

ABSTRAK

Proses pembuatan *test case* merupakan bagian krusial dalam pengujian perangkat lunak, namun metode generasi otomatis yang berbasis model bahasa besar (Large Language Model) seperti GPT masih menghadapi tantangan dalam hal konsistensi hasil *test case*. Ketika diberikan *requirement* yang telah ditetapkan namun metode generasi otomatis yang berbasis model bahasa besar (*Large Language Model*) seperti GPT masih menghadapi tantangan dalam hal konsistensi hasil. Ketika diberikan *requirement* yang serupa, model sering kali menghasilkan *test case* dengan format, jumlah, dan struktur yang tidak stabil. Penelitian ini mengusulkan pendekatan untuk meningkatkan konsistensi pembangkitan *test case* dengan meningkatkan pemahaman model GPT terhadap *requirement* melalui proses pelatihan ulang model bahasa besar atau *Large Language Model* yakni model GPT menggunakan data *requirement* yang telah diproses menggunakan teknik Natural Language Processing (NLP).

Dataset yang digunakan untuk melatih model berasal dari *requirement fungsional* sistem manajemen rumah sakit. Model yang telah di latih ulang dengan proses *continual pretraining* akan dibandingkan dengan model dasar ChatGPT menggunakan beberapa metrik evaluasi, yaitu *requirement coverage*, *Jaccard Similarity*, dan hasil eksekusi otomatis menggunakan Selenium. Hasil penelitian menunjukkan bahwa model yang diberi pelatihan ulang menggunakan *requirement* mampu mencapai *requirement coverage* 100%, nilai *Jaccard Similarity* 0.78, serta menghasilkan *test case* yang konsisten dan langsung dapat dijalankan. Sebaliknya, model dasar menunjukkan ketidakstabilan dalam jumlah dan struktur *test case*. Penelitian ini membuktikan bahwa peningkatan pemahaman *requirement* dengan melakukan pelatihan ulang model GPT menggunakan *requirement* dapat meningkatkan konsistensi pembangkitan *test case* dalam proses otomatisasi pengujian perangkat lunak.

Kata Kunci: *test case generation, GPT, continual pretraining, NLP, requirement coverage, Selenium*