

Ngân hàng dữ liệu: “mỏ vàng” kết nối AI và y sinh học

Trí tuệ nhân tạo (AI) đang ngày càng khẳng định vị thế quan trọng của mình trong nhiều lĩnh vực, đặc biệt là y sinh học. Với sự phát triển mạnh mẽ của AI, nhiều ứng dụng của công nghệ này đã mang lại những tiến bộ vượt bậc trong việc chăm sóc sức khỏe, điều trị bệnh tật và nghiên cứu khoa học. Một trong những yếu tố quan trọng giúp AI phát huy tối đa tiềm năng của mình chính là dữ liệu, đặc biệt là trong việc xây dựng các ngân hàng dữ liệu, nơi chứa đựng thông tin quý giá cho việc phát triển AI. Thông qua bài viết này, chúng ta sẽ tìm hiểu về tầm quan trọng của việc xây dựng ngân hàng dữ liệu như một “mỏ vàng” để khai thác và nghiên cứu AI, đồng thời khám phá những thách thức trong việc thu thập và xử lý dữ liệu, cũng như những cơ hội mà nó mang lại cho sự phát triển của y sinh học.



AI đã thực sự tạo nên những bước tiến ngoạn mục trong nhiều lĩnh vực, đặc biệt là trong y sinh học. Ứng dụng AI trong việc nghiên cứu và điều trị các bệnh lý đã mở ra một kỷ nguyên mới - kỷ nguyên của y học chính xác, hay còn gọi là y học cá thể hóa. Theo TS. Võ Sỹ Nam, Giám đốc Khoa học & Công nghệ CTCP GeneStory, trong Diễn đàn Trí tuệ nhân tạo Việt Nam 2024, AI đang giúp các nhà khoa học dự đoán cấu trúc ba chiều của protein, một công trình nghiên cứu đã giành giải Nobel Hóa học năm 2024. Cùng với đó, AI còn hỗ trợ trong việc phân loại bệnh nhân, tối ưu hóa phác đồ điều trị, từ đó nâng cao hiệu quả chữa bệnh và giảm thiểu rủi ro cho người bệnh.

Tuy nhiên, để AI có thể phát huy tối đa tác dụng trong y sinh học, một yếu tố then chốt không thể thiếu chính là dữ liệu. Dữ liệu chính là nền tảng để xây dựng và đào tạo các mô hình AI. Theo TS. Nam, một trong những vấn đề quan trọng nhất khi nghiên cứu AI chính là việc thu thập và gắn nhãn dữ liệu. Dữ liệu AI cần phải được làm sạch, gắn nhãn và xử lý một cách cẩn thận để đảm bảo tính chính xác và độ tin cậy khi đưa vào mô hình học máy. Nếu dữ liệu thu thập không đầy đủ, sai lệch hoặc không được xử lý đúng cách, kết quả mà AI mang lại sẽ thiếu chính xác và có thể gây hại thay vì giúp ích.

Việc xây dựng ngân hàng dữ liệu - đặc biệt là ngân hàng dữ liệu y sinh học - là một giải pháp hữu hiệu để giải quyết vấn đề này. Ngân hàng dữ liệu là nơi lưu trữ thông tin y tế, sinh học, gen, hình ảnh y khoa và các dữ liệu khác có liên quan, giúp các nhà nghiên cứu và ứng dụng AI có nguồn tài nguyên dồi dào để đào tạo các mô hình học máy. Nếu như những dữ liệu này được thu thập một cách có hệ thống, chất lượng và được gắn nhãn chính xác, nó sẽ trở thành "mỏ vàng" cho việc nghiên cứu và phát triển AI, đồng thời giúp Việt Nam có thể cạnh tranh với các quốc gia phát triển trong lĩnh vực này.

Việt Nam hiện nay đã có những bước tiến đáng kể trong việc ứng dụng AI vào y sinh học, như quản lý bệnh án điện tử, phân tích hình ảnh y khoa, dự đoán sự lây lan dịch bệnh và phát triển các loại vắc xin. Tuy nhiên, vẫn còn một số thách thức lớn cần được giải quyết. Đầu tiên, Việt Nam thiếu một khung pháp lý rõ ràng về việc sử dụng AI trong lĩnh vực y sinh học. Điều này dẫn đến việc khó có thể triển khai rộng rãi các ứng dụng AI một cách an toàn và hiệu quả. Thứ hai, vấn đề bảo mật dữ liệu và quyền riêng tư của bệnh nhân chưa được quan tâm đầy đủ, điều này có thể gây lo ngại cho người dân khi tham gia vào các nghiên cứu hoặc sử dụng dịch vụ y tế ứng dụng AI. Cuối cùng, thiếu nguồn nhân lực chất lượng cao trong lĩnh vực AI cũng là một thách thức lớn. Việc đào tạo và phát triển đội ngũ nhân lực có chuyên môn sẽ đóng vai trò quyết định trong việc thúc đẩy nghiên cứu AI tại Việt Nam.

Ngoài ra, việc tạo lập ngân hàng dữ liệu y sinh học sẽ giúp giải quyết một phần những vấn đề trên. Một ngân hàng dữ liệu lớn, được xây dựng một cách bài bản, không chỉ hỗ trợ nghiên cứu trong nước mà còn tạo cơ hội để Việt Nam hợp tác với các quốc gia khác trong việc phát triển và ứng dụng AI. Dữ liệu y tế ở Việt Nam, như các thông tin về bệnh tật, dịch tễ học, gen và sinh học, có thể là một nguồn tài nguyên vô giá, đặc biệt trong bối cảnh nhu cầu nghiên cứu AI ngày càng tăng cao.

Tạo lập và phát triển ngân hàng dữ liệu, đặc biệt là trong lĩnh vực y sinh học, không chỉ là một yếu tố quan trọng để thúc đẩy nghiên cứu AI, mà còn là chìa khóa để Việt Nam gia nhập và phát triển trong kỷ nguyên công nghệ 4.0. Mặc dù vẫn còn nhiều thách thức, nhưng với sự đầu tư đúng mức và sự phối hợp giữa các cơ quan, tổ chức, và cộng đồng nghiên cứu, Việt Nam hoàn toàn có thể tạo ra một "mỏ vàng" dữ liệu, từ đó phát triển các ứng dụng AI giúp cải thiện chất lượng cuộc sống và sức khỏe cộng đồng. Việc khai thác và nghiên cứu AI chính là con đường mở ra những cơ hội mới, đồng thời đẩy mạnh sự tiến bộ của xã hội trong thời đại số.

P.A.T (tổng hợp)

Nguồn: Cục Thông tin Khoa học và Công nghệ Quốc gia.