Chuẩn bị

1. Máy tính (dự trù 1 máy)
2. VGA (2 dây)
3. In slide
4. In báo cáo
5. In file note
6. Từ vựng tiếng anh chuyên môn
7. Demo code

Note:

* Đã từng thử chạy thực nghiệm LDA (phân lớp không giám sát)
* Luật đánh giá điểm dữ liệu
  + 10: lời khen hoàn toàn với nhiều khía cạnh.
  + 9,8: lời khen ở 1 vài khía cạnh. Các từ khóa phổ biến: tốt, dễ hiểu, vui vẻ, nhiệt tình, hiểu bài, thân thiện.v.v
  + 7,6,5: những đánh giá ở mức độ vừa phải có thể đi kèm 1 khía cảnh tiêu cực nhưng không phải yếu tố chính. Một số từ khóa phổ biến như: dễ hiểu, dễ tiếp thu, cần…để
  + <5: là những đánh giá tiêu cực, rõ ràng các hành động tiêu cực ảnh hưởng đến việc giảng dạy. Một số từ khóa như: không giải thích, khó khăn, áp lực, thiếu kiến thức, làm việc riêng, gây buồn ngủ, không công bằng, thiếu minh bạch.v.v
* Dán nhãn thủ công theo nguyên tắc:
  + Dưới 5đ: negative
  + Trên 5đ: positive
* Phương pháp đánh giá

|  |  |  |
| --- | --- | --- |
|  | Predicted as Positive | Predicted as Negative |
| Actual: Positive | TP | FP |
| Actual: Negative | FN | TN |

**Độ chính xác** là tỉ lệ dữ đoán cảm xúc chính xác trên toàn tập dữ liệu đầu vào được tính theo công thức như sau:

hay

**Độ bao phủ** là tỉ lệ số dự đoán đúng cảm xúc khi sử dụng bộ phân lớp được tính theo công thức sau:

hay

**F1** là giá trị trung hòa giữa 2 giá trị độ chính xác và độ bao phủ. Chúng ta cần tính F1 bởi vì nếu chỉ căn cứ vào giá trị độ chính xác và độ bao phủ, ta không thể so sánh và đánh giá các bộ phân lớp với nhau trong trường hợp bộ phân lớp này có độ chính xác cao, độ bao phủ thấp trong khi bộ phân lớp còn lại có độ chính xác thấp nhưng độ bao phủ cao. F1 được tính như sau: