

TRANG THÔNG TIN LUẬN ÁN TIẾN SĨ

Tên đề tài luận án tiến sĩ: “Rút gọn thuộc tính trong bảng quyết định không đầy đủ có dữ liệu thay đổi theo tiếp cận mô hình tập thô dung sai”

Chuyên ngành: Khoa học máy tính

Mã số ngành: 9480101

Họ và tên NCS: Nguyễn Anh Tuấn

Người hướng dẫn khoa học:

1. PGS.TS. Nguyễn Long Giang
2. GS.TS. Vũ Đức Thi

Cơ sở đào tạo: Trường ĐH Công nghệ thông tin và Truyền thông - ĐH Thái Nguyên

NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN

Luận án đã có những đóng góp mới sau:

1) Xây dựng công thức gia tăng cập nhật khoảng cách trong trường hợp bổ sung, loại bỏ tập thuộc tính, trên cơ sở đó xây dựng thuật toán gia tăng lọc - đóng gói tìm tập rút gọn trên bảng quyết định không đầy đủ trong trường hợp bổ sung tập thuộc tính: thuật toán FWIA_AA và trong trường hợp loại bỏ tập thuộc tính: thuật toán FWIA_DA.

2) Đề xuất công thức tính khoảng cách khi tập đối tượng thay đổi giá trị, trên cơ sở đó đề xuất thuật toán gia tăng lọc - đóng gói tìm tập rút gọn của bảng quyết định không đầy đủ trong trường hợp tập đối tượng thay đổi giá trị: thuật toán FWIA_U_Obj.

3) Đề xuất công thức tính khoảng cách khi tập thuộc tính thay đổi giá trị, trên cơ sở đó đề xuất thuật toán gia tăng lọc - đóng gói tìm tập rút gọn của bảng quyết định không đầy đủ trong trường hợp tập thuộc tính thay đổi giá trị: thuật toán FWIA_U_Attr.

CÁC ỨNG DỤNG, KHẢ NĂNG ỨNG DỤNG TRONG THỰC TIỄN VÀ NHỮNG VẤN ĐỀ CẦN TIẾP TỤC NGHIÊN CỨU

1. Các ứng dụng, khả năng ứng dụng trong thực tiễn

- Kết quả nghiên cứu của luận án cung cấp thêm cơ sở khoa học giúp các nghiên cứu toàn diện về tìm tập rút gọn của bảng quyết định không đầy đủ thay đổi trong tất cả các trường hợp về tập đối tượng, tập thuộc tính thay đổi.

- Xây dựng các thuật toán gia tăng theo hướng tiếp cận lọc - đóng gói nhằm giảm thiểu số thuộc tính tập rút gọn, từ đó nâng cao hiệu quả các mô hình phân lớp, học máy.

- Ứng dụng trong việc xử lý dữ liệu lớn luôn thay đổi và cập nhật, chẳng hạn, trong các hệ thống trực tuyến, trong tin sinh chuẩn đoán bệnh.

2. Những vấn đề còn bỏ ngỏ cần tiếp tục nghiên cứu

Cải tiến các thuật toán gia tăng lọc- đóng gói nhằm giảm thiểu thời gian thực hiện bằng giải pháp không thực hiện lặp lại các bộ phân lớp.

INFORMATION OF DOCTORAL DISSERTATION

Dissertation title: Attribute reduction in dynamic incomplete decision table according to tolerance rough set model approach.

Major: Computer Science

Code No: 9480101

Candidate: Nguyen Anh Tuan

Supervisors:

1. Assoc. Prof. Dr Nguyen Long Giang
2. Prof. Dr Vu Duc Thi

Training Institute: University of Education - Thai Nguyen University

Training unit: Thai Nguyen University of Information and Communication Technology

Training institution: Thai Nguyen University

NEW SCIENTIFIC FINDINGS OF THE DISSERTATION

The dissertation has following new scientific contributions:

1) Build an incremental formula to update the distance in the case of adding and removing the attribute set, on that basis build an incremental algorithms filter-wrapper to find the reduct on the incomplete decision table in case of adding attribute set: algorithm FWIA_AA and in case of removing attribute set: algorithm FWIA_DA.

2) Proposing a formula to calculate the distance when the object set changes its value, on that basis, proposed an incremental algorithms filter-wrapper to find the reduct of the incomplete decision table in case the object set changes value: algorithm FWIA_U_Obj.

3) Propose a formula to calculate the distance when the attribute set changes value, on that basis, proposed an incremental algorithms filter-wrapper to find the reduct of the incomplete decision table in case the attribute set changes value: algorithm FWIA_U_Attr.

APPLICATIONS IN PRACTICE

AND RECOMMENDATIONS FOR FURTHER STUDIES

1. Applications in Practice

- The research results of the thesis provide more scientific basis to help comprehensive studies on finding the reduct of the dynamic incomplete decision table in all cases of the object set, the attribute set changes.

- Building incremental algorithms filter-wrapper approach to minimize the number of reductive set attributes, thereby improving the efficiency of classification and machine learning models.

- Applications in processing big data that are constantly changing and updating, for example, in online systems, in diagnostic bioinformatics.

2. Recommendations for Further Studies

Improved incremental algorithms filter-wrapper to reduce execution time by not repeating classifiers.