

Abstrak— Bahasa Kailinese, yang digunakan di Provinsi Sulawesi Tengah, Indonesia, menghadapi tantangan karena penggunaan sehari-hari yang terbatas. Menurut artikel yang diterbitkan di situs brin.go.id, beberapa dialek Kailinese hanya memiliki empat penutur keluarga. Untuk mengatasi masalah ini, penelitian ini memperkenalkan model terjemahan mesin dan kumpulan data khusus untuk Kailinese. Sistem terjemahan kami menggunakan model IndoBART-V2, yang memungkinkan terjemahan yang lancar antara bahasa Indonesia dan Kailinese. Kami melakukan dua skenario pengujian: satu dengan kumpulan data ulasan yang beragam dan topik acak, dan satu lagi fokus pada kumpulan data tipe ulasan. Dengan menggunakan parameter default dan teknik prapemrosesan menggunakan "Kamus Leksikon Indonesia Kolokial," model terjemahan kami mencapai skor SacreBLEU yang mengesankan dibandingkan dengan tidak menggunakan pra-pemrosesan atau mengubah parameter default. Untuk skenario 1 (terjemahan dari bahasa Indonesia ke Kailinese), model ini mencapai skor 19,8, sedangkan untuk skenario 2 (terjemahan dari Kailinese ke bahasa Indonesia), skornya adalah 23,0. Pada skenario 2, di mana data latihan dan pengujian terdiri dari kalimat tipe ulasan, model ini mencapai skor 18,4 dan 22,7 untuk terjemahan dari bahasa Indonesia ke Kailinese dan dari Kailinese ke bahasa Indonesia, secara berturut-turut. Hasil ini menunjukkan efektivitas dan akurasi model yang dikembangkan. Selain itu, analisis kami mengungkapkan bahwa komposisi kalimat secara signifikan mempengaruhi kinerja model, tanpa perbedaan yang mencolok antara skenario 1 dan skenario 2. Hal ini menekankan pentingnya mempertimbangkan jenis kalimat dalam model terjemahan.

Kata kunci— mesin penerjemah, Bahasa indonesia, Bahasa kaili, indobart-v2