

ABSTRAK

Test-Driven Development (TDD) adalah pendekatan dalam pengembangan perangkat lunak yang fokus pada peningkatan kualitas sejak awal, dengan cara membuat pengujian terlebih dahulu sebelum menulis kode. Meski demikian, ada tantangan yang sering terjadi, seperti kesulitan dalam menulis pengujian yang efektif, sehingga menghambat proses TDD. Penelitian ini bertujuan untuk memanfaatkan *Large Language Models* (LLM) sebagai bantuan dalam proses TDD, terutama dalam menghasilkan pengujian dan kode berdasarkan deskripsi dalam bahasa alami. Penelitian ini mengeksplorasi kemampuan LLM, khususnya model LLaMA 3, dalam menghasilkan *test case* secara otomatis dari deskripsi kebutuhan. Untuk mengevaluasi efektivitas model, dua skenario digunakan, yaitu deskripsi pendek dan panjang, serta dilakukan penilaian berdasarkan metrik otomatis. Penilaian dilakukan menggunakan metrik CodeBLEU, BLEU, dan chrF untuk mengukur tingkat kesamaan antara *test case* hasil dan referensi manual. Hasil penelitian menunjukkan bahwa skenario dengan deskripsi yang lebih lengkap memberikan skor CodeBLEU sebesar 32,20% dan chrF sebesar 44,34%. Selain itu, model juga menghasilkan *output* yang konsisten dalam setiap percobaan. Penelitian ini menunjukkan bahwa LLM dapat digunakan untuk mendukung proses TDD secara langsung tanpa dilakukan pelatihan tambahan, dan tetap mampu menghasilkan *test case* yang relevan.

Kata Kunci: generasi tes, *large language model*, *test driven development*.