

## ABSTRAK

ANUGRAH AKBAR PRARAMADHAN, 10118969

### APLIKASI PEMODELAN BAHASA KAUSAL GPT-2 DALAM PEMILIHAN DIKSI BAHASA INDONESIA MENGGUNAKAN BAHASA PEMROGRAMAN PYTHON.

Tugas Akhir, Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Gunadarma, 2023.

Kata Kunci: Bahasa Pemrograman *Python*, *Benchmark Dataset Indo4B*, *Bilingual Language Understudy (BLEU)*, *GPT-2*, Pemodelan Bahasa Kausal.  
(xvi + 94 + Lampiran)

*Generative Pre-trained Transformer 2 (GPT-2)* ukuran *Medium* merupakan salah satu dari jenis arsitektur pemodelan bahasa besar (*large language model*) yang dikembangkan oleh OpenAI pada tahun 2019 dengan total parameter sejumlah 345 Juta. Model berbasis *transformer* ini terkenal akan kemampuan pembelajaran tanpa arahnya (*unsupervised learning*) dari suatu *unstructured data* dalam mengenali pola data dengan cukup baik. Oleh karena itu, penelitian ini bertujuan untuk membuat aplikasi berbasis web menggunakan bahasa pemrograman *Python* dalam memecahkan objektif pemodelan bahasa kausal terhadap pemilihan diksi bahasa Indonesia yang dilatih menggunakan *GPT-2 Medium* pada *benchmark dataset Indo4B*. Penelitian ini menggunakan metode *waterfall* dalam pembuatan aplikasi, terdiri dari tahap *requirement, analysis, design, development, testing, deployment*, dan *maintenance*. Berdasarkan hasil pengujian aplikasi didapatkan nilai *Bilingual Language Understudy (BLEU)* terhadap model dari objek uji A dengan nilai *mean* dalam persen sebesar 66.16. Penelitian ini berhasil mempertahankan nilai *BLEU* di ambang batas  $>50$  sebagai acuan bahwa model ini mampu dalam menghasilkan kalimat yang fasih dan mampu dalam memilih diksi bahasa Indonesia dengan sangat baik sesuai kaidah Kamus Besar Bahasa Indonesia (KBBI). Dikarenakan pelatihannya menggunakan *benchmark dataset Indo4B* yang bersifat umum, maka nilai kebaruan yang dihasilkan oleh penelitian ini berhasil dalam melakukan adaptasi kemampuan yang serupa dengan penelitian pendahulu. Model penelitian ini mampu dalam melakukan pembuatan artikel jurnalistik secara sebagian, memberikan informasi kalimat informal dan formal, menghasilkan langkah pembuatan makanan, melanjutkan suatu cerita, menguraikan definisi atau istilah umum, dan tanya jawab sederhana dengan catatan masih ke dalam objektif pemodelan bahasa kausal yang akan meneruskan dari suatu kalimat masukan.

Daftar Pustaka (2017-2023)