

NHỮNG ĐÓNG GÓP MỚI CỦA LUẬN ÁN

Tên đề tài luận án: *Nâng cao hiệu năng phân lớp dữ liệu trên cơ sở cải tiến thuật toán SVM*

Mã số: 9.48.01.01

Ngành: Khoa học máy tính

Họ và tên NCS: Nguyễn Thế Cường Khóa đào tạo: 12/2018

Chức danh, học vị, họ và tên người hướng dẫn: PGS. TS. Huỳnh Thế Phùng

Tên đơn vị đào tạo: Trường Đại học Khoa học, Đại học Huế

Nội dung:

- 1- Đề xuất các thuật toán phân lớp nhị phân đối với dữ liệu có cấu trúc phức tạp, bằng cách sử dụng chiến lược lớp-đối-cụm, tìm nghiệm toàn cục bằng phương pháp giải bài toán đối ngẫu của bài toán tối ưu hoặc phương pháp bình phương tối thiểu. Việc giải các bài toán quy hoạch toàn phương cỡ lớn được thay thế bởi các bài toán quy hoạch toàn phương có cỡ nhỏ hơn.
- 2- Đề xuất các thuật toán phân lớp nhị phân đối với dữ liệu có cấu trúc phức tạp, bằng cách sử dụng chiến lược cụm-đối-lớp, khai thác thông tin cấu trúc của từng cụm trong mỗi lớp và thông tin về số lượng điểm dữ liệu trong mỗi cụm để huấn luyện mô hình. Các thuật toán tìm nghiệm toàn cục bằng phương pháp giải bài toán đối ngẫu hoặc phương pháp dùng bình phương tối thiểu.
- 3- Cung cấp những cài đặt của các thuật toán đã có và các thuật toán đề xuất bằng ngôn ngữ lập trình Python. Có thể làm tài liệu tham khảo cho sinh viên đại học và học viên cao học ngành công nghệ thông tin thực hiện đề tài về phân lớp dữ liệu. Mở ra hướng nghiên cứu mới, áp dụng các chiến lược lớp-đối-cụm hay cụm-đối-lớp để giải quyết bài toán phân lớp với dữ liệu không cân bằng.

Chữ ký của người hướng dẫn

Chữ kí của NCS

NEW CONTRIBUTIONS OF THE DISSERTATION

Title of dissertation: *Enhancing data classification performance based on improving the SVM algorithm.*

Code: 9.48.01.01

Major: Computer Science

Full name of Ph.D.student: Nguyen The Cuong Course: 12/2018

Academic supervisors: Assoc. Prof., Dr. Huynh The Phung

Training institution: University of Sciences, Hue University

Content:

- 4- Propose binary classification algorithms for complex structured data, using the class-vs-cluster strategy, finding global solution by solving the dual problem of the optimization problem or the least square method. Solving large quadratic programming problem is replaced by smaller quadratic programming problems.
- 5- Propose binary classification algorithms for complex structured data, using the cluster-vs-class strategy, exploiting the structure information of each cluster in each class and the information of the number of data points in each cluster to train the model. The Algorithms find global solutions by solving dual problem or least squares method.
- 6- Provides implementations of existing algorithms and proposed algorithms by Python programming language. It can be used as a reference for university and graduate students in information technology to conduct data classification research. Open up new research directions, applying class-vs-cluster or cluster-vs-class strategies to solve classification problems with unbalanced data.

Academic Supervisor

Ph.D. Student