

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

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.

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.

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.

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 (Artificial Intelligence – AI)	3
1.2 Nền tảng Chatbot	3
1.3 Lý thuyết Ứng đáp Câu hỏi	3
1.3.1 Các mô hình đường cong đặc trưng của câu hỏi nhị phân	4
1.3.2 Quy trình ước lượng các tham số của câu hỏi trắc nghiệm	7
1.3.3 Điểm thực – đường cong đặc trưng của đề trắc nghiệm	9
1.3.4 Ước lượng năng lực của thí sinh	11
2 Xây dựng API xử lý	12
2.1 Tổng hợp ngân hàng câu hỏi chương Xác suất – Thống kê	12
2.2 Xây dựng API tạo đề kiểm tra	12
2.3 Xây dựng API xử lý dữ liệu	12
3 Thiết kế AI Chatbot	13
3.1 Nền tảng Facebook Chatbot và Chatfuel	13
3.2 Quá trình thiết kế Chatbot	13
4 Thực nghiệm Sư phạm	14
Kết luận	15

Danh sách hình vẽ

1.1	Đường cong DTCH một tham số (mô hình Rasch)	5
1.2	Các đường cong DTCH 2 tham số với các giá trị a khác nhau ($b = 0$) . . .	5
1.3	Các đường cong DTCH 3 tham số với $a = 2$, $c = 0.1$ và 0.2	6
1.4	Minh họa các tỉ lệ trả lời đúng ứng với mỗi nhóm năng lực	8
1.5	Đường cong DTCH 2 tham số trùng khớp tốt nhất với số liệu	8
1.6	5 đường cong DTCH theo mô hình 3 tham số	10
1.7	Đường cong đặc trưng của ĐTN gồm 5 CH và 5 đường cong DTCH tương ứng	11

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)
ML	Học máy (Machine Learning)
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 bố 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 dự 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 (Artificial Intelligence – AI)

1.2 Nền tảng Chatbot

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 kì 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 kì 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

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

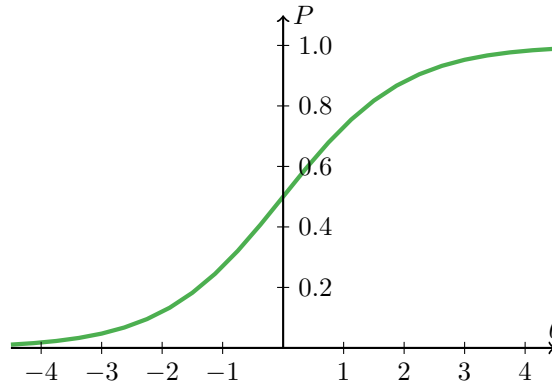
Suy ra

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

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 $b = 0$):

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 dụng



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

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.

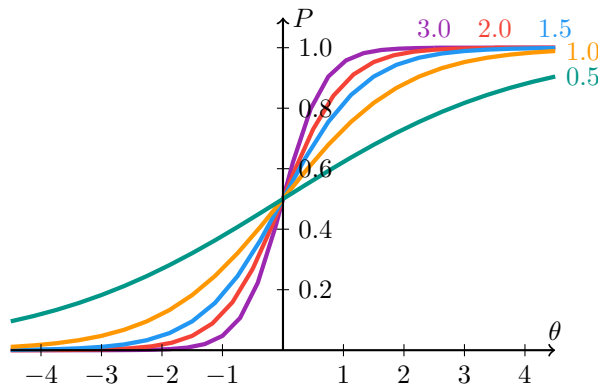
Từ công thức (1.4), ta thấy rõ khi trục hoành biểu diễn theo logit, độ dốc phần giữa đường cong được quyết định bởi hệ số ở số mũ của e , mà ở công thức (1.4), hệ số đó bằng 1. Từ đó, người ta đưa thêm tham số a liên quan đến độ phân biệt của CH vào hệ số ở số mũ của e , ta được:

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

(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 ĐTCH (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.

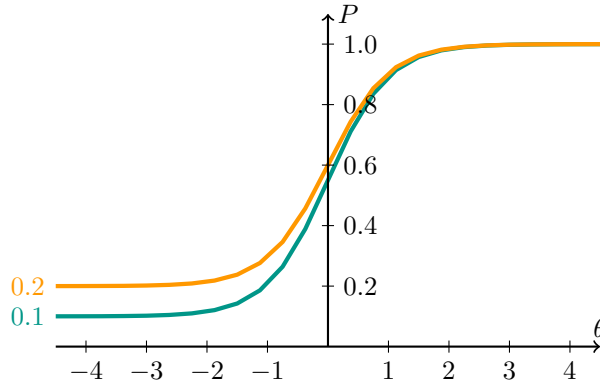
Tuy nhiên trong thực tế triển khai trắc nghiệm chúng ta đều biết có khi năng lực của TS rất thấp nhưng do đoán mò hoặc trả lời hù họa một CH nên TS vẫn có một khả năng nào đó trả lời đúng CH. Trong trường hợp đã nêu thì tung độ tiệm cận trái của đường cong không phải bằng 0 mà bằng một giá trị xác định 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 ĐTCH 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 ĐTCH 3 tham số khi $\theta = b$ ta có $P(\theta) = \frac{c+1}{2}$.

Hình 1.3 biểu diễn các đường cong ĐTCH 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 ĐTCH 3 tham số với $a = 2$, $c = 0.1$ và 0.2

Mô hình đường cong ĐTCH 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 ĐTCH 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 ĐTCH có dạng đường cong tích lũy vòm chuẩn và mô hình ĐTCH 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 ĐTCH 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).

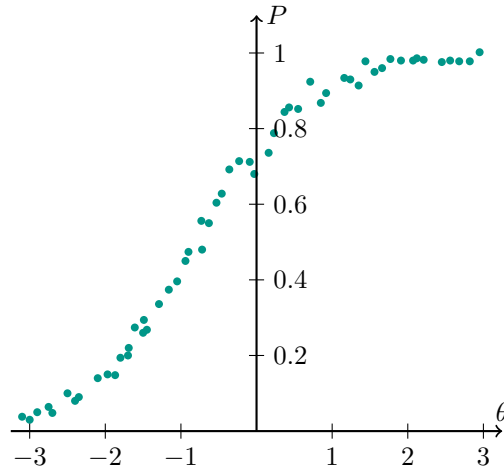
1.3.2 Quy trình ước lượng các tham số của câu hỏi trắc nghiệm

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.

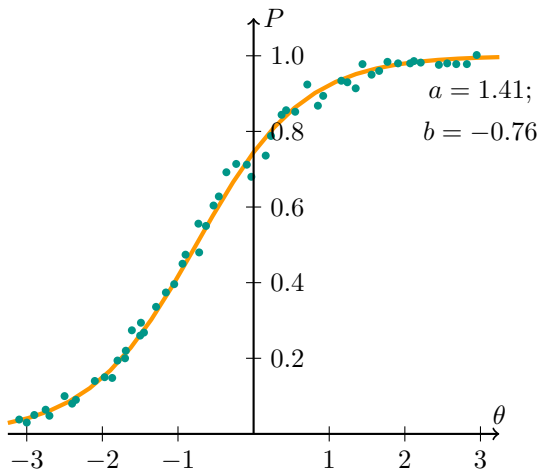
Nhiệm vụ được đặt ra là tìm một đường cong ĐTCH trùng khớp tốt nhất với các tỷ số trả lời đúng CH quan sát được. Muốn vậy, trước hết ta phải chọn một mô hình đường



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

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 priori) của các tham số đường cong, chẳng hạn $b = 0.0$ và $a = 1.0$ đối với mô hình hàm ĐTCH 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 ĐTCH sao cho đạt được một sự trùng khớp tốt hơn giữa đường cong ĐTCH 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 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 ĐTCH 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 ĐTCH 2 tham số trùng khớp tốt nhất với số liệu quan sát được ở hình 1.4.



Hình 1.5. Đường cong ĐTCH 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 ĐTCH 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:

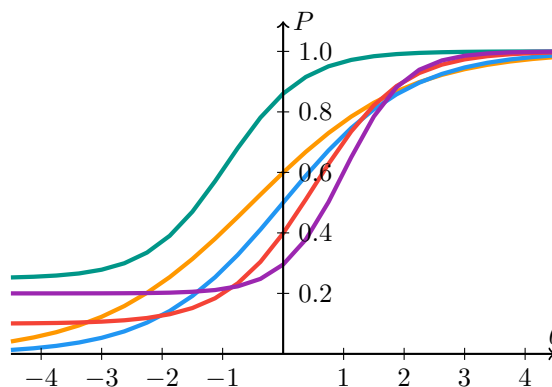
$$E(U_i) = 1 \cdot P_i(\theta) + 0 \cdot Q_i(\theta).$$

Cuối cùng ta có:

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

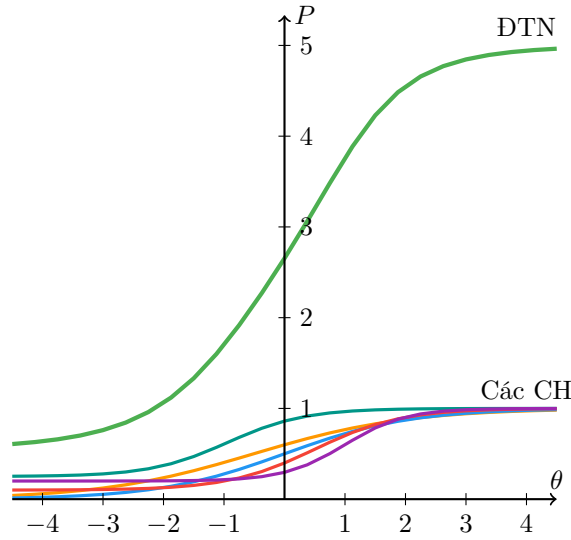
Tức là: điểm thực của một TS có năng lực θ là tổng của các xác suất trả lời đúng của mọi CH của ĐTN tại giá trị θ . Như vậy, đối với mọi giá trị θ nếu chúng ta tiến hành cộng tất cả mọi đường cong ĐTCH trong ĐTN chúng ta sẽ thu được đường cong đặc trưng của ĐTN, hoặc cũng gọi là đường cong điểm thực. 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.



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

Có thể mô tả các đặc điểm của đường cong đặc trưng DTN tương tự như mô tả các đường cong DTCH. Đường cong đặc trưng DTN không có biểu thức giải tích đơn giản nên không có các tham số đặc trưng. Độ nghiêng của đường cong đặc trưng DTN cho biết điểm thực phụ thuộc như thế nào vào năng lực, tức là liên quan đến *độ phân biệt của DTN*. Trong một số trường hợp đường cong đặc trưng DTN 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 DTN trên thang năng lực. Hoành độ của điểm đó xác định *độ khó của DTN*. Hai yếu tố độ dốc và mức năng lực ở trung điểm thang điểm thực mô tả khá rõ đặc tính của một DTN.

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

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*.
- Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. ERIC.
- Thiệp, L. Q. (2011). Đo lường trong giáo dục: Lý thuyết và ứng dụng. *NXB Đại học Quốc gia Hà Nội, Hà Nội*.