

Abstrak

Pada perkembangan teknologi yang pesat ini, masih sedikitnya ditemukan mesin penerjemah dari Bahasa daerah ke Bahasa Indonesia. Maka dari itu, Paper ini membahas tentang pembuatan mesin penerjemah statistik bahasa Muna ke bahasa Indonesia karena masih sedikitnya didapatkan mesin penerjemahan Bahasa Muna ke Bahasa Indonesia. Pendekatan yang digunakan berbasis statistik menggunakan corpus parallel. Pada penelitian ini, data yang diambil berasal dari buku berjudul Cerita Rakyat Buton Dan Muna di Sulawesi Tenggara dan beberapa artikel-artikel cerita rakyat yang ada di internet. Jumlah corpus parallel yang digunakan sebanyak 1050 baris kalimat dan corpus monolingual sebanyak 1351 baris kalimat. Adapun skenario yang akan dilakukan pada percobaan ini terbagi atas dua skenario. Skenario 1 adalah pengujian pada corpus parallel (training) yang diuji menggunakan baris kalimat yang tersedia dan baris kalimat tersebut akan ditambahkan pada setiap percobaan, sedangkan sisa dari baris kalimat yang dimiliki akan digunakan pada corpus parallel (testing). Pada skenario 2, pengujian dilakukan dengan melakukan perbandingan baris kalimat corpus monolingual setelah melakukan pengurangan dan penambahan baris kalimat. Agar skenario 2 berjalan, dibutuhkan akurasi pada skenario 1 yang paling terbaik. Pengujian dilakukan sebanyak 6 kali menggunakan tools BLEU (Bilingual Evaluation Understudy). Dari hasil pengujian yang dilakukan mendapatkan nilai akurasi terbaik yaitu sebesar 29,83%.

Kata kunci—Mesin Penerjemah Statistik, corpus parallel, corpus monolingual, BLEU