BAB I

PENDAHULUAN

1.1. Latar Belakang

Era digital memungkinkan akses mudah ke lautan informasi, tetapi justru memicu disinformasi dan bias media. Masyarakat sering terjebak dalam gelembung informasi yang membatasi paparan pada konten sesuai minat/pandangan mereka, memperkuat bias konfirmasi (Rodrigo-Ginés dkk., 2024). Media memiliki peran penting dalam membentuk persepsi publik, tidak jarang media menyampaikan berita dengan bias yang secara halus mempengaruhi opini publik tanpa disadari (Han dkk., 2022). Bias media didefinisikan oleh para peneliti sebagai liputan berita yang condong atau bias internal yang tercermin dalam artikel berita. Bias media dapat muncul dalam berbagai bentuk seperti pemilihan kata, penyajian fakta, hingga pengabaian informasi tertentu (Ramadani dkk., 2024). Deteksi bias media menjadi suatu hal yang semakin penting, sistem pakar berbasis *NLP* telah menjadi alat utama untuk menganalisis dan memahami kompleksitas konten media. Deteksi bias media dapat ditangani sebagai masalah klasifikasi biner atau masalah klasifikasi *multi-class*. Dalam klasifikasi biner, tujuan utamanya adalah untuk memprediksi apakah sebuah konten bias atau tidak. Dalam klasifikasi *multi-class*, tujuannya lebih beragam, seperti: (1) derajat bias atau polarisasi, (2) sikap terhadap suatu peristiwa, atau (3) kompas politik (Rodrigo-Ginés dkk., 2024).

Classical Machine Learning dan *Deep Learning* banyak digunakan dalam deteksi/analisis bias media. Dalam pendekatan *Deep Learning*, metode ini sering kali mengungguli metode klasik dalam tugas yang memerlukan pemahaman

mendalam tentang konteks dan semantik. Model seperti *RNN* dan *transformer* menunjukkan kemampuan yang luar biasa dalam memahami semantik rumit dari bahasa alami. Dalam tugas deteksi bias media, model *transformer* umumnya memperoleh hasil yang lebih baik (Rodrigo-Ginés dkk., 2024).

Beberapa penelitian sebelumnya menggunakan model *Deep Learning*. (Spinde dkk., 2021) melatih model transformer sebelum fine-tuning pada dataset berlabel lemah berdasarkan sumber berita. (Krieger dkk., 2022) mengusulkan metode *pre-training* berbasis domain adaptif pada model *transformer*, model diadaptasi secara khusus pada data yang sangat relevan dengan domain bias (Wiki Neutrality Corpus). (Horych dkk., 2024) memperkenalkan pendekatan *Multi-Task Learning* dengan *pre-finetuning* pada 59 tugas *Large Bias Mixture*. (Baly dkk., 2020) mengusulkan metode *Adversarial Media Adaptation* dan *Triple Loss* pada model *transformer*. (Menzner & Leidner, 2024) membandingkan kinerja model *GPT-3.5*, *GPT-4* dan *Meta Llama2* dalam deteksi bias berita. (Spinde dkk., 2022) mengusulkan *transformer* yang dilatih melalui *Multi-Task Learning* pada enam dataset terkait bias.

(Murayama dkk., 2021) mengusulkan metode *masking* berbasis wikidata untuk mengurangi bias nama orang, meningkatkan generalisasi model terhadap data masa depan. (Hamborg, 2023) mengusulkan *Person-Oriented Framing Analysis* untuk mengidentifikasi frame berita dengan menggunakan *NLP*. (Jiang dkk., 2020) membandingkan kinerja model *CNN*, *RNN* dan *LDA-Transformer* dalam deteksi bias berita. (Rodrigo-Ginés dkk., 2023) mengusulkan metode *cascade transformer* dalam mendeteksi bias media. (W.-F. Chen dkk., 2020)

mengusulkan model *RNN* dalam mendeteksi bias media pada tingkat granularitas teks. (Hong dkk., 2023) mengusulkan metode *Multi-Head Attention* yang memperhitungkan struktur retorika dokumen dan semantik tingkat kalimat.

Meskipun berbagai model canggih telah digunakan dalam deteksi bias media dengan pendekatan yang berbeda-beda, kompleksitas dan relevansi isu pekerjaan terkait serta kemampuan dan kekurangan metode yang ada masih menunjukkan kesenjangan penelitian, diantaranya yaitu keterbatasan ukuran dataset dan peningkatan kinerja model yang masih menjadi salah satu kendala dalam deteksi bias media (Rodrigo-Ginés dkk., 2024).

Penelitian ini mengembangkan pendekatan yang lebih optimal dalam deteksi bias media. Penyempurnaan dilakukan terhadap metode sebelumnya, yaitu *RoBERTa* dan *DA-RoBERTa* (Krieger dkk., 2022). Upaya ini melibatkan augmentasi data berbasis parafrase dengan model *T5* dalam deteksi bias media menggunakan dataset *BABE* sebagai klasifikasi biner. Dataset *BABE* (Bias Annotations By Experts) (Spinde dkk., 2021) dipilih karena merupakan dataset bias media yang lengkap, berisi 3700 kalimat yang mencakup berbagai topik dan artikel berita, dilabeli oleh lima ahli dengan kesepakatan annotator, yang jauh lebih tinggi dibandingkan dengan korpus lainnya (Krieger dkk., 2022). Metode *RoBERTa* dan *DA-RoBERTa* dipilih karena *DA-RoBERTa* mendapatkan kinerja terbaik dibandingkan metode sebelumnya yaitu *RoBERTa* dan *RoBERTa – distant supervision* (Spinde dkk., 2021). Namun, tidak melebihi metode dari penelitian (Horych dkk., 2024). Berdasarkan pertimbangan reproduktibilitas dan ketersediaan

sumber daya, sehingga penelitian ini memfokuskan eksperimen pada model *RoBERTa* dan *DA-RoBERTa*.

(J. Chen dkk., 2023) menunjukkan bahwa augmentasi data dapat meningkatkan jumlah dan keragaman dari kumpulan data tertentu, dengan cara memanipulasi data untuk membuat data augmentasi tambahan. (Bird dkk., 2023) mengonfirmasi bahwa parafrase merupakan teknik augmentasi yang efektif dalam mengatasi kekurangan data pada *NLP*. Selain itu, model *T5* (Text-to-Text Transfer Transformer) menawarkan pendekatan terpadu untuk tugas transformasi teks dalam *NLP* dan mampu meningkatkan kinerja model pada beberapa model transformer. Model ini menggunakan pendekatan *encoder-decoder* yang memungkinkan transformasi *text-to-text* secara fleksibel termasuk parafrase. *T5* menghasilkan parafrase dengan struktur sintaksis dan kosakata beragam namun tetap mempertahankan makna semantik.

Penelitian ini bertujuan untuk mengoptimalkan kinerja model *RoBERTa* dan *DA-RoBERTa* dengan teknik augmentasi data berbasis parafrase menggunakan model *T5* dan memperoleh metode yang telah disempurnakan yaitu *RoBERTa/DA-RoBERTa* + Augmentasi Data: *All Task*.

1.2. Rumusan Masalah

Rumusan masalah yang dapat diuraikan dari latar belakang tersebut:

1. Bagaimana teknik augmentasi data berbasis parafrase dengan model *T5* dapat meningkatkan jumlah & keragaman data pada dataset *BABE*?

2. Apakah augmentasi data berbasis parafrase dengan model *T5* dapat meningkatkan kinerja model *RoBERTa/DA-RoBERTa* dalam deteksi bias media?

1.3. Tujuan Penelitian

Hasil yang ingin dicapai dengan adanya penelitian ini adalah:

1. Menganalisis pengaruh augmentasi data berbasis parafrase dengan model *T5* terhadap jumlah & keragaman data pada dataset *BABE*.
2. Mengevaluasi dampak augmentasi data berbasis parafrase dengan model *T5* terhadap kinerja model *RoBERTa* dan *DA-RoBERTa* dalam deteksi bias media, termasuk perbandingan kinerja model dengan/tanpa augmentasi data.

1.4. Manfaat Penelitian

Penelitian ini memiliki manfaat yang signifikan, baik secara akademis maupun praktis, antara lain:

1. Secara akademis, penelitian ini memberikan kontribusi terhadap kemajuan ilmu pengetahuan dan teknologi, khususnya dalam konteks *NLP* yang diterapkan pada deteksi bias media. Selain itu, penelitian ini memperkaya pemahaman tentang cara mengoptimalkan kinerja model dalam deteksi bias media.
2. Secara praktis, hasil dari penelitian ini dapat diterapkan untuk meningkatkan metrik evaluasi model dalam deteksi bias media, membantu lembaga riset, jurnalis, atau organisasi independen untuk mendapatkan hasil analisis yang lebih objektif terkait isu-isu sensitif terkait bias media.

1.5. Batasan Masalah

Pendekatan dalam penelitian ini menggunakan model *RoBERTa* dan *DA-RoBERTa* yang digunakan untuk mendeteksi bias media. Implementasi *MAGPIE* dan *distant supervision* tidak dapat direproduksi secara penuh karena ketiadaan kode sumber atau dokumentasi teknis yang memadai. Kemudian, model *T5* digunakan untuk augmentasi data latih dengan cara mereformulasikan kalimat dalam dataset melalui parafrase. Augmentasi data dilakukan dengan rasio 1:1 (satu parafrase baru per satu sampel asli) setelah proses seleksi, langkah ini bertujuan untuk mengoptimalkan efisiensi komputasi karena keterbatasan sumber daya GPU pada platform Kaggle. Selain itu, dataset yang digunakan terbatas pada dataset *BABE* dan hanya mencakup berita dalam bahasa Inggris. Penelitian ini tidak di implementasikan secara langsung pada aplikasi produksi, namun kesesuaian pendekatan akan dinilai berdasarkan metrik evaluasi yang dihasilkan.

1.6. Sistematika Penulisan

Laporan tugas akhir ini terdiri atas 5 bab, yaitu:

1. Bab I Pendahuluan

Menjelaskan mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

2. Bab II Landasan Teori

Menjelaskan definisi Bias Media, *Natural Language Processing* (NLP), *Deep Learning*, *RoBERTa*, *Augmentasi Data*, *T5*, *BABE* (Bias Annotations By Experts), penelitian terkait dan matriks penelitian dalam ranah *NLP* deteksi/analisis bias media.

3. Bab III Metodologi

Menjelaskan kontribusi terhadap *Artificial Intelligence Siliwangi* (AIS) dan metodologi penelitian.

4. Bab IV Hasil dan Pembahasan

Berisi pemaparan *Hardware* dan *Software* yang digunakan serta hasil eksperimen dan analisis.

5. Bab V Kesimpulan dan Saran

Menjelaskan hasil penelitian dan rekomendasi penelitian selanjutnya berdasarkan pengalaman hasil pengujian untuk meningkatkan hasil uji di penelitian selanjutnya.