



**PENDEKATAN REKAYASA FITUR PADA PREDIKSI INTERAKSI
SENYAWA *Cyanthillium cinereum* SEBAGAI INHIBITOR
PROTEIN VENOM *Calloselasma rhodostoma***

SAID THAUFIK RIZALDI



**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Pendekatan Rekayasa Fitur pada Prediksi Interaksi Senyawa *Cyanthillium cinereum* sebagai Inhibitor Protein *Venom Calloselasma rhodostoma*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Said Thaufik Rizaldi
M0503241002

RINGKASAN

SAID THAUFIK RIZALDI. Pendekatan Rekayasa Fitur pada Prediksi Interaksi Senyawa *Cyanthillium cinereum* sebagai Inhibitor Protein Venom *Calloselasma rhodostoma*. Dibimbing oleh WISNU ANANTA KUSUMA, TOTO HARYANTO, dan FAJAR SOFYANTORO.

Gigitan ular berbisa dapat menyebabkan kerusakan jaringan, gangguan pembekuan darah, kelumpuhan, dan kematian. Antivenom merupakan terapi utama, tetapi memiliki keterbatasan cakupan spesies dan kemampuan menetralkan kerusakan lokal. Senyawa bioaktif dari tumbuhan berpotensi menjadi kandidat inhibitor terhadap protein bisa ular. Penelitian ini bertujuan mengevaluasi rekayasa fitur dan model pembelajaran mesin untuk prediksi *drug-target interaction* (DTI), mengidentifikasi interaksi senyawa *Cyanthillium cinereum* dengan protein venom *Calloselasma rhodostoma* dari empat populasi, serta menganalisis pasangan prioritas secara molekuler.

Data pelatihan diperoleh dari *Snakebite Envenoming Medicines Database* dan setelah praproses menghasilkan 102 pasangan interaksi positif. *Putative negative* dibangkitkan menggunakan GANATCpair untuk membentuk data pelatihan seimbang. Rekayasa fitur mencakup Morgan *fingerprint*, representasi graf senyawa, komposisi asam amino, *convolutional neural network* (CNN), dan *feature-wise linear modulation* (FiLM). Sembilan model dievaluasi menggunakan *nested cross-validation*, skenario *unseen*, *confusion matrix*, dan *ablation study*.

Model terbaik, yaitu GINE-FiLM untuk representasi senyawa dan CNN-FiLM untuk representasi protein, memberikan hasil paling konsisten. Pada *nested cross-validation*, model ini menghasilkan performa terbaik dengan *F1-score*, AUROC, dan AUPR masing-masing sebesar 0,920, 0,945, dan 0,965. Model tersebut digunakan untuk memprediksi DTI pada populasi Borneo, Jawa, Malaysia, dan Thailand. Pada masing-masing populasi, 100 pasangan dengan probabilitas tertinggi dipilih untuk *molecular docking*. Prediksi DTI dan *molecular docking* menghasilkan tujuh pasangan prioritas yang melibatkan N-Cyclohexylidene-2-carbamylcyclohex-1-enylamine (251559) dan 7-heptadecynoic acid (22600366). Skor *docking* paling negatif diperoleh pada pasangan protein phospholipase B-like A0A9F2NW96 dengan senyawa 251559, yaitu sebesar $-9,056$ kcal/mol.

Untuk simulasi *molecular dynamics*, dipilih dua pasangan dari tujuh pasangan prioritas tersebut, yaitu A0A9F2NW96–251559 dan T2HQ57–22600366. Pemilihan dilakukan berdasarkan skor *docking*, relevansi biologis protein target, serta kedekatan residu pada hasil visualisasi interaksi. Simulasi selama 100 ns menunjukkan bahwa kedua kompleks mengalami penyesuaian lokal tanpa perubahan besar pada struktur global. Perhitungan MM/PBSA menghasilkan energi pengikatan total masing-masing sebesar $-27,9813$ dan $-37,8049$ kcal/mol. Hasil tersebut menunjukkan bahwa kedua pasangan memiliki afinitas pengikatan yang baik secara komputasional dan berpotensi menjadi kandidat inhibitor protein bisa ular untuk divalidasi lebih lanjut secara eksperimental.

Kata kunci: *drug-target interaction*, pembelajaran mesin, *putative negative*, sawi langit, *venom* ular



SUMMARY

SAID THAUFIK RIZALDI. A Feature Engineering Approach to Predicting Interactions of *Cyanthillium cinereum* Compounds as Inhibitors of *Calloselasma rhodostoma* Venom Proteins. Supervised by WISNU ANANTA KUSUMA, TOTO HARYANTO, and FAJAR SOFYANTORO.

Venomous snakebites can cause tissue damage, blood coagulation disorders, paralysis, and death. Antivenom is the primary treatment, but it has limitations in species coverage and the neutralization of local tissue damage. Plant-derived compounds may serve as potential inhibitors of snake *venom* proteins. This study aimed to evaluate feature engineering and machine learning models for drug-target interaction (DTI) prediction, identify interactions between *Cyanthillium cinereum* compounds and *Calloselasma rhodostoma* venom proteins from four populations, and analyze the prioritized pairs at the molecular level.

The training data were obtained from the Snakebite Envenoming Medicines Database and yielded 102 positive interaction pairs after preprocessing. Putative negatives were generated using GANATCpair to construct a balanced training dataset. Feature engineering included Morgan fingerprints, compound graph representations, amino acid composition, convolutional neural networks (CNN), and feature-wise linear modulation (FiLM). Nine models were evaluated using nested cross-validation, unseen scenarios, a confusion matrix, and an ablation study.

The best-performing model, combining GINE-FiLM for compounds and CNN-FiLM for proteins, provided the most consistent results. In nested cross-validation, this model achieved an F1-score of 0.920, an AUROC of 0.945, and an AUPR of 0.965. The model was then used to predict DTI in the Borneo, Java, Malaysia, and Thailand populations. For each population, the 100 pairs with the highest probabilities were selected for molecular docking. DTI prediction and molecular docking identified seven priority compound–protein pairs involving N-Cyclohexylidene-2-carbamylcyclohex-1-enylamine (251559) and 7-heptadecynoic acid (22600366). The most negative docking score was obtained for the phospholipase B-like protein A0A9F2NW96 paired with compound 251559, with a score of -9.056 kcal/mol.

Among the seven pairs, A0A9F2NW96–251559 and T2HQ57–22600366 were selected for molecular dynamics simulations based on their docking scores, the biological relevance of the target proteins, and the proximity of residues observed in the interaction visualization. The 100 ns simulations showed that both complexes underwent local adjustments without major changes in global structure. MM/PBSA calculations yielded total binding energies of -27.9813 and -37.8049 kcal/mol, respectively. These findings suggest that both compound–protein pairs are promising candidates for further experimental validation as potential snake *venom* inhibitors.

Keywords: *Cyanthillium cinereum*, drug-target interaction, machine learning, putative negative, snake *venom*



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



PENDEKATAN REKAYASA FITUR PADA PREDIKSI INTERAKSI SENYAWA *Cyanthillium cinereum* SEBAGAI INHIBITOR PROTEIN VENOM *Calloselasma rhodostoma*

SAID THAUFIK RIZALDI

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Magister Ilmu Komputer

**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Tesis:
Dr. Sony Hartono Wijaya, S.Kom., M.Kom.



Judul Tesis : Pendekatan Rekayasa Fitur pada Prediksi Interaksi Senyawa *Cyanthillium cinereum* sebagai Inhibitor Protein *Venom Calloselasma rhodostoma*

Nama : Said Thaufik Rizaldi
NIM : M0503241002

Disetujui oleh

Pembimbing 1:
Prof. Dr.Eng. Wisnu Ananta Kusuma, S.T., M.T.



Pembimbing 2:
Dr. Toto Haryanto, S.Kom., M.Si.



Pembimbing 3:
Dr. Fajar Sofyantoro, S.Si., M.Sc.

Fajar Sofyantoro

Diketahui oleh

Ketua Program Studi:
Dr. Toto Haryanto, S.Kom., M.Si.
NIP 198211172014041001



Dekan Sekolah Sains Data, Matematika, dan Informatika:
Prof. Dr. Ir. Agus Bueno, M.Si., M.Kom.
NIP 196607021993021001



Tanggal Ujian:
31 Juli 2026

Tanggal Lulus:



PRAKATA

“Dan Dia ajarkan kepada Adam nama-nama semuanya, kemudian Dia perlihatkan kepada para malaikat, seraya berfirman, ‘Sebutkan kepada-Ku nama semua benda ini, jika kamu benar!’”

(QS. Al-Baqarah [2]: 31)

Segala puji hanya milik Allah *subhanahu wa ta’ala* atas segala pertolongan, hidayah, dan hikmah-Nya sehingga seluruh rangkaian penelitian hingga penyusunan karya ilmiah ini dapat diselesaikan dengan baik, *qadarullah*. Penelitian dalam bidang bioinformatika ini dilaksanakan sejak November 2025 hingga Agustus 2026 dengan judul “Pendekatan Rekayasa Fitur pada Prediksi Interaksi Senyawa *Cyanthillium cinereum* sebagai Inhibitor Protein *Venom Calloselasma rhodostoma*”.

Penulis menyampaikan terima kasih kepada Prof. Dr.Eng. Wisnu Ananta Kusuma, S.T., M.T., Dr. Toto Haryanto, S.Kom., M.Si., dan Dr. Fajar Sofyantor, S.Si., M.Sc. selaku komisi pembimbing yang telah memberikan bimbingan, arahan, serta dukungan teknis dan nonteknis yang sangat berguna bagi penulis selama penelitian dan penyusunan tesis. Secara khusus, penulis mengucapkan terima kasih sebesar-besarnya kepada program DAPT EQUITY IPB University melalui skema Riset Kolaborasi Indonesia (RKI) atas fasilitasi penggunaan *high performance computing* (HPC) NVIDIA DGX H100 yang mendukung kelancaran penelitian ini.

Ungkapan terima kasih penulis sampaikan kepada Abah Said Daryusman, Ibu Muryani, S.Pd., dan adik penulis, Syarifah Nurbaiti Azzahra, yang senantiasa memberikan kasih sayang, motivasi, dan doa sehingga penulis dapat menjalani pendidikan magister dengan lancar. Penghargaan juga disampaikan kepada Beasiswa Pendidikan Indonesia (BPI) melalui skema Beasiswa Indonesia Maju (BIM) dari Kementerian Pendidikan Tinggi, Sains, dan Teknologi yang bekerja sama dengan Lembaga Pengelola Dana Pendidikan (LPDP) atas dukungan pendidikan dan pengembangan diri selama penulis menempuh studi magister.

Terima kasih penulis sampaikan kepada Bapak Mustakim, S.T., M.Kom., selaku dosen pada masa studi sarjana, atas rekomendasi dan dukungan untuk melanjutkan pendidikan di IPB University. Penulis juga berterima kasih kepada sahabat *Puzzle Research Data Technology* (Predatech) Chapter IPB University, yaitu Bang Assad (M.Kom. 2023), Kak Helma (M.Kom. 2021), Zalana (M.Kom. 2022), dan Jeni (M.Kom. 2023), atas bantuan dan arahan selama tahun pertama penulis menempuh studi di Bogor. Ucapan terima kasih juga disampaikan kepada sahabat seperjuangan Program Studi Ilmu Komputer IPB University angkatan 2024, sahabat Padepokan lintas angkatan tahun 2023–2024, serta adik i, abang, kakak, dan teman-teman yang tidak dapat disebutkan satu per satu atas dukungan selama penyelesaian studi. Semoga seluruh pihak tersebut senantiasa diberikan kesehatan, kemudahan dalam setiap urusan, dan kesempatan untuk terus menjaga silaturahmi.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan turut mendukung pengembangan ilmu pengetahuan.

Bogor, Agustus 2026

Said Thaufik Rizaldi



DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xiii
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	4
2.1 Ular <i>C. rhodostoma</i>	4
2.2 Tanaman <i>C. cinereum</i>	5
2.3 <i>Drug-Target Interaction</i> (DTI)	6
2.4 <i>Graph Neural Network</i> (GNN)	9
2.5 <i>Graph Attention Network</i> (GAT)	10
2.6 <i>Graph Isomorphism Network</i> (GIN) dan GINE	11
2.7 <i>Graph Convolutional Network</i> (GCN)	13
2.8 <i>Convolutional Neural Network</i> (CNN)	14
2.9 <i>Feature-wise Linear Modulation</i> (FiLM)	15
III METODE	17
3.1 Akuisisi dan Praproses Data	17
3.2 Rekayasa Fitur Senyawa	18
3.3 Rekayasa Fitur Protein	22
3.4 Pembangkitan <i>Putative Negative</i>	24
3.5 Pemodelan <i>Machine Learning</i>	26
3.6 Evaluasi Model <i>Machine Learning</i>	27
3.7 Prediksi DTI Senyawa Sawi Langit	29
3.8 Analisis <i>Molecular Docking</i>	29
3.9 Analisis <i>Molecular Dynamics</i>	30
3.10 Peralatan Pelatihan	31
IV HASIL DAN PEMBAHASAN	32
4.1 Hasil Akuisisi dan Praproses Data Pelatihan	32
4.2 Hasil Akuisisi dan Praproses Data Prediksi	33
4.3 Hasil Rekayasa Fitur Data Pelatihan	39
4.4 Pembangkitan <i>Putative Negative</i>	47
4.5 Evaluasi Pelatihan Model <i>Machine Learning</i>	48
4.6 Prediksi DTI Senyawa Sawi Langit	59
4.7 Analisis <i>Molecular Docking</i>	62
4.8 Analisis <i>Molecular Dynamics</i>	75



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

V	SIMPULAN DAN SARAN	80
5.1	Simpulan	80
5.2	Saran	80
	DAFTAR PUSTAKA	81
	LAMPIRAN	89
	RIWAYAT HIDUP	104



DAFTAR TABEL

1	Efek envenomasi ular <i>C. rhodostoma</i>	4
2	Aktivitas farmakologis senyawa <i>C. cinereum</i>	6
3	Data penelitian	17
4	Komponen fitur <i>node</i> pada representasi graf senyawa	20
5	Komponen fitur <i>edge</i> pada representasi graf senyawa	21
6	Kombinasi pemodelan fitur rekayasa	26
7	Rancangan <i>ablation study</i> modulasi FiLM	28
8	Spesifikasi perangkat lokal	31
9	Spesifikasi server HPC	31
10	Data awal pelatihan dari <i>Snakebite Envenoming Medicines Database</i>	32
11	Data pelatihan terproses pasangan interaksi positif	33
12	Data awal protein <i>venom C. rhodostoma</i> populasi Borneo	33
13	Data terproses protein <i>venom C. rhodostoma</i> populasi Borneo	34
14	Data awal protein <i>venom C. rhodostoma</i> populasi Jawa	34
15	Data terproses protein <i>venom C. rhodostoma</i> populasi Jawa	35
16	Data awal protein <i>venom C. rhodostoma</i> populasi Thailand	35
17	Data terproses protein <i>venom C. rhodostoma</i> populasi Thailand	36
18	Data awal protein <i>venom C. rhodostoma</i> populasi Malaysia	36
19	Data terproses protein <i>venom C. rhodostoma</i> populasi Malaysia	37
20	Data awal senyawa metabolit <i>C. cinereum</i> hasil LC-HRMS	37
21	Data terproses senyawa metabolit <i>C. cinereum</i>	37
22	Data pasangan senyawa-protein populasi Borneo	38
23	Data pasangan senyawa-protein populasi Jawa	38
24	Data pasangan senyawa-protein populasi Malaysia	39
25	Data pasangan senyawa-protein populasi Thailand	39
26	Hasil Morgan <i>fingerprint</i> pada data pelatihan positif	40
27	Hasil Morgan <i>fingerprint</i> dengan <i>dense encoder</i> pada data pelatihan positif	40
28	Hasil representasi GNN Tsubaki pada data pelatihan positif	40
29	Hasil representasi GIN pada data pelatihan positif	41
30	Hasil graf GAT dengan modulasi FiLM pada data pelatihan positif	41
31	Hasil graf GINE dengan modulasi FiLM pada data pelatihan positif	42
32	Hasil graf GCN dengan modulasi FiLM pada data pelatihan positif	42
33	Hasil AAC pada data pelatihan positif	43
34	Hasil representasi CNN protein 64 dimensi pada data pelatihan positif	43
35	Hasil representasi CNN protein 128 dimensi pada data pelatihan positif	43
36	Hasil <i>embedding</i> protein Tsubaki pada data pelatihan positif	44
37	Hasil CNN dengan modulasi FiLM pada data pelatihan positif	44
38	Kombinasi fitur pada model 1–3	45
39	Kombinasi fitur pada model 4	45
40	Kombinasi fitur pada model 5	45
41	Kombinasi fitur pada model 6	46
42	Kombinasi fitur pada model 7	46

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



43	Kombinasi fitur pada model 8	46
44	Kombinasi fitur pada model 9	47
45	Data pelatihan seimbang pada model 1	47
46	Ringkasan data pelatihan seimbang setelah pembangkitan <i>putative negative</i>	47
47	Hasil evaluasi model menggunakan <i>nested cross validation</i>	48
48	Hasil evaluasi model pada skenario <i>unseen drug</i>	51
49	Hasil evaluasi model pada skenario <i>unseen target</i>	53
50	Hasil evaluasi model pada skenario <i>unseen drug-target</i>	55
51	<i>Confusion matrix</i> model pada skenario <i>unseen drug-target</i>	57
52	Hasil <i>ablation study</i> pada skenario <i>unseen drug-target</i>	58
53	Data prediksi DTI prioritas populasi Borneo	60
54	Data prediksi DTI prioritas populasi Jawa	60
55	Data prediksi DTI prioritas populasi Malaysia	61
56	Data prediksi DTI prioritas populasi Thailand	61
57	Hasil <i>molecular docking</i> populasi Borneo	62
58	Hasil <i>molecular docking</i> populasi Jawa	63
59	Hasil <i>molecular docking</i> populasi Malaysia	64
60	Hasil <i>molecular docking</i> populasi Thailand	64
61	Senyawa referensi pada analisis <i>molecular docking</i>	66
62	Perbandingan kandidat dan senyawa referensi pada populasi Borneo	66
63	Perbandingan kandidat dan senyawa referensi pada populasi Jawa	67
64	Perbandingan kandidat dan senyawa referensi pada populasi Malaysia	68
65	Perbandingan kandidat dan senyawa referensi pada populasi Thailand	68
66	Pasangan protein-senyawa terpilih berdasarkan hasil <i>molecular docking</i>	69
67	Komponen energi pengikatan dua kompleks berdasarkan MM/PBSA	79

DAFTAR GAMBAR

1	Ular <i>C. rhodostoma</i>	4
2	Tanaman <i>C. cinereum</i>	5
3	Proses DTI	7
4	Representasi fitur pada DTI	7
5	Pembentukan Morgan <i>fingerprint</i>	8
6	Proses <i>message passing</i> dan pembaruan fitur <i>node</i> pada GNN	10
7	Mekanisme <i>attention</i> dan pembaruan fitur <i>node</i> pada GAT	11
8	Proses agregasi dan pembaruan fitur <i>node</i> pada GINE	12
9	Proses normalisasi, agregasi, dan pembaruan fitur pada GCN	13
10	<i>Multi-kernel</i> 1D CNN pada representasi urutan protein	15
11	Lapisan FiLM pada jaringan CNN	16
12	Alur penelitian	17
13	Alur praproses data	18
14	Representasi GNN pada data senyawa	19
15	Representasi GIN pada data senyawa	19

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

16	Arsitektur pembentukan <i>embedding</i> graf senyawa dengan modulasi FiLM	22
17	<i>Multi-kernel</i> 1D CNN pada data protein	23
18	Arsitektur pembentukan <i>embedding</i> protein dengan CNN dan modulasi FiLM	24
19	Pembangkitan <i>putative negative</i> menggunakan GANATCpair	25
20	<i>Nested cross validation</i> pada pemodelan <i>machine learning</i>	26
21	Skenario evaluasi <i>unseen</i> pada prediksi DTI	28
22	Alur preparasi struktur sebelum <i>molecular docking</i> pada (a) protein sebagai reseptor dan (b) senyawa sebagai ligan	30
23	Alur pelaksanaan <i>molecular docking</i> menggunakan <i>AutoDock Vina</i>	30
24	Kurva ROC berdasarkan <i>pooling</i> seluruh prediksi <i>out-of-fold</i> pada <i>nested cross validation</i>	50
25	Kurva <i>precision-recall</i> berdasarkan <i>pooling</i> seluruh prediksi <i>out-of-fold</i> pada <i>nested cross validation</i>	50
26	Kurva ROC pada skenario <i>unseen drug</i>	52
27	Kurva <i>precision-recall</i> pada skenario <i>unseen drug</i>	52
28	Kurva ROC pada skenario <i>unseen target</i>	54
29	Kurva <i>precision-recall</i> pada skenario <i>unseen target</i>	54
30	Kurva ROC pada skenario <i>unseen drug-target</i>	56
31	Kurva <i>precision-recall</i> pada skenario <i>unseen drug-target</i>	56
32	Kinerja <i>ablation study</i> pada skenario <i>unseen drug-target</i>	59
33	Jumlah pasangan senyawa <i>C. cinereum</i> dengan skor <i>docking</i> ≤ -6 kcal/mol pada protein terpilih	65
34	Visualisasi interaksi A0A9F2NW96 dengan senyawa CID 251559 pada populasi Borneo	71
35	Visualisasi interaksi A0A6H0QVN7 dengan senyawa CID 251559 pada populasi Borneo	71
36	Visualisasi interaksi Q2LD49 dengan senyawa CID 251559 pada populasi Thailand	72
37	Visualisasi interaksi Q9W6M5 dengan senyawa CID 251559 pada populasi Thailand	73
38	Visualisasi interaksi J3SEZ3 dengan senyawa CID 251559 pada populasi Malaysia	73
39	Visualisasi interaksi T2HQ57 dengan senyawa CID 22600366 pada populasi Borneo dan Jawa	74
40	Visualisasi interaksi Q90ZI3 dengan senyawa CID 22600366 pada populasi Malaysia dan Thailand	75
41	RMSD protein dan ligan selama simulasi 100 ns. Panel (a) kompleks A0A9F2NW96–251559 dan panel (b) kompleks T2HQ57–22600366	76
42	RMSF residu protein selama simulasi 100 ns. Panel (a) kompleks A0A9F2NW96–251559 dan panel (b) kompleks T2HQ57–22600366	76
43	Rg kompleks selama simulasi 100 ns. Panel (a) kompleks A0A9F2NW96–251559 dan panel (b) kompleks T2HQ57–22600366	77



44	Jumlah ikatan hidrogen protein-ligan selama simulasi 100 ns. Panel (a) kompleks A0A9F2NW96–251559 dan panel (b) kompleks T2HQ57–22600366	77
45	Tumpang tindih struktur kompleks sebelum dan setelah simulasi 100 ns. Panel (a) kompleks A0A9F2NW96–251559 dan panel (b) kompleks T2HQ57–22600366	78

DAFTAR LAMPIRAN

1	Ruang pencarian dan nilai optimal <i>hyperparameter</i> model	90
2	Hasil <i>molecular docking</i> populasi Borneo	92
3	Hasil <i>molecular docking</i> populasi Jawa	95
4	Hasil <i>molecular docking</i> populasi Malaysia	98
5	Hasil <i>molecular docking</i> populasi Thailand	101