

Hướng dẫn ôn thi môn Học máy.

Đề thi 90', gồm lý thuyết và bài tập. Các bài tập chủ yếu tương tự các bài đã cho

Câu hỏi lý thuyết:

Trình bày các thuật toán xây dựng cây quyết định: ID3, C4.5, CART
Anh/chị hiểu thế nào hiện tượng phù hợp trội và các khắc phục khi sử dụng thuật toán ID3?

Hãy cho biết các tiêu chuẩn chọn thuộc tính và cách xử lý mẫu huấn luyện khi có giá trị thuộc tính bị mất để nâng cao hiệu quả ID3 .

Anh/chị hiểu thế nào là miền và hàm quyết định? Trình bày các thuật toán tìm hàm phân biệt tuyến tính, khoảng cách cực tiểu, máy vectơ tựa.

So sánh hai thuật toán tìm phân biệt tuyến tính: Perceptron và Widrow_Hoff

Cho biết các phương pháp ước lượng hàm mật độ thông dụng, tại sao các phương pháp của số Parzen và K-NN gọi là kỹ thuật phi tham số?

Trình bày các phương pháp phân cụm phân hoạch mà anh/chị biết

So sánh kiến trúc và phương pháp huấn luyện mạng MLP và mạng RBF

Trình bày thuật toán nghịch đảo khoảng cách và hồi quy tuyến tính địa phương.

Trình bày phương pháp lập luận dựa trên tình huống.

Trình bày về bài toán học tăng cường và phương pháp Q-learning cho bài toán đơn định
Giới thiệu các phương pháp tiếp cận xây dựng bộ phân lớp cơ sở, giải thích vì sao các phương pháp tập thể thường có hiệu quả hơn.

Giới thiệu các phương pháp bỏ phiếu, boosting, rừng ngẫu nhiên, cách sử dụng kiến trúc bậc thang để nâng cao hiệu quả các bộ nhận dạng.

Bài tập:

- 1) Xây dựng cây quyết định bằng thuật toán ID3, C4.5
- 2) Áp dụng thuật toán Find-S
- 3) Xây dựng mạng perceptron khi biết tập mẫu
- 4) Xác định hàm phân biệt nhờ phương pháp cực tiểu khoảng cách (Euclidean và Mahalanobis)
- 5) Phương pháp k-NN để phân lớp
- 6) Các bài toán phân lớp Bayes
- 7) Thiết kế kiến trúc mạng nơron khi biết tập mẫu