UNIVERSITAS
GADJAH MADA
Context Convolutional Neural Network dan Protein Language Model untuk Prediksi Struktur Sekunder Protein
Ririth Coshy Yah, Afiahayati, S.Kom., M.Cs., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN PENGESAHAN ..... **ii**
PERNYATAAN BEBAS PLAGIASI ..... **iii**
DAFTAR ISI ..... **iv**
DAFTAR TABEL ..... **vii**
DAFTAR GAMBAR ..... **viii**
INTISARI ..... **x**
ABSTRACT ..... **xi**
BAB I **1**
1.1 Latar Belakang 1
1.2 Perumusan Masalah 4
1.3 Batasan Penelitian 4
1.4 Tujuan Penelitian 5
1.5 Manfaat Penelitian 5
1.6 Sistematika Penulisan 5
BAB II **7**
BAB III **17**
3.1 Asam Amino 17
3.2 Protein 17
3.3 Struktur Sekunder Protein 18
3.4 Prediksi Struktur Sekunder Protein 19
3.5 Representasi dan *Embedding* Asam Amino 19
3.6 Model Bahasa Protein (*Protein Language Model*) 20
3.7 Model Ankh 21
3.7.1 Arsitektur Inti Ankh 21
3.7.1.1 Konfigurasi Layer (48 *Encoder* / 24 *Decoder*) 21
3.7.1.2 *Relative Positional Embedding* 21
3.7.1.3 Fungsi Aktivasi (Gated-GELU) 22
3.7.1.4 *Trade-off* Kedalaman dan Lebar Jaringan 22
3.7.2 Strategi Optimisasi Pra-Pelatihan 22
3.7.2.1 *Masking Strategy – 1-Gram Span De-Masking* 22
3.7.2.2 Probabilitas *Masking* 20% 23
3.8 Pembelajaran Mesin Mendalam 23

UNIVERSITAS GADJAH MADA
Context Convolutional Neural Network dan Protein Language Model untuk Prediksi Struktur Sekunder Protein
Ririth Coshy Yah, Afiahayati, S.Kom., M.Cs., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.9 *Context Convolutional Neural Network* 23
3.9.1 Lapisan Konvolusi 24
3.9.2 Konvolusi Dilatasi 26
3.9.3 Fungsi Aktivasi *Rectified Linear Unit* (ReLU) dan *Softmax* 26
3.9.4 Strategi Pelatihan Model 27
3.9.4.1 *Stochastic Gradient Descent (SGD) Optimizer* 27
3.9.4.2 *White Noise* 28
3.9.4.3 *Learning Rate* 28
3.9.4.4 Regularisasi L2 29
3.9.4.5 *Label Masking* 29
3.9.4.6 *Dropout* 29
3.9.4.7 AdamW *Optimizer* 30
3.10 Metrik Evaluasi 31
3.10.1 *Cross Enthropy Loss* 31
3.10.2 Akurasi Q3 dan Q8 32
3.10.3 *Confusion Matrix* 33
3.10.4 *Precision, Recall,* dan *F1-score* 33
BAB IV 35
4.1 Deskripsi Umum Penelitian 35
4.2 Analisis Permasalahan 36
4.3 Akuisisi dan Pengumpulan Data 37
4.4 Rancangan Arsitektur Metode 38
4.4.1 Pra-pemrosesan *Dataset* 39
4.4.2 Ekstraksi *Embedding* 40
4.4.3 Model *Context Convolutional Neural Network* 42
4.4.3.1 Layer Konvolusi Standar 42
4.4.3.2 Layer Konvolusi Dilatasi 43
4.4.3.3 *Concatenation* 44
4.4.3.4 Reduksi Dimensi 44
4.4.3.5 Layer Konvolusi 8 *Channel* 44
4.4.3.6 *Concatenation* Final 45
4.4.3.7 Reduksi Dimensi Kedua 45
4.4.3.8 Layer Klasifikasi 45
4.4.3.9 Pelatihan Model 46

UNIVERSITAS GADJAH MADA
Context Convolutional Neural Network dan Protein Language Model untuk Prediksi Struktur Sekunder Protein
Ririth Coshy Yah, Afiahayati, S.Kom., M.Cs., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

4.5 Rancangan Pengujian Model dan Analisis 48
BAB V 50
5.1 Lingkungan Penelitian 50
5.2 Instalasi dan Inisialisasi Pustaka 50
5.3 Inisialisasi Lingkungan Eksperimen 51
5.4 Inisialisasi Dataset dan Pra-Pemrosesan 52
5.5 Ekstraksi *Embedding* oleh Ankh 55
5.6 Pengembangan Model *Context Convolutional Neural Network* 59
5.7 Pelatihan Model 61
5.8 Pengujian Model 65
5.9 *White Noise* dan *Label Masking* 66
BAB VI 68
6.1 Ekstraksi *Embedding* Ankh 68
6.2 Analisis Hiperparameter 69
6.3 Pemilihan Konfigurasi Hiperparameter 75
6.4 Evaluasi Performa Model pada Dataset Uji 76
6.4.1 Performa Model pada Q8 dan Q3 76
6.4.2 Analisis Performa Per-Kelas dan Misklasifikasi 78
6.4.2.1 Analisis Pada Model Q8 78
6.4.2.2 Analisis Pada Model Q3 83
6.5 Pengaruh Penggunaan *White Noise* dan *Label Masking* 87
6.5.1 Performa Model pada Q8 dan Q3 87
6.5.2 Analisis Performa Per-Kelas dan Misklasifikasi dengan *White Noise* dan *Label Masking* 89
6.5.2.1 Analisis Pada Model Q8 89
6.5.2.2 Analisis Pada Model Q3 93
6.6 Waktu Inferensi pada Panjang Sekuens yang Berbeda 95
6.7 Perbandingan dengan Hasil Model Penelitian Lain 97
BAB VII 99
DAFTAR PUSTAKA 102