

ABSTRAK

Penggunaan AI sudah sangat marak digunakan, dan salah satu yang terdampak adalah pada ranah pengembangan aplikasi. Github Copilot yang merupakan salah satu *tools* berbasis AI yang dapat mempermudah pengembangan aplikasi serta meningkatkan efisiensi dalam pengembangan aplikasi. Namun, tingkat efisiensi pada Github Copilot juga bergantung dengan model AI yang digunakannya. Penelitian ini bertujuan untuk membandingkan secara kuantitatif efisiensi dari model AI GPT 4.1 dan Claude 3.7 Sonnet menggunakan Github Copilot dengan Mode Agent dengan studi kasus untuk melakukan pengembangan aplikasi *backend* REST API. Penelitian ini menggunakan pendekatan studi komparatif dengan melakukan eksperimen membuat beberapa fitur yang dikembangkan secara paralel pada masing-masing model. Pengukuran yang digunakan pada penelitian ini adalah menggunakan metode Goal Question Metric (GQM) dengan Metrik yang diukur adalah jumlah *prompt* dan durasi yang dibutuhkan oleh setiap *prompt* untuk menyelesaikan tiap tugas. Kesimpulan dari penelitian ini memperlihatkan adanya pengaruh antara kecepatan dengan model yang digunakan untuk *code generation*. GPT 4.1 menunjukkan keunggulan dari model ini yang dilihat dari durasi yang dibutuhkan pada model tersebut yang lebih cepat serta jumlah *prompt* yang dilakukan yang lebih sedikit. Akan tetapi, Claude 3.7 lebih efektif dalam menangani permintaan pembuatan kode yang lebih besar pada satu kali *prompt*. Dengan catatan, durasi *prompt* yang dilakukan cenderung lebih lama.

Kata Kunci: Generative AI, LLM4SE, Github Copilot, Rest API, Backend