

Evaluation of ChatGPT in Solving the Illustration Exams of the 2023 Vietnamese National High School Graduation Exam

Xuan-Quy Dao¹, Bac-Bien Ngo¹, Xuan-Dung Phan¹, Ngoc-Bich Le^{2*}

¹Eastern International University, Vietnam

²International University, Vietnam National University Ho Chi Minh City, Vietnam

*Corresponding author. Email: lnbich@hcmiu.edu.vn

ARTICLE INFO

Received: 21/03/2023
Revised: 30/03/2023
Accepted: 26/04/2023
Published: 28/04/2024

KEYWORDS

ChatGPT;
High School Graduation Exam;
Language Model;
Artificial Intelligence;
Performance Analysis.

ABSTRACT

This article presents the results of ChatGPT (Feb 13 version) in solving the illustration exams for the 2023 Vietnamese National High School Graduation Exam, released by the Ministry of Education and Training in Vietnam. ChatGPT was tested for the nine subjects, including Mathematics, Literature, English, Physics, Chemistry, Biology, History, Geography, and Civic Education. The results show that ChatGPT can pass the illustration essay in both natural science and social science combinations. For multiple-choice questions, ChatGPT achieved scores of 5.4, 5.75, 4.75, 6.00, 7.75, 6.75, 7.75, and 7.80 in Mathematics, Physics, Chemistry, Biology, History, Geography, Civic Education, and English, respectively, and was found to be potentially of achieve a score of 6.47 in Literature. However, the results also demonstrate that ChatGPT's scores in some subjects such as Mathematics, Physics, Chemistry, and Biology (scores 5-6) are lower than those in other subjects such as Literature, English, History, Geography, and Civic Education (scores 6-8). The data for ChatGPT's questions and answers have been publicly released. It is anticipated that ChatGPT possesses the requisite capabilities to successfully pass the 2023 Vietnamese National High School Graduation Examination.

Đánh Giá ChatGPT trong việc Giải Đề Thi Minh Họa Kỳ Thi Tốt Nghiệp Trung Học Phổ Thông 2023 tại Việt Nam

Đào Xuân Quy¹, Ngô Bắc Biên¹, Phan Xuân Dũng¹, Lê Ngọc Bích^{2*}

¹Trường Đại học Quốc tế Miền Đông, Việt Nam

²Trường Đại học Quốc tế, Đại học Quốc gia Thành phố Hồ Chí Minh, Việt Nam

*Tác giả liên hệ. Email: lnbich@hcmiu.edu.vn

THÔNG TIN BÀI BÁO

Ngày nhận bài: 21/03/2023
Ngày hoàn thiện: 30/03/2023
Ngày chấp nhận đăng: 26/04/2023
Ngày đăng: 28/04/2024

TỪ KHÓA

ChatGPT;
Kỳ thi THPT;
Mô hình ngôn ngữ;
Trí tuệ nhân tạo;
Phân tích năng lực.

TÓM TẮT

Bài báo giới thiệu kết quả thử nghiệm ChatGPT (phiên bản 13 tháng 2) giải đề thi minh họa kỳ thi tốt nghiệp Trung học Phổ thông năm (THPT) 2023 do Bộ GDĐT công bố. Thực hiện thử nghiệm cho các môn học: Toán, Văn, Anh, Lý, Hóa, Sinh, Sử, Địa và GDCD. Kết quả thử nghiệm cho thấy ChatGPT có khả năng vượt qua đề thi minh họa THPT 2023 trong cả hai tổ hợp xét tốt nghiệp: tự nhiên và xã hội. Đối với các môn trắc nghiệm: Toán, Lý, Hóa, Sinh, Sử, Địa, GDCD và Anh, ChatGPT đạt được điểm bài thi tương ứng là 5.4, 5.75, 4.75, 6.00, 7.75, 6.75, 7.75 và 7.80. Đối với môn Văn, ChatGPT có thể đạt điểm 6.47. Kết quả cũng thể hiện điểm số ChatGPT đạt được đối với Toán, Lý, Hóa và Sinh (5-6) thấp hơn số điểm đạt được đối với Văn, Anh, Sử, Địa và GDCD (6-8). Dữ liệu về câu hỏi và câu trả lời của ChatGPT được công bố công khai. Hơn nữa, dựa vào cấu trúc đề thi, có thể dự đoán rằng ChatGPT hoàn toàn có khả năng vượt qua kỳ thi tốt nghiệp THPT 2023.

Doi: <https://doi.org/10.54644/jte.2024.1365>

Copyright © JTE. This is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purpose, provided the original work is properly cited.

1. Giới thiệu

Trong những năm gần đây, sự phát triển của trí tuệ nhân tạo (AI) đã có tác động cách mạng đến lĩnh vực giáo dục. Nhiều mô hình ứng dụng AI trong giáo dục đã được đề xuất và nghiên cứu, bao gồm các nền tảng học trực tuyến với trợ lý ảo đàm thoại với học sinh và sinh viên [1], cũng như mô hình tạo nội dung bài giảng bằng AI dựa trên giọng nói và slide của giáo viên [2]. Tuy nhiên, những mô hình này vẫn còn một số hạn chế cần được giải quyết. Để giải quyết các hạn chế này, ChatGPT có thể được tích hợp vào các mô hình AI trong giáo dục do nó có khả năng trả lời một loạt các câu hỏi trong nhiều lĩnh vực khác nhau. **ChatGPT** [3] là một mô hình ngôn ngữ lớn (LLM), được phát triển bởi **OpenAI** [4], một công ty trí tuệ nhân tạo đóng tại San Francisco, California. ChatGPT được xây dựng dựa trên hệ thống máy học và đào tạo trên một tập dữ liệu văn bản lớn (Wikipedia, Twitter,...) nên nó có khả năng tạo ra các văn bản tự động và phức tạp một cách thông minh. ChatGPT là một trong những mô hình ngôn ngữ đàm thoại đầu tiên có khả năng trò chuyện với người dùng về nhiều chủ đề khác nhau.

Để tích hợp ChatGPT vào nền tảng học trực tuyến, một câu hỏi được đặt ra là liệu ChatGPT có khả năng trả lời các câu hỏi ở mức độ và lĩnh vực khác nhau hay không? Những nghiên cứu gần đây đã thử nghiệm phiên bản hiện tại của ChatGPT và đạt được kết quả tích cực trong lĩnh vực luật [5] và y học [6]. Tuy nhiên, trong lĩnh vực kiểm thử phần mềm [7], ChatGPT vẫn chưa đạt được kết quả như mong đợi. Điều này có thể được giải thích bởi ChatGPT là một mô hình mạng nơ-ron học sâu, chỉ hoạt động tốt khi được huấn luyện trên một lượng dữ liệu đủ lớn và đủ đa dạng để trả lời được nhiều loại câu hỏi. Tùy thuộc vào dữ liệu huấn luyện và cách thức mô tả câu hỏi, ChatGPT sẽ đạt được kết quả khác nhau. Áp dụng công nghệ như ChatGPT trong giáo dục là một xu hướng không thể tránh khỏi trong tương lai. Tuy nhiên, để áp dụng hiệu quả, chúng ta cần hiểu được khả năng của ChatGPT. Trong bài báo này, chúng tôi thực hiện thử nghiệm sử dụng ChatGPT để giải các đề thi minh họa THPT 2023 do Bộ GDĐT Việt Nam công bố [8]. Mục đích là xác định khả năng của ChatGPT trong việc giải các đề thi THPT, từ đó có những định hướng phù hợp cho việc giảng dạy và học tập.

Bài báo này được chia thành năm phần chính để trình bày nội dung. Mục 1 giới thiệu về đề tài nghiên cứu. Mục 2 trình bày về phương pháp nghiên cứu được áp dụng trong bài báo, tập trung vào phương pháp thực hiện được sử dụng để đánh giá hiệu quả của ChatGPT trong việc giải các đề thi THPT. Mục 3 trình bày kết quả thử nghiệm sử dụng ChatGPT để giải các đề thi trong kỳ thi THPT tại Việt Nam cho các môn học Toán, Văn, Anh, Lý, Hóa, Sinh, Sử, Địa và Giáo dục công dân (GDGD). Trong phần này, bài báo giới thiệu kết quả của việc sử dụng ChatGPT để giải các đề thi THPT và đưa ra bàn luận về khả năng của ChatGPT trong việc vượt qua kỳ thi THPT. Mục 4 trình bày các kết luận về khả năng của ChatGPT trong việc sử dụng trong giáo dục. Các phương pháp và kết quả được đánh giá trong bài báo có thể hữu ích trong việc nghiên cứu về sự phát triển của các công nghệ giáo dục.

2. Phương pháp nghiên cứu

Thử nghiệm ChatGPT giải các đề thi minh họa THPT 2023 được thực hiện gồm các bước:

1. Tìm hiểu đề thi THPT 2023: Cấu trúc của đề thi, các nội dung được đánh giá, độ khó của các câu hỏi và kiến thức được yêu cầu trong đề thi;
2. Thu thập dữ liệu: Thu thập các đề thi minh họa THPT 2023 [8] để sử dụng làm dữ liệu đầu vào cho ChatGPT;
3. Tiền xử lý dữ liệu: Làm sạch và chuẩn hóa các câu hỏi để đưa vào ChatGPT [9]. Trong đề thi có các công thức Toán, Lý, Hóa và Sinh được chuẩn hóa dưới dạng “*Latex*” text:

Công thức

$$\int f(x)dx = -\sin x + x^2 + C$$

Biểu diễn dưới dạng Latex

$$\int f(x) dx = -\sin(x) + x^2 + C$$

4. Lấy kết quả: nhập câu hỏi dưới dạng text và lấy kết quả từ ChatGPT;
5. Xử lý kết quả:

Câu hỏi trắc nghiệm: tổng hợp kết quả (đúng: giá trị 1, sai: giá trị 0), số câu trả lời đúng, phần trăm câu trả lời đúng, câu trả lời đúng theo mức độ và chủ đề ... Ví dụ về quá trình thực hiện:

| Câu hỏi trắc nghiệm | Nhập câu hỏi | Nhận câu trả lời | Kết quả |
|---------------------|--------------|------------------|---------|
|---------------------|--------------|------------------|---------|

| | | | |
|---|--|---|---|
| Toán, Lý, Hóa, Sinh, Sử, Địa, GDCD và Anh | Đồ thị của hàm số $y = x^3 - 3x + 2$ cắt trục tung tại điểm có tung độ bằng A. 0; B. 1; C. 2; D. -2 | Để tìm điểm cắt trục tung, ta thay $x = 0$ vào hàm số và tính giá trị của y : $y = 0^3 - 3(0) + 2 = 2$ Vậy, đồ thị của hàm số $y = x^3 - 3x + 2$ cắt trục tung tại điểm có tung độ bằng 2. Đáp án là C. | So sánh đáp án tham khảo: Đúng (giá trị 1) |
|---|--|---|---|

Câu hỏi tự luận: giáo viên chấm và đánh giá.

6. Đánh giá kết quả: Đánh giá ChatGPT trong việc giải các câu hỏi gồm độ chính xác, độ đa dạng và khả năng giải quyết các câu hỏi khó. Trả lời câu hỏi ChatGPT có thể vượt qua kỳ thi THPT 2023 hay không.

3. Kết quả thử nghiệm

Phần này trình bày các kết quả thử nghiệm thu được từ ChatGPT.

3.1. Toán

Trong đề thi minh họa THPT 2023 môn Toán, có tổng cộng 50 câu hỏi được chia thành các chủ đề và mức độ khác nhau, minh họa trong **Bảng 1**. Cụ thể, các chủ đề bao gồm: Hình học không gian (M11a) có 3 câu; Tổ hợp và xác suất (M11b) có 2 câu; Đạo hàm và ứng dụng (M12a) có 10 câu; Hàm số mũ và hàm số logarit (M12b) có 8 câu; Nguyên hàm và tích phân (M12c) có 7 câu; Số phức (M12d) có 6 câu; Khối đa diện (M12e) có 2 câu; Khối tròn xoay (M12f) có 3 câu; và Giải tích không gian (M12g) có 9 câu. Các mức độ bao gồm: Nhận biết (NB) có 28 câu; Thông hiểu (TH) có 10 câu; Vận dụng (VD) có 8 câu; Vận dụng cao (VDC) có 4 câu. Các mức độ có thể được phân loại theo thứ tự các câu hỏi, trong đó mức nhận biết từ câu 1 đến câu 30, mức TH từ câu 30 đến câu 40, và mức VD và VDC nằm từ câu 40 đến câu 50.

Bảng 1. Ma trận kiến thức môn Toán

| | | | Chủ đề | | | | | | | | |
|--------|-----|----|--------|------|----------------------|-----------------|---------------|-----------|--------|--------|-------------------|
| | | | M11a | M11b | M12a | M12b | M12c | M12d | M12e | M12f | M12g |
| | | | 3 | 2 | 10 | 8 | 7 | 6 | 2 | 3 | 9 |
| Mức độ | NB | 28 | 5 | 22 | 7, 9, 19, 20, 26, 27 | 2, 3, 4, 21, 28 | 8, 23, 24, 25 | 1, 12, 16 | 13, 14 | 17 | 6, 10, 11, 15, 18 |
| | TH | 10 | 30, 38 | 33 | 31, 32 | 34, 39 | | 35 | | | 36, 37 |
| | VD | 8 | | | 41 | | 29, 40, 44 | 45 | | 43, 48 | 46 |
| | VDC | 4 | | | 50 | 47 | | 42 | | | 49 |

Thực hiện thử nghiệm cho môn Toán như sau (Chi tiết xem [9]):

Lời nhắc (Prompt): Câu 1. Trên mặt phẳng tọa độ, điểm biểu diễn số phức $z=7-6i$ có tọa độ là

A. (-6;7)

B. (6;7)

C. (7;6)

D. (7; -6)

ChatGPT: Điểm biểu diễn số phức $z = 7 - 6i$ trên mặt phẳng tọa độ có tọa độ là (7, -6). Vì phần thực của số phức z là 7 và phần ảo của số phức z là -6i, vậy ta sẽ đặt điểm đó trên trục thực với giá trị là 7 và trục ảo với giá trị -6. Do đó, tọa độ của điểm biểu diễn số phức z trên mặt phẳng tọa độ sẽ là (7, -6). Vì vậy, đáp án là **D. (7; -6)**.

Lời nhắc (Prompt): Câu 7. Cho hàm số $y = \frac{ax + b}{cx + d}$ có đồ thị là đường cong trong hình bên. Tọa độ giao điểm của đồ thị hàm số đã cho và trục hoành là

A. (0; -2)

B. (2;0)

C. (-2;0)

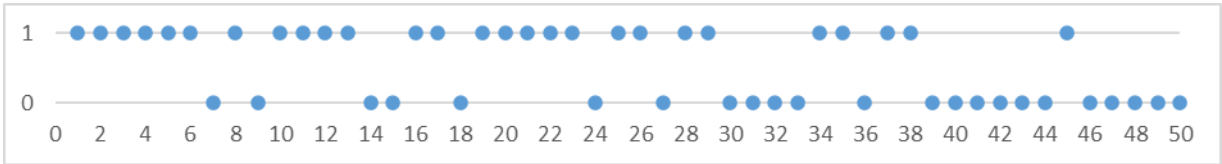
D. (0;2)

ChatGPT: Để tìm tọa độ giao điểm của đồ thị hàm số với trục hoành, ta cần tìm giá trị của x khi $y=0$. Thay $y=0$ vào phương trình hàm số, ta được: $\frac{ax+b}{cx+d}=0$ Do tử số bằng 0, nên $b=0$. Tiếp theo, ta giải phương trình $\frac{ax}{cx+d}=0$ để tìm giá trị của x khi mẫu số khác 0: $\frac{ax}{cx+d}=0$

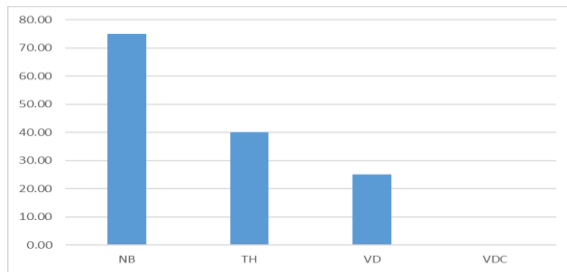
$\setminus\text{Leftrightarrow} ax=0 \setminus\text{Leftrightarrow} x=0$ Vậy tọa độ giao điểm của đồ thị với trục hoành là $(0,0)$. Vậy đáp án chính xác là: **không có đáp án phù hợp trong các lựa chọn.**

Kết quả thử nghiệm, ChatGPT đã trả lời đúng 27/50 câu hỏi (54%), trong đó có 7/50 câu hỏi không có câu trả lời và 16/50 câu trả lời sai. Các câu trả lời của ChatGPT phân bố theo mức độ nhận biết (NB), thông hiểu (TH), vận dụng (VD) và vận dụng cao (VDC) như sau: ChatGPT trả lời đúng nhiều câu hỏi từ câu 1 đến câu 30 (tập trung ở mức NB); số câu trả lời đúng giảm xuống từ câu 30 đến câu 40 (tập trung ở mức TH) và ChatGPT gần như không thể trả lời được những câu hỏi từ câu 40 đến câu 50 (tập trung ở mức VD và VDC). **Hình 1** minh họa sự phân bố này.

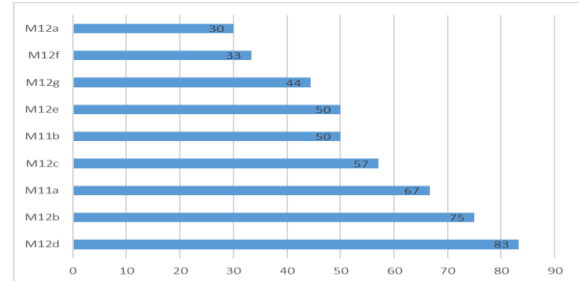
Câu trả lời của ChatGPT theo mức độ và chủ đề của các câu hỏi được minh họa trong **Hình 2**. Theo mức độ, **Hình 2(a)**, ChatGPT đã trả lời đúng 75% câu hỏi ở mức nhận biết (NB), 40% câu hỏi ở mức thông hiểu (TH) và 25% câu hỏi ở mức vận dụng (VD) và không trả lời được bất kỳ câu hỏi nào ở mức vận dụng cao (VDC). Theo chủ đề, **Hình 2(b)**, ChatGPT đã trả lời đúng 83% (5/6) câu hỏi chủ đề số phức (M12d), 75% (6/8) câu hỏi chủ đề hàm số mũ và hàm số logarit (M12b) và chỉ 30% (3/10) câu hỏi chủ đề đạo hàm và ứng dụng (M12a). Mặc dù một số câu hỏi chủ đề đạo hàm và ứng dụng thuộc mức độ nhận biết nhưng ChatGPT không thể trả lời được vì phương pháp khai báo không mô tả được bảng biến thiên của hàm số và một số câu hỏi yêu cầu thông tin trong hình vẽ. Nếu xét cả chủ đề và mức độ, **Hình 2(c-d)**, ChatGPT có thể trả lời đúng 50% câu hỏi mức NB ở tất cả các chủ đề. Tuy nhiên, tỷ lệ trả lời đúng giảm xuống ở mức TH, VD và ChatGPT không thể trả lời được bất kỳ câu hỏi nào ở mức VDC.



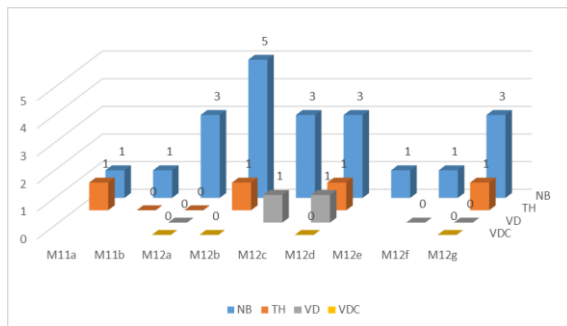
Hình 1. Mức độ chính xác trong câu trả lời của ChatGPT theo thứ tự câu hỏi cho môn Toán.



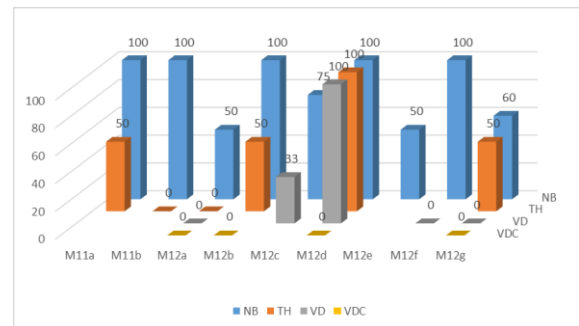
(a)



(b)



(c)



(d)

Hình 2. Câu trả lời đúng của ChatGPT theo mức độ và chủ đề. (a) theo mức độ, (b) theo chủ đề, (c) số câu trả lời đúng theo mức độ và chủ đề, (d) % câu trả lời đúng theo mức độ và chủ đề.

Kết luận: ChatGPT có 27/50 câu trả lời đúng, đạt điểm T = 5.4 (54%).

cấp thông tin và đưa ra một số bài tập tình huống. Theo cấu trúc đề thi, học sinh dễ dàng đạt được điểm nếu nắm vững các kiến thức cơ bản.

Thực hiện thử nghiệm cho các môn Sử, Địa và GDCD như sau (Chi tiết xem trong [9]):

Lời nhắc (Prompt): Câu 5. Trong thập niên 90 của thế kỉ XX, quốc gia nào sau đây có nền kinh tế đứng đầu thế giới?

- A. Hi Lạp. **B. Mĩ.** C. Triều Tiên. D. Ai Cập.

Chọn câu nào ?

ChatGPT: B. Mĩ. Trong thập niên 90 của thế kỉ XX, Mĩ có nền kinh tế phát triển vượt trội và đứng đầu thế giới

Câu 17. Trong giai đoạn 1945-1950, Mĩ có hoạt động nào sau đây để lôi kéo đồng minh?

- A. Viện trợ cho tất cả các nước Mĩ Latinh. B. Viện trợ cho tất cả các nước châu Phi.

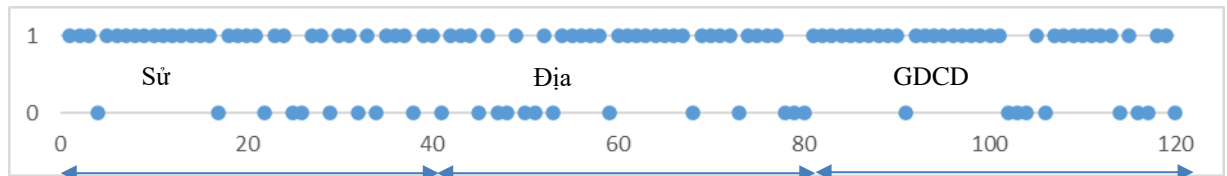
- C. Đưa ra và thực hiện Kế hoạch Mácsan.** D. Viện trợ cho tất cả các nước châu Á.

Chọn câu nào ?

ChatGPT: Chọn đáp án D. Viện trợ cho tất cả các nước châu Á.

Theo kết quả thử nghiệm, ChatGPT trả lời được tất cả các câu hỏi. Trong môn Sử, ChatGPT đã trả lời đúng 31/40 câu hỏi, chỉ sai 9/40 câu. Trong môn Địa, ChatGPT đã trả lời đúng 27/40 câu hỏi, và có 13/40 câu trả lời sai, những câu hỏi này tập trung các kiến thức liên quan đến bản đồ Atlas Việt Nam. Trong môn GDCD, ChatGPT đã trả lời đúng 31/40 câu hỏi, còn 9 câu trả lời sai liên quan đến các bài tập tình huống. Các câu hỏi trong đề thi các môn Sử, Địa và GDCD có tính chất khá rõ ràng và chủ yếu là cung cấp thông tin, do đó ChatGPT đã đạt kết quả tốt trong việc trả lời các câu hỏi.

Hình 4 thể hiện các câu trả lời của ChatGPT theo thứ tự các câu hỏi. Tương tự như các môn tự nhiên, câu trả lời đúng của ChatGPT giảm theo thứ tự câu hỏi nhưng nhìn chung ChatGPT có thể trả lời các câu hỏi VD và VDC tốt hơn khối tự nhiên do bản chất đề thi. Mặc dù, ChatGPT còn hạn chế trong một số câu hỏi trong môn Địa cần dữ liệu đọc bản đồ Atlas Việt Nam, hoặc những câu hỏi về bài tập tình huống trong môn GDCD.



Hình 4. Mức độ chính xác trong câu trả lời của ChatGPT theo thứ tự câu hỏi cho môn Sử (1-40), Địa (41-80) và GDCD (81-120).

ChatGPT là một mô hình ngôn ngữ lớn được huấn luyện trên dữ liệu văn bản rộng lớn, và đã được trang bị kiến thức về các lĩnh vực khác nhau, bao gồm lịch sử, địa lý và đạo đức, đạo lý, pháp luật, tôn giáo và văn hóa. Đối với lĩnh vực lịch sử, ChatGPT đã được huấn luyện trên các dữ liệu về lịch sử Việt Nam và các sự kiện quan trọng trong lịch sử thế giới để có khả năng phân tích các thông tin lịch sử và liên kết các sự kiện và người trong quá khứ để đưa ra các câu trả lời chính xác cho các câu hỏi lịch sử. Đối với lĩnh vực địa lý, ChatGPT có thể trả lời các câu hỏi về địa danh, địa hình, thủy văn học, địa chất học và các thông tin liên quan đến địa lý của một vùng đất cụ thể. Đối với lĩnh vực GDCD, ChatGPT đã được huấn luyện trên các dữ liệu về đạo đức, đạo lý, pháp luật, tôn giáo, văn hóa và các giá trị của xã hội để hiểu và đưa ra các câu trả lời phù hợp cho các câu hỏi trong lĩnh vực này. Vì vậy, ChatGPT có khả năng trả lời các nội dung trong đề thi Sử, Địa và GDCD một cách tương đối chính xác.

Kết luận: ChatGPT có **89/120** câu trả lời đúng, đạt điểm TBTH = **7.42** (74%).

3.4. Tiếng Anh

Phân tích đề thi môn Tiếng Anh cho thấy nội dung đề thi bao gồm kiểm tra các kỹ năng về ngữ âm, ngữ pháp, từ vựng, giao tiếp, viết và đọc. Đề thi có khoảng 84% câu hỏi ở mức độ NB, TH và chỉ khoảng 16% câu hỏi là các câu hỏi VD và VDC.

Thực hiện thử nghiệm cho môn Tiếng Anh như sau (Chi tiết xem [9]):

Lời nhắc (Prompt): Mark the letter A, B, C, or D to indicate the word whose underlined part differs from the other three in pronunciation in each of the following questions.

Question 1:

A. \underline{th}ank B. \underline{th}ick C. \underline{th}at D. \underline{th}ink

ChatGPT: Question 1: D. \underline{th}ink (pronounced with a voiced "th" sound)

Lời nhắc (Prompt): Mark the letter A, B, C, or D on your answer sheet to indicate the correct answer to each of the following questions.

Question 5: The 31st SEA Games, _____ in Vietnam in 2022, left a deep impression on the participants.

A. were held B. held C. holding D. were holding

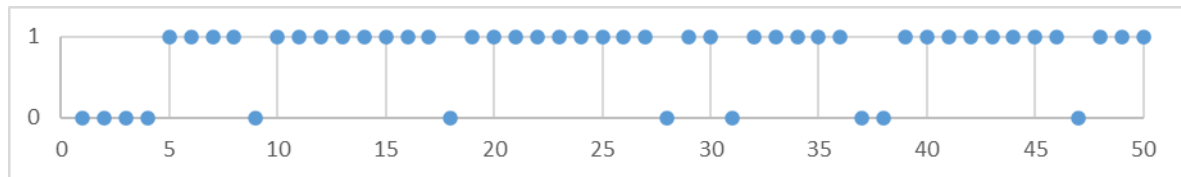
Question 6: Life in the city is _____ than life in the countryside.

A. exciting B. most exciting C. the most exciting D. more exciting

ChatGPT: Question 5: B. Held

Question 6: D. more exciting

Kết quả môn Tiếng Anh, ChatGPT trả lời đúng 39/50 (78%) câu, chỉ 11/50 câu sai. Kết quả thể hiện môn Tiếng Anh, ChatGPT đạt điểm cao hơn so với các môn khác trong kỳ thi minh họa THPT 2023 bởi vì ChatGPT là một mô hình ngôn ngữ tự nhiên, được đào tạo trên cơ sở dữ liệu rất lớn về văn bản Tiếng Anh. ChatGPT có khả năng phân tích và hiểu các cấu trúc ngôn ngữ phức tạp, phân biệt được các cách sử dụng từ và cấu trúc ngữ pháp trong câu. Ngoài ra, môn Tiếng Anh trong kỳ thi minh họa THPT 2023 có khoảng 84% câu hỏi ở mức độ NB, TH và chỉ khoảng 16% câu hỏi là các câu hỏi VD và VDC, điều này giúp ChatGPT có khả năng trả lời được các câu hỏi một cách hiệu quả. Hơn nữa, đối với đề thi Tiếng Anh, ChatGPT có câu trả lời đúng không giảm theo thứ tự câu hỏi, như trong **Hình 5**.



Hình 5. Mức độ chính xác trong câu trả lời của ChatGPT theo thứ tự câu hỏi cho môn tiếng Anh.

Kết luận: ChatGPT có 39/50 câu trả lời đúng, đạt điểm **A = 7.80** (78%).

3.5. Ngữ Văn

Đối với đề thi tự luận môn Văn, cấu trúc đề thi bao gồm hai phần I và II với tổng số điểm là 10. Phần I có ba câu hỏi, tập trung vào khả năng cung cấp thông tin và diễn giải ý nghĩa. Phần II bao gồm hai câu hỏi, trong đó câu hỏi 1 yêu cầu viết một đoạn văn và câu hỏi 2 yêu cầu phân tích một đoạn thơ. Để đánh giá câu hỏi phần II, các tiêu chí bao gồm: a) hình thức đoạn văn hoặc cấu trúc; b) xác định đúng vấn đề nghị luận; c) triển khai vấn đề nghị luận; d) chính tả, từ ngữ, ngữ pháp; và đ) sáng tạo. Thang điểm môn Văn được đưa ra trong **Bảng 2**, với cấu trúc đề thi môn Văn, không khó để đạt được điểm trung bình.

Thực hiện thử nghiệm cho môn Văn như sau (Chi tiết xem [9]):

Lời nhắc (Prompt): Đọc đoạn trích:

| | | |
|--|--|---|
| “chính mẹ đẻ anh hùng và truyền thuyết | dù uống nước đâu lòng vẫn nhớ nguồn | mùa hạ gió Lào quăng quạt mùa đông sắt se gió bắc |
| Từ túp lều lợp lá lợp tranh cắt cuống nhau bằng lưỡi liềm | thương từ cái kiến con ong tím ruột bằm gan thù bọn ác | dân tộc tôi khi đứng dậy làm người mồ hôi vã một trời sao trên đất |
| bàn chân thô quanh năm bùn lấm chưa một lần ướm qua sử sách | dân tộc tôi khi đứng dậy làm người là đứng theo dáng mẹ | trời sao lặn hóa thành muôn mạch nước |
| tập con bước vịn vào ca dao tục ngữ | "đòn gánh tre chín dạn hai vai" (1) | chạy âm thầm chạy dọc thời gian “ |

(1) Câu thơ của Nguyễn Du. Trích Những người đi tới biển, Thanh Thảo, NXB Quân đội Nhân dân, 2004, tr. 53-54)

Thực hiện các yêu cầu sau:

Câu 1. Xác định thể thơ của đoạn trích.

ChatGPT: Đoạn trích không phải là một bài thơ đầy đủ, tuy nhiên, nó được viết theo **thể thơ tự do** (free verse), không có quy luật về độ dài câu thơ, số lượng câu thơ, hoặc âm vần.

Kết quả thử nghiệm cho thấy rằng, với phần đọc hiểu, ChatGPT có khả năng cung cấp các thông tin phù hợp với nội dung câu hỏi (Câu 1, 3, và 4). Đối với phần tự luận, ChatGPT đáp ứng được các yêu cầu về hình thức của đoạn văn và xác định đúng vấn đề nghị luận.

Kết quả môn Ngữ Văn được đưa ra trong **Bảng 2** (trung bình từ ba giáo viên), ChatGPT đạt được điểm **6.47**. Với sự hỗ trợ của các thuật toán học sâu, ChatGPT có thể tự động rút trích các thông tin quan trọng trong đoạn văn, xác định văn phong và lối viết của tác giả, và đưa ra nhận xét về cảm nhận và suy nghĩ của tác giả trong phân tích và nhận xét văn học. Dựa trên nguồn dữ liệu khổng lồ được huấn luyện, ChatGPT có khả năng viết ra một đoạn văn phù hợp với yêu cầu. ChatGPT còn có thể phân tích thông tin và trình bày bài viết một cách xuất sắc, thậm chí trong việc xuất bản khoa học. Do cấu trúc và thang điểm của đề thi môn Văn, ChatGPT có thể đạt được điểm 6.47.

Bảng 2. Thang điểm môn Văn

| | Phần I. Đọc hiểu | | | | Phần II. Tự luận | | | | | | | | | |
|-----------------|------------------|------|------|------|------------------|------|------|------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 1.a | 1.b | 1.c | 1.d | 1.đ | 2.a | 2.b | 2.c | 2.d | 2.đ |
| Điểm thành phần | 0.75 | 0.75 | 1 | 0.5 | 0.25 | 0.25 | 1 | 0.25 | 0.25 | 0.25 | 0.5 | 3.5 | 0.25 | 0.5 |
| ChatGPT=6.47 | 0.75 | 0.55 | 0.73 | 0.37 | 0.23 | 0.23 | 0.58 | 0.20 | 0.10 | 0.20 | 0.50 | 1.83 | 0.20 | 0.30 |

Kết luận: ChatGPT hoàn toàn có thể đạt điểm **6.47**.

3.6. ChatGPT có thể vượt qua kỳ thi minh họa THPT 2023 hay không?

Trong phần này, chúng tôi đưa ra câu trả lời cho câu hỏi **ChatGPT có thể vượt qua kỳ thi minh họa THPT 2023 hay không?** Công thức tính điểm của Bộ Giáo dục và Đào tạo để xét tốt nghiệp:

$$\text{ĐXTN} = \frac{\left(\frac{\text{TĐ4M} + \text{TĐKK}}{4}\right) \times 7 + (\text{TB12}) \times 3}{10} + \text{ƯT} = \frac{\left(\frac{\text{T} + \text{V} + \text{A} + \text{TBTH}}{4}\right) \times 7 + (\text{TB12}) \times 3}{10}$$

Trong đó, TĐ4M= T+V+A+ TBTH với T: Toán, V: Văn, A: Anh, và TBTH: Điểm trung bình tổ hợp; TĐKK: Tổng điểm khuyến khích; TB12 Điểm trung bình cả năm lớp 12; ƯT: Điểm ưu tiên. Trường hợp không tính đến điểm ưu tiên, ĐXTN = 0.175(T+V+A+TBTH) + 0.3TB12. Với T, A, TBTH cố định, đặt giả định V và TB12 thay đổi (vì Văn là môn tự luận, nên không có kết quả chính xác như các môn trắc nghiệm), V có giá trị 5.5-7.5, TB12 có giá trị 5-9, ĐXTN có thể có giá trị như trong **Bảng 3**.

Bảng 3. ĐXTN theo Văn (5.5-7.5) và TB12 (5-9)

| | | Điểm môn Văn | | | | | | | | | |
|-----------------|---|--------------|------|------|------|---------------|------|------|------|------|------|
| | | 5.5 | 6 | 6.5 | 7 | 7.5 | 5.5 | 6 | 6.5 | 7 | 7.5 |
| Điểm TB12 | 5 | 5.74 | 5.82 | 5.91 | 6.00 | 6.09 | 6.07 | 6.16 | 6.25 | 6.33 | 6.42 |
| | 6 | 6.04 | 6.12 | 6.21 | 6.30 | 6.39 | 6.37 | 6.46 | 6.55 | 6.63 | 6.72 |
| | 7 | 6.34 | 6.42 | 6.51 | 6.60 | 6.69 | 6.67 | 6.76 | 6.85 | 6.93 | 7.02 |
| | 8 | 6.64 | 6.72 | 6.81 | 6.90 | 6.99 | 6.97 | 7.06 | 7.15 | 7.23 | 7.32 |
| | 9 | 6.94 | 7.02 | 7.11 | 7.20 | 7.29 | 7.27 | 7.36 | 7.45 | 7.53 | 7.62 |
| Tổ hợp Tự nhiên | | | | | | Tổ hợp Xã hội | | | | | |

Đối với tổ hợp tự nhiên và xã hội, ĐXTN của ChatGPT đều vượt qua kỳ thi THPT (tương ứng từ 5.91 và 6.25). Dựa trên cấu trúc đề thi THPT, phần NB và TH chiếm tỉ trọng cao trong đề thi (70-75%), và ChatGPT có khả năng trả lời đúng các câu NB và TH cao, nhận định **ChatGPT có thể vượt qua kỳ thi THPT 2023** với cấu trúc và nội dung đề thi tương tự đề thi minh họa THPT 2023.

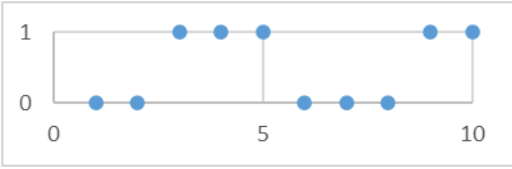
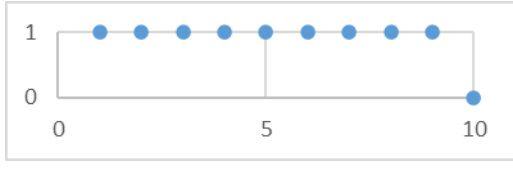
4. Thảo luận

4.1. ChatGPT đưa câu trả lời không ổn định, có cùng một câu hỏi nhưng nhiều câu trả lời khác nhau

ChatGPT đào tạo bằng dữ liệu lớn và được lập trình để đưa ra các câu trả lời và giải thích dựa trên dữ liệu đã học được từ các nguồn khác nhau. ChatGPT có thể đưa ra các câu trả lời không ổn định cho cùng một câu hỏi do một số nguyên nhân sau:

Cách mô tả câu hỏi: ChatGPT là một hệ thống học máy, có khả năng cung cấp các câu trả lời khác nhau cho cùng một câu hỏi hoặc các câu hỏi khác nhau. Để kiểm tra tính ổn định của ChatGPT trong trả lời câu hỏi Toán ở mức độ NB, hai thử nghiệm đã được tiến hành. Thử nghiệm một sử dụng đề thi Toán trực tiếp từ Bộ GDĐT, trong khi thử nghiệm hai chỉnh sửa nội dung bằng cách thay thế dấu chấm phẩy (“;”) bằng dấu phẩy (“,”) trong các đáp án. Mỗi thử nghiệm đã yêu cầu ChatGPT trả lời mười lần, và kết quả thu được được trình bày trong **Bảng 4** (Chi tiết xem [9]). Kết quả cho thấy, chỉ với một sự thay đổi nhỏ (“;” thành “,”) số lượng câu trả lời đúng đã tăng từ 5 lên 9. Nên lưu ý rằng trong thử nghiệm một, mặc dù ChatGPT đã cung cấp giải thích chính xác nhưng chọn sai đáp án cho 5 câu. Nếu được mô tả phù hợp, ChatGPT có thể cung cấp đáp án đúng và giải thích chính xác.

Bảng 4. Kết quả từ ChatGPT khi được hỏi 10 lần cùng một câu hỏi

| | Thử nghiệm một (TN1) | | Thử nghiệm hai (TN2) | |
|-------------------------|---|---------------------------------------|---|--------------------------------------|
| Lần thử | Lời nhắc (Prompt): Câu 1. Trên mặt phẳng tọa độ, điểm biểu diễn số phức $z=7-6i$ có tọa độ là A. (-6;7) B. (6;7) C. (7;6) D. (7;-6) | | Lời nhắc (Prompt): Câu 1. Trên mặt phẳng tọa độ, điểm biểu diễn số phức $z=7-6i$ có tọa độ là A. (-6,7) B. (6,7) C. (7,6) D. (7,-6) | |
| 1 | ChatGPT: Trên mặt phẳng tọa độ, số phức $z = 7 - 6i$ có phần thực bằng 7 và phần ảo bằng -6. Do đó, điểm biểu diễn số phức này trên mặt phẳng tọa độ là điểm có tọa độ $(7, -6)$. Vậy đáp án là \mathbf{D} . | | ChatGPT: Để biểu diễn số phức $z = 7 - 6i$ trên mặt phẳng tọa độ, ta cần lấy phần thực và phần ảo của z để làm tọa độ. Do đó, tọa độ của z sẽ là (7, -6) . Vậy đáp án là D . | |
| ... | ChatGPT: Đây là một câu hỏi lặp lại, câu trả lời đúng là: D. (7;-6). | | ChatGPT: Đáp án đúng là D. (7,-6) | |
| 10 | ChatGPT: Đáp án đúng là B. (6;7) <i>Giải thích:</i> Số phức $z=7-6i$ có phần thực bằng 7 và phần ảo bằng -6. Do đó, điểm biểu diễn số phức này trên mặt phẳng tọa độ là điểm có tọa độ $(7,-6)$. Tuy nhiên, câu hỏi yêu cầu tìm tọa độ của điểm này theo thứ tự (x,y) nên đáp án chính xác là (6,7) . | | ChatGPT: Đáp án đúng là C. (7,6). | |
| Kết quả theo số lần hỏi |  | |  | |
| So sánh | Giải thích chính xác, chọn đáp án đúng | Giải thích chính xác, chọn đáp án sai | Không có giải thích, chọn đáp án đúng | Không có giải thích, chọn đáp án sai |
| TN1 | 4 | 5 | 1 | 0 |
| TN2 | 8 | 0 | 1 | 1 |

Dữ liệu đào tạo không đầy đủ: ChatGPT có thể không có đủ thông tin để đưa ra câu trả lời chính xác hoặc đưa ra các câu trả lời không đồng nhất khi nội dung câu hỏi không được huấn luyện trong dữ liệu.

Độ tin cậy của nguồn dữ liệu để huấn luyện mô hình: độ tin cậy của các nguồn dữ liệu huấn luyện không đảm bảo, ChatGPT đưa ra các câu trả lời không chính xác.

Phương pháp đào tạo của ChatGPT: ChatGPT được huấn luyện bằng phương pháp học máy không giám sát và học từ dữ liệu được cung cấp. Do đó, ChatGPT có thể không hiểu hoặc đưa ra các câu trả lời không chính xác cho các câu hỏi phức tạp hoặc không rõ ràng. Đối với các câu hỏi đơn giản, ChatGPT có độ ổn định cao hơn so với những câu hỏi phức tạp hoặc không rõ ràng.

ChatGPT là mô hình ngôn ngữ tự nhiên không phải là một chương trình tính toán: Mặc dù ChatGPT có thể xử lý và giải quyết nhiều vấn đề phức tạp trong ngôn ngữ tự nhiên, nhưng việc giải các bài toán phức tạp vẫn là một thách thức đối với ChatGPT vì nó hiện tại chưa được thiết kế để giải quyết các bài toán trong lĩnh vực Toán, Lý, Hóa và Sinh.

4.2. ChatGPT “hiện tại” chưa thay thế được giáo viên

ChatGPT có thể mang lại nhiều tiềm năng trong việc giải quyết các vấn đề liên quan đến giáo dục, nhưng hiện tại nó vẫn chưa thể thay thế hoàn toàn vai trò của giáo viên. Trong quá trình giải đề thi minh họa THPT 2023, mặc dù ChatGPT đã thể hiện được khả năng giải quyết các câu hỏi trong nhiều lĩnh vực khác nhau, có kết quả khả quan và đưa ra các giải thích chính xác trong một số trường hợp. Tuy nhiên, ChatGPT chưa thể giải quyết các bài toán phức tạp cần suy luận vì để giải quyết các bài toán này, cần có sự sáng tạo, tư duy phân biện và khả năng tự học hỏi. Hiện tại, ChatGPT chỉ có thể dự đoán và đưa ra các phương án dựa trên những gì đã học từ dữ liệu được cung cấp. Ngoài ra, ChatGPT hiện vẫn còn hạn chế trong việc đưa ra các bình luận sâu sắc và đầy đủ về văn học và xã hội. Do đó, ChatGPT vẫn chưa thể thay thế được vai trò của giáo viên trong việc hướng dẫn và giải đáp thắc mắc của học sinh. ChatGPT và các công nghệ trí tuệ nhân tạo sẽ tiếp tục phát triển và cải tiến, có thể cung cấp cho giáo viên và học sinh các công cụ hữu ích để nâng cao chất lượng giảng dạy và học tập. Tuy nhiên, giáo viên vẫn đóng vai trò rất quan trọng trong việc truyền đạt kiến thức và kỹ năng cho học sinh, giúp họ phát triển toàn diện và đạt được thành công trong cuộc sống.

4.3. Tiềm năng công nghệ mang lại cho giáo dục trong tương lai

Các mô hình ngôn ngữ như ChatGPT đang mang lại tiềm năng trong việc cải thiện hiệu quả giảng dạy và trải nghiệm học tập. Bằng cách cung cấp cho giáo viên các công cụ và phương tiện tương tác đa dạng, công nghệ đang giúp cho việc trợ giúp giảng dạy và kiểm tra kiến thức của học sinh trở nên nhanh chóng và hiệu quả hơn. Việc giảm bớt các công việc lặp đi lặp lại cho phép giáo viên dành thời gian để tập trung vào việc nâng cao kiến thức, kỹ năng và chăm sóc đến học sinh hơn.

Ngoài ra, công nghệ cũng đang đem lại trải nghiệm học tập tiện lợi và dễ dàng hơn cho học sinh và giáo viên. Học sinh có thể tận dụng trợ lý ảo có khả năng trả lời và giải thích trong nhiều chủ đề mọi lúc mọi nơi, giúp tăng tính chủ động và động lực học tập. Việc sử dụng công nghệ giúp cho học sinh có thể tiếp cận với nội dung học tập một cách đa dạng và phong phú, đồng thời giúp cho giáo viên đưa ra phương pháp giảng dạy linh hoạt và hiệu quả hơn.

5. Kết luận

Bài báo giới thiệu kết quả thử nghiệm ChatGPT trong việc giải đề thi tốt nghiệp minh họa THPT 2023. Kết quả cho thấy rằng ChatGPT có khả năng vượt qua kỳ thi và đạt được điểm số khá (6-8) cho các môn Văn, Anh, Sử, Địa, GDCD, và đạt điểm trung bình (5-6) cho các môn Toán, Lý, Hóa và Sinh. ChatGPT là một mô hình ngôn ngữ tự nhiên nên có khả năng trả lời tốt hơn với các câu hỏi về xã hội (Văn, Anh, Sử, Địa và GDCD) so với các câu hỏi về tự nhiên (Toán, Lý, Hóa và Sinh). Mặc dù, hiện tại ChatGPT chưa thể thay thế được vai trò của giáo viên trong giáo dục hay giải những bài toán phức tạp cần khả năng suy luận cũng như phân tích chuyên sâu, nhưng những mô hình trí tuệ nhân tạo như ChatGPT hoàn toàn có tiềm năng lớn để trở thành công cụ hữu ích trong giáo dục.

Xung đột lợi ích

Các tác giả tuyên bố không có xung đột lợi ích trong bài báo này.

TÀI LIỆU THAM KHẢO

- [1] X. Q. Dao, N. B. Le, and T. M. T. Nguyen, "AI-Powered MOOCs: Video Lecture Generation," *ACM Int. Conf. Proceeding Ser.*, pp. 95–102, Mar. 2021, doi: 10.1145/3459212.3459227.
- [2] T. M. T. Nguyen, T. H. Diep, B. B. Ngo, N. B. Le, and X. Q. Dao, "Design of Online Learning Platform with Vietnamese Virtual Assistant," *ACM Int. Conf. Proceeding Ser.*, pp. 51–57, Feb. 2021, doi: 10.1145/3460179.3460188.
- [3] "ChatGPT." <https://chat.openai.com/chat> (accessed Mar. 14, 2023).
- [4] "OpenAI." <https://openai.com/> (accessed Mar. 14, 2023).
- [5] J. H. Choi, K. E. Hickman, and A. B. Monahan, "Chatgpt goes to law school," *Available SSRN*, pp. 1–16, 2022.
- [6] T. H. Kung *et al.*, "Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models," *medRxiv*, p. 2022.12.19.22283643, 2022, doi: 10.1371/journal.pdig.0000198.
- [7] S. Jalil, S. Rafi, T. D. LaToza, K. Moran, and W. Lam, "ChatGPT and Software Testing Education: Promises & Perils," Feb. 2023, doi: 10.48550/arxiv.2302.03287.
- [8] "Sample exam for the High School Graduation Examination in 2023." (in Vietnamese), <https://moet.gov.vn/tintuc/Pages/tin-tong-hop.aspx?ItemID=8425> (accessed Mar. 14, 2023).
- [9] D. X. Quy, "ChatGPT's Performance in Solving High/Middle School Exam Questions," 2023. <https://xdao85.github.io/ChatGPT-TH/> (accessed Mar. 18, 2023).



Dao Xuan Quy obtained the Ph.D. degree in signal processing and telecommunication at the French National Institute for Research in Computer Science and Automation in 2014. He received both the M.S. and B.S. degrees in automation engineering from the Grenoble Institute of Technology in France in 2010. He is currently a lecturer at the Eastern International University in Vietnam. From 2015 to 2017, he was a lecturer at Quang Binh University in Dong Hoi, Vietnam. His research interests include the development Online Learning Technology, Teaching-based Artificial Intelligence, Voice Cloning, and Speech-driven Face. Email: quy.dao@eiu.edu.vn



Ngô Bạc Biện received a Master of Science in Electronic Engineering at HCMC University of Technology and Education. He is currently a lecturer at the Eastern International University in Vietnam. His research interests include Power Converters, PWM Methodology, Online Learning and Data Collection. Email: bien.ngo@eidu.edu.vn/ngobacbiensp@gmail.com



Phan Xuan Dung received Master of Science in Automation Engineering at Ho Chi Minh University of Transport in 2010. He is currently a lecturer at the Eastern International University in Vietnam. From 2000 to 2015, he was a lecturer at Saigon Technology University, Ho Chi Minh City, Vietnam. His research interests include Smart Lighting, Internet of Things, Automation Engineering and Online Learning. Email: dung.phan@eiu.edu.vn



Le Ngoc Bích received his undergraduate degrees at Bach Khoa University, Vietnam, his Master and Ph.D. in Mechatronics Science from Southern Taiwan University of Science and Technology – Taiwan in 2004, 2007, 2010, respectively. He is currently a lecturer at School of Biomedical Engineering, International University, Vietnam National University Ho Chi Minh city, Thu Duc City, Ho Chi Minh City, Vietnam. Le published a number of papers in preferred Journals such as J. Biomedical Microdevices, J. Microfluidics and Nanofluidics, J. Sensors and Actuators, and many Vietnamese Engineering Books in automation, CAD, and mold design. He also presented various academic as well as research-based papers at several national and international conferences. His articles focus on Medical devices, MEMS, Microfluidics, Robotics, and AI. Email: lnbich@hcmiu.edu.vn. ORCID:  <https://orcid.org/0000-0001-7431-0157>