

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Penelitian ini memberikan sumbangsih berupa model *Machine Learning* untuk salah satu tugas NLP, yaitu klasifikasi teks dalam mendeteksi kekerasan yang terkandung dalam teks pada media sosial berbahasa Indonesia. Model tersebut ialah model *Bidirectional Encoder Representations from Transformers* (BERT). Penelitian ini melatih empat varian model *pre-trained* berbasis BERT yang mampu menangani data teks berbahasa Indonesia. Secara umum, keempat model berbasis BERT mampu melampaui beberapa penelitian sebelumnya pada dataset dari media sosial tunggal. Dari hasil percobaan, dari antara empat model *pre-trained* yang ada, diperoleh model *pre-trained* terbaik, yaitu IndoBERT versi standar. Dengan menggunakan dataset kompilasi dari seluruh dataset publik yang tersedia, model *pre-trained* IndoBERT versi standar memperoleh hasil akhir: 0,8434 untuk akurasi, 0,8376 untuk *Precision* Makro, 0,8375 untuk *Recall* Makro, dan 0,8332 untuk *F1* Makro untuk model yang menggunakan dataset ternormalisasi. Sedangkan untuk model yang menggunakan dataset tak ternormalisasi, hasil akhir yang diperoleh meliputi 0,8286 untuk akurasi, 0,8205 untuk *Precision* Makro dan *Recall* Makro, serta 0,8163 untuk *F1* Makro. Hasil dari setiap variasi percobaan dapat dilihat selengkapnya pada Bab IV.

Pada proses pemodelan dengan dataset yang dimiliki, model dengan dataset yang ternormalisasi menghasilkan performa yang lebih baik dibandingkan model

dengan dataset tak ternormalisasi. Kondisi ini berbalik pada tahap implementasi atau *deployment*, di mana model dengan dataset tak ternormalisasi dapat mendeteksi kekerasan dengan lebih baik ketika diuji dengan beberapa sampel teks baru. Model dengan dataset tak ternormalisasi mampu memahami teks media sosial yang sifatnya memang bahasa sehari-hari dan tidak baku dengan lebih baik. Hal ini terbukti pada tingkat *False Negative* yang lebih rendah pada hasil prediksi model dengan dataset tak ternormalisasi dibandingkan model dengan dataset ternormalisasi pada 25 sampel teks baru. Sampel teks diperoleh dari proses pemindaian OCR dari gambar tangkapan layar (*screenshot*) yang diambil secara mandiri oleh peneliti. *Screenshot* tersebut berisikan teks yang dianggap mengandung kekerasan (berlabel positif).

Terkait dataset pelatihan, kelangkaan dan ketersebaran data masih menjadi isu selama ini. Sebagian besar penelitian melakukan pekerjaannya secara independen. Mereka membangun datasetnya sendiri-sendiri dan berbeda-beda antara satu penelitian dengan yang lainnya, ada juga yang tidak menyediakan dataset penelitiannya secara publik, membuat keterkaitan dan kontinuitas perkembangan penelitian di topik ini kurang jelas. Selain itu seluruh penelitian hanya menggunakan satu media sosial, di mana karakteristik setiap media sosial yang ada dapat beragam. Dengan begitu, generalisasi dari model yang dihasilkan juga dipertanyakan. Oleh karena itu, penelitian ini sudah membangun dataset kompilasi dari seluruh dataset yang tersedia secara publik terkait topik kekerasan di media sosial dalam teks bahasa Indonesia, dengan harapan bahwa dataset ini bisa menjadi dataset *benchmark* bagi penelitian-penelitian selanjutnya. Dataset yang

dibangun sejauh ini berasal dari dua *platform* media sosial, yaitu Instagram dan Twitter, sehingga diharapkan model memiliki informasi yang lebih kaya dengan penambahan variasi media sosial.

Selain itu, penelitian ini juga menghasilkan purwarupa sistem sederhana untuk pelaporan *online* terhadap kekerasan di media sosial dengan memanfaatkan model yang sudah dibangun dan mekanisme OCR untuk mendeteksi kekerasan dari teks yang terkandung dalam gambar tangkapan layar atau *screenshot*. Purwarupa ini dapat menjadi referensi untuk pengembangan sistem pelaporan ini ke tahap yang lebih lanjut.

## **5.2. Saran**

Hasil dari penelitian ini masih jauh dari standar cukup bagi sebuah sistem komputer untuk dapat melakukan evaluasi yang mendekati kemampuan manusia, dalam hal mendeteksi kekerasan di media sosial dalam bahasa Indonesia. Salah satu hal yang menjadi perhatian paling utama pada penelitian ini adalah kualitas dari dataset. Kualitas data sangat mempengaruhi performa model, ditambah setiap dataset yang dikumpulkan memiliki metode pelabelan yang berbeda-beda. Meskipun ada penelitian yang menghasilkan model dengan hasil yang sangat tinggi dibandingkan yang lain, dataset yang dimiliki sangat terbatas dalam hal konteks, misalnya hanya menyorot topik tertentu atau indikator yang dinilai “kekerasan” itu sendiri yang masih sempit pengertiannya. Hal ini menyebabkan generalisasi model masih dipertanyakan untuk menangani data baru. Oleh karena itu, kualitas dari dataset ini perlu dianalisis lebih lanjut. Hal ini bisa dimulai dengan percobaan menggunakan metode atau algoritma lain menggunakan dataset yang sama dengan

penelitian ini. Diharapkan penelitian berikutnya mampu memberikan hasil yang lebih baik dari penelitian ini menggunakan dataset yang sama.

Pengembangan dataset juga dibutuhkan untuk semakin meningkatkan kualitas model. Penambahan konteks dari hal yang dinilai sebagai “kekerasan” dapat dilakukan sehingga kekerasan itu tidak hanya dinilai dari keberadaan kata-kata kasar atau sejenisnya. Tidak menutup kemungkinan juga untuk penelitian selanjutnya dapat mengumpulkan dataset dari media sosial lain.

Untuk kelanjutan dari penelitian ini, metode yang digunakan disarankan menggunakan dataset yang tak ternormalisasi, sehingga tidak mengubah teks secara berlebihan pada data pelatihan, agar jika nanti digunakan untuk tahap produksi, dapat memprediksi data dengan lebih baik. *Language model* yang digunakan juga tidak selalu harus yang berukuran besar, hal ini terbukti misalnya pada performa yang dihasilkan oleh model IndoBERT versi *lite* hampir menyamai bahkan dapat melampaui versi standar di beberapa metrik. Selain itu, purwarupa sistem web yang sudah dirancang selanjutnya dapat diimplementasi atau di-*deploy* ke publik untuk dilakukan pengujian langsung kepada penggunanya. Hal ini ditujukan untuk mengevaluasi apakah sistem pelaporan kekerasan di media sosial secara *online* ini memang benar-benar dibutuhkan dan bermanfaat bagi masyarakat.