

Sistem Pendukung Keputusan untuk Identifikasi Protein Kunci pada Kanker Darah: Integrasi MCDM, KMeans, dan Topologi Jaringan

Mohammad Hamim Zajuli Al Faroby¹, Muhammad Dzulfikar Fauzi²

^{1,2}School of Computing, Universitas Telkom
Surabaya, Indonesia

e-mail: alfarobymhz@telkomuniversity.ac.id, dzulfikarf@telkomuniversity.ac.id

Abstrak - Kanker darah merupakan salah satu penyakit kompleks yang dipicu oleh gangguan pada regulasi jalur pensinyalan seluler, salah satunya melibatkan mutasi pada protein JAK2. Mengingat pentingnya JAK2 dalam patogenesis kanker hematologi, diperlukan pendekatan sistemik berbasis data untuk mengidentifikasi protein-protein yang memiliki asosiasi fungsional dengannya secara menyeluruh. Penelitian ini bertujuan mengembangkan sistem pendukung keputusan (SPK) untuk mengidentifikasi protein kunci dalam jaringan interaksi protein (PPI) terkait JAK2, melalui integrasi algoritma KMeans *clustering*, fitur topologi graf, dan metode *Multi-Criteria Decision Making* (MCDM). Data PPI diperoleh dari STRING-DB dan divisualisasikan menggunakan Cytoscape. Delapan fitur topologi jaringan diekstraksi sebagai dasar analisis, di antaranya *degree*, *stress*, dan *neighborhood connectivity*. Proses klusterisasi menghasilkan tiga kelompok optimal yang divalidasi menggunakan *Silhouette Score* dan *Davies-Bouldin Index*. Selanjutnya, model MCDM diterapkan untuk mengevaluasi kontribusi masing-masing kluster secara agregat terhadap kestabilan jaringan. Hasil penelitian menunjukkan bahwa Cluster 2 memiliki skor MCDM tertinggi (0,6667), ditopang oleh nilai *stress* dan *degree* yang sangat tinggi, mengindikasikan peran strategis protein dalam kluster tersebut sebagai hub utama dalam jaringan. Temuan ini memberikan landasan kuat untuk eksplorasi kandidat target terapi baru yang potensial dalam konteks kanker darah, serta menegaskan efektivitas integrasi teknik analisis graf dan SPK berbasis MCDM dalam studi jaringan molekuler.

Kata Kunci: Multi-Criteria Decision Making, Kmeans, Centrality, Kanker Darah, PPI.

PENDAHULUAN

Kanker darah merupakan penyakit yang melibatkan deregulasi kompleks pada jalur pensinyalan seluler, salah satunya adalah mutasi pada protein Janus kinase 2 (JAK2) (Zon et al., 2024). Protein ini memainkan peran sentral dalam proliferasi dan diferensiasi sel hematopoietik, sehingga menjadi salah satu fokus utama dalam pengembangan terapi target. Meski demikian, pemahaman menyeluruh mengenai bagaimana JAK2 berinteraksi dan berasosiasi dengan protein lain dalam lanskap molekuler kanker darah masih terbatas. Kompleksitas hubungan antarprotein dalam jaringan interaksi protein (PPI) menuntut pendekatan sistemik dan berbasis data untuk mengungkap relasi fungsional yang tersembunyi (Amiroch et al., 2022).

Sebagian besar pendekatan konvensional dalam analisis jaringan interaksi protein (PPI) masih bersifat reduksionistik, dengan fokus utama pada individu protein atau relasi langsung antar-node. Hal ini mengabaikan kompleksitas struktural dan fungsional dari jaringan biologis yang bersifat dinamis dan hierarkis (Cook et al., 2018). Fitur-fitur topologis seperti *degree centrality*, *betweenness*, *closeness*, dan *clustering coefficient* telah terbukti mampu merepresentasikan kontribusi fungsional protein dalam konteks global jaringan. Meskipun analisis graf telah banyak digunakan dalam bidang

bioinformatika, penggunaannya masih jarang dikombinasikan secara sistematis dengan teknik *unsupervised learning* untuk eksplorasi awal kandidat target terapeutik (Kang et al., 2021). Khususnya dalam konteks kanker darah, belum banyak studi yang memanfaatkan integrasi penuh antara ekstraksi fitur graf dan teknik klusterisasi berbasis KMeans untuk mengidentifikasi kelompok protein fungsional yang berkaitan dengan JAK2 (Al Faroby, Alifah, et al., 2022).

Penelitian berbasis data jarang sekali sampai pada proses pengambilan keputusan berbasis data hasil eksplorasi jaringan. Mayoritas penelitian berhenti pada tahap visualisasi atau identifikasi kluster tanpa mekanisme terstruktur untuk menyaring dan memprioritaskan hasil. Untuk menjembatani gap tersebut, penelitian ini mengusulkan integrasi KMeans berbasis fitur graf dengan sistem pendukung keputusan (*Decision Support System/DSS*) berbasis metode *Multi-Criteria Decision Making* (MCDM) (Demir et al., 2024). Sistem ini tidak hanya mengelompokkan protein berdasarkan pola topologi yang mirip, tetapi juga mengevaluasi bobot kontribusi setiap fitur dalam konteks biologis, memungkinkan pemeringkatan kandidat protein secara objektif. Keunggulan pendekatan ini terletak pada kemampuannya untuk menggabungkan eksplorasi tidak terawasi dengan penilaian multi-kriteria yang



transparan, sistematis, dan kontekstual (Ellis & Kennedy, 2025). Dengan demikian, metodologi ini menawarkan kerangka kerja yang lebih utuh dalam menjawab tantangan eksplorasi awal target molekuler dalam jaringan kompleks seperti PPI.

Tujuan dari penelitian ini adalah mengembangkan sistem pendukung keputusan berbasis integrasi KMeans dan analisis graf untuk mengidentifikasi kelompok protein yang memiliki asosiasi erat dengan JAK2. Sistem ini diharapkan mampu memprioritaskan protein kunci sebagai kandidat target terapeutik baru secara objektif dan berbasis data, sehingga dapat memperkuat strategi eksplorasi awal dalam riset kanker darah berbasis biologi sistem.

METODE PENELITIAN

Penelitian ini diawali dengan pengumpulan dan persiapan data interaksi protein. Data diakuisisi pada basis data STRING-DB yang merupakan platform terpusat untuk data PPI (Szklarczyk et al., 2021). Fokus pengambilan data diarahkan pada subgraf yang berhubungan langsung dengan protein JAK2 dan protein-protein terkait lainnya yang relevan dalam konteks kanker darah. Setelah data terkumpul, dilakukan pra-pemrosesan untuk membentuk struktur graf yang merepresentasikan jaringan interaksi protein secara komprehensif. Data tersebut diimpor ke dalam perangkat lunak Cytoscape untuk memodelkan jaringan PPI secara visual dan komprehensif.

Selanjutnya, dilakukan ekstraksi fitur topologi dari graf PPI tersebut dengan menghitung berbagai metrik sentralitas penting, seperti *degree centrality*, *betweenness centrality*, *closeness centrality*, *neighborhood connectivity*, *radiality*, *stress*, *topological coefficient*, serta *clustering coefficient* (Amiroch et al., 2024). Setiap protein dalam jaringan direpresentasikan sebagai vektor fitur numerik berdasarkan nilai-nilai metrik tersebut, sehingga karakteristik topologis tiap simpul dapat dianalisis secara kuantitatif (Al Faroby, Fadhilah, et al., 2022).

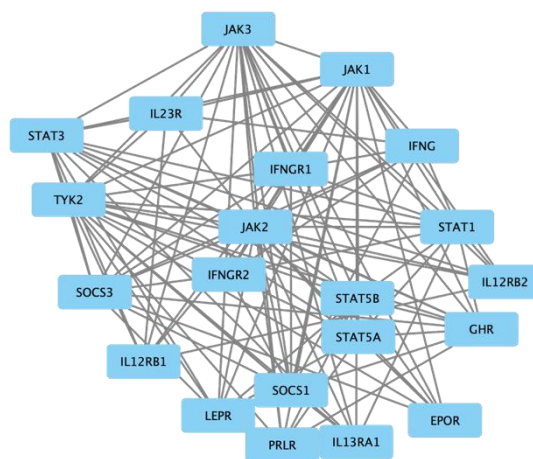
Tahap klusterisasi dilakukan dengan algoritma KMeans untuk mengelompokkan protein berdasarkan kesamaan pola fitur topologisnya (Shirmohammady et al., 2021). Penentuan jumlah kluster optimal dilakukan dengan evaluasi menggunakan metrik validasi kluster seperti *Silhouette Score* (Bagirov et al., 2023) dan *Davies-Bouldin Index* (Ashari et al., 2023), untuk memastikan bahwa pembagian kelompok memiliki kohesi yang baik dan terpisah dengan jelas. Hasil klusterisasi ini selanjutnya divisualisasikan untuk memberikan gambaran struktur fungsional protein dalam jaringan.

Setelah mendapatkan kluster dan skor topologi, dilakukan pengembangan sistem pendukung keputusan (*Decision Support System/DSS*) berbasis metode *Multi-Criteria Decision Making* (MCDM). Metode *Analytic Hierarchy Process* (AHP) digunakan untuk memberikan bobot pada masing-masing fitur topologi sesuai kepentingannya dalam konteks

biologis. Sistem DSS ini kemudian menghitung skor prioritas setiap *cluster*, sehingga dapat diidentifikasi kandidat target terapeutik yang paling potensial. Implementasi sistem dilakukan dalam bentuk skrip atau antarmuka yang memudahkan pengguna dalam pengambilan keputusan.

HASIL DAN PEMBAHASAN

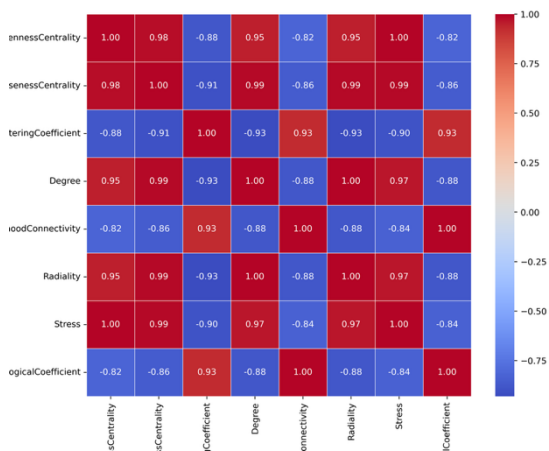
Hasil akuisisi data dari STRING-DB berhasil mengidentifikasi sebanyak 20 protein yang memiliki hubungan asosiasi dengan protein target JAK2. Jumlah ini diperoleh dengan menetapkan batas minimal skor kepercayaan interaksi sebesar 0,4 (*medium confidence*), yang merepresentasikan tingkat keyakinan moderat dalam hubungan biologis antarprotein berdasarkan berbagai sumber data, termasuk eksperimen laboratorium, basis data, dan prediksi komputasi. Parameter ini dipilih untuk menyeimbangkan antara cakupan interaksi dan kualitas data, agar hanya protein-protein dengan relevansi biologis yang cukup kuat terhadap JAK2 yang terlibat. Seluruh data interaksi ini kemudian divisualisasikan menggunakan Cytoscape, menghasilkan sebuah graf jaringan interaksi protein seperti ditampilkan pada Gambar 1. Visualisasi tersebut memperlihatkan bahwa JAK2 terhubung secara langsung maupun tidak langsung dengan protein-protein lain, membentuk jejaring interaksi yang kompleks. Keterhubungan ini mengindikasikan bahwa 20 protein tersebut kemungkinan memiliki keterlibatan dalam jalur sinyal atau proses biologis yang saling memengaruhi dengan fungsi JAK2, yang dikenal berperan penting dalam regulasi proliferasi sel darah dan telah dikaitkan dengan patogenesis berbagai jenis kanker darah.



Gambar 1. Visual interaksi protein JAK2 dengan aplikasi Cytoscape

Original data dari PPI merupakan graf tidak berarah, sehingga untuk mendapatkan fiturnya metode *centrality* digunakan. Hasil ekstraksi ciri tersebut membentuk matriks data 21×8 (21 baris data dan 8 kolom fitur). Sebelum dilakukan proses *clustering*, analisis korelasi fitur dilakukan untuk mengetahui apakah antara fitur memiliki hubungan satu dengan

yang lain. Hasil perhitungan korelasi disajikan pada Gambar 2, dimana menunjukkan *heatmap* yang merepresentasikan skor korelasi masing-masing fitur.



Gambar 2. Heatmap korelasi dari delapan fitur yang terdapat pada dataset.

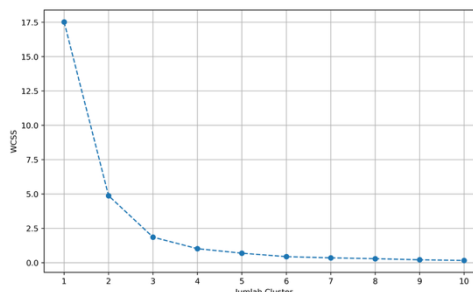
Gambar 2 merupakan heatmap korelasi yang menggambarkan hubungan antar delapan fitur jaringan (network features) hasil perhitungan centrality dan metrik topologi lainnya terhadap 20 protein yang terhubung dengan JAK2. Warna merah menunjukkan korelasi positif yang tinggi (mendekati +1), sedangkan warna biru menunjukkan korelasi negatif yang tinggi (mendekati -1). Skala warna di sisi kanan menunjukkan intensitas nilai korelasi, dengan nilai absolut mendekati 1 mengindikasikan hubungan linier yang kuat.

Dari heatmap terlihat bahwa hampir semua pasangan fitur memiliki nilai korelasi yang signifikan, baik secara positif maupun negatif. Contohnya, fitur *Closeness Centrality* dan *Stress* memiliki korelasi sangat tinggi ($r = 0.99$), demikian pula antara *Degree* dan *Radiality* ($r = 1.00$). Sebaliknya, terdapat korelasi negatif kuat antara *Clustering Coefficient* dengan *Closeness Centrality* ($r = -0.91$) dan *Radiality* ($r = -0.93$), yang menunjukkan bahwa protein dengan konektivitas lokal tinggi (clustering) cenderung memiliki nilai centrality global yang rendah.

Analisis ini memperkuat keputusan metodologis bahwa seluruh fitur yaitu *Betweenness Centrality*, *Closeness Centrality*, *Clustering Coefficient*, *Degree*, *Neighborhood Connectivity*, *Radiality*, *Stress*, dan *Topological Coefficient* layak digunakan bersama dalam proses *Multi-Criteria Decision Making* (MCDM). Meskipun terdapat korelasi tinggi antar beberapa fitur, keberadaan korelasi negatif maupun variasi korelasi antar fitur justru menunjukkan bahwa setiap metrik memberikan perspektif unik terhadap struktur jaringan.

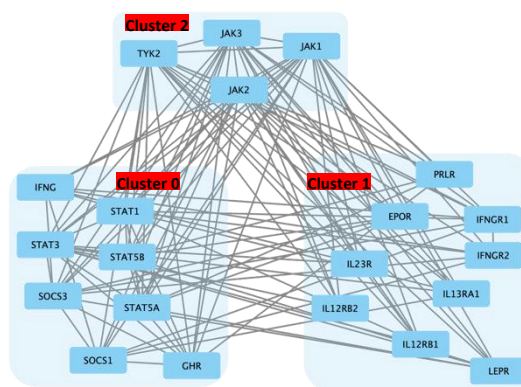
Dataset dilakukan proses standarisasi sebelum dimodelkan menggunakan kmeans. Hasil dari pemodelan kmeans menunjukkan *cluster* yang optimal untuk digunakan memodelkan dataset yakni berjumlah 3. Keputusan ini berdasarkan hasil analisis pada Gambar 3, yang merupakan visualisasi dari

Elbow Method. Selain itu jumlah 3 *cluster* dipilih karena nilai *Silhouette* sebesar 0.5575 lebih tinggi dari formasi *cluster* yang lain. Nilai *Davies-Bouldin Index* juga memiliki nilai paling kecil dibanding formasi *cluster* lain yakni 0.5678.



Gambar 3. Analisis *Elbow Method* untuk menentukan jumlah cluster optimal

Hasil pengelompokan 21 data protein ditunjukkan pada Gambar 4. Setiap kotak biru pada gambar merepresentasikan *cluster* protein yang memiliki kemiripan berdasarkan fitur topologi jaringan. Protein-protein di dalam satu *cluster* diasumsikan memiliki peran fungsional atau karakteristik topologis yang serupa dalam konteks interaksi molekuler. Pada gambar, terlihat bahwa protein JAK2, yang menjadi fokus utama dalam studi kanker darah, berada dalam satu kluster bersama protein JAK1, JAK3, dan TYK2, yang semuanya merupakan bagian dari keluarga *Janus kinase* (Al Faroby, Alifah, et al., 2022). Kluster ini merepresentasikan modul regulasi pensinyalan yang erat. Dua kluster lainnya tampak berisi protein-protein seperti STATs, SOCS, dan reseptor sitokin, yang merupakan bagian dari jalur pensinyalan JAK-STAT.



Gambar 4. Visualisasi *cluster* asosiasi protein JAK2 menggunakan kmeans clustering.

Sebelum membangun sistem pendukung keputusan untuk menentukan cluster mana yang paling berpengaruh, perlu ditentukan terlebih dahulu fitur penting yang dihasilkan model *cluster*. Hasil evaluasi pengaruh fitur topologi terhadap model KMeans menunjukkan bahwa fitur *Stress* memiliki

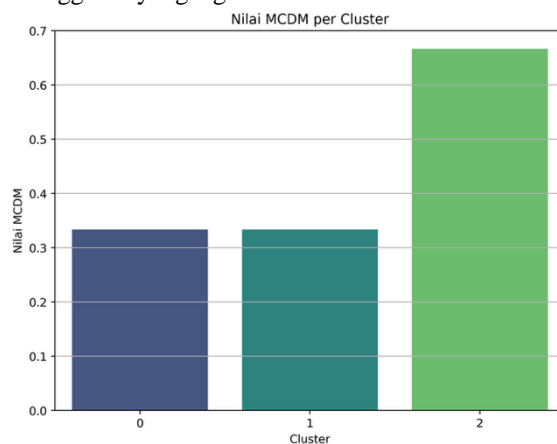
kontribusi paling signifikan dalam pembentukan klaster, dengan nilai deviasi antar centroid mencapai 84,46, jauh melampaui fitur-fitur lain seperti *Degree* (5,94) dan *Neighborhood Connectivity* (2,12). Nilai deviasi standar yang tinggi ini mengindikasikan bahwa distribusi nilai Stress antar klaster sangat bervariasi, sehingga fitur ini mampu secara efektif membedakan kelompok protein berdasarkan peran sentral mereka dalam jaringan. Sebaliknya, fitur-fitur seperti *Clustering Coefficient* (0,19), *Closeness Centrality* (0,17), dan *Betweenness Centrality* (0,046) menunjukkan deviasi yang sangat rendah, menandakan kontribusi yang minimal dalam proses segmentasi data oleh algoritma KMeans. Temuan ini memperkuat argumen bahwa struktur global jaringan—terutama yang berkaitan dengan lalu lintas informasi biologis—lebih menentukan dalam pemisahan kelompok protein daripada keterhubungan lokal. Oleh karena itu, fitur Stress dapat diprioritaskan sebagai parameter kunci dalam pengembangan sistem pendukung keputusan berbasis MCDM untuk identifikasi protein kunci atau target terapeutik potensial dalam konteks kanker darah.

Model *Multi-Criteria Decision Making* (MCDM) yang diterapkan pada hasil klasterisasi KMeans bertujuan untuk mengidentifikasi kelompok protein yang paling strategis berdasarkan tiga fitur topologi utama: Stress, Degree, dan Neighborhood Connectivity, dengan masing-masing diberikan bobot yang telah dinormalisasi sebesar 0,333. Penentuan bobot yang seimbang ini mencerminkan anggapan awal bahwa ketiga fitur tersebut sama pentingnya dalam membedakan peran biologis suatu protein dalam jaringan interaksi.

Berdasarkan nilai rata-rata fitur pada tiap klaster, terlihat bahwa Cluster 2 menunjukkan dominasi yang sangat kuat, terutama pada fitur Stress dengan nilai mencapai 163,00, dan Degree sebesar 19,00. Hal ini mengindikasikan bahwa protein-protein dalam klaster ini sering terlibat dalam lintasan informasi biologis (jalur terpendek) dan memiliki jumlah koneksi langsung yang tinggi, suatu karakteristik khas dari protein hub yang memiliki peran fungsional penting dalam stabilitas dan dinamika sistem biologis, khususnya dalam jaringan protein yang berkaitan dengan kanker darah. Sebaliknya, Cluster 1 memiliki nilai tertinggi pada Neighborhood Connectivity sebesar 15,07, yang mengindikasikan lingkungan lokal yang padat dan kohesif, meskipun nilai Stress dan Degree-nya jauh lebih rendah, masing-masing hanya 4,44 dan 7,33. Cluster 0 berada di posisi tengah untuk semua kriteria, dengan nilai Stress (33,25), Degree (11,25), dan Neighborhood Connectivity (12,87) yang relatif moderat, tanpa dominasi signifikan pada aspek manapun.

Ketika nilai rata-rata fitur tersebut diskalakan dalam rentang 0 hingga 1, hasilnya mempertegas dominasi Cluster 2. Nilai scaled Stress dan Degree

pada Cluster 2 masing-masing adalah 1,0, sementara nilai Neighborhood Connectivity-nya adalah 0,00, menandakan spesialisasi pada struktur jaringan global daripada lokal. Cluster 1 mendapat nilai tertinggi (1,0) hanya pada Neighborhood Connectivity, tetapi bernilai 0 pada dua fitur lainnya, yang menunjukkan keterbatasan pengaruh dalam jaringan secara keseluruhan. Cluster 0 memiliki nilai menengah: Stress (0,1817), Degree (0,3357), dan Neighborhood Connectivity (0,4826), namun tidak menunjukkan keunggulan yang signifikan secara relatif.



Gambar 5. Hasil Multi-Criteria Decision Making pada setiap cluster.

Melalui perhitungan skor akhir MCDM berdasarkan bobot dan skor terstandar (Gambar 5), Cluster 2 memperoleh skor tertinggi sebesar 0,6667, menandakan bahwa klaster ini adalah kelompok protein paling berpengaruh secara topologis dan fungsional dalam jaringan. Cluster 1 dan Cluster 0 masing-masing memperoleh skor 0,3333 dan 0,3333, dengan perbedaan sangat kecil di antara keduanya (selisih pada digit keempat desimal), mencerminkan bahwa keduanya kurang menonjol secara agregat. Temuan ini mengukuhkan bahwa Cluster 2 memiliki potensi paling tinggi untuk mengandung protein-protein kunci yang dapat dieksplorasi lebih lanjut sebagai kandidat biomarker atau target terapi, khususnya yang memiliki asosiasi dengan JAK2 dalam konteks kanker darah. Oleh karena itu, integrasi antara KMeans clustering dan MCDM tidak hanya memberikan pendekatan eksploratif yang kuat, tetapi juga menyediakan kerangka sistem pendukung keputusan yang terukur dan berbasis bukti kuantitatif dalam studi jaringan biologi molekuler.

Hasil analisis berbasis Multi-Criteria Decision Making (MCDM) yang menggabungkan tiga indikator topologis—Degree, Stress, dan Neighborhood Connectivity—memberikan gambaran menyeluruh tentang peran strategis masing-masing klaster dalam jaringan protein. Nilai akhir MCDM menunjukkan bahwa Cluster 2 unggul secara signifikan dibandingkan Cluster 0 dan 1, dengan skor 0,6667, mengindikasikan bahwa simpul-simpul dalam klaster ini memiliki kontribusi paling besar

terhadap kestabilan, konektivitas, dan efektivitas jaringan secara keseluruhan.

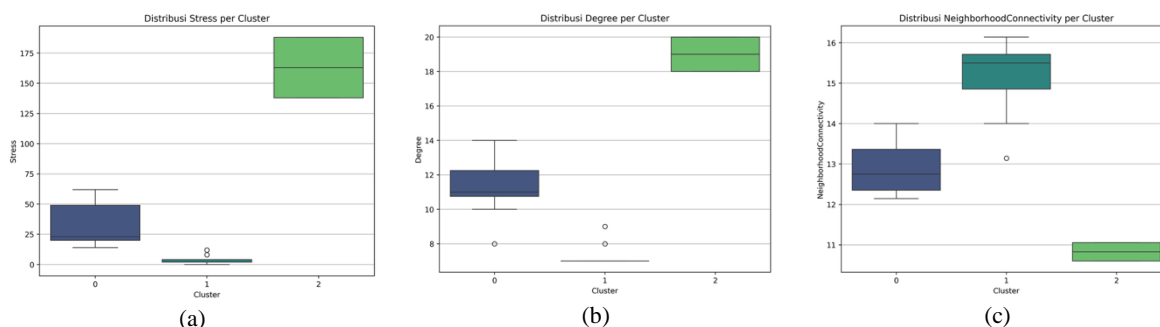
Jika dilihat secara terpisah seperti pada Gambar 6, Cluster 2 memiliki nilai median tertinggi pada dua metrik paling strategis dalam konteks global jaringan, yaitu Degree (~19) dan Stress (~160). Ini berarti node dalam Cluster 2 tidak hanya terhubung ke banyak node lain secara langsung (degree tinggi), tetapi juga menjadi jalur lintasan penting dalam rute terpendek di seluruh jaringan (stress tinggi). Kombinasi ini menunjukkan bahwa node dalam Cluster 2 berfungsi sebagai hub sentral yang menjaga efisiensi dan robust-nya arsitektur jaringan. Dalam konteks implementasi MCDM, hal ini sangat penting karena mengindikasikan bahwa Cluster 2 memiliki nilai strategis global, menjadikannya target utama dalam pengembangan terapi atau intervensi biologis berbasis jaringan.

Sementara itu, Cluster 1 memang unggul dalam Neighborhood Connectivity (median ~15.5), yang menunjukkan kohesi lokal yang kuat, namun performanya rendah pada Degree (~7) dan Stress (~7). Artinya, meskipun node di Cluster 1 berada dalam komunitas yang saling terhubung erat, mereka tidak memainkan peran penting secara global dalam

jaringan. Ini menjelaskan mengapa skor MCDM-nya tetap rendah, karena bobot kombinasi dari kriteria tersebut lebih condong ke parameter yang merepresentasikan pengaruh sistemik.

Cluster 0 berada di posisi tengah, dengan nilai Degree (~11) dan Neighborhood Connectivity (~12.8) yang cukup moderat, namun nilai Stress-nya (~30) masih jauh di bawah Cluster 2. Peran Cluster 0 dapat dikatakan sebagai node semi-pendukung—memiliki konektivitas sedang namun kurang vital dalam rute informasi utama jaringan.

Secara keseluruhan, MCDM berhasil mengintegrasikan kompleksitas jaringan ke dalam satu skor yang merepresentasikan kepentingan sistemik. Dalam implementasinya, skor ini sangat berguna untuk memprioritaskan simpul-simpul atau klaster dalam proses seperti drug targeting, identifikasi biomarker, atau pengembangan terapi sistemik. Oleh karena itu, hasil ini mendukung validasi bahwa Cluster 2 adalah kandidat utama yang layak menjadi fokus dalam penelitian lanjutan atau intervensi berbasis jaringan, karena posisinya sebagai core hub yang menopang kestabilan dan efisiensi jaringan secara keseluruhan.



Gambar 6. Hasil analisis distribusi kriteria yang digunakan pada MCDM pada setiap cluster, (a) distribusi untuk fitur Stress, (b) distribusi untuk fitur Degree, dan (c) distribusi untuk fitur NeighborhoodConnectivity

KESIMPULAN

Penelitian ini berhasil mengembangkan sebuah sistem pendukung keputusan berbasis integrasi metode KMeans clustering, fitur topologi jaringan, dan Multi-Criteria Decision Making (MCDM) untuk mengidentifikasi protein kunci dalam jaringan interaksi yang berasosiasi dengan JAK2. Hasil klusterisasi menggunakan delapan fitur topologi menunjukkan pembentukan tiga kelompok protein yang berbeda secara struktural, dengan validasi metrik Silhouette Score dan Davies-Bouldin Index yang menunjukkan pemisahan kluster yang optimal. Penerapan MCDM dengan bobot seimbang terhadap fitur stress, degree, dan neighborhood connectivity menghasilkan pemeringkatan yang objektif terhadap setiap kluster. Dari hasil perhitungan, Cluster 2 dinyatakan sebagai kluster paling strategis dengan skor MCDM tertinggi (0,6667), yang secara biologis menunjukkan dominasi protein dengan peran sentral sebagai hub informasi dalam jaringan PPI. Temuan ini menunjukkan bahwa pendekatan integratif yang digunakan mampu mengungkap struktur fungsional tersembunyi dalam jaringan biologis, serta memberikan rekomendasi berbasis data terhadap kandidat protein yang layak ditindaklanjuti sebagai target terapeutik potensial dalam riset kanker darah berbasis sistem.

REFERENSI

- Al Faroby, M. H. Z., Alifah, A. N., Aji, B. A. S., & Chaniago, I. O. (2022). Penerapan K-Means dan Analisis Jaringan untuk Identifikasi Protein-Meningitis Signifikan Berdasarkan Interaksi Antar Protein. *PROSIDING SEMINAR NASIONAL SAINS DATA*, 2(1), 39–50. <https://doi.org/10.33005/senada.v2i1.39>
- Al Faroby, M. H. Z., Fadhilah, H. N., & Sembiring, F. H. (2022). Identifikasi Interaksi Protein-Protein Meningitis Menggunakan ClusterONE dan Analisis Jaringan. *Journal of Advances in Information and Industrial Technology*, 4(1), 17–28. <https://doi.org/10.52435/jaiit.v4i1.180>
- Amiroch, S., Al Faroby, M. H. Z., & Fauzi, M. D. (2024). Target Baru Pengobatan Meningitis Berdasarkan Centrality Measure Jaringan Protein dan Self Organizing Map. *Limits: Journal of Mathematics and Its Applications*, 21(3), 397–416.
- Amiroch, S., Al Faroby, M. H. Z., Irawan, M. I., Mukhlash, I., & Nidhom, A. C. (2022). Analysis of protein-protein interaction to obtain significant protein in influenza virus type A/H9N2. *AIP Conference Proceedings*, 2498(1), 020021. <https://doi.org/10.1063/5.0083336>
- Ashari, I. F., Dwi Nugroho, E., Baraku, R., Novri Yanda, I., & Liwardana, R. (2023). Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta. *Journal of Applied Informatics and Computing*, 7(1), 89–97. <https://doi.org/10.30871/jaic.v7i1.4947>
- Bagirov, A. M., Aliguliyev, R. M., & Sultanova, N. (2023). Finding compact and well-separated clusters: Clustering using silhouette coefficients. *Pattern Recognition*, 135, 109144. <https://doi.org/10.1016/j.patcog.2022.109144>
- Cook, H. V., Doncheva, N. T., Szklarczyk, D., von Mering, C., & Jensen, L. J. (2018). Viruses.STRING: A virus-host protein-protein interaction database. *Viruses*, 10(10), 519. <https://doi.org/10.3390/v10100519>
- Demir, G., Chatterjee, P., & Pamucar, D. (2024). Sensitivity analysis in multi-criteria decision making: A state-of-the-art research perspective using bibliometric analysis. *Expert Systems with Applications*, 237, 121660. <https://doi.org/10.1016/j.eswa.2023.121660>
- Ellis, J. T., & Kennedy, P. J. (2025). Multi-criteria decision making and its application to in silico discovery of vaccine candidates for *Toxoplasma gondii*. *Vaccine*, 58, 127242. <https://doi.org/10.1016/j.vaccine.2025.127242>
- Kang, Z., Peng, C., Cheng, Q., Liu, X., Peng, X., Xu, Z., & Tian, L. (2021). Structured graph learning for clustering and semi-supervised classification. *Pattern Recognition*, 110, 107627. <https://doi.org/10.1016/j.patcog.2020.107627>
- Shirmohammady, N., Izadkhah, H., & Isazadeh, A. (2021). PPI-GA: A Novel Clustering Algorithm to Identify Protein Complexes within Protein-Protein Interaction Networks Using Genetic Algorithm. *Complexity*, 2021(1). <https://doi.org/10.1155/2021/2132516>
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, 49(D1), D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
- Zon, R. L., Sekar, A., Clapham, K., Oren, O., Niroula, A., Bick, A. G., Gibson, C. J., Griffin, G., Uddin, M. M., Neuberger, D., Natarajan, P., & Ebert, B. L. (2024). JAK2 mutant clonal hematopoiesis is associated with venous thromboembolism. *Blood*, 144(20), 2149–2154. <https://doi.org/10.1182/blood.2024024187>