

Systematic Literature Review : Penerapan dan Teknik Data Mining pada Sistem Pengambilan Keputusan Pemasaran di Perusahaan

Randy Isnaen Putra Zaeni^{*1}, Ramadhan Renaldy²

^{1,2} Program Studi Informatika, Universitas PGRI Semarang, Kota Semarang

*Email: randyisnaen7@gmail.com

Abstract

Data mining is a valuable resource in today's information age. These collected references discuss various aspects of data mining and how these techniques are used to support decision making in various fields. These articles show that data mining is utilized in the business world to improve marketing, project selection, and company development. Apart from that, data mining also plays a role in the health sector for disease research and clinical data analysis. In the field of education, data mining is used to improve the learning experience and the effectiveness of human resource decision support systems in the world of work. This literature study also discusses the challenges associated with data preparation and data quality, as well as the various methods used in the data mining process. By studying these references, we can gain a thorough understanding of how data mining turns data into actionable knowledge to produce better decisions. And the purpose of this Systematic Literature Review (SLR) is to find out which methods are most often used in data mining applications and techniques, what methods are most effective in data mining applications and techniques, and what datasets are most often used in data mining applications and techniques.

Keywords: Data mining, Marketing Decision, Predictive Analytics, Marketing.

Abstrak

Data mining adalah sumber daya yang berharga di era informasi saat ini. Referensi yang dikumpulkan ini membahas berbagai aspek penambangan data dan bagaimana teknik ini digunakan untuk mendukung pengambilan keputusan di berbagai bidang. Artikel-artikel ini menunjukkan bahwa penambangan data dimanfaatkan dalam dunia bisnis untuk meningkatkan pemasaran, pemilihan proyek, dan pengembangan perusahaan. Selain itu, penambangan data juga berperan dalam sektor kesehatan untuk penelitian penyakit dan analisa data klinis. Di bidang pendidikan, penambangan data digunakan untuk meningkatkan pengalaman belajar dan efektivitas sistem penunjang keputusan sumber daya manusia di dunia kerja. Studi literatur ini juga membahas tantangan yang terkait dengan persiapan data dan kualitas data, serta berbagai metode yang digunakan dalam proses penambangan data. Dengan mempelajari referensi ini, kita bisa memperoleh pemahaman yang menyeluruh tentang bagaimana data mining mengubah data menjadi pengetahuan yang dapat di tindak lanjuti untuk menghasilkan keputusan yang lebih baik.

Keywords: Data mining, Marketing Decision, Predictive Analytics, Marketing.

1. Pendahuluan

Di era digital ini, jumlah data yang dihasilkan setiap harinya mencapai proporsi yang mengagumkan di berbagai sektor industri. Data ini tidak hanya mencerminkan transaksi bisnis atau aktivitas sosial, tetapi juga mengandung potensi besar untuk mengungkapkan wawasan yang berharga yang dapat menginformasikan pengambilan keputusan yang lebih cerdas dan strategis [1][2]. Penambangan data, sebuah teknik kuat dalam ilmu data, memainkan peran

sentral dalam proses ini dengan memungkinkan ekstraksi pengetahuan dan pola yang tersembunyi dari kumpulan data yang besar dan kompleks [3][4].

Penelitian yang telah kami kumpulkan menggali berbagai aspek dari penambangan data dan aplikasinya di berbagai bidang, termasuk bisnis [5], kesehatan [6], pendidikan [7], dan keuangan [8][9]. Melalui referensi ini, kami menjelajahi bagaimana teknik penambangan data tidak hanya berfungsi sebagai alat untuk mengelola data [10], tetapi juga sebagai sarana untuk mendukung pengambilan keputusan yang berbasis bukti dan optimal di berbagai domain kehidupan [11][12]. Diskusi dalam referensi ini juga merangkum berbagai metode penambangan data yang digunakan secara luas, serta tantangan yang terkait dengan persiapan data yang tepat dan kualitas data yang diperlukan [13][14].

Dalam era persaingan bisnis yang ketat, pemahaman mendalam terhadap perilaku konsumen dan tren pasar menjadi krusial bagi keberhasilan pemasaran. Penambangan data tidak hanya memungkinkan identifikasi pola pembelian dan preferensi konsumen secara akurat [29], tetapi juga memfasilitasi pengembangan strategi pemasaran yang terarah dan personalisasi yang lebih efektif. Dengan memanfaatkan dataset yang mencakup media sosial, ulasan pelanggan, dan data transaksi, perusahaan dapat mengoptimalkan kampanye pemasaran mereka, mengurangi biaya akuisisi pelanggan, dan meningkatkan retensi dengan pendekatan yang lebih terfokus [30].

Beragam dataset digunakan dalam penelitian ini, mencakup data keuangan [15], operasional dan produksi [16], media sosial dan ulasan pelanggan [17], transformasi digital [18], pendidikan, serta pariwisata [19][20]. Pemahaman mendalam dari berbagai sumber ini diharapkan dapat memberikan gambaran yang komprehensif tentang peran krusial penambangan data dalam mengubah data mentah menjadi pengetahuan yang dapat diterapkan secara praktis [21][22].

Tujuan utama dari Systematic Literature Review (SLR) yang sedang kami lakukan adalah untuk mengidentifikasi metode yang paling sering digunakan dalam penerapan data mining [23][24], mengeksplorasi keefektifan relatif dari berbagai metode ini [25], serta mengidentifikasi dataset yang paling relevan dan sering digunakan dalam praktik penambangan data saat ini [26]. Dengan demikian, SLR ini diharapkan tidak hanya memberikan wawasan tentang tren dan praktik terkini dalam ilmu pengetahuan data, tetapi juga memberikan panduan praktis bagi para profesional dan peneliti untuk memilih pendekatan terbaik dalam menghadapi tantangan nyata dalam analisis data modern [27][28].

2. Metode

Metode yang digunakan penulis dalam melakukan penelitian ini adalah systematic literature review (SLR). Metode SLR digunakan untuk menganalisis informasi dari jurnal yang telah diteliti sebelumnya supaya mendapatkan tinjauan yang sistematis melalui berbagai tahap yang telah dipersiapkan.

Terdapat 3 proses dalam SRL yaitu Planning, Conducting, Reporting. Berikut ini adalah penjelasan mengenai masing-masing proses yang dilakukan dalam SRL:

1. Planning

Langkah pertama yang dilakukan dalam penulisan jurnal SLR yaitu pemilihan topik. Topik yang dipilih peneliti yaitu Penerapan dan teknik data mining pada Sistem Pengambilan Keputusan Pemasaran di Perusahaan. Penentuan topik ini digunakan penulis menentukan tujuan penelitian serta memilih Reserch Question(RQ). RQ tersebut akan menjadi acuan untuk penelitian dan pengumpulan referensi jurnal yang relevan dengan topik yang dipilih.

Setelah Menentukan RQ penulis mencari jurnal yang berkaitan dengan topik dengan kata kunci (keyword) yang dipakai adalah Data mining, Marketing Decision, Predictive Analytics, dengan tahun terbit 2019-2024. Peneliti menentukan 30 jurnal dari 3 sumber penerbit yaitu Google Scholar, ACM digital library dan Academia.edu dan dengan bantuan software Publish or Perish untuk memudahkan pencarian jurnal.

2. Conducting

Pada proses Conducting penulis melakukan analisis terhadap 30 jurnal dan menentukan metode, kriteria, implementasi serta tahun publikasi jurnal. Analisis tersebut dibuat di dalam Microsoft Excel. Berikut adalah tabel RQ beserta penjelasannya.

Tabel 2.1 Reserch Question

ID	Reserch Question	Explanation
RQ 1	Metode mana saja yang paling sering digunakan dalam penerapan data mining ?	Menganalisis Metode apa saja yang paling sering digunakan dalam penerapan dan Teknik data mining
RQ 2	metode dengan akurasi terbaik dalam penerapan data mining?	Menganalisis Metode apa saja yang memiliki akurasi terbaik dalam penerapan dan Teknik data mining
RQ 3	Dataset apa saja yang paling sering digunakan dalam penerapan data mining?	Menganalisis Dataset apa saja yang paling sering digunakan dalam penerapan dan Teknik data mining

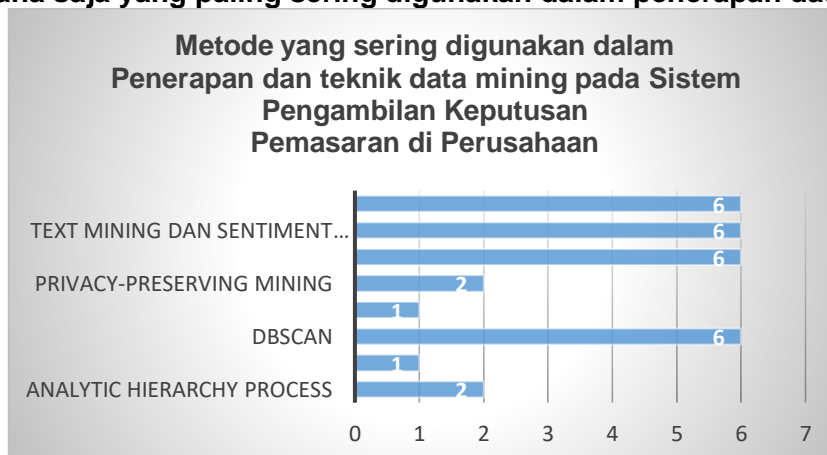
3. Reporting

Reporting adalah proses terakhir dalam pembuatan jurnal SLR. Pada proses ini peneliti mulai menulis jurnal yang diawali dengan penulisan abstrak, pendahuluan, metode, hasil dan pembahasan kesimpulan, dan referensi yang sesuai dengan topik Penerapan dan teknik data mining pada Sistem Pengambilan Keputusan Pemasaran di Perusahaan. Pada bagian abstrak memuat pengertian Penerapan dan teknik data mining pada Sistem Pengambilan Keputusan Pemasaran di Perusahaan dan gambaran singkat mengenai tujuan, metodologi dan hasil utama serta kesimpulan dari penelitian. Abstrak bertujuan mengetahui gambaran singkat mengenai topik yang peneliti.

Kemudian dilanjutkan dengan pendahuluan yang memuat tentang masalah yang dihadapi serta penyelesaian yang telah dilakukan oleh peneliti sebelumnya dan diakhiri dengan metode yang digunakan serta harapan yang diinginkan penulis kepada pembaca. Selanjutnya pembuatan metode yang memuat cara kerja jurnal itu dibuat. Lalu hasil dan pembahasan yang memuat tentang hasil review 30 jurnal yang dipilih, serta pembahasan hasil jurnal yang telah direview. Setelah itu dilanjutkan kesimpulan untuk menyimpulkan jurnal dari abstrak hingga hasil kesimpulan. Terakhir adalah referensi yang memuat semua referensi termasuk 30 jurnal yang direview. Penulisannya menggunakan perangkat lunak mandely agar lebih mudah dan sesuai dengan gaya IEEE.

3. Hasil dan Pembahasan

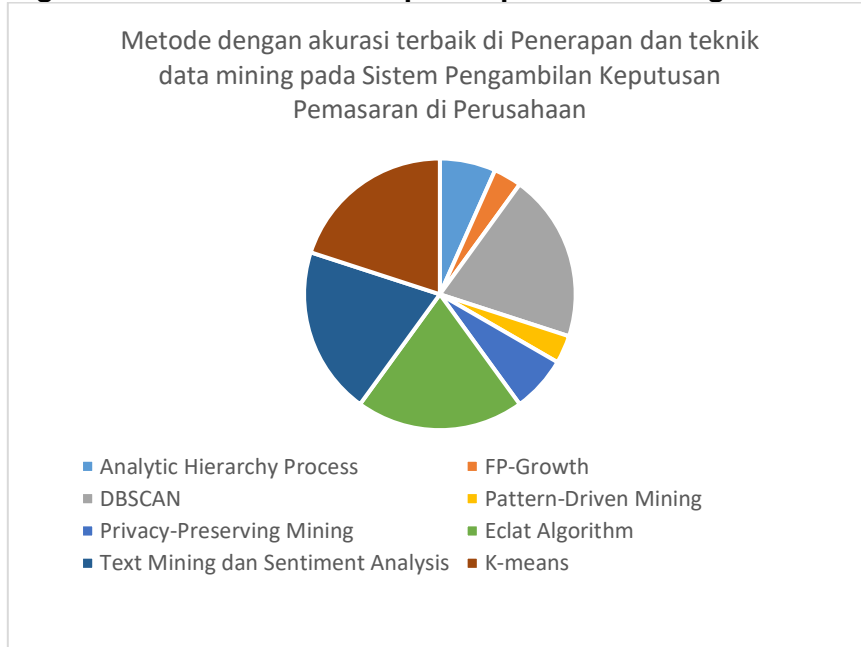
A. Metode mana saja yang paling sering digunakan dalam penerapan data mining



Gambar 3.1 Metode penerapan dan teknik mining

Pada gambar 3.1 dapat dilihat algoritma yang paling sering digunakan adalah K-MEANS, Eclat Algorithm, DBSCAN, dan Text Mining dan Sentiment Analysis dengan jumlah jurnal sebanyak 6 jurnal. Sedangkan diurutkan ke dua ada Analytic Hierarchy Process dan Privacy-preserving Mining masing masing sebanyak 2 jurnal. Urutan ketiga ada Analisi Big Data dan Pattern-Driven Mining dengan jumlah jurnal masing masing sebanyak 1 jurnal.

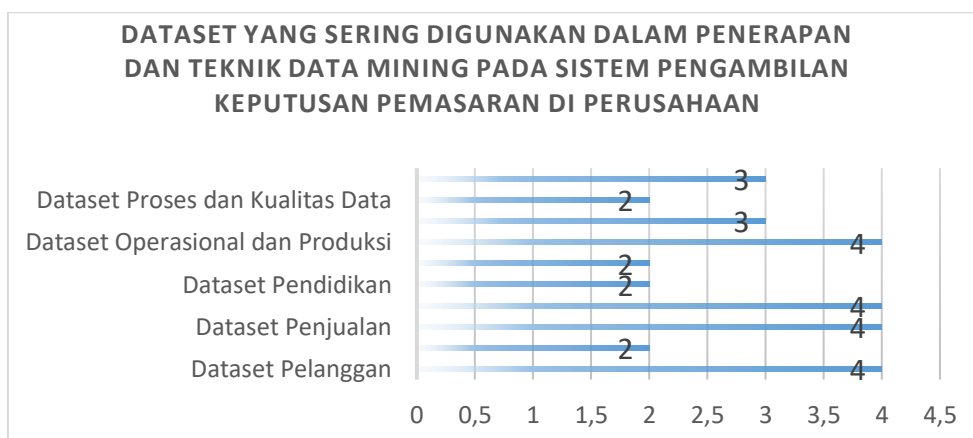
B. Metode dengan akurasi terbaik dalam penerapan data mining.



Gambar 3.2 Metode Yang Paling Efektif dalam penerapan dan teknik data mining

Melalui proses analisis dan pengumpulan data dari 30 jurnal dapat dijelaskan tentang penerapan dan Teknik data mining dengan akurasi terbaik untuk digunakan. Untuk penerapan dan Teknik data mining yang terbaik digunakan adalah adalah K-MEANS, Eclat Algorithm, DBSCAN, dan Text Mining dan Sentiment Analysis dengan persentase 20%. Dan ada juga beberapa metode yang kurang efektif sebagai contohnya adalah Metodologi Multi-criteria Decision Making, Privacy-preserving Mining, Analisi Big Data, dan Pattern-Driven Mining dengan persentase dibawah 20%.

C. Dataset paling sering digunakan



Gambar 3.3 Dataset Paling Banyak Digunakan

Melalui proses analisis dan pengumpulan data dari 30 jurnal diatas dapat disimpulkan bahwa dataset yang paling sering digunakan adalah dataset operasional produksi, dataset media sosial dan ulasan pelanggan, dataset penjualan, dan dataset pelanggan dengan masing masing berjumlah 4 jurnal dari 30 jurnal. adapun beberapa yang jarang digunakan seperti dataset pariwisata, pendidikan, dan sumber daya manusia.

4. Kesimpulan

Dari 30 jurnal yang telah di review didapatkan jawaban dari masing-masing RQ yang telah ditetapkan sebelumnya. Bisa kita simpulkan bahwa untuk RQ1, metode yang paling sering digunakan dalam aplikasi data mining untuk sistem pengambilan keputusan pemasaran di perusahaan adalah K-MEANS, Eclat Algorithm, DBSCAN, dan Text Mining dan Sentiment Analysis dengan jumlah jurnal sebanyak 6 jurnal. Sedangkan RQ2, metode dengan akurasi terbaik yang digunakan dalam aplikasi data mining untuk sistem pengambilan keputusan pemasaran di perusahaan adalah K-MEANS, Eclat Algorithm, DBSCAN, dan Text Mining dan Sentiment Analysis dengan jumlah jurnal sebanyak 6 jurnal. Dan untuk RQ3, Dataset yang paling sering digunakan adalah dataset pelanggan, dataset penjualan, dataset operasional dan produksi, serta dataset media sosial dan ulasan pelanggan.

5. Referensi

- [1]. Biagi, V., Bollati, M., & di Gravio, G. (2021). Decision Making and Project Selection: An Innovative MCDM Methodology for a Technology Company. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3447432.3447440>
- [2]. Cao, S. (2021). Opportunities and challenges of marketing in the context of big data. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3456389.3456390>
- [3]. Chen, Y., Fan, C., Li, Z., & Ren, L. (2020). Research on the Relationship between Precision Marketing and Company Development Ability. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3440054.3440062>
- [4]. Cwei, L. J., Sahoo, N., Srivastava, G., & Ding, W. (2021). Introduction to the Special Issue on Pattern-Driven Mining, Analytics, and Prediction for Decision Making, Part 1. In *ACM Transactions on Management Information Systems* (Vol. 13, Issue 1). <https://doi.org/10.1145/3486960>
- [5]. Desmet, C., & Cook, D. J. (2021). Recent Developments in Privacy-preserving Mining of Clinical Data. *ACM/IMS Transactions on Data Science*, 2(4). <https://doi.org/10.1145/3447774>
- [6]. Ding, B., Li, Y., Qiao, M., Sun, X., & Wang, P. (2022). Optimization Design of Electric Power Marketing System Based on Data Mining Technology. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3578339.3578351>
- [7]. Jian, L. (2022). Design of enterprise human resources decision support system based on data mining. *Soft Computing*, 26(20). <https://doi.org/10.1007/s00500-021-06659-4>
- [8]. Liu, S. (2021). The design of apriori-tids algorithm based on big data and its application in the information mining of college counselors' educational

- decision-making. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3482632.3483132>
- [9]. Liu, Y. (2022). Research on the Application of Big Data Technology in Enterprise Financial Decision Making. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3564665.3564675>
- [10]. Prathibha, S., Sarkar, S., Zynab, M., Harini, R., Shyam Kumar, M., Vibha, V., & Sathish, K. (2022). Synthesizing Data Analytics towards Intelligent Enterprises. *2022 International Conference on Advanced Computing Technologies and Applications, ICACTA 2022*. <https://doi.org/10.1109/ICACTA54488.2022.9753427>
- [11]. Sharma, M. P., Meena, U., & Sharma, G. K. (2022). Intelligent Data Analysis using Optimized Support Vector Machine Based Data Mining Approach for Tourism Industry. *ACM Transactions on Knowledge Discovery from Data*, 16(5). <https://doi.org/10.1145/3494566>
- [12]. Shirokova, S., Solovyov, L., Gnatenko, E., & Lohyeeta, N. (2020). Implementation of the digital transformation concept during decision-making process in a construction company: Digital transformation as a driver of strategic decision-making in a commercial organization. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3446434.3446464>
- [13]. ter Hofstede, A. H. M., Koschmider, A., Marrella, A., Andrews, R., Fischer, D. A., Sadeghianasl, S., Wynn, M. T., Comuzzi, M., de Weerd, J., Goel, K., Martin, N., & Soffer, P. (2023). Process-Data Quality: The True Frontier of Process Mining. *Journal of Data and Information Quality*, 15(3). <https://doi.org/10.1145/3613247>
- [14]. 伍思梦. (2022). Analysis of Opportunities and Challenges Faced by Marketing in the Context of Big Data. *Sustainable Development*, 12(06). <https://doi.org/10.12677/sd.2022.126205>
- [15]. Bou Rjeily, C., Badr, G., Hajjarm El Hassani, A., & Andres, E. (2019). Medical data mining for heart diseases and the future of sequential mining in medical field. In *Intelligent Systems Reference Library* (Vol. 149). https://doi.org/10.1007/978-3-319-94030-4_4
- [16]. Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology (Singapore)*, 12(4). <https://doi.org/10.1007/s41870-020-00427-7>
- [17]. Gupta, S. K., Khang, A., Somani, P., Dixit, C. K., & Pathak, A. (2023). Data Mining Processes and Decision-Making Models in the Personnel Management System. In *Designing Workforce Management Systems for Industry 4.0: Data-Centric and AI-Enabled Approaches*. <https://doi.org/10.1201/9781003357070-6>
- [18]. SuryaNarayana, G., Kolli, K., Ansari, M. D., & Gunjan, V. K. (2021). A Traditional Analysis for Efficient Data Mining with Integrated Association Mining into Regression Techniques. *Lecture Notes in Electrical Engineering*, 698. https://doi.org/10.1007/978-981-15-7961-5_127
- [19]. Tsui, K. L., Chen, V., Jiang, W., Yang, F., & Kan, C. (2023). Data Mining Methods and Applications. In *Springer Handbooks*. https://doi.org/10.1007/978-1-4471-7503-2_38
- [20]. Wu, W. T., Li, Y. J., Feng, A. Z., Li, L., Huang, T., Xu, A. D., & Lyu, J. (2021). Data mining in clinical big data: the frequently used databases, steps, and methodological models. In *Military Medical Research* (Vol. 8, Issue 1). <https://doi.org/10.1186/s40779-021-00338-z>

- [21]. Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2023). Data Mining and Machine Learning Retention Models in Higher Education. *Journal of College Student Retention: Research, Theory and Practice*, 25(1). <https://doi.org/10.1177/1521025120964920>
- [22]. Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining Big Data in Education: Affordances and Challenges. *Review of Research in Education*, 44(1). <https://doi.org/10.3102/0091732X20903304>
- [23]. Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1). <https://doi.org/10.3390/bdcc4010001>
- [24]. Jassim, M. A., & Abdulwahid, S. N. (2021). Data Mining preparation: Process, Techniques and Major Issues in Data Analysis. *IOP Conference Series: Materials Science and Engineering*, 1090(1). <https://doi.org/10.1088/1757-899x/1090/1/012053>
- [25]. Maus, V., Giljum, S., Gutschlhofer, J., da Silva, D. M., Probst, M., Gass, S. L. B., Luckeneder, S., Lieber, M., & McCallum, I. (2020). A global-scale data set of mining areas. *Scientific Data*, 7(1). <https://doi.org/10.1038/s41597-020-00624-w>
- [26]. Qi, C. chong. (2020). Big data management in the mining industry. In *International Journal of Minerals, Metallurgy and Materials* (Vol. 27, Issue 2). <https://doi.org/10.1007/s12613-019-1937-z>
- [27]. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3). <https://doi.org/10.1002/widm.1355>
- [28]. Shafiq, M., Tian, Z., Bashir, A. K., Jolfaei, A., & Yu, X. (2020). Data mining and machine learning methods for sustainable smart cities traffic classification: A survey. *Sustainable Cities and Society*, 60. <https://doi.org/10.1016/j.scs.2020.102177>
- [29]. Sun, J. hong, Sun, F., Yan, B., Li, J. yi, & Xin, D. li. (2020). Data mining and systematic pharmacology to reveal the mechanisms of traditional Chinese medicine in Mycoplasma pneumoniae pneumonia treatment. *Biomedicine and Pharmacotherapy*, 125. <https://doi.org/10.1016/j.biopha.2020.109900>
- [30]. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1). <https://doi.org/10.1109/TNNLS.2020.2978386>