

Abstrak

Peringkasan multi-dokumen dibutuhkan agar pencarian informasi dapat berjalan lebih efektif dan efisien. Sistem peringkasan ini menggunakan *framework* ekstraksi yang telah secara luas digunakan untuk meringkas dokumen, dimana hasil ringkasan merupakan kalimat- kalimat yang telah diekstraksi dari dokumen. Peringkasan dokumen dapat dilakukan dengan melihat nilai informasi suatu kalimat dari dokumen tersebut berdasarkan fitur yang diterapkan pada kalimat yang ada. Fitur tersebut diantaranya adalah fitur linguistik dan statistik seperti posisi kalimat. *Conditional Random Fields* (CRF) merupakan salah satu model probabilistik untuk mengatasi segmentasi dan pemberian label pada data sekuens. CRF digunakan untuk mengkombinasikan fitur- fitur untuk mendapatkan model yang akan digunakan untuk menilai tingkat kepentingan kalimat. Fitur yang digunakan adalah *basic feature* dan *complex feature*. Sistem akan dinilai akurasi menggunakan *F-Measure*, dengan membandingkan ringkasan yang dibuat oleh sistem dengan ringkasan yang dibuat oleh ahli. Hasil pengujian menunjukkan bahwa rata- rata akurasi yang dihasilkan sebesar 62,5% dengan menentukan nilai *threshold* sebesar 0.45 dimana nilai tersebut digunakan untuk mengklasifikasikan kalimat *summary* dan *non-summary*.

Kata kunci : Peringkasan multi-dokumen, *Conditional Random Fields*, *Basic Feature*, *Complex Feature*, *F-Measure*