

Abstrak

Saat ini banyak ahli dalam bidang teknologi informasi telah merancang dan mengembangkan algoritma untuk memecahkan masalah *stemming*, khususnya dalam bahasa arab. Namun dari sekian banyak analisa *stemming* dalam bahasa arab, belum ada standardisasi algoritma *stemming* yang baik dalam menganalisa akurasi teks pada Al-Quran. Pembangunan *stemming* pada Al-Quran merupakan suatu pekerjaan yang penting karena mendukung klasifikasi *sharaf* dalam Al-Quran guna memahami arti dari setiap kata pada Al-Qur'an. Salah satu *stemmer* atau algoritma *stemming* untuk mencari bentuk dasar dari suatu kata dalam bahasa arab ialah algoritma Khoja Stemmer. Cara kerja dari Khoja Stemmer ialah dengan mencoba untuk mencari akar pada suatu kata dalam bahasa arab dengan menghilangkan awalan terpanjang dan akhiran terpanjang suatu kata, lalu mencoba untuk menentukan akar dari kata yang tersisa menggunakan kamus akar kata. Dalam penelitian kali ini, Khoja Stemmer yang dibangun mampu menghitung rata-rata *stemming* pada Al-Quran sebesar 95,295%. Akan tetapi akar kata yang dihasilkan oleh Khoja Stemmer apabila di periksa secara manual masih terdapat beberapa kesalahan. Dengan demikian, dibutuhkan suatu kamus Al-Quran untuk menganalisa setiap hasil *stemming* yang dilakukan oleh Khoja stemmer dalam melakukan *stemming* pada Al-Quran.

Kata kunci: al-quran, stemming, khoja stemmer, sharaf, bahasa arab