





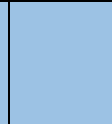









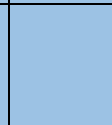





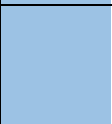











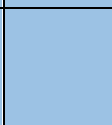

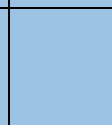
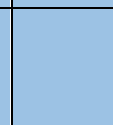
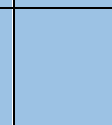

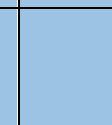

	VIETTEL AI RACE	Public 121
	PHƯƠNG PHÁP VÀ MỘT SỐ ĐỘ ĐO ĐÁNH GIÁ CÁC MÔ HÌNH PHÂN LỚP ĐA NHÃN	Lần ban hành: 1

1. Phương pháp đánh giá các mô hình phân lớp đa nhãn

Để đánh giá các mô hình phân lớp đa nhãn MLL, khóa luận đã sử dụng phương pháp k-fold cross validation tập dữ liệu ban đầu được chia ngẫu nhiên thành k tập con (fold) có kích thước xấp xỉ nhau $S_1, S_2 \dots S_k$. Quá trình học và kiểm tra được thực hiện tại k lần. Tại lần lặp thứ i, S_i là tập dữ liệu kiểm tra, các tập còn lại hợp thành dữ liệu huấn luyện. Có nghĩa là, đầu tiên chạy được thực hiện trên tập $S_2, S_3 \dots S_k$, sau đó test trên tập S_1 ; tiếp tục quá trình dạy được thực hiện trên tập $S_1, S_3, S_4 \dots S_k$, sau đó test trên tập S_2 ; và cứ tiếp tục như thế.

Ví dụ, $k = 10$, thì phương pháp k-fold cross validation được minh họa như hình dưới:

Bảng minh họa phương pháp k-fold cross validation

K=1										
K=2										
...										
K=10										

Trong đó:



Là dữ liệu kiểm tra



Là dữ liệu học

	VIETTEL AI RACE	Public 121
	PHƯƠNG PHÁP VÀ MỘT SỐ ĐỘ ĐO ĐÁNH GIÁ CÁC MÔ HÌNH PHÂN LỚP ĐA NHÂN	Lần ban hành: 1

2. Một số độ đo để đánh giá mô hình phân lớp đa nhãn

Đánh giá kết quả phương pháp phân lớp đa nhãn có sự khác biệt với đơn nhãn. Khóa luận đánh giá các phương pháp phân lớp đa nhãn dựa trên một số độ đo sau: Hamming Loss [15], One-error [15], Coverage [15], Ranking Loss [15], Average Precision, Mean Average Precision.

Cho một tập $S = \{(x_1, Y_1) \dots (x_n, Y_n)\}$ của n ví dụ kiểm tra. Cho $Y^*_i = h(x_i)$ là tập hợp nhãn dự đoán cho kiểm tra x_i , khi Y_i là tập nhãn cho x_i .

Hamming Loss: độ mất mát dữ liệu, được tính như sau:

$$hloss(h) = \frac{1}{n} \sum_{i=1}^n \frac{1}{Q} \sum_{q=1}^Q (\delta(q \in Y^*_i \wedge q \notin Y_i) + \delta(q \notin Y^*_i \wedge q \in Y_i))$$

Trong đó, δ là một hàm mà đầu ra là 1 nếu một nội dung đúng và 0 trong trường hợp ngược lại. Nhỏ hơn giá trị của $hloss(h)$, thực thi tốt hơn. Trong trường hợp lí tưởng, $hloss(h) = 0$.

One-error: đánh giá lỗi cho nhãn xếp hạng đầu:

$$one_error_{S(f)} = \frac{1}{n} \sum_{i=1}^n \delta([\arg \max_{y \in Y} f(x_i, y)] \notin Y_i)$$

Giá trị nhỏ hơn $one_err(f)$, thực thi tốt hơn. Chú ý, vấn đề phân loại cho đơn nhãn, một lỗi chính là phân loại lỗi thường.

Coverage: Để đánh giá hiệu suất của một hệ thống cho tất cả các nhãn của một mẫu (đếm số lượng tất cả các nhãn). Coverage được định nghĩa như khoảng cách trung bình cho tất cả các nhãn thích hợp được gán cho một ví dụ thử nghiệm:

	VIETTEL AI RACE	Public 121
	PHƯƠNG PHÁP VÀ MỘT SỐ ĐỘ ĐO ĐÁNH GIÁ CÁC MÔ HÌNH PHÂN LỚP ĐA NHÃN	Lần ban hành: 1

$$coverages_s(f) = \frac{1}{n} \sum_{i=1}^n \max_{y \in Y_i} rank_f(x_i, y) - 1$$

Ranking Loss: Tính phân bố trung bình của các cặp nhãn:

$$rloss(f) = \frac{1}{n} \sum_{i=1}^n \frac{1}{|Y_i| |\bar{Y}_i|} |\{(y_1, y_2) \in Y_i \times \bar{Y}_i \mid f(x_i, y_1) \leq f(x_i, y_2)\}|$$

Với \bar{Y}_i là phần bù của tập nhãn Y_i . Giá trị của $rloss(f)$ càng nhỏ thì càng tốt.

Average Precision: độ chính xác trung bình của $P@K$ tại các mức K có đối tượng đúng. Gọi $I(K)$ là hàm xác định đối tượng ở vị trí hạng K nếu đúng $I(K) = 1$ và ngược lại $I(K) = 0$, khi đó:

$$AP = \frac{\sum_{K=1}^n P@K \times I(K)}{\sum_{j=1}^n I(j)}$$

Với n là số đối tượng được xét, $P@K$ là độ chính xác của K đối tượng đầu bảng xếp hạng. Xác định số đối tượng đúng ở K vị trí đầu tiên của xếp hạng và gọi là $Match@K$ và ta có

$$P@K = \frac{Match@K}{K}$$

Mean Average Precision: Độ chính xác trung bình trên N xếp hạng. (N truy vấn, mỗi truy vấn có một thứ tự xếp hạng kết quả tương ứng).

$$MAP = \frac{\sum_{i=1}^N AP_i}{N}$$

Xét ví dụ:

	VIETTEL AI RACE	Public 121
	PHƯƠNG PHÁP VÀ MỘT SỐ ĐỘ ĐO ĐÁNH GIÁ CÁC MÔ HÌNH PHÂN LỚP ĐA NHÃN	Lần ban hành: 1

Giả sử có 5 đối tượng được xếp hạng tương ứng là: c, a, e, b, d

Một xếp hạng của đối tượng cần đánh giá là: **c, a, e**, d, b

Khi đó, $P@1 = 1/1$, $P@2 = 2/2$, $P@3 = 3/3$, $P@4 = 3/4$, $P@5 = 3/5$.

$$AP = \left(\frac{1}{1} + \frac{2}{2} + \frac{3}{3} \right)$$

$$MAP = \frac{1}{3} \left(\frac{1}{1} + \frac{2}{2} + \frac{3}{3} \right)$$