

Abstrak

Long Short-Term Memory (LSTM) sequence-to-sequence telah banyak digunakan untuk menyelesaikan sejumlah tantangan dalam peringkasan teks. Namun, model ini masih memiliki dua masalah utama: kemunculan kata diluar kosakata atau sering disebut *out-of-vocabulary (OOV)* dan perulangan kata. Pada makalah ini, *pointer generator* dan *coverage weighting* diusulkan untuk mengatasi kedua masalah tersebut. Hal ini dimulai dengan model *sequence-to-sequence* dasar. Kemudian, dikembangkan dengan *attention mechanism*, yang telah ditambahkan *coverage weighting* pada perhitungannya untuk mengurangi terjadinya perulangan kata, dan mengganti *encoder* menjadi *bi-directional LSTM*. Setelah itu, *pointer generator* diimplementasikan agar dapat menunjuk kembali ke kata-sumber dan menghasilkan kata jika bertemu dengan kata OOV. Menggunakan artikel berita bahasa Inggris CNN/Daily Mail sebagai *dataset* dan *ROUGE score* sebagai metrik evaluasi, model yang diusulkan menghasilkan sejumlah ringkasan yang lumayan mirip dengan ringkasan yang dibuat para ahli. Model ini memberikan peningkatan relatif dari 18% hingga 34% dari model *standard attention mechanism*. Bagaimanapun, model ini memerlukan kompleksitas relatif lebih tinggi hingga 40%, di mana model usulan perlu lebih banyak waktu proses (selama 7 hari) dibanding model standar (hanya 5 hari).

Kata Kunci: peringkasan teks abstraktif, *coverage weighting*, *long short-term memory*, *pointer generator*, *sequence-to-sequenc*

