Abstract

Research in the field of Natural Language Processing (NLP) is currently increasing especially with the arrival of a new term that is “big data”. The needs of the programming library that ready-to-use becomes very important to speed up the phases of research. Some libraries that have already been mature is available but generally for English language and its dependently. So, it can't be used for other languages. Stemming is one of the basic processes that exist in NLP. Indonesian stemming algorithm that often used is ECS (Enhanced Confix-Stripping). One of the libraries that already implemented the algorithm is Sastrawi. Results from the experiment show that the time of stemming processing by Sastrawi is still slow. Therefore, this research will optimize the speed of stemming processing using multiprocessing (MP). The data test are used in this research has manually taken from Wikipedia. The experiment results show that the MP technique can decrease the average time of stemming processing about 98.45%.
4. Abdi ansah A. and Winarko E., 2015, Question Classification Menggunakan Support Vector Machines dan Stemming, Seminar Nasional Aplikasi Teknologi Informasi (SNATI), 34

Index Terms

Computer Science Information Sciences

Keywords

Multiprocessing, Stemming, ECS, Sastrawi