J. Ris. & Ap. Mat. Vol. 5 No. 1 (2021) pp. 60-67
Jurnal Riset dan Aplikasi Matematika
e-ISSN: 2581-0154
URL: journal.unesa.ac.id/index.php/jram

ANALISIS HASIL *BICLUSTER* ALGORITMA POLS PADA INTERAKSI PROTEIN MANUSIA DAN HIV-1

TESDIQ PRIGEL KALOKA[1*], TITIN SISWANTINING[2], ALHADI BUSTAMAM[3]

[1]Program Studi Teknik Informatika Politeknik Hasnur, [2,3]Departemen Matematika FMIPA Universitas Indonesia

*tesdiq@sci.ui.ac.id

ABSTRAK

AIDS merupakan penyakit yang disebabkan oleh virus HIV-1. Protein merupakan bagian penting dari organisme yang memiliki beragam fungsi. HIV-1 dapat menyerang tubuh manusia karena adanya interaksi protein. Fungsi dan sifat interaksi protein dapat diketahui dengan mengelompokkan protein-protein yang saling berinteraksi. *Bicluster* merupakan metode yang dapat digunakan untuk menyelesaikan permasalahan interaksi protein. Algoritma POLS merupakan algoritma *bicluster* yang menggunakan pendekatan teori graf. Data interaksi protein manusia dan HIV-1 dibagi menjadi dataset 1 dan dataset 2. Setiap dataset dianalisis tingkat kepadatan hasil *bicluster* Algoritma POLS. Penelitian berfokus pada pencarian tingkat kepadatan hasil *bicluster* Algoritma POLS terhadap interaksi protein manusia dan HIV-1. Dari hasil penelitian ini diperoleh presentase *bicluster* Algoritma POLS dengan tingkat kepadatan 1 pada dataset 1 sebesar 1,01% dan pada dataset 2 sebesar 4,42%, selain itu *bicluster* Algoritma POLS yang berukuran kecil (2×2 dan 3×3) lebih optimal karena memiliki tingkat kepadatan 1.

Kata Kunci: Interaksi Protein, Algoritma POLS, *Bicluster*

ABSTRACT

AIDS is a disease caused by the HIV-1 virus. Proteins are an essential part of organisms that have various functions. HIV-1 can attack the human body due to protein interactions. The role and nature of protein interactions can be determined by grouping interacting proteins. Bicluster is a method that can be used to solve protein interaction problems. The POLS algorithm is a bicluster algorithm that uses a graph theory approach. Human protein interaction data and HIV-1 were divided into datasets one and dataset 2. Each dataset was analyzed for the density level of the POLS Algorithm bicluster. We focus on finding the density level of POLS Algorithm on the interaction of human and HIV-1 protein. This study obtained percentage of dataset 1 with density level 1 is 1,01% and dataset 2 with density 2 is 4,42%, beside that, the small bicluster from POLS Alogrihtm is more optimal because it has density level of 1.

Keyword: *Protein Interaction, POLS Algorithm, Bicluster*

1 Pendahuluan

Acquired Immunodeficiency Syndrome (AIDS) merupakan suatu penyakit yang disebabkan oleh virus *Human Immunodeficiency Virus* (HIV). HIV melemahkan kekebalan

2010 *Mathematics Subject Classification*: 92D20
Tanggal Masuk: 04-02-21; direvisi: 12-04-21; diterima: 29-04-21

tubuh manusia secara perlahan, sehingga diagnosa penyakit AIDS terlambat (kekebalan tubuh sudah sangat lemah) dan banyak penderita AIDS yang meninggal [1]. Menurut Frankenberg [2] virus HIV terdiri dari 2 jenis, yaitu HIV-1 dan HIV-2.

HIV-1 merupakan virus yang terdiri dari susunan RNA. HIV-1 dapat menyerang kekebalan tubuh manusia karena terjadi proses interaksi protein. Sifat dan fungsi suatu protein dapat diketahui dengan mengelompokkan interaksi protein [3]. Sifat dan fungsi protein HIV-1 yang telah diketahui akan memudahkan pihak kesehatan untuk dapat menemukan cara penyembuhan penderita penyakit AIDS.

Permasalahan yang ditemui pada bidang kesehatan dan biologi adalah proses penelitian yang memakan waktu dan biaya yang cukup besar. Permasalahan tersebut dapat diatasi dengan bantuan Bioinformatika. Bioinformatika merupakan ilmu pengetahuan yang menggabungkan bioligis, kedokteran, matematika, statistika, dan teknik komputasi. Secara sederhana, bioinformatika membantu menyelesaikan permasalahan biologi, kedokteran, dan *life science* dengan menggunakan teori-teori matematika dan statistika yang selanjutnya dieksekusi menggunakan bantuan komputer [4]. Salah satu cabang Bioinformatika yang dapat menyelesaikan permasalahan pengelompokkan protein adalah *bicluster*. *Bicluster* merupakan suatu metode yang digunakan untuk mengelompokkan data-data ekspresi gen dikarenakan metode clustering biasa tidak dapat mencari pola kesamaan yang cocok [5]. Pada umumnya *bicluster* disamakan dengan *clustering* dua kali.

Bicluster merupakan suatu metode analisis yang digunakan untuk data-data biologis [6]. Algoritma *bicluster* telah banyak digunakan, salah satu algoritma *bicluster* pertama yang terkenal adalah Algoritma CnC (Cheng and Chruch), diambil dari nama penemu algoritma tersebut [7]. Beberapa algoritma *bicluster* yang sering digunakan antara lain: *Qualitative Bicluster*ing (QUBIC), FLOC, dan *Binary Inclusion-Maximal* (BiMax). QUBIC dan BiMax merupakan algoritma *bicluster* yang menggunakan pendekatan teori graf [5]. FLOC merupakan algorima yang menggunakan konsep probabilitas [8].

Algoritma *Bicluster* terbaru dikemukakan oleh Wang [9] yang disebut dengan Algoritma POLS. Algoritma POLS menggunakan pendekatan teori graf dan teori *biclique* seimbang. Berdasarkan penelitian Bustamam [3], hasil *bicluster* Algoritma POLS menunjukkan adanya protein yang tidak berinteraksi pada satu *bicluster*. Kemampuan suatu Algortima *Bicluster* dalam mengelompokkan data diukur dengan tingkat kepadatannya, semakin tinggi tingkat kepadatannya, maka semakin baik algoritma tersebut. Sehingga penelitian ini berfokus pada tingkat kepadatan interaksi setiap *bicluster* Algoritma POLS terhadap interaksi protein manusia dan HIV-1. Harapan kedepannya penggunaan *bicluster* dapat membantu bidang kesehatan untuk melakukan penelitian lebih mendalam dari hasil *bicluster*.

2 Tinjauan Pustaka

2.1. *Bicluster*

Biclsuter pertama kali dikemukakan oleh Cheng & Church [8] dan lebih dikenal dengan *clustering* secara baris dan kolom. Pada awalnya, *biclsuter* digunakan untuk mencari suatu subset dari dari sebuah gen yang saling berkaitan [5]. Konsep pencarian *bicluster* adalah dengan mengelompokkan nilai-nilai yang memiliki keterhubungan menjadi suatu sebuah subset dari data yang ada.

Diberikan suatu matriks $A(D, C)$ dengan D adalah suatu himpunan data $D = \{I_1, I_2, I_3, \dots, I_D,\}$ dan C suatu himpunan fitur $C = \{J_1, J_2, J_3, \dots, J_C\}$. Sebuah *bicluster* didefinisikan sebagai:

Definisi 1

Sebuah *bicluster* merupakan submatriks $M(I,J) = [m_{i,j}]$ dengan $i \in I$, $j \in J$ adalah matriks $A(D,C)$ dengan $I \in D$ dan $J \in C$ [5]. Secara sederhana, *bicluster* merupakan pencarian submatriks dari data yang digunakan (Persamaan 1).

$$A(D,C) = \begin{bmatrix} & J_1 & J_2 & \cdots & J_c \\ I_1 & m_{11} & m_{12} & \cdots & m_{1C} \\ I_2 & m_{21} & m_{22} & \cdots & m_{2C} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_D & m_{D1} & m_{D2} & \cdots & m_{DC} \end{bmatrix} \quad (1)$$

2.2. Algoritma POLS

Algoritma POLS merupakan suatu algoritma pencarian *bicluster* yeng menggunakan pendekatan teori graf. Hasil *bicluster* Algoritma POLS merupakan sebuah *biclique* seimbang. *Biclique* seimbang merupakan suatu graf bipartit lengkap dengan jumlah partisi busur yang sama.

Bicluster hasil Algoritma POLS dibentuk dengan menggunakan beberapa tahap. Tahap-tahap tersebut dinamakan kandidat, *addset*, *dropset*, *pscore*, dan *addrule*. Kandidat merupakan tempat untuk menyimpan solusi hingga proses iterasi selesai. *Addset* merupakan himpunan suatu semesta yang akan ditambahkan ke kandidat. *Dropset* merupakan subset dati kandidat. *Pscore* merupakan nilai dari setiap pasang *addset*. *Addrule* merupakan proses pemilihan pasangan busur terbaik dari *addset* [9]. Penjelasan Algoritma POLS dapat lebih dipahami pada Tabel 1.

Tabel 1: Algoritma POLS

Masukan	:	Graf Bipartit
1	:	Inisialisasi *addset*, *dropset*, dan $S^* = S$
Ulangi		
2	:	Pilih sepasang simpul (u,v) dari *addset* menggunakan *addrule*
3	:	Masukkan pasangan simpul (u,v) ke $S = S \cup (u,v)$
Hingga	:	*Addset*=∅
4	:	Jika $\|S\| > \|S^*\|$, maka $S^* = S$
5	:	Hilangkan pasangan simpul (u,v) pada *dropset*
Keluaran	:	*Bicluster*

Jika kandidat solusi dinotasikan sebagai $S = (U^S, V^S, E^S)$. Misal Diberikan suatu graf bipartit $G = G(U,V,E)$ dengan simpul U dan V adalah $U = \{u_1, u_2, u_3, \dots, u_n\}$ dan $V = \{v_1, v_2, v_3, \dots, v_n\}$ serta busur E adalah $E = \{e_1, e_2, e_3, \dots, e_n\}$ maka menurut Wang [9] definisi *addset*, *dropset*, dan *pscore* adalah

$$Addset = \{(u,v) \in |u \notin U^S, u \in N(v'); \forall v' \in V^S, v \notin V^S, v \in N(u); \forall u' \in U^S\} \quad (2)$$

$$Dropset = \{(u,v) \in E | u \in U^S, v \in V^S\} \quad (3)$$

$$pscore(u,v) = score_{lb}(u,v) + \left\lfloor \frac{score_{ub}}{2} \right\rfloor \quad (4)$$

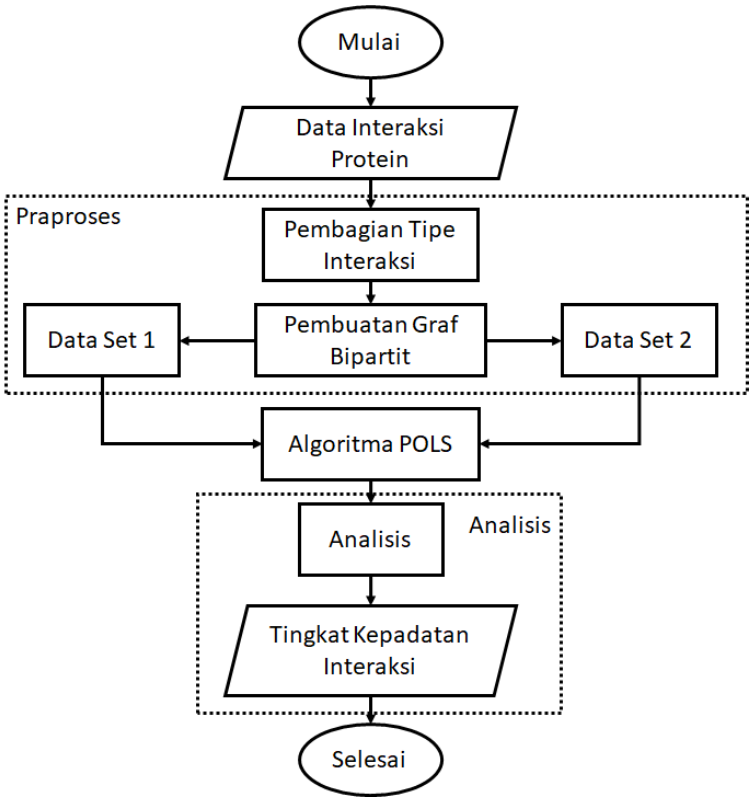
dengan $score_{lb}(u,v) = 1$ dan $score_{ub}(u,v) = \min\{|N(v)| - |N(v) \cap S|, |N(u)| - |N(u) \cap S|\}$. Nilai *pscore* terbesar disebut *add rule*.

2.3. Data

Data interaksi protein manusia dan HIV-1 merujuk pada [10] diperoleh melalui situs NCBI (*National Centre for Biotechnology Information*). Data dapat diakses melalui https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/browse/. Data tersebut merupakan sebuah tabel yang terdiri dari protein manusia, ID gen manusia, simbol gen manusia, protein HIV-1, ID gen HIV-1, symbol gen HIV-1, dan jenis interaksi. Data tersebut memiliki 16215 interaksi protein manusia dan HIV-1, 3797 protein manusia, 23 protein HIV, dan 130 jenis interaksi.

3 Metode Penelitian

Proses penelitian dibagi menjadi tiga tahap utama, yaitu: praproses data, implementasi, dan analisis. Praproses data merupakan tahapan transformasi data yang dimiliki menjadi sebuah dataset yang siap untuk diimplementasikan pada program. Tahap implementasi merupakan proses penggunaan Algoritma POLS pada dataset yang telah dimiliki. Tahap analisis merupakan proses untuk mengetahui tingkat kepadatan interaksi hasil *bicluster*. Keseluruhan proses penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitan

Berdasarkan Gambar 1, tahap praproses terdiri dari empat bagian. Pembagian tipe interaksi merupakan proses pembagian 130 jenis interaksi menjadi tiga tipe. Terdapat 69 jenis interaksi pada tipe pertama, 48 jenis interaksi pada tipe kedua, dan 13 jenis interaksi pada tipe ketiga. Pembagian seluruh jenis interaksi mengikuti Algoritma Dataset penelitian [11]. Tipe

pertama merupakan interaksi dari HIV-1 ke manusia. Tipe kedua merupakan interaksi dari manusia ke HIV-1. Tipe ketiga merupakan interaksi dua arah. Pembuatan graf bipartite dilakukan dengan memberikan nilai pada tipe interaksi. Berikan nilai 1 apabila tipe pertama, nilai -1 apabila tipe kedua dan X apabila tipe ketiga dan 0 apabila tidak ada interaksi. Dataset 1 merupakan kumpulan nilai 0, 1 dan X. Dataset 2 merupakan kumpulan nilai 0, -1, dan X.

Tahapan Algoritma POLS merupakan implementasi dari langkah-langkah Algoritma POLS dan teori graf. Tahapan analisis merupakan perhitungan kepadatan interaksi dengan menggunakan perbandingan sederhana.

4 Hasil dan Pembahasan

Hasil *bicluster* Algoritma POLS merupakan sebuah *biclique* seimbang. *Biclique* seimbang merupakan subset dari suatu graf bipartit dengan jumlah busur yang sama. Apabila $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{E})$ merupakan representasi dari dataset, maka hasil *bicluster* $\boldsymbol{H} = (\boldsymbol{U}', \boldsymbol{V}', \boldsymbol{E}')$ merupakan subset dari subset dari dataset $(\boldsymbol{H} \subseteq \boldsymbol{G})$, maka $|\boldsymbol{U}'| = |\boldsymbol{V}'|$. Pada penerapan *bicluster*, *biclique* seimbang mengharuskan terpenuhinya dua syarat: jumlah baris dan kolom sama; serta nilai baris dan kolom harus terisi (bukan 0). Sehingga submatriks merupakan matriks persegi. Selanjutnya entri-entri submatriks tersebut akan dianalisis tingkat kepadatan interaksi protein. Kepadatan interaksi merupakan perbandingan nilai 1 dan 0 pada submatriks hasil *bicluster*. Semakin tinggi tingkat kepadatan, maka hasil *bicluster* semakin baik.

Data penelitian ini dibagi menjadi 2, yaitu Dataset 1 dan Dataset 2, sehingga diperoleh hasil untuk dataset 1 dan dataset 2. Selanjutnya akan dijelaskan untuk hasil *bicluster* setiap dataset.

4.1. Dataset 1

Dataset 1 menghasilkan 297 bicluster dengan bicluster berukuran $\mathbf{2 \times 2}$ sebanyak 171, $\mathbf{3 \times 3}$ sebanyak 35, $\mathbf{4 \times 4}$ sebanyak 78, $\mathbf{5 \times 5}$ sebanyak 5, $\mathbf{6 \times 6}$ sebanyak 7, dan satu bicluster berukuran $\mathbf{7 \times 7}$. Hasil bicluster dataset 1 dapat dilihat pada Tabel 2. Terlihat bahwa seluruh hasil bicluster merupakan submatriks dari dataset 1 dengan jumlah baris dan kolom sama (matriks persegi).

Tabel 2: Hasil *Bicluster* Dataset 1

Ukuran *Bicluster*	**Jumlah**
2×2	171
3×3	35
4×4	78
5×5	5
6×6	7
7×7	1

Seluruh 297 *bicluster* kemudian dilihat tingkat kepadatan *bicluster*. Tingkat kepadatan *bicluster* diperoleh dengan cara:

$$D = \frac{n}{N} \tag{5}$$

dengan D merupakan tingkat kepadatan *bicluster*, n adalah jumlah entri *bicluster* yang tak 0 dan N adalah total entri *bicluster*.

Tabel 3: Tingkat Kepadatan *Bicluster* Dataset 1

Jumlah *Bilcuster*	Presentase	Tingkat Kepadatan
228	$\mathbf{76,77}\%$	$> \mathbf{0,5}$
22	$\mathbf{7,41}\%$	$> \mathbf{0,75}$
3	$\mathbf{1,01}\%$	$> \mathbf{1}$

Berdasarkan Persamaan 5 dan data Tabel 2, terdapat 228 *bicluster* dengan tingkat kepadatan lebih dari 0,5. Jumlah ini setara dengan 76,77% dari total hasil *bicluster*. Dari 297 *bicluster*, terdapat 22 *bicluster* (7,41%) yang memiliki tingkat kepadatan lebih dari 0,75 dan hanya terdapat 3 (1,01%) *bicluster* yang memiliki tingkat kepadatan 1. Rangkuman tingkat kepadatan *bicluster* dapat dilihat pada Tabel 3.

21 dari 22 *bicluster* dengan tingkat kepadatan lebih dari 0,75 masing-masing merupakan matriks $\mathbf{2 \times 2}$ dan hanya terdapat satu *bicluster* yang berukuran $\mathbf{3 \times 3}$. Selain itu, *bicluster* berukuran $\mathbf{3 \times 3}$ tersebut juga merupakan *bicluster* dengan tingkat kepadatan 1. Sehingga pada *bicluster* dengan tingkat kepadatan 1, terdapat 2 *bicluster* $\mathbf{2 \times 2}$ dan 1 *bicluster* $\mathbf{3 \times 3}$.

4.2. Dataset 2

Proses dan cara implementasi pada dataset 2 sama dengan dataset 1, sehingga diperoleh: 203 *bicluster* dengan *bicluster* berukuran $\mathbf{2 \times 2}$ sebanyak 110, $\mathbf{3 \times 3}$ sebanyak 24, $\mathbf{4 \times 4}$ sebanyak 59, $\mathbf{5 \times 5}$ sebanyak 5, $\mathbf{6 \times 6}$ sebanyak 4, dan satu *bicluster* berukuran $\mathbf{7 \times 7}$. Ukuran dan jumlah *bicluster* dapat lebih mudah dipahami melalui Tabel 4.

Tabel 4: Hasil *Bicluster* Dataset 2

Ukuran *Bicluster*	Jumlah
2×2	110
3×3	24
4×4	59
5×5	5
6×6	4
7×7	1

Terdapat 11 *bicluster* dengan tingkat kepadatan 1, 17 *bicluster* dengan tingkat kepadatan lebih dari 0,75, dan 182 *bicluster* dengan tingkat kepadatan lebih dari 0,5. Presentase *bicluster* dengan tingkat kepadatan lebih dari 0,5, lebih dari 0,75, dan 1 secara berturut-turut adalah 89,65%, 8,37%, dan 5,42%. Presentase tingkat kepadatan *bicluster* dapat pula dilihat pada Tabel 5. Seluruh *bicluster* dengan tingkat kepadatan 1 dan lebih dari 0,75 berukuran $\mathbf{2 \times 2}$.

Tabel 5: Tingkat Kepadatan *Bicluster* Dataset 2

Jumlah *Bicluster*	Presentase	Tingkat Kepadatan
182	$\mathbf{89,65\%}$	$> \mathbf{0,5}$
17	$\mathbf{8,37\%}$	$> \mathbf{0,75}$
11	$\mathbf{5,42\%}$	$\mathbf{1}$

4.3. Analisis

Berdasarkan Tabel 3 dan 5, presentase untuk $\boldsymbol{D = 1}$ tidak lebih dari $\mathbf{6\%}$, bahkan pada Dataset 1 (Tabel 3) hanya mencapai $\mathbf{1,01\%}$. Selain itu, berdasarkan penjelasan *bicluster* Dataset 1 dan 2, hampir seluruh *bicluster* dengan $\boldsymbol{D = 1}$ berukuran $\mathbf{2 \times 2}$. Hanya terdapat 1 *bicluster* berukuran $\mathbf{3 \times 3}$ pada Dataset 1. Sehingga dapat dikatakan bahwa, dari total 400 *bicluster*, hanya terdapat 14 *bicluster* yang memenuhi kriteria submatriks persegi dan memiliki tingkat kepadatan sama dengan 1. Berdasarkan Tabel 2 dan 4, hanya terdapat satu *bicluster* yang berukuran $\mathbf{7 \times 7}$ pada masing-masing Dataset 1 dan 2. Tingkat kepadatan *bicluster* tersebut tidak mencapai $\mathbf{0,5}$ sehingga tidak dapat dikatakan *bicluster* yang baik.

Pada permasalahan interaksi protein manusia dan HIV-1, jumlah (ukuran) protein manusia dan HIV-1 jauh berbeda. Protein manusia berjumlah 3797 dan HIV-1 berjumlah 23. Perbedaan jumlah yang signifikan tersebut menyebabkan hasil Algoritma POLS didominasi dengan *bicluster* berukuran $\mathbf{2 \times 2}$. Dilihat dari beberapa pemaparan, maka Aloritma POLS lebih baik menghasilkan *bicluster* berukuran kecil ($\mathbf{2 \times 2}$ dan $\mathbf{3 \times 3}$). Hal ini dikarenakan hanya *bicluster* berukuran kecil yang memiliki tingkat kepadatan 1. Selain itu Algoritma POLS lebih cocok untuk permasalahan dengan data yang seimbang (jumlah simpul yang sama).

5 Kesimpulan

Berdasarkan hasil penelitian, dari 400 *bicluster*, terdapat 14 *bicluster* yang memiliki tingkat kepadatan 1. Kemudian presentase hasil *bicluster* Algoritma POLS dengan tingkat kepadatan 1 pada dataset 1 sebesar 1,01% dan pada dataset 2 sebesar 4,42%. *Bicluster* Algoritma POLS yang berukuran kecil (2×2 dan 3×3) lebih optimal karena memiliki tingkat kepadatan 1.

Daftar Pustaka

[1] A. Trkola, “HIV – host interactions : vital to the virus and key to its inhibition,” 2004, doi: 10.1016/j.mib.2004.06.002.

[2] E. Frankenberg, “基因的改变NIH Public Access,” *Bone*, vol. 23, no. 1, pp. 1–7, 2012, [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf.

[3] A. Bustamam, T. Siswantining, T. P. Kaloka, and O. Swasti, “Application of BiMax, POLS, and LCM-MBC to Find Bicluster on Interactions Protein between HIV-1 and Human,” *Austrian J. Stat.*, vol. 49, no. 3, pp. 1–18, Feb. 2020, doi: 10.17713/ajs.v49i3.1011.

[4] G. Ardaneswari, A. Bustamam, and D. Sarwinda, "Implementation of plaid model biclustering method on microarray of carcinoma and adenoma tumor gene expression data," *J. Phys. Conf. Ser.*, vol. 893, no. 1, 2017, doi: 10.1088/1742-6596/893/1/012046.
[5] A. Mukhopadhyay, U. Maulik, and S. Bandyopadhyay, "On Biclustering of Gene Expression Data On Biclustering of Gene Expression Data," no. December 2013, 2010, doi: 10.2174/157489310792006701.
[6] G. Ardaneswari, A. Bustamam, and T. Siswantining, "Implementation of parallel k-means algorithm for two-phase method biclustering in Carcinoma tumor gene expression data," *AIP Conf. Proc.*, vol. 1825, 2017, doi: 10.1063/1.4978973.
[7] Y. Cheng and G. M. Church, "Biclustering of expression data.," *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, vol. 8, pp. 93–103, 2000.
[8] J. Yang, H. Wang, W. Wang, and P. Yu, "Enhanced biclustering on expression data," *Proc. - 3rd IEEE Symp. Bioinforma. Bioeng. BIBE 2003*, pp. 321–327, 2003, doi: 10.1109/BIBE.2003.1188969.
[9] Y. Wang, S. Cai, and M. Yin, "New heuristic approaches for maximum balanced biclique problem," *Inf. Sci. (Ny).*, vol. 432, pp. 362–375, 2018, doi: 10.1016/j.ins.2017.12.012.
[10] A. Mukhopadhyay, S. Ray, and U. Maulik, "Incorporating the type and direction information in predicting novel regulatory interactions between HIV-1 and human proteins using a biclustering approach," *BMC Bioinformatics*, vol. 15, no. 1, 2014, doi: 10.1186/1471-2105-15-26.
[11] T. P. Kaloka, A. Bustamam, D. Lestari, and W. Mangunwardoyo, "POLS algorithm to find a local bicluster on interactions between HIV-1 proteins and human proteins," *AIP Conf. Proc.*, vol. 2084, no. March, 2019, doi: 10.1063/1.5094280.