

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis sentimen merupakan proses sistematis untuk mengidentifikasi emosi maupun opini dalam kalimat yang dihasilkan oleh publik. Proses ini dilakukan dengan mengklasifikasikan kalimat ke dalam tiga kategori, yaitu positif, negatif, dan netral [1]. Kekuatan utama dari pendekatan ini terletak pada kemampuannya mengolah ribuan opini tidak terstruktur di media sosial seperti platform X menjadi data informasi yang terukur [2]. Pendekatan ini relevan untuk diterapkan pada berbagai isu publik yang sedang berkembang dan menyediakan data opini yang cukup banyak untuk diolah, salah satunya adalah polemik terkait sengketa royalti musik.

Royalti musik dalam praktiknya dipengaruhi oleh implementasi kebijakan yang mengatur kewajiban pembayaran royalti bagi pelaku usaha yang memanfaatkan karya musik dalam kegiatan komersial [3]. Dalam sistem tersebut, Lembaga Manajemen Kolektif Nasional (LMKN) berperan dalam menghimpun dan mendistribusikan dana royalti kepada pemilik hak cipta yang terdaftar [4]. Namun, implementasi regulasi ini masih menghadapi berbagai kendala, seperti transparansi penyaluran dana dan pemahaman pelaku usaha terhadap kewajiban royalti musik [4]. Permasalahan tersebut tercermin dalam beberapa kasus, mulai dari keluhan musisi terkait rendahnya nominal royalti yang diterima [5], hingga pelaku usaha yang baru mengetahui kewajiban royalti setelah menerima surat tagihan atau somasi [6] [7].

Kondisi tersebut memicu munculnya berbagai opini publik di media sosial, khususnya pada platform X, yang mencerminkan beragam persepsi masyarakat terhadap regulasi royalti musik di Indonesia. Informasi tersebut penting untuk dianalisis karena dapat memberikan gambaran mengenai respons publik terhadap kebijakan yang diterapkan. Namun, pengolahan data secara manual sulit dilakukan karena banyaknya jumlah data yang dihasilkan. Sebagai gambaran, identifikasi awal pada penelitian sebelumnya menunjukkan terdapat sekitar 1.513 unggahan terkait isu royalti musik dalam kurun waktu tertentu [8]. Secara teknis, apabila satu data dianalisis secara manual dengan estimasi waktu 2 menit per teks, seorang peneliti membutuhkan waktu total sekitar 50 jam kerja hanya untuk proses pelabelan. Hal ini membuktikan bahwa pendekatan manual tidak efisien secara waktu dan memiliki tingkat subjektivitas yang tinggi [9]. Oleh karena itu, diperlukan pendekatan

analisis sentimen berbasis komputasional untuk mengolah data opini publik secara sistematis dan mengklasifikasikannya ke dalam kategori sentimen positif, negatif, dan netral.

Salah satu pendekatan yang relevan adalah model berbasis Transformer seperti Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT), yang merupakan varian dari Bidirectional Encoder Representations from Transformers (BERT) yang dilatih menggunakan korpus bahasa Indonesia dari berbagai domain [10]. Model ini efektif dalam memahami struktur kalimat kompleks, keberagaman bahasa, serta penggunaan bahasa gaul yang umum di media sosial, lebih baik dibandingkan metode tradisional [11] [12]. Pada penelitian sebelumnya, IndoBERT menunjukkan performa yang unggul dalam tugas analisis sentimen dan klasifikasi teks berbahasa Indonesia. Hasil penelitian tersebut menunjukkan bahwa IndoBERT mampu mencapai nilai *f1-score* sebesar 88,1%, yang mengindikasikan tingkat *accuracy* klasifikasi yang tinggi dibandingkan model pembelajaran mesin lainnya [13]. Selain itu, penelitian lain juga menunjukkan bahwa penerapan *fine-tuning* pada model IndoBERT dengan dataset spesifik dapat meningkatkan performa klasifikasi teks pada data di platform X. Dalam penelitian tersebut, model IndoBERT yang dikombinasikan dengan Convolutional Neural Network (CNN) dan Bidirectional Long Short-Term Memory (Bi-LSTM) pada platform X mampu mencapai *accuracy* hingga 83%, sehingga menunjukkan efektivitas pendekatan *fine-tuning* dalam analisis teks berbahasa Indonesia [14]. Oleh karena itu, penerapan *fine-tuning* pada model IndoBERT diperlukan agar model dapat menyesuaikan diri dengan konteks diskusi publik terkait royalti musik.

Berdasarkan uraian diatas, maka judul tugas akhir adalah “**Analisis Sentimen Pengguna X terhadap Sengketa Royalti Musik di Ruang Publik dengan Menggunakan Model IndoBERT.**”

1.2 Rumusan Masalah

Berdasarkan pada latar belakang yang sudah di uraikan, maka rumusan masalah untuk tugas akhir ini adalah:

1. Pengolahan opini publik dalam jumlah besar secara manual memiliki risiko subjektivitas yang tinggi dan potensi ketidakkonsistenan pelabelan (*human error*) yang dapat mengurangi *accuracy* hasil pemetaan sikap publik terhadap isu royalti musik.
2. Model IndoBERT memerlukan proses *fine-tuning* khusus agar dapat memahami konteks spesifik dan nuansa bahasa dalam domain kebijakan publik, khususnya terkait isu sengketa royalti musik.

1.3 Tujuan

Tujuan dari Penelitian ini adalah:

1. Meningkatkan objektivitas dan konsistensi dalam pemetaan sikap publik terhadap kebijakan royalti musik melalui penerapan model IndoBERT.
2. Melakukan proses *fine-tuning* pada model IndoBERT dengan menggunakan dataset spesifik bertema kebijakan publik untuk meningkatkan kemampuan model dalam mengenali nuansa bahasa, istilah teknis, dan konteks sengketa royalti musik.

1.4 Manfaat

Manfaat dari penelitian ini dapat:

1. Memberikan gambaran umum mengenai kecenderungan sentimen masyarakat terhadap kebijakan royalti musik di ruang publik.
2. Memberikan pendekatan analisis sentimen berbasis IndoBERT yang dapat membantu proses identifikasi opini publik di media sosial secara sistematis pada isu royalti musik.

1.5 Ruang Lingkup

Agar pembahasan dalam tugas akhir ini lebih terarah dan mencegah terjadinya perluasan masalah dan pembahasan yang terlalu kompleks, maka diperlukan ruang lingkup sebagai berikut:

1. Data teks bahasa Indonesia yang diambil dari platform X terkait opini mengenai kebijakan royalti musik, dan pencarian menggunakan kata kunci spesifik ('royalti musik', 'LMKN', dan 'kebijakan royalti').
2. Pengumpulan data dilakukan selama periode Desember 2024 hingga Desember 2025, dimulai dari momentum pengungkapan royalti rendah oleh Piyu (Rp125 ribu) yang menjadi simbol kegagalan sistem kolektif [5], dan diakhiri dengan evaluasi kebijakan royalti musik terhadap pembebasan biaya royalti bagi pelaku (Usaha Mikro, Kecil, dan Menengah) UMKM [15] [16].
3. Analisis difokuskan pada klasifikasi opini ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.
4. Pada penelitian ini, data dibagi menjadi data latih, data uji, dan data validasi dengan rasio 80:10:10 yang mengacu pada penelitian sebelumnya yang menunjukkan bahwa komposisi pembagian data tersebut menghasilkan performa yang lebih stabil dibandingkan dengan variasi pembagian data lainnya [17].

5. Parameter pelatihan model tidak dibatasi pada satu nilai tertentu, melainkan akan diuji dalam beberapa variasi untuk memperoleh performa optimal pada dataset penelitian. Parameter yang digunakan dalam proses *fine-tuning* meliputi:

- Learning Rate: $1e-5$, $2e-5$, $3e-5$, $4e-5$, dan $5e-5$ [18].
- Batch Size: 16, dan 32 [18].
- Epoch: 10, dan 15 [19].

Rentang parameter tersebut mengacu pada praktik umum dalam *fine-tuning* IndoBERT, namun pemilihan model terbaik akan didasarkan pada hasil evaluasi performa pada data penelitian ini.

6. Evaluasi performa model dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, presisi, *recall*, dan *f1-score* untuk masing-masing kelas sentimen.

