

DAFTAR TABEL

Tabel 2.1.1 Tabel Perbandingan Penelitian.....	6
Tabel 2.1.2 Penelitian yang Dilakukan	9
Tabel 2.4.1 Waktu translasi dalam detik untuk newstest-2017 (3,004 kalimat, 76,501 source BPE tokens) untuk arsitektur dan batch size berbeda. (Junczys-Dowmunt et al., 2018).....	77
Tabel 3.2.1 List dalam <i>command</i> “marian” untuk proses dalam <i>encoder</i> yang digunakan	87
Tabel 3.2.2 List dalam <i>command</i> “marian-decoder” untuk proses dalam <i>decoder</i> yang digunakan	94
Tabel 4.2.1 Contoh isi korpus paralel training di sumber Opus (Wikimedia)	101
Tabel 4.2.2 Contoh isi korpus paralel training di sumber WMT’21 (QED & TED2020).....	102
Tabel 4.2.3 Contoh isi korpus paralel validasi “dev” di FLORES-101	103
Tabel 4.2.4 Contoh isi korpus paralel validasi “devtest” di FLORES-101.....	104

DAFTAR GAMBAR

Gambar 2.3.1 Overview dari arsitektur NMT	12
Gambar 2.3.2 GNMT	15
Gambar 2.3.3 ConvS2S	15
Gambar 2.3.4 Transformer	15
Gambar 2.3.5 Matriks term-document untuk 4 kata pada karya Shakespeare	17
Gambar 2.3.6 Vektor Co-occurrence untuk korpus Wikipedia	18
Gambar 2.3.7 Visualisasi spasial untuk vektor kata untuk <i>digital</i> dan <i>information</i>	18
Gambar 2.3.8 Demonstrasi grafis untuk <i>cosine similarity</i>	21
Gambar 2.3.9 Embeddings yang dipelajari oleh model skipgram	23
Gambar 2.3.10 Sebuah unit saraf	25
Gambar 2.3.11 <i>Feedforward network</i> sederhana dengan 2 layer	26
Gambar 2.3.12 <i>Forward inference</i> dalam sebuah <i>feedforward neural language model</i>	29
Gambar 2.3.13 Computation graph untuk fungsi $L(a,b,c) = c(a+2b)$	33
Gambar 2.3.14 <i>Backward computation</i> pada satu node e	34
Gambar 2.3.15 <i>Backward pass</i> untuk <i>computation graph</i> fungsi $L(a,b,c) = c(a+2b)$	35
Gambar 2.3.16 Contoh <i>computation graph</i> untuk 2-layer <i>neural network</i> sederhana	36
Gambar 2.3.17 Melakukan pembelajaran penuh kembali ke embeddings	38
Gambar 2.3.18 Simple recurrent neural network dari (Elman, 1990)	40
Gambar 2.3.19 <i>Simple recurrent neural network</i> diilustrasikan sebagai sebuah <i>feedforward network</i>	41
Gambar 2.3.20 <i>Forward inference</i> dalam sebuah simple recurrent network	42
Gambar 2.3.21 Recurrent neural network sederhana yang ditampilkan <i>unrolled</i> dalam waktu	43
Gambar 2.3.22 Training RNN sebagai Language Model	46
Gambar 2.3.23 Stacked Recurrent Networks	48
Gambar 2.3.24 Satu unit LSTM ditampilkan sebagai komputasi grafis	53
Gambar 2.3.25 <i>Basic neural units</i> yang digunakan dalam <i>feedforward</i> , <i>simple recurrent networks</i> (SRN), dan <i>long short-term memory</i> (LSTM)	54
Gambar 2.3.26 Arsitektur <i>encoder-decoder</i>	55
Gambar 2.3.27 Menerjemahkan satu kalimat (waktu inferensi) dalam versi RNN dasar dari pendekatan <i>encoder-decoder</i> ke terjemahan mesin	57
Gambar 2.3.28 Bentuk versi mendekati formal untuk menerjemahkan sebuah kalimat untuk arsitektur <i>encoder-decoder</i>	58
Gambar 2.3.29 Mengizinkan setiap <i>hidden state</i> dari <i>decoder</i>	58
Gambar 2.3.30 Melatih pendekatan <i>encoder-decoder</i> RNN dasar untuk terjemahan mesin	60
Gambar 2.3.31 Mekanisme <i>attention</i>	62
Gambar 2.3.32 Sebuah sketsa jaringan <i>encoder-decoder</i> dengan <i>attention</i>	63
Gambar 2.3.33 Pohon pencarian untuk menghasilkan string target $T = t_1, t_2, \dots$ dari <i>vocabulary</i> $V = \{\text{yes, ok, <s>}\}$	65
Gambar 2.3.34 <i>Beam search encoding</i> dengan <i>beam width</i> $k = 2$	66
Gambar 2.3.35 Penskoran untuk <i>decoding beam search</i> dengan <i>beam width</i> $k = 2$.	

.....	67
Gambar 2.3.36 Perbedaan bentuk kata dan <i>wordpieces</i>	69
Gambar 2.3.37 Contoh penyelarasan antara kalimat dalam bahasa Inggris dan Prancis	72
Gambar 2.4.1 Perbandingan RNNSearch & Nematus untuk algoritma <i>phase decoder</i>	78
Gambar 2.4.2 Visualisasi grafis untuk pencarian pada penerjemahan bahasa DE→EN "Hallo Welt!" dengan <i>beam size</i> 3.....	81
Gambar 2.5.1 Statistik dari data dalam FLORES-101	82
Gambar 3.2.1 Kebutuhan kompilasi MarianNMT dalam Linux.....	85
Gambar 3.2.2 <i>Cloning repository</i> MarianNMT	85
Gambar 3.2.3 Bagian pemroses dalam toolkit MarianNMT	86
Gambar 3.2.4 Skenario penggunaan <i>validation set</i> untuk evaluasi hasilnya dari <i>training set</i> . (Google Developer, 2022)	96
Gambar 3.2.5 Penggunaan <i>Sacrebleu</i> dengan file referensi lokal.	98
Gambar 4.2.1 Kode untuk menggabungkan file teks berbeda menjadi satu	101
Gambar 4.3.1 <i>Command</i> untuk melakukan pengecekan sederhana terhadap MarianNMT dan <i>Sentencepiece</i>	105
Gambar 4.3.2 Keluaran pengecekan dari gambar 4.3.1	105
Gambar 4.3.3 Tiga file yang digunakan untuk menggunakan MarianNMT	105
Gambar 4.3.4 Kode untuk inisialisasi file-path dari MarianNMT pada script ...	105
Gambar 4.4.1 Konfigurasi model dalam <i>encoder</i> dengan “marian-train” pada script “run-me.sh”	106
Gambar 4.4.2 <i>Logging</i> dari proses inisialisasi pada MarianNMT	107
Gambar 4.4.3 Proses training pada MarianNMT	107
Gambar 4.4.4 Grafik perbandingan hasil training dengan skor “bleu-detok”	109
Gambar 4.4.5 Grafik perbandingan hasil training dengan skor “ce-mean-words”	110
Gambar 4.5.1 Kode dalam file “decode-me.sh”	110
Gambar 4.6.1 Nilai BLEU yang didapat dari modul <i>Sacrebleu</i> untuk terjemahan model dari sumber data Wikimedia	111
Gambar 4.6.2 Nilai SpBLEU yang didapat dari modul <i>Sacrebleu</i> untuk terjemahan model dari sumber data Wikimedia.....	112
Gambar 4.6.3 Nilai BLEU yang didapat dari modul <i>Sacrebleu</i> untuk terjemahan model dari sumber data QED&TED	112
Gambar 4.6.4 Nilai SpBLEU yang didapat dari modul <i>Sacrebleu</i> untuk terjemahan model dari sumber data QED&TED	113
Gambar 4.6.5 Grafik perbandingan hasil terjemahan dengan skor “BLEU”	113
Gambar 4.6.6 Grafik perbandingan hasil terjemahan dengan skor “SpBLEU”..	114
Gambar 4.6.7 Perbandingan skor sebagai referensi terhadap skor dalam SpBLEU (WMT’21 “Large-Scale Multilingual Machine Translation”).....	114