

新聞稿

即時發放

## 香港大學科學家開發首個人工智能深度學習方法 預測聯繫疾病的金屬蛋白結合位點突變

2019 年 12 月 15 日



孫紅哲教授（右二）及其團隊

在過去幾年中，人工智能（AI）（以機器模仿人類行為的能力）已經成為藥物開發等高科技領域的關鍵參與者。人工智能工具可幫助科學家使用優化的計算方法來發現生物大數據背後的秘密。深層神經網絡這類的 AI 方法改善了生物和化學應用中的決策，如疾病相關蛋白的預測、新型生物標誌物的發現，以及小分子藥物前體的從頭設計。這些最先進的方法可幫助科學家更有效、更經濟地開發潛在的藥物。

由香港大學化學系孫紅哲教授領導的研究團隊，與美國亞利桑那州梅奧醫院的王俊文教授（前香港大學教授）合作，成功開發出強大的深度學習方法來預測與疾病相關的金屬結合位點突變。這是全球第一種用於預測金屬蛋白中與疾病相關的金屬相關位點突變的深度學習方法，為應對人類疾病提供了嶄新的平台。研究結果已於國際權威學術期刊《自然-機器智能》發表。

金屬離子在人類生命體系的（病理）生理學中起著結構或功能上的關鍵作用。鋅、鐵和銅等金屬對於所有生命都是必不可少的，必須嚴格控制其在細胞中的濃度，缺乏或過量均會導致人類嚴重疾病。科學家估計，人體蛋白質組中超過三分之一的蛋白質結合了各種生理金屬，以促進人體內不同的生理活動，例如誘發體內的酶產生各種催化作用等。

人類基因組中的突變與不同疾病密切相關，如果這些突變發生在 DNA 的編碼區域，則可能破壞蛋白質的金屬結合位點，從而引發嚴重的人類疾病。例如，鋅的結合位點出現突變，有機會引發乳房、肝臟、腎臟、免疫系統和前列腺的病變；鈣和鎂的結合位點突變分別與肌肉和免疫系統的病變有關；代謝病則多與鐵結合位點突變有關；錳和銅的結合位點突變則與心血管病變有關，後者更與神經系統的病變有關。所以，了解蛋白質金屬結合位點上與疾病相關的突變，將有助疾病的預防和新藥的開發。

研究團隊首先整合了來自不同數據庫的組學數據，以建立一個全面的數據集作培訓深度學習模型之用。他們使用一種新穎方法，基於能量的親和力網格圖從金屬結合位點提取空間特徵。這些空間特徵已與物理化學序列特徵結合以訓練模型。最終結果表明，使用空間特徵增強了預測性能，曲線下面積(AUC)為 0.90，準確度為 0.82，提升了預測聯繫疾病的金屬結合位點突變的技術。鑑於金屬學和金屬蛋白領域的先進技術和平台有限，今次研究開發的深度學習方法提供了一種強有力的方法來整合實驗數據與生物信息學分析。

孫教授說：「機器學習和人工智能在當前的生物和化學科學中發揮著重要作用。研究團隊使用整合組學的方法，包括金屬組學和金屬蛋白組學，來研究生物學和醫學領域中的金屬離子並且已經從體內、體外的實驗中獲得了大量有價值的數據。我們現在開發基於深度學習的人工智能方法將這些原始數據轉化為有價值的知識，從而發現疾病背後的秘密並與之抗爭。我相信這種新穎的深度學習方法可以用於我們實驗室正在進行的其他項目上。」

研究項目獲香港研究資助局（RGC）和香港大學葉志成范港喜基金的支持。

補充資料

文章連結: <https://www.nature.com/articles/s42256-019-0119-z>

相關報導連結: <https://www.nature.com/articles/s42256-019-0128-y>

Koohi-Moghadam M., Wang H.B., Wang Y.C., Yang, X.M., Li H.Y., Wang J.W.\*, **Sun H.Z.\*** (2019) "Predicting disease-associated mutation of metal binding sites of a protein using a deep learning approach" Nature Machine Intelligence, DOI: 10.1038/s42256-019-0119-z

### 孫紅哲教授及其團隊

孫紅哲教授是化學系講座教授和葉志成范港喜基金教授（生物無機化學）。他的研究重點是生物無機化學和無機化學生物學及藥物化學，尤其是金屬組學/金屬蛋白組學和多元組學的集成。孫教授發表了一系列有關了解金屬蛋白功能和克服抗生素抗藥性的論文和著作，並擁有數項專利。

傳媒查詢，請聯絡港大理學院助理傳訊總監陳詩迪（電話：3917-5286/ 6703-0212; 電郵：[cindycst@hku.hk](mailto:cindycst@hku.hk)）。

圖片下載：[www.scifac.hku.hk/press](http://www.scifac.hku.hk/press)

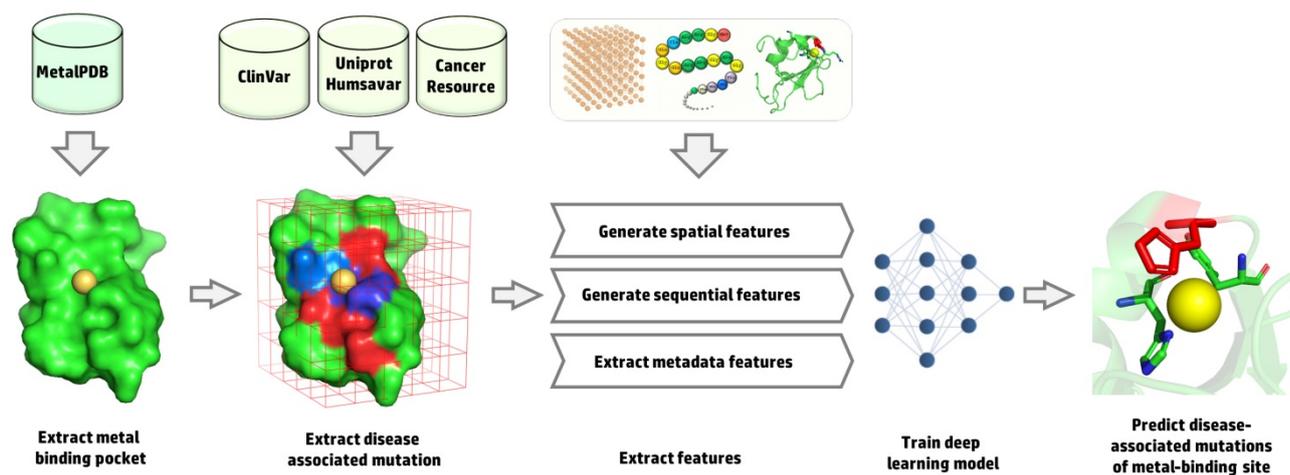


圖 1. 訓練深度學習模型的資料收集和特徵提取的工作流程。