

BÖLÜM 11

BİYOİNFORMATİK ALANINDA GRAF UYGULAMALARI

Fahriye GEMCİ¹

Bu bölümde, biyoinformatik alanında graf diğer bir adıyla çizge teorisinin kullanımını anlatılmıştır. Bu yüzden öncelikle biyoinformatik alanı, graf teorisi ve graf sinir ağları tanıtılmıştır. Ardından biyoinformatik alanında graf teorisinden ve graf sinir ağlarından yararlanılan makaleler ayrıntılı olarak incelenmiştir.

GİRİŞ

Günümüzde canlılara ait alanlarda verilerin oldukça hızlı bir şekilde artmasıyla büyük veri kümeleri ortaya çıkmıştır. Bu büyük verilerden anlamlı bilgiler çıkarabilmek amacıyla da bilgi teknolojileri ve biyoistatistik alanının kullanılmasıyla yeni bir alan olan biyoinformatik alanı ortaya çıkmıştır. Biyoinformatik alanı, hastalık tanısında ve ilaç keşfinde faydalandığı için oldukça önemli bir alandır (1). Canlılar için hayati önemi olan biyoinformatik alanı, pek çok karmaşık ve hesaplaması zaman alan problemler içermektedir. Karmaşık hesaplamaları geleneksel yöntemler ile çözmenin zor olması ve zaman alması sebebi ile biyoinformatik alanında graf sinir ağları yöntemlerine bir yönelim görülmektedir.

BİYOİNFORMATİK

Pek çok alanda olduğu gibi biyoloji alanında da bilgisayar teknolojileri kullanılmaya başlanmıştır. Biyoloji alanında artan verileri depolayabilmek ve istenildiğinde erişip analiz edebilmek amacıyla bilgisayar yazılımları ve algoritmaları kullanılmaktadır. Biyoinformatik, sadece biyoloji değil, canlıları ilgilendiren kimya, tıp gibi alanlarda da bilgisayar bilimlerinden faydalanılan alandır. Biyoinformatik, canlıları etkileyen alanlar ile birlikte bilgisayar bilimleri, matematik ve istatistik alanlarından da faydalanılan disiplinlerarası bir alandır. Biyoinformatik alanı protein fonksiyonu belirme ve gen fonksiyonu belirleme ve ilaç hedefi belirleme gibi pek çok alanı kapsamaktadır (1,2).

¹ Dr., Kahramanmaraş Sütçü İmam Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü fahriyegemci@ksu.edu.tr

SONUÇ

İlk olarak biyolojik dizilerin analizi için geliştirilmiş olan biyoinformatik alanı günümüzde, insanları etkileyen protein etkileşimleri, gen ifade çalışmaları, rasyonel yeni ilaç geliştirme, gibi pek çok alanı kapsamaktadır. Biyoinformatik alanı, canlılara ait karmaşık verilerin depolanması ve analizini gerçekleştirebilmek amacıyla, bilgisayar algoritmalarından faydalanılarak geliştirilmektedir.

Karmaşık canlı sistemini anlayabilmek için, canlının temel bileşenlerini anlayıp ardından bileşenler arasındaki bağlantıları çözmek gerekir. Bu amaçla bu çalışmada canlıların temel bileşenleri olan protein, DNA, RNA ve gen ile ilaç etkileşimleri analiz edilerek biyoinformatik alanı tanıtılmaya çalışılmıştır. Bu karmaşık canlı yapının bilgisayar ağları ile analiz etmenin iyi bir yol olduğu ve ağların da graflar ile temsili bu çalışmada açıklanmıştır. Graf teorisinin biyolojik ağların gizli özelliklerini keşfetmede kullanılabilecek başarılı bir yöntem olduğu gösterilmektedir.

KAYNAKLAR

1. G. Keklik, B. D. Özcan, Multidisipliner bir bilim dalı olarak biyoinformatiğe genel bir bakış, *Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 5(2), 1091-1082, 2022
2. A. C. Akın, B. Bürçe, B. Çevirici, B. Şahin, E. Şahin, Y. Şahin, Disiplinler arası bir bilim dalı: biyoinformatik, Erişim Tarihi: 20/12/2020 Web adresi: <http://tip.baskent.edu.tr/kw/upload/464/dosyalar/cg/sempozyum/ogrsmpzsnm16/16.P12.pdf>
3. J. Gao, T. Lyu, F. Xiong, J. Wang, W. Ke, Z. Li, MGNN: A multimodal graph neural network for predicting the survival of cancer patients, *43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Şian (Xi'an), Çin, 2020.
4. B. L. Hie, K. K. Yang, Adaptive machine learning for protein engineering, *Current Opinion In Structural Biology*, 72, 152-145, 2022.
5. P. J. Artymiuk, D. W. Rice, E. M. Mitchell, P. Willett. Structural resemblance between the families of bacterial signal-transduction proteins and of G proteins revealed by graph theoretical techniques, *Protein Engineering, Design and Selection*, 4(1), 43-39, 1990.
6. M. Habibi, C. Eslahchi, M. Sadeghi, H. Pezashk, The interpretation of protein structures based on graph theory and contact map, *Open Access Bioinformatics*, 2, 137-127, 2019.
7. N. Biggs, E. Lloyd, R. Wilson, Bölüm 1: Paths, *Graph Theory*, Oxford University Press.A, 1936-1736, 1986.
8. Simple English Wikipedia, the free encyclopedia, Seven Bridges of Königsberg Erişim tarihi: 10/12/2022, Web adresi: https://simple.wikipedia.org/wiki/Seven_Bridges_ofK%C3%B6nigsberg
9. H. C. Yi, Z. H. You, D. S. Huang, C. K. Kwoh, Graph representation learning in bioinformatics: trends, methods and applications, *Briefings in Bioinformatics*, 23(1), bbab340, 2022.
10. Z. Liu, J. Zhou, Introduction to graph neural networks, *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 14(2), 127-1, 2020.
11. Z. Ye, Y. J. Kumar, G. O. Sing, F. Song, J. Wang, A comprehensive survey of graph neural networks for knowledge graphs, *IEEE Access*, 10, 75741-75729, 2022.
12. M. Zhang, H. Luo, W. Song, H. Mei, C. Su, Spectral-Spatial offset graph convolutional networks for hyperspectral image classification, *Remote Sensing*, 13(21), 4342, 2021.
13. X. Zheng, Y. Liu, S. Pan, M. Zhang, D. Jin, P. S. Yu, Graph neural networks for graphs with heterophily: A survey, arXiv:2202.07082, Erişim Tarihi:15/12/2022, <https://arxiv.org/>

pdf/2202.07082.pdf.

14. E. Burgos, H. Ceva, L. Hernandez, RP, Perazzo, M, Devoto, D, Medan, Two classes of bipartite networks: nested biological and social systems, *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 78(4), 046113, 2008.
15. E. Khaledian, A. H. Gebremedhin, K. A. Brayton, S. L. Broshat, A network science approach for determining the ancestral phylum of bacteria, *2018 ACM International Conference On Bioinformatics, Computational Biology, and Health Informatics*, 403-398, 2018.
16. V. N. Ioannidis, A. G. Marques, G. B. Giannakis, Tensor graph convolutional networks for multi-relational and robust learning, *IEEE Transactions on Signal Processing*, 68, 6546-6535, 2020.
17. D. Davis, N. Chawla, Exploring and exploiting disease interactions from multi-relational gene and phenotype networks, *PLoS One*, 6, 2011.
18. C. Zhu, Z. Yang, X. Xia, N. Li, F. Zhong, L. Liu, Multimodal reasoning based on knowledge graph embedding for specific diseases, *Bioinformatics*, 38(8), 2245-2235, 2022.
19. M. Nickel, K. Murphy, V. Tresp, E. Gabrilovich, A review of relational machine learning for knowledge graphs, *Proceedings of the IEEE*, 104(1), 33-11, 2015.
20. M. Alshahrani, M. A. Thafar, M. Essack, Application and evaluation of knowledge graph embeddings in biomedical data, *PeerJ Computer Science*, 7, e341, 2021.
21. X. Zhu, Z. Li, X. Wang, X., Jiang, P. Sun, X. Wang, N. J. Yuan, Multi-Modal Knowledge Graph Construction and Application: A Survey, arXiv:2202.05786, Erişim Tarihi:15/12/2022, <https://arxiv.org/pdf/2202.05786.pdf>, 2022.
22. Q. Yuan, S. Chen, J. Rao, S. Zheng, H. Zhao, H, Y. Yang, AlphaFold2-aware protein–DNA binding site prediction using graph transformer, *Briefings in Bioinformatics*, 23(2), bbab564, 2022.
23. G. A. Pavlopoulos, P. I. Kontou, A. Pavlopoulou, C. Bouyioukos, E. Markou, P. G. Bagos. Bipartite graphs in systems biology and medicine: a survey of methods and applications, *GigaScience*, 7(4), giy014, 2018.
24. G. A. Pavlopoulos, M. Secrier, CN. Moschopoulos, TG. Soldatos, S. Kossida, J. Aerts, R. Schneider, P. G. Bagos, Using graph theory to analyze biological networks, *BioData Mining*, 4(10), 2011.
25. Z. Xie, J. Xu, Deep graph learning of inter-protein contacts, *Bioinformatics*, 38(4), 953-947, 2022.
26. Y. Wu, M. Gao, M. Zeng, J. Zhang, M. Li, BridgeDPI: a novel graph neural network for predicting drug–protein interactions, *Bioinformatics*, 38(9), 2578-2571, 2022.
27. X. Tang, J. Luo, C. Shen, Z. Lai, Multi-view multichannel attention graph convolutional network for miRNA–disease association prediction, *Brief in Bioinformatics*, 22(6), bbab174, 2021.
28. K. Han, E. E. Jeng, G. T. Hess, D. W. Morgens, A. Li, M. C. Bassik, Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions, *Nature Biotechnology*, 35(5), 474-463, 2017.
29. Y. Ma, S. Wang, C. C. Aggarwal, D. Yin, J. Tang, Multi-dimensional graph convolutional networks, *2019 SIAM International Conference on Data Mining*, 665-657, Alberta, Kanada, 2019.
30. A. Altabaa, D. Huang, C. Byles-Ho, H. Khatib, F. Sosa, T. Hu, geneDRAGNN: Gene Disease Prioritization using Graph Neural Networks, *19th IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*, Ottawa, Kanada 2022.
31. A. D. Perkins, M. A. Langston, M. A, Threshold selection in gene co-expression networks using spectral graph theory techniques, *BMC Bioinformatics*, 10(11), 11-1, 2009.
32. C. J. Wolfe, I. S. Kohane, A. J. Butte, Systematic survey reveals general applicability of “guilt-by-association” within gene coexpression networks, *BMC Bioinformatics*, 6(1), 2005.