berbagi opini, berdiskusi, hingga menyampaikan kritik terhadap berbagai isu. Namun, komunikasi melalui teks di platform ini sering kali mengandung ambiguitas makna, terutama dalam bentuk sarkasme. Sarkasme merupakan suatu fenomena dalam komunikasi di mana makna literal dari sebuah teks atau ucapan berbeda dengan maksud atau sentimen yang sebenarnya ingin disampaikan. Dengan kata lain, sarkasme sering kali mengandung makna yang berlawanan dari apa yang tertulis [1]. Seiring dengan perkembangan teknologi, model berbasis deep learning seperti *Bidirectional Encoder Representations from Transformers* (BERT) memberikan pendekatan yang lebih canggih dalam memahami konteks bahasa. BERT mampu memproses kata dalam konteks dua arah (*bidirectional*), sehingga berpotensi lebih efektif dalam mendeteksi sebuah kalimat seperti sarkasme. [2]. Sarkasme menambahkan bentuk yang menarik dari masalah linguistik yang menarik dan tentu saja merupakan tugas yang menantang. Misalnya, perbedaan interpretasi antara setiap individu adalah salah satu contoh bagaimana sarkasme mengharuskan seseorang untuk dapat menentukan makna sebenarnya dari kalimat. Telah diamati bahwa sarkasme dan ironi merupakan bentuk-bentuk elegan dari penggunaan bahasa non-harfiah atau simbolis, di mana seseorang mengatakan atau menulis hal yang bertolak belakang dengan apa yang sebenarnya ingin mereka sampaikan. Karena ambiguitas yang disengaja ini, mendeteksi sarkasme terutama dalam komentar *social media* di mana petunjuk-petunjuk khas seperti intonasi suara atau ekspresi wajah tidak tersedia dan merupakan hal yang cukup menantang [3].

Hambatan utama dalam deteksi sarkasme otomatis terletak pada penggunaan idiom dan penentuan pengaruhnya. Idiom adalah ungkapan kiasan yang maknanya tidak dapat ditentukan hanya dari komponen kata penyusunnya secara harfiah. Cakupan *scope* sarkasme sangat krusial untuk menentukan apakah unsur sarkastik tersebut membalikkan polaritas sentimen di seluruh kalimat atau hanya pada kalimat atau frasa tertentu saja [4]. Kesalahan dalam mendeteksi cakupan sering kali

menyebabkan model salah mengkategorikan opini yang negatif sebagai opini yang positif, pada akhirnya mengganggu akurasi pengambilan keputusan.

Berikut merupakan penelitian terdahulu yang menggunakan BERT untuk mendeteksi sarkasme. Utpol Kanti Das, Resma Sultana Ani menggunakan model BERT pada komentar sosial media terutama dengan bahasa Bengali mencapai akurasi 99.60% tanpa augmentasi [5]. Pradeep kumar roy juga mendapatkan hasil akurasi yang besar mencapai 96.29% dengan dataset yang berbahasa *Hindi English* [6].

1.2 Rumusan Masalah

Mengacu pada latar belakang yang sudah dijelaskan, permasalahan yang dirumuskan meliputi:

1. Bagaimana penerapan model BERT dalam mendeteksi sarkasme pada komentar Reddit?
2. Bagaimana tingkat akurasi model BERT dalam mendeteksi sarkasme pada data Reddit?

1.3 Batasan Permasalahan

Mengacu para Rumusan Masalah yang sudah dijelaskan, batasan permasalahan yang dibuat dapat meliputi:

1. Dataset yang digunakan sudah di label, yang tersedia pada kaggle. Kumpulan komentar hanya diambil dari satu platform yaitu Reddit.
2. Jumlah komentar Reddit pada dataset sekitar 1.010.826 komentar.
3. Komentar dalam bentuk berbahasa Inggris
4. Penelitian ini dilakukan untuk membantu penelitian di masa depan agar dapat diterapkan dengan hasil nyata.

1.4 Tujuan Penelitian

Mengacu para rumusan permasalahan yang sudah dijelaskan, tujuan penelitian memiliki beberapa tujuan yaitu:

1. Menerapkan model BERT untuk proses mendeteksi sarkasme.
2. Mengevaluasi hasil model BERT dalam mendeteksi sarkasme menggunakan metrik yang relevan.

1.5 Manfaat Penelitian

1. Memberikan informasi yang lebih akurat dalam memahami opini publik pada *Social Media*.
2. Meningkatkan akurasi sistem analisis pada komentar di *Social Media*.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

- Bab 1 PENDAHULUAN
Merangkum latar belakang permasalahan mengenai deteksi sarkasme pada komentar platform reddit dan model BERT yang digunakan untuk deteksi sarkasme.
- Bab 2 LANDASAN TEORI
Bab ini berisikan teori-teori yang menjadi landasan penelitian, meliputi dari sarkasme, dan juga model BERT yang menjadi tahapan evaluasi.
- Bab 3 METODOLOGI PENELITIAN
Bab ini dituliskan metodologi penelitian dalam bentuk *flowchart*, metodologi yang digunakan antara lain studi literatur dan perancangan sistem
- Bab 4 HASIL DAN DISKUSI
Bab ini berisikan hasil dan proses dalam pembangunan model BERT untuk mendeteksi sarkasme dan juga membahas hasil performa model.
- Bab 5 KESIMPULAN DAN SARAN
Bab ini berisikan kesimpulan hasil penelitian dan saran untuk penelitian selanjutnya untuk pengembangan lebih lanjut.