

TRƯỜNG ĐẠI HỌC CẦN THƠ
KHOA SƯ PHẠM
BỘ MÔN SƯ PHẠM TOÁN HỌC



Luận văn tốt nghiệp

Đề tài:
BƯỚC ĐẦU ỨNG DỤNG TRÍ TUỆ
NHÂN TẠO VÀO DẠY HỌC TOÁN:
Một nghiên cứu thực nghiệm trong
Chương Tổ hợp – Xác suất lớp 11

Giảng viên hướng dẫn:

ThS. Bùi Anh Tuấn

Sinh viên thực hiện:

Nguyễn Hiếu Thanh

MSSV: B1700039

Lớp: Sư phạm Toán học K43

Cần Thơ, 2021

Lời cảm ơn

Đầu tiên, em xin gửi lời cảm ơn đến Ban Giám hiệu trường Đại học Cần Thơ đã tạo cho em môi trường tốt để em có thể học tập và tiếp thu được những kiến thức quý báu trong những năm qua.

Em xin gửi lời cảm ơn sâu sắc đến Thầy giáo, Thạc sĩ Bùi Anh Tuấn đã nhiệt tình hướng dẫn, định hướng, hỗ trợ em trong suốt quá trình thực hiện luận văn, giúp đỡ em tiếp cận với cách tư duy, giải quyết và trình bày một vấn đề cần nghiên cứu. Những điều này đã giúp em khắc phục được những hạn chế của bản thân và những khó khăn để hoàn thành luận văn thành công, đúng thời hạn.

Em cũng gửi lời cảm ơn chân thành tới các thầy cô trong trường, đặc biệt các thầy cô trong Khoa Sư phạm đã giảng dạy em trong suốt thời gian học tập tại trường. Với những kiến thức, bài học có được sẽ là hành trang giúp em tự tin hơn trong công việc, cuộc sống và những mục tiêu trong tương lai.

Mặc dù em đã cố gắng hoàn thành luận văn bằng tất cả sự nỗ lực và khả năng của mình, nhưng chắc chắn vẫn còn nhiều hạn chế và thiếu sót. Em mong nhận được sự cảm thông và góp ý quý giá từ các thầy cô và các bạn.

Cần Thơ, tháng 04 năm 2021

Sinh viên

Nguyễn Hiếu Thanh

Mục lục

Mở đầu	1
1 Lý do chọn đề tài	1
2 Mục tiêu nghiên cứu	1
3 Nhiệm vụ nghiên cứu	1
4 Đối tượng nghiên cứu	2
5 Phạm vi nghiên cứu	2
6 Phương pháp nghiên cứu	2
1 Cơ sở lý luận	3
1.1 Trí tuệ nhân tạo	3
1.1.1 Định nghĩa	3
1.1.2 Thực trạng nghiên cứu lĩnh vực AI	3
1.1.3 Các xu hướng phát triển AI	4
1.1.4 Trí tuệ nhân tạo trong giáo dục	7
1.2 Máy trò chuyện (Chatbot)	8
1.2.1 Chatbot là gì?	8
1.2.2 Nền tảng Chatfuel	9
1.3 Lý thuyết Ứng đáp Câu hỏi	9
1.3.1 Các mô hình đường cong đặc trưng của câu hỏi nhị phân	9
1.3.2 Quy trình ước lượng các tham số của câu hỏi trắc nghiệm	13
1.3.3 Điểm thực – đường cong đặc trưng của đề trắc nghiệm	15
1.3.4 Ước lượng năng lực của thí sinh	17
2 Xây dựng API xử lý	19
2.1 Tổng hợp ngân hàng câu hỏi chương Xác suất – Thống kê	19
2.2 Xây dựng API tạo đề kiểm tra	19
2.3 Xây dựng API xử lý dữ liệu	19
3 Thiết kế AI Chatbot	20
3.1 Nền tảng Facebook Chatbot và Chatfuel	20
3.2 Quá trình thiết kế Chatbot	20
4 Thực nghiệm Sư phạm	21

Danh sách hình vẽ

1.1	Đường cong DTCH một tham số (mô hình Rasch)	10
1.2	Các đường cong DTCH 2 tham số với các giá trị a khác nhau ($b = 0$) . . .	11
1.3	Các đường cong DTCH 3 tham số với $a = 2$, $c = 0.1$ và 0.2	12
1.4	Minh họa các tỉ lệ trả lời đúng ứng với mỗi nhóm năng lực	14
1.5	Đường cong DTCH 2 tham số trùng khớp tốt nhất với số liệu	14
1.6	5 đường cong DTCH theo mô hình 3 tham số	16
1.7	Đường cong đặc trưng của ĐTN gồm 5 CH và 5 đường cong DTCH tương ứng	17

Danh sách bảng

Danh mục từ viết tắt

AI	Trí tuệ nhân tạo (Artificial Intelligence)
CH	Câu hỏi (item)
CTT	Lý thuyết Trắc nghiệm cổ điển (Classical Test Theory)
ĐTCH	Đặc trưng câu hỏi
DTN	Đề trắc nghiệm
ICC	Đường cong đặc trưng câu hỏi (Item Characteristic Curve)
ICF	Hàm đặc trưng câu hỏi (Item Characteristic Function)
IRT	Lý thuyết Ứng đáp Câu hỏi (Item Response Theory)
TS	Thí sinh (person)

Mở đầu

1 Lý do chọn đề tài

Cuộc cách mạng công nghiệp lần thứ tư hay còn gọi là cuộc cách mạng số diễn ra từ đầu thế kỷ XXI. Đặc trưng của cuộc cách mạng công nghiệp lần này là sẽ ngày càng phổ biến trí thông minh nhân tạo và máy móc tự động hóa, đem lại sự kết hợp giữa hệ thống ảo và thực tế. Cuộc cách mạng này tác động mạnh mẽ đến nhiều lĩnh vực, nhiều khía cạnh trong đời sống xã hội, trong đó đặc biệt không thể thiếu một nguồn nhân lực chất lượng cao; mà nguồn nhân lực lại là đối tượng trực tiếp của giáo dục – đào tạo.

Tuy nhiên, lĩnh vực AI ở Việt Nam chỉ mới được nghiên cứu và ứng dụng trên một số lĩnh vực như thông tin – truyền thông, y tế, du lịch... với một tỉ trọng rất nhỏ, thậm chí còn chưa được các doanh nghiệp quan tâm. Trong đó, việc đưa AI vào lĩnh vực giáo dục vẫn là một bài toán khó, thể hiện qua số lượng hạn chế các công trình nghiên cứu về AI trong giáo dục – đa số là nghiên cứu lý luận – và chưa có nhiều sản phẩm dành cho giáo dục.

Bên cạnh đó, các nền tảng mạng xã hội ngày càng phát triển và phổ biến, đặc biệt là với học sinh phổ thông. Đây là một môi trường tốt để hỗ trợ quá trình tự học của học sinh, cụ thể là cung cấp tài liệu bổ sung, bài tập về nhà... Song vẫn chưa được khai thác triệt để và đúng mức.

Trong chương trình Toán phổ thông, chương Xác suất, thống kê (Toán 11) có nhiều ứng dụng quan trọng và có nhiều mảng kiến thức mang tính hàn lâm, khó nắm bắt. Trong khi năng lực tiếp cận Toán học ở các học sinh thường không đồng đều nhau, đòi hỏi phương pháp tiếp cận khác nhau ở từng em. Tuy nhiên, trong các lớp học thực tế, một giáo viên thường phải quản lý 30 – 40 học sinh, tạo ra sự bất khả thi trong việc nắm bắt kịp thời mức độ tiếp nhận kiến thức của từng em.

Xuất phát từ những lý do trên, tôi chọn đề tài nghiên cứu ***"Bước đầu ứng dụng trí tuệ nhân tạo vào dạy học Toán: Một nghiên cứu thực nghiệm trong Chương Tổ hợp – Xác suất lớp 11"***.

2 Mục tiêu nghiên cứu

Mục tiêu của luận văn này là xây dựng một *máy trò chuyện* (chatbot) trên nền tảng mạng xã hội Facebook, trong đó cung cấp các câu hỏi trắc nghiệm, được tự động phân bổ theo năng lực học tập của học sinh, sử dụng vào phần bài tập về nhà.

3 Nhiệm vụ nghiên cứu

Luận văn thực hiện những nhiệm vụ sau:

- 3.1. Tìm hiểu vai trò và các ứng dụng của AI trong dạy học Toán học.
- 3.2. Tìm hiểu lý thuyết đánh giá và đo lường trong giáo dục.
- 3.3. Xây dựng thuật toán đánh giá và giao thức (API) xử lý thông tin.
- 3.4. Thiết kế một AI Chatbot trên nền tảng Facebook.
- 3.5. Vận dụng AI Chatbot vào đánh giá năng lực học sinh.
- 3.6. Thực nghiệm Sư phạm để đánh giá tính khả thi và xác định ưu nhược điểm khi sử dụng AI Chatbot trong đánh giá.

4 Đối tượng nghiên cứu

- 4.1. Ứng dụng của AI trong giáo dục.
- 4.2. Lý thuyết Ứng đáp Câu hỏi.

5 Phạm vi nghiên cứu

- 5.1. **Về phương pháp:** giới hạn sử dụng nền tảng Chatbot của mạng xã hội Facebook, ngôn ngữ lập trình PHP, một số thuật toán ML và lý thuyết ứng đáp câu hỏi IRT.
- 5.2. **Về chuyên môn:** giới hạn trong chương trình Toán 11 cơ bản chương Xác suất, thống kê.

6 Phương pháp nghiên cứu

6.1. Phương pháp nghiên cứu lý luận

Nghiên cứu các tài liệu về triết học, tâm lý học, giáo dục học lý luận dạy học, các phương pháp và ứng dụng công nghệ trong giáo dục nói chung và trí tuệ nhân tạo nói riêng.

6.2. Phương pháp thực nghiệm

Từ các nghiên cứu lý luận, sử dụng các công cụ để thiết kế Chatbot trên nền tảng Facebook.

6.3. Phương pháp điều tra, quan sát

Tiến hành dạy thực nghiệm và thu thập thông tin từ phiếu khảo sát về mức độ hứng thú của học sinh qua bài học.

6.4. Phương pháp thống kê Toán học

Phân tích định tính, định lượng, từ đó rút ra kết luận về tính khả thi cũng như ưu/nhược điểm của nền tảng Chatbot.

Chương 1

Cơ sở lý luận

1.1 Trí tuệ nhân tạo

Ngày nay, Trí tuệ nhân tạo (Artificial Intelligence – AI) đang góp phần thay đổi sâu sắc nhiều khía cạnh của cuộc sống, dần trở thành một yếu tố quan trọng trong hoạt động muôn màu muôn vẻ của nhân loại. Nhiều bức tranh về tương lai xán lạn do AI mang tới cho loài người đã được khắc họa.

1.1.1 Định nghĩa

Khái niệm về AI có thể được nhìn nhận theo nhiều cách khác nhau, chưa có định nghĩa nào được thừa nhận chung. Trên thế giới hiện có nhiều định nghĩa về trí tuệ nhân tạo (Nguyễn et al., 2018), cụ thể:

- AI là tự động hoá các hoạt động phù hợp với suy nghĩ con người, chẳng hạn các hoạt động ra quyết định, giải bài toán... (Bellman, 1978).
- AI là khoa học nghiên cứu xem làm thế nào để máy tính có thể thực hiện những công việc mà hiện con người còn làm tốt hơn máy tính (Rich and Knight, 1991).
- AI là lĩnh vực nghiên cứu các tính toán để máy có thể nhận thức, lập luận và tác động (Salin and Winston, 1992).

Mỗi khái niệm, định nghĩa đều có điểm đúng riêng, nhưng để đơn giản ta có thể hiểu AI là một ngành khoa học máy tính. Nó xây dựng trên một nền tảng lý thuyết vững chắc và có thể ứng dụng trong việc tự động hóa các hành vi thông minh của máy tính; giúp máy tính có được những trí tuệ của con người như: biết suy nghĩ và lập luận để giải quyết vấn đề, biết giao tiếp do hiểu ngôn ngữ, tiếng nói, biết học và tự thích nghi.

1.1.2 Thực trạng nghiên cứu lĩnh vực AI

Tính từ khi khởi đầu, nghiên cứu AI đã trải qua ba đợt sóng công nghệ. Làn sóng đầu tiên tập trung vào kiến thức thủ công, phát triển mạnh mẽ vào những năm 1980 trên các hệ chuyên gia dựa trên quy tắc trong các lĩnh vực được xác định rõ ràng, trong đó kiến thức được thu thập từ một người chuyên gia, được thể hiện trong quy tắc *nếu – thì* (if – then), và sau đó thực hiện trong phần cứng. Các hệ thống lập luận như vậy đã được áp dụng thành công các vấn đề hẹp, nhưng nó không có khả năng học hoặc đối phó với sự không chắc chắn. Tuy nhiên, chúng vẫn dẫn đến các giải pháp quan trọng, và các kỹ thuật phát triển vẫn được sử dụng hiện nay.

Làn sóng nghiên cứu AI thứ hai từ những năm 2000 đến nay được đặc trưng bởi sự phát triển của *máy học* (machine learning). Sự sẵn có một khối lượng lớn dữ liệu số, khả năng tính toán song song lớn tương đối rẻ, các kỹ thuật học cải tiến đã mang lại những tiến bộ đáng kể trong AI khi áp dụng cho các nhiệm vụ như nhận dạng hình ảnh và chữ viết, hiểu ngôn từ, và dịch thuật ngôn ngữ của người. Thành quả của những tiến bộ này có mặt ở khắp nơi: điện thoại thông minh thực hiện nhận dạng giọng nói, máy ATM thực hiện nhận dạng chữ viết tay, ứng dụng email lọc thư rác, và các dịch vụ trực tuyến miễn phí thực hiện dịch máy. Chìa khóa cho một số những thành công này là sự phát triển của *học sâu* (deep learning).

Các hệ thống AI giờ đây thường xuyên làm tốt hơn con người trong các nhiệm vụ chuyên môn. Các cột mốc quan trọng khi AI đầu tiên vượt qua năng lực của con người bao gồm: cờ vua (1997), giải câu đố (2011), trò chơi Atari (2013), nhận dạng hình ảnh (2015), nhận dạng giọng nói (2015), và Go (2016). Những thành tựu như vậy đã được thúc đẩy bởi một nền tảng mạnh mẽ của nghiên cứu cơ bản. Những nghiên cứu này đang mở rộng và có khả năng thúc đẩy tiến bộ trong tương lai.

Lĩnh vực AI hiện đang trong giai đoạn khởi đầu của làn sóng thứ ba, tập trung vào các công nghệ AI phổ quát và giải thích. Các mục tiêu của các phương pháp này là nâng cao mô hình học với sự giải thích và sửa giao diện, để làm rõ các căn cứ và độ tin cậy của kết quả đầu ra, để hoạt động với mức độ minh bạch cao, và để vượt qua AI phạm vi hẹp tới khả năng có thể khái quát các phạm vi nhiệm vụ rộng hơn. Nếu thành công, các kỹ sư có thể tạo ra các hệ thống xây dựng mô hình giải thích cho các lớp của hiện tượng thế giới thực, tham gia giao tiếp tự nhiên với người, học và suy luận những nhiệm vụ và tình huống mới gặp, và giải quyết các vấn đề mới bằng cách khái quát kinh nghiệm quá khứ. Các mô hình giải thích cho các hệ thống AI này có thể được xây dựng tự động thông qua các phương pháp tiên tiến. Những mô hình này có thể cho phép học tập nhanh chóng trong hệ thống AI. Chúng có thể cung cấp "ý nghĩa" hoặc "sự hiểu biết" cho hệ thống AI, sau đó có thể cho phép các hệ thống AI để đạt được những khả năng phổ quát hơn.

1.1.3 Các xu hướng phát triển AI

Cho đến thời điểm chuyển giao thiên niên kỷ, sự lôi cuốn của AI chủ yếu ở hứa hẹn cung cấp của nó, nhưng trong mười lăm năm qua, nhiều lời hứa đó đã được thực hiện. Các công nghệ AI đã thâm nhập vào cuộc sống của chúng ta. Khi chúng trở thành một lực lượng trung tâm trong xã hội, lĩnh vực này đang chuyển từ những hệ thống chỉ đơn giản là thông minh sang chế tạo các hệ thống có nhận thức như con người và đáng tin cậy.

Một số yếu tố đã thúc đẩy cuộc cách mạng AI. Quan trọng nhất trong số đó là sự trưởng thành của machine learning, được hỗ trợ một phần bởi nguồn tài nguyên điện toán đám mây và thu thập dữ liệu rộng khắp dựa trên web. Máy học đã đạt tiến bộ đáng kể bằng deep learning, một dạng đào tạo các mạng lưới thần kinh nhân tạo thích nghi sử dụng phương pháp gọi là lan truyền ngược. Bước nhảy vọt này trong việc thực hiện các thuật toán xử lý thông tin đã được hỗ trợ bởi các tiến bộ đáng kể trong công nghệ phần cứng cho các hoạt động cơ bản như cảm biến, nhận thức, và nhận dạng đối tượng. Các nền tảng và thị trường mới cho các sản phẩm nhờ vào dữ liệu, và các khuyến khích kinh tế để tìm ra các sản phẩm và thị trường mới, cũng góp phần cho sự ra đời của công nghệ

dựa vào AI.

Tất cả những xu hướng này thúc đẩy các lĩnh vực nghiên cứu "hot" được mô tả dưới đây.

Học máy quy mô lớn

Nhiều vấn đề cơ bản trong *máy học* đã được hiểu rõ. Trọng tâm chính của những nỗ lực hiện nay là mở rộng quy mô các thuật toán hiện có để làm việc với các tập dữ liệu rất lớn.

Học sâu

Khả năng để đào tạo thành công các mạng lưới thần kinh xoắn đã mang lại lợi ích nhiều nhất cho lĩnh vực thị giác máy tính, với các ứng dụng như nhận dạng đối tượng, ghi nhận video, nhận dạng hoạt động, và một số biến thể của nó.

Người máy

Kỹ thuật điều hướng *người máy* (robot), ít nhất là trong môi trường tĩnh, phần lớn đã được giải quyết. Những nỗ lực hiện tại tìm cách làm thế nào để đào tạo một robot tương tác với thế giới xung quanh theo các cách khái quát và dự đoán được. Một yêu cầu tự nhiên phát sinh trong môi trường tương tác là sự thao tác, một chủ đề quan tâm khác hiện nay. Cuộc cách mạng học sâu chỉ mới bắt đầu ảnh hưởng đến robot, chủ yếu là rất khó để có các bộ dữ liệu lớn có nhãn để thúc đẩy các lĩnh vực dựa trên học tập khác của AI.

Thị giác máy tính

Thị giác máy tính (Computer vision) hiện nay là hình thức nổi bật nhất của nhận thức máy. Nó là một phạm vi nhỏ của AI biến đổi nhiều nhất bởi sự xuất hiện của deep learning. Chỉ cách đây vài năm, các máy vectơ hỗ trợ là phương pháp được lựa chọn cho hầu hết các nhiệm vụ phân loại hình ảnh. Nhưng sự hợp lưu của máy tính quy mô lớn, đặc biệt là trên các *bộ xử lý đồ họa* (Graphics Processing Unit – GPU), sự sẵn có các tập dữ liệu lớn, đặc biệt là thông qua internet, và sàng lọc của các thuật toán mạng thần kinh đã dẫn đến những cải tiến đáng kể trong hiệu suất trên các nhiệm vụ chuẩn. Lần đầu tiên, các máy tính có thể thực hiện một số nhiệm vụ phân loại hình ảnh (hạn hẹp) tốt hơn so với con người. Nhiều nghiên cứu hiện nay đang tập trung vào tự động chú thích ảnh và video.

Xử lý ngôn ngữ tự nhiên

Thường kết hợp với nhận dạng giọng nói tự động, *Xử lý Ngôn ngữ tự nhiên* (natural language processing) là một khu vực rất tích cực khác về nhận thức máy. Nó nhanh chóng trở thành hàng hóa cho các ngôn ngữ chủ đạo với các tập dữ liệu lớn.

Các hệ thống hợp tác

Nghiên cứu về các hệ thống hợp tác tìm kiếm các mô hình và các thuật toán để giúp phát triển các hệ thống tự trị có thể hợp tác làm việc với các hệ thống khác và với con người. Nghiên cứu này dựa trên việc phát triển các mô hình hợp tác chính thức, và nghiên cứu các khả năng cần thiết cho hệ thống trở thành đối tác hiệu quả. Sự quan tâm ngày càng tăng đối với các ứng dụng có thể sử dụng các thể mạnh bổ sung của con người và máy móc – cho con người giúp hệ thống AI khắc phục những hạn chế của chúng, và cho các phần tử để tăng cường các khả năng và hoạt động của con người.

Tạo nguồn từ đám đông (crowdsourcing) và tính toán của con người

Do khả năng của con người vượt trội so với phương pháp tự động trong hoàn thành nhiều nhiệm vụ, nghiên cứu về tạo nguồn từ đám đông và tính toán của con người tìm kiếm các phương pháp để tăng cường các hệ thống máy tính bằng cách sử dụng trí tuệ của con người để giải quyết vấn đề mà một mình máy tính không thể giải quyết nổi. Được giới thiệu chỉ khoảng mười lăm năm trước, nghiên cứu này hiện nay có sự hiện diện vững chắc trong AI.

Kết quả trong lĩnh vực này đã hỗ trợ cho các tiến bộ trong các lĩnh vực nhánh khác của AI, bao gồm cả thị giác máy tính và xử lý ngôn ngữ tự nhiên, bằng cách cho phép một số lượng lớn dữ liệu huấn luyện được dán nhãn và các dữ liệu tương tác của con người được thu thập trong một khoảng thời gian ngắn.

Internet vạn vật (Internet of Things)

Đây là lĩnh vực nghiên cứu đang phát triển được tập trung vào ý tưởng rằng một loạt các thiết bị có thể được kết nối với nhau để thu thập và chia sẻ thông tin cảm biến của chúng. Các thiết bị này có thể bao gồm các đồ dùng, xe cộ, nhà cửa, máy ảnh, và những thứ khác. Trong khi vấn đề ở đây là công nghệ và mạng không dây để kết nối các thiết bị, AI có thể xử lý và sử dụng một lượng lớn dữ liệu thu được cho các mục đích thông minh và hữu ích.

Tính toán phỏng theo nơon thần kinh

Máy tính truyền thống thực hiện mô hình tính toán von Neumann, tách các mô-đun nhập/xuất, hướng dẫn – xử lý và bộ nhớ. Với sự thành công của các mạng lưới thần kinh sâu đối với một phạm vi rộng các nhiệm vụ, các nhà sản xuất đang tích cực theo đuổi các mô hình tính toán thay thế – đặc biệt là những mô hình lấy cảm hứng bởi những gì được biết về các mạng thần kinh sinh học – nhằm nâng cao hiệu quả phần cứng và sức mạnh của hệ thống máy tính. Tại thời điểm này, các máy tính "phỏng nơon" này chưa chứng tỏ thành công lớn, mới chỉ bắt đầu có khả năng thương mại. Nhưng có thể chúng sẽ trở thành thông dụng trong tương lai gần. Các mạng nơon sâu đã tạo ra một điểm nhấn trong bức tranh ứng dụng. Một làn sóng lớn hơn có thể ập đến khi các mạng này có thể được đào tạo và thực thi trên phần cứng phỏng nơon chuyên dụng.

1.1.4 Trí tuệ nhân tạo trong giáo dục

Mười lăm năm qua, ta đã chứng kiến những tiến bộ đáng kể của AI trong giáo dục. Các ứng dụng ngày nay được sử dụng rộng rãi bởi các nhà sư phạm và người học, với một số thay đổi giữa các hình thức ở trường phổ thông và trường đại học. Mặc dù chất lượng giáo dục sẽ luôn yêu cầu sự tham gia tích cực của các giáo viên, nhưng AI hứa hẹn sẽ tăng cường giáo dục ở tất cả các cấp, đặc biệt bằng cách cung cấp việc học ở quy mô cá nhân hóa. Tương tự như y tế, việc giải quyết làm thế nào để tích hợp tốt nhất sự tương tác của con người và học *mặt-đối-mặt* (face-to-face) với các công nghệ AI triển vọng vẫn còn là một thách thức lớn.

Các robot từ lâu đã là các thiết bị giáo dục phổ biến. Các hệ thống *day học* (tutoring) thông minh (Intelligent Tutoring Systems – ITS) cho khoa học, toán học, ngôn ngữ, và các môn học khác phù hợp sinh viên với các gia sư máy tương tác. Xử lý Ngôn ngữ tự nhiên, đặc biệt là khi kết hợp với máy học và tạo nguồn từ đám đông, đã đẩy mạnh học trực tuyến và cho phép giáo viên nhân bội quy mô của lớp học đồng thời giải quyết các nhu cầu và phong cách học tập của các cá nhân học sinh. Các bộ dữ liệu từ các hệ thống học trực tuyến lớn đã thúc đẩy sự tăng trưởng nhanh trong phân tích học tập.

Tuy nhiên, việc áp dụng các công nghệ AI trong các trường học (phổ thông và đại học) còn chậm, chủ yếu do thiếu ngân sách và thiếu bằng chứng vững vàng là chúng giúp học sinh đạt được mục tiêu học tập. Mười lăm năm tới ở một thành phố ở Bắc Mỹ điển hình, việc sử dụng các gia sư thông minh và công nghệ AI khác để hỗ trợ giáo viên trong lớp học và ở nhà có khả năng sẽ mở rộng đáng kể, cũng như sẽ học dựa trên các ứng dụng thực tế ảo. Nhưng hệ thống học tập dựa trên máy tính chưa có khả năng thay thế hoàn toàn giảng viên trong các trường học.

Robot dạy học

Ngày nay, nhiều công ty đã cung cấp các bộ dụng cụ tinh vi và đa dạng hơn sử dụng trong trường phổ thông cùng các robot với công nghệ cảm biến mới có thể lập trình bằng nhiều ngôn ngữ. Tuy nhiên, để các bộ dụng cụ như vậy trở nên phổ biến, sẽ cần phải có bằng chứng thuyết phục rằng chúng cải thiện thành tích học tập của học sinh.

Hệ thống gia sư thông minh (ITS) và học trực tuyến

Sự di chuyển nhanh của ITS từ giai đoạn thử nghiệm trong phòng thí nghiệm sang sử dụng thực tế là đáng ngạc nhiên và được hoan nghênh. Hệ thống giáo dục trực tuyến hỗ trợ đào tạo chuyên nghiệp cấp sau đại học và học tập suốt đời cũng đang phát triển nhanh chóng. Những hệ thống này có triển vọng rất lớn vì nhu cầu tương tác mặt-đối-mặt ít quan trọng đối với các chuyên gia và những người chuyển đổi việc làm. Tuy không phải là những người đi đầu trong các hệ thống và ứng dụng hỗ trợ TTNT hỗ trợ và các ứng dụng, nhưng họ sẽ trở thành những người tiếp nhận ban đầu khi các công nghệ được kiểm tra và xác nhận.

Các dự án hiện nay tìm cách lập mô hình các quan niệm sai lầm phổ biến của học sinh, dự đoán các học sinh có nguy cơ thất bại, và cung cấp ngay thông tin phản hồi cho học sinh liên quan chặt chẽ với kết quả học tập. Nghiên cứu gần đây cũng dành cho tìm

hiểu về quá trình nhận thức liên quan đến sự hiểu biết, viết, tiếp thu kiến thức, và trí nhớ, và áp dụng hiểu biết đó vào thực tế giáo dục bằng cách phát triển và thử nghiệm các công nghệ giáo dục.

1.2 Máy trò chuyện (Chatbot)

1.2.1 Chatbot là gì?

Chatbot là một lĩnh vực của AI. Chatbot là một hệ thống thực hiện sự trao đổi thông tin giữa hai hay nhiều đối tượng theo một quy chuẩn nhất định, quá trình trao đổi thông tin có thể bằng ngôn ngữ nói, ngôn ngữ viết hoặc kí hiệu (Hoàng, 2011).

Chatbot có thể hiểu đơn giản là một chương trình máy tính mà người dùng có thể giao tiếp với máy thông qua các ứng dụng nhắn tin. Một chatbot có thể nói và hiểu tiếng nói và sẽ phân tích những gì con người nói và cố gắng hiểu một yêu cầu đưa ra. Chatbot sau đó giao tiếp với các máy khác, truyền đạt câu hỏi sau đó trả lời con người.

Chatbot giúp cho con người tiết kiệm thời gian, chi phí thông qua ứng dụng trong việc chăm sóc khách hàng (tự động hóa quy trình...), hay nâng cao năng suất.lao động (các bot giúp đặt lịch...) hay thậm chí chăm sóc đời sống con người (các bot chăm sóc sức khỏe...).

Chatbot có thể được phân loại thành 3 loại chính (Hoàng, 2011):

- Chatbot giữa người với người.
- Chatbot giữa máy với máy.
- Chatbot giữa người và máy.

Mặc dù chatbot là chủ đề "hot" trong thời gian gần đây, nhưng thực ra chatbot đã có mặt từ cách đây 50 năm. Năm 1950, từ ý tưởng của Turing là đưa ra một thiết bị thông minh sẽ thay thế con người thực hiện các cuộc hội thoại. Ý tưởng này giúp hình thành nền tảng cho cuộc cách mạng chatbot. Sau đó, Eliza là chương trình chatbot đầu tiên được phát triển năm 1966. Chương trình được tạo ra để "đóng vai" nhà trị liệu trả lời các câu hỏi đơn giản với các cấu trúc câu xác định. Chương trình được phát triển bởi ông Joseph Weizenbaum, Viện Công nghệ Massachusetts, Mỹ.

Ngày nay với sự xuất hiện của máy tính ở mọi nơi và dựa trên kho *cơ sở dữ liệu* (database) đa dạng và đồ sộ được lưu trữ trên máy tính. Để có thể khai thác được kho dữ liệu đa dạng và đồ sộ này máy tính cần có khả năng xử lý thông tin trong quá trình trao đổi thông tin (hội thoại). Với khả năng hội thoại thông minh, chatbot có thể đáp ứng được yêu cầu trên để trở thành một chương trình tư vấn trợ giúp cho mọi người.

1.2.2 Nền tảng Chatfuel

1.3 Lý thuyết Ứng đáp Câu hỏi

Lý thuyết Ứng đáp Câu hỏi (Item Response Theory - IRT) là một lý thuyết của khoa học về đo lường trong giáo dục, ra đời từ nửa sau của thế kỷ XX và phát triển mạnh mẽ cho đến nay. Trước đó, Lý thuyết Trắc nghiệm cổ điển (Classical Test Theory – CTT), ra đời từ khoảng cuối thế kỷ XIX và hoàn thiện vào khoảng thập niên 1970, đã có nhiều đóng góp quan trọng cho hoạt động đánh giá trong giáo dục, nhưng cũng thể hiện một số hạn chế. Các nhà tâm lý học (psychometricians) cố gắng xây dựng một lý thuyết hiện đại sao cho khắc phục được các hạn chế đó. Lý thuyết trắc nghiệm hiện đại được xây dựng dựa trên mô hình toán học, đòi hỏi nhiều tính toán, nhưng nhờ sự tiến bộ vượt bậc của công nghệ tính toán bằng máy tính điện tử vào cuối thế kỷ XX – đầu thế kỷ XXI, nên nó đã phát triển nhanh chóng và đạt được những thành tựu quan trọng.

Trong phần này, ta quy ước gọi một con người có thuộc tính cần đo lường là *thí sinh* (person – TS) và một đơn vị của công cụ để đo lường (test) là *câu hỏi* (item – CH). Để đơn giản hóa cho mô hình nghiên cứu xuất phát có thể đưa ra các giả thiết sau đây:

- Tính đơn chiều: *Năng lực tiềm ẩn* (latent trait) cần đo chỉ có một chiều (unidimensionality), hoặc ta chỉ đo một chiều của năng lực đó.
- Tính độc lập: Các CH là *độc lập địa phương* (local independence), tức là việc trả lời một CH không ảnh hưởng đến các CH khác.

Khi thỏa mãn hai giả thiết nêu trên thì không gian năng lực tiềm ẩn đầy đủ chỉ chứa một năng lực. Khi đó, người ta giả định là có một *hàm đặc trưng câu hỏi* (Item Characteristic Function – ICF) phản ánh mối quan hệ giữa các biến không quan sát được (năng lực của TS) và các biến quan sát được (việc trả lời CH). Đồ thị biểu diễn hàm đó được gọi là *đường cong đặc trưng câu hỏi* (Item Characteristic Curve – ICC).

Trong phần này, ta chỉ khảo sát CH nhị phân, tức là CH mà câu trả lời chỉ có 2 mức: 0 (sai) và 1 (đúng).

1.3.1 Các mô hình đường cong đặc trưng của câu hỏi nhị phân

Đường cong đặc trưng câu hỏi nhị phân, một tham số (mô hình Rasch)

Mô hình Rasch chỉ biểu diễn CH qua tham số *độ khó* của CH. Phát biểu sau đây của Rasch có giá trị như một tiền đề làm cơ sở cho mô hình của ông:

"Một người có năng lực cao hơn một người khác thì xác suất để người đó trả lời đúng một câu hỏi bất kỳ phải lớn hơn xác suất của người sau; cũng tương tự như vậy, một câu hỏi khó hơn một câu hỏi khác có nghĩa là xác suất để một người bất kỳ trả lời đúng câu hỏi đó phải bé hơn xác suất để trả lời đúng câu hỏi sau."
(Rasch, 1993)

Với phát biểu trên, có thể thấy xác suất để một TS trả lời đúng một CH nào đó phụ thuộc vào tương quan giữa năng lực của TS và độ khó của CH. Chọn Θ để biểu diễn

năng lực của TS, và β để biểu diễn độ khó của CH. Gọi P là xác suất trả lời đúng CH, xác suất đó sẽ phụ thuộc vào tương quan giữa Θ và β theo một cách nào đó, do vậy ta có thể biểu diễn:

$$f(P) = \frac{\Theta}{\beta}, \quad (1.1)$$

trong đó f là một hàm nào đó của xác suất trả lời đúng.

Lấy logarit tự nhiên của (1.1) ta được:

$$\ln f(P) = \ln \left(\frac{\Theta}{\beta} \right) = \ln \Theta - \ln \beta = \theta - b. \quad (1.2)$$

Tiếp đến, để đơn giản, khi xét mô hình trắc nghiệm nhị phân, Rasch chọn hàm f chính là biểu thức *mức được thua* (odds) hoặc *khả năng thực hiện đúng* (likelihood ratio), tức là $f(P) = \frac{P}{1-P}$, qua đó biểu diễn tỉ số của khả năng xảy ra sự kiện khẳng định so với khả năng xảy ra sự kiện phủ định. Như vậy:

$$\ln \frac{P}{1-P} = \theta - b. \quad (1.3)$$

Biểu thức (1.2) được gọi là *logit* (log odds unit).

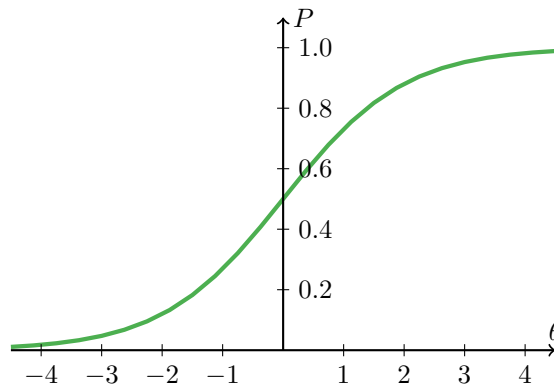
Từ (1.3) ta có thể viết

$$\frac{P}{1-P} = e^{\theta-b}.$$

Suy ra

$$P(\theta) = \frac{e^{\theta-b}}{1 + e^{\theta-b}}. \quad (1.4)$$

Hàm có dạng như biểu thức (1.4) thuộc loại hàm *logistic*. Biểu thức (1.4) chính là hàm đặc trưng của mô hình ứng đáp CH một tham số, hay còn gọi là *mô hình Rasch*, ta có thể biểu diễn như hình 1.1 (khi cho $b = 0$):



Hình 1.1. Đường cong DTCH một tham số (mô hình Rasch)

Mô hình đường cong đặc trưng của câu hỏi hai tham số

Đối với mô hình Rasch, chỉ một tham số của CH được sử dụng, đó là độ khó, nên được gọi là *mô hình một tham số*. Tuy nhiên, trong trắc nghiệm cổ điển, người ta còn sử

dùng một tham số quan trọng thứ hai đặc trưng cho CH là *độ phân biệt*. Do đó nhiều nhà tâm lý học mong muốn đưa độ phân biệt vào mô hình DTCH.

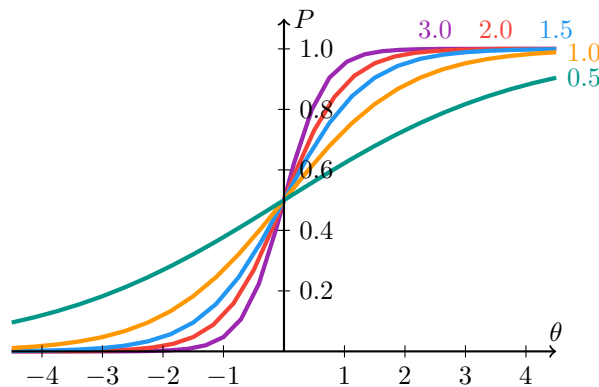
Từ công thức (1.4), ta thấy rõ khi trục hoành biểu diễn theo logit, độ dốc phần giữa đường cong được quyết định bởi hệ số ở số mũ của e , mà ở công thức (1.4), hệ số đó bằng 1. Từ đó, người ta đưa thêm tham số a liên quan đến độ phân biệt của CH vào hệ số ở số mũ của e , ta được:

$$P(\theta) = \frac{e^{a(\theta-b)}}{1 + e^{a(\theta-b)}}. \quad (1.5)$$

(1.5) chính là hàm DTCH 2 tham số. Hệ số a biểu diễn độ dốc của đường cong DTCH tại điểm có hoành độ $\theta = b$ và tung độ $P(\theta) = 0.5$.

Có thể thấy rõ độ dốc của đường cong DTCH phản ánh độ phân biệt của CH. Thật vậy, khi cho một biến đổi vi phân $\Delta\theta$ của năng lực thì sẽ thu được một biến đổi vi phân ΔP của xác suất trả lời đúng, giá trị ΔP này lớn hơn trên đường cong DTCH có độ dốc lớn so với trên đường cong có độ dốc nhỏ. Nói cách khác, đối với CH đã cho một sự khác biệt nhỏ về năng lực của TS cũng gây ra một độ chênh lớn về xác suất trả lời đúng. Đó chính là ý nghĩa của độ phân biệt.

Hàm DTCH 2 tham số trình bày trên đây và hàm DTCH 1 tham số theo mô hình Rasch có cùng dạng thức, chỉ khác nhau ở giá trị tham số a (đối với mô hình 1 tham số $a = 1$). Hình 1.2 biểu diễn các đường cong DTCH theo mô hình 2 tham số với $b = 0$, và a lần lượt bằng 0.5; 1.0; 1.5; 2.0; 3.0 nên độ dốc của các đường cong ở đoạn giữa tăng dần.



Hình 1.2. Các đường cong DTCH 2 tham số với các giá trị a khác nhau ($b = 0$)

Mô hình đường cong đặc trưng của câu hỏi ba tham số

Các hàm DTCH (1.4) và (1.5) chúng ta thấy tung độ tiệm cận trái của chúng đều có giá trị bằng 0, điều đó có nghĩa là nếu TS có năng lực rất thấp, tức $\Theta \rightarrow 0$ và $\theta \rightarrow -\infty$, thì xác suất trả lời đúng CH $P(\theta)$ cũng bằng 0.

Tuy nhiên trong thực tế triển khai trắc nghiệm chúng ta đều biết có khi năng lực của TS rất thấp nhưng do đoán mò hoặc trả lời hú họa một CH nên TS vẫn có một khả năng nào đó trả lời đúng CH. Trong trường hợp đã nêu thì tung độ tiệm cận trái của đường cong không phải bằng 0 mà bằng một giá trị xác định c nào đó, với $0 < c < 1$.

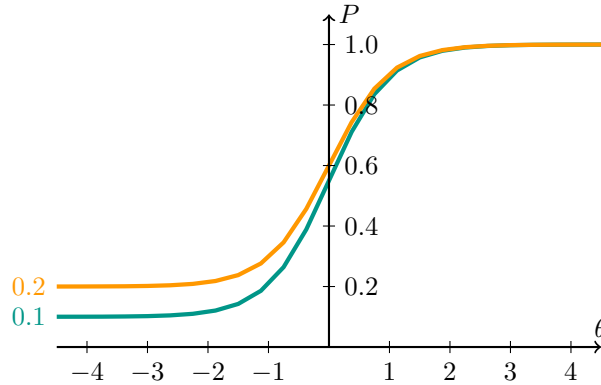
Từ thực tế nêu trên, người ta có thể đưa thêm tham số c phản ánh hiện tượng đoán mò vào hàm DTCH để thu được tung độ tiệm cận trái của đường cong khác 0. Kết quả

sẽ thu được biểu thức:

$$P(\theta) = c + (1 - c) \frac{e^{a(\theta-b)}}{1 + e^{a(\theta-b)}}.$$

Rõ ràng khi $\theta \rightarrow -\infty$, hàm $P(\theta) \rightarrow c$. Trong trường hợp mô hình đường cong DTCH 3 tham số khi $\theta = b$ ta có $P(\theta) = \frac{c+1}{2}$.

Hình 1.3 biểu diễn các đường cong DTCH theo mô hình 3 tham số với $a = 2$ và các tham số c có giá trị bằng 0.1 và 0.2.



Hình 1.3. Các đường cong DTCH 3 tham số với $a = 2$, $c = 0.1$ và 0.2

Mô hình đường cong DTCH 2 và 3 tham số do Birnbaum đề xuất đầu tiên, nên đôi khi được gọi là các mô hình Birnbaum (Birnbaum, 1968).

Mô hình đặc trưng của câu hỏi dạng đường cong tích lũy vòm chuẩn

Vì phân bố chuẩn xác suất là nền tảng của lý thuyết thống kê, nên từ lâu các nhà tâm lý học đã dùng *đường cong tích lũy vòm chuẩn* (normal ogive) làm mô hình để nghiên cứu việc trả lời CH. Tính hợp lý của việc sử dụng đường cong tích lũy vòm chuẩn làm đường cong DTCH được biện minh cả trên quan điểm thực dụng lẫn lý thuyết.

Biểu thức đường cong tích lũy vòm chuẩn đối với mô hình 2 tham số có dạng:

$$P(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a(\theta-b)} e^{-\frac{x^2}{2}} dx, \quad (1.6)$$

và đối với mô hình 3 tham số:

$$P(\theta) = c + (1 - c) \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a(\theta-b)} e^{-\frac{x^2}{2}} dx. \quad (1.7)$$

Biểu thức (1.6) và (1.7) cho thấy các hàm này là hàm xác suất tích lũy tính theo mật độ xác suất của phân bố chuẩn. Đó là các hàm của biến năng lực θ với các tham số a , b , c .

Khi khảo sát quan hệ định lượng giữa các mô hình DTCH có dạng đường cong tích lũy vòm chuẩn và mô hình DTCH có dạng logistic, nếu nhân tham số biểu thị độ dốc a

của hàm logistic cho hệ số $D = 1,702$ và sử dụng như ở biểu thức (1.5) thì sự sai khác tuyệt đối giữa các xác suất biểu diễn bởi biểu thức hàm dạng logistic (1.5) và biểu thức hàm dạng tích lũy vòm chuẩn (1.6) sẽ bé hơn 0.01 trên cả thang θ , nói cách khác, hai đường cong gần như trùng nhau.

Như vậy, đối với mọi ứng dụng thực tiễn hai mô hình hàm DTCH dạng logistic và dạng tích lũy vòm chuẩn là như nhau. Trong khi đó biểu thức toán học của hàm logistic đơn giản hơn nhiều và tốc độ tính toán thực tế đối với chúng giảm nhiều vì không phải tính tích phân, do đó thậm chí có thể tính chúng trên các máy tính giản đơn. Vì lý do đó, người ta thiên về sử dụng mô hình các đường cong logistic hơn là mô hình các đường cong tích lũy vòm chuẩn (Thiệp, 2011).

1.3.2 Quy trình ước lượng các tham số của câu hỏi trắc nghiệm

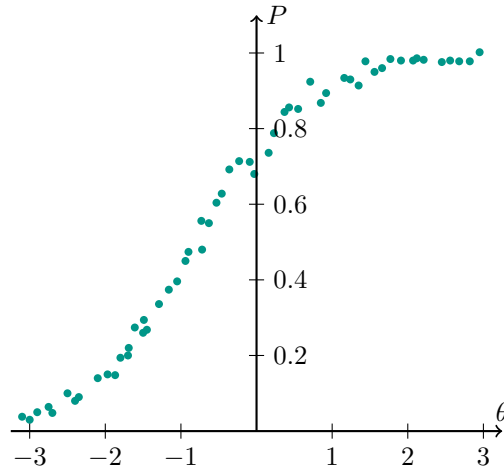
Trong các mô hình IRT, xác suất để trả lời đúng CH phụ thuộc vào năng lực θ của TS và các tham số đặc trưng cho CH. Tuy nhiên, cả hai loại tham số: năng lực của TS (θ) và đặc trưng của CH (a, b, c), đều không biết trước. Cái có thể biết được là việc trả lời các CH của các TS. Vấn đề của việc ước lượng là xác định các giá trị tham số năng lực θ của từng TS và các tham số a, b, c của từng CH từ các kết quả ứng đáp CH. Để áp dụng IRT cho số liệu trắc nghiệm, công việc đầu tiên và quan trọng nhất chính là ước lượng các tham số đặc trưng cho mô hình ứng đáp CH đã chọn. Thành công của áp dụng IRT là đưa ra được các quy trình thích hợp để ước lượng các tham số này.

Trước hết ta xem xét việc ước lượng tham số đặc trưng cho CH trắc nghiệm. Khi ước lượng các tham số này, ta giả thiết là đã biết các điểm năng lực của TS.

Xét tập hợp n TS làm một ĐTN có m CH. Các điểm năng lực của TS phân bố dọc theo một thang đo năng lực. Xét một CH _{i} xác định thứ i . Ta chia tập hợp TS trên thành I nhóm trên thang đo năng lực, sao cho các TS trong cùng một nhóm j nào đó có cùng một năng lực θ_j , cụ thể là có n_j TS trong nhóm j , với $j = \overline{1, I}$. Trong nhóm j giả sử có r_j TS trả lời đúng câu hỏi CH _{i} đã cho. Vậy ở mức năng lực θ_j , tỉ lệ trả lời đúng CH _{i} quan sát được là $p_j(\theta_j) = \frac{r_j}{n_j}$, đó là ước lượng xác suất trả lời đúng CH _{i} ở mức năng lực đã cho. Từ đó có thể thu được r_j và tính được $p_j(\theta_j)$ cho mỗi mức năng lực j dọc theo thang năng lực đã cho. Có thể biểu diễn các tỉ lệ trả lời đúng đối với mỗi nhóm năng lực như ở hình 1.4.

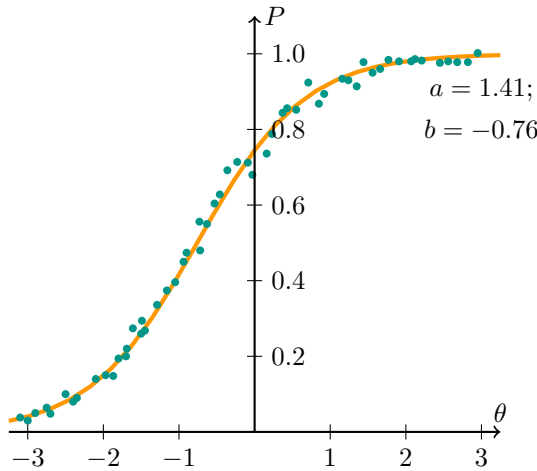
Nhiệm vụ được đặt ra là tìm một đường cong DTCH trùng khớp tốt nhất với các tỷ số trả lời đúng CH quan sát được. Muốn vậy, trước hết ta phải chọn một mô hình đường cong sao cho phù hợp. Quy trình sử dụng để tìm đường cong trùng khớp được dựa trên thuật toán *ước lượng theo biến cố hợp lý cực đại* (maximum likelihood estimation).

Trước hết, người ta cho các giá trị tiên nghiệm (a priory) của các tham số đường cong, chẳng hạn $b = 0.0$ và $a = 1.0$ đối với mô hình hàm DTCH 2 tham số. Sử dụng các ước lượng đó để tính các giá trị $P(\theta_j)$ đối với mọi nhóm năng lực nhờ công thức ứng với mô hình đường cong đã chọn. Sau đó theo một thuật toán xác định như đã nêu trên, người ta điều chỉnh các tham số ước lượng của đường cong DTCH sao cho đạt được một sự trùng khớp tốt hơn giữa đường cong DTCH tính theo các tham số ước lượng và các tỷ lệ trả lời đúng quan sát được. Quá trình tính lặp để điều chỉnh như vậy sẽ tiếp tục



Hình 1.4. Minh họa các tỉ lệ trả lời đúng ứng với mỗi nhóm năng lực

cho đến khi sự điều chỉnh không làm tăng mức trùng khớp một cách đáng kể. Lúc đó thì dừng chương trình tính lặp và các giá trị a và b đạt được cuối cùng chính là giá trị tham số của đường cong DTCH ước lượng được. Với các giá trị a và b thu được ta có thể tìm xấp xỉ đường cong $P(\theta)$ theo mô hình đã chọn, đó là đường cong trùng khớp tốt nhất với số liệu quan sát. Ví dụ trên hình 1.5 biểu diễn đường cong DTCH 2 tham số trùng khớp tốt nhất với số liệu quan sát được ở hình 1.4.



Hình 1.5. Đường cong DTCH 2 tham số trùng khớp tốt nhất với số liệu

Một câu hỏi quan trọng liên quan đến việc ước lượng tham số, đó là khi nào thì có thể xem một đường cong DTCH cụ thể là trùng khớp với số liệu trả lời một CH. Sự phù hợp giữa các tỷ số trả lời quan sát với các số liệu tính toán từ đường cong DTCH có thể xem là trùng khớp được đánh giá bằng chỉ số trùng khớp tốt *Chi-bình phương* (Chi-square goodness-of-fit index). Chỉ số đó được xác định ở công thức:

$$\chi^2 = \sum_{j=1}^I n_j \frac{[p(\theta_j) - P(\theta_j)]^2}{P(\theta_j)Q(\theta_j)}.$$

Trong đó:

- I : Số nhóm năng lực.

- θ_j : mức năng lực của nhóm thứ j .
- n_j : số TS của nhóm thứ j (có năng lực θ_j).
- $p(\theta_j)$: xác suất trả lời đúng của nhóm thứ j theo tính toán từ mô hình DTCH dùng để ước lượng tham số.
- $Q(\theta_j) = 1 - P(\theta_j)$.

Nếu giá trị của chỉ số thu được lớn hơn một giá trị tiêu chí quy định nào đó thì đường cong DTCH được xác định bởi các giá trị đã ước lượng của tham số CH_i là không trùng khớp với số liệu quan sát. Sự không trùng khớp này có thể do hai nguyên nhân. Thứ nhất, mô hình đường cong DTCH được chọn không phù hợp. Thứ hai, các giá trị của tỷ số trả lời đúng CH_i rất phân tán nên không thể thu được sự trùng khớp tốt đối với bất cứ mô hình đường cong DTCH nào. Thông thường khi phân tích một ĐTN có một ít CH không trùng khớp do nguyên nhân thứ hai thì người ta phải sửa chữa CH trắc nghiệm tương ứng hoặc loại bỏ nó khỏi ĐTN. Còn nếu có rất nhiều CH cho số liệu tính toán không trùng khớp với số liệu quan sát thì thường là do chọn mô hình đường cong DTCH không phù hợp, trong trường hợp đó người ta có thể thử nghiệm chọn một mô hình khác.

1.3.3 Điểm thực – đường cong đặc trưng của đề trắc nghiệm

Ở phần trước ta đã xét đặc trưng của từng CH trắc nghiệm và tương tác của từng CH với từng TS, nhưng trong thực tế các CH trắc nghiệm thường được tập hợp thành một ĐTN. Dưới đây ta sẽ xét đến một ĐTN bao gồm nhiều CH trắc nghiệm.

Giả sử CH trắc nghiệm chúng ta xét là CH nhị phân: trả lời đúng được 1 điểm, trả lời sai được 0 điểm. Điểm thô của một TS sẽ thu được bằng cách cộng các điểm của mọi CH trong ĐTN. Như vậy, điểm thô của ĐTN đối với một TS thường là một số nguyên nằm giữa 0 và n , trong đó n là số CH trong ĐTN. Giả sử một TS làm lại ĐTN (và khi làm lại người đó không nhớ những gì đã làm những lần trước), người đó sẽ được một điểm thô khác. Giả thiết là TS làm ĐTN nhiều lần và nhận được nhiều điểm thô khác nhau, các điểm này phân bố quanh một giá trị trung bình nào đó. Theo lý thuyết về đo lường, giá trị trung bình đó gần với một giá trị được gọi là điểm thực, và định nghĩa của nó phụ thuộc vào một lý thuyết đo lường xác định.

Với $U = (U_1, U_2, \dots, U_n)$ là vectơ ứng đáp, trong đó U_i ($i = \overline{1, n}$) có giá trị 0 (trả lời sai) hoặc 1 (trả lời đúng). Ta có thể biểu diễn điểm thô X tính theo số câu trả lời đúng bằng biểu thức:

$$X = \sum_{i=1}^n U_i.$$

Tiếp đến, ta biểu diễn điểm thực τ theo biểu thức kỳ vọng toán học của X như sau:

$$\tau = E(X) = E\left(\sum_{i=1}^n U_i\right) = \sum_{i=1}^n E(U_i),$$

trong đó E là toán tử kỳ vọng toán học và có tính chất tuyến tính.

Nếu một biến ngẫu nhiên Y lấy các giá trị y_1, y_2 với các xác suất tương ứng là P_1 và P_2 thì

$$E(Y) = y_1 P_1 + y_2 P_2.$$

Vì U_i có giá trị bằng 1 với xác suất $P_i(\theta)$ và giá trị bằng 0 với xác suất $Q_i(\theta) = 1 - P_i(\theta)$ nên:

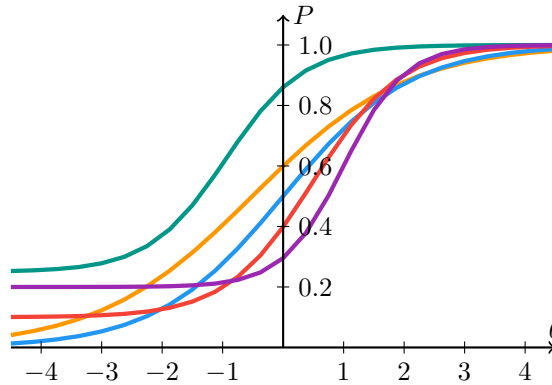
$$E(U_i) = 1.P_i(\theta) + 0.Q_i(\theta).$$

Cuối cùng ta có:

$$\tau = \sum_{i=1}^n P_i(\theta).$$

Tức là: điểm thực của một TS có năng lực θ là tổng của các xác suất trả lời đúng của mọi CH của ĐTN tại giá trị θ . Như vậy, đối với mọi giá trị θ nếu chúng ta tiến hành cộng tất cả mọi đường cong ĐTCH trong ĐTN chúng ta sẽ thu được đường cong đặc trưng của ĐTN, hoặc cũng gọi là đường cong điểm thực. Đường cong đặc trưng của ĐTN là quan hệ hàm số giữa điểm thực và thang năng lực: cho trước một mức năng lực bất kỳ có thể tìm điểm thực tương ứng qua đường cong đặc trưng ĐTN.

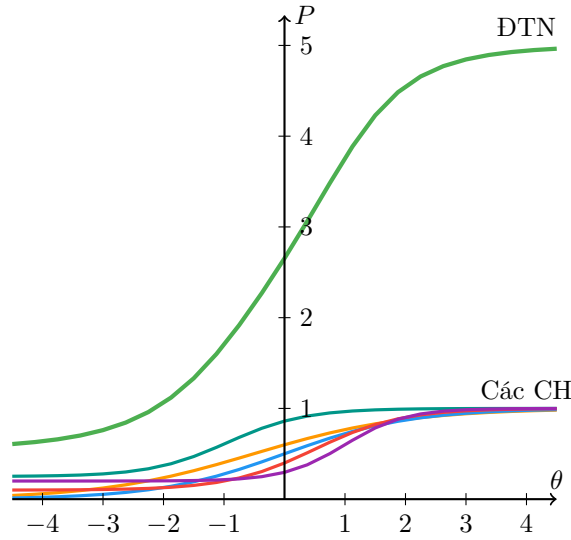
Giả sử một ĐTN bao gồm 5 câu hỏi với các đường cong ĐTCH tương ứng được biểu diễn ở hình 1.6.



Hình 1.6. 5 đường cong ĐTCH theo mô hình 3 tham số

Đường cong đặc trưng của ĐTN bao gồm 5 CH nói trên thu được bằng cách cộng 5 đường cong ĐTCH biểu diễn trên hình 1.7.

Có thể mô tả các đặc điểm của đường cong đặc trưng ĐTN tương tự như mô tả các đường cong ĐTCH. Đường cong đặc trưng ĐTN không có biểu thức giải tích đơn giản nên không có các tham số đặc trưng. Độ nghiêng của đường cong đặc trưng ĐTN cho biết điểm thực phụ thuộc như thế nào vào năng lực, tức là liên quan đến *độ phân biệt của ĐTN*. Trong một số trường hợp đường cong đặc trưng ĐTN có dạng gần đường thẳng trong một khoảng năng lực nào đó, nhưng nói chung nó có dạng một đường cong đồng biến. Mức năng lực ứng với trung điểm của thang điểm thực (ứng với $\frac{n}{2}$) xác định vị trí của ĐTN trên thang năng lực. Hoành độ của điểm đó xác định *độ khó của ĐTN*. Hai yếu tố độ dốc và mức năng lực ở trung điểm thang điểm thực mô tả khá rõ đặc tính của một ĐTN.



Hình 1.7. Đường cong đặc trưng của DTN gồm 5 CH và 5 đường cong DTCH tương ứng

Để biểu diễn điểm thực dưới dạng thập phân, người ta chia τ cho tổng số CH của DTN:

$$\pi = \frac{\tau}{n} = \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n P_i(\theta).$$

Khi θ ở trong khoảng $-\infty < \theta < +\infty$ thì π nằm giữa 0 và 1 (hoặc 0% và 100%). Đối với mô hình ứng đáp CH 3 tham số, giới hạn dưới của π là $\frac{1}{n} \sum_{i=1}^n c_i$.

1.3.4 Ước lượng năng lực của thí sinh

Trong phần này, để ước lượng năng lực của TS, ta giả thiết đã có biết giá trị các tham số của các CH trắc nghiệm. Với DTN gồm các câu hỏi nhị phân, ta thu được một vectơ ứng đáp bao gồm một dãy các giá trị 0 hoặc 1 đối với m CH trong DTN.

Để ước lượng tham số năng lực của TS, ta cũng dùng quy trình *biến cố hợp lý cực đại* như ở phần trước. Trước hết, ta gán một giá trị tiên nghiệm nào đó cho năng lực của một TS và sử dụng các tham số đã biết của các CH trong DTN để tính các xác suất ứng đáp đúng mỗi CH đối với TS đã chọn. Sau đó sử dụng một sự điều chỉnh giá trị ước lượng năng lực để làm tăng sự phù hợp của các xác suất ứng đáp CH tính được với vectơ ứng đáp CH của TS. Quá trình điều chỉnh được lặp lại nhiều lần cho đến khi có một bước điều chỉnh cho giá trị đủ bé, tức là không tạo một sự thay đổi đáng kể của giá trị năng lực được ước lượng. Kết quả ước lượng đó được xem là giá trị tham số năng lực của TS.

Giả sử một TS nào đó được chọn cách ngẫu nhiên có năng lực θ ứng đáp một nhóm m CH nhị phân với kiểu ứng đáp được biểu diễn bởi vectơ U sau đây:

$$U = (U_1, U_2, \dots, U_i, \dots, U_m),$$

trong đó $U_i = 0$ (ứng đáp đúng) hoặc $U_i = 1$ (ứng đáp sai) đối với CH thứ i . Với giả thiết về tính *độc lập địa phương*, ta có thể biểu diễn xác suất ứng đáp nhóm CH của TS có năng lực θ là tích của các xác suất trả lời từng CH:

$$P(U_1, U_2, \dots, U_i, \dots, U_m | \theta) = P(U_1 | \theta) \cdot P(U_2 | \theta) \dots P(U_i | \theta) \dots P(U_m | \theta),$$

hay

$$P(U|\theta) = \prod_{i=1}^m P(U_i|\theta).$$

Vì U_i bằng 0 hoặc 1 nên viết:

$$P(U|\theta) = \prod_{i=1}^m P(U_i|\theta)^{U_i} [1 - P(U_i|\theta)]^{1-U_i} = \prod_{i=1}^m P_i^{U_i} Q_i^{1-U_i}, \quad (1.8)$$

trong đó, $P_i = P(U_i|\theta)$ và $Q_i = 1 - P(U_i|\theta)$. Dạng thức 1.8 biểu diễn xác suất của kiểu ứng đáp nhóm CH nói trên. Khi kiểu ứng đáp nhóm CH đã quan sát được, tức đã có các giá trị $U_i = u_i$, thì sử dụng từ xác suất sẽ không thích hợp nữa, nên xác suất đó được gọi là *biến cố hợp lý* (likelihood) và được biểu diễn bởi hàm $L(u_1, u_2, \dots, u_i, \dots, u_m|\theta)$, trong đó u_i là sự ứng đáp đối với CH thứ i , tức là:

$$L(u_1, u_2, \dots, u_i, \dots, u_m) = \prod_{i=1}^m P_i^{u_i} Q_i^{1-u_i}. \quad (1.9)$$

Vì P_i và Q_i là các hàm của θ và các tham số của CH nên L cũng là hàm của các tham số đó.

Đơn giản hóa việc tính toán bằng cách lấy logarit tự nhiên của biểu thức (1.9):

$$\ln L(u|\theta) = \sum_{i=1}^m [u_i \ln P_i + (1 - u_i) \ln 1 - P_i],$$

trong đó u là vectơ các ứng đáp các CH của TS. Giá trị θ làm cho hàm biến cố hợp lý (hoặc tương ứng, ln của hàm biến cố hợp lý) đối với một TS đạt cực đại được định nghĩa là *ước lượng* của năng lực θ theo *biến cố hợp lý cực đại* đối với TS đó.

Việc tìm giá trị cực đại của L hoặc $\ln L$ là một quá trình phức tạp khi có nhiều TS và nhiều CH. Giá trị tạo cực đại của hàm có thể tìm bằng quy trình tìm kiếm nhờ máy tính. Một trong các cách tìm có hiệu quả là dựa vào tính chất đạo hàm bậc nhất của L hoặc $\ln L$ bằng 0 ở vị trí cực đại. Người ta thiết lập được các phương trình từ tính chất đó và giải giải bằng phương pháp giải tích trực tiếp hoặc phương pháp xấp xỉ, điển hình là phương pháp xấp xỉ Newton – Raphson.

Chương 2

Xây dựng API xử lý

- 2.1 Tổng hợp ngân hàng câu hỏi chương Xác suất – Thống kê
- 2.2 Xây dựng API tạo đề kiểm tra
- 2.3 Xây dựng API xử lý dữ liệu

Chương 3

Thiết kế AI Chatbot

- 3.1 Nền tảng Facebook Chatbot và Chatfuel**
- 3.2 Quá trình thiết kế Chatbot**

Chương 4

Thực nghiệm Sư phạm

Kết luận

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus iaculis quis quam a mollis. Maecenas vulputate viverra dui, vitae luctus elit porta sed. In fringilla eu augue ac pharetra. Donec sodales sem tortor, luctus fermentum est semper eu. Morbi ac leo vel sapien interdum commodo. Vivamus non aliquam leo. Aliquam ac enim et sapien imperdiet mollis pretium id nisi. Mauris non sagittis tortor. Vestibulum commodo, ante vel bibendum pharetra, neque enim venenatis tellus, sed dictum lectus justo vel diam.

Nulla facilisi. Etiam quis sapien vel dui tempus volutpat. Cras ut turpis non turpis posuere volutpat luctus eu magna. Curabitur ornare tellus felis, non hendrerit nisi luctus tristique. Aenean mollis faucibus scelerisque. Aenean commodo feugiat quam, hendrerit fringilla arcu feugiat in. Aliquam cursus luctus ex. Aenean aliquet varius nibh sit amet ullamcorper. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Pellentesque posuere leo ac vestibulum laoreet. Sed molestie, tellus vel vulputate malesuada, nulla leo imperdiet turpis, in facilisis libero mauris at odio.

Integer tincidunt sagittis turpis, at tincidunt sapien viverra ut. Fusce maximus est et nulla consectetur euismod. Maecenas accumsan vestibulum vehicula. Cras molestie odio ac ex tincidunt, in ullamcorper lectus tristique. Curabitur luctus sagittis arcu eget viverra. Curabitur eget justo odio. Duis vel neque sollicitudin, eleifend elit sit amet, facilisis lacus. Nam tempor elementum convallis. Aliquam id magna sed purus porta mollis vitae quis quam. Aenean ac dolor euismod, pellentesque quam eget, efficitur magna.

Tài liệu tham khảo

- Birnbaum, A. L. (1968). Some latent trait models and their use in inferring an examinee's ability. *Statistical theories of mental test scores*.
- Bellman, R. (1978). *An introduction to artificial intelligence: Can computers think?* Thomson Course Technology.
- Rich, E., & Knight, K. (1991). Artificial intelligence.
- Salin, E., & Winston, P. (1992). Machine learning and artificial intelligence: An introduction. *Analytical chemistry (Washington, DC)*, 64(1), 49A–60A.
- Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. ERIC.
- Hoàng, Đ. T. (2011). *Ứng dụng mã nguồn mở aiml xây dựng hệ thống chatbot trợ giúp phương pháp học tập cho sinh viên ngành kỹ thuật* (Doctoral dissertation). Đại học Đà Nẵng.
- Thiệp, L. Q. (2011). *Đo lường trong giáo dục: Lý thuyết và ứng dụng*. NXB Đại học Quốc gia Hà Nội, Hà Nội.
- Nguyễn, T. T., Hà, Q. T., Phan, X. H., & Nguyễn, T. T. (2018). Trí tuệ nhân tạo trong thời đại số: Bối cảnh thế giới và liên hệ với Việt Nam. *Tạp chí Công Thương – Các kết quả nghiên cứu khoa học và Ứng dụng công nghệ*.