

ABSTRAK

Jumlah dokumen berita di situs web universitas semakin banyak dan *up-to-date*. Kebutuhan pengunjung akan informasi tersebut semakin tinggi, sedangkan di sisi lain pengguna membutuhkan penemuan informasi yang cepat dan akurat. Penelitian ini menggunakan metode *cosine similarity* untuk menghitung kemiripan antar dokumen. Selain itu agar pencarian dokumen mudah ditemukan, maka dokumen diklasifikasikan berdasarkan pada kelas-kelas yang sudah ditentukan sebelumnya (*predefined*) menggunakan metode k-NN (*k-nearest neighbour*). Untuk mempercepat penyampaian hasil pencarian kepada pengguna maka pada aplikasi ini digunakan teknik *caching* terhadap dokumen-dokumen yang pernah dikembalikan sebelumnya kepada pengguna. Teknik ini dimaksudkan untuk meningkatkan transfer dokumen dengan menyimpan dokumen yang pernah diakses pada *cache* tersebut, sehingga apabila ada dokumen yang ingin diakses adalah dokumen yang sama maka akses akan dapat dilakukan lebih cepat. Hasil pengujian memperlihatkan bahwa akurasi klasifikasi k-NN untuk dokumen teks berita mampu mencapai 95% saat dokumen diwakili oleh judul dan paragraf pertama ($k=20, 25$ dan 45) tetapi hanya maksimal 40% saat dokumen diwakili oleh judul dan isi lengkap. Adanya klasifikasi terbukti meningkatkan presisi hasil pencarian dan mempersingkat proses pencarian tersebut. Kehadiran *cache* menjadikan waktu pemrosesan *query* yang telah ada sebelumnya menjadi jauh lebih cepat, rata-rata hanya memerlukan 7 mili detik.

Kata kunci : *k-Nearest Neighbour*, kemiripan cosinus, klasifikasi teks, *cache*, paragraf pertama