

BAB V

SIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini menghasilkan sebuah sistem agregasi berita berbasis web yang secara otomatis mengklasifikasikan, mengelompokkan, dan meringkas artikel berita dari berbagai sumber. Sistem ini dibangun melalui serangkaian tahapan meliputi pengumpulan data dengan metode scraping, klasifikasi kategori berita menggunakan model BLSTM-2DCNN, pengelompokan topik dengan algoritma k-means, serta peringkasan menggunakan model BART yang telah dilakukan *fine tune*.

Model klasifikasi yang dikembangkan berhasil mencapai skor akurasi 86% dan F1 sebesar 0.85, yang menunjukkan bahwa model mampu mengklasifikasikan artikel berita dengan tingkat akurasi yang cukup tinggi. Namun, masih terdapat ketimpangan prediksi pada kategori tertentu, seperti kategori “inspirasi” yang tidak muncul pada hasil prediksi karena tidak seimbangya distribusi data pada saat pelatihan.

Evaluasi hasil pengelompokan topik menggunakan CH index menunjukkan nilai yang bervariasi, mulai dari 4.43 hingga 525.44, dengan rata-rata sebesar 80.87 dan median bernilai 40.25. Distribusi CH index yang condong ke kanan dapat disebabkan oleh rendahnya jumlah artikel atau kemiripan topik antar artikel dalam satu periode tertentu.

Untuk tahap peringkasan, proses *fine-tuning* yang dilakukan pada model BART berhasil meningkatkan skor ROUGE secara konsisten. Skor ROUGE-1 meningkat dari 0.5983 menjadi 0.6434, ROUGE-2 dari 0.5303 menjadi 0.5510, dan ROUGE-L dari 0.5599 menjadi 0.6017. Meskipun terjadi peningkatan, nilai tersebut masih tergolong moderat dan menunjukkan bahwa kualitas ringkasan belum sepenuhnya optimal untuk konteks berita berbahasa Indonesia.

Secara umum, sistem yang dibangun telah berhasil mengintegrasikan teknik klasifikasi, clustering, dan peringkasan dalam satu kesatuan proses yang saling mendukung. Sistem ini mampu menyajikan informasi berita secara terstruktur dan

ringkas, sehingga membuka peluang untuk eksplorasi berita secara lebih efisien oleh pengguna.

5.2 Saran

Terdapat beberapa hal yang dapat dilakukan sebagai pengembangan pada penelitian selanjutnya.

1. Performa klasifikasi dapat ditingkatkan dengan menyeimbangkan distribusi data pelatihan, khususnya pada kategori yang jarang muncul. Penggunaan model transformer seperti IndoBERT juga dapat menjadi alternatif untuk meningkatkan akurasi.
2. Meskipun terjadi peningkatan ROUGE setelah *fine-tuning* BART, hasilnya belum terlalu signifikan. Disarankan untuk mengeksplorasi model peringkasan khusus untuk bahasa Indonesia untuk meningkatkan kualitas ringkasan, serta memperluas pada aspek peringkasan multi dokumen.
3. Untuk melihat manfaat sistem secara utuh, perlu dilakukan uji coba sistem pada pengguna sesungguhnya, agar memperoleh umpan balik terhadap antarmuka, kualitas hasil, serta kemudahan pengguna.