**PHIẾU ĐỀ XUẤT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CẤP CƠ SỞ**

**1. Tên đề tài:**

**HỌC KHÁI NIỆM ĐỐI VỚI CÁC CƠ SỞ TRI THỨC  
TRONG LOGIC MÔ TẢ DỰA VÀO MÔ PHỎNG HAI CHIỀU**

**2. Lĩnh vực nghiên cứu:**

Tự nhiên: Kinh tế, XHNV: Giáo dục:

**X**

Kỹ thuật: Nông Lâm: Y Dược:

Môi trường: ATLĐ: Sở hữu trí tuệ:

Khác:

**3. Tính cấp thiết**

Học máy là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc thiết kế và phát triển các thuật toán cho phép máy tính phân tích được các thông tin dựa trên các dữ liệu thực nghiệm. Học máy được áp dụng trong nhiều lĩnh vực như chuẩn đoán y tế, phân tích thị trường chứng khoán, phân lớp dữ liệu văn bản, … Một trong những nhánh của học máy là học có giám sát, tức là phải học một hàm dự đoán từ các mẫu cho trước. Cụ thể là cho một tập các mẫu của một hàm , xây dựng một hàm sao cho xấp xỉ với . Trong học máy, mỗi đối tượng thường được đặc tả bởi một tập các thuộc tính và được xem như một vector các giá trị thuộc tính. Nếu là một hàm nhị phân thì bài toán học máy được gọi là bài toán phân lớp nhị phân.

Logic mô tả là một họ các ngôn ngữ hình thức thích hợp cho việc biểu diễn và suy luận tri thức. Trong logic mô tả một miền được mô tả thông qua các các thể, khái niệm và vai trò. Một khái niệm đại diện cho một tập các đối tượng, một vai trò đại diện cho một quan hệ hai ngôi giữa các đối tượng.

Bài toán phân lớp nhị phân trong ngữ cảnh logic mô tả được gọi là học khái niệm, hàm phân loại ở đây được đặc trưng bởi một khái niệm. Học khái niệm trong logic được nghiên cứu bởi nhiều nhà nghiên cứu. Badea và Nienhuys-Cheng[[1]](#footnote-1) nghiên cứu học khái niệm trong logic mô tả thông qua các toán tử làm mịn. Cohen và Hirsh[[2]](#footnote-2) đã đề xuất một thuật toán học khái niệm LCSLearn dựa trên “bao hàm phổ biến nhỏ nhất”. Lambrix và Larocchia[[3]](#footnote-3) đưa ra thuật toán học khái niệm đơn giản dựa trên sự chuẩn hóa khái niệm.

Tuy nhiên, những nghiên cứu trong các công trình đã đề cập trên áp dụng cho các hệ thống thông tin trong ngữ cảnh logic mô tả với bộ ký tự và các cấu tử khái niệm còn hạn chế. Do vậy việc tổng quát hóa và mở rộng phương pháp học khái niệm cho các cơ sở tri thức trong ngữ cảnh logic mô tả dựa trên mô phỏng hai chiều là một vấn đề cấp thiết hiện nay.

**4. Mục tiêu:**

Xây dựng lý thuyết và mở rộng phương pháp học khái niệm cho cơ sở tri thức trong ngữ cảnh logic mô tả dựa trên mô phỏng hai chiều.

**5. Nội dung chính:**

- Nghiên cứu và khái quát hóa phương pháp học máy trong ngữ cảnh logic mô tả.

- Mở rộng phương pháp học khái niệm với bộ từ vựng tổng quát dựa trên mô phỏng hai chiều.

- Xây dựng các quan hệ mô phỏng hai chiều lớn nhất nhằm tối ưu phương pháp học khái niệm và xấp xỉ khái niệm.

**6. Sản phẩm và kết quả dự kiến:**

a) Sản phẩm khoa học:

- Số công trình khoa học đăng tạp chí/hội thảo quốc tế: 01 công trình

- Số công trình khoa học đăng tạp chí/hội thảo trong nước: 01 công trình

b) Sản phẩm đào tạo (số lượng CN, ThS, TS): 02 Cử nhân

c) Sản phẩm ứng dụng: *(Mô tả tóm tắt về sản phẩm dự kiến, phạm vi, khả năng và địa chỉ ứng dụng)*:

- Các công trình khoa học đăng trên các tạp chí/hội thảo trong nước và quốc tế.

- Phương pháp học khái niệm dựa trên mô phỏng hai chiều để áp dụng vào các bài toán học máy.

d) Sản phẩm khác:

**7. Thời gian nghiên cứu dự kiến:** Từ năm 2013 đến năm 2014.

**8. Nhu cầu kinh phí dự kiến:** 60.000.000 đồng *(Sáu mươi triệu đồng chẵn)*

*Ngày 02 tháng 05 năm 2012*

Tổ chức/Cá nhân đề xuất

**Trần Thanh Lương**

1. L. Badea and S.-H. Nienhuys-Cheng, “A reﬁnement operator for description logics,” in Proceedings of ILP’2000, ser. LNCS, vol. 1866. Springer, 2000, pp. 40–59. [↑](#footnote-ref-1)
2. W. Cohen and H. Hirsh, “Learning the Classic description logic: Theoretical and experimental results,” in Proceedings of KR’1994, 1994, pp. 121–133. [↑](#footnote-ref-2)
3. P. Lambrix and P. Larocchia, “Learning composite concepts,” in Proceedings of DL’1998, 1998. [↑](#footnote-ref-3)