

ABSTRAK

FAUZAN AZHARI RAHMAN, Pemeriksaan keaslian kode pada tugas *Capstone Project* umumnya masih dilakukan secara manual, sehingga tidak efisien dan rentan melewatkan kasus *plagiarisme* yang disamarkan melalui *refactoring* atau penggantian nama variabel. Penelitian ini bertujuan untuk menerapkan kombinasi metode *CodeBERT* dan *Winnowing Algorithm* dalam mendeteksi kemiripan kode sumber secara semantik dan tekstual, serta mengintegrasikan fitur tersebut ke dalam sistem pengumpulan tugas *Capstone Project* mahasiswa Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar. *CodeBERT* digunakan untuk mengukur kemiripan semantik melalui ekstraksi *embedding* vektor dan perhitungan *cosine similarity*, sedangkan *Winnowing Algorithm* digunakan untuk mengukur kemiripan tekstual berbasis *fingerprint k-gram* dengan indeks Jaccard. Kedua skor digabungkan menggunakan *weighted average* dengan bobot seimbang ($a = 0,5$), dan pasangan proyek diklasifikasikan ke dalam empat kategori yaitu *Plagiarisme Kuat*, Mirip Semantik, Mirip Tekstual, dan Normal. Pengujian dilakukan terhadap 17 proyek valid yang menghasilkan 136 pasangan unik. Hasil analisis menunjukkan 16 pasangan (11,76%) termasuk kategori *Plagiarisme Kuat*, 30 pasangan (22,06%) Mirip Semantik, 30 pasangan (22,06%) Mirip Tekstual, dan 60 pasangan (44,12%) Normal. Seluruh 11 skenario *black box testing* lulus dengan tingkat keberhasilan 100%, dan fitur analisis berhasil ditampilkan secara lengkap pada *dashboard* dosen maupun mahasiswa.

Kata kunci: Deteksi *plagiarisme*, kode sumber, *CodeBERT*, *Winnowing Algorithm*, kemiripan semantik, kemiripan tekstual, *Capstone Project*.