

Adaptasi Pretrained Language Model untuk Pembangkitan Komentar dari Data Replay Counter-Strike = Adapting Pretrained Language Model for Commentary Generation from Counter-Strike Replay Data

Eko Julianto Salim, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920541983&lokasi=lokal>

Abstrak

Pembangkitan komentar dari data terstruktur adalah tugas yang menantang dalam bidang pembangkitan bahasa alami dengan berbagai potensi aplikasi dalam bidang olahraga dan esports. Penelitian-penelitian sebelumnya sering berfokus pada penyesuaian model datake-teks khusus, namun eksperimen awal penulis menemukan bahwa pretrained language model menunjukkan performa yang lebih baik dalam pembangkitan komentar. Penelitian ini berfokus pada bagaimana pretrained language model dapat diadaptasi untuk menghasilkan performa yang lebih baik pada permasalahan pembangkitan komentar menggunakan data replay Counter-Strike. Penulis menemukan bahwa penyesuaian pretrained language model dapat meningkatkan kualitas semantik dan alami dari pembangkitan komentar. Namun, penulis juga menemukan bahwa peningkatan kualitas semantik dan alami ini didampangi dengan penurunan pada kualitas leksikal akibat tantangan intrinsik yang ada dalam pembangkitan komentar.

.....Generating commentary from structured data poses a significant challenge in natural language generation, with extensive applications in sports and esports domains. Previous research has predominantly centered on tailoring specialized data-to-text models for this task. However, our preliminary investigation indicates that pretrained language models exhibit superior performance in commentary generation. This study focuses on further adapting pretrained language models to enhance their suitability for esports commentary generation, specifically utilizing Counter-Strike replay data. Our findings reveal that through this adaptation, pretrained language models can augment the semantic quality and naturalness of generated commentary. Nevertheless, we observed a trade-off, as the improvement in semantic quality and naturalness was accompanied by a decline in lexical quality, owing to the inherent complexities involved in commentary generation.