



ABSTRACT

Indonesian is the national language of Indonesia. However, the use of local languages in daily communication is still occur. Javanese is a local language that has a big number of speakers. There are so many people who come from outside Java living in Java. It causes the cross-language communication. An inability to communicate in different language can cause anxiety, which can disturb the concentration.

Translation machine is media that is expected to be the connector of cross-language communication. Rule-based approach is still being an option in developing translation machine. However, the approach has some lacks. Therefore, this study uses statistics-based approach to develop Indonesian and Javanese translation machine. Statistical approach is translation machine approach that use a probability scale based on the Bayes method.

Based on the testing, the highest evaluation value which is obtained by using BLEU is 55.09% for the Indonesian to Javanese translation and 49.68% for the Javanese to Indonesian translation. The evaluation value is also the highest value obtained after process of finding the best weight values in the parameter phrase translation table, language model, reordering model, and word penalty.

Keywords : Machine Translation, Statistical-Based, Bayes Formulation



INTISARI

Bahasa Indonesia merupakan bahasa nasional di Indonesia. Meski demikian, penggunaan bahasa daerah dalam proses komunikasi sehari-hari masih kerap terjadi. Dan Bahasa Jawa adalah bahasa daerah dengan penutur terbanyak. Hal ini disebabkan salah satunya karena banyak penduduk luar Pulau Jawa yang bermukim di Pulau Jawa. Hal tersebut memicu terjadinya komunikasi lintas bahasa. Dan ketidakmampuan berkomunikasi dengan bahasa yang berbeda disinyalir dapat menimbulkan kecemasan, yang pada akhirnya mengganggu konsentrasi.

Mesin penerjemah adalah media yang diharapkan mampu menjadi penghubung komunikasi lintas bahasa. Pendekatan berbasis aturan masih menjadi pilihan dalam mengembangkan mesin penerjemah. Namun, pendekatan tersebut memiliki beberapa kekurangan. Oleh karenanya, penelitian ini menggunakan pendekatan berbasis statistik untuk mengembangkan mesin penerjemah Bahasa Indonesia dan Bahasa Jawa. Pendekatan statistik adalah pendekatan pada mesin penerjemah yang menggunakan skala probabilitas yang berlandaskan pada metode Bayes.

Dari pengujian, nilai evaluasi tertinggi yang diperoleh dengan menggunakan BLEU adalah sebesar 55,09% untuk penerjemahan Bahasa Indonesia ke Bahasa Jawa dan 49,68% untuk penerjemahan Bahasa Jawa ke Bahasa Indonesia. Nilai evaluasi tersebut juga merupakan nilai tertinggi yang diperoleh setelah dilakukan proses pencarian nilai bobot terbaik pada parameter *phrase translation table*, *language model*, *reordering model*, dan *word penalty*.

Kata kunci -- Mesin Penerjemah, Berbasis Statistik, Formula Bayes