

BÁO CÁO KẾT QUẢ TÌM KIẾM VÀ ĐÁNH GIÁ THÔNG TIN HỌC THUẬT

Ứng dụng của học máy trong phát hiện và phát triển thuốc

Nguyễn Hoàng Vũ

Table of contents

1	Giới thiệu	1
2	Quá trình tìm kiếm thông tin	1
2.1	Các nguồn thông tin	1
2.2	Tiêu chí đánh giá	2
3	Đánh giá độ tin cậy của các nguồn thông tin	2
	Danh mục tài liệu tham khảo	4

1 Giới thiệu

Trong những năm gần đây, học máy (Machine Learning - ML) đã trở thành một công cụ mang tính cách mạng trong ngành dược phẩm. Quy trình nghiên cứu và phát triển (R&D) một loại thuốc truyền thống thường kéo dài trên 10 năm với chi phí lên đến hàng tỷ đô la. Việc ứng dụng các thuật toán học máy giúp tối ưu hóa toàn bộ quy trình này, từ đó giảm thiểu chi phí, rút ngắn thời gian và gia tăng tỷ lệ thành công.

2 Quá trình tìm kiếm thông tin

2.1 Các nguồn thông tin

Loại nguồn	Nguồn cụ thể	Vai trò trong nghiên cứu
Cơ sở dữ liệu sinh học và hóa học	PubChem, ChEMBL, Protein Data Bank (PDB), UniProt	Cung cấp cấu trúc phân tử, hoạt tính sinh học, và sơ đồ 3D của protein đích.
Tạp chí khoa học chuyên ngành AI/Y sinh	Nature Biotechnology, Journal of Medicinal Chemistry, Bioinformatics, Briefings in Bioinformatics	Cập nhật các mô hình học máy mới nhất, thuật toán học sâu và kết quả thử nghiệm lâm sàng.
Sách chuyên khảo và giáo trình	Thư viện Springer, Elsevier, Wiley (các đầu sách về Cheminformatics và AI-driven Drug Discovery)	Cung cấp kiến thức nền tảng về tin sinh học, tin hóa học và lý thuyết thuật toán.
Báo cáo và nguồn mở từ các tập đoàn lớn	Website/Sách trắng của Insilico Medicine, AlphaFold (Google DeepMind), BenevolentAI	Thu thập các nghiên cứu ứng dụng thực tiễn, đột phá công nghệ và xu hướng công nghiệp mới nhất.

2.2 Tiêu chí đánh giá

Các nguồn thông tin được đánh giá dựa trên 5 tiêu chí sau để xếp hạng độ tin cậy từ 5 (Rất cao) đến 1 (Rất thấp).

Tiêu chí	Mô tả về đánh giá
Tác giả	Nghiên cứu được thực hiện bởi các chuyên gia liên ngành (Tin sinh học, Hóa tin học, Khoa học dữ liệu) có thâm niên.
Cơ quan Xuất bản / Hội nghị uy tín	Tài liệu xuất bản trên tạp chí có Impact Factor cao hoặc các hội nghị AI hàng đầu (NeurIPS, ICML, ISMB).
Phương pháp và Xác thực (Validation)	Mô hình AI phải được xác thực bằng thực nghiệm (In Vitro/In Vivo) hoặc sử dụng các tập dữ liệu thử nghiệm chuẩn (Benchmark datasets).
Trích dẫn và Chia sẻ mã nguồn	Bài báo có lượng trích dẫn cao; mã nguồn (Source Code) và bộ dữ liệu được công khai công khai (như trên GitHub).
Tính Cập nhật	Ưu tiên tài liệu mới nhất (2022-2026) do tốc độ phát triển và thay đổi của các mô hình AI/Học sâu diễn ra cực kỳ nhanh.

3 Đánh giá độ tin cậy của các nguồn thông tin

STT	Tài liệu tham khảo	Loại nguồn	Tiêu chí đánh giá chi tiết	Xếp hạng độ tin cậy
1	(Vamathevan et al. 2019)	Bài báo khoa học (Review)	Tác giả: Chuyên gia đầu ngành về Tin hóa học. Cơ quan: Tạp chí uy tín Elsevier. PPNC: Tổng quan toàn diện, được bình duyệt chặt chẽ. Trích dẫn: Rất cao >1.000 lượt. Cập nhật: Tương đối mới (2019).	5 (Rất cao)
2	(Schneider 2018)	Bài báo khoa học (Review)	Tác giả: Giáo sư danh tiếng về Thiết kế thuốc phân tử. Cơ quan: Thuộc hệ thống Nature danh giá. PPNC: Phân tích chuyên sâu về tính tự động hóa bằng AI. Trích dẫn: >2.500 lượt. Cập nhật: Tương đối mới (2018).	5 (Rất cao)
3	(Ekins et al. 2019)	Bài báo khoa học (Perspective)	Tác giả: Nhà khoa học y sinh nổi tiếng có thâm niên. Cơ quan: Nature Materials. PPNC: Hệ thống hóa toàn bộ lý thuyết từ cơ bản đến nâng cao. Cập nhật: Tương đối mới (2019).	5 (Rất cao)
4	(Jumper et al. 2021)	Bài báo khoa học (Method)	Tác giả: Nhóm nghiên cứu DeepMind. Cơ quan: Tạp chí Nature. PPNC: Thực nghiệm đột phá áp dụng Deep Learning giải mã protein (AlphaFold). Trích dẫn: >44.000 lượt. Cập nhật: Tương đối mới (2021).	5 (Rất cao)
5	(Stokes et al. 2020)	Bài báo khoa học (Clinical)	Tác giả: Nhóm nghiên cứu liên ngành từ MIT và Harvard. Cơ quan: Tạp chí danh tiếng Cell. PPNC: Ứng dụng mô hình AI phát hiện kháng sinh mới (Halicin) và đã xác thực in vivo. Trích dẫn: Rất cao. Cập nhật: Tương đối mới (2020).	5 (Rất cao)

STT	Tài liệu tham khảo	Loại nguồn	Tiêu chí đánh giá chi tiết	Xếp hạng độ tin cậy
6	(Liu et al. 2026)	Bài báo khoa học (Review)	Tác giả: Nhóm nghiên cứu máy tính tiềm năng. Cơ quan: Nature Signtrans. PPNC: Tổng hợp thuật toán mới nhất. Cập nhật: Rất mới (2026).	5 (Rất cao)
7	(He et al. 2023)	Bài báo khoa học (Phân tích kỹ thuật)	Tác giả: Đội ngũ chuyên gia từ các tập đoàn dược phẩm lớn. Cơ quan: SemCancer. PPNC: Phân tích sâu: multi-omics, AI . Cập nhật: Mới (2023)	4 (Cao)
8	(Khanh et al. 2025)	Bài báo khoa học (Review)	Tác giả: Các Tiến sĩ, Phó Giáo sư ngành Y Dược trong nước. Cơ quan: Tạp chí Y học Cộng đồng. PPNC: Cung cấp góc nhìn chiến lược phù hợp với bối cảnh Việt Nam. Cập nhật: Rất mới (2026).	4 (Cao)
9	(team 2025)	Nguồn mở có chọn lọc	Tác giả: Google Deepmind. Cơ quan: Google Deepmind. Đặc điểm: Thiếu bình duyệt học thuật. PPNC: Bài viết tổng hợp. Cập nhật: Rất mới (2025).	3 (Trung bình)
10	(VinBrain Improves Quality of Healthcare with AI-infused Solutions, n.d.)	Nguồn mở có chọn lọc	Tác giả: Đội ngũ chuyên gia công nghệ và dược liệu. Cơ quan: Website tập đoàn công nghệ lớn tại Việt Nam. PPNC: Phân tích giải pháp thương mại. Đặc điểm: Thiếu bình duyệt học thuật. Cập nhật: Mới (2024).	2 (Thấp)

Danh mục tài liệu tham khảo

Ekins, Sean, Ana C. Puhl, Kimberley M. Zorn, et al. 2019. “Exploiting Machine Learning for End-to-End Drug Discovery and Development.” *Nature Materials* 18 (5): 435–41. <https://doi.org/10.1038/s41563-019-0338-z>.

- He, Xiuqing, Xiaowei Liu, Fengli Zuo, Hubing Shi, and Jing Jing. 2023. “Artificial Intelligence-Based Multi-Omics Analysis Fuels Cancer Precision Medicine.” *Seminars in Cancer Biology* 88 (January): 187–200. <https://doi.org/10.1016/j.semcancer.2022.12.009>.
- Jumper, John, Richard Evans, Alexander Pritzel, et al. 2021. “Highly Accurate Protein Structure Prediction with AlphaFold.” *Nature* 596 (7873): 583–89. <https://doi.org/10.1038/s41586-021-03819-2>.
- Khanh, Van Cong, Phan Thi Thao Trang, Nguyen Bao Uyen, Tran Ba Kien, and Duong Thi Thuan. 2025. “APPLICATION OF ARTIFICIAL INTELLIGENCE IN PHARMACEUTICAL RESEARCH AND DEVELOPMENT.” *Vietnam Journal of Community Medicine* 66 (CĐ2-NCKH). <https://doi.org/10.52163/yhc.v66iCĐ2.2211>.
- Liu, Yuqing, Kun Zhu, Weijun Peng, Zhaoqian Liu, and Xiaoyuan Mao. 2026. “Multi-Omics and Artificial Intelligence for Precision Drug Discovery and Potential Clinical Applications.” *Signal Transduction and Targeted Therapy* 11 (1): 210. <https://doi.org/10.1038/s41392-026-02631-6>.
- Schneider, Gisbert. 2018. “Automating Drug Discovery.” *Nature Reviews Drug Discovery* 17 (2): 97–113. <https://doi.org/10.1038/nrd.2017.232>.
- Stokes, Jonathan M., Kevin Yang, Kyle Swanson, et al. 2020. “A Deep Learning Approach to Antibiotic Discovery.” *Cell* 180 (4): 688–702.e13. <https://doi.org/10.1016/j.cell.2020.01.021>.
- team, Pushmeet Kohli and Anna Koivuniemi on behalf of the AlphaFold, John Jumper. 2025. “AlphaFold: Five Years of Impact.” Science in *Google DeepMind*. <https://deepmind.google/blog/alphafold-five-years-of-impact/>.
- Vamathevan, Jessica, Dominic Clark, Paul Czodrowski, et al. 2019. “Applications of Machine Learning in Drug Discovery and Development.” *Nature Reviews Drug Discovery* 18 (6): 463–77. <https://doi.org/10.1038/s41573-019-0024-5>.
- VinBrain Improves Quality of Healthcare with AI-infused Solutions. n.d. <https://www.nvidia.com/en-us/news/vinbrain-improves-quality-of-healthcare-with-ai-infused-solutions/>.