BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis komparatif yang telah dilakukan terhadap Analisis Sentimen Berbasis Aspek (ABSA) pada diskursus bencana alam di Sumatera, penelitian ini menghasilkan beberapa kesimpulan utama sebagai berikut:

1. Terdapat ketidakseimbangan kelas (imbalanced data) yang ekstrem pada dataset observasi, di mana sentimen negatif mendominasi secara absolut (933 data), sedangkan representasi kelas netral (59 data) dan positif (17 data) berada pada proporsi yang sangat marjinal. Fakta empiris ini merepresentasikan polarisasi opini yang sangat rentan memicu bias mayoritas (majority class bias), sehingga secara metodologis membutuhkan penanganan komputasional yang spesifik agar model klasifikasi tidak terdistorsi oleh kecenderungan overfitting pada kelas dominan.
2. Model arsitektur *Deep Learning* mutakhir (IndoBERT *Optimized*) terbukti superior dalam membedah dan mengklasifikasikan aspek kebencanaan (Bantuan, Korban, dan Pemerintah) dengan capaian F1-Macro 0.9546. Kemampuan mekanisme *self-attention* dalam menangkap relasi semantik kontekstual terbukti jauh lebih andal dibandingkan pendekatan partisi ekstraksi fitur TF-IDF pada model *Support Vector Machine* (SVM) yang hanya memperoleh skor 0.8800.
3. Terjadi kejanggalan yang menolak asumsi superioritas absolut *Deep Learning*. Pada klasifikasi sentimen yang sangat timpang, model konvensional SVM yang diintervensi dengan penyeimbangan data sintetis (SMOTE) justru mengungguli IndoBERT *Optimized* (F1-Macro 0.7151 berbanding 0.6719). Hal ini menyimpulkan bahwa arsitektur *Transformer* sangat rentan mengalami *overfitting* dan kelumpuhan prediksi pada dataset *imbalanced* tanpa adanya intervensi di tingkat arsitektur data, meskipun optimasi hiperparameter (penyesuaian *learning rate* dan *batch size*) telah diupayakan. Intervensi kualitas distribusi data terbukti lebih determinan daripada sekadar peningkatan kompleksitas algoritma.
4. Implementasi metode SHAP (*SHapley Additive exPlanations*) berhasil mendobrak sifat *black-box* dari model *Deep Learning*. Atribusi skor marjinal pada setiap fitur (kata) mampu memberikan justifikasi logis terhadap prediksi sentimen, sehingga faktor linguistik kritis yang memicu sentimen negatif publik

dapat teridentifikasi secara transparan dan terukur untuk mendukung evaluasi kebijakan mitigasi bencana.

5.2 Saran

Meninjau keterbatasan analitis dan temuan evaluatif selama proses penelitian, beberapa rekomendasi yang diusulkan untuk pengembangan pada penelitian selanjutnya adalah sebagai berikut:

1. Berangkat dari kelemahan model IndoBERT dalam menghadapi kelas minoritas, disarankan agar penelitian selanjutnya mengeksplorasi teknik *data augmentation* berbasis generatif atau mengadaptasi algoritma SMOTE ke dalam ruang vektor *embedding* (*latent space*) sebelum proses *fine-tuning* pada arsitektur *Transformer* dilakukan.
2. Untuk meminimalisir bias distribusi kelas secara natural, disarankan untuk memperluas rentang waktu akuisisi data dan menambah volume korpus. Selain itu, penggabungan sumber data lintas platform (seperti ekstraksi opini dari forum berita atau platform video pendek) dapat memberikan komparasi wacana publik yang lebih representatif dan menyeluruh.
3. Penelitian lanjutan dapat menyertakan kerangka kerja *Explainable AI* alternatif, seperti LIME (*Local Interpretable Model-Agnostic Explanations*), untuk dikomparasikan dengan SHAP. Hal ini bertujuan untuk memvalidasi konsistensi identifikasi faktor kritis dan menguji efisiensi komputasi dalam menafsirkan fitur linguistik pada teks bahasa Indonesia yang informal.