

Harry Yang (*chủ biên*)

**KHOA HỌC DỮ LIỆU,  
TRÍ TUỆ NHÂN TẠO  
VÀ HỌC MÁY TRONG  
PHÁT TRIỂN THUỐC**

Hoài Sơn dịch  
PGS TS Lê Văn Truyền *hiệu đính*



## **BAN CỐ VẤN**

PGS TS DS Lê Văn Truyền

TS BS Nguyễn Trọng Khoa

TS BS Lý Quốc Trung

ThS Lê Hồng Minh

TS BS Ông Thế Duệ

ThS DS Nguyễn Đức Thọ

TS DS Kiều Thị Tuyết Mai

TS DS Nguyễn Thị Hà

TS DS Vũ Thị Thu Hương

ThS DS Nguyễn Thị Phương Lan

## **ĐỐI TÁC ĐỒNG HÀNH**

Trung tâm Nghiên cứu và Đánh giá Kinh tế Y tế (HERAC), thành lập năm 2017, là tổ chức nghiên cứu và tư vấn khoa học độc lập, hoạt động trong lĩnh vực kinh tế y tế, đánh giá công nghệ y tế và chính sách y tế tại Việt Nam. Trung tâm đóng vai trò tiên phong trong việc thúc đẩy ứng dụng bằng chứng khoa học vào hoạch định chính sách và thực hành lâm sàng, góp phần xây dựng hệ thống y tế bền vững, công bằng và hiệu quả.



## **ĐƠN VỊ HỢP TÁC TRUYỀN THÔNG**



**TRẠM ĐỌC**

<https://www.facebook.com/tramdocvn>  
website: <http://tramdoc.vn/>

**HARDI+**

VIỆN ĐÁNH GIÁ CHẤT LƯỢNG VÀ PHÁT TRIỂN NGUỒN NHÂN LỰC Y TẾ

## Một số nhận xét về cuốn sách

Để một thuốc mới ra đời cần có sự tham gia của các nhà khoa học trên nhiều lĩnh vực. Bắt đầu từ các hiểu biết về nguyên nhân gây bệnh, cơ chế bệnh sinh, đích tác dụng ở mức độ phân tử: gene, protein, quá trình tìm kiếm các phân tử hóa học gắn được vào đích để phát huy tác dụng dược lý. Rất nhiều nghiên cứu trong phòng thí nghiệm và trên lâm sàng được tiến hành để có dữ liệu đầy đủ về an toàn và hiệu quả của phân tử thuốc. Quá trình này kéo dài từ 10–15 năm và tốn hàng tỷ đô-la với xác suất thành công cực kỳ thấp. Cuốn sách này là cẩm nang đầy đủ, cập nhật mô tả quá trình nghiên cứu-phát triển thuốc mới kết nối với nguồn dữ liệu lớn thông qua AI nhằm rút ngắn thời gian và tiết kiệm nguồn lực.

Cuốn sách vô cùng bổ ích cho các dược sĩ, bác sĩ và những người làm việc trong lĩnh vực nghiên cứu-phát triển thuốc.

PGS TS PHẠM THỊ VÂN ANH

*Giám đốc Trung tâm Dược lý – Dược lâm sàng*

*Trường Đại học Y – Hà Nội*

...Cuốn sách là tài liệu chuyên sâu, cập nhật, có tính định hướng cao về ứng dụng khoa học dữ liệu, trí tuệ nhân tạo (AI) và học máy (ML) trong nghiên cứu-phát triển thuốc, cả ở giai đoạn trước và sau khi cấp phép. Từ góc độ là một chuyên gia thẩm định hồ sơ lâm sàng và thành viên Hội đồng tư vấn cấp phép lưu hành thuốc (Bộ Y tế), tôi đặc biệt đánh giá cao các chương bàn về thử nghiệm lâm sàng hỗ trợ bởi AI, y học chính xác, bằng chứng đời thực (RWE) và cảnh giác dược dựa trên dữ liệu lớn – những yếu tố ngày càng được cơ quan quản lý thuốc các quốc gia quan tâm trong đánh giá lợi ích-nguy cơ và hiệu quả-chi phí của thuốc. Cuốn sách

cũng đề cập đến các thay đổi trong quy chế quản lý nhằm thích ứng với sự phát triển của công nghệ, bao gồm việc xây dựng hệ thống pháp lý có khả năng hướng dẫn, phân tích và diễn giải dữ liệu lớn bằng cách sử dụng các công nghệ hiện đại, bao gồm AI và ML.

Với vai trò là Trưởng đơn vị Dược lâm sàng tại bệnh viện, tôi đặc biệt quan tâm đến các nội dung liên quan đến cá thể hóa điều trị, học tăng cường trong y học chính xác, cảnh giác dược dựa trên dữ liệu lớn. Những ứng dụng này không chỉ giúp tối ưu hóa lựa chọn thuốc và liều dùng cho từng người bệnh mà còn nâng cao an toàn trong sử dụng thuốc. Cuốn sách cung cấp các ví dụ thực tiễn và mô hình ứng dụng AI, ML trong phân tích hồ sơ sức khỏe điện tử (EHR), giúp dược sĩ lâm sàng nâng cao vai trò trong tư vấn điều trị, giám sát an toàn thuốc và phối hợp đa ngành. Đây là tài liệu tham khảo giá trị cho đội ngũ dược sĩ lâm sàng trong bối cảnh bệnh viện đang chuyển đổi số và hướng tới y học chính xác, lấy người bệnh làm trung tâm.

Đáng chú ý, các tác giả đã đề cập đến nhu cầu cấp bách cần thay đổi quy chế quản lý nhằm thích ứng với công nghệ mới – một nội dung hết sức thiết thực đối với công tác quản lý thuốc ở Việt Nam trong bối cảnh chuyển đổi số sâu rộng ở tất cả các lĩnh vực. Với ngôn ngữ rõ ràng, ví dụ thực tiễn và cách tiếp cận liên ngành, cuốn sách góp phần thúc đẩy chuyển đổi số trong ngành dược, hướng tới một nền y-dược học chính xác, hiệu quả và nhân văn.

PGS TS PHẠM THỊ THÚY VÂN  
*Trưởng Khoa Dược lý – Dược lâm sàng  
 Trường Đại học Dược, Hà Nội*

Trong thời đại bùng nổ của cách mạng khoa học, công nghệ và đổi mới sáng tạo, chúng ta đang chứng kiến cuộc cách mạng chuyển đổi số trên tất cả các lĩnh vực ở quy mô toàn cầu. Với việc hình thành của thế giới số và sự phát triển mãnh liệt của trí tuệ nhân tạo, cần tiếp cận sớm tri thức trong các lĩnh vực này để khoa

học sức khỏe có thể phục vụ con người hiệu quả hơn. Trí tuệ nhân tạo có thể giúp tăng hiệu suất làm việc, nếu được kết hợp với sự thấu cảm, niềm say mê và sáng tạo mà chỉ con người mới có thì sẽ mang lại nhiều lợi ích to lớn, đặc biệt trong lĩnh vực khoa học sức khỏe lấy người bệnh làm trung tâm.

Cuốn sách giúp người đọc tiếp cận với khoa học dữ liệu và trí tuệ nhân tạo, đồng thời cũng là tài liệu tham khảo cho các nhà khoa học đang khai thác lĩnh vực này trong hoạt động nghiên cứu-phát triển và sử dụng thuốc an toàn, hợp lý và hiệu quả.

GS TS PHẠM THỊ MINH HUỆ  
*Trường Đại học Dược, Hà Nội*

# Mục lục

Lời giới thiệu.....	11
Danh mục từ viết tắt .....	14
Lời nói đầu .....	27
<b>01</b> Chuyển đổi ngành Dược với khoa học dữ liệu, AI và học máy.....	37
<b>02</b> Quan điểm quản lý về dữ liệu lớn, AI và học máy .....	63
<b>03</b> Xây dựng một tổ chức khoa học dữ liệu linh hoạt và có khả năng mở rộng.....	101
<b>04</b> AI và học máy trong khám phá thuốc.....	128
<b>05</b> Dự đoán hoạt tính hiệp đồng chống ung thư thông qua học máy và xử lý ngôn ngữ tự nhiên .....	173
<b>06</b> Thử nghiệm lâm sàng được hỗ trợ bởi AI.....	210
<b>07</b> Học máy cho y học chính xác .....	246
<b>08</b> Học tăng cường trong y học cá thể hóa .....	292
<b>09</b> Tận dụng học máy, xử lý ngôn ngữ tự nhiên và học sâu trong an toàn thuốc và cảnh giác dược .....	316
<b>10</b> Sản xuất và cung ứng dược phẩm sinh học thông minh.....	371
<b>11</b> Đổi mới Ban Y khoa trong kỷ nguyên dữ liệu lớn và phân tích .....	390
<b>12</b> Học sâu với hồ sơ sức khỏe điện tử.....	418
<b>13</b> Bằng chứng đời thực cho quyết định tiếp cận điều trị và thanh toán .....	448
Mục từ tra cứu.....	474



## Lời giới thiệu

Trong kỷ nguyên số hóa mạnh mẽ, ngành công nghiệp dược phẩm đang chứng kiến một cuộc cách mạng chưa từng có nhờ sự phát triển vượt bậc của khoa học dữ liệu (Data Science), sự bùng nổ của dữ liệu lớn (Big Data), trí tuệ nhân tạo (Artificial Intelligence – AI), học máy (Machine Learning – ML) và công nghệ thông tin nói chung.

Cuốn sách *Khoa học dữ liệu, Trí tuệ nhân tạo và Học máy trong phát triển thuốc* của Harry Yang ra đời như một lời đáp cho nhu cầu cấp thiết: làm thế nào để khai thác sức mạnh của công nghệ thông tin nhằm tăng tốc quá trình nghiên cứu và phát triển thuốc – vốn rất lâu dài, tốn kém – đồng thời hướng tới cá nhân hóa điều trị thay vì cách tiếp cận “một đơn thuốc cho tất cả mọi người”.

Trong kỷ nguyên số, hàng trăm triệu hồ sơ bệnh án, kết quả xét nghiệm, dữ liệu gen và dữ liệu về phản ứng có hại của thuốc được lưu trữ mỗi ngày trên khắp thế giới. Nhưng dữ liệu chỉ thực sự có giá trị khi được phân tích và chuyển hóa thành tri thức để phục vụ người bệnh. Cuốn sách này dẫn dắt người đọc qua hành trình biến dữ liệu thô thành những phát hiện khoa học có thể cứu sống con người. Tác giả cho thấy cách các thuật toán học máy (ML) có thể phát hiện mối liên hệ tiềm ẩn giữa các phân tử thuốc và bệnh tật, dự đoán hiệu quả của thuốc, thậm chí nhận diện tác dụng bất lợi trước khi thuốc được đưa ra thị trường – những công việc mà trước đây phải tốn rất nhiều thời gian và chi phí.

Không chỉ là công cụ, trí tuệ nhân tạo (AI) ngày nay đang trở thành đối tác chiến lược trong phòng thí nghiệm và bệnh phòng. Cuốn sách giới thiệu những ứng dụng thực tế của AI trong thiết kế thuốc mới, tối ưu hóa thử nghiệm lâm sàng và cá nhân hóa

điều trị dựa trên đặc điểm di truyền của từng bệnh nhân. Từ học sâu (Deep Learning – DL) đến mạng nơ-ron nhân tạo (Artificial Neural Networks – ANN), độc giả sẽ được khám phá những công nghệ tiên tiến đang định hình tương lai của y học nói chung và dược học nói riêng.

Điểm đặc biệt của cuốn sách nằm ở cách nó kết nối hai lĩnh vực tưởng chừng như cách biệt: công nghệ thông tin và sinh học phân tử. Với ngôn ngữ dễ hiểu, hình minh họa trực quan và ví dụ thực tiễn, sách không chỉ dành cho chuyên gia công nghệ thông tin hay các nhà khoa học y-dược, sinh học, mà còn phù hợp với bác sĩ, dược sĩ, nhà hoạch định chính sách, quản lý y tế, sinh viên y-dược và tất cả những ai quan tâm đến sự phát triển của y học hiện đại trong thời đại lấy người bệnh làm trung tâm, hướng tới một nền y học chính xác và nhân văn.

Đồng thời, cuốn sách cũng khơi gợi những suy ngẫm sâu sắc về đạo đức, quyền riêng tư và trách nhiệm xã hội khi ứng dụng AI trong chăm sóc sức khỏe con người. Trong một thế giới mà dữ liệu là tài nguyên quý giá, việc sử dụng công nghệ một cách minh bạch và nhân văn là điều cốt lõi, không thể thiếu.

Có thể nói, *Khoa học dữ liệu, Trí tuệ nhân tạo và Học máy trong phát triển thuốc* của Harry Yang mở ra cánh cửa đến tương lai, nơi công nghệ không thay thế con người mà nâng tầm khả năng của chúng ta trong việc chữa lành và bảo vệ sự sống. Đây là cuốn sách dành cho những ai tin rằng khoa học và lòng nhân ái có thể song hành.

Với tính chất vừa là tài liệu tổng quan toàn diện, vừa là kim chỉ nam thực tiễn, cuốn sách sẽ giúp độc giả – từ chuyên gia công nghệ đến nhà nghiên cứu y dược – nắm bắt cách thức khai thác sức mạnh của công nghệ thông tin để đẩy nhanh tiến trình khám phá, thử nghiệm và sớm đưa ra thị trường những loại thuốc mới an toàn, hiệu quả, phục vụ sứ mệnh bảo vệ sức khỏe nhân loại.

Do trong bản tiếng Anh, các tác giả đã sử dụng rất nhiều thuật ngữ viết tắt trong lĩnh vực y, dược và công nghệ thông tin, chúng tôi đã bổ sung Danh mục từ viết tắt để bạn đọc tiện theo dõi.

Xin trân trọng cảm ơn Thương hiệu sách MEDINSIGHTS và dịch giả Hoài Sơn đã chuyển ngữ và mang đến cho độc giả Việt Nam một tác phẩm quý giá, đặc biệt là cho đội ngũ dược sĩ nhân Ngày Dược sĩ Thế giới 25/9.

PGS TS LÊ VĂN TRUYỀN

*Chuyên gia cao cấp (Dược học)*

*Nguyên Thứ trưởng Bộ Y tế (1992–2002)*

*Chủ tịch Hội đồng Khoa học & Đào tạo – Trung tâm Nghiên cứu và  
Đánh giá Kinh tế Y tế (HERAC)*