

MESIN PENERJEMAH BAHASA INDONESIA-JAWA MENGGUNAKAN METODE STATISTIK BERBASIS FRASA

JOHAN PRANATA

*Program Studi Teknik Informatika - S1, Fakultas Ilmu
Komputer, Universitas Dian Nuswantoro Semarang*

URL : <http://dinus.ac.id/>

Email : 111201207245@mhs.dinus.ac.id

ABSTRAK

Bahasa Jawa yang dulu merupakan bahasa yang besar, dengan bertambahnya waktu, penggunaannya semakin berkurang. Untuk mencegah terjadinya kepunahan bahasa atau biasa disebut dengan istilah kematian bahasa, maka sudah seharusnya masyarakat menyadari pentingnya pemeliharaan bahasa daerah. Di antaranya meliputi pemekaran kosakata dan kodifikasi berupa penyusunan kamus. Dalam era ini dengan pendekatan teknologi, kamus dapat dikembangkan dalam bentuk mesin penerjemah. Penelitian ini mengembangkan mesin penerjemah statistik berbasis frasa, yaitu yang merupakan suatu paradigma dari mesin penerjemah dimana penerjemahan dilakukan berbasis model statistik dengan parameter-parameter yang diturunkan dari analisis paralel korpus. Penelitian ini menggunakan 4.500 pasang kalimat sebagai korpus paralel yang bersumber dari Alkitab. Hasil dari implementasi mesin penerjemah statistik berbasis frasa ini diuji dan dievaluasi menggunakan evaluasi otomatis, BLEU yang menghasilkan nilai evaluasi sebesar 42,78%.

Kata Kunci : mesin penerjemah, statistik, bahasa Indonesia, bahasa Jawa, basis frasa

INDONESIAN-JAVANESE TRANSLATOR MACHINE USING STATISTICAL METHOD BASED ON PHRASE

JOHAN PRANATA

*Program Studi Teknik Informatika - S1, Fakultas Ilmu
Komputer, Universitas Dian Nuswantoro Semarang*

URL : <http://dinus.ac.id/>

Email : 111201207245@mhs.dinus.ac.id

ABSTRACT

Java language, which used to be a great language, with time, its use decreases. To prevent the extinction of a language or commonly referred to as the death of a language, then it should be public aware of the importance of maintenance of regional languages. Among these include the expansion of vocabulary and codification in the form of a dictionary compilation. In this era of technological approach, the dictionary can be developed in the form of machine translation. This study developed a phrase-based statistical machine translation, which is a paradigm of machine translation where translation is done based on a statistical model with parameters derived from the analysis of the parallel corpus. This research using 4,500 pairs of sentences as a parallel corpus derived from the Bible. The results of the implementation of phrase-based statistical machine translation is tested and evaluated using automated evaluation, BLEU which generates the evaluation value of 42.78%.

Keyword : translator machine, statistical, Indonesian, Javanese, phrase-based