

## فصل دهم: جمع بندی یادگیری نظارت شده

ارزیابی طبقه بندی‌های نظارت شده:

بنا به طبقه بندی  $M$  کلاس، را در نظر بگیریم.

هدف: تخمین احتمال حصدی طبقه بندی بر اساس اینکه طبقه بندی از  $N$  دسته چه تعداد را درست یا غلط طبقه بندی می کند.

$$p_i \text{ احتمال حصدی طبقه بندی } \omega_i, \quad \sum_{i=1}^M N_i = N$$

احتمال اینکه  $k_i$  بردارد از کلاس  $\omega_i$  به اشتباه طبقه بندی شده باشد به شرط استقلال برداشته شدن از توزیع در جمله اول

$$\text{prob} \{ k_i \text{ خفا} \} = \binom{N_i}{k_i} p_i^{k_i} (1-p_i)^{N_i-k_i} \quad (10.1) \quad \text{به دست می آید:}$$

$$\max_{p_i} (10.1) = \hat{p}_i = \frac{k_i}{N_i}$$

$$\text{total error prob. } \hat{P} = \sum_{i=1}^M p(\omega_i) \frac{k_i}{N_i}$$

$$E[\hat{P}] = P \quad \text{تخمین بدون بایاس}$$

میانگین نده

$$S_{\hat{P}}^2 = \frac{\sum_{z=1}^M p^2(w_z) \frac{P_z(1-P_z)}{n_z}}{n}$$

consistent  $n_z \rightarrow \infty$

اگر  $n_z$  (دوره ها) به قدر کافی بزرگ باشند تخمین  $\hat{P}$  تخمین مناسب است.

# ماتریس درهم‌آزمایی Confusion Matrix

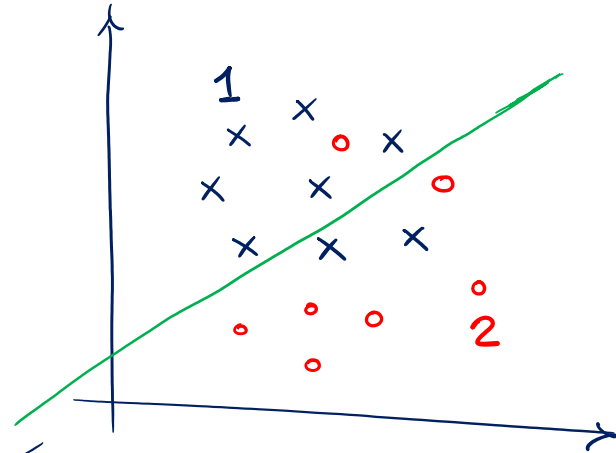
دقت، یادآوری، دقت

$$A = \begin{bmatrix} A(0,0) & & & \\ & A(1,1) & & \\ & & \dots & \\ & & & \dots \end{bmatrix}$$

این  $A(i,j)$  تعداد داده‌هایی که برچسب درست آنها  $i$  و برچسب خرابی طبقه بندی برای آنها  $j$  است.  
 $j$ : predicted

$i$ : true

6	2	$\frac{6}{8} \times 100 = 75\%$
1	6	$\frac{6}{7} \times 100 = 86\%$
$\frac{6}{7} = 86\%$	$\frac{6}{8} = 75\%$	$\frac{12}{15} \times 100 = 80\%$



$$Ac = \frac{1}{N} \sum_{i=1}^m A(i,i) \quad \text{دقت}$$

$(R_i)$  Recall: تعداد داده‌هایی که برچسب  $i$  دارند در دسته درست  $i$  در همین صنف طبقه بندی شده‌اند.  
 $R_1 = \frac{6}{8}$        $R_2 = \frac{6}{7}$        $R_1 = \frac{A(1,1)}{A(1,1) + A(1,2)}$

$(P_i)$  Precision: تعداد داده‌هایی که در صنف  $i$  طبقه بندی شده‌اند و برچسب درست نیز  $i$  بود.  
 $P_1 = \frac{A(1,1)}{A(1,1) + A(2,1)}$