

## **Abstrak**

*Code clones* adalah porsi kode yang identic atau mirip dengan yang lain. Mendeteksi *code clones* dapat membantu menemukan redundansi di dalam kode. Masalah dalam mendeteksi *code clones* secara manual adalah hasil *code clone* yang dilaporkan tidak selalu konsisten. Salah satu cara untuk meningkatkan *F1-score* pada proses mendeteksi *code clones* adalah dengan otomasi proses mendeteksi *code clones*. Mendeteksi *code clones* otomatis menggunakan metode pendekatan token merepresentasikan *source codes* sebagai urutan token yang memungkinkan untuk mendeteksi *code clones* dengan struktur baris berbeda. Metode pendekatan token termasuk metode kelas berat karena memerlukan *language parser* khusus untuk setiap bahasa pemrograman berbeda. Penelitian ini bertujuan untuk membuat sebuah sistem pendeteksi *code clones* otomatis menggunakan metode pendekatan token untuk menganalisis dan membandingkan nilai *F1-score* dalam mendeteksi *code clones*. Penelitian ini dilakukan pada tiga *source codes* aplikasi yaitu *Apache*, *Postgre*, dan *Python* dengan hasil *F1-score* secara berturutan adalah “0,94”, “0,55”, “0,89”. Hasil penelitian ini mendapatkan *F1-score* tertinggi dibandingkan dengan tiga *tools* pendeteksi *code clones* lain, termasuk Simcad, Nicad, dan MeCC. Implementasi sistem pendeteksi *code clones* pada penelitian ini menunjukkan sistem dapat mendeteksi *code clones* dengan baik dibandingkan dengan tiga metode lain.

**Kata kunci :** *code clones, sistem pendeteksi code clones otomatis, metode pendekatan token.*