

ABSTRAKSI

Perkembangan teknologi saat ini menghasilkan perkembangan data yang sangat pesat, mulai dari bidang ekonomi, industri, pendidikan serta berbagai bidang lainnya. Institusi pendidikan saat ini juga memiliki kumpulan data diantaranya data mahasiswa, data pegawai, data tugas akhir dan data lainnya. Universitas Telkom memiliki kumpulan data tugas akhir terdiri atas dokumen tugas akhir dari seluruh jurusan yang ada di Universitas Telkom. Dokumen tersebut memiliki jumlah yang sangat besar sehingga sulit bagi mahasiswa atau dosen untuk mencari secara tepat dokumen yang mana yang diinginkan.

Oleh karena ini dibutuhkan suatu cara pengorganisasian dari dokumen-dokumen tersebut agar lebih terstruktur, yaitu dengan pengelompokan dokumen. *Clustering*, Salah satu algoritma yang dapat digunakan untuk mengelompokkan dokumen adalah *COBWEB*. *COBWEB* merupakan salah satu contoh algoritma *hierarchical* karena mengorganisasikan data menjadi *classification tree*. Berdasarkan hasil pengujian sistem pengelompokan dokumen menggunakan algoritma *COBWEB* didapat nilai *internal similarity* total akan berkurang jika jumlah dokumen yang dikelompokkan bertambah. Sedangkan nilai *internal similarity* pada suatu *cluster* akan bertambah jika jumlah dokumen yang dikelompokkan bertambah.

Kata Kunci: *Clustering, COBWEB, Hierarchical, Internal Similarity*

ABSTRACT

Current technological developments resulted in the development of data very rapidly, ranging from economics, industry, education and various other fields. Educational institutions today have a data set including student data, employee data, the final task data and other data. Telkom University have final project data set consists of the final project documents from all departments in the University of Telkom. The Document has a very large number making it difficult for students or lecturer to locate precisely which documents are required.

Therefore, it needed a way to organize these documents to be more structured, with grouping documents. Clustering, is one algorithm that can be used to grouping the document was a cobweb. Cobweb is one example of hierarchical algorithm for organizing data into a tree classification. Based on document grouping test results using cobweb algorithms, total internal similarity value obtained will be reduced if the number of grouped documents increases. While the internal similarity value to a cluster will increase if the number of grouped documents increases.

Keywords : *Clustering, COBWEB, Hierarchical, Internal similarity*