

TRANG THÔNG TIN LUẬN ÁN TIẾN SĨ

Tên đề tài luận án tiến sĩ:

NGHIÊN CỨU, PHÁT TRIỂN MỘT SỐ PHƯƠNG PHÁP KHAI PHÁ DỮ LIỆU TRÊN DỮ LIỆU CÓ CẤU TRÚC

Chuyên ngành: Hệ thống thông tin

Mã số: 09.48.01.04

Họ và tên nghiên cứu sinh: **Hoàng Minh Quang**

Người hướng dẫn khoa học:

- GS. TS. Vũ Đức Thi**
- GS. TSKH. Nguyễn Ngọc San**

Đơn vị đào tạo: Học viện Công nghệ Bru chính Viễn thông

Cơ sở đào tạo: Học viện Công nghệ Bru chính Viễn thông

NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN

- Phát hiện thuật toán tìm một rút gọn thuộc tính trong thời gian đa thức không sử dụng heuristic.
- Phát hiện thuật toán tìm một rút gọn đối tượng trong thời gian đa thức mà vẫn bảo toàn quá trình tìm tất cả các rút gọn thuộc tính.
- Cải tiến phương pháp sinh cây quyết định thực hiện nhanh hơn quá trình sinh cây quyết định của thuật toán ID3.
- Phát hiện thuật toán PSICFSM giải quyết đẳng cấu đồ thị con trong thời gian đa thức trong khai phá đồ thị con thường xuyên đóng.
- Xây dựng độ đo trên dàn giao khái niệm khác với độ đo dựa trên vector áp dụng cho phân loại đa nhãn đồ thị theo lý thuyết Dempster-Shafer.

CÁC ỨNG DỤNG, KHẢ NĂNG ỨNG DỤNG TRONG THỰC TIỄN HOẶC NHỮNG VẤN ĐỀ CÒN BỎ NGỎ CẦN TIẾP TỤC NGHIÊN CỨU

- Thuật toán tìm rút gọn đối tượng, tìm một rút gọn thuộc tính, sinh cây quyết định được ứng dụng để giảm kích thước dữ liệu, tăng tính hiệu quả khai phá dữ liệu trên bảng quyết định nhất quán.
- Khai phá đồ thị con thường xuyên đóng được ứng dụng trong việc trợ giúp hữu ích trong lĩnh vực y tế như khám phá các thành phần thuốc đặc trị, hoặc hỗ trợ khám phá quy luật biến đổi của tế bào có khả năng gây bệnh chẳng hạn như ung thư.
- Phân loại đa nhãn đồ thị được ứng dụng hỗ trợ cho dữ liệu sinh trắc học phân loại các mẫu gen đồng thời gây nhiều bệnh, phân loại thuốc mới có thể chữa được nhiều bệnh đồng thời.

**Xác nhận của tập thể
người hướng dẫn khoa học**

Nghiên cứu sinh