

ÔN TẬP KẾT THÚC MÔN HỌC

Câu 4: Machine Learning (3.0 điểm)

Cho tập mẫu dữ liệu huấn luyện gồm 8 mẫu như sau. Để dự đoán xem một người nào đó có bị Cháy nắng hay Không cháy nắng dựa vào các thuộc tính dẫn xuất như Tóc, Chiều cao, Dùng Kem?. Thuộc tính Kết quả là thuộc tính phân loại một người nào đó bị cháy nắng hay không. Thuộc tính Tên chỉ là cho biết thông tin về mẫu dữ liệu; Có nghĩa là một người cháy nắng hay không cháy nắng không phụ thuộc vào tên người được xét là gì.

Tên	Tóc	Chiều cao	Dùng kem?	Kết quả
Sarah	Vàng	T.Bình	Không	Cháy
Dana	Vàng	Cao	Có	Không
Alex	Nâu	Thấp	Có	Không
Annie	Vàng	Thấp	Không	Cháy
Emilie	Đỏ	T.Bình	Không	Cháy
Peter	Nâu	Cao	Không	Không
John	Nâu	T.Bình	Không	Không
Kartie	Vàng	Thấp	Có	Không

- Liệt kê các giá trị có thể có của từng thuộc tính dẫn xuất dựa trên tập dữ liệu đã cho. (0.5đ)
- Hãy áp dụng giải thuật Naive Bayes để tính khả năng bị Cháy nắng từng thuộc tính dẫn xuất. (1.0đ)
- Hãy áp dụng giải thuật Naive Bayes để tính khả năng bị Cháy nắng của người nào đó có sử dụng sửa lỗi Laplace. (1.0đ)
- Hãy cho biết một mẫu dữ liệu mới gồm các thông tin <Tóc=Vàng, Chiều cao = Cao, Dùng Kem=No> thuộc nhóm phân loại Kết quả là Cháy hay Không sử dụng đối với trường hợp Câu c. (0.5đ)

ĐÁP ÁN:

Câu 2: (3.0 điểm): Cho các câu được phát biểu bằng ngôn ngữ tự nhiên như sau:

- a. Tất cả các quả bóng đều được làm bằng cao su.
- b. Tất cả các cây hoa được trồng ở nhà của Lan có màu đỏ.

2.1. Viết các câu đã cho dưới dạng logic vị từ cấp 1

2.2. Chuyển câu a, b sang dạng mệnh đề (Clausal Form)

ĐÁP ÁN
