

Abstrak

Saat ini sudah banyak digunakan berbagai jenis metode untuk ekstraksi frasa kunci. Terdapat dua kategori untuk algoritma frasa kunci diantaranya ekstraksi frasa kunci untuk dokumen tunggal dan ekstraksi frasa kunci untuk sekelompok dokumen.

Kategori ekstraksi frasa kunci yang dibahas pada tugas akhir ini adalah ekstraksi frasa kunci untuk sekelompok dokumen. Metode yang digunakan adalah metode *CorePhrase*. Diharapkan dengan menggunakan metode *CorePhrase* ini didapat frasa kunci yang sesuai dengan topik dan memiliki performansi yang baik.

Metode *CorePhrase* bekerja dengan mengiriskan dokumen menggunakan *graph-based model* kemudian *feature* dari tiap kandidat *keyphrase* yang diekstraksi dihitung lalu dilakukan perhitungan ranking terhadap *phrase feature* yang didapat untuk kemudian diperoleh nilai *top rank* nya.

Pada tugas akhir ini performansi dari tiap frasa kunci dapat diukur dengan berdasarkan *overlap* dan *precision*-nya. Berdasarkan percobaan yang dilakukan didapat nilai performansi *overlap* yang paling tinggi adalah sebesar 1 dan nilai performansi *precision* yang paling tinggi adalah sebesar 1. Dapat dikatakan bahwa performansinya sangat baik. *Keyphrase* yang dihasilkan sesuai dengan topik dokumen.

Kata kunci: *CorePhrase* , *ekstraksi*, *phrase feature*, *overlap*, *precision*