

Abstrak

Perkembangan teknologi informasi dan komunikasi yang pesat mengakibatkan data semakin berlimpah, namun kekayaan data tersebut justru membuat user kesulitan untuk memilah informasi yang berguna. Oleh sebab itu, saat ini user membutuhkan sistem peringkasan teks yang handal, yaitu sistem yang secara efektif mampu meringkas informasi yang ditemukan pada beberapa dokumen menjadi lebih pendek, namun tetap tidak kehilangan makna dari dokumen tersebut. Sehingga user tidak perlu menghabiskan waktu banyak untuk mencari informasi yang mereka butuhkan dalam sebuah artikel yang panjang, cukup dengan membaca *summary* dari artikel yang mereka temukan. *Centroid-Based Summarization* (CBS) adalah salah satu metode yang digunakan untuk peringkasan teks, baik itu untuk *single-document* maupun *multi-document*. Metode ini menghitung centroid dari setiap kalimat dalam dokumen dalam cluster, semakin tinggi nilai centroid yang dimiliki suatu kalimat berarti semakin mirip ia dengan topik dalam dokumen tersebut. Berkiblat dari penelitian yang dilakukan Dragomir R. Radev, seorang professor di bidang *Computer Science* di University of Michigan, penelitian ini mencoba menerapkan metode CBS untuk membuat *summary* dari sekumpulan dokumen berbahasa Indonesia dengan teknik *extractive summarization*. Sekumpulan dokumen tersebut akan dikelompokkan terlebih dahulu berdasarkan topiknya menggunakan teknik Single Pass Clustering menjadi cluster-cluster. Setelah itu setiap cluster dibuat ringkasannya menggunakan metode CBS. Yang kemudian akan dievaluasi dengan membandingkan hasil *summary* dari CBS dengan hasil *summary* oleh manusia menggunakan metode ROUGE-2 dan *Relative Utility*.

Kata kunci : *Centroid-Based Summarization, Single Pass Clustering, ROUGE-2, Relative Utility, multi-document, extractive summarization.*