

BAB II

TINJAUAN PUSTAKA

2.1. Aplikasi LLMs dalam Domain Hukum

Penelitian terkini menunjukkan beragam pemanfaatan Large Language Models (LLMs) di bidang hukum. Sebagai contoh, dalam upaya meningkatkan akses bantuan hukum publik, model "LegalBot-EC" berbasis BERT berhasil mendemokratisasi informasi hukum di Ekuador dengan mencapai tingkat akurasi hingga 94% [21]. Di sektor hukum pajak dan tata kelola AI, evaluasi terhadap GPT-4 mengungkapkan potensi peningkatan efisiensi serta pengurangan biaya operasional, meskipun model tersebut masih memiliki keterbatasan dalam penalaran otonom yang dapat diandalkan sepenuhnya [22]. Studi lain yang membandingkan berbagai model (termasuk GPT-4 dan Mixtral) untuk otomatisasi klasifikasi persyaratan kepatuhan keamanan pangan dan GDPR menemukan bahwa BERT memberikan performa terbaik dengan F-score mencapai 87% [23]. Secara keseluruhan, temuan-temuan ini mengindikasikan bahwa LLMs mampu meningkatkan efisiensi secara signifikan di berbagai aspek hukum, namun masih memerlukan pengembangan lanjutan agar mencapai tingkat otonomi dan reliabilitas yang memadai untuk aplikasi kritis.

2.2. Tantangan dan Keterbatasan Model Generatif Saat Ini

Walaupun model generatif skala besar seperti GPT-4 menunjukkan kemampuan zero-shot yang mengesankan, sifat generatifnya sering kali memunculkan hallucination terhadap fakta hukum penting, seperti vonis pengadilan atau nomor perkara, yang berpotensi menimbulkan risiko malpraktik hukum [18]. Tantangan ini semakin diperparah oleh bias budaya dan linguistik akibat pre-training yang didominasi bahasa Inggris, sehingga model kurang optimal dalam menangani konteks hukum Indonesia yang kaya akan nuansa lokal [19]. Kondisi tersebut semakin rumit karena sistem yurisprudensi Indonesia bersifat hibrida menggabungkan hukum sipil Belanda, hukum Adat, dan hukum Islam yang menuntut ketepatan tinggi pada terminologi spesifik (misalnya

'Majelis Hakim', 'Pertimbangan Hukum', atau 'Amar Putusan') yang jarang tercakup secara memadai dalam benchmark umum.

2.3. Perbandingan Arsitektur Model Hukum: Encoder vs. Decoder

Penelitian ini membedakan diri dari studi sebelumnya melalui pilihan arsitektur model. Berbeda dengan pendekatan berbasis encoder seperti LegalBERT dan Lawformer yang utamanya difokuskan pada tugas klasifikasi dalam bahasa Inggris atau Mandarin [14], [20], Legal-Case LLM yang diusulkan menggunakan arsitektur decoder-only (Gemma 3) yang telah di-instruction-tuned untuk mendukung penalaran generatif. Pendekatan ini secara khusus dirancang untuk mengatasi tantangan unik yurisprudensi Indonesia, termasuk struktur dokumen formal dan terminologi hibrida. Secara metodologis, penelitian ini menerapkan pipeline data sintesis yang efisien secara parameter guna mengurangi risiko hallucination faktual. Hasil evaluasi menunjukkan peningkatan signifikan pada skor ROUGE (~0.555) dibandingkan baseline zero-shot (~0.11–0.24), serta menghasilkan model open-source pertama di domain ini yang transparan dan dapat direproduksi secara penuh.