



PENGEMBANGAN ALGORITMA DETEKSI KOMUNITAS DAN *CENTRALITY* DALAM MENENTUKAN PROTEIN TARGET PADA KANKER

HERU CAHYA RUSTAMAJI



PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024



@Hak cipta milik IPB University

IPB University

- 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.

PERNYATAAN MENGENAI DISERTASI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa disertasi dengan judul “Pengembangan Algoritma Deteksi Komunitas dan *Centrality* dalam Menentukan Protein Target pada Kanker” 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 disertasi ini. Seluruh dataset yang digunakan dalam penelitian ini berasal dari pangkalan data yang bersifat publik serta dapat diakses secara bebas. Semua sumber pangkalan data telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir disertasi ini.

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

Bogor, Juni 2024

Heru Cahya Rustamaji
G6601202015

- 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 menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

RINGKASAN

HERU CAHYA RUSTAMAJI. Pengembangan Algoritma Deteksi Komunitas dan *Centrality* dalam Menentukan Protein Target pada Kanker. Dibimbing oleh WISNU ANANTA KUSUMA, SRI NURDIATI dan IRMANIDA BATUBARA

Komunitas dan *centrality* sangat penting untuk memahami sistem yang kompleks seperti jaringan protein dalam analisis jaringan. Masalah deteksi komunitas adalah bagaimana membagi jaringan yang direpresentasikan dalam bentuk *graph* menjadi beberapa kelompok simpul yang mempunyai kepadatan yang tinggi, memiliki karakteristik atau sifat yang berbeda. Kualitas komunitas yang terbentuk oleh algoritma deteksi komunitas dapat ditentukan dengan nilai *modularity*. Salah satu algoritma terkenal dan banyak menjadi rujukan dalam menyelesaikan masalah deteksi komunitas adalah algoritma *greedy modularity*. Algoritma ini melakukan eksplorasi untuk mencapai nilai *modularity* tertinggi pada setiap iterasi untuk mendapat komunitas terbaik.

Kelemahan algoritma *greedy modularity* pada deteksi komunitas adalah dapat terjebak pada solusi lokal optimal. Untuk itu penelitian ini mengembangkan strategi *disassembly greedy modularity* untuk meningkatkan *modularity* tersebut dengan eksplorasi yaitu lepas simpul dan pecah komunitas. Terdapat empat strategi melepas simpul yaitu melepas simpul secara *random*/acak dengan distribusi *uniform*, melepas simpul lemah, melepas simpul yang mempunyai nilai *embeddedness* rendah, serta melepas simpul yang tidak membentuk *triad*. Terdapat lima strategi membongkar komunitas yaitu membongkar komunitas secara acak dengan distribusi *uniform*, membongkar komunitas yang lemah, membongkar komunitas dengan *internal edge density* yang rendah, membongkar komunitas dengan *triad participation ratio* yang rendah serta membongkar komunitas dengan *conductance* yang tinggi.

Data yang digunakan dalam penelitian adalah data *real-world* dan data sintetis. Data *real-world* yang digunakan meliputi data karate Zachary, *Dolphins*, *Les Miserables*, *Polbooks*, *Adjnoun* serta *Football*. Sedangkan data sintetis yang digunakan adalah Lancichinetti–Fortunato–Radicchi (LFR). Hasil yang diperoleh menggunakan *disassembly greedy modularity* berupa peningkatan *modularity* dibandingkan dengan *greedy modularity* untuk semua *dataset* tersebut. Bahkan dibandingkan dengan 25 algoritma lainnya, algoritma *disassembly greedy modularity* menduduki mayoritas menduduki urutan pertama berdasarkan *modularity*.

Di lain pihak, *centrality* adalah konsep yang digunakan untuk mengukur pentingnya suatu simpul dalam jaringan. Terdapat beberapa jenis *centrality* yang umum digunakan dalam analisis jaringan, di antaranya adalah *degree centrality*, *closeness centrality*, dan *betweenness centrality*. Namun demikian, *degree*, *closeness* dan *betweenness centrality* tersebut dihitung berdasarkan struktur jaringan, belum memperhatikan komunitas di dalamnya. Sehingga, *centrality* tersebut perlu dikembangkan agar dapat mempertimbangkan komunitas dalam jaringan, yang dalam penelitian ini disebut dengan *community consideration degree centrality*, *community consideration closeness centrality*, dan *community consideration betweenness centrality*. Variabel α dengan rentang $[0,1]$ digunakan sebagai pertimbangan komunitas. Semakin rendah nilai α akan lebih mempertimbangkan komunitas, sementara semakin tinggi nilai α akan lebih

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.

mempertimbangkan jaringan secara keseluruhan. Hasil yang diperoleh menggunakan *dataset* karate Zachary dan *dolphins* dapat melibatkan simpul pada komunitas yang lebih beragam.

Secara khusus, algoritma *disassembly greedy modularity* dan *community consideration centrality* dalam penelitian ini digunakan untuk mengelompokkan protein target pada kanker dan mengungkap protein yang menjadi *hallmark of cancer*. Data genetik pada kanker paru-paru adenokarsinoma diambil dari empat pangkalan data: TCGA, IntOGen, Bioportal, dan Cosmic, yang selanjutnya data tersebut dipergunakan untuk memperoleh jaringan interaksi protein melalui pangkalan data *String*. Algoritma *disassembly greedy modularity* digunakan untuk mendeteksi stuktur komunitas pada jaringan interaksi protein tersebut. Hasil dari algoritma ini terbentuk 11 komunitas dengan ukuran berbeda dengan dua komunitas besar, empat komunitas menengah, dan lima komunitas kecil yang sebagian besar mempunyai asosiasi fungsional melalui *enrichment analysis*, lebih baik daripada hasil yang diperoleh dengan algoritma *greedy modularity*.

Hasil berikutnya adalah, bahwa *community consideration degree*, *closeness* dan *betweenness centrality* dengan $\alpha = 0,1$ dan $\alpha = 0,2$ dapat mengungkap protein *hallmark of cancer* yang lebih banyak daripada *centrality* yang tidak mempertimbangkan komunitas. *Centrality* ini juga dibandingkan dengan berbagai *centrality* yang lain, seperti *eigen vector centrality*, *Katz centrality*, *information centrality*, *load centrality*, *subgraph centrality*, *harmonic centrality*, *second order centrality*, *distinctive centrality* maupun *viral rank centrality*. Hasilnya adalah *community consideration centrality* mengungguli kesemua *centrality* pembandingan dalam mengungkapkannya banyaknya protein *hallmark of cancer*.

Lebih lanjut, penentuan nilai α juga dipengaruhi struktur komunitas yang diperoleh. Apabila terdapat struktur komunitas terdapat pada jaringan yang dicirikan dengan *modularity* yang tinggi, maka α cenderung bernilai kecil. Sebaliknya apabila tidak ada struktur komunitas, maka nilai α cenderung mendekati 1, dalam arti bahwa *centrality* tidak dipengaruhi oleh komunitas.

Mengingat potensi yang ada, metodologi ini mempunyai implikasi umum dan memberikan peluang untuk diterapkan di berbagai bidang yang lebih luas. Ukuran *centrality* seringkali mengabaikan struktur yang rumit dalam suatu jaringan. Dengan melibatkan komunitas, pendekatan *community consideration centrality* memungkinkan analisis jaringan yang lebih rinci dan akurat, sehingga meningkatkan pemahaman tentang pentingnya keberadaan simpul secara individu maupun pengaruhnya secara kolektif pada sebuah komunitas. *Community consideration centrality* mempunyai potensi yang tidak hanya terbatas pada jaringan biologis tetapi juga berlaku pada bidang lain seperti analisis jaringan sosial, organisasi, dan jaringan komunikasi maupun jaringan lainnya yang ada dalam sebuah sistem yang kompleks.

Kata kunci: *community consideration centrality*, deteksi komunitas, *disassembly greedy modularity*, kanker, protein

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

SUMMARY

HERU CAHYA RUSTAMAJI. *Development of a Community Detection and Centrality Algorithm and Its Use in Determining Target Proteins in Cancer*. Supervised by WISNU ANANTA KUSUMA, SRI NURDIATI and IRMANIDA BATUBARA

Community and centrality are critical to understanding complex systems, such as protein networks, in network analysis. The community detection problem involves dividing a network in graph form into several groups of nodes with a high density and different characteristics or properties. Modularity can determine the quality of the community formed by the community detection algorithm. The greedy modularity algorithm is a well-known algorithm with many references for solving community detection problems. This algorithm performs exploration to achieve the highest modularity value at each iteration to obtain the best community.

The weakness of the greedy modularity algorithm in community detection is that it can become stuck in a locally optimal solution. Therefore, this study develops a disassembly greedy modularity strategy to increase modularity through exploration: disassembly nodes and disassembly communities. There are four strategies for disassembling nodes: releasing nodes randomly with a uniform distribution, releasing weak nodes, releasing nodes with low embeddedness values, and releasing nodes that do not form a triad. There are five strategies for disassembling communities: random with uniform distribution, weak communities, communities with low internal density, communities with low triad participation ratio, and communities with high conductance.

The data used in the research were real-world and synthetic data. The real-world data used includes data on Zachary's karate, Dolphins, Les Miserables, Polbooks, Adjnoun and Football. The synthetic data used was LFR. The results obtained using disassembly greedy modularity show an increase in modularity compared to greedy modularity for all of these datasets. Compared to the 25 other algorithms, the disassembly greedy modularity algorithm occupies the first place in the majority.

On the other hand, centrality is a concept used to measure the importance of a node in a network. Several types of centralities are commonly used in network analysis, including degree, closeness, and betweenness centralities. However, degree, closeness, and betweenness centrality are calculated based on the network structure without paying attention to the community within it. Therefore, centrality needs to be developed to consider the community in the network, which in this research is called community consideration degree centrality, community consideration closeness centrality, and community consideration betweenness centrality. Variable α in the range $[0,1]$ is used as a community consideration. The lower the α value, the more the community will be considered, whereas the higher the α value, the more the network as a whole will be considered. The results obtained using the Zachary Karate and Dolphin datasets can involve nodes in more diverse communities.

Specifically, the disassembly greedy modularity and community consideration centrality algorithms in this research were used to group target proteins in cancer and obtain proteins that are hallmarks of cancer. Data on genetic

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

mutations in lung adenocarcinoma were obtained from four databases: TCGA, IntOGen, Bioportal, and Cosmic, which were then used to obtain protein interaction networks via the String database. Greedy modularity disassembly algorithm to detect community structure in protein interaction network. The results of this algorithm formed 11 communities of different sizes, with two large communities, four medium communities, and five small communities, most of which had functional associations through enrichment analysis.

The following result is that community consideration degree, closeness, and betweenness centrality with $\alpha = 0.1$ and $\alpha = 0.2$ can reveal more cancer protein hallmarks than centrality that does not consider the community. Next, it was compared with other centralities, such as eigenvector centrality, Katz centrality, information centrality, load centrality, subgraph centrality, harmonic centrality, second-order centrality, distinctive centrality, and viral rank centrality. The result is that community consideration centrality outperforms all in revealing the protein hallmarks of cancer.

Furthermore, determining the α value was also influenced by the community structure obtained. If a community structure in a network is characterized by high modularity, α tends to have a small value. However, if there is no community structure, the value of α tends to be close to 1 because the community does not influence centrality.

Given the existing potential, this methodology has general implications and the potential to be applied in a wider variety of fields. Traditional centrality measures often ignore the complex substructures within a network. By involving the community, the community consideration centrality approach allows for more detailed and accurate network analysis, thereby increasing the understanding of the importance of individual nodes and their collective influence on a community. Community consideration centrality has the potential not only to be limited to biological networks but also to other fields, such as social network analysis, organizational studies, communication networks, and other networks that exist in a complex system.

Keywords: cancer, community consideration centrality, community detection, disassembly greedy modularity, protein

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 University

IPB University



- 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 2024
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.



PENGEMBANGAN ALGORITMA DETEKSI KOMUNITAS DAN *CENTRALITY* DALAM MENENTUKAN PROTEIN TARGET PADA KANKER

HERU CAHYA RUSTAMAJI

Disertasi
sebagai salah satu syarat untuk memperoleh gelar
Doktor pada
Program Studi Ilmu Komputer

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

IPB University

Penguji Luar Komisi Pembimbing pada Ujian Tertutup Disertasi:

- 1 Dr. Mushthofa, S.Kom., M.Sc.
- 2 dr. Irandi Putra Pratomo, Ph.D, Sp.P(K), FAPSR

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.



Judul Disertasi: Pengembangan Algoritma Deteksi Komunitas dan *Centrality*
dalam Menentukan Protein Target pada Kanker

Nama : Heru Cahya Rustamaji
NIM : G6601202015

Disetujui oleh

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

Pembimbing 2:
Prof. Dr. Ir. Sri Nurdianti, M.Sc.

Pembimbing 3:
Prof. Dr. Irmanida Batubara, M.Si.



Diketahui oleh

Plh. Ketua Program Studi:
Dr. Hendra Rahmawan S.Kom, M.T
NIP 198205012009121004

Dekan FMIPA IPB:
Dr. Berry Juliandi, S.Si., M.Si.
NIP 197807232007011001



Tanggal Ujian:
4 Juni 2024

Tanggal Lulus:

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.

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah SWT atas segala nikmat dan karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Mei 2022 sampai bulan Juni 2024 ini berjudul “Pengembangan Algoritma Deteksi Komunitas dan *Centrality* dalam Menentukan Protein Target pada Kanker”.

Terima kasih penulis ucapkan kepada para promotor, Dr. Wisnu Ananta Kusuma, S.T., M.T, Prof. Dr. Ir. Sri Nurdiati, M.Sc, dan Prof. Dr. Irmanida Batubara, M.Si yang telah mencurahkan ilmu, membimbing, memotivasi, dan banyak memberi saran demi terselesaikannya disertasi ini. Demikian pula ucapan terima kasih juga disampaikan kepada:

1. Rektor IPB, Pimpinan Sekolah Pascasarjana IPB, Dekan FMIPA IPB, Ketua Program Studi Ilmu Komputer, dosen dan tenaga kependidikan IPB, para penguji luar komisi pembimbing, para moderator seminar yang telah memberikan banyak ilmu pengetahuan, pelajaran, pengalaman dan layanan administrasi dengan sangat baik.
2. Rektor UPN Veteran Yogyakarta, Dekan Fakultas Teknologi Industri, serta teman teman dosen Jurusan Informatika UPN “Veteran” Yogyakarta.
3. Istri Bekti Astuti dan anak-anak tersayang, Fikri Astama Putra dan Aisha Rahmania Putri yang telah memberikan dukungan dan do’a yang tulus.
4. Teman-teman Program Studi S-3 Ilmu Komputer IPB, terutama angkatan 2020 atas bantuan dan dukungannya.
5. Pihak-pihak lain yang tidak bisa saya sebutkan satu persatu, telah banyak membantu dalam penelitian ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang memerlukan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juni 2024

Heru Cahya Rustamaji

DAFTAR ISI

DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	6
1.4 Manfaat	6
1.5 Ruang Lingkup	6
1.6 Kebaruan	6
1.7 Kerangka Penulisan	6
II TINJAUAN PUSTAKA	8
2.1 Deteksi Komunitas	8
2.2 Algoritma <i>Greedy Modularity</i>	11
2.3 <i>Centrality</i>	12
2.4 Kanker Paru-paru	15
2.5 <i>Hallmark of Cancer</i>	18
III METODE PENELITIAN	21
3.1 <i>Dataset</i>	21
3.2 Pengembangan Algoritma Deteksi Komunitas <i>Disassembly Greedy Modularity</i>	24
3.3 Pengembangan <i>Centrality</i> yang Mempertimbangkan Komunitas	39
3.4 Penentuan kelompok fungsional dan protein target yang menjadi ciri kanker	44
IV PENGEMBANGAN ALGORITMA DISASSEMBLY GREEDY MODULARITY	46
4.1 Evaluasi Hasil Deteksi Komunitas	50
4.2 Hasil pada <i>Dataset</i> Sintetis	51
4.3 Kompleksitas komputasi	52
V PENGEMBANGAN ALGORITMA <i>CENTRALITY</i> YANG MEMPERTIMBANGKAN KOMUNITAS	53
5.1 Jaringan contoh	53
5.2 Karate Zachary	55
5.3 <i>Dolphins</i>	56
VI IMPLEMENTASI PADA PROTEIN KANKER	58
6.1 <i>Dataset</i>	58
6.2 Deteksi komunitas	59
6.3 Enrichment Analysis	60
6.4 <i>Community Consideration Degree Centrality</i>	62
6.5 <i>Community-consideration Closeness Centrality</i>	64
6.6 <i>Community-consideration Betweenness Centrality</i>	66
6.7 Perbandingan antar <i>centrality</i>	68

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.



6.8	Analisis Hasil <i>Hallmark of Cancer</i> Melalui <i>Community Consideration Centrality</i>	69
6.9	Optimasi Parameter	71
6.10	Percobaan menggunakan <i>dataset</i> NCBI	71
6.11	Penggunaan <i>Community Consideration Centrality</i> pada Domain Lain	73
PENUTUP		75
6.1	Simpulan	75
6.2	Saran	75
DAFTAR PUSTAKA		76
RIWAYAT HIDUP		88

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.

DAFTAR TABEL

1	Penelitian terkait deteksi komunitas	9
2	Penelitian terkait <i>centrality</i>	14
3	Penelitian terkait <i>hallmark of cancer</i> secara komputasi	19
4	Perbandingan strategi antara algoritma <i>greedy modularity</i> dan <i>disassembly greedy modularity</i>	25
5	Kombinasi strategi pada algoritma DGM	47
6	Kombinasi strategi terbaik di beberapa kumpulan data referensi.	48
7	<i>Modularity</i> DGM dibanding dengan algoritma lain pada berbagai <i>dataset</i>	49
8	NMI antara DGM dengan berbagai metode lainnya, dengan berbagai <i>dataset</i>	51
9	Perbandingan <i>modularity</i> pada <i>dataset</i> sintetik dengan variasi μ dengan algoritma deteksi komunitas yang berbeda.	52
10	Hasil <i>community-consideration centrality</i> pada jejaring karate	56
11	Hasil <i>community-consideration centrality</i> pada jaringan lumba lumba	57
12	Komunitas yang terbentuk beserta jumlah simpulnya	60
13	Asosiasi fungsional kelompok protein berdasarkan deteksi komunitas dengan algoritma <i>Disassembly Greedy Modularity</i>	60
14	Asosiasi fungsional kelompok protein berdasarkan deteksi komunitas dengan algoritma <i>Greedy Modularity</i>	61
15	Perbandingan antara algoritma <i>greedy modularity</i> dan <i>disassembly greedy modularity</i>	62
16	Urutan protein berdasarkan <i>community-consideration degree centrality</i>	62
17	Perbandingan perhitungan TP53 dan SRC	63
18	Urutan protein berdasarkan <i>community-consideration closeness centrality</i>	64
19	Perbandingan perhitungan EP300, SRC dan TP53	65
20	Urutan protein berdasarkan <i>community-consideration betweenness centrality</i>	66
21	Perbandingan perhitungan TP53, EP300, CTNNB1, SMAD3 serta AKT1	67
22	<i>Hallmark of cancer</i> berdasarkan <i>community-consideration centrality</i>	70
23	Asosiasi fungsional kelompok protein berdasarkan deteksi komunitas menggunakan <i>dataset</i> NCBI	72

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.

DAFTAR GAMBAR

1	Kasus dan kematian penyakit kanker di Indonesia tahun 2020 per 100.000 penduduk (<i>International Agency for Research in Cancer 2019</i>)	5
2	Kerangka penulisan	7
3	Jaringan contoh terdiri dari dua komunitas, merah dan biru.	8
4	Jaringan berdasarkan kata kunci dan terkait penelitian mengenai deteksi komunitas	9
5	Pembentukan komunitas melalui <i>Greedy Modularity</i>	12
6	(a) Protein KRAS (Radzioch 2010) dan (b) interaksi protein (Rosenberger <i>et al.</i> 2016)	17
7	Ilustrasi <i>hallmark of cancer</i> (Čižmaríková <i>et al.</i> 2023)	20
8	Tahapan penelitian	21
9	<i>Dataset Zachary Karate</i> (Zachary 1977)	22
10	<i>Dataset Dolphins</i> (Lusseau <i>et al.</i> 2003)	22
11	<i>Dataset Les Miserables</i> (Knuth 1993)	22
12	Ilustasi <i>Disassembly Greedy Modularity</i>	24
13	Ilustrasi Algoritma Louvain (Liu <i>et al.</i> 2019)	34
14	Ilustrasi Algoritma Label Propagation (Mittal dan Bhatia 2021)	35
15	Ilustrasi Algoritma Genetika (a) Model jaringan; (b) representasi genotipe berbasis lokus; (c) struktur genotipe berbasis jaringan. Di sebelah kanan <i>crossover</i> dari dua individu, genotipe mereka, representasi berbasis jaringan, dan anak yang dihasilkan (Pizzuti 2008)	36
16	Ilustrasi <i>betweenness centrality</i> pada simpul <i>i</i> yang berada di komunitas biru. Sisi kiri garis vertikal adalah simpul yang berada dalam satu komunitas dengan <i>i</i> , dan sisi kanan adalah simpul yang berbeda komunitas dengan <i>i</i> .	43
17	Nilai <i>modularity</i> algoritma <i>DGM</i> pada <i>dataset Zachary</i> ($n=34$), $c=30$	46
18	Visualisasi deteksi komunitas pada (a) <i>ground-truth Zachary</i> (b) <i>Greedy Modularity</i> (c) <i>Disassembly Greedy Modularity</i>	48
19	Nilai NMI dan F1 divisualisasikan menggunakan <i>heatmap</i>	50
20	Visualisasi deteksi komunitas pada <i>dataset</i> sintetis	52
21	<i>Community consideration degree centrality</i> pada jaringan contoh. Simpul 5 yang menjadi urutan pertama pada <i>degree centrality</i> level jaringan, menjadi urutan ke 5 pada level komunitas.	53
22	<i>Community consideration closeness centrality</i> pada jaringan contoh. Terdapat perbedaan urutan dengan variasi α	54
23	<i>Community consideration betweenness centrality</i> pada jaringan contoh. Simpul 6 dan 5 mendominasi di urutan pertama dan kedua untuk $\alpha > 0,2$	54
24	Karate <i>ground-truth</i>	55
25	Jaringan lumba lumba	57

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.

- 26 Diagram Venn Mutasi Genetik Adenokarsinoma: Menampilkan kontribusi data mutasi yang mempunyai irisan dari pangkalan data TCGA, *IntOGen*, *Bioportal*, dan *Cosmic*, menunjuk data gabungan yang komprehensif 58
- 27 Jaringan Interaksi Protein-Protein menunjukkan interaksi kompleks antara protein yang terlihat dari data yang diperoleh dari mutasi genetik pada adenokarsinoma dari pangkalan data STRING. 59
- 28 Komunitas Protein Adenokarsinoma menunjukkan komunitas protein, berbagai ukuran dari subkelompok besar hingga kecil, dari jaringan kanker paru paru adenokarsinoma 59
- 29 Perubahan nilai *community consideration degree centrality* pada TP53 dan SRC 63
- 30 Kurva *Jackknife* terhadap *hallmark of cancer* yang terungkap terhadap protein target melalui *community-consideration degree centrality* 64
- 31 Perubahan nilai *community consideration closeness centrality* pada TP53, SRC dan EP300 65
- 32 Kurva *Jackknife* terhadap *hallmark of cancer* yang terungkap terhadap protein target melalui *community-consideration closeness centrality* 66
- 33 Perubahan nilai *community consideration betweenness centrality* pada beberapa protein pada berbagai variasi α 67
- 34 Kurva *Jackknife* terhadap *hallmark of cancer* yang terungkap terhadap protein target melalui *community-consideration betweenness centrality* 67
- 35 Kurva *Jackknife* terhadap *hallmark of cancer* yang terungkap melalui *community-consideration centrality* dan berbagai algoritma pembandingan *centrality* lainnya 69

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.