

0.1. Analisis Hasil

Dari pengujian yang dilakukan untuk *datasets no obfuscation*, *random obfuscation*, *translation*, dan *summary obfuscation*, sistem yang dibangun masih mengalami masalah dalam proses *merging*. Hal ini dikarenakan oleh banyak bagian pada dokumen yang diolah yang tidak termasuk plagiat dideteksi sebagai plagiat. Hal ini disebabkan oleh beberapa hal. Seperti fitur yang dihasilkan terlalu banyak dan saling berdekatan sehingga pada proses *merging* banyak fitur yang tidak diinginkan ikut terseret menjadi fitur yang dianggap plagiat. Jarak minimal yang terlalu besar juga membuat sistem yang dibangun terlalu sensitif, sehingga fitur yang di-*merge* semakin menumpuk. Hal ini dapat dilihat dari hasil pengujian dimana nilai *false positive* pada *datasets* diatas yang tinggi.

Untuk *datasets no plagiarism* sistem sudah mampu menangani tipe plagiat ini. Karena dokumen yang dideteksi plagiat hanya sedikit. Adanya dokumen yang dianggap plagiat diakibatkan karena masih terlalu banyak fitur yang dihasilkan oleh sistem.

Sedangkan untuk *datasets summary obfuscation* fitur yang dihasilkan oleh sistem untuk *datasets* ini terlalu sedikit, sehingga saat proses *merge* sistem tidak mampu fitur dari dokumen plagiat kedalam *cluster*. Sehingga nilai *false negative*, atau bagian plagiat yang dapat dideteksi oleh sistem kecil.