

Nhận diện đặc điểm ngôn ngữ trong văn bản được viết bởi trí tuệ nhân tạo (AI) với văn bản được viết bởi con người và đánh giá hiệu suất của công cụ phát hiện văn bản do AI viết

Trong bối cảnh trí tuệ nhân tạo (AI) phát triển nhanh chóng và học sinh, sinh viên có xu hướng lạm dụng các công cụ chatbot như ChatGPT để tạo lập văn bản, vấn đề phân biệt văn bản do con người viết và văn bản do AI viết càng trở nên cấp thiết. Nghiên cứu được thực hiện bằng phương pháp thu thập dữ liệu với ba nhóm ngữ liệu và sử dụng phần mềm R để xử lý số liệu, tạo ra các biểu đồ mang tính trực quan. Kết quả nghiên cứu cho thấy một số đặc điểm ngôn ngữ không có giá trị phân biệt văn bản do người viết và AI, đồng thời chỉ ra các đặc điểm nhận diện hai loại văn bản này bao gồm: từ chỉ ngôi, từ chỉ yếu tố bản địa, câu đơn, câu phức, câu ghép và câu tỉnh lược, câu hỏi tu từ, hình thức đoạn văn, phương tiện liên kết câu, lỗi chính tả và lỗi ngữ pháp. Trong giới hạn nghiên cứu, GPTZero đạt hiệu suất cao nhất và là công cụ hỗ trợ đáng tin cậy trong việc phát hiện văn bản AI.

Thông tin các tác giả

Trịnh Khánh Hiền: Sinh viên, đang học tập tại Trường Đại học Sư phạm Hà Nội, 136 Xuân Thủy, Cầu Giấy, Hà Nội, Việt Nam, email: khanhhien816@gmail.com

Từ khóa

AI, ChatGPT, tạo sinh văn bản, văn bản tiếng Việt

Auteur: HIỀN, Trịnh Khánh (Trường Đại học Sư phạm Hà Nội)

Classification de Session: Tiểu ban 1

Classification de thématique: Tiểu ban khoa học xã hội, giáo dục