

## BAB 5

### SIMPULAN DAN SARAN

#### 5.1 Simpulan

Berdasarkan serangkaian proses implementasi dan evaluasi yang telah dilakukan dalam penelitian ini, maka diperoleh beberapa kesimpulan sebagai berikut:

1. Analisis performa model IndoBERT dalam mengklasifikasikan sentimen publik menunjukkan bahwa model bekerja dengan sangat baik, sehingga penelitian ini berhasil mengimplementasikan model IndoBERT dalam analisis sentimen publik terhadap kolaborasi NCT DREAM dan TosTos. Implementasi ini dilakukan melalui serangkaian tahapan komprehensif, mulai dari pra-pemrosesan data yang terverifikasi, hingga pengujian model dengan *dataset* baru yang tidak masuk dalam proses pembelajaran model untuk mengukur kemampuannya dalam klasifikasi sentimen.
2. Hasil evaluasi menunjukkan bahwa model terbaik yang dipilih adalah Model B yang dilatih dengan rasio 70:15:15, tanpa *hyperparameter tuning*, dan *class weight*. Model ini memiliki performa yang optimal dan stabil, dengan akurasi keseluruhan sebesar 83% pada data validasi. Model menunjukkan kinerja yang kuat pada sentimen negatif (*F1-score* 0.87) dan sentimen positif (*F1-score* 0.84), namun masih menghadapi tantangan pada kelas netral (*F1-score* 0.76). Kinerja yang lebih rendah pada kelas netral dapat dijelaskan oleh kompleksitas dan ambiguitas bawaan dari bahasa netral itu sendiri. Hal ini membuat model lebih sulit untuk menemukan pola yang jelas dibandingkan dengan sentimen positif atau negatif yang memiliki indikator linguistik lebih kuat. Terlepas dari distribusi *dataset* pelatihan yang tidak sepenuhnya seimbang, tantangan ini tetap menjadi faktor utama penyebab penurunan performa model pada kelas netral.
3. Analisis tren pelatihan lebih lanjut menguatkan pemilihan model B yang memiliki kestabilan *validation loss*, menunjukkan model berhasil mengelola potensi *overfitting*. Hal ini menjadi pertimbangan utama, karena model yang stabil lebih dapat diandalkan untuk melakukan generalisasi pada data yang belum pernah dilihat. Analisis kesalahan prediksi mengungkapkan bahwa meskipun kelas positif memiliki *F1-score* yang baik, ia memiliki jumlah

kesalahan prediksi terbanyak. Fenomena ini menunjukkan bahwa beberapa komentar positif mungkin mengandung nuansa bahasa atau konteks yang ambigu, sehingga model cenderung kesulitan membedakannya dari sentimen lain.

4. Setelah model dilatih dan dievaluasi, model digunakan untuk memprediksi sentimen pada *dataset* komentar yang tidak berlabel (total 5.030 komentar). Kemudian dikumpulkan 100 sampel acak dari hasil *dataset* yang sudah diberi label otomatis oleh model B untuk diberikan label manual oleh penulis dan *annotator* kedua. Kemudian dua hasil *labelling* ini dibandingkan. Performa model pada 100 sampel acak yang tidak berlabel menunjukkan adanya kelemahan lain pada kelas negatif, yang mengindikasikan sensitivitas model terhadap distribusi data yang berbeda.
5. Hasil prediksi ini mengungkapkan bahwa opini publik terhadap kolaborasi ini didominasi oleh sentimen positif (44.5%), diikuti oleh sentimen negatif (36.1%) dan netral (19.4%). Isu yang memicu kenaikan jumlah sentimen negatif berkaitan dengan besarnya jumlah *photocard* *zonk* berisi *quotes* yang ditemui oleh para konsumen. Selain itu, yang menjadi isu utama dalam kenaikan jumlah sentimen negatif adalah kontroversi terkait pengundian *fanmeet* yang terjadi pada bulan Agustus 2024, yang menggarisbawahi pentingnya komunikasi yang jelas dalam promosi.

## 5.2 Saran

Berikut beberapa saran yang dapat diberikan untuk peneliti selanjutnya.

1. Mengembangkan model sentimen untuk mendeteksi nuansa emosi yang lebih halus, seperti sarkasme, ironi, atau ambiguitas, terutama pada kelas netral.
2. Mengeksplorasi penggunaan *ensemble learning* atau perbandingan dengan model *deep learning* lain (selain IndoBERT) untuk melihat potensi peningkatan performa klasifikasi.
3. Memperkaya *dataset* dengan komentar dari platform media sosial lain (misalnya TikTok atau Twitter/X) untuk mendapatkan gambaran sentimen yang lebih komprehensif.

4. Melakukan analisis faktor-faktor yang secara spesifik memengaruhi *recall* dan presisi pada kelas sentimen minoritas untuk perbaikan model yang lebih terarah.
5. Mengintegrasikan analisis sentimen dengan analisis topik secara lebih mendalam untuk mengidentifikasi isu-isu spesifik yang memicu sentimen tertentu.

