Clustering dokumen teks banyak diteliti karena peranan pentingnya dalam bidang text-mining dan information retrieval. Dalam algoritma clustering pemilihan fungsi jarak atau fungsi similaritas antar objek menjadi kunci keberhasilan algoritma. Pada fungsi jarak, jarak euclidean paling sering digunakan. Fungsi ini memiliki kelemahan jika digunakan untuk vektor berdimensi sangat tinggi yang menyebabkan kinerja clustering menurun. Alternatif dari fungsi jarak adalah fungsi similaritas, antara lain jaccard, dice, cosine dan pearson. Penelitian ini melakukan kajian tentang unjuk kerja fungsi jarak euclidean dengan empat fungsi similaritas tersebut di atas jika diterapkan untuk melakukan clustering dokumen teks berbahasa Indonesia. Dua pendekatan clustering yang dicobakan adalah pendekatan hierarchi dan partisi. Untuk pendekatan hierachi digunakan teknik aglomeratif dengan 2 metode similaritas cluster yaitu GroupAverage dan CompleteLink. Untuk pendekatan partisi juga dicobakan 2 metode, yaitu Bisecting K-Mean dan Buckshot. Koleksi dokumen yang digunakan 12 koleksi dokumen teks berita, yaitu dengan cacah dokumen 50, 100, 200, 300, 400, 500, 600, 700, 800, 1009, 1270 dan 1370 dokumen. Semua koleksi telah dilakukan clustering secara manual. Kriteria kinerja clustering diukur berdasarkan waktu komputasi dan validitas clustering. Untuk validitas digunakan nilai F-measure, yaitu nilai yang diturunkan dari Recall dan Precision yang mengukur kemampuan algoritma melakukan klasifikasi secara benar. Hasil penelitian menunjukkan bahwa hasil clustering terbaik adalah jika digunakan fungsi Cosine dengan rata-rata F-measure untuk seluruh koleksi 0,9313; sementara yang terburuk adalah jika digunakan fungsi jarak eucledian dengan rata-rata F-measure 0,4668. Secara waktu komputasi fungsi cosine juga memiliki kinerja tercepat dengan rata-rata 12,9 detik sedangkan terjelek adalah pearson dengan rata-rata 58,2 detik.