Abstrak

*Plagiarism (*plagiarisme) adalah menyalin sepenuhnya suatu paper milik orang lain tanpa menyertakan sumbernya (penulis aslinya) dan kemudian diakui sebagai paper miliknya. *Plagiarism* ini akan sangat merugikan bagi pemilik jika sampai dipublikasikan oleh orang yang tidak berhak.

Deteksi otomatis suatu plagiatisme adalah sesuatu yang sangat dibutuhkan dalam suatu lingkup akademisi. *Plagiarism* ini akan menjadi lebih kompleks jika dibuat dalam bentuk multilingual, yaitu teks asli ditulis ke dalam bahasa yang berbeda. Dari sebuah *cross-lingual perspective*, suatu fragmen teks dalam satu bahasa dianggap plagiat dari suatu teks dalam bahasa lain jika isi teks tersebut secara semantik sama.

Dalam Tugas Akhir ini, analisis *cross-lingual plagiarism* berdasarkan pada eksploitasi dari kamus bilingual statistical. Kamus ini dibentuk pada basis corpus paralel yang berisi original fragmen ditulis dalam satu bahasa dan versi plagiat dari fragmen tersebut ditulis dalam bahasa lain. Pembangunan kamus bilingual statistical yaitu dengan menggunakan *alignment* IBM Model-1 dan Algoritma Expectation Maximization. Kamus bilingual statistical merupakan inti dari penelitian ini.

Kata kunci: Plagiarism, Cross-Lingual Plagiarism, Corpus Paralel, Kamus Bilingual Statistical, Fragmen, *alignment* IBM Model-1, Algoritma Expectation Maximization.